Monitoring the variability of cement compressive strength using Multivariate Exponentially Weighted Moving Covariance Matrix (MEWMC) control chart

M Qori’atunnadyah¹, Wibawati¹*, M Ahsan¹ and H Khusna¹

¹Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

*Corresponding author: wibawati@statistika.its.ac.id

Abstract. Cement quality plays an important role in ensuring the infrastructure development of a nation. PT “X” is a business entity engaged in cement production. The high demand for cement has driven PT “X” to constantly improve the quality under the established standards. In general, the quality of cement is determined by the compressive strength of cement measure on the 3rd, 7th, and 28th days. In this paper, the MEWMC control chart is used to monitor the compressive strength of cement. MEWMC is an exponentially weighted control chart for monitoring the stability of the covariance matrix of a process. Using the optimum weighting value \( \lambda = 0.1 \) in phase two, the monitoring results show that the cement quality has been statistically controlled or there is no process shift.

1. Introduction

Indonesia is a developing country that carries out infrastructure development in various fields. Since 2015, the government has shifted subsidy expenditure to the productive area, namely infrastructure development, health, and education. The infrastructure budget continues to increase from IDR 155 trillion in 2014 to around IDR 410 trillion in 2018. The community's need for infrastructure is still very large. The government targets Indonesia's competitiveness to be ranked 40th in the world. In 2019, the Government's focus will be on developing Indonesian human resources by continuing to develop infrastructure. The infrastructure budget in 2019 increases by 1.04% to Rp 415 trillion with the target of building/reconstruction/widening of the road by 2,007 km, bridge construction and rehabilitation of 27,067 m, construction of 4 new airport units, construction/completion of railroads by 415.2 km's, an irrigation network of 162 thousand hectares and 48 dam units.

Maximum infrastructure development is produced from quality raw materials. One of the raw materials used in construction is cement which is manufactured by PT "X". The high demand for cement has driven PT "X" to always improve the quality of its products following established standards. Good cement quality will add value and the trust of the community. Cement quality is determined by its compressive strength at the final grinding stage. The compressive strength of cement is the ability of the cement to receive high pressure. If the compressive strength of cement is high, then the components made from cement will be stronger to support heavy loads. Consequently, they are not easily damaged. Cement compressive strength can be divided into three, namely cement compressive strength measured on the 3rd, 7th, and 28th days. The three variables are correlated, the longer the age of the cement, the higher the compressive strength. Therefore, it is very important to do quality testing at the final grinding stage before the cement enters the packaging stage.
The Statistical Process Control (SPC) method can be used to monitor the compressive strength of the cement. The main purpose of this method is to quickly detect causes of the out-of-control samples so that investigation and corrective actions can be carried out before many defective units are produced. Control charts are one of the Statistical Process Control tools that are effective in reducing processes variability [1]. The most popular multivariate control chart is $T^2$ Hotelling introduced by Harold Hotelling in 1947 [2]. The advantage of employing this method in monitoring the manufacturing process is easy for application. However, this chart is less sensitive to monitor a small shift in the mean processes. Therefore, the Multivariate Exponentially Weighted Moving Average (MEWMA) (proposed by reference [3]) can be utilized for this issue due to its high sensitiveness of a small shift. The recent development of the MEWMA chart can be found in reference [4] [5] [6] [7] [8].

The shift not only can be occurred in the mean of the process but also happens on the variability of the correlated multivariate characteristics. Hawkins and Maboudou-Tchao [9] proposed the Multivariate Exponentially Weighted Moving Covariance Matrix (MEWMC) chart in monitoring shift in the covariance matrix. In this research, the MEWMC control chart is chosen due to its effectiveness to monitor the changes in the covariance matrix. This chart has an excellent performance to inspect cases with the changes in variance or correlation. Also, MEWMC is robust to monitor the process with the non-multivariate normal distribution. Therefore, in this research, the MEWMC control chart is employed to monitor the variability of cement compressive strength.

2. Methods
This section presents the literature related to the method used for controlling the compressive strength of cement.

2.1. Multivariate normal test
The multivariate normal density is a generalization of the univariate normal density with $p \geq 2$. The univariate normal distribution has a probability distribution function

$$f(x) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}.$$  \hfill (1)

Following the normal distribution function notation with mean $\mu$ and variance $\sigma^2$ obtained

$$\left(\frac{x-\mu}{\sigma}\right)^2 = (x-\mu)^T (\sigma^2) (x-\mu).$$  \hfill (2)

Based on the exponent of the univariate distribution function, it can be generalized that for a $p \times 1$ vector $X$ of observations on several variables as

$$(x-\mu)^T \Sigma^{-1} (x-\mu).$$  \hfill (3)

The vector $\mu_{[p \times 1]}$ represents the expected value of the vector $X$, and the matrix $\Sigma_{(p \times p)}$ is the variance-covariance matrix of the matrix $X$, which $p$ is the number of observations variables, so

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu)}.$$  \hfill (4)

Let $X_1, X_2, \ldots, X_p$ has a multivariate normal distribution, then $(x-\mu)^T \Sigma^{-1} (x-\mu)$ distributed as $\chi^2_p$. A multivariate normal test can be performed using the correlation coefficient test. Using the following
correlation coefficient test can be seen whether the data meets the assumption of normal multivariate distribution or not.

Hypothesis:

\( H_0 \): The data are a multivariate normal distribution

\( H_1 \): The data are not a multivariate normal distribution

Test statistics:

\[
R_q = \frac{\sum_{j=1}^{n}(x_{j} - \bar{x})(q_{j} - \bar{q})}{\sqrt{\sum_{j=1}^{n}(x_{j} - \bar{x})^2} \sqrt{\sum_{j=1}^{n}(q_{j} - \bar{q})^2}},
\]

where \( q_j = \frac{j-1}{n}, j = 1, 2, \ldots, n \).

Critical Region:

Reject \( H_0 \) if \( R_q < r_{(\alpha,n)} \).

If the results of the test statistics have fewer results than the normal table of the correlation coefficient probability, it can be concluded that the data are not following the multivariate normal distribution and otherwise [10].

2.2. Bartlett test

The Bartlett sphericity test aims to determine whether the variance is homogeneous between variables in multivariate cases. If the variables \( X_1, X_2, \ldots, X_p \) are independent of each other, then the correlation matrix between variables is the same as the identity matrix [11]. The Bartlett sphericity test can be stated in the following hypothesis:

\( H_0 \): \( R=I \) (there is no correlation between variable)

\( H_1 \): \( R \neq I \) (there is a correlation between variable)

Test Statistics:

\[
\chi^2 = -\left[ n - 1 - \frac{2p + 5}{6} \right] \ln |R|,
\]

where \( n \) is the number of observations, while \( p \) is the number of variables, and \( R \) is the correlation matrix of each variable and \( \chi^2_{\left(\alpha, \frac{1}{2}p(p-1)\right)} \) is the value of the chi-square distribution. If the significance level is set \( \alpha \) at 0.05, then \( H_0 \) is rejected if \( \chi^2 > \chi^2_{\left(\alpha, \frac{1}{2}p(p-1)\right)} \) or it can be concluded that the correlation matrix is not the same as the identity matrix or there is a correlation between variables [11].

2.3. Multivariate Exponentially Weighted Moving Covariance Matrix (MEWMC)

Suppose that measurement is \( X \), a p-component vector, is assumed to follow a multivariate normal distribution. Assume that the process will be monitored by taking a sample of \( n \geq p \) independent observation vectors at the sampling point. Let \( X_{jk} \) represent observation \( j \) (\( j = 1, 2, \ldots, n \)) for variable \( k \) (\( k = 1, 2, \ldots, p \)). The use of Multivariate Exponentially Weighted Moving Covariance Matrix (MEWMC) is easier when using standardized vector data than using raw data. Therefore, the matrix \( A \) is calculated using the following equation
\[ \mathbf{A} \Sigma \mathbf{A}^T = \mathbf{I}_p. \]  

(7)

Then transformed using the formula as follows

\[ \mathbf{U}_j = \mathbf{A}(\mathbf{X}_j - \mathbf{\mu}_j), \]  

(8)

where when in control \( \mathbf{U}_j \sim N(0, \mathbf{I}_p) \).

Matrix A is the lower triangular (inverse-Cholesky root) matrix, which corresponds to the “cascade” regression adjustment of Hawkins (1993). The notation \( \mathbf{S} \) is standard for an estimated covariance matrix, we recycle notation and define the MEWMC by the recursion \( \mathbf{S}_0 = \mathbf{I}_p \) and for \( j = 1, 2, ..., n \), then

\[ \mathbf{S}_j = (1 - \lambda)\mathbf{S}_{j-1} + \lambda \mathbf{U}_j \mathbf{U}_j^T, \]  

(9)

When the process is in control, the expectation \( E(\mathbf{S}_j) = \mathbf{I}_p \).

The test statistics used in the MEWMC control chart are as follows

\[ c_j = \text{tr}(\mathbf{S}_j) - \log |\mathbf{S}_j| - p \]  

(10)

The control limit (\( h \)) is chosen based on the weighted value (\( \lambda \)), Average Run Length (ARL), many dimensions, and simulation length. After getting the \( c_j \) and \( h \) values, a control chart can be formed which displays a plot of \( c_j \) against \( j \) and indicates out of control if \( c_j > h \) [9].

3. Numerical result

This section discusses the numerical result of controlling the compressive strength of cement.

3.1. Multivariate normal test

A multivariate normal test is used to determine whether the research data is following the multivariate normal distribution or not. In the multivariate normal test, the testing is carried out based on the variables used. The correlation between \( d_j^2 \) and \( q_j \) is 0.993 which is a test statistic from the multivariate normal test. This value is then compared to the critical point from the normal probability plot correlation coefficient (PPCC) distribution. For alpha equal to 5%, the critical point is 0.9955. Therefore, the obtained correlation coefficient value is smaller than the critical point. Consequently, \( H_0 \) is rejected and the data are not following the multivariate normal distribution.

3.2. Bartlett test

Bartlett test is performed to determine whether there is a correlation between the variables used in the multivariate case. In this test, the value \( \chi^2 \) is 22,229. For alpha equal to 5% and degrees of freedom equal to 3, it is obtained that the value of \( \chi^2_{[0.05, 3]} \) equal to 7.8147 (smaller than \( \chi^2 \) was 22,229). In the other words, it can be decided that \( H_0 \) is rejected. In a conclusion, the correlation matrix of the variables is not an identity matrix or there is a correlation between the variables used.

3.3. MEWMC control charts

After the test the assumptions, the next step is monitoring the compressive strength of cement using a control chart. The type of control chart used in this research is the Multivariate Exponentially Weighted Moving Covariance Matrix (MEWMC) control chart where can be used to monitor variability through changes in the covariance matrix. Cement quality control in this research was carried out using two phases. Phase one is carried out to get the optimum weighting value. Meanwhile, phase two is carried out to monitor cement quality when there is a low demand for ALM production for easy combustion, which is then raised again to save iron sand (Fe) usage. In this research, the 1st to
255th production observations are used in the first phase. Process variability control in phase one is carried out to obtain the optimum weighting value. The estimated optimum weight is used to monitor the process in phase two. In this research, three weight values were used, that is 0.1, 0.2, and 0.3. These values were chosen because using larger weight values can produce a biased result [9]. Furthermore, \( c_t \) is the test statistic of the MEWMC control chart. The following steps are detailed to create a MEWMC control chart.

1. Calculate the matrix \( A \) from the cement compressive strength data using equation 7. The calculated matrix \( A \) is presented as follows:

\[
A = \begin{bmatrix}
0.0789 & 0 & 0 \\
-0.0795 & 0.0858 & 0 \\
0.0010 & -0.0544 & 0.0608
\end{bmatrix}
\]

2. Transform the data using the formula in equation 8. The test calculations for the 1st and 2nd production data are obtained as follows:

For \( j=1 \), then

\[
U_j = A(X_j - \mu)
\]

\[
= \begin{bmatrix}
0.0789 & 0 & 0 \\
-0.0795 & 0.0858 & 0 \\
0.0010 & -0.0544 & 0.0608
\end{bmatrix}
\begin{bmatrix}
190 \\
251 \\
361
\end{bmatrix}
= \begin{bmatrix}
204.0216 \\
270.7808 \\
360.5717
\end{bmatrix} - \begin{bmatrix}
1.1063 \\
0.5825 \\
1.0881
\end{bmatrix}
\]

For \( j=2 \), then

\[
U_j = A(X_j - \mu)
\]

\[
= \begin{bmatrix}
0.0789 & 0 & 0 \\
-0.0795 & 0.0858 & 0 \\
0.0010 & -0.0544 & 0.0608
\end{bmatrix}
\begin{bmatrix}
190 \\
251 \\
361
\end{bmatrix}
= \begin{bmatrix}
204.0216 \\
270.7808 \\
360.5717
\end{bmatrix} - \begin{bmatrix}
1.1063 \\
0.5825 \\
1.0881
\end{bmatrix}
\]

For \( U_j \) where \( j=3,4, \ldots, 255 \) is calculated in the same way as \( U_1 \) and \( U_2 \).

3. Calculate the value of \( S_j \) using the formula in equation 9. The calculation of \( S_j \) for the 1st and 2nd production data using \( \lambda = 0.3 \) is presented as follows:

For \( j=1 \), then

\[
S_1 = (1 - \lambda) S_0 + \lambda U_1 U_1
\]

\[
= (1 - 0.3) \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
+ 0.3 \begin{bmatrix}
-1.1063 \\
-0.5825 \\
1.0881
\end{bmatrix}
= \begin{bmatrix}
1.0672 & 0.1933 & -0.3611 \\
0.1933 & 0.8018 & -0.1901 \\
-0.3611 & -0.1901 & 1.0552
\end{bmatrix}
\]

For \( j=2 \), then

\[
S_2 = (1 - \lambda) S_1 + \lambda U_2 U_2
\]

\[
= (1 - 0.3) \begin{bmatrix}
1.0672 & 0.1933 & -0.3611 \\
0.1933 & 0.8018 & -0.1901 \\
-0.3611 & -0.1901 & 1.0552
\end{bmatrix}
+ 0.3 \begin{bmatrix}
0.0784 \\
0.0772 \\
-0.4057
\end{bmatrix}
= \begin{bmatrix}
0.8976 & 0.1517 & -0.3390 \\
0.1517 & 0.5630 & -0.1425 \\
-0.3390 & -0.1425 & 0.7880
\end{bmatrix}
\]

and so on until you get \( S_{255} \).
4. Calculate the test statistics of the MEWMC control chart $c_j$. The calculation process of $c_j$ for the 1st and 2nd production data using $\lambda = 0.3$ is detailed as follows:

For $j=1$, then

$$c_1 = tr(S_1) - \log |S_1| - p$$

$$= (0.0672 + 0.1933 + (-0.3611)) - \log \begin{pmatrix} 0.0672 & 0.1933 & -0.3611 \\ 0.1933 & 0.8018 & -0.1901 \\ -0.3611 & -0.1901 & 1.0552 \end{pmatrix} - 3$$

$$= 2.9241 - \log(0.7468) - 3 = 0.2160$$

For $j=2$, then

$$c_2 = tr(S_2) - \log |S_2| - p$$

$$= (0.8976 + 0.5630 + 0.7880) - \log \begin{pmatrix} 0.8976 & 0.1517 & -0.3390 \\ 0.1517 & 0.5630 & -0.1425 \\ -0.3390 & -0.1425 & 0.7880 \end{pmatrix} - 3$$

$$= 2.2486 - \log(0.3118) - 3 = 0.413959$$

in the same way for $j=3,4,\ldots, 255$ until all the values for $c_j$ are found.

5. Calculate the control limit $h$ of MEWMC control chart for each value of lambda $\lambda$, ARL$_0$ of 370, $p$ is 3, and the number of simulations is 1000.

6. Plot the calculated MEWMC statistics $c_j$ and control limits in one chart.

3.4. Results and discussions

- Phase one

Figure 1 shows a plot of $c_j$ for each production data with the same control limit value. The control limit value is by a weighting value $\lambda = 0.1$ so that the upper control limit is 0.934 with a lower control limit of 0. Based on this weighting value, the MEWMC control chart of 255 observations is found to be uncontrolled in variability where seven observations are out of upper control limits (UCL), which is 14th, 15th, 16th, 17th, 18th, 50th, and 51st observations. Monitoring is out of control which occurs because the company has an adjustment of the calibration curve. Therefore, the process has not been controlled statistically.
For the weighting value $\lambda = 0.2$, the estimated upper control limit value is 2.0930 with a lower control limit is 0. Based on this weighting value, there are five production data found outside of the upper control limit (see Figure 2) namely $10^{th}$, $14^{th}$, $48^{th}$, $49^{th}$, and $97^{th}$ observations. Monitoring is out of control occurs because the company had a calibration curve adjustment at the $10^{th}$, $14^{th}$, $48^{th}$, and $49^{th}$ observations. Whereas in the $97^{th}$ observation there was a clinker experiment with low ALM and $C_3A$. Therefore, it can be concluded that the process has not been controlled statistically.

For weighting value $\lambda = 0.3$, the estimated upper control limit value is 3.4173 and the lower control limit is 0. Based on this weighting value, three production data are out of the upper control limit (see Figure 3) namely $10^{th}$, $48^{th}$, and $97^{th}$ observations. Monitoring is out of control occurs because the company had a calibration curve adjustment at the $10^{th}$ and $48^{th}$ observations. Whereas in the $97^{th}$ observation there was a clinker experiment with low ALM and $C_3A$. Therefore, the process has not been controlled statistically.

The first step is used to obtain the optimum weighting value after plotting is to look at the control chart for each $\lambda$. The control chart that has more out of control observations is selected because it is more sensitive than the control chart that does not have out of control data. For $\lambda = 0.1$, $\lambda = 0.2$, and $\lambda = 0.3$ it can be found 7, 5 and 3 out of control observations, respectively. Based on these results, it can be seen that $\lambda = 0.1$ (see Figure 1) produces more out-of-control data (7 data) compared to the other weighting values. Thus, $\lambda = 0.1$ is chosen as the most optimum weighting for phase one of the monitoring process using the MEWMC control chart.

For the weighting value $\lambda = 0.1$, it can be seen in Figure 1 that seven production data are out of control observation, which one of them is the $17^{th}$ observations. After technical inspection at the $17^{th}$ observations, which is the highest out of control data, an event occurred where the ALM (Alumina Modulus) is not positive for the compressive strength of cement. After removing all of the out of control data, the process was found to be in control.
control production data, Figure 4 shows that there are no points that fall on the outside of the control limit.

![Figure 4](image1.png)

**Figure 4.** MEWMC control chart in phase one using $\lambda = 0.1$ (in control)

- **Phase two**
  After the monitoring process in phase one complete, process variability control is carried out in phase two which aims to monitor the production process on the 256th to 320th production data. In the production data, there is a low demand for ALM production for ease of combustion which is then raised again to save iron sand (Fe) usage. The weighting value used in phase two is the optimum weighting value in phase one ($\lambda = 0.1$).

![Figure 5](image2.png)

**Figure 5.** MEWMC control chart in phase two using $\lambda = 0.1$

Figure 5 depicts phase two of the monitoring variability process of cement production data. Using the weighted value $\lambda = 0.1$ and control limit of 0.9340, it can be concluded that phase two of the monitoring process is controlled statistically (there are no data that are out of control).

4. **Conclusion**
In this paper, the compressive strength of cement is monitored using the MEWMC control chart. First, the optimum weighting value for the MEWMC control chart needs to be determined to obtain the best performance. From the numerical results, the variability process of phase one is still not statistically controlled when using $\lambda = 0.1, 0.2$, and 0.3. Further, due to the more out of control data produced, the optimum weighting value $\lambda = 0.1$ is chosen. This value is then applied in phase two of the monitoring process and found that there is no shift in the process. Thus, according to the phase two results, it can be concluded that the variability of the compressive strength of cement is statistically controlled. The CUSUM type chart in monitoring covariance matrix such as in [12] can be considered for future work. The use of a squared trace correlation matrix as demonstrated in [13] and robust method as proposed
in [14] can improve the performance of the MEWMC control chart. Also, Kernel Density Estimation (KDE) and Bootstrap methods [15] can be employed to estimate the control limit of the proposed chart.

References

[1] Montgomery D C 2013 Introduction to Statistical Quality Control (7th ed) (United States of America: John Wiley & Sons, Inc.)
[2] Hotelling H 1947 Multivariate quality control (New York: McGraw-Hill)
[3] Cynthia A L, William H W, Champ C W and Steven E R 1992 A Multivariate Exponentially Weighted Moving Average Control Chart Technometrics 34 46-53
[4] Haq A and Khoo M B C 2019 Memory-type multivariate control charts with auxiliary information for process mean Quality and Reliability Engineering International 35 192-203
[5] Haq A and Khoo M B C 2019 An adaptive multivariate EWMA chart Computers & Industrial Engineering 127 549-557
[6] Hidayatul K, Mashuri M, Suhartono, Prastyo D D and Ahsan M 2018 Multioutput Least Square SVR Based Multivariate EWMA Control Chart Journal of Physics: Conference Series 1028
[7] Hidayatul K, Mashuri M, Suhartono, Prastyo D D and Ahsan M 2018 Multioutput least square SVR-based multivariate EWMA control chart: The performance evaluation and application Cogent Engineering 5
[8] Steven E R and William H W 2017 Using the predictive distribution to determine control limits for the Bayesian MEWMA chart Communications in Statistics - Simulation and Computation 46 7818-7826
[9] Hawkins D M and Maboudou-Tchao E M 2008 Multivariate Exponentially Weighted Moving Covariance Matrix Technometrics 50 55-166
[10] Johnson R A and D W Wichern 2007 Applied Multivariate Statistical Analysis (6th ed) (New Jersey: Prentice Hall)
[11] Morrison D F 2005 Multivariate Statistical Methods (4th ed) (USA: McGRAW-Hill Book Company)
[12] Lai K C and Jian Z 2001 Cumulative Sum Control Charts for the Covariance Matrix Statistica Sinica 11 767–790
[13] Mashuri M, Haryono H, Aksioma D F, Wibawati W, Ahsan M and Khusna H 2019 Tr (R2) Control Charts based on Kernel Density Estimation for Monitoring Multivariate Variability Process Cogent Engineering 6 1665949
[14] Ahsan M, Mashuri M, Lee M H, Kuswanto H and Prastyo D D 2020 Robust Adaptive Multivariate Hotelling’s T² Control Chart based on Kernel Density Estimation for Intrusion Detection System Expert Systems with Applications 145 113105
[15] Ahsan M, Mashuri M and Khusna H 2018 Intrusion Detection System Using Bootstrap Resampling Approach Of T² Control Chart Based On Successive Difference Covariance Matrix Journal of Theoretical and Applied Information Technology 96 2128–38