Probabilistic Anomaly Trend Detection for Cable-Supported Bridges Using Confidence Interval Estimation

Xiang Xu; Zhen-Dong Qian; Qiao Huang; Yuan Ren; and Bin Liu

Abstract: To rate uncertainties within anomaly detection course for large span cable-supported bridges, a probabilistic approach is developed based on confidence interval estimation of extreme value analytics. First, raw signals from structural health monitoring system are pre-processed, including missing data imputation using moving time window mean imputation approach, and thermal response separation through multi-resolution wavelet-based method. Then, an energy index is extracted from time domain signals to enhance robust of detection performance. A resampling-based method, namely the bootstrap, is adopted herein for confidence interval estimation. Four confidence levels are defined for the anomaly trend detection in this study, namely 95%, 80%, 50% and 20%. Finally, the effectiveness of the proposed anomaly trend detection methodology is validated by using in-situ cable force measurements from the Nanjing Dashengguan Yangtze River Bridge. As a result, the four-level anomaly detection triggers are determined by using the confidence interval estimation based on cable force measurements in 2007, which are 58671, 48862, 42499 and 39035, respectively. Subsequently, three cases are presented, which are spike detection, overloading vehicle detection and snow disaster detection. Through the spike detection, it is verified that energy index is capable to tolerate signal spikes. Three overloading events are simulated to conduct overloading vehicle detections. As a result, the three overloading events are detected successfully associated with different confidences. Snow disaster is detected with a more than 80% confidence based on the field measurements during the snow storm time window.

Author keywords: large span bridges, anomaly detection, structural health monitoring, extreme value analysis, confidence interval estimation

1. Introduction

Large span bridges face potential threats every day owing to sophisticated outer loadings and aggressive environments (Liu et al., 2021). Although regular inspections are scheduled, it is difficult to deal with sudden incidents timely, where the sudden incidents include significant changes in

operational loadings (e.g., overloading vehicles) and structural capacity (structural damage) (Liu et al., 2020). On the other hand, structural health monitoring (SHM) systems are adopted to provide real-time screenings of structural responses, outer loadings and environmental factors. Moreover, SHM systems are often devised and installed on large-scale civil infrastructures worldwide, especially large span bridges, to afford quantitative information for further analytics (Fujino et al., 2019; Xu and Xia, 2011; Ou and Li, 2010). This makes it possible to conduct timely anomaly detection for large span bridges based on real-time measurements from SHM systems to ensure the operational and structural safety.

In order to avoid sudden structural failures, anomalous events, including sudden structural damages and accidents, are supposed to be detected as early as possible. Since SHM data lay a solid foundation for anomaly detection investigations, related studies were broadly implemented all over the world within the last few decades. Initially, dynamic fingerprints of structures (e.g., frequency) were used to detect structural damages since the dynamic characteristics were stable structural parameters, which in theory reflect changes in condition of structures. The natural frequencies, damping values, mode shapes and curvature mode shapes were explored to detect structural damages (Pandey et al., 1991; Salawu, 1997; Das et al., 2016). Although the effectiveness of dynamic parameter-based damage detection methods was verified in both theory and laboratory, challenges posed when applied to large-scale sophisticated structures. The most two prime factors are the contamination of noise and influence of environmental variations (Peeters et al., 2001). It is proved that variations of dynamic indexes caused by temperature effects were larger than those induced by structural damages for a cable-stayed bridge (Xu and Wu, 2007). Peeters et al. (2001) suggested to distinguish thermal effects from damage events when detecting damages.

In view of the limitations of dynamic characteristics in practical applications, researchers explored to use static indexes (e.g., strain) to carry out anomaly detection for infrastructures. Yu et al. (2016) took advantages of deflection time history of a beam for damage detection using wavelet transform and Lipschitz exponent, where the effectiveness of the method was verified by a model experiment. Hua et al. (2009) used the changes in cable forces for damage detection of a cable-stayed bridge based on the fact that damage occurring in the bridge girders would cause a redistribution of forces in stay cables, where the validity of the approach is illustrated by numerical studies. Ni et al. (2020) adopted expansion joint displacements to detect damages of expansion joints under Bayesian

context, where in-situ measurements acquired from a cable-stayed bridge were employed to validate the effectiveness of the proposed method. Similar to dynamic characteristics, static parameter-based anomaly detection is affected by environmental factors, especially temperature effects. Plenty of methodologies for thermal effect modeling and separation are developed, such as regression models (Ren et al., 2019; Kromanis and Kripakaran, 2014), wavelet transform (Xu et al., 2020a; Wu et al., 2014), blind source separation (Zhu et al., 2018), numerical model-based methods (Xu et al., 2019; Zhou and Sun, 2019), and Bayesian dynamic linear model (Wang et al., 2019). Considering the influence of thermal effects to static parameter-based anomaly detection, various temperature-driven anomaly detection approaches are developed. Xu et al. (2020b) proposed a two-level anomaly detection for a suspension bridge using girder deflection measurements, where multi-resolution wavelet-based method was used to separate thermal responses from recorded signals. Zhu et al. (2019) introduced the blind source separation to improve the performance of moving principal component analysis for anomaly detection by using strain data, where three cases were used to verify the effectiveness of the proposed methodology. Tome et al. (2020) and Fan et al. (2020) presented a strategy for early damage detection for a large cable-stayed bridge based on multivariate cointegration analysis and statistical process control, where the effects of environmental and operational variations were suppressed using cointegration analysis. Huang et al. (2020) proposed a strain-based anomaly detection method for bridge main girders, where the correlation relationship between temperature and strain of the main girder was established.

Majority of the aforementioned anomaly detection methodologies are based on deterministic manners, where the uncertainty inherent is not taken into account. However, uncertainties inevitably exist in the monitoring data induced by environmental variability, measurement noise and the estimated parameters. In this context, confidence interval estimation, instead of the point estimation, is devoted to determining the trigger for anomaly detection to consider the uncertainty inherent in the detection course. Compared with the point estimation, the confidence interval estimation specifies instead a range within which the parameter is estimated to lie (Efron, 1987; DiCiccio and Efron, 1996), which has been widely employed to between-subject designs (Loftus and Masson, 1994), clinical research (Schober and Vetter, 2020; Cho et al., 2020), distribution locational marginal price (Wei et al., 2020) etc. To the best of the author's knowledge, however, no investigation has been observed to investigate the anomaly detection of large span bridges in the context of confidence

interval estimation of extreme value analytics.

In this paper, a probabilistic anomaly detection approach for large span bridges is developed in the context of confidence interval estimation of extreme value analytics. The measured static response data from SHM systems are first pre-processed, where missing data are filled to achieve the continuity and thermal response is separated. Subsequently, the signal energy within a determined time window is extracted as the index to detect anomalous trend. The trigger associated with confidence coefficient for anomaly detection is determined based on confidence interval estimation of generalized Pareto distribution (GPD). Finally, the effectiveness of the proposed anomaly detection method is verified by using cable force monitoring data of the Nanjing Dashengguan Yangtze River Bridge, where three cases are presented.

2. Methodology for probabilistic anomaly trend detection

The general flowchart for probabilistic anomaly trend detection is demonstrated in Fig. 1. In the data pre-processing, raw signals from SHM systems are first processed using the moving time window method for missing data imputation, which enhances the continuity of recorded signals (Kalaycioglu et al., 2016; Nevalainen et al., 2009). Subsequently, the multi-resolution wavelet-based approach is adopted to address thermal effects based on the distinguished frequency bandwidths (Ni et al., 2012; Xu et al., 2020a). Once obtaining the pre-processed signals, the anomaly detection index is extracted as the energy within a certain time window, where two principal parameters (i.e., length of window and number of overlaps) need to be determined. Based on relatively long-term monitoring data, the GPD, one of extreme value analysis, is used to determine the trigger for anomaly trend detection, where confidence interval estimation is employed to rate the uncertainty.

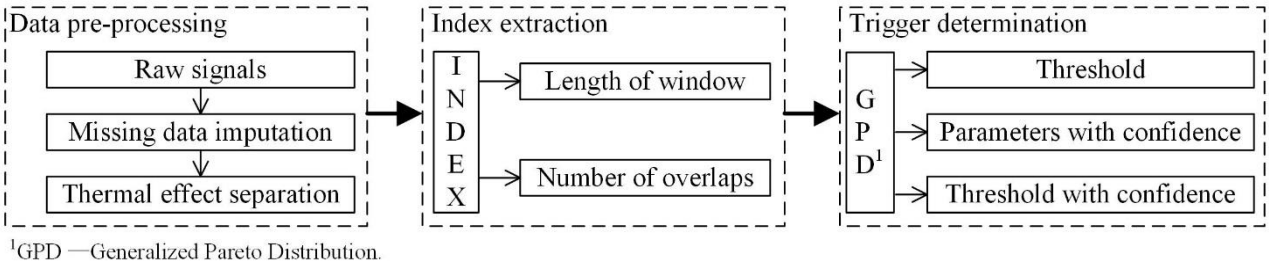


Fig. 1 Flowchart for probabilistic anomaly trend detection

2.1 Data pre-processing

Data missing is a common issue in data mining of SHM signals, leading to information loss or even

algorithm failures. According to the scale of missing data, methods subject to missing data imputation could be classified into two categories: large portion missing data imputation and small portion discrete missing data imputation. The maximum likelihood (Enders, 2001), artificial neural network (Martinez-Luengo et al., 2019), Bayesian inference approach (Lai et al., 2019) etc. are preferred to be used in the case of large portion missing data. For the small portion discrete missing data, the moving time window imputation (Hawthorne et al., 2005) is mostly employed in practice due to its brevity. During the course of pre-processing, only small portion missing data were observed. Thus, the moving time window mean imputation is adopted to address the data missing. The imputation value x_k subject to the missing location k is calculated as

$$x_k = \frac{1}{2n}\left(\sum_{i=k-n}^{k-1} x_i + \sum_{j=k+1}^{k+n} x_j\right) \tag{1}$$

where $2n$ is the length of moving time window, and x_i, x_j are the neighboring values of the missing point.

The multi-resolution wavelet-based method is applied to separate thermal responses from recorded data on the foundation of their distinguished frequency bandwidths. The decomposition level n is determined by

$$n \le \log_2\left(\frac{365\times 24\times 60\times 60 f_s}{2}\right) \tag{2}$$

where f_s is the sampling rate of the signal. The specific interpretation regarding the multi-resolution wavelet-based approach is described in our previous paper (Xu et al., 2020a).

2.2 Index extraction

In existing static response-based anomaly detection investigations, physical quantities (e.g., deflection) and their changes in form (e.g., cointegration residual) were always adopted as anomaly detection indexes owing to their straightforwardness and practicability (Xu et al., 2020b; Tome et al., 2020). However, these indexes are sensitive to signal spikes which are common phenomena for measurements of SHM systems, leading to false detection. Thus, an energy anomaly detection index is extracted from the measured data, which has tolerance to signal spikes.

The anomaly detection index is defined as the average energy within a determined time window, which is expressed as

$$DI(1)=\frac{\sum_{t=1}^{m}x^2(t)}{m},\qquad i=1$$
$$DI(i)=\frac{\sum_{t=(i-1)m+1-p}^{mi-p}x^2(t)}{m},\quad i>1 \tag{3}$$

where $x(t)$ is the time history, m is the length of window, and p is the number of overlaps.

There are two problems when determining the energy index in practical applications, including determination of the overlap and length of window. Considering the extracted indexes will be applied to predict anomaly detection trigger by using GPD that requires samples to be independent identically distributed, none overlap is employed. The length of window will influence the effectiveness of the anomaly detection. If the length of window is too short, the tolerance of the energy index to spikes is insufficient, while if the length is too large, the effectiveness of anomaly detection is weakened owing to the peak clipping of anomalous signals. In this context, the length of window should be determined by trade-off discussions subject to the time attribute of anomalous events. In theory, the optimal length of window is the minimum duration of all the potential anomalous events, which could be defined as

$$\triangle T=\min(t_1,t_2,\ldots,t_n) \tag{4}$$

where t_i is the duration of anomalous event i, and n is the number of potential anomalous events.

2.3 Trigger determination

In general, the trigger is the extreme value of the defined index (Liu et al., 2015). In this study, GPD is used to estimate the extreme values of indexes. Compared with the typical block maximum method, GPD takes full advantages of limited available information (Deng et al., 2018).

The cumulative GPD of a variate x takes the form

$$G(x;\sigma,\zeta)=\begin{cases}1-\left(1+\zeta\frac{x}{\sigma}\right)^{-1/\zeta}, & \zeta\neq 0\\ 1-\exp\left(-\frac{x}{\sigma}\right), & \zeta=0\end{cases}, \tag{5}$$

where σ is the scale parameter and ζ is the shape parameter.

To determine the quantile value corresponding to a T-year return period, data are required to be resampled to meet the request of GPD analysis in independent identically distribution. In this paper,

maxima of the index within 24 hours are adopted.

Based on the extracted daily maxima, an appropriate threshold is then determined for the GPD analytic. If the threshold is set too high, the number of out-of-sample is small, resulting in statistical uncertainty. On the other hand, if the threshold is too low, the excess quantity differs significantly from the maximum value, leading to a biased estimator. The mean excess function of the GPD is introduced to determine a proper threshold, which is

$$e(u)=\frac{1}{N_u}\sum_{i=1}^{N_u}(x_i-u)=\frac{\zeta}{1-\zeta}u+\frac{\sigma}{1-\zeta} \tag{6}$$

in which N_u denotes the number of excesses over the trigger, x_i is the i^{th} excess, and u is the threshold. The mean excess function is supposed to be a linear function of the excess quantity (Gilli, 2006). However, the transition from the curve to the straight line is not a point but an interval. To overcome the shortcoming, the standardized residual is introduced (Zhou et al., 2017) as

$$S_r(u)=e(u)-L_e(u)+S_d\left[e(u)\right] \tag{7}$$

where $L_e(u)$ is the linear fit of the mean residual life above the threshold, and $S_d[e(u)]$ is the standard deviation of the $e(u)$. There is an optimal threshold corresponding to the lowest value of standardized residual to achieve the balance.

To consider the uncertainty in the parameter determination process, the confidence interval estimation is used to estimate probabilistic triggers (Kysely, 2010). A re-sampling method, the bootstrap, is adopted for interval estimation (Castillo and Hadi, 1997). The specific steps using the bootstrap to estimate the shape parameter ζ are summarized as follows (Chen et al., 2017):

(1) estimated shape and scale parameters $\hat{\zeta}$ and $\hat{\sigma}$ are first predicted using a point estimation method (e.g., the maximum likelihood estimation, the probability moment estimation or the L moment method) based on a given dataset of size n;

(2) form the estimated standard error of $\hat{\zeta}$, denoted as $\hat{S}_e\left(\hat{\zeta}\right)$;

(3) generate B bootstrap samples each with size n from GPD($\hat{\zeta}$, $\hat{\sigma}$);

(4) for each bootstrap sample, estimate ζ using the point estimation approach as $\hat{\zeta}_b$ and compute

$t_b = \dfrac{\left(\hat{\zeta}_b - \hat{\zeta}\right)}{\hat{S}_e\left(\hat{\zeta}_b\right)}$, where $\hat{S}_e\left(\hat{\zeta}_b\right)$ is the estimated standard error of $\hat{\zeta}_b$;

(5) the equal-tailed 100(1-α)% confidence interval for ζ is $\left[\hat{\zeta} - t_{1-\alpha/2}\hat{S}_e\left(\hat{\zeta}\right), \hat{\zeta} - t_{\alpha/2}\hat{S}_e\left(\hat{\zeta}\right)\right]$, where t_α is the α percentile point of t_b's.

The procedure of estimation of scale parameter confidence interval is similar.

Finally, the quantile value is estimated as the anomaly detection trigger. Within the reference period of T years, the cumulative probability p corresponding to a certain guarantee rate P_r is

$$p = 1 - \sqrt[T]{P_r} \tag{8}$$

The quantile value subject to a 100(1-α)% confidence coefficient is

$$x_{p,1-\alpha} = u_0 + \frac{\hat{\sigma}_{1-\alpha}}{\hat{\zeta}_{1-\alpha}}\left[\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\zeta}_{1-\alpha}} - 1\right] \tag{9}$$

where u_0 is the threshold, $\hat{\sigma}_{1-\alpha}$ is the estimated scale parameter with 100(1-α)% confidence coefficient, $\hat{\zeta}_{1-\alpha}$ is the estimated shape parameter with 100(1-α)% confidence coefficient, n is the number of samples, and N_u is the number of excesses.

Generally, the trigger is defined as the quantile value corresponding to a 95% guarantee rate within a 100-year reference period (i.e., a return period of 1950 years).

3. Case study

3.1 The Nanjing Dashengguan Yangtze River Bridge and its SHM System

The Nanjing Dashengguan Yangtze River Bridge, a vital transportation link, crosses the Yangtze River and connects Liuhe District with Nanjing City, whose site-plan is shown in Fig. 2. The steel cable-stayed bridge has a total length of 1288m, where the main span is 648m. The configuration of the Nanjing Dashengguan Yangtze River Bridge is shown in Fig. 3. The bridge deck is supported by a total of 167 stay cables, and each cable consists of 109 to 241 wires of a 7mm diameter.

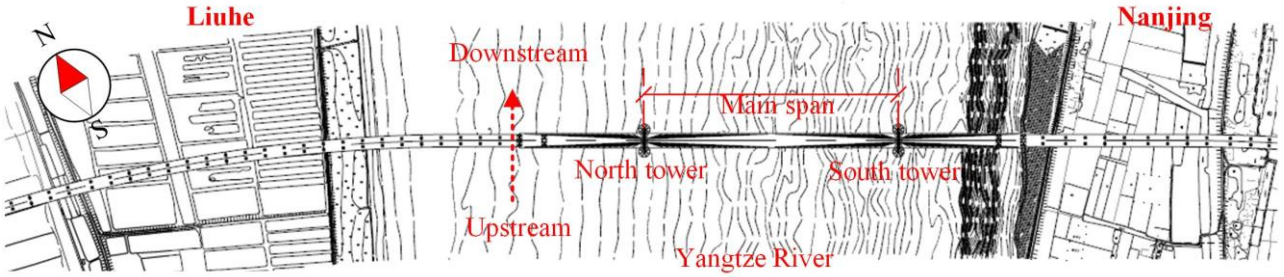


Fig. 2 Site plan of the Nanjing Dashengguan Yangtze River Bridge

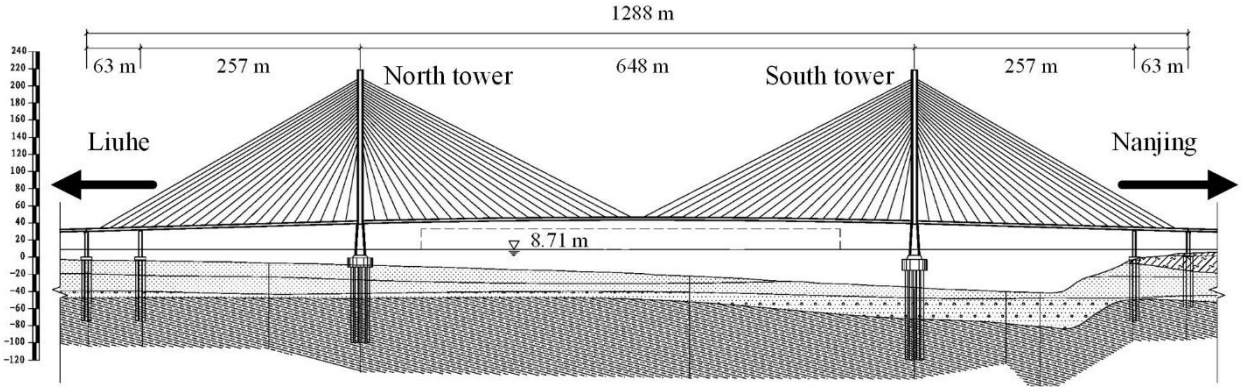


Fig. 3 Configuration of the Nanjing Dashengguan Yangtze River Bridge

A SHM system was devised and installed to monitor the environmental, loading and response information in the second year after the completion of the bridge in 2005. A total of 1078 sensors were employed, including anemometers, temperature sensors, vehicle weighing systems, anchor load cells and others. The cable forces of all the 167 stay cables were recorded by using the JC1-type anchor load cells with a sampling frequency of 10Hz and a relative measurement error of ±1%. In this paper, cable force measurements of the stay cable NJX21 as highlighted in Fig. 4 are adopted for the anomaly trend detection analysis.

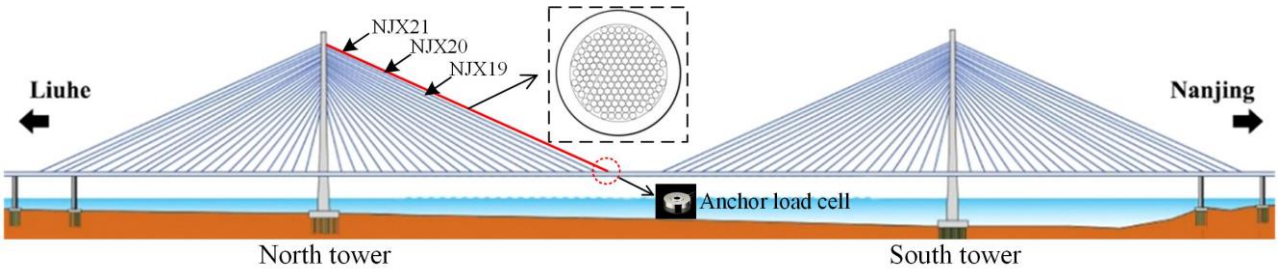


Fig. 4 The studied stay cable NJX21

3.2 Data pre-processing

Raw cable force measurements of the three stay cables (i.e., NJX21, NJX20 and NJX19 as shown in Fig. 4) in 30 seconds are plotted in Fig. 5, where the data missing phenomenon is widespread. As

discussed earlier, the moving time window mean imputation is employed to address the data missing issue with a window length of 2 seconds. The cable force signals after missing data imputation are plotted in Fig. 6, which achieve a good continuity.

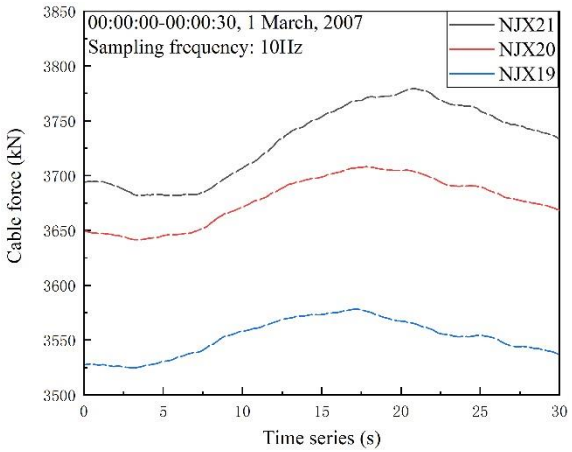


Fig. 5 Raw cable force signals of the three stay cables in 30s

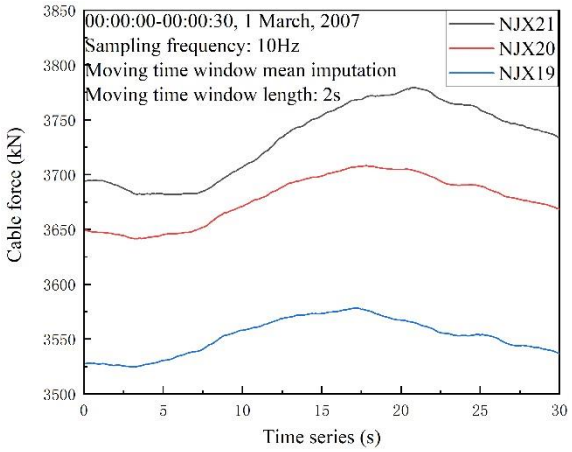


Fig. 6 Signals after missing data imputation

Cable force signals of the stay cable NJX21 in 24 hours, as shown in Fig. 7, are taken as the example to explain the thermal separation procedure using the multi-resolution wavelet-based approach. The decomposition level is set as 27 according to Eq. (2), and the wavelet basis function is 'coif5'. The daily thermal effect lies on the 19th detail level, whose energy, as shown in Fig. 8, is not quite large because the daily temperature variation is limited to 4°C. The extracted temperature-induced cable force is plotted in Fig. 7, whose variation trend is in line with that of the monitored temperatures. Fig. 9 is the cable force signal without the influence of thermal actions, which is used for index extraction, trigger determination and anomaly trend detection.

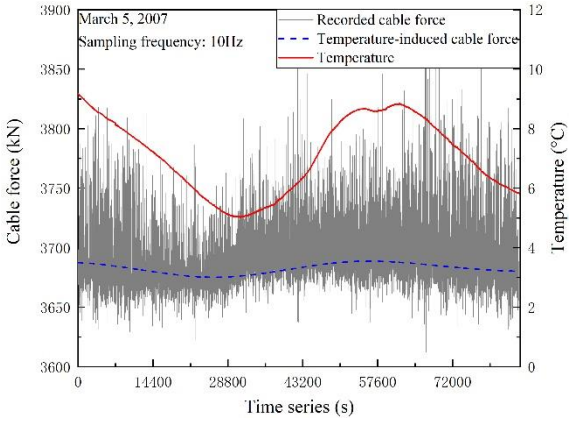


Fig. 7 Thermal effect separation procedure

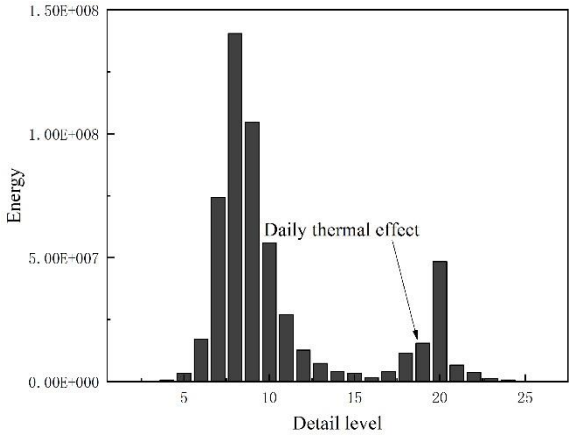


Fig. 8 Energy in each detail level

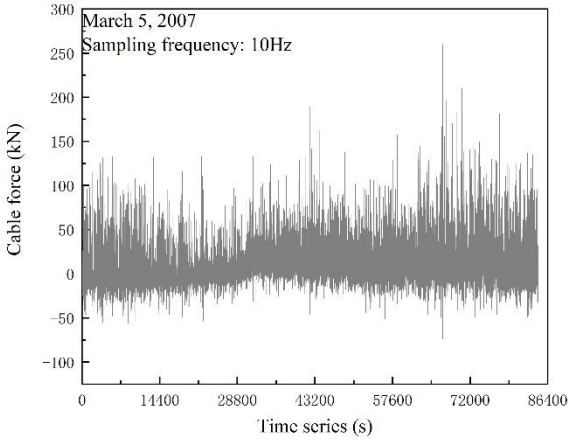


Fig. 9 Signals after thermal effect separation

3.3 Index extraction

The length of window should be determined based on the minimum duration of anomalous events. Considering the relatively long duration of structural damages (from the instant of occurrence of damage to maintenance), damage-induced events are not the controllable cases for determination of window length. In this study, overloading vehicle events are adopted as the dominated cases to

determine the length of window.

The overall length of the Nanjing Dashengguan Yangtze River Bridge is 1288m, and the significant positive influence line of the studied stay cable shown in Fig. 10 almost crosses 905m of the bridge. Moreover, based on the statistical results of heavy vehicle speeds in this bridge as shown in Fig. 10, the average speed of heavy vehicles is 52.6km/h. The average active time window of overloading vehicle events is calculated as 61.9 seconds, i.e., 619 data points for the sampling frequency of 10Hz. In this paper, the length of time window is determined as 60s subject to the sampling frequency of 10Hz.

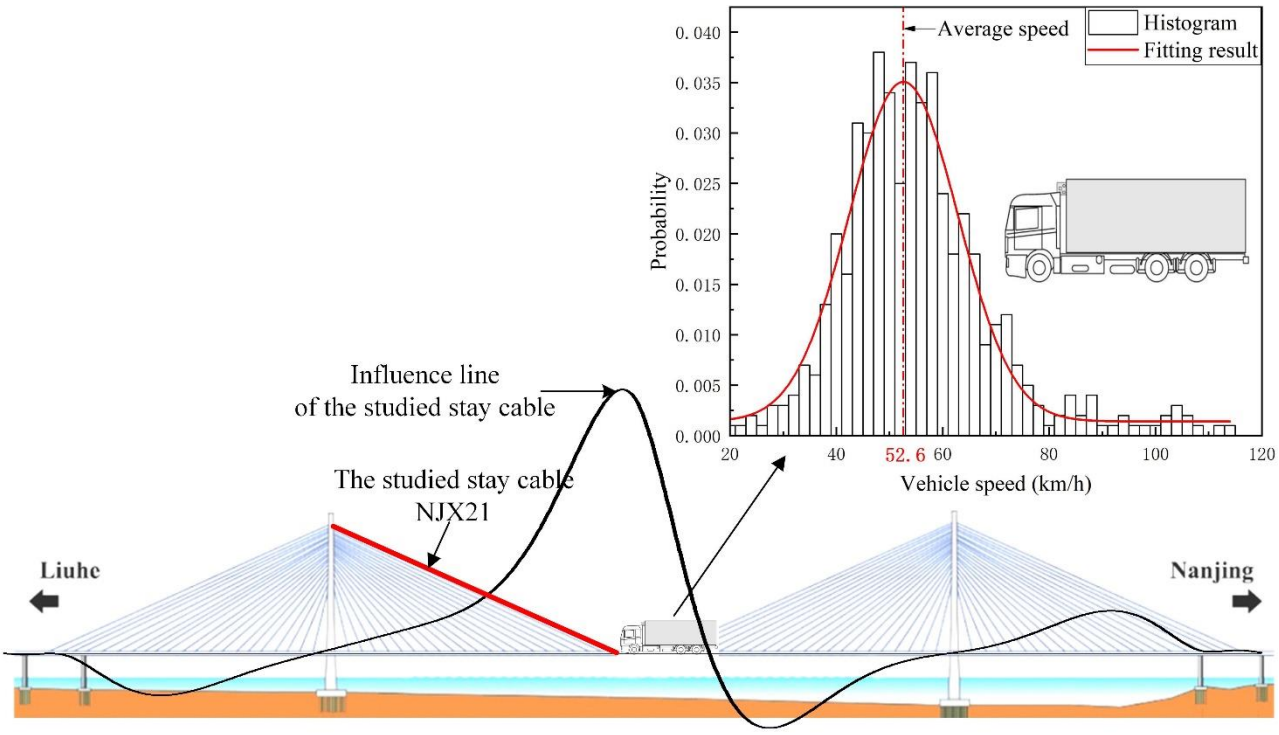


Fig. 10 Duration time window discussion subject to overloading vehicle events

According to Eq. (3), the energy anomaly detection index along the timeline is extracted from the time history as shown in Fig. 9, which is plotted in Fig. 11. As requested by the GPD analysis in independent identically distribution, daily maximum index highlighted in Fig. 11 is adopted for the trigger estimation. The extracted daily maximum indexes in 2007 are plotted in Fig. 12, which are the database for the GPD analysis.

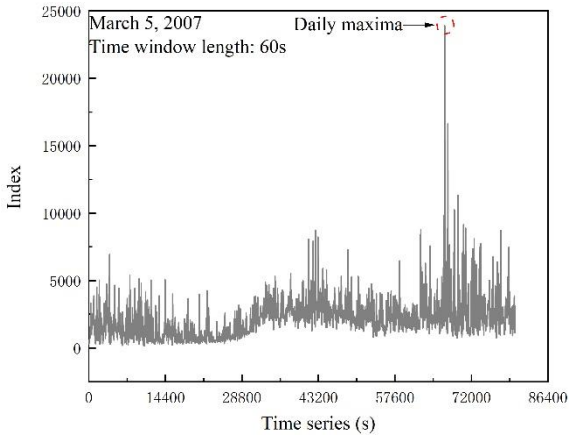


Fig. 11 Extracted energy index in March 5, 2007

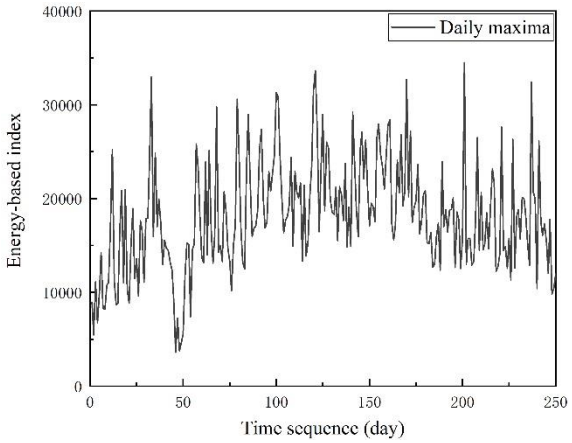


Fig. 12 Daily maximum energy index sequence in 2007

3.4 Trigger determination

To predict the anomaly detection trigger, the threshold needs to be first determined based on the characteristics of the mean excess function and standardized residual. According to Eqs. (6) and (7), the mean excess function and standardized residual derived from the daily maximum indexes in 2007 are plotted in Fig. 13. It is observed that if the threshold is 17 000, the corresponding standardized residual approaches its lowest point, and the mean residual life plot approximately follows a straight line. Therefore, the optimal threshold of the daily maximum energy indexes for the GPD discussions is set as 17 000.

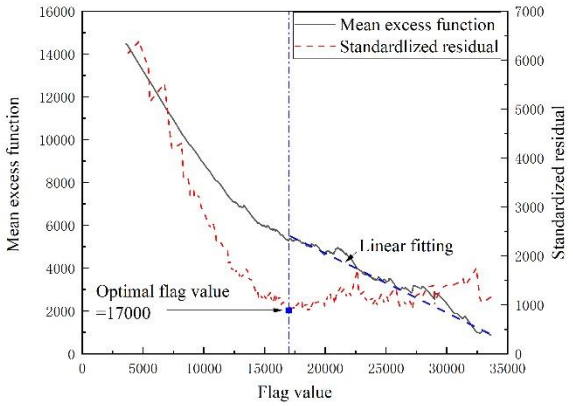
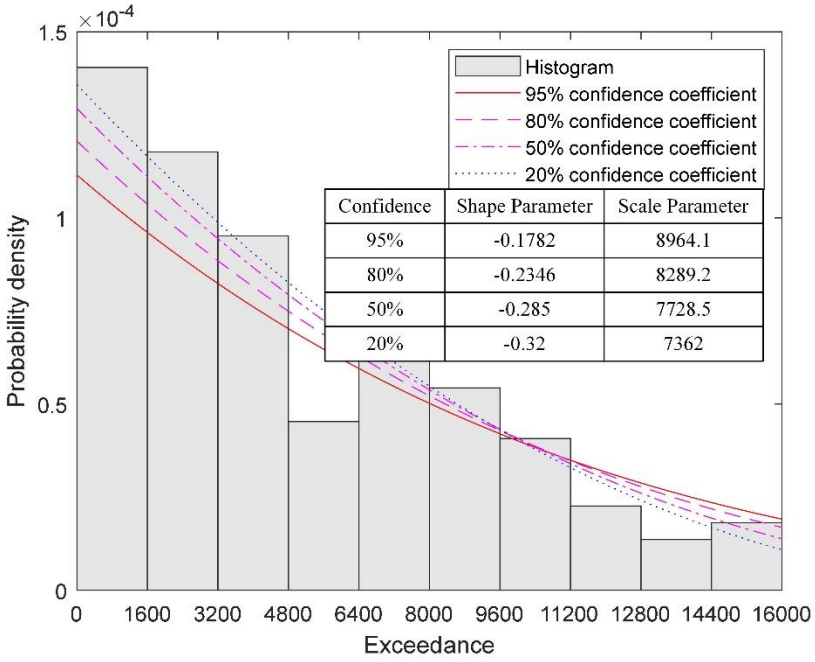


Fig. 13 Mean excess function and standardized residual

Following the method of confidence interval estimation (i.e., the bootstrap), shape and scale parameters subject to 95%, 80%, 50% and 20% confidence coefficients are estimated as (-0.1782, 8964.1), (-0.2346, 8289.2), (-0.285, 7728.5), and (-0.32, 7362), respectively. Histograms of the exceedance and GPD fitting results corresponding to the four confidence coefficients are plotted in Fig. 14.



Confidence	Shape Parameter	Scale Parameter
95%	-0.1782	8964.1
80%	-0.2346	8289.2
50%	-0.285	7728.5
20%	-0.32	7362

Fig. 14 Pareto distribution with various confidence coefficients

According to Eq. (9), the triggers of anomaly trend detection with 95%, 80%, 50% and 20% confidence coefficients are predicted as 58671.47, 48862.53, 42499.48 and 39035.2, respectively.

3.5 Case study 1: spike detection

Signal spikes are always observed in measured signals from SHM systems, which are caused by electrical transients in voltage, current, or transferred energy in an electrical circuit. Spikes are not

signals of interest for anomaly detection, and even coupled with signals induced by anomalous scenarios, resulting in false detection. For instance, cable force signals of the studied stay cable on Jan. 17, 2008 are influenced by the spike as shown in Fig. 15. False detection is observed if the index of absolute cable force value is employed. The specific detection result is illustrated in Fig. 16 by using the absolute cable force index. Based on Fig. 16, anomaly event is detected with more than 95% confidence, however, the actual situation is that the anomaly detection is triggered by the spike signal, and the structure operated as usual.

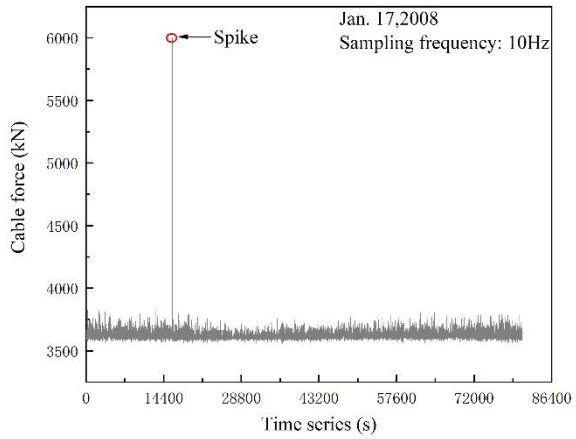


Fig. 15 Measured cable force signals on Jan. 17, 2008

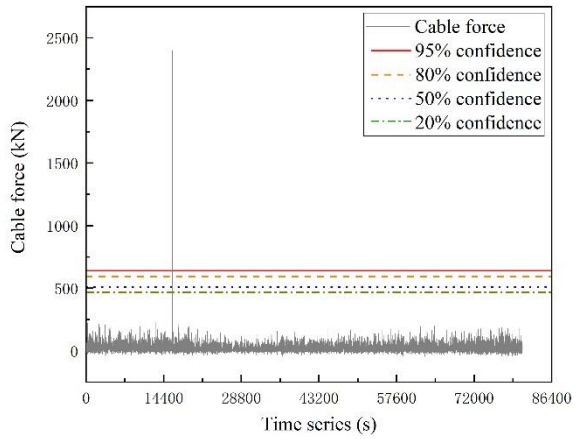


Fig. 16 Detection result using absolute cable force index

In addition, the cointegration residual of the monitored cable force served as index to detect spikes. The details using cointegration techniques to detect anomalies are presented in our previous paper (Fan et al., 2020). The detection result is shown in Fig. 17. As a result, the spike triggered the anomaly detection, resulting in a false alarm.

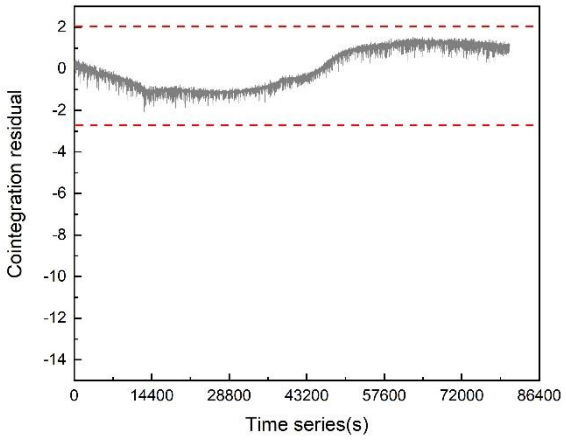


Fig. 17 Detection result using cointegration residual

Considering the characteristics of spikes in short duration and large absolute value, the energy index is proposed in this paper to improve the tolerance of index to spikes. Similarly, the raw signals of the cable force shown in Fig. 15 are used to conduct the anomaly trend detection. The detection result is demonstrated in Fig. 18. Compared with the absolute cable force index and cointegration residual, the energy index has more tolerance to spikes. Thus, when using the energy index, it is not necessary to delete spikes in the data pre-processing section.

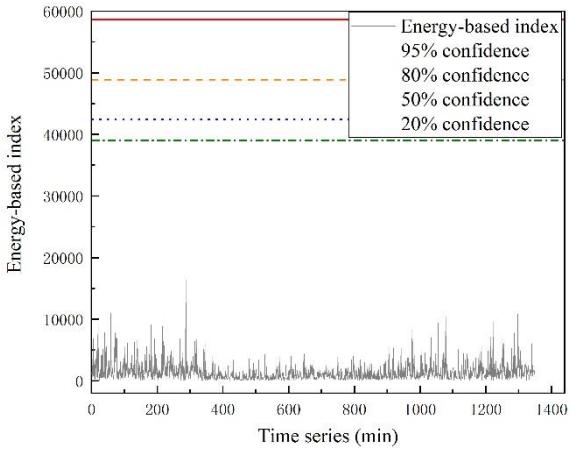


Fig. 18 Detection result using energy index

3.6 Case study 2: overloading vehicle detection

Since no confidential overloading vehicle scenarios were recorded, simulated cases were introduced in this study. It is assumed that three 100-ton overloading vehicles simultaneously go through the bridge with a constant speed of 60 km/h, and the corresponding simulated cable force signal of the stay cable NJX 21 is shown in Fig. 19. Then, the signal is merged into the real-time monitored cable force data at instant *T* on Jan. 22, 2008 to simulate the event that the overloading vehicles involve in

the normal operational traffic flow. The simulated cable force signals after data pre-processing are shown in Fig. 20, where the simulated overloading event is highlighted. Moreover, it is observed that during the time window between 3:00 a.m. and 6:00 a.m., signals are concentrated on the horizontal axis of zero since there are rare traffic volumes during this time window, which further verifies the effectiveness of the proposed thermal response separation methodology.

Similarly, the energy indexes are calculated and plotted in Fig. 21 together with triggers of point estimation and the four confidence levels. The details regarding calculation of the trigger using point estimation method is presented in our previous paper (Xu et al., 2020b). The trigger derived from the point estimation is lower than that of 20% confidence level, which is prone to raise false detection. Based on Fig. 21, the anomalous event is detected by using the point estimation, while it is detected with more than 95% confidence using confidence interval estimation.

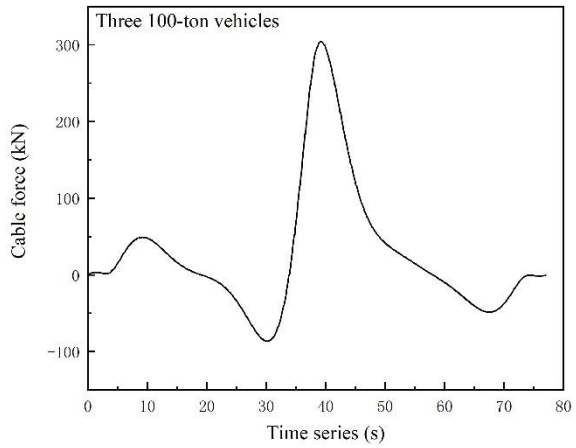


Fig. 19 Simulated cable force signal with three 100-ton vehicles going through the bridge

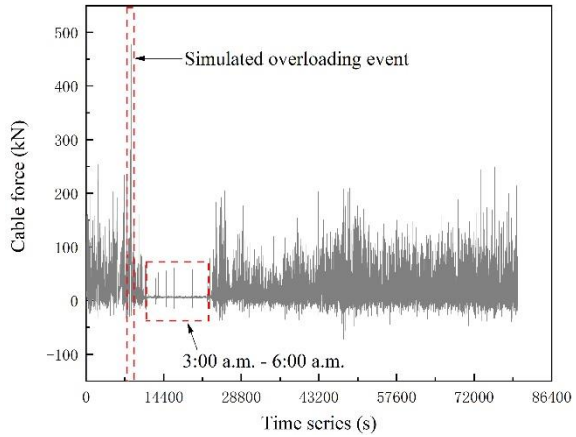


Fig. 20 Simulated cable force signals after data pre-processing

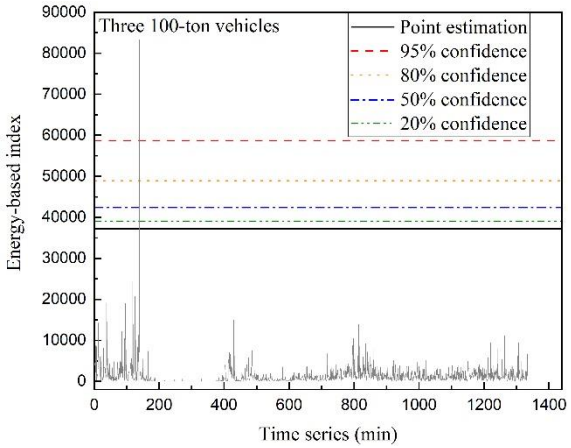


Fig. 21 Detection result of the three 100-ton vehicles through the bridge

Additional two overloading events are assumed herein, which are two 100-ton vehicles and single 100-ton vehicle going through the bridge, respectively. The detection results of the two overloading events are shown in Fig. 22 and Fig. 23. Two 100-ton vehicles through the bridge is detected with a more than 80% confidence, and the single 100-ton vehicle through the bridge is detected with an almost 20% confidence. As a result, the level of detection confidence increases with the weight of overloading vehicles. Although the two overloading cases could be detected by using the point estimation, the detection results from the confidence interval estimation give more information for decision-makings.

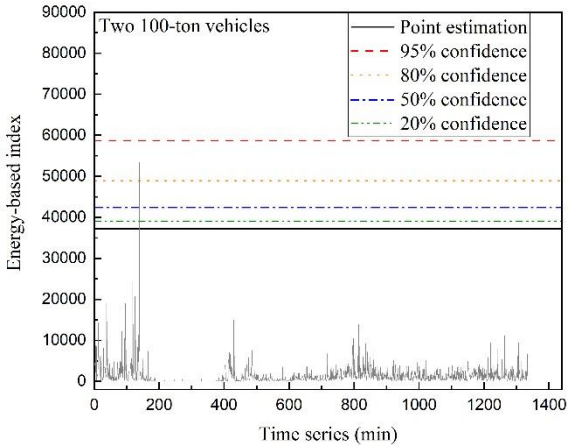


Fig. 22 Detection result of the two 100-ton vehicles through the bridge

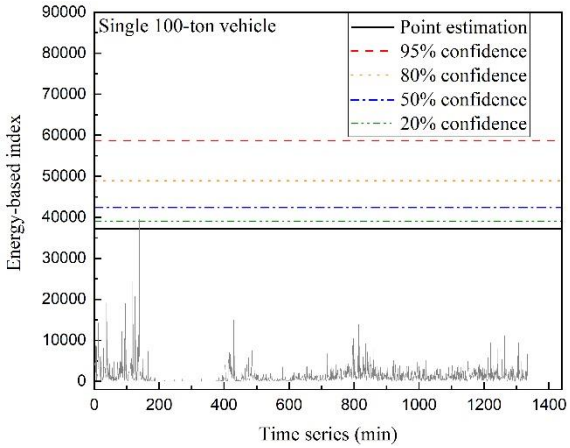


Fig. 23 Detection result of the single 100-ton vehicle through the bridge

3.7 Case study 3: snow disaster detection

Nanjing city suffered from a heavy snow storm at the end of Jan. in 2008. With the accumulation of snow on the pavement of the Nanjing Dashengguan Yangtze River, the bridge gradually carried extra snow loads. The recorded cable force data of the stay cable NJX21 during the snow storm time window (Jan. 26, 2008) are plotted in Fig. 24. After the data pre-processing and index extraction, the energy index on Jan. 26, 2008 is plotted in Fig. 25.

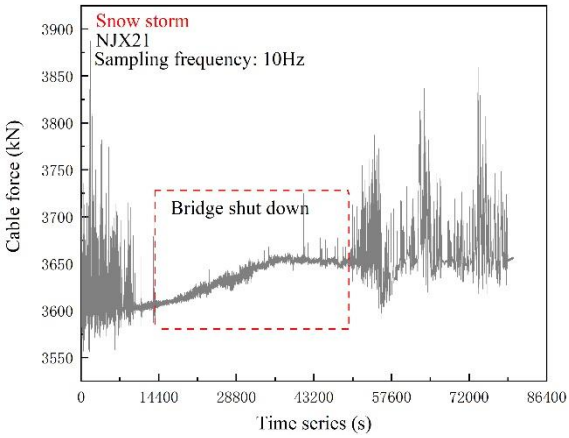


Fig. 24 Raw cable force signal during the snow storm

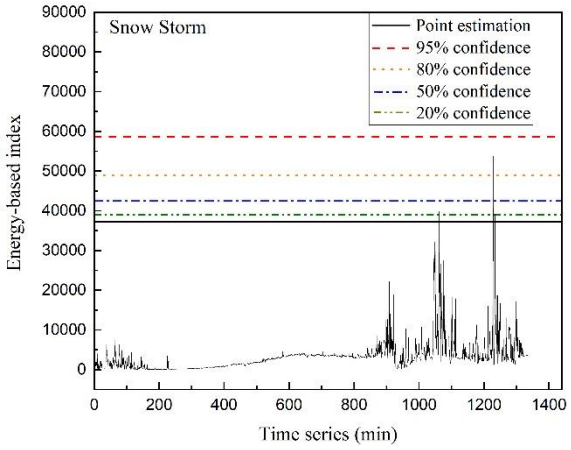


Fig. 25 Detection result of the snow storm

Based on Fig. 25, the snow storm induced anomalous scenario is detected via the point estimation, and it is detected with a more than 80% confidence using confidence interval estimation. The anomalous scenario resulted from two facts: (1) the bridge took extra snow loads with the accumulation of snow on the pavement; and (2) the traffic volume was extremely large when the bridge was re-opened since the short-term shutdown of the bridge generated large number of waiting vehicles. In view of the structural safety, the bridge was shut down for the whole day on Jan. 27, 2008.

4. Conclusions

In this paper, a probabilistic anomaly trend detection method is developed for large span cable-supported bridges, where energy index is proposed to achieve robust detection performance and confidence interval estimation is used to measure the uncertainty within the anomaly detection procedure. The concluding remarks are summarized as follows:

(1) Data pre-processing (i.e., missing data imputation and thermal response separation) is a critical step in the anomaly detection process. Moving time window mean imputation is adopted for missing data imputation, and the multi-resolution wavelet-based approach is applied to separate thermal effects from the recorded structural response signals.

(2) An energy index in the time domain is proposed for the probabilistic anomaly detection. Compared with the absolute value-based index, the energy index has more tolerance to spikes. Thus, spike detection is not necessary in the data pre-processing when using the energy index.

(3) Confidence interval estimation is used to predict triggers with different confidences in the

GPD analysis for anomaly detection, where four confidence levels (i.e., 95%, 80%, 50% and 20% confidence) are defined.

(4) The effectiveness of the anomaly detection methodology is verified by using field cable force measurements of the Nanjing Dashengguan Yangtze River Bridge. The triggers with different confidences are derived based on the measurements in 2007. Three cases are presented in this study, which are spike detection, overloading vehicle detection and snow disaster detection. The spike is not detected when using the energy index. The simulated overloading events are all detected, where different overloading vehicles have different confidence levels. The snow disaster is detected on Jan. 26, 2008 with more than 80% confidence.

In this study, the single index, cable force, is used for the anomaly detection discussions. In the future, a multi-index detection method is promising in the practical applications to ensure the structural and operational safety of bridges.

Acknowledgements

The authors thank the anonymous reviewers for their constructive comments that greatly improved the quality of this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported in this article was supported in part by the Fundamental Research Funds for the Central Universities under Grant No. 2242021k30034, the Natural Science Foundation of Jiangsu Province under Grant No. BK20181278, Transportation Science Research Project in Jiangsu under Grant No. 2019Z02, and Special Research Fund for Academicians of China Communications Group under Grant No. YSZX-03-2020-01-B.

References

Castillo E and Hadi AS (1997) Fitting the generalized Pareto distribution to data. *Journal of American Statistical Association* 92(440):1609-20.

Chen P, Ye ZS and Zhao XQ (2017) Minimum distance estimation for the generalized Pareto distribution. *Technometrics* 59(4):528-41. doi: 10.1080/00401706.2016.1270857

Cho J, Seo DM and Uh Y (2020) Clinical application of overlapping confidence intervals for monitoring changes in serial clinical chemistry test results. *Annals of Laboratory Medicine* 40(3):201-08. doi: 10.3343/alm.2020.40.3.201

Das S, Saha P and Patro SK (2016) Vibration-based damage detection techniques used for health monitoring of structures: a review. *Journal of Civil Structural Health Monitoring* 6(3):477-507. doi: 10.1007/s13349-016-0168-5

Deng Y, Li AQ, Liu Y, et al. (2018) Investigation of temperature actions on flat steel box girders of long-span bridges with temperature monitoring data. *Advances in Structural Engineering* 21(14):2099-113. doi: 10.1177/1369433218766946

DiCiccio TJ and Efron B (1996) Bootstrap confidence intervals. *Statistical Science* 11(3):189-212.

Efron B (1987) Better bootstrap confidence intervals. *Journal of the American statistical Association* 82(397): 171-185. doi: 10.2307/2289144

Enders CK (2001) A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling* 8(1):128-41. doi: 10.1207/s15328007sem0801_7

Fan ZY, Huang Q, Ren Y, et al. (2020) A cointegration approach for cable anomaly warning based on structural health monitoring data: An application to cable-stayed bridges. *Advances in Structural Engineering* 23(13): 2789-802.

Fujino Y, Siringoringo DM, Ikeda Y, et al. (2019) Research and implementations of structural monitoring for bridges and buildings in Japan. *Engineering* 5(6): 1093-119. doi: 10.1016/j.eng.2019.09.006

Gilli M (2006) An application of extreme value theory for measuring financial risk. *Computational Economics* 27(2-3):207-28.

Hawthorne G, Hawthorne G and Elliott P (2005) Imputing cross-sectional missing data: comparison of common techniques. *Australian & New Zealand Journal of Psychiatry* 39(7):583-90.

Hua XG, Ni YQ, Chen ZQ, et al. (2009) Structural damage detection of cable-stayed bridges using changes in cable forces and model updating. *Journal of Structural Engineering* 135(9):1093-106. doi: 10.1061/(asce)0733-9445(2009)135:9(1093)

Huang HB, Yi TH, Li HN, et al. (2020) Strain-based performance warning method for bridge main girders under variable operating conditions. *Journal of Bridge Engineering* 25(4):04020013. doi: 10.1061/(asce)be.1943-5592.0001538

Kalaycioglu O, Copas A, King M, et al. (2016) A comparison of multiple-imputation methods for handling missing data in repeated measurements observational studies. *Journal of the Royal Statistical Society Series A-Statistics in Society* 179(3):683-706. doi: 10.1111/rssa.12140

Kromanis R and Kripakaran P (2014) Predicting thermal response of bridges using regression models derived from measurement histories. *Computers and Structures* 136:64-77. doi: 10.1016/j.compstruc.2014.01.026

Kysely J (2010) Coverage probability of bootstrap confidence intervals in heavy-tailed frequency models, with application to precipitation data. *Theoretical and Applied Climatology* 101(3-4):345-61. doi: 10.1007/s00704-009-0190-1

Lai WY and Kuok KK (2019) A study on Bayesian principal component analysis for addressing

missing rainfall data. *Water Resources Management* 33(8):2615-28. doi: 10.1007/s11269-019-02209-8

Liu Y, Deng Y and Cai CS (2015) Deflection monitoring and assessment for a suspension bridge using a connected pipe system: a case study in China. *Structural Control and Health Monitoring* 22(12):1408-25. doi: 10.1002/stc.1751

Liu Z, Guo T, Han D, et al. (2020) Experimental study on corrosion-fretting fatigue behavior of bridge cable wires. *Journal of Bridge Engineering* 25(12): 04020104. doi: 10.1061/(ASCE)BE.1943-5592.0001642

Liu Z, Guo T, Yu X, et al. (2021) Corrosion fatigue and electrochemical behaviour of steel wires used in bridge cables. *Fatigue & Fracture of Engineering Materials & Structures* 44(1): 63-73. doi: 10.1111/ffe.13331

Loftus GR and Masson MEJ (1994) Using confidence-intervals in within-subject designs. *Psychonomic Bulletin and Review* 1(4):476-90. doi: 10.3758/bf03210951

Martinez-Luengo M, Shafiee M and Kolios A (2019) Data management for structural integrity assessment of offshore wind turbine support structures: data cleansing and missing data imputation. *Ocean Engineering* 173:867-83. doi: 10.1016/j.oceaneng.2019.01.003

Nevalainen J, Kenward MG and Virtanen SA (2009) Missing values in longitudinal dietary data: A multiple imputation approach based on a fully conditional specification. *Statistics in Medicine* 28(29):3657-69. doi: 10.1002/sim.3731

Ni Y, Wang Y and Zhang C (2020) A Bayesian approach for condition assessment and damage alarm of bridge expansion joints using long-term structural health monitoring data. *Engineering Structures* 212:110520.

Ni YQ, Xia HW, Wong KY, et al. (2012) In-service condition assessment of bridge deck using long-term monitoring data of strain response. *Journal of Bridge Engineering* 17(6):876-85. doi: 10.1061/(asce)be.1943-5592.0000321

Ou JP and Li H (2010) Structural health monitoring in mainland China: review and future trends. *Structural Health Monitoring* 9(3): 219-31. doi: 10.1177/1475921710365269

Pandey AK, Biswas M and Samman MM (1991) Damage detection from changes in curvature mode shapes. *Journal of Sound and Vibration* 145(2):321-32. doi: 10.1016/0022-460x(91)90595-b.

Peeters B, Maeck J and De Roeck G (2001) Vibration-based damage detection in civil engineering: excitation sources and temperature effects. *Smart Materials & Structures* 10(3):518-27. doi: 10.1088/0964-1726/10/3/314

Ren Y, Xu X, Huang Q, et al. (2019) Long-term condition evaluation for stay cable systems using dead load-induced cable forces. *Advances in Structural Engineering* 22(7):1644-56. doi: 10.1177/1369433218824486

Salawu OS (1997) Detection of structural damage through changes in frequency: A review. *Engineering Structures* 19(9):718-23. doi: 10.1016/s0141-0296(96)00149-6

Schober P and Vetter TR (2020) Confidence intervals in clinical research. *Anesthesia and Analgesia* 130(5):1303-03. doi: 10.1213/ane.0000000000004731

Tome ES, Pimentel M and Figueiras J (2020) Damage detection under environmental and operational effects using cointegration analysis - Application to experimental data from a cable-stayed bridge. *Mechanical Systems and Signal Processing* 135:14. doi: 10.1016/j.ymssp.2019.106386

Wang H, Zhang YM, Mao JX, et al. (2019) Modeling and forecasting of temperature-induced strain of a long-span bridge using an improved Bayesian dynamic linear model. *Engineering*

Structures 192:220-32 doi: 10.1016/j.engstruct.2019.05.006
Wei W, Shen ZQ, Wu L, et al. (2020) Estimating DLMP confidence intervals in distribution networks with AC power flow model and uncertain renewable generation. *IET Generation Transmission and Distribution* 14(8):1467-75. doi: 10.1049/iet-gtd.2019.0958
Wu BJ, Li ZX, Chan THT, et al. (2014) Multiscale features and information extraction of online strain for long-span bridges. *Smart Structures and Systems* 14(4):679-97. doi: 10.12989/sss.2014.14.4.679
Xu X, Huang Q, Ren Y, et al. (2019) Modeling and separation of thermal effects from cable-stayed bridge response. *Journal of Bridge Engineering* 24(5):04019028. doi: 10.1061/(asce)be.1943-5592.0001387
Xu X, Ren Y, Huang Q, et al. (2020a) Thermal response separation for bridge long-term monitoring systems using multi-resolution wavelet-based methodologies. *Journal of Civil Structural Health Monitoring* 10(3):527-41. doi: 10.1007/s13349-020-00402-7
Xu X, Ren Y, Huang Q, et al. (2020b) Anomaly detection for large span bridges during operational phase using structural health monitoring data. *Smart Materials and Structures* 29(4):045029. doi: 10.1088/1361-665X/ab79b3
Xu YL and Xia Y (2011) *Structural health monitoring of long-span suspension bridges*. Boca Raton: CRC Press.
Xu ZD and Wu Z (2007) Simulation of the effect of temperature variation on damage detection in a long-span cable-stayed bridge. *Structural Health Monitoring* 6(3):177-89. doi: 10.1177/1475921707081107
Yu Z, Xia H, Goicolea JM, et al. (2016) Bridge damage identification from moving load induced deflection based on wavelet transform and Lipschitz exponent. *International Journal of Structural Stability and Dynamics* 16(5):1550003. doi: 10.1142/s0219455415500030
Zhou GD, Yi TH, Chen B, et al. (2017) A generalized Pareto distribution-based extreme value model of thermal gradients in a long-span bridge combining parameter updating. *Advances in Structural Engineering* 20(2):202-13. doi: 10.1177/1369433216660010
Zhou Y and Sun LM (2019) Insights into temperature effects on structural deformation of a cable-stayed bridge based on structural health monitoring. *Structural Health Monitoring* 18(3):778-91. doi: 10.1177/1475921718773954
Zhu YJ, Ni YQ, Jesus A, et al. (2018) Thermal strain extraction methodologies for bridge structural condition assessment. *Smart Materials and Structures* 27(10):15. doi: 10.1088/1361-665X/aad5fb
Zhu YJ, Ni YQ, Jin H, et al. (2019) A temperature-driven MPCA method for structural anomaly detection. *Engineering Structures* 190:447-58. doi: 10.1016/j.engstruct.2019.04.004