



This is a repository copy of *Machine learning for structural design models of continuous beam systems via influence zones*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/211256/>

Version: Published Version

---

**Article:**

Gallet, A. [orcid.org/0000-0003-4939-9916](https://orcid.org/0000-0003-4939-9916), Liew, A., Hajirasouliha, I. [orcid.org/0000-0003-2597-8200](https://orcid.org/0000-0003-2597-8200) et al. (1 more author) (2024) Machine learning for structural design models of continuous beam systems via influence zones. *Inverse Problems*, 40 (5). 055011. ISSN 0266-5611

<https://doi.org/10.1088/1361-6420/ad3334>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

PAPER • OPEN ACCESS

## Machine learning for structural design models of continuous beam systems via influence zones

To cite this article: Adrien Gallet *et al* 2024 *Inverse Problems* **40** 055011

View the [article online](#) for updates and enhancements.

### You may also like

- [Preliminary assessment of the earth's surface movement in the impact zone of a subway tunnel construction](#)  
G V Alekseev, E A Medvedev and A A Tereshin
- [Morphodynamics of Irkutsk water storage basin influence zone and principles of its geoinformation mapping](#)  
V P Stupin, L A Plastinin and B N Olzoev
- [Risk Assessment of Debris Flow in Longchi Area of Dujiangyan based on GIS and AHP](#)  
Hai Xiao, Xingsheng Tang and Hongtao Zhang

# Machine learning for structural design models of continuous beam systems via influence zones

Adrien Gallet<sup>1,\*</sup> , Andrew Liew<sup>2</sup> , Iman Hajirasouliha<sup>1</sup>   
and Danny Smyl<sup>3</sup> 

<sup>1</sup> Department of Civil and Structural Engineering, The University of Sheffield, Sheffield, United Kingdom

<sup>2</sup> Unipart Construction Technologies, Rotherham, United Kingdom

<sup>3</sup> School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, United States of America

E-mail: [agallet1@sheffield.ac.uk](mailto:agallet1@sheffield.ac.uk)

Received 23 August 2023; revised 26 February 2024

Accepted for publication 12 March 2024

Published 2 April 2024



CrossMark

## Abstract

This work develops a machine learned structural design model for continuous beam systems from the inverse problem perspective. After demarcating between forward, optimisation and inverse machine learned operators, the investigation proposes a novel methodology based on the recently developed influence zone concept *which represents a fundamental shift in approach compared to traditional structural design methods*. The aim of this approach is to conceptualise a non-iterative structural design model that predicts cross-section requirements for continuous beam systems of arbitrary system size. After generating a dataset of known solutions, an appropriate neural network architecture is identified, trained, and tested against unseen data. The results show a mean absolute percentage testing error of 1.6% for cross-section property predictions, along with a good ability of the neural network to generalise well to structural systems of variable size. The CBeamXP dataset generated in this work and an associated python-based neural network training script are

\* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

available at an open-source data repository to allow for the reproducibility of results and to encourage further investigations.

Keywords: machine learning, structural design models, neural networks, influence zone, inverse problems.

## 1. Introduction

It was recently argued that structural design is an *inverse problem* [1], in which one estimates the *model parameters* (the causal factors) of possible structural solutions from a set of *structural utilisations* (the observations). This inverse problem perspective, highlighted in figure 1, is underscored by the ill-posed characteristics structural design shares with other inverse problems [2], which in civil and structural engineering include subject areas such as structural health monitoring [3, 4], self-sensing smart materials [5, 6] and forensic blast engineering [7, 8].

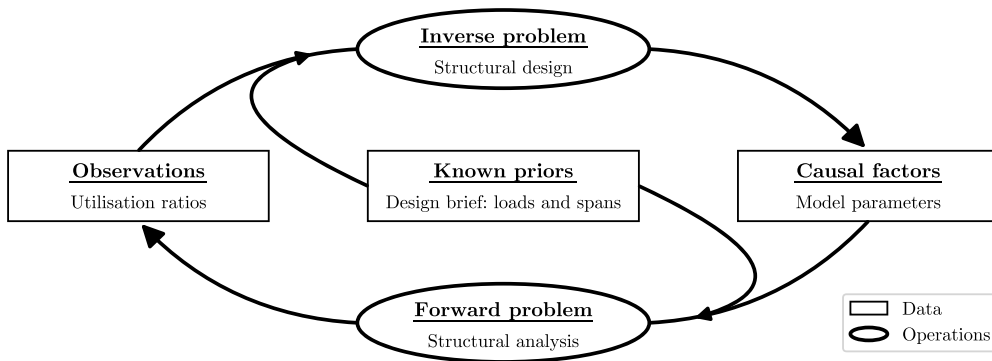
Inverse problems are predominantly solved iteratively [9], and unsurprisingly so is structural design [10], often with the help of structural optimisation such as size [11, 12], shape [13, 14], topology [15, 16] and layout optimisation [17, 18]. Provided that a clear objective function exists, these techniques are the state of the art for solving the structural design inverse problem *iteratively*.

However, in industry, the uptake of iteration based design approaches face certain barriers, including high computational costs [19], complex outputs that require additional post-rationalisation [20], and demand a particular expertise from practising design engineers that can be absent from engineering curriculums [21]. These challenges have encouraged researchers to investigate the use of machine learning (ML) methodologies for structural design [22]. This parallels a similar development of using ML within the domain of inverse problems [23], with exemplary applications in areas such as structural health monitoring [24, 25], that aid or replace the optimisation problem with *learned* components.

The earliest application of such machine learned components for structural design occurred in 1989 [26] with simplified perceptron models. This research was followed in the 1990s by more advanced feed-forward neural networks for simple reinforced concrete beam depth estimations [27] as well as cross-sectional area predictions of trusses [28, 29]. Whilst other machine learning modalities such as support vector machines [30] have also been studied, neural networks tend to outperform other ML models archetypes in terms of prediction error [31].

More recently, deep learning techniques have been investigated for structural design. These include convolutional [32, 33] and generative adversarial networks [34, 35] to accelerate topology optimisation, and the application of variational auto encoders for structural design space exploration [36]. A common limitation across such investigations is the inability for the same machine learned model to generalise to differently sized topologies and structural arrangements. These two challenges, highlighted by design ill-posedness and the inability of previous machine learning models to generalise to structural arrangements of arbitrary size, have motivated the work presented here.

This investigation has two objectives. The first objective is to reconcile the relationship between structural design, inverse problems and machine learning by developing a non-iterative structural design model for continuous beam systems using a multi-layer neural network. The authors believe that this perspective could serve as a framework to distinguish between different types of machine learning applications within the field of structural engineering in the future. The second objective is to address the inherent issue of generalisability in



**Figure 1.** The inverse problem perspective for structural design, which relies on known priors such as design brief details of loading and span requirements along with observations of utilisation ratios that represent structural adequacy to evaluate the model parameters of a solution, such as size, shape and topology of a viable structure. Structural analysis is treated as the forward problem.

respect to system size by taking advantage of a recently developed concept known as a continuous beam's influence zone [37]. This technique could potentially form the basis to generalise a design model for continuous structural systems of arbitrary topology, and might complement other techniques that attempt to address the generalisability issue such as graph neural networks [38, 39].

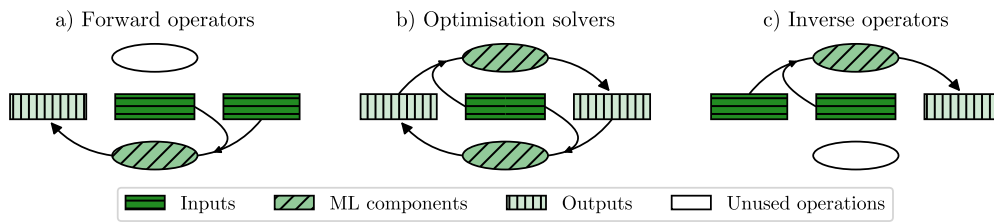
The paper is structured as follows: section 2 explores the problem statement from the inverse problem perspective and provides the rationale for machine learned design models, section 3 explains the methodology employed to develop the generalisable structural design model, section 4 presents the step-by-step process of the neural network development process, and section 5 discusses the model's generalisability and prediction variability, along with suggestions for further research.

## 2. Problem statement

### 2.1. A novel perspective

The inverse problem perspective for structural design as shown in figure 1 consists out of two operations, the *forward* and *inverse problem* (shown as the bottom and top ellipses, respectively) and three sets of data: *observations*, *known priors* and *causal factors* (shown as rectangles from left to right, respectively). One of the underpinning features of the inverse problem perspective is the clear demarcation between structural analysis and structural design, a distinction often re-iterated in engineering philosophy [40, 41], yet never linked to the corresponding nature of forward and inverse problems, respectively.

Both the forward (structural analysis) and the inverse problem (structural design) rely on *known priors*, shown centrally in figure 1, which can be thought of as constraints set by a design brief such as load and span requirements. During design they inform and regularise the search space of *causal factors* (model parameters such as section properties and topologies), and in analysis they allow the evaluation of *observations* (utilisation ratios such as ultimate (ULS) and serviceability limit states [42]). Unlike traditional inverse problems, the observations are not measured physically, yet are expressed theoretically based on the utilisation ratios that could



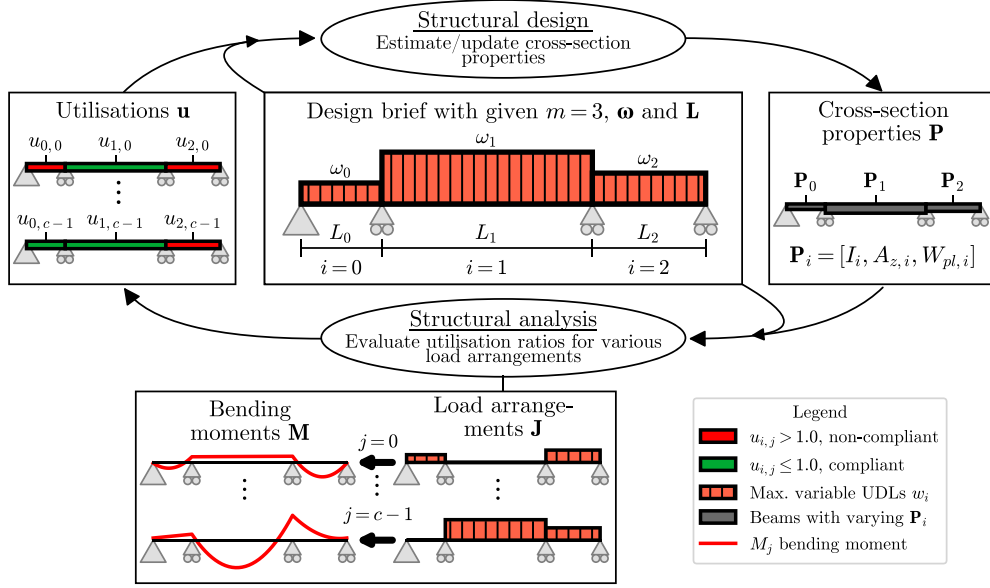
**Figure 2.** Types of machine learning (ML) components from the inverse problem perspective. Shapes in each sub-figure; ellipses: inverse problem (top), forward problem (bottom); rectangles: observations (left), known priors (middle), causal factors (right).

be measured from a compliant design solution which the set of causal factors correspond with; inverse problems are not defined by the physicality of the observations.

Within this context, the application of machine learning in structural engineering can be split into three categories based on the type of operations the machine learned components replace. These categories help distinguish between fundamentally different types of machine learning applications that occur within the context of structural engineering and are identifiable across different decades of the literature:

- (a) *ML forward operators*: machine learned components that aid or accelerate solving the forward problem (structural analysis) to inform or validate design decisions. Examples include neural network like models as quick re-analysis tools for optimum design (1991) [43] and machine learning models to determine the buckling behaviour and model decomposition of thin-walled members required for structural analysis (2023) [44].
- (b) *ML optimisation solvers*: machine learned components entirely motivated by the traditional iterative solution process to arrive at structural designs. Examples include ‘neural dynamic models’ developed as an alternative structural design optimisation technique (1995) [45] and a physics informed neural energy-force network that replaces both the structural design and analysis steps (2023) [46].
- (c) *ML inverse operators*: machine learned components which solve the inverse problem (structural design) by mapping a set of structural utilisations and known priors to model parameters directly. Examples include estimating cross-sectional properties for simple trusses directly based on known optimum examples using neural networks (1994) [29] and approximating topological optimised structures in real-time using convolutional neural networks (2022) [47].

These three categories can also be differentiated visually as shown in figure 2. It is worth noting that the field of ML forward operators has likely received the most research attention in the form of ‘surrogate models’ [48, 49]. In this respect, machine learned optimisation solvers and inverse operators are less common. Furthermore, the machine learned forward operators and optimisation solvers identified above typically require some form of iteration to achieve structural design; machine learned inverse operators on the other hand can be non-iterative [29, 47]. The ability to provide real-time design feedback is of particular interest to address the limitations of current iterative structural design approaches. To this end, and in support of the inverse problem perspective, this paper will focus on developing a non-iterative structural design model for continuous beam systems.



**Figure 3.** Design process of a continuous beam system from the inverse problem perspective.

## 2.2. Design problem: continuous beam systems

Continuous beam systems arise in structural engineering when rigid connections between members are required or unavoidable due to design or material considerations. The support fixity and structural connectivity render the system statically indeterminate. This poses a challenge from a design perspective, since the compliance of cross-sectional properties cannot be evaluated without knowledge of their magnitudes; this results in an iterative design process, especially for complex design scenarios with heterogeneous loading and span conditions [50].

Figure 3 highlights the design problem for continuous beam systems from the inverse problem perspective. The known priors, which are shown centrally as the design brief, include the number of members  $m$  in the system indexed by  $i$  with span length  $L_i$  from vector  $\mathbf{L} = [L_i]_{0 \leq i < m}$ , subjected to uniformly distributed loads (UDLs)  $\omega_i$  from vector  $\boldsymbol{\omega} = [\omega_i]_{0 \leq i < m}$ . These known priors and the utilisation ratios  $\mathbf{u}$  of the members, shown on the left in figure 3, are needed to evaluate the causal factors, shown on the right as the cross-section property vector  $\mathbf{P} = [\mathbf{P}_i]_{0 \leq i < m}$ .

The design problem is complicated due to the existence of  $c$  potentially critical load arrangements  $J$  indexed by  $j$  from set  $\mathbf{J} = [J_j]_{0 \leq j < c}$  shown at the bottom of figure 3. The size  $c$  of  $\mathbf{J}$  was studied in [37]. Each of these load arrangements cause different structural responses such as bending moments  $\mathbf{M}$ , and will give rise to a matrix of utilisation ratios  $u_{i,j}$  to form matrix  $\mathbf{u} = [u_{i,j}]_{0 \leq i < m, 0 \leq j < c}$  that can be evaluated with structural analysis to check for structural compliance ( $u_{i,j} \leq 1.0$ ). Instead of repeatedly assuming cross-section properties  $\mathbf{P}$  and conducting structural analysis calculations until the matrix of utilisation ratios  $\mathbf{u}$  are compliant, a machine learned inverse operator relies solely on the known priors and the utilisation ratios  $\mathbf{u}$  to directly evaluate the cross-section properties  $\mathbf{P}$ .

For the purpose of the continuous beam system considered in this work, several assumptions will be made: members are made out of S355 steel, are considered laterally restrained (and

hence not susceptible to lateral instability), Timoshenko–Ehrenfest beam theory is used to model this system and the structure will be analysed elastically yet designed against ULS plastic cross-section property checks as allowed by Eurocode EN 1993-1-1 5.4.2 (2) [51]. The cross-sectional properties to be evaluated include the major axis second moment of area  $I$ , the major axis shear area  $A_z$  and the major axis plastic section modulus  $W_{pl}$  for each member  $i$ . Together they form the member-based cross-sectional property vector  $\mathbf{P}_i$ :

$$\mathbf{P}_i = [I_i, A_{z,i}, W_{pl,i}]. \quad (1)$$

The structural analysis operation in figure 3 is defined by a forward operator:

$$\mathbf{u} = O_{\text{forw}}(m, \boldsymbol{\omega}, \mathbf{L}, \mathbf{P}) \quad (2)$$

and similarly the structural design operation by an inverse operator:

$$O_{\text{inv}}(m, \boldsymbol{\omega}, \mathbf{L}, \mathbf{u}) = \mathbf{P} \quad (3)$$

where both  $O_{\text{forw}}$  and  $O_{\text{inv}}$  rely on the same known priors, the design brief information  $m$ ,  $\boldsymbol{\omega}$  and  $\mathbf{L}$  that define the structural system and design problem.

### 2.3. The need for machine learned inverse operators

Defining an explicit non-iterative inverse operator for equation (3) is challenging due to the difficulty of inverting the forward operator and is directly linked to the ill-posed nature common across most inverse problems [2]. A quantitative evaluation of the extent of ill-posedness in structural design is not obvious, however it is possible to describe why the structural design problem shown in figure 3 is ill-posed, namely due to the infinite number of viable solutions, and:

- (a) The physical limitations introduced by yielding, buckling, serviceability that arise when combining the forward model with structural codes, resulting in a discontinuous relationship between the observations and causal factors.
- (b) The indeterminacy of the continuous beam system which increases with the number of members of the system.

Furthermore, any structural analysis forward operators themselves are approximations of the true behaviour of structures, and dealing with this associated uncertainty is a key challenge in design. For example, engineers need to decide if the assumptions and simplifications of structural analysis models, such as the material response (e.g. perfectly elastic) and underlying beam theory (e.g. Euler–Bernoulli theory), are representative of the structure’s true behaviour.

The difficulty of inverting a forward operator can be shown mathematically. Typically, the  $O_{\text{forw}}$  operator contains two steps. The first step, defined by  $O_{\text{forw},1}$  would evaluate the structural response of the system when subjected to a set of external forces  $\boldsymbol{\omega}$  in terms of deflections and internal forces, and the second step, defined by  $O_{\text{forw},2}$ , would take these structural response observations to evaluate the utilisation ratios based on design codes. Consider for example building a  $O_{\text{forw},1}$  operator using the stiffness matrix method to evaluate the internal force vector  $[f_p]_i$  for member  $i$  defined as:

$$[f_p]_i = [V_{1,i}, M_{1,i}, V_{2,i}, M_{2,i}]^T \quad (4)$$



where  $V$ , and  $M$  represent the internal shear forces and bending moments within member  $i$  at the start (index 1) and end of the member (index 2). Let us also assume, for simplicity, that the members consist out of steel with  $E$  and  $G$  for the Youngs and shear modulus, respectively, with a maximum yield stress of  $\sigma_y$ . In this case, the internal forces  $[f_p]_i$  for each member could be evaluated using a simplified Timoshenko–Ehrenfest beam theory for a *single* load arrangement by equation (5). To achieve this,  $[k_{pq}]_i$  is defined as the local stiffness matrix shown in equation (6),  $[K_{pq}]$  as the global stiffness matrix in equation (7),  $[d_q]$  as the nodal displacement vector in equation (8) with  $F_p([\omega_i])$  as the external force vector, where rows and columns of all matrices are indexed by  $p$  and  $q$ , respectively:

$$[f_p]_i = [k_{pq}]_i [d_q] \quad (5)$$

$$[k_{pq}]_i = \frac{EI_i}{L_i^3(1-\varphi)} \begin{bmatrix} 12 & 6L_i & -12 & 6L_i \\ 6L_i & 4L_i^2 & -6L_i & 2L_i^2 \\ -12 & -6L_i & 12 & -6L_i \\ 6L_i & 2L_i^2 & -6L_i & 4L_i^2 \end{bmatrix}, \quad \varphi = \frac{12EI_i}{A_zGL_i^2} \quad (6)$$

$$[K_{pq}] = [k_{pq}]_0 + [k_{pq}]_1 + \dots + [k_{pq}]_{m-1} = \sum_{i=0}^{m-1} [k_{pq}]_i \quad (7)$$

$$[d_q] = [K_{pq}]^{-1} [F_p([\omega_i])]. \quad (8)$$

These operations can be succinctly written to transform the cross-section vector  $\mathbf{P}$  with help of the known priors  $m$ ,  $\omega$ ,  $\mathbf{L}$  into the internal forces vector for each member  $i$ :

$$O_{\text{forw},1,i} = [f_p]_i = [k_{pq}]_i \left[ \sum_{i=0}^{m-1} [k_{pq}]_i \right]^{-1} [F_p([\omega_i])]. \quad (9)$$

Inverting this equation to yield  $O_{\text{inv}}$  is difficult since it would require separating or decomposing the individual cross-section properties  $\mathbf{P}_i$  out of the stiffness matrices  $[k_{pq}]_i$ . This cannot be done without, at minimum, making some assumptions about the relative proportions of the cross-section properties from one member to another. Inverting the second step of the forward operator  $O_{\text{forw},2}$  poses further challenges. Suppose  $O_{\text{forw},2}$  transforms the internal member forces  $[f_p]_i$  to evaluate the governing (critical) utilisation ratios indexed by  $r$  for  $t$  design checks for a single load arrangement  $J$ . For example, using the steel design code EN 1993-1-1 [51]:

$$u_i = O_{\text{forw},2,i}(O_{\text{forw},1,i}) = \max \left( [u_{ir}]_{0 \leq i < m, 0 \leq r < t} \right) = \max(u_{i,0}, u_{i,1}, \dots, u_{i,t-1})$$

where:

$$\begin{aligned} u_{i,0} &= \frac{V_{1,i}}{A_{z,i} \sigma_y / \sqrt{3}}, & u_{i,1} &= \frac{M_{1,i}}{W_{\text{pl},i} \sigma_y} \\ u_{i,2} &= \frac{V_{2,i}}{A_{z,i} \sigma_y / \sqrt{3}}, & u_{i,3} &= \frac{M_{2,i}}{W_{\text{pl},i} \sigma_y} \end{aligned} \quad (10)$$

⋮

$u_{i,t-1}$  for other compliance checks.

The difficulty here is that the governing utilisation ratio can change according to the known priors of the problem statement. This means that an individual equation for each possible critical design check would need to be derived. For example different design equations exist for the same structural check depending on the type of cross-section (Class 1 vs. Class 4) [51] a final design solution might contain, which is not known ahead of time. Note also that the equations above do not even consider the serviceability limit state, the multiple load arrangements  $\mathbf{J}$  which may be critical, nor the need to sufficiently discretise individual beam members.

It is because of the challenges identified above that machine learned inverse operators are particularly appealing, since they can approximate a relationship between a set of variables that may be difficult to encode explicitly [52]. Given a dataset generated by the  $O_{\text{forw}}$  operator that maps a set of cross-section properties  $\mathbf{P}$  to compliant utilisations ratios  $\mathbf{u}$ , one can train a probabilistic machine learning model  $O_{\text{inv}}^\dagger$  with parameters  $\theta$  to map the set of bounded utilisation ratios  $\mathbf{u}$  back to the cross-sectional properties  $\mathbf{P}$  with known priors  $m$ ,  $\omega$  and  $\mathbf{L}$ :

$$O_{\text{inv}}^\dagger(\theta, m, \omega, \mathbf{L}, \mathbf{u}) \approx \mathbf{P}. \quad (11)$$

By generating a dataset of valid structural designs with the help of existing optimisation approaches that contain the forward operator, a supervised machine learning model can be trained to learn the mapping of known priors and utilisations to cross-sectional properties directly. This represents a fundamental shift from traditional approaches employed in structural design that rely on engineering expertise and computationally expensive structural analysis or optimisation models at the point of design application. Machine learned inverse operators create non-iterative structural design models for which there currently exist no explicitly defined equivalents. Instead of focusing on accelerating forward models, computational resources can be invested in generating a dataset using physically complex yet realistic modelling assumptions. These machine learned structural design models aim to provide significantly greater generalisability than typical rules of thumb employed in design whilst still providing real-time feedback, benefit non-expert stakeholders whose own decision making relies on structural design outcomes and improves design knowledge permanence which can be difficult to attain due to industry turnover.

### 3. Methodology

#### 3.1. Choosing an appropriate machine learning model archetype

The aim of the inverse operator  $O_{\text{inv}}^\dagger$  is to predict the cross-section property vector defined by equation (1) numerically; therefore  $O_{\text{inv}}^\dagger$  will be a regression model. This restricts the types of supervised machine learning models of interest. The complexity and size of the design space are likely to demand a large dataset size discouraging the use of instance-based models such as the  $k$ -nearest neighbour algorithm that store similarity measurements in memory [53]. Similarly, support and relevance vector machines become impractical for datasets containing more than 3000 samples [54]. The non-linearity of the design problem voids the applicability of linear regression models, and decision trees (including the ensembled variants such as random forests) perform better at classification tasks [55].

These reasons motivated the use of neural networks, in particular multilayer neural networks (MLPs), a choice which is supported by evidence that suggests neural networks outperform other data-driven approximation algorithms in structural engineering applications [31]. Although various archetypes exist ranging from convolutional (CNNs), recurrent (RNNs) and graph-based types (GNNs), MLPs are commonly used in literature [28, 29], and the results

within this work could prove useful as a comparative performance measure for more advanced deep learning architectures [32, 34, 36] in future studies.

Multilayer neural networks have a fixed-dimensional input vector  $\mathbf{x}_0$  of size  $n$  that map to the output vector  $\mathbf{x}_D$  of size  $o$  with  $D$  layers. In this study, a network has  $D - 1$  hidden layers of height  $H$  each indexed by  $d$  and are defined by  $f_d(\mathbf{x}_d)$ , which contains a (non-linear) activation function  $a_d$  with weight matrix  $\mathbf{w}_d$  and bias vector  $\mathbf{b}_d$ . The weight matrices and bias vectors of each layer form the model's parameters  $\boldsymbol{\theta} = (\mathbf{w} = [\mathbf{w}_d]_{0 \leq d < D}, \mathbf{b} = [\mathbf{b}_d]_{0 \leq d < D})$ . Multiple hidden layers give form to the neural network  $f$  through a function composition defined as:

$$\begin{aligned} f: \mathbb{R}^n &\rightarrow \mathbb{R}^o \\ f(\mathbf{x}_0) &\rightarrow \mathbf{x}_D : f_{D-1} \circ \dots \circ f_1 \circ f_0(\mathbf{x}_0) \\ f_d(\mathbf{x}_d) &= a_d(\mathbf{w}_d \mathbf{x}_d + \mathbf{b}_d). \end{aligned} \quad (12)$$

The exact choice of architecture in terms of depth  $D$ , height  $H$  and activation functions  $a_d$  of the network as indicated in equation (12) will require experimentation to achieve acceptable performance with a good bias-variance trade-off [56]. More importantly though, the features used for the input vector  $\mathbf{x}_0$  will require careful consideration to create a generalisable inverse operator  $O_{\text{inv}}^\dagger$  as set out in equation (11).

### 3.2. Selecting appropriate neural network features

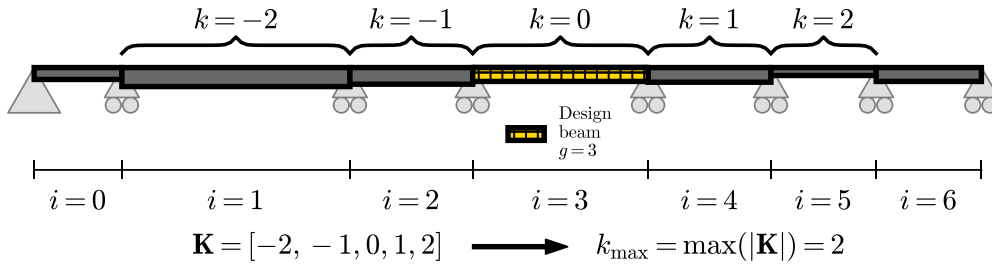
Feature selection, the process of choosing appropriate inputs, is essential for a machine learning model to generalise well to unseen data points. Unnecessary or irrelevant features can cause a model to learn a relationship with target variables that are not representative of the physical behaviour of the system, and thereby lead to worse results when interpolating within or extrapolating beyond the training set.

Previous studies of neural network based design models selected features relevant to the singular topology of the structural system at hand [29, 31]. Such approaches expose the largest limitation of MLPs: the fixed-dimensionality of the input vector [38]. These models may perform well for the particular topology they were trained against, yet the same model tends to perform worse or may not be applicable for differently sized structural systems, which severely limits their utility.

To address this limitation, this work takes advantage of a recently developed concept known as the *influence zone* [37]. The influence zone  $k_{\text{max}}$  is a measure of the extent to which surrounding design information is relevant for the utilisation evaluation of members. Whilst  $k_{\text{max}}$  differs for each member within a continuous structural system as shown in figure 4, for well defined design constraints and error thresholds, the maximum value of  $k_{\text{max}}$  within continuous beam systems converges towards a non-negative integer. The influence zone of member  $g$  is found when the following two conditions are met:

$$\begin{aligned} \left| 1 - \frac{u_{g,\text{cap}}}{u_{g,\text{true}}} \right| &\leq \epsilon_{\text{max}} \\ u_{g,\text{cap}} &= \max \left( \sum_{i=-k_{\text{max}}}^{k_{\text{max}}} \mathbf{u}_{g,i,j}(\boldsymbol{\omega}, \mathbf{L}, \mathbf{P}, \mathbf{J}) \right). \end{aligned} \quad (13)$$

In equation (13),  $\epsilon_{\text{max}}$  represents the maximum error threshold due to the difference between  $u_{g,\text{cap}}$ , the captured utilisation ratio of the design beam  $g$  for a given value of  $k_{\text{max}}$ , and  $u_{g,\text{true}}$ ,



**Figure 4.** A figurative influence zone of  $k_{\max} = 2$  for design beam  $g = 3$  within a  $m = 7$  continuous beam system with  $\epsilon_{\max} = 0.02$  limit.

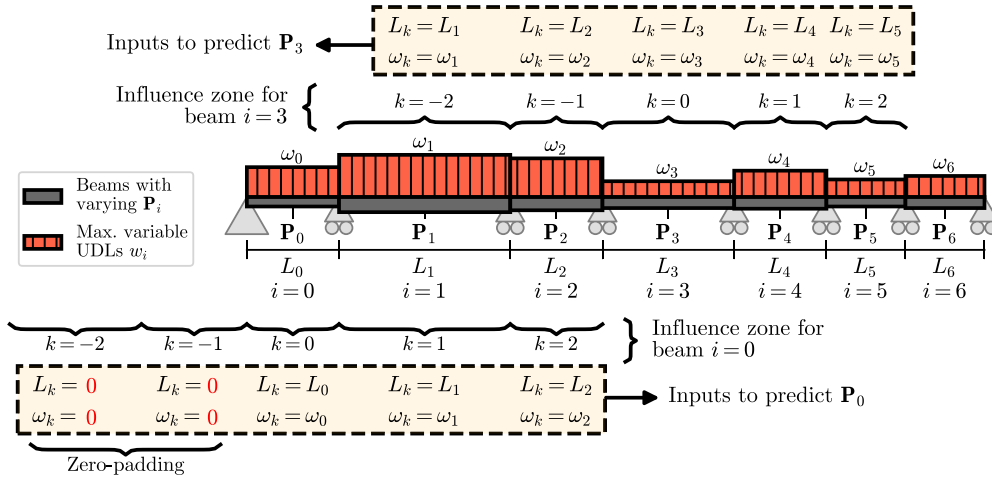
the true utilisation ratio of the design beam  $g$  if the contribution of all members of the continuous beam system had been considered.  $\mathbf{u}_{g,i,j}$  is the utilisation ratio contribution function towards the design beam  $g$  by member  $i$  based on the UDLs  $\omega$ , spans  $\mathbf{L}$ , structural properties  $\mathbf{P}$  and load arrangements  $\mathbf{J}$ . If the requirement for  $\epsilon_{\max}$  is sufficiently relaxed, the maximum influence zone  $k_{\max}$  can be determined for any potential continuous beam system arising under the specified design constraints [37]. This is extremely useful to ensure the relevant inputs are fed to a machine learning model. The influence zone thereby acts as a mechanics-driven feature selection process, and provides the basis to generalise to a continuous beam system of arbitrary size  $m$ .

### 3.3. Structuring features for arbitrary system size $m$

Zero-padding, the process of adding zero-valued inputs, arises in the context of convolutional neural networks to allow trained kernel filters to parse through the edges and corners of an input space [56]. This technique can also be applied to continuous beam systems to conceptualise a design model that parses over a structural system to make localised predictions for each member  $i$ . If the design information, here the UDLs  $\omega$  and span  $\mathbf{L}$  that fall within the influence zone are provided as inputs to the network, then this would result in an input vector  $\mathbf{x}_0$  of size  $n = 4k_{\max} + 2$ , as shown in figure 5 for member  $i = 3$  and  $i = 0$ . These inputs should, based on the principle of the influence zone, contain the relevant information to predict the cross-section properties of member  $d$  with an accuracy of up to  $\epsilon_{\max}$ .

It is now conceivable that the same neural network could be used to make a prediction for any other member using a fixed-dimensional input vector  $\mathbf{x}_0$  by structuring the inputs relative to the position of the design beam's influence zone. This would include end-span beams by using zero-padding as shown in figure 5 for member  $i = 0$ . Zero-padding in this instance is also logically consistent, since it corresponds with a beam that does not in fact exist; that is a beam of zero length  $L$  and zero UDL load  $\omega$ . Therefore, instead of structuring the neural network based on the absolute position of a beam within the entire continuous beam system (as indexed by  $i$ ), the inputs are structured relative to the influence zone of a design beam  $g$  to predict the cross-section properties of that design beam  $\mathbf{P}_g$ .

Whilst such an approach will require  $m$  forward passes (inferences) to predict the cross-section properties of an  $m$  sized system (one prediction per beam), it enables the same neural network to be applied to continuous beam systems of any size  $m$  for which the maximum influence zone value  $k_{\max}$  that determined the size of the input vector  $\mathbf{x}_0$  applies to. Based on the principle of influence zones, the neural network will be able to make predictions for continuous beam systems of size greater than the fixed-dimensional input vector size  $m > 2k_{\max} + 1$ , since



**Figure 5.** An illustration demonstrating the structuring of the neural network inputs using influence zones and zero-padding with  $k_{\max} = 2$ , leading to  $n = 4k_{\max} + 2 = 10$  inputs.

**Table 1.** Ranges and intervals of know priors used for the influence zone evaluation and data generation.

Property	Min.	Interval	Max.
$\omega$ (kN m <sup>-1</sup> )	5	5	325
$L$ (m)	0.5	0.5	20.0

any information outside the influence zone should by definition not be relevant (for an assumed  $\epsilon_{\max}$ ). On the other hand, zero-padding allows the same neural network to predict along system edges as well as continuous beam systems of sizes smaller than the influence zone.

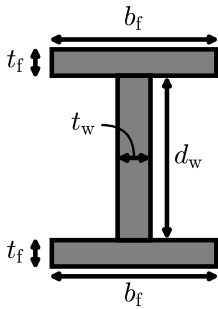
### 3.4. Generating an appropriate dataset

As explained previously, the maximum influence zone  $k_{\max}$  size depends on the design constraints and an assumed error threshold  $\epsilon_{\max}$ . These design constraints can be defined by setting minimum and maximum ranges on the known priors, UDLs  $\omega$  and spans  $L$ , as well as the cross-section properties within vector  $\mathbf{P} = [I, A_z, W_{pl}]$ :

$$\begin{aligned}
 \omega_{\min} &< \omega_i < \omega_{\max} \\
 L_{\min} &< L_i < L_{\max} \\
 I_{\min} &< I_i < I_{\max} \\
 A_{z,\min} &< A_{z,i} < A_{z,\max} \\
 W_{pl,\min} &< W_{pl,i} < W_{pl,\max}.
 \end{aligned} \tag{14}$$

Constraints for each of these variables were chosen generously to cover the entire range of potential continuous beam systems that arise in structural design (from fixed framed multi-storey buildings to continuous bridge decks). Table 1 highlights the ranges chosen for the UDLs and spans, along with the interval at which these inputs were sampled at using a random uniform distribution.

**Table 2.** Cross-section properties comparison between universal Beams (UB) from BS EN 10365:2017 and custom generated I-sections. Note in particular that mean dimension ratios ( $d_w/t_w$  etc) are identical for both groups of cross-sections.

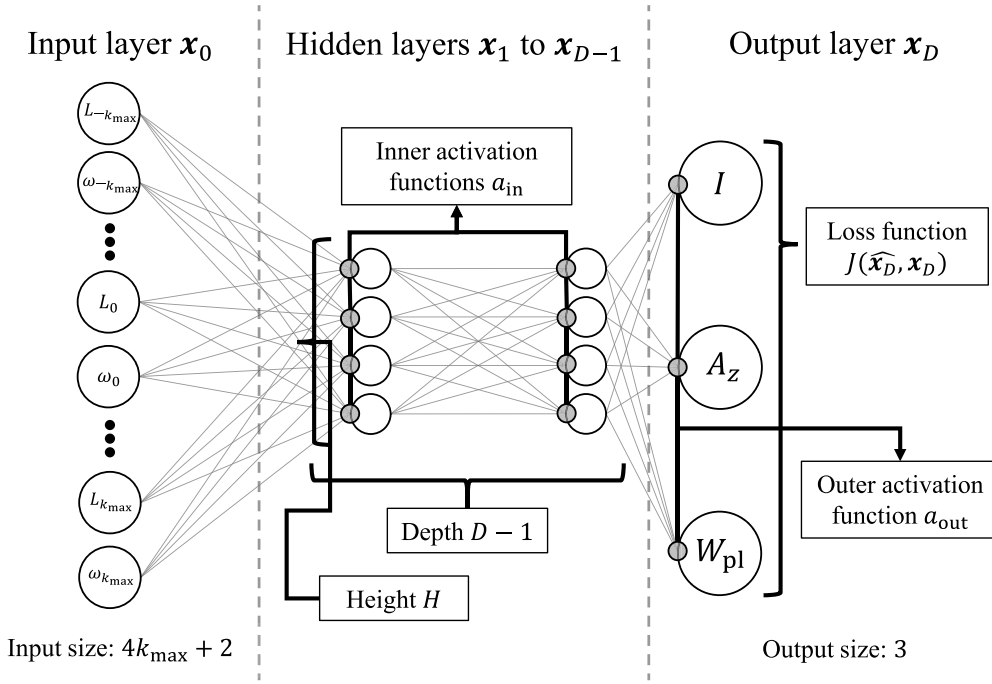
Diagram	Property	UB BS EN10365			Custom I-sections		
		Min	Mean	Max	Min	Mean	Max
	$t_w$ (mm)	4.0	13.0	36.1	3.0	24.0	45.0
	$d_w$ (mm)	112	550	928	30.3	243	455
	$t_f$ (mm)	6.8	21.4	65.0	4.8	38.6	72.4
	$b_f$ (mm)	76.0	220	421	55.5	444	833
	$d_w/t_w$	23.9	43.9	59.9	43.9	43.9	43.9
	$t_f/t_w$	1.19	1.61	1.90	1.61	1.61	1.61
	$b_f/t_w$	8.7	18.5	27.5	18.5	18.5	18.5
	$A_z$ (cm <sup>2</sup> )	4.47	87.9	334	4.1	329.2	922
	$I$ (cm <sup>4</sup> )	473	$233 \times 10^3$	$1.25 \times 10^6$	305	$3.34 \times 10^6$	$15.5 \times 10^6$
	$W_{pl}$ (cm <sup>3</sup> )	84.2	6110	$28.0 \times 10^3$	49.4	$44.9 \times 10^3$	$167 \times 10^3$

Although arbitrary cross-section property combinations could have been chosen for  $I$ ,  $A_z$  and  $W_{pl}$ , using cross-section properties from an explicitly defined set ensures the predicted cross-section properties are physically realistic. Initially, the standardised UB cross-sections from BS EN 10365 [57] were considered. However, the minimum and maximum cross-section properties from this set were not sufficient for the lightest and heaviest loading conditions possible under the design constraints set by table 1. For this reason, a set of custom I-sections were generated and used exclusively for all members.

These custom I-sections were generated by averaging the geometrical ratios between the web depth  $d_w$ , flange thickness  $t_f$ , flange breadth  $b_f$  and the web thickness  $t_w$  that arise in BS EN 10365 [57]. Aside from ensuring that they share commonalities with the UB BS EN 10365, this process also ensured at minimum Class 2 sections [51] to allow the use of plastic cross-section properties. 1000 individual cross-sections were generated that ensured equal spacing across these ratios. The resulting granularity (as opposed to the 91 within BS EN 10365) meant that the utilisation ratio precision achievable during data-generation was significantly higher. The custom I-sections and associated cross-section properties are shown in table 2.

Together, these efforts ensure that the dataset on which the neural network is trained on covers sufficient breadth in terms of the input and output space to generalise for a wide variety of continuous beam systems. The dataset generated based on the aforementioned design constraints, the concept of influence zones, and the technique of zero-padding were chosen with the aim to maximise the generalisability of the inverse operator for any system size  $m$ , UDLs  $\omega$  and spans  $\mathbf{L}$ . This leaves only the utilisation ratios  $\mathbf{u}$  as the remaining input variable in equation (11). Instead of passing utilisation ratios as explicit inputs to the network, it was decided that the dataset will be generated so that all beams closely correspond to the target utilisation ratio  $u_{target}$ . The network will therefore implicitly learn the  $u_{target}$  from the data itself.

The dataset was generated by designing continuous beam systems of size  $m = 2k_{max} + 1$  with each member having a span  $L$  and UDL  $\omega$  value drawn from a random uniform distribution based on the discretised ranges and intervals specified in table 1. These heterogeneous structural systems were modelled and optimised using third-party software (Rhino3D<sup>®</sup>, Grasshopper<sup>®</sup> and Karamba3D<sup>®</sup> [58]) after having identified the influence zone  $k_{max}$  for the



**Figure 6.** Generalised neural network structure with known priors from the influence zone  $k_{max}$  as the input layer  $\mathbf{x}_0$  and cross-section properties of the beam as the output layer  $\mathbf{x}_D$ .

design constraints in tables 1 and 2. The beams were optimised for minimum depth against ULS cross-section checks from EN 1993-1-1 6.2 [51] using a coupled analysis and design procedure [50] with a target utilisation ratio of  $u_{target} = 0.99$ .

### 3.5. Neural network training procedure

The generalised neural network structure developed in this work is shown in figure 6. Identifying an appropriate architecture in terms of height  $H$ , depth  $D$  and activation functions  $a_d$  requires experimentation. The choice of loss function  $J$  to compare predicted targets  $\hat{\mathbf{x}}_D$  against true targets  $\mathbf{x}_D$  also form part of the experimentation process.

**3.5.1. Loss functions and performance metrics.** In this study, four different loss functions were investigated as shown in table 3. These include the mean absolute error (MAE) and mean square error (MSE) loss functions that are commonly used for regression models. One limitation associated with both is that their derivatives (in respect to predicted targets) back-propagate the model parameters  $\theta$  with no regards what the relative size of the error is in relation to the magnitude of the output variables  $I, A_z$  and  $W_{pl}$ .

This is problematic given the orders of magnitude difference between the largest and smallest section properties of the custom I-sections as shown in table 2. An error of  $100\text{ cm}^4$  for  $I$  would cause the same back propagation adjustment using MAE or MSE regardless if the true second moment of area value target is  $305\text{ cm}^4$  or  $305 \times 10^5\text{ cm}^4$ . As a consequence, both

**Table 3.** Loss functions to be tested with  $p$  predicted targets  $\widehat{x}_D$ ,  $p$  true targets  $x_D$  and small  $\epsilon$  to avoid division by zero errors.

Name	Formula	Derivative
Mean absolute error	$J_{MAE} = \frac{1}{p} \sum_{i=1}^p  \widehat{x}_{D,i} - x_{D,i} $	$J_{MAE}' = \frac{1}{p} \sum_{i=1}^p \begin{cases} -1, & \widehat{x}_{D,i} < x_{D,i} \\ 1, & \widehat{x}_{D,i} > x_{D,i} \end{cases}$
Mean squared error	$J_{MSE} = \frac{1}{2p} \sum_{i=1}^p (\widehat{x}_{D,i} - x_{D,i})^2$	$J_{MSE}' = \frac{1}{p} \sum_{i=1}^p \widehat{x}_{D,i} - x_{D,i}$
Mean absolute percentage error	$J_{MAPE} = \frac{1}{p} \sum_{i=1}^p \left  \frac{\widehat{x}_{D,i} - x_{D,i}}{x_{D,i} + \epsilon} \right $	$J_{MAPE}' = \frac{1}{p} \sum_{i=1}^p \begin{cases} \frac{-1}{ x_{D,i} + \epsilon }, & \widehat{x}_{D,i} < x_{D,i} \\ \frac{1}{ x_{D,i} + \epsilon }, & \widehat{x}_{D,i} > x_{D,i} \end{cases}$
Mean squared percentage error	$J_{MSPE} = \frac{1}{2p} \sum_{i=1}^p \frac{(\widehat{x}_{D,i} - x_{D,i})^2}{ x_{D,i} + \epsilon }$	$J_{MSPE}' = \frac{1}{p} \sum_{i=1}^p \frac{\widehat{x}_{D,i} - x_{D,i}}{ x_{D,i} + \epsilon }$

MAE and MSE would prioritise minimising the absolute error, which mathematically favours target values of large magnitudes at the expense of smaller ones.

To address the above mentioned issue, percentage-based versions of both MAE and MSE were tested, defined in table 3 as the mean absolute percentage error (MAPE) and the mean squared percentage error (MSPE). Whilst MAPE is commonly used, MSPE is not tested in practice. Both MAPE and MSPE ensure that during back-propagation, the optimiser updates model parameters in proportion to the relative deviation between predicted  $\widehat{\mathbf{x}}_D$  and true outputs  $\mathbf{x}_D$ , which should be a better performance criterion to address the orders of magnitude difference in the output space that arise in these particular continuous beam systems.

Regardless of the choice of loss function, MAPE will be used as a comparison metric between different networks. However to study the dispersion of prediction errors, an accuracy metric  $M$  will also be evaluated with minimum, 0.5%, 2.5%, 50% (median), 97.5%, 99.5% and maximum percentile values. This will allow the evaluation of the 95% and 99% confidence intervals (CI) and help identify the range of over and under prediction of outputs, which is important in the context of safe structural design:

$$M(\widehat{\mathbf{x}}_D, \mathbf{x}_D) = \frac{1}{p} \sum_{i=1}^p \frac{\widehat{x}_{D,i}}{x_{D,i}} = \frac{1}{3} \left( \frac{\widehat{I}}{I} + \frac{\widehat{A}_z}{A_z} + \frac{\widehat{W}_{pl}}{W_{pl}} \right). \quad (15)$$

**3.5.2. Activation functions.** The activation functions tested in this work are listed in table 4, and includes the commonly used rectified linear unit (ReLU) function amongst others [56]. A distinction is drawn between the inner activation functions  $a_{in}$  within the hidden layers, and the outer activation function  $a_{out}$ , that evaluate the target values  $\mathbf{x}_D$ . All inputs and outputs were scaled between 0 and 1 by dividing the values by the maximum magnitude of the features and targets within the training and validation set, respectively. Therefore, it is important to choose only output activation functions compatible with the scaled values of the targets as reflected by the range of  $a_{out}$  functions listed in table 4.

**3.5.3. Height and depth analysis.** The appropriate size of a neural network in terms of height  $H$  and depth  $D$  was found by finding a suitable trade-off between under and over fitting the model parameter space. The design complexity of continuous beam systems will likely be



**Table 4.** Table of inner  $a_{in}$  and outer  $a_{out}$  activation functions to be tested with weight vector  $\mathbf{w}$ , bias vector  $\mathbf{b}$  and layer vector  $\mathbf{x}$ .

Name	Formula	$a_{in}$	$a_{out}$
Rectified linear unit (ReLU)	$a_{ReLU} = \max(\mathbf{w}\mathbf{x} + \mathbf{b}, 0)$	✓	✓
Sigmoid	$a_{sigm} = 1/(1 + e^{\mathbf{w}\mathbf{x} + \mathbf{b}})$	✓	✓
Hyperbolic tangent	$a_{tanh} = \tanh(\mathbf{w}\mathbf{x} + \mathbf{b})$	✓	
Exponential	$a_{exp} = e^{\mathbf{w}\mathbf{x} + \mathbf{b}}$		✓

**Table 5.** Options and/or ranges of neural network learning parameters and hyperparameters tested, along with selected parameters for all training runs presented in results.

Neural network training aspect	Options/range considered	Selected
Optimiser	SGD, RMSprop, Adam, Nadam	Nadam
Learning rate	$\alpha = [0.0001, 1.0000]$	$\alpha = 0.0005$
Initialiser	Gaussian with $\mu = 0, \sigma = [0.00, 1.00]$	$\sigma = 0.05$
Batch-size	$[128, 8192]$	1024

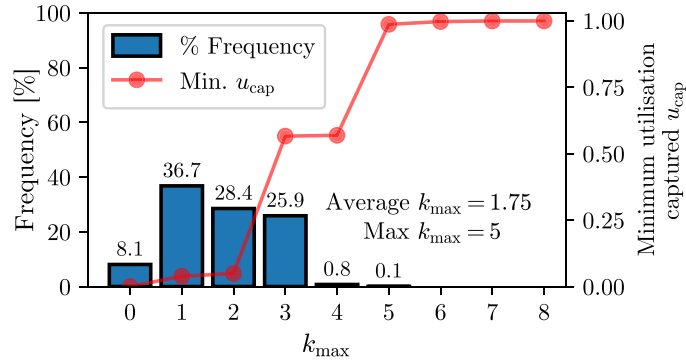
reflected in deeper and wider neural networks than those considered in previous literature [31] due to the large number of load arrangements that may be critical, the numerous design criteria that govern the design, and the variety of viable cross-sections. For this reason, a wide range of heights and depths were tested. The size of the networks were denoted by a simple syntax based on the architecture of the hidden layers. For example, ‘50–50–50’ refers to a neural network with three hidden layers with 50 nodes each.

**3.5.4. Other neural network parameters and hyperparameters.** Given the large dataset size and computational resources required for training, a simple hold-out strategy was deemed appropriate as opposed to other validation strategies [59], and hence the final dataset was randomised and split into training, validation and testing sets using a 70%, 15%, 15% split, respectively. The testing set was only used once after an appropriate neural network architecture was found experimentally. A robustness check with various initialiser seeds was carried out on the final architecture. Other neural network training aspects, such as optimisers, types of initialisers, learning rates and batch-sizes were chosen empirically based on MAPE performance and qualitative comparison of learning behaviour. The options/ranges for each of these are summarised in table 5. All stochastic elements were controlled through explicit initialiser seeds.

### 3.6. Summary of methodology

The following procedure was adopted to develop the machine learned inverse operator:

1. Evaluate the maximum influence zone  $k_{max}$  for the continuous beam system using the procedure from [37] based on the design constraints specified in tables 1 and 2.
2. Design continuous beam systems of size  $m = 2k_{max} + 1$  using a coupled analysis and design approach [50] with a target utilisation ratio  $u_{target} = 0.99$ . Each beam within the continuous beam system will correspond with one data point, with zero-padding for edge or near-edge beams as shown in figure 5. Finally, split and normalise the data into a training, validation and testing set as explained in section 3.5.4.



**Figure 7.** Influence zone results for a  $m = 17$  system with  $\epsilon_{\max} = 0.02$  based on the design constraints established by tables 1 and 2 using the methodology from [37].

3. Develop the neural network model using the following steps:
  - 3.1. Assume a standard 50–50 architecture and test out the various combinations of loss and activation functions as identified tables 3 and 4, respectively, based on 100k training data points.
  - 3.2. Test various height  $H$  and depth  $D$  variations as explained in section 3.5.3 based on 100k training data points.
  - 3.3. For the best architecture (height, depth and activation function), test the performance against different training set sizes.
4. Evaluate the performance of the final neural network against the testing dataset and conduct a robustness test using various initialiser seeds for the weights and biases.

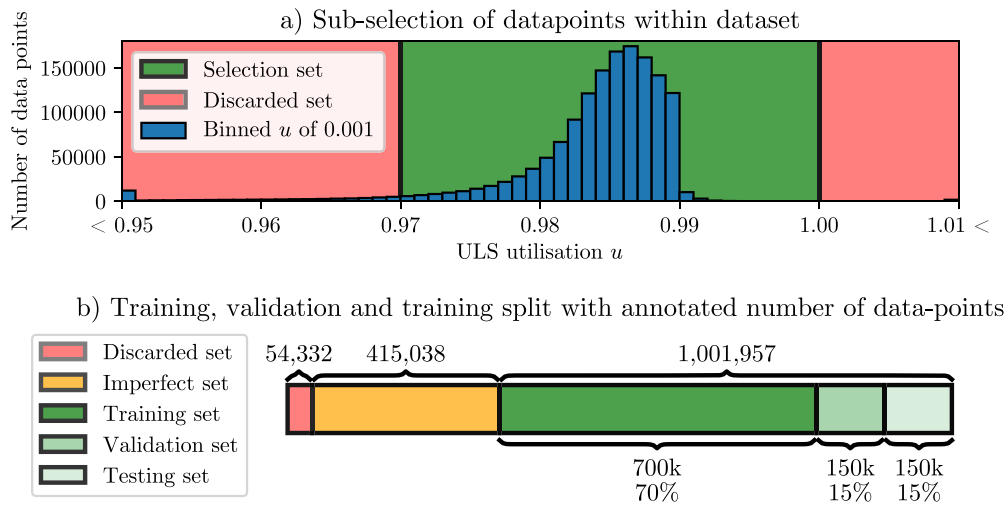
## 4. Model development and results

### 4.1. Influence zone size estimation

The maximum influence zone  $k_{\max}$  of continuous beam systems subject to design constraints specified by tables 1 and 2 was established. Using the procedure from [37], 25 random UDL and span distributions were generated for a  $m = 17$  sized system and designed against ULS checks from EN 1993-1-1 [51] using the custom I-sections specified in table 2 with a target utilisation ratio  $u_{\text{target}} = 0.99$ . This led to the creation of 10 625 continuous beams ( $25 \times 25 \times 17$ ). Each beam's influence zone value was evaluated using an error threshold of  $\epsilon_{\max} = 0.02$ . This threshold was selected based on the expected MAPE performance achievable with the multi-layered neural network, with the results shown in figure 7. The results indicate that the average and maximum influence zone size is  $k_{\max} = 1.75$  and  $k_{\max} = 5$ , respectively. This suggests the system size required for the dataset generation is  $m = 2k_{\max} + 1 = 11$ , and the required input layer size is  $4k_{\max} + 2 = 22$ . This influence zone evaluation took 5 hours of computation time.

### 4.2. Data generation, visualisation and pre-processing

Drawing from uniform distributions for spans and UDL values identified in table 1, two datasets were created. The first consists out of 266 unique UDL  $\omega$  and span  $\mathbf{L}$  permutations



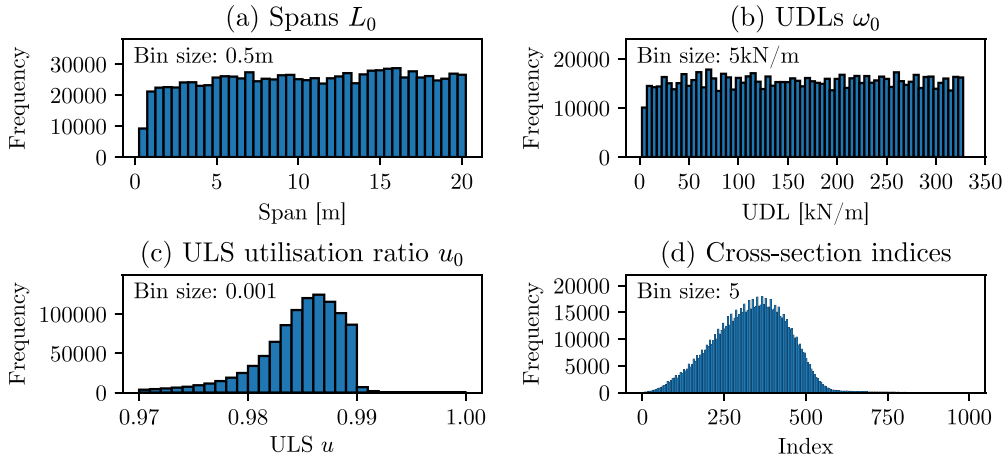
**Figure 8.** Sub-selection of data points from the initial 1471 327 dataset based on a ULS utilisation ratio  $u$  range of 0.97–1.00. The 1000k CBeamXP dataset was drawn from the green sets.

of  $m = 11$  sized continuous beam systems, and the second out of 251 unique permutations. A coupled analysis and design optimisation approach [50] with a target utilisation ratio  $u_{\text{target}} = 0.99$  based on ULS cross-section checks [51] and all critical load arrangements [37] was implemented to find the appropriate custom I-section from table 2 for each beam within the system. This process resulted in 1471 327 individual data-points ( $11 \times (266^2 + 251^2)$ ) that took 3.5 days to generate.

The distribution of utilisation ratios achieved for the specified target utilisation ratio  $u_{\text{target}} = 0.99$  are shown in figure 8(a). Since the design space is limited to discretized cross-section properties the utilisation target ratio  $u_{\text{target}} = 0.99$  was rarely met exactly. Therefore, a sub-selection of this dataset took place, discarding all of the beams that fell outside of utilisation ratio range  $0.97 \leq u < 1.00$ .

Although only 54 322 data-points belonged to the discarded set, the dataset was further stripped of all data points which had beam members within their system that belonged to the discarded set, even if those beams themselves fell within the selected utilisation ratio range. These data points are defined as the ‘Imperfect set’ in figure 8(a). This process removed another 415 038 data-points. This left 1001 957 data points ( $1471\ 327 - 415\ 038 - 54\ 322$ ), each representing a beam within a  $m = 11$  member system and the surrounding design information from the influence zone of valid structural designs under ULS conditions. This set was randomised, further stripped of another 1957 data points, to yield a dataset size of exactly 1 million (1000k).

This 1000k dataset was named CBeamXP: continuous beam cross-section predictors [60] and represents ULS compliant beam systems of system size  $m = 11$  with utilisation ratios between  $0.97 \leq u < 1.00$ . The CBeamXP dataset was split into a training, validation and testing set using a 70%, 15%, 15% split as shown in figure 8(b). Histograms of the spans, UDLs, utilisation ratios and cross-section indices (corresponding to one of the 1000 custom I-sections in ascending stiffness order) are shown in figure 9. For pre-processing, all inputs and outputs were divided by the maximum value within the training and validation set.



**Figure 9.** Frequency distributions for various descriptor variables of the CBeamXP dataset. Spans and UDL values are uniformly distributed, whilst the selected cross-section indices of the optimised beam systems follow a normal distribution.

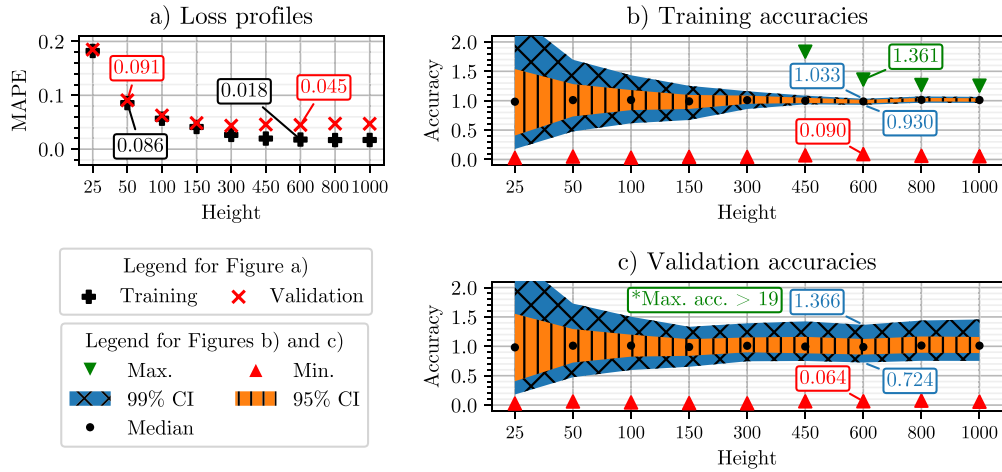
**Table 6.** Validation MAPE metrics at epoch 1000 for different combinations of loss  $J$ , inner  $a_{in}$  and outer  $a_{out}$  activation functions for a 50–50 architecture using 100k training and 150k validation data points. MAPE values of less than 0.100 (10%) are in bold.

	$J_{MAE}$			$J_{MSE}$		
	$a_{out,ReLU}$	$a_{out,sigm}$	$a_{out,exp}$	$a_{out,ReLU}$	$a_{out,sigm}$	$a_{out,exp}$
$a_{in,ReLU}$	0.105	0.164	0.143	0.116	0.168	0.146
$a_{in,sigm}$	0.558	0.158	0.168	0.568	0.250	0.243
$a_{in,tanh}$	0.205	0.143	0.133	0.274	0.160	0.147
	$J_{MAPE}$			$J_{MSPE}$		
	$a_{out,ReLU}$	$a_{out,sigm}$	$a_{out,exp}$	$a_{out,ReLU}$	$a_{out,sigm}$	$a_{out,exp}$
$a_{in,ReLU}$	<b>0.072</b>	<b>0.090</b>	<b>0.091</b>	<b>0.089</b>	<b>0.093</b>	<b>0.089</b>
$a_{in,sigm}$	0.707	0.129	0.124	0.707	0.147	0.137
$a_{in,tanh}$	0.205	<b>0.083</b>	<b>0.083</b>	0.235	<b>0.082</b>	<b>0.087</b>

#### 4.3. Network architecture development results

**4.3.1. Loss and activation function variations.** The neural network development began by evaluating the MAPE performance of 50–50 architectures for different loss functions  $J$ , inner  $a_{in}$  and outer  $a_{out}$  activation functions. These networks were trained for 1000 epochs using 100k datapoints from the training set, yet validated against the entire 150k validation set. The total training time was 5 h with results shown in table 6.

The results clearly indicate that the percentage-based loss functions  $J_{MAPE}$  and  $J_{MSPE}$  typically outperform their non-percentage-based counter-parts. Performance between either  $J_{MAPE}$  and  $J_{MSPE}$  was relatively similar, with all MAPE values of less than 10% (0.100) bolded.  $J_{MAPE}$  was chosen as the loss function for this investigation due to it being more commonly used. The five best inner and outer activation function combinations in table 6 under  $J_{MAPE}$  (bolded values) were subsequently qualitatively analysed.

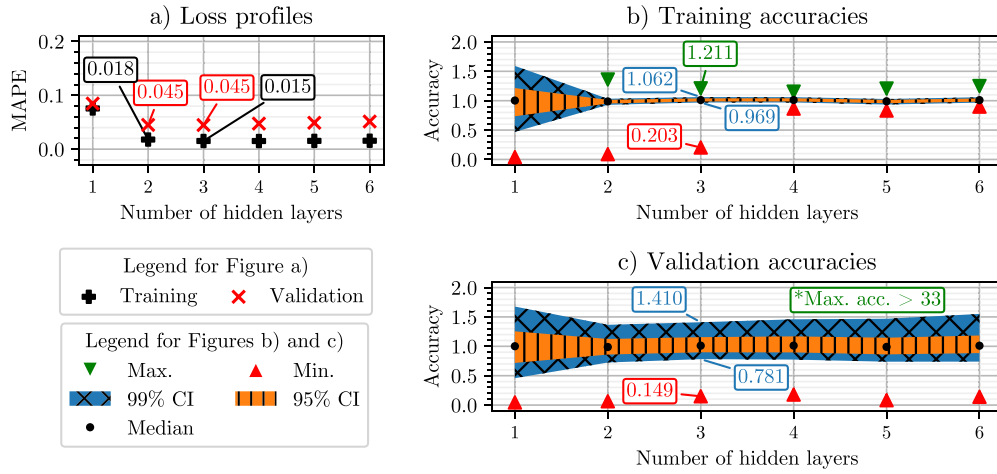


**Figure 10.** Loss and accuracy profiles for  $a_{in,ReLU}$  and  $a_{out,exp}$  networks at epoch 1000 with  $J_{MAPE}$  with two hidden layers of equal height. Training set size of 100k and validation set size of 150k.

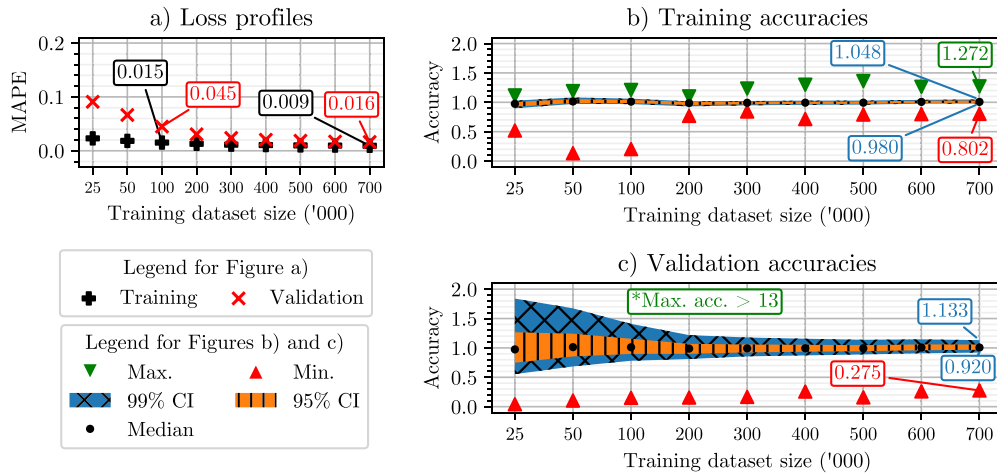
This qualitative analysis highlighted that using ReLU as an outer activation function allows the prediction of null cross-section properties. This is an invalid prediction since the model operates on the basis that a beam with some minimum cross-section properties must exist in the context of this structural system. A similar limitation applied for the sigmoid activation function which asymptotically approaches the value of one at positive infinity. This limits the network's ability of predicting cross-sections larger than those found within the training and validation dataset. For these reasons, the  $a_{out,exp}$  function was selected for this particular network architecture since  $a_{out,exp}$  does not result in zero-valued cross-section properties and also does not impose an upper limit on the outputs. From the remaining viable networks, the  $a_{in,ReLU}$  and  $a_{out,exp}$  architecture converged the quickest and was therefore chosen for further development in this study.

**4.3.2. Height and depth variations.** The architecture of the hidden layers (height  $H$  and depth  $D$ ) needs to be sufficiently expressive to reflect the design complexity of continuous beam systems, and need to avoid under- and over-fitting the model. Therefore, comparison between training and validation performance is needed. The 100k training and 150k validation sets from section 4.3.1 were re-used for this purpose. Figure 10 compares the performance of various networks containing two hidden layers of varying heights at epoch 1000. The combined training time of these networks was 22 h. The 600–600 network was identified as the point at which the performance transitioned from under-fitting to slight over-fitting. Figures 10(b) and (c) further indicate the accuracy profiles for both training and validation, respectively. Note that the maximum validation accuracy values greatly exceed the value of at least 19 (1900%) for all networks, regardless of height, meaning the network predicted cross-sectional properties 19 times larger than the target value.

The 'optimal' number of hidden layers for height 600 was investigated, with results shown in figure 11. This resulted in a combined training time of 12 h. Networks with more than three hidden layers showed no major improvements in either training or validation performance except in minimum training accuracy. However, this was not associated with an improvement



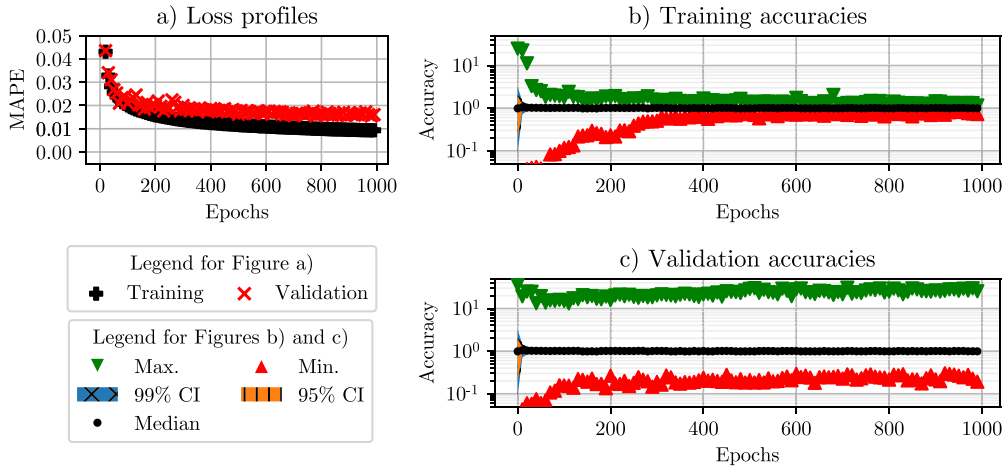
**Figure 11.** Loss and accuracy profiles for  $a_{in,ReLU}$  and  $a_{out,exp}$  networks at epoch 1000 with  $J_{MAPE}$  with hidden layers of height  $H = 600$ . Training set size of 100k and validation set size of 150k.



**Figure 12.** Loss and accuracy profiles for 600–600–600  $a_{in,ReLU}$  and  $a_{out,exp}$  network at epoch 1000 with  $J_{MAPE}$  for various training dataset sizes, with a validation set size of 150k.

in minimum validation accuracy as shown in figure 11(c). For these reasons, a depth of three hidden layers was deemed appropriate.

**4.3.3. Dataset size variations.** Figure 12 shows the change in performance as a function of the training dataset size, from 25k to 700k data points, with the same 150k validation dataset as in the previous sections. The combined training time was 1.5 d. Except for slight variations in the minimum and maximum accuracy values, the performance of the neural network naturally improved with a larger training set.



**Figure 13.** Loss and accuracy profiles for 600–600–600  $a_{in,ReLU}$  and  $a_{out,exp}$  network at epoch 1000 with  $J_{MAPE}$ , 700k training and 150k validation sets. Note use of logarithmic y-axis to show the full range of maximum validation accuracies.

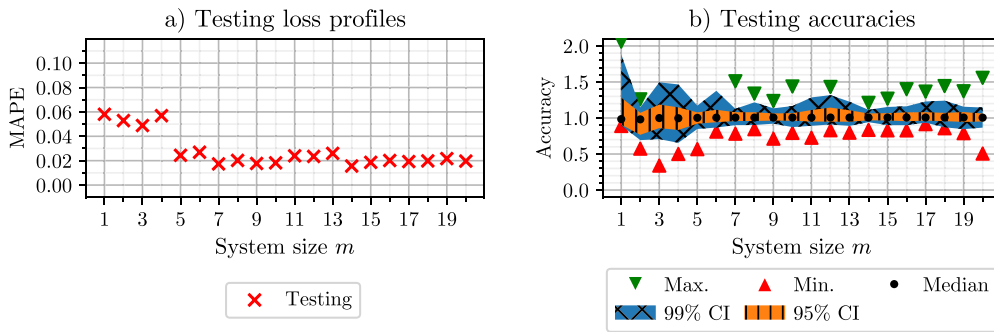
**Table 7.** Loss and accuracy profiles for 600–600–600  $a_{in,ReLU}$  and  $a_{out,exp}$  network at epoch 1000 with  $J_{MAPE}$ , 700k training, 150k validation sets and 150k testing set.

Dataset	Data type	MAPE	Accuracy percentiles						
			Min	0.5%	2.5%	Median	97.5%	99.5%	Max
Training	Results	0.009	0.802	0.980	0.990	1.007	1.030	1.048	1.272
	$\sigma_{initialiser}$	0.000	0.161	0.013	0.010	0.007	0.010	0.014	0.138
Validation	Results	0.016	0.275	0.920	0.968	1.006	1.058	1.133	26.811
	$\sigma_{initialiser}$	0.001	0.058	0.014	0.010	0.007	0.012	0.018	4.977
Testing	Results	0.016	0.313	0.917	0.967	1.006	1.058	1.138	12.282
	$\sigma_{initialiser}$	0.001	0.070	0.012	0.010	0.007	0.012	0.018	2.189

#### 4.4. Model performance: testing and robustness

The final neural network model consists of a 600–600–600 architecture with  $a_{in,ReLU}$  and  $a_{out,exp}$  activation functions trained using the  $J_{MAPE}$  loss function based on a training and validation dataset size of 700k and 150k data points, respectively, with learning graphs shown in figure 13. The neural network at epoch 1000 was also evaluated against the testing set created in figure 8, and checked for robustness by re-training the same network a further 9 times using different kernel initialiser seeds, which took 4.5 d. The general model performance results and standard deviations  $\sigma_{initialiser}$  due to these different initialiser seeds is summarised in table 7.

The similar performance between the testing and validation set in table 7 strongly suggests that the model is likely to generalise well to new data-points. The impact of changing initialiser seed is minimal except for the minimum and maximum accuracy values.



**Figure 14.** Generalisability performance of the final 600–600–600 neural network in terms of MAPE and accuracy for unseen structures of varying system size  $m$ .

## 5. Discussion

### 5.1. Model generalisability

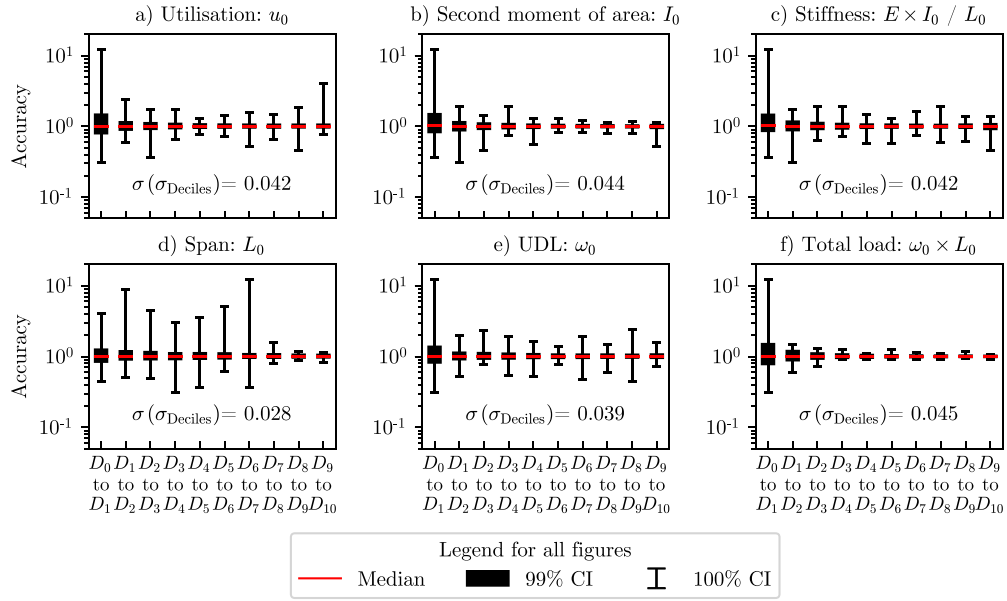
One of the fundamental objectives of this work was to develop a machine learned structural design model capable of generalising beyond the system size  $m = 11$  it was trained on. To achieve this, the influence zone concept was leveraged with zero-padding to theoretically allow the neural network to make localised predictions for continuous beams of arbitrary system size  $m$ . To test this, over 1000 additional testing data-points were generated using the same methodology as described in section 3.4 and sub-selection process as shown in section 4.2 for each system size  $1 \leq m \leq 20$  (including  $m = 11$ ). The MAPE and accuracy performance are shown in figure 14.

Figure 14 indicates that the machine learned inverse operator demonstrates strong generalisation capability for continuous beam system sizes  $m \geq 5$  with  $\text{MAPE} \approx 2\%$ . System sizes  $m < 5$  saw a slightly deteriorating MAPE values of 5%–6%. These are encouraging results given that the neural network was never trained on system sizes less or greater than  $m = 11$ . The greatest variations in performance were typically in the maximum and minimum accuracy values; in fact the model performed often better in terms of maximum performance for system size other than  $m = 11$ .

These results provide merit to the novel implementation of the influence zone concept [37] as a mechanics driven feature selector and using zero-padding to build machine learned inverse operators capable of generalising to differently sized continuous structural systems. This opens the possibility of investigating the applicability of this methodology for two or three dimensional frames. Furthermore, these results also provide a solution to the limitation of fixed-dimensional input vectors of multi-layer neural networks [38].

In recent years, other researchers have investigated the development of generalisable machine learning models; most of these efforts have focused on machine learned forward operators [38, 39, 61]. Within the realm of structural design inverse operators, researchers have noted that the question of generalisability remains typically under-investigated [32]. Whilst previous works studied the ability of neural networks to generalise under different boundary conditions [33, 34], this work distinguishes itself on generalising across differently sized





**Figure 15.** Custom box plots of accuracies vs. variables in binned deciles. Total load (f) correlates with the greatest dispersion  $\sigma(\sigma_{\text{Deciles}}) = 0.045$ , which can also be identified visually.

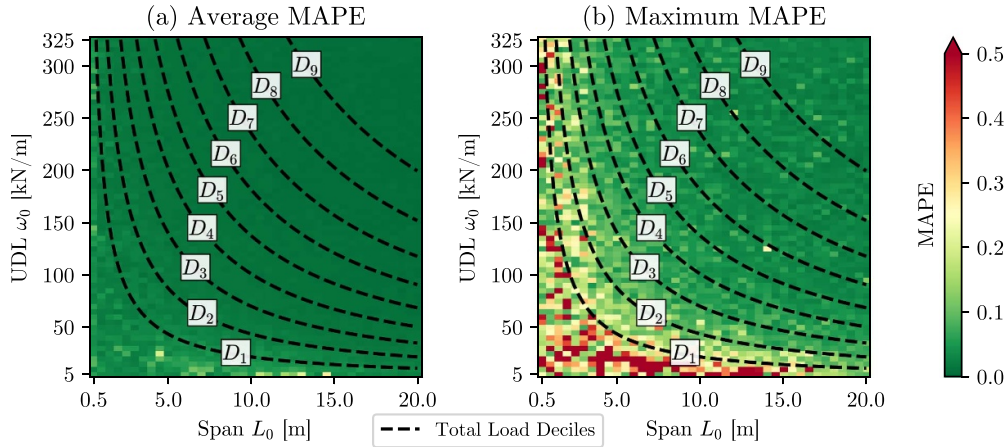
systems. Combining the underlying techniques behind these studies may allow one to train a generalisable model of arbitrary size and arbitrary boundary conditions.

## 5.2. Performance variability

This investigation also differentiated itself from previous works by measuring the variability of predictions in terms of accuracies. Notably, this allowed one to identify the range of over- and under-predictions, which are not captured by average loss function metrics such as MAE or MAPE. Despite gradual improvement within the 95% and 99% confidence intervals, the final performance graph in figure 13 indicates that the confidence intervals of the validation set lag those of the training set. The same can also be said for the testing set, especially for maximum and minimum accuracies as shown in table 7.

To identify potential causes of this divergence in performance between the training and testing set, custom box plots of testing accuracies were generated for a number of variables that describe the dataset, ordered based on ascending deciles ( $D_0$ – $D_{10}$ ). By evaluating the standard deviation of each decile’s accuracy values, and taking the standard deviation of those standard deviations  $\sigma(\sigma_{\text{Deciles}})$ , one can quantify numerically which variable causes the greatest dispersion of the accuracy values. These results are shown in figure 15.

By studying figure 15 in detail, it was identified that the total load variable  $\omega_0 \times L_0$  caused the greatest  $\sigma(\sigma_{\text{Deciles}})$  dispersion as seen in figure 15(f). Figure 15(f) also showed the most identifiable demarcation between low and high accuracy results. The prediction variability of cross-section properties of a beam is the worst when the combined product of both the UDL



**Figure 16.** Total load heatmaps which evaluates the (a) average MAPE and (b) maximum MAPE values for all UDL  $\omega_0$  and span  $L_0$  combinations. Notice how maximum MAPE errors occur at low total load combinations ( $\omega_0 \times L_0$ ) within the first decile up to  $D_1$ .

load  $\omega_0$  and span  $L_0$  fell in the lowest Decile ( $<D_1$ ). This pattern can also be identified by studying heat-maps of the average and maximum MAPE performance which occurred at each  $\omega_0$  and span  $L_0$  combination within the dataset as shown in figure 16.

Using structural engineering intuition, one infers that the design of short and lightly loaded spans is more likely to be influenced by the UDLs of the surrounding members within a continuous system. Whilst the influence zone concept ensures the pertinent design information is contained within the inputs, providing that information solely in the form of an input vector may not be sufficient to make accurate predictions under all circumstances. The fact that there were also wide prediction variabilities for the smallest deciles for both the second-moment of area and stiffness values as shown in figures 15(b) and (c) suggests that exposing the machine learning model to additional physics knowledge (other than influence zones) of the structural system may lead to further improvements.

In one of the early studies, Berke *et al* [28] noted that despite achieving relatively low prediction errors on average, neural network predictions can occasionally vary significantly. In more recent works, the presence of large error predictions (over  $>40\%$ ) were noted and manually removed from the final reported average prediction error metric [33]. The results from this study highlight that this variability issue needs to be further addressed. So far, the authors have identified only a single study that investigated error variability when evaluating machine learning performance for civil and structural engineering applications [31]. The use of the accuracy metric along with its minimum, maximum, 95% and 99% metrics could provide a framework to study error variability in more detail.

### 5.3. Other neural network performance observations

Despite using a simple multi-layer neural network for the structural design inverse operator  $O_{\text{inv}}^\dagger$ , the development procedure successfully lowered the validation error from MAPE values of  $\approx 10\%$  in table 6 to 1.6% in table 7, a performance that was matched by the testing dataset as well. This was attributable to numerous factors, a notable one being the use of the percentage

based loss function, which as anticipated in section 3.5.1, was more suitable for the dataset given the orders of magnitude differences in the targets. The use of the exponential output activation function  $a_{\text{exp}}$  may also have positively contributed to dealing with target values that vary greatly in magnitude.

The lack of literature on machine learned structural design models for continuous beam systems means that a direct comparison of the 1.6% MAPE performance is not possible at present. However, one can compare this performance with performance metrics of structural design models developed for different applications. For example, the network developed in this work outperformed previous multi-layer neural network regression models; an early concrete beam prediction model achieved a MAPE value of 10.17% [27], whilst a cross-section predictor of aerospace components averaged out at a MAPE value of 5% [28]. The network presented in this study also performed well when compared to more advanced network architectures such as convolutional neural networks for topologically optimised truss structures that achieved voxel value errors of 5.63% [33]. Comparison with further works that developed machine learned structural design models was not possible for studies which reported performance with non-percentage based metrics such as MAE [29] or MSE [31, 32, 47].

This study also differentiates itself by the quantity of data it was trained on (up to 700 000 data-points), which based on figure 12 helped improve validation performance. Early works from the 1990s had training set sizes smaller than 100 data-points [27–29], and even more recent literature only trained using 600 [31], 12 000 [32] 28 000 [33] or just under 40 000 [34] data-points. Whilst large datasets significantly increase computational cost, the combination of big data and more advanced neural network architectures may improve performance further, both in terms of average error and prediction variability.

#### 5.4. Limitations and scope for future works

There are multiple limitations that restrict practical use of the proposed design model. The first is the fact that the structural systems within the dataset were designed against ULS constraints only, and made other assumptions on the nature of the design problem listed in section 2.2. The generalisability of the model, specifically for system sizes  $m < 5$  also requires further work, and the issue of prediction variability will also require additional investigation in terms of either model architectures or generating larger datasets. Furthermore, there likely exist a wide range of mathematical techniques from inverse problems that could aid in developing and assessing operators for structural design problems, by for example estimating the Lipschitz coefficient of the mapping. These limitations provide a clear basis for further works in the future.

On another note, table 8 summarises the total computation time required for the entirety of the results section. Whilst the computation time could have been accelerated through parallelisation, improved computation resources and simplification of metric evaluations algorithms, the purpose of table 8 is to indicate the relative proportion of time spent at each stage. Greater computational resources may allow investigations using alternative validation strategies such as  $k$ -fold cross-validation [59] and automated hyperparameter selection procedures [62] that could result in improved performance. A significant portion of the computation effort was spent simply generating the data-points for training, validation and testing. In light of encouraging reproducibility studies [63] and to encourage research that improves the predictive capability of the machine learned structural design model presented here, the CBeamXP dataset along with an associated python-based neural network training script are made available at an open-source data repository [60].

**Table 8.** Computation time for each neural network development stage.

section	Stage	Computation time		Proportion [%]
		In hours (h)	In days (D)	
Section 4.1	Influence zone evaluation	5	0.2	33
Section 4.2	Data generation and pre-processing	84	3.5	
Section 4.3.1	Loss and activation function study	5	0.2	27
Section 4.3.2	Height and depth study	34	1.4	
Section 4.3.3	Dataset size study	36	1.5	
Section 4.4	Testing and robustness study	108	4.5	40
Total computation time:		272	11.3	100

## 6. Conclusions

This work developed a new neural network based structural design model to predict cross-section property requirements of continuous beam systems non-iteratively. The major contributions of this investigation include:

- Framing structural design as an inverse problem, and using this novel perspective to identify three distinct types of machine learning applications. One of these types, machine learned inverse operators, were investigated in this work to develop a non-iterative structural design model. This presents a fundamental shift from traditional design approaches.
- Developing a non-iterative structural design model for continuous beam systems of arbitrary member size through the novel use of influence zones [37] to provide a mechanics-driven feature selection process that enhanced the model's generalisability.
- Achieving a mean absolute percentage error of 1.6% which was lower than machine learned structural design models from comparative literature. This performance was attributable to the careful consideration of the network architecture in terms of height and depth of the hidden layers, the selection of loss and activation functions that were appropriate to address the challenges posed by continuous beam system, and a dataset size of 700 000 data points.
- Identifying the importance of measuring and reducing prediction error variability. In this study the 99% confidence interval for testing accuracy was between 91.7% and 113.8%. Reducing prediction variability is a significant knowledge gap in literature, especially in regards to machine learning applications within safety critical systems such as structural design.

The CBeamXP dataset generated in this work containing one million data-points along with an associated python-based neural network training script were published at an open-source data repository [60]. Aside from allowing results to be reproduced, sharing this data will hopefully encourage future research towards machine learned structural design models that improve the mean absolute percentage error, generalisability, or prediction variability achieved in this investigation.

## Data availability statement

The data that support the findings of this study as well as a neural network training script are openly available at the following URL/DOI: <https://doi.org/10.15131/shef.data.23945562>.

## Conflict of interest

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

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article. The authors gratefully acknowledge the support of the Engineering Physical Sciences Research Council's (EPSRC) Doctoral Training Partnership Studentship and the Ramboll Foundation. For the purpose of open access and funding stipulations, the authors have applied a Creative Commons Attribution (CC-BY) licence to any Accepted Manuscript (AM) version arising.


## ORCID authorship contribution statement

**Adrien Gallet:** Conceptualisation, Methodology, Software, Investigation, Formal analysis, Validation, Data curation, Visualisation, Writing—Original Draft, Project administration. **Andrew Liew:** Writing—Review & Editing. **Iman Hajirasouliha:** Validation, Writing—Review & Editing. **Danny Smyl:** Supervision, Validation, Writing—Review & Editing.

## ORCID iDs

Adrien Gallet  <https://orcid.org/0000-0003-4939-9916>

Andrew Liew  <https://orcid.org/0000-0001-8229-9605>

Iman Hajirasouliha  <https://orcid.org/0000-0003-2597-8200>

Danny Smyl  <https://orcid.org/0000-0002-6730-5277>

## References

- [1] Gallet A, Rigby S, Tallman T N, Kong X, Hajirasouliha I, Liew A, Liu D, Chen L, Hauptmann A and Smyl D 2022 Structural engineering from an inverse problems perspective *Proc. R. Soc. A* **478** 20210526
- [2] Alifanov O M 1983 Methods of solving ill-posed inverse problems *J. Eng. Phys.* **45** 1237–45
- [3] Friswell M I 2007 Damage identification using inverse methods *Phil. Trans. R. Soc. A* **365** 393–410
- [4] Smyl D, Pour-Ghaz M and Seppänen A 2018 Detection and reconstruction of complex structural cracking patterns with electrical imaging *NDT & E Int.* **99** 123–33
- [5] Tallman T N and Smyl D J 2020 Structural health and condition monitoring via electrical impedance tomography in self-sensing materials: a review *Smart Mater. Struct.* **29** 123001
- [6] Zhao L, Tallman T and Lin G 2021 Spatial damage characterization in self-sensing materials via neural network-aided electrical impedance tomography: a computational study *ES Mater. Manuf.* **12** 78–88

- [7] Rigby S E, Lodge T J, Alotaibi S, Barr A D, Clarke S D, Langdon G S and Tyas A 2020 Preliminary yield estimation of the 2020 Beirut explosion using video footage from social media *Shock Waves* **30** 671–5
- [8] van der Voort M, van Wees R, Brouwer S, van der Jagt-Deutekom M and Verreault J 2015 Forensic analysis of explosions: inverse calculation of the charge mass *Forensic Sci. Int.* **252** 11–21
- [9] Mueller J L and Siltanen S 2012 *Linear and Nonlinear Inverse Problems With Practical Applications* (Society for Industrial and Applied Mathematics)
- [10] Mason J and Manning M 2011 *Structural Design: The Engineer's Role* (Institution of Structural Engineers)
- [11] Pollini N 2021 Gradient-based prestress and size optimization for the design of cable domes *Int. J. Solids Struct.* **222–223** 111028
- [12] Altay O, Cetindemir O and Aydogdu I 2023 Size optimization of planar truss systems using the modified salp swarm algorithm *Eng. Optim.* **1–17**
- [13] Daxini S D and Prajapati J M 2017 Parametric shape optimization techniques based on Meshless methods: a review *Struct. Multidiscipl. Optim.* **56** 1197–214
- [14] Upadhyay B D, Sonigra S S and Daxini S D 2021 Numerical analysis perspective in structural shape optimization: a review post 2000 *Adv. Eng. Softw.* **155** 102992
- [15] Hajirasouliha I, Pilakoutas K and Moghaddam H 2011 Topology optimization for the seismic design of truss-like structures *Comput. Struct.* **89** 702–11
- [16] Sigmund O and Maute K 2013 Topology optimization approaches: a comparative review *Struct. Multidiscipl. Optim.* **48** 1031–55
- [17] Weldeyesus A G, Gondzio J, He L, Gilbert M, Shepherd P and Tyas A 2019 Adaptive solution of truss layout optimization problems with global stability constraints *Struct. Multidiscipl. Optim.* **60** 2093–111
- [18] Fairclough H E and Gilbert M 2022 Layout optimization of long-span structures subject to self-weight and multiple load-cases *Struct. Multidiscipl. Optim.* **65** 197
- [19] Zaheer Q, Manzoor M M and Ahamad M J 2023 A review on developing optimization techniques in civil engineering *Eng. Comput.* **40** 348–77
- [20] He L and Gilbert M 2015 Rationalization of trusses generated via layout optimization *Struct. Multidiscipl. Optim.* **52** 677–94
- [21] Coelho R F, Tysmans T and Verwimp E 2014 Form finding & structural optimization: A project-based course for graduate students in civil and architectural engineering *Struct. Multidiscipl. Optim.* **49** 1037–46
- [22] Málaga-Chuquitaype C 2022 Machine learning in structural design: an opinionated review *Front. Built Environ.* **8** 815717
- [23] Arridge S, Maass P, Öktem O and Schönlieb C-B 2019 Solving inverse problems using data-driven models *Acta Numer.* **28** 1–174
- [24] Chen L, Gallet A, Huang S-S, Liu D and Smyl D 2022 Probabilistic cracking prediction via deep learned electrical tomography *Struct. Health Monit.* **21** 1574–89
- [25] Hamilton S J and Hauptmann A 2018 Deep D-bar: real-time electrical impedance tomography imaging with deep neural networks *IEEE Trans. Med. Imaging* **37** 2367–77
- [26] Adeli H and Yeh C 1989 Perceptron learning in engineering design *Comput.-Aided Civil Infrastruct. Eng.* **4** 247–56
- [27] Vanluchene R D and Sun R 1990 Neural networks in structural engineering *Comput.-Aided Civil Infrastruct. Eng.* **5** 207–15
- [28] Berke L, Patnaik S and Murthy P 1993 Optimum design of aerospace structural components using neural networks *Comput. Struct.* **48** 1001–10
- [29] Kang H-T and Yoon C J 1994 Neural network approaches to aid simple truss design problems *Comput.-Aid. Civil Infrastruct. Eng.* **9** 211–8
- [30] Hanna S 2007 Inductive machine learning of optimal modular structures: Estimating solutions using support vector machines *Artif. Intell. Eng. Des. Anal.* **21** 351–66
- [31] Tseranidis S, Brown N C and Mueller C T 2016 Data-driven approximation algorithms for rapid performance evaluation and optimization of civil structures *Autom. Constr.* **72** 279–93
- [32] Behzadi M M and Iliş H T 2021 Real-time topology optimization in 3D via deep transfer learning *Comput.-Aided Design* **135** 103014



- [33] Xiang C, Wang D, Pan Y, Chen A, Zhou X and Zhang Y 2022 Accelerated topology optimization design of 3D structures based on deep learning *Struct. Multidiscipl. Opt.* **65** 99
- [34] Nie Z G, Lin T, Jiang H L and Kara L B 2021 TopologyGAN: topology optimization using generative adversarial networks based on physical fields over the initial domain *J. Mech. Design* **143** 031715
- [35] Liao W, Lu X, Huang Y, Zheng Z and Lin Y 2021 Automated structural design of shear wall residential buildings using generative adversarial networks *Autom. Constr.* **132** 103931
- [36] Danhaive R and Mueller C T 2021 Design subspace learning: Structural design space exploration using performance-conditioned generative modeling *Autom. Constr.* **127** 103664
- [37] Gallet A, Liew A, Hajirasouliha I and Smyl D 2023 Influence zones for continuous beam systems (arXiv:2305.02211)
- [38] Whalen E and Mueller C 2022 Toward reusable surrogate models: graph-based transfer learning on trusses *J. Mech. Design* **144** 021704
- [39] Bleker L, Pastrana R, Ohlbrock P O and D'Acunto P 2023 Structural form-finding enhanced by graph neural networks *Towards Radical Regeneration* ed C Gengnagel, O Baverel, G Betti, M Popescu, M R Thomsen and J Wurm (Springer) pp 24–35
- [40] Koen B V 2003 *Discussion of the Method: Conducting the Engineer's Approach to Problem Solving* (Oxford University Press)
- [41] Bulleit W, Schmidt J, Alvi I, Nelson E and Rodriguez-Nikl T 2015 Philosophy of engineering: what it is and why it matters *J. Professional Issues Eng. Education Practice* **141** 02514003
- [42] European Committee for Standardisation 2002 BS EN 1990:2002+A1:2005 Eurocode *Basis of Structural Design* (BSI British Standards) (<https://doi.org/10.3403/03202162>)
- [43] Hajela P and Berke L 1991 Neurobiological computational models in structural analysis and design *Comput. Struct.* **41** 657–67
- [44] Mojtabaei S M, Becque J, Hajirasouliha I and Khandan R 2023 Predicting the buckling behaviour of thin-walled structural elements using machine learning methods *Thin-Walled Struct.* **184** 110518
- [45] Adeli H and Park H S 1995 Optimization of space structures by neural dynamics *Neural Netw.* **8** 769–81
- [46] Mai H T, Mai D D, Kang J, Lee J and Lee J 2023 Physics-informed neural energy-force network: a unified solver-free numerical simulation for structural optimization *Eng. Comput.* (<https://doi.org/10.1007/s00366-022-01760-0>)
- [47] Yan J, Zhang Q, Xu Q, Fan Z R, Li H J, Sun W and Wang G Y 2022 Deep learning driven real time topology optimisation based on initial stress learning *Adv. Eng. Inform.* **51** 101472
- [48] Wang G G and Shan S 2007 Review of metamodeling techniques in support of engineering design optimization *J. Mech. Design* **129** 370–80
- [49] Koziel S and Leifsson L 2013 *Surrogate-Based Modeling and Optimization: Applications in Engineering* (Springer)
- [50] Saka M P and Geem Z W 2013 Mathematical and metaheuristic applications in design optimization of steel frame structures: an extensive review *Math. Problems Eng.* **2013** 1–33
- [51] European Committee for Standardisation 2015 BS EN 1993-1-1:2005+A1:2014 Eurocode 3. Design of steel structures *General Rules and Rules for Buildings* (BSI British Standards) (<https://doi.org/10.3403/03270565>)
- [52] Mitchell T M 1997 *Machine Learning, Mcgraw-Hill Series in Computer Science* (McGraw-Hill)
- [53] Kubat M 2021 *An Introduction to Machine Learning* 3rd edn (Springer)
- [54] Smola A J and Schölkopf B 2004 A tutorial on support vector regression *Stat. Comput.* **14** 199–222
- [55] Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32
- [56] Murphy K P 2022 *Probabilistic Machine Learning: An Introduction (Adaptive Computation and Machine Learning Series)* (The MIT Press)
- [57] European Committee for Standardisation 2017 BS EN 10365:2017 hot rolled steel channels, I and H sections *Dimensions and Masses* (BSI British Standards) (<https://doi.org/10.3403/30318327>)
- [58] Preisinger C 2013 Linking structure and parametric geometry *Architect. Design* **83** 110–3
- [59] Kim J-H 2009 Estimating classification error rate: repeated cross-validation, repeated hold-out and bootstrap *Comput. Statist. Data Anal.* **53** 3735–45
- [60] Gallet A and Smyl D 2024 ORDA CBeamXP: continuous beam cross-section predictors dataset (<https://doi.org/10.15131/shef.data.23945562>)

- [61] Dennis A A and Rigby S E 2023 The direction-encoded neural network: a machine learning approach to rapidly predict blast loading in obstructed environments *Int. J. Protect. Struct.* (<https://doi.org/10.1177/20414196231177364>)
- [62] Vakharia V, Shah M, Nair P, Borade H, Sahlot P and Wankhede V 2023 Estimation of lithium-ion battery discharge capacity by integrating optimized explainable-AI and stacked LSTM model *Batteries* **9** 125
- [63] Sun H, Burton H V and Huang H 2021 Machine learning applications for building structural design and performance assessment: state-of-the-art review *J. Build. Eng.* **33** 101816