Smart Structures and Systems, *Vol. 33, No. 3 (2024) 189-199*
https://doi.org/10.12989/sss.2024.33.3.189

Operational performance evaluation of bridges using autoencoder neural network and clustering

Huachen Jiang [1], Liyu Xie [2], Da Fang [3], Chunfeng Wan*[3], Shuai Gao [3], Kang Yang [4], Youliang Ding [3] and Songtao Xue**[2,5]

[1] *Shanghai Key Laboratory of Engineering Structure Safety, SRIBS, Shanghai 200032, China*
[2] *Department of Disaster Mitigation for Structures, Tongji University, Shanghai 200092, China*
[3] *Southeast University, Key Laboratory of Concrete and Prestressed Concrete Structure of Ministry of Education, Nanjing 210096, China*
[4] *School of Railway Transportation, Shanghai Institute of Technology, Shanghai 201418, PR China*
[5] *Department of Architecture, Tohoku Institute of Technology, Sendai, Miyagi 982-8577, Japan*

(*Received December 29, 2021, Revised January 6, 2023, Accepted January 25, 2024*)

Abstract. To properly extract the strain components under varying operational conditions is very important in bridge health monitoring. The abnormal sensor readings can be correctly identified and the expected operational performance of the bridge can be better understood if each strain components can be accurately quantified. In this study, strain components under varying load conditions, i.e., temperature variation and live-load variation are evaluated based on field strain measurements collected from a real concrete box-girder bridge. Temperature-induced strain is mainly regarded as the trend variation along with the ambient temperature, thus a smoothing technique based on the wavelet packet decomposition method is proposed to estimate the temperature-induced strain. However, how to effectively extract the vehicle-induced strain is always troublesome because conventional threshold setting-based methods cease to function: if the threshold is set too large, the minor response will be ignored, and if too small, noise will be introduced. Therefore, an autoencoder framework is proposed to evaluate the vehicle-induced strain. After the elimination of temperature and vehicle-induced strain, the left of which, defined as the model error, is used to assess the operational performance of the bridge. As empirical techniques fail to detect the degraded state of the structure, a clustering technique based on Gaussian Mixture Model is employed to identify the damage occurrence and the validity is verified in a simulation study.

Keywords: autoencoder; early warning; Gaussian mixture model; strain measurement; structural health monitoring; strain measurement

1. Introduction

With increasingly structural health monitoring (SHM) systems being employed on bridges in recent decades (Bao *et al*. 2019, Cross *et al*. 2013, Fan *et al*. 2019, Wan and Ni 2018, Peng *et al*. 2022), managers and stakeholders are able to better understand the in-service condition of the bridge structure, so that timely warnings can be issued and proper maintenance plan will be implemented. The in-service operational performance of bridge has long been the focus for both researchers and managers because each bridge is subjected to varying load conditions in the long term (e.g., ambient temperature load and vehicle load). Any degradation accumulated will inevitably impair the structure to a certain extent, and to timely identify and quantify the degradation level is regarded of great importance.

Among all the measurements collected by in situ SHM systems, acceleration and strain responses are most widely adopted to detect the damage occurrences or to evaluate the structure condition (Amezquita-Sanchez and Adeli 2016, Doebling *et al*. 1998, Nagarajaiah and Yang 2017, Noel *et al*. 2017, Ko and Ni 2005). In practice, modal characteristics are more sensitive to the environmental and operational variations (Sohn 2007, Xia *et al*. 2012), thus many researches have been devoted to investigate the relationships between dynamic properties and varying load conditions. For example, Li *et al*. (2010) employed the nonlinear principal component analysis and neural network technique to model the relationships between environmental loadings and structural modal parameters. Zhang *et al*. (2017) utilized Gaussian process regression and ridge regression to correlate the modal frequency and varying conditions of a cable-supported bridge. Shajihan *et al*. (2022) respectively adopted time series-based and image-based deep learning technique, i.e., LSTM and CNN, to mitigate the problem of sensor fault or malfunctioning of accelerometer. Although many successful cases have been implemented based on acceleration responses to evaluate the structure condition, many researchers also suggested some problems exist. As is often the case, modal parameters fail to be accurately identified because of sensor selection (Kuok and Yuen 2016), nonlinear vibration of the structure

*Corresponding author, Ph.D., Professor,
E-mail: wan@seu.edu.cn
**Co-corresponding author, Ph.D., Professor,
E-mail: xue@tongji.edu.cn

ISSN: 1738-1584 (Print), 1738-1991 (Online)

(Zapico-Valle *et al*. 2013) and some other reasons. Furthermore, modal analysis represents the integral health status of the structure rather than the local properties (Farrar *et al*. 2000, Huang *et al*. 2018), which poses a challenge when the exact location of the degradation of the structure is concerned.

Gradually, strain measurement partly took the place of vibration response in condition assessment owing to its wealth of information about stress experiences during the operation process of the structure, where vibration measurement is not able to offer (Cardini and DeWolf 2009, Costa and Figueiras 2012, Ni *et al*. 2012). For instance, Ding (2017) investigated the correlation between strain measurement and temperature with the aim of detecting and localizing the degraded member of a steel arch bridge. Wang *et al*. (2015) correlated the static strains and temperature field via multivariate linear regression combined with principal component analysis to evaluate the static performance of a truss bridge. Huang *et al*. (2018 and 2020) proposed an early warning method for a cable-stayed bridge based on a novel representation of strain measurement, namely canonically correlation model. Xia *et al*. (2017) evaluated the condition of a suspension bridge when collision happened using temperature-induced strain data. Yarnold and Moon (2015) created a unique numerical strain baseline caused by thermal loads through long term observations in normal operation to ensure the serviceability of the structure. In addition, a comparison study with vibration-based approach was conducted and results suggested the strain-based approach is more sensitive for normal operational changes.

In this article, an early warning method using strain measurement is proposed to assess the operational performance of a real concrete bridge. In contrast to methods mentioned above which only consider thermal actions imposed on the structure, each strain component (i.e., temperature-induced strain and vehicle-induced strain) is identified and quantified with great emphasis in the rest of the paper. Generally, vehicle-induced response contributes a major part to the strain variation, however in practice, vehicle load is difficult to measure or infer, and so does the vehicle-induced responses like strain or deflection (Huang *et al*. 2015, Zhao *et al*. 2020, Chen *et al*. 2019). Therefore, conventional researches inclined to only analyze temperature-induced strain and ignore other factors like wind and vehicles which would make analysis complex if taken into account. Considering the importance of the live load actions, i.e., vehicle load, strain component caused by the vehicle is extracted through a latent variable model based on a deep learning technique, namely autoencoder. The extraction is followed by the elimination of the thermal effect at first, realized by wavelet packet decomposition technique. Then, an early warning method for identifying and classifying the degradation state is proposed based on clustering analysis. Subsequently, the proposed early warning identification method is verified in a simulation study and is compared with some other statistical methods. Finally, the study is concluded with more details and emphases contained in this paper. The flowchart of the paper is presented in Fig. 1.

2. Problem formulation and method proposed

During the normal operation of bridges, measurement recorded by a strain sensor reflects the strain variation under varying excitations like daily temperature change and live load forces of moving vehicles. Hence, the raw strain measurement can be decomposed and evaluated according to the source of excitation. Only after comprehending the composition of the strain measurement, can a reasonable and reliable early warning method be put forward. Consequently, let $S = S_{Tem} + S_{Veh} + e$ denote any strain measurement in SHM system, where S_{Tem} represents temperature-induced strain, S_{Veh} represents vehicle-induced strain and e represents model error that is least influenced by varying conditions, which is the core of degradation identification in later chapters (Huang *et al*. 2020).

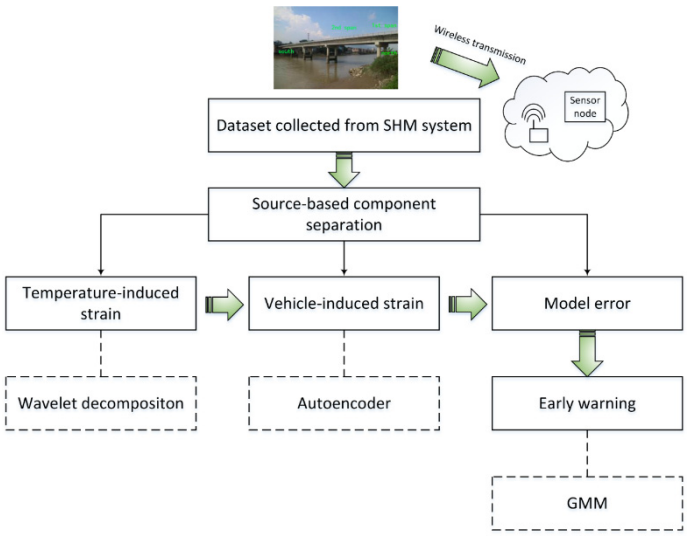

Fig. 1 Flowchart of the approach proposed in the paper

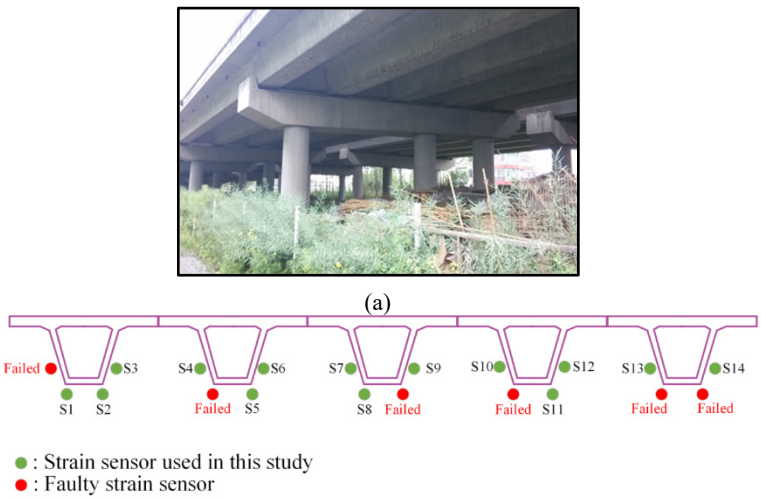

Fig. 2 Target structure and sensor deployment in this article: (a) photo of investigated bridge; and (b) sensor deployment of the investigated bridge

As described above, raw strain measurement is composed of three parts, and how to properly evaluate or extract each strain component is of great importance to structural condition assessment. Generally speaking, temperature is the primary and quasi-static load imposed on the bridge, thus strain stemming from it can be quantified using many conventional smoothing tools like empirical mode decomposition (Kim *et al*. 2021), Kalman filter (Petersen *et al*. 2018) and wavelet transform (Ni *et al*. 2012). In this work, the trend variation of strain sensors caused by temperature load is estimated and then eliminated by wavelet packet decomposition (Mallat 1989).

After S_{Tem} being eliminated, the remaining strain is composed of S_{Veh} and e, which is more similar to the vibration response as moving vehicle plays a leading role. Because both vehicle load and corresponding strain component are hard to measure or infer directly as alluded before, some latent variable models are adopted to extract strain component caused by moving vehicles, and the most common strategy is principal component analysis (PCA) technique (Huang *et al*. 2020, Yan *et al*. 2005a, b). The prime purpose of PCA is extracting the main information, i.e., vehicle-induced strain from original multivariate data by restricting the dimensionality to a finite number of components. The rationale behind this method is a linear transformation which projects the multivariate data into another space. Such an analysis is incompetent when deciphering highly nonlinear and complex model, furthermore, the combination of certain principal components is hard to explain what exactly have been learned in the latent representation. Consequently, in this study, another latent variable model, namely autoencoder, is used to extract the most important feature in the remaining strain component, i.e., vehicle-induced strain. Autoencoder has an encoder-decoder architecture, in which encoder learns the latent representation of the input and decoder reconstructs the input. Trained by back-propagation, in the latent representation, the feature of vehicle-induced strain is mostly captured and noise is ignored. Actually, it shares the same function with the PCA: dimensionality reduction, although autoencoder offers a more accurate and explicable feature extraction and reconstruction and will be demonstrated in the rest of the paper.

The evaluation of the strain components under varying operational conditions is conducted on a real concrete bridge located in China, seen in Fig. 2(a). The target structure is a 5-span box-girder bridge with each span in length of 25 meters. The middle section in the main span of the bridge is installed with SHM system, with the aim of monitoring the in-service condition of the newly constructed bridge. The sensor deployment is demonstrated in Fig. 2(b), and 14 resistance strain sensors in total, are used to validate the proposed method in this paper. The sampling frequency of the wireless system is 1 Hz, which is able to capture both static and vehicle-induced responses.

3. Extraction of temperature-induced strain

It is believed that temperature-induced strain component is the variation trend of the raw measurement, because temperature is the most important quasi-static external load imposed on the bridge and its thermal effect has been investigated thoroughly. Hence, the evaluation of temperature-induced strain is implemented via wavelet packet decomposition. Wavelet analysis is a popular tool in non-stationary signals because of its powerful multiresolution capacity and has been embedded in many signal analysis toolboxes. Owing to the quasi-static property of the thermal effect, temperature-induced strain actually resides in the low-frequency band and is what wavelet analysis should focus on. In each level of decomposition, approximation coefficient, which represents low-frequency part, and detail coefficient, which represents high-frequency part, are separated. Then, the low-frequency part is decomposed into next level, and so forth. Finally, with proper selection of the decomposition level, the lowest-frequency part which reflects temperature-induced strain is

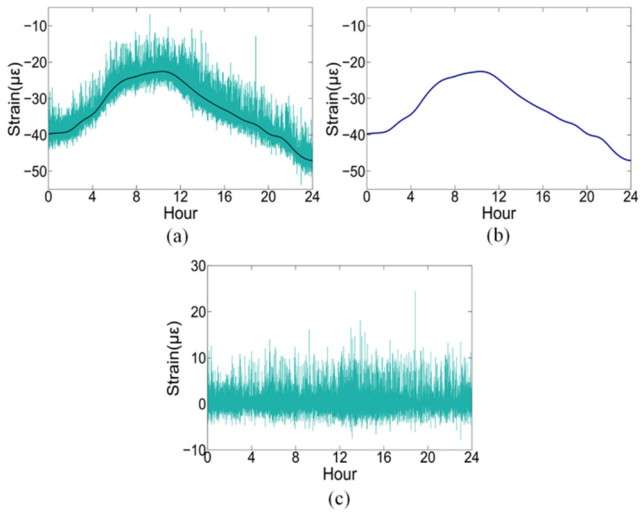

Fig. 3 Extraction of temperature-induced strain from raw measurements: (a) raw measurement of Sensor 1; (b) temperature-induced strain; and (c) remaining strain after the elimination of temperature-induced strain

extracted.

In this study, temperature-induced strain is successfully extracted when the number of decomposition level is 8 (where the temperature-induced strain component is relatively smooth), and the result is shown in Fig. 3. As the temperature effect presents prominent diurnal cycle (Yang *et al*. 2019), daily measured strain is investigated, and Fig. 2(a) shows a sample of daily strain measurement of Sensor 1. The corresponding extracted temperature-induced strain is demonstrated in Fig. 3(b), which is a smooth curve representing the variation trend of daily strain measurement. The remaining strain component which contains vehicle-induced strain and model error is depicted in Fig. 3(c), and in the next chapter, strain component under the effect of moving vehicles will be properly extracted.

In order to validate the feasibility of the wavelet analysis on the extraction of temperature-induced strain, some typical examples of field temperature measurements and extracted temperature-induced strain curves are illustrated in Fig. 4. Two curves both have a similar trend which corresponds to the variation of the temperature. However, the correlation is not complete linear, and there is some phase shift between these curves. This phenomenon most commonly occurs in the concrete bridges which was referred to as the time-lag effect in the literature. Researchers attributed this phenomenon to the material property of the concrete and the non-uniform temperature distribution. Generally speaking, strains measured at different locations will display distinct time lag albeit with similar trend. More documents related to the time lag effect are referred to (Jiang *et al*. 2021, Taysi and Abid 2015, Wang *et al*. 2020). Thus, in such a case, temperature-induced strain estimated by wavelet analysis, which shares similar trend with the temperature and differs merely on the phase is expected.

4. Extraction of vehicle-induced strain

After the successful elimination of temperature-induced strain, how to extract vehicle-induced strain is always troublesome. In traditional ways, such component is identified using manually setting threshold as shown in Fig. 5, however, the exact bound of the empirical threshold lacks theoretical derivation. If the threshold setting is too large, some actual responses caused by vehicles, though small, will be ignored. On the other hand, if the threshold is set to small, some noise will be introduced and will inevitably impede consequent analysis. In some cases, the ideal

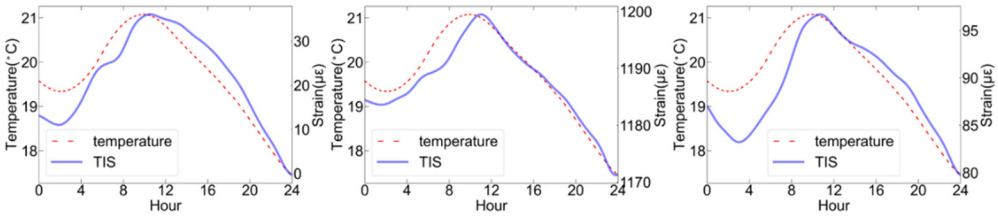

Fig. 4 Typical examples of temperature curve and extracted temperature induced strain: red line represents temperature curve and purple line represents temperature-induced strain (TIS)

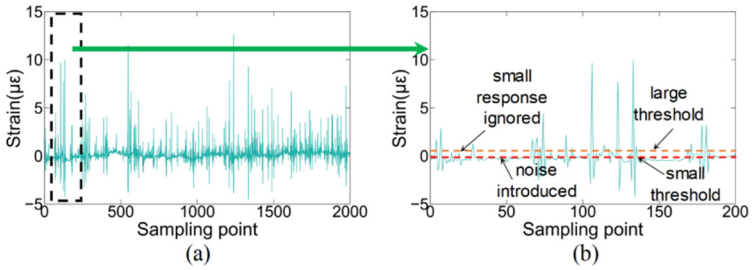

Fig. 5 Conventional method of extracting vehicle-induced strain: (a) remaining strain after the elimination of temperature-induced strain; and (b) threshold-based method of extracting vehicle-induced strain

threshold does not even exist, and the optimal threshold is given by minimizing misidentification of the vehicle-induced strain. Furthermore, it is impossible to accurately distinguish responses from noise if the response is too subtle when the moving vehicle is too far away from the measuring point.

Therefore, only focusing on the readings of a single strain sensor is ineffective to extract vehicle-induced strain. A more comprehensive and reasonable method is required. As vehicles cross bridges, wheels forces will cause a global vibration of a certain section or even the whole bridge. As a result, strain measurements at different locations are inter-connected with each other, and the closer the sensor locations are, the more resemblances they share in dynamic vibrations. Guided by this principle, it is believed that the actual response caused by vehicles resides in the strain field of the bridge, and the more the sensors are used, the more precise the result will be. Therefore, vehicle-induced strain can be regarded as a latent variable, also at the same time, the most significant variable mixed in the response. A latent variable analysis (LVA) is required to infer such component. In this study, a deep learning based LVA technique named autoencoder, is presented to extract vehicle-induced strain component.

4.1 Brief overview of autoencoder neural network

Autoencoder is an unsupervised deep learning method that maps the data into hidden layer and then reconstructs the input back (Bengio *et al*. 2007, Kingma and Welling 2014). The main purpose of the autoencoder is to learn the underlying manifold, or the latent representation of the data and reconstructs the inputs as the outputs. Typically, an autoencoder is composed of two parts: encoder architecture, which encodes the original high-dimensional data to low-dimensional latent space; and decoder architecture, which decodes the information learned in the latent space to original input space. A fundamental architecture of an autoencoder is schematically shown in Fig. 6. It is also noted that the input data shares the same size with the output data and an autoencoder does not need labels to enable training process. The mathematical background of the encoder-decoder architecture is

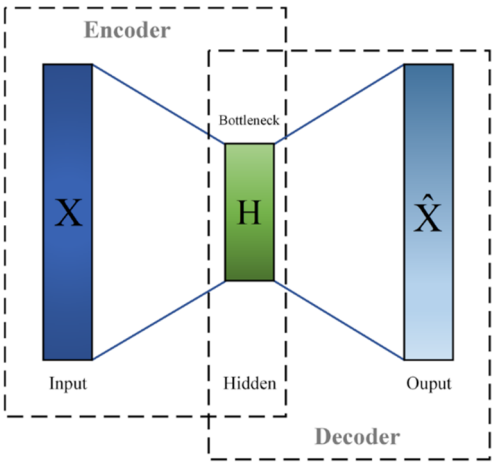

Fig. 6 Scheme of a basic autoencoder

$$\phi: X \to H \quad \text{(Encoder)} \tag{1}$$

$$\psi: H \to X \quad \text{(Decoder)} \tag{1}$$

$$\phi, \psi = arg \min_{\phi,\psi} \|X - (\psi \circ \phi)X\|^2 \tag{2}$$

In the above equations, ϕ is the encoder function which reduces the original data X to a latent space H, and ψ is the decoder function which reconstructs the latent space H to the output. Because the essential operation in a vanilla neural network is composition operation, denoted as $(\psi \circ \phi)$, and so does the autoencoder (Eq. (3)). If there are more than one hidden layer in the encoder and decoder, then, ϕ and ψ are also composition functions in each part.

4.2 Proposed autoencoder architecture

The crucial point of the autoencoder is not a simple copy from the input data to the output. Actually, the most significant features of the original data are what an autoencoder should learn from. In this article, vehicle-induced strain is the only concerned feature hidden in the strain component, and is what should be learned in the encoder architecture and reconstructed in the decoder architecture. In another word, the pure vehicle-induced strain without any other component is expected to be extracted. Considering the nonlinearity of the vehicle-induced strain field, a multilayer autoencoder, also known as stacked autoencoder is proposed to capture the highly nonlinear feature. In both the encoder and decoder architecture, there are two hidden layers to gradually transform the original data to the latent space or to map back.

In order to capture the vehicle-induced strain in accuracy, 14 sensors located in the same section of the bridge are used to implement the algorithm. To determine the hyper parameters of the hidden layers and hidden neurons, several parameter settings were experimented. A concise 2-layer architecture with high accuracy and efficiency was chosen to be the final model. The hidden features in the hidden layers are 8 and 4 respectively, both for the encoder and decoder part. It is noted that only the vehicle-induced strain feature is expected to be reconstructed, so in the latent space, only one hidden neuron which represents the latent representation of the vehicle-induced strain is considered. The proposed autoencoder architecture is illustrated in Fig. 7. Hitherto, the autoencoder shows a great superiority over other latent variable analysis like PCA, where the physical meaning of the latent representation (principal component) is hard to explain. In addition, how many principal components are needed is only based on the accumulated contribution of variance and also lacks intuitive comprehension.

In the encoder neural network, the arithmetic operation is like that in a vanilla neural network

$$h = \sigma(Wx + b), \tag{4}$$

where h is the hidden layer representation, $x \in \mathbb{R}^{14}$ is a realization of the original data X, σ is the element-wise

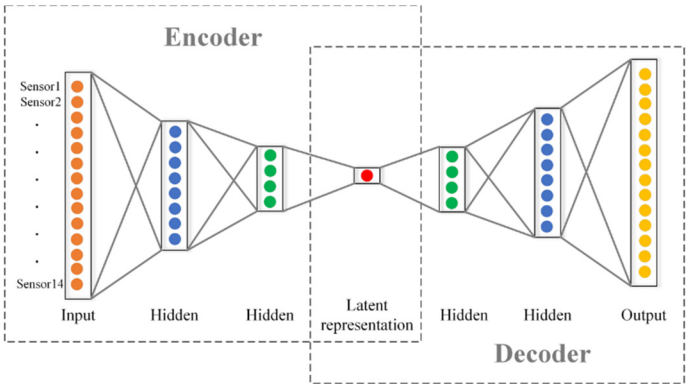

Fig. 7 Proposed deep stacked autoencoder architecture in this article

activation function (sigmoid function is used in the study), W is a weight matrix, and b is a bias vector. Eq. (4)) plies to the operation from the encoder to the latent representation.

Similarly, in the decoder architecture, the latent representation is mapped back to reconstruct X

$$x' = \sigma'(W'h + b'), \tag{5}$$

where σ', W' and b' share the same meanings with those in the encoder part, although with different values and functions being taken, and x' is a sample of reconstructed signal X' which shares the same size with X. Physically speaking, X' is the strain field of the bridge that only contains vehicle-induced component.

The proposed autoencoder is trained to minimize the reconstruction errors, so the loss function L can be denoted as

$$L(x, x') = \|x - x'\|^2 = \|x - (\psi \circ \phi)x\|^2 \tag{6}$$

To train the autoencoder is actually to select a proper encoder function ϕ and decoder function ψ through standard backpropagation procedure. Then, the latent representation of the vehicle-induced strain is $\phi(x)$, and the reconstructed component concerned about is $(\psi \circ \phi)x$.

4.3 Intelligent extraction of vehicle-induced strain by autoencoder

The vehicle-induced strain can be successfully extracted using aforementioned autoencoder technique, and the result for sensor 1 is shown in Fig. 8. It is obvious in Fig. 8(b) that the autoencoder captures all peaks and valleys which represent the actual responses caused by the vehicles. For small fluctuations which represent noise or condition related responses, the reconstructed values are almost 0. Due to the identification of the vehicle-induced strain is based on the inter-relationships among the strain field of the bridge through an autoencoder neural network, hence, for any strain sensor, the expected vehicle-induced strain is identified at the same time, in a similar pattern.

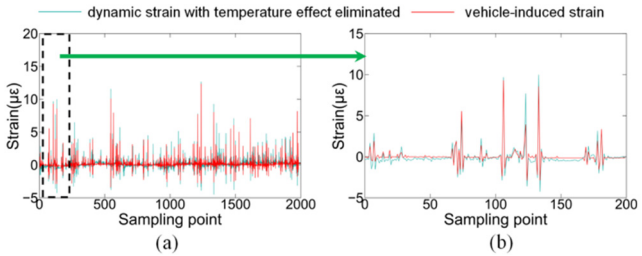

Fig. 8 Extraction of vehicle-induced strain from the remaining strain components: (a) 2000 sampling points demonstration; and (b) 200 sampling points demonstration in detail

More detailed demonstrations of the extracted vehicle-induced strain are shown in Fig. 9. Generally speaking, exact vehicle-induced strain extraction should be compared with the record of vehicle passing time. However, the SHM system deployed on this bridge does not include the vehicle tracking. Even without exact passing time, the vehicle-induced strain under normal operation can still be estimated since it is treated as a learnable feature instead of a threshold-based data. As data only collected from the same cross section is concerned, the occurrence of the vehicle-induced strain is the common feature shared by each sensor and can be cross validated. There is sufficient evidence to show that all strain sensors that constitute the strain field of the bridge share a similar pattern of variation. The autoencoder outputs all major fluctuations which represent responses caused by vehicles, and outputs zeros (or very approximate zeros) to eliminate responses regarded as noise. As is illustrated, sensors are closely connected or interrelated to each other although the interrelationship is mixed with linearity and nonlinearity as well. Because of the non-uniform variation among each sensor which reflects the nonlinearity, traditional correlation analysis which only focuses on the linearity is not recommended in vehicle-induced signals. Also, the disadvantages of the PCA method have also been listed above. As a result, the proposed autoencoder neural network, which takes the whole strain field into account, shows its capability in capturing complex and nonlinear feature among sensors. Furthermore, such a technique also solves the problem in distinguishing the actual small response from the environmental noise based on the interrelationships of the strain field, which may be difficult for manual distinguishing. Another interesting phenomenon in Fig. 9 is it seems that autoencoder does not fully reconstruct the peak values at some points. A plausible explanation is that autoencoder tries its best to reconstruct

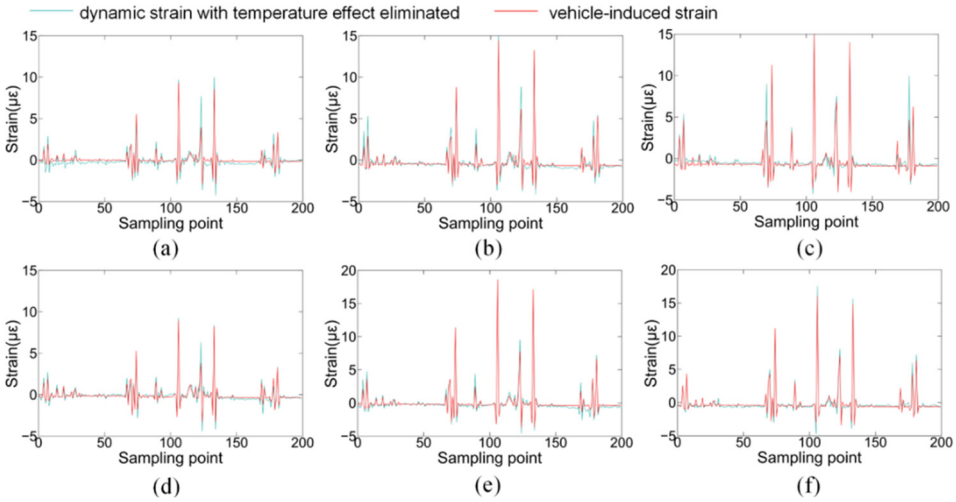

Fig. 9 More demonstrations of the extraction of vehicle-induced strain at a randomly given same interval: (a) Sensor 1; (b) Sensor 3; (c) Sensor 5; (d) Sensor 7; (e) Sensor 9; and (f) Sensor 11

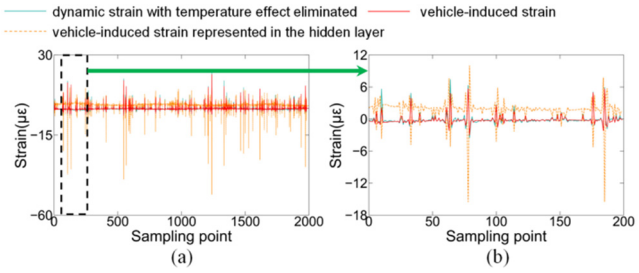

Fig. 10 Latent representation of vehicle-induced strain of Sensor 2 in proposed autoencoder framework: (a) 2000 sampling points demonstration, and (b) 200 sampling points demonstration in detail

original signals based on the implicit correlations among sensors. While having learned the correlations from historical data, the model outputs the estimated vehicle-induced strain which the particular sensor deserves. The remnant of the dynamic strain is then considered as uncertainty of the bridge performance for further processing in later chapters.

In order to exploit more information about how the autoencoder learns the significant feature (vehicle-induced strain) in the latent space, the latent representation of the strain field is studied and demonstrated in Fig. 10. As is can be seen, the latent representation generates much greater amplitude than the original signals, and in the meantime, the noise remains little variation. In another word, the latent representation actually gives more weight to the responses caused by vehicles, and less to the noise. Generally speaking, the autoencoder has "learned" the most important feature and what's shared among the strain field after weighing different strain component. Finally, the autoencoder outputs the reconstructed signals with minimal errors when giving more weight to the vehicle-induced strain and less to the noise.

5. Operational performance evaluation and early warning method

In the last two sections, responses caused by the temperature and vehicles have been studied and the corresponding strain components have been extracted. Hitherto, the left strain component is noise and condition-related response, and is named model error in the remaining part of this paper. Model error plays a key role in identifying the performance degradation of the bridge, because it is hardly influenced by the external loads like temperature and vehicles, and only reflects the internal condition of the bridge and environmental noise. Therefore, the model error which is the core in the operational performance evaluation is further investigated in the later analysis. Fig. 11 demonstrates some samples of model error after eliminating the effects of temperature and moving vehicles. As is illustrated, model error fluctuates around 0 and is similar to the white noise, which indicates the well-performing in-service condition of this newly constructed bridge. However, as is often the case, the condition-related signal is mixed within the environmental noise, and to accurately identify the degradation state from the environmental noise is a main obstacle. In addition, the normal condition state being misclassified to the degradation state is another annoying problem, and these issues are discussed and solved in this paper.

As operational performance of the bridge degrades, strain response of each sensor will correspondingly increase. Thus, the strain expression can be simulated as follows when degradation occurs (Huang *et al*. 2020)

$$S_{deg} = S + \varepsilon \cdot (S_{max} - S_{min}), \tag{7}$$

where S is the original reading of the strain measurement, S_{max} and S_{min} are the maximum value and the minimum value of the strain being investigated in one day, ε represents the degradation rate and reflects the extent of the degradation stage, and S_{deg} is the simulated strain that undergoes a manually controlled degradation.

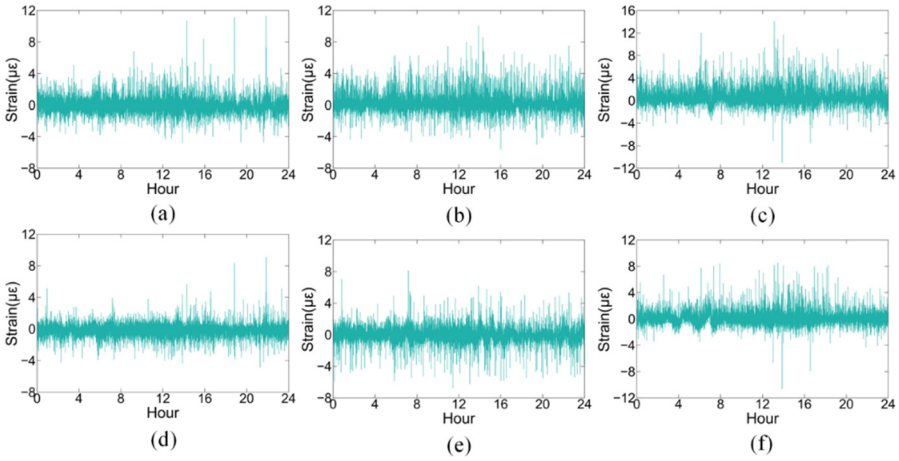

Fig. 11 The remaining model error after removing both temperature and vehicle effect for: (a) Sensor 1; (b) Sensor 3; (c) Sensor 5; (d) Sensor 7; (e) Sensor 9; and (f) Sensor 11

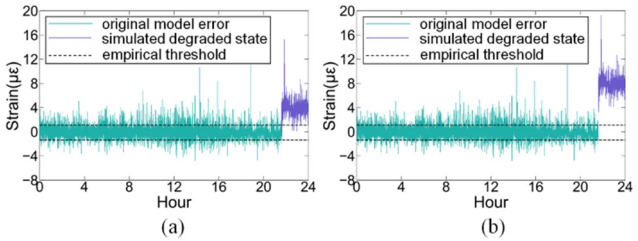

Fig. 12 Simulation of the degradation of the operational performance and empirical early warning method: (a) ε = 0.25; and (b) ε = 0.5

In order to investigate the internal degradation state of the structure, varying load effects, i.e., temperature and moving vehicles, should be first eliminated. Then, the model error is what's left to identify the degradation, and if the original strain increases because of the degradation, the corresponding model error will also increase. In Fig. 12, two examples of the simulated model error in the degraded state are demonstrated, with $\varepsilon = 0.25$ and $\varepsilon = 0.5$ respectively. As is seen from both examples, there is an obvious deviation from the original model error and depicted in purple. In order to distinguish the degraded state from the normal state, the empirical threshold for the early warning of the degradation is also illustrated. The range of the empirical setting includes most points of the normal state and excludes most points of the simulated degraded state, which seems efficient to identify the degradation. However, some spikes (almost 2.5% of the normal state data) are misclassified to the degraded state and will cause thousands of false alarms every day which is a very annoying problem. Therefore, not only accurately identify the degraded state, but reduce the number of false alarms is also important.

5.1 Gaussian mixture model for identifying the degradation state

The mixed model error is composed of two components or clusters, i.e., normal state and degradation state, and it is hard to distinguish them using threshold method because of the existence of fluctuations and noise. Due to the intrinsic character of the data distribution, it is intuitive to suppose that the two clusters are from different Gaussian distributions. Therefore, in this study, a clustering technique, namely Gaussian mixture model (GMM) is adopted to fit the data and then identify the degradation state.

For one dimensional data, the probability density function (pdf) for a Gaussian distribution is expressed as Eq. (8)

$$p(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{8}$$

where μ and σ respectively denote the mean and variance of the distribution. For a multivariate Gaussian distribution, the pdf is expressed as Eq. (9)

$$p(x|\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)|\Sigma|}} exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)) \tag{9}$$

where Σ is the covariance matrix of the distribution. Assume a GMM with K components (K = 2 in this article), then the pdf of the GMM is defined as a linear combination of all K distributions

$$p(x) = \sum_{k=1}^{K} \pi_k p(x|\mu_k,\Sigma_k) \tag{10}$$

where π_k is the mixing coefficient for each k-th component with the constraint that $\sum_{k=1}^{K}\pi_k = 1$,with the aim of normalization, and μ_k and Σ_k are estimated μ and Σ for each k-th distribution. Since the number of clusters K is already known, expectation maximization (EM) technique is commonly used to train the model and the iterative process converges with optimal parameters μ_k, Σ_k and π_k finally obtained.

5.2 Results of early warning application and comparison study

To demonstrate the applicability of the proposed GMM method, several cases with different degradation rate are investigated, and results are shown in Fig. 13. The tested rate is set to 10%, which means the proportion of the degraded state to the whole dataset is 10%. It can be concluded that (1) most points reside in the degradation state have been successfully identified, similar as traditional method based on the threshold setting; (2) with the increase of the degradation severity, the accuracy of the GMM method also increases; and (3) compare to the traditional method, GMM overwhelmingly reduces the number of false alarms and only falsely alarms with the probability of 0.05% around and it will inevitably reduce the labor cost to exclude the enormous false alarms.

It is noted that in Fig. 13(a), a strange pattern of classification result exists. In this case, the degradation rate is 0 which means the degradation does not occur. Actually, GMM outputs the probability of each point belonging to each state, and the corresponding probability is almost 50% for each point belonging to each state, and because of the arrangement of the dataset, such classification pattern forms. In another word, GMM cannot distinguish them and regard the whole dataset as a normal state. Furthermore, it is hard to distinguish the degradation state when the severity is less than 0.25 ($\varepsilon \leq 0.25$), and GMM outputs a similar pattern of classification result with Fig. 13(a) when $\varepsilon \leq 0.25$.

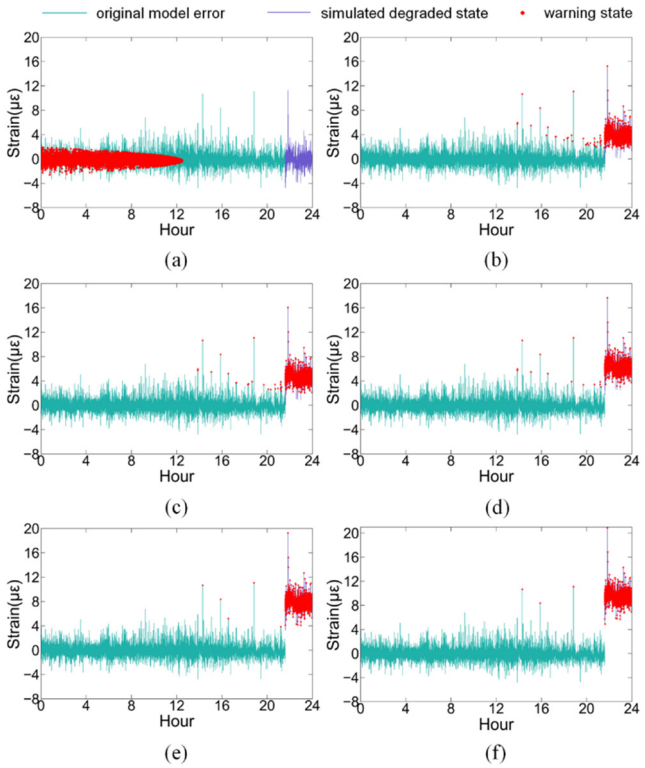

Fig. 13 Early warning results based on proposed gaussian mixture method when degradation occurs: (a) ε = 0; (b) ε = 0.25; (c) ε = 0.3; (d) ε = 0.4; (e) ε = 0.5; and (f) ε = 0.6

Table 1 Comparison of the proposed GMM method with empirical threshold, KNN, HB, and IF for detecting degraded state

Method	ε	TPR (%)	FPR (%)	Accuracy (%)
Proposed GMM method	0.25	99.79	0.05	99.94
	0.3	99.90	0.03	99.96
	0.5	99.99	6.43×10^{-5}	99.99
Empirical threshold	0.25	99.91	2.35	97.87
	0.3	99.94	2.35	97.88
	0.5	≈1	2.34	97.88
KNN	0.25	16.74	9.25	83.35
	0.3	17.01	9.22	83.40
	0.5	17.85	9.13	83.57
HB	0.25	6.59	6.17	85.11
	0.3	3.76	10.10	81.29
	0.5	6.17	0.38	90.27
IF	0.25	59.61	4.49	91.92
	0.3	61.44	4.28	92.29
	0.5	70.63	3.26	94.13

In order to highlight the superiority of GMM, comparison studies with other commonly used techniques are conducted. To better understand the performance of each method, some indexes and notations should first be defined: (1) TP is the number of points in degradation state successfully classified as in degradation; (2) FP is the number of points in normal state while classified as in degradation; (3) TN is the number of points in normal state successfully classified as in normal; (4) FN is the number of points in degradation state while classified as in normal; (5) TPR = TP/(TP+FN) is the TP relative to all points in degradation state, which indicates the ability to raise warnings when the degradation occurs, and this index should be as high as possible; (6) FPR = FP/(FP+TN) is the FP relative to all points in normal state, which indicate the ability to falsely alarm, and this index should be as low as possible; and (7) Accuracy = (TP+FN)/(TP+TN+FP+FN) is the overall accuracy of the model and should be as high as possible.

Results of the classification performance for different methods, namely GMM method, empirical threshold method, K-nearest neighbors (KNN) method, histogram-based (HB) method, and isolation forest (IF) method are demonstrated in Table 1 (binary classification with default parameter setting predetermined by built-in program). It is obvious that GMM and empirical threshold method share the highest score in TPR and Accuracy index. However, the major difference is the FPR index, where GMM shows its great advantage in few false alarms, similar to the conclusions drawn previously. The FPR of the empirical threshold is almost 47 times as that of GMM when ε = 0.25, and almost 3600 times when ε = 0.5. The result for the empirical threshold is unacceptable because its waste of considerable human resources to handle false alarms. The performance of GMM method gradually becomes better as

the severity of the degradation increases. When ε is 0.5, the TPR and Accuracy are almost 100% and the FPR is negligibly small. Finally, it can be concluded that GMM method is the optimal technique to identify the degraded state and to be applied in practical engineering.

6. Conclusions

In order to better understand the operational performance of the bridge, strain measurement, which has a great ability to decipher the health state of the bridge is thoroughly investigated in this paper. Therefore, strain components under varying load conditions, i.e., temperature and moving vehicles, are respectively identified and extracted. After eliminating the external effects, model error is what's left to evaluate the degradation state of bridge and the results have been discussed. The following conclusions can be drawn from this paper:

- Temperature-induced strain which represents the trend variation is extracted using wavelet packet decomposition technique. As the temperature-induced strain resides in the low-frequency band, it is properly reconstructed using approximation coefficient when the decomposition level is set to 8.
- Because the vehicle-induced strain is hard to measure or infer, a deep learning-based latent variable analysis, namely autoencoder is adopted. The autoencoder actually takes the interrelationships of the strain field into account, and reconstructs the strain field only based on the most significant feature learned, i.e., vehicle-induced strain and ignore the noise. The latent representation of the strain field is also demonstrated to verify the feature learned in the autoencoder.
- After the temperature and vehicle-induced strain being eliminated, the condition-related response, i.e., model error is the remaining strain component that is least influenced by the varying conditions. Model error plays a key role in detecting degradation state and is fully discussed in this paper. A GMM method is proposed to distinguish the degradation state from the normal one when the severity rate is over 0.25. Some simulation cases show that GMM-based method efficiently identifies the degradation state and significantly reduces the false alarms, which outperforms other commonly used techniques.

Acknowledgments

This work is supported by the National Key Research and Development Program of China (2021YFE0112200), Shanghai Municipal Transportation Commission (JT2023-KY-003), Key Research Support Project of SRIBS (KY10000038.20230065), the Japan Society for Promotion of Science (Kakenhi No. 18K04438), the Tohoku Institute of Technology research Grant, and the Housing & Urban-Rural Construction Commission of Shanghai Municipality (2023-002-029).

References

Amezquita-Sanchez, J.P. and Adeli, H. (2016), "Signal processing techniques for vibration-based health monitoring of smart structures", *Arch. Computat. Methods Eng.*, **23**(1), 1-15. https://doi.org/10.1007/s11831-014-9135-7

Bao, Y., Chen, Z., Wei, S., Xu, Y., Tang, Z. and Li, H. (2019), "The state of the art of data science and engineering in structural health monitoring", *Engineering*, **5**, 234-242. https://doi.org/10.1016/j.eng.2018.11.027

Bengio, Y., Lamblin, P., Popovici, D. and Larochelle, H. (2007), "Greedy layer-wise training of deep networks", *Adv. Neural Inform. Process. Syst.*, **19**, 153. https://www.iro.umontreal.ca/~lisa/pointeurs/BengioNips2006All.pdf

Cardini, A.J. and DeWolf, J.T. (2009), "Long-term structural health monitoring of a multi-girder steel composite bridge using strain data", *Struct. Health Monitor.*, **8**(1), 47-58. https://doi.org/10.1177/1475921708094789

Chen, Z.C., Li, H. and Bao, Y.Q. (2019), "Analyzing and modeling inter-sensor relationships for strain monitoring data and missing data imputation: a copula and functional data-analytic approach", *Struct. Health Monitor.*, **18**(4), 1168-1188. https://doi.org/10.1177/1475921718788703

Costa, B.J.A. and Figueiras, J.A. (2012), "Evaluation of a strain monitoring system for existing steel railway bridges", *J. Constr. Steel Res.*, **72**, 179-191. https://doi.org/10.1016/j.jcsr.2011.12.006

Cross, E.J., Koo, K.Y., Brownjohn, J.M.W. and Worden, K. (2013), "Long-term monitoring and data analysis of the Tamar Bridge", *Mech. Syst. Signal Process.*, **35**(1-2), 16-34. https://doi.org/10.1016/j.ymssp.2012.08.026

Ding, Y.L., Wang, G.X., Hong, Y., Song, Y.S., Wu, L.Y. and Yue, Q. (2017), "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), 04017082. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001075

Ding, Y., Yang, K., Jiang, H., Wang, M. and Wan, C. (2022), "Data abnormal detection using bidirectional long-short neural network combined with artificial experience", *Smart Struct. Syst., Int. J.*, **29**(1), 117-127. https://doi.org/ 10.12989/sss.2022.29.1.117

Doebling, S.W., Farrar, C.R. and Prime, M.B. (1998), "A summary review of vibration-based damage identification methods", *Shock Vib. Digest*, **30**(2), 91-105. https://doi.org/10.1177/058310249803000201

Fan, G., Li, J. and Hao, H. (2019), "Lost data recovery for structural health monitoring based on convolutional neural networks", *Struct. Control Health Monitor.*, **26**(10), e2433. https://doi.org/10.1002/stc.2433

Farrar, C.R., Cornwell, P.J., Doebling, S.W. and Prime, M.B. (2000), "Structural health monitoring studies of the alamosa canyon and I-40 Bridges", Los Alamos National Lab., NM, USA.

Huang, H.B., Yi, T.H., Li, H.N. and Liu, H. (2018), "New representative temperature for performance alarming of bridge expansion joints through temperature-displacement relationship", *J. Bridge Eng.*, **23**(7), 04018043. https://ascelibrary.org/doi/10.1061/%28ASCE%29BE.1943-5592.0001258

Huang, H.B., Yi, T.H., Li, H.N. and Liu, H. (2020), "Strain-based performance warning method for bridge main girders under variable operating conditions", *J. Bridge Eng.*, **25**(4), 04020013. https://ascelibrary.org/doi/10.1061/%28ASCE%29BE.1943-5592.0001538

Jiang, H., Wan, C., Yang, K., Ding, Y. and Xue, S. (2021),

"Modeling relationships for field strain data under thermal effects using functional data analysis", *Measurement*, **177**, 109279. https://doi.org/10.1016/j.measurement.2021.109279

Kim, S.W., Yun, D.W., Park, D.U., Chang, S.J. and Park, J.B. (2021), "Estimation of live load distribution factor for a PSC I girder bridge in an ambient vibration test", *Appl. Sci.*, **11**(22), 11010. https://doi.org/10.3390/app112211010

Kingma, D.P. and Welling, M. (2014), "Auto-encoding variational bayes", *Proceedings of the 2nd International Conference on Learning Representations*. https://arxiv.org/abs/1312.6114v10

Ko, J.M. and Ni, Y.Q. (2005), "Technology developments in structural health monitoring of large-scale bridges", *Eng. Struct.*, **27**(12), 1715-1725. https://doi.org/10.1016/j.engstruct.2005.02.021

Kuok, S.C. and Yuen, K.V. (2016), "Investigation of modal identification and modal identifiability of a cable-stayed bridge with Bayesian framework", *Smart Struct. Syst., Int. J.*, **17**(3), 445-470. https://doi.org/10.12989/sss.2016.17.3.445

Li, H., Li, S., Ou, J. and Li, H. (2010), "Modal identification of bridges under varying environmental conditions: Temperature and wind effects", *Struct. Control Health Monitor.*, **17**(5), 495-512. https://doi.org/10.1002/stc.319

Mallat, S.G. (1989), "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Transact. Pattern Anal. Mach. Intell.*, **11**(7), 674-693. https://doi.org/10.1109/34.192463

Nagarajaiah, S. and Yang, Y.C. (2017), "Modeling and harnessing sparse and low-rank data structure: a new paradigm for structural dynamics, identification, damage detection, and health monitoring", *Struct. Control Health Monitor.*, **24**(1), e1851. https://doi.org/10.1002/stc.1851

Ni, Y.Q., Xia, H.W., Wong, K.Y. and Ko, J.M. (2012), "In-service condition assessment of bridge deck using long-term monitoring data of strain response", *J. Bridge Eng.*, **17**(6), 876-885. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000321

Noel, A.B., Abdaoui, A., Elfouly, T., Ahmed, M.H., Badawy, A. and Shehata, M.S. (2017), "Structural health monitoring using wireless sensor networks: a comprehensive survey", *IEEE Commun. Surveys Tutorials*, **19**(3), 1403-1423. https://doi.org/10.1109/COMST.2017.2691551

Peng, Z., Li, J. and Hao, H. (2022), "Long-term condition monitoring of cables for in-service cable-stayed bridges using matched vehicle-induced cable tension ratios", *Smart Struct. Syst., Int. J.*, **29**(1), 167-179. https://doi.org/10.12989/sss.2022.29.1.167

Petersen, Ø.W., Øiseth, O., Nord, T.S. and Lourens, E. (2018), "Estimation of the full-field dynamic response of a floating bridge using Kalman-type filtering algorithms", *Mech. Syst. Signal Process.*, **107**, 12-28. https://doi.org/10.1016/j.ymssp.2018.01.022

Shajihan, S.A.V., Wang, S., Zhai, G. and Spencer Jr, B.F. (2022), "CNN based data anomaly detection using multi-channel imagery for structural health monitoring", *Smart Struct. Syst., Int. J.*, **29**(1), 181-193. https://doi.org/ 10.12989/sss.2022.29.1.181

Sohn, H. (2007), "Effects of environmental and operational variability on structural health monitoring", *Philosoph. Transact. Royal Soc. A-Mathe. Phys. Eng. Sci*, **365**(1851), 539-560. https://doi.org/10.1098/rsta.2006.1935

Taysi, N and Abid, S. (2015), "Temperature distributions and variations in concrete box-girder bridges: experimental and Finite Element Parametric Studies", *Adv. Struct. Eng.*, **18**(4), 469-486. https://doi.org/10.1260/1369-4332.18.4.469

Wan, H.P. and Ni, Y.Q. (2018), "Bayesian modeling approach for forecast of structural stress response using structural health monitoring data", *J. Struct. Eng.*, **144**(9), 04018130. https://doi.org/10.1061/(ASCE)ST.1943-541X.0002085

Wang, G.X., Ding, Y.L., Sun, P., Wu, L.L. and Yue, Q. (2015), "Assessing static performance of the dashengguan yangtze bridge by monitoring the correlation between temperature Field and Its Static Strains", *Mathe. Probl. Eng.*, **2015**, 946907. https://doi.org/10.1155/2015/946907

Wang, H., Zhu, Q., Zou, Z., Xing, C., Feng, D. and Tao, T. (2020), "Temperature distribution analysis of steel box-girder based on long-term monitoring data", *Smart Struct. Syst., Int. J.*, **25**(5), 593-604. https://doi.org/ 10.12989/sss.2020.25.5.593

Xia, Y., Chen, B., Weng, S., Ni, Y.Q. and Xu, Y.L. (2012), "Temperature effect on vibration properties of civil structures: A literature review and case studies", *J. Civil Struct. Health Monitor.*, **2**(1), 29-46. https://doi.org/10.1007/s13349-011-0015-7

Xia, Q., Cheng, Y., Zhang, J. and Zhu, F. (2017), "In-service condition assessment of a long-span suspension bridge using temperature-induced strain data", *J. Bridge Eng.*, **22**(3), 04016124. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001003

Yan, A.M., Kerschen, G., DoBoe, P. and Golinval, G. (2005a), "Structural damage diagnosis under varying environmental conditions - Part I: A linear analysis", *Mech. Syst. Signal Process.*, **19**(4), 847-864. https://doi.org/10.1016/j.ymssp.2004.12.002

Yan, A.M., Kerschen, G., DoBoe, P. and Golinval, G. (2005b), "Structural damage diagnosis under varying environmental conditions - Part II: local PCA for non-linear cases", *Mech. Syst. Signal Process*, **19**(4), 865-880. https://doi.org/10.1016/j.ymssp.2004.12.003

Yang, K., Ding, Y., Sun, P., Zhao, H. and Geng, F. (2019), "Modeling of temperature time-lag effect for concrete box-girder bridges", *Appl. Sci.*, **9**(16), 3225. https://doi.org/10.3390/app9163255

Yarnold, M.T. and Moon, F.L. (2015), "Temperature-based structural health monitoring baseline for long-span bridges", *Eng. Struct.*, **86**, 157-167. https://doi.org/10.1016/j.engstruct.2014.12.042

Zapico-Valle, J.L., Garcia-Diegues, M. and Alonso-Camblor, R. (2013), "Nonlinear modal identification of a steel frame", *Eng. Struct.*, **56**, 246-259. https://doi.org/10.1016/j.engstruct.2013.04.026

Zhang, Y.L., Kurata, M. and Lynch, J.P. (2017), "Long-term modal analysis of wireless structural monitoring data from a suspension bridge under varying environmental and operational conditions: system design and automated modal analysis", *J. Eng. Mech.*, **143**(4), 04016124. https://ascelibrary.org/doi/10.1061/%28ASCE%29EM.1943-7889.0001198

Zhao, H., Ding, Y., Li, A., Ren, Z. and Yang, K. (2020), "Live-load strain evaluation of the prestressed concrete box-girder bridge using deep learning and clustering", *Struct. Health Monitor.*, **19**(4), 1051-1063. https://doi.org/10.1177/1475921719875630