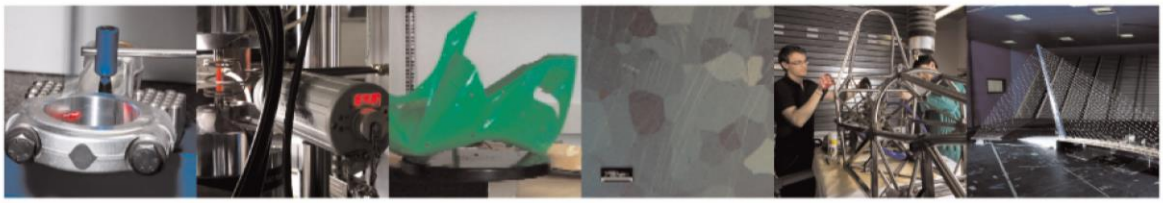

Using Curve-Registration Information for Profile Monitoring

Grasso, M.; Menafoglio, A.; Colosimo, B.M.; Secchi, P.

This is an Accepted Manuscript of an article published by Taylor & Francis in JOURNAL OF QUALITY TECHNOLOGY on 21 Nov 2017, available online: https://doi.org/10.1080/00224065.2016.11918154.

Using Curve Registration Information for Profile Monitoring

MARCO GRASSO, ALESSANDRA MENAFOGLIO,
BIANCA M. COLOSIMO and PIERCESARE SECCHI[1]
Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

The quality characteristics in manufacturing processes are often represented in terms of spatially or time ordered data, called "profiles", which are characterized by amplitude and phase variability. In this context, curve registration plays a key role, as it allows separating the two kinds of between-profiles variability, and to reduce any undesired inflation of the natural phase variability. In the mainstream literature, registration warping functions are not generally considered in the monitoring process, even though this may cause a significant information loss. We propose a novel approach for profile monitoring, which combines the Functional Principal Component Analysis and the use of parametric warping functions. The key idea is to jointly monitor the stability over time of the registered profiles (i.e., the information related to amplitude variability) and the registration coefficients (i.e., the information related to phase variability). This allows improving the capability of detecting unnatural pattern modifications, thanks to a better characterization of the overall natural variability. The benefits of a proper management of functional data registration, together with the advantages over the most common approaches used in the literature, are demonstrated by means of Monte Carlo simulations. The proposed methodology is finally applied to a real industrial case study relying on a dataset acquired in waterjet cutting processes under different health conditions of the machine tool.

Keywords: Warping Functions, Curve Registration, Functional Data Analysis, Profile Monitoring

1. Introduction

In several manufacturing applications of practical interest, the signals acquired during the process and/or the geometric features measured on the product may be represented by spatially or time ordered data known as "profiles". In the presence of quality characteristics consisting of cyclically repeating patterns, the term "profile monitoring" refers to a suite of methods that provides the natural framework to evaluate the stability over time of process quality (Woodall *et al*., 2004; Kang and Albin, 2000). Different techniques have been proposed to deal with complex shape profiles acquired from sensor signals. Some contributions in this field include the studies of Jin and Shi (1999; 2001), Zhou *et al*. (2005), Ding *et al*. (2006), Chang and Yadama (2010) and Grasso *et al*. (2014a; 2014b). An overview of parametric and nonparametric

Dr. Marco Grasso (corresponding author) is a Post Doc in the Department of Mechanical Engineering. His email address is marcoluigi.grasso@polimi.it.
Dr. Alessandra Menafoglio is a Post Doc in the Department of Mathematics. Her email is alessandra.menafoglio@polimi.it.
Prof. Bianca Maria Colosimo is a full professor in the Department of Mechanical Engineering. Her email is biancamaria.colosimo@polimi.it.
Prof. Piercesare Secchi is a full professor in the Department of Mathematics. His email is piercesare.secchi@polimi.it.

approaches for profile monitoring can be found in Noorossana *et al*. (2012).

The recorded profiles generally present two kinds of natural variability: a *phase variability* and an *amplitude variability* (Ramsay and Li, 1998; Vantini, 2012). The natural phase variability may be inflated by undesired misalignments caused by random durations of cyclical processes or by other disturbance factors. An inflated phase variability results in an inflation of the overall natural variability, which may mask the effects of unnatural deviations associated to assignable causes. In addition, a curve misalignment makes the salient features of different profiles to be shifted in phase, which reduces the capability of detecting local pattern modifications.

As far as the curve registration problem is concerned, the mainstream literature devoted to profile monitoring (see Section 2) does not deal with registration, as it is deemed an unnecessary task or simply treated as a pre-processing step to be neglected when profile monitoring is applied. However, registration is a very important and complex task in many application domains, and it plays a relevant role when monitoring is applied either to geometric profiles and surfaces (Colosimo and Pacella, 2007, Dryden and Mardia, 1998) or signal profiles. As a matter of fact, simple re-sampling or triggering methods, which are sometimes used in practice for signal misalignment reduction, do not guarantee a proper separation of amplitude and phase variability. Second, monitoring only the registered curves, without keeping any explicit track of the registration parameters, implies a loss of information about the phase variability, which may reduce the control chart performances. The present study is a first contribution in the direction of integrating curve registration algorithms into the profile monitoring framework. We propose a novel approach to jointly monitor the stability over time of both the registered profiles and the warping functions used to align them. To this end, we propose a three-step procedure consisting of: (i) applying a parametric warping registration (Eilers, 2004; Ramsay and Silverman, 2005) of the acquired profiles to decouple the phase and the amplitude variability, (ii) performing a dimensionality reduction of the registered profiles through Functional Principal Component Analysis (FPCA) (Ramsay and Silverman, 2005), and (iii) monitoring both the phase and the amplitude variability of the signals through appropriate control charts. Specifically, the coefficients of the warping functions singled out in Step (i) are used to describe the natural phase variability of profile data. A FPCA of registered curves allows reducing the number of features representing the amplitude variability. Step (iii) finally controls the signals through both the phase and the amplitude variability (in terms of warping coefficients and principal components scores), together with the FPCA residuals, which are used to detect possible deviations involving principal directions orthogonal to those of the retained components. This approach avoids information losses, besides allowing a proper management of both kinds of shape variability. A rationale for the use of this approach, together with a brief state of the art on curve registration in profile monitoring applications, is discussed in Section 2.

We demonstrate the benefits of the proposed approach through Monte Carlo simulations and by means of a real industrial test case in waterjet cutting (Kovacevic *et al*., 1997). Two competitor implementations of a FPCA-based profile monitoring approach are considered: (i) FPCA on raw data without registering the curves in advance, and (ii) FPCA on the registered curves, monitoring the FPCA scores only. The former approach is representative of what is done in practice when only a rough alignment (or no alignment at all) is applied before analysing the data, whereas the latter approach is representative of what is usually proposed in the mainstream literature.

The paper is organized as follows: Section 2 describes the curve registration issue in profile monitoring applications and presents two motivating examples; Section 3 reviews the time warping approach for functional data registration; Section 4 describes the proposed approach; Section 5 discusses the simulation analysis; Section 6 presents the real case study in waterjet cutting processes; Section 7 discusses the possible shape distortion effects caused by the time warping operation and some possible ways to deal with them; Section 8 concludes the paper.

2. The Curve Registration Issue and Motivating Examples

Despite of a wide literature devoted to the registration of 2D/3D point clouds in coordinate metrology for shape analysis (Chen and Hung, 1999; Okello and Ristic, 2003; Guo *et al*., 2011; Senin *et al*., 2013; Del Castillo and Colosimo, 2011), only few authors emphasize the critical role of registration for profile monitoring of signal data (Woodall *et al*. 2004; Mosesova *et al*., 2007, Colosimo and Pacella, 2007). As a matter of fact, most of the proposed methods

assume that profiles to be monitored are already registered or do not need registration at all. Unfortunately, this hypothesis is often unrealistic when signal profiles have to be monitored, because manufacturing processes and industrial machinery never exhibit an exact cyclical behaviour, the result being that the salient features of acquired profiles are seldom naturally aligned.

Signal triggering and/or synchronous resampling procedures are sometimes used to reduce the misalignment effects (e.g., Zhou *et al.*, 2005, Jin and Shi, 2001). These methods are also often exploited in commercial monitoring toolkits used in industry. However, they may be insufficient to achieve reliable registration performances, as illustrated by the following examples. Two kinds of signal profiles acquired in two different machining operations are shown in Fig. 1. Fig. 1 a) shows three torque profiles acquired under natural process conditions during an M8 threads tapping operation on a mild steel part. Fig. 1 b) shows three profiles representing the natural water pressure ripples in a waterjet cutting process, acquired during three consecutive pumping cycles.

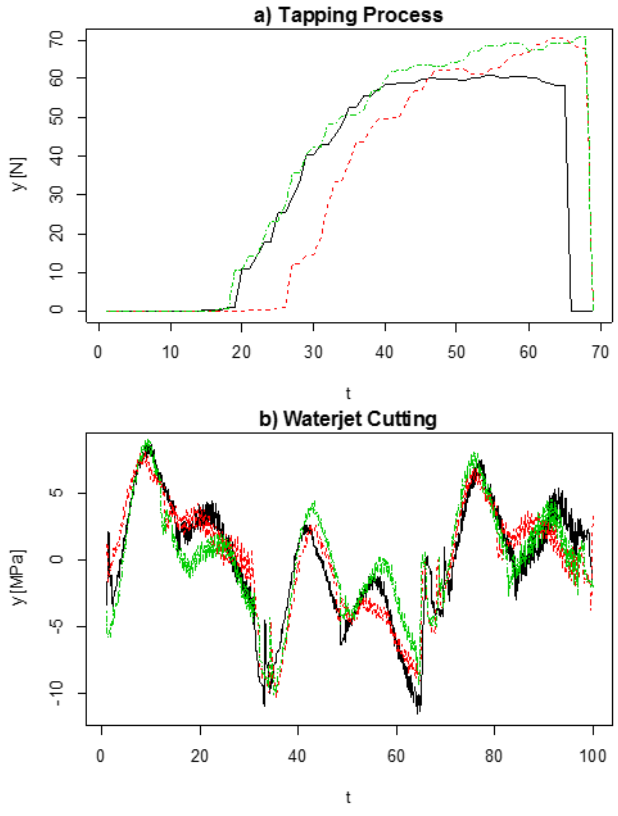

FIGURE 1. Examples of misaligned signal profiles acquired during a) a tapping process and b) a waterjet cutting process

In both cases, the profile acquisition was automatically started and stopped by means of embedded trigger commands. Concerning the waterjet case, a synchronous re-sampling procedure (Gao, 2012) was also applied to get profiles of equal length. The examples in Fig. 1 show that automated triggering and synchronous re-sampling might not be sufficient to guarantee a proper alignment of repeating patterns. Re-sampling procedures may enhance the analysis, but in most cases they cannot avoid undesired time-shifts of the salient shape features.

Other approaches used in the profile monitoring literature include shift registration (Colosimo and Pacella, 2007) and landmark registration (Mosesova *et al.*, 2007), possibly coupled with a trigger-based segmentation and re-sampling operations. Shift registration allows aligning periodic profiles when the misalignment error consists only of a pure translation along the time axis. Thus, it can be useful only in particular cases. Landmark registration, instead, allows dealing with more complex misalignment conditions, by accurately matching pre-selected features. However, the automated identification of the salient landmarks may be a troublesome task, which makes this approach more suitable for post-process monitoring applications.

Regardless of the method used for reducing the misalignment effects, the registration operation is usually carried out during the signal pre-processing steps: after registration, only the registered profiles are considered for monitoring purposes. As a consequence, the information about the phase variability is partially lost. Furthermore, the registration procedure implies that each newly observed profile is aligned to an in-control reference pattern before being classified as either an in-control or an out-of-control profile. Thus, the registration operation viewed as a pre-processing step may even mitigate the effects of actual out-of-control behaviours, as it forces the new pattern to resemble the reference one. This motivates the construction of an effective method to keep under statistical control not only the amplitude variability but also the phase variability of the signals during the process monitoring.

3. Functional data registration via warping functions

Different curve registration methods have been proposed in the literature. One research stream in this

field consists of aligning the curves by maximizing a similarity index: this approach is followed by Ramsay and Li (1998), Ramsay and Silverman (2005), James (2007), Kaziska and Srivastava (2007) and Sangalli *et al*. (2009a). Apart from the simplest case of shift registration, two popular methods are based on landmark registration (Gasser *et al*., 1990; Kneip and Gasser, 1992) and time warping (Eilers, 2004; Tang and Muller, 2008; Tucker *et al*., 2013; Zhou *et al*., 2014). Other approaches have been studied by different authors (Altman and Villareal, 2004; Lindstrom and Bates, 1990; Ke and Wang, 2001; Gervini, 2014).

The use of a continuous monotone registration allows overcoming the limitations of landmark registration and involves a transformation of the time axis, $h: \mathbb{R} \to \mathbb{R}$. Either parametric (Silverman, 1995), semi-parametric (Gervini and Gasser, 2004) or non-parametric (Ramsay and Li, 1998) methods have been proposed to define the transformation h.

In this study, we propose the use of parametric warping functions (i) to increase the computational efficiency, in order to comply with industrial implementation requirements, and (ii) to control the number of warping coefficients to be monitored. We remark that a quadratic or cubic model is usually sufficient for the transformation function, h, to lead to a good registration in practical applications.

The formalization of the time warping problem is briefly reviewed hereafter. Let $\{y_j(t_i), i = 1, \dots, n;\ j = 1, \dots, M\}$ be the sample of raw profiles acquired in the measurement process. Following the FDA approach, each raw datum is here assumed to be a collection of points sampled from an underlying curve $y_j(t), t \in [0, T]$. Hereafter, we assume each profile y_j, $j = 1, \dots, M$, to be a square-integrable function. This assumption is not restrictive for practical applications, since any bounded function on a bounded domain fulfils this requirement.

To set the notation, we consider the following general model for the observations:

$$y_j(t_i) = g_j\left(\tilde{h}_j^{-1}(t_i)\right) + \varepsilon_{ij}, \qquad i = 1, \dots, n;\ j = 1, \dots, M \tag{1}$$

where, $\tilde{h}_j: [0, T] \to [0, T]$, $j = 1, \dots, M$, are invertible functions which capture the phase variability; $g_j, j = 1, \dots, M$, are squared integrable functions, featured by amplitude variability only; ε_{ij}, $i = 1, \dots, n$, $j = 1, \dots, M$, are i.i.d. zero-mean random errors with finite variance σ^2. The latter is generally assumed to be small, since FDA methods usually rely upon the assumption of a high signal-to-noise ratio. We note that the assumption of variance homogeneity across profiles is justified whenever the instrument used to acquire the profiles is the same for each observation.

As a first step of the statistical analysis one needs to represent the acquired raw profiles in a functional form. In the presence of a high signal-to-noise ratio, either a smoothing or an interpolating basis can be used. In the following, we represent the functional profiles – denoted with y_j for the sake of notation simplicity – via a cubic B-spline basis as:

$$y_j(t) = \sum_{q=1}^{Q+L-1} c_{q,j}\Phi_q(t, \boldsymbol{\tau}), \qquad t \in [0, T];\ j = 1, 2, \dots, M \tag{2}$$

where $Q = 4$ is the order of the B-spline functions, L is the number of subintervals separated by $L - 1$ interior knots, $\boldsymbol{\tau}$ is the knot sequence $\boldsymbol{\tau} = \{\tau_l, l = 1, 2, \dots, L\}$, $c_{q,j}$ are the B-spline coefficients, and Φ_q are the B-spline basis functions. Hereafter, we will use the matrix notation $\boldsymbol{y}(\cdot) = \boldsymbol{C}\boldsymbol{\Phi}(\cdot)$, where $\boldsymbol{y}(\cdot) = \left(y_1(\cdot), \dots, y_M(\cdot)\right)'$ is the array of functional data, $\boldsymbol{C} = \left(c_{q,j}\right) \in \mathbb{R}^{Q+L-1,M}$ is the coefficient matrix and $\boldsymbol{\Phi} = \left(\Phi_1, \dots, \Phi_{Q+L-1}\right)'$ is the basis function array. Note that, in case of interpolating B-splines, one has $\boldsymbol{\tau} = \{0, 1, \dots, T\}$.

Given a square-integrable reference profile $\tilde{y}$ (e.g., the sample mean of the g_j's) and the j-th curve y_j, the registration of y_j to $\tilde{y}$ via a functional time warping approach consists of estimating a function h_j within a given class of warping functions such that the similarity between the registered function $y^*(t) = y_j\left(h_j(t)\right)$, $t \in [0, T]$, and the reference function $\tilde{y}(t)$, $t \in [0, T]$, is maximized. The parametric time warping approach (Eilers, 2004; Ramsay and Silverman, 2005) involves the use of parametric warping functions induced by polynomials of degree K. Here, the generic warping function h is assumed to be a monotonic increasing function such that $h(0) = 0$ and $h(T) = T$, and is obtained as

$$h(t) = \frac{\int_{[0,t]} \exp\left(h^K(z)\right) dz}{\int_{[0,T]} \exp\left(h^K(z)\right) dz}, \qquad t \in [0, T] \tag{3}$$

where h^K is a polynomial function of degree K:

$$h^K(t) = \sum_{k=0}^{K} w_k t^k, \quad t \in [0, T] \qquad (4)$$

The degree K can be fixed ex-ante, or estimated, e.g., by a stepwise estimation method. The choice of the parameter K is mostly problem-dependent, and needs to represent a compromise between the capability of dealing with arbitrary phase shifts and the avoidance of undesired shape distortions induced by the registration itself (see Section 7). We propose a method for the selection of the degree K in Appendix A. Note that even if $K = 1$, transformation (3) induces a non-linear warping function. Nevertheless, in case of anchored values at the boundaries of the domain (as in this case), the use of non-linear warping functions is a natural choice for the registration problem. In general, non-linear warping functions are likely to be required when the misalignment involves local effects, as in the real case study motivating this work.

To determine the optimal function h_j within a given class of warpings – which in the following is the class defined in (4) –, Ramsay and Silverman (2005) advocates the use of the *continuous fitting criterion*, that is based on the minimization of the following statistics:

$$MINEIG(h) = \lambda_2[\mathbf{T}(h)] \qquad (5)$$

where

$$\mathbf{T}(h) = \begin{bmatrix} \int \{\tilde{y}(t)\}^2 dt & \int \tilde{y}(t) y_j[h(t)] dt \\ \int \tilde{y}(t) y_j[h(t)] dt & \int \{y_j[h(t)]\}^2 dt \end{bmatrix} \qquad (6)$$

and $\lambda_2[\mathbf{T}(h)]$ is the size of the second eigenvalue of $\mathbf{T}(h)$. The size of the smallest eigenvalue of $\mathbf{T}(h)$ quantifies the dissimilarity between the original curve and the aligned curve induced by a pure difference in phase (Ramsay and Silverman, 2005). In order to provide an intuitive justification for the Ramsay and Silverman's procedure, suppose two curves, $\tilde{y}(t)$ and $y_1(t)$, differ only in amplitude but not in phase (Fig. 2, panel a). Then, if we plot their function values against each other (Fig. 2, panel b), we observe a straight line (i.e., a one-dimensional set of points), such that the magnitude of the amplitude difference is reflected in the slope of the line. If the curves differ both in amplitude and in phase (Fig. 2, panel c), by plotting their function values against each other we observe a bidimensional curve (Fig. 2, panel d), i.e., a departure from unidimensionality, where the slope of semi-major axis reflects the difference in amplitude. Matrix $\mathbf{T}(h)$ is such that the second eigenvalue measures departures from unidimensionality, and hence the minimization of the size of $\lambda_2[\mathbf{T}(h)]$ corresponds to minimizing the dissimilarity imposed by a pure difference in phase. Because of this, the $MINEIG(h)$ statistic is more effective than a least squares criterion, as the latter is intrinsically designed to minimize differences in amplitude, rather than in phase (Ramsay and Silverman, 2005). Other approaches different from the continuous fitting criterion proposed by Ramsay and Silverman are known. Indeed Sangalli *et. al*. (2013) and Vantini (2012) point out that the choice of the similarity index and the class of warping functions is problem-specific and needs to be done in agreement to a minimal set of requirements that guarantee coherence. By coherence they refer particularly (a) to the invariance of the similarity between two profiles if these are registered via the same warping function, and (b) to the invariance of similarity gained by registering a first curve to a second one or the second to the first. For the purposes of this paper, we set aside this deeper analysis and decided to resort to the similarity induced by MINEIG, since this criterion is already well established in the literature on functional data analysis. Further details are discussed in Section 7.

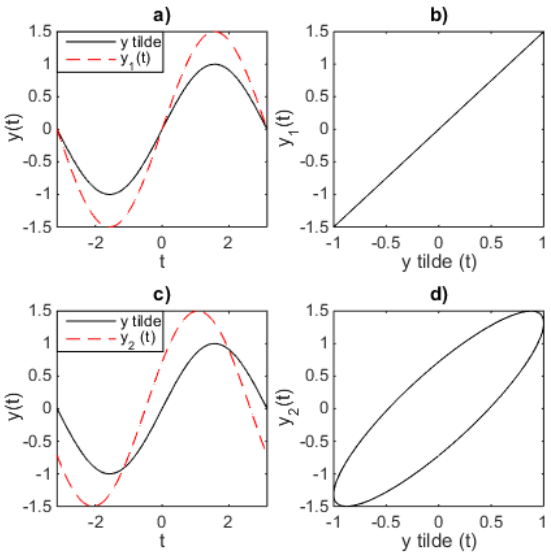

FIGURE 2. Example of two curves that differ only in amplitude (top panels) and two curves that differ both in amplitude and phase (bottom panels)

Once the optimal h_j has been estimated, the registered function, y_j^*, is calculated by applying the two following steps (Ramsay and Silverman, 2005):

1. Estimation of the inverse warping function h_j^{-1} such that $h_j^{-1}[h_j(t)] = t$ for all $t \in [0,T]$;
2. Interpolation of the relationship between $h_j^{-1}(t)$ and $y_j(t)$, for $t \in [0,T]$.

The warping function h_j can be estimated by using the Newton – Raphson algorithm to minimize the $MINEIG(h)$ statistics. The stopping rule can be based on a tolerance threshold (e.g., $tol = 10^{-4}$).

4. Methodology

The profile monitoring approach we propose involves three major steps: (i) profile registration via parametric time warping, (ii) analysis and dimensionality reduction of the registered data via FPCA, and (iii) application of multivariate control charts to monitor both the registered curves and the warping coefficients. Fig. 3 schematically depicts the proposed approach. The three major steps are discussed in the next Sub-Sections.

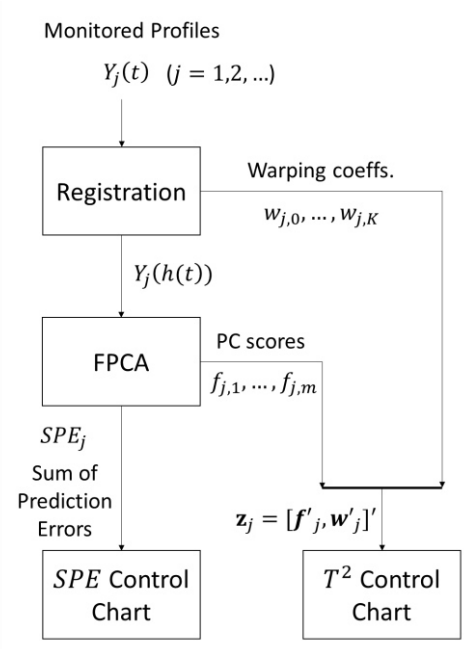

FIGURE 3. Scheme of the proposed approach

4.1 Profile registration

Let $\{y_j(t), t \in [0,T]\}$ be the j^{th} acquired profile under natural process conditions, where $j = 1, \dots, M$. These M profiles represent the Phase I dataset, as they are used to design the control charts. Without loss of generality, we assume all the profiles to be sampled at $T+1$ equispaced data points, within the common temporal domain $[0,T]$, and represented via a (smoothing or interpolating) cubic B-spline basis with equispaced knots as in (2).

The proposed approach for the registration of the Phase I historical dataset consists of a coarse registration, followed by a fine registration, which represent the two stages of a Procrustes fitting criterion (Ramsay and Silverman, 2005):

Coarse registration: Phase I profiles $y_j(t), t \in [0,T]$, $j = 1,2, \dots, M$, are registered by using the sample mean profile $\bar{y}(t), t \in [0,T]$, as the reference pattern. To choose the degree K of the polynomial h^K, the following two options are available: (i) ex-ante selection of a degree K (e.g., as a rule of thumb, a cubic model is assumed to be a good compromise in most practical applications), or (ii) stepwise selection of the degree K. In the latter case, the method proposed in Appendix A, based on the average $MINEIG(h)$ statistics, can be used.

Fine registration: Once all the Phase I profiles have been preliminarily registered, one can compute the new sample mean profile $\overline{y^*}$, which is a better estimation of the reference pattern. Thus, a second stage registration is applied, by registering all the original profiles y_j with respect to $\overline{y^*}$; the same degree K is applied in this stage.

This procedure represents a Procrustes approach because it involves estimating a transformation by registering to an iteratively updated reference curve. In many practical applications, the first registration stage may be sufficient (Ramsay and Silverman, 2005). However, in the absence of a good reason for applying only the coarse registration stage (e.g., the need to reduce the computation time), we suggest to apply the two sequential stages, to guarantee a better estimation of the reference pattern. Indeed, an iterative refinement of this two-stage procedure would lead to the 1-mean alignment advocated by Sangalli *et al*. (2010).

The nomenclature y_j^* is used in this study to identify the final registered profiles. The registration of Phase II profiles simply consists of aligning each newly observed curve to the reference pattern $\overline{y^*}$ by applying the continuous fitting criterion (Section 3) to choose the optimal warping function within the same class (3) used to register Phase I profiles.

4.2 Functional PCA on registered curves

The output of the profile registration procedure consists of a collection of registered curves and $K+1$ warping coefficients $w_{j,k},\ k=0,1,\dots,K,\ j=1,2,\dots,M$. An effective approach to control the stability over time of the registered patterns consists of monitoring a limited set of functional features representing the largest portion of data variability, as a result of a FPCA. The use of the Principal Component Analysis (PCA) (Jolliffe, 2002) for profile monitoring applications was proposed by Colosimo and Pacella (2007; 2010), and some variants were discussed by other authors (Kim *et al*., 2006; Paynabar *et al*., 2013). Colosimo and Pacella (2007; 2010) apply the PCA on the discretized profile matrix obtaining loadings (i.e., the eigenvectors of the sample covariance matrix) of the same length as the profiles. Instead, we propose an FPCA approach leading to functional loadings, i.e., the eigenfunctions of the sample covariance operator (Ramsay and Silverman, 2005). In the following, the terms *loading, functional principal component (FPC)* and *eigenfunction* are used interchangeably.

The reader is referred to Appendix B for a brief review of the FPCA methodology for squared integrable profiles. The FPCs, denoted by $\xi_i,\ i=1,\dots,m$, play the role of the loadings in the classic multivariate PCA, whereas the scores, $f_{j,i},\ ,\ j=1,\dots,M, i=1,\dots,m$, are obtained by projecting the original functions onto a functional subspace spanned by the first m eigenfunctions, where $m \ll M$ is chosen to capture a given portion of the original data variability.

The score vector, $\boldsymbol{f}_j=[f_{j,1},\dots,f_{j,m}]',\ j=1,\dots,M$, associated to the aligned curves, represents a multivariate quality characteristic to assess the stability over time of the amplitude variability of the monitored profiles. Thus, the score vector on the one hand, and the vector of warping coefficients on the other hand, capture two complementary portions of information about the process.

4.3 Control chart design and utilization

The scores along the FPCs of the registered functions $y_j^*,\ j=1,2,\dots,M$, exhibit only (or mainly) the amplitude variability. Nevertheless, it is not possible to assume, in general, a zero-correlation between FPCs and warping functions (see Appendix C). To the best of our knowledge, finding classes of warping functions and associated similarities that guarantee the independence or zero-correlation between phase and amplitude variability is still an open problem. Because of this, to achieve a simultaneous monitoring of the stability over time of both the amplitude and the phase variability, we propose to control the $(m+K+1)$-dimensional vectors $\boldsymbol{z}_j, j=1,\dots,M$, obtained by concatenating the warping coefficients $w_{j,k},\ k=0,1,\dots,K$, and the principal component scores $f_{j,i}\ i=1,2,\dots,m$, namely, $\boldsymbol{z}_j=[f_{j,1},\dots,f_{j,m},w_{j,0},\dots,w_{j,K}]'$. These can be monitored by means of a Hotelling's T^2 control chart. In this case, the control statistics is:

$$T_j^2(m,K)=(\mathbf{z}_j-\bar{\mathbf{z}}_j)\mathbf{S}_z^{-1}(\mathbf{z}_j-\bar{\mathbf{z}}_j)', \qquad (7)$$

$$j=1,2,\dots,M$$

where $\bar{\mathbf{z}}_j$ is the Phase I sample mean of the coefficient vector, and $\mathbf{S}_z$ is its Phase I sample variance-covariance matrix. We remark that the block-structure of the matrix $\boldsymbol{S}_z$ can be exploited to efficiently compute the T^2 statistics in case a high number m of eigenfunctions is required to accurately describe the profiles:

$$\mathbf{S}_z=$$

$$\begin{bmatrix} \begin{bmatrix} \sigma_{f_{11}}^2 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \sigma_{f_{mm}}^2 \end{bmatrix} & \begin{bmatrix} \sigma_{f_1 w_0} & \dots & \sigma_{f_1 w_K} \\ \dots & \dots & \dots \\ \sigma_{f_m w_0} & \dots & \sigma_{f_m w_K} \end{bmatrix} \\ \begin{bmatrix} \sigma_{f_1 w_0} & \dots & \sigma_{f_m w_0} \\ \dots & \dots & \dots \\ \sigma_{f_1 w_K} & \dots & \sigma_{f_m w_K} \end{bmatrix} & \begin{bmatrix} \sigma_{w_{00}}^2 & \dots & \sigma_{w_{K0}} \\ \dots & \dots & \dots \\ \sigma_{w_{K0}} & \dots & \sigma_{w_{KK}}^2 \end{bmatrix} \end{bmatrix} \qquad (8)$$

where $\sigma_{f_{\cdot\cdot}}^2$ is the FPC score variance, $\sigma_{w_{\cdot\cdot}}^2$ is the warping coefficient variance, and $\sigma_{f.w.}$ is the covariance between FPC scores and warping coefficients. In addition, the FPCA model residuals can be monitored to detect deviations along directions orthogonal to the first m directions. The sum of prediction error (SPE):

$$SPE_j(m)=$$

$$\sum_{i=1}^{n}(\hat{y}_j^*(t_i,m)-\overline{y^*}(t_i))(\hat{y}_j^*(t_i,m)-\overline{y^*}(t_i)), \qquad (9)$$

$$j=1,\dots,M$$

can be used for such a task.

Two control charts are designed to monitor respectively the $T_j^2(m,K)$ and the $SPE_j(m)$ statistics. The control limits may be estimated as $100(1-\alpha')\%$

percentiles of the empirical distributions of the two statistics, where α is the overall Type I error, and $\alpha = 1-(1-\alpha')^{(1/2)}$ is the Type I error associated to each chart, computed by using the Sidak correction (Montgomery, 2008). During Phase II, i.e., the actual monitoring phase, each newly observed profile $y_j(t_i)$, $j = M+1, M+2, ..., t_i = 0, \dots, T$, is represented in a functional form by using the same basis introduced in Phase I, it is registered with respect to the Phase I reference profile $\overline{y^*}$, and the corresponding warping coefficients are estimated. The Phase I FPCA model is applied by using the number m of principal components chosen during the control chart design phase. The result is a projection of the new registered profile y_j^* onto the m-dimensional functional subspace spanned by the eigenfunctions estimated during the Phase I. Then, the two statistics $T_j^2(m)$ and $SPE_j(m)$ are computed and their values are compared with the corresponding control limits. If at least one statistic violates the limit, an alarm signal is issued.

5. Simulation study

5.1 Data generation

The performances of the proposed approach are firstly evaluated by means of Monte Carlo simulations. A benchmark signal inspired by the work of Tang and Muller (2008) is used to simulate the patterns. Two different scenarios are considered by using the same kind of signal: (i) *Scenario A*, characterized by a natural amplitude variability that predominates over the natural phase variability, and (ii) *Scenario B*, characterized by a natural phase variability that predominates over the amplitude variability. Scenario A is representative of practical situations where a rough alignment is applied (e.g., by using triggers or a synchronous resampling), but a fine registration may still be required. Scenario B is representative of situations where no preliminary alignment step is applied on raw data. The in-control models are generated as follows:

Scenario A:

$$y_j(t_i) = \sum_{i=1}^{5} \beta_{i,j} \exp\left\{\gamma_{i,j}\left(t_i + \omega_{i,j}\right)^2\right\} + \varepsilon_{ij}, \quad t_i = 0,1, \dots, 100 \text{ and } j = 1, \dots, M \tag{10}$$

where:

(i) $\boldsymbol{\beta}_j = [\beta_{1,j}, \dots, \beta_{5,j}] \sim MN[\boldsymbol{\mu}_{\beta,A}, \boldsymbol{\Sigma}_{\beta,A}]$,
(ii) $\boldsymbol{\gamma}_j = [\gamma_{1,j}, \dots, \gamma_{5,j}] \sim MN[\boldsymbol{\mu}_{\gamma,A}, \boldsymbol{\Sigma}_{\gamma,A}]$, and
(iii) $\boldsymbol{\omega}_j = [\omega_{1,j}, \dots, \omega_{5,j}] \sim MN[\boldsymbol{\mu}_{\omega,A}, \boldsymbol{\Sigma}_{\omega,A}]$,

with the following parameters:

- $\boldsymbol{\mu}_{\beta,A} = [0.88, -0.5, 0.6, 0.6, -0.5]$, $\boldsymbol{\Sigma}_{\beta,A} = diag[(8.8,5,6,6,5)10^{-2}]$
- $\boldsymbol{\mu}_{\gamma,A} = [-20, -50, -100, -150, -200]$, $\boldsymbol{\Sigma}_{\gamma,A} = diag[2,5,10,15,20]$
- $\boldsymbol{\mu}_{\omega,A} = [-0.5, -0.45, -0.3, 0.7, -0.45]$, $\boldsymbol{\Sigma}_{\omega,A} = diag[(5,4.5,3,2,1.5)10^{-2}]$
- ε_{ij} $i.i.d. \sim N(0, \sigma_\varepsilon^2)$ (noise term), where $\sigma_\varepsilon = 0.05$

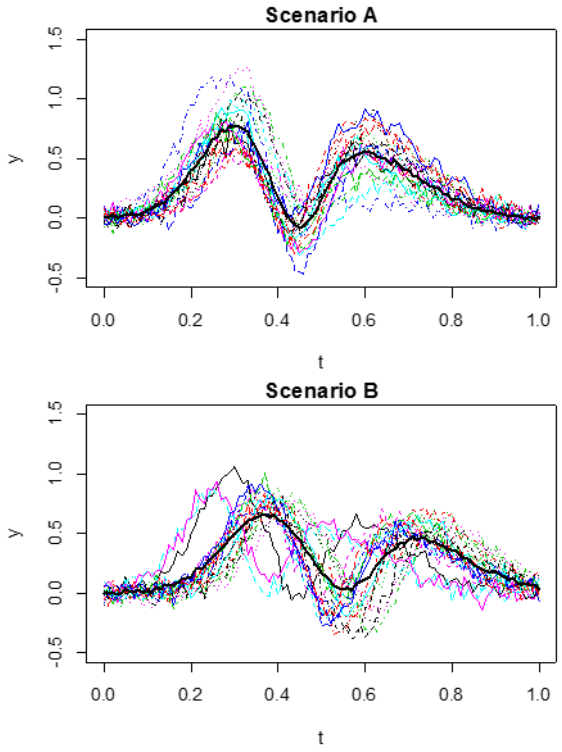

FIGURE 4. Simulated in-control signals in Scenario A and Scenario B, and their sample mean (thick black line)

Scenario B:

$$y_j(t_i) = \sum_{i=1}^{5} \beta_{i,j} \exp\left\{\gamma_{i,j}\left(t_i + \tau_j \omega_{i,j}\right)^2\right\} + \varepsilon_{ij}, \quad t_i = 0,1, \dots, 100, \text{ and } j = 1, \dots, M \tag{11}$$

where:

(i) $\boldsymbol{\beta}_j = [\beta_{1,j}, \dots, \beta_{5,j}] \sim MN[\boldsymbol{\mu}_{\beta,B}, \boldsymbol{\Sigma}_{\beta,B}]$,
(ii) $\boldsymbol{\gamma}_j = [\gamma_{1,j}, \dots, \gamma_{5,j}] \sim MN[\boldsymbol{\mu}_{\gamma,B}, \boldsymbol{\Sigma}_{\gamma,B}]$,
(iii) $\boldsymbol{\omega}_j = [\omega_{1,j}, \dots, \omega_{5,j}] \sim MN[\boldsymbol{\mu}_{\omega,B}, \boldsymbol{\Sigma}_{\omega,B}]$, and

(iv) $\tau_j \sim N[\mu_\tau, \sigma_\tau]$,

with the following parameters:

- $\boldsymbol{\mu}_{\beta,B} = \boldsymbol{\mu}_{\beta,A}, \boldsymbol{\Sigma}_{\beta,B} = d\boldsymbol{\Sigma}_{\beta,A}$
- $\boldsymbol{\mu}_{\gamma,B} = \boldsymbol{\mu}_{\gamma,A}, \boldsymbol{\Sigma}_{\gamma,B} = \boldsymbol{\Sigma}_{\gamma,A}$
- $\boldsymbol{\mu}_{\omega,B} = \boldsymbol{\mu}_{\omega,A}, \boldsymbol{\Sigma}_{\omega,B} = diag[(5,4.5,3,2,1.5)10^{-3}]$
- $\mu_\tau = 1.2$ and $\sigma_\tau = 0.15$
- $\varepsilon_{ij}\ i.i.d. \sim N(0, \sigma_\varepsilon^2)$ (noise term), where $\sigma_\varepsilon = 0.05$

Fig. 4 shows $M = 50$ realizations of the in-control profiles in Scenario A and Scenario B.

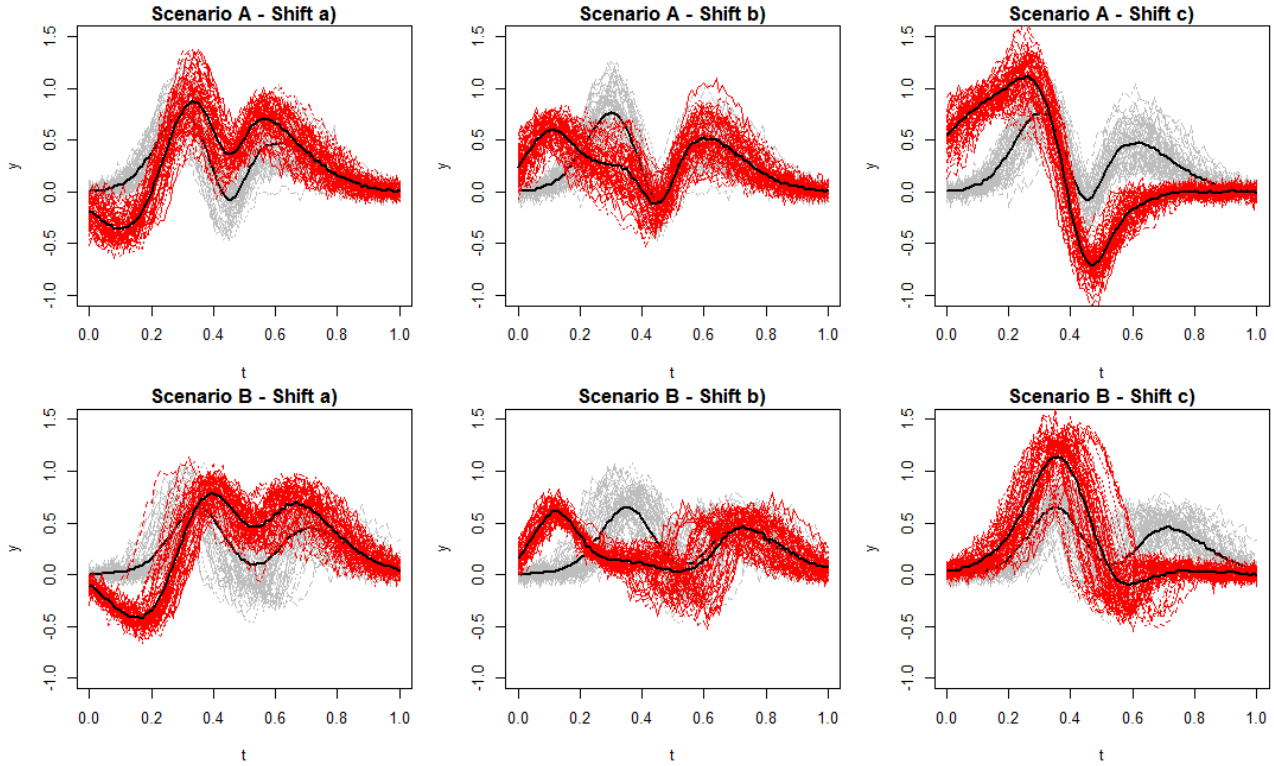

FIGURE 5. Effects of simulated out-of-control shifts (at largest severity level) in Scenario A and Scenario B. Each panels shows: in-control profiles (background grey lines), out-of-control profiles (foreground red lines) and the corresponding sample mean (thick black lines)

TABLE 1. Simulated out-of-control shift conditions in Scenario A and Scenario B

Scenario A		Scenario B	
Shift a)	$\mu^s_{\omega,A,2} = \mu_{\omega,A,2}/\delta_{a,A}$	Shift a)	$\mu^s_{\omega,B,2} = \mu_{\omega,B,2}/\delta_{a,B}$
Shift b)	$\mu^s_{\omega,A,3} = \mu_{\omega,A,3}/\delta_{b,A}$	Shift b)	$\mu^s_{\omega,B,3} = \mu_{\omega,B,3}/\delta_{b,B}$
Shift c)	$\mu^s_{\omega,A,1} = \mu_{\omega,A,1}/\delta_{c,A}$	Shift c)	$\mu^s_{\omega,B,1} = \mu_{\omega,B,1}/\delta_{c,B}$

TABLE 2. Severity levels associated to each out-of-control shift conditions

Scenario A		Scenario B	
Shift a)	$\delta_{a,A} \in [1.5,2,2.5,3]$	Shift a)	$\delta_{a,B} \in [1.5,2,2.5,3]$
Shift b)	$\delta_{b,A} \in [1.5,2,2.5,3]$	Shift b)	$\delta_{b,B} \in [1.5,2,2.5,3]$
Shift c)	$\delta_{c,A} \in [1.5,2,2.5,3]$	Shift c)	$\delta_{c,B} \in [1.15,1.2,1.3,1.4]$

In order to simulate unnatural departures from the in-control pattern, we considered three kinds of signal perturbation that affect both the amplitude and the phase variability. They are representative of faults that modify the shape of the reference pattern, as in many practical situations. The shift conditions were simulated as shown in Table 1, where $\boldsymbol{\mu}^{S}_{\omega,A}$ and $\boldsymbol{\mu}^{S}_{\omega,B}$ are the sample mean of the model parameters in Scenario A and B in presence of a shift, respectively.

The shifts are applied on one term of $\boldsymbol{\mu}_{\omega,A}$ or $\boldsymbol{\mu}_{\omega,B}$ at a time, i.e., on $\mu_{\omega,A,u}$ or $\mu_{\omega,B,u}$, where $u = 1, \ldots, 5$. The severity levels are shown in Table 2. The effect of the different shifts on the profile pattern, corresponding to the largest severity level, is shown in Fig. 5. Note that the model used to generate both the in-control and the out-of-control profiles can be interpreted in the light of the general model (Eq. 1).

5.2 Analysis of data under in-control conditions

The degree of the polynomial warping functions was assessed by computing the mean values $\overline{MINEIG}_K(h)$ of the MINEIG statistic at different values of the warping function degree K for $M = 50$ realizations of in-control profiles, represented via interpolating cubic B-splines (Fig. 6, top panels). According to the criterion proposed in Appendix A, the choice of $K = 3$ appears to be the most appropriate in both scenarios. In addition, Fig. 6 reports the registration results for $K = 3$, namely the registered profiles (central panels) and the warping functions used to align the data (bottom panels).

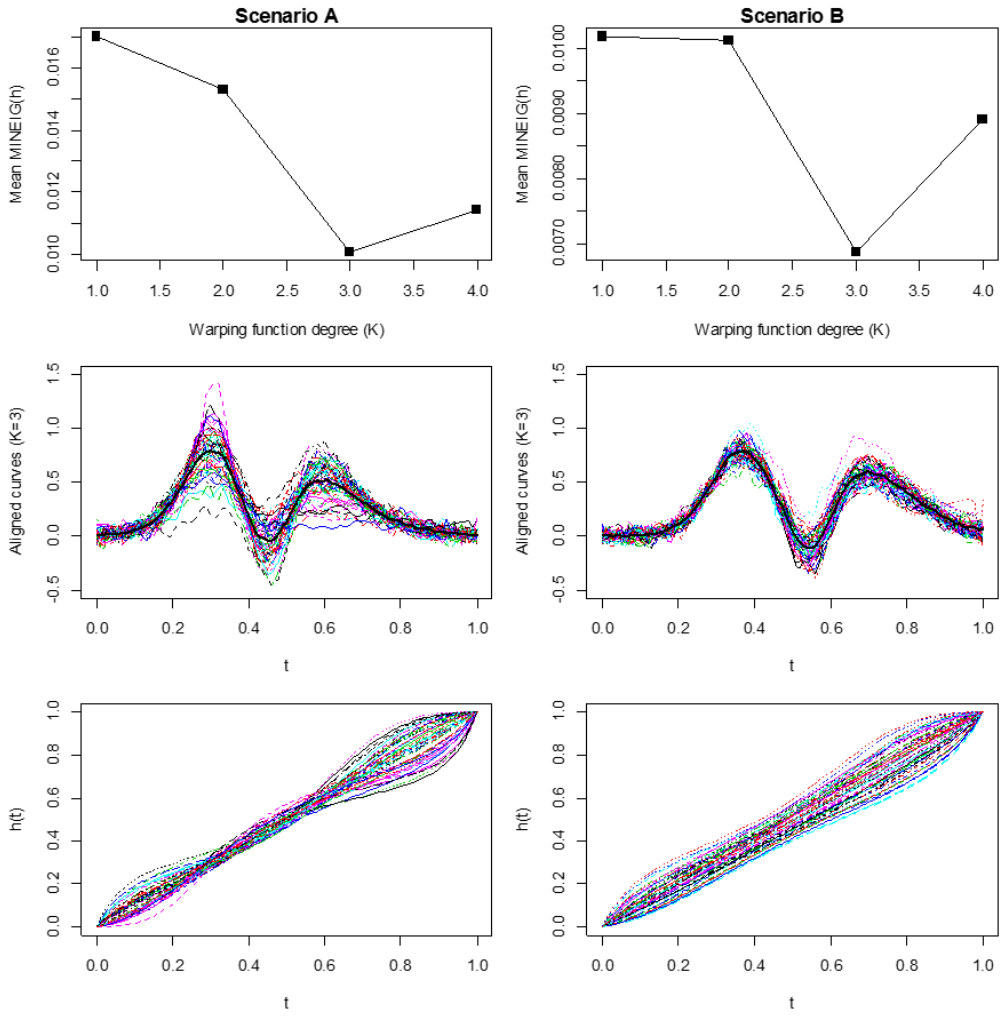

FIGURE 6. $\overline{MINEIG}_K(h)$ statistic for different degrees K of the polynomials h's (top panels), aligned curves by using $K = 3$ (central panels), and corresponding warping functions (bottom panels).

We recall that, in Scenario A, the amplitude variability was predominant: this still affects the profiles variability after the registration. Instead, in Scenario B, the phase variability was predominant: a higher similarity among the curves is achieved after the registration, with a corresponding higher variability of the warping functions $h_j(t)$, $j = 1,2,\dots,M$. To evaluate how the phase and the amplitude variability affect the principal components, we now focus on the pattern presented by the loadings. Fig. 7 and Fig. 8 show, for Scenario A and B respectively, the first $m = 2$ eigenfunctions estimated by applying the FPCA on the unregistered curves, and on the registered profiles in the central panels of Fig. 6 (bottom panels).

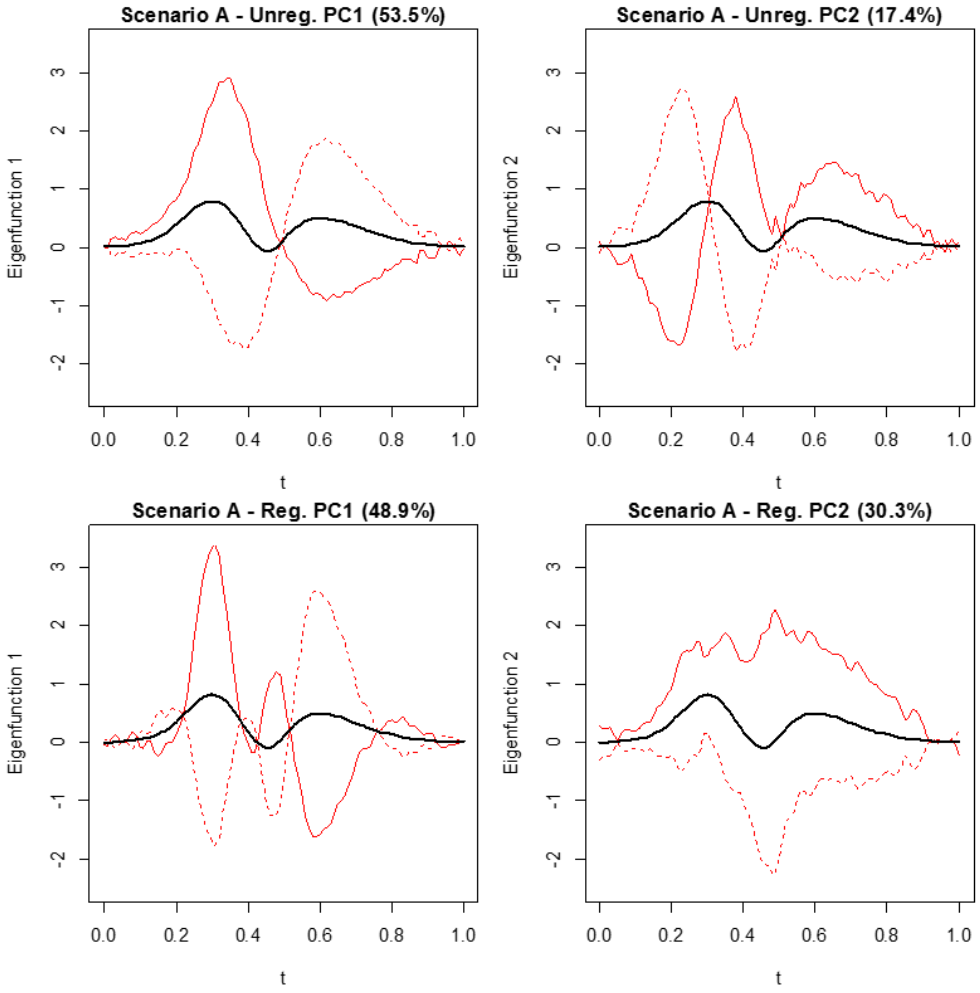

FIGURE 7. First $m = 2$ PCs in Scenario A, for unregistered (top panels) and aligned (bottom panels) profiles. Each panel shows the mean curve (solid thick line), the mean curve plus the eigenfunction (solid fine line) and the mean curve minus the eigenfunction (dashed line)

With regard to Scenario A, when the FPCA is applied to unregistered data, there is some misalignment between the eigenfunction peaks and the salient features of the profile. Such a shift is minimized by applying the FPCA on the registered curves, meaning that the registration allows enhancing the localizations of the points of largest local variability.

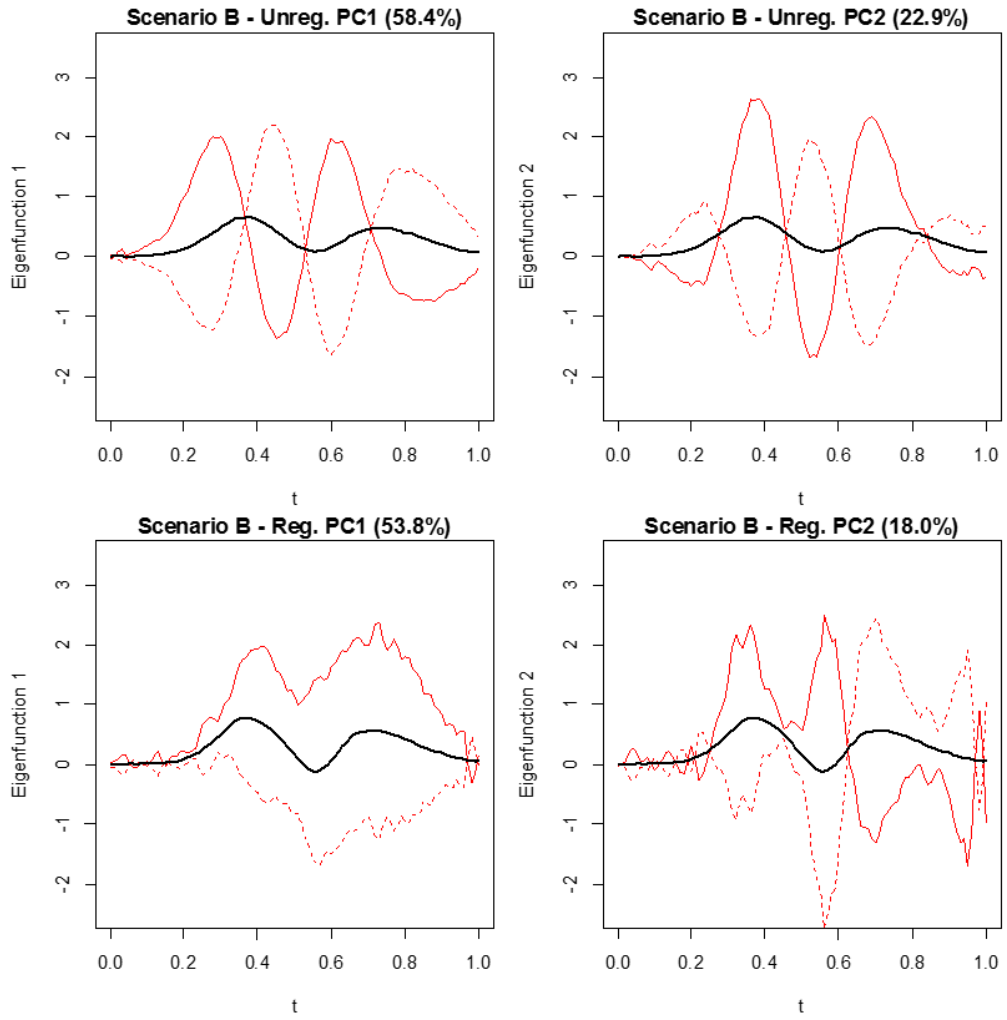

FIGURE 8. First $m = 2$ PCs in Scenario B, for unregistered (top panels) and aligned (bottom panels) profiles. Each panel shows the mean curve (solid thick line), the mean curve plus the eigenfunction (solid fine line) and the mean curve minus the eigenfunction (dashed line)

With regard to Scenario B, when the FPCA is applied to unregistered data, the second FPC is nearly a phase-shifted version of the first one. This means that those FPCs are considerably influenced by the phase variability. After the registration, the FPCs resemble the ones of Scenario A (Fig. 7, bottom panels), but in inverted order. Thus, after the misalignment errors are removed by the warping functions, the FPCs mainly capture the amplitude variability. In all those cases, the first two FPCs account for about 80% of the global variability.

5.3 Simulation results

We compare three different implementations of the FPCA-based control charting approach: (i) profile monitoring without registration (hereafter denoted by *Unreg* approach), (ii) profile monitoring after warping-based registration, but without including the warping coefficients into the monitored statistics (denoted by *Reg* approach) and (iii) the proposed approach for profile monitoring (denoted by *RegWarp* approach). The *Unreg* approach is representative of what is done in practice when no alignment (or only a rough alignment) is applied before analysing the data, whereas the *Reg* approach is representative of what is usually done in the mainstream literature.

The Phase II performances are compared in terms of the Average Run Length, ARL, for a targeted Type I error $\alpha = 0.01$. Without loss of generality, an in-control ARL (ARL_0) equal to 100 is used to keep affordable the duration of the overall simulation tests.

In each simulation scenario, we performed 100 runs. In each run, we used as Phase I dataset a set of $M = 50$ randomly generated in-control profiles. The Phase II dataset comprises a tuning set and a testing set. The tuning dataset, consisting of $M_1 = 1000$ in-control profiles, is generated to estimate the empirical distribution of the monitored statistics (i.e., the T^2 and the SPE statistics) and to guarantee that the Phase II performance comparison is made under statistically equivalent values of ARL_0. To guarantee the achievement of the targeted ARL_0, we estimate the empirical percentiles by applying the Kernel Density Estimation (KDE) technique (Chou *et al*., 2001). To this end, we resort to the KDE algorithm by Bowman and Azzalini (1997), with normal kernel and 1000 equally spaced points. We based the estimation of the Phase II performances on a testing set consisting of $N = 2000$ profiles.

To avoid the inclusion of outlier profiles within the in-control dataset, we imposed the 97.5% percentile as cut-off threshold for the Mahalanobis distance of the retained FPCs (Varmuza and Filzmoser, 2009). To perform registration and FPCA, we exploited the routines available in the R package FDA (Ramsay *et al*., 2012). Finally, the number of FPCs to be retained was chosen according to an 80% cut-off of the explained data variability.

TABLE 3. Out-of-control detection results in Scenario A: ARLs and 95% Confidence Intervals

	Severity	**Scenario A - ARLs and 95% Confidence Intervals**					
		Unreg		**Reg**		**RegWarp**	
IC	-	*101.17*	*[91.00, 111.34]*	*101.29*	*[83.45, 119.14]*	*100.85*	*[92.03, 109.67]*
Shift a)	1.5	**3.24**	**[2.91, 3.56]**	**4.10**	**[3.56, 4.64]**	4.52	[4.16, 4.88]
	2.0	**1.74**	**[1.60, 1.87]**	2.53	[2.35, 2.72]	2.65	[2.45, 2.85]
	2.5	**1.64**	**[1.52, 1.77]**	2.31	[2.06, 2.57]	2.29	[2.08, 2.51]
	3.0	**1.71**	**[1.56, 1.88]**	**1.90**	**[1.71, 2.09]**	**1.99**	**[1.81, 2.16]**
Shift b)	1.5	7.82	[5.99, 9.64]	26.06	[18.83, 33.30]	**2.01**	**[1.90, 2.11]**
	2.0	6.46	[5.31, 7.61]	15.05	[13.15, 16.94]	**1.07**	**[1.06, 1.08]**
	2.5	5.65	[4.28, 7.02]	5.89	[5.28, 6.50]	**1.04**	**[1.03, 1.05]**
	3.0	5.48	[4.70, 6.25]	3.46	[3.19, 3.73]	**1.06**	**[1.05, 1.08]**
Shift c)	1.5	8.13	[7.35, 8.90]	**6.03**	**[5.45, 6.61]**	**5.50**	**[5.20, 5.80]**
	2.0	5.16	[4.85, 5.47]	**4.09**	**[3.81, 4.37]**	**3.82**	**[3.68, 3.97]**
	2.5	2.78	[2.66, 2.89]	2.38	[2.29, 2.48]	**2.20**	**[2.12, 2.27]**
	3.0	1.87	[1.80, 1.94]	1.69	[1.65, 1.73]	**1.57**	**[1.54, 1.61]**

Table 3 and Table 4 show the simulation results, in Scenario A and B respectively, in terms of Phase II performances for the different kinds of unnatural pattern modification. These are also graphically depicted in Fig. 9. For each simulated condition, a 95% confidence interval for the estimated *ARL* values is shown.

In Scenario A (Table 3 and Fig. 9 - left), the *RegWarp* and the *Reg* approaches provide statistically comparable results for shift *a* and shift *c*, where the pattern modification mainly involves the amplitude variability. The *RegWarp* outperforms the *Reg* approach for shift *b*, where the pattern modification largely affects the phase variability.

In the out-of-control condition of shift *a*, the best detection results are achieved by the *Unreg* approach, although this is outperformed by the *RegWarp* approach for shift *b* and shift *c*. This means that the out-of-control effect on the unregistered curves may be sometimes more evident than the effect after the warping-based registration, especially when the original variability of the data is dominated by the amplitude term. This can be caused by a shape distortion introduced by the time warping operation, as discussed in Section 7.

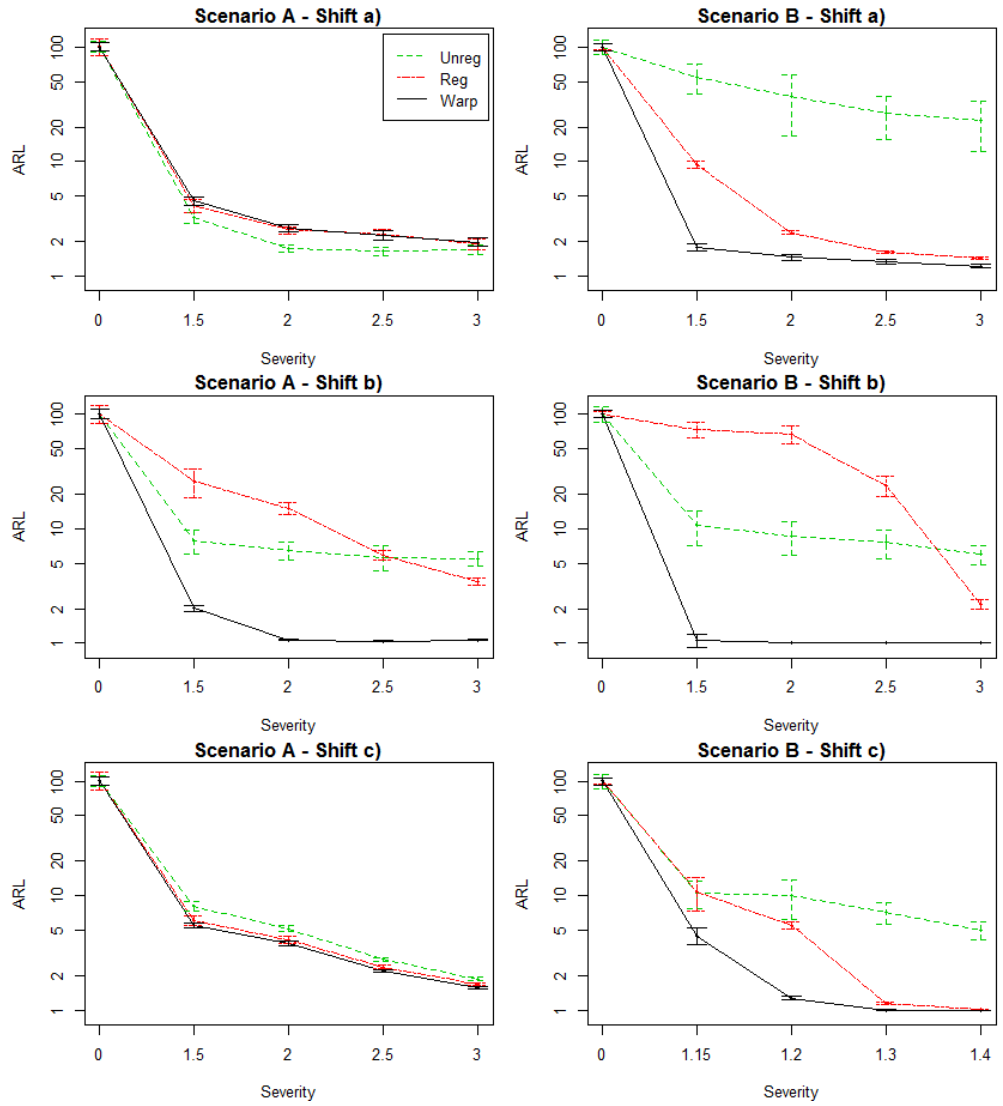

FIGURE 9. ARLs and 95% Confidence Intervals for different out-of-control shifts in Scenario A (left) and Scenario B (right)

TABLE 4 – Out-of-control detection results in Scenario B: ARLs and 95% Confidence Intervals

	Severity	Scenario B - ARLs and 95% Confidence Intervals					
		Unreg		Reg		RegWarp	
IC	-	*100.43*	*[85.46, 115.40]*	*100.30*	*[94.32, 106.29]*	*100.32*	*[92.85, 107.78]*
Shift a)	1.5	55.18	[38.86, 71.50]	9.45	[8.80, 10.11]	**1.78**	**[1.64, 1.92]**
	2.0	36.80	[16.73, 56.87]	2.40	[2.32, 2.49]	**1.46**	**[1.38, 1.55]**
	2.5	26.47	[15.59, 37.35]	1.63	[1.59, 1.67]	**1.32**	**[1.26, 1.39]**
	3.0	23.00	[12.33, 33.67]	1.43	[1.39, 1.47]	**1.22**	**[1.17, 1.27]**
Shift b)	1.5	10.69	[7.16, 14.22]	72.98	[61.93, 84.02]	**1.06**	**[1.00, 1.20]**
	2.0	8.63	[5.90, 11.36]	67.07	[54.82, 79.32]	**1.00**	**[1.00, 1.00]**
	2.5	7.60	[5.49, 9.70]	23.86	[19.07, 28.65]	**1.00**	**[1.00, 1.00]**
	3.0	5.96	[4.83, 7.10]	2.18	[1.96, 2.39]	**1.01**	**[1.00, 1.01]**
Shift c)	1.15	10.60	[7.76, 13.45]	10.89	[7.33, 14.45]	**4.46**	**[3.71, 5.22]**
	1.2	9.95	[6.17, 13.74]	5.52	[5.09, 5.96]	**1.27**	**[1.22, 1.33]**
	1.3	7.16	[5.57, 8.76]	1.14	[1.13, 1.16]	**1.00**	**[1.00, 1.00]**
	1.4	5.00	[4.13, 5.87]	1.01	[1.01, 1.01]	**1.00**	**[1.00, 1.00]**

In Scenario B (Table 4 and Fig. 9 - right), where the misalignment of original profiles is larger, the *RegWarp* approach considerably outperforms the two competitor implementations of the FPCA for shift *a*, shift *b* and shift *c*. The performance enhancement produced by the *RegWarp* approach over the *Reg* approach is caused by the fact that the registration operation mitigates the effect of the fault when only the amplitude variability is considered, as it forces the current out-of-control observations to resemble the in-control reference pattern. Such a dissimilarity reduction is desired when the process operates under in-control conditions, but it may have detrimental effects on fault detection capabilities when traditional control charting methods are applied only on registered curves. The performance enhancement over the *Unreg* approach is instead a consequence of the separation of the shape variability sources: the proposed approach allows reducing the variability inflation caused by profile misalignment, which may mask the effect of faults, without losing any information about the warping functions.

A discussion about the possible shape distortion effects introduced by the time warping operation, together with further analysis of different time warping algorithms, is reported in Section 7.

6. A real case study

We present a real case study to demonstrate the performances of the proposed approach in an actual industrial application. It consists of a waterjet cutting operation, where water pressure fluctuations are monitored during the process itself, to detect faults affecting either the Ultra High Pressure (UHP) pump or the cutting head components.

Due to challenging pressure conditions, cyclic pressure loadings, aggressiveness of abrasives and other factors, most of the machine components are subject to wear and unpredictable faults. Thus, the in-process monitoring of machine health conditions is of great industrial interest, as it allows implementing condition-based maintenance strategies, and providing an automatic reaction to critical faults. Fig. 10 shows the dynamic pressure profiles, i.e., the pressure fluctuations around the static level, acquired during a cutting process by using a 45 kW positive-displacement pump that includes three single-acting pistons (we refer to Annoni *et al*., 2008 for a description of the plant). The water pressure set value was 350 MPa, by using a 0.25 mm orifice. The acquired profiles came from a synchronous re-sampling procedure, leading to a fixed profile length of 820 data points.

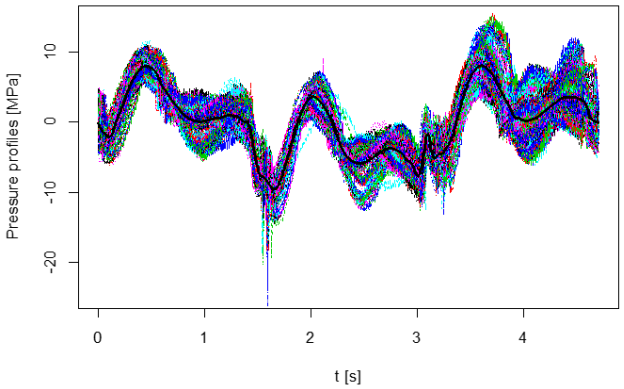

FIGURE 10. Dynamic pressure profiles under in-control conditions

The pressure signal is a suitable source of information for monitoring purposes, as it is featured by fluctuations that are influenced by both upstream and downstream flow rate modifications. The cutting process is characterized by repeating pressure profiles, one for each pumping cycle. Real data were acquired under normal health conditions and in the presence of actual faults. The following faults scenarios were considered, as they involve the most critical components and refer to common contingencies in WJ/AWJ shop floors: A) cracked high pressure cylinder, B) cracked discharge check valve, C) worn discharge check valve seat, and D) broken orifice. With regard to faults A, B and C, different faulty components coming from actual faults were available, and they were ranked on the basis of the severity of the crack or the wear level. With regard to fault D, three different broken orifices were tested. The effects of those faults on the pressure profile are shown in Fig. 11. We refer to Grasso *et al*. (2013; 2014a) for further details about the fault scenarios.

6.1 Implementation details

The Phase I dataset consists of $M = 130$ profiles corresponding to the in-control pumping cycles. Each profile, after the synchronous-resampling step (see Section 2), consists of 820 data points, and the duration of a pumping cycle is about 4.7 s. In the frame of in-process SPC, one has to deal with a trade-off decision about the accuracy of the functional model fitting and the time devoted to registration algorithm: the higher the number of basis functions,

n_b, the higher the computational time. An interpolating cubic B-spline basis would require $n_b = 820 + 2$ basis function, considerably slowing the registration procedure. By using a regression B-spline basis, the computational time can be reduced, but some smoothing is attained. Smoothing filters out high frequency components which might be due to measurement noise. This may cause a potential information loss, but, it may enhance the shape variability characterization when large-scale pattern features are of predominant interest. Fig. 12 shows the CPU time[2] for registering the M in-control profiles for different numbers of cubic B-spline basis functions with equally spaced knots, and the corresponding sum of squared error, SSE, i.e., the sum of squared error between the estimated regression B-spline curves and the original observations.

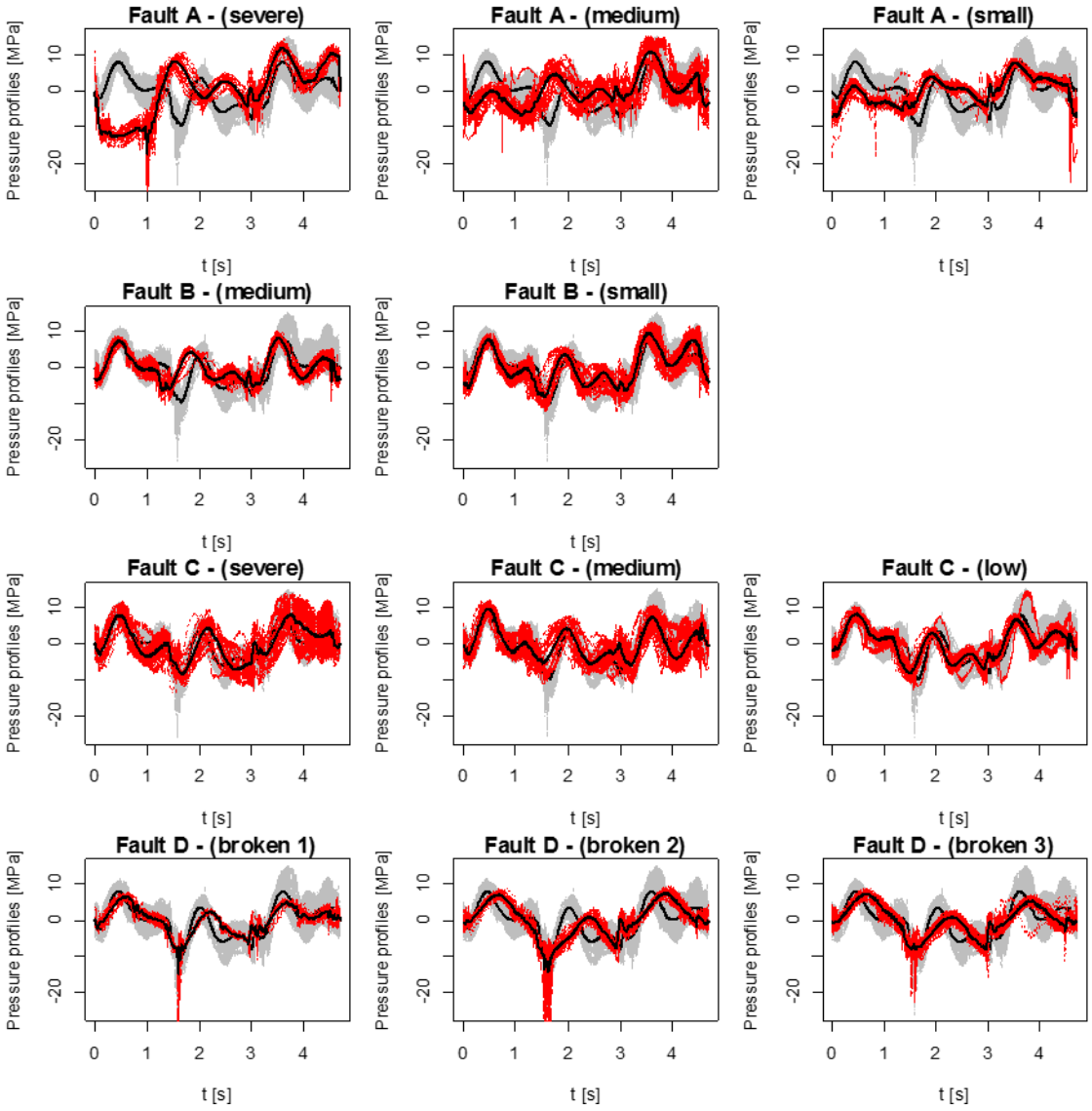

FIGURE 11. Effect of real faults (foreground red lines) on the dynamic pressure signal, for different severity levels; the in-control pattern is shown in background, grey lines

[2] CPU time refers to an Intel® Core™ i7-3740QM CPU @ 2.70 GHz

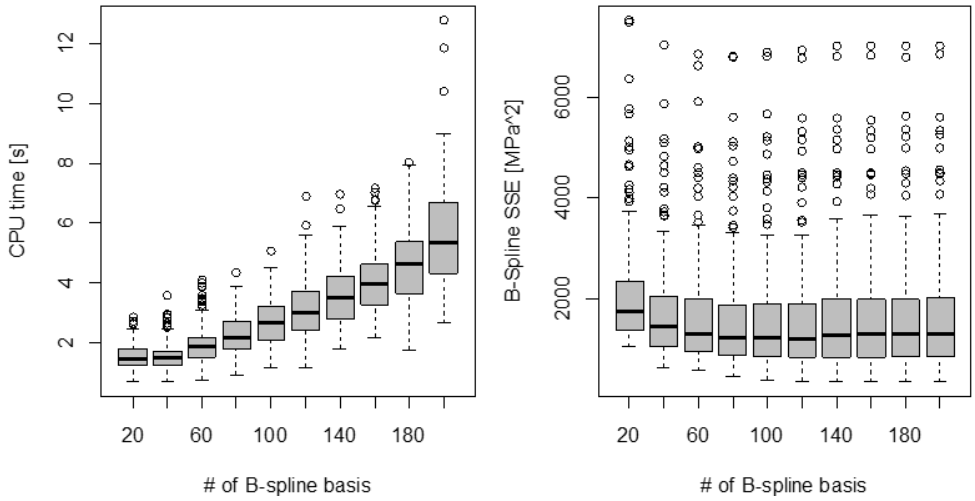

FIGURE 12. CPU computational time for single profile registration under in-control conditions, and corresponding SSE of the regression B-spline model

For the present analysis, we compare the results achieved by using an interpolating basis and a regression B-spline basis consisting of $n_b = 100$ cubic functions with equispaced knots. The regression B-spline basis appears to be a good balance between the CPU time and the SSE. A more effective smoothing may be achieved by using different knot selection strategies (e.g., Goldenthal and Bercovier, 2004, Zhou and Shen, 2001; Molinari *et al.*, 2004, Sangalli *et al.*, 2009b), but the additional computational effort for the automatic selection of the knot sequence has to be taken into account. If the high frequency and transient components filtered out by the smoothing process are deemed relevant for shape characterization, an additional control chart may be added to monitor the B-spline residuals.

Fig. 13 shows the aligned in-control profiles and the corresponding warping functions using the regression B-spline basis, which highlight an initial moderate misalignment (the parameter K controlling the complexity of the warping function being set to $K = 3$).

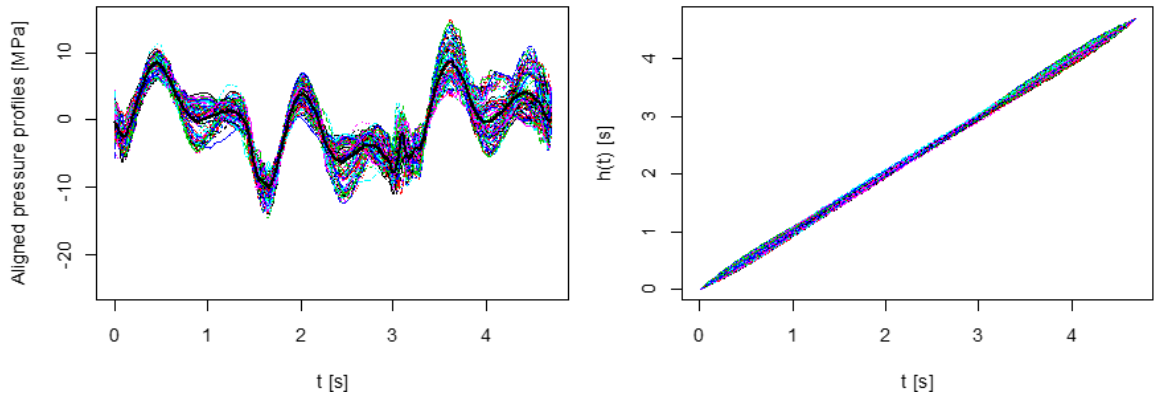

FIGURE 13 – Aligned pressure profiles and the corresponding warping functions

By way of example, Fig. 14 shows the first $m = 2$ FPCs resulting from applying the FPCA to unregistered and registered data respectively. Since the misalignment was moderate, the eigenfunctions

look very similar, but after the registration the phase error between the FPCs and the salient features of the curve is minimized. In both cases, the first $m = 2$ FPCs account for about 75% of the overall variability.

6.2 Analysis of the results

To collect signal data, the same type of cutting process – consisting of cutting 75 x 75 x 2.5 mm square aluminium plates – was performed under each machine health condition, both in-control and with faulty components installed. The fault detection percentages provided by the three competitor methods are shown in Table 5. The empirical control limits were estimated by using the KDE technique (Chou *et al*., 2001), with a target Type I error $\alpha = 0.0027$ as per common industrial practice. Table 5 shows that the use of a regression B-spline basis not only drastically reduces the computational effort, but slightly improves the detection percentage yielded by the *RegWarp* approach. The *Unreg* and the *Reg* methods, instead, perform slightly better when an interpolation basis is used. A smoothing basis enhances the registration by removing the high frequency noise component: the result is an increase of the fault mitigation affecting the *Reg* method, and an improvement of the *RegWarp* method that is robust against such an effect. On the other hand, the removal of high frequency components may reduce the fault detectability when the *Unreg* method is used.

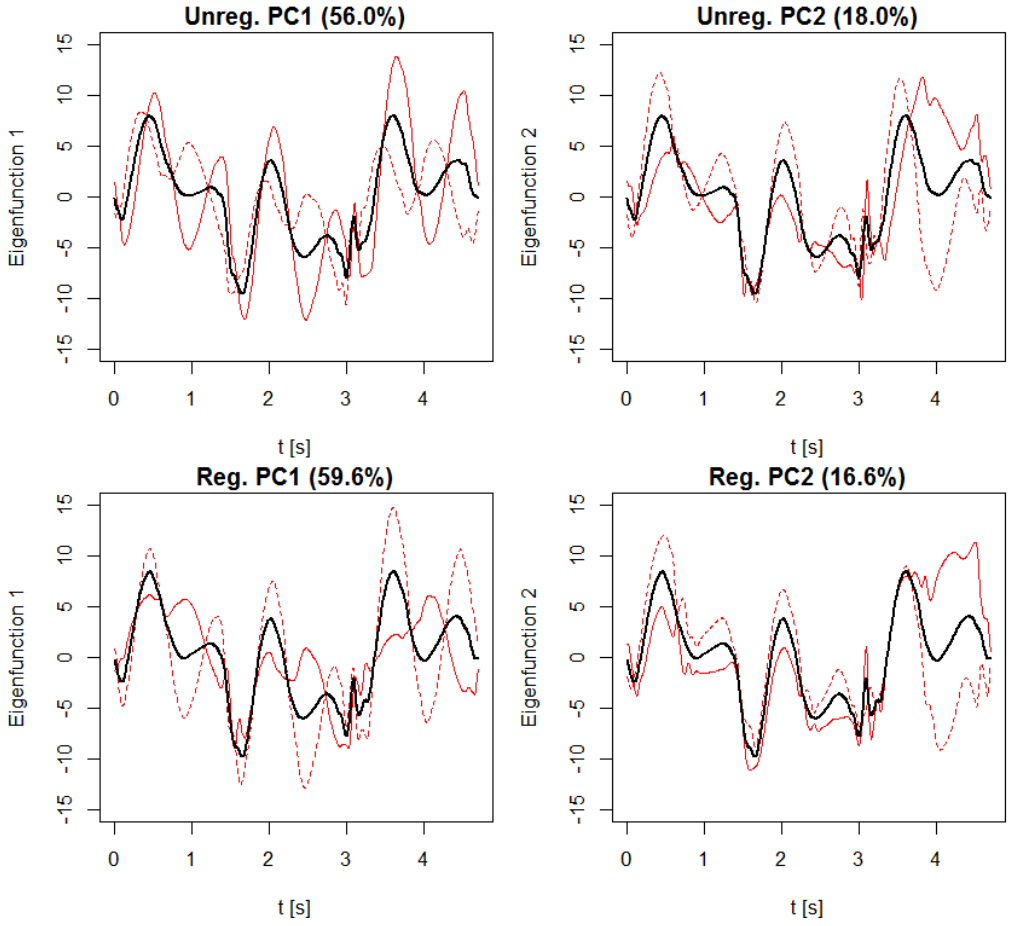

FIGURE 14. First $m = 2$ FPCs of the pressure signal, for unregistered curves (top panels) and aligned curves (bottom panels). Each panel shows the mean curve (solid thick line), the mean curve plus the eigenfunction (solid fine line) and the mean curve minus the eigenfunction (dashed line)

Regardless of the selected basis, Fault A is always detected by all the methods. With regard to fault B, C and D, the *RegWarp* approach is the one that provides the highest performances, i.e., a 100% detection capability for fault B and fault D, and a slightly lower detection percentage for fault C, at low and medium

wear levels. The *Reg* approach performs worse than the *Unreg* one in the presence of fault B, which mainly affects the phase variability (see Fig. 13), and in presence of fault C, when a barely worn salve seat in installed (also in this case the fault effect mainly involves the phase variability). Both the *Reg* and the *Unreg* approaches yield poor detection performances when the "broken 1" orifice is installed and better performances for the other two faulty orifices (this is more evident when a smoothing basis is used). Globally, the *Reg* and the *Unreg* approaches allow detecting about the 83-84% of out-of-control profiles with a smoothing basis, and about 86-88% with an interpolating basis. The *RegWarp* approach, instead, yields a detection rate of about 98-99%, with a slight performance improvement when smoothed profiles are monitored.

TABLE 5. Fault detection percentages of *Unreg*, *Reg* and *RegWarp* approaches, for different kinds of fault, different severity levels and different choices of the B-spline basis

Fault	Fault severity	Fault detection percentage (%)					
		Interpolating basis			Smoothing basis		
		Unreg	Reg	RegWarp	Unreg	Reg	RegWarp
A	Severe	**100**	**100**	**100**	**100**	**100**	**100**
	Medium	**100**	**100**	**100**	**100**	**100**	**100**
	Small	**100**	**100**	**100**	**100**	**100**	**100**
B	Medium	**100**	75.86	**100**	**100**	82.76	**100**
	Small	96.55	63.64	**100**	89.66	81.82	**100**
C	Severe	**100**	**100**	**100**	**100**	**100**	**100**
	Medium	92.31	76.92	92.31	84.62	84.62	**96.15**
	Low	69.23	30.77	88.46	53.85	3.85	**92.31**
D	Broken 1	23.53	98.84	94.12	17.65	35.29	**100**
	Broken 2	**100**	**100**	**100**	91.43	**100**	**100**
	Broken 3	71.43	**100**	**100**	68.54	**100**	**100**
Total		*88.40*	*86.01*	*97.95*	*83.96*	*83.28*	***98.98***

7. Discussion: on the distortion effect imposed by the warping operation

The time warping operation aims at minimizing a dissimilarity measure between a reference pattern and any newly observed patterns. This operation improves the profile variability characterization when the process is in-control, but leads to a shape distortion when the process is out-of-control. Indeed, the out of control profiles are forced in this stage to resemble the in-control reference one, leading to a "*fault mitigation effect*", which can be dealt with by monitoring both the registered profiles and the warping functions used to align them. Moreover, when the process is out-of-control, the Phase I reference profile, $\overline{y^*}(t)$, is no longer a consistent reference for the new observations. This may lead the registration procedure to converge to a highly distorted profile by inflating or denaturing the outcome of the fault itself, resulting into a "*fault alteration effect*". As illustration of this, we here consider the same Phase I data as that in Scenario B of Section 5, and a fault effect consisting of a pure shift of the time reference t, such that $t^s = t - \delta_d$ in equations (8) and (9). Even though a pure shift is unlikely produced by an actual fault, it represents an event that may occur in practice, especially when profile monitoring is applied to signal data. Indeed, it may be caused by a wrong profile segmentation or by an error involving the triggering signals used to define the profile time window.

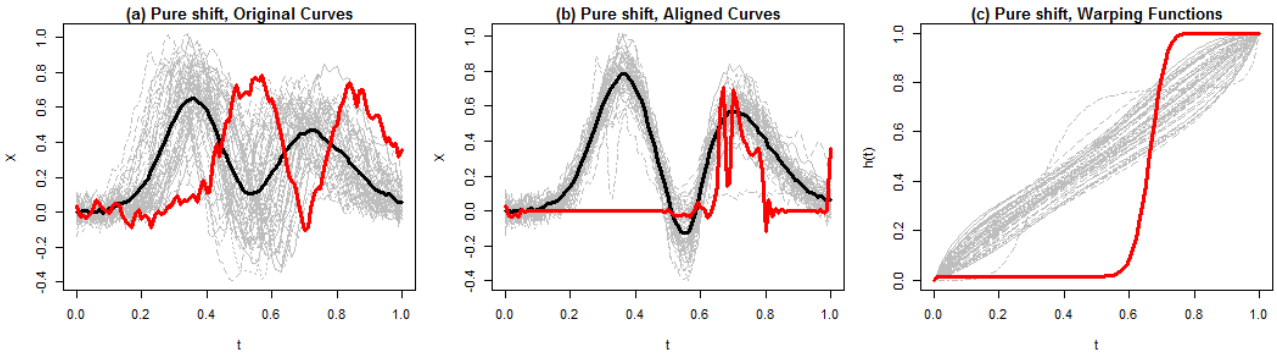

FIGURE 15. Effect of pure shift on unregistered profiles (a), aligned profiles (b) and warping functions (c) for a severity level equal to 0.15. Each panels shows: in-control profiles (background grey lines) and a realization of the out-of-control profiles (red line); in-control sample mean profile is depicted with thick black lines

Fig. 15 a) depicts the effect of such a kind of shift for a severity level of $\delta_d = 0.15$ (red line), compared to the 50 in-control profiles (grey line). In this case, registration based on the MINEIG criterion and the family of polynomial warping functions of order $K = 3$, fails to converge to the aligned profile, resulting into the distorted pattern shown in Fig. 15 b). We note that the largest the severity of the shift is, the more likely the pattern alteration, as the first peak of the out-of-control profiles may be erroneously aligned to the second peak of the in-control profiles. The unnatural shape alteration, due to a local minimum entrapment, is worsened in this case by a border effect caused by the truncation of the out-of-control profiles within the $[0,1]$ time window.

The problem of alteration effect is well-known in the literature on functional data registration. To cope with such a problem, different approaches can be adopted, depending on the problem at hand. A possible parameter to be tuned is the order of the polynomial determining the warping functions. Indeed, the higher the order of the warping functions is, the most likely severe deformations occur. For instance, Fig. 16 shows the results of applying the registration procedure with the MINEIG criterion when $K = 1,2$. In this case, the use of polynomials of lower degree allows to solve the fault alteration effect, eventually avoiding Phase I profiles alteration for $K = 1$.

A different approach to cope with the alteration effect is to adopt a different registration criterion for the registration procedure. A relatively large body of literature has been recently devoted to this topic with particular focus on the characterization of the most suitable classes of warping functions to be used for registration, depending on the minimization criteria (Sangalli *et al*., 2013; Vantini, 2012). As a way of example, we here illustrate the results of two alternative methods: (a) the 1-mean alignment (Sangalli *et al*., 2010) and (b) the registration based on the Fisher-Rao metric (Srivastava *et al*., 2011). We performed the 1-mean alignment with similarity measure determined by the Pearson metric and affine warping functions – advocated by Sangalli *et al*. (2013); results are reported in Fig. 17 a) and b). The solution obtained with the Fisher-Rao metric (Srivastava *et al*., 2011) and warping functions class of diffeomorphism is displayed in Fig. 17 c) and d).

The results appear pretty similar and in both cases the fault alteration effect seems overcome. From a monitoring viewpoint, the 1-mean alignment is performed with parametric warping functions at the expense of modifying the profile domain. Here the monitoring of the phase variability could be performed by monitoring the two parameters of shift and dilation determining the affine warping. The monitoring of the diffeomorphisms singled out when registering according to the Fisher-Rao metrics would instead require a phase variability dimensionality reduction. In both cases, an extensive study to assess the performance of both methods in profile monitoring will be scope of future work.

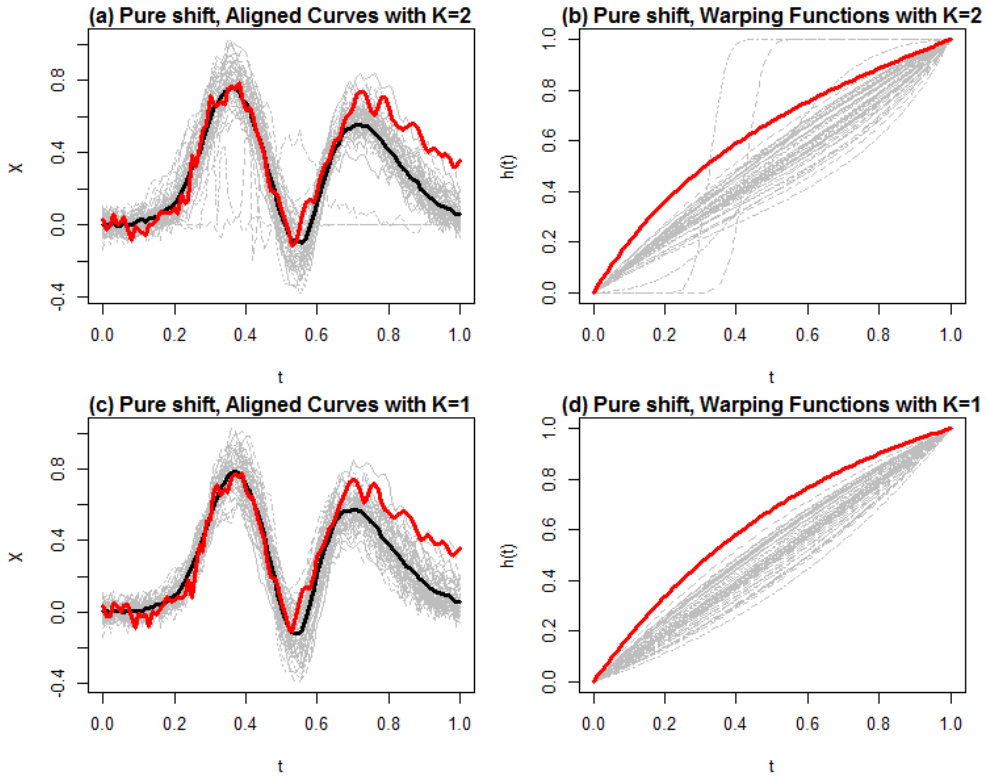

FIGURE 16 - Registration of an out-of-control profile with warping functions determined by polynomials of order $K = 1$ ((a) and (b)) and $K = 2$ ((c) and (d)): registered profiles ((a) and (c)) and warping functions ((b) and (d))

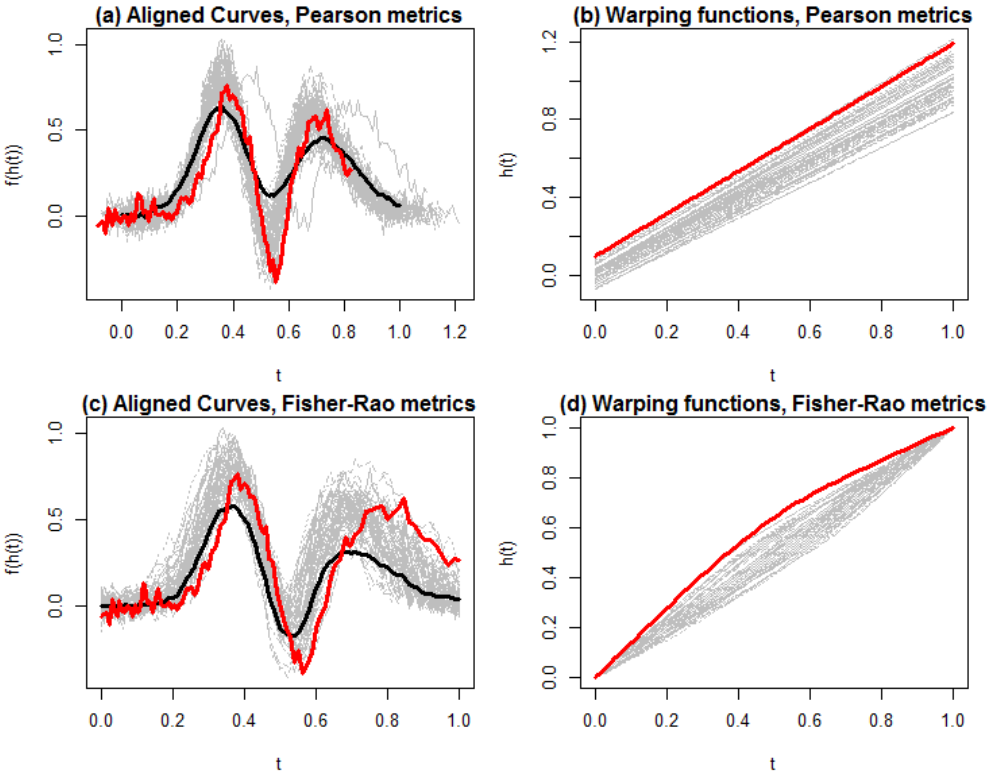

FIGURE 17 - Registration of an out-of-control profile based on alternative criteria. Results with 1-mean alignment based on the Pearson Metrics with affine functions ((a) and (b)) and Fisher-Rao Metrics with diffeomorphism ((c) and (d), with smoothing parameter $\lambda = 0.5$): registered profiles ((a) and (c)) and warping functions ((b) and (d))

It is worth to notice that the avoidance (or mitigation) of the shape alteration effect might be achieved at the expense of a worst registration (this is evident in Fig. 17 c)). Generally speaking, the choice of the warping approach may be problem-dependent and it involves a compromise between the goodness of the registration itself and the avoidance of undesired shape distortions. To this aim, both the flexibility of the warping function and the registration criterion play a relevant role.

We finally remark that, from a profile monitoring viewpoint, when a shape distortion occurs only under out-of-control conditions, the result is likely to be an increase of the fault detectability when the *Reg* method or the *RegWarp* method are applied. If it occurs when the process is in-control, outlier data will appear in the Phase I historical dataset, with a detrimental effect on the performances of both the *Reg* and the *RegWarp* methods. For this reason, one has to pay close attention to the alteration effect, particularly during Phase I, when a careful analysis – and possibly a decontamination operation – may be needed to tune the monitoring parameters and choose the most appropriate registration criterion.

8. Conclusions

Functional data registration, in the frame of profile monitoring applications, is required to guarantee a proper decoupling of the amplitude variability from the phase variability, and to minimize the undesired inflation of the phase variability due to disturbance factors. Even though an effective method to keep under statistical control both the phase and the amplitude variability of the process is essential for profile monitoring, the mainstream literature typically focuses the control only on the amplitude variability. Moreover, some signal processing techniques commonly employed, including automatic triggering or synchronous re-sampling, are not completely effective, and do not guarantee an actual registration of the profile data.

Our study is a first contribution aimed at investigating the integration of time warping algorithms into a profile monitoring framework. We proposed a novel approach to jointly monitor the stability over time of the registered profiles and of the warping functions used to align them.

The simulated scenarios show that the inclusion of the warping coefficients into the monitored statistics improves in most cases the process monitoring performances with respect to monitoring only the aligned profiles. The larger the original misalignment of the acquired profiles is, the larger the provided improvement. Furthermore, a profile monitoring without registration is likely to be more reliable than a profile monitoring of aligned curves without keeping track of the registration parameters, as the registration itself may even mitigate the effect of the fault. The proposed approach is expected to be more effective in detecting small shape modifications, thanks to an improved shape variability characterization. The real case study concerning an actual waterjet cutting process confirmed the results achieved in simulated scenarios. In particular, the information loss is minimized by including the warping coefficients into the monitored statistics. This allows enhancing the fault detection capability, even when the registration procedure produces a mitigation of the actual fault effects. However, when the original phase variability is very large and/or in the presence of strong departures from the natural pattern, the registration operation may introduce some shape distortion that may alter the nature of the occurred events. We discussed possible ways to cope with the fault alteration effect, while highlighting that the determination of the most appropriate solution appear to be highly problem-driven. Future research streams should address the development of registration criteria coupled with a proper choice of warping functions to enhance the monitoring performances.

In the end, time warping algorithms may be computationally expensive, and their use in practical applications should be driven by a trade-off analysis between the desired enhancement of profile characterization and the reactivity of the process monitoring tool. The computational time of functional warping algorithms may be influenced by the number of basis functions used to represent the sampled profiles as functional forms, and hence this parameter can be controlled to find the best compromise between monitoring performances and computational efficiency.

Nomenclature

ARL, ARL_0	Average Run Length, Average Run Length under in-control conditions
$\boldsymbol{b}_j$	Vector of the j^{th} eigenfunction B-spline coefficients
$\boldsymbol{C} = (c_{q,j})$	B-spline coefficient matrix
$\bar{\boldsymbol{c}}$	Vector of the sample mean B-spline coefficients
$D_{K,i}$	Relative improvement (or worsening) by passing from a degree K to a degree $K+i$ warping function
$\boldsymbol{f}_j$	j^{th} score vector associated to the m retained FPCs ($\boldsymbol{f}_j = [f_{j,1}, \dots, f_{j,m}]'$) , $j = 1,2, \dots$
FPC	Functional Principal Component
FPCA	Functional Principal Component Analysis
$h(t)$	Warping function
$\mathbf{J}$	$(Q+L-1) \times (Q+L-1)$ matrix used in equation (B7)
K	Degree of the warping function
KDE	Kernel Density Estimation
L	Number of B-spline sub-intervals ($L-1$ is the number of internal knots)
m	Number of retained FPCs
M	Number of profiles in Phase I dataset
$MINEIG(h)$	Similarity index used for registration
n_b	Number of B-spline basis functions
PC	Principal Component
PCA	Principal Component Analysis
Q	Order of the B-spline function
Reg	Profile monitoring based on monitoring only the registered curves
$RegWarp$	Proposed approach for profile monitoring
$SPE_j(m)$	j^{th} realization of the sum of prediction error statistic
SSE	Sum of squared errors statistic
$\mathbf{S}_z$	Sample variance-covariance matrix of multivariate variable $\boldsymbol{z}_j$
T+1	Number of data points of monitored profiles ($t, t_i \in [0,T]$)
$\mathbf{T}(h)$	2×2 matrix used to compute the $MINEIG(h)$ statistics
$T_j^2(m,K)$	j^{th} realization of the Hotelling's T^2 statistics for the multivariate variable $\boldsymbol{z}_j$
UHP	Ultra High Pressure
$\boldsymbol{u}_j$	j^{th} eigenvector in FPCA eigenfunction
$Unreg$	Profile monitoring without any registration
$v(s,t)$	Sample covariance function
V	Sample covariance operator (with kernel $v(s,t)$)
$\boldsymbol{w}_j$	j^{th} warping coefficient vector ($\boldsymbol{w}_j = [w_{j,1}, \dots, w_{j,K}]'$) , $j = 1,2, \dots$
WJ/AWJ	Waterjet/Abrasive Waterjet
$\bar{y}(t)$	Sample mean profile in Phase I before registration
$\overline{y^*}(t)$	Sample mean profile in Phase I after registration

$y_j^*(t)$	j^{th} registered profile, $j = 1,2, \ldots$
$y_j(t_i)$	j^{th} sampled profile, $j = 1,2, \ldots$
$y_j(t)$	j^{th} functional profile, $j = 1,2, \ldots$
$\hat{y}_j^*(t,m)$	j^{th} reconstructed profile by using the first m FPCs
$\boldsymbol{z}_j$	j^{th} monitored vector that includes the PC scores and the warping coefficients ($\boldsymbol{z}_j = [f_{j,1}, \ldots, f_{j,m}, w_{j,0}, \ldots, w_{j,K}]'$)
α, α'	Type I error
$\beta_{i,j}, \gamma_{i,j}, \mu_{i,j}, \mu_{i,j}^S, \omega_{i,j}, \tau_j$	Parameters used to generate the simulate datasets, $j = 1,2, \ldots; i = 1, \ldots, 5$
$\delta_{\cdot,A}, \delta_{\cdot,B}$	Shift parameter used to simulated out-of-control shape modifications
$\varepsilon_j(t)$	j^{th} error term in simulated scenarios ($\varepsilon_j(t) \sim N(0, \sigma_\varepsilon)$)
ξ_j	j^{th} FPC (eigenfunctions), $j = 1,2, \ldots$
λ_2	Size of the second eigenvalue of a 2×2 matrix
ρ_j	j^{th} eigenvalue in FPCA eigenfunction
$\sigma_{f\cdot\cdot}^2$	FPC score variance
$\sigma_{w\cdot\cdot}^2$	Warping coefficient variance
$\sigma_{f.w.}$	Covariance between FPC scores and warping coefficients
τ	Knot sequence (B-spline basis)
$\boldsymbol{\Phi}$	B-spline basis functions array

References

Altman, N. S., and Villarreal, J. C. (2004), Self-Modelling Regression for Longitudinal Data With Time-Invariant Covariates, The Canadian Journal of Statistics, 32, 251–268

Annoni M., Cristaldi L., Lazzaroni M. (2008), Measurements, Analysis And Interpretation of The Signals From A High-Pressure Waterjet Pump, IEEE Transactions on Instrumentation and Measurement, 57:1, 34-47

Bowman, A. W., and A. Azzalini (1997), Applied Smoothing Techniques for Data Analysis, New York: Oxford University Press

Chang S. I., Yadama S. (2010), Statistical Process Control for Monitoring Non-Linear Profiles Using Wavelet Filtering and B-Spline Approximation, International Journal of Production Research, 48:4, 1049-1068

Chen, C.S., Hung, Y.P., (1999), RANSAC-based DARCES: a new approach to fast automatic registration of partially overlapping range images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 21, 1229–1234

Chou, Y-M., Mason, R.L., and Young J.C. (2001), The Control Chart for Individual Observations from A Multivariate Non-normal Distribution. Communications in Statistics, Simulation and Computation, 30(8–9), 1937–1949

Colosimo, B.M., Pacella, M. (2007), On the Use of Principal Component Analysis to Identify Systematic Patterns in Roundness Profiles, Quality and Reliability Engineering International, 23, 925 - 941

Colosimo, B.M., Pacella, M. (2010), A Comparison Study of Control Charts for Statistical Monitoring of Functional Data, International Journal of Production Research, 23, 707 – 725

Del Castillo, E., Colosimo, B.M. (2011), Statistical Shape Analysis of Experiments for Manufacturing Processes, Technometrics, 53(1), 1-15

Ding, Y., Zeng, L., & Zhou, S. (2006). Phase I analysis for monitoring nonlinear profiles in manufacturing processes. Journal of Quality Technology, 38:3, 199-216

Dryden, I. L., Mardia, K. V. (1998), Statistical shape analysis (Vol. 4), New York: John Wiley & Sons.

Eilers, P. H. (2004), Parametric time warping. Analytical Chemistry, 76:2, 404-411.

Gao X. (2012), On-line Monitoring of Batch Process with Multiway PCA/ICA, Chapter 13 in Principal Component Analysis – Multidisciplinary Application, Intech Ed.

Gasser, T., Kneip, A., Ziegler, P., Largo, R. Prader, A. (1990), A method for determining the dynamics and intensity of average growth, Annals of Human Biology, 17, 459–474

Gervini, D., Gasser, T. (2004), Self-modelling warping functions, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66(4), 959-971

Gervini, D. (2014), Dynamic Retrospective Regression for Functional Data, Technometrics, DOI: 10.1080/00401706.2013.879076

Goldenthal, R., Bercovier, M. (2004), Spline curve approximation and design by optimal control over the knots. Computing, 72:1-2, 53-64

Grasso, M., Goletti, M., Annoni, M., Colosimo, B.M. (2013), A New Approach for Online Health Assessment of Abrasive Waterjet Cutting Systems, International Journal of Abrasive Technology, 6:2, 158-181

Grasso, M., Pennacchi, P., Colosimo, B. M. (2014a), Empirical mode decomposition of pressure signal for health condition monitoring in waterjet cutting, The International Journal of Advanced Manufacturing Technology, 72:1-4, 347-364

Grasso, M., Colosimo B.M., Pacella, M., (2014b), Profile Monitoring via Sensor Fusion: the Use of PCA Methods for Multi-Channel Data, International Journal of Production Research, 52:20, 6110-6135

Guo, Y., Gu, Y., Zhang, Y., (2011), Invariant feature point based ICP with the RANSAC for 3D registration. Information Technology Journal, 10, 276–284

Horváth, L., & Kokoszka, P. (2012), Inference for functional data with applications (Vol. 200), Springer

James, G. M. (2007), Curve alignment by moments, The Annals of Applied Statistics, 1, 480–501

Jin J., Shi, J. (1999), Feature-Preserving Data Compression of Stamping Tonnage Information Using Wavelets, Technometrics, 41:4, 327 – 339

Jin, J., Shi, J. (2001), Automatic Feature Extraction of Waveform Signals for In-process Diagnostic Performance Improvement, Journal of Intelligent Manufacturing, 12, 257 – 268

Jolliffe I. T. (2002), Principal Component Analysis, 2nd Edition, Springer Series in Statistics

Kang, L., & Albin, S. L. (2000), On-Line Monitoring When the Process Yields a Linear Profile, Journal of Quality Technology, 32(4), 418-426

Kaziska, D. Srivastava, A. (2007), Gait-Based Human Recognition by Classification of Cyclostationary Processes on Nonlinear Shape Manifolds, Journal of the American Statistical Association, 102, 1114–1128

Ke, C., Wang, Y. (2001), Semiparametric Nonlinear Mixed-Effects Models and Their Applications (with comments and a rejoinder by the authors), Journal of the American Statistical Association, 96, 1272–1298

Kim, J., Huang, Q., Shi, J., & Chang, T. S. (2006), Online multichannel forging tonnage monitoring and fault pattern discrimination using principal curve, Journal of manufacturing science and engineering, 128(4), 944-950

Kneip, A. Gasser, T. (1992), Statistical tools to analyze data representing a sample of curves, Annals of Statistics, 20, 1266–1305

Lindstrom, M. J., and Bates, D. M. (1990), Nonlinear Mixed Effects Models for Repeated Measures Data, Biometrics, 46, 673–687

Molinari, N., Durand, J. F., & Sabatier, R. (2004), Bounded optimal knots for regression splines, Computational statistics & data analysis, 45:2, 159-178

Montgomery D. C. (2008), Introduction to Statistical Quality Control, John Wiley & Sons, 6th Ed

Mosesova, S. A., Chipman, H. A., MacKay, R. J., Steiner, S. H. (2007), Profile monitoring using mixed-effects models. BISRG Report RR-06-06, URL: http://www.bisrg.uwaterloo.ca

Noorossana R., Saghaei A., Amiri A. (2012), Statistical Analysis of Profile Monitoring, John Wiley & Sons

Okello, N., Ristic, B. (2003), Maximum likelihood registration for multiple dissimilar sensors. Aerospace and Electronic Systems, IEEE Transactions on, 39(3), 1074-1083

Paynabar K., Jin J., Pacella M. (2013), Analysis of Multichannel Nonlinear Profiles Using Uncorrelated Multilinear Principal Component Analysis with Applications in Fault Detection and Diagnosis, IEE Transactions, 45:11, 1235 – 1247

Kovacevic, R., Hashish, M., Mohan, R., Ramulu, M., Kim, T.J., Geskin, E.S. (1997), State Of The Art Of Research And Development In Abrasive Waterjet Machining, Journal of Manufacturing Science and Engineering, 119:4, 776 – 785

Ramsay J.O., Wickham H., Graves S., Hooker G. (2012), fda: Functional Data Analysis. R package version 2.4.0, URL http://CRAN.R-project.org/package=fda

Ramsay, J. O. Silverman, B. W. (2005), Functional Data Analysis, Springer New York NY, 2nd ed.

Ramsay, J. O., Li, X. (1998), Curve registration. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 60:2, 351-363.

Sangalli, L. M., Secchi, P., Vantini, S., Veneziani, A. (2009a), A case study in exploratory functional data analysis: geometrical features of the internal carotid artery, Journal of the American Statistical Association, 104:485, 37 – 48

Sangalli, L. M., Secchi, P., Vantini, S., Veneziani, A. (2009b), Efficient estimation of three-dimensional curves and their derivatives by free-knot regression splines, applied to the analysis of inner carotid artery centrelines, Journal of the Royal Statistical Society, Series C (Applied Statistics), 58:3, 285 – 306

Sangalli, L. M., Secchi, P., Vantini, S., Vitelli, V., (2010), K-mean alignment for curve clustering, Computational Statistics & Data Analysis, 54(5), 1219-1233

Sangalli, L. M., Secchi, P., Vantini, S., (2013), Analysis of AneuRisk65 data: K-mean Alignment, MOX-Report No. 45/2013, http://mox.polimi.it/it/progetti/pubblicazioni/quaderni/45-2013.pdf

Senin, N., Colosimo, B. M., Pacella, M. (2013), Point Set Augmentation Through Fitting For Enhanced ICP Registration Of Point Clouds In Multisensor Coordinate Metrology, Robotics and Computer-Integrated Manufacturing, 29(1), 39-52.

Silverman B. W. (1995), Incorporating parametric effects into functional principal components analysis, Journal of the Royal Statistical Society, Series B, 57 , 673–689

Srivastava A., Wu W., Kurtek S., Klassen E., Marron J. S. (2011), Registration of functional data using Fisher-Rao metric, arXiv preprint arXiv:1103.3817

Tang, R., Müller, H. G. (2008), Pairwise curve synchronization for functional data, Biometrika, 95(4), 875-889

Tucker, J. D., Wu, W., Srivastava, A. (2013), Generative Models for Function Data using Phase and Amplitude Separation, Computational Statistics and Data Analysis, 61, 50-66

Vantini, S. (2012), On the definition of phase and amplitude variability in functional data analysis, TEST, 21:4, 676-696

Varmuza K. and Filzmoser P. (2009), Introduction to Multivariate Statistical Analysis in Chemometrics, CRC Press, Boca Raton, FL

Woodall, W.H., Spitzner, D.J., Montgomery, D.C., Gupta, S. (2004), Using Control Charts to Monitor Process and Product Quality Profiles, Journal of Quality Technology, 36:3, 309 - 320

Zhou, S., Jin, N., Jin, J. (2005), Cycle-based Signal Monitoring Using a Directionally Variant Multivariate Control Chart System, IEE Transactions, 37, 971 - 982

Zhou, S., Shen, X. (2001), Spatially adaptive regression splines and accurate knot selection schemes, Journal of the American Statistical Association, 96:453, 247-259

Zhou, R. R., Serban, N., Gebraeel, N., Muller, H-G., (2014), A Functional Time Warping Approach to Modeling and Monitoring Truncated Degradation Signals, Technometrics, 56:1, 67:77, DOI: 10.1080/00401706.2013.805661

Appendix A: Selection of the warping function degree, *K*

An iterative procedure for the automatic selection of the most appropriate warping function degree, K, should be based on the same statistic adopted by the warping algorithm. Thus, we propose the following procedure based on the average $MINEIG(h)$ statistics, which can be computed at increasing values of K:

$$\overline{MINEIG}_K(h) = \frac{1}{M}\sum_{j=1}^{M} MINEIG_{j,K}(h), \quad K = 1,2,\dots \tag{A1}$$

The relative improvement (or worsening) produced by passing from a degree K function to a degree $K+i$ function, can be expressed as follows:

$$D_{K,i} = 100\left(\frac{\overline{MINEIG}_K(h) - \overline{MINEIG}_{K+i}(h)}{\overline{MINEIG}_K(h)}\right)\%, \tag{A2}$$

$$i = 1,2,\dots; K = 1,2,\dots$$

We propose to choose the degree K for the warping function when: (i) the relative improvement $D_{K,i}$ is lower than a given threshold (e.g., $D_{K,i} < 5\%$) for at least two consecutive steps ($i = 1,2$), or (ii) the relative change $D_{K,1}$ is negative. Notice that a negative value of $D_{K,1}$ is possible. A negative value is likely to occur for high values of K, when a too complex warping function introduces a detrimental curve distortion.

Appendix B: Brief review of the FPCA methodology

Principal Component Analysis (PCA) is a multivariate statistical technique aiming to search the directions of maximum variability (i.e., the main modes) of the dataset. In the context of statistical process control, PCA allows monitoring complex signal patterns without selecting a model in advance, thanks to the identification of a limited number of features (i.e., the principal components), which are mutually uncorrelated linear combinations of the measured profiles. In the presence of functional observations (e.g., profiles observed along time), the FPCA plays the same role as PCA for multivariate data.

Given a (possibly registered) functional dataset $\{y_j(t), t \in [0,T]\}$, $j = 1,\dots,M$, the first FPC is its direction of maximum variability. In the following, we will assume each datum y_j to be a square integrable real valued functions, and we will denote this with $y_j \in L^2([0,T])$. In this case, the first FPC is found by maximizing, over $\xi \in L^2$, with $||\xi||^2 = \int_0^T |\xi(t)|^2 dt = 1$:

$$\frac{1}{M}\sum_{j=1}^{M}\int_0^T y_j(t)\xi(t)dt \tag{B1}$$

The remaining FPCs, $\xi_i, i = 2,\dots,M$, capture the remaining modes of variability subject to be mutually orthogonal, and are thus obtained by solving problem (B1) with the additional orthogonality constraint $\int_0^T \xi_k(t)\xi_i(t)dt = 0,\ k < i$.

We recall that the sample covariance function $v(\cdot,\cdot)$ of the functional dataset can be expressed as:

$$v(s,t) = \frac{1}{M-1}\sum_{j=1}^{M}\Big(y_j(s)-\bar{y}(s)\Big)\Big(y_j(t)-\bar{y}(t)\Big), \quad (B2)$$

$$s,t \in [0,T]$$

where $\bar{y}(t),\ t \in [0,T]$, is the sample mean. The covariance operator is then defined as the kernel operator $V: L^2([0,T]) \to L^2([0,T])$, acting on $x \in L^2$ as:

$$Vx = \int v(\cdot,t)x(t)dt \quad (B3)$$

As proved, e.g., in Horvàth and Kokoszka (2012), and analogously to the multivariate case, the FPCs $\xi_i, i = 1,\dots,M$, are the eigenfunctions of the sample covariance operator V, which are obtained by solving the eigenequation:

$$V\xi_i = \rho_i\xi_i, \quad i = 1,2,\dots,M \quad (B4)$$

where ρ_i is the i^{th} eigenvalue, or, equivalently,

$$\int v(t,s)\,\xi_i(s)ds = \rho_i\xi_i(t), \quad t \in [0,T], i = 1,2,\dots,M \quad (B5)$$

We note that the eigenfunctions can be expressed on the same basis $\{\Phi_k(\cdot)\}_{k\geq 1}$ as that used in (2) for the data:

$$\xi_i(t) = \boldsymbol{b}_i'\Phi(t), \quad t \in [0,T], i = 1,2,\dots,M \quad (B6)$$

$\boldsymbol{b}_i \in \mathbb{R}^{Q+L-1}$ being the vector of coefficients corresponding to i^{th} the eigenfunction. Hence, equation (B5) becomes:

$$\frac{1}{M-1}\Phi'(t)(\mathbf{C}-\bar{\mathbf{C}})'(\mathbf{C}-\bar{\mathbf{C}})\int \Phi(s)\Phi'(s)ds\,\boldsymbol{b}_i = \boldsymbol{\rho}\Phi'(t)\boldsymbol{b}_i, i = 1,2,\dots,M \quad (B7)$$

where $\bar{\mathbf{C}}_{\cdot i} = \bar{\boldsymbol{c}}$ for $i = 1,\dots,M$, if $\bar{y}(t) = \bar{\boldsymbol{c}}'\Phi(t)$, $t \in [0,T]$. Expression (B7) reduces to the symmetric eigenequation:

$$\frac{1}{M-1}\mathbf{J}^{\frac{1}{2}}(\mathbf{C}-\bar{\mathbf{C}})'(\mathbf{C}-\bar{\mathbf{C}})\mathbf{J}^{\frac{1}{2}}\boldsymbol{u}_i = \rho\boldsymbol{u}_i, \quad i = 1,2,\dots,M \quad (B8)$$

subject to $\boldsymbol{u}_i'\boldsymbol{u}_i = 1$, with $\mathbf{J} = (\int \Phi_k(s)\Phi_l'(s)ds) \in \mathbb{R}^{Q+L-1,Q+L-1}$ and having defined $\boldsymbol{u}_i = \mathbf{J}^{1/2}\boldsymbol{b}_i \in \mathbb{R}^{Q+L-1}$, $i = 1,2,\dots,M$. This allows using standard software tools to find the eigenvectors $\boldsymbol{u}_i$ of the modified eigenequations (B8), eventually back-transforming the result to obtain the coefficients $\boldsymbol{b}_i$ of the target eigenfunctions ξ_i, for $i = 1,\dots,M$.

The dimensionality reduction task is then accomplished by retaining the first $m \ll M$ FPCs, where m is such that $\xi_1, \xi_2, \dots, \xi_m$ explain a sufficient percentage of the original data variability. For this purpose, one may proceed analogously to the classical multivariate PCA, e.g., by associating a threshold to the cumulative explained variance or to the eigenvalues (e.g., the 80%). When the first m principal components are retained, each profile can be expressed as:

$$\hat{y}_j(t,m) = \bar{y}(t) + \sum_{i=1}^{m} f_{j,i}\xi_i(t), \quad j = 1,2,\dots,M \quad (B9)$$

where:

$$f_{j,i} = \int_0^T (y_j(t)-\bar{y}(t))\xi_i(t)dt = (\boldsymbol{C_j}-\bar{\boldsymbol{c}})'\mathbf{J}\boldsymbol{b}_j \quad (B10)$$

for $j = 1,\dots,M, i = 1,\dots,m$, are the FPC scores. The scores $f_{j,i}$ can be then monitored to detect any unnatural behaviour that affects the amplitude variability of the observed profiles along the first m principal directions.

Appendix C: On the correlation between FPCs and warping coefficients

Fig. C1 shows the scatterplot of the first 3 FPC scores, i.e., $\{f_{j,1}, f_{j,2}, f_{j,3}, j = 1,\dots,50\}$ and the 3 warping coefficients $\{w_{j,1}, w_{j,2}, w_{j,3}, j = 1,\dots,50\}$ of 50 in-control profiles generated in Scenario B. Fig. C1 shows that a non-zero correlation exists between some FPCs and warping coefficients. This justifies the choice of jointly monitoring phase and amplitude coefficients via a Hotelling's T^2 statistic.

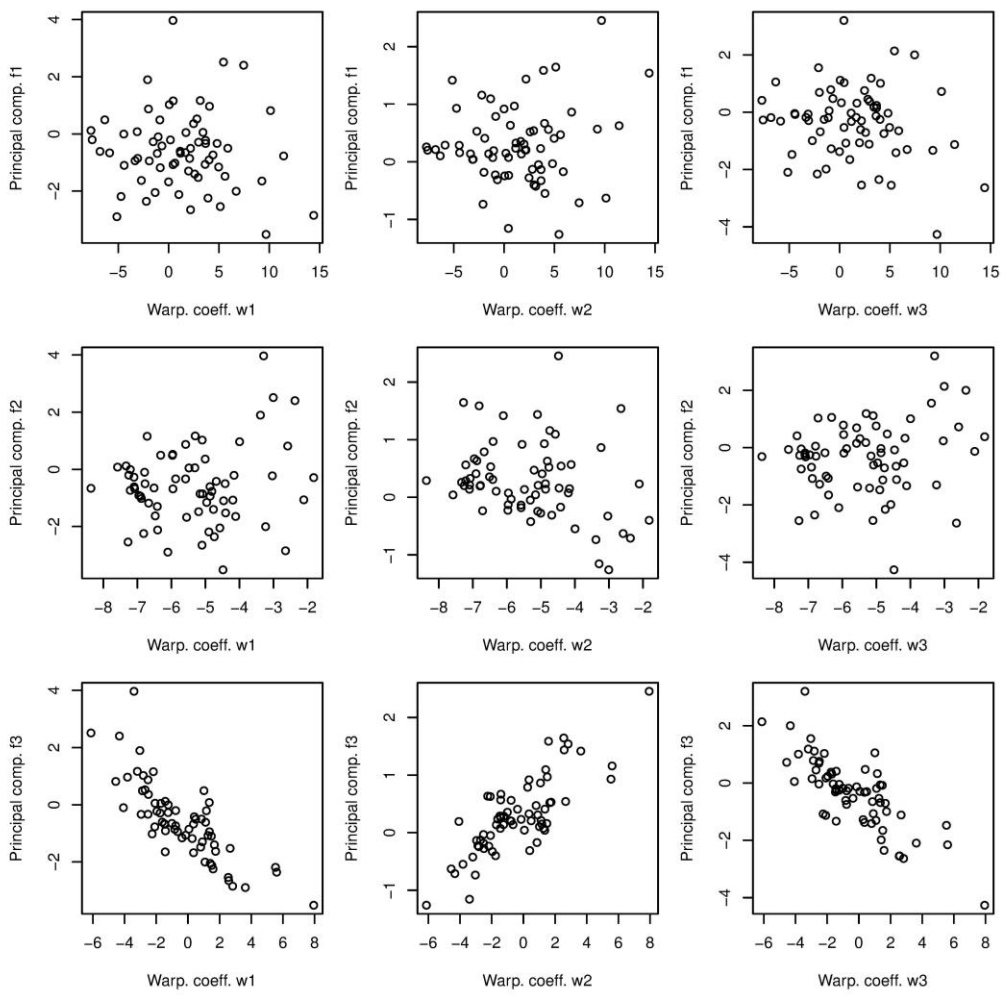

FIGURE C1 – Scatterplot of first three FPCs vs warping coefficients, Scenario B, $M = 50$