Check for updates

Received: 26 September 2024 | Accepted: 16 January 2026

DOI: 10.1111/bmsp.70035

the british psychological society
promoting excellence in psychology

ARTICLE

Detecting association changes in intensive longitudinal data in real time: An exponentially weighted moving average procedure

Evelien Schat | **Sarah Schrevens** | **Francis Tuerlinckx** | **Eva Ceulemans**

Quantitative Psychology and Individual Differences, Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

Correspondence
Evelien Schat, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102 Box 3713, B-3000 Leuven, Belgium.
Email: evelien.schat@kuleuven.be

Funding information
Research Council of KU Leuven, Grant/Award Number: C14/23/062 and PDMT2/23/019; Research Foundation Flanders (FWO) and F.R.S.-FNRS under the Excellence of Science (EOS) program, Grant/Award Number: EOS 40007528/G0I2422N; Research Foundation Flanders (FWO), Grant/Award Number: 12AA826N

Abstract

Within-person changes in linear associations may indicate worsening well-being and maladaptive functioning. We investigated whether such changes can be detected in real time using the exponentially weighted moving average (EWMA) procedure. Specifically, we investigated the effectiveness of first calculating association strength within time windows, considering several association measures (i.e. Pearson correlation, Spearman correlation, Pearson covariance, Penrose shape distance, Euclidean distance, Lorentzian distance, Manhattan distance and squared Euclidean distance), and then monitoring mean-level changes in these scores using EWMA. Additionally, we examined how changes in the mean and variance in the observed time series (with or without a correlation change) influence the detection performance of EWMA when applied to association scores. Our simulation results show that monitoring Pearson and Spearman correlation scores is advised, when no additional information is available about the presence of additional mean and/or variance changes in the observed time series. However, using other association measures, which are sensitive to more types of changes apart from the correlation (i.e. mean and/or variance), can improve detection performance given specific combinations of mean, variance and correlation changes. Using other measures can thus be valuable when the presence of such a combination of changes can be predicted before monitoring begins.

KEYWORDS

association measures, exponentially weighted moving average, real-time monitoring

1 | INTRODUCTION

In psychological research, intensive longitudinal data are collected to gain insight into dynamic psychological processes and to understand how differences in these processes, both between and within individuals, are related to well-being and maladaptive functioning (Gable & Reis, 1999; Hamaker & Wichers, 2017; Lafit et al., 2021). An important feature of these dynamic psychological processes is both the presence and strength of associations between the variables that are tracked across time (Cramer et al., 2016; Pe et al., 2015; Scheffer et al., 2018; Shin et al., 2022; Van De Leemput et al., 2014). Given that most research to date has focused on linear associations, we largely restrict our scope to these associations in this study. Research has indeed shown that the strength of these within-person linear associations differs between (groups of) individuals, on top of differences in mean levels and variances. For instance, a stronger bipolar relationship between positive and negative affect is associated with greater severity of depressive symptoms (Dejonckheere et al., 2018). Additionally, it has been shown that such within-person associations may shift over time in response to changes in contexts. For example, the correlation between momentary positive and negative affect becomes more negative as a result of personally meaningful negative events, such as receiving examination results (Dejonckheere et al., 2021).

Given these indications that changes in linear associations between positive and negative affect may signal changes in well-being, it is crucial to monitor for the occurrence of such changes, as detecting them in *real time* is essential for timely intervention. For example, when the relation between negative and positive affect changes significantly as a result of anxiously waiting for examination results, one can intervene immediately to help alleviate the stress experienced. This could involve using just-in-time adaptive interventions (Spruijt-Metz & Nilsen, 2014).

Detecting changes in linear association in real time is not trivial, however. A first challenge we face is determining which measure of linear association is most appropriate to monitor. In the above-mentioned studies on depression and examination results, the association between positive and negative affect was quantified using the Pearson correlation—a measure that is often relied upon as the primary measure of association (Allik et al., 2015; Sels et al., 2020). However, changes in correlation often go hand in hand with changes in mean and/or variance (e.g. depressed individuals tend to experience more negative affect and greater affect variability compared to healthy individuals; Houben et al., 2015). This raises the question whether it is better to use a measure that focuses solely on linear association, irrespective of means and variances, or one that is sensitive to multiple types of changes simultaneously. Indeed, emerging research underscores the need to explore and consider alternative measures of association. Carlier et al. (2024) examined a wide range of measures and their applicability in psychological research, where the measures varied in their sensitivity to different types of changes (i.e. correlation, variance and mean changes).[1]

A second challenge is *how* we can detect changes in the strength of association in real time. Indeed, most analysis techniques for modelling changes in intensive longitudinal data (e.g. change-point detection methods: Albers & Bringmann, 2020; Cabrieto et al., 2018; Sels et al., 2022; regime-switching models: Chow et al., 2013; Chow & Zhang, 2013; Crayen et al., 2017; Stifter & Rovine, 2015) are designed for retrospective analyses, once all data have been collected. However, the exponentially weighted moving average (EWMA) procedure (Roberts, 1959)—originally developed for monitoring the mean level of a quality characteristic in industrial settings—appears to be particularly useful. While other real-time monitoring procedures exist for detecting mean changes, such as the Shewhart procedure (Shewhart, 1931) and the cumulative sum procedure (Page, 1954), we focus on EWMA because it is well suited to detect smaller, sustained changes, which are often observed in psychological processes, and is less sensitive to parameter tuning (Hawkins & Wu, 2014; Schat et al., 2023). The EWMA procedure first captures the natural variation present in a set of in-control data, by estimating the mean and standard deviation of the scores of a particular individual. These estimates are then used to compute a person-specific upper and lower control limit. These

[1]Correlation, variance and mean changes are sometimes also referred to as changes in shape, scatter and level similarity, respectively.

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

limits can afterwards be used to monitor incoming data over time. As long as the monitored scores stay within the control limits, the monitored process is considered to be 'normal' (or in control). When a score goes beyond one of the limits, the process deviates from what is normal (or goes out of control), which may call for intervention. Given that EWMA is primarily designed to detect mean changes over time, the question arises: How can it be adapted to monitor for changes in linear associations? Note that several correlation-based monitoring procedures have been proposed in the literature (e.g. MEWMS, MEWMV and MEWMC; Hawkins & Maboudou-Tchao, 2008; Huwang et al., 2007; Reynolds Jr & Cho, 2006); however, to the best of our knowledge, no existing real-time monitoring procedure focuses solely on detecting correlation changes without also being sensitive to other types of changes (i.e. variance, mean). Developing such a procedure would be highly valuable for research that primarily focuses on detecting within-person changes in linear associations.

The goal of this study is to address the two abovementioned challenges by applying EWMA not to the original observed scores, but to scores of association measures. In order words, building on work by Cabrieto et al. (2018) and Schat et al. (2023, 2024), we first calculate association measure scores within time windows and then monitor these scores to detect mean-level changes in association strength using EWMA. We aim to investigate how well this approach works when plugging in different measures of association. Additionally, we investigate how changes in mean and variance (with or without a correlation change) influence detection performance. In this study, we focus on eight association measures that were studied in Carlier et al. (2024): Euclidean distance, Lorentzian distance, Manhattan distance, Penrose shape distance, Squared Euclidean distance, Pearson correlation, Spearman correlation and Pearson covariance.[2] While this is not an exhaustive list of all possible measures, we focus on a subset of frequently used association measures that can be calculated with a small number of data points, making them suitable for the shorter time series typically encountered in real-time monitoring applications in psychological research (Schreuder et al., 2024; Smit et al., 2019; Snippe et al., 2023) and serving as a first step in exploring the feasibility of our approach.

The remainder of this paper is organized as follows. First, we describe a dataset in which a pilot's mental load is measured while performing various tasks. These data serve as an illustrative example to explain the association measures. Next, we discuss the different association measures examined in this study. We then explain the EWMA procedure and illustrate its application using the mental load data. Following this, we report on a simulation study where we evaluate EWMA applied to the different association measures. Finally, we present a discussion of the results and suggest directions for future research.

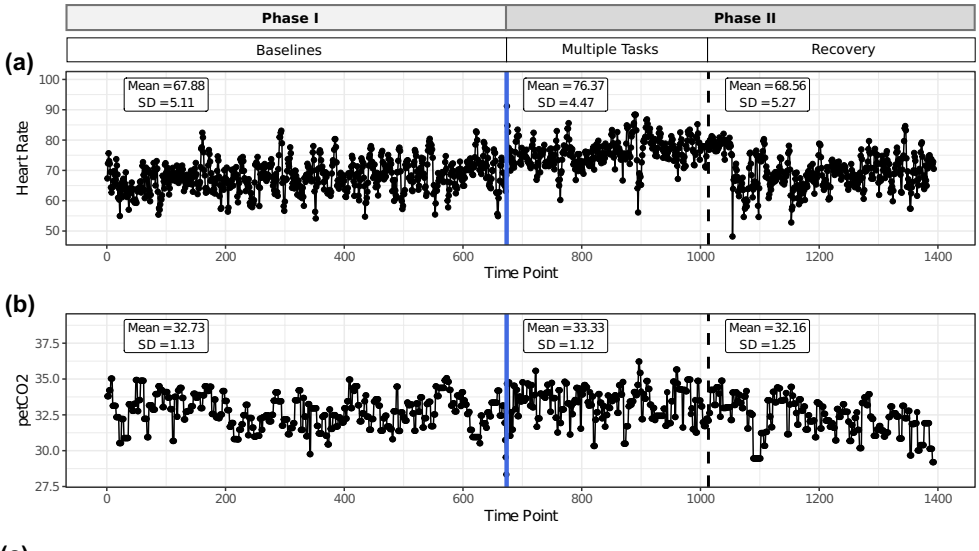
2 | EWMA APPROACH FOR DETECTING ASSOCIATION CHANGES

2.1 | Mental load data

The experimental data on the mental load of an aviation pilot were collected by Grassmann et al. (2016) and are publicly available through the `kcpRS` R package (Cabrieto et al., 2022). During the experiment, physiological measures were monitored to evaluate the pilot's mental load while performing various highly demanding tasks. Here, we focus on the physiological measures heart rate and $petCO_2$, both measured on a continuous scale. The experiment was divided into four periods:

[2]We excluded several of the distance measures studied by Carlier et al. (2024): Avg(L_1, L_n), Burrows' delta, Chebyshev's distance, Gower distance, half-range standardized distance, maximum scaled difference and scaled Euclidean distance. These measures were excluded because they require the calculation of summary statistics such as the maximum or range on data from a single time point (i.e. two observations—one from each variable), which makes them highly susceptible to outliers. We also excluded Cattell's r_p and mean censored Euclidean distance, as they assume certain data preprocessing steps, such as setting variable means to zero. Finally, we do not report on the following remaining measures: mean character distance (also known as signal matching), Penrose size distance and Spearman covariance. The first two measures produced the same simulation results as the Manhattan distance, while Spearman covariance yielded results identical to those of the Spearman correlation.

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



(c)

Phase	Period	Pearson correlation	Spearman correlation	Pearson covariance	Penrose shape distance	Euclidean distance	Lorentzian distance	Manhattan distance	Squared Euclidean distance
I		-.12	-.11	-.26	12.30	111.23	35.40	347.86	12513.22
II	Multiple task	.07	.04	-.11	9.97	135.56	37.71	427.23	18498.75
	Recovery	-.05	-.13	-.17	12.98	116.35	36.13	365.19	13649.43

FIGURE 1 Raw data and average association scores of heart rate and $petCO_2$. The raw data of heart rate and $petCO_2$ are shown in (a) and (b), respectively. The blue vertical line marks the boundary between phases I and II. The dashed vertical line marks the boundary between the experimental periods in phase II: Multiple tasks and recovery. The mean and SD are given for phase I and for the two periods in phase II. (c) The averages of the association scores for phases I and II, calculated in non-overlapping windows of 10 time points. The phase II averages are shown separately for the multiple task and recovery periods.

resting baseline, vanilla baseline, multiple tasks and recovery. In the resting baseline period, the participant was asked to focus on a cross displayed on the screen. During the vanilla baseline, the participant engaged in a minimally demanding vigilance task intended to minimize anticipatory arousal and enhance baseline measure accuracy (Jennings et al., 1992). The multiple task period required the participant to simultaneously perform three tasks involving perceptual speed, spatial orientation and working memory (for more details, see Grassmann et al., 2016). Lastly, in the recovery period, the participant watched a calming underwater movie. Each period lasted approximately 6 min, and the periods were separated by a break of 3 min. We used the data of the first two periods (i.e. resting and vanilla baselines) to calculate the control limits; these two periods will be labelled phase I. The remaining data (i.e. multiple task and recovery periods; labelled phase II) were used for monitoring. To mitigate the impact of artefacts caused by speech or movement during period transitions, the first and last 30 s of each phase were removed, resulting in 1393 observations per physiological measure. Furthermore, we introduced jitter to the data (sampled from $N(0, .0001)$) to ensure that each association measure could be computed.[3]

The raw data of heart rate and $petCO_2$ are shown in Figure 1a,b, respectively, revealing some mean level and variance changes. For instance, the mean level of heart rate and $petCO_2$ both increase during the multiple-task period and decrease again during the recovery period.

[3]In some time windows, the standard deviation was equal to 0, which prevented calculation of the measures. We dealt with this by adding a very small amount of random noise to the data.

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Overview of association measures.

Measures	Formula	Type of change(s) the measure is sensitive to
Pearson correlation	$\frac{\sum_{i=1}^{n}(a_i-\bar{a})(b_i-\bar{b})}{\sqrt{\sum_{i=1}^{n}(a_i-\bar{a})^2}\sqrt{\sum_{i=1}^{n}(b_i-\bar{b})^2}}$	Correlation
Spearman correlation	$\frac{\sum_{i=1}^{n}(R(a_i)-\overline{R(a)})(R(b_i)-\overline{R(b)})}{\sqrt{\sum_{i=1}^{n}(R(a_i)-\overline{R(a)})^2}\sqrt{\sum_{i=1}^{n}(R(b_i)-\overline{R(b)})^2}}$, with $R()$ = the rank of the score on the variable	Correlation
Pearson covariance	$\frac{\sum_{i=1}^{n}(a_i-\bar{a})(b_i-\bar{b})}{n-1}$	Correlation, variance
Penrose shape distance	$\sqrt{\sum_{i=1}^{n}\left((a_i-\bar{a})-(b_i-\bar{b})\right)^2}$	Correlation, variance
Euclidean distance	$\sqrt{\sum_{i=1}^{n}(a_i-b_i)^2}$	Correlation, variance, mean
Lorentzian distance	$\sum_{i=1}^{n}\ln(1+\lvert a_i-b_i\rvert)$	Correlation, variance, mean
Manhattan distance	$\sum_{i=1}^{n}\lvert a_i-b_i\rvert$	Correlation, variance, mean
Squared Euclidean distance	$\sum_{i=1}^{n}(a_i-b_i)^2$	Correlation, variance, mean

2.2 | Association measures

To detect association changes over time, the next step is to calculate association measures within specified windows. We used non-overlapping windows of 10 time points, which per measure results in 68 scores in phase I, and 38 and 31 in multiple task and recovery periods in phase II, respectively, and are summarized in Figure 1c.

We used the following eight association measures: Pearson correlation, Spearman correlation, Pearson covariance, Penrose shape distance, Euclidean distance, Lorentzian distance, Manhattan distance and squared Euclidean distance (see Table 1 for an overview). These measures vary in their sensitivity to changes in correlation, variance and/or mean, as demonstrated in Figure 2. Each panel in Figure 2 displays hypothetical scores of two variables over time, with blue representing one variable (*a*) and yellow representing the other (*b*). Note that *a* and *b* represent heart rate and $petCO_2$ in the mental load data. Figure 2a shows our hypothetical baseline data, where the Pearson and Spearman correlations between the two variables equal 1, the standard deviations of both variables are 1.73, and the means of *a* and *b* are −1 and 0, respectively. In Figure 2b–h, the data for variable *a* remain constant, while variable *b* undergoes a change—the scores are permuted, leading to a lower Pearson correlation with *a*, an increase in mean, and/or an increase in standard deviation. The impact of these changes on the eight association measures is illustrated. In the following subsections, we will discuss how the different measures are affected by the changes (an overview can be found in Table 1).

2.2.1 | Measures only sensitive to correlation changes

The Pearson and Spearman correlations are *only* sensitive to correlation changes. The Pearson correlation measures the strength of the linear relationship between the observed scores of two variables *a* and *b*.

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

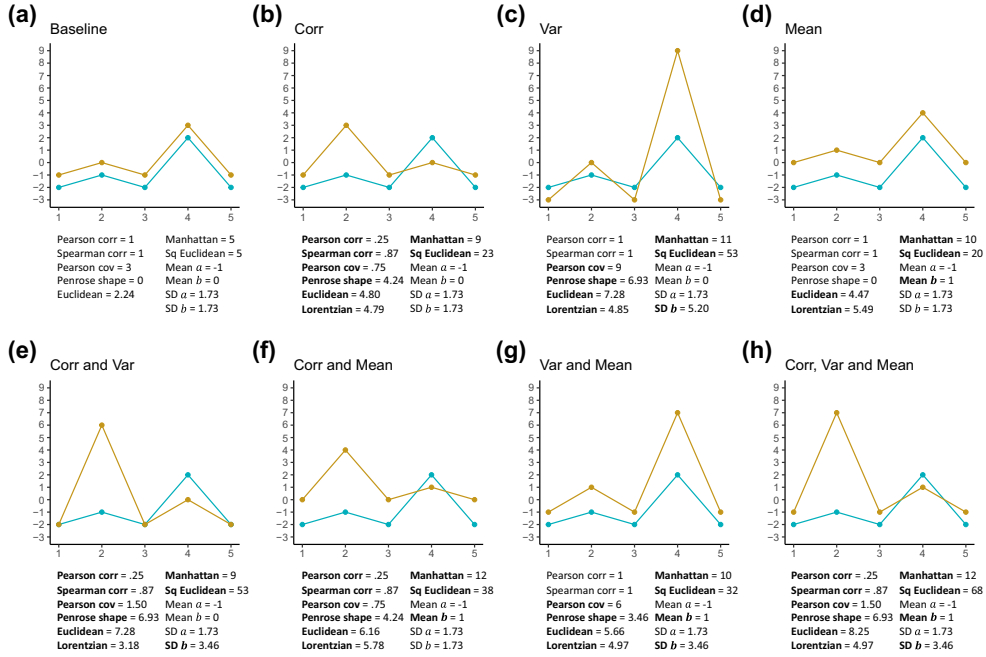


FIGURE 2 Influence of changes in correlation, variance and/or mean on the association measures. The yellow and blue lines correspond to variables *a* and *b*, respectively. Measures that change due to the introduced change (i.e. correlation, variance and/or mean), in comparison with the baseline data in (a), are highlighted in bold. Panels b–h show specific combinations of changes. Corr, correlation; Sq Euclidean, squared Euclidean; Var, variance.

The Spearman correlation measures the strength of the linear association between the rank order of the scores on the two variables. In Figure 2, we observe a decrease in both the Pearson and Spearman correlations, when only a correlation change is introduced (Figure 2b), as well as when it is combined with a variance and/or mean change (Figure 2e,f,h). Only changing the variance (Figure 2c) and/or the mean (Figure 2d,g), does not impact the correlations.

2.2.2 | Measures sensitive to correlation and variance changes

The Pearson covariance and Penrose shape distance are sensitive to both correlation and variance changes. The Pearson covariance is the unstandardized form of the Pearson correlation, making it sensitive to changes in variance as well. To calculate the Penrose shape distance, we take the squared difference between the variable-centred scores of *a* and *b* at each time point *i*. The Penrose shape distance is then calculated by taking the square root of the sum of these squared differences. Hence, decreasing the correlation only leads to a higher Penrose shape distance score, as we see in Figure 2b.

In our example, both correlation and variance changes—whether occurring separately, together, or in combination with a mean change (Figure 2b,c,e–h)—lead to Pearson covariance and Penrose shape distance scores that differ from the baseline scores (Figure 2a), illustrating their sensitivity to these changes. However, if we only introduce a mean change (Figure 2d), the scores of these two measures are unaffected.

2.2.3 | Measures sensitive to correlation, variance and mean changes

The remaining measures—the Euclidean, Lorentzian, Manhattan and squared Euclidean distances—are sensitive to all three types of changes (i.e. correlation, variance and mean). To calculate the Euclidean

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

distance, we compute the squared difference in the scores of variables *a* and *b* at each time point *i*, and then take the square root of the sum of these squared differences. To calculate the Lorentzian distance, we first compute the absolute difference between the scores of variables *a* and *b* at each time point *i*. We then apply the logarithmic function to these differences, after adding one to the differences,[4] and take the sum of these logarithmic values. The Manhattan distance measures similarity by summing the absolute differences between the scores of variables *a* and *b* at each time point *i*. Finally, for the squared Euclidean distance, we first calculate the Euclidean distance and then square the resulting value. For all four distance measures, decreasing only the correlation leads to higher distance scores, as we observe in Figure 2b. Since these four measures are sensitive to changes in variance and mean as well, we see however in Figure 2 that the association measures yield different scores whenever any type of change is introduced as compared to the baseline scores (Figure 2a), complicating interpretation.

2.3 | EWMA procedure

After calculating the association scores within specified windows, we apply the EWMA procedure (Roberts, 1959) to monitor for mean-level changes in these scores. First, we use the association scores from the phase I data to estimate the in-control mean (μ_1) and standard deviation (σ_1). These estimates are then used to calculate the upper (UCL) and lower (LCL) control limits of the EWMA procedure:

$$\mathrm{UCL} = \hat{\mu}_1 + L\hat{\sigma}_1\sqrt{\frac{\lambda}{(2-\lambda)}} \tag{1}$$

and

$$\mathrm{LCL} = \hat{\mu}_1 - L\hat{\sigma}_1\sqrt{\frac{\lambda}{(2-\lambda)}} \tag{2}$$

The constant L determines the width of the control limits and is set based on the desired performance of the EWMA procedure (see Section Performance evaluation). The parameter λ, which also influences the phase II EWMA scores, is explained in the next paragraph. These control limits are statistically well established and allow us to keep the type I error rate of the decision taken each window under control (for a more in-depth explanation, see Schat et al., 2023).

Next, in phase II, we monitor the association scores by transforming them into EWMA scores and comparing these with the phase I control limits. The EWMA procedure monitors a weighted average of the current and all previous association scores, where the highest weight is given to the current score and exponentially decreasing weights are assigned to the previous scores over time. Specifically, at each measurement occasion t (i.e. each window in which an association measure is calculated), we compute the exponentially weighted moving average score z_t for $t = [1, \ldots, w]$:

$$z_t = \lambda x_t + (1-\lambda)z_{t-1} \tag{3}$$

x_t represents the association score at measurement occasion t, and z_{t-1} refers to the EWMA score at the previous measurement occasion. The starting value z_0 equals the estimated mean $\hat{\mu}_1$, based on the phase I data. The constant λ ($0 < \lambda \leq 1$) is the weight[5] given to x_t. Based on earlier applications, we set λ to .10 (Schat et al., 2023, 2024; Smit et al., 2023; Smit & Snippe, 2022; Snippe et al., 2023).

[4]Adding one to the absolute difference before applying the logarithmic function ensures that small differences are accounted for and prevents the logarithm from becoming undefined when the difference between a_i and b_i is equal to zero.

[5]A weight between .05 and .25 is typically recommended, where lower values of λ are used to detect smaller changes in the monitored process (Montgomery, 2009). Simulation and empirical results suggest that EWMA performance is fairly robust across different small λ values (i.e. ≤ .15), with only minor differences observed (Schat et al., 2024; Snippe et al., 2023).

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.3.1 | Performance evaluation

The performance of EWMA is usually evaluated in terms of the run length (RL), which is the measurement occasion in phase II at which the first EWMA score goes out of control (i.e. beyond one of the control limits; Montgomery, 2009). However, the RL, which depends on L and λ, can vary a lot across replicates from the same monitored process (for more details, see Schat et al., 2024). Therefore, the average run length (ARL) is commonly reported, where we make a distinction between the ARL_0 and ARL_1. The ARL_0 denotes the average run length in case the process remains in control in phase II. This value is ideally large, as any out-of-control score is a false positive, where an ARL_0 of 370 is a commonly chosen value for EWMA. Based on an ARL_0 of 370 and a λ of .10, we can determine the L value with the `spc` package in R (Knoth, 2020); given $ARL_0 = 370$ and $\lambda = .10$, $L = 2.70$. The ARL_1, on the other hand, denotes the average run length in case the process changes at the start of phase II. This value is ideally low, as it indicates the power of detecting the change. Fast detection of changes allows for early interventions.

2.4 | Application of EWMA to the mental load data

Figure 3 shows the control charts resulting from applying EWMA to the association scores of heart rate and $petCO_2$. Red dots indicate out-of-control EWMA scores that exceed the control limits, and the dashed vertical line marks the boundary between the multiple task and recovery periods. Around this boundary, the EWMA control charts for the Pearson and Spearman correlations show out-of-control scores, indicating a short-lived increase in the correlation between heart rate and $petCO_2$. During the recovery period, the EWMA scores return within the control limits, reflecting a decrease in correlation between the two variables. EWMA applied to the Penrose shape distance scores results in out-of-control scores immediately after the period boundary, indicating greater similarity between the variables. In contrast, the EWMA control chart for the Pearson covariance shows no out-of-control scores. The EWMA control charts for the Euclidean, Lorentzian, Manhattan and squared Euclidean distances exhibit many out-of-control scores during the multiple task period, likely reflecting the increase in mean level of heart rate during this period (Figure 1a). During the recovery period, the EWMA scores decrease and return within the control limits. However, towards the end of this period, the scores again exceed the upper control limit (except for the Lorentzian distance), reflecting another increase in the mean level of heart rate.

The different patterns of out-of-control EWMA scores across the different association measures underscore the need for a more thorough investigation. It is crucial to explore how specific combinations of changes and their directions may improve or worsen detection performance and to examine which measure can be recommended for which types of data.

3 | SIMULATION STUDY

We conducted a simulation study to evaluate and compare the performance of EWMA applied to the association measures[6] in detecting correlation changes. Additionally, we explored the impact on the results when a change in the mean and/or variance is introduced, with or without an accompanying correlation change. We assessed their performance in detecting these changes by computing ARL values. The R code to reproduce the simulation study is available at https://osf.io/egjfx.

[6]We conducted an additional simulation study in which first the Pearson and Spearman correlations were calculated, then these correlations were Fisher Z-transformed to obtain approximately normally distributed scores, and finally, EWMA was applied to these transformed scores. The results show that EWMA applied to the Fisher Z-transformed Pearson and Spearman correlation scores is slightly more sensitive but less specific than when applied to the untransformed Pearson and Spearman correlation scores. The results can be found at https://osf.io/egjfx.

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

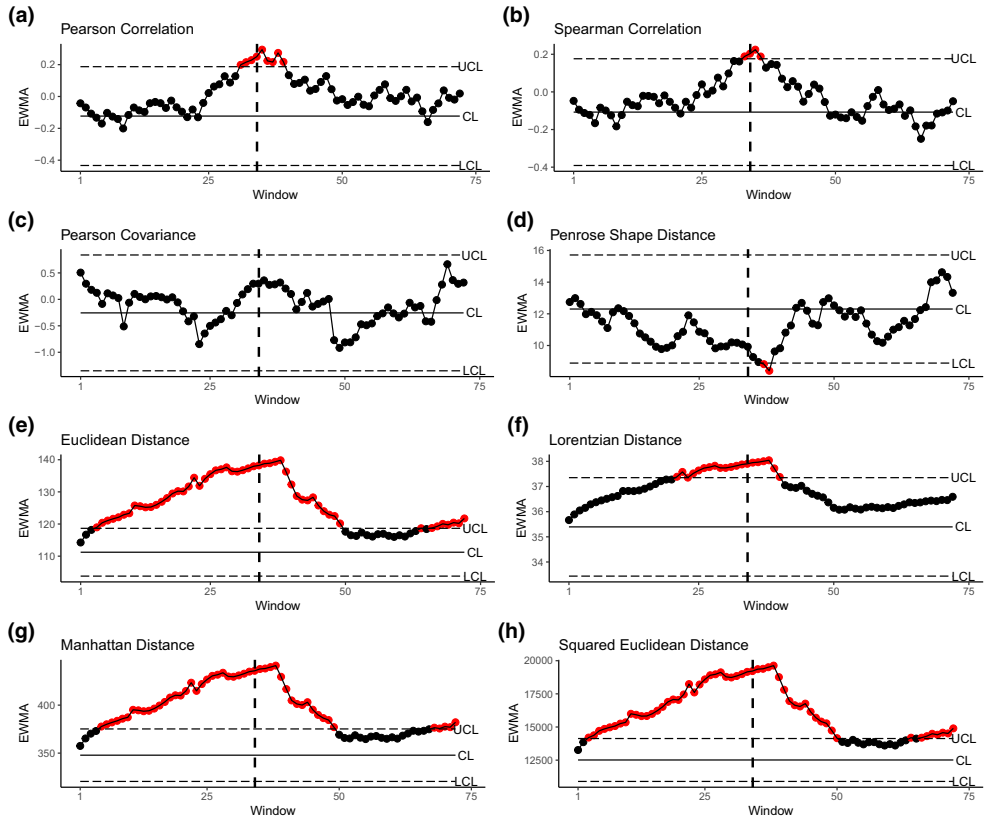


FIGURE 3 EWMA control charts of the association scores between heart rate and $petCO_2$. The dashed vertical line marks the boundary between the multiple task period and recovery period in phase II. The dashed horizontal lines indicate the upper control limit (UCL) and lower control limit (LCL). The solid horizontal line denotes the centre line (CL), which is equal to the estimated phase I mean $\hat{\mu}_1$. The red dots indicate the out-of-control EWMA scores exceeding control limits.

Since EWMA applied to Pearson and Spearman correlations is sensitive to changes in correlation, we expect these procedures to perform well in detecting such changes. In contrast, for the other measures that are sensitive to a broader range of changes beyond the correlation, we expect varying effects on the ARL scores depending on the type of change(s) introduced.

3.1 | Design

3.1.1 | Data characteristics

The phase I scores were drawn from a standard bivariate normal distribution (i.e. $\mu_1 = 0$ and $\sigma_1 = 1$), with no correlation between the two variables ($\rho_1 = 0$). The number of windows[7] in phase I was set to 100. The number of measurement occasions per window in both phases I and II was set to 10. In phase II, the scores were also drawn from a bivariate normal distribution, with varying values for

[7]We investigated the impact of reducing the amount of phase I data by using 100, 50 or 20 windows and either 5 or 10 measurement occasions per window. Overall patterns were similar to the main simulation results (100 windows with 10 measurement occasions per window). While ARL_0 values, reflecting type I error, remained stable, ARL_1 values increased with less phase I data, indicating reduced power to detect changes. The results can be found at https://osf.io/egjfx.

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the mean, standard deviation and correlation depending on the specific setting. We introduced either no change or a change in the mean, standard deviation, correlation or any combination of these statistics. When a change in a statistic was introduced, it was either positive (randomly drawn from a uniform distribution between .10 and .60) or negative (randomly drawn from a uniform distribution between −.10 and −.60). For mean and standard deviation changes, we introduced the change in either one or both variables, consistently across both types of changes. Across all settings, this resulted in 54 possible combinations of changes: 3 options for the mean (negative, zero, positive) × 3 options for the standard deviation (negative, zero, positive) × 3 options for the correlation (negative, zero, positive) x 2 options regarding the number of variables the changes in the mean and standard deviation were applied to. Note that when both the mean and standard deviation were zero, the results are identical regardless of whether the change was applied to one or both variables. The change(s) were introduced at the start of phase II. For each combination of settings, 100,000 replicates were generated. The size of the positive and negative changes varied across replications, as they were randomly sampled.

3.1.2 | EWMA settings and performance measures

We applied the EWMA procedure eight times to each simulated dataset, that is, to the different association scores calculated in windows: Pearson correlation, Spearman correlation, Pearson covariance, Penrose shape distance, Euclidean distance, Lorentzian distance, Manhattan distance and squared Euclidean distance. We set λ to .10 and the control limits such that the ARL_0 equals 370 with $L = 2.70$. We assessed the run length of each analysis—that is, in which phase II window the first out-of-control score occurred. In case there was no out-of-control score within the 10,000 phase II windows, the run length was set to 10,001.

The performance of EWMA applied to the different association measures was then measured in terms of the ARL values. We expect the ARL to vary based on the combination of introduced changes, the association measure (and the type of change(s) it is sensitive to) and the number of variables (one or two) in which the mean change was introduced.

3.2 | Results

3.2.1 | Detection of correlation changes (without mean or variance changes)

Figure 4 shows the *ARL* curves of EWMA applied to the association measures for positive, negative or no change in correlation. There is a clear effect of the correlation change: When no change is introduced, the ARL_0 values for all EWMA procedures are around 370, while the ARL_1 values for both positive and negative correlation changes are lower. The two procedures only sensitive to correlation changes, namely EWMA applied to the Pearson and Spearman correlations, perform best (along with the Pearson covariance), with the Pearson correlation slightly outperforming Spearman correlation (slightly higher ARL_0 and lower ARL_1).

3.2.2 | Influence of mean and variance changes on top of the correlation change

Figure 5 shows the ARL curves of EWMA applied to the association measures for positive, negative or no change in mean, variance and correlation, as well as for changes occurring in one or both variables. Note that when there is no change in the mean or standard deviation, the results are identical for one and two variables, as the correlation change is by definition introduced in both variables. Overall, the ARL curves shows three measure clusters: The Pearson and Spearman

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

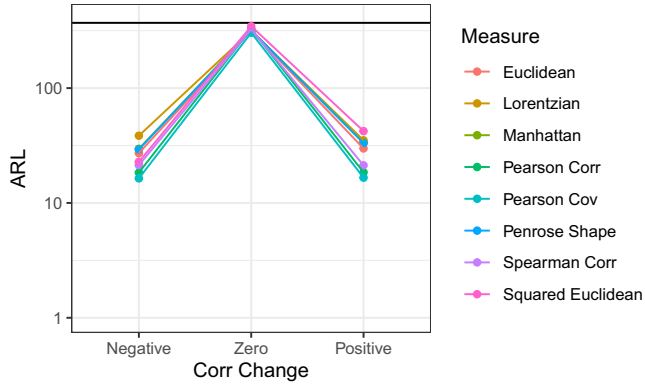


FIGURE 4 ARL curves of EWMA applied to the association measures for detecting correlation changes. The ARL values are displayed on a logarithmic scale, with the horizontal line indicating the set ARL_0 value of 370.

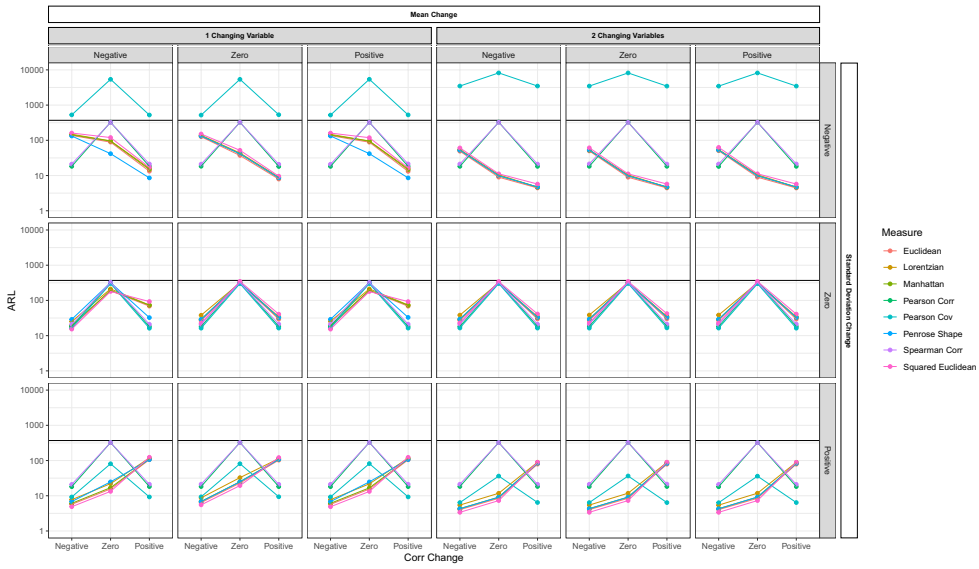


FIGURE 5 ARL curves of EWMA applied to the association measures. The first three columns display the results where mean and variance changes were introduced in one variable, while the last three columns show results for changes introduced in both variables. Within these settings, the columns indicate the presence/direction of the mean change, and the rows indicate the presence/direction of the standard deviation change. The ARL values are displayed on a logarithmic scale, with the horizontal line indicating the set ARL_0 value of 370.

correlations show similar performance, the Pearson covariance exhibits a unique pattern, while the remaining measures (i.e. Euclidean, Lorentzian, Manhattan, squared Euclidean and Penrose shape distances) show comparable results.

Does an additional mean and/or variance change on top of a correlation change improve EWMA performance in terms of detecting (correlation) changes? Since the Pearson and Spearman correlations are insensitive to mean and variance changes, we expected there to be no influence from these changes on the results. Indeed, the ARL curves of EWMA applied to the Pearson and Spearman correlations are consistent across all mean and variance change conditions.

For the other association measures, which are sensitive to more than just correlation changes, the addition of a mean and/or variance change appears to either improve or worsen EWMA performance.

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

For example, the Pearson covariance performs better when there is an increase in the standard deviation (i.e. positive change), as this leads to EWMA scores exceeding the control limits, resulting in shorter ARL values. On the other hand, a decrease in the standard deviation (i.e. negative change) keeps EWMA scores within the limits, leading to higher ARL values. This effect is more pronounced when the standard deviation change occurs in both variables rather than just one.

EWMA applied to the remaining measures (i.e. Euclidean, Lorentzian, Manhattan, squared Euclidean and Penrose shape distances) shows an interaction effect between the direction of the standard deviation change and the correlation change. Specifically, a negative standard deviation change combined with a negative correlation change results in higher ARL values, while a positive change in both results in lower ARL values. In the absence of a standard deviation change, introducing a mean change alongside a positive correlation change leads to higher ARL values (for all measures except the Penrose shape distance, which remains unaffected as it is not sensitive to mean changes). This is not the case for a negative correlation change. When mean and standard deviation changes occur in both variables, the results of the Penrose shape distance align more closely with the other measures. Here, the effect of the mean change is no longer visible, as the level is increased/decreased for both variables, and thus does not lead to larger distance scores.

3.3 | Follow-up and proof-of-concept simulations

We conducted four additional simulation studies to examine the impact of (1) outliers, and (2) underlying autocorrelation and to assess the applicability of the EWMA approach for detecting association changes in (3) ordinal and (4) binary data. Details of the simulation designs and results can be found at https://osf.io/egjfx.

First, we introduced artificial outliers into 5% or 10% of the measurement occasions, either in one or both variables, with outliers present in both phases I and II (thus also affecting control limit estimation). Results show that EWMA applied to the Spearman correlation outperforms EWMA applied to the Pearson correlation, with the advantage becoming more pronounced as the proportion of outliers increased and when both variables contained outliers.

Second, we introduced underlying autocorrelation in both variables across phases I and II, with the same autocorrelation strength of .1, .3 or .5 being applied to each variable. The overall pattern of results was similar to that of the main simulation. Methods that are only sensitive to changes in association—EWMA applied to the Pearson and Spearman correlations—performed best across autocorrelation levels, along with the Pearson covariance. Overall, higher underlying autocorrelation led to worse performance, with lower ARL_0 and higher ARL_1 values.

Third, to demonstrate the feasibility of this approach beyond continuous data, we conducted a proof-of-concept simulation using ordinal data (symmetrically distributed on a 1–7 Likert scale, with probabilities .05, .10, .20, .30, .20, .10, and .05). The overall pattern of results resembled that of the main simulation with bivariate normal data: When only a correlation change was introduced, the procedures solely sensitive to correlation changes—EWMA applied to the Pearson and Spearman correlations—perform best (along with the Pearson covariance), with the Pearson correlation showing slightly better performance (higher ARL_0 and lower ARL_1) than the Spearman correlation.

Finally, we conducted a proof-of-concept simulation using binary data, in which Cohen's Kappa was included as an additional association measure alongside those used in the main study. Overall, the results show high ARL_0 values and low ARL_1 values, with only minor differences between the different association measures. EWMA applied to the Euclidean distance achieved the lowest ARL_1 for positive correlation changes, while EMWA applied to Lorentzian, Manhattan and squared Euclidean distances had the lowest ARL_1 values for negative correlation changes, performing equally overall. EWMA applied to the Pearson and Spearman correlations also demonstrated identical performance.

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4 | DISCUSSION

Real-time detection of within-person changes in linear associations is important, as such changes may indicate worsening well-being or maladaptive functioning. Although several correlation-based monitoring procedures have been proposed in the literature (e.g. MEWMS, MEWMV, MEWMC; Hawkins & Maboudou-Tchao, 2008; Huwang et al., 2007; Reynolds Jr & Cho, 2006), these procedures are sensitive not only to correlation changes but also to changes in the variance and, in some cases, the mean. In the current study, we proposed a simple yet flexible approach that allows for real-time monitoring of a wide range of association measures using EWMA. This approach addresses an existing gap by allowing researchers to focus monitoring on changes in correlations, independent of mean or variance changes. To demonstrate its feasibility, we investigated how well it works to first calculate association measure scores (i.e. Pearson and Spearman correlations, Pearson covariance, Penrose shape, Euclidean, Lorentzian, Manhattan and squared Euclidean distances) within time windows and monitor these scores for mean-level changes using EWMA. Additionally, we studied how changes in the mean and variance (with or without a correlation change) influence the performance of EWMA applied to association scores.

We conducted a simulation study in which we manipulated the correlation, standard deviation and mean change (i.e. positive, negative, or no change). The simulation results indicate that EWMA applied to the Pearson and Spearman correlations perform best when only a correlation change is introduced. For the Pearson covariance, an additional increase in standard deviation improves performance whereas a decrease in standard deviation worsened performance. For the remaining association measures (i.e. Penrose shape, Euclidean, Lorentzian, Manhattan and squared Euclidean distances), more intricate patterns emerged.

Thus, we recommend researchers to carefully consider both the types of changes they expect to observe and the direction of those changes. If only correlation changes are expected, we recommend using an association measure that is only sensitive to correlation changes (i.e. the Pearson or Spearman correlations). Not only were these measures good at detecting correlation changes, one does not risk missing relevant changes due to the masking effects of other changes (which lead to higher ARL values). However, if other types of changes are expected and their direction is known, researchers can opt for an association measure that benefits from that specific combination of changes.

4.1 | Directions for future research

The simulation results showed that EWMA applied to the Pearson correlation slightly outperforms EWMA applied to the Spearman correlation. Carlier et al. (2024) recommended to be cautious when using the Pearson correlation, as it is sensitive to certain data artefacts and does not always yield consistent results across different datasets. For more robustness, the Spearman correlation is recommended. This is illustrated by our follow-up simulation in which outliers were introduced. Under these conditions, EWMA applied to the Spearman correlation outperformed EWMA applied to the Pearson correlation, with the advantage becoming more pronounced as the proportion of outliers increased and when both variables contained outliers. Future research could further investigate how EWMA applied to association measures performs with more complex or 'messy' data—for example, datasets with missing values or nonlinear association changes. Additionally, strategies for handling high underlying autocorrelation warrant further study. When autocorrelation is not too high (e.g. .10 or .30), accounting for it may not be necessary, but for higher autocorrelation (e.g. >.30), additional preprocessing steps may be required to ensure sufficient performance. An option that has been proposed in the EWMA literature is to fit a time series model to the phase I data and to apply the EWMA procedure to the residuals for both phase I and phase II data based on this model (Lu & Reynolds Jr, 1999; Schat et al., 2023).

In this study, we focused on a subset of association measures that are frequently used and can be calculated with a small number of data points. The selected measures reflect a pragmatic choice aligned with typical data constraints in psychological applications, rather than an exhaustive overview of all

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

possible measures. Nonetheless, other association measures remain relevant. For instance, autocorrelation is a specific form of correlation which is of prime interest in psychopathology research (Cabrieto et al., 2019; Wichers et al., 2016) and can be monitored in a similar way (i.e. calculating autocorrelation within time windows and monitoring these using EWMA). However, future research is needed, because we expect additional challenges: issues with missing data (Little & Rubin, 2019; Somer et al., 2022), insufficient data leading to biased autocorrelation estimates (Huitema & McKean, 1991; Marriott & Pope, 1954). With larger amounts of data and consequently more observations within each time window, it becomes feasible to use more complex association measures, such as dynamic time warping (Berndt & Clifford, 1994). Passive sensing through smartphones and wearable devices indeed allow to collect sufficient intensive longitudinal data for this purpose (Harari et al., 2016; Insel, 2018; Torous et al., 2021). In principle, such complex measures can also be applied when fewer measurement occasions are available overall; however, this will result in fewer association scores in phase I, which would reduce the accuracy of the control limits.

The EWMA procedure assumes that the monitored scores are normally distributed. However, this assumption will not always hold in our approach. Even if the raw data themselves are normally distributed, the resulting association scores are often distributed differently (see supplementary material for example distributions; https://osf.io/egjfx). For example, the Squared Euclidean scores were right skewed. It may thus be worthwhile to investigate whether performance can be improved by using real-time monitoring procedures that do not rely on the normality assumption inherent to EWMA. In particular, future research could build on the growing literature on non-parametric monitoring procedures (Liang et al., 2022; Mukherjee & Chakraborti, 2012; Qiu, 2018; Zou & Tsung, 2010). Although promising, these methods are more difficult to implement in practice, as they are not yet widely available in R and require additional tuning of the method to the problem at hand. For example, non-parametric procedures based on data categorization and categorical data analysis require users to decide on the number of categories to use (Qiu & Li, 2011). Clear guidance on how to make such choices for intensive longitudinal data in psychology is therefore needed. Furthermore, it remains to be investigated whether non-parametric procedures truly outperform the EWMA procedure at detecting changes in association scores.

5 | CONCLUSION

Applying the EWMA procedure to association scores is promising for detecting linear association changes in psychological research. Monitoring Pearson and Spearman correlation scores is advised when no additional information is available about the presence of additional mean and/or variance changes. Using association measures that are sensitive to more types of changes apart from the correlation (i.e. mean and/or variance) can improve detection performance given specific combinations of mean, variance and correlation changes. Using other measures can thus be valuable when the presence of such a combination of changes can be predicted before starting the monitoring.

AUTHOR CONTRIBUTIONS

Evelien Schat: Conceptualization; methodology; formal analysis; visualization; writing – original draft; writing – review and editing; funding acquisition; software. **Sarah Schrevens:** Conceptualization; methodology; writing – review and editing. **Francis Tuerlinckx:** Funding acquisition; supervision; conceptualization; methodology; writing – review and editing. **Eva Ceulemans:** Funding acquisition; supervision; writing – review and editing; conceptualization; methodology.

FUNDING INFORMATION

The current research is funded by grant C14/23/062 from the Research Council of KU Leuven, grant EOS 40007528/G0I2422N from the Research Foundation Flanders (FWO) and F.R.S.-FNRS under the Excellence of Science (EOS) program, grant PDMT2/23/019 from the Research Council of KU Leuven and grant 12AA826N from the Research Foundation Flanders (FWO). The computational

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation—Flanders (FWO) and the Flemish Government.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare regarding the content of this article.

DISCLOSURE OF ARTIFICIAL INTELLIGENCE-GENERATED CONTENT (AIGC) TOOLS

No generative AI tools were used in the creation of this article.

DATA AVAILABILITY STATEMENT

R code and supplementary material can be found at https://osf.io/egjfx. The data used as illustrative example are publicly available through the `kcpRS` R package (Cabrieto et al., 2022). The data that support the findings of this study are openly available in the R package kcpRS at https://doi.org/10.32614/CRAN.package.kcpRS.

ORCID

Evelien Schat https://orcid.org/0000-0003-1169-3984
Francis Tuerlinckx https://orcid.org/0000-0002-1775-7654

REFERENCES

Albers, C. J., & Bringmann, L. F. (2020). Inspecting gradual and abrupt changes in emotion dynamics with the time-varying change point autoregressive model. *European Journal of Psychological Assessment*, *36*(3), 492–262499. https://doi.org/10.1027/a000001

Allik, J., Borkenau, P., Hrebícková, M., Kuppens, P., & Realo, A. (2015). How are personality trait and profile agreement related? *Frontiers in Psychology*, *6*, 785. https://doi.org/10.3389/fpsyg.2015.00785

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (AAAIWS'94)* (pp. 359–370). www.aaai.org

Cabrieto, J., Adolf, J. K., Tuerlinckx, F., Kuppens, P., & Ceulemans, E. (2019). An objective, comprehensive and flexible statistical framework for detecting early warning signs of mental health problems. *Psychotherapy and Psychosomatics*, *88*(3), 184–186. https://doi.org/10.1159/000494356

Cabrieto, J., Meers, K., Adolf, J. K., Schat, E., Kuppens, P., Tuerlinckx, F., & Ceulemans, E. (2022). kcpRS: An R package for performing kernel change point detection on the running statistics of multivariate time series. *Behavior Research Methods*, *54*(3), 1092–1113. https://doi.org/10.3758/s13428-021-01603-8

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Wilhelm, F. H., Liedlgruber, M., & Ceulemans, E. (2018). Capturing correlation changes by applying kernel change point detection on the running correlations. *Information Sciences*, *447*, 117–139. https://doi.org/10.1016/j.ins.2018.03.010

Carlier, C., Karch, J. D., Kuppens, P., & Ceulemans, E. (2024). A comparison of measures for assessing profile similarity in dyads. *Psychologica Belgica*, *64*(1), 72–84. https://doi.org/10.5334/pb.1297

Chow, S. M., Grimm, K. J., Filteau, G., Dolan, C. V., & McArdle, J. J. (2013). Regime-switching bivariate dual change score model. *Multivariate Behavioral Research*, *48*(4), 463–502. https://doi.org/10.1080/00273171.2013.787870

Chow, S. M., & Zhang, G. (2013). Nonlinear regime-switching state-space (RSSS) models. *Psychometrika*, *78*(4), 740–768. https://doi.org/10.1007/s11336-013-9330-8

Cramer, A. O. J., Van Borkulo, C. D., Giltay, E. J., Van Der Maas, H. L. J., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PLoS One*, *11*(12), 1–20. https://doi.org/10.1371/journal.pone.0167490

Crayen, C., Eid, M., Lischetzke, T., & Vermunt, J. K. (2017). A continuous-time mixture latent- state-trait Markov model for experience sampling data: Application and evaluation. *European Journal of Psychological Assessment*, *33*(4), 296–311. https://doi.org/10.1027/1015-5759/a000418

Dejonckheere, E., Mestdagh, M., Houben, M., Erbas, Y., Pe, M., Koval, P., Brose, A., Bastian, B., & Kuppens, P. (2018). The bipolarity of affect and depressive symptoms. *Journal of Personality and Social Psychology*, *114*(2), 323–341. https://doi.org/10.1037/pspp0000186

Dejonckheere, E., Mestdagh, M., Verdonck, S., Lafit, G., Ceulemans, E., Brock, B., & Kalokerinos, E. K. (2021). The relation between positive and negative affect becomes more negative in response to personally relevant events. *Emotion*, *21*(2), 326–336. https://doi.org/10.1037/emo0000697

Gable, S. L., & Reis, H. T. (1999). Now and then, them and us, this and that: Studying relationships across time, partner, context, and person. *Personal Relationships*, *6*(4), 415–432. https://doi.org/10.1111/j.1475-6811.1999.tb00201.x

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Grassmann, M., Vlemincx, E., von Leupoldt, A., & Van den Bergh, O. (2016). The role of respiratory measures to assess mental load in pilot selection. *Ergonomics*, *59*(6), 745–753. https://doi.org/10.1080/00140139.2015.1090019

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science*, *26*(1), 10–15. https://doi.org/10.1177/0963721416666518

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science*, *11*(6), 838–854. https://doi.org/10.1177/1745691616650285

Hawkins, D. M., & Maboudou-Tchao, E. M. (2008). Multivariate exponentially weighted moving covariance matrix. *Technometrics*, *50*(2), 155–166. https://doi.org/10.1198/004017008000000163

Hawkins, D. M., & Wu, Q. (2014). The CUSUM and the EWMA head-to-head. *Quality Engineering*, *26*(2), 215–222. https://doi.org/10.1080/08982112.2013.817014

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, *141*(4), 901–930. https://doi.org/10.1037/a0038822

Huitema, B. E., & McKean, J. W. (1991). Autocorrelation estimation and inference with small samples. *Psychological Bulletin*, *110*(2), 291–304. https://doi.org/10.1037/0033-2909.110.2.291

Huwang, L., Yeh, A. B., & Wu, C. W. (2007). Monitoring multivariate process variability for individual observations. *Journal of Quality Technology*, *39*(3), 258–278. https://doi.org/10.1080/00224065.2007.11917692

Insel, T. R. (2018). Digital phenotyping: A global tool for psychiatry. *World Psychiatry*, *17*(3), 276–277. https://doi.org/10.1002/wps.20550

Jennings, J. R., Kamarck, T., Stewart, C., Eddy, M., & Johnson, P. (1992). Alternate cardiovascular baseline assessment techniques: Vanilla or resting baseline. *Psychophysiology*, *29*(6), 742–750. https://doi.org/10.1111/j.1469-8986.1992.tb02052.x

Knoth, S. (2020). *spc: Statistical process control – Calculation of ARL and other control chart performance measures* (0.6.4). https://cran.r-project.org/package=spc

Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial for performing power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–24. https://doi.org/10.1177/2515245920978738

Liang, W., Mukherjee, A., Xiang, D., & Xu, Z. (2022). A new nonparametric adaptive EWMA procedures for monitoring location and scale shifts via weighted Cucconi statistic. *Computers and Industrial Engineering*, *170*, 108321. https://doi.org/10.1016/j.cie.2022.108321

Little, R., & Rubin, D. (2019). *Statistical analysis with missing data* (3rd ed.). Wiley. https://doi.org/10.1002/9781119482260

Lu, C. W., & Reynolds, M. R., Jr. (1999). EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology*, *31*(2), 166–188. https://doi.org/10.1080/00224065.1999.11979913

Marriott, F. H. C., & Pope, J. A. (1954). Bias in the estimation of autocorrelations. *Biometrika*, *41*(3), 390–402. https://doi.org/10.2307/2332719

Montgomery, D. C. (2009). *Introduction to statistical quality control* (8th ed.). John Wiley & Sons, Inc.

Mukherjee, A., & Chakraborti, S. (2012). A distribution-free control chart for the joint monitoring of location and scale. *Quality and Reliability Engineering International*, *28*(3), 335–352. https://doi.org/10.1002/qre.1249

Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, *41*(1/2), 100–115. https://doi.org/10.2307/2333009

Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., Mata, J., Jaeggi, S. M., Buschkuehl, M., Jonides, J., Kuppens, P., & Gotlib, I. H. (2015). Emotion-network density in major depressive disorder. *Clinical Psychological Science*, *3*(2), 292–300. https://doi.org/10.1177/2167702614540645

Qiu, P. (2018). Some perspectives on nonparametric statistical process control. *Journal of Quality Technology*, *50*(1), 49–65. https://doi.org/10.1080/00224065.2018.1404315

Qiu, P., & Li, Z. (2011). On nonparametric statistical process control of univariate processes. *Technometrics*, *53*(4), 390–405. https://doi.org/10.1198/TECH.2011.10005

Reynolds, M. R., Jr., & Cho, G. Y. (2006). Multivariate control charts for monitoring the mean vector and covariance matrix. *Journal of Quality Technology*, *38*(3), 230–253. https://doi.org/10.1080/00224065.2006.11918612

Roberts, S. W. (1959). Control chart tests based on geometric moving averages. *Technometrics*, *1*(3), 239–250. https://doi.org/10.2307/1266443

Schat, E., Tuerlinckx, F., De Ketelaere, B., & Ceulemans, E. (2024). Real-time detection of mean and variance changes in experience sampling data: A comparison of existing and novel statistical process control approaches. *Behavior Research Methods*, *56*, 1459–1475. https://doi.org/10.3758/s13428-023-02103-7

Schat, E., Tuerlinckx, F., Smit, A. C., de Ketelaere, B., & Ceulemans, E. (2023). Detecting mean changes in experience sampling data in real time: A comparison of univariate and multivariate statistical process control methods. *Psychological Methods*, *28*(6), 1335–1357. https://doi.org/10.1037/met0000447

Scheffer, M., Elizabeth Bolhuis, J., Borsboom, D., Buchman, T. G., Gijzel, S. M. W., Goulson, D., Kammenga, J. E., Kemp, B., van de Leemput, I. A., Levin, S., Martin, C. M., Melis, R. J. F., van Nes, E. H., Michael Romero, L., & Olde Rikkert, M. G. M. (2018). Quantifying resilience of humans and other animals. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(47), 11883–11890. https://doi.org/10.1073/pnas.1810630115

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Schreuder, M. J., Schat, E., Smit, A. C., Snippe, E., & Ceulemans, E. (2024). Monitoring emotional intensity and variability in real-time to forecast depression recurrence in remitted adults. *Journal of Consulting and Clinical Psychology*, *92*(8), 505–516. https://doi.org/10.1037/ccp0000871

Sels, L., Cabrieto, J., Butler, E., Reis, H., Ceulemans, E., & Kuppens, P. (2020). The occurrence and correlates of emotional interdependence in romantic relationships. *Journal of Personality and Social Psychology*, *119*(1), 136–158. https://doi.org/10.1037/pspi0000212

Sels, L., Schat, E., Verhofstadt, L., & Ceulemans, E. (2022). Introducing change point detection analysis in relationship research: An investigation of couples' emotion dynamics. *Journal of Social and Personal Relationships*, *39*(10), 3133–3154. https://doi.org/10.1177/02654075211070558

Shewhart, W. A. (1931). *Economic control of quality of manufactured product*. Macmillan and Co Ltd.

Shin, K. E., Newman, M. G., & Jacobson, N. C. (2022). Emotion network density is a potential clinical marker for anxiety and depression: Comparison of ecological momentary assessment and daily diary. *British Journal of Clinical Psychology*, *61*, 31–50. https://doi.org/10.1111/bjc.12295

Smit, A. C., Schat, E., & Ceulemans, E. (2023). The exponentially weighted moving average procedure for detecting changes in intensive longitudinal data in psychological research in real-time: A tutorial showcasing potential applications. *Assessment*, *30*(5), 1354–1368. https://doi.org/10.1177/10731911221086985

Smit, A. C., & Snippe, E. (2022). Real-time monitoring of increases in restlessness to assess idiographic risk of recurrence of depressive symptoms. *Psychological Medicine*, *53*, 5060–5069. https://doi.org/10.1017/S0033291722002069

Smit, A. C., Snippe, E., & Wichers, M. (2019). Increasing restlessness signals impending increase in depressive symptoms more than 2 months before it happens in individual patients. *Psychotherapy and Psychosomatics*, *88*(4), 249–251. https://doi.org/10.1159/000500594

Snippe, E., Smit, A. C., Kuppens, P., Burger, H., & Ceulemans, E. (2023). Recurrence of depression can be foreseen by monitoring mental states with statistical process control. *Journal of Psychopathology and Clinical Science*, *132*(2), 145–155. https://doi.org/10.1037/abn0000812

Somer, E., Gische, C., & Miočević, M. (2022). Methods for modeling autocorrelation and handling missing data in mediation analysis in single case experimental designs (SCEDs). *Evaluation & the Health Professions*, *45*(1), 36–53. https://doi.org/10.1177/01632787211071136

Spruijt-Metz, D., & Nilsen, W. (2014). Dynamic models of behavior for just-in-time adaptive interventions. *IEEE Pervasive Computing*, *13*(3), 13–17. https://doi.org/10.1109/MPRV.2014.46

Stifter, C. A., & Rovine, M. (2015). Modeling dyadic processes using hidden markov models: A time series approach to mother-infant interactions during infant immunization. *Infant and Child Development*, *24*(3), 298–321. https://doi.org/10.1002/icd.1907

Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurhold-Jepsen, M., Whelan, P., Carvalho, A. F., Keshavan, M., Linardon, J., & Firth, J. (2021). The growing field of digital psychiatry: Current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry*, *20*(3), 318–335. https://doi.org/10.1002/wps.20883

Van De Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., Van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen, S. H., Derom, C., Jacobs, N., Kendler, K. S., Van Der Maas, H. L. J., Neale, M. C., Peeters, F., Thiery, E., Zachar, P., & Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(1), 87–92. https://doi.org/10.1073/pnas.1312114110

Wichers, M., Groot, P. C., & Psychosystems, ESM Group, & EWS Group. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, *85*(2), 114–116. https://doi.org/10.1159/000441458

Zou, C., & Tsung, F. (2010). Likelihood ratio-based distribution-free EWMA control charts. *Journal of Quality Technology*, *42*(2), 174–196. https://doi.org/10.1080/00224065.2010.11917815

How to cite this article: Schat, E., Schrevens, S., Tuerlinckx, F., & Ceulemans, E. (2026). Detecting association changes in intensive longitudinal data in real time: An exponentially weighted moving average procedure. *British Journal of Mathematical and Statistical Psychology*, *79*, 362–378. https://doi.org/10.1111/bmsp.70035

20448317, 2026, 2, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70035 by Ku Leuven, Wiley Online Library on [02/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License