Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P5913-cd

Monitoring Confounder-adjusted Principal Component Scores with an Application to Load Test Data

Lizzie Neumann

Dept. of Mathematics and Statistics, Helmut Schmidt University Hamburg, Germany.
E-mail: neumannl@hsu-hh.de

In structural health monitoring (SHM), measurements from various sensors are collected and often reduced to damage-sensitive features. Diagnostic values for damage detection are then obtained through statistical analysis of the measurements or features. However, the system outputs, i.e., sensor measurements or extracted features, depend not only on damage but also on confounding factors (environmental or operational variables). These factors affect the mean and the covariance. The latter is particularly important because the covariance is often used as an essential building block in damage detection tools. This paper discusses a nonparametric kernel estimator for estimating the conditional covariance matrix, allowing it to vary based on the confounding variable. This improves the understanding of how factors, such as temperature, influence system outputs. Additionally, a method for calculating confounder-adjusted scores using conditional principal component analysis is described, thus adjusting not only the mean but also the covariance. The technique is applied to monitor real-world data from the Vahrendorfer Stadtweg bridge in Hamburg, Germany, using a MEWMA control chart.

Keywords: Bootstrapping, Conditional covariance, Confounder-adjusted scores, Kernel method, MEWMA control chart, Monitoring, Principal component analysis, Structural Health Monitoring, Temperature removal.

1. Introduction

Structural health monitoring (SHM) is the process of utilizing sensors to automate the assessment of damage in engineered systems within the fields of civil, mechanical, and aerospace engineering (Farrar and Worden, 2013). It includes acquiring measurement data, extracting damage-sensitive features, and their statistical evaluation, for example, with unsupervised methods such as principal component analysis (Reynders et al., 2014; Wang et al., 2022). However, system outputs, i.e., sensor measurements or extracted damage-sensitive features, change due to damage as well as environmental and operational parameters (EOP). Therefore, the effects of the EOPs need to be eliminated. In SHM, this process is referred to as “data normalization” (Farrar and Worden, 2013), and there are various comprehensive literature reviews available (e.g., Avci et al., 2021; Han et al., 2021; Wang et al., 2022). Some methods for diagnosing damage, such as neural networks (Avci et al., 2021), include data normalization during the training phase, while other approaches require data normalization as a separate step. Following the data normalization, the system outputs are utilized in other post-processing methods, such as statistical process control (SPC) (e.g., Magalhães et al., 2012; Wang and Ong, 2008; Wittenberg et al., 2024) or model updating approaches.

All data normalization methods mentioned in the literature, e.g., the cited literature reviews, have in common that they only focus on the mean, i.e., the expected value of the system outputs. However, Neumann et al. (2025) demonstrated that the EOPs affect not only the mean but also the (co-)variances. To extract confounder-adjusted scores, this paper uses a conditional version of the principal component analysis (Neumann et al., 2025) that can also adjust for confounding effects on second-order statistical moments. These scores are then utilized for monitoring real-world load test data from the Vahrendorfer Stadtweg bridge (Köhncke et al., 2024) in Hamburg, Germany.

The remainder of this article is structured as follows. In Section 2, the Vahrendorfer Stadtweg bridge is presented. Section 3 discusses the confounder-adjusted covariance estimator, conditional principal component analysis, and control charts. Section 4 illustrates the application to the data, and Section 5 concludes.

2. Vahrendorfer Stadtweg Bridge

The Vahrendorfer Stadtweg bridge is a 50-meter-long and 10-meter-wide prestressed concrete bridge over Highway ("Autobahn") A7 in the south of Hamburg, Germany. Constructed in 1972, it features a single lane for agricultural traffic and a pedestrian walkway on the southeastern side of the bridge. It was built with an open frame design and a box girder cross-section. Figure 1 shows the bridge from different angles. For the

Fig. 1. Vahrendorfer Stadtweg bridge view from north (top) and northeast (bottom left), and three additional masses at the middle of the bridge (bottom right).

analysis, measurements of six strain sensors (in z-direction) and one material temperature sensor are used. The sampling frequency of 200 Hz (strain) and 10 Hz (temperature) was downsampled (averaged) to one measurement per hour, compare Han et al. (2021). Recording of measurements started in April 2023. In the following year, between February 22nd and March 23rd, 2024, load tests were carried out. Those consisted of three additional masses (680 kg, 740 kg, 740 kg), which were placed in the middle of the bridge in a stepwise increasing procedure. For ten days, the first mass was placed (scenario A). Then, the second mass was added for another ten days (scenario B), and finally, the third mass was added for another ten days (scenario C). Scenario C is shown in Figure 1 (bottom right). For a more comprehensive description of the bridge and further load tests, see Köhncke et al. (2024).

3. Theoretical Background

Principal component analysis (PCA) is a widely-used mathematical tool in SHM, for example, for feature extraction (Tibaduiza et al., 2016; Zhu et al., 2019). Neumann et al. (2025) introduced a novel approach for conditional principal component analysis using a conditional covariance estimator. The conditional eigenvalues and principal components obtained through conditional PCA can be used to calculate confounder-adjusted scores. This will be revisited in Section 3.1. Section 3.2 presents the MEWMA control chart with a monitoring scheme for the scores.

3.1. *Conditional Principal Component Analysis and Feature Extraction*

Let $\mathbf{x} = (x_1, \dots, x_p)^\top \in \mathbb{R}^p$ be a p-dimensional random (output) vector and $z \in \mathbb{R}$ a potential confounder. Furthermore, let $\mathbf{\Sigma}(z)$ denote the *conditional* covariance matrix of $\mathbf{x}$ given z. Then, the Nadaraya-Watson kernel estimator used in this paper has the following form (Yin et al., 2010; Neumann et al., 2025):

$$\hat{\mathbf{\Sigma}}(z;h) = \left(\sum_{i=1}^{n} K_h(z_i - z) \left[\mathbf{x}_i - \hat{\mathbf{m}}(z_i)\right]\left[\mathbf{x}_i - \hat{\mathbf{m}}(z_i)\right]^\top \right) \left(\sum_{i=1}^{n} K_h(z_i - z) \right)^{-1}, \tag{1}$$

where $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^\top$, $i = 1, \dots, n$, are observations of $\mathbf{x}$, and z_i is the associated confounder variable (e.g., temperature). $K_h(\cdot)$ is a kernel function with bandwidth h, and $\hat{\mathbf{m}}(z_i)$ is an estimate of the mean of $\mathbf{x}$ at z_i. The conditional mean vector $\mathbf{m}(z)$ can be estimated utilizing methods such as penalized regression splines

(Eilers and Marx, 1996; Neumann and Gertheiss, 2022), local polynomial regression (Cleveland et al., 2017) or a Nadaraya-Watson kernel estimator (Yin et al., 2010; Neumann et al., 2025). The optimal bandwidth h can be estimated through cross-validation as, e.g., described in Yin et al. (2010) or Neumann et al. (2025). The bandwidth is the smoothing parameter of the kernel, i.e., the higher the bandwidth, the wider the kernel, and the smoother the estimate. The case $h \to \infty$ corresponds to the marginal covariance. In this paper, a Gaussian kernel is used, which is equivalent to a normal density with a mean of zero. The estimation was done with the statistical software R (R Core Team, 2024), and an example code for estimating the conditional covariance is available from Neumann (2024).

Conditional PCA (Neumann et al., 2025) can then be performed by applying the eigendecomposition to the conditional output covariance matrix $\mathbf{\Sigma}(z)$, Eq. (1),

$$\mathbf{\Sigma}(z) = \mathbf{A}(z)\mathbf{\Lambda}(z)\mathbf{A}(z)^\top, \tag{2}$$

where the matrix $\mathbf{\Lambda}(z) = \text{diag}(\lambda_1(z), \ldots, \lambda_p(z))$ holds the conditional eigenvalues in decreasing order and $\mathbf{A}(z) = [\mathbf{a}_1(z) \ \ldots \ \mathbf{a}_p(z)]$ the corresponding eigenvectors, the *principal components*. As the conditional covariance requires the confounder to be measured, it makes conditional principal component analysis a supervised method rather than an unsupervised one.

Once the eigenvalues and principal components are estimated on the in-control data (denoted by $\hat{\lambda}_j$ and $\hat{\mathbf{a}}_j$), they can be used for *conditional* feature extraction. "In-control" means that only common causes but no special causes, such as damage, are affecting the system. For that purpose, we extract the corresponding scores (Neumann et al., 2025)

$$\mathbf{s}_i = (\mathbf{x}_i - \hat{\mathbf{m}}(z_i))^\top \hat{\mathbf{A}}(z_i)(\hat{\mathbf{\Lambda}}(z_i))^{-1/2}, \tag{3}$$

$i = 1, \ldots, n$, with $\hat{\mathbf{A}}(z_i) = [\hat{\mathbf{a}}_1(z_i) \ \ldots \ \hat{\mathbf{a}}_p(z_i)]$, $(\hat{\mathbf{\Lambda}}(z_i))^{-1/2} = \text{diag}(\hat{\lambda}_1^{-1/2}(z_i), \ldots, \hat{\lambda}_p^{-1/2}(z_i))$, and $\hat{\mathbf{m}}(z_i)$ being an estimate of the conditional mean. Using the conditional mean, eigenvalues, and principal components removes the effect of the confounder z from the component scores. For in-control data, those are uncorrelated, standardized quantities, each with mean zero and variance one for any given z-value.

3.2. *Control charts*

The Multivariate Exponentially Weighted Moving Average (MEWMA) control chart (Lowry et al., 1992; Wittenberg et al., 2024) is used to monitor the scores $\mathbf{s}_i \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Gamma})$ in Eq. (3). In reference to Knoth (2017), a mean vector is defined following a change point model, i.e., $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ if $i < \tau$ and $\boldsymbol{\mu} = \boldsymbol{\mu}_1$ if $i \geq \tau$ for an unknown time point τ, is introduced. As mentioned above, the in-control scores should be of zero mean, hence, $\boldsymbol{\mu}_0 = 0$, and are assumed to be uncorrelated with variances $\nu_1, \ldots, \nu_p$ equal to one. The covariance matrix Γ is diagonal with $\Gamma = \text{diag}(\nu_1, \ldots, \nu_p)$. Then, a smoothing procedure is applied to estimate the MEWMA statistic

$$\boldsymbol{\omega}_i = (1-\kappa)\boldsymbol{\omega}_{i-1} + \kappa \mathbf{s}_i, \quad \boldsymbol{\omega}_0 = 0, \tag{4}$$

with time point $i = 1, 2, \ldots$ and smoothing parameter $0 < \kappa \leq 1$. The smoothing parameter κ controls the sensitivity of the shift to be detected; specifically, smaller values such as $\kappa \in \{0.1, 0.2, 0.3\}$, are usually selected to detect smaller shifts (Hunter, 1986), whereas $\kappa = 1$ results in the Hotelling chart (Hotelling, 1947). The control statistic is the Mahalanobis distance

$$T_i^2 = (\boldsymbol{\omega}_i - \boldsymbol{\mu}_0)^\top \mathbf{\Gamma}_{\boldsymbol{\omega}}^{-1} (\boldsymbol{\omega}_i - \boldsymbol{\mu}_0), \tag{5}$$

with the asymptotic covariance matrix $\mathbf{\Gamma}_{\boldsymbol{\omega}}$ of $\boldsymbol{\omega}_i$, $\mathbf{\Gamma}_{\boldsymbol{\omega}} = \lim_{i\to\infty} \text{Cov}(\boldsymbol{\omega}_i) = \left(\frac{\kappa}{2-\kappa}\right)\mathbf{\Gamma}$. The MEWMA control chart will trigger an alarm if the control statistic T_i^2 exceeds the threshold value h_4. If the process is in-control, the average run length (ARL_0), i.e., the number of observations until a signal occurs, should be high to minimize false alarms. This ARL_0 can be computed following the procedure outlined in Knoth (2017) and is available in the R-package spc (Knoth, 2022; R Core Team, 2024). One requirement for its application is that the scores must not be auto-correlated, i.e., correlated with themselves over time. As this is the case for the data considered in this paper, a *block bootstrapping* procedure will be used to determine the average run length and control limit, following the procedure outlined below:

(i) In-control score data $\{\mathbf{s}_i,\, i = 1,\ldots,n\}$ is divided into blocks $\{B_k,\, k = 1,\ldots,m\}$ of fixed size, where m is the maximum number of available days. Then, block B_k contains the estimated scores of day k.
(ii) These blocks are resampled with replacement to generate a new dataset $\{\tilde{\mathbf{s}}_i,\, i = 1,\ldots,n\}$ while maintaining the original data's temporal structure.
(iii) The average run length (ARL_0) is computed using Eq. (4) and (5), and 10^5 repetitions, and the control limit h_4 is determined by a grid search.

4. Application & Results

The methods from Section 3 are applied in the following to the Vahrendorfer Stadtweg data described in Section 2. The strain measurements from July 2nd, 2023, to February 17th, 2024, are used as Phase I data. Thus, Phase I consists only of in-control data and is used for training. The measurements from February 18th to March 23rd, 2024, are used as Phase II data, i.e., a combination of in-control and out-of-control data (the three added mass scenarios), and are used for validation and monitoring. First, the conditional mean was estimated using penalized regression splines (Eilers and Marx, 1996; Neumann and Gertheiss, 2022), and the optimal bandwidth h for the conditional covariance matrix was estimated for each entry separately following Neumann et al. (2025) and is between 0.1 and 0.2. Additional smoothing was necessary due to sparse data between $10°$C and $15°$C and to ensure that the covariance matrices for all z-values are positive semi-definite. Therefore, a bandwidth of 1.2 was chosen. Subsequently, the conditional covariance was estimated following Eq. (1) for each sensor pair separately.

In the following, three different methods to estimate the scores are compared, each using a different type of temperature adjustment:

(i) *Standard*: The marginal covariance is used for PCA, and the scores are estimated using the standard PCA, but only considering the second to sixth principal components. The first score is excluded from the estimation of the MEWMA control chart, assuming that the first principal component mainly accounts for operational and environmental effects (Cross et al., 2012).
(ii) *Residuals*: Residual data are calculated, i.e., the conditional mean is subtracted from the measurements to remove the temperature influence. This refers to the methods of "data normalization" through mean adjustment discussed in the introduction. The standard PCA is then applied based on the marginal covariance. All principal components are used to estimate the scores.
(iii) *Conditional*: The eigendecomposition, following Eq. (2), is applied, and the conditional scores are estimated as in Eq. (3) to remove the temperature influence in the mean *and* in the covariance. As with the previous method, all principal components are used to estimate the scores.

4.1. *Scores*

First, all six scores are calculated using the three described methods (i)-(iii). Figure 2 shows the marginal distribution of the Phase I scores for each of the three methods (columns) in form of histograms. In theory, the scores should follow a standard normal distribution, so the standard normal density is plotted in black for comparison. The first standard score (first column) follows a bimodal distribution. As described above, a possible reason for this could be the impact of temperature; hence, it was excluded from further analysis. Furthermore, the first standard scores using the residual data (second column) and the first conditional score (third column) are not normally distributed. The bimodal distribution of the first conditional score could be due to a second, unobserved confounder. The second to sixth scores approximately follow standard normal distributions, with some slightly skewed, indicating a tail on the right or left side, respectively.

The influence of the temperature can also be seen in the mean of the scores. Figure 3 shows the first three scores as a function of the temperature with the conditional mean in black. The first standard score is highly correlated with the tem-

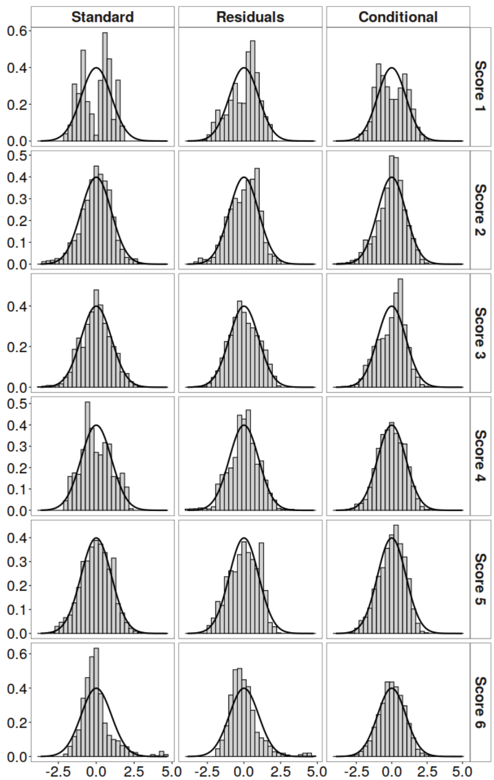

Fig. 2. Distributions (gray histograms) of the scores (rows) for Phase I data for the standard, standard using residual data, and conditional method (columns) in comparison with the standard normal density (black).

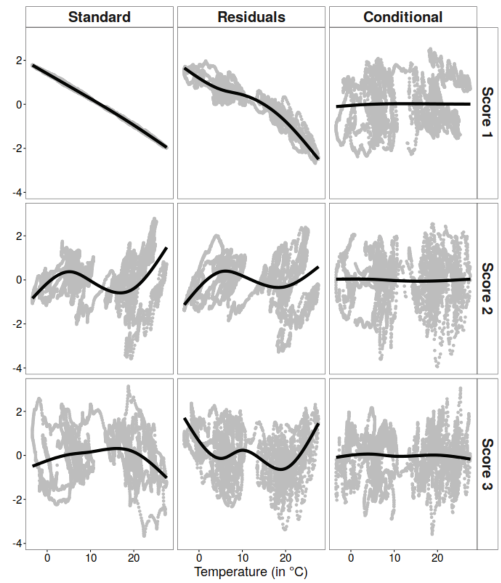

Fig. 3. First three standardized scores (gray dots, rows) of the three methods (columns) with their conditional mean (black) as a function of temperature.

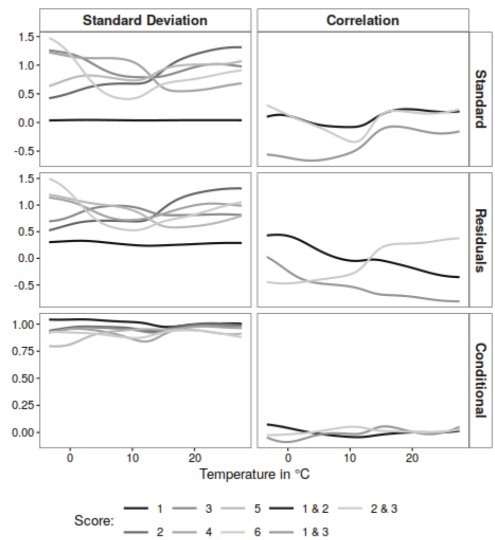

Fig. 4. Conditional standard deviation (left column) and correlation (right column) of the standard scores (top), standard scores using residual data (middle), and conditional scores (bottom).

perature, but the dependence is also visible in the second and third scores, however, in a non-linear way. As seen from the second column of Figure 3, simply subtracting the estimated, temperature-dependent mean from the strain data is not sufficient either to remove the temperature influence in the scores. However, using conditional PCA from Section 3.1 eliminates the temperatures' effect on the scores' mean as can be seen in the third column of Figure 3. All conditional scores have a constant mean of zero.

Furthermore, as described in Section 3.1, the scores should be uncorrelated and have a standard deviation of one for all temperature values. Therefore, Figure 4 shows the conditional standard deviation (left column) and correlation (right column) for the scores for the standard method (top row), the standard method using the residual data (middle row), and the conditional approach (bottom row). The standard deviation and correlation vary across different temperature values for the two standard methods while showing minimal to no variation for the conditional method, as they are (nearly) constant around one and zero, respectively. The most noticeable variation occurs in the standard deviation of the fifth score for negative temperatures. This demonstrates that the covariance of scores from both standard methods continues to be influenced by temperature, whereas the covariance of the conditional scores is nearly constant across all temperature values. Thus, only the conditional scores are uncorrelated and standardized, with a mean of zero and a variance of one.

4.2. *Monitoring*

To apply a monitoring procedure to the adjusted scores of the previous section, a MEWMA control chart is set up with $\kappa = 0.2$ for all three methods (i)-(iii). Figure 5 shows the control chart of the standard scores using only the second to sixth principal components (top), the standard scores using the residual data (middle), and the conditional scores (bottom). A logarithmic scale is used for display. The black vertical dotted line indicates the start of the Phase II data. The light gray, gray, and dark gray shaded areas indicate the stepwise placement of the additional masses one to three (scenario A-C), respectively. The black horizontal lines are the control limits $h_4 = 117.12$ (standard), $h_4 = 119.92$ (residuals), and $h_4 = 103.9$ (conditional) and were estimated via block bootstrapping with an ARL_0 of 370.4. The control statistic is lighter gray above the control limit and darker gray below. Furthermore, the figure shows that for both standard methods (i) and (ii), false alarms occurred at the beginning of the control chart and in December 2023. This could be reduced using the conditional method (iii). In particular, the false alarms in December 2023 were eliminated.

Moreover, only the conditional method triggers the alarm for all three added mass scenarios. For the standard method (i), the control limit was only exceeded in the middle of the second and third added mass scenarios. The standard method

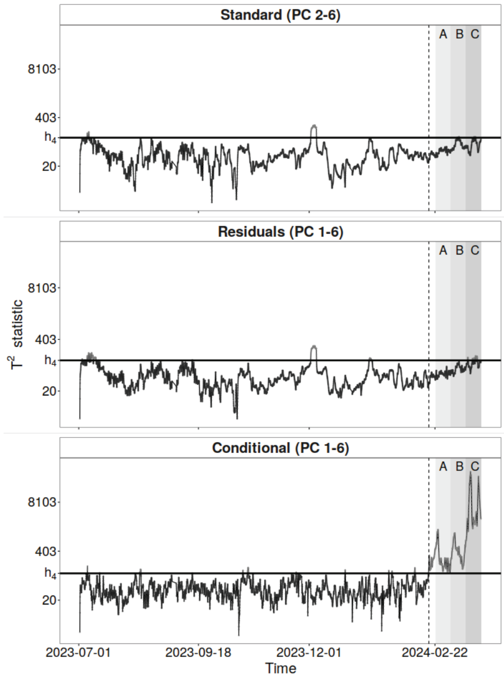

Fig. 5. MEWMA control chart in logarithmic scale of the standard scores (top, PC 2-6), standard scores using residual data (middle, PC 1-6) and conditional scores (bottom, PC 1-6) with $\kappa = 0.2$ and $h_4 = 117.12$, $h_4 = 119.92$, and $h_4 = 103.9$ (black horizontal line), respectively. The dotted vertical line marks the beginning of the Phase II data, and the shaded areas correspond to the different added mass scenarios: A (light gray), B (gray), and C (dark gray).

using the residual data (ii) even reduces the alarms during the second added mass scenario (B, gray), but during the third added mass scenario (C, dark gray), the control limit is exceeded almost continuously. However, the test statistic is still lower than for the false alarms in December 2023. The peak of the control statistic for the conditional method (iii) at the beginning of each added mass scenario can be attributed to the car crossing the bridge while towing a trailer that carried the additional mass. This was further influenced during the first added mass scenario (A, light gray) by a wheel loader that needed to reposition the added mass. However, the peaks did not occur with either standard method.

Table 1 summarizes the false alarms and probability of detection for each of the three methods and different values of parameter κ. As mentioned in Section 3.2, the parameter κ controls the sensitivity of the shift to be detected. Therefore, the values $\kappa = 0.1, 0.3, 1$ were added for comparison. Phase I consists of 5429 data points. As can be seen in the table, the false alarms for the standard method are lower than for the standard method using the residual data but still are rather high for both methods, up to 269 and 457, respectively. Also, the probability of detection is higher for the standard method using the residual data. Although, the probability of detection is quite low for both standard methods with values between 0% and 25.4%. However, the conditional method produces fewer false alarms than the standard methods, only between 77 and 124, and has a probability of detection between 93% and 100% for different values of κ. Therefore, using the confounder-adjusted scores for monitoring appears to be the most appropriate option.

Table 1. Summary of false alarms and probability of detection (POD) per method and parameter $\kappa \in \{0.1, 0.2, 0.3, 1\}$.

		Standard (PC 2-6)	**Residuals** (PC 1-6)	**Conditional** (PC 1-6)
κ	h_4	false alarms		
0.1	192.82	291 [5.4%]		
	192.71		457 [8.4%]	
	162.21			124 [2.3%]
0.2	117.12	119 [2.2%]		
	119.92		269 [5.0%]	
	103.9			94 [1.7%]
0.3	79.73	115 [2.2%]		
	82.47		224 [4.2%]	
	70.87			100 [1.8%]
1	15.65	112 [2.1%]		
	16.67		157 [2.9%]	
	16.11			77 [1.4%]
κ	h_4	probability of detection [%]		
0.1	1192.82	19.8		
	192.71		25.4	
	162.21			100
0.2	117.12	3.3		
	119.92		8.6	
	103.9			100
0.3	79.73	1.4		
	82.47		6.1	
	70.87			99.7
1	15.65	0		
	16.67		3.1	
	16.11			93.2

5. Conclusion

The main contribution of this paper is the discussion of a method for estimating and monitoring confounder-adjusted scores of sensor measurements. This method requires the measurement of the confounding variable (such as temperature or operational loads) and estimating the conditional covariance matrix to account for the influence of confounding variables in the covariance. The estimate is utilized for conditional principal component analysis. Subsequently, using the conditional mean, eigenvalues, and principal components, conditional scores can be calculated and monitored through a MEWMA control chart. In a comparative study, the conditional method and two standard approaches – one that uses the measurements while omitting the first principal component and another that uses the residual data – were applied to strain data from the Vahrendorfer Stadtweg bridge in Hamburg, Germany. These three methods illustrated three different techniques to adjust for temperature influence. The presented confounder-adjusted method for estimating the scores ensured that their mean and covariance were no longer influenced by temperature, in contrast to both standard methods. Furthermore, the conditional method could substantially reduce the number of false alarms while increasing the probability of detection compared to the standard methods. Reducing false alarms is crucial because each bridge closure resulting from a false alarm undermines trust in the monitoring system and leads to user dissatisfaction. In this paper, only the influence of temperature was considered, while other confounding effects were disregarded. Future research will involve extending this work to account for multiple confounders using neural networks.

Acknowledgement

This research paper out of the project 'SHM – Digitalisierung und Überwachung von Infrastrukturbauwerken' is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU. Thanks to Philipp Wittenberg for providing support in finding the optimal control limit.

References

Avci, O., O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. J. Inman (2021). A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mechanical Systems and Signal Processing 147*, 107077.

Cleveland, W. S., E. Grosse, and W. M. Shyu (2017). Local regression models. In *Statistical models in S*, pp. 309–376. Routledge.

Cross, E. J., G. Manson, K. Worden, and S. G. Pierce (2012). Features for damage detection with insensitivity to environmental and operational variations. In *Proceedings of the Royal Society A*, Volume 486, pp. 4098–4122.

Eilers, P. H. C. and B. D. Marx (1996). Flexible smoothing with B-splines and penalties. *Statistical Science 11*(2), 89–121.

Farrar, C. and K. Worden (2013). *Structural Health Monitoring: A Machine Learning Perspective*. John Wiley & Sons, Ltd.

Han, Q., Q. Ma, J. Xu, and M. Liu (2021). Structural health monitoring research under varying temperature condition: a review. *Journal of Civil Structural Health Monitoring 11*, 149–173.

Hotelling, H. (1947). Multivariate quality control illustrated by the air testing of sample bombsights. In C. Eisenhart, M. Hastay, and W. Wallis (Eds.), *Techniques of Statistical Analysis*, New York, pp. 111–184. McGraw-Hill.

Hunter, J. S. (1986). The exponentially weighted moving average. *Journal of Quality Technology 18*(4), 203–210.

Knoth, S. (2017). ARL numerics for MEWMA Charts. *Journal of Quality Technology 49*(1), 78–89.

Knoth, S. (2022). *spc: Statistical Process Control – Calculation of ARL and Other Control Chart Performance Measures*. R package version 0.6.7.

Köhncke, M., Y. Jaelani, A. Mendler, L. Neumann, P. Wittenberg, A. Rode-Klemm, and S. Keßler (2024). Static and Dynamic Load Tests on the Bridge Vahrendorfer Stadtweg. arXiv:2412.15713.

Lowry, C. A., W. H. Woodall, C. W. Champ, and S. E. Rigdon (1992). A Multivariate Exponentially Weighted Moving Average Control Chart. *Technometrics 34*(1), 46–53.

Magalhães, F., A. Cunha, and E. Caetano (2012). Vibration based structural health monitoring of an arch bridge: From automated oma to damage detection. *Mechanical Systems and Signal Processing 28*, 212–228.

Neumann, L. (2024). *Estimation of the Conditional Mean, Covariance and Correlation of the Frequency Data from Railway Bridge KW51*. https://github.com/neumannLizzie/conditional-covariance.

Neumann, L. and J. Gertheiss (2022). Covariate-adjusted association of sensor outputs for structural health monitoring. In D. Schulz, A. Fay, W. Matiaske, and M. Schulz (Eds.), *dtec.bw-Beiträge der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg*, Volume 1, pp. 287–291.

Neumann, L., P. Wittenberg, A. Mendler, and J. Gertheiss (2025). Confounder-adjusted covariances of system outputs and applications to structural health monitoring. *Mechanical Systems and Signal Processing 224*, 111983.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reynders, E., G. Wursten, and G. D. Roeck (2014). Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification. *Structural Health Monitoring 13*(1), 82–93.

Tibaduiza, D. A., L. E. Mujica, J. Rodellar, and A. Güemes (2016). Structural damage detection using principal component analysis and damage indices. *Journal of Intelligent Material Systems and Structures 27*(2), 233–248.

Wang, Z. and K. Ong (2008). Autoregressive coefficients based hotelling's t2 control chart for structural health monitoring. *Computers & Structures 86*(19), 1918–1935.

Wang, Z., D.-H. Yang, T.-H. Yi, G.-H. Zhang, and J.-G. Han (2022). Eliminating environmental and operational effects on structural modal frequency: A comprehensive review. *Structural Control and Health Monitoring 29*(11), e3073.

Wittenberg, P., L. Neumann, A. Mendler, and J. Gertheiss (2024). Covariate-Adjusted Functional Data Analysis for Structural Health Monitoring. arXiv:2408.02106.

Yin, J., Z. Geng, R. Li, and H. Wang (2010). Nonparametric covariance model. *Statistica Sinica 20*, 469–479.

Zhu, Y., Y.-Q. Ni, D. Inaudi, and I. Laory (2019). A temperature-driven MPCA method for structural anomaly detection. *Engineering Structures 190*, 447–458.