INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 02, Issue 08, August 2022, pp : 137-144

e-ISSN : 2583-1062

Impact Factor : 2.265

A REVIEW ON STATISTICAL MACHINE LEARNING APPROACHES FOR SEISMIC DAMAGES ON BRIDGES

Aman Gautam[1], Prof. Rahul Sharma[2]

[1]Research Scholar, Department of Civil Engineerin, Prashanti Institute Of Technology & Science, Ujjain, India.

[2]Assistant Professor, Department of Civil Engineering, Prashanti Institute Of Technology & Science, Ujjain, India.

ABSTRACT

With rapid advances in statistical modelling and data analysis, analysis of bridge failure and damage is also being considered under the purview of statistical models. It is ubiquitous that manual inspection of bridge damage and estimations are time consuming and may not be practically feasible in cases where the experts are not readily available and/or sites of huge mass calamities such as earthquake prone cites. Thus a quick estimation and evaluation of the damages is necessary so as to immediately stop the cascading effect of the possible failure and casualties. In this paper, a comprehensive review on the various statistical techniques being employed for the evaluation of bridge failure have been discussed along with the salient points. Previous work in the domain has also been cited so as to garner attention towards the recent trends in the domain. The evaluation metrics commonly employed for the analysis of the statistical systems have also been defined. It is expected that this comprehensive review would pave the path for further research domains in the field.

Keywords: Seismic Damage Estimation, statistical models. Machine learning, mean absolute percentage error, mean squared error.

1. INTRODUCTION

To take an informed decision for the recovery process, it is critical to identify the damage state of a bridge in the aftermath of an earthquake [1]. The simple and widely strategy adopted for the rapid damage assessment of the bridge is the implementation of fragility curves in earthquake alerting systems [2]. Various researchers have generated fragility curves for infrastructure portfolios. However, the variation in the geometric, structural, and material properties across the bridges in a region necessitates the grouping of bridge classes for the generation of fragility curves[3]. Such a grouping often leads to the fragility curves of bridge classes rather than bridge-specific fragility curves, i.e., it merely accounts for important attributes of a specific bridge's structural design [4].

Fig. 1 Common causes of bridge failure

Earthquakes are sudden movements across faults that release elastic energy stored in rocks and radiate seismic waves that travel throughout Earth [5]. Every day there are about fifty earthquakes worldwide that are strong enough (magnitude > 2.5) to be felt locally, and every few days an earthquake occurs that is capable of damaging structures [6]. Seismology is a data-rich and data-driven science, and the rate of data acquisition is accelerating as seismic sensors get steadily less costly [7]. The massive and rapidly growing amount of data highlights the need for more effective tools for the efficient processing and extraction of as much useful information as possible to enable scientist to realize the full potential to gain new insights into earthquake processes from them [8].

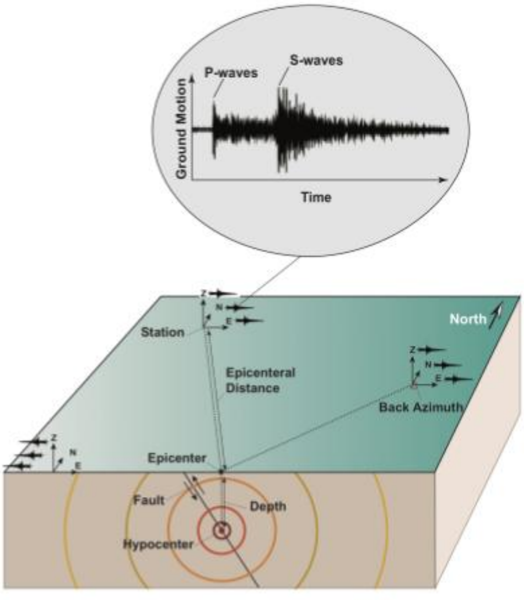

Fig. 2 Propagation of Seismic Waves

Seismologists use only a portion of the recorded data to understand the physics of earthquakes and learn about Earth's deep interior, where direct observations are impossible. Most seismic data sets have not been fully analyzed and important discoveries can result from reanalysis of data sets using new data analysis tools [9]. Machine learning (ML) techniques have been shown to be powerful tools for processing and exploring seismic data [10]-[11]. The success of these MLbased methods in achieving state-of-the-art performance is mainly due to availability of large-scale and accurately labeled training data sets. Although, hundreds of terabytes of archived seismic waveform data and tens of millions of human picked parameters are available, a large and highquality-labeled benchmark data set for seismic waveforms does not yet exist [12]. This is attributable to several technical issues regarding reliable synchronization of metadata and waveform data and a lack of comprehensive and efficient quality control mechanism [13].

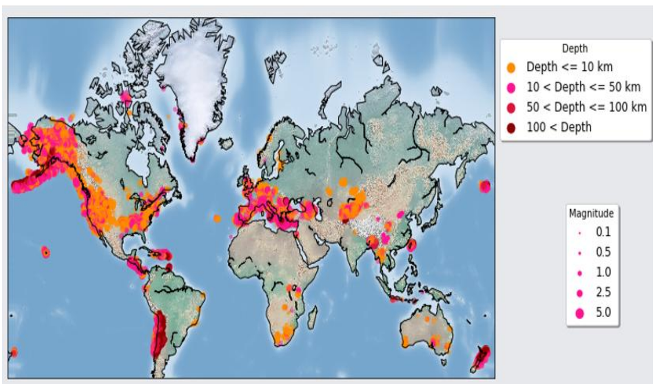

Fig. 3 Location, Size and Depth of recorded earthquakes.

Unfortunately, the uncertainties in location, depth, and origin time estimates are not uniformly reported for all events by our resources and it is difficult to estimate them. Several statistical and machine learning modes are being used off late for the quick response and rapid estimation of damages. The followng section highlights such techniques [14]

2. STATISTICAL MODELS

Evolutionary statistical algorithms are a set of such algorithms which show the aforesaid characteristics.

Evolutionary Statistical algorithms try to mimic the human attributes of thinking which are [15]:

1) Parallel data processing
2) Self-Organization
3) Learning from experiences

Some of the commonly used techniques are discussed below [16]-[19]:

1) Statistical Regression: These techniques are based on the time series approach based on the fitting problem that accurately fits the data set at hand. The approach generally uses the auto-regressive models and means statistical measures. They can be further classified as:

a) Linear

b) Non-Linear

Mathematically:

Let the time series data set be expressed as:

$$\mathbf{Y} = \{\mathbf{Y1}, \mathbf{Y2} \ldots\ldots\ldots\ldots\ldots\ldots..\mathbf{Yt}) \qquad \textbf{(1)}$$

Here,

Y represents the data set

t represents the number of samples

Let the lags in the data be expressed as the consecutive differences.

The first lag is given by:

$$\Delta\mathbf{Y_1} = \mathbf{Y_{t-1}} \qquad \textbf{(2)}$$

Similarly, the j^{th} lag is given by:

$$\Delta\mathbf{Y_j} = \mathbf{Y_{t-j}} \qquad \textbf{(3)}$$

2) Correlation based fitting of time series data: The correlation based approaches try to fit the data based on the correlation among the individual lags. Mathematically it can be given by:

$$\mathbf{A_t} = \mathbf{corr}(\mathbf{Y_t}, \mathbf{Y_{t-1}}) \qquad \textbf{(4)}$$

Here, Correpresents the auto-correlation (which is also called the serial correlation)

Y_t is the t^{th} lagged value

Y_{t-1} is the $(t-1)^{st}$ lagged value

The mathematical expression for the correlation is given by

$$\mathbf{corr}(\mathbf{Y_t}, \mathbf{Y_{t-1}}) = \frac{\mathbf{conv}(\mathbf{Y_t}, \mathbf{Y_{t-1}})}{\sqrt{\mathbf{varY_t}, \mathbf{varY_{t-1}}}} \qquad \textbf{(5)}$$

Here, Correpresents convolution given by:

$$\mathbf{conv}\{\mathbf{x}(\mathbf{t}), \mathbf{h}(\mathbf{t})\} = \int_{\mathbf{t=1}}^{\infty} \mathbf{x}(\boldsymbol{\vartheta})\mathbf{h}(\mathbf{t} - \boldsymbol{\vartheta})\mathbf{d}\boldsymbol{\vartheta} \ \textbf{(6)}$$

Here,

ϑ is a dummy shifting variable for the entire span of the time series data

t represents time

Y_t is the t^{th} lagged value

Y_{t-1} is the $(t-1)^{st}$ lagged value

X is function 1

H is function 2

Var represents the variance given by:

$$\mathbf{var}(\mathbf{X}) = \mathbf{X_i} - \mathbf{E}(\mathbf{X}) \qquad \textbf{(7)}$$

Here,

Xi is the random variable sample

E represents the expectation or mean of the random variable X

3) Finite Distribution Lag Model (FDL): This model tries to design a finite distribution model comprising of lags fitted to some distribution such as the normal or lognormal distributions. Mathematically:

$$\mathbf{Y_t} = \boldsymbol{\alpha_t} + \boldsymbol{\delta_1}\mathbf{z_1} + \cdots \ldots\ldots\ldots\ldots.\boldsymbol{\delta_t}\mathbf{z_t} + \boldsymbol{\mu_t} \ \textbf{(8)}$$

Here,

Y_t is the time series data set

α_t is a time dependent variable

δ_1 is a time-varying co-efficient

z is the variable (time variable)

t is the time index

μ_t is the time dependent combination-coefficient

4) Artificial Neural Networks (ANN): In this approach, the time series data is fed to a neural network resembling the working of the human based brain architecture with a self-organizing memory technique.

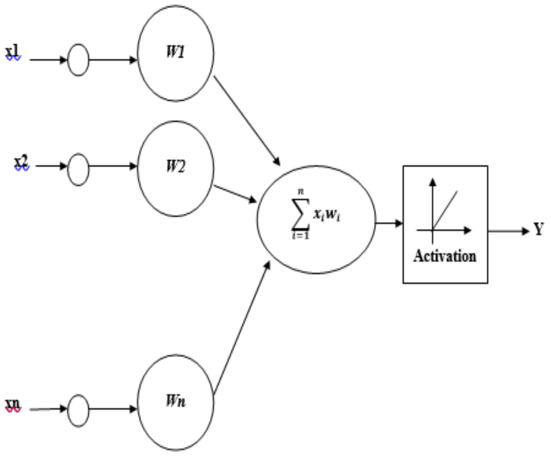

Fig.4 Mathematical Model of Neural Network

The approach uses the ANN and works by training and testing the datasets required for the same. The general rule of the thumb is that 70% of the data is used for training and 30% is used for testing. The neural network can work on the fundamental properties or attributes of the human brain i.e. parallel structure and adaptive self-organizing learning ability. Mathematically, the neural network is governed by the following expression:

$$\mathbf{Y = f(\textstyle\sum_{i=1}^{n} X_i . W_i \; + \; \theta_i)} \qquad \textbf{(9)}$$

Here,

X_i represents the parallel data streams

W_i represents the weights

θ represents the bias

f represents the activation function

The second point is critically important owing to the fact that the data in time series problems such as sales forecasting may follow a highly non-correlative pattern and pattern recognition in such a data set can be difficult. Mathematically:

$$\mathbf{x = f(t)} \qquad \textbf{(10)}$$

Here,

x is the function

t is the time variable.

The relation f is often difficult to find being highly random in nature.

The neural network tries to find the relation f given the data set (D) for a functional dependence of x(t).

The data is fed to the neural network as training data and then the neural network is tested on the grounds of future data prediction. The actual outputs (targets) are then compared with the predicted data (output) to find the errors in prediction. Such a training-testing rule is associated for neural network. The conceptual mathematical architecture for neural networks is shown in the figure below where the input data is x and fed to the neural network.

Fuzzy Logic

Another tool that proves to be effective in several prediction problems is fuzzy logic. It is often termed as expert view systems. It is useful for systems where there is no clear boundary among multiple variable groups. The relationship among the inputs and outputs are often expressed as membership functions expressed as [6]:

A membership function for a fuzzy set A on the universe of discourse (Input) X is defined as:

$$\mathbf{\mu A: X \rightarrow [0, 1]} \qquad \textbf{(11)}$$

Here,

Each element of X is mapped to a value between 0 and 1. It quantifies the degree of membership of the element in X to the fuzzy set A.

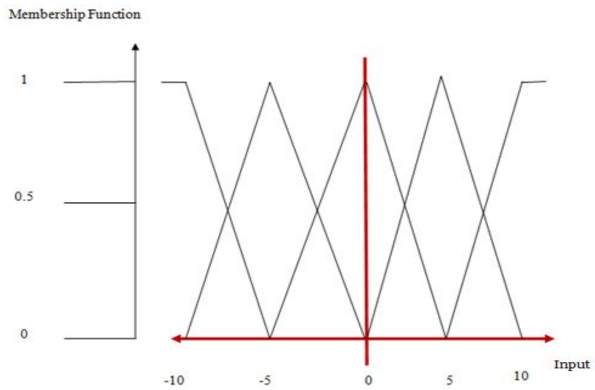

Fig.5 Graphical Representation of Membership Functions

Here,

x axis represents the universe of discourse (Input).

y axis represents the degrees of membership in the [0, 1] interval.

The final category is neuro fuzzy expert systems which governs the defining range of the membership functions.

3. PREVIOUS WORK

This section presents a brief summary of the salient points in the proposed work.

Mangalathu, et al. in [1] proposed that the damage state of a bridge has significant implications on the post-earthquake emergency traffic and recovery operations and is critical to identify the post-earthquake damage states without much delay. Currently, the damage states are identified either based on visual inspection or pre-determined fragility curves. Although these methodologies can provide useful information, the timely application of these methodologies for large scale regional damage assessments is often limited due to the manual or computational efforts. This paper proposes a methodology for the rapid damage state assessment (green, yellow, or red) of bridges utilizing the capabilities of machine learning techniques. Contrary to the existing methods, the proposed methodology accounts for bridge-specific attributes in the damage state assessment. The proposed methodology is demonstrated using two-span box-girder bridges in California. The prediction model is established using the training set, and the performance of the model is evaluated using the test set. It is noted that the machine learning algorithm called Random Forest provides better performance for the selected bridges, and its tagging accuracy ranges from 73% to 82% depending on the bridge configuration under consideration. The proposed methodology revealed that input parameters such as span length and reinforcement ratio in addition to the ground motion intensity parameter have a significant influence on the expected damage state.

Malekjafarian et al. in [2] proposed that a machine learning algorithm for bridge damage detection using the responses measured on a passing vehicle. A finite element (FE) model of vehicle bridge interaction (VBI) is employed for simulating the vehicle responses. Several vehicle passes are simulated over a healthy bridge using random vehicle speeds. An artificial neural network (ANN) is trained using the frequency spectrum of the responses measured on multiple vehicle passes over a healthy bridge where the vehicle speed is available. The ANN can predict the frequency spectrum of any passes using the vehicle speed. The prediction error is then calculated using the differences between the predicated and measured spectrums for each passage. Finally, a damage indicator is defined using the changes in the distribution of the prediction errors versus vehicle speeds. It is shown that the distribution of the prediction errors is low when the bridge condition is healthy. However, in presence of a damage on the bridge, a recognisable change in the distribution will be observed. Several data sets are generated using the healthy and damaged bridges to evaluate the performance of the algorithm in presence of road roughness profile and measurement noise. In addition, the impacts of the training set size and frequency range to the performance of the algorithm are investigated.

Guo et al. in [3] proposed that bridge health monitoring system has been widely used to deal with massive data produced with the continuous growth of monitoring time. However, how to effectively use these data to comprehensively analyze the state of a bridge and provide early warning of bridge structure changes is an important topic in bridge engineering research. This paper utilizes two algorithms to deal with the massive data, namely Kohonen neural network and long short-term memory (LSTM) neural network. The main contribution of this study is using the two algorithms for health state evaluation of bridges. The Kohonen clustering method is shown to be effective for getting classification pattern in normal operating condition and is straightforward for outliers detection. In addition, the LSTM prediction method has an excellent prediction capability which can be used to predict the future deflection values with good accuracy and mean square error. The predicted deflections agree with the true deflections, which indicate that the LSTM method can be utilized to obtain the deflection value of structure. What's more, we can observe the changing trend of bridge structure by comparing the predicted value with its limit value under normal operation.

Bao et al. in [4] proposed that Structural health monitoring (SHM) is a multi-discipline field that involves the automatic sensing of structural loads and response by means of a large number of sensors and instruments, followed by a diagnosis of the structural health based on the collected data. Because an SHM system implemented into a structure automatically senses, evaluates, and warns about structural conditions in real time, massive data are a significant feature of SHM. The techniques related to massive data are referred to as data science and engineering, and include acquisition techniques, transition techniques, management techniques, and processing and mining algorithms for massive data. This paper provides a brief review of the state of the art of data science and engineering in SHM as investigated by these authors, and covers the compressive sampling-based data-acquisition algorithm, the anomaly data diagnosis approach using a deep learning algorithm, crack identification approaches using computer vision techniques, and condition assessment approaches for bridges using machine learning algorithms. Future trends are discussed in the conclusion.

Silva et al. in [5] showed that The structural health monitoring (SHM) field is concerned with the increasing demand for improved and more continuous condition assessment of engineering infrastructures to better face the challenges presented by modern societies. Thus, the applicability of computer science techniques for SHM applications has attracted the attention of researchers and practitioners in the last few years, especially to detect damage in structures under operational and environmental conditions. In the SHM for bridges, the damage detection can be seen as the end of a process to extract knowledge regarding the structural state condition from vibration response measurements. In that sense, the damage detection has some similarities with the Knowledge Discovery in Databases (KDD) process. Therefore, this chapter intends to pose damage detection in bridges in the context of the KDD process, where data transformation and data mining play major roles. The applicability of the KDD for damage detection is evaluated on the well-known monitoring data sets from the Z-24 Bridge, where several damage scenarios were carried out under severe operational and environmental effects.

Mei et al. in [6] proposed that Bridge health monitoring is a very important part for infrastructure maintenance. Traditional bridge health monitoring techniques require sensors to be installed on bridges, which is costly and time consuming. In order to resolve this issue, new damage detection techniques by installing sensors on passing-by vehicles on bridges and considering vehicle bridge interaction (VBI) have gained much attention from researchers in last decade. In this paper, a novel damage detection technique utilizing data collected from sensors mounted on a large number of passing-by vehicles is developed. First, an approach based on Mel-frequency cepstral coefficients (MFCCs) is introduced. Then, an improved version based on MFCCs and principal component analysis (PCA) taking advantage of mobile sensor network is proposed to overcome the deficiencies in the approaches that utilize single measurement. In the improved approach, the acceleration data is first collected from all the vehicles within a certain period. Then, the transformed features that are related to bridge damage are extracted from MFCCs and PCA. The damage can be identified by comparing the distributions of these transformed features. The results from the numerical analysis and lab experiments show that the approach not only identifies the existence of the damage, but also provides useful information about severity.

Worden et al. in [7] proposed that if the structure of interest is subject to changes in its environmental or operational conditions, one must understand the effects of these changes in order that one does not falsely claim that damage has occurred when changes in measured quantities are observed. This problem – the problem of confounding influences – is particularly pressing for civil infrastructure where the given structure is usually openly exposed to the weather and may be subject to strongly varying operational conditions. One approach to understanding confounding influences is to construct a data-based response surface model that can represent measurement variations as a function of environmental and operational variables. The models can then be used to remove environmental and operational variations so that change detection algorithms signal the occurrence of damage alone. The current paper is concerned with such response surface models in the case of SHM of bridges. In particular, classes of response surface models that can switch discontinuously between regimes are discussed.

Goulet et al. in [8] proposed a bayesian dynamic linear models (BDLMs) are traditionally used in the fields of applied statistics and machine learning. This paper performs an empirical validation of BDLMs in the context of structural health monitoring (SHM) for separating the observed response of a structure into subcomponents. These subcomponents describe the baseline response of the structure, the effect of traffic, and the effect of temperature. This utilization of BDLMs for SHM is validated with data recorded on the Tamar Bridge (United Kingdom). This study is performed in the context of large-scale civil structures in which missing data, outliers, and nonuniform time steps are present. The study shows that the BDLM is able to separate observations into generic subcomponents to isolate the baseline behavior of the structure.

Vagnoli et al. in [9] proposed that one of the most important aspects of evaluation of the reliability of the overall railway transport system is bridge structural health monitoring, which can monitor the health state of the bridge by allowing an early detection of failures. Therefore, a fast, safe and cost-effective recovery of the optimal health state of the bridge, where the levels of element degradation or failure are maintained efficiently, can be achieved. In this article, after an introduction to the desired features of structural health monitoring, a review of the most commonly adopted bridge fault detection methods is presented. Mainly, the analysis focuses on model-based finite element updating strategies, non-model-based (data-driven) fault detection methods, such as artificial neural network, and Bayesian belief network–based structural health monitoring methods. A comparative study, which aims to discuss and compare the performance of the reviewed types of structural health monitoring methods, is then presented by analysing a short-span steel structure of a railway bridge.

Neves et al. in [10] proposed that bridges, in particular, are critical links in today's transportation networks and hence fundamental for the development of society. In this context, the demand for novel damage detection techniques and reliable structural health monitoring systems is currently high. This paper presents a model-free damage detection approach based on machine learning techniques. The method is applied to data on the structural condition of a fictitious railway bridge gathered in a numerical experiment using a three-dimensional finite element model. Data are collected from the dynamic response of the structure, which is simulated in the course of the passage of a train, considering the bridge in healthy and two different damaged scenarios. In the first stage of the proposed method, artificial neural networks are trained with an unsupervised learning approach with input data composed of accelerations gathered on the healthy bridge. Based on the acceleration values at previous instants in time, the networks are able to predict future accelerations.

5. PERFORMANCE METRICS

The performance evaluation of the statistical machine learning based techniques are evaluated in terms of the common metrics discussed here [20]-[21]:

Mean Square Error:

It is mathematically defined as:

$$\mathbf{mse} = \frac{1}{n}\sum_{I=1}^{N}(\mathbf{X} - \mathbf{X}')^2 \qquad \mathbf{(12)}$$

Here,

X is the predicted value and

X' is the actual value and

n is the number of samples.

Mean Absolute Percentage Error (MAPE):

It is mathematically defined as:

$$\mathbf{MAPE} = \frac{100}{M}\sum_{t=1}^{N}\frac{E - E_t|}{E_t} \qquad \mathbf{(13)}$$

Here,

E_t and $E_t^{\sim}$ stand for the predicted and actual values respectively.

Iterations:

The iterations denote the number of cycles of training needed to reach convergence.

Accuracy:

The accuracy is defined as:

$$\mathbf{Accuracy} = \mathbf{100} - \mathbf{MAPE\%} \qquad \mathbf{(14)}$$

6. CONCLUSION

It can be concluded from previous discussions that it is mandatory to make a quick estimation and evaluation of the damages is necessary so as to immediately stop the cascading effect of the possible failure and casualties. This paper presents the basis of the employment of statistical and machine leaning based techniques for the estimation of bridge failure in the presence of seismic effects. The salient points pertaining to the contemporary literature in the domain has bene cited along with the brief description of the model being used. Finally the evaluation metrics for the techniques employed have been discussed.

7. REFERENCES

[1] S Mangalathu, SH Hwang, E Choi, JS Jeon, "Rapid seismic damage evaluation of bridge portfolios using machine learning techniques", Engineering Structures, Elsevier 2020, vol. 201, pp.109785

[2] A Malekjafarian, C Moloney, F Golpayegani, "Drive-by bridge health monitoring using multiple passes and machine learning", European Workshop on Structural Health Monitoring, Springer 2020, pp 695-703.

[3] A Guo, A Jiang, J Lin, X Li, "Data mining algorithms for bridge health monitoring: Kohonen clustering and LSTM prediction approaches", The Journal of Supercomputing, Springer 2020, vol. 76, pp.932–947.

[4] Y Bao, Z Chen, S Wei, Y Xu, Z Tang, H Li, "The state of the art of data science and engineering in structural health monitoring", Journal of Structural Engineering, Elsevier 2019, vol. 5, issue 2, pp. 234-242.

[5] M Silva, A Santos, E Figueiredo, "Damage detection for structural health monitoring of bridges as a knowledge discovery in databases process", Data Mining in Structural Dynamic Analysis, Springer 2019, pp 1-24.

[6] Q Mei, M Gül, M Boay, "Indirect health monitoring of bridges using Mel-frequency cepstral coefficients and principal component analysis", Mechanical Systems and Signal Processing, Elsevier 2019, vol.119, pp.523-546.

[7] K Worden, EJ Cross, "On switching response surface models, with applications to the structural health monitoring of bridges", Mechanical Systems and Signal Processing,, Elsevier 2018, vol.98, pp. 139-156.

[8] JA Goulet, K Koo, "Empirical validation of bayesian dynamic linear models in the context of structural health monitoring", Journal of Bridge Engineering 2018, vol.23, issue.2.

[9] M Vagnoli, R Remenyte-Prescott, "Railway bridge structural health monitoring and fault detection: State-of-the-art methods and future challenges", Journal of Structural Health Monitoring, SAGE Publications 2018, Vol 17, Issue 4, pp. 971-1007.

[10] AC Neves, I Gonzalez, J Leander, "Structural health monitoring of bridges: a model-free ANN-based approach to damage detection", Journal of Civil Structural Health Monitoring, Springer 2017, vol. 7, pp.689–702.

[11] J. Peng, S. Zhang, D. Peng and K. Liang, "Application of machine learning method in bridge health monitoring," 2017 Second International Conference on Reliability Systems Engineering (ICRSE), 2017, pp. 1-7.

[12] EK Chalouhi, I Gonzalez, C Gentile, R Karoumi -, "Damage detection in railway bridges using machine learning: application to a historic structure", Procedia Engineering, Elsevier 2017, vol.199, pp. 1931-1936.

[13] M Alipour, DK Harris, LE Barnes, "Load-capacity rating of bridge populations through machine learning: Application of decision trees and random forests", Journal of Bridge Engineering 2017, vol.22, issue.10.

[14] A Anaissi, NLD Khoa, S Mustapha, "Adaptive one-class support vector machine for damage detection in structural health monitoring", Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2017: Advances in Knowledge Discovery and Data Mining, Springer 2017, pp 42-57.

[15] A Diez, NLD Khoa, MM Alamdari, Y Wang, "A clustering approach for structural health monitoring on bridges", Journal of Civil Structural Health Monitoring, Springer 2016, vol. 6, pp.429–445.

[16] M Malekzadeh, FN Catbas, "A machine learning framework for automated functionality monitoring of movable bridges", Dynamics of Civil Structures, Springer 2016, vol. 2, pp 57-63.

[17] H Li, J Ou, "The state of the art in structural health monitoring of cable-stayed bridges", Journal of Civil Structural Health Monitoring, Springer 2016, vol. 6, pp.43–67.

[18] R Ghiasi, P Torkzadeh, M Noori, "A machine-learning approach for structural damage detection using least square support vector machine based on a new combinational kernel function", Journal of Health Monitoring, SAGE publications, vol. 15, issue. 3, pp.302-316.

[19] D. Sun, V. C. S. Lee and Y. Lu, "An intelligent data fusion framework for structural health monitoring," 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), 2016, pp. 49-54.

[20] I Gonzalez, R Karoumi, "BWIM aided damage detection in bridges using machine learning", Journal of Civil Structural Health Monitoring, Springer 2015, vol. 5, pp.715–725.

[21] C Cappello, D Zonta, M Pozzi, B Glisic, "Impact of prior perception on bridge health diagnosis", Journal of Civil Structural Health Monitoring, Springer, vol. 5, pp.509–525.