

## Article

# Semi-Empirical Prediction of the Carrying Capacity of Rack Structure Elements: An ML-Driven Approach

Mirjana Piskulic <sup>1</sup>, Rodoljub Vujanac <sup>1,\*</sup>, Snezana Vulović <sup>2</sup>, Nenad Miloradovic <sup>1</sup>, Nenad Kostic <sup>1</sup>,  
Milos Matejic <sup>1</sup> and Ivan Miletic <sup>1</sup>

<sup>1</sup> Faculty of Engineering, University of Kragujevac, 34000 Kragujevac, Serbia; mirjana.piskulic@kg.ac.rs (M.P.); mnenad@kg.ac.rs (N.M.); nkostic@kg.ac.rs (N.K.); mmatejic@kg.ac.rs (M.M.); imiletic@kg.ac.rs (I.M.)

<sup>2</sup> Institute of Information Technologies, University of Kragujevac, 34000 Kragujevac, Serbia; vsneza@kg.ac.rs

\* Correspondence: vujanac@kg.ac.rs; Tel.: +381-34-335-990

## Abstract

Reliable prediction of carrying capacity is essential for efficient design of all carrying structures, including pallet racks. Precise or simplified analytical models are surely indispensable in engineering practice, but when considering racking structures, they still remain limited primarily by specific beam-to-column connections. This study proposes a semi-empirical, machine learning-driven approach for predicting the behavior of rack elements, i.e., upright frames, as a function of three structurally distinct connection types and appropriate racking element properties. Random Forest was selected for feature screening based on comparative analysis of six ML algorithms, and the retained parameters were used to develop interpretable power-law models. Within this research, upright cross-section area and beam-level height were retained as main influential parameters for the first two considered connection types, whereas upright cross-section moment of inertia, beam-level height and beam cross-section height governed the third type. These differences indicate that the contribution of each input parameter to capacity prediction varied among connection types. Grouped correction factors were introduced to account for deviations associated with upright profile and beam-level height, with additional differences specifically for the third connection type. The adopted models achieved near-exact agreement with the upright frame carrying capacities within the calibration domain. Although the resulting coefficients are specific to the analyzed rack structures, the proposed sequence provides a transferable framework for converting large structural datasets into simplified and interpretable engineering models.

**Keywords:** cold-formed steel structures; pallet racks; beam-to-column connections; carrying capacity; machine learning; regression; power-law model



Academic Editors: George Taranu,  
Viorel Ungureanu and Zsolt Nagy

Received: 28 August 2026

Revised: 14 September 2026

Accepted: 21 September 2026

Published: 24 September 2026

**Copyright:** © 2026 by the authors.

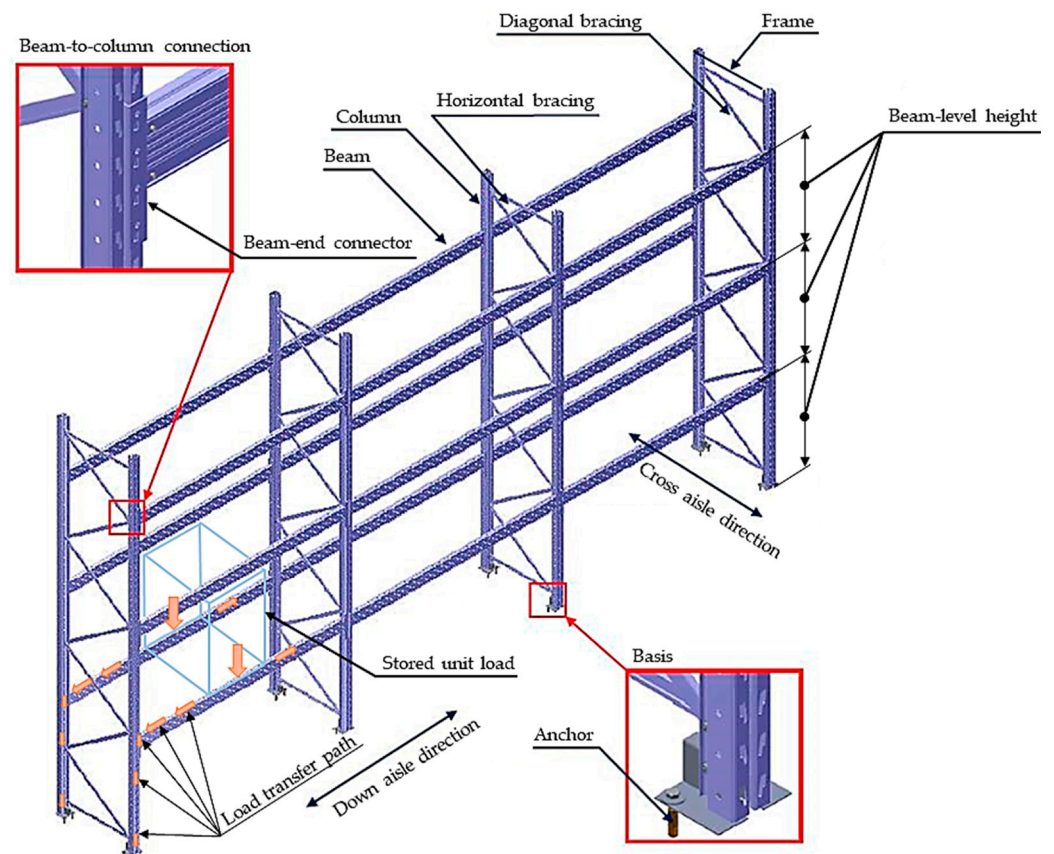
Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

## 1. Introduction

Generally, racking structures are widely used nowadays in warehouses because they use cold-formed thin-walled steel profiles that provide structural efficiency, low self-weight and easy assembly. All actions from stored unit loads are transferred through beams and beam-end connectors (BECs) to perforated uprights. Due to such a mode of load transfer, as a logical consequence, beam-to-column connections (BCCs) showing semi-rigid behavior influence the specific structural response of beams, uprights and the complete racking structure at all, with influential elements shown in Figure 1.



**Figure 1.** Configuration of rack structure elements. The arrows indicate the load transfer path. Adapted from Ref. [1] under the CC BY 4.0 license.

Research on pallet racks, primarily oriented on BCCs, relies on experimental testing to evaluate connection stiffness, moment capacity and failure modes. Due to the complex interaction between perforated uprights and BECs, which causes significant local deformations, analytical solutions were limited to simplified assumptions and could not accurately represent the actual structural response. Rapid development in computational techniques in recent decades has enabled the widespread application of numerical methods, such as finite element method (FEM) analysis, to the global behavior of pallet racks and their structural elements. Thus, prior studies [2–4] have experimentally and numerically analyzed the influence of upright slenderness, length and perforations on buckling behavior, and a data-driven decision-making model was developed for the optimization of thin-walled perforated sections [5]. More recently, data-driven modeling for local damage of cold-formed steel uprights was also analyzed [6]. Additionally, numerous previous studies have demonstrated that beam and upright dimensions affect the ultimate moment capacity and rotational stiffness of boltless BCCs [1,7–10]. Besides that, these connections provide a partial degree of rotation, not as in limiting rigid or pinned ones, but consequently, that influences the specific global structural response. Thus, FEM significantly reduced the need for extensive experimental testing and provided insight into stress distribution, local deformation and stiffness. However, these models still require experimental validation and remain computationally demanding for practical applications.

In parallel with the development of computational techniques for the design process, pallet racks have also undergone continuous structural evolution, aiming to increase carrying capacity while reducing labor and material costs. As a result, each further generation of pallet racks has a different impact on the structural behavior of uprights, beams, BECs and BCCs. Initial racking structures consisted of relatively strong but mostly standard (C

or U cross-section) profiles of perforated uprights, a beam cross-section mostly made of standard or common (I, U or rectangular tube cross-section) profiles simply welded to the simpler BEC. Subsequent developments, for example, introduced cold-formed, optimized upright and beam cross-sections and modified BECs with multiple rows of hooks, which consequently required additional perforations of uprights. Also, the connection between the beam and the BEC can be modified to a bolted one. All these modifications were made in order to simplify the production and reduce labor costs while still keeping flexible assembly. However, this development caused a reduction in the carrying capacity of rack elements through two independent effects: the bolted BEC exhibited lower rotational stiffness than the welded one, as confirmed by earlier experimental research [1], and the additional rows of perforations simultaneously reduced the effective cross-sectional area of the upright. Therefore, this configuration obviously represented a temporary solution in the structural development of a rack. In order to enhance carrying capacity without incurring further major changes and, therefore, higher production costs, the same type of upright and BEC was retained, but the improvement of carrying capacity was obtained by the new modification of the beam cross-section and its return to a welded joint with BEC. In this context, Yazici et al. [11] demonstrated that weld position significantly affects moment capacity, initial rotational stiffness, ductility and the governing failure mode. Consequently, it has been reaffirmed that the structural performance of recent pallet rack elements is determined not solely by individual elements but also by the interaction between all rack elements. Previous research by the authors experimentally and numerically compared different types of connections but at the local behavior level [1]. However, the impact of differences in these connections on the carrying capacity of overall rack structures has not been examined. Further, partially available commercial data on carrying capacity, therefore, can provide a common analytical basis for comparing considered BCC types.

Therefore, various attempts in analytical and semi-empirical models have also been proposed to reflect the local behavior of BCCs. Among the most widely applied are the Frye–Morris polynomial model [12], later extended in [13] through the Ramberg–Osgood function for three-dimensional structures. Subsequent developments evaluated the flexibility of cold-formed boltless BCCs and proposed a Frye–Morris polynomial and a three-parameter power model [14]. Later, Prabha et al. [15] modified the original Frye–Morris model by introducing an “air gap” parameter, thus achieving better agreement with experiments. Component-based and modified Ramberg–Osgood approaches have also been proposed to estimate connection stiffness and moment capacity [9,16]. More recently, Biemba et al. [17] used the Richard–Abbott model [18] to reproduce the  $M-\theta$  response of BCCs and developed regression equations. Across these developments, each model was calibrated and validated against a specific BCC type or experimental dataset, and their coefficients remain specific to the analyzed geometries and experimental conditions. Beyond racking structures, reduced mathematical representations have also been shown to efficiently capture complex structural responses while retaining a compact formulation [19], providing a methodological parallel to the present approach.

The increasing availability of experimental and numerical data, together with data provided in technical handbooks for pallet rack elements, determined according to standardized design and testing procedures specified in EN 15512 [20] and EN 1993-1-3 [21], has enabled the development and wider application of machine learning (ML) techniques in structural engineering [22]. Unlike traditional approaches, ML can efficiently identify complex nonlinear relationships and rank the relative importance of some parameters [23]. However, ML models are often treated as black-box predictions and therefore provide limited physical interpretation [7]. Recent applications have further demonstrated the use of several ML algorithms and interpretability methods for evaluating complex structural

responses, including the seismic performance of steel concrete composite columns and the confinement behavior of reinforced concrete columns [24,25]. ML techniques have been increasingly applied not only in structural research but also in racking structures to predict overall structural and connection behavior [26,27]. Ensemble learning algorithms, together with Artificial Neural Networks (ANNs), have demonstrated good predictions for non-linear connection behavior [28]. However, like the aforementioned analytical approaches, ML techniques have been primarily focused on predicting the  $M-\theta$  characteristics of BCCs. Their use for predicting the carrying capacity of rack structure elements and its representation through explicit analytical equations is lacking. But recent research by Mammadli et al. [29] extended ML applications from only the prediction of  $M-\theta$  characteristics to the prediction of axial upright buckling capacity. The trained ML model was retained as the final tool for predicting upright buckling capacity. Since the analysis focused on uprights, the effects of different BCCs and beam types on the carrying capacity were not considered. Furthermore, the ML findings were not transformed into explicit analytical equations for predicting axial upright buckling capacity.

Technical handbooks generally provide carrying capacities for a limited set of the most commonly used rack configurations in practice. Determining the carrying capacity of every geometrically feasible rack configuration, generated by a combination of different beam-level heights and beam profiles, would require substantial effort in experimental or numerical analyses.

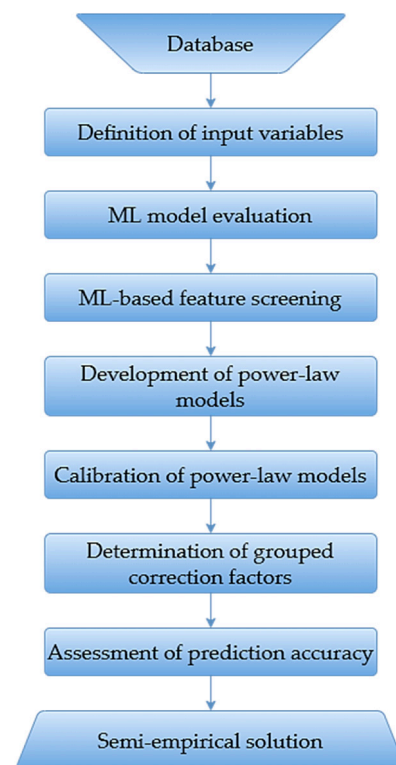
Against this background, this research analyzed whether relationships between the available carrying capacity data and the geometric variables of rack elements can be identified and expressed analytically or semi-empirically. Such relationships could enable the estimation of carrying capacity for additional configurations not explicitly provided in technical handbooks and, ultimately, extend the available design information to a broader range of rack configurations. A semi-empirical methodology combining ML-based feature screening, power-law regression and grouped correction factors was proposed. ML is used to quantify the contribution of structural variables to the prediction of carrying capacity rather than as the final predictive model. Retained variables are then introduced into power-law regression to obtain explicit and interpretable semi-empirical equations. The aim is to examine whether the contribution of structural variables to prediction of carrying capacity differs among connection types and whether it is necessary to develop a separate model for each BCC. This extends previous research, which has predominantly addressed  $M-\theta$  characteristics, rotational stiffness and moment capacity, where direct estimation of carrying capacity has not received attention. The broader contribution lies in a generalized methodological sequence that transforms large datasets into simplified engineering models and can be adapted to other connection types or engineering applications following calibration and validation specific to different purposes.

## 2. Overview of the Proposed Methodology

The overall carrying capacity of pallet rack elements depends on many different parameters, among the most important being characteristics of uprights, beams, BCCs, BECs and beam-level heights. Accordingly, after attempting a single common equation, which proved insufficient for representing all connection types, the proposed methodology combines ML-based feature screening and power-law regression to obtain simplified and physically interpretable predictive equations. As shown in Figure 2, the methodology consists of four stages:

- Construction of a database containing all relevant structural variables;
- Evaluation of ML models and feature screening using the selected ML model to identify variables that contribute the most to the prediction of carrying capacity;

- Development of separate power-law models for considered BCC types;
- Determination of correction factors from deviations associated with rack elements.



**Figure 2.** Workflow of the proposed methodology.

### 2.1. Dataset and Analyzed Connection Types

The dataset for this research was obtained from commercial racks designed according to EN 15512 [20] and EN 1993-1-3 [21] for BCC types denoted as T-1, T-2 and T-3. Data were verified through experimental testing and numerical analysis conducted in accordance with the applicable standardized procedures. Within each BCC type, the dataset was generated by combining the available upright profiles, beam profiles and beam-level heights. Each sample, therefore, represents a specific combination of these structural elements. Since several beam profiles may be compatible with the same upright profile and beam-level height, different configurations may have the same carrying capacity but differ in beam profile cross-sectional characteristics. Due to a broader range of available upright and beam profiles for each BCC type, the resulting number of configurations increased substantially, as summarized in Table 1.

**Table 1.** Number of rack configurations per BCC type.

| Connection Type | No. of Upright Profiles | No. of Beam Profiles | No. of Beam-Level Heights | No. of Samples |
|-----------------|-------------------------|----------------------|---------------------------|----------------|
| T-1             | 6                       | 6                    | 5                         | 180            |
| T-2             | 13                      | 23                   | 5                         | 1495           |
| T-3             | 13                      | 52                   | 5                         | 3380           |

Analyzed connection types represent stages in the improvement of pallet rack design. The earliest type, T-1, where the beam cross-section is shown in Figure 3a, consists of an upright with a single row of perforations, shown in Figure 3b, including a BEC with a corresponding single row of hooks, while the beam is welded to the BEC, also shown in Figure 3a. In the subsequent T-2 type, with a beam cross-section shown in Figure 3c,

the upright profile was modified, as shown in Figure 3e, and consists of a double row of perforations to take on a BEC with two rows of hooks, thus changing the mechanism of transferring the load between the beam and the upright. However, this modification reduces the effective cross-section area of the upright. Consequently, modifications in geometry affect not only the connection itself but also the relative contribution of structural parameters to the global response of a rack. The joint between the beam and the BEC in type T-2 relies on bolts, as shown in Figure 3c. The last T-3 type, shown in Figure 3d, retains the same upright profile and double-hooked BEC as in the previous type, with modifications to the beam profile and BEC dimensions, while reverting to a welded joint between beam and BEC. All these changes shown in Figure 3 illustrate efforts to balance structural efficiency, ease of assembly and carrying capacity with direct consequences for influential parameters and the overall structural response.

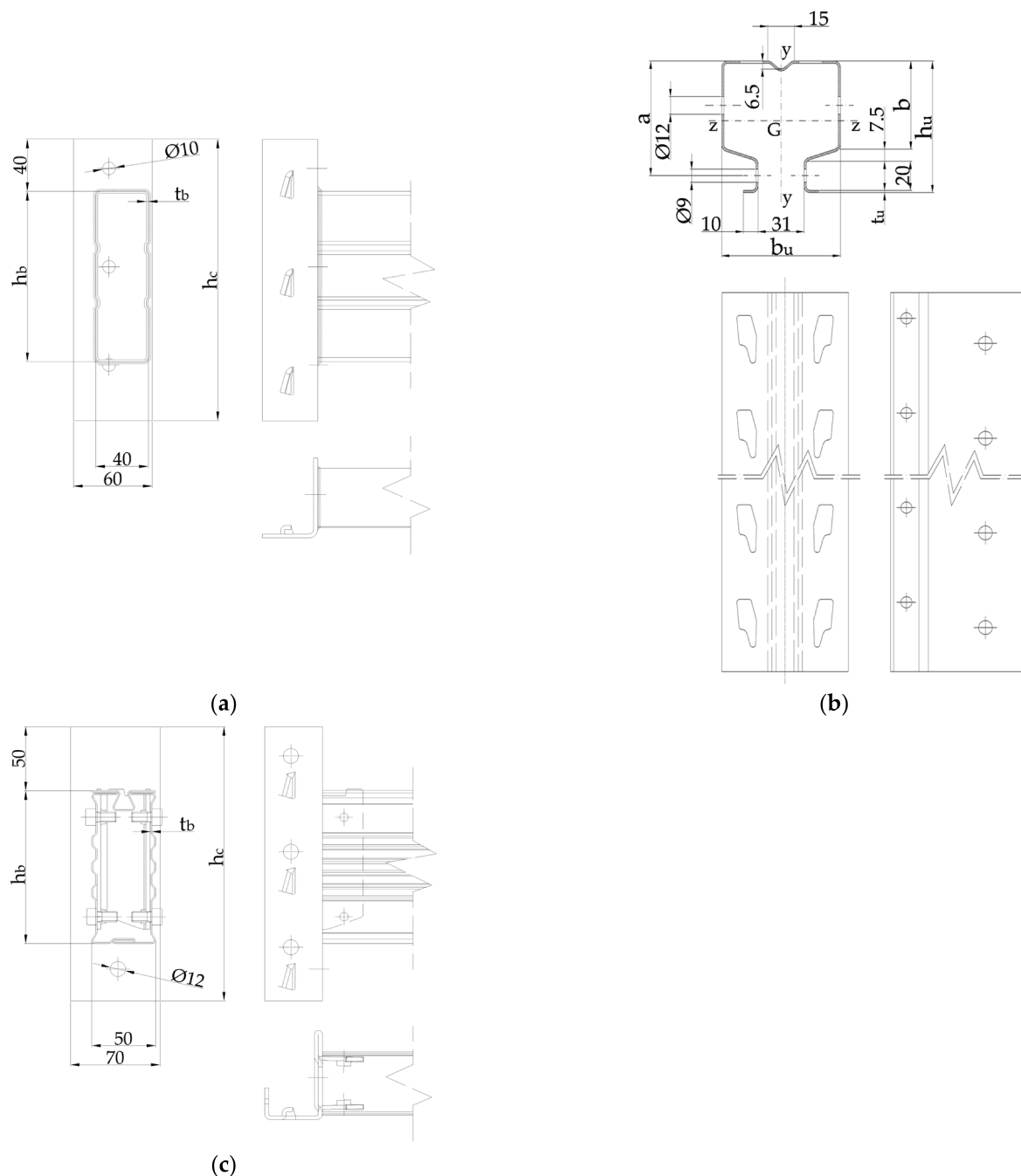
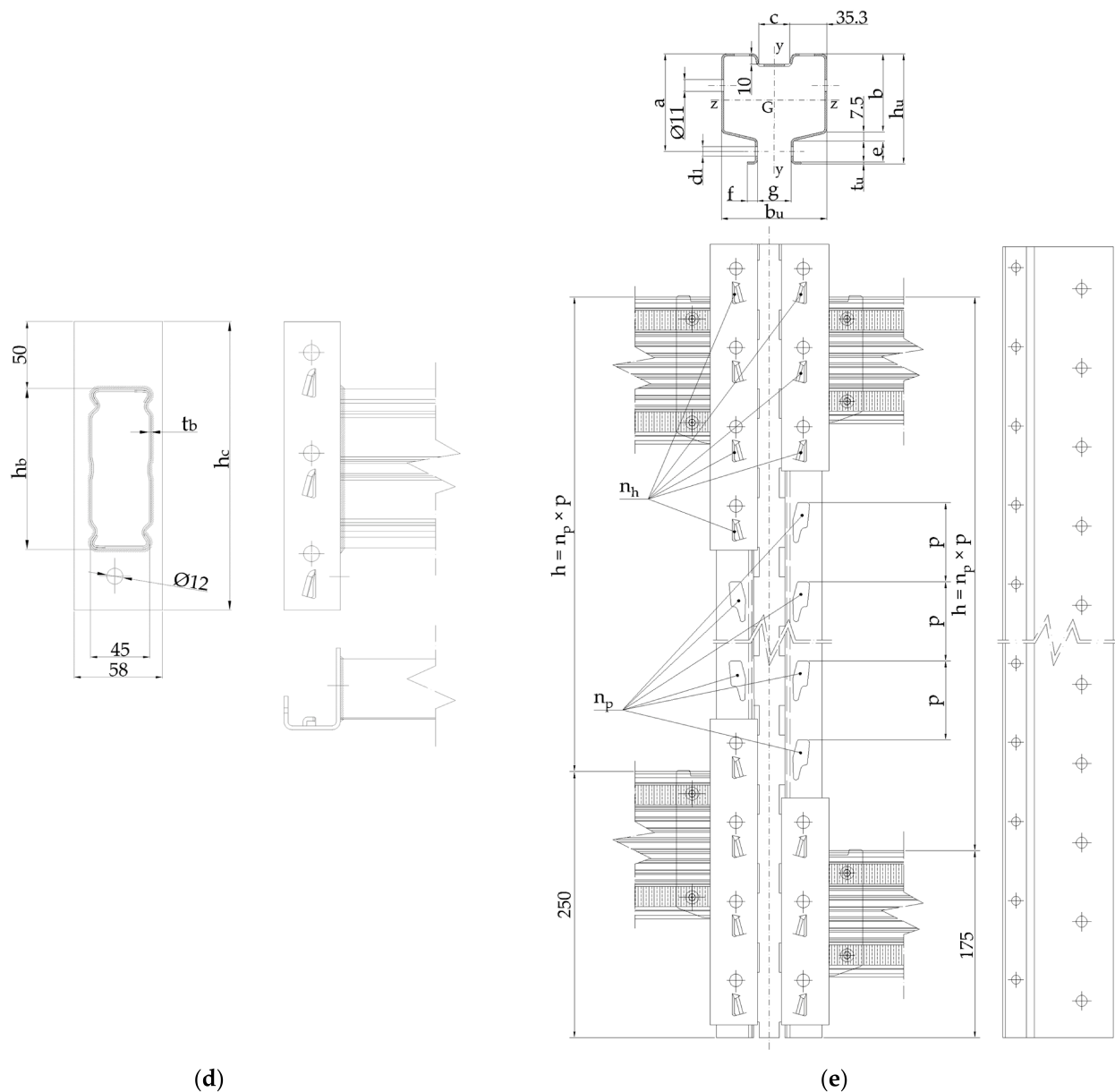


Figure 3. Cont.



**Figure 3.** Characteristic dimensions of BCCs: (a) beam-BEC joint of type T-1, (b) upright profile for type T-1, (c) beam-BEC joint of type T-2, (d) beam-BEC joint of type T-3 and (e) upright profile for types T-2 and T-3.

Input variables included geometric dimensions of the upright, beam and BEC, as shown in Figure 3, namely, the upright cross-section height ( $h_u$ ), width ( $b_u$ ) and wall thickness ( $t_u$ ), the beam cross-section height ( $h_b$ ) and wall thickness ( $t_b$ ) and the BEC height ( $h_c$ ). Corresponding cross-section properties are the cross-section area ( $A$ ), moments of inertia about principal axes ( $I$ ), and section moduli ( $W$ ) of the upright and the beam. Beam-level height ( $h$ ), number of hooks ( $n_h$ ) and BCC type complete the initial set of variables. Additionally,  $n_p$  denotes the number of perforations on uprights, while  $p$  is the perforation step. Consistent with the scope of the data provided in technical handbooks, the total upright height and the number of beam levels were not considered as input variables. Upright height is closely related to the number of beam levels and the available height of the warehouse building, which most often reflects a customer request arising in practice. Also, beam profiles can vary considerably within one rack configuration depending on the

specific storage requirements. As a result, the variety of possible combinations is very large and therefore not considered within this analysis.

The carrying capacity of upright frames (hereinafter: carrying capacity) from technical handbooks was defined as the target variable and therefore was not included among the input variables. Table 2 illustrates carrying capacities for upright frames for type T-1 as a function of available upright profiles (U-1 ÷ U-6) and the specific beam-level heights.

**Table 2.** Carrying capacities in kg, obtained from technical handbook for type T-1 of BCCs.

| Upright Profiles<br>for T-1 | Beam-Level Height, h [mm] |        |        |        |        |
|-----------------------------|---------------------------|--------|--------|--------|--------|
|                             | 1200                      | 1500   | 1800   | 2100   | 2400   |
| U-1                         | 8500                      | 7800   | 7100   | 5900   | 5400   |
| U-2                         | 11,200                    | 10,400 | 9400   | 8200   | 7400   |
| U-3                         | 14,000                    | 12,900 | 11,700 | 10,300 | 9300   |
| U-4                         | 15,600                    | 14,700 | 13,700 | 12,500 | 11,600 |
| U-5                         | 19,300                    | 18,200 | 16,900 | 15,500 | 14,300 |
| U-6                         | 26,000                    | 24,200 | 22,200 | 20,000 | 18,300 |

Note: Additional results for other BCC types are available from the authors upon reasonable request.

## 2.2. Analytical Background

The semi-rigid behavior of BCCs has primarily been determined through experimental testing, characterized through the  $M-\theta$  curve. Experiments have demonstrated that rotational stiffness and moment capacity are dependent on the upright, beam, BEC, number of hooks and interaction among these elements [10]. Despite their high accuracy, experimental testing remains expensive, time-consuming and limited to analyzed geometries and loading conditions. To reduce the dependence on extensive experiments, some analytical and semi-empirical solutions have been developed to describe the obtained  $M-\theta$  characteristics with a limited number of parameters. One of the earliest and most widely applied solutions is the Frye–Morris polynomial model [12], given by Equation (1) as follows:

$$\phi = C_1 \cdot (K \cdot M)^1 + C_2 \cdot (K \cdot M)^3 + C_3 \cdot (K \cdot M)^5 \quad (1)$$

where

- $C_1, C_2, C_3$ —constants for fitting the  $M-\theta$  curve;
- $M$ —moment of rotation;
- $K$ —coefficient that scales the ordinate of curves, defined as

$$K = \prod_{j=1}^m p_j^{a_j} \quad (2)$$

where

- $p_j$ —numerical value of  $j$  parameter;
- $a_j$ —exponent that shows the effect of the numerical value of  $j$  parameter on the  $M-\theta$  relation.

This formulation enables several parameters to be included in a single scaling factor. However, coefficients and exponents are calibrated for analyzed connections, which limits their application to configurations outside the experimental range.

Building on the same procedure, Ang and Morris [13] extended this approach to three-dimensional frames by replacing the polynomial model with the Ramberg–Osgood

function, which provides a more accurate representation of the nonlinear M- $\theta$  response. The standardized form of the function is given by

$$\frac{\phi}{\phi_0} = \frac{K \times M}{[K \cdot M]_0} \cdot \left[ 1 + \left( \frac{K \times M}{[K \cdot M]_0} \right)^{n-1} \right] \quad (3)$$

where

- $\phi_0$ —constant in standardized M- $\theta$  function;
- $[KM]_0$ —constant in standardized M- $\theta$  function;
- $n$ —parameter that controls the shape of the curve;
- $K$ —dimensionless factor that scales the ordinates of the M- $\theta$  curve, evaluated by the same procedure as in the Frye–Morris model:

$$K = \prod_{j=1}^m q_j^{a_j} \quad (4)$$

Although this formulation provides a more flexible representation of the nonlinear response, its scaling factor and curve parameters must also be calibrated for the considered connection types. Subsequent studies expanded these formulations by introducing additional geometric parameters. Prabha et al. [14] evaluated the flexibility of cold-formed rack connections using polynomial and power-law models. As a broader example of parameterization of connections, the modified Frye–Morris formulation was proposed in [15], retaining the same base Equation (1). Still, parameter  $K$  within the base equation included the “air gap” parameter between the BEC and the upright, described by the geometry of the gap, as follows:

$$K = d_a^{a_1} \cdot t_a^{a_2} \cdot g^{a_3} \cdot g_a^{a_4} \quad (5)$$

where

- $d_a$ —depth of the angle;
- $t_a$ —thickness of the angle;
- $g$ —gauge distance;
- $g_a$ —air gap distance;
- $a_1, a_2, a_3, a_4$ —corresponding regression exponents.

In addition to polynomial formulations, several analytical models have been proposed to reflect the nonlinear response of BCCs with a small number of influential geometric parameters. Another widely used analytical representation of nonlinear connection behavior is the Richard–Abbott model [18], expressed as

$$M = \frac{(R_{ki} - R_{kp}) \cdot \theta}{\left[ 1 + \left( \frac{(R_{ki} - R_{kp}) \cdot \theta}{M_0} \right)^n \right]^{\frac{1}{n}}} + R_{kp} \cdot \theta \quad (6)$$

where

- $M$ —the moment at the connection;
- $\theta$ —the relative rotation between beam and upright;
- $R_{ki}$ —the initial stiffness of the connection;
- $R_{kp}$ —the post-yield stiffness;
- $M_0$ —the reference moment, typically set equal to ultimate moment capacity  $M_u$ ;
- $n$ —shape parameter that controls transition between initial and post-yield response.

Biemba [17] successfully predicted M- $\theta$  characteristics using this model as a function of beam depth and number of hooks within parameters  $R_{ki}$ ,  $R_{kp}$  and  $M_0$ . Despite their different formulations, these models follow a common principle: complex structural behavior is represented by an explicit function whose coefficients depend on a reduced set of structural parameters. However, their main objective is to express local M- $\theta$  characteristics of BCCs rather than to directly estimate the carrying capacity of pallet rack elements.

This research follows a similar principle as previous analytical and semi-empirical formulations, namely, assuming that complex structural behavior can be represented by simplified equations. However, instead of describing the rotation of BCCs, this methodology focuses on predicting the carrying capacity of rack components, such as upright frames. Hence, rather than selecting parameters from experiments and calibrating a model for the M- $\theta$  response, this approach uses ML to identify influential variables and extends the analysis from local connection behavior toward the direct prediction of carrying capacity. Thus, combining the predictive capability of modern computational techniques with analytical solutions makes them suitable for engineering practice.

### 2.3. Machine Learning-Based Feature Screening

The initial dataset contained all geometric and cross-section variables describing the aforementioned rack elements. A preliminary power-law regression was initially applied to the complete dataset. However, a single formulation was not suitable for accurately representing different connection types, while several input variables overlapped, because they derived from the same profile. This preliminary model was therefore used only as an attempt and was not retained or reported as a predictive result. Unlike conventional models, which require a connection between input and output variables to be specified in advance, ML models learn predictive patterns directly from the available data. Therefore, they were used to identify variables with the highest prediction of carrying capacity. The retained variables were then used to develop simplified power-law equations for each connection type.

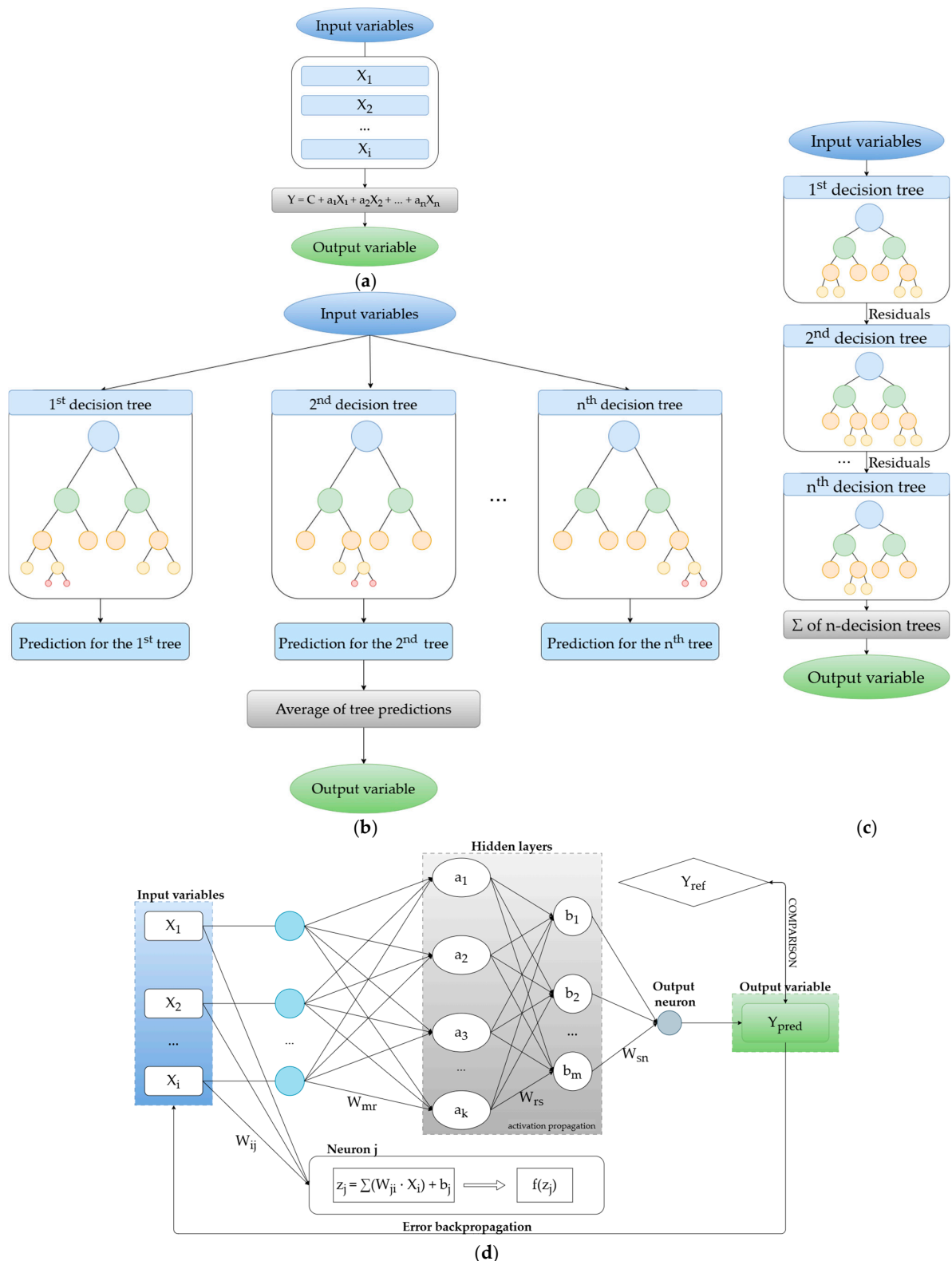
#### 2.3.1. Selection and Classification of ML Algorithms

In order to identify variables with the highest importance in the prediction of carrying capacity, ML algorithms representing different learning principles were evaluated:

- Linear regression (LR);
- Gradient boosting (GB);
- Extreme gradient boosting (XGBoost);
- Categorical boosting (CatBoost);
- Random forest (RF);
- Multilayer perceptron neural network (MLP NNs).

The selected algorithms cover three broad learning approaches: linear regression, tree-based ensemble learning and Artificial Neural Networks. LR estimates a single global relationship between the input and target variables and was therefore used as a linear basis. Ensemble algorithms considered in this study are tree-based but differ in how individual trees are constructed and combined. RF trains multiple trees independently using random data subsets and input variables and averages their predictions. In contrast, GB, XGBoost and CatBoost construct trees sequentially, with each tree attempting to reduce the prediction errors of the previous ensemble. GB represents the conventional boosting algorithm. XGBoost represents a regularized GB with complexity control. CatBoost is also a GB algorithm with ordered boosting and regularization. MLP NN represents a nonlinear approach that is not based on decision trees. This model learns nonlinear relationships by transforming input variables through multiple interconnected neurons. Because most of the

considered nonlinear ML models do not directly provide explicit engineering equations, the selected ML algorithm was used for feature screening rather than as the final model [7,23]. General learning principles of these algorithms are illustrated in Figure 4.



**Figure 4.** Schematic representation of (a) LR, (b) ensemble bagging, (c) ensemble boosting and (d) MLP NN. Blue and green shapes denote input and output elements, respectively. Gray shapes represent intermediate computational steps. Circles represent model nodes, and arrows indicate the direction of information flow.

### 2.3.2. Data Preprocessing and Model Training

All models were implemented in Python 3.12.1 software. The model configurations used in the analysis are summarized in Table 3. The hyperparameters were predefined and held constant throughout the analysis. Systematic hyperparameter optimization was not performed because the primary purpose of the ML analysis was comparative evaluation and feature screening rather than development of a final ML predictive model.

**Table 3.** Adopted hyperparameters of evaluated ML models.

| ML Model | Adopted Hyperparameters                                                                                                        |
|----------|--------------------------------------------------------------------------------------------------------------------------------|
| LR       | Default parameters                                                                                                             |
| GB       | n_estimators = 500; learning_rate = 0.03; max_depth = 4                                                                        |
| XGBoost  | n_estimators = 800; learning_rate = 0.03; max_depth = 4; subsample = 0.9; colsample_bytree = 0.9; objective = reg:squarederror |
| CatBoost | iterations = 800; learning_rate = 0.03; depth = 5; loss_function = RMSE                                                        |
| RF       | n_estimators = 500; max_depth = none; min_samples_leaf = 2                                                                     |
| MLP NN   | hidden_layer_sizes = (80, 40); activation = relu; max_iter = 5000                                                              |

The separate BCC datasets were combined into a single dataset that consists of 5055 configurations. BCC type was included as a categorical input variable and transformed using one-hot encoding. The same encoded representation was used for all algorithms, including CatBoost, to ensure consistent model comparison.

Numerical input variables were standardized for LR and MLP NN. This placed the LR input variables on a common numerical scale and supported stable and efficient training of the MLP NN. Standardization was not required for RF, GB, XGBoost or CatBoost because tree-based models are not sensitive to changes in variable scale. All preprocessing operations were fitted using the training set and then applied unchanged to the test set. The pooled dataset was randomly divided into an 80% training set and a 20% test set. A fixed random seed of 42 was used for the train–test split to ensure that all algorithms were evaluated using the same data partition. The same seed was also fixed within the relevant ML algorithms to ensure reproducible model training.

For model comparison, all algorithms were evaluated using 30 repeated group-aware 80/20 splits. Each group was assigned exclusively to either the training or test subset. Mean values and standard deviations of the performance metrics were calculated across the splits. Feature rankings were compared across repeated splits for each connection type.

### 2.3.3. Model Evaluation Criteria

The ML models were evaluated on the same test subset using four metrics:

- Coefficient of determination ( $R^2$ ) quantifies the overall agreement between the predicted and reference values, and it is given as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_{i,\text{ref}} - Y_{i,\text{base}})^2}{\sum_{i=1}^n (Y_{i,\text{ref}} - \bar{Y}_{\text{ref}})^2} \quad (7)$$

- Mean Absolute Error (MAE) represents the average absolute prediction error expressed in the same unit as the target variable and is calculated as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_{i,\text{ref}} - Y_{i,\text{base}}| \quad (8)$$

- Root Mean Square Error (RMSE), where the squared differences give greater weight to configurations with larger prediction errors, is determined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{i,\text{ref}} - Y_{i,\text{base}})^2} \quad (9)$$

- Mean Absolute Percentage Error (MAPE) expresses the average prediction error relative to the corresponding target variable, calculated as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_{i,\text{ref}} - Y_{i,\text{base}}}{Y_{i,\text{ref}}} \right| \cdot 100\% \quad (10)$$

where

- $n$ —number of samples in the evaluated data subset;
- $Y_{i,\text{ref}}$ —reference target value for  $i$ th sample;
- $Y_{i,\text{base}}$ —predicted target value for  $i$ th sample;
- $\overline{Y_{\text{ref}}}$ —the mean reference target value.

All reference target values were positive, so the percentage error was defined for each configuration. MAPE was used as the main measure for comparing model performance because it expresses the average prediction error as a percentage of the reference value, while MAE, RMSE and  $R^2$  provided additional measures of model performance. Lower MAE, RMSE and MAPE values indicate better agreement with reference data, whereas values of  $R^2$  closer to one indicate stronger overall agreement.

#### 2.3.4. Feature Screening and Importance Analysis

A suitable ML model was selected to identify the variables further evaluated in the power-law regression. Feature importance quantified each variable's relative contribution to the prediction, with all variables summing to 1 (or 100%). These numerical importance values were not used in the regression equations.

Since several geometric and cross-sectional variables derived from the same profiles contain overlapping information, Spearman correlation analysis was first used to reduce the initial feature set. The reduced set was not determined solely by statistical criteria. The structural meaning of the variables and their suitability for inclusion in an explicit semi-empirical equation were also considered. After calculating the feature importance using the selected ML model, SHapley Additive exPlanations (SHAP) and permutation importance were applied to test subsets as additional checks of the obtained feature rankings.

The selected ML model was then applied using the same reduced set of input variables. Each dataset was independently divided into an 80% training set and a 20% test set. The same random seed of 42 was used. Resulting values describe how much each variable influences the ML predictions. They should therefore not be interpreted as direct evidence that any variable causes a corresponding change in the target variable.

#### 2.4. Power-Law Regression Analysis

The reduced set of variables obtained by ML-based feature screening was used to develop explicit semi-empirical equations for predicting the target variable. Since the analyzed BCCs differ in cross-section characteristics, as well as in mechanisms of transferring the load, separate power-law models were created for each type. Because of the ability to

represent nonlinearities while retaining a simple form, a power-law model was selected. The general form of the power-law model is defined as follows:

$$Y_{\text{base}} = C \cdot \prod_{i=1}^m X_i^{a_i} \quad (11)$$

where

- $Y_{\text{base}}$ —predicted value of the target variable;
- $C$ —regression constant;
- $X_i$ — $i$ th input variable included in the model;
- $a_i$ —regression exponent associated with  $X_i$ ;
- $m$ —number of input variables included in the considered formulation.

The coefficients were estimated through logarithmic transformation. Since all input and target values were positive, taking the natural logarithm of Equation (11) converts the multiplicative model into a linear form:

$$\ln(Y_{\text{ref}}) = \ln(C) + \sum_{i=1}^m a_i \cdot \ln(X_i) \quad (12)$$

Multiple linear regression was subsequently performed separately for each BCC type in a spreadsheet-based program, allowing direct inspection of regression coefficients and their statistical significance. For each BCC type, the coefficients of the power-law models were estimated using only the 80% training subset. The equations and their coefficients were subsequently applied unchanged to the corresponding 20% test subset. Transformed regression parameters are defined as follows:

- $\ln(Y_{\text{ref}})$ —logarithmically transformed target variable;
- $\ln(C)$ —regression intercept;
- $a_i$ —regression coefficients that correspond to the power-law exponents.

The regression coefficients correspond directly to the exponents  $a_i$ . Their signs indicate the direction of the relationship, while their magnitudes describe the sensitivity of the predicted target variable to changes in the corresponding input variable. For each BCC type, the complete power-law model was compared with simpler variants in which one variable was omitted. Model selection considered  $R^2$ , MAE, RMSE, MAPE and statistical significance of regression coefficients. A reduced formulation was retained when removing a variable did not result in a meaningful loss compared to the initial model.

## 2.5. Determination of Grouped Correction Factors

Following calibration of the power-law models, the ratio between the reference target value ( $Y_{i,\text{ref}}$ ) and the corresponding one predicted by the base model ( $Y_{i,\text{base}}$ ) was calculated for each configuration denoted as  $k_i$ , but only from the training subset, as follows:

$$k_i = \frac{Y_{i,\text{ref}}}{Y_{i,\text{base}}} \quad (13)$$

The grouped correction factor  $k_{0,g}$  was calculated only from the training subset as the mean of the individual correction factors within the corresponding group, as follows:

$$k_{0,g} = \frac{1}{n_g} \sum_{i \in G} k_i \quad (14)$$

where  $G$  is the group of the same influential variables and  $n_g$  the number of configurations belonging to group  $G$ . The corrected target variable was calculated as

$$Y_{i,\text{corr}} = k_{0,g} \cdot Y_{i,\text{base}} \quad (15)$$

The obtained  $k_{0,g}$  values calibrated using the training subset were applied unchanged to the corresponding configurations in the test subset. The resulting corrected carrying capacities were then compared with the corresponding reference values for validation. Additionally, these correction factors are specific to the analyzed profiles and application of the methodology to other similar profiles would require recalibration using the corresponding data.

The uncertainty of the retained regression coefficients was evaluated using 95% confidence intervals. Residual diagnostics included the Breusch–Pagan test for heteroscedasticity and Cook’s distance for identifying potentially influential observations. When heteroscedasticity was detected, HC3 robust confidence intervals were reported.

### 3. Results

This section presents the results obtained through stages defined in the proposed methodology. A preliminary power-law model of the complete dataset was unable to represent deviation patterns related to the frame capacity of all BCC types using a single set of coefficients. Differences among BCC types demonstrated the need for separate formulations. The predictive performance of ML algorithms and the obtained feature importance using the chosen model are first presented. The influential variables are then evaluated through power-law regression, followed by the determination of grouped correction factors and assessment of their effect on the prediction of carrying capacity.

#### 3.1. ML Model Evaluation

The predictive performance of the analyzed ML models on the fixed test subset is presented in Table 4.

**Table 4.** Predictive performance of the ML models evaluated on the fixed test subset.

| ML Model | R <sup>2</sup> [-] | MAE [kg] | RMSE [kg] | MAPE [%] |
|----------|--------------------|----------|-----------|----------|
| RF       | 0.999540           | 47.165   | 151.316   | 0.347    |
| CatBoost | 0.999586           | 110.397  | 143.575   | 0.907    |
| XGBoost  | 0.999294           | 139.401  | 187.413   | 1.081    |
| GB       | 0.998446           | 207.220  | 278.181   | 1.582    |
| MLP NN   | 0.993050           | 318.267  | 588.226   | 2.677    |
| LR       | 0.946449           | 1262.942 | 1632.845  | 10.924   |

LR exhibited notably larger prediction errors than the nonlinear models. On the original fixed split, RF achieved the lowest MAE and MAPE, whereas CatBoost achieved the highest R<sup>2</sup> and lowest RMSE. Across 30 repeated group-aware splits, CatBoost and MLP NN achieved the lowest mean MAPE values ( $2.46 \pm 0.42\%$  and  $2.94 \pm 0.54\%$ ), while RF achieved  $6.35 \pm 1.38\%$  with a mean R<sup>2</sup> of  $0.969 \pm 0.014$ . Since predictive accuracy was not the only criterion for feature screening, RF was retained for its adequate performance and direct feature importance assessment. Its rankings were further verified using SHAP and permutation importance.

#### 3.2. RF-Based Feature Importance

The resulting importance analysis of RF model features is given in Table 5.

**Table 5.** Performance of independently trained RF models used for feature importance analysis for each BCC type.

| Connection Type | R <sup>2</sup> [-] | MAE [kg] | RMSE [kg] | MAPE [%] | No. of Samples | Training Samples | Test Samples |
|-----------------|--------------------|----------|-----------|----------|----------------|------------------|--------------|
| T-1             | 0.999127           | 135.033  | 157.969   | 1.030    | 180            | 144              | 36           |
| T-2             | 0.982920           | 377.566  | 956.216   | 3.391    | 1495           | 1196             | 299          |
| T-3             | 0.993829           | 216.171  | 563.664   | 1.770    | 3380           | 2704             | 676          |

Among connection types, T-1 exhibited the lowest prediction error, with a MAPE of 1.030%, followed by T-3 with 1.770%. The highest errors were obtained for T-2, which produced the highest MAE and RMSE along with a MAPE of 3.391%.

The Spearman correlation analysis indicated that several geometric and cross-section variables contained overlapping information. This was confirmed by the strong correlations between upright cross-section moment of inertia  $I_{xu}$  and cross-section modulus  $W_{xu}$  ( $\rho = 1$ ) and between  $I_{xu}$  and upright cross-section area  $A_u$  ( $\rho = 0.991$ ). Considering these relationships, together with the structural meaning of variables and their ability to be expressed in semi-empirical equations, the initial set of input variables was reduced to the four most influential structural variables, as follows:

- Upright cross-section area,  $A_u$  [mm<sup>2</sup>];
- Upright cross-section moment of inertia,  $I_{xu}$  [mm<sup>4</sup>];
- Beam-level height,  $h$  [mm];
- Beam cross-section height,  $h_b$  [mm].

Relative feature importance was evaluated using RF, SHAP and permutation importance separately for each BCC type, as presented in Table 6.

**Table 6.** Relative feature importance obtained using RF, SHAP and permutation importance for each BCC type.

| Connection Type | Method      | $A_u$ [%] | $I_{xu}$ [%] | $h$ [%] | $h_b$ [%] |
|-----------------|-------------|-----------|--------------|---------|-----------|
| T-1             | RF          | 44.65     | 44.35        | 10.97   | 0.03      |
|                 | SHAP        | 35.82     | 36.28        | 27.73   | 0.17      |
|                 | Permutation | 38.24     | 39.09        | 22.67   | 0         |
| T-2             | RF          | 51.22     | 40.55        | 8.07    | 0.16      |
|                 | SHAP        | 44.46     | 33.75        | 21.27   | 0.52      |
|                 | Permutation | 49.12     | 33.61        | 17.27   | 0         |
| T-3             | RF          | 11.99     | 73.10        | 11.16   | 3.75      |
|                 | SHAP        | 20.91     | 55.46        | 13.92   | 9.710     |
|                 | Permutation | 19.74     | 54.99        | 10.53   | 14.74     |

For T-1, upright cross-section area and moment of inertia exhibited nearly equal RF feature importance, accounting for 44.65% and 44.35%, respectively. Upright cross-section area was more dominant for T-2, with the feature importance at 51.22%, followed by upright moment of inertia at 40.55%. A different distribution was obtained for T-3, where upright moment of inertia accounted for 73.10% of the total feature importance. Beam-level height contributed between 8% and 11% across connection types. Beam cross-section height had negligible importance for T-1 and T-2, but reached nearly 4% for T-3 type. These variables were then evaluated in power-law regression analysis.

SHAP and permutation importance analyses confirmed main patterns obtained using RF importance, although they produced some differences. For T-1,  $A_u$  and  $I_{xu}$  remained nearly equally influential. For T-2,  $A_u$  remained the leading variable, whereas  $I_{xu}$  remained dominant for T-3. Both analyses also indicated a greater influence of beam-level height and

additionally, for T-3, beam cross-section height. Retained variables were further examined through power-law regression.

Across the 30 group-aware splits, mean  $R^2$  values of  $0.888 \pm 0.114$ ,  $0.916 \pm 0.055$  and  $0.964 \pm 0.014$  and mean MAPE values of  $9.48 \pm 2.73\%$ ,  $13.36 \pm 3.33\%$  and  $6.78 \pm 1.54\%$  were obtained for T-1, T-2 and T-3, respectively. Despite the stricter separation of the training and test subsets, the original feature rankings were preserved for all BCCs, supporting the robustness of the RF-based feature screening.

### 3.3. Development of Power-Law Regression Models

Four variables retained after RF screening were evaluated using the general power-law formulation defined by Equation (11). Upright cross-section area and moment of inertia describe properties of the same upright profile and may overlap. In order to separate the influences of upright cross-sectional properties, alternative formulations were compared:

$$\begin{cases} \text{(I)} \rightarrow Q_{\text{base}} = C \cdot I_{xu}^a \cdot A_u^b \cdot h^c \cdot h_b^d, \\ \text{(II)} \rightarrow Q_{\text{base}} = C \cdot A_u^a \cdot h^b \cdot h_b^c \text{ and} \\ \text{(III)} \rightarrow Q_{\text{base}} = C \cdot I_{xu}^a \cdot h^b \cdot h_b^c. \end{cases} \quad (16)$$

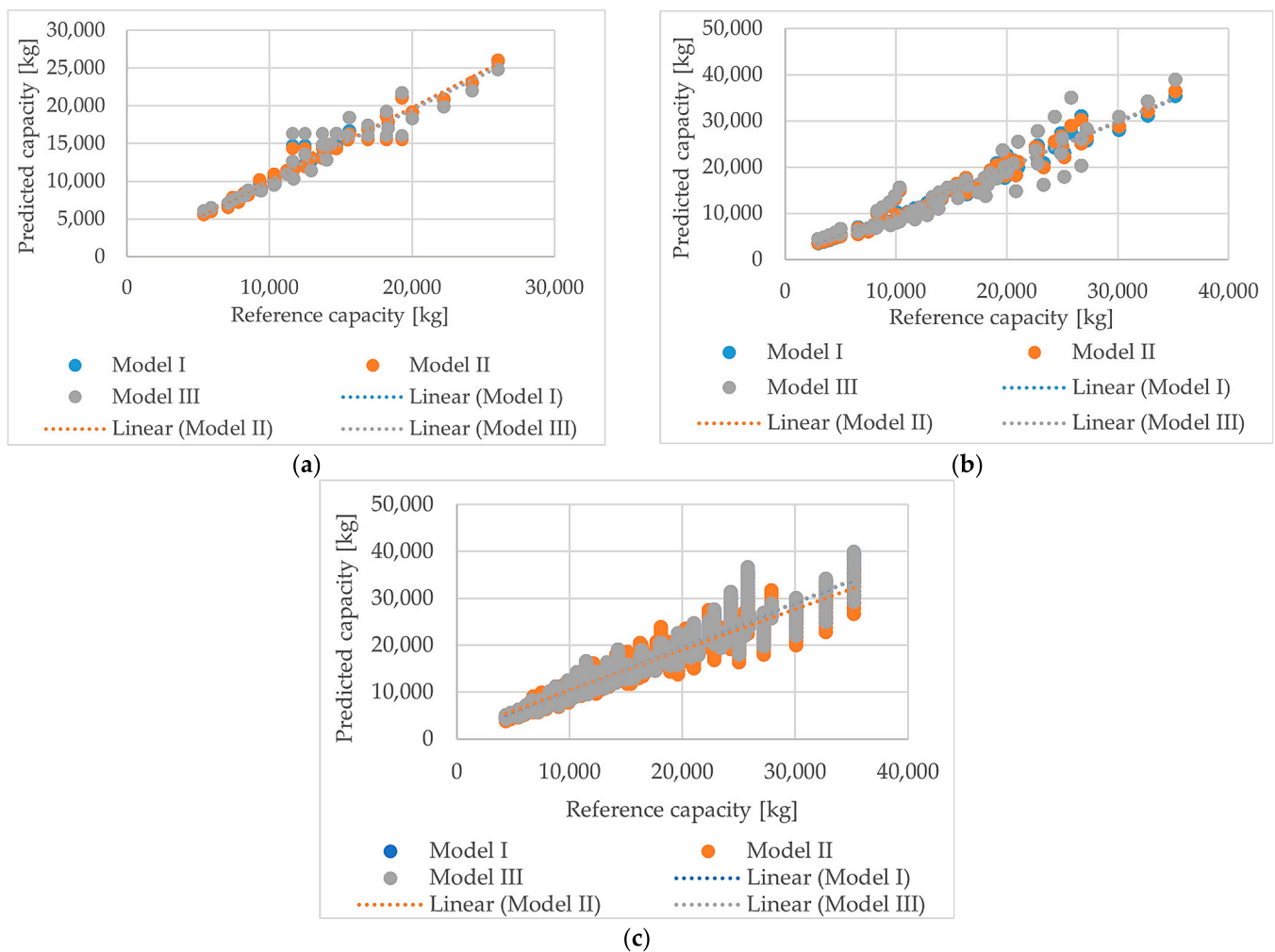
Model I included all influential parameters, while models II and III excluded the upright cross-section moment of inertia and area, respectively. Table 7 summarizes the agreement between reference carrying capacities and values predicted by regression solutions.

**Table 7.** Comparison of possible regression solutions with corresponding agreements.

| Connection Type | Regression Models | $R^2$ [-] | MAE [kg] | RMSE [kg] | MAPE [%] |
|-----------------|-------------------|-----------|----------|-----------|----------|
| T-1             | (I)               | 0.9859079 | 455.862  | 629.836   | 3.401    |
|                 | (II)              | 0.986310  | 485.812  | 620.774   | 3.816    |
|                 | (III)             | 0.9447709 | 1015.237 | 1246.876  | 7.195    |
| T-2             | (I)               | 0.9581651 | 1259.628 | 1567.923  | 9.332    |
|                 | (II)              | 0.9575847 | 1276.234 | 1578.909  | 9.71     |
|                 | (III)             | 0.8509216 | 2176.068 | 2959.804  | 16.069   |
| T-3             | (I)               | 0.9220791 | 1350.695 | 1965.859  | 8.498    |
|                 | (II)              | 0.8882692 | 1696.511 | 2354.031  | 10.638   |
|                 | (III)             | 0.9299999 | 1311.123 | 1863.266  | 8.483    |

For T-1 type, models I and II produced similar results. Model II had slightly higher MAE and MAPE but marginally higher  $R^2$  and lower RMSE, whereas model III showed significantly higher errors. A similar pattern was obtained for T-2 type, where excluding  $I_{xu}$  resulted in only a limited change in the error, while excluding  $A_u$  considerably reduced agreement with reference capacities, indicating that this variable should be retained for types T-1 and T-2.

For T-3 type, models I and III provided similar results, whereas excluding  $I_{xu}$  in model II significantly increased all errors and reduced  $R^2$ . This comparison supported the retention of  $I_{xu}$  and the omission of  $A_u$  for this connection type. The predictions obtained from the adopted and alternative formulations are compared in Figure 5.



**Figure 5.** Comparison of carrying capacities obtained by all power-law models for type: (a) T-1, (b) T-2 and (c) T-3. The dotted line represents perfect agreement between the predicted and reference carrying capacities.

After selecting the upright cross-section properties for each BCC type, the beam cross-section height was evaluated for inclusion in the final model. For T-1 and T-2, the estimated exponent of  $h_b$  had a  $p$ -value equal to 1, indicating a negligible contribution to the regression model. Therefore,  $h_b$  was omitted from their final equations. For T-3,  $h_b$  was statistically significant, with a  $p$ -value  $< 0.001$  and was retained. The retained equations defined for each BCC type with appropriate coefficients are presented in Table 8.

**Table 8.** Retained power-law regression models and corresponding coefficients with HC3 CI.

| Connection Type | Regression Equations                                | C [95% HC3 CI]             | a [95% HC3 CI]       | b [95% HC3 CI]          | c [95% HC3 CI]       |
|-----------------|-----------------------------------------------------|----------------------------|----------------------|-------------------------|----------------------|
| T-1             | $Q_{base} = C \cdot A_u^a \cdot h^b$                | 426.763 [334.340, 544.736] | 1.166 [1.144, 1.187] | −0.542 [−0.571, −0.513] | –                    |
| T-2             | $Q_{base} = C \cdot A_u^a \cdot h^b$                | 180.525 [147.350, 221.170] | 1.324 [1.310, 1.337] | −0.568 [−0.594, −0.544] | –                    |
| T-3             | $Q_{base} = C \cdot I_{xu}^a \cdot h^b \cdot h_b^c$ | 123.6 [107.972, 141.490]   | 0.604 [0.599, 0.610] | −0.620 [−0.635, −0.605] | 0.241 [0.231, 0.251] |

Notes: All geometric and cross-sectional variables are expressed in millimeters, as specified in the Nomenclature. The coefficient C is valid only for these units.

The coefficient estimates and corresponding HC3 robust 95% confidence intervals are presented in Table 8. The Breusch–Pagan test indicated no significant heteroscedasticity for T-1 with  $p = 0.104$ , whereas it was detected for T-2 and T-3 with  $p < 0.001$ . Although 3.33%, 3.08% and 4.85% of observations exceeded the  $4/n$  Cook's distance threshold, respectively, all values remained below 0.026, indicating that no individual observation exerted excessive influence on the fitted models.

Nevertheless, the retained equations were still not precise enough. Systematic deviations remained after applying these equations and were subsequently expressed through grouped correction factors.

### 3.4. Grouped Correction Factors

After the development of power-law models, remaining deviations defined by Equation (13) were examined. For T-1 and T-2, deviations were associated with upright profile and beam-level height, so these configurations were grouped, and denoted as  $k_{0,(u,h)}$ . Accordingly, the grouping applied to T-3 additionally included the beam group (c), and therefore was denoted as  $k_{0,(u,h,c)}$ . These deviations showed a notable difference between two ranges of beam cross-section heights. Beam profiles were therefore divided into two groups:

- Profiles with beam cross-section heights below or equal to 120 mm;
- Profiles with beam cross-section heights greater than 120 mm.

Regarding the observed dependencies, variables given in Section 2.4, as also by Equations (14) and (15), are presented in Table 9, and defined separately as follows:

- $G$ —group of the same influential variables, where

$$\left\{ \begin{array}{l} G_{u,h} \text{—group of T-1 and T-2, representing configurations with the same upright} \\ \text{profile (u) and beam-level height (h);} \\ G_{u,h,c} \text{—group of T-3, representing configurations with the same upright} \\ \text{profile (u), beam-level height (h) and beam group (c).} \end{array} \right.$$

- $n_g$ —number of configurations belonging to group  $G$ , where

$$\left\{ \begin{array}{l} n_{u,h} \text{—number of configurations for T-1 and T-2, belonging to group } G_{u,h}; \\ n_{u,h,c} \text{—number of configurations for T-3, belonging to group } G_{u,h,c}. \end{array} \right.$$

**Table 9.** Final semi-empirical equations per connection type.

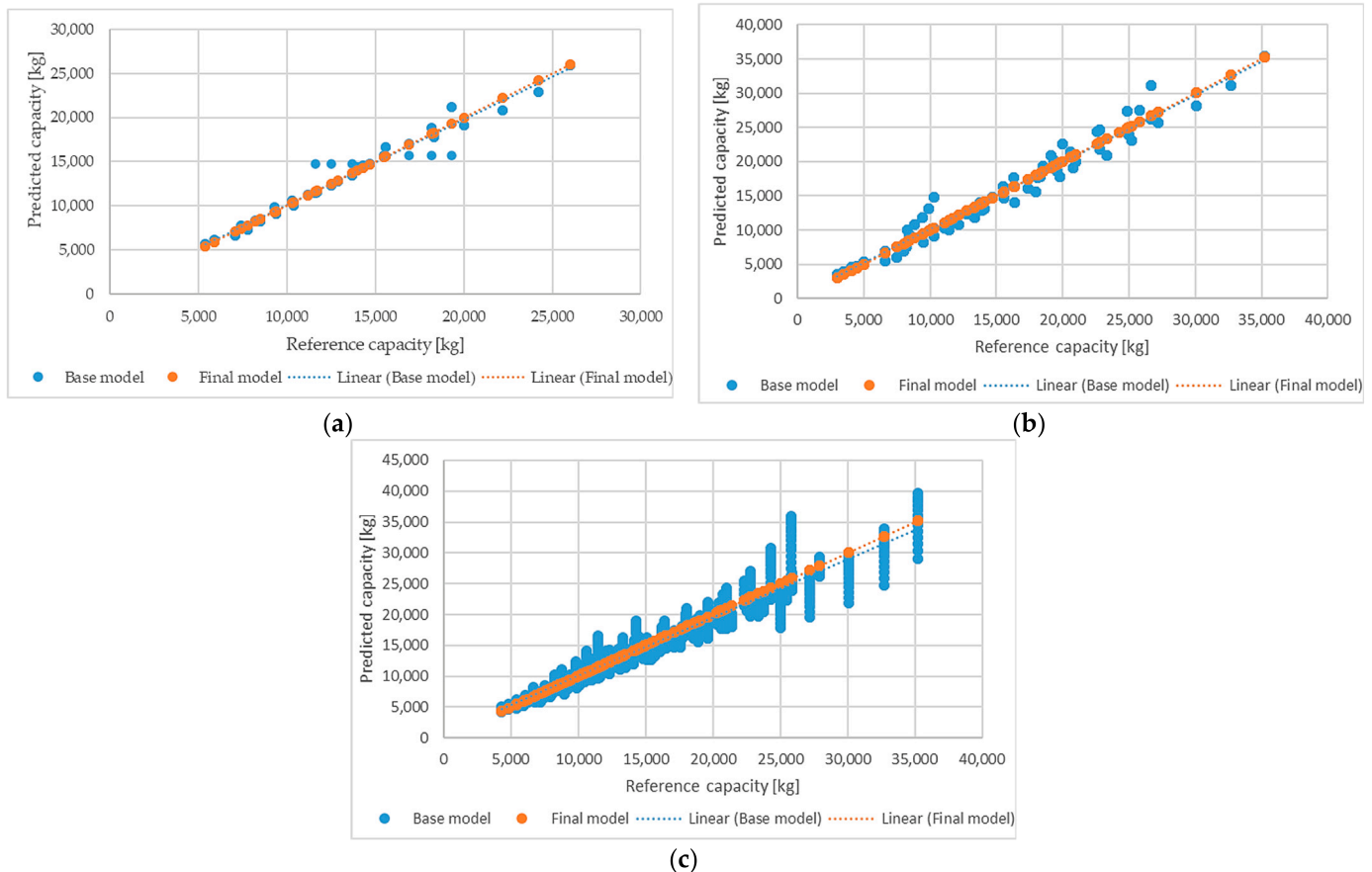
| Connection Type | Semi-Empirical Equations                                                                                                                                                              |
|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| T-1             | $Q_{i,corr} = k_{0,g} \cdot Q_{base} = k_{0,(u,h)} \cdot C \cdot A_u^a \cdot h^b = k_{0,(u,h)} \cdot 426.763 \cdot A_u^{1.166} \cdot h^{-0.542}$                                      |
| T-2             | $Q_{i,corr} = k_{0,g} \cdot Q_{base} = k_{0,(u,h)} \cdot C \cdot A_u^a \cdot h^b = k_{0,(u,h)} \cdot 180.525 \cdot A_u^{1.324} \cdot h^{-0.568}$                                      |
| T-3             | $Q_{i,corr} = k_{0,g} \cdot Q_{base} = k_{0,(u,h,c)} \cdot C \cdot I_{xu}^a \cdot h^b \cdot h_b^c = k_{0,(u,h,c)} \cdot 123.5 \cdot I_{xu}^{0.604} \cdot h^{-0.62} \cdot h_b^{0.241}$ |

Notes: The appropriate values of  $k_{0,(u,h)}$ ,  $k_{0,(u,h,c)}$ ,  $A_u$ ,  $I_{xu}$ ,  $h$  and  $h_b$  used in this study are available upon request.

Values of  $k_{0,g} > 1$  indicate the underprediction by the base model, whereas values of  $k_{0,g} < 1$  indicate the overprediction. This variation represents the upright/beam-level height-dependent deviations remaining after application of the corresponding power-law equation. The grouped correction factors were then applied to the corresponding predictions of the base model using Equation (15). Using Equations (7)–(10), the comparison metrics before and after applying group correction factors are given in Table 10 and Figure 6.

**Table 10.** Comparison of predicted carrying capacities before and after the application of the grouped correction factors.

| Connection Types |       | $R^2$       | MAE [kg]    | MAPE [%]    | RMSE [kg]   |
|------------------|-------|-------------|-------------|-------------|-------------|
| T-1              | base  | 0.986       | 485.812     | 3.816       | 620.774     |
|                  | final | $\approx 1$ | $\approx 0$ | $\approx 0$ | $\approx 0$ |
| T-2              | base  | 0.958       | 1276.234    | 9.710       | 1578.909    |
|                  | final | $\approx 1$ | $\approx 0$ | $\approx 0$ | $\approx 0$ |
| T-3              | base  | 0.930       | 1311.123    | 8.483       | 1863.266    |
|                  | final | $\approx 1$ | $\approx 0$ | $\approx 0$ | $\approx 0$ |

**Figure 6.** Comparison of predicted carrying capacities before and after the application of the grouped correction factors for types: (a) T-1, (b) T-2 and (c) T-3.

#### 4. Discussion

Previous experimental, numerical and analytical studies have provided a basis for understanding the behavior of pallet rack elements, particularly through the characterization and modeling of M- $\theta$  characteristics of BCCs, rotational stiffness and moment capacity [9,12–18]. Analytical and semi-empirical formulations have further demonstrated that complex connection behavior can be represented by explicit relationships with a limited number of structural parameters. This research follows the same general principle but shifts the focus from local connection behavior to the prediction of carrying capacity based on numerous characteristics of rack elements. Here, ML is used to identify structural variables that affect the prediction of carrying capacity the most. These variables are then used to develop explicit semi-empirical equations, while the ML model itself is not retained further as the final predictive tool. Differences identified by RF, as a chosen ML model, in feature importance were also reflected in the subsequent regression analysis and in the

variables retained in the final equations. This consistency supports a basis for developing formulations for BCC types separately and examining how the structural modifications affect the prediction of carrying capacity.

The general agreement between RF, SHAP and permutation analysis confirmed that the identified dominant variables were not dependent on a single importance method. Differences in their influence, particularly for the correlated  $A_u$  and  $I_{xu}$ , reflect their overlapping structural information. Nevertheless, all considered methods identified upright properties as the principal variables, with the additional relevance of  $h_b$  for T-3.

For the T-1 type, upright cross-section area and moment of inertia exhibited nearly equal RF feature importance, accounting for 44.65% and 44.35%, respectively. Nevertheless, the alternative power-law formulations showed that the cross-section moment of inertia could be omitted without a notable reduction in overall model performance. This result should not be interpreted as evidence that upright stiffness is mechanically unimportant. Instead, it indicates that the analyzed T-1 type dataset contained partly overlapping information because both properties derive from the same cross-sectional geometry.

For T-2, upright cross-section area had the highest RF feature importance of 51.22%, followed by moment of inertia at 40.55%. The regression analysis supported this ranking, as excluding  $I_{xu}$  caused only a limited change in model agreement, while removing  $A_u$  notably increased prediction errors. Compared with T-1, the T-2 type introduced an additional row of upright perforations, a bolted joint between beam and BEC and a modified beam profile. These modifications resulted in a shift from the nearly equal feature importance of  $A_u$  and  $I_{xu}$  in T-1 toward a greater contribution of  $A_u$  in T-2.

A different pattern was obtained for T-3 type. The upright cross-section moment of inertia accounted for 73% of the total RF feature importance and was retained in the final power-law equation. The power-law model without the upright cross-section area,  $A_u$ , produced results comparable to those containing both  $A_u$  and  $I_{xu}$ . Therefore, the final equation retained only  $I_{xu}$  as the characteristic of upright. Compared with type T-2, this type retained the same upright profiles but reintroduced a welded joint between beam and BEC and included a broader range of slightly modified beam profiles. The observed shift from area- to inertia-based prediction indicates a stronger alignment between carrying capacity and upright flexural behavior within the dataset of T-3 type. Nevertheless, direct confirmation that this change resulted from connection rigidity would require experimentally or numerically calibrated rotational stiffness data. It would also connect this semi-empirical finding with earlier experimental and numerical comparisons of types T-1 and T-2 [1].

Exponents obtained by regression analysis provide an additional quantitative interpretation of the final equations. When other parameters remain constant, a 1% increase in  $A_u$  corresponds approximately to a 1.166% increase in the predicted capacity for T-1 and a 1.324% increase for T-2 type. For T-3, a 1% increase in  $I_{xu}$  corresponds approximately to a 0.604% increase in predicted capacity.

When it comes to the height of the beam level, it is directly related to the free buckling length of the upright. In the analyzed configurations, increasing the height of the first beam level raises the lowest unbraced segment of the upright, thereby reducing its stability and carrying capacity [30]. Consistent with this structural behavior, the exponent of beam-level height was negative for each BCC type:  $-0.542$  for T-1,  $-0.568$  for T-2, and  $-0.62$  for T-3. Therefore, among considered connections, T-3 type exhibited the greatest sensitivity to beam-level height within the analyzed range. These exponents describe changes in predicted carrying capacity per regression equations from Table 8, when the other parameters remain unchanged. The relevance of beam-level height and upright cross-section variables is also consistent with the findings from [29], where upright length, thickness, cross-sectional area and width were identified as important variables for axial buckling

capacity of uprights. In this research, this perspective was extended to configurations that additionally include BCCs and beam characteristics.

Beam cross-section height  $h_b$  showed a different pattern. Its contribution was negligible for T-1 and T-2 and it was therefore omitted from their final equations. In contrast,  $h_b$  showed statistical significance for T-3 type, and was retained in the final equation with a positive exponent of 0.241. The wider range of beam profiles represented in T-3 type may have made this influence easier to identify statistically. Earlier experimental results have also shown that beam cross-section height affects the moment capacity and rotational stiffness of boltless rack connections [1,10], which once again confirms its influence and also the reliability of the obtained results.

Although the base power-law models showed close agreement with the overall carrying capacities, the remaining deviations were consistent with changes in upright profile and beam-level height. Accordingly, for T-1 and T-2 the grouped correction factor is  $k_{0,(u,h)}$ . For T-3, an additional difference remained between the two beam cross-section height groups, despite the inclusion of  $h_b$  in the base power-law equation. The correction factor was therefore extended to  $k_{0,(u,h,c)}$ , where (c) denotes beam cross-section profiles with  $h_b \leq 120$  mm and  $h_b > 120$  mm. Thus, in regression equations, there exists a general relationship between beam cross-section height and carrying capacity, while the grouped correction factor accounts for remaining differences between these two beam groups.

Regarding the results given in Table 10 and illustrated in Figure 6, the base power-law models achieved  $R^2$  values ranging from 0.929 to 0.986 and MAPE values from 3.816% to 9.71%. The grouped correction factors were calibrated exclusively using the 80% training subset and were subsequently applied unchanged to the corresponding configurations in the 20% test subset. After their application, predicted carrying capacities yielded an  $R^2$  of 1 and minimal MAE, RMSE and MAPE across all connection types. This near-exact agreement indicates that the deviations were highly consistent within the defined groups. However, since both subsets originate from the same technical handbook, application to new profiles or BCCs outside the analyzed dataset would require recalibration following the proposed procedure.

Beyond near-exact agreement with the available reference capacities, the developed procedure has an important practical application. Carrying capacities are usually reported only for specific beam-level heights ( $h$ ), as given for T-1 in Table 2. Since beam-level height is explicitly included in the developed power-law equations with a constant regression exponent, the base carrying capacity can be estimated for every geometrically feasible height corresponding to the upright perforation step of 75 mm within the analyzed range. Each introduced height is assigned the available carrying capacities applicable within the corresponding height interval, after which the grouped correction factor is recalibrated using the procedure defined in Section 2.5. This extends the available carrying capacities to configurations that are not explicitly defined in technical handbooks but can be found in practice. At present, the database expansion and recalibration can be performed manually by changing or adding the required beam-level height values in the existing database. Future development could automate this process.

Because of that, several limitations should be considered when interpreting these results. Differences observed between T-2 and T-3 can be more directly attributed to the modified joint between beam and BEC and beam profiles, since the upright and BEC geometry remained unchanged. In contrast, differences involving T-1 reflect the combined effect of all modified elements and cannot be attributed to a single structural change.

In addition, each analyzed configuration combines one upright profile and one beam profile at a time. Configurations that use different beam profiles at different levels on the same upright were not considered, although they are common in practice. For instance,

appropriate loads and their redistribution along the rack, considering both directions, then need to be included. Consequently, their direct application to rack configurations should be treated with caution. As a result, numerical modeling, such as FEM, would be suitable for predicting such configurations because it can accurately model the entire rack structure, including all connections and boundary conditions.

Within these limitations, the main transferable contribution is the methodological sequence: database construction, ML-based feature screening of a large structural dataset, development of interpretable power-law equations and grouped correction factors for residual effects. The obtained exponents and correction factors remain specific to analyzed structures, while the procedure can be adapted to other racks that share similar prediction problems and, more broadly, to engineering issues with sufficiently large datasets.

## 5. Conclusions

The main findings of this research, together with their practical implications and limitations, are summarized as follows:

- (1) An ML-guided semi-empirical methodology that transforms available commercial rack data into explicit equations for the prediction of carrying capacity was developed. The methodology combines ML-based feature screening, power-law regression and development of grouped correction factors. The main contribution is the conversion of a large structural dataset into simplified and interpretable semi-empirical equations, while keeping the identified importance of upright, beam, BCC and beam-level heights. Thus, the transferable outcome of this analysis is the methodological sequence itself, which can be applied to other racking systems using appropriate databases.
- (2) An initial attempt to form a single analytical equation capable of describing the carrying capacity of all BCC types proved insufficient. The equations were formed separately for each BCC type. The ML-based feature importance analysis supported this conclusion. Therefore, upright cross-section area and beam-level height were retained for T-1 and T-2, whereas upright cross-section moment of inertia, beam-level height and beam cross-section height were retained for T-3. Additionally, the regression exponents quantify this influence. Change toward inertia-based prediction for T-3 indicates a stronger association between carrying capacity and upright flexural behavior within the analyzed dataset.
- (3) Beam-level height showed a “negative influence” on carrying capacity for all equations. This is consistent with the reduction in carrying capacity associated with increasing the free buckling length of the upright. Furthermore, the direct inclusion of beam-level height gives the equations an important practical advantage. The base models can generate initial carrying-capacity estimates for every geometrically feasible beam-level height corresponding to the 75 mm upright perforation step within the analyzed range (between 1200 and 2400 mm). For these additional heights, the existing database can be manually expanded. The available carrying capacity is assigned according to the corresponding height interval, and the correction factor is recalibrated using the developed procedure. This extends the information about carrying capacity to additional geometrically feasible configurations that are not explicitly listed in technical handbooks.
- (4) Deviations that remained after the regression analysis were associated primarily with upright profile and beam-level height. The wider range of beam profiles additionally revealed systematic differences between two beam cross-section height groups, but only for the last T-3 type. Grouped correction factors accounted for all these deviations, improving the base model agreement ( $R^2 = 0.929\text{--}0.986$ ) to a near-exact agreement across all BCC types. The equations, coefficients and correction factors

remain specific to the analyzed rack structures and to variable ranges represented within the analyzed data.

- (5) The upright height and the number of beam levels, as well as the use of various beam profiles on the same upright, were not analyzed within this research. As discussed in Section 2.1, this reflects the growth of possible combinations. Also, it goes beyond the data included in technical handbooks. The equations developed in this research are intended to support rapid, approximate estimation of carrying capacity for configurations that combine a single upright and beam profile, consistent with the analyzed dataset. Broader rack configurations would require appropriately extended data and may be more suitably assessed using structural or parametric analysis supported by numerical modeling, for example, such as FEM, which is capable of analyzing the complete rack structure under relevant boundary conditions, remaining a more suitable approach.

Based on these findings and the current scope of the analysis, future research can proceed in the following directions:

- Future experimental and numerical research on T-3 should include rotational stiffness as a measured variable. This would help determine how the influence of the upright cross-section moment of inertia and beam geometry is related to local connection behavior. It would also allow the influence of BCC flexibility to be separated from that of upright and beam geometry. Finally, it would connect the present findings with the earlier experimental and numerical comparisons of T-1 and T-2 [1].
- The methodology should be applied to the databases from other technical handbooks, including a wider range of upright and beam profiles and connection types. This would enable its transferability to other racking systems to be assessed and support the development of additional simplified models.
- The expansion of the existing database with beam-level heights corresponding to the upright perforation step and the recalibration of  $k_0$  should be automated. It would extend their application to a larger number of geometrically feasible configurations.
- The dataset should be expanded to include the upright height and the number of beam levels. Further, different beam profiles should be combined within the same rack configuration. Given the growth of such configurations, this would likely require structural or parametric analysis, supported by numerical modeling, such as FEM, thus providing a more realistic representation of racking structures in practice.
- The methodology could be extended to the buckling analysis of upright frames in the down-aisle and cross-aisle directions, quantifying the influence of bracing, different connection types and their characteristics on rack stability.

Overall, the developed equations provide practical tools for estimating carrying capacity within the analyzed racking systems, while the methodological sequence offers a transferable basis for future application to other rack and structural configurations as well.

**Author Contributions:** Conceptualization, M.P. and R.V.; methodology, M.P. and R.V.; software, M.P., N.K., M.M. and S.V.; validation, M.P., R.V., N.K., M.M. and I.M.; formal analysis, M.P., R.V. and I.M.; investigation, M.P. and R.V.; resources, R.V.; data curation, M.P.; writing—original draft preparation, M.P.; writing—review and editing, M.P., R.V., N.K., M.M. and N.M.; visualization, M.P. and R.V.; supervision, R.V., S.V., N.K., M.M. and I.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Ministry of Science, Technological Development and Innovation under Grant Nos. 451-03-33/2026-03/200107 and 451-03-34/2026-03/200107.

**Data Availability Statement:** The underlying dataset is not publicly available at this stage because it forms part of the ongoing doctoral research of the first author M.P. Further inquiries can be directed to the corresponding author.

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

## Abbreviations

The following abbreviations are used in this manuscript:

|                             |                                                                    |
|-----------------------------|--------------------------------------------------------------------|
| BCC                         | Beam-to-Column Connection                                          |
| ML                          | Machine Learning                                                   |
| RF                          | Random Forest                                                      |
| BEC                         | Beam-End Connector                                                 |
| FEM                         | Finite Element Method                                              |
| M- $\theta$                 | Moment-Rotation                                                    |
| ANN                         | Artificial Neural Network                                          |
| LR                          | Linear Regression                                                  |
| GB                          | Gradient Boosting                                                  |
| XGBoost                     | Extreme Gradient Boosting                                          |
| CatBoost                    | Categorical Boosting                                               |
| MLP NN                      | Multilayer Perceptron Neural Network                               |
| MAE                         | Mean Absolute Error                                                |
| RMSE                        | Root Mean Square Error                                             |
| MAPE                        | Mean Absolute Percentage Error                                     |
| SHAP                        | SHapley Additive exPlanations                                      |
| Nomenclature                |                                                                    |
| T-1, T-2 and T-3            | analyzed beam-to-column connection types                           |
| $h_u$ [mm]                  | the upright cross-section height                                   |
| $b_u$ [mm]                  | the upright cross-section width                                    |
| $t_u$ [mm]                  | the upright wall thickness                                         |
| $h_b$ [mm]                  | the beam cross-section height                                      |
| $t_b$ [mm]                  | the beam wall thickness                                            |
| $h_c$ [mm]                  | the BEC height                                                     |
| $A_u$ [mm <sup>2</sup> ]    | the upright cross-sectional area                                   |
| $I_{xu}$ [mm <sup>4</sup> ] | the upright cross-section moment of inertia about principal axis x |
| $W_{xu}$ [mm <sup>3</sup> ] | the upright cross-section modulus about the principal axis x       |
| $h$ [mm]                    | beam-level height                                                  |
| $n_h$ [-]                   | number of hooks                                                    |
| $n_p$ [-]                   | the number of perforations on uprights                             |
| $p$ [mm]                    | the upright perforation step                                       |
| $R^2$ [-]                   | coefficient of determination                                       |
| $n$ [-]                     | number of samples in the evaluated data subset                     |
| $Y_{i,ref}$ [kg]            | reference target value for $i^{th}$ sample                         |
| $Y_{i,base}$ [kg]           | predicted target value for $i^{th}$ sample                         |
| $\overline{Y}_{ref}$ [kg]   | the mean reference target value                                    |
| $Y_{i,corr}$ [kg]           | the corrected target variable                                      |
| $C$                         | unit-dependent power-law regression constant                       |
| $X_i$                       | $i^{th}$ input variable included in the power-law model            |
| $a_i$ [-]                   | regression exponent associated with $X_i$                          |
| $m$ [-]                     | number of input variables included in the power-law model          |
| $k_i$ [-]                   | individual correction ratio for $i^{th}$ configuration             |
| $k_{0,g}$ [-]               | the grouped correction factor for group G                          |
| $G$ [-]                     | the group of the same influential variables                        |
| $n_g$ [-]                   | the number of training configurations belonging to group G         |

|                   |                                                                                                                                   |
|-------------------|-----------------------------------------------------------------------------------------------------------------------------------|
| $u$ [-]           | upright-profile identifier                                                                                                        |
| $c$ [-]           | beam cross-section height group identifier for T-3                                                                                |
| $G_{u,h}$ [-]     | group of T-1 and T-2, representing configurations with the same upright profile ( $u$ ) and beam-level height ( $h$ )             |
| $G_{u,h,c}$ [-]   | group of T-3, representing configurations with the same upright profile ( $u$ ), beam-level height ( $h$ ) and beam group ( $c$ ) |
| $n_{u,h}$ [-]     | number of training configurations for T-1 and T-2, belonging to group $G_{u,h}$                                                   |
| $n_{u,h,c}$ [-]   | number of training configurations for T-3, belonging to group $G_{u,h,c}$                                                         |
| $Q_{base}$ [kg]   | carrying capacity predicted by the retained base power-law model                                                                  |
| $Q_{i,corr}$ [kg] | carrying capacity predicted by the corrected semi-empirical model                                                                 |

## References

1. Vujanac, R.; Miloradović, N.; Vulović, S.; Pavlović, A. A Comprehensive Study into the Boltless Connections of Racking Systems. *Metals* **2020**, *10*, 276. [\[CrossRef\]](#)
2. Elias, G.C.; De Almeida Neiva, L.H.; Sarmanho, A.M.C.; Alves, V.N.; Barbosa E Castro, A.F. Ultimate Load of Steel Storage Systems Uprights. *Eng. Struct.* **2018**, *170*, 53–62. [\[CrossRef\]](#)
3. Neiva, L.H.D.A.; Sarmanho, A.M.C.; Faria, V.O.; Souza, F.T.D.; Starlino, J.A.B. Numerical and Experimental Analysis of Perforated Rack Members under Compression. *Thin-Walled Struct.* **2018**, *130*, 176–193. [\[CrossRef\]](#)
4. Taranu, G.; Iacob, S.; Taranu, N. Buckling Behavior of Perforated Cold-Formed Steel Uprights: Experimental Evaluation and Comparative Assessment Using FEM, EWM, and DSM. *Buildings* **2025**, *15*, 1561. [\[CrossRef\]](#)
5. Lyu, Z.-J.; Lu, Q.; Song, Y.; Xiang, Q.; Yang, G. Data-Driven Decision-Making in the Design Optimization of Thin-Walled Steel Perforated Sections: A Case Study. *Adv. Civ. Eng.* **2018**, *2018*, 6326049. [\[CrossRef\]](#)
6. Lyu, Z.-J.; Chen, Q.; Ji, M.; Sun, W.; Li, H. Physics and Data Driven Modeling for Local Damage Evolution of Cold-Formed Steel Rack Uprights: Cases Study. *Int. J. Steel Struct.* **2025**, *25*, 755–772. [\[CrossRef\]](#)
7. Shah, S.N.R.; Ramli Sulong, N.H.; El-Shafie, A. New Approach for Developing Soft Computational Prediction Models for Moment and Rotation of Boltless Steel Connections. *Thin-Walled Struct.* **2018**, *133*, 206–215. [\[CrossRef\]](#)
8. Dai, L.; Zhao, X.; Rasmussen, K.J.R. Cyclic Performance of Steel Storage Rack Beam-to-Upright Bolted Connections. *J. Constr. Steel Res.* **2018**, *148*, 28–48. [\[CrossRef\]](#)
9. Mourya, U.T.; Jayachandran, S.A. Moment-Rotation Model for a Range of Beam-to-Upright Connections in Cold-Formed Steel Storage Racks. *Structures* **2024**, *66*, 106842. [\[CrossRef\]](#)
10. Mohamad Rosli, N.A.; Tan, C.G.; Hashim, H.; Ramli Sulong, N.H. Experimental Investigation and Parameter Analysis of Boltless Steel Pallet Rack Beam-to-Column Connection. *J. Constr. Steel Res.* **2025**, *231*, 109578. [\[CrossRef\]](#)
11. Yazici, C.; Gusella, F.; Özkal, F.M. Effect of Beam-to-Connector Weld Position on the Flexural Response of Boltless Rack Connections. *J. Constr. Steel Res.* **2026**, *237*, 110141. [\[CrossRef\]](#)
12. Frye, M.J.; Morris, G.A. Analysis of Flexibly Connected Steel Frames. *Can. J. Civ. Eng.* **1975**, *2*, 280–291. [\[CrossRef\]](#)
13. Ang, K.M.; Morris, G.A. Analysis of Three-Dimensional Frames with Flexible Beam–Column Connections. *Can. J. Civ. Eng.* **1984**, *11*, 245–254. [\[CrossRef\]](#)
14. Prabha, P.; Marimuthu, V.; Saravanan, M.; Arul Jayachandran, S. Evaluation of Connection Flexibility in Cold Formed Steel Racks. *J. Constr. Steel Res.* **2010**, *66*, 863–872. [\[CrossRef\]](#)
15. Prabha, P.; Rekha, S.; Marimuthu, V.; Saravanan, M.; Palani, G.S.; Surendran, M. Modified Frye–Morris Polynomial Model for Double Web-Angle Connections. *Int. J. Adv. Struct. Eng.* **2015**, *7*, 295–306. [\[CrossRef\]](#)
16. Zhao, X.; Dai, L.; Wang, T.; Sivakumaran, K.S.; Chen, Y. A Theoretical Model for the Rotational Stiffness of Storage Rack Beam-to-Upright Connections. *J. Constr. Steel Res.* **2017**, *133*, 269–281. [\[CrossRef\]](#)
17. Biemba, B.; Yazici, C.; Shah, S.N.R.; Özkal, F.M. Predictive Analytical Framework for Industrial Steel Pallet Rack Connection Behavior. *Adv. Civ. Eng.* **2026**, *2026*, 9097355. [\[CrossRef\]](#)
18. Richard, R.M.; Abbott, B.J. Versatile Elastic-Plastic Stress-Strain Formula. *J. Eng. Mech. Div.* **1975**, *101*, 511–515. [\[CrossRef\]](#)
19. Jafarian Kafshgarkolaei, H.; Lotfollahi-Yaghin, M.A.; Mojtahedi, A. A Modified Orthonormal Polynomial Series Expansion Tailored to Thin Beams Undergoing Slamming Loads. *Ocean Eng.* **2019**, *182*, 38–47. [\[CrossRef\]](#)
20. EN 15512:2009; Steel Static Storage Systems—Adjustable Pallet Racking Systems—Principles for Structural Design. CEN: Brussels, Belgium, 2009.
21. EN 1993-1-3:2006; Eurocode 3: Design of Steel Structures, Part 1-3: General Rules—Supplementary Rules for Cold Formed Thin Gauge Members and Sheeting. CEN: Brussels, Belgium, 2006.
22. Shah, S.N.R.; Syed, F.H. An Optimized Machine Learning Based Moment-Rotation Analysis of Steel Pallet Rack Connections. *Struct. Eng. Mech.* **2021**, *79*, 499–506. [\[CrossRef\]](#)

23. Tapeh, A.T.G.; Naser, M.Z. Artificial Intelligence, Machine Learning, and Deep Learning in Structural Engineering: A Scientometrics Review of Trends and Best Practices. *Arch. Comput. Methods Eng.* **2023**, *30*, 115–159. [[CrossRef](#)]
24. Lai, B.-L.; Bao, R.-L.; Zheng, X.-F.; Vasdravellis, G.; Mensinger, M. Machine-Learning Assisted Analysis on the Seismic Performance of Steel Reinforced Concrete Composite Columns. *Structures* **2024**, *68*, 107065. [[CrossRef](#)]
25. Lai, B.; Zhou, X.; Zheng, X.; Li, S.; Venkateshwaran, A. Theoretical Reevaluation and Machine Learning Analysis on the Concrete Confinement Effect of Square Reinforced Concrete Columns. *Struct. Concr.* **2025**, *26*, 4215–4239. [[CrossRef](#)]
26. Salehi, H.; Burgueño, R. Emerging Artificial Intelligence Methods in Structural Engineering. *Eng. Struct.* **2018**, *171*, 170–189. [[CrossRef](#)]
27. Dimiduk, D.M.; Holm, E.A.; Niezgoda, S.R. Perspectives on the Impact of Machine Learning, Deep Learning, and Artificial Intelligence on Materials, Processes, and Structures Engineering. *Integr. Mater. Manuf. Innov.* **2018**, *7*, 157–172. [[CrossRef](#)]
28. Elhadary, M.; Bediwy, A.; Elshaer, A. Machine Learning-Based Prediction of the Structural Performance of a Long-Bolted Steel Connection. *AI Civ. Eng.* **2026**, *5*, 6. [[CrossRef](#)]
29. Mammadli, B.; Yazici, C.; Gürbüz, M.; Kocaman, İ.; Domínguez-Gutiérrez, F.J.; Mehmet Özkal, F. A Data-Driven Machine Learning Approach for Predicting Axial Load Capacity in Steel Storage Rack Columns. *Results Eng.* **2025**, *28*, 107475. [[CrossRef](#)]
30. Dai, L.; Chen, Y.; Cheng, L.; Ren, C. Buckling Behavior and Design of Cold-Formed Thin-Walled Steel Storage Rack Upright Frames. *Structures* **2025**, *75*, 108685. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.