Measurement 177 (2021) 109279

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Modeling relationships for field strain data under thermal effects using functional data analysis

Huachen Jiang [a], Chunfeng Wan [a,*], Kang Yang [a], Youliang Ding [a], Songtao Xue [b,c]

[a] Southeast University, Key Laboratory of Concrete and Prestressed Concrete Structure of Ministry of Education, Nanjing 210096, China
[b] Research Institute of Structural Engineering and Disaster Reduction, College of Civil Engineering, Tongji University, Shanghai 200092, China
[c] Department of Architecture, Tohoku Institute of Technology, Sendai, Miyagi 982-8577, Japan

ARTICLE INFO

Keywords:
Structural health monitoring
Functional data analysis
Time warping
Correlation analysis
Thermal effects
Time-lag effect

ABSTRACT

In the field of bridge health monitoring, it is ubiquitous to model the relationship for monitoring data. However, in many cases, especially for concrete bridge structures, field response data collected from sensors, such as strains and deflections, presents hysteresis phenomena due to the cyclic temperature variation. Such phenomena refer to the time-lag effect of the structure and impede the following correlation analysis because it will weaken the linearity among strains. In conventional methods, correlation analysis will be conducted after eliminating time-lag effect by using averaging method or simple phase shifting. In this paper, correlations between temperature-induced strain responses are investigated after the time-lag effect is accurately quantified by nonlinear phase shifting. In order to capture the features of nonlinear phase variation between responses, a novel approach, namely functional data analysis (FDA), is then proposed. Phase component is extracted through warping functions in square-root slope framework (SRSF) under Bayesian inference, in which the warping functions reveal the time delay effect between quasi-static strain data. The deformation pattern of warping function is further studied through functional principal component analysis (FPCA) in tangent space. After the phase difference is eliminated by the warping function, inter-relationship between strain data exhibits highly strong linearity instead of original hysteresis loop. To seek the underlying features of time-lag effect, hundreds of pairs of daily field measurements acquired from different locations on a long-span bridge have been analyzed to disclose statistical regularities. It is found that after the application of warping function, the time-lag effect has been remarkably eliminated and strong linearity of the responses emerges.

1. Introduction

With a rapid technological progress in structural health monitoring (SHM), increasingly more bridges have been implemented with SHM systems to gain practical understanding of bridge operation conditions [1–4]. Thermal actions, which act as a continuous and quasi-static external loading, lead to periodic variation of strain measurements on bridges. It is widely recognized that thermal actions play a crucial role in life cycle operation of bridges [5–8]. Once the correlations between quasi-static strain measurements being correctly modeled, bridge managers will have a better understanding of the physical significance of the strain measurements, state assessment of the bridge and other more practical knowledge. A few studies have revealed the unsynchronized i.e., nonlinear variation between different strain sensor data on the bridge under thermal actions [5,6,9], which poses a challenge to properly model the relationships. With the aim of parsimoniously characterizing the relationships between strain measurements induced by ambient temperature, the methods or techniques to well model the relationships is imperative.

Correlations between temperature-induced strain responses are very informative and have encouraged many pioneering researchers to devote themselves to investigating its variation mode. Catbas et al. [8] utilized linear regression model as well as a sinusoidal component to model the temperature-induced strain data via large amounts of monitoring data. Chen et al. [11] established a univariate linear regression for temperature and temperature-induced strain for the purpose of separation from vehicle-induced strain. Wang et al. [10] modeled the temperature and temperature-induced strain by using multivariate linear regression after extracting principal components of the temperature distribution. Ding et al. [5] conducted correlation analysis on strain and

* Corresponding author.
E-mail address: wan@seu.edu.cn (C. Wan).

https://doi.org/10.1016/j.measurement.2021.109279
Received 10 December 2020; Received in revised form 23 February 2021; Accepted 10 March 2021
Available online 19 March 2021

uniform temperature at night with the aim of damage detection and localization. Duan et al. [13] also pointed out the well-performed linear model for temperature and strain under thermal effects. Li et al. [14] mentioned the evident linear relationships between average temperature and strain measurements at the sampling frequency of 62.5 Hz.

Aforementioned researches didn't take time-lag effect into account , although it is a common case in a concrete bridge. Some researches attenuated the time-lag impacts by averaging temperature distribution over the bridge or adopting dimension reduction technique. Furthermore, generalization ability of the linear models presented above actually varies, and basically, linearity is moderate rather than sharp. Yang et al. [6] studied the time-lag effect between ambient temperature and temperature-induced strain and then enhanced the linearity via phase shifting by Fourier series expansion. Yang et al. [15] mitigated the time-lag effect for quasi-static deflection measurements by utilizing the cross-correlation technique, followed by a linear phase shifting process and correlation analysis. Guo et al. [16] found that a phase shifting of 45 min will amplify the correlation between displacement data and ambient air temperature. Brownjohn et al. [17] conceptualized the time-lag phenomena as thermal inertia effect, which provided a theoretical explanation rather than only experimental observation. Actually, heat transfer in concrete is very complicated due to its heterogeneity and forms a non-uniform temperature field in the structure. As the temperature field can be deemed as ambient quasi-static load on the structure, the stress distribution caused by the temperature distribution is often nonlinear which is represented by the time-lag phenomenon. All of these studies have considered the time-lag effect caused by temperature before conducting correlation analysis, and the general method is to use linear phase shifting, which will improve linearity undoubtedly. However, the direct phase shifting wouldn't properly characterize the essence of time-lag effect, therefore improvements in linearity are still needed. In addition, most researchers were inclined to concentrate on the relationships between input and output responses, i.e., ambient air temperature and strain data or deflection data. However, inter-relationships between sensor data were seldom analyzed although they also provide valuable information for condition assessment, sensor fault diagnostics and other more practical knowledge.

Monitoring data, at its core, is a real-valued and continuous function or curve on a certain observation interval. Thus, statistical analysis on such issues can be addressed by functional data analysis (FDA) in principle [18,19]. FDA has been widely applied in signal processing, image processing, neuroscience and geometry, which has a promising prospect to be a data-driven method in SHM. Chen et al. [9] investigated the correlations between strain measurements using the tool of phase-amplitude separation by FDA with the aim of data loss recovery. However, linearity was yet conspicuous after phase shifting, and statistical properties haven't been explored further although time-lag effect has been mitigated explicitly. Chen et al. analyzed the correlation for probability distributions of strain data utilizing the functional principal component (FPCA), which is the most powerful and useful tool in functional data analysis [12].

The main objective of this paper is to model the relationships between temperature-induced strains via eliminating the time-lag effect. A well-performed model will be proposed to successfully eliminate time-lag effect caused by ambient air temperature, and is represented by warping functions in SRSF space. Statistical properties of warping functions will be further discussed by using FPCA tool in tangent space. Each component represents how temperature-induced strains correlate in one direction. By utilizing first several principal components, correlation model of strain monitoring data will be finally established in original space.

2. Inter-relationship modeling of field strains with functional data analysis

As mentioned above, time-lag effect is actually the result of the non-synchronized variation of heat transfer in concrete because of the heterogeneity of the material. Temperature at different part of the bridge will not vary in a synchronous manner every day, hence there is a nonlinear time-lag phenomenon or phase difference between different locations, and so does the strain induced by such non-uniform temperature field. How to quantify and eliminate the time-lag effect is the core of the research and will be studied in this paper.

In this paper, strain monitoring data of a structure will be treated as functional data and later analysis will be made based on FDA. In order to eliminate the time-lag effect caused by temperature, function registration or alignment is adopted. The essence of function alignment is to separate amplitude variation from phase variation, i.e., align peak features and valley features of the function observations.[20]. Quantities of research have further improved the theory, especially of automatically aligning target function curves based on shape analysis [20–23]. However, most of these researches didn't apply Bayesian approach. Bayesian-based function registration provides a more robust and accurate solution to this problem, for the reason that it explores more information from warping function and incorporates prior knowledge into the model [24,25]. Hence, a model to eliminate time-lag effect based on the Bayesian registration is proposed and detailed explanation is also presented.

2.1. Mathematical representation of monitoring data

Daily strain measurement pairs, which obtained from different locations on the bridge, are used to conduct FDA. In this paper, wavelet filtering is first adopted to extract temperature-induced stain, and daily temperature-induced strain pair is denoted as f_1 and f_2. These strains, f_1 and f_2, are defined on the domain of time points determined by sampling frequencies of sensors. However, in order to meet the corresponding arrangement of FDA, we normalize the time period from [0 h, 24 h] to [0, 1] with no loss of generality. To obtain non-linear phase variation, also known as time lag between f_1 and f_2, warping function is to be estimated, which maps from the interval [0,1] to itself, although with different values taken: $\gamma : [0,1] \rightarrow [0,1]$. Warping function is the core of aligning f_2 and f_1, via composition operation: $f_2 \circ \gamma$, where the notation $\circ$ represents composition operation. Through the composition operation with γ, f_2 takes values from itself in a nonlinear increment which indicates nonlinear phase shifting of f_2. Consequently, time-lag effect has been eliminated by utilizing warping function, and f_1 holds and f_2 is converted to $f_2 \circ \gamma$.

In this paper, we denote $\mathscr{F}$ to be the set of all functional data observations like f_1 and f_2: $\mathscr{F} = \{f : [0,1] \rightarrow \mathbb{R}\}$, where $\mathbb{R}$represents real numbers. Likewise, let Γ be the set of all warping functions γ: $\Gamma = \{\gamma : [0,1] \rightarrow [0,1] | \gamma(0) = 0, \gamma(1) = 1\}$. In topology, Γ is termed Lie group and is an orientation-preserved diffeomorphism with identity mapping element $\gamma_{id}(t) = t$, which suggests no time lag exists. To estimate $\gamma, \underset{\gamma \in \Gamma}{inf} \|f_1, f_2 \circ \gamma\|$ is often considered as the cost function term. However, there exists a severe problem of asymmetry of the alignment, i.e.,

$$\underset{\gamma_1, \gamma_2 \in \Gamma}{inf} \|f_1 \circ \gamma_1, f_2 \circ \gamma_2\| \neq \underset{\gamma_1}{inf} \|f_1 \circ \gamma_1, f_2\| \neq \underset{\gamma_2}{inf} \|f_1, f_2 \circ \gamma_2\|$$

where $\|\cdot\|$ stands for $\mathbb{L}^2$ norm. Physically speaking, γ provides asymmetric solution to f_2 aligning to f_1 and f_1 aligning to f_2, which is unacceptable. This problem was managed to be addressed since square-root slope function (SRSF) framework was introduced [20,21].

2.2. Functional space transformation

With the advent of SRSF framework, monitoring functional observations f_1 and f_2 can be transformed to the corresponding SRSF representations, followed by:

$$q(f) : [0,1] \to \mathbb{R}, q(t) = sign(\dot{f}(t))\sqrt{|\dot{f}(t)|} \tag{1}$$

Similarly, SRSF of $f^{\circ}\gamma$ is denoted as

$$(q,\gamma) = q(f^{\circ}\gamma) = q(f)(\gamma(t))\sqrt{|\dot{\gamma}(t)|} \tag{2}$$

Actually, the mapping process implies an important property of isometry, followed by:

$$\|q_1 - q_2\| = \|(q_1,\gamma) - (q_2,\gamma)\| \tag{3}$$

in which q_1, $q_2$are the corresponding SRSF of f_1 and f_2. With SRSF framework, warping function γ now provides a symmetric solution to aligning q_2 with q_1 and aligning q_1 with q_2. As long as γ is found to eliminate the time lag between q_1 and q_2, it is also the same function that eliminates the time lag between f_1 and f_2 in original space, also, both for aligning f_1 with f_2 and aligning f_2 with f_1. Set of functions q is termed quotient space$\mathbb{L}^2/\Gamma$. Hence, we now can align q_1 with q_2 and vice versa in quotient space, followed by an inverse mapping of Eq. (1), which would be a precise definite integral, to original space $\mathscr{F}$, rather than seek an ill-conditioned warping function γ in $\mathscr{F}$ directly.

In order to exploit more information about warping function γ(Γ is a nonlinear manifold, thus working on this space becomes difficult), and to incorporate Bayesian inference into the model, warping function γ is required to be projected onto another space, which should be a linear space. We denote $\psi = q(\gamma) = \sqrt{\dot{\gamma}}$, and it can be easily verified that ψ is actually the SRSF of γ, and the mapping process is also invertible. One principal property of ψ is it has $\mathbb{L}^2$ norm which means:

$$\|\psi\|^2 = \int_0^1 \psi(t)^2 dt = \int_0^1 \dot{\gamma}(t)^2 dt = \gamma(1) - \gamma(0) = 1 \tag{4}$$

Actually, the resulting space $\mathbb{S}_\infty$ ($\psi \in \mathbb{S}_\infty$) is termed unit sphere in the Hilbert space $\mathbb{L}^2$. The corresponding distance defined on this space is arc-length, and the mapping process also has the property of isometry. However, $\mathbb{S}_\infty$ is yet to be linear (only in linear space can FPCA and Bayesian inference be conducted), and the last transformation is needed to be applied.

Tangent space is what satisfies all our requirements and can be mapped back to $\mathbb{S}_\infty$ and then to Γ.Tangent space is defined on a specified point residing on $\mathbb{S}_\infty$, one common choice is $1 \in \mathbb{S}_\infty$, which is adopted in this paper. This very tangent space is denoted as follows:

$$T_1(\mathbb{S}_\infty) = \left\{ g : [0,1] \to \mathbb{R} \middle| \int_0^1 g(t)dt = 0 \right\} \tag{5}$$

The following equations, which is termed exponential map and inverse exponential map, express the mapping process:

$$T_1(\mathbb{S}_\infty) \to \mathbb{S}_\infty : \exp_1(g) = cos(\|g\|) + \frac{sin(\|g\|)}{\|g\|} g, \tag{6}$$

$$\mathbb{S}_\infty \to T_1(\mathbb{S}_\infty) : \exp_1^{-1}(\psi) = \frac{\theta}{sin(\theta)}(\psi - cos(\theta)), \tag{7}$$

$$\theta = arccos(\langle 1, \psi \rangle), \ g \in T_1(\mathbb{S}_\infty). \tag{8}$$

Warping function in $\mathbb{S}_\infty$ is transformed to tangent space $T_1(\mathbb{S}_\infty)$ based on Eq. (7), and can be mapped back via Eq. (6). In the above equations, gis the representation of warping function γ in tangent space, and θ is the vectorial angle between 1 and ψ. Till now, all later data analysis could be performed on the tangent space, which is a linear space with property of isometry and invertibility. Bayesian model will be established and FPCA will be performed in the rest of this section.

2.3. Warping function estimated under Bayesian inference

Bayesian model is based on the difference of $q(f_1)(t) - q(f_2{}^{\circ}\gamma)(t)$ in the quotient space, which is actually a standard multivariate normal distribution. After being transformed into tangent space, the following reparametrized expression is obtained:

$$q(f_2{}^{\circ}\gamma)(t) = q(f_2)\left(\int_0^t exp_1^2(g)(s)ds \right) exp_1(g)(t) \tag{9}$$

where $\exp_1$ is the exponential map and has been defined in Eq. (6).

Finally, the Bayesian model for warping function is established as follows:

$$q(f_1)(t) - q(f_2)\left(\int_0^t exp_1^2(g)(s)ds \right) exp_1(g)(t) | g, \sigma_1^2 \tilde{N}(0, \sigma_1^2 I) \tag{10}$$

$$g \tilde{GP}(0, C_g, I), \sigma_1^2 \tilde{I}G(a,b) \tag{11}$$

In Eq. (11), *GP* represents Gaussian process with determined covariance function C_g. *IG* (a, b) represents inverse gamma distribution with determined constants a and b, $\sigma_1^2 I$ represents the covariance matrix of the normal distribution, where *I*is the identity matrix. More details are referred to Lu Y et al. [24] where how to update prior distribution of (g, σ_1^2) is fully explained. Since we have the Bayesian model, represented in Eq. (11), function g can be updated after each iteration using Metropolis-Hastings algorithm. Furthermore, warping function γ has the direct connection with g via exponential mapping and its inverse, hence γ is also updated after each iteration along with g to find a best solution to eliminate the time-lag effect.

2.4. Functional principal component analysis

So far, warping function γ has been determined via aforementioned SRSF model under Bayesian inference. Hence the way to eliminate time-lag effect has been utterly defined by γ. However, correlation model for strain measurements is implicitly expressed by warping function, thus an explicit and parsimonious correlation is imperative. To this aim, we perform functional principal component analysis (FPCA) on warping function γ in order to have a better understanding of how warping functions eliminate the time-lag effect and enhance the linearity of field monitoring strain measurements. Each principal component (PC) stands for an optimal direction to correlate two strain measurements, same as conventional principal component analysis (PCA), and at its core is a dimension reduction technique (infinite dimension to finite dimension). Therefore, by analyzing a long-term monitoring data, and extracting some main PCs of γ can adequately represent its variation mode. Since $\Gamma$is a nonlinear manifold where FPCA can not be performed, FPCA will be performed on the tangent space (tangent space is a linear space).

Let $X_i(t)$ be a sample of $X(t)$ restricted on the tangent space $\mathbb{S}_\infty$, $1 \leqslant i \leqslant n$. As mentioned in part 2.2, $X(t)$ is exactly a set of warping functions γ projected onto the tangent space. Thus, without detailed verification (which is beyond the topic of this research), $X(t)$ can be decomposed and represented by finite combination of principal components:

$$X(t) \approx X_m(t) = \mu(t) + \sum_{k=1}^{m} \xi_k \varphi_k(t) \tag{12}$$

where $\mu(t) = E(X(t))$, and ξ_k is the principal component defined by the *k*-th eigenfunction φ_k:

$$\xi_k = \int (X(t) - \mu(t))\varphi_k(t)dt \tag{13}$$

$\varphi_1, \varphi_2, \ldots$ stand for orthonormal eigenfunctions. φ_k indicates the dominant variation mode of $\varphi_1, \varphi_2, \ldots, \varphi_{k-1}$:

$$\varphi_k = \underset{\|\varphi\| = 1, \langle \varphi, \varphi_j \rangle = 0, j=1,2,\ldots,k-1}{argmax} \left\{ Var\left(\int X(t) - \mu(t) \right) \varphi(t)dt \right\} \tag{14}$$

where $\|\varphi\| = \left(\int \varphi(t)^2 dt \right)^{1/2}$ and $\langle \varphi, \varphi_j \rangle = \int \varphi(t)\varphi_j(t)dt$,

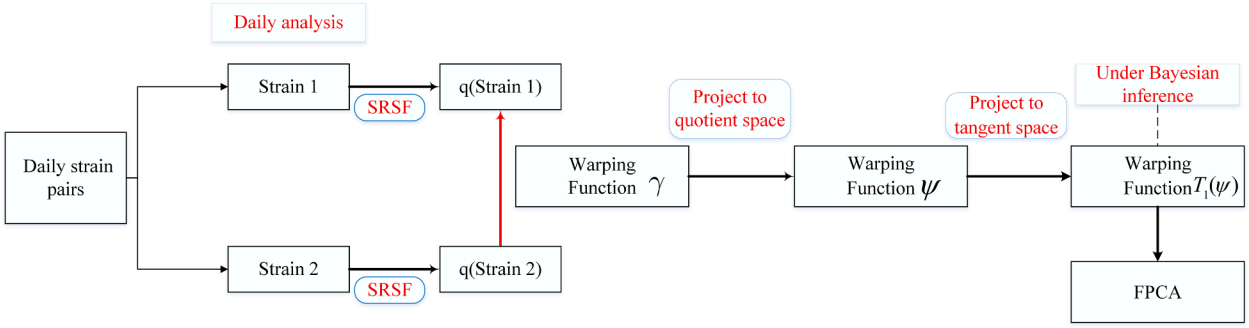
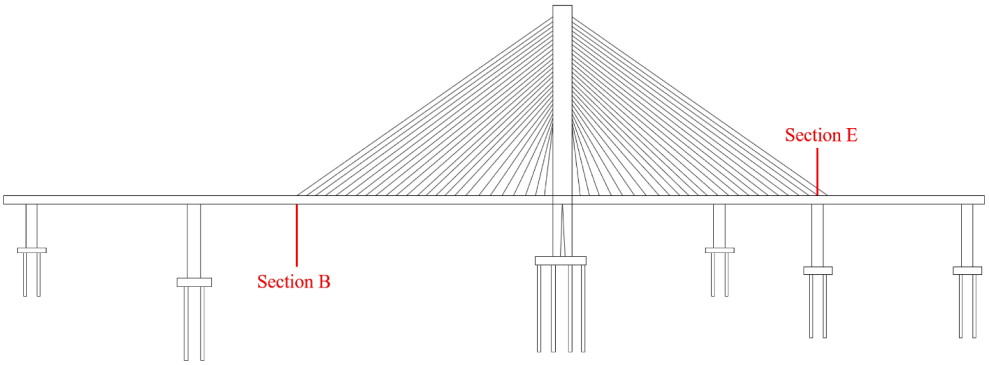


Fig. 1. Diagram of the algorithm implementation.

(a) Photo of target bridge



(b) Diagram of deployment of strain sensor studied in this paper.

Fig. 2. Illustration of target bridge and sensor deployment.

for j = 1, 2, ..., k - 1 .

Hitherto, main and finite functional principal components have been estimated, the model for warping function γ can be reconstructed and represented by the combination of these PCs. Once the model for γ being established and generalized, the correlation among strain measurements has been finally determined. The aforementioned algorithm process is described in Fig. 1.

3. Real application and analysis

In this section, a dataset obtained from a real bridge located in China is used to validate the proposed FDA analysis. The method to eliminate the time-lag effect is verified and the warping function which contains the information of time-lag is modeled via FPCA. Finally, the correlation model defined by warping function is built and presented.

3.1. Data description

The monitoring strain data is acquired from a real asymmetric cable-stayed bridge in China (Fig. 2(a)). Although a sound SHM system has been implemented on this bridge to continuously and closely monitor the operation condition of the bridge, only monitoring strain data is considered in this paper. Therefore, other sensors installed on the bridge will be ignored in the diagram (Fig. 2(b)) of sensor deployment. Two sensors located on the sections B and E respectively are of interest and analyzed as an example. In each section, only one sensor is needed to

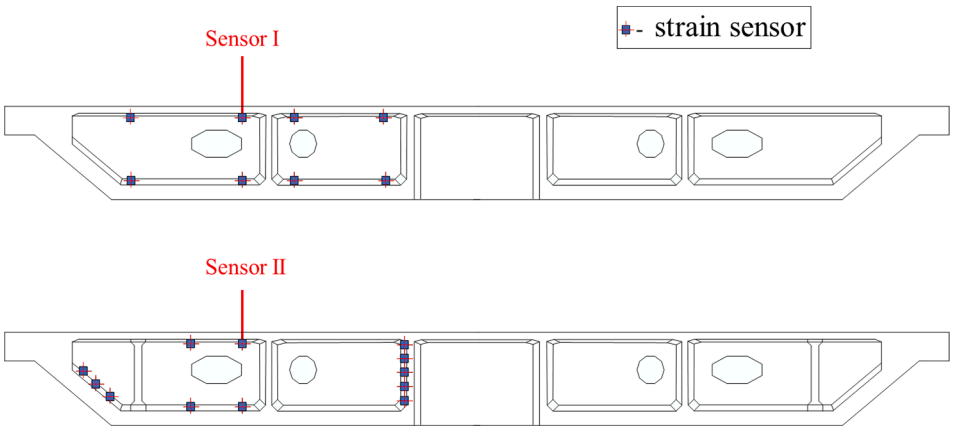


Fig. 3. Deployment of strain sensors on section B and Section E.

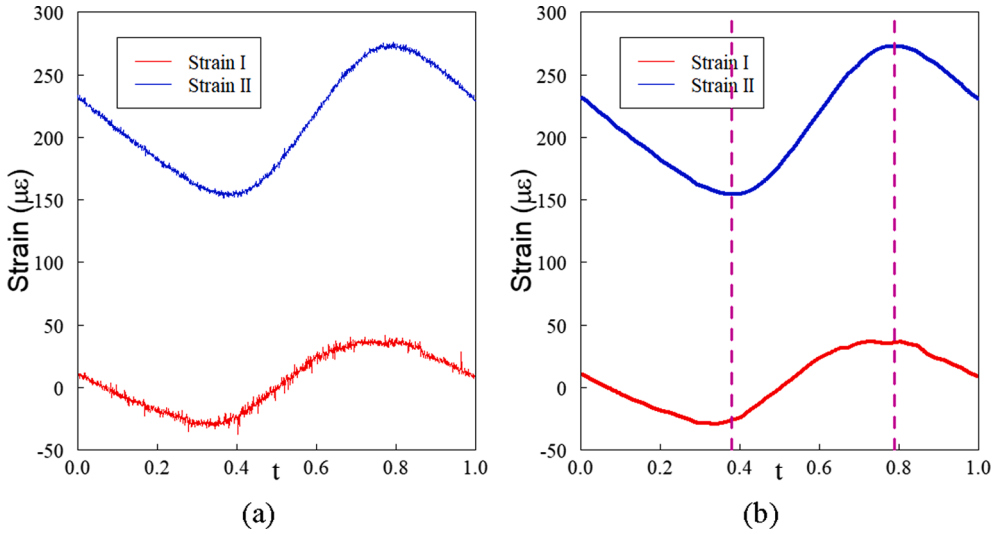


Fig. 4. Extraction of temperature-induced strain: (a) original strain measurements; (b) temperature-induced strain.

model the correlation model in the later analysis, and is depicted in Fig. 3. Sensors that are selected to demonstrate the algorithm are denoted as Sensor I and Sensor II.

3.2. Extracting temperature-induced strain

As mentioned above, the goal of this research is to properly model the correlation between strain measurements under thermal effects, thus original data should be processed in order to extract the temperature-induced strain, i.e., quasi-static strain measurements. In order to perform FDA, time domain in x-axis should be restricted to [0, 1] without loss of generality (see Section 2.1). The sampling frequency of Sensor I and II are 1/60 Hz which is quite low, thus some of other dynamic load effects have been eliminated, e.g. traffic load and wind load. It is generally accepted that the stationary or smooth parts represent the strain induced by temperature and the non-stationary parts represent the strain induced by vehicles or wind. However, noise (referred to other variation loads) still exists, shown in Fig. 4(a), a method to separate pure temperature-induced strain is also required. Many techniques are able to effectively meet this requirement, like wavelet transformation, low-pass filter and empirical mode decomposition (EMD). In this paper, wavelet decomposition is adopted to extract temperature-induced strain [26] as shown in Fig. 4(b). It should also be noted that there exits apparent time-lag in time domain between Strain I and Strain II, in another word, peak and valley of the curves for Strain I and II are not aligned.

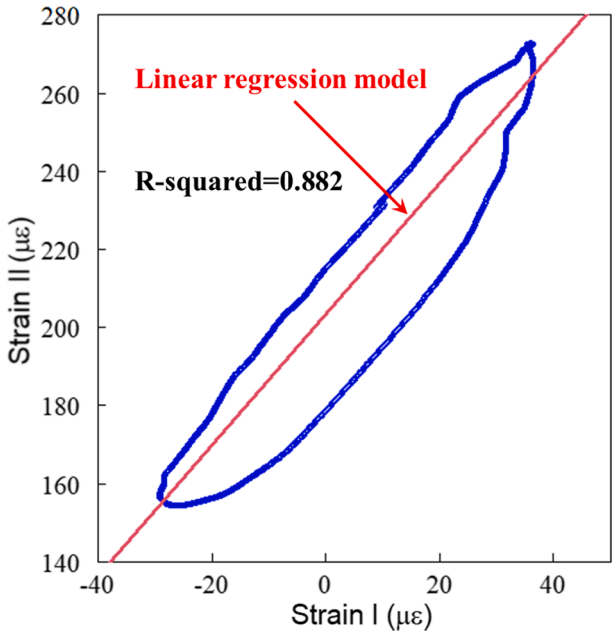


Fig. 5. Correlation plot of Strain I vs. Strain II and estimated linear model.

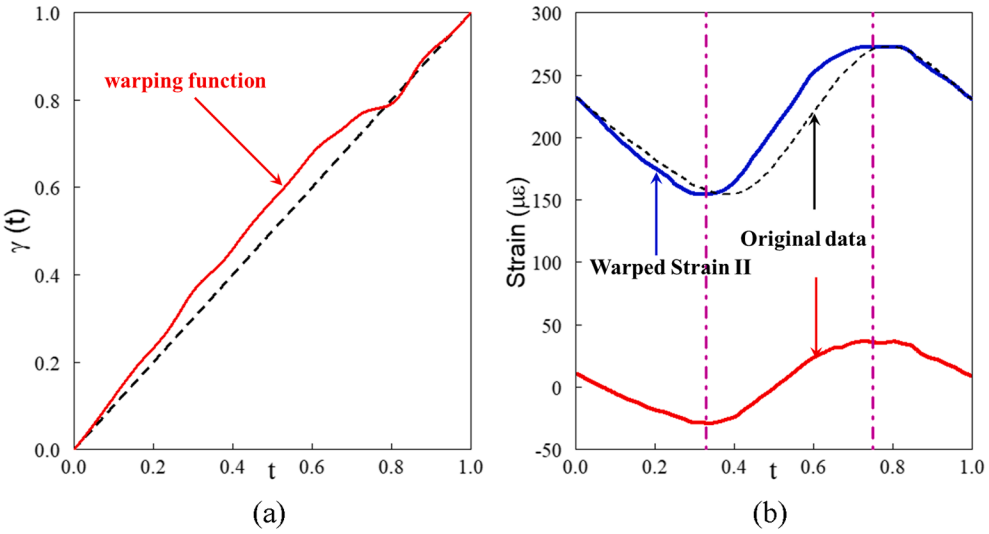


Fig. 6. Elimination of time-lag effect using warping function: (a) warping function; (b) time-lag effect elimination process.

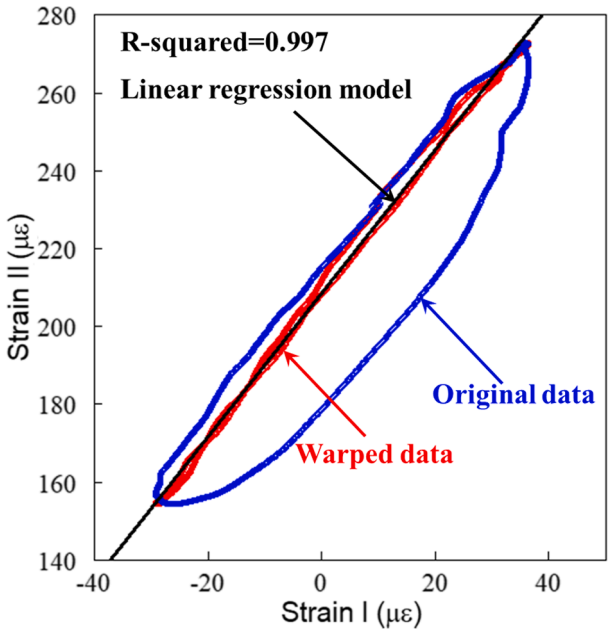


Fig. 7. Correlation plot of Strain I vs. warped Strain II and estimated linear model.

3.3. Correlation analysis for quasi-static strain using time warping

As previously mentioned, cyclic temperature variation is the main reason that causes the time-lag effect. An archetypal example of correlation plot of Strain I and II is presented in Fig. 5, the scatter plot forms a hysteresis loop in a single day, which further indicates the time-lag effect in the correlation domain. A statistical measure to interpret the strength of the correlations between sensor measurements is also adopted, named R-squared. The correlation between these measurements should have strong linearity in theory if thermal effects are eliminated properly. R-squared of the estimated linear regression model in Fig. 5 is 0.882, although it is a high linear relationship, a stronger linearity is still masked by the time-lag effect.

In order to eliminate time-lag effect such that a stronger linear correlation will emerge, warping function plays a crucial role. In FDA analysis, Strain I and II are deemed as functional sample observations, and are transformed into the quotient space in order to estimate the warping function. According to the theory described in sections 2.2 and 2.3, warping function is estimated via SRSF framework under Bayesian inference. The corresponding result is demonstrated in Fig. 5. The solid red line in Fig. 6(a) represents the estimated warping function, which also implies the quantification of nonlinear phase variation between Strain I and II. The dashed black line represents non-phase variation between two curves, i.e., unnecessary to align two curves. After performing composition operation of Strain II$^{\circ}\gamma$, Strain I and II are optimally aligned in Fig. 6(b). Original and warped Strain II are plotted in dashed black line and solid blue line respectively, and the manifest phase difference emerges in the time domain which also known as time-lag effect.

Hence the warping function has been estimated, time-lag effect is consequently eliminated, and the linearity between Strain I and II becomes much stronger than before, seen in Fig. 7. R-squared is up to 0.997, which is a huge breakthrough compared with 0.882. The essential relationship between strain sensors has been revealed, and the correlation between them will inevitably become the key component for further analysis and condition assessment of the bridge.

More daily archetypal examples to verify the method proposed are presented in Fig. 8, and one of the biggest features could be noticed from the diagram above is that all correlation scatter plots are transformed from hysteresis elliptical loop to almost a line. Moreover, statistical measure R-squared, which determines how well the linear regression model performs, all approaches 1. In this sense, warping function effectively reveal the relationship between monitoring strain measurements. However, in order to seek the pattern how the warping function changes, warping function is modeled via FPCA in the rest of the section for this very purpose.

3.4. Feature extraction using functional principal component analysis

With the aim of obtaining deformation pattern of the warping function, FPCA technique is used to extract features of warping function, and 131 daily strain measurement pairs are fed into FPCA algorithm. Warping functions estimated from these strain measurements are demonstrated in Fig. 9(a) and (c), and because warping functions can only be performed by FPCA in linear space, and tangent space is the optimal choice to represent these warping functions shown in Fig. 9(b) and (d).

So far, warping functions can be performed by FPCA algorithm to extract deformation mode, and first 9 functional principal components are presented in Fig. 10. Therefore, according to the theory discussed in Section 2.4, warping functions can be decomposed and reconstructed by

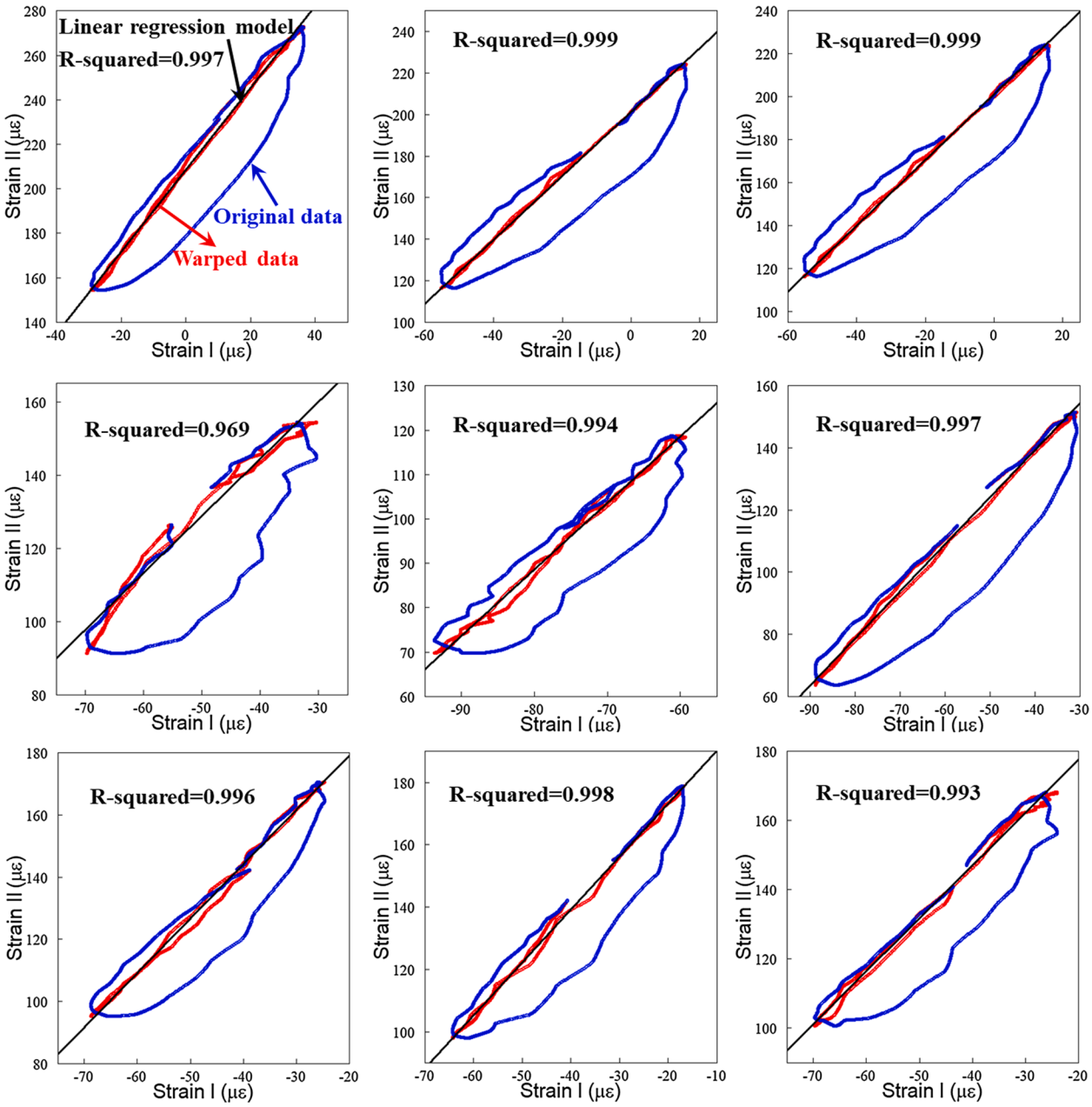


Fig. 8. More archetypal examples of correlation plot before and after warping.

finite combination of the principal components. Although the physical essence of each component is hardly interpretable (intrinsic drawback) in functional data analysis, these components provide an informative tool to effectively model the deformation mode of warping functions.

In order to capture most of the variation mode of warping function, first 16 principal components are used to represent them. Furthermore, the quantification of variation captured by each component is measured by proportion of variance explained: $\pi_j = \lambda_j / \sum_{j=1}^{\infty} \lambda_j$, $\mathrm{j} = 1, 2, \ldots, 16$, λ_j is the eigenfunction of the warping functions, which has almost the same meaning with conventional PCA method. As can be seen in Fig. 11 and Table 1, first 16 PCs have explained 95.3% variation of the warping function, which interprets most of information contained in them. Thus, any warping function can be represented and reconstructed by these 16 components properly. Otherwise, those can't be properly fitted by principal components implies bridge operation condition has changed or sensor has failed or any other problem, which should be analyzed further. In another word, thermal effect, as well as time-lag effect can be practically modeled and eliminated properly the with the technique proposed.

In order to validate the generalization of warping function reconstructed by principal components, some test experiments are conducted shown in Fig. 12. The first column represents the true warping function in tangent space and the corresponding estimated one. There does exist some difference in tangent space, however, most features have been captured, i.e., peaks and valleys of the curve. When projected to original parameter space via exponential mapping, ground true warping functions and estimated ones show well consistency. Finally, the estimated correlation line (time-lag effect has been eliminated) also superimposes on the ground true one. That is to say, any given a pair of temperature-induced strain measurements, correlation model is based on the warping function. And if the warping function can be properly fitted by finite combination of the functional principal components, this pair of quasi-static strain measurement is linearly correlated principally under thermal effects. Hence, temperature-induced strain can be accurately and

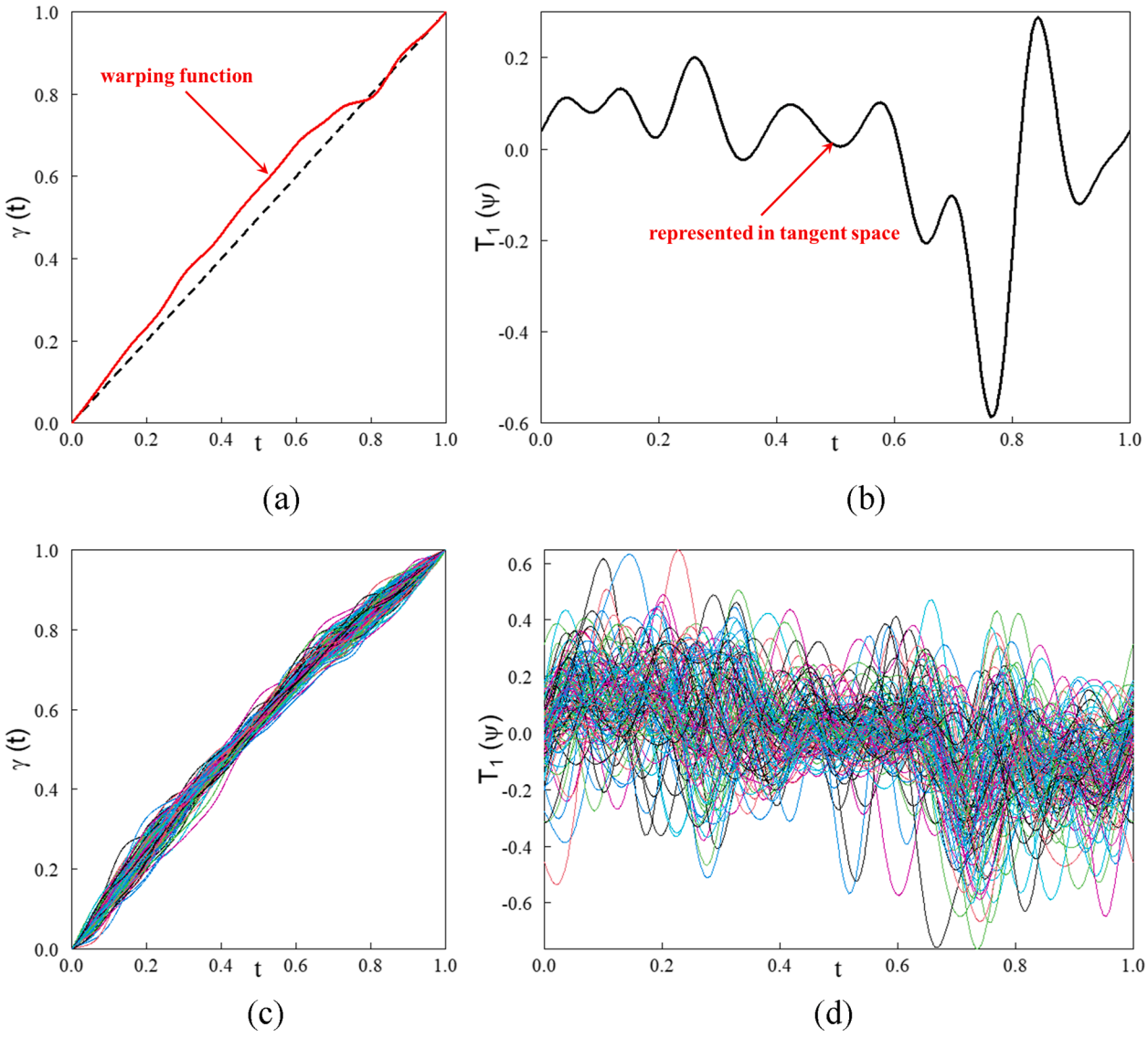


Fig. 9. Warping functions for 131 daily strain measurements and corresponding representation in tangent space: (a) warping functions for one daily strain measurements; (b) warping functions projected onto tangent space for one daily strain measurements; (c) warping functions for 131 daily strain measurements; (d) warping functions projected onto tangent space for 131 daily strain measurements.

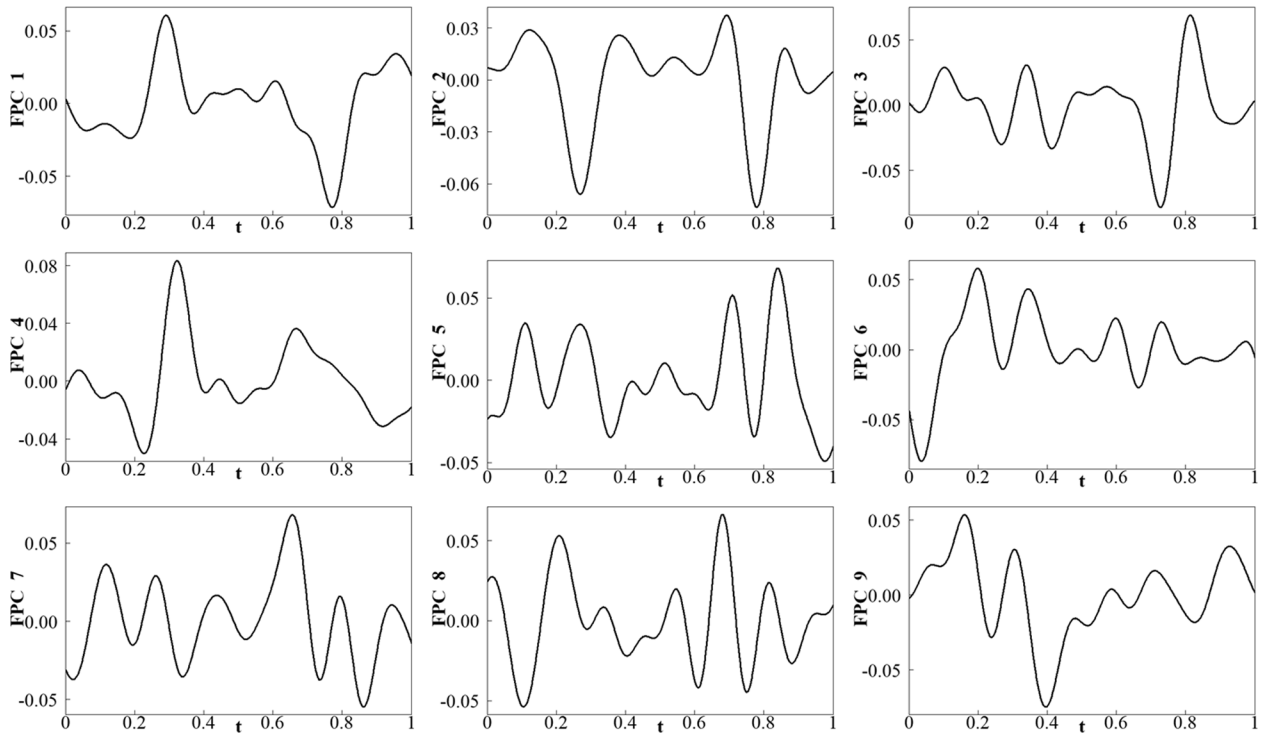


Fig. 10. Visualization for first 9 principal components of warping functions.

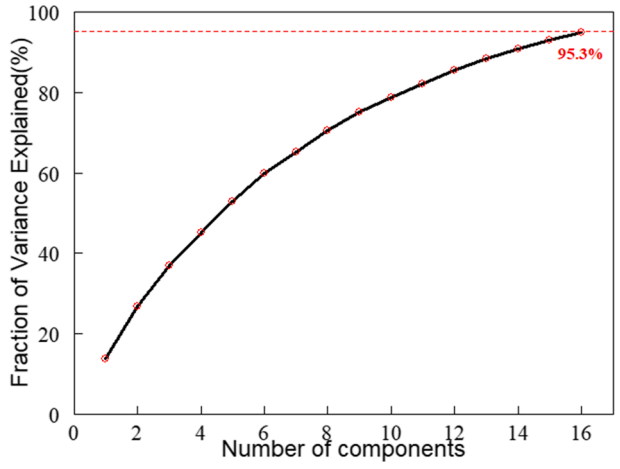


Fig. 11. Cumulative fraction of variance explained by principal components.

Table 1
Variation explained by each principal component.

PC	Fraction (%)	PC	Fraction (%)	PC	Fraction (%)	PC	Fraction (%)
1	13.8	**5**	7.6	**9**	4.7	**13**	2.8
2	12.9	**6**	6.9	**10**	3.6	**14**	2.6
3	10.3	**7**	5.5	**11**	3.5	**15**	2.3
4	8.3	**8**	5.2	**12**	3.3	**16**	2.1

elegantly eliminated, and further analysis such as correlations for traffic-induced strain and early warning for the bridge can be performed.

4. Discussion

Properly and parsimoniously modeling the relationship for strain measurement is imperative, especially when there are no decent ways to estimate the time-lag effect which weakens the linearity among strains and disturbs the structural analysis. It is the foundation for many other studies in SHM and should be paid adequate attention. In this paper, correlation for temperature-induced strain is modeled based on SRSF framework under Bayesian inference via functional data analysis. Time-lag effect is estimated by warping function and the deformation pattern of time-lag effect has also been discussed by utilizing the tool of FPCA. Conclusions can be summarized as follows:

1. First, field strain measurements can be regarded as functional data observations and FDA can be performed. The main goal is to seek the warping function, which represents the time-lag effect, so that the phase variation between a pair of temperature-induced strains can be eliminated. SRSF framework under Bayesian inference are then introduced, and corresponding warping function is searched in tangent space. After the warping function being optimally searched, time-lag between temperature-induced strain disappears, and a strong linearity correlation emerges.
2. In order to understand the deformation pattern of the warping function, FPCA is performed in tangent space. Hundreds of pairs of daily strain measurements are used to conduct this dimension reduction technique, and finally the variation mode of warping function can be represented and reconstructed by first several PCs, which explain over 95% of variance. In this way, any warping function between a pair of strain measurements can be fitted by these PCs. If the estimated fitting curve superimposes the ground true correlation plot (time-lag effect has been eliminated) or warping function curve, it implies this pair of strain measurements is only

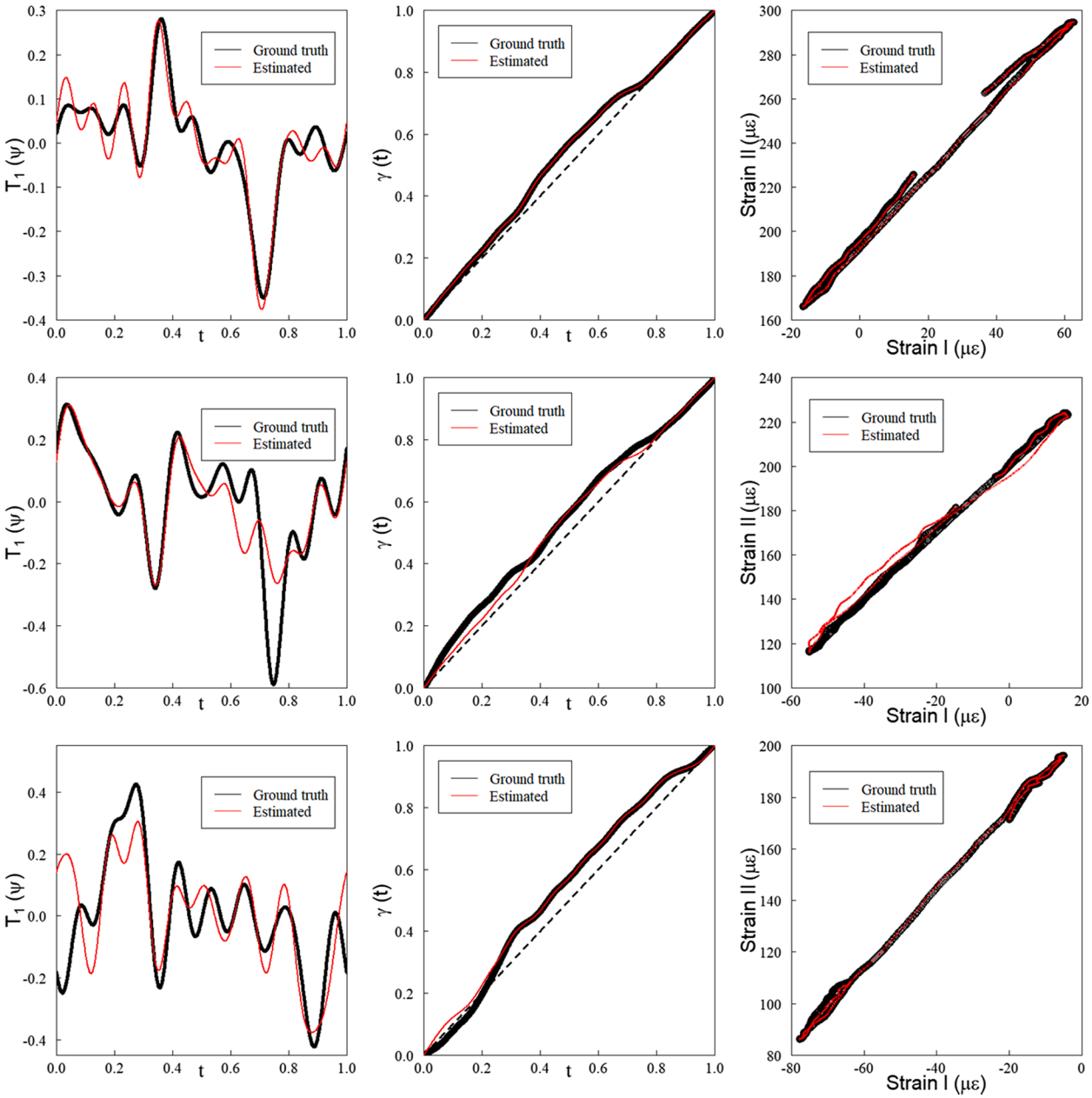


Fig. 12. Ground truth warping functions and correlations compared to corresponding estimated warping functions and correlations determined by FPCA.

affected by the thermal effects, so that thermal effect and time-lag effect can be accurately and elegantly modeled, eliminated and further analysis can be carried out.

CRediT authorship contribution statement

Huachen Jiang: Methodology, Software, Writing - original draft. **Chunfeng Wan:** Writing - review & editing, Supervision, Conceptualization, Resources. **Kang Yang:** Validation. **Youliang Ding:** Funding acquisition. **Songtao Xue:** Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by the Fund for Distinguished Young Scientists of Jiangsu Province (BK20190013).

References

[1] H. Li, J.P. Ou, The state of the art in structural health monitoring of cable-stayed bridges, J. Civil Struct. Health Monit. 6 (2016) 43–67.
[2] H. Li, J.P. Ou, X.G. Zhang, et al., Research and practice of health monitoring for long-span bridges in the mainland of China, Smart Struc. Syst. 15 (2015) 555–576.
[3] Y.Q. Bao, Z.C. Chen, S.Y. Wei, et al., The state of the art of data science and engineering in structural health monitoring, Engineering. 5 (2019) 234–242.
[4] M. Azimi, A. Eslamlou, G. Pekcan, Data-driven structural health monitoring and damage detection through deep learning: state-of-the-art review, Sensors. 20 (10) (2020).
[5] Y.L. Ding, G.X. Wang, Y. Hong, et al., Detection and localization of degraded truss members in a steel arch bridge based on correlation between strain and temperature, J. Perform. Constr. Facil 31 (5) (2017).
[6] K. Yang, Y.L. Ding, P. Sun, et al., Modeling of temperature time-lag effect for concrete box-girder bridges, Appl. Sci. 9 (16) (2019).

[7] Y.Q. Ni, H.W. Xia, K.Y. Wong, et al., In-service condition assessment of bridge deck using long-term monitoring data of strain response, J. Bridge Eng. 17 (6) (2012) 876–885.
[8] F.N. Catbas, M. Susoy, D.M. Frangopol, Structural health monitoring and reliability estimation: long span truss bridge application with environmental monitoring data, Eng. Struct. 30 (9) (2008) 2347–2359.
[9] Z.C. Chen, H. Li, Y.Q. Bao, Analyzing and modeling inter-sensor relationships for strain monitoring data and missing data imputation: a copula and functional data-analytic approach, Struct. Health Monit. 18 (4) (2019) 1168–1188.
[10] G.X. Wang, Y.L. Ding, P. Sun, et al., Assessing static performance of the Dashengguan Yangtze bridge by monitoring the correlation between temperature field and its static strains, Math. Prob. Eng. 2015 (2015) 946907.
[11] C. Chen, Z.L. Wang, Y.H. Wang, et al., Reliability assessment for PSC box-girder bridges based on SHM strain measurements, J. Sensors. 2017 (2017) 8613659.
[12] Z.C. Chen, Z.Y. Tang, J.H. Chen, et al., Clarifying and quantifying the geometric correlation for probability distributions of inter-sensor monitoring data: a functional data analytic methodology, Mech. Syst. Signal Pro. 138 (2020) 106540.
[13] Y.F. Duan, Y. Li, Y.Q. Xiang, Strain-temperature correlation analysis of a tied arch bridge using monitoring data, in: 2011 International Conference on Multimedia Technology, IEEE, New York, 2011, pp. 6025–6028.
[14] S.L. Li, H. Li, J.P. Ou, et al., Integrity strain response analysis of a long span cable-stayed bridge, Key Eng. Mater. 413–414 (2009) 775–783.
[15] D.H. Yang, T.H. Yi, H.N. Li, et al., Correlation-based estimation method for cable-stayed bridge girder deflection variability under thermal action, J. Perform. Constr. Facil 32 (5) (2018) 04018070.
[16] T. Guo, J. Liu, Y.F. Zhang, et al., Displacement monitoring and analysis of expansion joints of long-span steel bridges with viscous dampers, J. Bridge Eng. 20 (9) (2015).
[17] J.M.W. Brownjohn, P. Kripakaran, B. Harvey, et al., Structural Health Monitoring of short to medium span bridges in the United Kingdom, Struct. Monit. Maint. 3 (3) (2016) 259–276.
[18] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, Springer.
[19] J.O. Ramsay, G. Hooker, S.Graves, Functional Data Analysis with R and Matlab, Springer.
[20] J.D. Tucker, W. Wu, A. Srivastava, Generative models for functional data using phase and amplitude separation, Comput. Stat. Data Anal. 61 (2013) 50–66.
[21] A. Srivastava, E. Klassen, S. Joshi, et al., Shape analysis of elastic curves in Euclidean spaces, IEEE Trans. Pattern Anal. Mach. Intell. 33 (7) (2011) 1415–1428.
[22] S. Kurtek, A. Srivastava, E. Klassen, Statistical modeling of curves using shapes and related features, J. Am. Stat. Assoc. 107 (499) (2012) 1152–1165.
[23] A. Kneip, J.O. Ramsay, Combining registration and fitting for functional models, J. Am. Stat. Assoc. 103 (483) (2008) 1155–1165.
[24] Y. Lu, R. Herbei, S. Kurtek, Bayesian registration of functions with a Gaussian process prior, J. Comput. Graph. Stat. 26 (4) (2017) 894–904.
[25] S. Kurtek, A geometric approach to pairwise Bayesian alignment of functional data using importance sampling, Electronic J. Stat. 11 (1) (2017) 502–531.
[26] H.W. Zhao, Y.L. Ding, S. Nagarajaiah, et al., Behavior analysis and early warning of girder deflections of a steel-truss arch railway bridge under the effects of temperature and trains: case study, J. Bridge Eng. 24 (1) (2019) 05018013.