DOI: 10.3217/978-3-99161-057-1-135

Advancing PEAR: Development of a Bridge Benchmark Datasets for PBSHM Research

Connor O'Higgins[1, 0000-0001-7034-005X], Tristan Gowdridge[2, 0009-0008-9979-8396], David Hester[1, 0009-0007-2280-4340], Keith Worden[2, 0000-0002-1035-238X], Daniel S. Brennan[2, 0009-0001-7223-3431]
[1]School of Natural and Built Environment, Queen's University of Belfast, Stranmillis Road, Belfast, BT9 5AG, UK
[2]Dynamics Research Group, School of Mechanical, Aerospace, and Civil Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3TD, UK
email: c.ohiggins@qub.ac.uk

ABSTRACT: Population-Based Structural Health Monitoring (PBSHM) is an emerging field in Structural Health Monitoring that leverages data from multiple structures to enhance the assessment of individual structures. Unlike traditional SHM, which generally relies on data from a single structure, PBSHM utilises collective knowledge from a population to facilitate increasing the knowledge on an individual structure. Transfer learning enables the inference from a source structure to a target structure within the population. One of the limitations of this method is that a lot of transfer-learning methods require data models that are trained using substantial amounts of high-quality data which can be difficult to obtain. To support PBSHM research, the concept of the Population-based SHM Engineered Asset Resource (PEAR) has been introduced. PEAR is conceptualised as a benchmark dataset containing semi-realistic structures and associated data intended to drive the development and validation of PBSHM methodologies. This work advances the PEAR prototype by developing complete populations for two types of bridges, along with their associated data. The pipelines for generating these populations are presented, detailing how they produce structural data and PBSHM-specific models. Additionally, a simple analysis of the generated populations is conducted, demonstrating their utility in PBSHM research and showcasing the potential of PEAR as a resource for current and future PBSHM research.

KEY WORDS: SHMII-13; Population-based Structural Health Monitoring, Benchmark Dataset, Irreducible Element Model, Bridges

1 INTRODUCTION

In traditional Structural Health Monitoring (SHM), benchmark datasets such as the S101 [1] and Z24 [2] bridges have been pivotal in advancing the field. These datasets have provided researchers with common platforms to test, validate, and compare new SHM methods and algorithms. However, in the emerging field of Population-Based Structural Health Monitoring (PBSHM), data are leveraged from multiple structures to enhance the assessment of individual assets, no equivalent benchmark datasets currently exist. This gap is not only a reflection of the relative recency of PBSHM but also the inherent complexity of gathering multi-structure data necessitated to form a meaningful population.

To address this challenge, the Population-based SHM Engineered Asset Resource (PEAR) was devised as a potential solution. The foundational principles and envisaged structure of PEAR were outlined in a previous conference paper [3], laying the groundwork for a benchmark dataset that integrates curated synthetic populations. The envisioned PEAR dataset aims to serve as a standard for evaluating and advancing PBSHM methodologies by providing researchers with readily-accessible, semi-realistic data representative of various structural populations.

The fundamentals of the PEAR database have been established, this paper extends that work by developing two specific bridge populations for inclusion in the PEAR dataset. This work not only demonstrates the feasibility of generating synthetic populations but also how these populations can serve as a practical resource for PBSHM research.

The remainder of this paper is structured as follows. First, a background section provides an overview of PBSHM, detailing existing databases, schemas, and benchmark datasets, and highlighting the motivation for this work. This section is followed by an overview of the foundations of PEAR and the specific requirements for creating a benchmark dataset that is applicable for a population-based approach. Next, the "Developing Bridge Dataset" section describes the process of generating the two initial bridge populations, focussing on the design of semi-realistic bridge structures and the simulation of their structural response data. Subsequently, the "Bridge Datasets" section presents the developed populations, including a simple analysis to showcase their utility in PBSHM methods. Finally, the paper concludes with a discussion of the current dataset's limitations, potential avenues for future expansion, and the broader implications for advancing PBSHM research.

2 BACKGROUND

Population-Based Structural Health Monitoring (PBSHM) represents a shift from traditional SHM by focussing on the analysis of data from multiple structures within a related population. This approach not only provides valuable insights into the collective behaviour of the population but also enhances the understanding of individual structures. In contrast to conventional SHM methods that typically concentrate on a single structure, PBSHM enables the application of advanced techniques such as transfer learning, where knowledge gained from one task or structure is leveraged to improve performance on a related task in another structure. This methodology allows models to be adapted rather than built from scratch, thereby

DOI: 10.3217/978-3-99161-057-1-135

enhancing efficiency and potentially making inferences about the condition of a structure that would otherwise be missed.

A critical aspect of implementing transfer learning in PBSHM is the careful identification of structural similarities across the population. Without a clear understanding of these similarities, there is an increased risk of negative transfer, where inappropriate model adaptation could degrade performance. To combat this risk, techniques such as Irreducible Element (IE) models combined with graph-based approaches have been developed [4], [5]. These methods are instrumental in identifying and quantifying both the similarities and variations among structures, ensuring that transfer learning is applied only where it is most appropriate. This framework for similarity assessment forms the foundation for effective knowledge transfer across structures.

Despite these methodological advancements, a significant challenge in PBSHM remains: the scarcity of comprehensive data representing populations of structures. Previous studies have attempted to address this gap by collecting data from similar populations, for example, one study collected data from four beam-and-slab bridges and two pedestrian footbridges [6]. Other research efforts have simulated populations using models with 10 degrees of freedom [7] or even toy structures to validate PBSHM transfer-learning methods and graph-matching algorithms [8]. However, these approaches have been limited in scope, so the developed datasets cannot always be used to test novel PBSHM methods.

It is within this context of data scarcity and the development of new PBSHM methods that the need for benchmark datasets in PBSHM becomes evident. A robust benchmark dataset would not only facilitate the testing and validation of new methodologies but also drive forward the development of PBSHM research. This motivation underpins the development of the Population-based SHM Engineered Asset Resource (PEAR) dataset, which aims to provide the PBSHM community with a dedicated set of synthetic populations that represent real-world structures and realistic structural behaviours for algorithm development and benchmarking.

3 PEAR OVERVIEW

This section gives an overview of the PEAR database, including its key requirements (Section 3.1), how the database is structured (Section 3.2), and the stages involved with producing populations of structures and data (Section 3.3).

3.1 Requirements

Requirement 1:

The primary objective for the dataset is to function as a robust testbed for both established and emerging PBSHM methods. It must support the development and evaluation of methods and techniques across the entire remit of PBSHM. This remit includes the calculation of similarity scores and transfer-learning methods. Moreover, the dataset should facilitate structural comparisons via similarity metrics, including IE models and graph-matching algorithms, which are essential for guiding successful knowledge transfer. The datasets should also be compatible with various machine-learning techniques, ensuring high-quality data are available for training and validation purposes.

Requirement 2:

The dataset should be designed for ease of searchability, enabling users to quickly locate and extract relevant data subsets. To ensure this level of accessibility, the dataset must be thoroughly indexed so that key variables, such as the type of structure or, in the case of bridges, the number of spans, are easily queried and filtered.

Requirement 3:

The methodologies employed to generate the structures within the dataset must be transparent, clearly documented, and, whenever possible, grounded in real-world structural design practices. This approach guarantees that the simulated structures accurately reflect their real-world counterparts. Additionally, the dataset should capture the natural variability seen in practice, for instance, differences in span lengths, beam dimensions, and deck thicknesses in bridge designs, thus representing the diversity found in actual structural stocks.

Requirement 4:

A shared-data domain is fundamental to the success and broad adoption of PBSHM practices. Building on the work of Brennan et al. [5], the dataset will adhere to the PBSHM schema, a standardised format that ensures consistency in data storage and interpretation for PBSHM. All data and future additions to the dataset must comply with this schema. In cases where the current PBSHM schema does not accommodate certain data components, it will be necessary to propose and integrate an extension to the PBSHM schema to incorporate this data.

3.2 Dataset Structure

Designing the dataset's structure is important to ensure effective data retrieval and utilisation. A well-structured dataset will facilitate users to quickly access the specific data they need, making it easier to perform tasks with the datasets and subsets of the data. In the context of PEAR, the structure is crafted not only to organise the data logically for easy retrieval but also to accommodate the addition of new data without disrupting existing records.

Figure 1 illustrates the four-level hierarchical organisation of the PEAR dataset: root category, subtype, dataset, and scenario. To illustrate how these levels interact, consider the analogy of a file system. At the highest level is a main "PEAR" folder. Inside this folder, there is a separate folder for each root category. Each root category folder contains folders for its various subtypes, which differentiate structures based on design or purpose. Within each subtype folder, there are individual dataset folders; each dataset represents a collection of structures forming the population. At this level of the database a descriptive file outlining the general forms of the population and allowable variations of these structures will be stored. The IE models and the meta data about each structure will also be stored at this level. The final level, scenario, is analogous to a load case in a structural model, for example, applying a 40-ton load to the mid-span of a bridge would constitute a single scenario.

DOI: 10.3217/978-3-99161-057-1-135

```
Root Category
└──── Subtype
        └──── Dataset
                └──── Scenario
```

Figure 1: The hierarchical structure of the PEAR dataset

In this hierarchical structure, a root category groups structures that are typically classified together, such as aeroplanes and bridges. Within each root category, subtypes further differentiate the structures. For example, within the aeroplane category, subtypes may be single-engine and multi-engine, while the bridge category could be divided into truss bridges and suspension bridges.

During the initial planning phase for PEAR, three primary root categories were identified for inclusion: bridges, masts, and wind turbines. Figure 2 shows the initial dataset organised into root and subtype levels. This work specifically generates data for two bridge populations: a beam-and-slab bridge population and a ladder-deck bridge population, each accompanied by a corresponding scenario.

```
PEAR Database
├──── 1. Bridges
│       ├──── 1.1 beam-and-slab
│       ├──── 1.2 ladder-deck
│       └──── 1.3 truss
├──── 2. Wind Turbines
│       ├──── 2.1 monopole-tower
│       └──── 2.2 lattice-tower
└──── 3. Masts
        ├──── 3.1 monopole-tower
        └──── 3.2 lattice-tower
```

Figure 2: The initial ‘root categories’ and ‘subtypes’ for PEAR

3.3 Dataset Stages

PEAR has been designed from the ground up to develop over time by iterations and expansion of the data included within. Defined stages serve as clear milestones, enabling collaborative efforts across the diverse disciplines involved in PBSHM. The initial design of the structure within any population is undertaken in Stage 0, the development of IE models in Stage 1 and the simulation of structural responses in Stage 2+. By breaking the development into stages, different teams can contribute depending on their area of research. For example, one team can focus on constructing the IE models needed to derive similarity metrics (Stage 1) while another develops simulations for scenario data (Stage 2). Importantly, each stage builds exclusively on the data from the preceding stages; for instance, any Stage 2 processes rely only on the outputs from Stages 0 and 1.

For a dataset to be integrated into PEAR, it must, at a minimum, complete Stage 0 and Stage 1. Once a dataset reaches this milestone, it is assigned a unique reference within PEAR, and the structures it contains are fixed to ensure consistency and reproducibility of results. The specific processes of Stage 2 are left flexible, allowing researchers to select the most relevant structural responses for their research as well as the most appropriate simulation method. The overall staging structure is designed to be generic, applying to all datasets regardless of their root category or subtype.

Stage 0: Design of the Dataset Population

In this initial stage, realistic structures are generated with pseudo-random properties drawn from a predefined range of parameters. These parameters, along with the methods used for their selection, are detailed and stored in the dataset. For each structure, a Structural Information (SI) model is created alongside a structural report. The report offers a detailed, human-readable description, while the SI model provides a computer-interpretable format that facilitates indexing and querying.

Stage 1: IE Models

At this stage, Irreducible Element (IE) models are developed for each structure in the population, i.e. one IE model for each SI model in Stage 0. The typical workflow for producing IE models will be using the parameters defined in the SI for each structure to create detailed IE models. After the completion for Stage 1 the dataset may be included in PEAR.

Stage 2+: Scenario Data

Using the data from Stages 0 and 1, simulated scenarios are then produced using the information in the generated IE models. Although PEAR does not prescribe specific simulation methods, the simulation outputs must be saved back into the dataset as valid PBSHM Schema data. Scenarios might include, but are not limited to, static-load displacements, natural frequencies, mode shapes, or frequency response functions. Moreover; a scenario can simulate conditions where the structure is considered 'healthy' or introduce 'damage' prior to simulation, with the results documented accordingly.

4 DEVELOPING BRIDGE DATASETS

This section outlines the process by which the two initial bridge populations have been developed for the PEAR dataset. The development process is divided into two major phases. The first phase focusses on designing semi-realistic bridges and creating their associated Irreducible Element (IE) models, corresponding to Stage 0 and 1 of the dataset. The second phase involves simulating the structural responses of these bridges, which represents Stage 2. This section describes these processes to create both a beam-and-slab bridge population and a ladder-deck bridge population.

4.1 General Form of Population

For each population included in PEAR, the first step is to define the general form of the structure. This general form specifies the primary structural components and their arrangement, ensuring a consistent yet flexible framework for generating individual structures. Once the general form is established, a set of rules and parameters is defined to guide the creation of each structure within the population. These rules, grounded in engineering principles, ensure that the generated structures are

DOI: 10.3217/978-3-99161-057-1-135

both realistic and representative of real-world variations. The diversity within each population is achieved by systematically varying these parameters, allowing for a range of structures that share a common form while exhibiting meaningful differences. The following subsections detail the general forms and parameterisation approaches for the beam-and-slab bridge population and the ladder-deck bridge population.

4.1.1 Beam-and-Slab Bridge

The general form of a beam-and-slab bridge consists of precast concrete beams that serve as the primary structural components of the bridge. These beams are typically lifted into place, where they are supported by piers or abutments that have been constructed in advance. Once the beams are positioned, an *in situ* concrete deck is poured over them to form the main deck of the bridge. Below are the main components and a brief description of a typical beam-and-slab bridge:

- Precast Prestressed Bridge Beams: These are the primary load-bearing components of the bridge. Precast off-site and prestressed to increase their load-carrying capacity, these beams are designed to resist the main traffic loads of the bridge.
- Reinforced Concrete Deck: The reinforced concrete deck is poured *in situ* on top of the precast beams, forming the main surface of the bridge.
- Diaphragm: The diaphragm is a transverse structural element placed between the bridge beams. Its primary function is to distribute loads evenly across the bridge and provide lateral stability, and help transfer load to the supporting structure of the bridge.
- Columns: Columns are vertical structural supports that transfer the load from the bridge deck and beams down to the foundation.

4.1.2 Ladder-Deck Bridge

The general form of a ladder-deck bridge consists of steel girder beams that serve as the primary structural components of the bridge. These beams are situated at the edge of the bridge with smaller steel girder beams spanning transversely, connecting the main beams. Once the beams are positioned and connected, an *in situ* concrete deck is poured over them to form the main deck of the bridge. Below are the main components and a brief description of a typical beam-and-slab bridge:

- Longitudinal Girder Beams: These are primary load-carrying members that run along the length of the bridge (parallel to the roadway). They bear the main loads from the deck and transfer them to the piers or abutments. These are typically made of steel and are placed at the two edges of the bridge.
- Transverse Girder Beams (Cross Beams): These are secondary beams that span between the longitudinal girders, providing lateral support and distributing loads from the deck to the longitudinal girders. These are also typically made of steel.
- Concrete Deck: The reinforced-concrete deck is poured *in situ* on top of the longitudinal and transverse beams, forming the main surface of the bridge.
- Columns: Columns are vertical structural supports that transfer the load from the bridge deck and beams down to the foundation.

4.2 Structural Parameters

This section outlines the parameters that can be adjusted within the general forms described in the previous section. These parameters are the tools used to create a varied population of structures while ensuring that each model remains realistic and grounded in sound engineering principles. Methods have been developed to select these parameters, ensuring that any variations still adhere to the constraints of real-world structural behaviour. For clarity, the parameters are divided into two groups:

1. Generic Bridge Variables: These parameters are common to all bridges, such as the number of spans and overall bridge length.
2. Subtype-Specific Parameters: These parameters are unique to each bridge subtype. For instance, in a beam-and-slab bridge, a key parameter might be the selection of the precast beam geometry.

By systematically varying these parameters, the PEAR dataset is able to generate diverse yet realistic bridge structures.

Generic Bridge Parameters

The generic bridge parameters define common characteristics shared by both bridge populations, such as the number of spans, span lengths, deck dimensions, column details, and material properties. The selection process for these parameters combines random selection from predefined ranges with engineering constraints to ensure that the resulting structures remain realistic and consistent with real-world practices. For example, the number of spans is determined by randomly choosing a value within a range that typically mirrors actual bridge stocks, usually between one and five spans. Once the number of spans is set, the span lengths are similarly selected from a defined range. However, to avoid unrealistic configurations, such as pairing an exceptionally long span with an extremely short one, an additional constraint is imposed. All selected span lengths must fall within 70% of each other, ensuring a realistic design.

Other parameters, like the width and thickness of the deck, are also chosen from ranges that reflect standard practices in bridge construction. The inclusion of columns is treated as a variable feature; whether columns are present is determined randomly, and if they are included, further details, such as their height and quantity, are specified. These column characteristics are based on established engineering principles; for instance, the minimum column height adheres to government standards for bridge clearance. Material properties, too, are selected from realistic ranges that reflect common construction materials, contributing to the overall authenticity of the generated structures while proving realistic variation in the population.

Subtype-Specific Parameters

Subtype-specific parameters are the parameters that are only relevant to each bridge subtype. For the beam-and-slab population, some key parameters include the beam centre-to-centre distance, the number of primary beams, the geometry of both primary and edge beams, and the diaphragm geometry. For the ladder-deck bridges, the parameters include the number of transverse beams and the geometry of both the primary and secondary girders. The methods for selecting these parameters tend to be more involved than those used for generic bridge variables to ensure that each choice results in realistic engineering practices.

DOI: 10.3217/978-3-99161-057-1-135

The process of selecting these parameters is highly dependent on the subtype and the accepted process for designing the structure. For instance, when determining the primary bridge beam geometry for each beam-and-slab bridge, the process begins with a list of common bridge beam types. Beam types that are incompatible with the given span length or bridge width are eliminated using span tables provided by bridge beam manufacturers. From the remaining viable options, a beam type is then randomly selected. This method ensures that the chosen beam geometry is suitable for that span length while allowing for variation in the population.

Similarly, for the ladder-deck subtype, the main structural component is determined using a range of acceptable span-to-depth ratios sourced from the encyclopedia for UK steel construction (SteelConstruction.info). A random ratio is selected from this range and, together with the predetermined span length, used to set the depth of the girder. Additional properties such as flange width and thickness are subsequently determined using comparable ratios from the same source.

The parameterisation strategy for the two bridge datasets employs a total of 17 distinct parameters for beam-and-slab bridges and 23 for ladder-deck bridges. By integrating randomness while adhering to sound engineering principles, the established general form and corresponding parameter sets create a robust framework for generating a diverse yet realistic population of bridge structures. This approach not only mirrors real-world variability but also enables the production of thousands of unique and plausible structures, significantly enhancing the utility of the PEAR dataset for PBSHM research.

4.3 Dataset Scenarios

A scenario is defined as a specific set of actions applied to the generated structures within a given population. These scenarios are designed to simulate various conditions or loads to determine the structural responses. The scope of the scenarios is flexible and can vary considerably based on the intended requirements. There is no single prescribed workflow for transforming the generated population into simulated structural responses, provided that the outputs remain compliant with PBSHM standards. In this work, we implement a simple scenario for both bridge types by employing finite-element models (software: LUSAS) to simulate their structural responses. However, alternative approaches, such as computational fluid dynamics, might be more appropriate for other applications, leaving the choice of methodology to the scenario authors.

4.3.1 Description of the Implemented Scenario

In this study, two simple scenarios were implemented to evaluate the structural response of the two generated bridge populations. Scenario 1 was the application of only the dead load on the bridge. Scenario 2 was the application of a point load of 40 kN at the mid-span of every span on each structure. After the mid-span load is applied, the maximum displacement and its corresponding location are extracted and recorded in a PBSHM-compliant format. This scenario is intentionally simple, which facilitates basic validation, such as comparing the self-weight of the structures with reaction forces of the FE models.

The implementation of this scenario can be defined using the following three actions.

1. The Irreducible Element (IE) models are converted into Finite Element (FE) models. This conversion uses the detailed information contained within the IE models to generate FE model files.
2. Loading parameters are defined by extracting necessary information from the Structural Information (SI) models. This step involves producing scripts for each load case, with one script generated per span for each bridge.
3. The generated loading scripts are integrated with the FE models to run simulations. Once the simulations are complete, the mid-span vertical displacements and reaction forces are extracted from the simulation results and stored in the output folder in the correct PBSHM format.

4.3.2 Validation of Structures and Data

Validation of the populations ensures that the generated structures and associated data are reliable, accurate, and true to the design intentions. By confirming that both the individual models and the overall dataset behave as expected, researchers can trust the integrity of the PEAR dataset and confidently use it for their PBSHM research.

During Stage 0 (design of structures), there are two main validation checks. The first is to check if the produced IE models comply with the required PBSHM schema. They are required to comply with the format to ensure consistency and that the developed PBSHM methods can be used on the IE models. The second validation step during Stage 0 is the examination of dataset statistics. These statistics verify that the composition of the populations aligns with the predetermined design process. Analysing parameter distributions and other statistical metrics ensures that the variability within the population matches what is expected based on the defined ranges and engineering principles. A discussion of the two populations developed for this work can be found in Section 5.1.

Further validation is carried out during the scenario stage by comparing simulated structural responses with theoretical predictions. The specifics of this will depend on the workflow to obtain the structural responses of the structure. For illustration purposes, the validation methods that were used in this work will be described. These methods are expected to be applicable when FE models are used as part of the workflow. After converting the IE models to FE models, two key validation methods are employed. First, dead-load reaction forces are extracted from the FE models and compared with the calculated weights of the structures as determined in Stage 0. This comparison confirms that the transition from the IE model to the FE model has maintained the integrity of the original design, ensuring that the overall weight remains consistent. Figure 3 presents the percentage difference between the design self-weight and the reaction forces from the produced FE models. With the average percentage difference being 1.69% and 1.55% for the beam-and-slab population and the ladder-deck population, respectively, there is very good agreement between the design stage and the FE model stage. For the second validation, a point load was applied to the structures, and displacement across the deck was recorded from the FE

 DOI: 10.3217/978-3-99161-057-1-135

simulation. These displacements were checked to ensure there were no discontinuities in the deformed mesh. This comparison serves to verify that the FE models are simulating sensible structural behaviour.

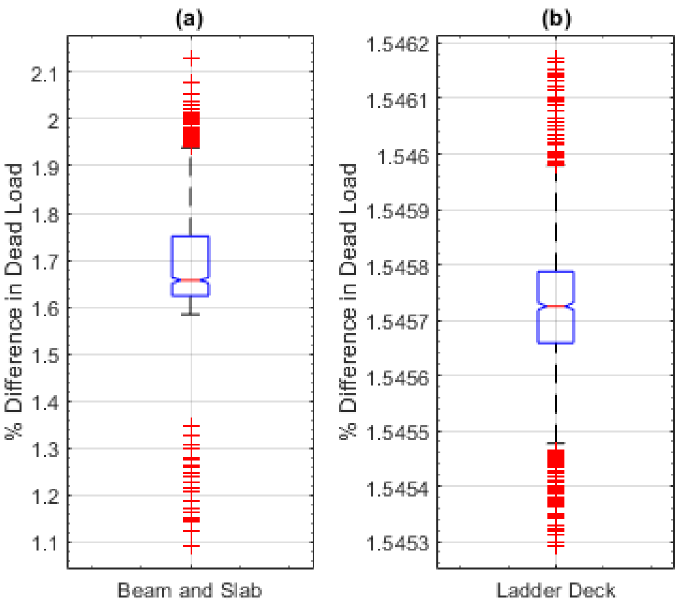

Figure 3: Percentage difference between the design self-weight of the bridges and the reaction forces from the FE models for (a) beam-and-slab population and (b) Ladder-deck population

As a further validation step data from all stages will go through a testing phase here the data will be accessible to the public for review and comments for a period of time. Following this review period it will be uploaded to the published version of the PEAR database.

5 BRIDGE DATASETS

In this section, the populations for both beam-and-slab bridges and ladder-deck bridges are presented. Firstly the structural composition of each population is examined, detailing how the individual structures come together to form a representative population of each bridge subtype. Following this, a simple analysis of the outputs from the scenario described in Section 4.3 is presented to demonstrate the practical application of these populations within the PBSHM framework. Finally, a similarity analysis of the populations is presented in section 5.3. This analysis serves to highlight how the PEAR dataset can be used as a valuable benchmark resource for evaluating transfer learning and other advanced PBSHM methods.

5.1 Dataset Statistics

Examining the overall composition of the two generated bridge populations provides insights and validation at the population level. By presenting a range of statistics, including the distributions of key parameters and variability, this analysis provides an insight into the composition of the populations. The main purpose of examining the dataset statistics is to ensure that the variations in the dataset reflect realistic engineering principles and expected real-world trends. Additionally, these statistical insights serve as a validation step, confirming that the methods used to select the parameters yielded the expected set of structures.

For illustration, Figure 4 below displays the distributions of four key parameters for the beam-and-slab bridge dataset: the number of spans, the centre-to-centre spacing of the beams, the number of primary beams, and the width of the deck (subplots (a) to (d), respectively). The distribution of the number of spans is approximately even between one and five, which aligns with expectations, given that this value is selected randomly without influence from other design factors. In contrast, the distributions for the centre-to-centre spacing, the number of primary beams, and the deck width are less uniform because of their interdependent selection processes.

In Figure 4(b), the distribution of the centre-to-centre spacing reflects the specific weighting applied during parameter selection, following a ratio of 4:2:3:1:1, which is clearly visible in the resulting histogram. Figure 4(c), which shows the number of primary beams, indicates a skew towards higher numbers. Although the number of beams is randomly chosen between 4 and 10, this selection is further refined by ensuring that the combination of the number of beams and the centre-to-centre spacing produces a bridge width within the range of 6 to 20 meters. This constraint necessitates reselecting the number of beams when the initial combination falls outside the acceptable range, thereby skewing the distribution toward larger values. Finally, the distribution seen in Figure 4(d) for the deck width is a direct consequence of the deck width being determined by using the centre-to-centre spacing and the number of primary beams.

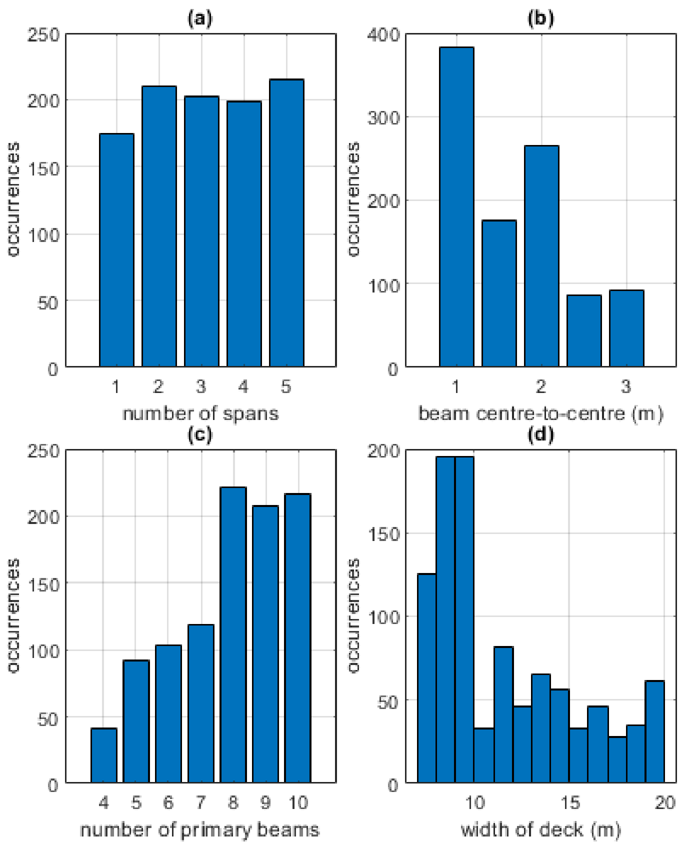

Figure 4: Beam-and-slab population statistics (a) the number of spans (b) the centre-to-centre spacing of the beams (c) the number of primary beams and (d) the width of the deck

DOI: 10.3217/978-3-99161-057-1-135

Figure 5 presents the distributions of four key parameters for the ladder-deck bridge dataset: the number of spans, the width of the deck, the height of the main girder, and the height of the cross girders (subplots (a) through (d), respectively). The number of spans is distributed approximately evenly between one and five, similar to the statistics for beam-and-slab bridges. In contrast, the deck width, shown in Figure 5(b), displays a roughly even distribution as well; this is because of its independence from other design parameters in ladder-deck bridges, unlike in the beam-and-slab case.

The height of the main girders, illustrated in Figure 5(c), is determined using defined span-to-girder depth ratios and so is dependent on the length of the spans, with the span limitations setting clear upper and lower bounds for the distribution. Similarly, the height of the cross girders, depicted in Figure 5(d), is selected based on specific span-to-girder depth ratios, but with the additional constraint that they must not exceed the depth of the main girder.

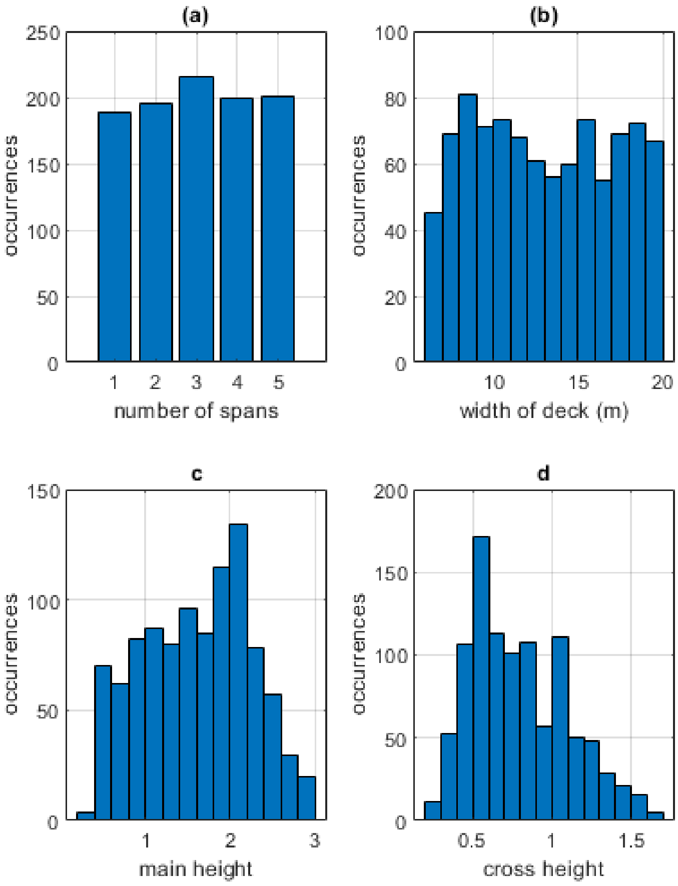

Figure 5: Ladder-deck population statistics (a) the number of spans (b) the width of the deck (c) the height of the main girder and (d) the height of the cross girders

5.2 *Scenario Analysis*

In this section, the results from the two implemented scenarios and conduct a simple analysis are presented. The aim is to provide the reader with a clear example of the types of data that can be generated from the populations, illustrating the applications of the PEAR dataset within PBSHM research.

Figure 6 illustrates the displacement responses of a representative ladder-deck bridge under the various loading conditions simulated. In Figure 6(a) the vertical displacement of the bridge under dead load is presented. Figure 6(b) to (d) present the displacement responses when a 400 kN load is applied at the mid-span of each of the three spans, respectively. From the displacement results shown in Figure 6, the maximum displacement values for each load case were extracted. These values were then compared against the corresponding maximum displacements from the rest of the population.

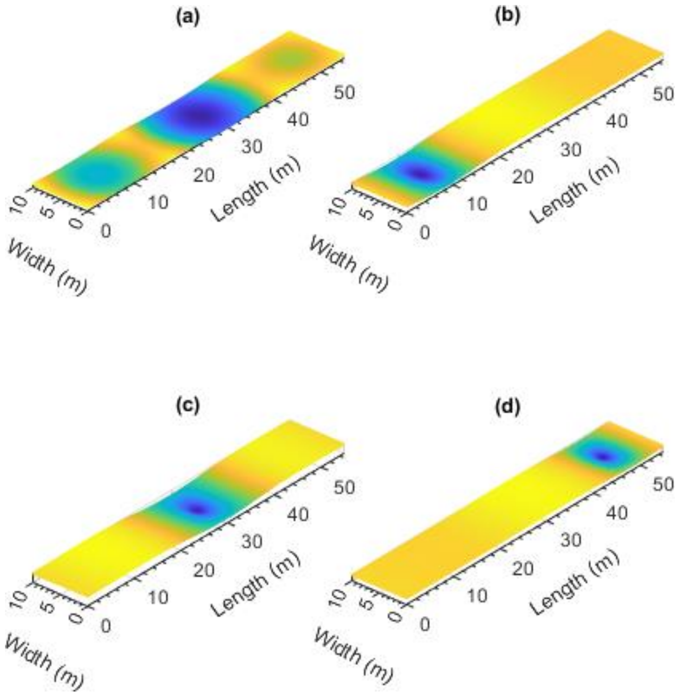

Figure 6: Vertical displacement output from FE model for (a) dead load (b) 400 kN applied to span 1 (c) 400 kN applied to span 2 (d) 400 kN applied to span 3

Figure 7 presents the combined maximum displacement values extracted from all of the FE models from the ladder-deck population, plotted against the span length where the point load was applied. In this figure, the colour of each data point represents the height of the main beam, with blue indicating the smallest beam depths and yellow indicating the largest. The displacements vary from 0.0018 m to 0.0653 m. This figure illustrates the relationship between the span length, the main beam depth and the displacement of the bridge. One of the key observations from the figure is the presence of distinctive bands. These bands show that as the span length increases, there is a corresponding increase in the depth of the beams, a trend that is consistent with design principles used for the ladder-deck bridges.

DOI: 10.3217/978-3-99161-057-1-135

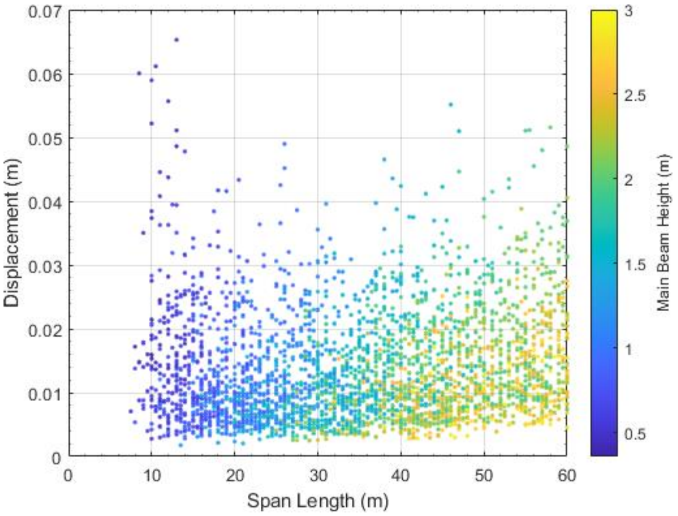

Figure 7: Span length vs the max displacement with the datapoints coloured with respect to main beam depth

5.3 Similarity Analysis

In this section, a basic similarity analysis between the populations will be presented. As stated in the prior, similarity assessments are key to PBSHM as they show where it is appropriate to undertake transfer learning between structures. In this work, the Jaccard index, a method used in graph theory [9], is employed to measure the similarity between IE models. The method used here will be slightly amended by embedding a geometry attribute of the elements in the IE model, specific details of this method can be found in [5]. The goal of this section is to highlight the suitability of the data in the PEAR being used for similarity analysis. Figure 8 shows the similarity matrix for a subset of 125 ladder-deck bridges. In the figure each pixel represents the similarity of one ladder-deck structure with another with dark blue representing least similar and yellow representing most similar. In this figure the structures are grouped by how many spans the bridges have; one-span bridges being grouped in the first 25 structures (indicated by the red box) and two-span bridges grouped in the next 25 positions and so on. Grouping in this way means that pattens can be observed about the relationship between the number of spans a bridge has and the similarity of other bridges with the same number of spans. In the case presented in Figure 8 it can be seen that the one-span bridges show a higher level of similarity with each other than with bridges of differing span. This pattern is repeated with the two-span bridges but the pattern seems to be diminished as the numbers of spans increase.

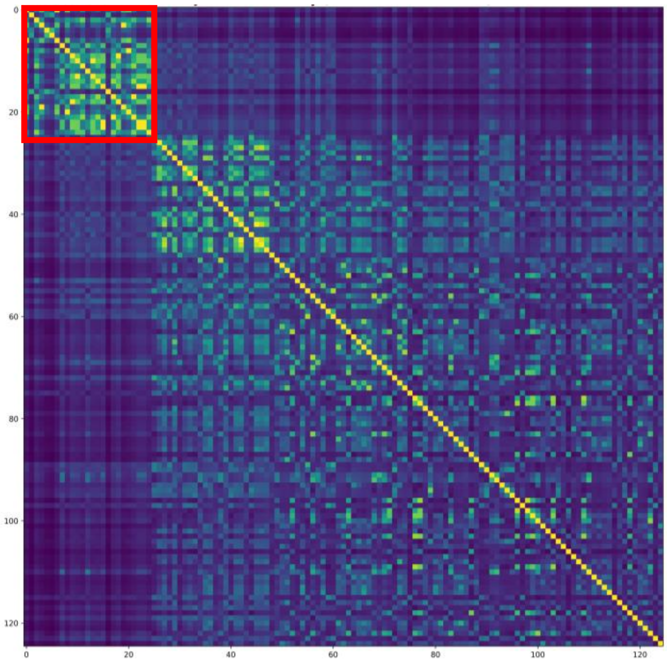

Figure 8: Similarity matrix for a subset of 125 ladder-deck bridges grouped by number of spans

To investigate this pattern further the similarity scores can be averaged across bridges with the same number of spans. This average is presented in Figure 9 and confirms the pattern seen in Figure 8. The bridges with the highest similarity are those with one span and as the number of spans increases the level of similarity decreases.

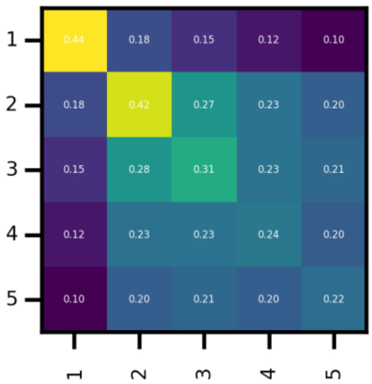

Figure 9: Similarity matrix for a subset of 125 ladder-deck bridges averaged by number of spans

This process was repeated for the beam-and-slab bridge population and the results are presented in Figure 10. This population largely follows the same pattern however the variability seems to be higher than that compared to the ladder-deck population this is most likely because of the other variations (other than number of spans) in the population affecting the similarity score.

DOI: 10.3217/978-3-99161-057-1-135

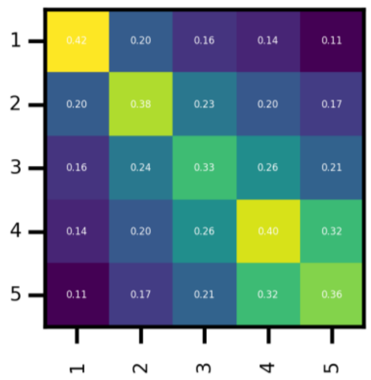

Figure 10: Similarity matrix for a subset of 125 beam-and-slab bridges averaged by number of spans

The similarity analysis in this section is the result of the question of how the number of spans of structures effects the similarity score between other members of the same population? This was chosen as an illustrative example of how a similarity analysis may be undertaken; however, it is only a single example of what may be considered in a similarity analysis. This produced dataset allows for a bespoke similarity assessment based on the needs of the specific research.

6 CONCLUSION

A fundamental challenge in Population-Based Structural Health Monitoring (PBSHM) is the need for data spanning multiple structures within a population. Ideally, these data would be sourced from real-world structures; however, acquiring comprehensive datasets across similar populations is challenging because of practical constraints, including data availability, monitoring costs, and access limitations. To address this issue, the database presented in this paper provides a synthetic dataset specifically tailored to the requirements of PBSHM research.

This work serves as a proof of concept for generating realistic populations of structures and associated data within a structured database. Two distinct bridge populations, beam-and-slab bridges and ladder-deck bridges, have been developed using engineering principles to maximise their realism. These structures have been validated via statistical analysis and comparison with theoretical expectations obtained from FE simulations. The FE models enabled structural responses under different loading conditions to be obtained.

The processes outlined in this work have been developed to be generalisable, meaning that aspects of engineering design can be incorporated into the population, such as varying material properties to adjust the material's strength. More complex design aspects, such as the amount of reinforcement or prestressed reinforcement, can be incorporated if the numerical simulation (e.g., an FE model) allows for it.

The results presented in this study demonstrate that it is feasible to create synthetic bridge populations that exhibit realistic structural behaviours, making them suitable for PBSHM development and validation. Although the analysis performed in this paper is relatively simple, it highlights the potential of the database for a range of PBSHM methodologies, including data-driven condition assessment and transfer-learning applications.

The approach outlined in this paper has been intentionally designed to be flexible, allowing for the generation of both bespoke structural populations and corresponding datasets depending on specific research needs. Future work will build on this foundation by expanding the database with additional structural populations and scenarios, further enhancing its applicability for PBSHM research and facilitating broader adoption within the field.

ACKNOWLEDGMENTS

The authors of this paper gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) via grant reference EP/W005816/1. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

REFERENCES

[1] M. Döhler, F. Hille, L. Mevel, and W. Rücker, 'Structural health monitoring with statistical methods during progressive damage test of S101 Bridge', *Eng. Struct.*, vol. 69, pp. 183–193, Jun. 2014, doi: 10.1016/j.engstruct.2014.03.010.

[2] B. Peeters and G. De Roeck, 'One-year monitoring of the Z24-Bridge: environmental effects versus damage events', *Earthq. Eng. Struct. Dyn.*, vol. 30, no. 2, pp. 149–171, Feb. 2001, doi: 10.1002/1096-9845(200102)30:2<149::AID-EQE1>3.0.CO;2-Z.

[3] C. O'Higgins, D. Hester, K. Worden, and D. S. Brennan, 'Towards PEAR: a Benchmark Dataset for Population-based SHM', presented at the IMAC-XLIII, Orlando, Florida, Feb. 2025.

[4] J. Gosliga, P. A. Gardner, L. A. Bull, N. Dervilis, and K. Worden, 'Foundations of Population-based SHM, Part II: Heterogeneous populations – Graphs, networks, and communities', *Mech. Syst. Signal Process.*, vol. 148, p. 107144, Feb. 2021, doi: 10.1016/j.ymssp.2020.107144.

[5] D. S. Brennan, J. Gosliga, E. J. Cross, and K. Worden, 'Foundations of population-based SHM, Part V: Network, framework and database', *Mech. Syst. Signal Process.*, vol. 223, p. 111602, Jan. 2025, doi: 10.1016/j.ymssp.2024.111602.

[6] A. Bunce *et al.*, 'On population-based structural health monitoring for bridges: Comparing similarity metrics and dynamic responses between sets of bridges', *Mech. Syst. Signal Process.*, vol. 216, p. 111501, Jul. 2024, doi: 10.1016/j.ymssp.2024.111501.

[7] A. J. Hughes, R. J. Barthorpe, N. Dervilis, C. R. Farrar, and K. Worden, 'A probabilistic risk-based decision framework for structural health monitoring', *Mech. Syst. Signal Process.*, vol. 150, p. 107339, Mar. 2021, doi: 10.1016/j.ymssp.2020.107339.

[8] Daniel S Brennan, Timothy J Rogers, Elizabeth J Cross, and Keith Worden, '"On calculating structural similarity metrics in population-based structural health monitoring', *Submitt. Data Centric Eng.*, 2024.

[9] P. Jaccard, 'Distribution de la flore alpine dans le Bassin des Dranses et dans quelques régions voisines', 1901, doi: 10.5169/SEALS-266440.