Chapter 12
From Short-Term Monitoring to Long-Term Insight: VAE Driven Synthetic Data for SHM Applications

M. Conti, G. Battista, S. Pavoni, and M. Vanali

Abstract Structural Health Monitoring (SHM) is essential for ensuring the safety and durability of structures, buildings, and industrial components. In recent years, data-driven SHM techniques have gained increasing attention due to their versatility for different structures and their ability to detect damage and assess structural conditions. However, their main limitations lies in the need for large dataset to effectively train the algorithms. Acquiring such data requires continuous monitoring over extended periods (sometimes years) making the process costly, time-consuming, and often impractical.

To overcome these challenges, recent studies have explored synthetic data generation to supplement real-world datasets. Among the most promising approaches is the use of Variational Autoencoders (VAEs), a class of generative models capable of learning compact representations of input data and producing realistic new samples based on learned distributions. VAEs are probabilistic generative models that combine encoder-decoder networks to learn the characteristics of input data and generate novel coherent data samples.

In this study, VAEs are employed to generate synthetic structural response parameters using only a limited set of real-world measurements from a short monitoring period. The benchmark is a laboratory truss girder subjected to environmental variations (temperature drift and thermal gradient evolution). The goal is the augmentation of the available data and the improvement of effective training of SHM algorithms, even when long-term monitoring data is unavailable.

Keywords Structural health monitoring · Variational autoEncoders · Machine learning · Data augmentation

Introduction

Ensuring the safety and durability of civil and mechanical structures is a major challenge in modern engineering. Structural Health Monitoring (SHM) has emerged as a key discipline to address this challenge, providing diverse methods to evaluate the structure's condition. These strategies involve the continuous observation of the structural characteristics and the extraction of damage-sensitive features used for the evaluation of the system's state [1]. In recent years, strategies based on vibration measurements have gained a central role, due to their easy implementation in sensor installation and their non-destructive nature [2]. Several studies have demonstrated that modal parameters are found to be particularly effective in the detection of the condition variations of the structures investigated [3]. However, these strategies present a critical drawback, since the employed damage-sensitive features can be strongly influenced by environmental and operational factors too [4]. For this reason, especially in the data-driven methods, the collection of a large number of observations is necessary to both remove the external impact [5] and train machine learning algorithms [6]. In fact, data-driven algorithms usually require the collection of very large datasets to capture all variations caused by daily and seasonal environmental variations. In many cases, acquiring such datasets requires continuous monitoring over extended periods (sometimes spanning years), making the process costly, time-consuming, and often impractical [7].

To address these challenges, recent research has progressively focused on developing synthetic data generation strategies in order to supplement existing real-world datasets. These generative models (GMs) typically learn the distribution of a specific given dataset and generate new data coherent with the original ones [8]. Common architectures of GMs employed in this field are the Generative Adversarial Networks (GANs) [9], the Diffusion Models (DMs) [10], and the AutoEncoders

M. Conti · G. Battista · S. Pavoni · M. Vanali
Università di Parma – Parco Area Delle Scienze,181/A – 43124 Parma
e-mail: matteo.conti@unipr.it; gianmarco.battista@unipr.it; stefano.pavoni@unipr.it; marcello.vanali@unipr.it

Kristin B. Zimmerman, PhD (eds.), *Structural Health Monitoring, Metamaterials, Metastructures and Additive Manufacturing and Dynamics of Biomedical and Bioinspired Systems, 2026, Vol. 10*, Conference Proceedings
of the Society for Experimental Mechanics Series, https://doi.org/10.1007/978-8-743-81449-8_12

(AEs) [11]. Among the three techniques mentioned, AEs provide a simpler and more stable training process through direct reconstruction loss, avoiding the instability of adversarial training. Additionally, their feed-forward decoders allow for rapid data generation [12]. Particularly useful in this context are the Variational AutoEncoders (VAEs), which, unlike the classic AEs, map the input image in a probabilistic distribution (usually a Gaussian) instead of a single point in the latent space.

In this work, Convolutional Neural Network (CNN) [13] based VAEs are employed to generate new data to implement a pre-existing dataset collected to develop novel SHM strategies. The research aims to increase the observations of the modal parameters affected by environmental influence, used to improve the characterisation of the ambient variability in the baseline of an already employed SHM strategy. The vibration measurements are collected from a laboratory truss girder, realised ad hoc in the noise and vibration laboratory at the University of Parma, for SHM purposes [14]. As the structure is subjected to a white-noise excitation, the generation of vibration measurements directly is impractical, because VAEs are not able to detect patterns in such a case. Thus, the data generated are the power and cross-spectra evaluated between the acceleration measurements, which present a more consistent and perceptible pattern. In the following section, a brief introduction to the benchmark is proposed.

Benchmark and Experimental Setup

The benchmark structure is a laboratory Pratt truss girder. It is a 2D truss anchored to the floor with rigid supports and Teflon roller connections, spanning 4.2 m in length and 0.8 m in height. The truss consists of aluminium diagonal and vertical rods, and two horizontal steel H-beams. Stability is preserved by seven 15 kg masses at the base and four struts with roller cylinders, allowing only vertical movement.

Since ambient excitation is insufficient, the structure is vertically forced using a PCB 2007E electrodynamic shaker with a band-limited white-noise (10-250 Hz). The response is measured with six MEMS accelerometers (TE4030, ±2 g), while four thermocouples monitor ambient conditions. Tests are repeated every 30 minutes to track changes in modal and environmental parameters. In Figure 1 the experimental setup layout is presented.

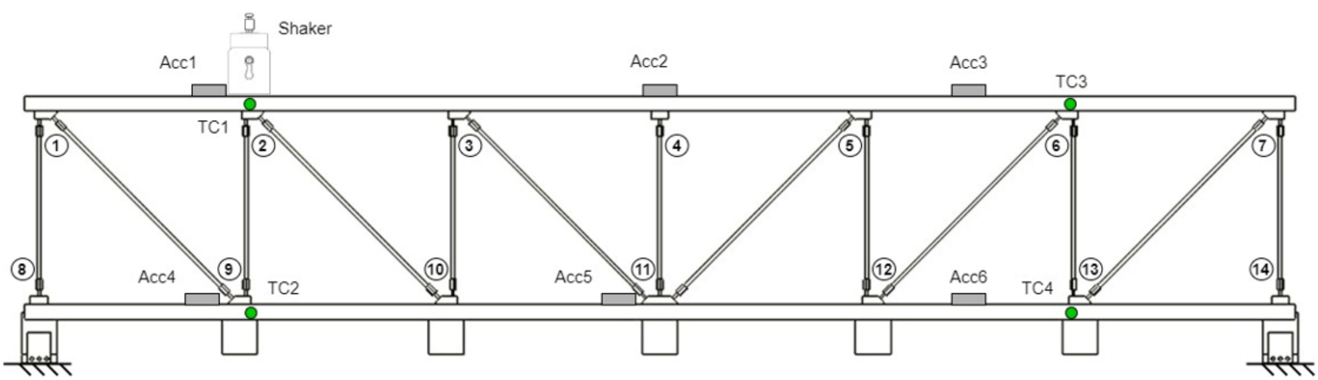

Fig. 1 Truss sketch with node numeration and data acquisition system layout with shaker, accelerometers (grey rectangles) and thermocouples (green circles) location

Damage-Sensitive Features Extraction

The damage-sensitive parameters employed in already improved SHM strategies are the modal parameters, in particular, natural frequency and mode shapes. However, in this work, only the enhancement of the natural frequencies dataset is investigated. Therefore, natural frequencies are extracted using the power and cross-spectra by exploiting an enhanced Frequency Domain Decomposition (FDD) method. In more depth, the response power spectra are processed using Singular Value Decomposition (SVD) as follows

$$G_{xx}\left(\omega\right) = U\left(\omega\right) S(\omega) U(\omega)^{H} \tag{1}$$

where $S(\omega)$ is a diagonal matrix containing the singular values, which are employed to identify the resonances. Natural frequencies are extracted by applying the Peak Picking (PP) method to $S(\omega)$ [15]. To reduce the effects of the frequency

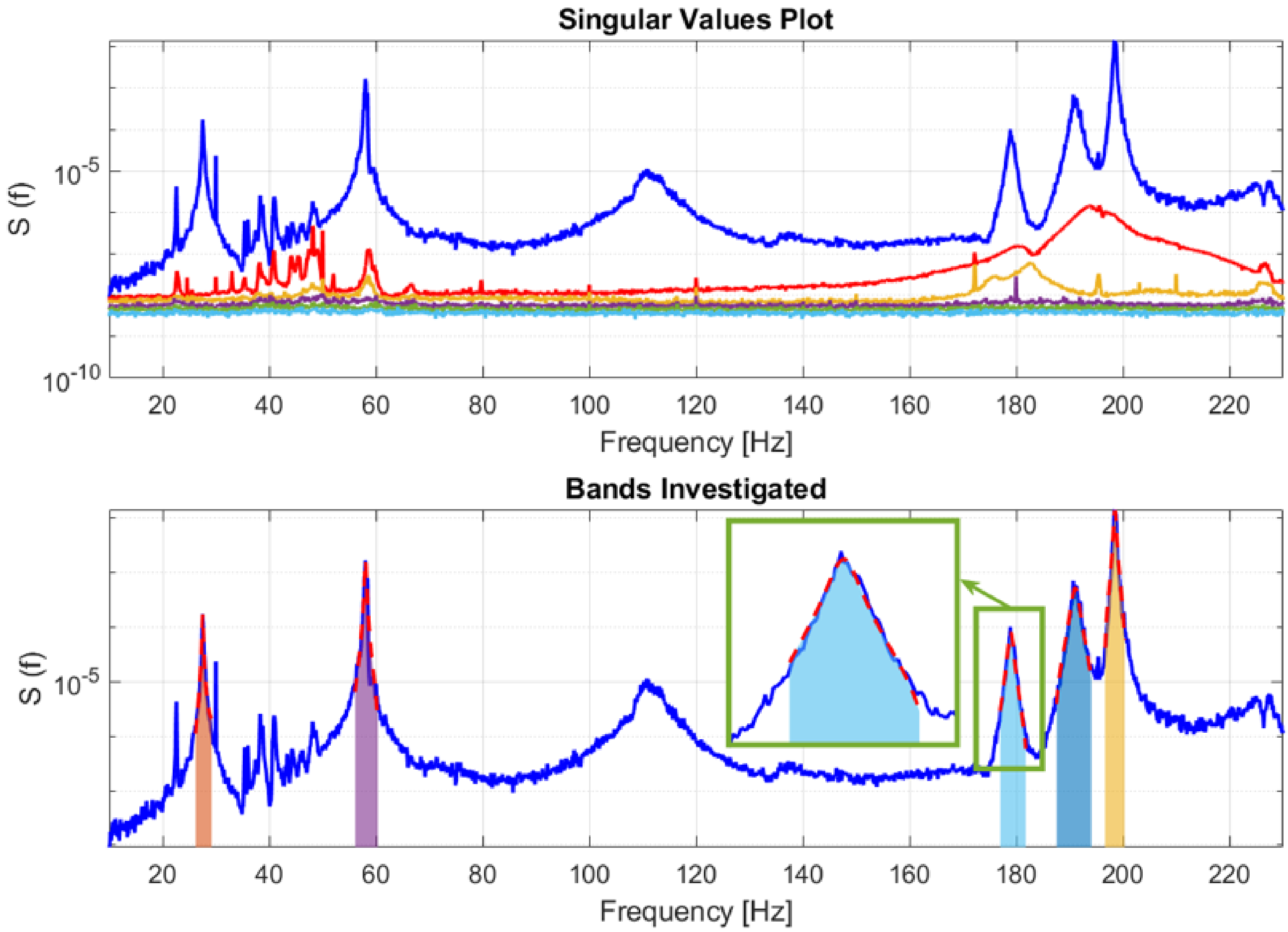

Fig. 2 Enhanced FDD performed on the truss girder data with bands analysed and reconstruction of resonance peaks

resolution, the singular value is minimised using a Lorentzian Peak-Shape function [16]. An example of the procedure is shown in Figure 2.

The natural frequencies investigated are listed in Table 1.

Table 1 Natural frequencies of the truss girder

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Natural Frequency [Hz]	27	58	179	191	198

Natural frequencies collected while the structure was in a healthy condition and used as the baseline set, encompass the period from 6th April to 22nd July and consist of 3014 observations. In Figure 3, the baseline relationship between natural frequencies is shown.

These observations are employed to train a generative model for enhancing the baseline set. Details of that process are introduced in the following section.

Generative Model and Generation of New Samples

Variational Autoencoders (VAEs) Introduction

VAEs are generative models that learn how to represent the features of a dataset in a structured latent space. In contrast to classic AEs, which compress inputs into fixed vectors, VAEs encode data into probability distributions (usually Gaussian).

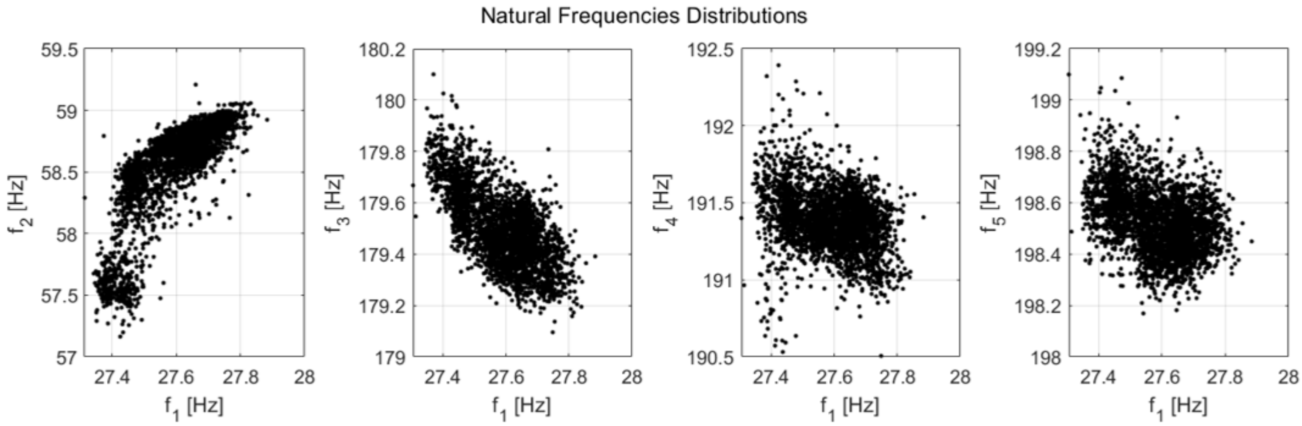

Fig. 3 Relationship between natural frequencies collected during the baseline period

Indeed, VAEs consist of an encoder that maps inputs into a probability distribution using a latent space, and a decoder that reconstructs the initial data. VAE's generic architecture is shown in Figure 4.

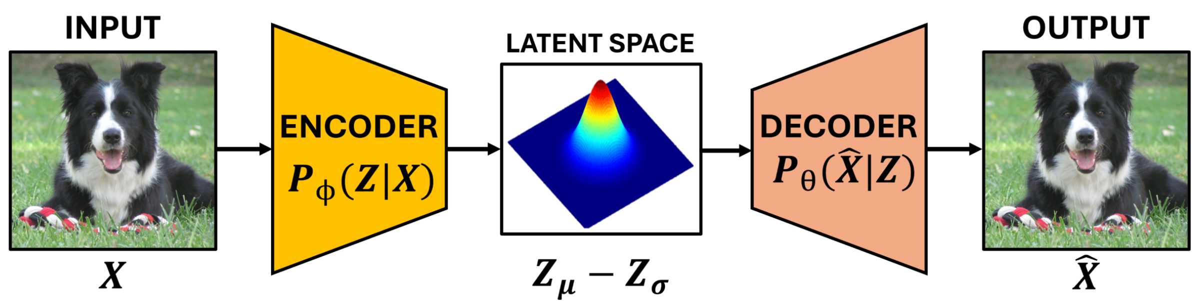

Fig. 4 VAEs architecture

Both the encoder and decoder can be formed by either Fully Connected Neural Networks (FCNN) or Convolutional Neural Networks (CNN). This enables the network to learn specific and characteristic patterns in the images used as input data. Particularly effective are the CNNs, which employ filters (called Kernels) to recognise image characteristics.

Through this architecture, it is possible to train the network with target images and then generate new coherent samples by using the decoder structure fed with a random Gaussian distribution.

VAEs Input Data

VAEs are able to detect determined patterns in a set of images. For this reason, in the present work, acceleration measurements are not used due to their random nature. Thus, the functions employed for data generation are the power and cross-spectra for two reasons:

- spectra converted into images present recognisable specific patterns.
- spectra are used for modal parameters extraction.

However, some consideration must be given regarding the power and cross-spectra employed as input:

- Spectra need to be processed in terms of real and imaginary parts, as the magnitude and phase differ by several orders of magnitude.
- Spectra must be normalised to allow a conversion into images. In this work, the normalisation was performed taking into account the maximum and minimum values of all spectra.
- Spectra need to be raised to a power (in this study, -5) to enhance characteristic patterns, making them more identifiable by the network (see Figure 6).

- Spectra have to be rearranged to obtain a squared image. This is made by neglecting the complementary functions, eliminating the imaginary part of the power spectra (zero), and dividing the functions into strips. With six accelerometers, 36 functions are used to construct the image, as shown in Figure 5, where the power and cross-spectra exploited are deepened.

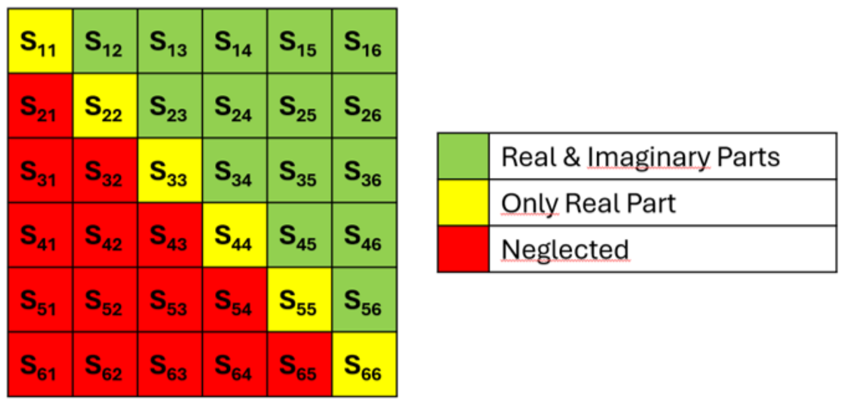

Fig. 5 Power and cross-spectra employed

Since the image generated by the process is square (a better format for CNN), it has been arranged to have a resolution of 288 x 288 pixels, covering the final frequency range of 10-240.3 Hz.

In Figure 6, the image conversion process is shown in more detail.

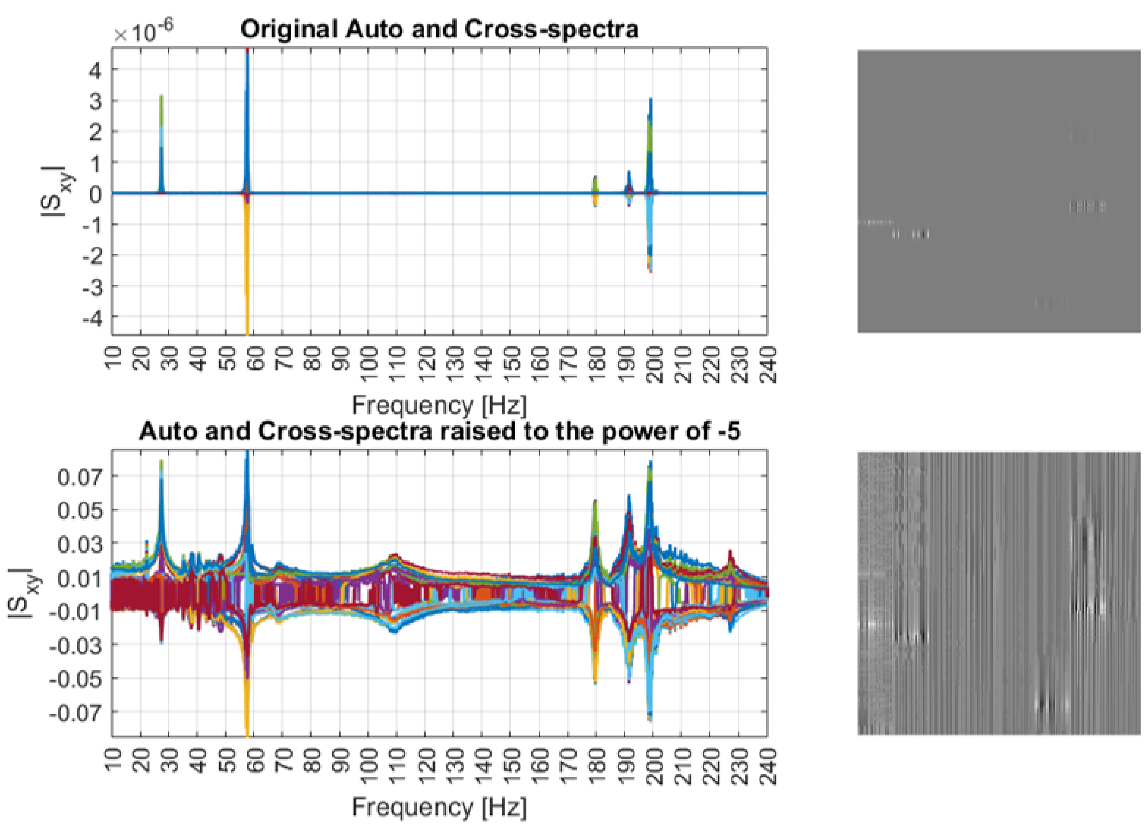

Fig. 6 Power and cross-spectra processed and converted into images. Above the original spectra, below the processed spectra

VAEs Architecture

Several layouts were tested to obtain a network with generated data as coherent as possible with the training test. The network's performance was, first of all, evaluated through the ELBO Loss function (Evidence Lower Bound), which indicates simultaneously the good reconstruction of the input data and the preservation of a well-structured latent space. The following formula represents the ELBO function

$$\mathcal{L}_{ELBO}\left(x\right) = \mathbb{E}_{q(z|x)}\left[\log p\left(x \mid z\right)\right] - D_{KL}(q(z|x)||p(z)) \tag{2}$$

where the first term $\mathbb{E}_{q(z|x)}\left[\log p\left(x \mid z\right)\right]$ measures the reconstruction fidelity, while the second one $D_{KL}(q(z|x)|\,|p\left(z\right))$ is the Kullback-Leibler divergence and serves to approximate posterior $q(z|x)$ to match the prior $p(z)$, typically a Gaussian [17].

Finally, the networks were evaluated based on damage-related features by analysing the generated natural frequency distributions and assessing how closely they matched the training set.

The final VAE architecture consists of:

- 4 convolutional layers with 6 kernels (3x3) with 2-element stride
- 1 fully connected layer with 1944 elements
- a latent space dimension of 50
- a latent vector (82944x1) converted into a feature map (18x18x256)
- 4 deconvolutional layers with 6 kernels (3x3) with 2-element stride.

The final architecture of the VAE employed is shown in Figure 7.

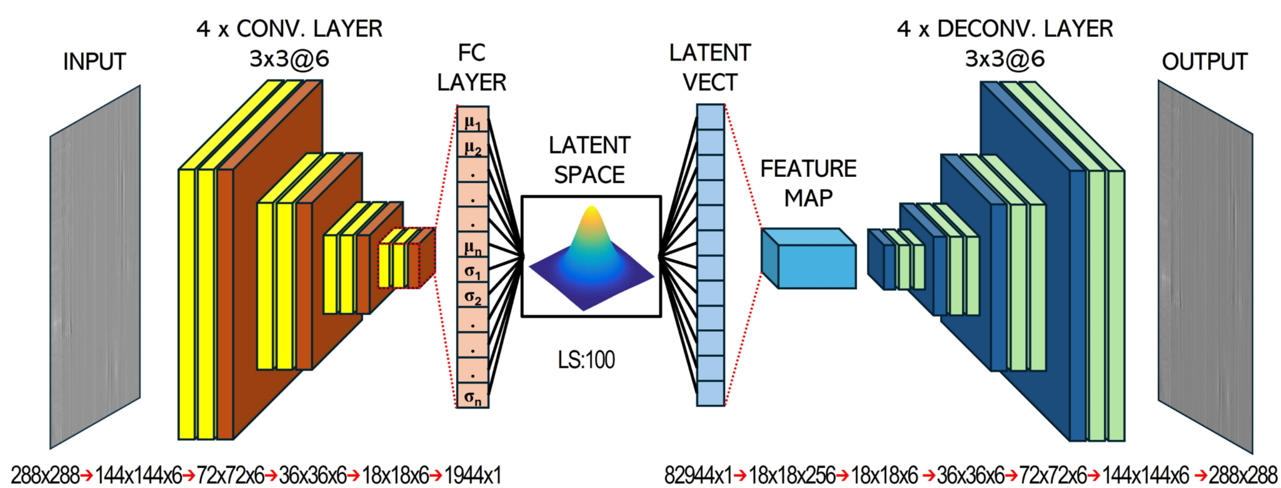

Fig. 7 VAE architecture with layers and dimensions below

Results

The entire process of conversion between the images generated and the spectra functions (and vice versa) has been deepened in order to assess the reliability of the procedure.

An example of spectral functions and the relative image generated with a trained VAE is shown in Figure 8.

First preliminary approach

An initial attempt was carried out using either the entire baseline set or portions of it (30, 50 and 70 %) to train the VAE network. The reconstruction of the spectral functions appeared to be consistent with the experimental ones. However, by

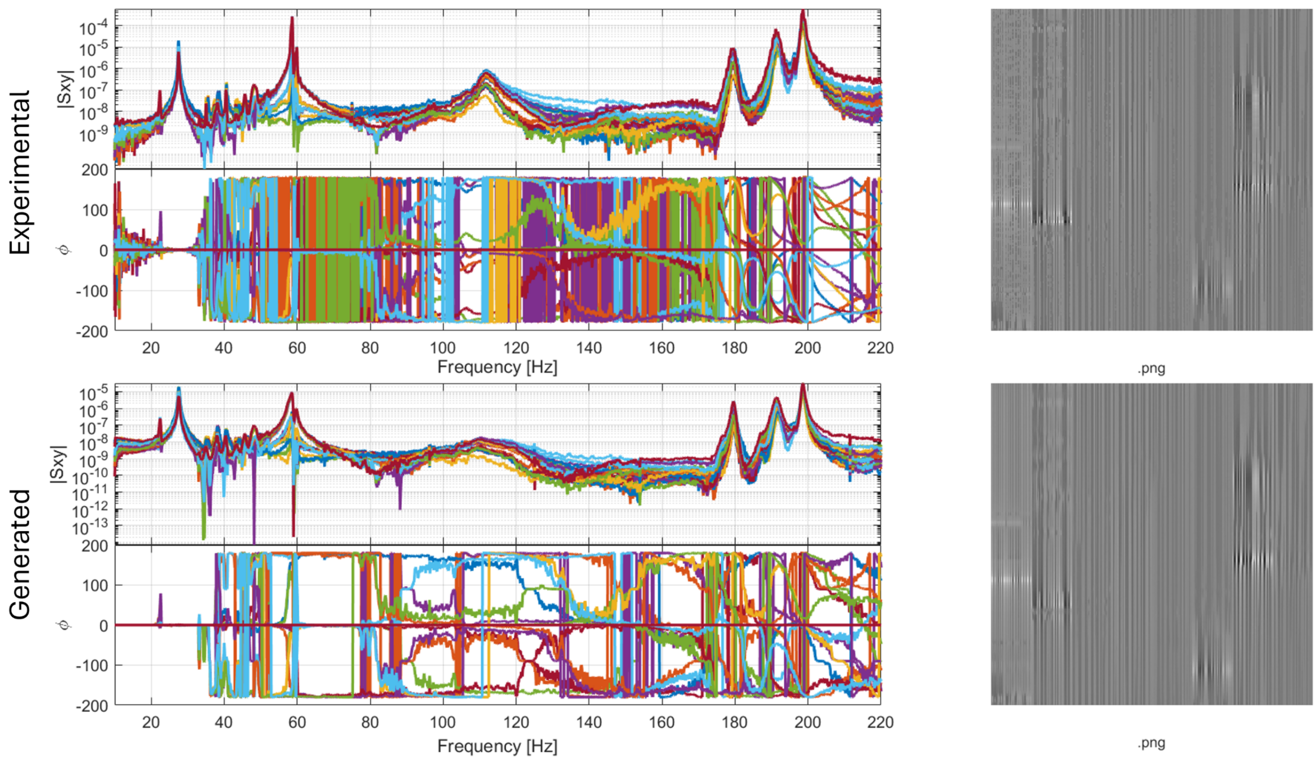

Fig. 8 Spectra from an experimental test (above) and generated with VAE (below) with the corresponding images on the right

observing the distributions of the natural frequencies extracted from the generated data, the variability of the baseline observations seemed not to be properly represented. As shown in Figure 9, natural frequencies from generated data likely tend to capture the average dynamic behaviour of the structure.

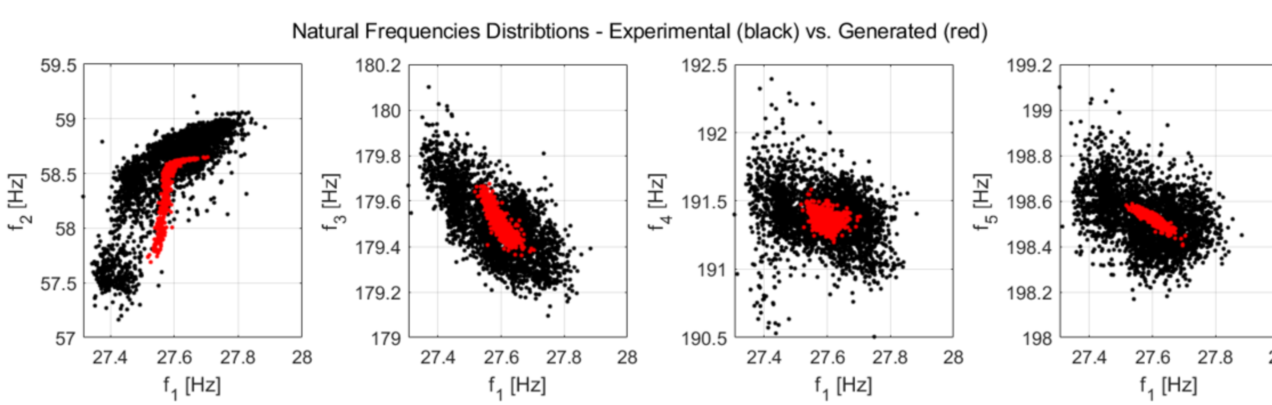

Fig. 9 Comparison between experimental and generated natural frequencies distributions

Such a result cannot be considered reliable, since the variability of the structural dynamic behaviour does not seem to be adequately represented by the generated data.

Final approach

A different approach was considered to obtain more reliable results. In order to obtain newly generated observations coherent with the experimental variability, the baseline set was split into subsets according to a weekly grouping. Then, a series of individual VAEs was trained using the data as input on a week-by-week basis. Unfortunately, the number of observations for each group was not enough to train a VAE in a proper way, which led to underfitting. Therefore, to address this problem, a pilot VAE was pre-trained using the entire baseline set. Subsequently, the network parameters were optimised by retraining the principal VAE with the weekly observations. The final procedure is shown in Figure 10.

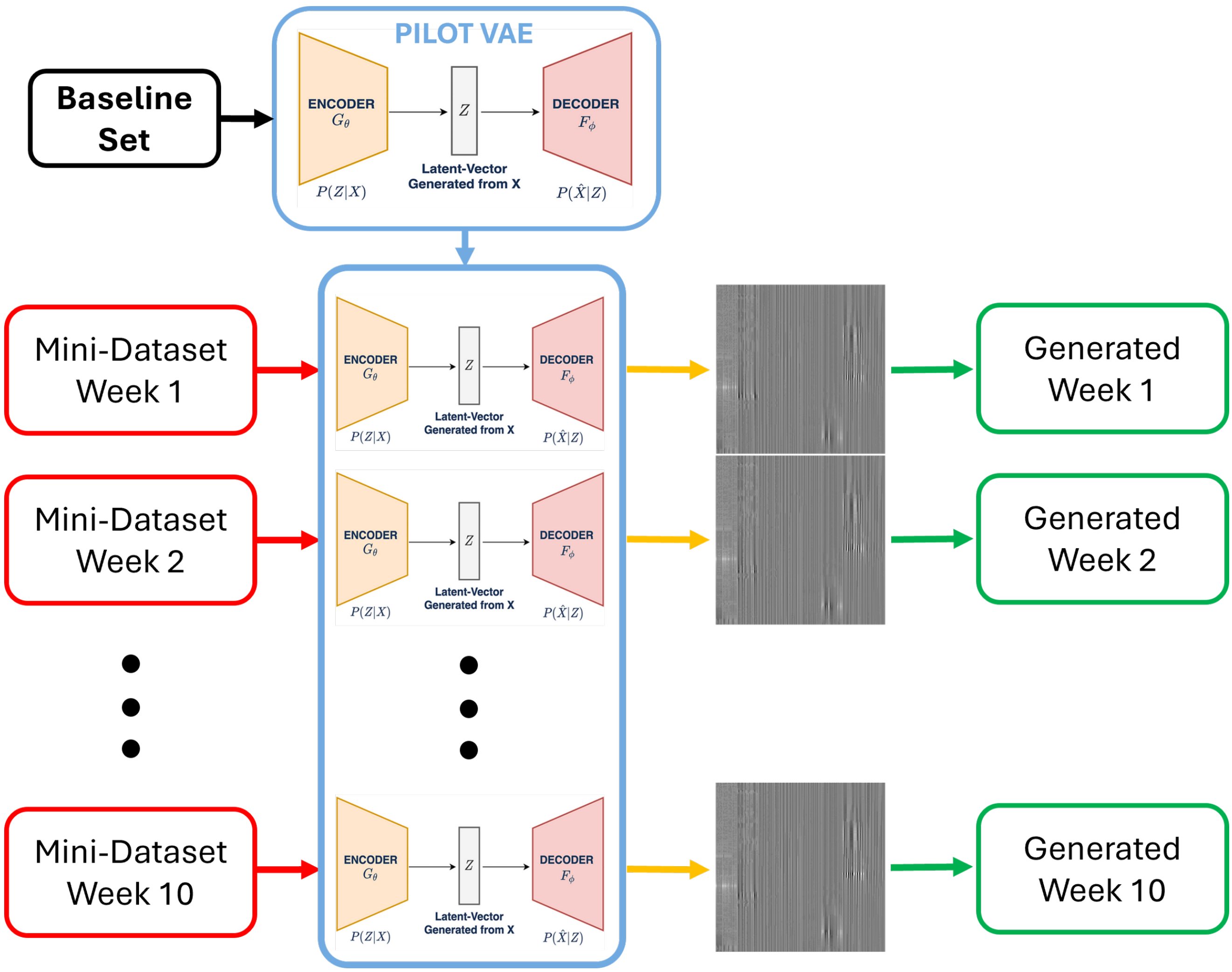

Fig. 10 Definitive procedure for new data generation

This approach led to a better representation of the variability of the dynamic behaviour of the structure, as shown in Figure 11.

A total of 10000 new observations were generated through the trained decoders. To avoid possible generation failures, the outcomes were filtered using a probabilistic strategy. Only a few dozen observations were removed.

Damage detection enhancement

To illustrate the potential benefits of the proposed strategy, a damage detection strategy based on the Principal Component Analysis (PCA) and Mahalanobis Squared Distance (MSD) exploiting the natural frequencies is evaluated using both original and enhanced baseline sets. This method is further deepened in [18]. Results in Figure 12 demonstrate the improvement in the damage detection (1 kg placed in several points of the structure) using the generated data for baseline augmentation.

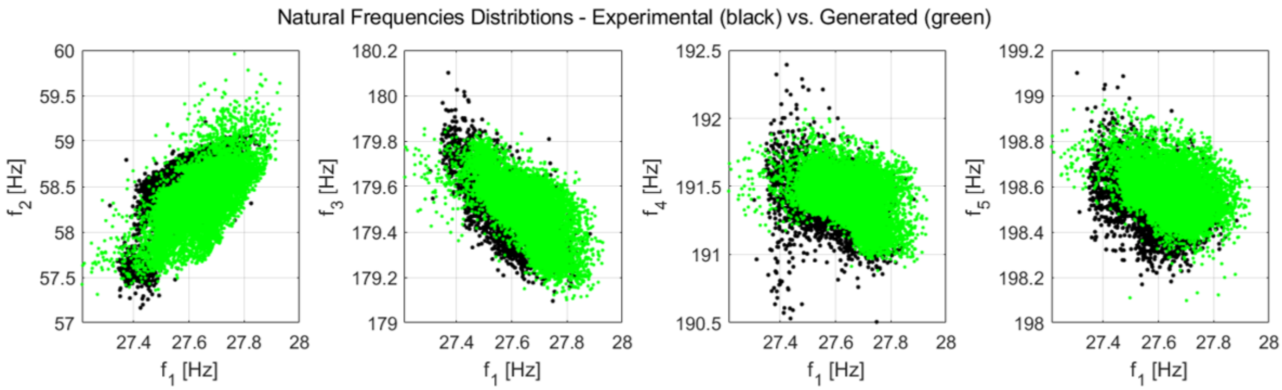

Fig. 11 Comparison between experimental and generated natural frequencies distributions with the final approach

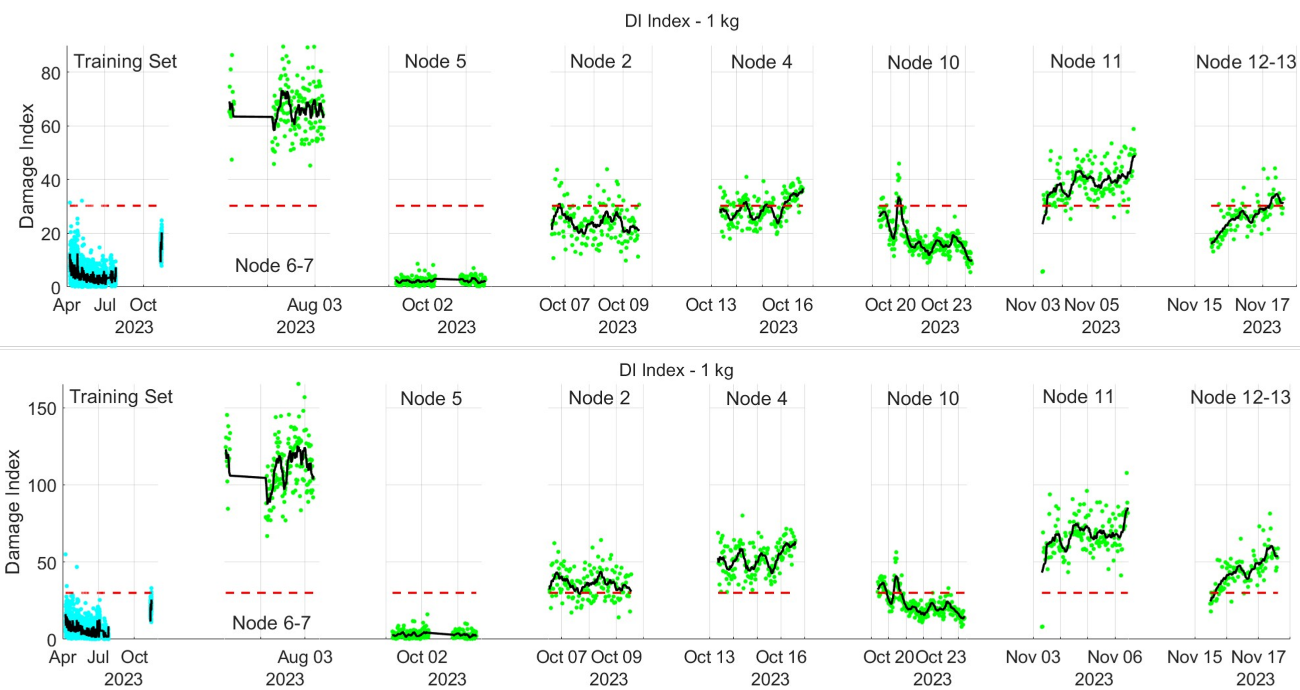

Fig. 12 Damage detection strategy performed with the original baseline (above) and the enhanced baseline (below)

Conclusion

This work demonstrates the improvement of a damage detection technique by enhancing a limited baseline dataset using generative models. The data used for the generative network consisted of the experimental power and cross-spectra from a laboratory truss girder subjected to environmental variations. The damage-sensitive features adopted in this work were the natural frequencies extracted from the spectral functions.

A series of VAEs was trained by weekly observations to properly capture the variability of the baseline set due to different ambient conditions. Subsequently, the trained decoders were employed to generate new spectra and, consequently, new natural frequencies consistent with the experimental ones. These synthetic observations were employed to augment the limited baseline set, with the aim of improving an already developed damage detection technique.

The results demonstrate a benefit in the detection of induced damage to the laboratory structure. Possible future developments may involve the employment of different modal parameters, such as mode shapes, to explore alternative damage detection strategies.

References

1. C. R. Farrar and K. Worden, *Structural Health Monitoring*. Wiley, 2012. doi: 10.1002/9781118443118.
2. Wei Fan and Pizhong Qiao, "Vibration-based Damage Identification Methods: A Review and Comparative Study," *Struct Health Monit*, vol. 10, no. 1, pp. 83–111, Jan. 2011, doi: 10.1177/1475921710365419.
3. C. Zhang, S.-X. Lai, and H.-P. Wang, "Structural Modal Parameter Recognition and Related Damage Identification Methods under Environmental Excitations: A Review," *Structural Durability & Health Monitoring*, vol. 19, no. 1, pp. 25–54, 2025, doi: 10.32604/sdhm.2024.053662.
4. C. Rainieri and F. Magalhaes, "Challenging aspects in removing the influence of environmental factors on modal parameter estimates," *Procedia Eng*, vol. 199, pp. 2244–2249, 2017, doi: 10.1016/j.proeng.2017.09.210.
5. K. Lakshmi and A. Rama Mohan Rao, "Handling Environmental and Operational Variations in Structural Health Monitoring," 2019, pp. 1111–1122. doi: 10.1007/978-981-13-0362-3_88.
6. K. Worden and G. Manson, "The application of machine learning to structural health monitoring," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 515–537, Feb. 2007, doi: 10.1098/rsta.2006.1938.
7. A. Entezami, H. Sarmadi, B. Behkamal, and S. Mariani, "Big Data Analytics and Structural Health Monitoring: A Statistical Pattern Recognition-Based Approach," *Sensors*, vol. 20, no. 8, p. 2328, Apr. 2020, doi: 10.3390/s20082328.
8. G.-H. Gwon and H.-J. Jung, "A survey of generative models for image-based structural health monitoring in civil infrastructure," *Journal of Infrastructure Intelligence and Resilience*, vol. 4, no. 2, p. 100138, Jun. 2025, doi: 10.1016/j.iintel.2025.100138.
9. I. Goodfellow *et al.*, "Generative adversarial networks," *Commun ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.
10. H. Zhang, N. Chen, M. Li, and S. Mao, "The Crack Diffusion Model: An Innovative Diffusion-Based Method for Pavement Crack Detection," *Remote Sens (Basel)*, vol. 16, no. 6, p. 986, Mar. 2024, doi: 10.3390/rs16060986.
11. X. Li *et al.*, "Mechanics-informed autoencoder enables automated detection and localization of unforeseen structural damage," *Nat Commun*, vol. 15, no. 1, p. 9229, Oct. 2024, doi: 10.1038/s41467-024-52501-4.
12. H. Wang, "Comparative Analysis of GANs and Diffusion Models in Image Generation," *Highlights in Science, Engineering and Technology*, vol. 120, pp. 59–66, Dec. 2024, doi: 10.54097/9gba9v27.
13. L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
14. S. Pavoni, "Development of a damage detection technique with a strong immunity to environmental influence implemented on a laboratory truss girder subjected to ambient variations," University of Parma, 2024.
15. R. Brincker and C. E. Ventura, *Introduction to Operational Modal Analysis*. Wiley, 2015. doi: 10.1002/9781118535141.
16. G. Cowan, *Statistical Data Analysis*. Oxford University PressOxford, 1998. doi: 10.1093/oso/9780198501565.001.0001.
17. D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational Inference: A Review for Statisticians," *J Am Stat Assoc*, vol. 112, no. 518, pp. 859–877, Apr. 2017, doi: 10.1080/01621459.2017.1285773.
18. M. Berardengo, F. Lucà, S. Manzoni, S. Pavoni, and M. Vanali, "A PCA/Natural Frequency-Based Approach for Damage Detection: Implementation on a Laboratory Structure Subjected to Environmental Variability," 2024, pp. 133–139. doi: 10.1007/978-3-031-68901-7_17.