Mechanical Systems and Signal Processing 241 (2025) 113478

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Full length article

Machine learning algorithm for rivet-squeezing force estimation based on the dynamic response of the joint

Tim Vrtač, Miha Pogačar, Miha Kodrič, Gregor Čepon *

Faculty of Mechanical Engineering, University of Ljubljana, Aškerčeva 6, 1000 Ljubljana, Slovenia

ARTICLE INFO

Communicated by M. Brake

Keywords:
Rivet
Rivet-squeezing force
Dynamic substructuring
Machine learning
Quality control

ABSTRACT

The rivet-squeezing force is one of the most influential parameters affecting the performance of a riveted joint. It directly affects the stress distribution near the joint, influencing mechanical properties, fatigue life, and failure mechanisms. Ensuring control over this parameter is therefore essential for product quality. Our prior research has shown that the squeezing force can be estimated post-production by analyzing the dynamic response of the joint—an approach that enables quality control even during later stages of a product's lifecycle. Adequacy of the riveted joint's mechanical performance is then estimated through comparison of the riveting force estimate to the desired reference value. Traditional methods rely on the similarity between the dynamic response of the observed joint and those of reference joints with known rivet-squeezing forces. This paper proposes an improved methodology that replaces the similarity-based criterion with a Machine Learning (ML) algorithm to enhance estimation robustness. As in the conventional method, dynamic substructuring is used to isolate the joint's response from the surrounding assembly, enabling the application of the same ML model across different assemblies—provided that the material, geometric, and frictional properties near the joint remain consistent. The proposed method is validated through a laboratory case study and benchmarked against the existing LAC-based estimation approach. The case study results demonstrate improved robustness and broader generalization against the LAC-based approach. This indicates that the advanced inference capabilities allow the ML model to better distinguish the effects of the rivet-squeezing force from variations caused by material inconsistencies, minor changes in process parameters, or sensor placement compared to the LAC-based approach. The key challenge for the ML-based approach is to acquire a sufficiently large and representative dataset for the ML model training.

1. Introduction

Riveting is a process for joining metal sheet plates, widely used in sectors such as the aerospace industry [2,3] and civil construction [4]. The main advantages of riveting compared to other component-joining methods include process simplicity, the ability to join dissimilar materials, and high reliability [5]. To maintain its function, it is crucial that the riveted joint retains adequate mechanical properties throughout its lifespan under varying operational loads and environmental conditions. This requires appropriate mechanical properties of the components (i.e., rivet and sheet metal), as well as optimized process parameters. The mechanical performance depends on the material and geometrical properties of the components, while key process parameters include the riveting force, speed, and die configuration [6].

* Corresponding author.
E-mail address: gregor.cepon@fs.uni-lj.si (G. Čepon).

https://doi.org/10.1016/j.ymssp.2025.113478
Received 29 April 2025; Received in revised form 6 August 2025; Accepted 3 October 2025

Available online 10 October 2025

List of Abbreviations

CV: Cross-Validation
DT: Decision Tree
DoF: Degree of Freedom
EoL: End-of-Line
FBS: Frequency-Based Substructuring
FEM: Finite Element Method
FRF: Frequency Response Function
IDM: Interface Displacement Matrix
kNN: k-Nearest Neighbors
LAC: Local Amplitude Criterion
LDA: Linear Discriminant Analysis
LM FBS: Lagrange Multiplier Frequency-Based Substructuring
ML: Machine Learning
OC: Ordinal Classification
PCA: Principal Component Analysis
QWK: Quadratic Weighted Kappa
RBF: Radial Basis Function
SVM: Support Vector Machine
TSVD: Truncated Singular Value Decomposition
VP: Virtual Point
VPT: Virtual Point Transformation

List of Symbols

A	Set of unique classes in classification problem
a	Number of observed response DoFs to be transformed using VPT
$\mathbf{B}_c$, $\mathbf{B}_e$	Boolean matrices corresponding to compatibility and equilibrium condition, respectively
b	Number of VP response DoFs
C	Amount of regularization
C_i	ith unique class in classification problem
$d_{i,f}$	Distance between the actual and approximated value for the fth feature of the ith sample
$d_{\mathrm{f,max}}$	Limit for distance value $d_{i,f}$ in outlier detection algorithm
$\boldsymbol{f}$	Vector of external loads
F_{sq}	Squeezing force
$\boldsymbol{F}_{\mathrm{sq}}$	Vector of squeezing forces
$\tilde{f}_f$	Function approximating the relation between fth column of the dataset $\mathbf{X}_f$ and the riveting squeezing forces $\boldsymbol{F}_{\mathrm{sq}}$
$\boldsymbol{g}$	Vector of interface loads
$\boldsymbol{g}_i^\star$	Vector of interface loads at group of DoFs i at substructure $\star$
$\mathbf{I}$	Unit matrix
J	Set of impedance matrix-squeeze force pairs
J_{v}	Set of impedance matrix-squeeze force pairs for validation samples
k	number of unique classes in classification problem
$\mathbf{L}_{\mathrm{c}}$, $\mathbf{L}_{\mathrm{e}}$	Localization matrices corresponding to compatibility and equilibrium condition, respectively
$\boldsymbol{m}$	Vector of VP loads
$m_\star$	Point of mass as substructure $\star$
N_Ω	Number of mutually exclusive subsets of set Ω
n	Number of samples
n_{resp}, n_{exc}	number of observed/controlled response and excitation DoFs
n_{f}, $n_{\mathrm{f,red}}$	Number of features in dataset $\mathbf{X}$, and number of features in the reduced dataset $\mathbf{X}_{\mathrm{red}}$
$\boldsymbol{q}$	Vector of VP responses
$\mathbf{R}_{\mathrm{u}}$, $\mathbf{R}_{\mathrm{f}}$	IDM matrices corresponding to the responses and loads, respectively
$\mathbf{T}_{\mathrm{u}}$, $\mathbf{T}_{\mathrm{f}}$	Displacement ad force transformation matrix responsively
$\boldsymbol{u}$	Vector of displacements
$\boldsymbol{u}_i^\star$	Vector of displacement responses at group of DoFs i at substructure $\star$
$\mathbf{X}$, $\mathbf{X}_{\mathrm{red}}$	Full dataset and dataset with reduced number of features
$\boldsymbol{X}_f$	fth column (feature) of dataset $\mathbf{X}$
$\boldsymbol{X}'$	Data for sample not used in ML model training
$\boldsymbol{x}_i$	Data for ith sample of the dataset $\mathbf{X}$
$\mathbf{Y}$	Global admittance matrix
$\widetilde{\mathbf{Y}}$	Admittance matrix of the decoupled system
$\mathbf{Y}^\star$	Admittance matrix of substructure $\star$
$\boldsymbol{y}$	Vector of labels
$\boldsymbol{y}'$	Label for the sample not used in the ML model training
$\tilde{\boldsymbol{y}}'$	Prediction for the sample not used in the ML model training
y_i^*	Transformed label of jth sample in the simple approach to ordinal classification [1]
$\mathbf{Z}^\star$	Impedance matrix of substructure $\star$
γ	Parameter of the RBF kernel
λ	Vector of Lagrange multipliers

μ_i	ith friction-affected contact area
$\mu(\boldsymbol{a})$	Mean value of the vector $\boldsymbol{a}$
$\boldsymbol{\mu}$	Vectors of residual displacements in VPT
ρ	Correlation
$\rho_{\min}$	Minimum correlation value in the outlier detection algorithm
$\sigma(\boldsymbol{a})$	Standard deviation of the vector $\boldsymbol{a}$
Ω	Set of discrete frequency points
ω	Angular frequency
ω_i	ith discrete frequency point

Although many of these parameters can be monitored during production, comprehensive joint quality assessment often requires additional inspection and is relevant both for End-of-Line (EoL) testing and for evaluations throughout the product's service life. Ensuring reliable procedures to evaluate the integrity of riveted joints is critical for the safety, performance, and longevity of products. With industry shifting towards lighter structures, efficient assembly, and improved lifecycle management, the need for effective riveted joint quality assessment is becoming increasingly important.

Riveted joint properties can be evaluated using two main types of methods: destructive and non-destructive testing. Destructive tests allow us to gain a detailed insight into the joint's mechanical properties. Quasi-static tensile and shear tests provide information about the tensile and shear strength of the joint, the critical static load as well as the fracture mechanisms [7]. Special testing devices were designed to test the effect of the loading angle on the mechanical behavior of the riveted joint [8]. Pull-off tests are performed to test the pull-off resistance of the riveted joint [9]. As the mechanical response of the riveted joint also depends on the loading speed the quasi-static tests do not always suffice. Riveted joints' properties under variable loading can thus be tested using various dynamic tests. Komorek and Godzimmirski [10] designed a pendulum hammer impact test adapted for testing of the riveted joints. Wang et al. [11] investigated the mechanical response of the solid riveted joint under different loading speeds and different combinations of tensile and shear loading using the split Hopkinson tensile bar system. To determine the fatigue strength of the riveted joint, fatigue tests are performed [12]. The described tests provide us with a substantial amount of information regarding the mechanical properties of the riveted joint, however, because they irreversibly damage the tested component, they are unsuitable for EoL inspection or for monitoring throughout the product's lifecycle.

On the other hand, the non-destructive methods, such as ultrasonic testing [13], enable testing of riveted joints without compromising the product's integrity. Stamm et al. [14] discuss non-destructive techniques for inspecting aircraft fuselage rivets: pulse thermography, the Eddy current technique [15], magneto-optic imaging [16], and others. The paper also researched the application of a thermal inspection procedure that combines vibrational excitation of the specimen with the thermal detection of flaws using Ultrasound Lock-In Thermography. The established non-destructive methods allow us to detect internal flaws. However, they also have one key drawback as they do not provide direct information about mechanical parameters.

In practice, the quality of the riveted joint is frequently evaluated indirectly [17]. A common approach to do this is by measuring the rivet's driven head dimensions. However, the mechanical properties of the riveted joint can vary substantially at the same driven head size, as it is affected by the type of rivet (e.g. solid, semi-tubular, or tubular rivet), rivet length, and material of the rivet [17]. On the other hand, the rivet-squeezing force changes along with these parameters. In the process of riveting the rivet-squeezing force is applied to the rivet in the axial direction to deform the rivet's tail and form a shop head. Besides axial deformations, the rivet's shank also expands radially. Thus, different parts of the assembly (i.e. metal sheet plates and the rivet) press on one another, resulting in a residual stress field that depends on the size of the squeezing force and influences the joint's mechanical properties [17–20]. Previous research has shown that the higher rivet-squeezing force results in increased static failure load, high-cycle fatigue life, and low-cycle fatigue life [21–23]. In addition, it also affects the crack shape, initiation position, and propagation path [24].

The described effects of the rivet-squeezing force make it a key parameter of the riveting process, but also a suitable parameter for joint quality evaluation [25]. Substantial efforts were put into estimation of the appropriate squeezing force in the early design stages via numerical simulations. Cheraghi [26] utilized Finite Element Simulations (FEM) to study the effects of the riveting parameters on the quality of the riveted joint - among others, the range of the squeezing force that can be applied without violating the quality requirements. Chang et al. [27] used FEM to improve control over the riveting deformations, paying special attention to local plastic deformations and testing the effect of different squeezing forces in the process. Zhang et al. [28] researched the application of mathematic model of the riveting process and FEM simulations for accurate analysis of the deformation of thin-walled sheet-metal parts. Nevertheless, numerical simulations only tell us what the values of the squeezing force are supposed to be and thus do not ensure a quality riveted joint without monitoring the force during or after the riveting process. Therefore, non-destructive methods for its estimation could provide valuable insights for structural assessment.

Squeezing force can be measured using the load-cells directly during the riveting process, however such measurement are not always possible. Therefore, multiple algebraic models were developed for estimating the squeezing force after the completed riveting process. Müller [17] studied the relations between the driven head dimensions and the riveting squeeze force. Based on this relations, De Rijck et al. [25] proposed widely used equations for the squeezing force estimation. Later, these equations have been updated taking account of parameters other than driven head dimensions in order to increase accuracy of estimations. Zeng et al. [29] took into account the flow of rivet material into the metal-sheet hole, resulting in better agreement with the experimental results, compared to the original equations. Figueira and Trabasso [30] proposed a revised model that (in addition to existing parameters)

takes into account the grip thickness and the friction among assembly components. Kang et al. [31] extended the model even further by considering frictional effects, non-uniform rivet/sheet interference, the barreling effect, the spring-back of the driven head's dimensions, and the modified hole diameter. While these analytical approaches allow for estimation of the squeezing force directly after riveting, they do not account for the changes in joint properties due to the loading of the joint which would be required for evaluation of the joint quality throughout the product's service life.

To address this gap, in our previous work [32], we introduced a novel approach for estimating the rivet-squeezing force based on the joint's dynamic response. The method employed Frequency-Based Substructuring (FBS) to isolate a dynamic model of the joint by decoupling it from neighboring components [33,34]. The key advantage of this approach is that the dynamic model obtained through FBS is independent of the surrounding assembly, allowing the method to generalize across different assembly configurations. However, this generalization is only valid when the local material, geometric, and frictional conditions near the joint remain consistent. The dynamic model of a joint with unknown parameters was classified by identifying the closest match within a reference dataset constructed from joints with known rivet-squeezing forces. Specifically, the comparison was performed using the Local Amplitude Criterion (LAC), which compares the amplitude spectra of Frequency Response Functions (FRFs)—hereafter referred to as the LAC-based approach. While the LAC-based approach showed promising results, its broader application revealed limitations in robustness. The method proved sensitive to minor variations in material properties, friction conditions, joint geometry, and experimental noise. As a result, not all changes observed in the dynamic response could be directly attributed to variations in the rivet-squeezing force, and under such conditions, similarity-based estimation using LAC may lack sufficient accuracy. These limitations represent a significant research gap and motivate the exploration of machine learning-based alternatives.

This paper aims to advance the previously proposed approach by replacing the LAC-based comparison with a Machine Learning (ML) algorithm. This enhancement addresses the identified research gap and is expected to reduce the sensitivity of rivet-squeezing force estimation to minor variations in material and frictional properties of assembly components, differences in riveting process parameters (e.g., squeezing speed, die configuration), and experimental uncertainty.

ML algorithms have the potential to disentangle the influence of such confounding factors and more reliably isolate the effect of the rivet-squeezing force. Moreover, while the LAC-based method relies on a single sample to represent each force class, ML approaches utilize multiple samples per class during training. This not only improves generalization but also mitigates the impact of sample-specific deviations and measurement noise. Although the ML-based approach does not eliminate the need for comparable material and frictional properties across samples, it can accommodate limited variability within the bounds of the training dataset. A key challenge of this method, however, lies in acquiring a sufficiently large and representative dataset to enable effective model training.

The study includes a case example in which a training dataset is experimentally acquired under controlled conditions. The proposed ML-based methodology is trained and validated on this dataset, and its performance is evaluated on the laboratory test structure. Results are compared to those of the previously developed LAC-based approach, demonstrating improved robustness and potential for broader generalization.

2. Joint identification

The first step in developing an ML model capable of estimating the rivet-squeezing force is to obtain the data used for training. Training data consists of data from multiple samples obtained by joining the metal sheet plates together by deforming rivets, applying different rivet-squeezing forces to the rivet's driven head (Fig. 1). The result of the riveting process is an assembly on which sensors can be attached, and the dynamic response can be measured. Assuming that the dynamic response is related to the rivet-squeezing force, the measured dynamic response can be utilized for the estimation of the squeezing force. However, in such a case, an estimation is limited to the observed assembly. To achieve a more general approach, we want these dynamic responses to be independent of the components hold together by the joint. To achieve such an isolation of the riveted joint from the assembly, FBS can be utilized [33,34]. FBS allows us to dynamically decouple a component from a multi-component system and thus obtain the dynamic response of the remaining part (Fig. 2). In this manner a joint admittance $\mathbf{Y}^{\mathrm{J}}(\omega)$[1] is obtained by decoupling all the adjacent substructures (components A and B in Fig. 2b).

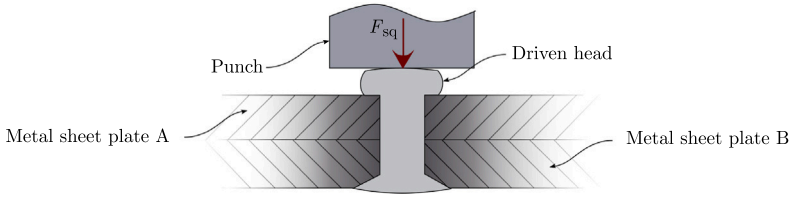


Fig. 1. Riveting process: driven head of the rivet is deformed by applying the squeezing force F_{sq} in order to join metal sheet plates A and B.

[1] Admittance matrix $\mathbf{Y}(\omega)$ is a matrix describing connections between excitations and responses to these excitations, stored in form of FRFs, at selected degrees of freedom (DoFs) of the object.

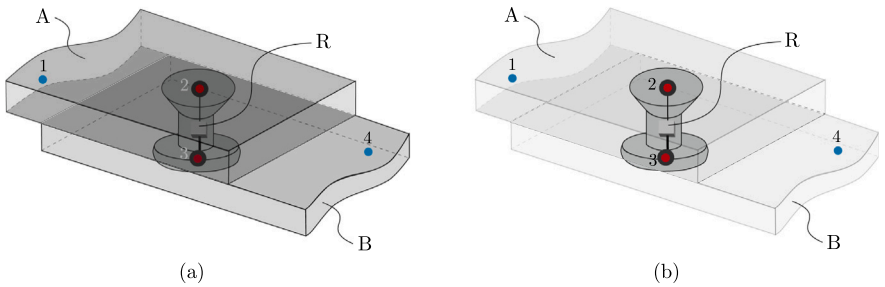


Fig. 2. Dynamic decoupling: assembly ARB in (a) coupled state and (b) after decoupling of substructures A and B, resulting in dynamic model of the riveted joint R. (•) Interface DoFs, (•) internal DoFs. Interface DoFs are DoFs located at the boundary between the substructures, while internal DoFs are the DoFs located away from the interface.

To decouple the joint-adjacent components, two constraints are enforced at the contacts between the joint and the neighboring components. The first constraint is the compatibility condition that states that responses at the contact points (i.e. interface DoFs; $\star_2$ and $\star_3$ in Fig. 2) must align ($\boldsymbol{u}_2^{\mathrm{A}}(\omega) = \boldsymbol{u}_2^{\mathrm{R}}(\omega)$ and $\boldsymbol{u}_3^{\mathrm{R}}(\omega) = \boldsymbol{u}_3^{\mathrm{B}}(\omega)$ in Fig. 3). This can be written as follows[2] [35]:

$$\mathbf{B}_{\mathrm{c}}\boldsymbol{u} = \mathbf{0}; \quad \mathbf{B}_{\mathrm{c}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix}, \tag{1}$$

where $\mathbf{B}_{\mathrm{c}}$ is a boolean matrix connecting corresponding interface responses at response DoFs between adjacent substructures, stored in the vector of responses $\boldsymbol{u}$. For the definition of $\mathbf{B}_{\mathrm{c}}$ as well as vector $\boldsymbol{u}$ in Eq. (1), the following order of DoFs is assumed: internal response DoFs at substructure A ($\star_1$ in Fig. 2), followed by internal DoFs at substructure B ($\star_4$ in Fig. 2) of the coupled assembly ARB, interface DoFs at the contact between substructure A and joint R ($\star_2$ in Fig. 2), interface DoFs at the contact between joint R and substructure B ($\star_3$ in Fig. 2), internal DoFs followed by interface DoFs at substructure A (when A is treated individually) and, internal DoFs followed by interface DoFs at substructure B (when B is treated individually). Each group of DoFs corresponds to an individual column of $\mathbf{B}_{\mathrm{c}}$ in Eq. (1).

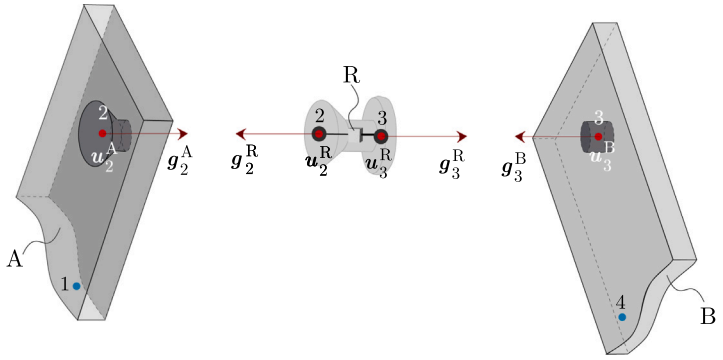


Fig. 3. Assembly ARB in exploded view, displaying forces $\boldsymbol{g}$ and responses $\boldsymbol{u}$ at interface DoFs. (•) Interface DoFs, (•) internal DoFs.

The second constraint is the equilibrium condition, which states that the pairs of interface forces in vector of interface loads $\boldsymbol{g}$ (i.e. forces holding substructures together; Fig. 3) must have the same amplitudes and opposite directions ($\boldsymbol{g}_2^{\mathrm{A}} = -\boldsymbol{g}_2^{\mathrm{R}}$ and $\boldsymbol{g}_3^{\mathrm{R}} = -\boldsymbol{g}_3^{\mathrm{B}}$) [35]:

$$\boldsymbol{g} = -\mathbf{B}_{\mathrm{e}}^{\mathrm{T}}\lambda; \quad \mathbf{B}_{\mathrm{e}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix}, \tag{2}$$

where $\mathbf{B}_{\mathrm{e}}$ is a boolean matrix connecting corresponding interface loads at excitation DoFs between adjacent substructures. The order of excitation DoFs assumed in Eq. (2) is the same as the one of the response DoFs in Eq. (1).

By considering these two conditions in the equation connecting the system responses to the excitations (i.e. $\boldsymbol{u} = \mathbf{Y}(\boldsymbol{f} + \boldsymbol{g})$ where $\mathbf{Y}$ is a global admittance matrix) the Lagrange Multiplier Frequency-Based Substructuring (LM FBS) equation can be derived [36]:

$$\widetilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{Y}\mathbf{B}_{\mathrm{e}}^{\mathrm{T}}\left(\mathbf{B}_{\mathrm{c}}\mathbf{Y}\mathbf{B}_{\mathrm{e}}^{\mathrm{T}}\right)^{-1}\mathbf{B}_{\mathrm{c}}\mathbf{Y}, \tag{3}$$

[2] An explicit dependency on the frequency is omitted to improve the readability of the notation, as will be for the remainder of the paper.

where $\widetilde{\mathbf{Y}}$ is an admittance matrix of the decoupled system. For decoupling joint adjacent substructures A and B from assembled system ARB, the global admittance matrix $\mathbf{Y}$ is structured as follows [33,34]:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}^{\mathrm{ARB}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{Y}^{\mathrm{A}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{Y}^{\mathrm{B}} \end{bmatrix}. \tag{4}$$

The $\widetilde{\mathbf{Y}}$ matrix resulting from Eq. (3) contains all the DoFs included in the global admittance matrix $\mathbf{Y}$ which means that after the dynamic decoupling not all DoFs correspond to locations on the remaining components of the system (i.e. joint R) - internal DoFs of A and B (i.e. $\star_1$ and $\star_4$ in Fig. 2, respectively) have no physical meaning. To obtain the joint admittance $\mathbf{Y}^{\mathrm{R}}$, the non-physical DoFs have to be extracted from $\widetilde{\mathbf{Y}}$. This can be done by following equation [37]:

$$\mathbf{Y}^{\mathrm{R}} = \mathbf{L}_{\mathrm{c}}\widetilde{\mathbf{Y}}\mathbf{L}_{\mathrm{e}}; \quad \mathbf{L}_{\mathrm{c}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \ \mathbf{L}_{\mathrm{c}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}^{\mathrm{T}}, \tag{5}$$

where $\mathbf{L}_{\mathrm{c}}$ and $\mathbf{L}_{\mathrm{e}}$ are localization matrices allowing for extraction of joint DoFs. DoFs of $\mathbf{Y}^{\mathrm{R}}$ actually occur twice in $\widetilde{\mathbf{Y}}$, but only one set is extracted when using Eq. (5). The definitions of $\mathbf{L}_{\mathrm{c}}$ and $\mathbf{L}_{\mathrm{e}}$ in Eq. (5) are valid for the order of response and excitation DoFs matching the DoF order in Eqs. (1) and (2), respectively.

2.1. Considerations regarding the dynamic decoupling techniques

The decoupling process outlined above is referred to as the standard interface decoupling. This designation arises because the compatibility condition (Eq. (1)) and the equilibrium condition (Eq. (2)) are imposed on the DoFs at the interface between the decoupled component and the neighboring components (interface DoFs). This formulation can result in spurious peaks in the resulting FRFs [38–40]. These peaks arise from residual dynamics in the original system that cannot be observed or controlled in the selected set of interface DoFs [40] and measurement errors when dealing with experimentally obtained FRFs [38]. These effects can be mitigated using extended interface decoupling. Unlike in the standard interface decoupling, in extended interface formulation, the compatibility and equilibrium conditions are imposed on both, the interface and internal DoFs of the decoupled components. However, the interface matrix $\mathbf{B}_{\mathrm{c}}\mathbf{Y}\mathbf{B}_{\mathrm{e}}^{\mathrm{T}}$ in the extended interface formulation is often ill-conditioned. This presents a challenge when calculating the inverse of the interface matrix in the LM FBS Eq. (3) [39]. Commonly, this ill-conditioning is tackled by applying the Moore–Penrose inverse in combination with Truncated Singular Value Decomposition (TSVD) to obtain the interface matrix inverse [41].

2.2. Considerations regarding the joint model DoFs

Successful identification of the rivet-squeezing force requires understanding how it affects the properties of the joint. In the course of riveting the rivet's tail is deformed, forming a shop head. Simultaneously, the rivet's shank expands in the radial direction, resulting in a stress field at the interface between the shank and the hole. To some lesser degree, the metal sheet plates are also deformed. After the squeezing process concludes, elastic relaxation of the rivet and metal sheet plates takes place. If the rivet's axial relaxation (z-axis in Fig. 4) is smaller than the relaxation of the metal sheet plates, this results in a clamping force pressing the metal sheet plates together [17].

These phenomena manifests as frictional forces at the contact surfaces between the rivet and the hole, between rivet heads and outer surfaces of the metal sheet plates, and between the metal sheet plates themselves. For a single rivet, this friction is especially evident in the rotations around the z-axis (Fig. 4) [17,32]. In the dynamic model of the joint, variations in frictional forces can be discerned as differences in the dynamic stiffness of the joint, especially in rotational or moment-related dynamics around the z-axis.

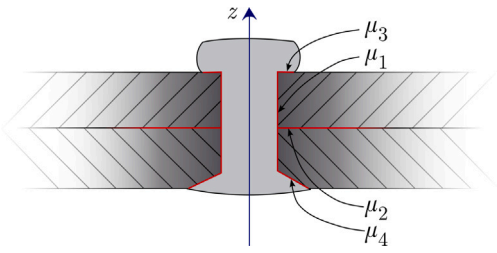


Fig. 4. Friction-affected areas (denoted by —) in vicinity of the riveted joint: contact between the rivet shank (μ_1), contact between the metal sheet plates (μ_2) and contacts between rivet heads and outer surfaces of the metal sheet plates (μ_3 and μ_4).

Frictional effects at the contact surfaces are difficult to model within numerical simulations. Consequently, the FRFs required for FBS-based joint identification are typically obtained through experimental measurements. In such setups, rotational interface DoFs are often unavailable, primarily due to two reasons: (1) only translational motion is typically measured, using uni- or tri-axial accelerometers that capture translations in up to three directions; (2) even when rotational accelerometers are available, it is generally infeasible to position them precisely at the contact points between substructures. The same limitations apply to

excitations—applying moments directly at interface DoFs is often impractical. To overcome these challenges, the Virtual Point Transformation (VPT) can be employed. VPT enables the transformation of responses measured near the interface into a single point called a Virtual Point (VP), which can be placed exactly at the contact location [42]. Additionally, VPT facilitates the conversion of translational DoFs into rotational ones, and even into flexible DoFs such as extension, bending, and torsion [43]. Excitations can be transformed using VPT in similar manner as responses.

During joint identification, both metal sheet plates are decoupled from the assembly. Hence, there are two interfaces: an interface between the metal sheet plate A and the rivet, and the interface between the metal sheet plate B and the rivet (Fig. 5). As a result, VPT must be applied twice on the assembly admittance (once per interface) and once on the admittance of each metal sheet plate. This results in a set of transformed admittance matrices of assembly ARB ($\mathbf{Y}^{\mathrm{ARB}}_{\mathrm{VPT}}$), component A ($\mathbf{Y}^{\mathrm{A}}_{\mathrm{VPT}}$), and component B ($\mathbf{Y}^{\mathrm{B}}_{\mathrm{VPT}}$).

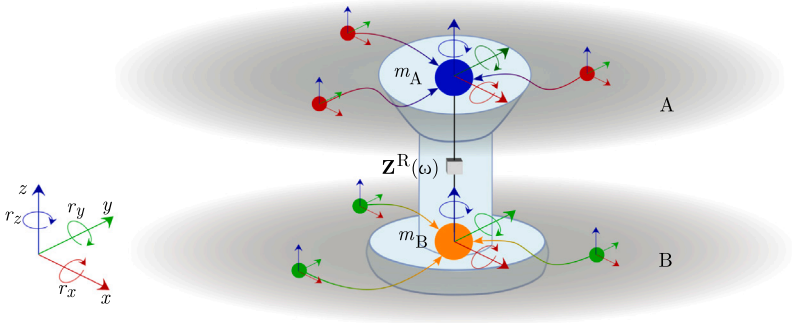


Fig. 5. Responses/excitations from multiple 3 DoF nodes at substructures A and B are transformed into two 6 DoF VPs. Next, components A and B are decoupled resulting in a model of the joint consisting of points of mass m_{A} and m_{B} connected over mutual stiffness properties captured within impedance matrix $\mathbf{Z}^{\mathrm{R}}$. (•) Observed/controlled DoFs close to the interface on substructure A, (•) VP on substructure A, (•) observed/controlled DoFs close to the interface on substructure B, (•) VP on substructure B. Local coordinate systems at each point denote the DoFs observed/controlled at point's location.

Finally, metal sheet plates A and B are decoupled from the assembly. This results in a riveted joint model consisting of two point masses m_{A} and m_{B} (Fig. 5). Each point of mass corresponds to one of the VPs with a selected number of DoFs (usually 3 translations and 3 rotations), resulting in a 12-DoF model of the joint displayed in Fig. 5.

2.3. VPT

VPT [42] leverages the existence of kinematic relationships between responses and loads applied at different points on a structure. Given a vector of displacements $\boldsymbol{u} \in \mathbb{C}^a$ (where a is the number of observed response DoFs), these displacements can be transformed into the displacements $\boldsymbol{q} \in \mathbb{C}^b$ at a VP, where b denotes the number of VP response DoFs and $b \leq a$:

$$\boldsymbol{u} = \mathbf{R}_{\mathrm{u}}\boldsymbol{q} + \boldsymbol{\mu}, \tag{6}$$

where $\mathbf{R}_{\mathrm{u}}$ is the Interface Displacement Matrix (IDM), which contains the kinematic relationships between the displacements $\boldsymbol{u}$ and $\boldsymbol{q}$, while $\boldsymbol{\mu}$ represents the residual displacements not captured by the selected IDM.[3] By introducing the constraint $\mathbf{R}_{\mathrm{u}}^{\mathrm{T}}\boldsymbol{\mu} = 0$, Eq. (6) can be solved for $\boldsymbol{q}$ in the least-squares sense:

$$\boldsymbol{q} = \mathbf{T}_{\mathrm{u}}\boldsymbol{u}; \quad \mathbf{T}_{\mathrm{u}} = \left(\mathbf{R}_{\mathrm{u}}^{\mathrm{T}}\mathbf{R}_{\mathrm{u}}\right)^{-1}\mathbf{R}_{\mathrm{u}}^{\mathrm{T}}, \tag{7}$$

where $\mathbf{T}_{\mathrm{u}}$ is a displacement-transformation matrix. A separate set of IDMs stored in IDM matrix $\mathbf{R}_{\mathrm{f}}$ can be used to describe the relationship between the observed set of loads $\boldsymbol{f}$ and the VP loads $\boldsymbol{m}$:

$$\boldsymbol{f} = \mathbf{T}_{\mathrm{f}}^{\mathrm{T}}\boldsymbol{m}; \quad \mathbf{T}_{\mathrm{f}}^{\mathrm{T}} = \mathbf{R}_{\mathrm{f}}\left(\mathbf{R}_{\mathrm{f}}^{\mathrm{T}}\mathbf{R}_{\mathrm{f}}\right)^{-1}. \tag{8}$$

Using the above transformations, the original admittance matrix $\mathbf{Y}$ can be transformed into the admittance containing VP DoFs $\mathbf{Y}_{\mathrm{VPT}}$ in a single step using the following equation:

$$\mathbf{Y}_{\mathrm{VPT}} = \mathbf{T}_{\mathrm{u}}\mathbf{Y}\mathbf{T}_{\mathrm{f}}^{\mathrm{T}}. \tag{9}$$

3. Procedure for the rivet-squeezing force estimation using machine learning algorithms

In this section, we introduce a ML-based approach for estimating the rivet-squeezing force. The method exploits the strong correlation between the dynamic response of a riveted joint and the applied rivet-squeezing force, as discussed in Section 2.2. Accordingly, the joint's dynamic response serves as the input to the ML model. This methodology offers several key advantages. Most

[3] The IDM matrix maps the translations in $\boldsymbol{u}$ to a chosen set of VP DoFs stored in $\boldsymbol{q}$. It can include rigid-body motions (translations and rotations), as well as flexible modes such as extension, torsion, and bending [43].

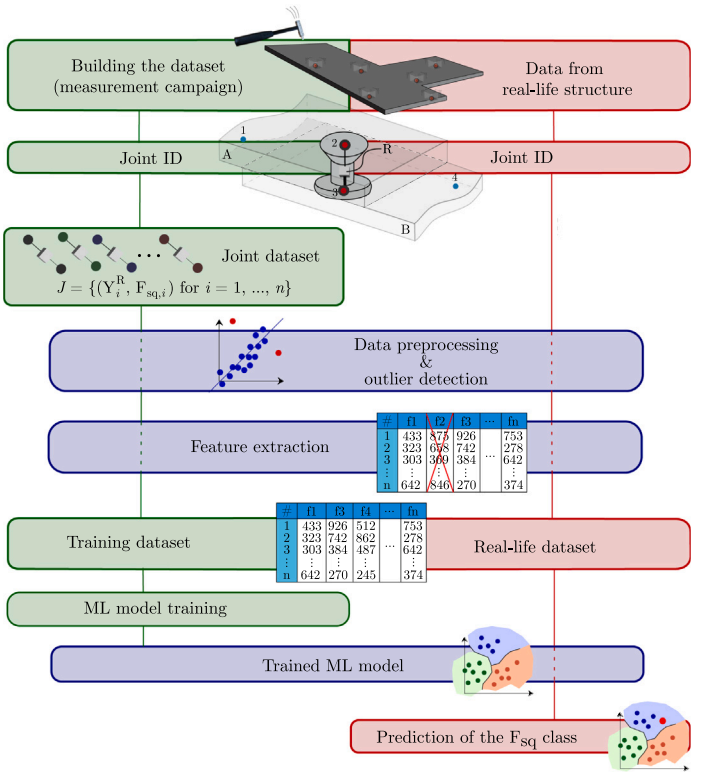


Fig. 6. Procedure of training and using the ML algorithms for identification of the rivet-squeezing force.

notably, FBS enables the isolation of joint dynamics from the overall dynamic behavior of the assembled structure. This decoupling allows the ML model to generalize across different assemblies, provided that the material, geometric, and frictional properties in the vicinity of the joint remain sufficiently similar. Further advantages stem from the use of ML algorithms. Their advanced inference capabilities allow them to account for variability in the joint's dynamic response caused by minor changes in material or frictional properties, variations in riveting process parameters (excluding the rivet-squeezing force itself), and measurement errors. On the other hand, the LAC-based methodology [32] estimates the rivet-squeezing force based on overall similarity rather than based on understanding of complex relations between the dynamic response within different frequency regions and the squeezing force. Therefore, more robust squeezing force estimations are anticipated from the ML-based approach, which will be demonstrated within the Case study section (Section 4).

The proposed workflow for estimating the rivet-squeezing force from the joint's dynamic response is illustrated in Fig. 6. The first step is to collect training data for the ML algorithm. These data may be acquired experimentally or through numerical simulations. While experimental data are generally preferred due to their realistic representation of actual structures, obtaining large volumes of such data can be challenging. In cases where experimental data are limited, additional training samples can be generated numerically. However, it is difficult to accurately capture frictional conditions at the contact surfaces around the rivet in simulations. Therefore, the required data in this study are collected experimentally. When the available dataset consists of a relatively small number of samples (typically on the order of a few hundred or fewer), special care must be taken in data preprocessing, feature extraction, and algorithm selection to ensure reliable model performance.

Data is collected by measuring the dynamic responses of the assemblies with different corresponding rivet-squeezing forces. In addition to the dynamic models of assemblies, the models of joint-adjacent substructures must be measured. Joint admittance for each sample is then extracted by utilizing the FBS-based joint identification approach (Section 2). Joint dynamic properties are most commonly interpreted in the impedance (dynamic stiffness) form, therefore joint admittances are inverted to obtain the dynamic stiffness matrices $\mathbf{Z}^R$ ($\mathbf{Z}^R = (\mathbf{Y}^R)^+$).

The output from the ML model for the task at hand is accurately defined (the rivet-squeezing force), therefore, a supervised ML approach will be applied. For supervised ML, the rivet-squeezing force corresponding to each training sample is logged. Joint models $\mathbf{Z}^R$, together with corresponding rivet-squeezing forces $\boldsymbol{F}_{sq}$ are stored into a set $\boldsymbol{J} = \left\{(\mathbf{Z}_i^R, \boldsymbol{F}_{sq,i}) \text{ for } i = 1, \ldots, n\right\}$, where n is a number of training set samples (Fig. 6).

The dataset $\boldsymbol{J}$ samples may vary in the quality of the riveted joint. For the vast majority of samples, a satisfactory quality is anticipated. Nevertheless, individual samples may contain larger errors due to faults in rivet material or irregularities in the riveting

process, causing large deviations in the joint's dynamic response that are unrelated to the rivet-squeezing force. Such samples do not provide relevant information for the training of the ML algorithm. In fact, their involvement in the training process may result in the ML model's poorer performance. Therefore, it is reasonable to exclude these samples (commonly referred to as outliers) from the dataset J. A simple approach for detecting the outliers for the observed problem will be presented in Section 3.1. Errors can also occur in the samples for which we would like to predict the rivet-squeezing force using the trained ML model. Inserting these samples into the ML algorithm will most likely result in erroneous predictions, therefore, a separate algorithm must be used to detect those samples. Dataset J joints passing through the outlier detection algorithm, represent the training set (Fig. 6).

In their original form, dynamic responses often contain large numbers of features, given by the product of the number of excitation DoFs, response DoFs, and frequency lines (which can be several thousand). Commonly, not all the features contain the same amount of information. To make the ML training process efficient, it is key to implement an efficient feature selection, that discards the features containing little or no information about the problem of interest (Fig. 6). Involving such features in the ML training may result in lower ML model accuracy due to the poor ratio between the number of training samples and the number of features.

With the final reduced training set prepared, a suitable ML algorithm can be selected, trained, and, finally, used to infer the rivet-squeezing force for new riveted joints.

In Sections 3.1 to 3.4, individual steps from the data preprocessing to the training of the ML algorithm will be further elaborated.

3.1. Data preprocessing

Raw experimental data often contain noise, outliers, and a large number of features that offer little or no information relevant to the task. Through careful preprocessing, these issues can be significantly mitigated, resulting in a more efficient and effective ML training process. The following section outlines the preprocessing steps applied to the dataset J for ML-based estimation of the rivet-squeezing force.

3.1.1. Impedance segmentation

Let $\mathbf{Z}_i^{\mathrm{R}} \in \mathbb{C}^{n_{\mathrm{resp}} \times n_{\mathrm{exc}} \times |\Omega|}$ be a sample of the training set J (where n_{resp} is a number of joint's response DoFs, n_{exc} is a number of joints excitation DoFs, and Ω is a set of discrete frequency points ω_j). Such a sample contains $n_{\mathrm{resp}} \times n_{\mathrm{exc}} \times |\Omega|$ values, each representing an individual dataset feature. Taking into account the complex-valued nature of the impedance matrix, the number of features doubles when treating the real and imaginary parts separately, or when transforming the data into amplitude and phase components. For a joint model with 12 response DoFs, 12 excitation DoFs, and 1000 frequency lines, this results in 288000 features. To reduce the dimensionality, one can exploit a key characteristic of the dynamic stiffness: its gradual variation with frequency. Typically, the rate of change is high in the lower frequency range (up to a few tens of Hz) and decreases at higher frequencies (Fig. 7). This means that neighboring frequency lines often carry redundant information. To take advantage of this, the frequency axis can be segmented, and representative values — such as the average amplitude and phase within each segment — can be computed. This reduces the feature count significantly while preserving essential information. For instance, segmenting a 12 × 12 × 1000 impedance matrix into 50 Hz segments reduces the feature count from 288000 to 5760.

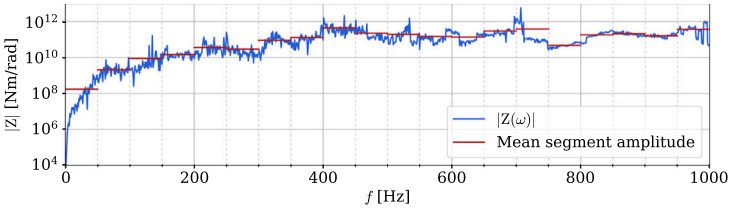


Fig. 7. Segmentation of the impedance Z connecting the moment around z-axis at joint's point of mass m_{A} to the rotations around z-axis at point of mass m_{B} (Fig. 5) for a riveted joint with corresponding rivet-squeezing force 5.80 kN.

The process of segmentation is conducted as follows. First, the complex impedance matrix of ith joint $\mathbf{Z}_i^{\mathrm{R}}$ (for $i = 1 \dots n$) is separated to the amplitude part $\left|\mathbf{Z}_i^{\mathrm{R}}\right|$ and phase part $\angle \mathbf{Z}_i^{\mathrm{R}}$. Then each of these parts is segmented into smaller frequency regions - set Ω is divided into N_Ω mutually exclusive subsets:

$$\Omega = \bigcup_{j}^{N_\Omega} \Omega_j \quad \text{with } \Omega_j \cap \Omega_k = \emptyset \ \forall j \neq k. \tag{10}$$

Amplitude and phase averages are then calculated as:

$$\mathbf{Z}^{\mathrm{R}}_{i,\mathrm{seg,ampl}} = \left[\frac{1}{|\Omega_k|} \sum_{\omega_j \in \Omega_k} \left|\mathbf{Z}_i^{\mathrm{R}}(\omega_j)\right| \quad \text{for } \Omega_k \text{ in } \Omega \right], \tag{11}$$

$$\mathbf{Z}^{\mathrm{R}}_{i,\mathrm{seg,phase}} = \left[\frac{1}{|\Omega_k|} \sum_{\omega_j \in \Omega_k} \angle \mathbf{Z}_i^{\mathrm{R}}(\omega_j) \quad \text{for } \Omega_k \text{ in } \Omega \right], \tag{12}$$

respectively. The resulting matrices $\mathbf{Z}^{\mathrm{R}}_{i,\mathrm{seg,ampl}} \in \mathbb{R}^{n_{\mathrm{resp}} \times n_{\mathrm{exc}} \times N_\Omega}$ and $\mathbf{Z}^{\mathrm{R}}_{i,\mathrm{seg,phase}} \in \mathbb{R}^{n_{\mathrm{resp}} \times n_{\mathrm{exc}} \times N_\Omega}$ are flattened into vectors and concatenated to form a single feature vector $\boldsymbol{x}_i$, where each element represents a dataset feature. For most ML algorithms, the

specific order of flattening is not important, as long as it is applied consistently across all samples. By obtaining feature vectors $\boldsymbol{x}_i$ for each of n dataset joints, this results in the dataset $\mathbf{X} \in \mathbb{R}^{n \times n_\mathrm{f}} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix}^\mathrm{T}$, where n_f is a number of features ($n_\mathrm{f} = 2 n_\mathrm{resp} n_\mathrm{exc} N_\Omega$).

3.1.2. Outlier detection

Next, potential outliers are identified and removed from the dataset. Due to inconsistencies in the riveting process, some riveted joints may exhibit lower quality. These joints — considered outliers — can show dynamic responses that differ significantly from the rest of the dataset for reasons unrelated to the applied rivet-squeezing force. Including such samples in the training process may negatively impact the performance of the ML model. Therefore, it is essential to detect and exclude them from the training set.

In this study, outliers are identified using the following procedure. First, a correlation between each column $\boldsymbol{X}_f$ (for $f = 1, \dots, n_\mathrm{f}$) of $\mathbf{X}$ (i.e. individual feature) and the squeezing force vector $\boldsymbol{F}_\mathrm{sq}$ is calculated:

$$\rho_f = \mathrm{corr}(\boldsymbol{X}_f, \boldsymbol{F}_\mathrm{sq}) \quad \text{for } f = 1, \dots, n_\mathrm{f}. \tag{13}$$

Only the features $\boldsymbol{X}_f$ that exhibit a correlation above a specified threshold ρ_min are used for outlier detection; i.e. features for which holds $\rho_f > \rho_\mathrm{min}$. This results in a reduced dataset $\mathbf{X}_\mathrm{red}$ with $n_\mathrm{f,red}$ columns.

Next, for each column $\boldsymbol{X}_{\mathrm{red},f}$ of the reduced dataset $\mathbf{X}_\mathrm{red}$, a linear function $\tilde{f}_f(\boldsymbol{F}_\mathrm{sq})$ between the rivet-squeezing force values $\boldsymbol{F}_\mathrm{sq}$ and the column values $\boldsymbol{X}_{\mathrm{red},f}$ is approximated:

$$\boldsymbol{X}_{\mathrm{red},f} \approx \tilde{f}_f(\boldsymbol{F}_\mathrm{sq}) \boldsymbol{F}_\mathrm{sq} \quad \text{for } f = 1, \dots, n_\mathrm{f,red}. \tag{14}$$

The use of a linear model is motivated by the fact that correlation quantifies the strength of a linear relationship between variables. Since the purpose of this function is solely to identify samples that significantly deviate from the trend (i.e., outliers), it is not necessary for the approximation to model the actual relationship with high accuracy. To assess how well each sample aligns with the estimated trends $\tilde{f}_f(\boldsymbol{F}\mathrm{sq})$, the absolute distance d between the actual and approximated feature values is computed for each sample and each retained feature:

$$d_{i,f} = \left| X_{\mathrm{red},i,f} - \tilde{f}_f(\boldsymbol{F}_\mathrm{sq}) F_{\mathrm{sq},i} \right| \quad i = 1, \dots, n; \; f = 1, \dots, n_\mathrm{f,red} \tag{15}$$

Thus, for each sample, an array of distances is obtained. For each segment, a maximum allowed distance $d_{f,\mathrm{max}}$ between the actual value and approximation is defined as the mean distance plus the standard deviation of the segment distances of all samples, that is, $d_{f,\mathrm{max}} = \mu(\boldsymbol{d}_f) + \sigma(\boldsymbol{d}_f)$ for $f = 1, \dots, n_\mathrm{f,red}$, where $\boldsymbol{d}_f$ is a vector of distances for fth feature. This way, the limit value automatically adapts to the size and the scatter of distance values for each feature. It is then determined, for each feature of each sample, whether the distance is smaller than the maximum allowed distance (i.e. $d_{i,f} < d_{f,\mathrm{max}}$ for $f = 1, \dots, n_\mathrm{f,red}$). Each sample is not required to meet this criterion for all segments, but rather for a majority of them. For example, a sample identified as acceptable must meet the $d_{i,f} < d_{f,\mathrm{max}}$ criterion in 90% (or some other empirically determined percentage) of segments. On the other hand, the samples not meeting this condition are identified as outliers and discarded. The output of the outlier detection step is the dataset $\mathbf{X}$, retaining all n_f features but excluding outlier samples.

3.1.3. Normalization

Another important aspect of data preparation is normalization. Whether normalization is required depends on the specific ML algorithm and the numerical solvers used during model training. For instance, algorithms that rely on gradient-based optimization typically benefit from normalized data, as it can improve convergence speed and overall training stability.

3.2. Feature extraction

Now that initial preprocessing has been completed and outliers have been removed, the next step is to analyze the features of dataset $\mathbf{X}$ and extract those most relevant to the task. In this work, two classical feature extraction techniques are applied — Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) — along with a third, custom approach based on the correlation between features and the rivet-squeezing force.

PCA is an unsupervised feature extraction technique, meaning it does not require class labels. It transforms the original feature space into a set of orthogonal components (principal components), where the first few components capture the directions of highest variance across the dataset. By discarding components with low variance, PCA allows for dimensionality reduction while preserving most of the relevant information [44]. However, this method has a key limitation: since it is unsupervised, the identified directions of variance may not correspond to meaningful variations in the rivet-squeezing force.

LDA, on the other hand, is a supervised feature extraction method typically used for classification problems. It projects the original feature space into a lower-dimensional subspace where the separation between class means is maximized, and within-class variance is minimized [44]. An important advantage of the LDA over PCA is that it classifies as a supervised feature extraction approach. This means that such a set of features is extracted that is relevant for the provided set of labels (i.e. squeezing forces corresponding to the training data). However, LDA also has a notable limitation: it reduces the dataset to the (up to) number of classes minus one. In regression tasks, where the number of unique target values may be large, this can result in excessive dimensionality reduction and potential loss of important information.

To address the drawbacks associated with PCA and LDA, a third approach is tested within our study. This method, referred to here as the segment-correlation approach, relies on the correlation between individual features and the target squeezing forces—similar

to the strategy used for outlier detection in Section 3.1.2. Since each feature in $\mathbf{X}$ represents either the mean amplitude or phase within a frequency segment of $\mathbf{Z}^{\mathrm{R}}$, computing the correlation between each feature $\boldsymbol{X}_f$ and the squeezing force vector $\boldsymbol{F}_{\mathrm{sq}}$ provides a straightforward means of evaluating feature relevance:

$$\rho_f = \mathrm{corr}(\boldsymbol{X}_f, \boldsymbol{F}_{\mathrm{sq}}) \text{ for } f = 1, \ \ldots, \ n_{\mathrm{f}}. \tag{16}$$

Features with correlation coefficients below a specified threshold $\rho_{\min}$ are discarded (i.e., those for which $\rho_f < \rho_{\min}$), resulting in a further reduced feature set. This approach offers considerable flexibility: a lower threshold preserves more features and, therefore, more information—though potentially at the cost of including less relevant data. Conversely, a higher threshold results in fewer, more strongly correlated features, which can be beneficial when only a limited number of training samples are available. Selecting an appropriate threshold $\rho_{\min}$ thus represents a trade-off between preserving information and controlling model complexity.

3.3. Considerations regarding the selection of the ML model

The dataset, now consisting of a refined set of relevant features and cleared of outliers, can be used to train an ML model capable of inferring the rivet-squeezing force based on joint dynamic behavior. However, a wide range of ML algorithms is available—each with its own characteristics, advantages, and limitations depending on the specific nature of the problem.

To proceed, let us first specify the ML task. The input to the ML model is the dataset $\mathbf{X}$, containing a few hundred or fewer training samples and anywhere from a few to several hundred features, depending on the parameters chosen during preprocessing and feature extraction. The model output is a single, continuous value representing the rivet-squeezing force. Given this clearly defined mapping between inputs and outputs, the problem is categorized as a supervised learning task.

Due to the relatively small number of training samples, it is preferable for the ML problem to be as simple as possible. Therefore, before selecting a specific ML approach, potential simplifications should be considered. The problem described above is a regression problem with infinite possible outputs. This problem can be simplified by transforming it into a classification problem with a limited number of output classes. Each class represents an interval of the squeezing force values. The number of ML model outputs is thus substantially reduced. However, this comes at the expense of the precision with which the ML model predicts the rivet-squeezing force. The trade-off between the simplicity of the problem and the precision of the squeezing force prediction can be observed even within the classification domain. The larger the number of classes, the higher the precision of the prediction, and vice versa.

Next, a suitable classification algorithm is chosen. For a small number of training samples, complex ML models with large numbers of parameters are likely to overfit. Nevertheless, studies in different fields have shown that a successful classification on small datasets is possible. Cilia et al. [45] used the DARWIN dataset [46] (174 samples, 451 features) to classify individuals as either Alzheimer's disease patients or healthy controls based on various handwriting characteristics. Different classifiers were tested, achieving accuracies ranging from 78.57 and 88.29%. Yilmaz & Sekeroglu [47] used the Higher Education Students Performance Evaluation dataset [48] (145 samples, 31 features) to predict student performance (8 grades as possible classification outputs) based on information like former education data, accommodation type, transportation type, etc. Radial-Basis Function Neural Network was applied, achieving up to 88% accuracy. Tussel-Rey et al. [49] used the Non Verbal Tourists dataset (73 samples, 22 features) to predict tourists' communication preferences (6 classes) based on the questionnaire data. A Customized Naïve Associative Classifier was utilized, achieving 80.76% accuracy.

When selecting the classification algorithm, the most straightforward options are the classical classification approaches like Decision Trees (DT) and their ensembles, Support Vector Machines (SVMs), k-Nearest Neighbors (kNN), etc. What these algorithms lack is the ability to take into account that the force intervals are related to each other (e.g. 5–6 kN force interval is closer to the 6–7 kN interval than to the 7–8 kN interval). To take advantage of these relations, E. Frank and M. Hall [1] proposed an extension of classical classification algorithms to the ordinal classification (OC) problems,[4] resulting in improved performance in the classification tasks. Furthermore, the proposed classification approach divides the described classification problem into multiple binary classification problems and thus additionally reduces the complexity of the problem. The proposed approach to ordinal classification will be summarized in the following section. Another way to deal with the interval problem is to train a regression model and perform an interval mapping. This means a continuous value is obtained from the ML model and then assigned to the corresponding force interval.

In the case-study (Section 4), different approaches will be tested to identify the most suitable one.

3.3.1. A simple approach to ordinal classification

The approach to the ordinal classification proposed by E. Frank and M. Hall [1] allows standard classification algorithms to utilize class order in the ordinal classification problems. In the process, the problem is divided into multiple binary classification problems. Therefore, the selection of the classification algorithms is limited to the binary classifiers allowing for the calculation of the probability for each output class.

Assume the goal of the multi-class problem is to classify a sample into one of k unique classes that can be logically ordered one after another. Let these output classes be stored in a set $A \in \{\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_k\}$ where index of each class represents the location of

[4] Ordinal classification problems are a superset of interval classification problems where there is an inherent ordering between the classes, but there is not necessarily a meaningful numeric difference between them [50].

the class in the order. For training the ML algorithm, there are n samples available, comprising the dataset $\mathbf{X}$ and the corresponding labels $\boldsymbol{y}$.

Such multi-class problem can be transformed into $k-1$ binary classification problems. Labels $\boldsymbol{y}$ of ith binary problem are transformed such that jth sample ($j = 1, \ldots, n$) is assigned a new label y_j^* with value 0 if $y_j \leq \mathrm{C}_i$ and value 1 if $y_j > \mathrm{C}_i$. Thus, $k-1$ datasets are obtained on which $k-1$ binary classifiers are trained on.

The $k-1$ trained classifiers are used to make a prediction for a new sample described by data $\boldsymbol{X}'$. Inserting $\boldsymbol{X}'$ into each of $k-1$ classifiers, each classifier returns a prediction in form of probability for each output class (for ith classifier i.e. $\Pr(y' \leq \mathrm{C}_i)$ for class 0 and $\Pr(y' > \mathrm{C}_i)$ for class 1 where y' is a label for the sample of interest). Using these outputs, the probability for each original class label in set A is computed as follows:

$$\Pr(C_1) = 1 - \Pr(y' > C_1), \tag{17}$$

$$\Pr(C_i) = \Pr(y' > C_{i-1}) - \Pr(y' > C_i); \quad 1 < i < k, \tag{18}$$

$$\Pr(C_k) = \Pr(y' > C_{k-1}). \tag{19}$$

The prediction $\tilde{y}'$ then equals the class with highest probability:

$$\tilde{y}' = \arg\max_{c \in A} \Pr(c). \tag{20}$$

In the following this approach will be denoted with OC ($\star$), where $\star$ denotes the classification algorithm (e.g. logistic regression, LDA, SVM etc.) applied within the described classification approach.

3.4. Considerations regarding evaluation of the ML model

To obtain meaningful insight into the ML model's performance, a suitable metric must be utilized for the model validation. For the interval classification problem at hand, the following demands were set: (1) the metric must allow for the evaluation of multi-class classification models, (2) metric must account for the imbalanced class distribution (i.e. there are more samples in some classes than in others within the validation set), (3) the metric must account for interval relations between classes (i.e. some classes are more distanced from the reference class than others), and (4) the metric's output must be restricted to some limited interval for easier interpretation of the result. In this case, simple metrics like classification accuracy do not suffice. Hence, some more advanced metrics are required. While commonly used classification metrics like precision, recall, and F1 score can be adapted to account for the imbalanced class distribution and extended to multi-class problems, these metrics do not account for the distance between the prediction and the reference class. To account for the distance between classes, regression metrics like the weighted mean absolute error and weighted mean squared error[5] can be adapted for the interval classification problems. However, these metrics do not provide results on a limited interval, which complicates interpretation of model performance.

One metric that satisfies all these criteria is the Quadratic Weighted Kappa (QWK). QWK allows us to assess the performance of the ordinal classifier. The class imbalance and the distance between the predictions and reference classes for individual validation samples are accounted for. The quadratic version of weighted kappa penalizes the misclassifications, weighted by the squared distance from the true class. In addition, output values of QWK are limited to the interval between −1 and 1, where 1 represents perfect agreement between predictions and labels, 0 represents the agreement expected by chance, and values below 0 represent agreement worse than expected by chance.

Besides the evaluation metric, it is important to take a suitable approach to the model validation. When dealing with a small dataset, a simple train-test split may not be optimal as we end up with only a small number of training samples or, on the other hand, only a small number of testing samples. This would result in an unrealistic assessment of the ML model's performance. A common way to deal with this problem is the utilization of Cross-Validation (CV). In CV, the dataset is split into multiple parts or folds. Validation is then performed iteratively. In each iteration, an ML model is trained on all but one fold and tested on the remaining one. The number of iterations equals the number of folds.[6] If the number of folds is smaller than the number of samples and larger than 2, the validation process is called a k-Fold Cross-Validation (k-Fold CV). This approach will be used for the validation of the model trained for the estimation of the rivet-squeezing force.

4. Case study

The ML-based approach for estimating the rivet-squeezing force was evaluated using a laboratory case study. This case study encompasses the entire workflow—from data acquisition through experimental measurements, to data preprocessing, ML model training, and finally, validation of the model and comparison with an existing methodology for rivet-squeezing force identification.

Throughout the process, several key decisions had to be made, including the selection of outlier detection parameters, the feature extraction method, and the choice of ML algorithm. These decisions, along with the rationale behind them, will be discussed in the following sections. Before that, however, details of the experimental structure and the setup used for data collection will be presented.

[5] Weighted versions of these metrics need to be used to account for imbalanced class distribution.

[6] In the extreme case when number of folds equals the number of training set samples, the ML model is validated on only one sample in each iteration. Such special case of CV is referred to as Leave-One-Out Cross-Validation (LOOCV).

4.1. Test structure

The test structure consists of two metal sheet plates, labeled A and B. These plates are joined using a rivet to create two distinct assembly configurations, each connected through a different set of holes (Fig. 8). The first configuration — denoted as ARB in Fig. 8 — is used for training the ML model and evaluating its performance. It will be referred to as the original assembly in the following text. Note that only a larger number of samples from this assembly is required for the ML model training, regardless of the assembly for which we would like to predictions of the rivet-squeezing force. The second configuration, denoted as ARB', is thus used for testing and validating the trained ML model only. This configuration will be referred to as the modified assembly.

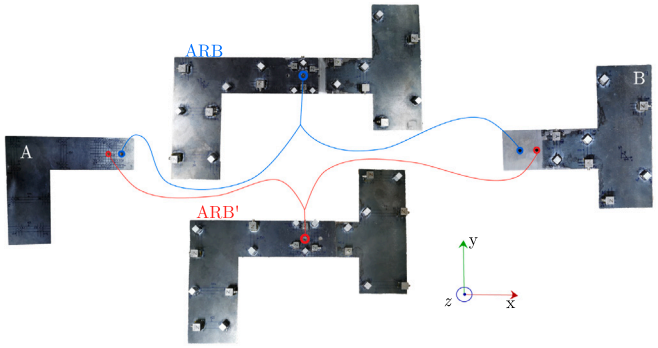


Fig. 8. Assembling the metal sheet plates A and B into assemblies ARB and ARB'.

Although configurations ARB and ARB' may appear similar at first glance, they differ significantly from a dynamic perspective. These differences are evident in the locations of the resonance peaks (i.e., natural frequencies), as shown in Fig. 9.[7] These differences in assembly dynamics may cause issues when estimating rivet-squeezing force for joints obtained from different assemblies (within the aforementioned limitations, i.e. the material, geometrical, and frictional properties close to the joint must remain similar). These challenges arise from the inherent limitation of dynamic decoupling: the dynamics of the parent structure (i.e., the full assembly) may not be completely removed. Although the extended interface decoupling formulation significantly reduces residual parent dynamics, it does not fully eliminate them in most cases. Consequently, despite involving the same set of substructures, the residual dynamics captured within the joint impedances of ARB and ARB' can differ substantially. This difference poses a major obstacle when attempting to apply an ML model trained on ARB joint data to joints from the ARB' configuration.

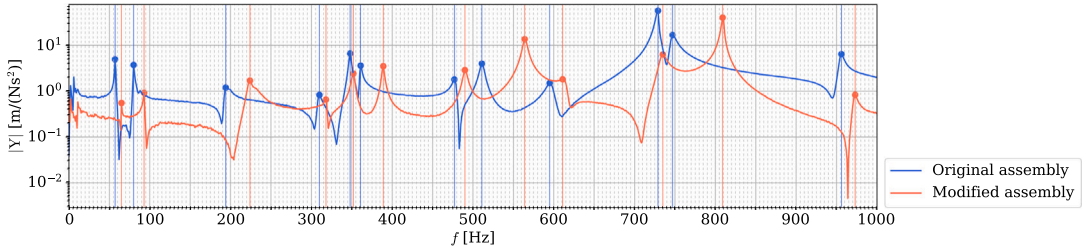


Fig. 9. Comparison of FRFs of the original assembly ARB and modified assembly ARB'.

The rationale for using the same set of metal sheet plates A and B was to ensure that the material, frictional, and geometrical properties in the vicinity of the joint remained consistent across all samples. However, the contact surfaces for each training set sample were prepared manually, which introduced some degree of variability in the surface conditions between individual samples.

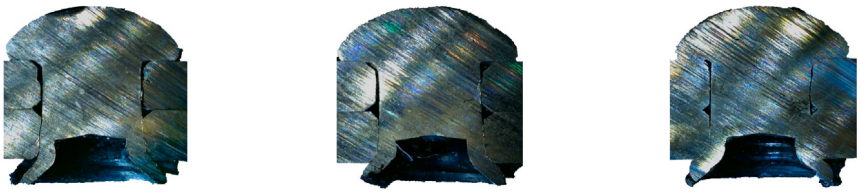
The metal sheet plates were joined using steel semi-tubular rivets (DIN 660B), with a diameter of 5 mm and a length of 10 mm (Fig. 10a). The riveting process was carried out using a hydraulic riveting device equipped with an integrated pressure sensor (Fig. 10b). The rivet-squeezing forces were calculated from the measured pressure values, using the piston cross-sectional area derived from the manufacturer-specified piston diameter of 56 mm.

Three pairs of metal sheet plates were riveted together with different corresponding squeeze forces of 4.39 kN, 6.51 kN, and 8.82 kN, which represent the lowest, middle, and the highest rivet-squeezing forces used to build the dataset in the following steps of the case study. To indicate the potential differences in joint geometry due to the rivet-squeezing force, the cross-section images of these samples are displayed in Fig. 11. It can be observed that the increased squeezing force results in a higher radial extension of the rivet's stem in the hole. Thus, at higher rivet-squeezing forces, the gaps between the rivet and the hole walls reduce, and the rivet fills the hole better. This changes the contact between the rivet and the metal sheet plates, which contributes to the differences in the dynamic response of the riveted joint at different rivet-squeezing forces.

[7] FRFs in Fig. 9 connect different sets of DoFs for original and modified assembly. Therefore, amplitudes of FRFs cannot be compared, however, differences in the natural frequencies give an insight into how different assemblies are.

(a) (b)

Fig. 10. Components of the riveting setup: (a) semi-tubular rivet (DIN 660B), (b) hydraulic riveting device with pressure sensor.



(a) (b) (c)

Fig. 11. Cross-sections of riveted joints riveted with: (a) 4.39 kN, (b) 6.51 kN, and (c) 8.82 kN squeezing force.

4.2. Experimental setup

The FBS-based joint identification requires dynamic models (i.e., admittance matrices) of the assembly and the joint-adjacent assembly components. VPT was applied on each model to obtain rotational DoFs at the interface. To measure the dynamic responses close to the joint, required for the VPT, three three-axial accelerometers (Dytran 3273A2, PCB 356A32, and PCB 356A45) were placed close to the riveted joint at each metal sheet plate (A and B in Fig. 9). Besides, four additional three-axial accelerometers were attached to each metal sheet plate away from the interface for the purpose of the extended interface decoupling. Metal blocks were attached to assembly components (Fig. 12), allowing us to excite the object in directions parallel to the surface of the plates. Impact blocks are distributed similarly to accelerometers, i.e., three blocks close to the interface and four away from the interface at each assembly component. Modal hammer PCB 086E80 was used to apply force at different impact locations. During the measurements, structures were suspended from elastic strings to imitate the free boundary conditions (Fig. 12).

It is expected that the rotations around the z-axis (Fig. 8) are particularly important for the rivet-squeezing force estimation (see Section 2.2). To account for this, the structures were excited, and responses were observed only in directions parallel to the metal plate surfaces, utilizing two of the three accelerometer axes. Additionally, this considerably reduced the measurement efforts and the amount of measurement data. All accelerometers were oriented such that the active axes coincided with the global x- or y-axis (Fig. 8). Excitation forces were applied at an angle of 45° relative to the global x and y axes (Fig. 8), ensuring that both active accelerometer channels received sufficient excitation, resulting in a high signal-to-noise ratio.

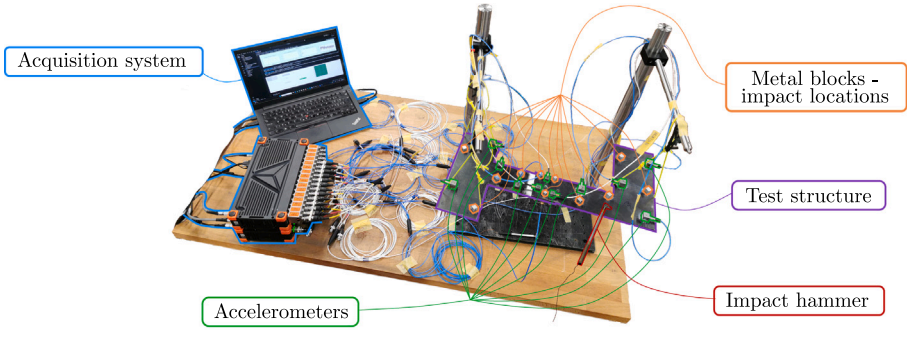


Fig. 12. Experimental setup.

4.3. Building the joint dataset

The FRFs required to build the dataset were measured using the experimental setup described in Section 4.2. Three admittance matrices, containing a set of internal and interface DoFs, were needed for an individual sample: $\mathbf{Y}^{\mathrm{A}} \in \mathbb{C}^{14\times14\times2000}$ (14 response channels, 14 excitation channels and 2000 discrete frequency values), $\mathbf{Y}^{\mathrm{B}} \in \mathbb{C}^{13\times14\times2000}$ and $\mathbf{Y}_i^{\mathrm{ARB}} \in \mathbb{C}^{27\times28\times2000}$ (for $i = 1 \dots 120$). The measured response/impact DoFs are displayed in Fig. 13. On these admittances, VPT was performed to obtain the rotational DoFs at the interface. VPT was applied once for each of components A and B and twice for the assembled system - a VP at the contact between the rivet and substructure A and a VP at the contact between the rivet and substructure B. 6 DoFs were observed in the vicinity of the interface (3 in the x-direction and 3 in the y-direction; Fig. 8) to be transformed into a single VP with 3 response DoFs (translation in the x-direction, translation in the y-direction, rotation around the z-axis) and 3 excitation DoFs (force in the x-direction, force in the y-direction, moment around the z-axis). The following transformed admittances were obtained: $\mathbf{Y}^{\mathrm{A}}_{\mathrm{VPT}} \in \mathbb{C}^{11\times11\times2000}$, $\mathbf{Y}^{\mathrm{B}}_{\mathrm{VPT}} \in \mathbb{C}^{10\times11\times2000}$ and $\mathbf{Y}^{\mathrm{ARB}}_{\mathrm{VPT},i} \in \mathbb{C}^{21\times22\times2000}$ (for $i = 1 \dots 120$).

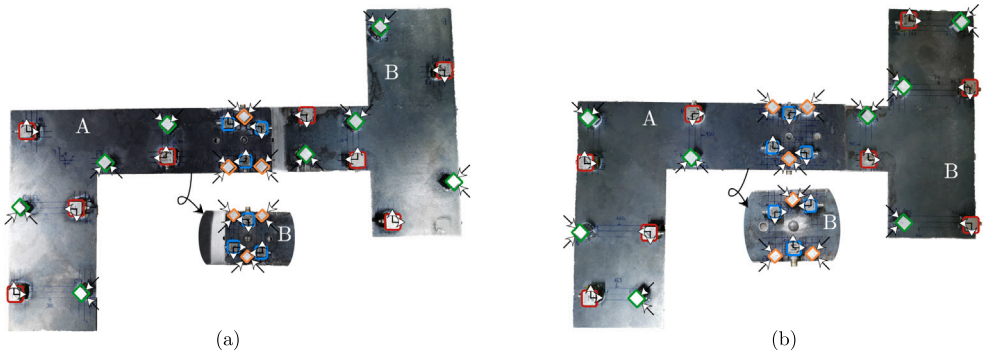


Fig. 13. The observed internal response locations (•), internal impact locations (•), interface response locations (•), and interface impact locations (•) at: (a) original assembly ARB, (b) modified assembly ARB'. Arrows at each response/impact location denote the observed/controlled DoFs at that location. Note that the cutout at each structure represents the mirrored image of the bottom side of the structure, so that the DoFs are oriented as they were viewed from above.

Next, the FBS-based joint identification was performed on the transformed admittances. Substructures A and B are decoupled from the assembly ARB using the extended interface decoupling formulation. The ill-conditioning of the interface matrix $\mathbf{B}_{\mathrm{e}}\mathbf{Y}\mathbf{B}_{\mathrm{e}}^{\mathrm{T}}$ in the LM FBS Equation (3) was coped with by utilizing the Moore–Penrose inverse to determine the interface matrix inverse. In the process, TSVD was applied to discard insignificant singular values, i.e., singular values for factor 1e2 (or more) smaller than the highest singular value at each discrete frequency point. Thus, only a few dominant singular values were kept. Finally, the joint admittance $\mathbf{Y}_i^{\mathrm{R}} \in \mathbb{C}^{6\times6\times2000}$ is extracted from the decoupled matrix $\widetilde{\mathbf{Y}}_i \in \mathbb{C}^{42\times44\times2000}$ using Eq. (5). The resulting joint admittance matrices were twice integrated (i.e., divided by $-\omega^2$) to obtain receptance from accelerance, and then inverted to obtain the dynamic stiffness matrix of the joint $\mathbf{Z}_i^{\mathrm{R}}$.

For the supervised ML algorithms, a record of the squeezing forces F_{sq} must be kept. Squeezing forces were calculated based on the hydraulic pressure measured on the riveting device (see Section 4.1). A balanced distribution of the squeezing forces within the predetermined interval was ensured by determining the approximate sample pressure values in advance. Samples were drawn from the uniform distribution within the 20 to 35 bar interval (i.e., squeezing forces from 4.93 kN to 8.62 kN). The actual pressures measured during riveting deviated slightly from the predetermined values. The pressures measured during the riveting represented the reference values. Finally, the resulting joint dynamic stiffness matrices and the corresponding reference squeezing force values form a dataset $J = \left\{(\mathbf{Z}_i^{\mathrm{R}} \in \mathbb{C}^{6\times6\times2000};\ \mathrm{F}_i) \text{ for } i = 1 \dots 120\right\}$.

The validation set for joints from the modified assembly ARB' (Fig. 8) was obtained using the same procedure as for the original assembly, resulting in a validation dataset $J_{\mathrm{v}} = \left\{(\mathbf{Z}^{\mathrm{R}}_{\mathrm{v},i};\ \mathrm{F}_{\mathrm{v},i}) \text{ for } i = 1 \dots 20\right\}$.

4.4. Data preprocessing, outlier detection and feature selection

The data collected and stored in dataset J were processed as proposed in Sections 3.1 and 3.2. The following steps were carried out: (1) Joint matrices were segmented to frequency intervals of size of choice, separated into the amplitude and phase part, and flattened (see Section 3.1.1). (2) Outliers were detected and discarded (see Section 3.1.2). Within this step, two parameters, the minimum correlation $\rho_{\mathrm{min},1}$ between each feature and the rivet-squeezing force, and the minimum percentage of features per sample that satisfy the criterion $d_f < d_{f,\max}$ (for $f = 1, \dots, n_{\mathrm{f}}$), had to be selected. (3) Feature space was reduced using the three feature extraction approaches described in Section 3.2 (PCA, LDA, and segment-correlation approach). For PCA, the principal components to keep were determined based on the minimum allowed correlation $\rho_{\mathrm{min},2}$ between each principal component and the rivet-squeezing force. The minimum correlation parameter $\rho_{\mathrm{min},2}$ also had to be determined for the segment-correlation approach. (4) The resulting dataset $\mathbf{X}$ was normalized by setting the mean value of each feature to 0, scaling each feature to the range [−1, 1]:

$$\boldsymbol{X}_{f,\mathrm{norm}} = \frac{\boldsymbol{X}_f - \mu(\boldsymbol{X}_f)}{\max(|\boldsymbol{X}_f|)}; \quad \text{for } f = 1, \dots, n_{\mathrm{f}}. \tag{21}$$

The parameters of each of the described steps were determined by conducting a grid search. Table 1 shows lists of tested values for each parameter. Different combinations of parameter values resulted in training sets with varying numbers of training samples and features. For each parameter configuration, multiple ML models were trained and evaluated. Suitable parameters were selected for each ML algorithm separately. The following sections describe the training process, model evaluation, parameter selection based on the grid search, and the final selection of the most suitable ML algorithm.

Table 1
Data preprocessing and feature extraction grid search values.

Parameter	List of values
Frequency interval size [Hz]	25, 50, 100, 200
$\rho_{\min,1}$	0.05, 0.1, 0.15, 0.2
Min. percentage of features meeting condition: $d_f < d_{f,\max}$ [%][a]	60, 70, 75, 80, 83, 87, 90, 93
$\rho_{\min,2}$[a]	0.1, 0.15, 0.2, 0.25

[a] Parameters used in segment-correlation approach and PCA, but not in LDA.

4.5. ML hyperparameter selection

Different ML approaches were tested: SVM with radial basis function (RBF) kernel, DT, kNN, OC (⋆) approach [1] with different binary classifiers (⋆), and ridge regression with interval mapping. Each ML algorithm has some model-tuning parameters (referred to as hyperparameters). To get the most out of each algorithm and to ensure a fair comparison of ML algorithms, the grid search was extended by testing different hyperparameter setups besides the data preprocessing and feature extraction parameters. Table 2 shows the hyperparameters of the tested ML algorithms. Table 3 displays the tested values of individual parameters. The *scikit-learn* [51] implementations of the tested ML algorithms were utilized.

Table 2
ML algorithm hyperparameters.

ML algorithm	Hyperparameters
SVM with RBF kernel	Regularization parameters[a], RBF kernel's γ parameter
DT	None[b]
kNN	Number of nearest neighbors
Logistic regression	Regularization parameters[a]
LDA	None
Ridge regression with interval mapping	Regularization parameters[a]

[a] Type of regularization, amount of regularization C, and ratio between L1 and L2 regularization.
[b] Training is finished when all leaves are pure[c].
[c] Leave is pure when it contains only samples of a single class.

Table 3
ML algorithm hyperparameter values used within grid search.

Hyperparameter	List of values
Type of regularization	L1, L2, or elastic net
Ratio between L1 and L2 regularization	0.2, 0.4, 0.6, and 0.8[a]
Amount of regularization C	0.001, 0.005, 0.01, 0.05, 0.1, 0.5, and 1
RBF kernel's γ parameter	selected automatically: $\gamma = 1/(n_t \sigma^2(\mathbf{X}))$
Number of nearest neighbors	1, …, 10

[a] Only for elastic net regularization.

4.6. ML model evaluation and validation

In each iteration of the grid search, multiple ML models were trained and evaluated. Fig. 14 illustrates a single iteration. First, the data — joint admittances from both the original and modified assemblies (ARB and ARB', respectively) — were preprocessed according to the parameter set corresponding to that iteration. This included outlier detection and feature extraction. The performance of each ML model, configured with a specific set of hyperparameters, was then assessed using 5-fold CV with 10 repetitions. For each repetition, the data were randomly split into new folds. Each fold contains approximately 20% of the dataset samples (depending on the divisibility of the number of dataset samples). During each CV iteration, the model was trained on four of the five folds (80% of the dataset samples) and tested on the remaining one (20% of the dataset samples; Fig. 14). In addition to the CV, model performance was also evaluated using the validation dataset from the modified assembly. After completing the CV, the mean QWK was computed across all iterations. Importantly, QWK values for the original and modified assemblies were calculated separately to evaluate the model's performance across different structural configurations.

Thus, ML algorithms were trained on joints from the original assembly and tested on joints from both the original and modified assemblies. The ML algorithm was then selected based on its QWK score on the test data. Although the test data were not used during training, they were involved in parameter tuning and model selection. As a result, the test QWK does not provide an unbiased estimate of model performance, and further evaluation is necessary. To address this, the data were additionally split to create a separate group of samples, referred to as the validation set (Fig. 15). The validation data were processed using the same parameters as those used for the training and test sets. For final model validation, 5 joints from the original assembly and 5 from the modified

assembly were set aside. The validation samples were selected to ensure that each rivet-squeezing force class was represented in the validation set for both assemblies.

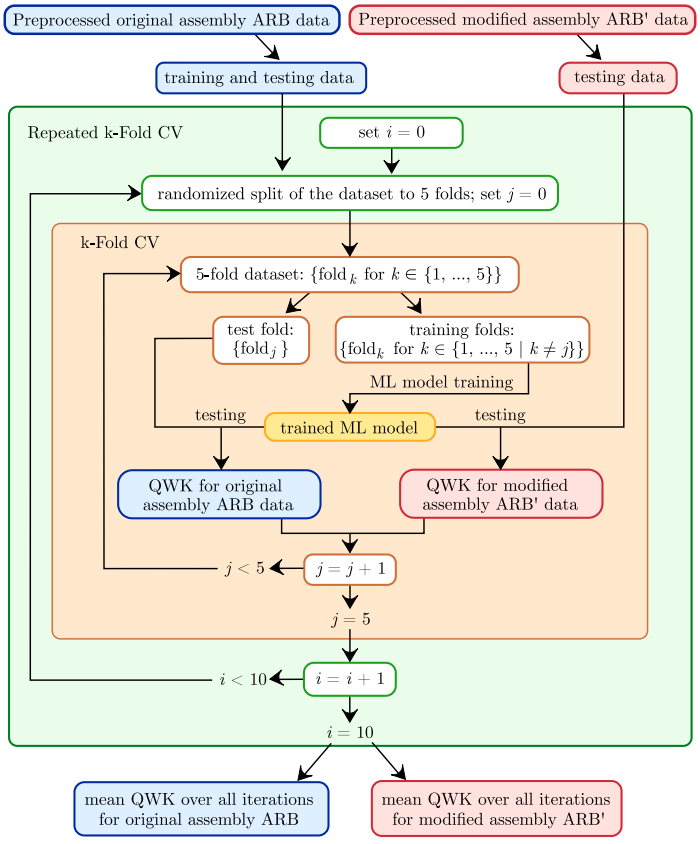


Fig. 14. Evaluation of ML model performance for individual parameter configuration within grid search.

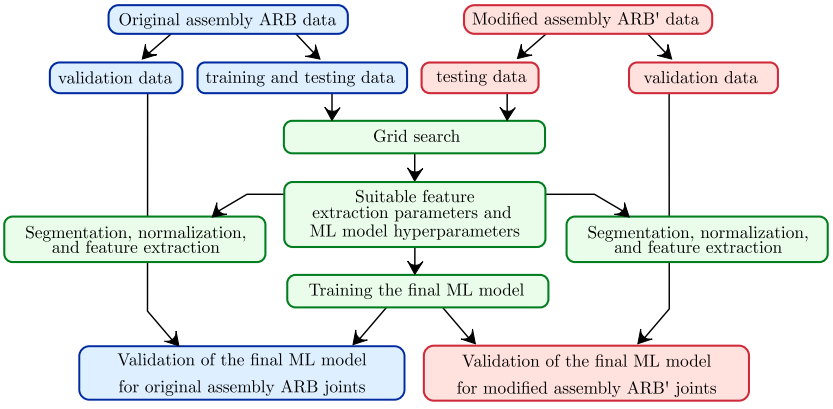


Fig. 15. Process of obtaining and validation of the ML model for identification of the rivet-squeezing force.

4.7. Results

Three feature extraction approaches were tested within the grid search (segment-correlation approach, PCA, and LDA). Fig. 16 shows the performance of the OC (logistic regression) model[8] for the tested methods. Feature extraction parameters for each ML model in Fig. 16, along with the resulting training set sizes, are provided in Table 4. Fig. 16a shows the test QWK values for models

[8] Note that the use of other ML algorithms resulted in different QWK values; however, the same conclusions regarding the suitability of individual feature extraction approaches were observed.

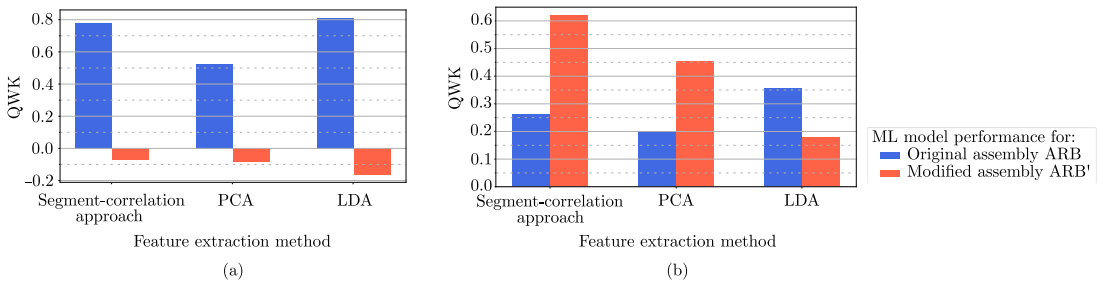


Fig. 16. ML model, i.e. OC (log. reg.) performance for different feature extraction approaches. ML models (a) with highest QWK for original assembly joints, and (b) with highest QWK for modified assembly joints.

Table 4
Segment-corelation, PCA, and LDA configurations returning highest QWK for either original or modified assembly joints in grid search OC (logistic regression) classifiers.

Highest QWK for	Original assembly			Modified assembly		
Feature extraction approach	Segment-correlation	PCA	LDA	Segment-correlation	PCA	LDA
Frequency interval size [Hz]	25	200	200	200	200	200
Min. correlation 1	0.05	0.05	0.1	0.15	0.15	0.2
Min. correlation 2[a]	0.2	0.15	/	0.25	0.2	/
Penalty	L2	L2	elastic net	L2	L2	L2
C	1.0	1.0	0.5	0.5	1.0	0.01
L1 ratio	/	/	0.4	/	/	/
Min. percentage of features meeting condition: $d_f < d_{f,\max}$ [%][a]	87	70	75	83	83	83
No. training samples	82	119	118	94	94	83
No. features	258	8	4	2	2	4

[a] Parameters used in segment-correlation approach and PCA, but not in LDA.

that performed best on joints from the original assembly ARB, while Fig. 16b shows the QWK values for models that performed best on the modified assembly ARB'. In Fig. 16a, it can be observed that these models perform poorly on the modified assembly. Conversely, the models in Fig. 16b perform considerably better on the modified assembly, but their QWK values for the original assembly are significantly lower.

These results suggest that the ML model's predictive performance on joints from the modified assembly can be improved at the expense of its performance on the original assembly. Therefore, it is essential that the QWK values are evaluated separately for the original and modified assemblies, as joint properties may vary across different configurations. This is due to the fact that, despite using the extended interface decoupling technique, residual parent dynamics may persist, and these differ between assemblies. Consequently, the preprocessing and feature extraction configuration that yields the best results for the original assembly does not necessarily ensure accurate force prediction for joints from other assemblies. Note that this does not necessarily indicate model overfitting. Rather, the improved performance may result from using a larger portion of the features, which contain more information specific to the original assembly. To develop a generally applicable ML model, the preprocessing and feature extraction parameters should be selected based on grid search results in such a way that the returned feature set is not dependent on the parent assembly.

Both charts in Figs. 16a and 16b confirm the concerns regarding PCA and LDA that were outlined in Section 3.2. PCA transforms the original feature space such that the first principal component corresponds to the linear combination of original features exhibiting the highest variance across the dataset. The variance captured by each subsequent component decreases. However, the features with the highest variance are not necessarily correlated with variations in the rivet-squeezing force, which may limit the effectiveness of selecting the first few principal components for this task. Additionally, since some features may contain residual effects from the parent assembly dynamics, the PCA transformation may distribute this non-generalizable information across all components. As a result, poorer performance for the modified assembly was anticipated. These drawbacks are evident in Figs. 16a and 16b, where PCA consistently produced the lowest QWK scores for joints from the original assembly. Fig. 16b further shows that, although PCA slightly outperformed LDA for the modified assembly, both were outperformed by the segment-correlation approach.

On the other hand, LDA-based feature extraction has the drawback of performing very aggressive dimensionality reduction. In addition, LDA is a supervised method, meaning that the extracted features are highly correlated with the specific training data—in this case, the original assembly. For the current problem, only four features were retained after applying LDA (Table 4). As a result, LDA yielded strong performance on the original assembly data (Figs. 16a and 16b). However, even the highest QWK obtained for the modified assembly (Fig. 16b) was notably lower than those achieved by the other two feature extraction methods. Based on these results, the segment-correlation approach was identified as the preferred feature extraction method.

Next, the most suitable ML algorithm, along with its corresponding hyperparameters, was selected. Figs. 16a and 16b illustrate two extreme cases: one in which the model performs well on the original assembly and another in which it performs well on the

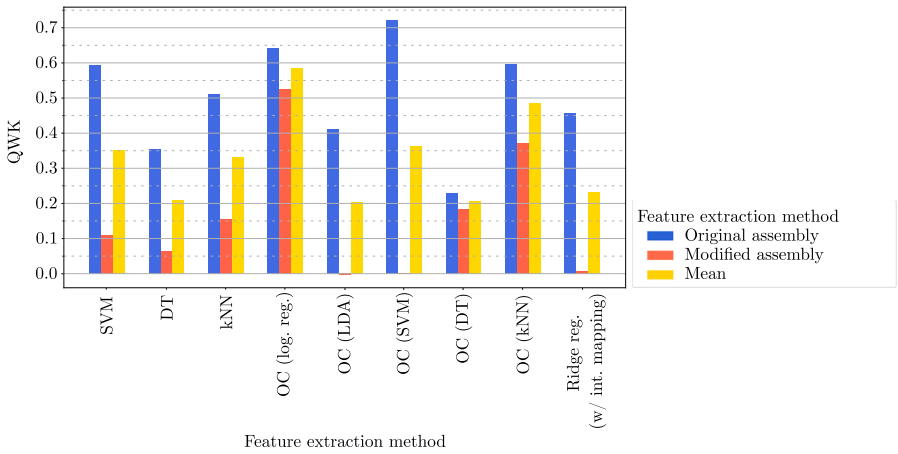


Fig. 17. Performance of various ML models in task of rivet-squeezing force identification.

Table 5
Grid search results for ML models exhibiting highest mean QWK for segment-corelation feature extraction.

Classifier	SVM	DT	kNN	OC (log. reg)	OC (LDA)	OC (SVM)	OC (DT)	OC (kNN)	Ridge reg. with interval mapping
Frequency interval size [Hz]	25	100	25	25	25	25	100	25	25
Min. correlation 1	0.05	0.15	0.2	0.2	0.2	0.05	0.2	0.2	0.2
Min. correlation 2	0.2	0.25	0.25	0.25	0.25	0.15	0.25	0.25	0.25
Min. percentage of features meeting condition: $d_f < d_{f,\max}$ [%]	80	87	83	83	83	87	80	83	83
Regularization type	L2	/	/	L2	/	L2	/	/	L2
C	1.0	/	/	0.5	/	1.0	/	/	0.01
γ	0.075	/	/	/	/	0.012	/	/	/
No. neighbors	/	/	10	/	/	/	/	10	/
No. training samples	91	68	91	91	98	77	91	91	91
No. features	47	23	47	47	224	539	4	47	26

modified assembly. These models are likely overfitted to their respective datasets and are not generally applicable. Therefore, a model exhibiting balanced performance across both assemblies was required. This was achieved by selecting the model with the highest average QWK across both assemblies. The performance of the tested multi-class classification algorithms, selected based on this criterion, is shown in Fig. 17.

The preprocessing parameters, feature extraction configurations, model hyperparameters, and resulting training set sizes for the ML models with the highest mean QWK values are summarized in Table 5. For most models, only a few dozen features were retained (Table 5), representing a reduction compared to the model that achieved the highest QWK for the original assembly alone (Table 4). The increase in QWK for the modified assembly joints suggests that assembly-specific features were removed. Conversely, the number of retained features in the balanced models is generally higher — except in the case of the OC (DT) model — compared to the model that achieved the highest QWK for the modified assembly using the segment-correlation approach (Table 4). This indicates that more generalizable information was preserved.

The described findings highlight the need for the proposed approach to utilize a smaller number of reference samples (with known squeezing forces) from the assembly different from the one used to obtain the training set samples, to select a set of features that is generalizable across different assemblies. Nevertheless, the ML model is still trained on the data from single a assembly. Note that, including the alternative assembly samples into the parameter selection process, as proposed, is more efficient than including these samples directly into the ML model training. Including a small number of modified assembly samples into the ML model training without prior efficient feature selection would increase the complexity of the ML problem, as the ML model would be required to comprehend the inconsistencies in assembly-specific features. To train the ML model on such training data a much higher number of samples would be required.

In Fig. 17, a drawback of the proposed balanced ML model selection approach can be observed. For some models, the QWK values for the original assembly joints are so high that the overall best-performing model becomes poorly balanced. This issue is particularly evident for the OC (LDA), OC (SVM), and ridge regression classifiers, which achieve high QWK values for the original

assembly joints (approximately 0.4, 0.7, and 0.45, respectively; Fig. 17) but yield very low QWK values for the modified assembly joints (close to zero in all cases). However, since none of these models outperformed better-balanced alternatives in terms of mean QWK, this drawback was considered negligible. Based on the results shown in Fig. 17, the OC (log. reg.) model was selected as the best-performing algorithm for the task. It outperformed all other tested models in terms of mean QWK and QWK for the modified assembly joints. For the original assembly joints, only the poorly balanced OC (SVM) model achieved a higher QWK.

Fig. 18 presents the testing and validation results for the OC (log. reg.) model in the form of confusion matrices. The number in each cell of the displayed heatmaps represents the number of samples corresponding to the reference class represented by the heatmap row for which the class, represented by the corresponding heatmap column, was predicted. The results of cross-validation testing are shown in Figs. 18a and 18b. For the original assembly joints (Fig. 18a), the model was able to capture the increasing trend of the rivet-squeezing force. Prediction errors rarely deviated by more than a single class from the reference. In contrast, prediction accuracy for the modified assembly joints (Fig. 18b) was somewhat lower. Nevertheless, the model successfully recognized the trend of increasing rivet-squeezing force, though it exhibited a tendency toward overestimation.

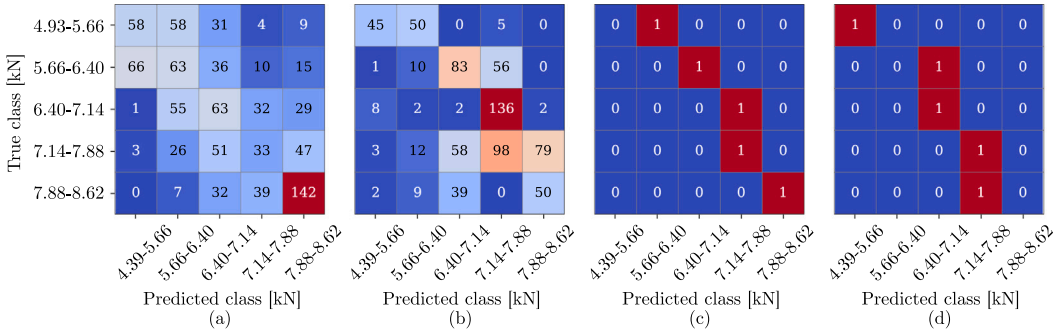


Fig. 18. Results for OC (logistic regression) model for: (a) testing within CV on the original assembly data, (b) testing within CV on the modified assembly data, (c) validation on the original assembly data, (d) validation on the modified assembly data.

The ML model was also validated using the separate validation dataset. The results are shown in Figs. 18c and 18d. For the original assembly validation samples, the predictions for the first three classes were overestimated by one class, while the samples from the fourth and fifth classes were correctly classified (Fig. 18c). +For the modified assembly validation samples, the predictions were correct for three out of five cases—specifically, the samples belonging to the first, third, and fourth classes. The sample from class 2 was misclassified as class 3 (one class too high), and the sample from class 5 was classified as class 4 (one class too low) (Fig. 18d).

The authors believe that the main reason underlying the inaccuracies of the ML model for both test and validation data is a combination of substantial variability among individual samples (i.e., training, testing, and validation samples) with a limited available training set. There are two main sources of variability. The first one is introduced intentionally, i.e., the manual preparation of the contact surfaces Section 4.1. The second reason stems from measurement errors, i.e., inaccuracies in sensor placement and impact locations where minor deviations can lead to substantial differences in response amplitudes. The ML model trained on a considerably larger training set may be able to handle this variability more successfully, however, obtaining a large enough training set is one of the main challenges of the proposed approach to the rivet-squeezing force estimation.

4.8. Comparison to the LAC-based rivet-squeezing force identification

With the ML-based squeezing force identification results established, they can now be compared to the existing methodology for estimating the rivet-squeezing force—specifically, the LAC-based approach [32]. In this method, a reference dataset must be constructed, comprising joints with known rivet-squeezing force values. Each force class is represented by a single joint impedance $\mathbf{Z}^{\mathrm{R}}$, resulting in as many reference impedances as there are rivet-squeezing force classes.

Five impedance–force pairs were collected and stored in the set $J_{\mathrm{ref}} = \left\{(\mathbf{Z}_i^{\mathrm{R}}, \mathrm{F}_{\mathrm{sq},i}) \text{ for } i = 1, \ldots, 5\right\}$. Each reference joint was selected such that its squeezing force value lies approximately at the midpoint of the corresponding class interval (e.g., a squeezing force of 5.3kN was selected to represent the 4.93–5.66kN range). The reference samples were selected from the dataset J used in the ML model training described in previous sections.

Following the selection of reference samples, the frequency regions influenced by the squeezing force were identified. The dynamic impedance matrices were segmented into frequency intervals of a specific size. Since the LAC criterion relies solely on the amplitude of the input FRFs, the segment selection was guided by the correlation between the mean amplitude of each frequency segment in each FRF of $\mathbf{Z}^{\mathrm{R}}$ and the corresponding squeezing force. Two parameters were defined for this process: the segment size and the minimum correlation threshold. For the FRF in the ith row and jth column of $\mathbf{Z}^{\mathrm{R}}$, the selected frequency segments were stored in the set $\Omega'_{i,j}$, where $\Omega'_{i,j} \subseteq \Omega$, and Ω denotes the complete set of discrete frequency lines.

The rivet-squeezing force for an unknown joint $\mathbf{Z}^{\mathrm{R}}\mathrm{u}$ was then estimated using a Nearest Neighbor Search. The estimation was based on the similarity between $\mathbf{Z}^{\mathrm{R}}\mathrm{u}$ and the reference impedances, evaluated within the selected frequency regions. The LAC

criterion was used as the similarity measure:

$$\mathbf{LAC}^{\mathrm{u},i} = \mathrm{LAC}\left(\mathbf{Z}_{\mathrm{u}}^{\mathrm{R}},\ \mathbf{Z}_{\mathrm{J},i}^{\mathrm{R}}\right) = \frac{2\left|\mathbf{Z}_{\mathrm{u}}^{\mathrm{R}*} \odot \mathbf{Z}_{\mathrm{J},i}^{\mathrm{R}}\right|}{\left(\mathbf{Z}_{\mathrm{u}}^{\mathrm{R}*} \odot \mathbf{Z}_{\mathrm{u}}^{\mathrm{R}}\right) \odot \left(\mathbf{Z}_{\mathrm{J},i}^{\mathrm{R}*} \odot \mathbf{Z}_{\mathrm{J},i}^{\mathrm{R}}\right)}; \quad i = 1, \ldots, 5, \tag{22}$$

where $\mathbf{LAC}^{\mathrm{u},i} \in \mathbb{R}^{n_{\mathrm{resp}} \times n_{\mathrm{exc}} \times |\Omega|}$ is the matrix of LAC values computed between the FRFs of the unknown joint impedance $\mathbf{Z}_{\mathrm{u}}^{\mathrm{R}}$ and $\mathbf{Z}_{\mathrm{J},i}^{\mathrm{R}}$. Finally, the joint of interest was assigned to class K with the highest mean LAC value between the $\mathbf{Z}_{\mathrm{u}}^{\mathrm{R}}$ and $\mathbf{Z}_{\mathrm{J},k}^{\mathrm{R}}$, calculated over the selected frequency segments of all FRFs of the joint impedance:

$$K = \underset{k \in 1,\ldots,5}{\arg\max} \left[\frac{1}{n_{\mathrm{resp}}} \frac{1}{n_{\mathrm{exc}}} \sum_{i=1}^{n_{\mathrm{resp}}} \sum_{j=1}^{n_{\mathrm{exc}}} \frac{1}{\left|\Omega'_{i,j}\right|} \sum_{\omega' \in \Omega'_{i,j}} \mathbf{LAC}_{i,j}^{\mathrm{u},k}(\omega') \right]. \tag{23}$$

To determine the optimal frequency segment size and minimum correlation threshold for the LAC-based rivet-squeezing force identification, a grid search was performed. The FRFs were segmented using intervals of 25, 50, 100, 150, and 200 Hz, while the minimum correlation threshold varied from 0.01 to 0.99 in steps of 0.01. Each configuration was evaluated using data from the original assembly J (excluding the five samples used for the reference set) and the modified assembly J_{v}. The performance of each configuration was assessed using the QWK metric. Fig. 19a presents the confusion matrices for parameter configurations that performed best for original assembly joints (25 Hz interval size, minimum correlation of 0.96, resulting in QWK of 0.22), and Fig. 19b shows results for the configuration that returned the highest QWK for modified assembly joints (150 Hz interval size, minimum correlation of 0.75, resulting in QWK of 0.22).

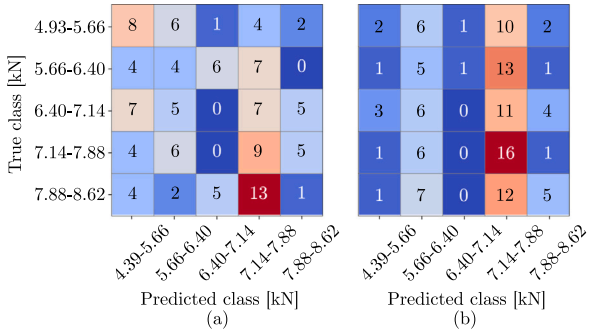


Fig. 19. Results for LAC-based rivet-squeezing force identification for parameter configuration that returns: (a) the highest QWK for the original assembly joints, and (b) the highest QWK for modified assembly joint and the mean QWK.

The results indicate that the LAC-based approach was able to detect the general trend of increasing rivet-squeezing force for the original assembly joints, as higher force values were typically predicted for samples with higher reference rivet-squeezing forces. However, even the best-performing configuration for the modified assembly resulted in poor predictive accuracy. Since the test samples were involved in the configuration selection process, a formal validation set would typically be required for an unbiased evaluation. Nonetheless, given that even the test results under optimal conditions do not match the performance of the ML-based approach, it is reasonable to assume that validation performance would not exceed these outcomes. A comparison between the cross-validation results for the balanced OC (logistic regression) model (Figs. 18a and 18b) and the best-case results of the LAC-based approach (Figs. 19a and 19b) clearly demonstrates that the ML-based method outperforms the LAC-based approach for joints from both assemblies.

5. Conclusions

This paper proposed an ML-based approach for estimating the rivet-squeezing force based on the dynamic response of the riveted joint. The method extends the previously established LAC-based identification approach [32], offering improved robustness and generalization capabilities. The training dataset consists of dynamic models of riveted joints obtained using the FBS-based joint identification method, which decouples the joint's dynamic behavior from that of the surrounding assembly. This decoupling is critical to the general applicability of the proposed method, as it allows the resulting joint models to be independent of the specific assembly configuration. Consequently, an ML model trained on joints from a single assembly can be applied to predict the rivet-squeezing force across joints from different assemblies, provided that the material, frictional, and geometric properties near the joint remain comparable. Nevertheless, due to the potential residual parent dynamics in the joints resulting from the decoupling process, several samples from assemblies different from the one used to obtain the dataset are required to obtain a set of features that generalizes across different assemblies.

Compared to the simpler LAC criterion, the application of ML algorithms enhances the robustness of rivet-squeezing force estimation under typical process variability. The ML-based approach accounts for minor variations in riveting parameters (i.e. riveting tool alignment, riveting speed etc.), component properties and their alignment, as well as measurement noise. Moreover, by incorporating multiple training samples per force class, it reduces sensitivity to individual measurement errors—something not

addressed in the LAC-based method. However, the main challenge remains the need for a sufficiently large dataset to effectively train ML models.

The proposed method was validated through a laboratory case study in which various feature extraction techniques and ML algorithms were evaluated. The optimal configuration — including preprocessing steps, feature selection, and model hyperparameters — was identified via grid search. Among the feature extraction strategies tested, PCA and LDA were found to be less effective than the segment-correlation approach, which leverages the correlation between individual features and the target rivet-squeezing force. Among the classification algorithms evaluated, ordinal classification using logistic regression (OC with log. reg.) yielded the most promising results.

Cross-validation testing of the final OC (log. reg.) model yielded a QWK of 0.64 for joints from the original assembly and 0.53 for those from the modified assembly. The model successfully captured the trend of increasing rivet-squeezing force but tended to overestimate the values for modified assembly joints. Validation with an independent dataset further supported the model's performance: for original assembly joints, the first three class samples were overestimated by one class, while the remaining two were correctly classified. For the modified assembly, three predictions were accurate, and two deviated by one class. The authors believe that the main reason for the limited accuracy of the ML model lies in the variability among samples unrelated to changes in rivet-squeezing force, combined with the limited amount of the training data.

Despite the encouraging results and clear advantages over the existing methodology, there is still room for further research. The primary challenge remains the generation of a sufficiently large and diverse training set. Realistically modeling the riveted joint's dynamic response in numerical simulations would be a major advancement in this regard. Additionally, the applicability of the proposed method to joints involving multiple rivets — a common scenario in industrial products — should be investigated.

CRediT authorship contribution statement

Tim Vrtač: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Miha Pogačar:** Writing – original draft, Validation. **Miha Kodrič:** Writing – review & editing. **Gregor Čepon:** Validation, Resources, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge financial support from the European Union's Horizon Europe research and innovation programme under the grant agreement No. 101138182 (CREDIT project) and the partial financial support from the Slovenian Research Agency (research core funding No. P2-0263 and research project L2-60145).

Data availability

Data will be made available on request.

References

[1] E. Frank, M. Hall, A simple approach to ordinal classification, in: Machine Learning: ECML 2001: 12th European Conference on Machine Learning Freiburg, Germany, September 5–7, 2001 Proceedings 12, Springer, 2001, pp. 145–156.

[2] H. Zhao, J. Xi, K. Zheng, Z. Shi, J. Lin, K. Nikbin, S. Duan, B. Wang, A review on solid riveting techniques in aircraft assembling, Manuf. Rev. 7 (2020) 40.

[3] Y. Liu, Y. Ma, M. Lou, E. Zhou, B. Zhang, Z. Fan, E. Ge, H. Zhao, Y. Li, Effects of rotation speed on microstructure and mechanical properties of non-prehole riveting process for high-strength aluminium alloys, Materials & Design 249 (2025) 113539, http://dx.doi.org/10.1016/j.matdes.2024.113539.

[4] L. Blaga, J. dos Santos, R. Bancila, S. Amancio-Filho, Friction riveting (FricRiveting) as a new joining technique in GFRP lightweight bridge construction, Constr. Build. Mater. 80 (2015) 167–179, http://dx.doi.org/10.1016/j.conbuildmat.2015.01.001.

[5] J. Wu, C. Chen, Y. Ouyang, D. Qin, H. Li, Recent development of the novel riveting processes, Int. J. Adv. Manuf. Technol. 117 (2021) 19–47, http://dx.doi.org/10.1007/s00170-021-07689-w.

[6] J. Yuan, Y. Ma, Y. Liu, W. Xu, N. Ma, Y. Li, Fatigue behavior and failure mechanism of friction self-piercing riveted aluminum alloy 2060-t8 joints, Int. J. Fatigue 197 (2025) 108940, http://dx.doi.org/10.1016/j.ijfatigue.2025.108940.

[7] J. Mucha, W. Witkowski, Mechanical behavior and failure of riveting joints in tensile and shear tests, Strength Mater. 47 (2015) 755–769, http://dx.doi.org/10.1007/s11223-015-9712-5.

[8] R. Porcaro, A. Hanssen, M. Langseth, A. Aalberg, An experimental investigation on the behaviour of self-piercing riveted connections in aluminium alloy AA6060, Int. J. Crashworthiness 11 (5) (2006) 397–417.

[9] Z. Yang, R. Jiang, Y. Zuo, Riveting damage behavior and mechanical performance assessments of CFRP/CFRP single-lap gasket-riveted joints, Eng. Fail. Anal. 149 (2023) 107253, http://dx.doi.org/10.1016/j.engfailanal.2023.107253.

[10] A. Komorek, J. Godzimirski, Modified pendulum hammer in impact tests of adhesive, riveted and hybrid lap joints, Int. J. Adhes. Adhes. 104 (2021) 102734, http://dx.doi.org/10.1016/j.ijadhadh.2020.102734.

[11] C. xian Wang, T. Suo, H. mai Gao, P. Xue, Determination of constitutive parameters for predicting dynamic behavior and failure of riveted joint: Testing, modeling and validation, Int. J. Impact Eng. 132 (2019) 103319, http://dx.doi.org/10.1016/j.ijimpeng.2019.103319.
[12] J. Maljaars, D. Leonetti, C. Maas, Fatigue life prediction of hot riveted double covered butt joints, Int. J. Fatigue 124 (2019) 99–112, http://dx.doi.org/10.1016/j.ijfatigue.2019.03.002.
[13] Z. Fan, K. Bai, C. Chen, Ultrasonic testing in the field of engineering joining, Int. J. Adv. Manuf. Technol. 132 (9) (2024) 4135–4160.
[14] M. Stamm, P. Krüger, H. Pfeiffer, B. Köhler, J. Reynaert, M. Wevers, In-plane heatwave thermography as digital inspection technique for fasteners in aircraft fuselage panels, Appl. Sci. 11 (1) (2021) http://dx.doi.org/10.3390/app11010132.
[15] Y. Le Diraison, P.-Y. Joubert, D. Placko, Characterization of subsurface defects in aeronautical riveted lap-joints using multi-frequency eddy current imaging, NDT E Int. 42 (2) (2009) 133–140, http://dx.doi.org/10.1016/j.ndteint.2008.10.005.
[16] Y. Deng, Z. Zeng, A. Tamburrino, L. Udpa, W. Shih, Automatic classification of magneto-optic image data for aircraft rivet inspection, Int. J. Appl. Electromagn. Mech. 25 (1–4) (2007) 375–382.
[17] R.P.G. Müller, An Experimental and Analytical Investigation on the Fatigue Behaviour of Fuselage Riveted Lap Joints (Ph.D. thesis), TU Delft, 1995.
[18] H. Huan, M. Liu, Effects of squeeze force on static behavior of riveted lap joints, Adv. Mech. Eng. 9 (5) (2017) http://dx.doi.org/10.1177/1687814016686891.
[19] M. Li, W. Tian, J. Hu, C. Wang, W. Liao, Effect of hole perpendicularity error and squeeze force on the mechanical behaviors of riveted joints, Microsc. Res. Tech. 85 (3) (2022) 1075–1088.
[20] J. Yuan, Y. Ma, H. Wan, W. Xu, P. Li, N. Ma, Y. Li, Fatigue life evaluation model of friction self-piercing riveted aluminum alloy sheets based on the structural stress method, Int. J. Fatigue 201 (2025) 109198, http://dx.doi.org/10.1016/j.ijfatigue.2025.109198.
[21] G. Lv, D. Jia, X. Zhang, C. Li, C. Zhao, H. Zhang, Effect of squeezing force on stress and fatigue of aircraft riveted joints, J. Aircr. (2025) 1–11.
[22] M. Skorupa, T. Machniewicz, A. Skorupa, A. Korbel, Effect of load transfer by friction on the fatigue behaviour of riveted lap joints, Int. J. Fatigue 90 (2016) 1–11, http://dx.doi.org/10.1016/j.ijfatigue.2016.04.005.
[23] C. Zeng, J.T. Xue, X.Y. Liu, W. Tian, Design variables influencing the fatigue of al 2024-T3 in riveted aircraft lap joints: Squeeze force and initial fit tolerance, Int. J. Fatigue 140 (2020) 105751, http://dx.doi.org/10.1016/j.ijfatigue.2020.105751.
[24] M. Skorupa, T. Machniewicz, A. Skorupa, A. Korbel, Fatigue life predictions for riveted lap joints, Int. J. Fatigue 94 (2017) 41–57, http://dx.doi.org/10.1016/j.ijfatigue.2016.09.007.
[25] J. de Rijck, J. Homan, J. Schijve, R. Benedictus, The driven rivet head dimensions as an indication of the fatigue performance of aircraft lap joints, Int. J. Fatigue 29 (12) (2007) 2208–2218, http://dx.doi.org/10.1016/j.ijfatigue.2006.12.010.
[26] S.H. Cheraghi, Effect of variations in the riveting process on the quality of riveted joints, Int. J. Adv. Manuf. Technol. 39 (11) (2008) 1144–1155.
[27] Z. Chang, Z. Wang, L. Xie, Y. Kang, M. Xu, Z. Wang, Prediction of riveting deformation for thin-walled structures using local-global finite element approach, Int. J. Adv. Manuf. Technol. 97 (5) (2018) 2529–2544.
[28] K. Zhang, H. Cheng, et al., Riveting process modeling and simulating for deformation analysis of aircraft's thin-walled sheet-metal parts, Chin. J. Aeronaut. 24 (3) (2011) 369–377.
[29] C. Zeng, W. Tian, W.-H. Liao, Improved model concerning driven rivet head dimensions based on material flow characteristics, J. Aircr. 53 (4) (2016) 1180–1185.
[30] J.A.N. Figueira, L.G. Trabasso, Riveting squeezing force estimation: a revised algebraic model, J. Aircr. 59 (4) (2022) 1005–1019.
[31] Y. Kang, S. Song, T. Wang, S. Kou, G. Li, Y. Chen, Analytical modeling of riveting squeezing force considering non-uniform deformation of rivets in aeronautical structures, Materials 17 (11) (2024) 2756.
[32] T. Vrtač, M. Kodrič, M. Pogačar, G. Čepon, Dynamic substructuring-based identification of the rivet-squeezing force, Mech. Syst. Signal Process. 229 (2025) 112487, http://dx.doi.org/10.1016/j.ymssp.2025.112487.
[33] M. Haeussler, S. Klaassen, D. Rixen, Experimental twelve degree of freedom rubber isolator models for use in substructuring assemblies, J. Sound Vib. 474 (2020) 115253, http://dx.doi.org/10.1016/j.jsv.2020.115253.
[34] M. Kreutz, F. Trainotti, V. Gimpl, D. Rixen, On the robust experimental multi-degree-of-freedom identification of bolted joints using frequency-based substructuring, Mech. Syst. Signal Process. (2023) http://dx.doi.org/10.1016/j.ymssp.2023.110626.
[35] D. De Klerk, D.J. Rixen, S.N. Voormeeren, General framework for dynamic substructuring: history, review and classification of techniques, AIAA J. 46 (5) (2008) 1169–1181.
[36] S. Voormeeren, D. Rixen, A family of substructure decoupling techniques based on a dual assembly approach, Mech. Syst. Signal Process. 27 (2012) 379–396, http://dx.doi.org/10.1016/j.ymssp.2011.07.028.
[37] M. van der Seijs, Experimental Dynamic Substructuring: Analysis and Design Strategies for Vehicle Development (Ph.D. thesis), Delft University of Technology, 2016, http://dx.doi.org/10.4233/uuid:28b31294-8d53-49eb-b108-284b63edf670.
[38] D.J. Rixen, How measurement inaccuracies induce spurious peaks in frequency based substructuring, in: Proceedings of the Twenty Sixth International Modal Analysis Conference, Orlando, FL. Society for Experimental Mechanics, Bethel, CT, 2008.
[39] W. D'Ambrogio, A. Fregolent, The role of interface DoFs in decoupling of substructures based on the dual domain decomposition, Mech. Syst. Signal Process. 24 (7) (2010) 2035–2048, http://dx.doi.org/10.1016/j.ymssp.2010.05.007.
[40] S.W. Klaassen, M.V. van der Seijs, D. de Klerk, System equivalent model mixing, Mech. Syst. Signal Process. 105 (2018) 90–112, http://dx.doi.org/10.1016/j.ymssp.2017.12.003, URL https://www.sciencedirect.com/science/article/pii/S0888327017306301.
[41] P.C. Hansen, Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank, SIAM J. Sci. Stat. Comput. 11 (3) (1990) 503–518, http://dx.doi.org/10.1137/0911028.
[42] M.V. Van Der Seijs, D.D. van den Bosch, D.J. Rixen, D. de Klerk, An improved methodology for the virtual point transformation of measured frequency response functions in dynamic substructuring, in: 4th ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, vol. 4, 2013, http://dx.doi.org/10.7712/120113.4816.C1539.
[43] E.A. Pasma, M.V.v.d. Seijs, S.W.B. Klaassen, M.W.v.d. Kooij, Frequency based substructuring with the virtual point transformation, flexible interface modes and a transmission simulator, in: A. Linderholt, M.S. Allen, R.L. Mayes, D. Rixen (Eds.), Dynamics of Coupled Structures, vol. 4, Springer International Publishing, 2018, pp. 205–213.
[44] I. Guyon, S. Gunn, M. Nikravesh, L.A. Zadeh, Feature extraction: foundations and applications, vol. 207, Springer, 2008.
[45] N.D. Cilia, G. De Gregorio, C. De Stefano, F. Fontanella, A. Marcelli, A. Parziale, Diagnosing Alzheimer's disease from on-line handwriting: A novel dataset and performance benchmarking, Eng. Appl. Artif. Intell. 111 (2022) 104822, http://dx.doi.org/10.1016/j.engappai.2022.104822.
[46] F. Fontanella, DARWIN, 2022, http://dx.doi.org/10.24432/C55D0K, UCI Machine Learning Repository.
[47] N. Yılmaz, B. Sekeroglu, Student performance classification using artificial intelligence techniques, in: International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions, Springer, 2019, pp. 596–603.
[48] N. Yilmaz, B. Şekeroğlu, Higher education students performance evaluation, 2019, http://dx.doi.org/10.24432/C51G82, UCI Machine Learning Repository.
[49] C.C. Tusell-rey, R. Tejeida-Padilla, O. Camacho-Nieto, Y. Villuendas-Rey, C. Yáñez-Márquez, Improvement of tourists satisfaction according to their non-verbal preferences using computational intelligence, Appl. Sci. (2021) URL https://api.semanticscholar.org/CorpusID:233683589.
[50] L. Gaudette, N. Japkowicz, Evaluation methods for ordinal classification, in: Y. Gao, N. Japkowicz (Eds.), Advances in Artificial Intelligence, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 207–210.
[51] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830.