



## Article

# Online Hybrid Learning Methods for Real-Time Structural Health Monitoring Using Remote Sensing and Small Displacement Data

Alireza Entezami <sup>1,\*</sup>, Ali Nadir Arslan <sup>2</sup>, Carlo De Michele <sup>1</sup> and Bahareh Behkamal <sup>1</sup>

<sup>1</sup> Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milano, Italy; carlo.demichela@polimi.it (C.D.M.); bahareh.behkamal@polimi.it (B.B.)

<sup>2</sup> Finnish Meteorological Institute (FMI), Erik Palménin aukio 1, FI-00560 Helsinki, Finland; ali.nadir.arslan@fmi.fi

\* Correspondence: alireza.entezami@polimi.it

**Abstract:** Structural health monitoring (SHM) by using remote sensing and synthetic aperture radar (SAR) images is a promising approach to assessing the safety and the integrity of civil structures. Apart from this issue, artificial intelligence and machine learning have brought great opportunities to SHM by learning an automated computational model for damage detection. Accordingly, this article proposes online hybrid learning methods to firstly deal with some major challenges in data-driven SHM and secondly detect damage via small displacement data from SAR images in a real-time manner. The proposed methods contain three main parts: (i) data augmentation by Hamiltonian Monte Carlo and slice sampling for addressing the problem of small displacement data, (ii) data normalization by an online deep transfer learning algorithm for removing the effects of environmental and/or operational variability from augmented data, and (iii) feature classification via a scalar novelty score. The major contributions of this research include proposing two online hybrid unsupervised learning methods and providing effective frameworks for online damage detection. A small set of displacement samples extracted from SAR images of TerraSar-X regarding a long-term monitoring scheme of the Tadcaster Bridge in United Kingdom is applied to validate the proposed methods.

**Keywords:** structural health monitoring; pre-collapse prediction; online learning; remote sensing; small data; bridge



**Citation:** Entezami, A.; Arslan, A.N.; De Michele, C.; Behkamal, B. Online Hybrid Learning Methods for Real-Time Structural Health Monitoring Using Remote Sensing and Small Displacement Data. *Remote Sens.* **2022**, *14*, 3357. <https://doi.org/10.3390/rs14143357>

Academic Editor: Fabio Tosti

Received: 6 June 2022

Accepted: 10 July 2022

Published: 12 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In civil engineering, structural health monitoring (SHM) is a new technology for evaluating the integrity and the safety of civil structures and infrastructures by using data acquired from diverse sensors and computational techniques. On this basis, an SHM strategy intends to obtain insightful information about the current or unknown condition of a civil structure from measured responses over time. The key role of SHM is to detect any abnormal change in structural behavior and measured responses of the structure. This is because the performance of important civil structures (e.g., bridges, high-rise buildings, dams, etc.) under operational and environmental conditions may decrease stemming from aging, material deterioration, extreme events (e.g., earthquakes, strong wind, flood, hurricane, etc.), overloads, impact of a foreign object, etc. Accordingly, a long-term SHM system for anomaly/damage detection is mandatory for such structural systems that significantly assist civil engineers in predicting and avoiding catastrophic events, such as partial or global collapse, and also governmental authorities in reducing economic and human losses [1–3].

Generally, there are numerous methods for implementing an SHM project depending upon the type of sensing technology and data acquisition system (i.e., contact-based vs. non-contact-based sensors) [4–6], the type of data (i.e., acceleration time histories, displacements,

strains, images, videos, etc.) [5,7], the type of civil structure and built environment, the type of data processing, the type of computational/statistical method (i.e., model-driven vs. data-driven) [8,9], and the number of measurements (i.e., short-term vs. long-term [10–12]), etc. Despite most of the aforementioned topics have been evaluated properly, the issue of data processing of an SHM system needs further research.

Depending upon the type of data processing, an SHM system can be categorized as batch/offline processing and stream/online processing. The key rationale behind batch processing is to process a relatively large volume of data, all at once, through an offline learning strategy. Actually, it presents an efficient technique to perform an SHM project when large amounts of data can be collected over a period of time. Under such circumstances, it can be considered that there is sufficient training data in an effort to develop a statistical/computational model based on the concept of machine learning. On the contrary, stream processing relies on the premise that an initial model constructed from early data can instantaneously be analyzed and updated by receiving new data in an online fashion [13].

More precisely, offline or batch learning aims at representing the training of a machine learning model so that the data is initially collected over a period of time and then the model of interest is learned with the same collected data. Despite the popularity and the applicability of this technique, particularly in SHM [14–20], it cannot take advantage of incremental learning from stream data. In some cases, measuring long-term vibration data as well as all possible environmental and/or operational conditions in a fixed period is attempted. The main idea lies in the fact that one can learn a promising model in a fixed and long training time if it is possible to measure sufficient environmental/operational and structural data. In this case, one can ensure that all possible uncertainty conditions are incorporated. On the other hand, one assumes that the structure does not suffer from damage during the training period. Although this assumption is plausible, two major limitations emerge. First, it cannot be ensured that the same pattern of variability resulting from the environmental and/or operational conditions is available in new data measured after the training period, even in an undamaged structure. Simply speaking, it is natural to consider that newly unpredictable environmental and/or operational conditions occur out of the training period. As such conditions are often unknown and unpredictable, and they do not incorporate into the learning process, false alarms (false positive errors) are the main errors in decision-making. Second, the structure may suffer from any kind of damage, e.g., a strong collision with the pier of a bridge or the occurrence of an earthquake, during collecting the training samples in the training period. Therefore, it is illogical to suppose that the structure does not sustain any damage or abnormal change during the training time.

For these reasons, online or incremental learning under stream processing can fulfill a more reliable and realistic condition of real-time SHM. In this regard, the expression “real-time” mentions the implementation of data analysis, update of the learned model, and an online decision by receiving new data. In online learning, the learning process is performed in an incremental or online fashion by continuously feeding data, when it arrives, and updating the initial model through such new data. As usual, an online learning procedure starts by using prior small data to train an initial machine learner (model). After measuring or receiving new data, it feeds into the model so as to update it or make a decision. Accordingly, one needs to consider a criterion for distinguishing between updating and decision-making in the problem of SHM. In case of ensuring the undamaged condition of the structure, the learning process continues by updating the model of interest; otherwise, this process terminates. In this case, the last model is considered the final trained model and any new measured information should serve as test data. In the field of SHM, several studies exploited the benefits of online learning for detecting damage in civil and mechanical systems [21–24].

Although the online learning approach has succeeded in real-time SHM, some drawbacks should be appropriately considered. One of them pertains to the issue that if the

data used in developing the online model contains uncertainty or variability, the outputs of the model may be erroneous. In the context of SHM, this challenge refers to environmental and/or operational variability (EOV). This is due to the fact that such variations can produce changes in the inherent properties of a civil structure (e.g., mass and stiffness) similar to damage [12,25]. Accordingly, the structure may sustain damage; however, the EOV conditions mask it and do not allow the technical method to alarm the emergence of damage. This phenomenon often occurs when the magnitude of the EOV condition is larger than the severity of the damage, particularly in minor damage cases, which refers to false negative (i.e., false detection or Type II) error. On the other hand, the structure may mainly be undamaged; however, the high rate of EOV leads to another erroneous result in which case the method incorrectly alarms the occurrence of damage. This phenomenon often occurs in sudden or harsh EOV conditions leading to a false positive (i.e., false alarm or Type I) error. Because both false negative and false positive errors are concerned with human and economic losses, it is indispensable to remove any variability conditions from measured data either in an offline learning or an online learning strategy.

The other limitation of online learning methods pertains to the sensing technology and the type of measured data. In most cases, the measurement process is based on traditional contact-based sensors (e.g., accelerometers, strain gages, etc.). Alongside a computer-vised technology for some SHM applications, such as surface crack detection [26,27] corrosion [28] and bolt loosening detection [29], the use of satellite-based synthetic aperture radar (SAR) to SHM, called here SAR-based SHM, has recently received increasing attention from civil engineers due to its great advantages [30–32]. First, this technology can provide prior information on the structure (i.e., based on the date of the satellite launch), particularly previously baseline data. Second, it is highly cost-effective compared to many ground-based sensing technologies without some of their limitations, such as dense sensor networks, fault sensors, sensor malfunction systems, sensor placement optimization, etc. Third, the SAR-based SHM through remote sensing can supply surface displacements over time with high-resolution satellite images (i.e., up to 2–5 m operating at C-band in roughly 5.6 cm wavelength and approximately 1 m for X-band in roughly 3 cm wavelength). In other words, this technology allows civil engineers to have surface displacements with a precision of up to millimeter-scale [33]. For this reason, it is necessary to identify some point targets on a civil structure in an effort to show stable scattering properties along a time series of satellite images.

The majority of SAR-based SHM applications have been based on direct analyses of displacement data obtained from some satellite images. In this regard, Selvakumaran et al. [34] utilized and analyzed 22 SAR images from 8 target points to investigate the scour failure of a moderate-span masonry bridge in Tadcaster, United Kingdom. Zhu et al. [35] considered a set of 57 COSMO-SkyMed stripmap SAR images acquired from 17 February 2013 to 14 January 2018 in order to monitor the settlement of a building in Foshan, China. Qin et al. [36] analyzed 18 displacement samples at 3 target points from a bridge. In another research, Qin et al. [37] assessed the structural performance and behavior of an arch bridge using a set of 96 SAR images from different satellites. Milillo et al. [38] used 201 SAR images of some satellites in a long-term SHM strategy from 2010 to 2015 to monitor a dam in Italy. Haung et al. [39] analyzed 29 SAR images at 6 target points in order to demonstrate the correlation between the displacement samples from these images and temperature variations. Milillo et al. [40] evaluated the global collapse on the Morandi Bridge, Genova, Italy through a large number of SAR images from different satellites to conduct a long-term SHM. In their research, the authors only analyzed the displacement data in the vicinity of the collapsed area of the bridge.

In contrast to vibration-based SHM techniques by traditional sensing technologies and popular data (e.g., acceleration time histories, strain gages, etc.), which can contain adequate vibration data, the SAR-based SHM is often performed by a small number of data, as the above-mentioned references confirm this claim. This means that despite the possibility of

implementing a long-term monitoring assessment (e.g., five years [38]) in a SAR-based SHM project, a few satellite images are considered to extract displacement samples. On the other hand, compared to the vibration- and vision-based SHM projects, the provision of a large number of SAR images may be problematic due to some restrictions such as extremely large sizes of satellite images and memory space limitations for collecting numerous images. Therefore, SAR-based SHM is a process under small data. The other limitation of the classical SAR-based SHM techniques is to directly analyze the small displacement samples. There is no doubt that civil structures are complex and expensive systems that play crucial roles in every society. In the problems of construction, maintenance, and reconstruction, all of them require huge investments [33]. On this basis, it may not be logical to use simple and unreliable SHM techniques despite applying a cost-effective sensing technology (i.e., remote sensing). As mentioned earlier, the false positive and false negative errors cause economic and human losses, respectively. Thus, it is essential to develop a more rigorous approach to SHM of civil structures via SAR images.

For all the challenges discussed here, this article aims to propose automated online hybrid learning methods compatible with an unsupervised learning class. These methods consist of three main steps, including: (i) data augmentation by two Markov Chain Monte Carlo (MCMC) algorithms, that is, Hamiltonian Monte Carlo (HMC) and slice sampling (SLS), for dealing with the problem of small data, (ii) data normalization by an unsupervised artificial neural network (ANN)-based strategy in an online deep transfer learning (ODTL) algorithm using sequential auto-associative neural networks for addressing the problem of EOV conditions, and (iii) feature classification by a novelty or decision-making index based on the Euclidean norm of the Mahalanobis-squared distances (MSDs), called here EMSD. The main differences between the online hybrid learning methods pertain to the type of augmented displacement samples gained by two different MCMC algorithms. On this basis, the ODTL process yields different outputs depending on the augmented data. The major contributions and innovations of this article can be summarized as: (i) proposing two innovative online hybrid learning methods called HMC-ODTL-EMSD and SLS-ODTL-EMSD for health monitoring of civil structures using SAR images, (ii) proposing two data augmentation algorithms based on HMC and slice sampling, (iii) proposing a novel online data normalization by using the concept of DTL through sequential auto-associative neural networks, and (iv) defining a novelty index called EMSD for providing scalar scores for decision-making. A small set of displacement samples extracted from satellite images of TerraSar-X regarding a long-term monitoring scheme of the Tadcaster Bridge in the United Kingdom is utilized to validate the proposed methods.

## 2. Online Hybrid Learning Methods

### 2.1. Data Augmentation by MCMC

The proposed methods in this article aim at dealing with the problems of small data and EOV conditions. For this first problem, these methods exploit a data augmentation strategy based on the MCMC and the HMC algorithm to augment small displacement samples. The augmented data samples are applied to visualize the EOV. In probability theory, MCMC is a computer-driven sampling method that allows one to characterize a probability distribution model by randomly sampling values out of the distribution of interest without any entire knowledge about its mathematical properties [41]. The term “Monte Carlo” is the practice of estimating the properties of a distribution by examining random samples from the distribution of interest. Moreover, the term “Markov Chain” refers to a sequential process of random sample generation in such a way that each new sample depends only on the one before itself, i.e., new samples do not depend on any samples before the previous [42]. Recently, Entezami et al. [43] utilized the MCMC method to develop a probabilistic framework for a threshold estimator by small output samples regarding the problem of damage localization in SHM. This article attempts to take advantage of this technique for data augmentation.

### 2.1.1. Hamiltonian Monte Carlo Sampling

The HMC sampling is a gradient-based MCMC method that aims to generate samples from a target probability density [44], i.e., the multivariate Gaussian distribution in this study. Assume that the vector  $\mathbf{d} \in \mathbb{R}^{p \times 1}$  contains  $p$ -dimensional displacement samples. Hence, one attempts to generate an  $N$ -dimensional multivariate set of displacement samples. The HMC sampling is based on a logarithmic function of the target distribution, its gradient, and a momentum vector  $\lambda$ . Using these details, one needs to define a Hamiltonian function  $H(\mathbf{d}, \boldsymbol{\Omega})$  based on Hamiltonian dynamics as follows:

$$H(\boldsymbol{\delta}, \boldsymbol{\lambda}) = U(\boldsymbol{\delta}) + V(\boldsymbol{\lambda}) \quad (1)$$

where  $\boldsymbol{\delta} \in \mathbf{d}$ ;  $U(\boldsymbol{\delta})$  denotes the logarithmic function of the probability of interest and  $V(\boldsymbol{\lambda}) = \frac{1}{2} \boldsymbol{\lambda}^T \mathbf{M}^{-1} \boldsymbol{\lambda}$  in which  $\mathbf{M}$  is a symmetric, positive definite matrix, (i.e., called the mass matrix) that is typically diagonal, and is often a scalar multiple of the identity matrix. On this basis,  $V(\boldsymbol{\lambda})$  can be equivalent to minus the logarithmic probability density of the zero-mean Gaussian distribution with covariance matrix  $\mathbf{M}$  [44]. Therefore, Hamiltonian dynamics operate on  $\boldsymbol{\delta}$  and  $\boldsymbol{\lambda}$  to develop Hamiltonian equations that aim at determining how the vectors  $\boldsymbol{\delta}$  and  $\boldsymbol{\lambda}$  change over time,  $t$ , in the following forms:

$$\frac{d\delta_i}{dt} = \frac{\partial H}{\partial \lambda_k} = \frac{\partial V(\boldsymbol{\lambda})}{\partial \lambda_k} \quad (2)$$

$$\frac{d\lambda_i}{dt} = \frac{\partial H}{\partial \delta_k} = \frac{\partial U(\boldsymbol{\delta})}{\partial \delta_k} \quad (3)$$

where  $k = 1, 2, \dots, p$ . Based on Equations (2) and (3), and the initial values of the parameter  $\mathbf{d}_0$  and the momentum  $\lambda_0$  at time  $t_0$ , it is possible to simulate the parameter and the momentum at any time  $t = t_0 + \Delta t$ , where  $\Delta t$  is the duration time or step size, by the leapfrog method [44]. Accordingly, the main objective is to predict  $\mathbf{d}$  under pre-defined chain ( $C$ ) and sampling ( $N$ ) numbers. As such, the HMC algorithm is performed to draw  $N$  samples of  $\mathbf{d}$  from the target probability distribution, designated here as  $\mathbf{D} \in \mathbb{R}^{p \times N}$ , and  $C$  chains under an iterative strategy (i.e.,  $i = 1, \dots, C$  and  $j = 0, \dots, N - 1$  for sampling  $\mathbf{D}_{j+1}^{(i)}$ ) and the acceptance probability criterion from Gelman–Rubin convergence statistic [45]. If the convergence statistic holds, the simulated parameter at the  $(j + 1)$ th iteration is fixed; otherwise, the simulated parameter of the previous,  $j$ th, iteration should be selected. Once the  $C$  sets of  $\mathbf{D}$  have been determined, the average of these  $C$  sets is considered as the final multivariate data. It should be noted that the same HMC sampling procedure should be carried out for any new small data set. In the context of machine learning, the process of decision-making is based on defining training and test data sets. Accordingly, the augmented data matrix  $\mathbf{D}$  is divided into two parts  $\mathbf{X} \in \mathbb{R}^{p \times n}$  and  $\mathbf{Z} \in \mathbb{R}^{p \times m}$ , where  $n < N$  and  $m < N$ , referring to the training and test datasets. These matrices are fed into the proposed data normalization algorithm to remove any influence of any variability conditions from the augmented data.

### 2.1.2. Slice Sampling

Slice sampling (called here SLS) is one of the MCMC algorithms for drawing random samples from a statistical distribution. The basic idea behind this technique comes from the fact that any distribution can be sampled by selecting uniformly spaced points under a probability distribution curve. The SLS algorithm generates random samples based on their previous states. The SLS originates with the observation that to sample from a univariate distribution in which case it is possible to sample points uniformly from the region under the curve of its density function and then observe only at the horizontal coordinates of the sample points. A Markov chain that converges to this uniform distribution can be constructed by alternately sampling uniformly from the vertical interval defined by the density at the current point and from the union of intervals that constitutes the horizontal

slice using the plot of the density function defined by this vertical position. The great advantage of SLS is its applicability to a wide variety of distributions. This technique is often simpler to implement than Gibbs sampling and more efficient than simple Metropolis updates due to its ability to adaptively select the magnitude of changes.

The SLS algorithm draws samples from the region under the density function using a sequence of vertical and horizontal steps. At first, it selects a height at random from zero to a PDF. Then, it chooses a new point randomly by sampling from the horizontal slice of the density above the selected height. The SLS can generate random numbers from a distribution with an arbitrary form of the PDF provided that an efficient numerical process is available to find an interval, which is the slice of the PDF. Because the full details of the SLS along with its theoretical aspects are available in Neal [46], its main steps are only mentioned here:

- (1) Assume an initial value  $x_i$  within the domain of the target PDF  $f(x)$ ;
- (2) Draw a real value  $y$  uniformly from  $(0, f(x_i))$ , thereby defining a horizontal “slice” as  $S = \{x: y < f(x)\}$ ;
- (3) Find an interval around  $x_i$  that contains all, or much of the slice  $S$ ;
- (4) Draw the new point  $x_{i+1}$  within this interval;
- (5) Increment  $i \rightarrow i + 1$  and repeat Steps 2–4 until obtain the desired number of samples.

In contrast to HMC sampling, it is not necessary to perform any gradient operation, estimate any parameter, and consider chains. Nonetheless, in a similar manner, the main objective of the SLS is to generate a new multivariate dataset (matrix)  $\mathbf{D} \in \mathbb{R}^{p \times N}$  from the initial small set of the displacement set (vector)  $\mathbf{d} \in \mathbb{R}^{p \times 1}$ . As such, in the following, the augmented displacement matrix  $\mathbf{D}$  is decomposed into two sub-matrices  $\mathbf{X} \in \mathbb{R}^{p \times n}$  and  $\mathbf{Z} \in \mathbb{R}^{p \times m}$ , where  $n < N$  and  $m < N$ , referring to the training and test datasets.

## 2.2. Data Normalization by ODTL

### 2.2.1. Auto-Associative Neural Network

The auto-associative neural network is one of the tried-and-tested unsupervised ANNs that mainly aims to reconstruct the input data at the output layer. This network makes a feed-forward configuration with linear or sigmoid transfer functions as well as back-propagation training algorithms [47]. The architecture of auto-associative neural network entails an input layer, three hidden layers called *mapping*, *bottleneck*, and *de-mapping* and an output layer. The network learns a mapping from given inputs to desired output values by adjusting internal weights to minimize a least-square error objective function. Figure 1 shows a graphical configuration of this network. The input data is fed into the input layer. By contrast, the output layer reconstructs the input data with the same dimension. In the network topology, the bottleneck layer plays a key role in the functionality of the auto-associative neural network, as it enforces an internal encoding and compression of the input data. Hence, according to Kramar’s recommendation [47], the bottleneck layer should have smaller neurons than the other hidden layers. The great advantage of the auto-associative neural network is its unsupervised learning aspect and ability to remove the noise, outliers, and any variability condition in data (features) [47–49].

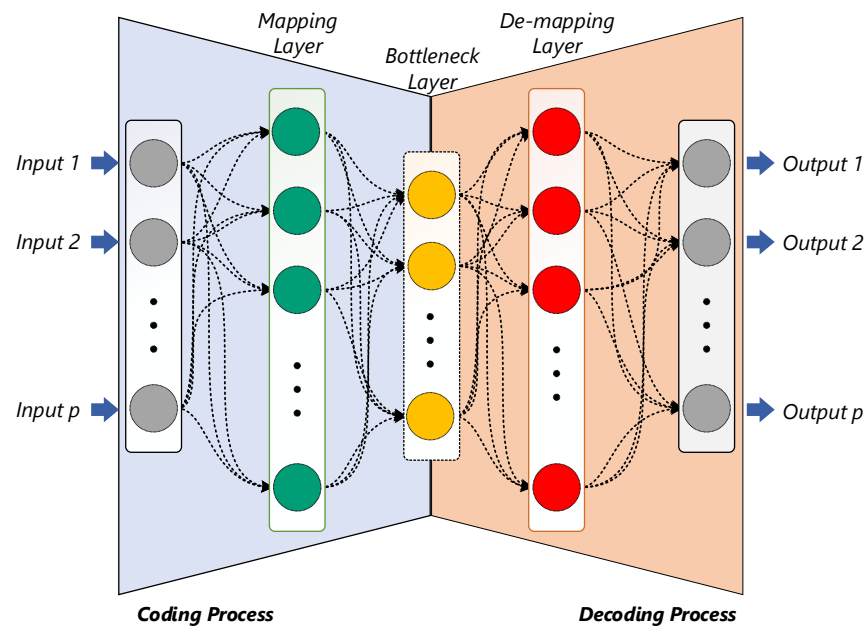
Suppose that  $x_1, \dots, x_p$  are  $p$  data points fed into the input layer in which case it contains  $p$  neurons. The auto-associative neural network is viewed as a serial combination of two single-hidden layer networks. In other words, it contains a coding process in the first network and a de-coding process in the second. For this reason, it highly resembles the autoencoder neural network [50]. The input, mapping, and bottleneck layers represent the nonlinear function  $G$  aiming at projecting the feature samples  $x_1, \dots, x_p$  to a lower dimension space of  $l_m$  neurons of the mapping layer and then  $l_b$  neurons of the bottleneck layer. The mapping layer often uses a nonlinear transfer function (i.e., the sigmoid function in this article) to map the input data onto the bottleneck layer. The transfer function in this layer can be linear or nonlinear, without affecting the generality of the network. This mapping is expressed as follows:

$$b_i = G_i(\mathbf{x}), \quad i = 1, 2, \dots, l_b \quad (4)$$

where  $b_i$  denotes the output of the  $i$ th bottleneck neuron;  $l_b$  represents the number of neurons of the bottleneck layer; and  $\mathbf{x} = [x_1 \dots x_p]^T$  is the vector of the  $p$ -dimensional feature samples or inputs. Subsequently, the bottleneck, de-mapping, and output layers decode the outputs of the first network through another nonlinear function  $H$  that reproduces an approximation of the inputs from the factors at the output of the bottleneck layer in the following form:

$$\hat{x}_j = H_j(\mathbf{b}), \quad j = 1, 2, \dots, l_m \quad (5)$$

where  $\hat{x}_j$  represents the output or reconstructed data at the  $j$ th output layer.



**Figure 1.** Graphical configuration of an auto-associative neural network.

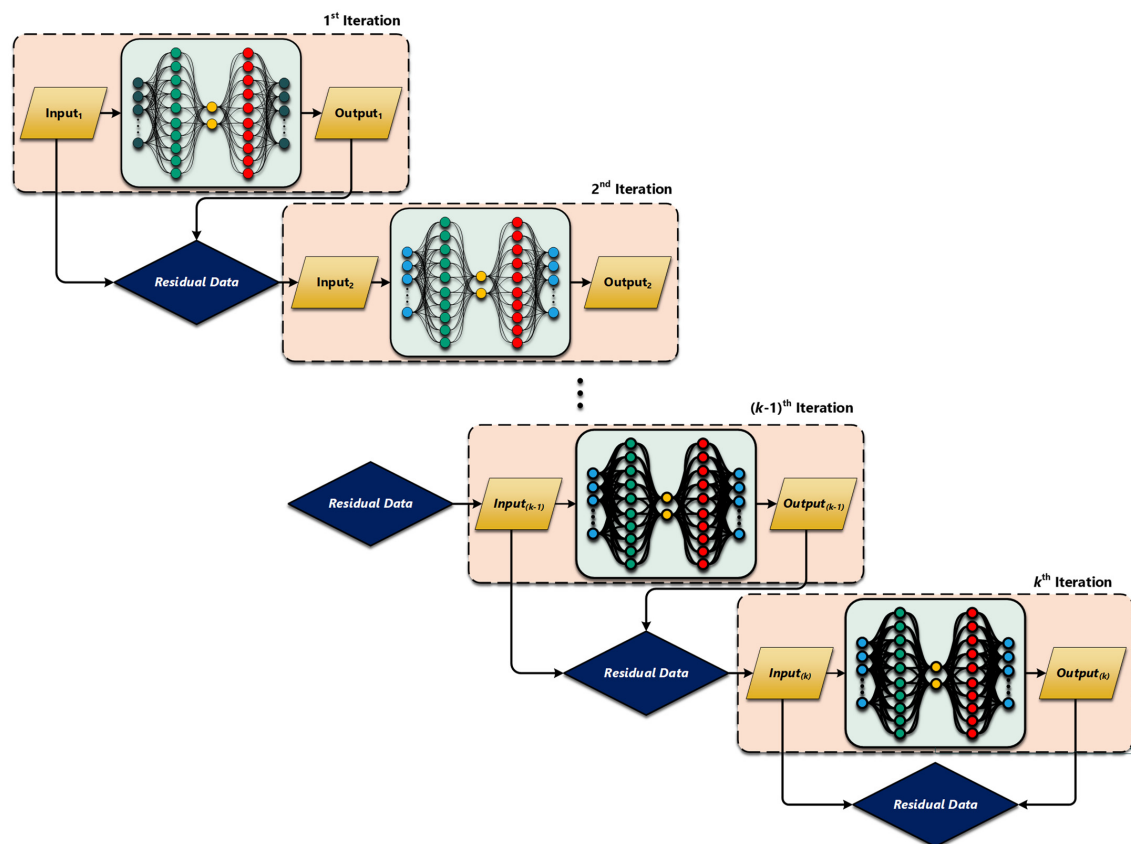
### 2.2.2. Deep Transfer Learning

In this section, it is attempted to introduce a novel method for real-time early damage detection of civil structures based on the concept of online learning or incremental learning. In the context of machine learning, the online learning aims to update the learning process, or the trained model by using new information (training data). In this regard, if the information of interest confirms that it came from the abnormal (damaged) condition of the structure, the learning process terminates and the information is considered as the test data.

The procedure of feature normalization by the proposed ODTL algorithm is the key part of the proposed method. The main purpose of this step is to learn a deep network of sequential auto-associative neural networks. On the other hand, the deep transfer learning (DTL) aims at addressing one of the major demanding issues in machine learning regarding the problem of hyperparameter optimization. It is a strategy for estimating any unknown parameter of a machine learning model that has direct influence on the model's performance. For example, the number of neurons of hidden layers of an ANN is a hyperparameter. According to the concept of transfer learning, one can utilize part of the full information of a well-established model in future tasks. Transfer learning is a new branch of machine learning, particularly deep learning and ANNs, that aims at transforming knowledge and information or the learning process of a related and successful domain (i.e., the source domain) for applications to other areas [51]. In this research, the main hyperparameter optimization pertains to determining the number of hidden layers of the mapping, bottleneck, and de-mapping layers. Even though there are some techniques

to solve this problem, they may not be suitable for SHM and its major challenge, that is, the negative effects of the EOV conditions.

The main idea behind the proposed method is that the auto-associative neural network with pre-defined neurons of the mapping, bottleneck, and de-mapping layers may be sufficient for removing the EOV conditions, without any hyperparameter selection. As mentioned in Figueiredo et al. [52], an auto-associative neural network with 10 and 2 neurons for the mapping (de-mapping) and the bottleneck layers is reasonable for vibration-based applications. In particular, the choice of two neurons for the bottleneck layer can capture the nonlinear relationship between the structural features and the variability conditions. Although this idea provided correct results on Figueiredo et al. [52], Sarmadi et al. [16] demonstrated that the use of one auto-associative neural network by using the aforementioned neuron sizes cannot be effective for SHM under severe variability conditions. Hence, a deep network of sequential ANNs is applied to address this drawback. First, the general concept of this deep network is described and then its online version is discussed in detail. Figure 2 indicates the graphical representation of a deep network of sequential auto-associative neural networks. In essence, the DTL is an iterative algorithm consisting of more than one ANN. The key rationale behind the DTL for removing the EOV conditions is to determine the residuals between inputs and outputs applied to learn an auto-associative neural network at each iteration. In the first step (iteration), the input data is the augmented displacement set. By computing the residuals between the inputs and the outputs, the next ANNs are re-trained by the residual samples of the previous step (iteration). This procedure continues until a stopping condition satisfies at the last iteration. Finally, the last residual set treats as the normalized augmented data.



**Figure 2.** Graphical representation of the proposed DTL based on the sequential auto-associative neural networks.

The only requirement of this method is to find the stopping condition of its iterative algorithm. Although the numbers between 2–10 can provide reasonable results, it is better



to develop an effective approach to find the optimal stopping condition. This approach is based on examining several iteration numbers and computing an error based on the outputs of each iteration. In other words, after each iteration, it is possible to determine a residual set (matrix). Accordingly, one can compute its MSD values and estimate a threshold limit. Using the concept of a false positive, the error rate is calculated at each iteration. Finally, the iteration number with the minimum error is selected as the optimal number of ANNs for the DTL algorithm. It should be clarified that since the determination of the stopping condition is only carried out by the augmented training data, the mentioned error is the false positive of the MSD values.

### 2.2.3. Online Deep Transfer Learning

The use of sufficient (augmented) data is one of the major requirements of the deep network in the DTL algorithm. Simply speaking, it relies upon an offline or batch learning process. Here, one attempts to propose its improved version under an online learning algorithm, which includes two stages. First, an online learning strategy is implemented until the stopping condition ends the learning process. In this stage, the algorithm decides that the new data no longer belongs to the undamaged or normal condition and it is most likely concerned with a damaged state. In other words, the algorithm triggers the emergence of damage. Thus, the second stage begins by learning a one-step-ahead sequential auto-associative neural network. For simplicity, Figure 3 shows the flowchart of the proposed ODTL method in both stages.

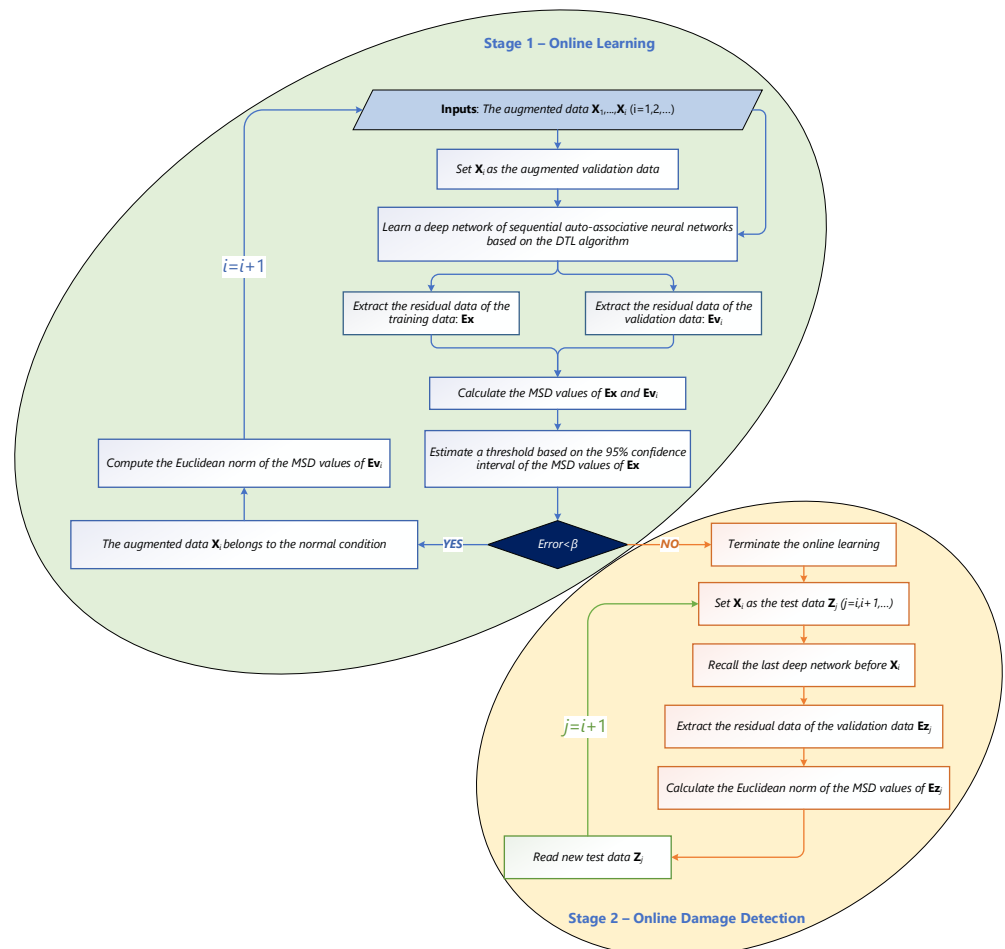


Figure 3. The flowchart of the proposed ODTL method.

More precisely, the proposed ODTL algorithm starts with the first augmented displacement data from the first SAR image. The major goal is to learn a deep network of sequential

auto-associative neural networks and extract the residual dataset. Since the first stage contains only one set of augmented data, this augmented data acts as both the training and the validation sets. In the following, the MSD measure is applied to calculate the distance values of these datasets. Naturally, both datasets yield similar distance quantities. The MSD values of the training data are then considered to determine an alarming threshold (i.e., a criterion for the stopping condition). In this study, a 95% confidence interval of the distance values of the training data is applied to obtain the threshold. The main idea behind the first stage of the proposed ODTL method is to incorporate only the error in decision-making related to the validation data. This is because any new data in the first stage of online learning is defined as validation data. If this set belongs to the normal condition (i.e., the false positive error becomes zero or insubstantial), it is added to the previous datasets to update the learning process; otherwise, it sends to the second stage as the test data related to the damaged condition. In the online learning procedure, one supposes that any new information (augmented validation data) originates from the undamaged or normal condition. Therefore, their distance values should fall under the threshold. According to the number of false positive errors (i.e., the number of samples whose distance values are over the threshold) or its percentage (i.e., the ratio of the number of distances over the threshold per the total number of all distance values), it is possible to make two main decisions:

- (1) If the error is smaller than a pre-defined criterion ( $\beta$ ), it makes sense that the hypothesis of the validation data (i.e., this dataset pertains to the normal condition) is *accurate*. Therefore, one should add it to the previous data (after the first iteration) and a new deep network of sequential auto-associative networks should be learned and updated. This emphasizes that the online learning should be continued. Furthermore, the Euclidean norm of the MSD quantities of the validation dataset are calculated to provide an output at each step of the online learning for real-time damage detection;
- (2) If the error is larger than a pre-defined criterion ( $\beta$ ), it makes sense that the hypothesis of the validation data (i.e., this dataset is related to the normal condition) is *inaccurate*. On this basis, it is necessary to terminate the online learning procedure and start the second stage of the ODTL algorithm regarding real-time or online damage detection. Hence, the label of the augmented validation data is changed to the test data. To provide a novelty score of the test data, the one-step-ahead deep network of sequential auto-associative neural networks from the test data is applied to extract the residuals of the test samples. Subsequently, the Euclidean norm of the MSD values of the test residual samples are calculated to store as novelty scores. When a new data sample is received, it is labeled as a new test point and the aforementioned levels concerning the previous test data are repeated, once again.

### 2.3. Feature Classification by EMSD

To make the final decision about the current state of the structure in the real-time damage detection scheme, it only suffices to determine a novelty score. Unlike most of the unsupervised anomaly detection techniques [10,11,25], the proposed online learning methods do not explicitly use an alarming threshold to compare novelty score with this limit and make a decision. In other words, the threshold estimation is incorporated into the algorithm of the ODTL algorithm, as shown in Figure 3. On the other hand, the proposed novelty score differs from those techniques. In this research, rather than vector-style novelty scores, a scalar index is defined to detect damage in an online manner. This index is the Euclidean norm of the MSD values of the residual samples at each iteration step of the ODTL algorithm.

Assume that  $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_n] \in \mathbb{R}^{p \times n}$  refers to the matrix of residual samples extracted from the deep network of sequential auto-associative neural network. This matrix can be representative of each of the augmented training, validation, and test data. The main objective is to estimate its mean vector  $\mathbf{m}_e \in \mathbb{R}^p$  and covariance matrix  $\Sigma_e \in \mathbb{R}^{p \times p}$ . Having considered  $n$  vectors of  $\mathbf{E}$ , the MSD is derived as:

$$D_{e_i} = (\mathbf{e}_i - \mathbf{m}_e)^T \Sigma_e^{-1} (\mathbf{e}_i - \mathbf{m}_e) \quad (6)$$

Using both Equation (6), one can determine  $n$  distance values  $D_{e_1}, \dots, D_{e_n}$ . Hence, the EMSD of these distances is given by:

$$d_e = \sqrt{(D_{e_1})^2 + \dots + (D_{e_n})^2} \quad (7)$$

Therefore, the scalar novelty score  $d_e$  is considered to only present the outputs of the online learning and online damage detection. It should be noted that when the second stage of the OTSL algorithm starts, this means that the structure suffered from damage.

### 3. Application: The Tadcaster Bridge

In this section, a set of small displacement samples extracted from a few SAR images is considered to demonstrate the applicability and effectiveness of the proposed methods. This dataset belongs to the SAR-based SHM on the Tadcaster Bridge, Tadcaster, North Yorkshire, United Kingdom, which was provided by Selvakumaran et al. [34]. Before presenting the SHM results, it is important to clarify the reasons for incorporating this case-study. This case-study provides a realistic condition of a partial collapse of a full-scale civil structure. Hence, it is very suitable for assessing SHM before a global collapse. It should be noted that the main aim of an SHM strategy is to predict any global collapse. This case study contains a long-term monitoring strategy with a few SAR images that is able to demonstrate the importance of our idea and proposed methods. It considers appropriate distributed target points to capture or to measure displacement data. Such measurements are equivalent to a dense sensor network of contact-based sensors with optimum sensor placement. It utilizes high-resolution SAR images.

#### 3.1. A Brief Description of the Bridge

The Tadcaster Bridge is a historic nine-arch masonry bridge over the River Wharfe in Tadcaster, United Kingdom. The road bridge is believed to date from around 1700. It is the main route connecting the two sides of the town and one of two road crossings in the town, the other being the bridge for the A64 bypass. The bridge is approximately 100 m long and 10 m wide, carrying a single lane of vehicular traffic in each direction and a pedestrian walkway on each side. The present bridge (prior to collapse) comprises two structures of different dates, built side by side to expand the width of the original structure. Documentary evidence suggests that it was built from 1698 to 1699 replacing an earlier bridge on the same site that had been swept away by flood. Flooding events in recent years prior to the collapse meant that the bridge was inspected by divers to detect movement of the riverbed that may have resulted in scour. The Tadcaster Bridge partially collapsed on 29 December 2015 after flooding that followed Storm Eva, and reopened on 3 February 2017 after a major re-building of the damaged area. Figure 4a illustrates this bridge, while Figure 4a,c indicates the damaged area (i.e., partial collapse).

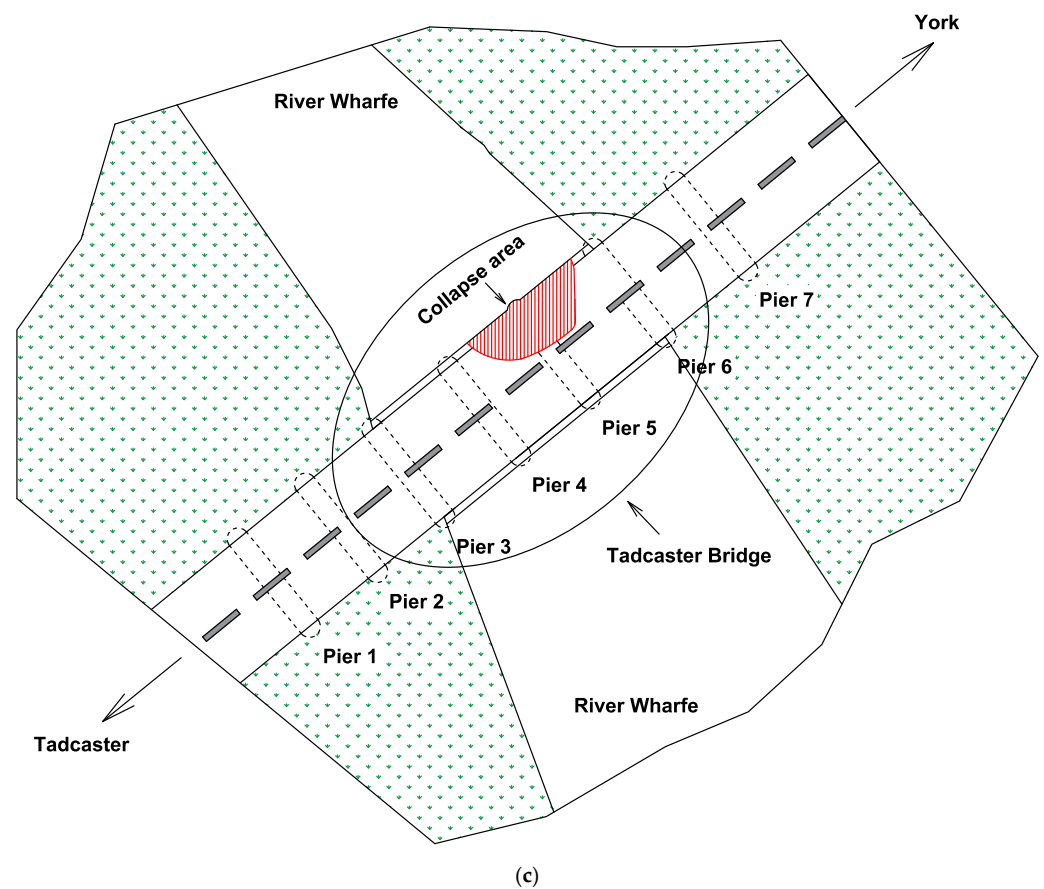


(a)



(b)

Figure 4. Cont.



**Figure 4.** (a) The Tadcaster Bridge, (b) the image of the collapsed area [34], (c) the plan view, pier labels, and collapse area.

To analyze the deformation behavior in the period preceding collapse, 45 TerraSAR-X Stripmap mode images ( $3 \times 3$  m ground resolution) taken prior to the collapse in the period from 9 March 2014 to 26 November 2015 were analyzed. The final acquisition in November was the last image available prior to the bridge collapse on 29 December 2015. These image acquisitions were taken at 11-day intervals, where possible. Based on the DInSAR methodology and Small Baseline Subset (SBAS) technique, a small set of 45 displacement samples at eight points called “a”, “b”, “c”, “d”, “e”, “f”, “g”, and “h” in Figure 5 is extracted to directly analyze them for pre-collapse prediction [34]. On this basis, the target point at “b” is the location of the partial collapse (i.e., the damaged area of the Tadcaster Bridge). As such, the other points are the undamaged areas of the bridge. Figure 6 shows the variations in the displacements of the undamaged and the damaged areas. Indeed, these displacements are the movements of the line of sight (LOS) of the SAR satellite over time, and they were plotted as movements relative to the position of the bridge at the first acquisition, taken on 9 March 2014 (i.e., the first image). Moreover, the Small Baseline Subset (SBAS) technique was utilized to detect eight different distributed scatterer locations across the bridge and then to obtain the displacement data [34]. Notice that each of these target points is equivalent to a classical contact-based sensor. The other important note is that although the SAR images can cover a large area in the unit of km, the process of extraction of displacement samples is based on selecting many target points (scatterers) on the area that one needs to extract displacement samples. This is one of the initial steps in the InSAR techniques. In other words, the interferometric processing generates interferograms to identify stable distributed scatterers. One of the advantages of the InSAR techniques is to distribute scatterers over a large area and also a specific area with the small scale. For example, in the displacement measurements of the Tadcaster

Bridge, Selvakumaran et al. [34] utilized the scatterers over Tadcaster city, the bridge, and the collapse area.

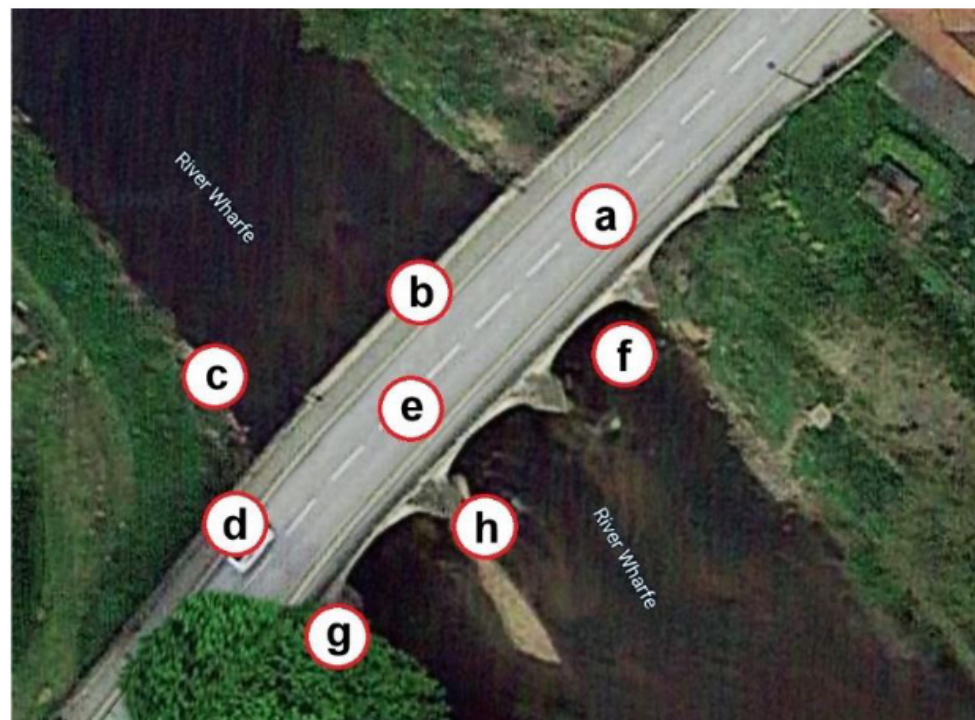


Figure 5. The eight points of displacement samples (i.e., the point “b” is the location of the partial collapse) [34].

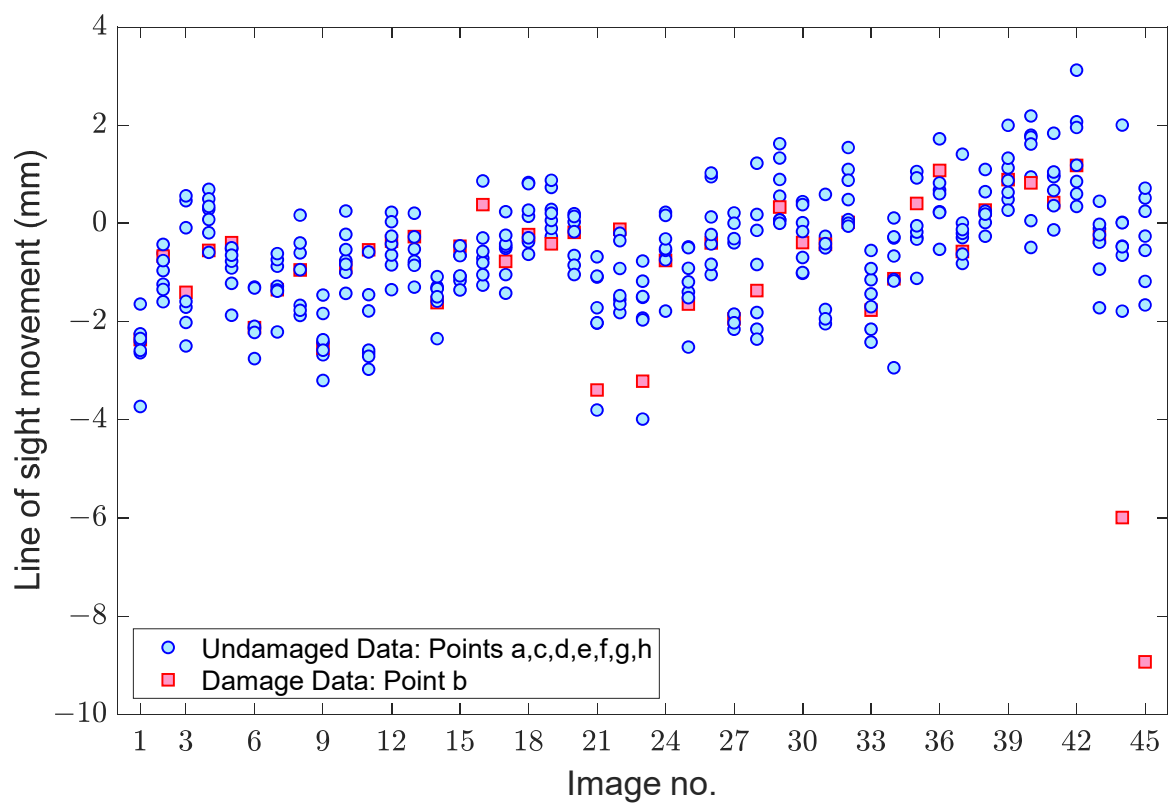


Figure 6. Displacement samples of the eight points regarding the undamaged and damaged areas [34].

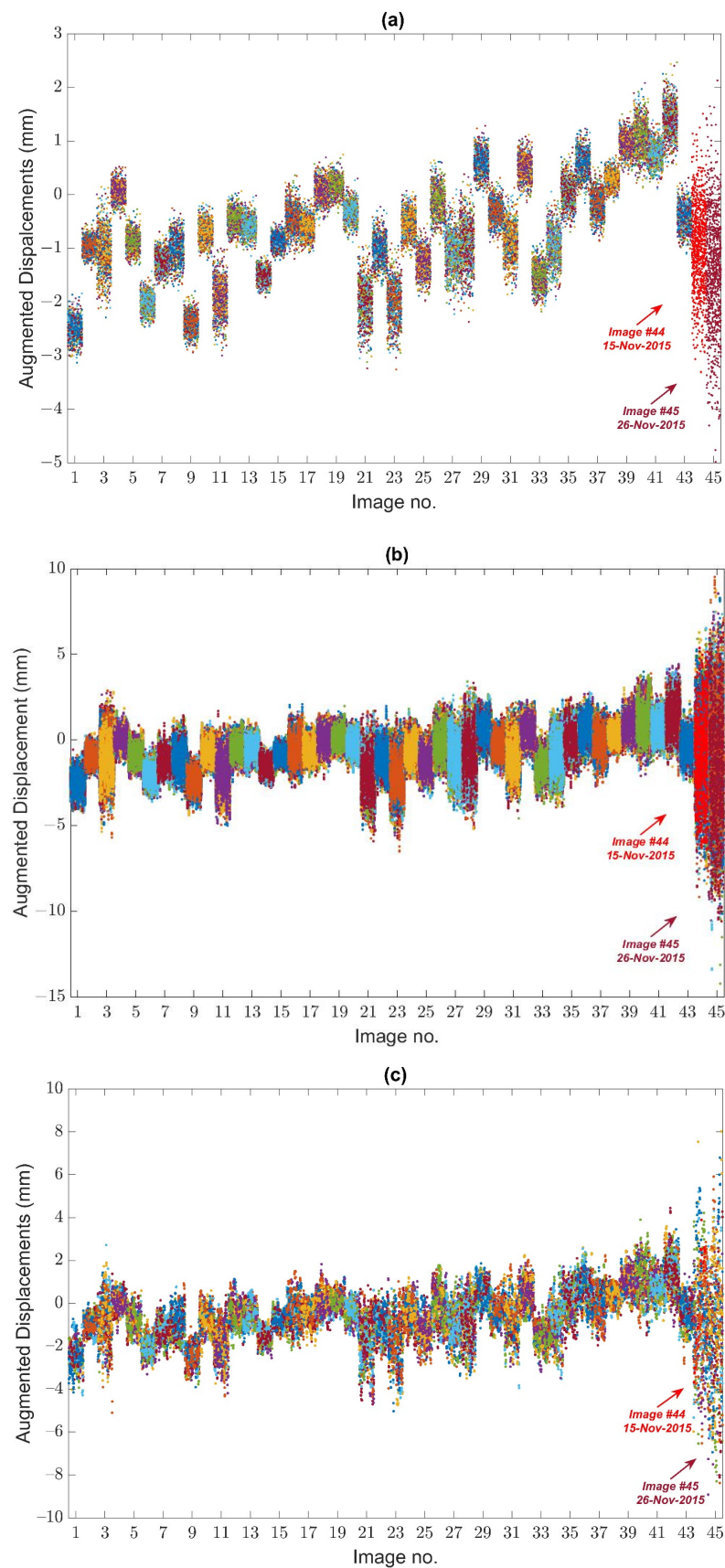
From Figure 6, it is observed that the small set of displacement samples in a long-term monitoring scheme includes randomly unpredictable variations. Apart from the last two displacement samples regarding Images #44 and #45 at the point “b”, there are large variations in the other displacement samples. Such variations are related to the EOV conditions. Moreover, the amounts of displacement samples include positive and negative values. This means that the direct analysis of changes in displacements for decision-making is slightly questionable. Therefore, it would not be reasonable to only lie in a direct data analysis or graphical interpretation for SHM of an important civil structure.

### 3.2. Data Augmentation and Variability Assessment

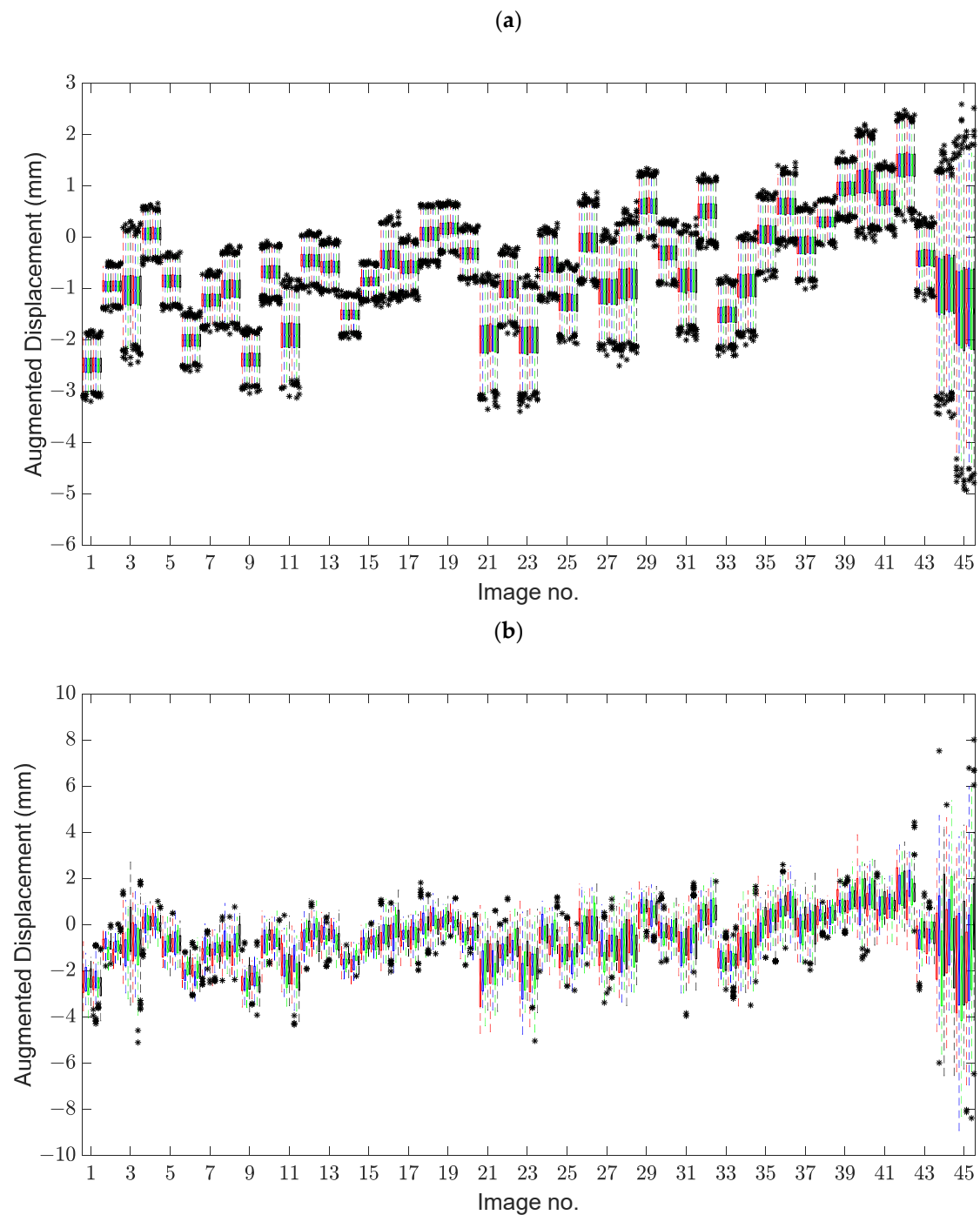
In order to further investigate the effect of EOV conditions, the small displacement data is augmented by using HMC and slice sampling techniques to prepare augmented data. In HMC sampling, the numbers of extended samples  $N$  and chains are set as 100 and 10, respectively. Furthermore, the number of extended samples  $N$  regarding the slice sampling is set as 1000. Accordingly, a small set of each displacement data of the eight target points regarding each image is converted to larger sets of 100 and 1000 augmented data samples.

Figure 7a–c shows the augmented data associated with HMC and slice sampling, respectively. A simple graphical comparison between Figure 6, which shows the real small displacement samples, and Figure 7a reveals that the use of HMC sampling can better observe the EOV conditions. However, as Figure 7b illustrates, this conclusion is not valid for slice sampling. Although this technique significantly helps to recognize variations in the augmented displacement samples of the damaged conditions (i.e., Images #44 and #45), the variations in the other images are very small. To provide a fair comparison between HMC and slice sampling, Figure 7c indicates the augmented displacement data with  $N = 100$ . The same conclusion in Figure 7b is also observable. Another important note is Figure 7 pertains to the detection of abnormal change in the bridge with the direct observation of the augmented data. Although the augmented samples of Images #44 and #45 differ from the other images, there are three limitations of directly applying such information for a critical event, such as pre-collapse detection. At this moment, it is assumed that the condition of the bridge is unknown. First, the augmented data samples of some images roughly resemble the corresponding samples of Images #44 and #45. This means that the direct use of such samples increases the probability of a false alarm. Second, some sets of the augmented data have positive values, whereas others have negative values. This property makes it difficult to accurately decide when the bridge suffered from damage. Third, this study proposes two online learning methods for real-time SHM. This means that it is not necessary to collect all data during a fixed training period. In other words, before the occurrence of damage, it is essential to consider the other data samples in an effort to avoid false alarms or false detection errors.

For further evaluation of the EOV conditions, Figure 8 depicts the box plots of the augmented displacement data obtained from the HMC and slice sampling. In Figure 8a, it is clear that the utilization of data augmentation helps to better observe the EOV conditions by using HMC sampling. However, it is slightly difficult to get the same conclusion via the augmented data from slice sampling. Hence, one can conclude that HMC sampling outperforms slice sampling in order to *graphically* evaluate the rate of variations caused by environmental and/or operational conditions. Nonetheless, the same conclusion in the aforementioned figures is that the occurrence of damage leads to significant increases in the augmented data and their variances, as can be observed in both box plots in Figure 8. In this regard, one can observe that the last image has the largest variance. It should be noted that although the small displacement points at Images 44 and 45, see Figure 6, indicate variances in data, it is difficult to reach the conclusion as good as the augmented data. This indicates how the aforementioned theory interprets variations in real data appropriately.



**Figure 7.** Augmented displacement samples: (a) the HMC sampling  $N = 100$ , (b) the slice sampling  $N = 1000$ , (c) the slice sampling  $N = 100$ .

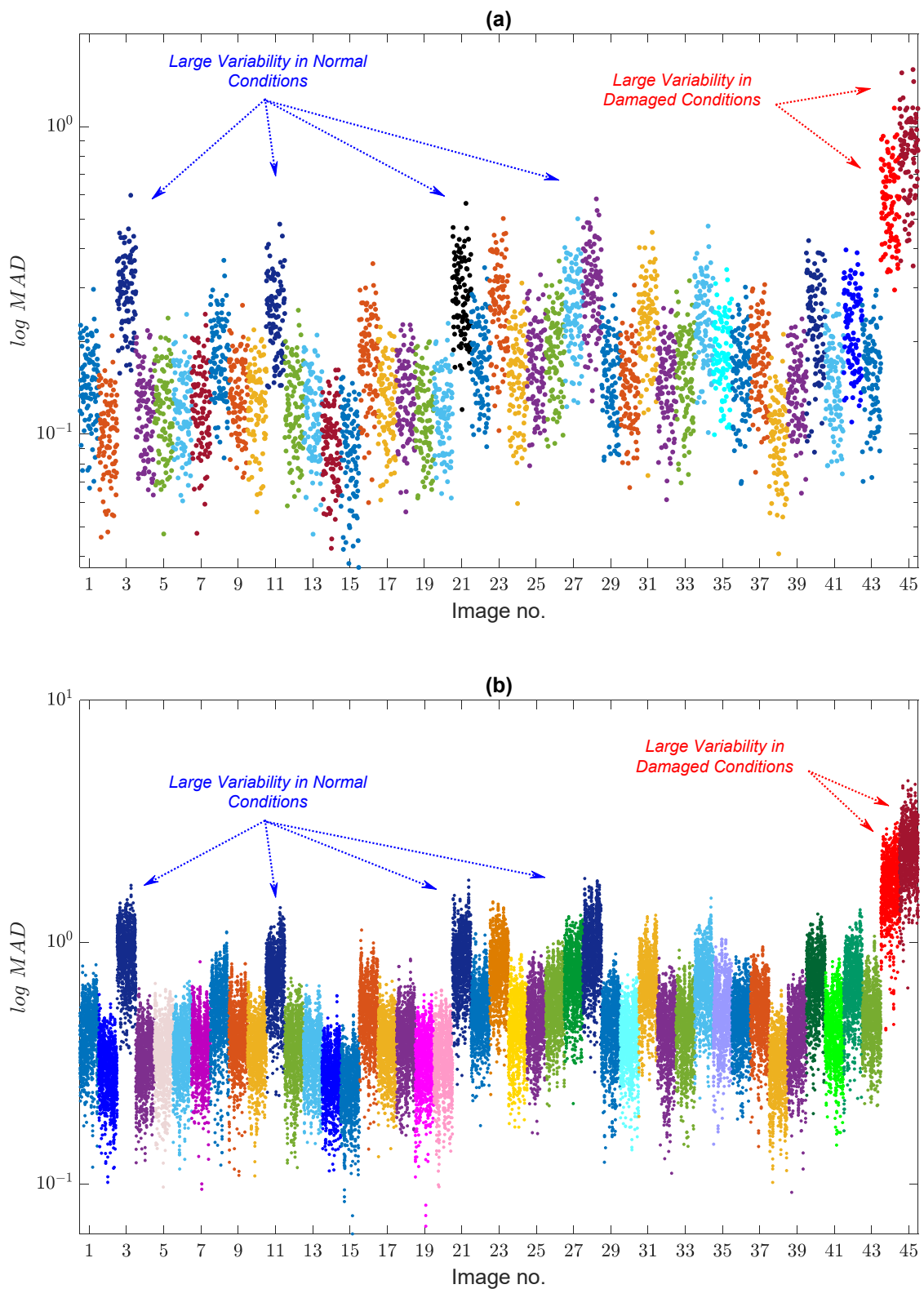


**Figure 8.** Box plot of the augmented displacement samples: (a) the HMC sampling, (b) the slice sampling.

For more rigorous evaluation, Figure 9a,b shows the logarithmic values of the medium absolute deviation (MAD) of the augmented displacement samples obtained from the HMC and slice sampling, respectively. In statistics, MAD is a robust measure for assessing the variability in data. Accordingly, a large MAD quantity is representative of large variability and vice versa. For this reason, the MAD values of the augmented data are computed to investigate the effect of the EOv in a numerical manner. From Figure 9, one can observe that the augmented displacement samples from both sampling techniques could properly indicate the influence of the EOv conditions. As can be seen, along with the augmented data of the damaged conditions (i.e., Images 44 and 45), there are large variations in the augmented displacement samples of some normal conditions. An important note is that



the augmented data for the slice sampling could also indicate the EOV effect. This conveys the importance of numerical analyses rather than graphical assessment.

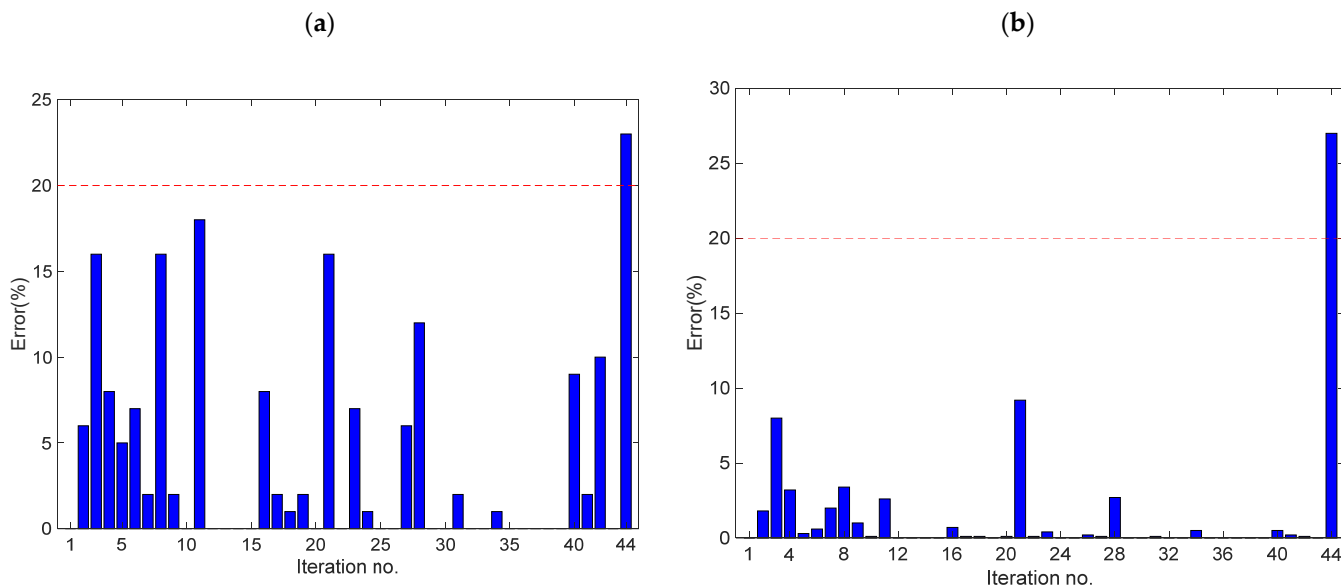


**Figure 9.** Variations in the MAD values of the augmented displacement samples: (a) the HMC sampling, (b) the slice sampling.

### 3.3. Data Normalization and Feature Classification

As mentioned earlier, one of the advantages of the proposed online hybrid learning methods is to simultaneously implement the procedures of data normalization and feature classification. For the first procedure, the main aim is to train an online model based on the ODTL algorithm of sequential auto-associative neural networks. Hence, the iterative online algorithm begins with the augmented displacement set from the first SAR image and  $\beta = 20\%$ . This amount makes sense that one can suppose that the augmented data with the error rate larger than this value is most likely concerned with the damaged condition of the structure.

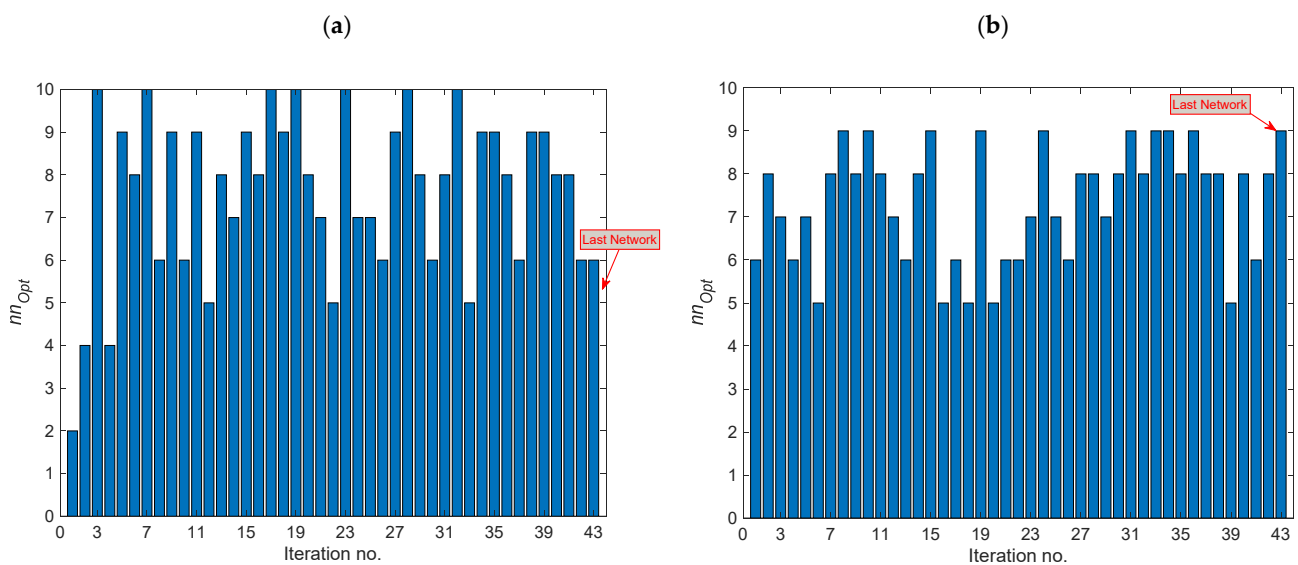
In this regard, Figure 10 shows the variations in the error rates by receiving any new SAR image or new augmented displacement samples. In this figure, the horizontal axis depicts the number of iterations, which is equivalent to the number of augmented data or images. In the iterative algorithm, the sample number of neural networks is set as 10. From Figure 10, one can observe that the error rates are equal to zero at some iterations, while the others are relatively large values. Nonetheless, the largest error value belongs to the 44th iteration, regarding Image #44. This means that the process of ODTL and updating the deep neural network should be terminated. Afterward, this model is applied to extract the test residual samples of any new data. Hence, the neural network updated at the 43rd iteration is set as the final model or last network. The comparison between Figure 10a,b regarding the HMC-ODTL-EMSD and the SLS-ODTL-EMSD reveals that the second approach provides smaller errors in the normal condition.



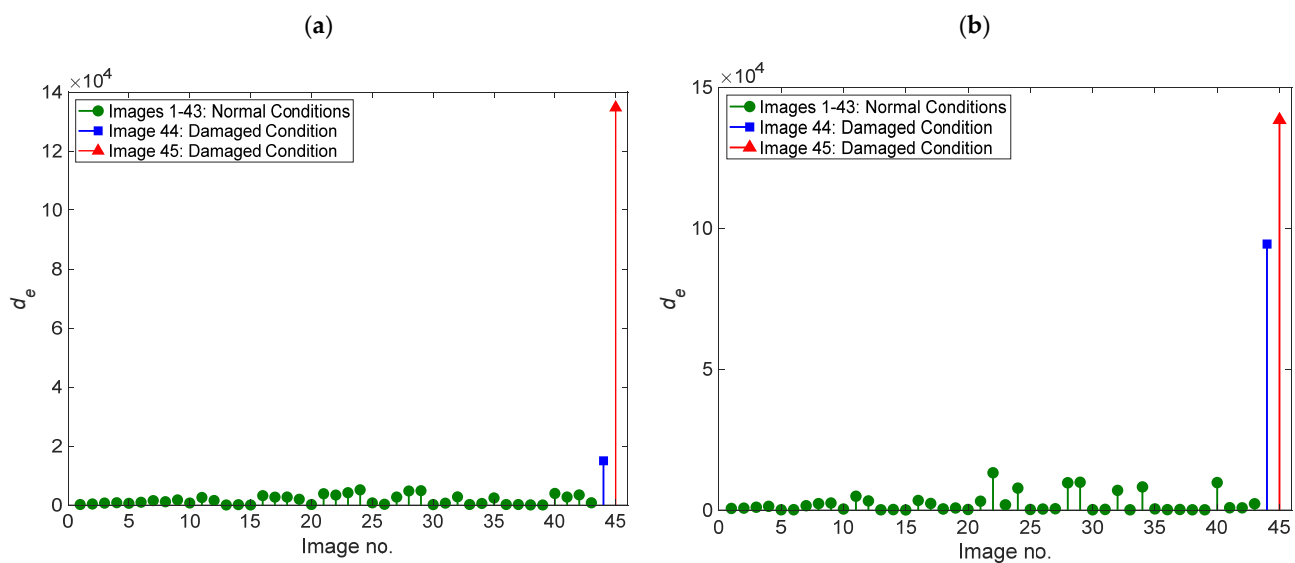
**Figure 10.** Evolution error rates in the iterative process of the ODTL algorithm: (a) HMC-ODTL-EMSD, (b) SLS-ODTL-EMSD.

In addition, Figure 11a,b indicates the optimal numbers of the sequential auto-associative neural networks at each iteration of the HMC-ODTL-EMSD and the SLS-ODTL-EMSD methods, respectively. As such, in Figure 11a, the last network for the first method contains six sequential auto-associative neural networks, while the second method in Figure 11b includes nine networks for constructing the final online model. One of the interesting properties of the proposed online hybrid learning methods is the lack of final decision-making by comparing any novelty score with any threshold limit. In other words, each output of the ODTL-EMSD algorithm can help to make a decision. For example, the evolution of the error rate in Figure 10 demonstrates that the information extracted from Image #44 is dependent on the damaged condition and the bridge may be threatened by dangerous situations, such as partial or global collapse. However, in order to further assess the process of damage detection, the EMSD values on each date or image are shown in Figure 12. As

can be seen, the EMSD quantities on 15 November 2015 regarding Image #44 is larger than the previous images. The comparison between the HMC-ODTL-EMSD and the SLS-ODTL-EMSD methods demonstrates that the latter provides better and larger output and decision-making than the former regarding Image #44. The other important conclusion is that the EMSD values on 26 November 2015 associated with Image #45 suddenly increase so that these are the largest distance quantities among the other values. This clearly indicates the growth of damage or the level of the damage severity. Therefore, one can state that the proposed methods could quantitatively estimate the severity of damage. In Figure 12b, one can observe that the EMSD values of the two last images differ from the other quantities. In contrast, the only EMSD amount of the last image in Figure 12a is different from the EMSD values of the undamaged conditions (i.e., the images 1–43). Hence, it can be concluded that the proposed SLS-ODTL-EMSD method can provide more appropriate results of early damage detection compared to the HMC-ODTL-EMSD technique.



**Figure 11.** Selecting the optimal number of sequential auto-associative neural networks ( $nm_{Opt}$ ) at each iteration before terminating the iterative process of the ODTL algorithm: (a) HMC-ODTL-EMSD, (b) SLS-ODTL-EMSD.



**Figure 12.** Decision-making for early damage detection: (a) HMC-ODTL-EMSD, (b) SLS-ODTL-EMSD.

The other important note in Figure 12 is that the detection of the undamaged conditions before Image #44 means that the proposed methods work well. More precisely, one of the main objectives of SHM is to declare the current state of a structure in terms of being undamaged or being damaged before the occurrence of a catastrophic event, such as collapse. If it is assumed that the current status of the Tadcaster Bridge is unknown, the main goal is to collect a SAR image, extract displacement data, augment small displacement samples, and implement the proposed online learning methods, as shown in Figure 3. Accordingly, the first set of the augmented data from the first image is selected as the baseline data ( $X_1$ ) and the next image (on a different date) is considered as the validation data (e.g.,  $X_2$ ). If the methods declare that this set belongs to the undamaged (normal) condition, it can be concluded that the structure operates normally, and the validation data can be added to the training set and the process continues. In Figure 12, as the proposed methods yielded the accurate decision about the undamaged condition of the bridge before Image #44, one can deduce that these methods are successful in health monitoring of the bridge. However, when the augmented data of Image #44 is used, the methods terminate the first stage (online learning) and arrive at the second stage (online damage detection). Afterward, all measured (augmented) data should be considered as the test data. From Figure 12, one can conclude that the methods accurately detect the undamaged and the damaged states of the bridge before the occurrence of the partial collapse.

Here, it should be justified that although the flooding caused the partial collapse of one part of the Tadcaster Bridge, the scour was the main factor. On the other hand, although the bridge partially collapsed on 29 December 2015; both proposed methods seriously trigger the existence of an adverse variation on 15 November 2015, i.e., regarding Image #44, and a worse change on 26 November 2015, i.e., regarding Image #45. This means that these methods can help authorities and the bridge owners to prevent an adverse phenomenon such as a partial collapse, which can threaten human life and induce considerable economic losses due to rehabilitation or even rebuilding of seriously damaged structures.

#### 4. Conclusions

The main objective of this article was to propose new online hybrid learning methods called HMC-ODTL-EMSD and SLS-ODTL-EMSD for early damage detection and pre-collapse prediction of civil structures. The key novel element of the proposed methods in this article was their online learning and online damage detection procedures. Moreover, new data normalization based on the ODTL algorithm and a new feature classification metric, EMSD, were proposed. The small displacement samples of the Tadcaster Bridge obtained from some satellite images of TerraSAR-X were applied to validate the proposed methods. The main conclusions of this study can be summarized as follows:

- (1) The proposed idea for the augmentation of the small displacement data through the HMC and slice sampling provided better observations of structural responses, structural behavior, and EOVS conditions. Regarding the last item in the graphical evaluation, the HMC sampling better indicated the variations caused by the environmental and/or operational conditions as well as damaged cases. However, the numerical assessment via the MAD measure revealed that both the HMC and slice sampling succeeded in demonstrating the EOVS conditions. Generally, the augmented displacement data outperforms the small data for this issue;
- (2) The second stage of the proposed methods regarding the online data normalization and online damage detection accurately detected the damaged state of the bridge and predicted the pre-collapse condition on 15 November 2015 (i.e., the 14 days before the partial collapse on 29 November 2015). In other words, the online hybrid learning methods correctly alarmed the emergence of damage and the hazard of collapse before its occurrence. This alarm triggered before 26 November 2015 (i.e., the 3 days before the partial collapse), when the outputs of the proposed methods indicated the growth of damage and abnormal changes in the bridge;

- (3) In both online hybrid learning methods, the EMSD value of Image #45 regarding 26 November 2015 was larger than the corresponding value of Image #44 concerning 15 November 2015. This means that the proposed methods could correctly and quantitatively estimate the level of damage severity;
- (4) Although both HMC-ODTL-EMSD and SLS-ODTL-EMSD gave reasonable results of early damage detection and succeeded in alarming the occurrence of damage at the earliest time (i.e., 15 November 2015), the latter outperformed the former due to smaller error variations in the normal conditions (i.e., the images 1–43) and more discriminative EMSD values between the undamaged and damaged conditions.

For further research in this category, it is essential to develop the online learning method with a more rigorous and robust stopping condition for  $\beta$ . As mentioned earlier, the main reason for applying the displacement samples of the Tadcaster Bridge is an appropriate target point distribution on the structure similar to contact-based sensor networks. Hence, new machine learning methods for locating the damaged area of the structure can be proposed. Accordingly, any new method should be able to identify the location of the point “b” as the damaged zone of the structure.

**Author Contributions:** Conceptualization, A.E., C.D.M. and B.B.; methodology, A.E., C.D.M. and B.B.; software, A.E. and B.B.; validation, A.E. and B.B.; formal analysis, A.E., C.D.M. and A.N.A.; investigation, A.E. and C.D.M.; resources, A.E.; data curation, A.E. and B.B.; writing—original draft preparation, A.E. and B.B.; writing—review and editing, A.E., C.D.M., A.N.A. and B.B.; visualization, A.E., C.D.M., A.N.A. and B.B.; supervision, C.D.M. and A.N.A.; project administration, A.E., C.D.M. and A.N.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partially funded by the European Space Agency (ESA) under ESA Contract No. 4000132658/20/NL/MH/ac.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Li, H.-N.; Ren, L.; Jia, Z.-G.; Yi, T.-H.; Li, D.-S. State-of-the-art in structural health monitoring of large and complex civil infrastructures. *J. Civ. Struct. Health Monit.* **2016**, *6*, 3–16. [[CrossRef](#)]
2. Deng, L.; Wang, W.; Yu, Y. State-of-the-Art Review on the Causes and Mechanisms of Bridge Collapse. *J. Perform. Constr. Facil.* **2016**, *30*, 04015005. [[CrossRef](#)]
3. Rizzo, P.; Enshaiean, A. Challenges in Bridge Health Monitoring: A Review. *Sensors* **2021**, *21*, 4336. [[CrossRef](#)] [[PubMed](#)]
4. Wang, M.L.; Lynch, J.P.; Sohn, H. *Sensor Technologies for Civil Infrastructures: Applications in Structural Health Monitoring*; Woodhead Publishing (Elsevier): Cambridge, UK, 2014.
5. Sony, S.; Laventure, S.; Sadhu, A. A literature review of next-generation smart sensing technology in structural health monitoring. *Struct. Contr. Health Monit.* **2019**, *26*, e2321. [[CrossRef](#)]
6. Shen, N.; Chen, L.; Liu, J.; Wang, L.; Tao, T.; Wu, D.; Chen, R. A Review of Global Navigation Satellite System (GNSS)-Based Dynamic Monitoring Technologies for Structural Health Monitoring. *Remote Sens.* **2019**, *11*, 1001. [[CrossRef](#)]
7. Feng, D.; Feng, M.Q. Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection—A review. *Eng. Struct.* **2018**, *156*, 105–117. [[CrossRef](#)]
8. An, D.; Kim, N.H.; Choi, J.-H. Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. *Reliab. Eng. Syst. Saf.* **2015**, *133*, 223–236. [[CrossRef](#)]
9. Malekloo, A.; Ozer, E.; AlHamaydeh, M.; Girolami, M. Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights. *Struct. Health Monit.* **2022**, *21*, 1906–1955. [[CrossRef](#)]
10. Daneshvar, M.H.; Sarmadi, H. Unsupervised learning-based damage assessment of full-scale civil structures under long-term and short-term monitoring. *Eng. Struct.* **2022**, *256*, 114059. [[CrossRef](#)]
11. Entezami, A.; Shariatmadar, H.; De Michele, C. Non-parametric empirical machine learning for short-term and long-term structural health monitoring. *Struct. Health Monit.* **2022**, 14759217211069842. [[CrossRef](#)]
12. Sarmadi, H.; Yuen, K.-V. Structural health monitoring by a novel probabilistic machine learning method based on extreme value theory and mixture quantile modeling. *Mech. Syst. Sig. Process.* **2022**, *173*, 109049. [[CrossRef](#)]
13. Hoi, S.C.H.; Sahoo, D.; Lu, J.; Zhao, P. Online learning: A comprehensive survey. *Neurocomputing* **2021**, *459*, 249–289. [[CrossRef](#)]
14. Shi, H.; Worden, K.; Cross, E.J. A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring. *Mech. Syst. Sig. Process.* **2018**, *103*, 381–397. [[CrossRef](#)]
15. Entezami, A.; Shariatmadar, H.; Mariani, S. Early damage assessment in large-scale structures by innovative statistical pattern recognition methods based on time series modeling and novelty detection. *Adv. Eng. Softw.* **2020**, *150*, 102923. [[CrossRef](#)]

16. Sarmadi, H.; Entezami, A.; Saeedi Razavi, B.; Yuen, K.-V. Ensemble learning-based structural health monitoring by Mahalanobis distance metrics. *Struct. Contr. Health Monit.* **2021**, *28*, e2663. [[CrossRef](#)]
17. Sarmadi, H.; Yuen, K.-V. Early damage detection by an innovative unsupervised learning method based on kernel null space and peak-over-threshold. *Comput. Aided Civ. Inf.* **2021**, *36*, 1150–1167. [[CrossRef](#)]
18. Sarmadi, H.; Entezami, A.; Salar, M.; De Michele, C. Bridge health monitoring in environmental variability by new clustering and threshold estimation methods. *J. Civ. Struct. Health Monit.* **2021**, *11*, 629–644. [[CrossRef](#)]
19. Entezami, A.; Sarmadi, H.; Salar, M.; De Michele, C.; Nadir Arslan, A. A novel data-driven method for structural health monitoring under ambient vibration and high dimensional features by robust multidimensional scaling. *Struct. Health Monit.* **2021**, 1475921720973953. [[CrossRef](#)]
20. Sarmadi, H.; Entezami, A.; Behkamal, B.; De Michele, C. Partially online damage detection using long-term modal data under severe environmental effects by unsupervised feature selection and local metric learning. *J. Civ. Struct. Health Monit.* **2022**, 1–24. [[CrossRef](#)]
21. Krishnan, M.; Bhowmik, B.; Hazra, B.; Pakrashi, V. Real time damage detection using recursive principal components and time varying auto-regressive modeling. *Mech. Syst. Sig. Process.* **2018**, *101*, 549–574. [[CrossRef](#)]
22. Jin, S.-S.; Cho, S.; Jung, H.-J. Adaptive reference updating for vibration-based structural health monitoring under varying environmental conditions. *Comput. Struct.* **2015**, *158*, 211–224. [[CrossRef](#)]
23. Jin, S.-S.; Jung, H.-J. Vibration-based damage detection using online learning algorithm for output-only structural health monitoring. *Struct. Health Monit.* **2018**, *17*, 727–746. [[CrossRef](#)]
24. Nguyen, L.H.; Goulet, J.-A. Real-time anomaly detection with Bayesian dynamic linear models. *Struct. Contr. Health Monit.* **2019**, *26*, e2404. [[CrossRef](#)]
25. Sarmadi, H.; Karamodin, A. A novel anomaly detection method based on adaptive Mahalanobis-squared distance and one-class kNN rule for structural health monitoring under environmental effects. *Mech. Syst. Sig. Process.* **2020**, *140*, 106495. [[CrossRef](#)]
26. Liu, Z.; Cao, Y.; Wang, Y.; Wang, W. Computer vision-based concrete crack detection using U-net fully convolutional networks. *Autom. Constr.* **2019**, *104*, 129–139. [[CrossRef](#)]
27. Zhang, Y.; Yuen, K.-V. Crack detection using fusion features-based broad learning system and image processing. *Comput. Aided Civ. Inf.* **2021**, *36*, 1568–1584. [[CrossRef](#)]
28. Yao, Y.; Yang, Y.; Wang, Y.; Zhao, X. Artificial intelligence-based hull structural plate corrosion damage detection and recognition using convolutional neural network. *Appl. Ocean. Res.* **2019**, *90*, 101823. [[CrossRef](#)]
29. Kong, X.; Li, J. Image Registration-Based Bolt Loosening Detection of Steel Joints. *Sensors* **2018**, *18*, 1000. [[CrossRef](#)]
30. Biondi, F.; Addabbo, P.; Ullo, S.L.; Clemente, C.; Orlando, D. Perspectives on the Structural Health Monitoring of Bridges by Synthetic Aperture Radar. *Remote Sens.* **2020**, *12*, 3852. [[CrossRef](#)]
31. Bakon, M.; Czikhardt, R.; Papco, J.; Barlak, J.; Rovnak, M.; Adamisin, P.; Perissin, D. remotIO: A Sentinel-1 Multi-Temporal InSAR Infrastructure Monitoring Service with Automatic Updates and Data Mining Capabilities. *Remote Sens.* **2020**, *12*, 1892. [[CrossRef](#)]
32. Amoroso, N.; Cilli, R.; Bellantuono, L.; Massimi, V.; Monaco, A.; Nitti, D.O.; Nutricato, R.; Samarelli, S.; Taggio, N.; Tangaro, S.; et al. PSI Clustering for the Assessment of Underground Infrastructure Deterioration. *Remote Sens.* **2020**, *12*, 3681. [[CrossRef](#)]
33. Macchiarulo, V.; Milillo, P.; Blenkinsopp, C.; Giardina, G. Monitoring deformations of infrastructure networks: A fully automated GIS integration and analysis of InSAR time-series. *Struct. Health Monit.* **2022**, *21*, 14759217211045912. [[CrossRef](#)]
34. Selvakumaran, S.; Plank, S.; Geiß, C.; Rossi, C.; Middleton, C. Remote monitoring to predict bridge scour failure using Interferometric Synthetic Aperture Radar (InSAR) stacking techniques. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 463–470. [[CrossRef](#)]
35. Zhu, M.; Wan, X.; Fei, B.; Qiao, Z.; Ge, C.; Minati, F.; Vecchioli, F.; Li, J.; Costantini, M. Detection of Building and Infrastructure Instabilities by Automatic Spatiotemporal Analysis of Satellite SAR Interferometry Measurements. *Remote Sens.* **2018**, *10*, 1816. [[CrossRef](#)]
36. Qin, X.; Liao, M.; Yang, M.; Zhang, L. Monitoring structure health of urban bridges with advanced multi-temporal InSAR analysis. *Ann. GIS* **2017**, *23*, 293–302. [[CrossRef](#)]
37. Qin, X.; Zhang, L.; Yang, M.; Luo, H.; Liao, M.; Ding, X. Mapping surface deformation and thermal dilation of arch bridges by structure-driven multi-temporal DInSAR analysis. *Remote Sens. Environ.* **2018**, *216*, 71–90. [[CrossRef](#)]
38. Milillo, P.; Perissin, D.; Salzer, J.T.; Lundgren, P.; Lacava, G.; Milillo, G.; Serio, C. Monitoring dam structural health from space: Insights from novel InSAR techniques and multi-parametric modeling applied to the Pertusillo dam Basilicata, Italy. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 221–229. [[CrossRef](#)]
39. Huang, Q.; Crosetto, M.; Monserrat, O.; Crippa, B. Displacement monitoring and modelling of a high-speed railway bridge using C-band Sentinel-1 data. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 204–211. [[CrossRef](#)]
40. Milillo, P.; Giardina, G.; Perissin, D.; Milillo, G.; Coletta, A.; Terranova, C. Pre-collapse space geodetic observations of critical infrastructure: The Morandi Bridge, Genoa, Italy. *Remote Sens.* **2019**, *11*, 1403. [[CrossRef](#)]
41. Hashemi, F.; Naderi, M.; Jamalizadeh, A.; Bekker, A. A flexible factor analysis based on the class of mean-mixture of normal distributions. *Comput. Stat. Data Anal.* **2021**, *157*, 107162. [[CrossRef](#)]
42. Van Ravenzwaaij, D.; Cassey, P.; Brown, S.D. A simple introduction to Markov Chain Monte-Carlo sampling. *Psychon. Bull. Rev.* **2018**, *25*, 143–154. [[CrossRef](#)] [[PubMed](#)]
43. Entezami, A.; Sarmadi, H.; De Michele, C. Probabilistic damage localization by empirical data analysis and symmetric information measure. *Measurement* **2022**, *198*, 111359. [[CrossRef](#)]

44. Neal, R.M. MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo*; CRC Press: Boca Raton, FL, USA, 2011.
45. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **1992**, *7*, 457–472. [[CrossRef](#)]
46. Neal, R.M. Slice sampling. *Ann. Stat.* **2003**, *31*, 705–767. [[CrossRef](#)]
47. Kramer, M.A. Autoassociative neural networks. *Comput. Chem. Eng.* **1992**, *16*, 313–328. [[CrossRef](#)]
48. Sarmadi, H. Investigation of machine learning methods for structural safety assessment under variability in data: Comparative studies and new approaches. *J. Perform. Constr. Facil.* **2021**, *35*, 04021090. [[CrossRef](#)]
49. Entezami, A.; Sarmadi, H.; Behkamal, B.; Mariani, S. Health Monitoring of Large-Scale Civil Structures: An Approach Based on Data Partitioning and Classical Multidimensional Scaling. *Sensors* **2021**, *21*, 1646. [[CrossRef](#)]
50. Charte, D.; Charte, F.; García, S.; del Jesus, M.J.; Herrera, F. A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines. *Inf. Fusion* **2018**, *44*, 78–96. [[CrossRef](#)]
51. Yang, Q.; Zhang, Y.; Dai, W.; Pan, S.J. *Transfer Learning*; Cambridge University Press: Cambridge, UK, 2020.
52. Figueiredo, E.; Park, G.; Farrar, C.R.; Worden, K.; Figueiras, J. Machine learning algorithms for damage detection under operational and environmental variability. *Struct. Health Monit.* **2011**, *10*, 559–572. [[CrossRef](#)]