



Robust profile alignment based on penalised-spline smoothing

Yangyang Zang & Kaibo Wang

To cite this article: Yangyang Zang & Kaibo Wang (2019) Robust profile alignment based on penalised-spline smoothing, International Journal of Production Research, 57:10, 2966-2983, DOI: 10.1080/00207543.2018.1519263

To link to this article: <https://doi.org/10.1080/00207543.2018.1519263>



Published online: 17 Sep 2018.



Submit your article to this journal [↗](#)



Article views: 152



View related articles [↗](#)



View Crossmark data [↗](#)

Robust profile alignment based on penalised-spline smoothing

Yangyang Zang and Kaibo Wang*

Department of Industrial Engineering, Tsinghua University, Beijing 100084, People's Republic of China

(Received 17 October 2017; accepted 23 August 2018)

In quality engineering practice, profiles that are used for quality monitoring or evaluation are sometimes unaligned due to engineering constraints. In such cases, profiles have to be registered (aligned) through shifting, time warping or coordinate alignment such that samples are comparable and easy to handle. Among the different registration algorithms, time warping, or alignment of profiles with unequal lengths, is a challenging task. In quality engineering, a typical phenomenon observed in profile alignment is that neighbours of an aligned pair have a high possibility of being similar, which means that a large jump in a warping path is less likely. In this article, a penalised-spline smoothing method is proposed for profile alignment to handle this problem. The newly proposed nonparametric alignment strategy attempts to capture the smoothness and spatially correlated features of warping shifts, and is proven more robust than existing algorithms. A dynamic programming algorithm is developed to obtain the optimal path. Both simulation studies and a real example demonstrate the effectiveness of the proposed method.

Keywords: dynamic time warping; quality control; profile alignment; profile monitoring; robust method; statistical process control

1. Introduction

Due to the rapid development of sensing and information technology, signal profiles that characterise the functional trajectory of interested quality or process variables have been widely used in quality engineering practice for process monitoring or analysis. Over the past years, statistical analyses of profile data have been attracting increasingly more attention. In many manufacturing applications, profile monitoring has been widely investigated to detect the unusual variability and identify the stability of processes. Over the past decade, there has been increasing research interest in profile monitoring procedures, which subsequently provide an effective approach for fault diagnosis, variability reduction and capability improvement of processes.

Usually, raw profile data collected from industrial applications possess two types of variability: amplitude variation and phase variability (Ramsay and Silverman 2005; Grasso et al. 2016). Amplitude variability, also called vertical variability, occurs when profiles are different only in the amplitude direction. For example, in Figure 1(a), profiles 1 and 2 differ in vertical amplitude only. Conversely, phase variability represents the differences between two profiles along the phase direction, such as the difference between profiles 1 and 3 in Figure 1(a) represents.

In the literature, a wide collection of profile monitoring methods have been seen focusing on the detection of amplitude variability. For profiles that can be characterised by particular functional forms, parameters of the functions are usually used for profile monitoring. For example, Kang and Albin (2000) and Jensen, Birch, and Woodall (2008) proposed the monitoring of linear coefficients to detect profile changes, Gupta, Montgomery, and Woodall (2006) compared the performance of charts when monitoring linear calibration profiles. Jensen and Birch (2009) monitored parameters in a nonlinear model to trigger process changes. In cases that profiles are too complex to be characterised by single functional forms, nonparametric approaches can be utilised. For example, Chicken, Pignatiello, and Simpson (2009) investigated the monitoring of profiles of military radar signatures, which is formulated by the change-point wavelet, Walker and Wright (2002) and Williams, Woodall, and Birch (2007) studied the monitoring of vertical density profiles (VDP) of particle boards using spline smoothing model. Lee et al. (2012) proposed the integrated use of wavelet transformation and a CUSUM chart when monitoring complex signals. The recent progress in the area of profile monitoring is referred to Colosimo and Pacella (2010), Woodall (2007), Noorossana, Saghaei, and Amiri (2011) and Qiu (2013). However, it should be noted that in most existing

*Corresponding author. Email: kbwang@tsinghua.edu.cn

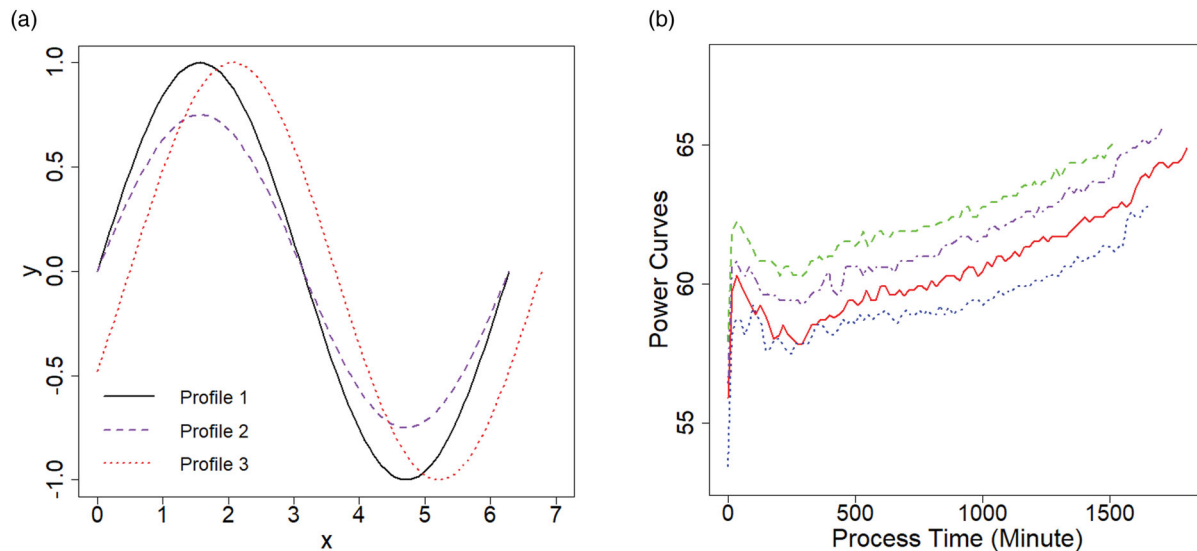


Figure 1. (a) Illustrative examples of profile variability. Phase variability: differences between solid curve and dotted curve; vertical variation: differences between solid curve and dashed line. (b) Real examples of four heating power profiles in an ingot growth process.

profile monitoring researches, profiles are sampled using the same observing epoch and have an equal length, which means that these profiles are well aligned and have no phase variability.

However, in some engineering processes, phase variability does exist. The existence of phase variability results in the irregular expansion of the overall variability of profiles, and consequently increases the difficulty and complexity of fault detection; profiles with phase variability such as unequal length cannot be monitored by the aforementioned algorithms directly. Therefore, it is important to recognise phase variation in profiles and remove it through alignment operations.

Take heating power curves collected from an ingot growth process as an illustrative example, as shown in Figure 1(b). Due to the different amounts of raw materials used in each production run and uncertain environmental noises during long production cycles, the manufacturing production processes are not repeated exactly and identical features of profiles do not occur at the same moments or locations. Therefore, these collected curves exhibit similar features but with different phases and lengths.

In the literature, we can see many applications in which alignment operations are needed before data analysis. For example, in the applications of monitoring multiple profiles by Zhang et al. (2018), due to the inherent fluctuations in the fabrication process, signal profiles from a reaction chamber to measure key process variables had different lengths, and exhibited un-synchronisation for profile features. As the automatically controlling of the system, torque profiles (Grasso et al. 2016) during an M8 threads-tapping operation on a mild steel part are misaligned, that is, signals start and end at different times. Clearly, to monitoring these profiles, the phase shifts have to be removed so that abnormal profiles can be correctly identified.

Profile alignment, also known as curve registration in the statistics area, is a technique for addressing profile misalignment, and it distinguishes these two types of variability through registration transformations (Ramsay and Silverman 2005; Tucker, Wu, and Srivastava 2013; Gervini and Carter 2014; Panaretos and Zemel 2016). Different approaches have been proposed to achieve profile alignment. Shift registration aligns profiles through a pure transformation along the phase direction (Grasso et al. 2016). Landmark registration aligns profiles by taking important features (such as maximal points, minimal points) as landmarks and estimating other alignment parts through linear interpolation. In some cases, landmarks are difficult to define when features are not obvious or even missing.

To achieve a global optimal alignment path, dynamic time warping (DTW), proposed by Berndt and Clifford (1994), has been widely studied over the past two decades. DTW-based methods estimate warping functions by minimising the sum of distances of all points between two series, such as derivative DTW investigated by Keogh and Pazzani (2001) to improve the alignment performance on profile features; weighted DTW presented by Jeong, Jeong, and Omiaomu (2011), which takes into account the information of the phase difference between two points to be aligned to improve the accuracy of curve classification and clustering; and pairwise DTW presented by Arribas-Gil and Müller (2014) to extend the basic method to pairwise alignment maps. Although these DTW-based methods are useful for addressing more complex profile alignment cases, the paths are sensitive to spikes, shifts or jumps in profiles, which are common features of abnormal profiles in quality engineering.

Rather, smooth monotone registration, investigated by Ramsay and Li (1998) and extended by Grasso et al. (2016), is a nonparametric approach to reduce the impact of random measurement noises in profiles by adding a second-order penalty. The smooth monotone registration methods could obtain a smooth continuous registration path; however, the method only considers the situation when profiles have an equal length, and has limitations in aligning profiles with different lengths or incongruous warping sizes over time, which is inappropriate in some applications, including the heating power profiles illustrated above.

Figures 2 and 3 compares the above typical alignment methods using two simulated profile curves. In the warping paths shown in Figure 2, the dashed lines are the ideal warping path, which is the true path calculated from the simulated functions directly for comparison. The solid lines represent the estimated warping paths by various methods. Figure 3 displays the alignment pairs in the given two profiles. The upper line is the profile to be aligned, while the lower line is the reference, and oblique lines between them indicate the alignment pairs, which link points on the unaligned profile to their counterparts on the reference. It is obvious from Figures 2 and 3 that shift registration gives a constant movement along phase direction; landmark registration finds a more accurate path, but the path is not smooth in space; registration by DTW severely fluctuates along the ideal path due to the influence of noises or shifts; and smooth monotone registration gives poor the performance near the right end of the profile.

Hence, the purpose of this work is to propose a novel and robust profile alignment algorithm. An important feature of real profiles collected from engineering processes is that they are generally contaminated by noises, process shifts or failures.

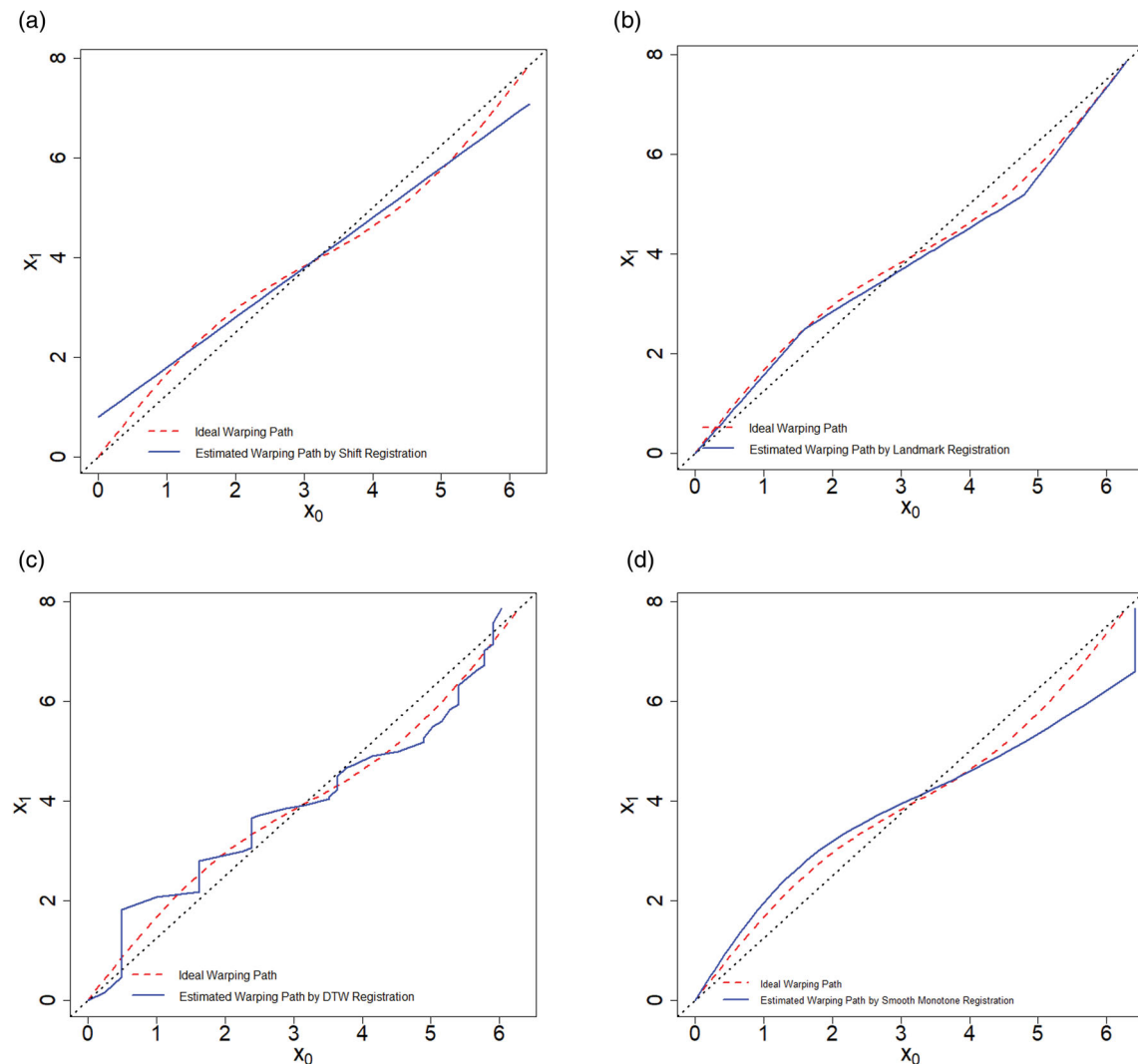


Figure 2. Example of warping paths estimated by (a) shift registration, (b) landmark registration, (c) DTW, and (d) smooth monotone registration.

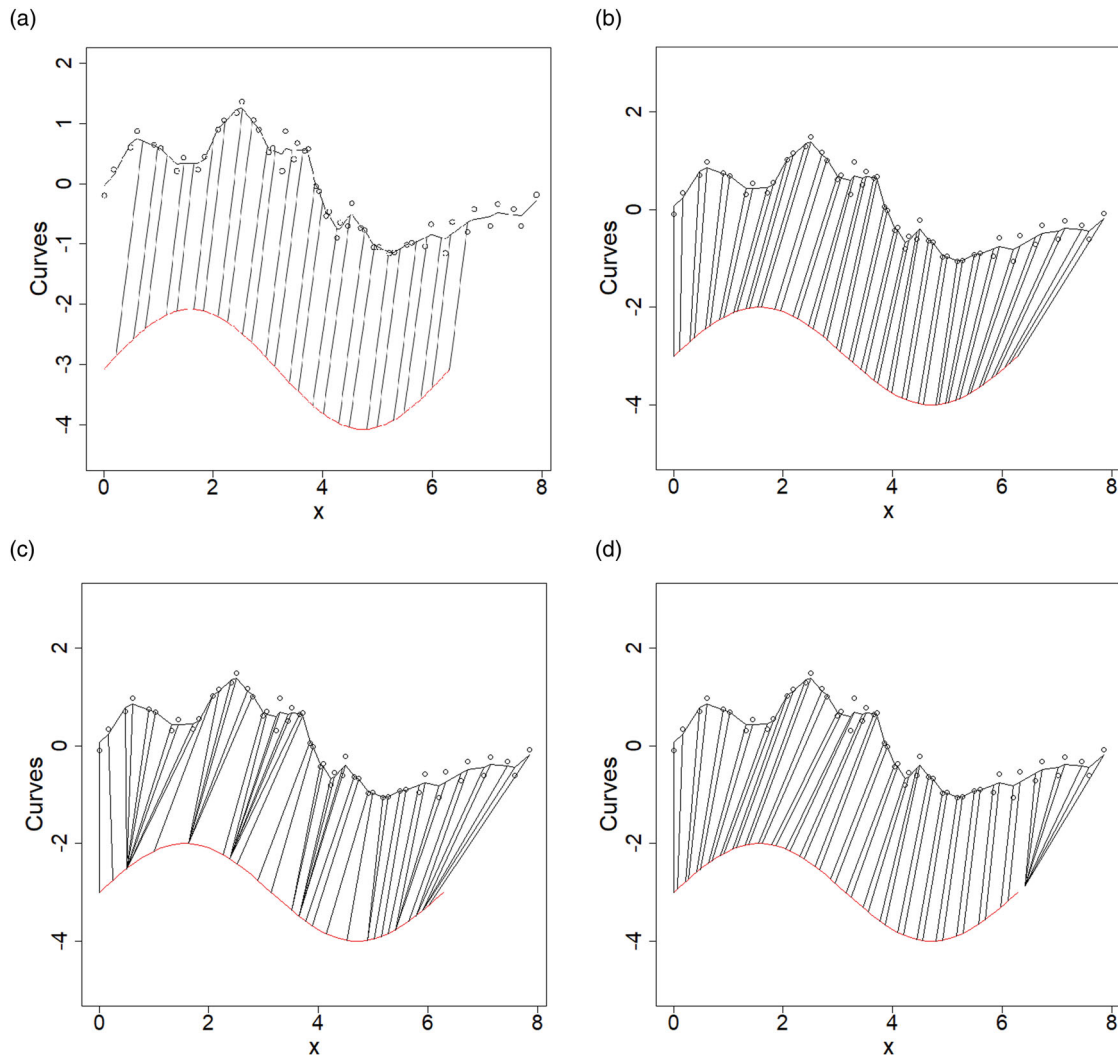


Figure 3. Example of alignment pairs in the given two profiles estimated by (a) shift registration, (b) landmark registration, (c) DTW, and (d) smooth monotone registration.

When such profiles are aligned and used for process monitoring, conventional alignment algorithms are easily misled by contaminated signals. However, it is observed that such engineering signals, although misaligned, follow similar trends. Therefore, the phase shifts in the alignment path should be relatively smooth because the shift distance of one point should be close to the shift distances of its neighbours. To guarantee the smoothness of the alignment path and phase shifts, this paper proposes limiting the path by applying a penalised distance measure. The penalty is focused on the relative distances of neighbouring phase shifts. The proposed method is expected to be more robust than existing methods, and it also better fits with the engineering implication behind the real signals.

This paper has two main contributions for the unaligned free-form profiles. First, a penalised-spline registration method is proposed to transform measuring profiles, where the new strategy attempts to capture the smooth and spatially correlated features of warping shift during alignment. The designed registration strategy is more effective and robust in applications where there are noises, mean shifts of profiles. Second, a dynamic programming algorithm is accepted to solve the optimal path. The new alignment algorithm attempts to estimate the aligning parameters in a nonparametric manner.

The remainder of this paper is organised as follows. In Section 2, we present the framework used for curve registration. The curve registration model and assumption are first introduced, and then the penalised-spline method is proposed for profile alignment. Next, we discuss the dynamic programming strategy with penalty-based alignment, and then we propose our strategy for parameter estimations. In Section 3, we present the performance of our curve registration method through simulations. In Section 4, a real example is presented to demonstrate the use of the penalised alignment procedures. Finally, we conclude this work and make suggestions for future research in Section 5.

2. Profile alignment on penalised-spline approach

2.1. Model description

Let (x_{ij}, y_{ij}) be the j th observation of the i th profile under consideration for $j = 1, 2, \dots, N_i$ and $i = 1, 2, \dots$. Then, a nonparametric model for describing the observed data is

$$y_{ij} = g_i(h_i(x_{ij})) + \varepsilon_{ij}, \quad \text{for } j = 1, 2, \dots, N_i, \quad i = 1, 2, \dots, \quad (1)$$

where g_i is the true regression function for describing the i th profile curve and ε_{ij} s are random errors with mean 0 and variance σ^2 . Let $g_0(x_0), x_0 \in [0, x_{N_0}]$ be the regression function for the reference profile, which is considered to be the baseline in profile registration. $h_i(x)$ is a warping function that transforms an observation x to a warped x , and it makes certain profile features occurring on observation x_i of profile i match with their counterparts occurring on a warped x of the baseline by stretching or contracting x , that is, $h_i(x) = x_0$. Thus, warping function $h_i(x)$ eliminates the phase shift between the i th profile and baseline on phase point x . In general, two boundary conditions of $h_i(x)$ need to be satisfied, that is, $h_i(0) = 0$ and $h_i(x_{iN_i}) = x_{N_0}$ for starting and ending points, respectively. Under the assumption that the warping path should be continuous and monotonic, a local constraint should be added to Equation (1) that $h_i(x) \leq h_i(x + t), \forall t \geq 0$. Model (1) can be considered as a general regression model to describe profiles along with profile registration, and many known registration algorithms can be considered to predict the warping function $h_i(x)$ based on comparing the regression functions g_i and g_0 estimated by the observed data of profile i and the baseline.

In this part, we assume that profile g_i and warping function h_i are continuous in the application; thus, the Nadaraya-Watson local constant kernel (LCK) smoothing procedure (Qiu 2005) can be applied to estimate $g_i(h_i(x))$. For a given location x , we consider its symmetrical neighbourhood $O(x) = \{u : |u - x| \leq l\}$, where $l > 0$ is a bandwidth parameter that needs to be predetermined. Then, according to Model (1), the LCK estimator of $g_i(h_i(x))$, denoted as $\hat{f}_i(x)$, is defined as

$$\hat{f}_i(x) = \frac{\sum_{x_{ij} \in O(x)} y_{ij} K\left(\frac{x_{ij} - x}{l}\right)}{\sum_{x_{ij} \in O(x)} K\left(\frac{x_{ij} - x}{l}\right)}, \quad \text{for } i = 1, 2, \dots, \quad (2)$$

where K is a one-dimensional symmetric kernel function.

The estimator of warping function $h_i(x)$ can be obtained by functional least squares by minimising the squared distance of two profiles as

$$\min_{h_i} \left\{ \int_{x_{i1}}^{x_{iN_i}} [\hat{f}_i(x) - g_0(h_i(x))]^2 dx \right\}. \quad (3)$$

In Model (3), x is the independent variable of profile i before alignment, while $h_i(x)$ is its counterpart in reference profile g_0 . Model (3) means the squared distance of points on profile g_i and their counterparts on profile g_0 .

The warping function transforms an observation x to a warped x , and it makes certain profile features occurring on observation x_i of profile i match with their counterparts occurring on a warped x of the baseline. The warping function is applied to express the phase variability between profile i and the baseline. In traditional profile monitoring procedures (such as Xu et al. 2012; Paynabar, Zou, and Qiu 2016; Li et al. 2018), the assumption is that there is no phase variability between two profiles, that is, profile features on profile i and their counterparts on the baseline occur on the same x , thus the warping function is $h_i(x) = x_0 = x$.

However, when phase variability does exist, the warping function can be used to stretch or contract x to match certain profile features. Thus, $h_i(x)$ has a phase shift on x , and it is no longer equivalent to x everywhere. Generally, different parts of a profile have various degrees of stretching or contracting to align different features. Assume that the phase shift of x is $\delta(x)$, which represents the degree of stretching or contracting of x corresponding to the baseline profile; then, the warping function is defined as

$$h_i(x) = x + \delta(x). \quad (4)$$

Therefore, predicting the warping function $h_i(x)$ in Model (1) can be considered as the problem of parameter estimation for $\delta(x)$ in Model (4). To simplify the expression, in the following context, assume that δ_{ij} is the symbol of $\delta(x_{ij})$. Because of the conditions of $h_i(x)$, there are global and local constraints for Equation (4) derived from Model (1) as follows.

Global constraints: The two ends of the profile and reference should be aligned, respectively:

$$\delta_{i1} = 0; \delta_{iN_i} = x_{N_0} - x_{iN_i}.$$

Local constraints: The warping function should be monotonic:

$$\delta(x + t) - \delta(x) \geq -t, \quad \forall t \geq 0.$$

The warping function Model (4) along with these two constraints is a general form that can be used to represent any alignment functions. In this function, the core part is the $\delta(x)$ term, which reflects the mapping relationship between the unaligned and the baseline profiles. Specifically, if there is no phase variability, the phase shift $\delta(x)$ is 0. For several periodic processes (such as Rønn 2001), x is transformed by individual rigid shifts, and the phase shift $\delta(x)$ is a constant. Olsen, Markussen, and Rakêt (2018) extended Model (4) by treating $\delta(x)$ as random effects, and applying a Gaussian process model to estimate latent warp variables $\delta(x)$ at specified anchor points.

In practice, we use the sampling points for computing the integration in Model (3). However, in applications, when sampling points are not ideal, applicators can discretise the related functions to be equally spaced, or even unequally spaced if needed, covering the entire space of x for profile i . Then, we can consider the following discretised version of (3):

$$\min_{\delta_{i1}, \delta_{i2}, \dots, \delta_{iN_i}} \left\{ \sum_{j=1}^{N_i} [\widehat{f}_i(x_{ij}) - g_0(x_{ij} + \delta_{ij})]^2 \right\}, \quad \text{for } i = 1, 2, \dots \quad (5)$$

There are N_i parameters in Model (5). This path can be found by dynamic programming (Keogh and Pazzani 2001; Keogh 2002; Jeong, Jeong, and Omitaomu 2011, and so on).

2.2. Curve registration with adaptive spline penalty

The warping path of Model (5) is discrete and stochastic because of being completely motivated by the amplitudes of the global distances of the two profiles. However, in general, the alignment of subjects will not be satisfactory for the situation in which profiles are continuous, and smooth warping functions are generally a better choice in this application. Ramsay and Li (1998) introduced estimating the warping function h by adding an penalised squared error to the cost function of Model (3)

$$\min_{h_i} \left\{ \int_{x_{i1}}^{x_{iN_i}} [\widehat{f}_i(x) - g_0(h_i(x))]^2 dx + \lambda \int_{x_{i1}}^{x_{iN_i}} \left[\frac{h_i''(x)}{h_i'(x)} \right]^2 dx \right\}, \quad (6)$$

where $h_i'(x)$ and $h_i''(x)$ are the first derivative and second derivative, respectively, of warping function $h_i(x)$. Then they developed a nonparametric curve registration approach by representing the penalty part by a linear combination of B -spline bases, and hence, they expressed the warping function in a closed form.

Although the smoothness and monotonicity of the warping function could be achieved by this method, it does not consider the correlation within the warping path. In fact, in the application of curve registration, it is reasonable that a large time shift of one point probably leads to a relatively large time shift of its neighbours, and a tiny warping scale also seldom appears among sharp warping regions. In other words, warping value $\delta(x)$ in Model (4) is spatially correlated or time-dependent, and large jumps of warping shift need to be avoided.

To impose the correlation of $\delta(x)$ and choose a more reasonable warping path, inspired by Guo et al. (2016), Simpkin and Newell (2013), and Eilers and Marx (1996), a second-order difference penalty acting on shift value $\delta(x)$ is considered to be added to the cost function in Model (5),

$$\begin{aligned} \min_{\delta_{i1}, \delta_{i2}, \dots, \delta_{iN_i}} & \left\{ \sum_{j=1}^{N_i} [\widehat{f}_i(x_{ij}) - g_0(x_{ij} + \delta_{ij})]^2 \right. \\ & \left. + \lambda \sum_{j=1}^{N_i-2} \left[\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1} - \delta_{i,j}}{x_{i,j+1} - x_{i,j}} \right]^2 \right\}, \quad \text{for } i = 1, 2, \dots, \quad (7) \\ S.T. & \quad \delta_{i1} = 0; \\ & \quad \delta_{iN_i} = x_{N_0} - x_{iN_i} \\ & \quad \delta(x + t) - \delta(x) \geq -t, \quad \forall t \geq 0. \end{aligned}$$

where the smoothing parameter $\lambda \geq 0$ controls the roughness and correlation of the warping shift. If λ is large, then the alignment path will generally be very smooth. Additionally, warping shift coefficients λ encourage smoothness and a higher

correlation in the spatial or temporal region; in other words, a jump of transformation value $\delta(x)$ is preferred to fade away. However, if $\lambda = 0$ and the warping path is under none of the added constraints, then the profile alignment will fluctuate in space/time and thus have high variance. Note that the optimality in Model (7) is with respect to the global and local constraints. The second term of Model (7) is penalising the L_2 norm of the discrete version on the differences between neighbouring normalised difference of warped parameters, and the normalised form is chosen because of the potential unequal sampling. Model (7) can be viewed as adding the following constraint of warping parameters to Model (5),

$$\sum_{j=1}^{N_i-2} \left[\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1} - \delta_{i,j}}{x_{i,j+1} - x_{i,j}} \right]^2 \leq s, \tag{8}$$

where $s \geq 0$ is a predetermined parameter. Thus, from Model (8), the optimal warping path needs to meet the requirement that the sum of the square differences of adjoining regularised differences of phase shifts is less than s . In fact, the objective function in Model (7) can be considered as a Lagrangian function of Model (5) along with Equation (8), and λ is the Lagrangian multiplier.

The penalised registration is designed to overcome the three difficulties that we mentioned previously. First, it attempts to capture the smoothly varying and spatially/temporally correlated features of warped values when registering two profiles. This penalised method is straightforward and has attractive applicability for providing smoothness of warped values while maintaining a discrete cost function with a conceivable unequal sampling strategy. Second, penalty conjugating with local constraints guarantees the need for the simultaneous smoothness and monotonicity of the warping path. Penalization in this model ensures the smoothness of phase shift δ , which also yields the smoothness of warping path h , while the local constraints that Model (7) are subject to ensure the monotonicity of the warping path. Third, it is computationally easy since the cost function with the L_2 penalty term can easily be demonstrated by a recursion process; thus, we can make full use of the dynamic programming algorithm to align profiles.

2.3. Warping path estimation for curve registration

In this part, we discuss estimating the transformations $\delta(x)$ along the warping path $h_i(x)$ on Model (7). The derivation of the dynamic programming algorithm is investigated first. Although the recurrence formula is straightforward, it is reasonable and necessary to show the implementation process in the application.

First, regarding the global constraint for the ending points, the objective function in Model (7) is equivalent to

$$\begin{aligned} & \min_{\delta_{i,N_i-1}} \left\{ \left[\widehat{f}_i(x_{i,N_i-1}) - g_0(x_{i,N_i-1} + \delta_{i,N_i-1}) \right]^2 \right. \\ & + \min_{\delta_{i1}, \dots, \delta_{i,N_i-2}} \left\{ \sum_{j=1}^{N_i-2} \left[\widehat{f}_i(x_{ij}) - g_0(x_{ij} + \delta_{ij}) \right]^2 \right. \\ & \left. \left. + \lambda \sum_{j=1}^{N_i-2} \left[\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1} - \delta_{i,j}}{x_{i,j+1} - x_{i,j}} \right]^2 \right\} \right\}, \tag{9} \end{aligned}$$

which is defined as $\Gamma_{N_i}(\delta_{iN_i})$.

Then, the second minimisation part is defined as $\Gamma_{N_i-1}(\delta_{iN_i}, \delta_{i,N_i-1})$, which represents a function of shift parameters δ_{iN_i} and δ_{i,N_i-1} on the parameter spaces of $\delta_{i1}, \dots, \delta_{i,N_i-2}$,

$$\begin{aligned} & \Gamma_{N_i-1}(\delta_{iN_i}, \delta_{i,N_i-1}) \\ & = \min_{\delta_{i1}, \dots, \delta_{i,N_i-2}} \left\{ \sum_{j=1}^{N_i-2} \left[\widehat{f}_i(x_{ij}) - g_0(x_{ij} + \delta_{ij}) \right]^2 + \lambda \sum_{j=1}^{N_i-2} \left[\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1} - \delta_{i,j}}{x_{i,j+1} - x_{i,j}} \right]^2 \right\} \\ & = \min_{\delta_{i,N_i-2}} \left\{ \left[\widehat{f}_i(x_{i,N_i-2}) - g_0(x_{i,N_i-2} + \delta_{i,N_i-2}) \right]^2 + \left[\frac{\delta_{iN_i} - \delta_{i,N_i-1}}{x_{iN_i} - x_{i,N_i-1}} - \frac{\delta_{i,N_i-1} - \delta_{i,N_i-2}}{x_{i,N_i-1} - x_{i,N_i-2}} \right]^2 \right\} \end{aligned}$$

$$\begin{aligned}
 &+ \min_{\delta_{i1}, \dots, \delta_{i, N_i-3}} \left\{ \sum_{j=1}^{N_i-3} [\widehat{f}_i(x_{ij}) - g_0(x_{ij} + \delta_{ij})]^2 \right. \\
 &\left. + \lambda \sum_{j=1}^{N_i-3} \left[\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1} - \delta_{ij}}{x_{i,j+1} - x_{ij}} \right]^2 \right\}. \tag{10}
 \end{aligned}$$

Thus, the minimisation is further iterated, and functions $\Gamma_{N_i-2}(\delta_{i, N_i-1}, \delta_{i, N_i-2})$, $\Gamma_{N_i-3}(\delta_{i, N_i-2}, \delta_{i, N_i-3})$, \dots , $\Gamma_2(\delta_{i3}, \delta_{i2})$ are defined similarly. Furthermore, the function $\Gamma_j(\delta_{i,j+1}, \delta_{ij})$, for $j \in [2, N_i - 1]$ is a function of shift parameters $\delta_{i,j+1}$ and δ_{ij} on the parameter spaces of $\delta_{i1}, \dots, \delta_{i,j-1}$. Meanwhile, the recursion function of $\Gamma_j(\delta_{i,j+1}, \delta_{ij})$ is implemented as

$$\begin{aligned}
 \Gamma_{j+1}(\delta_{i,j+2}, \delta_{i,j+1}) = \min_{\delta_{ij}} &\left\{ [\widehat{f}_i(x_{ij}) - g_0(x_{ij} + \delta_{ij})]^2 \right. \\
 &\left. + \left[\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1} - \delta_{ij}}{x_{i,j+1} - x_{ij}} \right]^2 + \Gamma_j(\delta_{i,j+1}, \delta_{ij}) \right\}. \tag{11}
 \end{aligned}$$

Note that the optimality of δ_{ij} for Model (11) should be under its parameter space, which is $\Omega_{ij} = \{\delta_{ij} : \delta_{ij} \in [x_{i1} - x_{ij}, \delta_{i,j+1} + (x_{i,j+1} - x_{ij})]\}$ deviated from local constraints. For $j = N_i - 1$, the recursion function is revised as

$$\Gamma_{N_i}(\delta_{iN_i}) = \min_{\delta_{i, N_i-1}} \left\{ [\widehat{f}_i(x_{i, N_i-1}) - g_0(x_{i, N_i-1} + \delta_{i, N_i-1})]^2 + \Gamma_{N_i-1}(\delta_{i, N_i}, \delta_{i, N_i-1}) \right\}.$$

Assume that $\widehat{\delta}_{ij}$ is the locally optimal warping shift value for x_{ij} for Model (11), which is a variable determined by its following neighbouring parameters $\delta_{i,j+1}$ and $\delta_{i,j+2}$,

$$\widehat{\delta}_{ij} = \arg \min_{\delta_{ij}} \left(\Gamma_{j+1}(\delta_{i,j+2}, \delta_{i,j+1}) \right) \quad \text{for } j = 1, 2, \dots, N_i - 2. \tag{12}$$

Then, the variable $\widehat{\delta}_{ij}$ plays a role in the backward pass of the program; in other words, optimal sequences of $\delta(x)$ are determined through a recursion from Model (12). Thus, the global optimal phase shifts are obtained in a final backtrace. To this end, we consider using the following iterative searching algorithm.

- Set the initial value of Γ to be $\Gamma_1(\delta_{i2}, \delta_{i1}) = 0$ because of the starting alignment condition.
- In the j th iteration, $j = 1, 2, \dots, N_i - 2$, for the k th combination of $(\delta_{i,j+2}, \delta_{i,j+1})$ from their parameter spaces, marked as $(\delta_{i,j+2}^k, \delta_{i,j+1}^k)$, evaluate the objective function

$$\left[\widehat{f}_i(x_{ij}) - g_0(x_{ij} + \delta_{ij}) \right]^2 + \left[\frac{\delta_{i,j+2}^k - \delta_{i,j+1}^k}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1}^k - \delta_{ij}}{x_{i,j+1} - x_{ij}} \right]^2 + \Gamma_j(\delta_{i,j+1}^k, \delta_{ij})$$

at possible values of δ_{ij} from space Ω_{ij} with the component obtained in the previous iteration. The minimum value is denoted as $\Gamma_{j+1}(\delta_{i,j+2}^k, \delta_{i,j+1}^k)$. Note that there is an inner loop procedure to compute every potential combination of $(\delta_{i,j+2}, \delta_{i,j+1})$ because of the comparisons of the objective functions at various δ_{ij} , which illustrates the functional relationship of $\Gamma_{j+1}(\delta_{i,j+2}, \delta_{i,j+1})$.

- The iteration terminates when $j = N_i - 1$, and at this step, the objective function is evaluated by

$$\left[\widehat{f}_i(x_{i, N_i-1}) - g_0(x_{i, N_i-1} + \delta_{i, N_i-1}) \right]^2 + \Gamma_{N_i-1}(\delta_{i, N_i}, \delta_{i, N_i-1})$$

at possible values of δ_{i, N_i-1} from space Ω_{i, N_i-1} . The optimal value is marked as $\Gamma_{N_i}^*(\delta_{i, N_i})$, which is the global minimisation of the penalised distance of two profiles.

- Backtrack to obtain the optimal warping parameter sequences $\delta^*(x)$. Start at δ_{i, N_i-1}^* , where the penalised objective function reaches its minimal value $\Gamma_{N_i}^*(\delta_{i, N_i})$; then, for $j = N_i - 2, N_i - 3, \dots, 1$, δ_{ij}^* is the values that make function $\Gamma_{j+1}(\delta_{i,j+2}^*, \delta_{i,j+1}^*)$ reach its optimised form, $\Gamma_{j+1}^*(\delta_{i,j+2}^*, \delta_{i,j+1}^*)$.

Although the implementations of this algorithm will yield excellent performance for random profiles, the computational time is mainly relevant to the profile sequence's length N_i and searching scale space that determines the scale of k in the iteration procedure. In the numerical examples presented in the next section, if we assume an equal sampling interval, which is denoted as $d = x_{i,j+1} - x_{ij}$, $d > 0$, and set the search interval as $d/2$, then the asymptotic time complexity of this search procedure is $O(n^3)$.

2.4. Remarks regarding penalised-spline alignment

In this section, we present several remarks regarding our proposed penalised-spline alignment method.

If d is relatively small, that is, $d \rightarrow 0$, where d has the same definition as in the previous section, then

$$\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} \approx \delta'_{i,j+1}|_{x=x_{i,j+1}};$$

$$\frac{\delta_{i,j+1} - \delta_{ij}}{x_{i,j+1} - x_{ij}} \approx \delta'_{ij}|_{x=x_{ij}}.$$

Assuming that $\lambda' = \lambda d$, the approximation of the penalty part of the proposed registration method is

$$\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1} - \delta_{ij}}{x_{i,j+1} - x_{ij}} \approx \delta''_{ij}|_{x=x_{ij}}.$$

Under this assumption, the penalised-spline alignment in Model (7) is approximated as

$$\min_{\delta_{i1}, \delta_{i2}, \dots, \delta_{iN_i}} \left\{ \sum_{j=1}^{N_i} [\hat{f}_i(x_{ij}) - g_0(x_{ij} + \delta_{ij})]^2 + \lambda' \sum_{j=1}^{N_i-2} (\delta''_{ij}|_{x=x_{ij}})^2 \right\}.$$

Hence, subject to monotonicity constraints and under certain conditions, this model has a similar estimation as the one proposed by Ramsay and Li (1998). In addition, if parameter λ' is sufficiently large, then it shrinks $\delta''(x)$ to 0, and therefore, it shrinks δ to γx , where γ is a constant satisfying $\gamma \geq -1$, which leads to shrinking the warping path $h(x)$ to $(1 + \gamma)x$. Therefore, smoothing parameter λ is required to be designed and appropriately selected before implementing the proposed registration procedure. Figure 4 shows the profile alignment pairs between a given profile to be aligned (upper line) and the reference profile (lower line) by the DTW method and our investigated penalised-spline method under different values of λ . Figure 5 explains the registration results under these cases, where the solid curve is the true inverse warping function $h^{-1}(x)$ and the other curves represent the estimations of $h^{-1}(x)$ by different procedures. Figures 4 and 5 show that when $\lambda = 0$, the penalised-spline method has similar results with DTW because of the non-use of the penalty part. Under this case, the alignment path is only affected by the global distance (the first part in Model (7)) of these two profiles, and the global pattern of profiles to be aligned is neglected. Thus, one point on the reference may link to several points on the unaligned profile (upper curve), shown in Figure 4(a) and 4(b), and the estimated paths by these methods are tortuous and randomly fluctuate, shown in Figure 5. However, with the increase of the penalty parameter λ , the correlation of phase shift becomes higher and the possibility that several points align to one point decreases. That is to say, when a larger λ is selected, the global profile's pattern rather than the process noises or small variations weighs more on the warping path, and a smoother path is estimated, as shown in Figure 5. Thus, the proposed penalised-spline alignment method reduces the detrimental effects of process variations and leads to a more robust warping path.

Figure 6 extends the effects of the choice of parameter λ . Figure 6(a) shows the squared distance between aligned profiles and the reference under certain values of λ , which is obtained from the first part in Model (7). As shown in this figure, this distance tends to approach stability from $\log(\lambda + 1) = 0.5$. The warping error, which is defined as the difference between the estimated warping function and its true value, is shown in Figure 6(b). As shown, the optimal parameter λ is located between values satisfying $\log(\lambda + 1) = 0.5$ and $\log(\lambda + 1) = 1.5$. Hence, in this context, the recommended smoothing and penalty parameter λ is in $[0.5, 3]$.

2.5. Guidelines for design and implementation

This subsection provides guidelines on how to enforce the proposed approach for practitioners. In particular, several practical issues are discussed, including the selection of kernel function and bandwidth in Model (2), determination of penalty parameter λ in Model (7), and the effects of differences in profiles' lengths and sampling discretization.

A smoothing procedure is utilised to create smooth curves and decrease the impact of noise in data on alignment purpose. However, construction of profile features does not depend on the special structure of the smoothing procedure. For the kernel function, many kernels are available to meet the requirements, such as uniform, Epanechnikov, quadratic, and Gaussian kernels. We find that the performance of the alignment methods is mostly unaffected by the choice of kernel functions

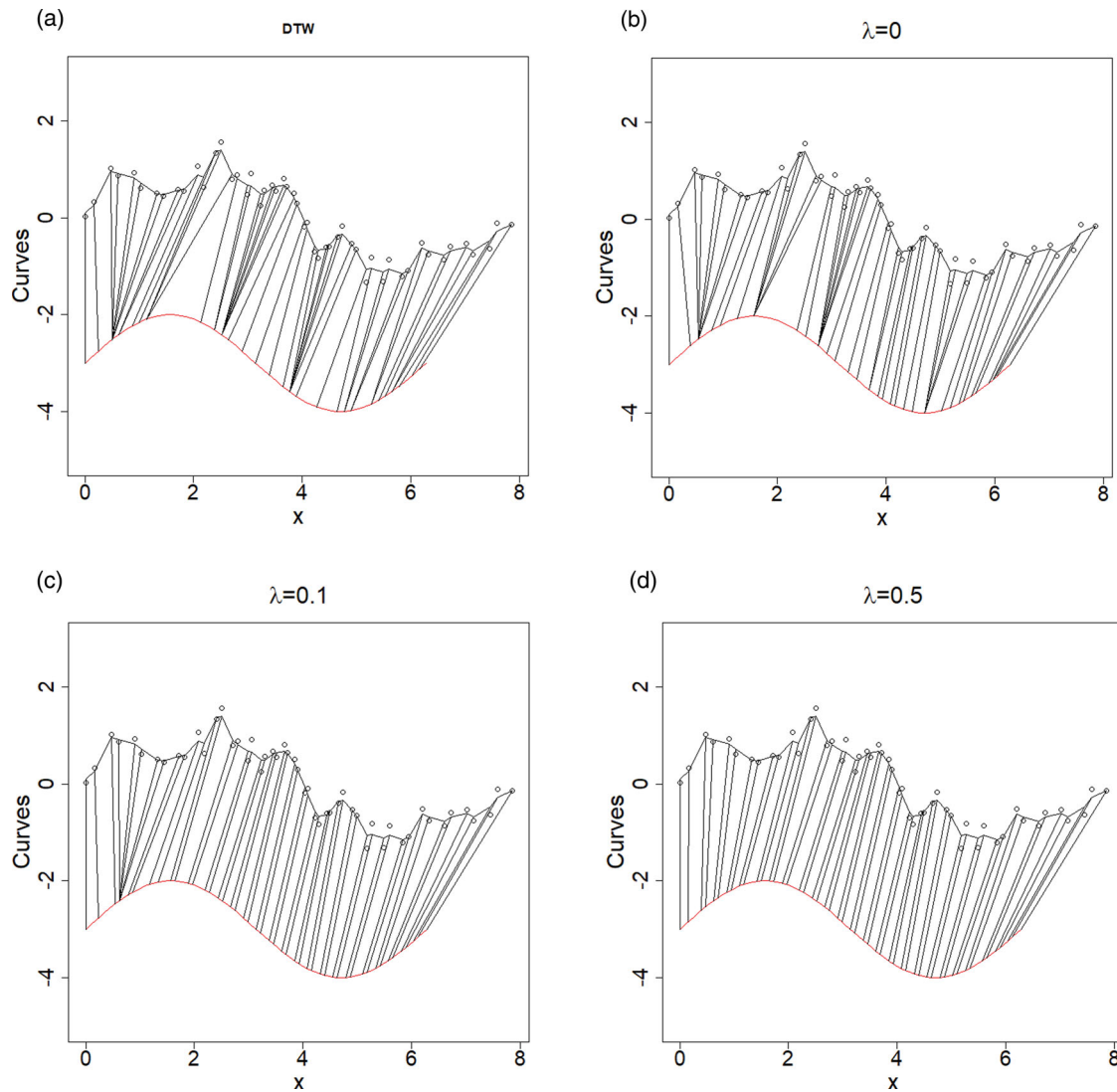


Figure 4. Profile alignments by DTW and penalised-spline methods under various sets of λ .

according to our simulations. For simplicity, we recommend to use the Epanechnikov kernel, which is

$$K_E(u) = 3/4 (1 - u^2) I(|u| \leq 1).$$

Generally speaking, the size of the optimal bandwidth would be expected to be proportional to the smoothness of the underlying function. However, we usually do not have specific information about the profiles from manufacturing processes, so we cannot choose an optimal l for actual engineering conditions before profile alignment and monitoring. In this work, we accept the recommendation by Zou, Tsung, and Wang (2008) to set the bandwidth for profile i as

$$l = \frac{1}{2} \times \left(\frac{1}{N_i} \sum_{j=1}^{N_i} (x_{ij} - \bar{x}_i)^2 \right)^{1/2} N_i^{-1/5},$$

where $\bar{x}_i = \sum_{j=1}^{N_i} ij$.

For the penalty parameter λ , in practice, data-driven approaches such as cross-validation and generalised cross-validation could be conducted to search for the best λ over a parameter space. However, for profile monitoring applications, these methods may be inappropriate. Roughly speaking, if parameter λ is sufficiently large, the warping path $h(x)$ will shrink to $(1 + \gamma)x$. If parameter λ is close to 0, the warping path $h(x)$ is similar to what is obtained from the traditional DTW. Within

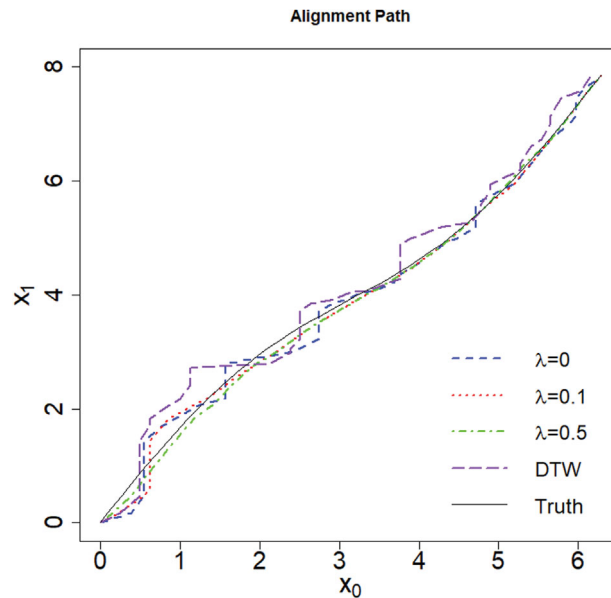


Figure 5. Alignment paths by DTW and penalised-spline methods.

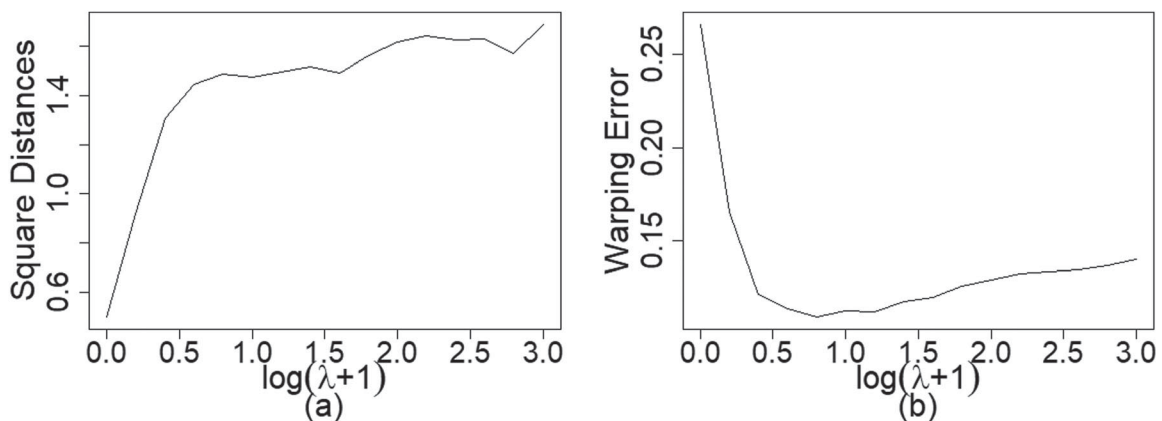


Figure 6. Effects of choice of parameter for penalised-spline method. (a) Squared distances of aligned profiles and the baseline under various values of λ , (b) the difference between the estimated warping function and its designed value under various choices of λ .

a certain range, an increase of λ could improve the estimation accuracy of the warping path, but at the expense of further distance. In manufacturing applications, we usually do not know the actual warping function because of the complex and flexible processes. Therefore, a flexible λ is recommended when the profiles' distance grows slowly or stays stable as λ increases within a parameter space.

The proposed alignment method is adequate to unequal profiles' lengths and non-equispaced sampling. The differences in length between the reference curve and the one to be aligned is less important, because the performance of the alignment procedure indeed is insensitive about the length. However, sampling sizes and discretization are important issues to be addressed. For the proposed alignment to perform well, a relatively large sample size is needed, especially with complicated profiles. This is because the alignment procedure is conducted based on discretised curves, and alignment results are shown by point pairs between two curves. A set of sampling points contain important features of profiles to be aligned. In addition, re-sampling operations or other pre-processing techniques could help improve the alignment result.

The proposed alignment methods could be applied to reduce the misalignment effects of profiles, and also be extended to applications in manufacturing, sociology and many other areas, where signal profiles have both amplitude variation and phase variation. Usually, the goal of profile analysis is to study the variations in or between profiles and explore connotative information, such as reasons cause the variations or essential characteristics in curves. Hence, the proposed method can be used to estimate the common shape function or the template curve from a series of curves. The alignment step is also needed

for clustering or classifying unaligned functional observations to explore similar or various data features, such as in image pattern recognition, spam identification, medical diagnosis. Nevertheless, in numerous manufacturing applications, profile monitoring has been widely investigated to detect the unusual variability and identify the stability of processes. Applying the proposed method to capture two types of variations, and leaving useful features behind is an effective approach to improve profile monitoring procedures, which then contributes to fault diagnosis, variability reduction and capability improvement of processes.

3. Simulation study

In this section, we present some numerical results regarding the performance of the proposed penalty-based smoothing warping method, which is marked as *PSW*. In addition to the proposed alignment approach, we also consider the *DTW* applied by Jeong, Jeong, and Omitaomu (2011), Arribas-Gil and Müller (2014), and the *B-spline-based (BSP)* method, one form of smooth monotonous approach, proposed by Ramsay and Li (1998) as the comparison methods. Curve smoothing is also a prior treatment before profile alignment for these three methods by setting the kernel bandwidth as $l = 0.3$. The performance of *DTW* was conducted by the *dtw* package of the R software environment. Another approach for functional alignment is to use metrics with better invariance properties. *SRV F*-based approach is an elastic registration or shape analysis considered in the functional data analysis literature, thus, we also compare with *SRV F*-based alignment investigated by Srivastava et al. (2011). *SRV F*-based approach is an extension of *DTW*, by replacing the distance of profiles with the warping invariant distance of two square-root velocity functions that are proportional to the derivatives of the original profiles. Hence, the optimal warping function of *SRV F*-based approach is achieved feature-to-feature.

In the simulation, we consider the reference profile as

$$y = \sin(x), \quad x \in [0, 2\pi] \quad (13)$$

and the following unaligned profile:

$$y = \sin(h(x)) + be^{-ux} \sin(vx) + \varepsilon, \quad x \in [0, 2\omega\pi] \quad (14)$$

where $b = 0.5$, $u = 0.2$, $v = 4$, and ω is a parameter that defines the lengths of the profiles to be aligned. In this context, ω is set as $\omega = 1.25$ when the profiles to be aligned are longer than the reference, whereas $\omega = 0.8$ when the profiles to be aligned are shorter than the reference. The first part in Model (14) displays the main shapes of the profiles, and the second part shows the periodic vibrations in applications. In these settings, the unaligned profiles are different from the reference surface both in amplitude and phase. The observed profiles are generated with the random errors ε generated i.i.d. from the normal distribution $N(0, \sigma^2)$ with $\sigma = 0.2$ and $\sigma = 1$ for a stable process and severe noises, respectively. For these sampling situations and the analysis in Section (2.4), the penalty parameter λ used in Model (7) was $\lambda = 0.5$.

Each of these sample functions was generated with 51 points in its space. Because of the potential unequally spaced sampling strategy, observation locations are first generated uniformly and then moved vertically by s_v , where $s_v = 0.02$ and $s_v = -0.03$ for odd and even locations, respectively, except for the starting and ending points. For example, for the reference profile in Equation (13), its sampling locations are $x = 0, (2\pi/50) - 0.03, (4\pi/50) + 0.02, (6\pi/50) - 0.03, (8\pi/50) + 0.02, \dots, (98\pi/50) - 0.03, 2\pi$.

The performance of *BSP* by Ramsay and Li (1998) was obtained using the 'fda' package of the R software environment, while *SRV F* was conducted using the 'fdastrvf' package. Because both *BSP* and *SRV F* methods are only available for alignments of profiles that have equal lengths, profiles that are shorter are added to the same length with the longer one, and the added values are set the same as the last sampling values. For example, in the case of $w = 1.25$ and the reference profile is shorter than the profiles to be aligned, we extended the profiles from point 2π to point 2.5π by setting all of their amplitudes as 0. However, these extending parts are not used in the stage of performance evaluation for the comparisons with other methods.

To compare the performances of different warping methods, we consider three scenarios of real warping function $h(x)$ by setting their inverse functions $h^{-1}(x)$: (1) Linear warping function, $h^{-1}(x) = \omega x$; (2) Sine-increasing warping function, $h^{-1}(x) = \omega x + 0.5 \sin(x)$; (3) Parabolic warping function, $h^{-1}(x) = 2\omega\pi - (\omega/2\pi)(x - 2\pi)^2$. The figures in Table (1) illustrate the relationships of x_0 (horizontal ordinate in figures) of the reference profile and its inverse-warped value x_i of the i th profile for these three scenarios, respectively, where the solid curves represent situations of $\omega = 1.25$ when the profiles to be aligned are longer than the reference, while the dashed lines show the situations of $\omega = 0.8$ when the profiles to be aligned are shorter than the reference.

To evaluate the performance of each alignment method, we apply three measurements. The first one is the squared distance (S_{Dist}) of the aligned profile with its reference, as defined in Model (3). A lower S_{Dist} means a lower global distance of

two profiles, which represents a better alignment estimation when choosing the global distance as the correlation measurement of two profiles. The second one is the root mean squared error (W_{Error}) of the estimated warping function and the true one, which is

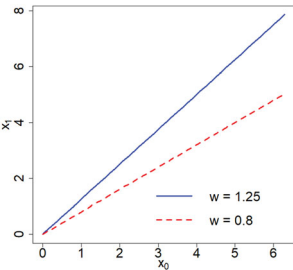
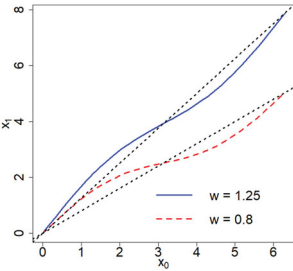
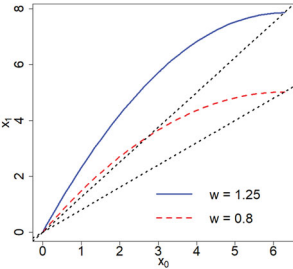
$$W_{Error} = \sqrt{\int_x [\hat{h}_i(x) - h(x)]^2 dx}$$

A lower W_{Error} means a more accurate estimation of warping function. The third one is the correlation (S_{Corr}) of the warping phase shift δ , which is defined as the sum of square differences of adjoining regularised differences of phase shifts δ . That is, S_{Corr} is the summation part of Inequation (8), which is

$$S_{Corr} = \sum_{j=1}^{N_i-2} \left[\frac{\delta_{i,j+2} - \delta_{i,j+1}}{x_{i,j+2} - x_{i,j+1}} - \frac{\delta_{i,j+1} - \delta_{i,j}}{x_{i,j+1} - x_{i,j}} \right]^2.$$

A lower S_{Corr} means a higher global correlation of neighbouring phase shifts (that is, two neighbouring points have more similar alignment path), which represents a stronger spatial or temporal autocorrelation of warping function. Each index in this section is computed based on 1000 repeated simulations, and the results are presented in Table (1). The numbers in parentheses are the standard errors of these measured values, and bold values are the lowest number in each row, which means the corresponding method has the best performance.

Table 1. Performances of the four profile alignment methods.

Scenario	$h^{-1}(x)$	Parameters		Indices	PSW	DTW	BSP	SRVF
		w	σ					
1	ωx 	1.25	0.2	S_{Dist}	1.38 (0.011)	1.10 (0.008)	4.36 (0.017)	1.65 (0.015)
				W_{Error}	0.177 (0.001)	5.12 (0.001)	0.407 (0.001)	1.08 (0.001)
				S_{Corr}	1.55 (0.005)	34.7 (0.174)	2.00 (0.002)	83.8 (0.710)
		1.25	1.0	S_{Dist}	4.93 (0.178)	18.9 (0.120)	10.4 (0.185)	14.6 (0.140)
				W_{Error}	0.488 (0.004)	5.05 (0.004)	0.492 (0.003)	1.06 (0.004)
				S_{Corr}	1.66 (0.012)	83.1 (0.748)	2.00 (0.012)	72.2 (0.676)
		0.8	0.2	S_{Dist}	1.42 (0.010)	0.481 (0.004)	12.6 (0.227)	2.28 (0.017)
				W_{Error}	0.170 (0.001)	3.36 (0.001)	0.384 (0.004)	1.03 (0.001)
				S_{Corr}	1.98 (0.012)	66.3 (0.351)	6.68 (2.58)	70.3 (0.587)
2	$\omega x + 0.5 \sin(x)$ 	1.25	0.2	S_{Dist}	1.27 (0.006)	1.12 (0.008)	3.36 (0.018)	4.01 (0.054)
				W_{Error}	0.127 (0.001)	4.93 (0.001)	0.403 (0.001)	0.773 (0.004)
				S_{Corr}	1.46 (0.008)	49.2 (0.267)	1.92 (0.001)	113 (0.995)
		1.25	1.0	S_{Dist}	1.27 (0.006)	17.7 (0.118)	10.6 (0.249)	21.3 (0.241)
				W_{Error}	0.127 (0.001)	4.87 (0.003)	0.423 (0.002)	0.968 (0.007)
				S_{Corr}	1.46 (0.008)	111 (1.11)	1.97 (0.011)	158 (2.17)
		0.8	0.2	S_{Dist}	0.894 (0.005)	0.545 (0.004)	9.22 (0.033)	6.68 (0.066)
				W_{Error}	0.353 (0.001)	3.15 (0.001)	9.16 (0.003)	1.15 (0.005)
				S_{Corr}	2.39 (0.004)	185 (1.81)	0.054 (0.001)	237 (3.80)
3	$2\omega\pi - \frac{\omega}{2\pi}(x - 2\pi)^2$ 	1.25	0.2	S_{Dist}	4.04 (0.013)	1.27 (0.010)	6.85 (0.026)	2.12 (0.050)
				W_{Error}	0.295 (0.001)	6.55 (0.001)	0.373 (0.001)	1.52 (0.014)
				S_{Corr}	6.97 (0.020)	148 (5.12)	22.0 (0.066)	106 (0.885)
		1.25	1.0	S_{Dist}	4.03 (0.012)	17.4 (0.120)	3.09 (0.205)	17.0 (0.282)
				W_{Error}	0.295 (0.001)	6.47 (0.006)	4.94 (0.037)	1.27 (0.013)
				S_{Corr}	6.93 (0.020)	203 (5.54)	3.36 (0.258)	180 (5.47)
		0.8	0.2	S_{Dist}	4.34 (0.014)	0.607 (0.005)	7.64 (0.037)	6.61 (0.038)
				W_{Error}	0.453 (0.002)	4.20 (0.001)	0.275 (0.001)	0.571 (0.003)
				S_{Corr}	7.11 (0.018)	513 (16.4)	0.002 (0.000)	328 (1.85)

From the table, we can draw the following conclusions.

(1) In most cases, our proposed penalised-spline method performs the best. It has a perfect ability to discover the correct warping functions and rather better performances in terms of profiles' distances and correlations of phase shifts. The penalised-spline method attempts to rectify the stretching or shrinking of horizontal axis x , and it has the lowest registration differences between the estimated warping functions and their desired values compared with the other three methods. Meanwhile, the proposed penalised-spline method results in better accuracy than *BSP* and *SRV F* depending on the profile distances, and it has a comparable ability with *DTW*. Also, The proposed method *PSW* also has a competitive performance on S_{Corr} in most cases. Therefore, a spline-based penalty is efficient in finding a good match.

(2) The penalised-spline method has a robust alignment ability for different profile noises and different profile lengths when comparing cases of $\sigma = 0.2$ with $\sigma = 1.0$ and comparing cases of $\omega = 1.25$ with $\omega = 0.8$, respectively. Compared to other methods, the penalised-spline approach is more effective and noise-robust in profile alignment.

(3) If only the profiles' distances are considered, the *DTW* method performs well in most cases for profile alignment. However, it loses this ability when samples have larger noise with $\sigma = 1.0$. The reason is that *DTW* attempts to find the minimum distances of two given sequences, and in the cases that sequences are smoother, it can perform well. However, for the situation of $\sigma = 1.0$ where violent fluctuations exist in the profiles, *DTW* has to find the potential optimal path in a limited search space due to the discreteness and inflexibility of the warping path, which leads to large mistakes in global optimisation in the profiles' distances. Similar results have also been reported by Keogh and Pazzani (2001) and Zang, Wang, and Jin (2016). Compared with *PSW* and *BSP* alignment methods, *DTW* performs the worst in terms of S_{Corr} . Because *DTW* tries to find the optimal alignment path by solving Model (3) without additional constraints, the alignment pairs by *DTW* are flexible, and thus the neighbouring correlation of warping path is weak.

(4) The *BSP* method performs better than *DTW* in terms of warping errors but worse in terms of profiles' distances. However, *BSP* has a competitive performance on S_{Corr} with *PSW*, because of the smoothing penalty in Model (6). Meanwhile, *BSP* has better performance when the setting warping function is parabolic.

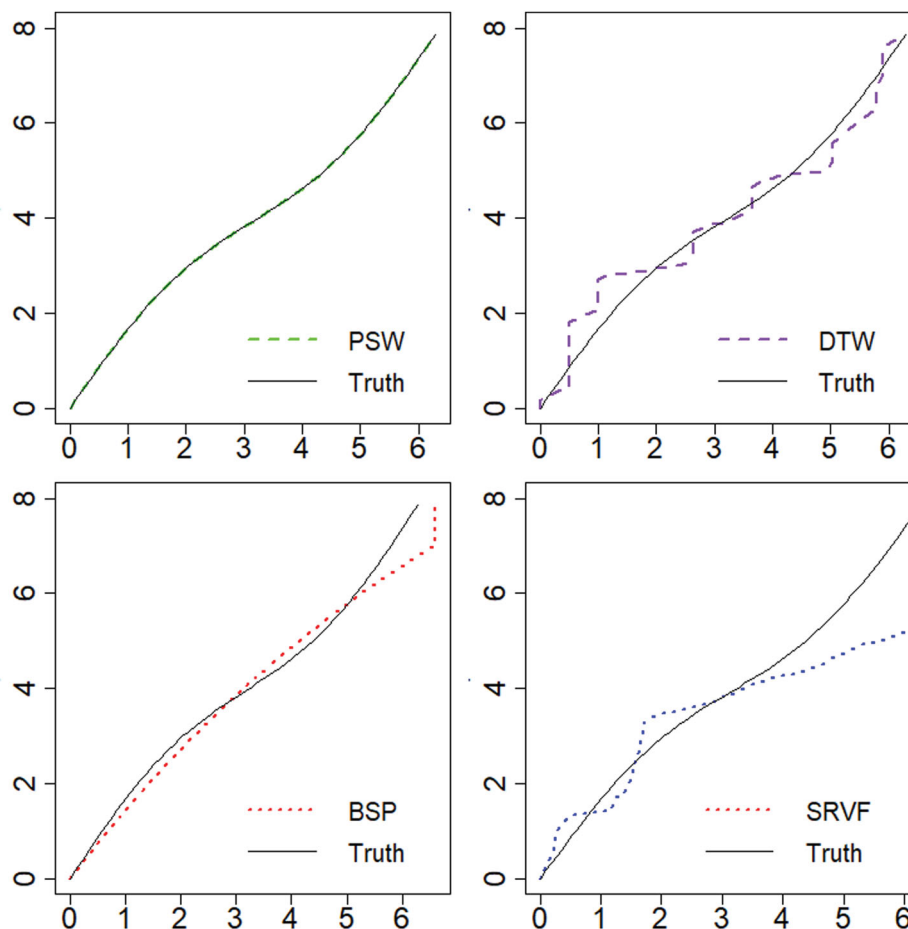


Figure 7. An example of alignment paths estimated by penalised-spline, *DTW*, *BSP* and *SRV F* methods.

(5) The *SRV F*-based method performs better than *DTW*, but worse than proposed *PSW* and *BSP* when considering warping errors. Meanwhile, in most cases, it has a similar profiles' distance value with *BSP*, and a similar global correlation of neighbouring phase shifts with *DTW*, but nearly all of them are worse than the performances of *PSW*. A possible reason behind these is that those profiles have high noises or severe fluctuations. Because *SRV F* is based on the derivative of curves, original profiles are assumed absolutely continuous and smooth, which is usually not the case in profile monitoring application.

Figure 7 is an example of the warping function estimated by these four methods for the sine-increasing case when $\omega = 1.25$ and $\sigma = 0.2$. The vertical axis is the independent variable x_i in profile i , and the horizontal axis is the estimated warped value of x_i in the reference profile. In these figures, solid curves are the inverse of the predesigned warping function. As shown in this figure, our proposed penalised-spline method outperforms both *DTW*, *BSP* and *SRV F* in capturing the smoothness and correlated features of the warping function. As expected, the warping function estimated by *DTW* is

Table 2. Performances of the four profile alignment methods when profiles have heterogeneous noises.

Scenario	$\sigma(x)$	Indices	<i>PSW</i>	<i>DTW</i>	<i>BSP</i>	<i>SRVF</i>
1	$e^{-0.2x}$	S_{Dist}	3.68 (0.034)	6.57 (0.062)	10.6 (0.249)	11.1 (0.157)
		W_{Error}	0.336 (0.004)	4.93 (0.001)	0.423 (0.002)	1.01 (0.008)
		S_{Corr}	2.23 (0.013)	66.9 (0.515)	1.97 (0.011)	139 (2.02)
2	$\frac{1}{3}[\sin(h(x)) + 0.5e^{-0.2x} \sin(4x) + 1.5]$	S_{Dist}	3.34 (0.033)	5.94 (0.056)	6.30 (0.112)	9.46 (0.150)
		W_{Error}	0.334 (0.004)	4.94 (0.002)	0.389 (0.001)	1.12 (0.008)
		S_{Corr}	2.35 (0.015)	69.3 (0.567)	1.97 (0.004)	144 (2.39)

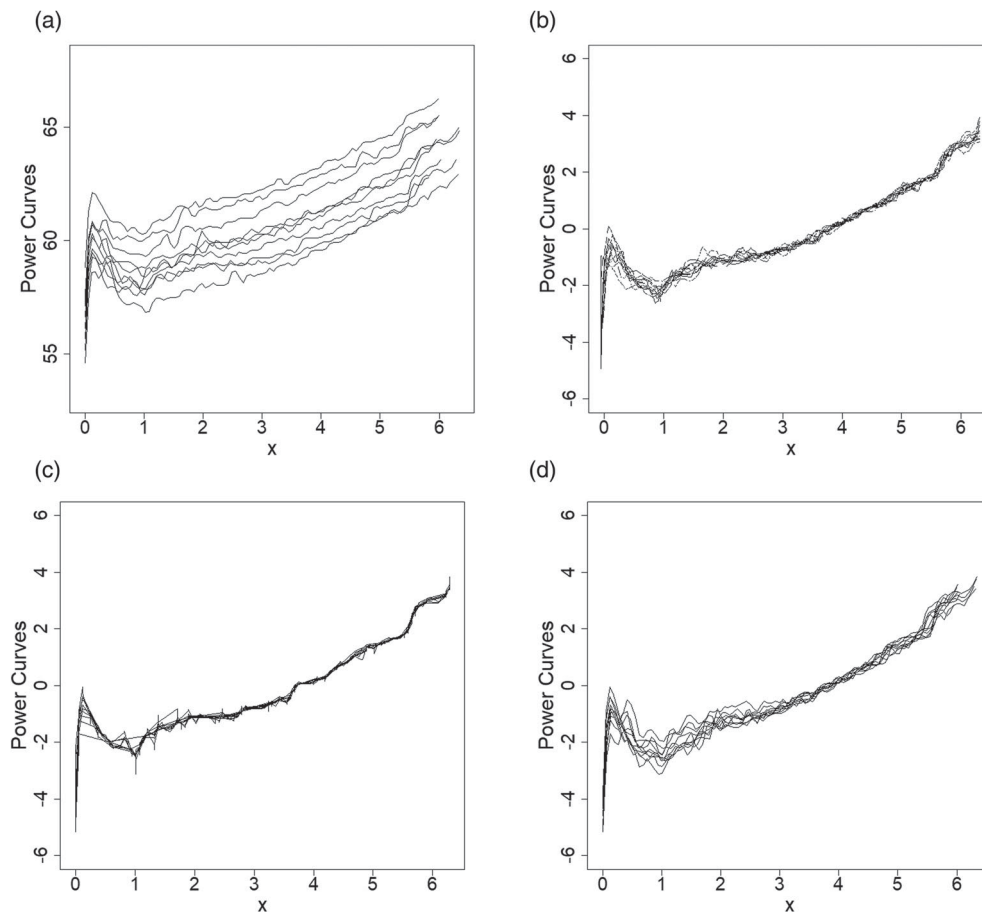


Figure 8. Alignment of heating power profiles: (a) unaligned profiles, (b) profiles aligned by the penalised-spline method, (c) profiles aligned by *DTW*, and (d) profiles aligned by *BSP*.

discrete. Meanwhile, although smoothness is achieved by *BSP*, it completely loses the sine tendency. For *SRV F*-based alignment, the warping path is slightly smoother than *DTW*, but worse than *BSP*.

In all above examples, the noises of profiles are homogeneous. However, in some manufacturing applications, this assumption may be violated; noises at different time or locations have different variances, which means such profiles have heterogeneous noises. Next, we consider examples with the same setup in Table 1 when $\omega = 1.25$ and $h^{-1}(x) = \omega x + 0.5 \sin(x)$, except that the standard deviation σ of profiles changes with x . To investigate the effects of heterogeneous noise, we consider two scenarios: (1) σ decreases as x , as in the heating power from the ingot growth process shown in Figure 1(b). Set σ at x be $\sigma(x) = e^{-0.2x}$. (2) σ varies as mean values of profiles, that is, σ is larger where the amplitude of profile is higher, and vice versa. Set σ at x be $\sigma(x) = \frac{1}{3}[\sin(h(x)) + 0.5e^{-0.2x} \sin(4x) + 1.5]$. Then the simulated results for heterogeneous noises are presented in Table 2. From the table, it can be seen that similar conclusions to those from Table 1 can be made for heterogeneous noises. More specially, the proposed method *PSW* performs the best for seeking more accurate warping function and obtaining the shortest global distances between profiles. *BSP* performs slightly better than *PSW* on S_{Corr} , but a little worse in terms of W_{Error} . For *DTW* and *SRV F* methods, they do not perform well in all cases considered. Specifically, the *SRV F*-based method outperforms *DTW* in the estimation of warping function, but it performs the worst in S_{Dist} and S_{Corr} .

4. A real-data example

In this section, we illustrate the application of the proposed profile alignment method using a real example of an ingot growth process. The profiles that need to be aligned are descriptions of heating power profiles, as studied by Dai, Wang, and Jin (2014) and Zang, Wang, and Jin (2016). Heating power profiles have both phase and amplitude variation. However, a difference in profile length is not an indication of process deterioration or failure. Important profile features occurring in different times/locations are just phase variation of profiles, which are not process shifts that a quality engineer intends to detect. Instead, downward or upward amplitudes in the aligned profiles may reflect shifts in the component or equipment

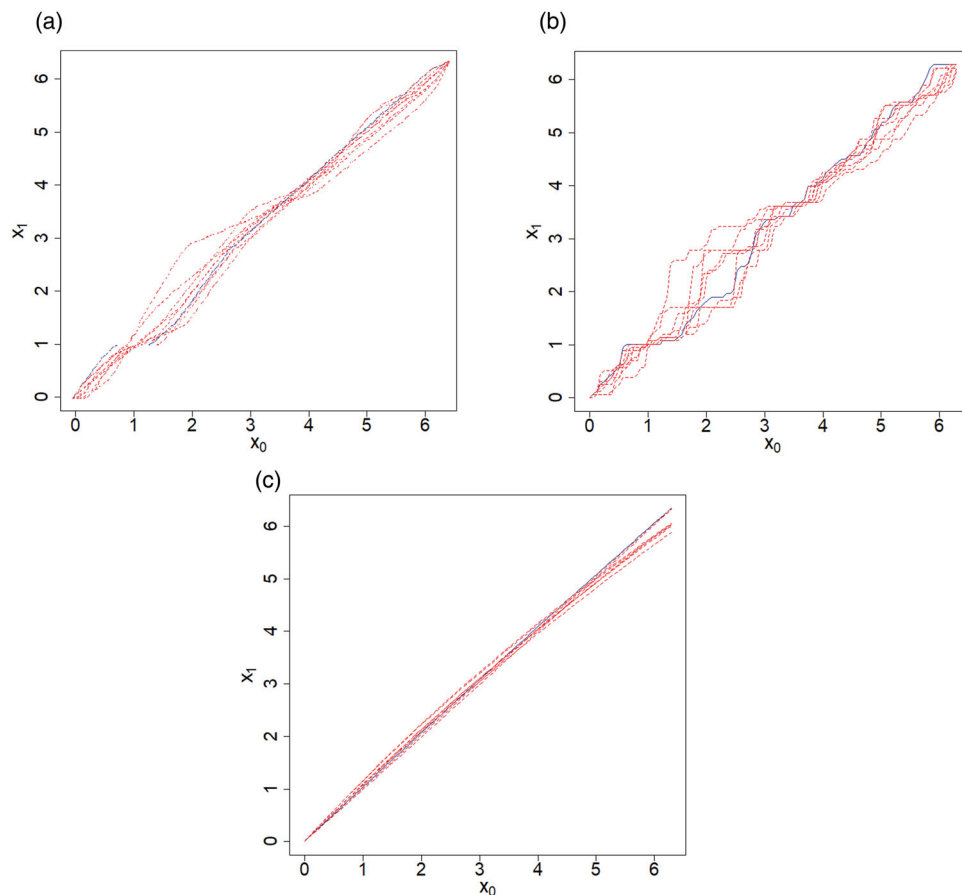


Figure 9. Warping functions of profile alignments by (a) penalised-spline, (b) *DTW* and (c) *BSP*.

status, and thus, in most SPC applications, the registration task is mainly acting on the phase variation, and leaving amplitude variation behind as large as possible. Therefore, a control chart could be carried out to detect those amplitude variations that reflect process shifting signals.

Here, we collected 11 heating power profiles and smoothed them using a kernel smoothing procedure, as shown in Figure 8(a). Then, we chose the third smoothed profile (dashed line in Figure 8(a)) as the reference, and the other 10 profiles were aligned to this chosen one using our proposed penalised-spline method, *DTW*, and *BSP*. Note that we transformed the independent variable (the processing time) of the reference profile to the range of $[0, 2\pi]$, and the other profiles were condensed in the same compression ratio. Before curve registration, power data were centralised with their own profile mean value.

Figures 8(b–d) are the registration results obtained using the penalised-spline method, *DTW*, and *BSP*, respectively, while the warping functions of each heating power profile are shown in Figure 9. From the plots in these figures, it can be observed that the penalised-spline method has a smoother warping function than *DTW* and a more flexible warping solution than *BSP*, which is consistent with the results in the previous section.

After profile alignment, the phase variations of profiles have been reduced, and only amplitude variations which may include potential amplitude shifts are left behind in profiles. Then the aligned profiles could be used by traditional monitoring methods, such as Zou, Tsung, and Wang (2008); Qiu, Zou, and Wang (2010); McGinnity, Chicken, and Pignatiello (2015), to detect profile shifts.

5. Conclusions and future work

In this paper, we have shown that the analysis of misaligned profiles is challenging because the profile variation is a combination of phase variation and vertical variation, and then identical features in different profiles do not occur at the same phase. Therefore, we proposed an alignment procedure based on a penalised-spline smoothing approach to separate phase variation and amplitude variation appropriately. In the proposed procedure, correlation of neighbour phase shifts is taken into consideration, and a penalised L_2 norm of differences between neighbour phase difference is added to the traditional cost function. Then, a new profile alignment path is obtained using the dynamic programming algorithm. Numerical examples show that the proposed robust alignment method performs better in discovering the correct warping functions compared with the traditional methods *DTW* and a *B*-spline-based approach.

In this work, we considered the alignment of one pair of profiles. In certain applications, process or product quality may be represented by multiple profile variables. The robust alignment of multiple profiles is an interesting topic that deserves future research efforts.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Natural Science Foundation of China (NSFC) under [grant number 71072012] and a key NSFC project under [grant number 71731008].

References

- Arribas-Gil, A., and H.-G. Müller. 2014. "Pairwise Dynamic Time Warping For Event Data." *Computational Statistics & Data Analysis* 69: 255–268.
- Berndt, D. J., and J. Clifford. 1994. "Using Dynamic Time Warping to Find Patterns in Time Series." In *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining, AAAIWS'94*, 359–370. Newport Beach, CA: AAAI Press.
- Chicken, E., Joseph J. Pignatiello, and J. R. Simpson. 2009. "Statistical Process Monitoring of Nonlinear Profiles Using Wavelets." *Journal of Quality Technology* 41 (2): 198–212.
- Colosimo, B. M., and M. Pacella. 2010. "A Comparison Study of Control Charts for Statistical Monitoring of Functional Data." *International Journal of Production Research* 48 (6): 1575–1601.
- Dai, C., K. Wang, and R. Jin. 2014. "Monitoring Profile Trajectories with Dynamic Time Warping Alignment." *Quality and Reliability Engineering International* 30 (6): 815–827.
- Eilers, P. H., and B. D. Marx. 1996. "Flexible Smoothing with B-Splines and Penalties." *Statistical Science* 11: 89–121.
- Gervini, D., and P. A. Carter. 2014. "Warped Functional Analysis of Variance." *Biometrics* 70 (3): 526–535.
- Grasso, M., A. Menafoglio, B. M. Colosimo, and P. Secchi. 2016. "Using Curve-Registration Information for Profile Monitoring." *Journal of Quality Technology* 48 (2): 99.

- Guo, J., J. Hu, B.-Y. Jing, and Z. Zhang. 2016. "Spline-lasso in High-Dimensional Linear Regression." *Journal of the American Statistical Association* 111 (513): 288–297.
- Gupta, S., D. Montgomery, and W. Woodall. 2006. "Performance Evaluation of Two Methods for Online Monitoring of Linear Calibration Profiles." *International Journal of Production Research* 44 (10): 1927–1942.
- Jensen, W. A., and J. B. Birch. 2009. "Profile Monitoring via Nonlinear Mixed Models." *Journal of Quality Technology* 41 (1): 18–34.
- Jensen, W. A., J. B. Birch, and W. H. Woodall. 2008. "Monitoring Correlation Within Linear Profiles Using Mixed Models." *Journal of Quality Technology* 40 (2): 167–183.
- Jeong, Y.-S., M. K. Jeong, and O. A. Omitaomu. 2011. "Weighted Dynamic Time Warping for Time Series Classification." *Pattern Recognition* 44 (9): 2231–2240.
- Kang, L., and S. L. Albin. 2000. "On-line Monitoring When The Process Yields a Linear Profile." *Journal of Quality Technology* 32 (4): 418–426.
- Keogh, E. 2002. "Exact Indexing of Dynamic Time Warping." In *Proceedings of the 28th International Conference on Very Large Data Bases*, 406–417. Hong Kong: VLDB Endowment.
- Keogh, E. J., and M. J. Pazzani. 2001. "Derivative Dynamic Time Warping." In *Proceedings of the 2001 SIAM International Conference on Data Mining*, 1–11. Chicago, IL: SIAM.
- Lee, J. J., Y. Hur, S. -H. Kim, and J. R. Wilson. 2012. "Monitoring Nonlinear Profiles Using a Wavelet-Based Distribution-Free Cusum Chart." *International Journal of Production Research* 50 (22): 6574–6594.
- Li, Y., Q. Zhou, X. Huang, and L. Zeng. 2018. "Pairwise Estimation of Multivariate Gaussian Process Models with Replicated Observations: Application to Multivariate Profile Monitoring." *Technometrics* 60 (1): 70–78.
- McGinnity, K., E. Chicken, and J. J. Pignatiello. 2015. "Nonparametric Change-point Estimation for Sequential Nonlinear Profile Monitoring." *Quality and Reliability Engineering International* 31 (1): 57–73.
- Noorossana, R., A. Saghaei, and A. Amiri. 2011. *Statistical Analysis of Profile Monitoring*. Vol. 865. Hoboken, NJ: John Wiley & Sons.
- Olsen, N. L., B. Markussen, and L. L. Rakêt. 2018. "Simultaneous Inference for Misaligned Multivariate Functional Data." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. doi:10.1111/rssc.12276.
- Panaretos, V. M., and Y. Zemel. 2016. "Amplitude and Phase Variation of Point Processes." *The Annals of Statistics* 44 (2): 771–812.
- Paynabar, K., C. Zou, and P. Qiu. 2016. "A Change-Point Approach for Phase-i Analysis in Multivariate Profile Monitoring and Diagnosis." *Technometrics* 58 (2): 191–204.
- Qiu, P. 2005. *Image Processing and Jump Regression Analysis*. Vol. 599. Hoboken, NJ: John Wiley & Sons.
- Qiu, P. 2013. *Introduction to Statistical Process Control*. Boca Raton, FL: CRC Press.
- Qiu, P., C. Zou, and Z. Wang. 2010. "Nonparametric Profile Monitoring by Mixed Effects Modeling." *Technometrics* 52 (3): 265–277.
- Ramsay, J. O., and X. Li. 1998. "Curve Registration." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 60 (2): 351–363.
- Ramsay, J., and B. Silverman. 2005. *Functional Data Analysis* (Springer Series in Statistics). New York: Springer.
- Rønn, B. 2001. "Nonparametric Maximum Likelihood Estimation for Shifted Curves." *Journal of The Royal Statistical Society Series B-statistical Methodology* 63: 243–259.
- Simpkin, A., and J. Newell. 2013. "An Additive Penalty P-Spline Approach to Derivative Estimation." *Computational Statistics & Data Analysis* 68: 30–43.
- Srivastava, A., W. Wu, S. Kurtek, E. Klassen, and J. S. Marron. 2011. "Registration of Functional Data Using Fisher-Rao Metric." *ArXiv e-prints*.
- Tucker, J. D., W. Wu, and A. Srivastava. 2013. "Generative Models for Functional Data Using Phase and Amplitude Separation." *Computational Statistics & Data Analysis* 61: 50–66.
- Walker, E., and S. P. Wright. 2002. "Comparing Curves Using Additive Models." *Journal of Quality Technology* 34 (1): 118–129.
- Williams, J. D., W. H. Woodall, and J. B. Birch. 2007. "Statistical Monitoring of Nonlinear Product and Process Quality Profiles." *Quality and Reliability Engineering International* 23 (8): 925–941.
- Woodall, W. H., and H. Williams, 2007. "Current Research on Profile Monitoring." *Produçã* 17 (3): 420–425.
- Xu, L., S. Wang, Y. Peng, J. Morgan, Marion R. Reynolds Jr, and W. H. Woodall. 2012. "The Monitoring of Linear Profiles with a glr Control Chart." *Journal of Quality Technology* 44 (4): 348–362.
- Zang, Y., K. Wang, and R. Jin. 2016. "Unaligned Profile Monitoring Using Penalized Methods." *Quality and Reliability Engineering International* 32 (8): 2761–2776.
- Zhang, C., H. Yan, S. Lee, and J. Shi. 2018. "Weakly Correlated Profile Monitoring Based on Sparse Multi-Channel Functional Principal Component Analysis." *IIEE Transactions* 50 (10): 1–29.
- Zou, C., F. Tsung, and Z. Wang. 2008. "Monitoring Profiles Based on Nonparametric Regression Methods." *Technometrics* 50 (4): 512–526.