BRNO UNIVERSITY OF TECHNOLOGY

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF CIVIL ENGINEERING

FAKULTA STAVEBNÍ

INSTITUTE OF METAL AND TIMBER STRUCTURES

ÚSTAV KOVOVÝCH A DŘEVĚNÝCH KONSTRUKCÍ

MACHINE LEARNING APPROACHES FOR THE RELIABLE DESIGN OF STEEL JOINTS

METODY STROJOVÉHO UČENÍ PRO SPOLEHLIVÝ NÁVRH STYČNÍKŮ OCELOVÝCH KONSTRUKCÍ

HABILITATION THESIS

HABILITAČNÍ PRÁCE

AUTHOR

AUTOR PRÁCE

Ing. Martin Vild, Ph.D.

BRNO 2025

Abstract:

This thesis presents a framework for developing machine-learned prediction models design of steel connections, with a focus on maintaining the reliability level required by Eurocode design procedures. The work responds to the increasing use of finite element analysis (FEA) in structural joint design, where traditional analytical or empirical design equations are often insufficient to describe nonlinear and interaction-dependent behavior. The study outlines the motivation to replace simplified analytical expressions with data-driven models trained on validated numerical simulations, while preserving reliability in line with the requirements of structural engineering design practice.

A column web in transverse compression is selected as a representative example to illustrate the proposed workflow. The procedure begins with automated generation of a large dataset through parametric numerical design calculations using IDEA StatiCa Connection, where more than 9 000 FE models were computed. The dataset covers a broad range of geometric and material parameters, and each simulation was evaluated according to a newly developed equivalent plastic strain limit criterion. The definition of this strain limit is one of the central contributions of the thesis, as it directly influences the calculated resistance and the consistency of the resulting machine learning models.

A detailed discussion is provided on the identification of a reliable equivalent plastic strain limit for steel structural components. The thesis compares recommendations given in different parts of Eurocode 3 and highlights the inconsistencies that currently exist. A new plastic strain limit proposal is derived based on numerous physical tests, numerical simulations and numerical design calculations for the net section failure of plates with notches and bolt holes, which is deemed to be the governing case. The proposed limit is clear and simple; it enables reproducible resistance determination in numerical design calculations and serves as a solid basis for subsequent data-driven modeling.

The thesis further introduces a complete process for the verification, validation, and reliability assessment of machine learning models in structural engineering applications. The process is based on the principles of EN 1990 and the draft standard FprEN 1993-1-14, ensuring that predictive models meet equivalent reliability levels to traditional design methods. Neural networks were selected as the primary modeling technique due to their ability to approximate complex nonlinear relationships. Using Python-based automation and systematic hyperparameter tuning, models were trained to predict resistances with a coefficient of variation below 2 % compared with FEA results. The study also investigates the influence of dataset coverage, feature scaling, and model regularization on prediction accuracy and reliability.

The results demonstrate that when trained on sufficiently rich and consistent data, machine learning models can replicate the results of detailed numerical simulations with no bias and negligible coefficient of variation. However, the reliability of such models remains limited to the range of input parameters represented in the training dataset. The thesis emphasizes that careful dataset generation, transparent preprocessing, and continuous verification are prerequisites for trustworthy machine learning applications in structural engineering. A discussion is provided on extending the range of validity through geometric similarity scaling.

The thesis concludes with recommendations for how machine learning can be integrated into structural design practice while keeping the same level of safety and transparency required by engineering standards. Several directions for future research are suggested: (i) improving the quality of finite element models used for training, including possible expansion through geometric similarity scaling or the use of simplified models with coarser meshes trained to produce high-quality results; (ii) enhancing reliability assessment by considering objectively safer FEA modifications, such as reducing the plastic strain limit or neglecting strain hardening; and (iii) creating well-documented datasets of experimental studies and numerical simulations, and encouraging researchers to share their results – a practice increasingly important in today's data-driven research environment.

The developed framework helps to connect advanced numerical simulations with practical design procedures and opens the path toward the next generation of efficient and reliable design tools.

Keywords:

finite element analysis; steel connections; numerical design calculation; plastic strain limit; Eurocode 3; reliability analysis; neural networks; machine learning; structural engineering; IDEA StatiCa

Bibliographic Citation of the Thesis:

Vild, Martin: *Machine Learning Approaches for the Reliable Design of Steel Joints*. Habilitation thesis. Brno University of Technology, Faculty of Civil Engineering, Institute of Metal and Timber Structures, Brno, Czech Republic, 2025.

Brno University of Technology
Faculty of Civil Engineering
Veveří 331/95, 602 00 Brno, Czech Republic

Acknowledgments

I express my deepest gratitude to my family – my wife Gabriela, my two sons Jaremiáš and Kilián, and daughter Judit – for allowing me to work long hours. I can see the deep truth in the viral post: "Twenty years from now, the only people who will remember that you worked late are your kids." Thank you to my children for teaching me the patience I never thought myself capable of, and to my wife for her unwavering patience with me despite my many flaws. I also thank my parents for raising me and supporting me in every step I have taken.

I would never have climbed the ladder in my professional life without my mentors, Lubomír Šabatka and Miroslav Bajer, who constantly challenged me with deadlines and thought-provoking questions. Luboš loves the saying "People only become great with great tasks."

I am grateful to my superiors, Juraj Šabatka and Milan Šmak, for allowing me flexibility in my everyday responsibilities so that I could focus on my research and thesis. My thanks also go to Petr Červinka, who replaced me as Product Owner at IDEA StatiCa, taking over the burdens of day-to-day management and freeing my hands to concentrate on research.

I thank František Wald for including me in his remarkable research group and for enabling me to see the broader landscape of European steel research at the highest level.

I also thank Andreas Müller and Andreas Taras for guiding me to this topic of machine learning; showing me how useful this tool can be. My gratitude extends to Michael Konečný and Jaromír Kabeláč for guiding me through the intricacies and pitfalls of machine learning.

My thanks go to our staff of the Institute of Metal and Timber Structures, particularly Alena Sklenářová for her daily assistance in navigating the administrative maze, and Jiří Veselý, Jiří Fischer, and Miloslav Ledina from our laboratory for conducting numerous experiments. I also acknowledge Ivan Balázs, the driving force behind the column web panel in compression experiments and related papers, and Ondřej Pešek, who developed the ANSYS numerical models. I am grateful to the members of the Laboratory of Numerical Structural Design – Mykola Lastovetskyi for his unparalleled enthusiasm, gym sessions, ice cream deliveries, and Python discussions, and Amina Hajdarević for her thorough literature review on web panel in transverse compression and speedy review of this thesis, during which she revealed how many missing English articles I had overlooked.

All of these people take selfishly without any trace of envy my success as their own.

The core of this thesis lies in the programming of neural networks. I have only very basic programming skills in Python, I am no programmer at all. The developing of codes relied almost entirely on large language models (LLMs), such as ChatGPT or Microsoft Copilot (both using OpenAI's GPT-4, GPT-4 Turbo, GPT-4o, GPT-4o mini or GPT-5). I have utilized LLMs also in writing itself. However, all the codes and texts were reviewed and modified, and I take full responsibility for them.

The experimental works at Brno University of Technology (BUT) would not be possible without the financial support of projects Inter Excellence, LUAUS23114, supported by the Ministry of Education, Youth and Sports of the Czech Republic, and internal BUT grant FAST-S-25-8879, Advanced methods for the design of load-bearing structural elements made of composite materials and joints of steel structures.

Contents

Chapter 1

Introduction

The aim of this thesis is to develop a framework for creating machine-learned predictions for components of steel connections. It does not attempt to provide these precise predictions or formulas, as their derivation will likely require several more years of research.

The thesis shows a workflow and highlights the critical areas that should be addressed on an example of a column web in a transverse compression component. The process includes the generation of data for machine learning, whose quality must be better than the current code formulas, the establishment of a neural network, and the reliability assessment of predictions. It is assumed that the data will be generated by finite element analyses.

The thesis focuses on two main areas, which are aligned with the projects and research papers by the author:

- Failure criterion of plastic strain limit for numerical design calculation of steel connections by finite element analysis.
- The process of data generation by numerical design calculation and machine learning, with recommendations for the best performance.

1.1 Problem statement

There are two methods in the use of current design codes, such as EN 1993-1-8 [1]:

- Analytical methods
- Curve-fitting method

Analytical methods are mathematical equations derived from physical phenomena based on predefined assumptions. In contrast, curve-fitting methods identify the variables that influence the results but do not explicitly define the underlying assumptions or dependencies. These methods are typically calibrated to a specific set of experiments or validated numerical simulations and are commonly used, e.g., in the design of hollow section joints. The curve-fitting methods are not popular in the scientific community because the mechanisms and theory behind them are not apparent. The curve-fitting methods are only as good as the input data, i.e., the experiments, in which regard the curve-fitting method is similar to machine learning. Although there was an attempt to generalize the design of hollow section joints using the component method [2], this effort was never completed.

There are two major drawbacks of these traditional methods:

1. The phenomenon is too complex and cannot be correctly described by simple formulas. The assumptions and dependencies of variables are no longer apparent from the formulas.
2. The assumptions, e.g., first-order analysis, may be insufficient to achieve a good fit with the real-world behavior.

Therefore, in many cases of the design of steel joints, the formulas are too complex, and regular engineers no longer recognize the mechanics and theory behind the design formulas. Often, there is no background to the code development, such as EN 1993-1-8, and scientific papers are not

available to everyone. Despite their complexity, code formulas often provide very large dispersion to the experimental results or numerical simulations.

The range of validity of code formulas are not stated except for isolated cases, e.g., hollow section joints. However, the code formulas, used for decades, were derived from a limited number of experiments containing specimens of sizes reasonable for experimental testing. For example, the formulas for the column web panel in transverse compression were derived from a database of tests with the load resistance between 175 and 980 kN [3]. These tests obviously did not cover the whole range of available hot-rolled cross-sections in terms of size or web slenderness. The sizes of welded cross-sections may wildly vary from those tested. As a result, the engineering practice believes that the code formula is reliable for any H-shaped cross-sections while the code was developed only for a reasonable range of validity. Obviously, since 1989, many more tests were performed, and the range of validity was expanded, but concerns of misuse of code formulas for very different cases still remain.

With the advancements of computer performance, numerical simulations, and artificial intelligence, it is nowadays possible to generate enough data to create a machine-learned algorithm with a much lower coefficient of variation to numerical simulations and experimental results. Even though the numerical models can be validated only on the available set of experiments, numerical simulations are able to expand the range of validity with higher precision than simplified formulas that neglect some physical phenomena. However, it should be kept in mind that the resulting set of weights and biases in the neural network or derived formulas also has drawbacks:

- They are totally ineligible to humans
- They provide good results only in the range of validity used for training. Extrapolation cannot be used.

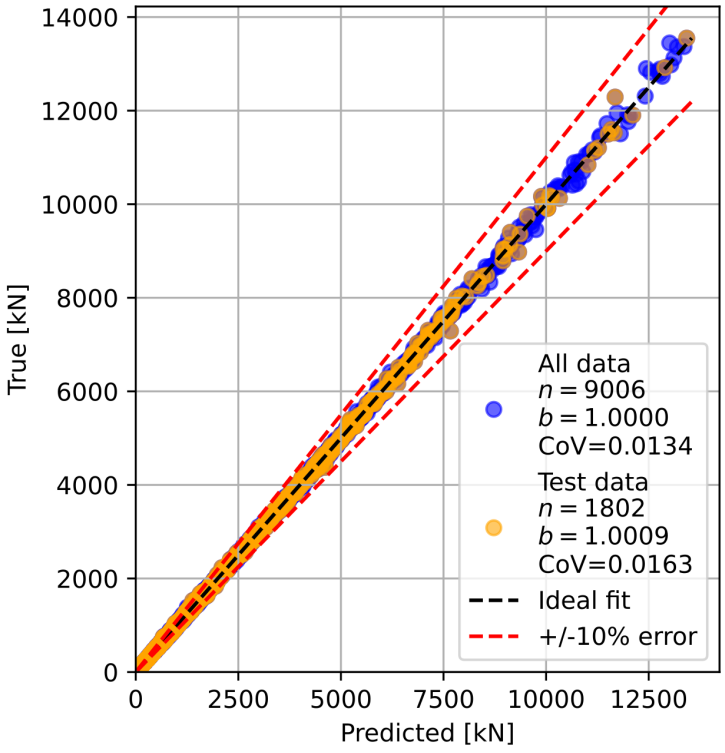


Figure 1.1: *Comparison of machine-learning prediction example for column web panel in transverse compression near the unstiffened end*

The machine-learned formulas can be seen as a modern form of curve-fitting, where the dependencies are no longer searched for by humans but by the activation functions of deep neural networks. The accuracy of predictions within the range of validity where the sufficient amount of data is provided is striking, surpassing any reasonable prediction by human-derived curve-fitting

formulas; see the reliability graph in Fig. 1.1. The coefficient of variation (CoV) in this example is only slightly higher than the expected numerical error. In essence, the quality of machine-learned prediction is nearly the same as the quality of finite element analysis that was used to generate the dataset.

Deep neural networks are able to predict the datasets extremely closely, but two issues need to be tackled:

- Creation of a reliable dataset with a sufficient range of validity;
- Building an understanding of artificial intelligence in the engineering community and creating a set of rules preventing misuse and dangerous design.

1.2 Finite Element Analysis

There is an ever-increasing reliance on numerical calculations to replace traditional design procedures prescribed in design standards such as EN 1993-1-1 [4] and EN 1993-1-8 [1]. These standards primarily address basic and idealized cases; however, real structures often involve irregular geometries or loading conditions that fall outside their scope. For decades, finite element analysis (FEA) has been widely employed to determine internal forces in structural members. These members and their joints are subsequently verified using code-based formulas, typically integrated into structural design software.

Nowadays, FEA has become the dominant method for determining internal forces in structural systems, with alternative approaches being virtually obsolete. Nevertheless, it will always be crucial to validate numerical results through simplified hand calculations and by inspecting the deformed shape of the structure. However, the checks of members and connections are often performed by code formulas despite the fact that finite element analysis may provide much smaller variability than these simple formulas. The creation of reliable numerical model for design of a substructure, members, or connections is a complicated process that was formalized in FprEN 1993-1-14:2024 [5]. Generally, the following steps are necessary:

1. Mesh sensitivity
2. Verification
3. Validation
4. Reliability analysis

Each step will be explained in detail in the following chapters. Only thoroughly investigated models can be used for reliable structural design in practice.

Important to note is also the distinction between:

- Numerical simulation
- Numerical design calculation

Numerical simulation attempts to replace the physical experiment. After numerical simulation is proven to coincide well with the experiment, its parameters may be slightly modified to create a large database of simulated experiments that are much cheaper to produce. Furthermore, they are not encumbered by the uncertainties of physical experiments and may provide a more straightforward correlation to physical parameters. Such simulations may be further analyzed to produce design formulas. To obtain design resistance, numerical simulation must be subjected to reliability analysis according to EN 1990 – Annex D [6] that produces a partial safety factor γ_{FE}. Numerical simulation typically contains solid finite elements that allow for precise modeling of geometry. True-stress true-strain material models are typically used for validation with the experiment. However, the parametric studies may transition to statistically confirmed material models, such as [7], [8] for hot-rolled steels or [9] for cold-formed steels.

Numerical design calculation directly provides the design resistance. It is usually a much simpler model allowing fast calculation of a large number of load effects. Typically shell or beam elements are used and the geometry is simplified with nominal dimensions. Bi-linear material models are used with perfectly plastic or only slight hardening, which is meant to improve convergence of the iteration process rather than to utilize the reserve of steel.

1.2.1 Failure Criteria

Numerical design calculations should have clearly defined criteria that determine model load resistance. Failure criteria of numerical models are [5]:

- C1: The maximum load of the load-deformation response – typically obtained by buckling
- C2: The load corresponding to a limiting deformation or strain criterion

Failure criterion C2 is nowadays heatedly discussed in technical committees for Eurocode development. The new Eurocode FprEN 1993-1-14:2024 *Eurocode 3 – Design of steel structures – Part 1-14: Design assisted by finite element analysis* [5] did not bring a clear definition. The issue of the plastic strain limit for numerical design calculation was moved to the relevant annexes of other parts of Eurocode 3. That may be ideal provided that every committee will invest research capacities into determining the highest strain limit for the given steel grade and failure mode that still provides sufficient reliability.

However, a clear criterion that can be used in algorithms is needed. For decades, in mechanical, aerospace, and civil engineering, the plastic strain limit of 5% has been used for structural steel with sufficient ductility (at least 15% for standardized tensile tests [10]). This limit is not only for the material ductility but also for the finite element formulation that may produce increasing error with increasing plastic strains. This is especially true for commonly used finite elements and analysis types:

- Some solid finite elements are poor in plasticity and sensitive to mesh distortion. The well-known culprit is the linear tetrahedral element that should be avoided for regions where yielding is expected. On the other hand, this element allows for easy meshing of complicated geometries.
- Shell finite elements with commonly used settings remain with constant thickness neglecting the Poisson effect (necking) in one direction.
- Small displacement analysis (geometrically linear) is sufficient for many engineering problems. However, for increasing deformations and strains, only geometrically nonlinear analysis can provide precise results.

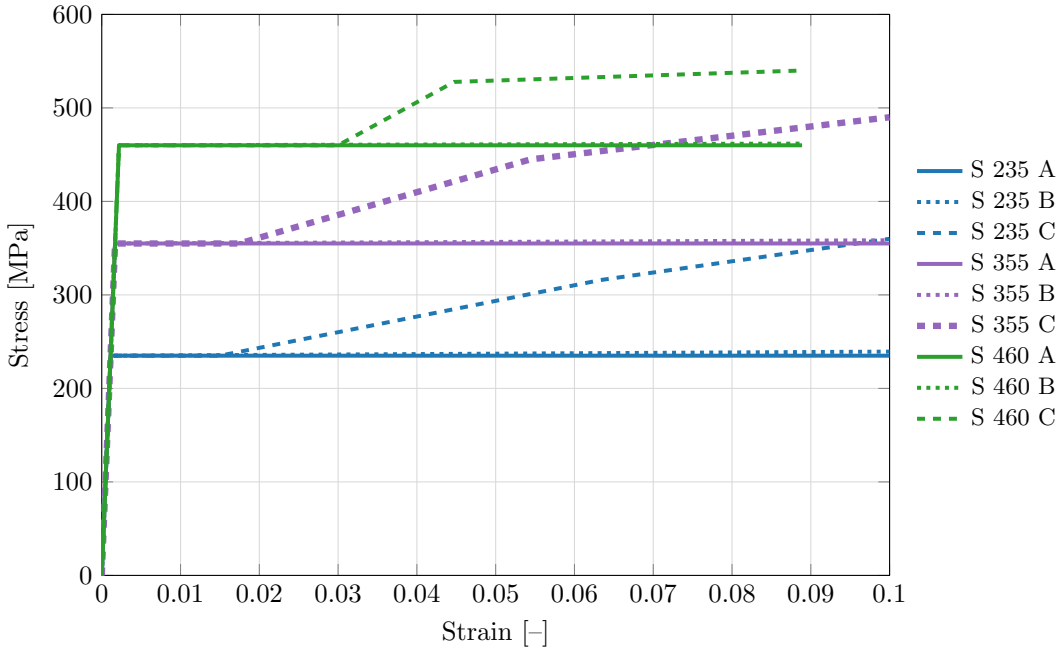


Figure 1.2: *Design material models according to EN 1993-1-14 [5] for S 235, S 355, and S 460*

Furthermore, the chosen material model and failure criteria should consider material and geometry tolerances given in EN 1090-2 [11]. Tolerances in this standard may be assumed as the

maximum imperfections. Anything worse should be discarded. Eurocode assumes normal distribution for geometrical and material properties for structures within these tolerances; see Annex E of FprEN 1993-1-1 [12]. It must be shown that using nominal material properties, nominal geometry (while plate thickness is typically the most important with the highest coefficient of variation) without manufacturing tolerances with the selected failure criteria provide the reliable design resistance in terms of EN 1990 [6]; see Eq. (2.30).

For the failure modes of yielding and tensile rupture, the governing parameter of numerical design calculation is the plastic strain limit. It is assumed that geometry is modeled as perfect, material model with nominal properties listed in EN 1993-1-14, Cl. 5.3.2 [5] is used (see Fig. 1.2):

- linear elastic – perfectly plastic without strain hardening (A)
- linear elastic – perfectly plastic with a nominal plateau slope for numerical stability (B)
- linear elastic – linear hardening plastic material model (quad-linear material model with strain hardening) [7] (C)

If the plastic strain limit is constant, it can be assumed that a linear elastic – perfectly plastic material model without strain hardening will always be the safest option. The material model with a nominal plateau slope for numerical stability as specified in EN 1993-1-14 [5] (with strain-hardening slope $E_{sh} = E/10000$) provides only an extremely small increase in strength. It can be assumed that the quad-linear material model (C) provides the most dangerous results if the plastic strain limit is expected around 5 %.

The goal is to determine a reliable plastic strain limit for plates and hot-rolled steel profiles. The most dangerous case is assumed to be the net section failure with the highest stress concentrators allowed in execution standard EN 1090-2 [11] with the largest net-section reduction allowed in EN 1993-1-8 [1], and therefore, it was selected for thorough investigation. It can be argued that strain limit is also important in welds and in the heat-affected area, however, the resistance of welds and their imminent vicinity may be calculated by code formulas and is not investigated in this work.

Finding a reliable strain limit for numerical design calculation is the central topic of project Inter Excellence LUAUS23114 of the research team from Czech Technical University in Prague, Brno University of Technology, and University of Tennessee, Knoxville. The outcomes of this research are shown primarily in Chapter 3.

1.3 Reliability

In recent years, the level of detail in numerical modeling has significantly increased. Not only are global structural analyses performed using FEA, but members and joints are now frequently modeled with shell elements to capture local effects. The outputs of such analyses—primarily stresses and strains—are directly utilized for design verification, enabling the construction of structures based on these advanced simulations. For instance, IDEA StatiCa Connection alone performs over one million numerical design calculations each month. These models encompass both standard configurations, such as end-plate connections, and highly complex cases where an engineer would need to bend the code clauses and use engineering judgment extensively. The need to assess the reliability of FEA that is used for such a vast number of designs is clear.

Eurocodes were developed with a reliability target $\beta = 3.8$ with the reference period of 50 years, which corresponds to the yearly probability of failure approximately 10^{-6}. The same reliability should be ensured by numerical design calculations, which is almost never done and the guidelines on how to assess reliability are not well-known and publicly available. This thesis presents a step-by-step procedure based on several documents – Eurocode background [13], SAFEBRICTILE project Final report [14], and most importantly SAFEBRICTILE deliverable D1.1 [15].

The reliability assessment of numerical design calculation requires a set of physical experiments. The design resistance is drastically affected if the number of specimens is low; see Fig. 1.3. Ideally, more than 100 specimens are available, which can be assumed as sufficient to use the minimum value of k_{dn}, which is used as a multiplier of the coefficient of variation V_X. The reliability target β is divided between the probability of loads by the factor $\alpha_E = 0.7$ and resistances by the factor $\alpha_R = 0.8$. This way, the reliability may be tackled separately. According to Tab. D2 of EN 1990 [6], $k_{dn} = \alpha_R \cdot \beta = 0.8 \cdot 3.8 = 3.04$ for an infinite number of specimens (which can be assumed 100).

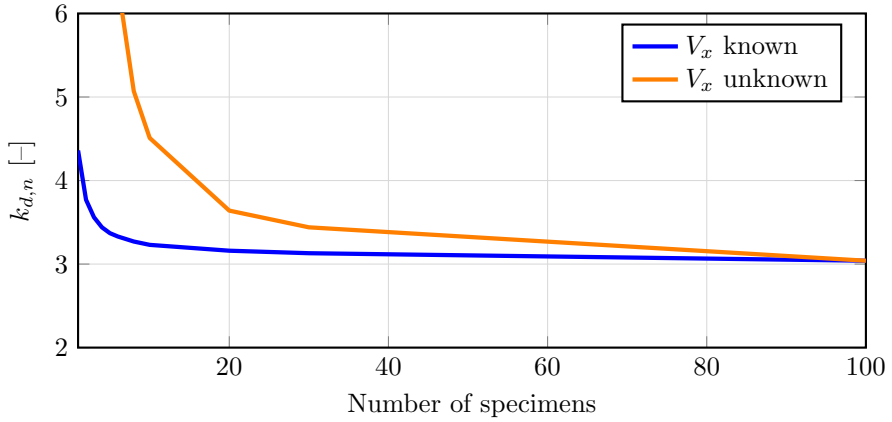


Figure 1.3: *Values of k_{dn} based on number of specimens*

On the other hand, the specimens and their corresponding numerical design calculations should be divided into batches, which typically decreases the coefficient of variation of experimental resistance to model resistance ratio.

The reliability assessment procedure is described in detail in Section 2.6 and evaluated for Eurocode design procedure, a new proposed analytical method, and machine-learned predictions in Section 4.6.

1.4 Machine Learning

Recent development of neural networks providing machine-learned formulas allows very precise model predictions taking into account all input variables. Machine learning is included in many open-source packages, such as PyTorch [16] for the Python programming language. The programming of neural networks is well documented, and simple programs can therefore be constructed easily with the help of large language models (LLMs), such as ChatGPT [17]. Anyone can nowadays simply define a neural network that is able to provide predictions as an alternative to a design formula. However, several aspects need to be kept in mind:

- Training data should be high-quality. In cases where finite element analysis is used to generate them, the reliability of training data should be at or above the Eurocode target.
- Due to the nature of activation functions, the machine learning predictions can be used only inside a defined range of validity (where training data are available in sufficient density). The range of validity must be stated and strictly adhered to. If any generality is required from the neural network, the number of data points must typically be in tens of thousands.
- Machine-learning predictions will introduce additional error to the error of finite element analysis. It is expected that the reliability of predictions will be lower than the reliability of the original training data. Reliability analysis must be repeated for the machine-learning predictions.

The ultimate goal is to provide a framework to enhance prediction accuracy while maintaining Eurocode-level reliability in structural engineering design. All tools are at our disposal now. The following chapters provide examples and the author's experience with this goal in mind.

Three important topics are discussed in detail in Sections 2.4, 2.5 and Chapter 4.7:

- Data generation and scaling
- Neural network architecture to obtain the best predictions
- Interpretation and explainability of prediction behavior

Chapter 2

Methods

The author works as a Product Owner at IDEA StatiCa and his responsibility is the reliability of component-based finite element model (CBFEM) calculations. He is involved in a large number of comparisons to Eurocode [1], AISC [18] and the other six design codes across the world [19], [20], [21]. In the author's work, differences between design formulas and component-based finite element models are apparent in several cases. This indicates that either the numerical model or the design formula is wrong. The author's job is to find out which is the case and if needed to improve the numerical model. Two cases are presented here:

- Component Column web in transverse compression
- Tensile rupture of a weakened plate

The column web in transverse compression project started in 2022 with a project by colleagues from Brno University of Technology [22]. Here, CBFEM was confirmed by more detailed ANSYS [23] model with solid finite elements SOLID 186. The author and his colleagues continue with this project; several dozen experiments were performed and articles published [24], [25], [26]. This work continued with a joint project of IDEA StatiCa and ISISE, Portugal, with a focus on column web in shear. Inevitably, the component column web in transverse compression is also present [27]. All these studies show significant differences between design formulae in the first [1] and second [28] generation of Eurocode, and AISC Specification [18]. In the case of significant compressive force in the column, the design codes even provide unsafe results compared to the design numerical model in Abaqus. Therefore, the definition of the component column web in transverse compression can be further improved.

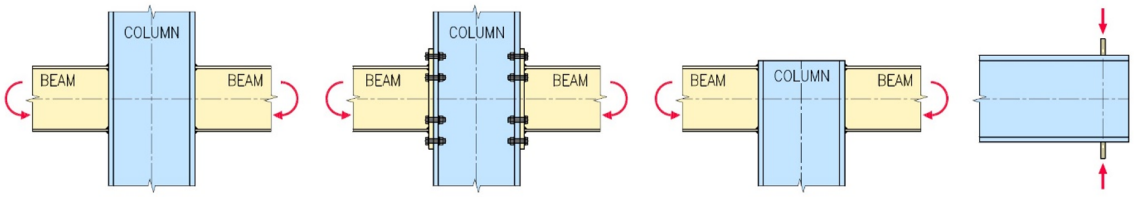
The tensile rupture (net section failure) of a weakened plate shows a smaller deviation between code predictions and the experiments. The experimental resistance is typically simply equal to the net section area multiplied by ultimate strength with small variations caused by stress concentrations and the presence of shear stresses causing the disruption of the uniaxial stress state. The primary concern is then the calculation of net section area [29] for staggered bolt holes, which seem to provide much less conservative values than the code formula for regular bolt-grid, as will be shown later. The experiments of weakened plates are relatively simple, which allows performing large sets with multiple parameters like steel grade, plate thickness, or stress concentrations. Furthermore, the variability of crucial parameters was investigated and is nowadays codified in EN 1993-1-1:2024 [12] with the boundaries given by manufacturing tolerances in EN 1090-2 [11].The primary objective of these tests was to find a reliable plastic strain limit for structural steel that could be used for any type of steel structures and joints for the failure modes of yielding and tensile rupture. The improvement of the formulas for net-section failure is outside the scope of this thesis and rather a numerical-analytical method is used to generate samples for Monte-Carlo simulation.

2.1 Design codes and state-of-the-art

2.1.1 Web panel in transverse compression

The component column web panel in transverse compression is present in unstiffened beam-to-column connections or at locations of points of load concentrations, i.e., patch loading. This

component is described in Eurocode EN 1993-1-8:2005, Cl. 6.2.6.2 [1], EN 1993-1-8:2024, Cl. A.6 [28], or AISC 360-22, Cl. J.10 [18]. The state-of-the-art is described in detail in papers [24], [30] and it will be shortly also introduced here.



(a) (b) (c) (d)

Figure 2.1: *Examples of web panel in transverse compression active component: welded (a) and bolted (b) beam-to-column connection, roof connection (c), and load near the unstiffened end(d) [30]*

The elastic behavior could be mathematically derived, but the phenomenon of yielding is actually complicated and parameters are set empirically from a limited number of experiments [31].

In EN 1993-1-8 [28], the CWC resistance is obtained as the minimum of web yielding and web buckling expressions based on an effective width into which the load is dispersed; the formulation does not explicitly account for proximity to a free member end. EN 1993-1-5 [32] adopts an effective-length concept that embeds plate-buckling effects and distinguishes between loading at one flange, at both flanges, and near an unstiffened end. While broadly safe, parts of these models rely on empirical parameters calibrated over limited slenderness ranges, and simplifications in the assumed boundary conditions (primarily simply supported web-to-flange) may lead to conservatism. [33, 34]

The American Specification AISC 360-22 [18] treats web local yielding with expressions equivalent to Eurocode away from member ends, and provides separate formulations for web buckling (for load acting from both sides) and crippling (for load acting from one side only). The resistances are halved when the load is applied near the unstiffened end. For web local yielding, the boundary is at a distance from the member end that is less than or equal to the full nominal depth of the member, h. For web local crippling and compression buckling, the boundary is at a distance from the member end that is less than $h/2$.

Recent experimental campaigns and validated GMNIA studies have sharpened the understanding of the actual mechanisms [24]. For loading away from member ends, plastic strains concentrate at the flange–web transition and, for typical rolled IPE sections, local buckling seldom governs the resistance when realistic fixity at the junction is present. Interaction with axial force can, however, significantly diminish column web panel resistance—an effect not fully captured by the current EN 1993-1-8 [28] reduction factor k_{wc}.

Balazs [30] addressed a case still weakly codified: transverse compression applied close to an unstiffened end. Combining full-scale tests on IPE 200 specimens with GMNIA across IPE 100–600, the end-distance effect and its dependence on web slenderness was quantified. When the load is very close to the unstiffened end (e.g., $e \approx 0.1h$), the resistance may reduce to about 60% of the far-from-end case, and buckling becomes more likely due to the loss of restraint at the free edge. An analytical procedure consistent with current design philosophies is proposed to capture these effects.

This phenomenon is therefore ideal as a test case for finite element analysis and subsequent machine learning predictions.

Only the procedure of FprEN 1993-1-8:2024, Cl. A.5 [28] is described below, because the other code implementations are very similar and older.

The design resistance is the minimum of the yielding and buckling resistances:

$$F_{c,wc,Rd} = \frac{\omega\, k_{wc}\, b_{eff,w}\, t_{wc}\, f_{y,wc}}{\gamma_{M0}} \tag{2.1}$$

$$F_{c,wc,Rd} \leq \frac{\omega\, k_{wc}\, \rho\, b_{eff,w}\, t_{wc}\, f_{y,wc}}{\gamma_{M1}} \tag{2.2}$$

The spreading angle for the determination of $b_{eff,w}$ starts at the weld outer edge and is 1:1 through an end plate and 2.5:1 through the flange and web-to-flange radius (of hot-rolled) or leg size of fillet weld (of welded) of the compressed member; see Fig. 2.2.

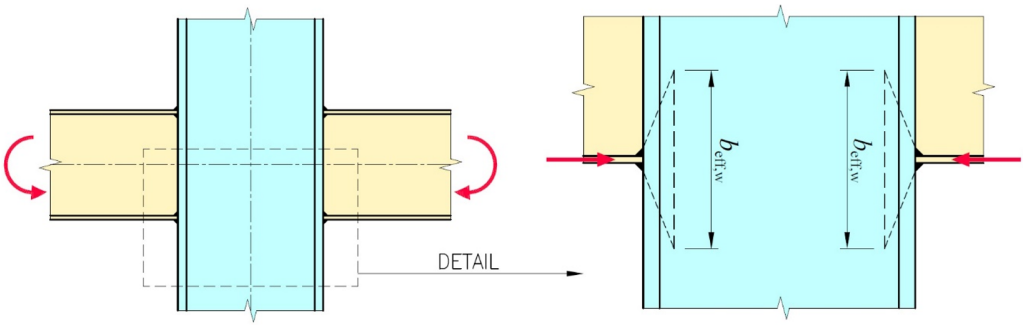


Figure 2.2: *Spreading angle through the rolled transversely compressed member [30]*

Reduction factor ω depends on the ratios of bending moments of the beams on the opposite side of the column, the transformation parameter β that can be taken from Cl. 7.2.3 [28]. In the dataset, the forces coming from the opposite sides will be equal, and therefore, $\omega = 1$ as can be seen from Tab. A.1. [28]

Further variables are t_{wc}, column web thickness, $f_{y,wc}$, column web yield strength, and $\gamma_{M0} = 1.0$, partial safety factor.

Buckling failure mode is determined by the buckling reduction factor ρ. Buckling of the web in transverse compression is based on the following assumptions:

- Web is pinned into the flange [35]
- The effective width, $b_{eff,w}$ is irrelevant for the calculation of critical force
- The web plate is assumed to be infinitely long. Or in other words, the end is far enough that it does not affect the buckling resistance nor the critical force.
- If any stiffeners are nearby, but not directly under the point of compression, their contribution is disregarded.

The buckling curve was modified in the second generation of Eurocode [28] to reflect Winter curve [36], which actually made buckling even more likely to govern the design.

$$\rho = \begin{cases} 1.0, & \text{if } \bar{\lambda}_p \leq 0.673 \\ \frac{\bar{\lambda}_p - 0.22}{\bar{\lambda}_p^2}, & \text{if } \bar{\lambda}_p > 0.673 \end{cases} \tag{2.3}$$

Relative slenderness is calculated:

$$\bar{\lambda}_p = 0.932 \sqrt{\frac{\omega\, k_{wc}\, b_{eff,w}\, t_{wc}\, f_{y,wc}}{E\, t_{wc}^2}} \tag{2.4}$$

The following paragraphs are not included in the code [28], but they are included in the paper in preparation [30]. There an analytical approach is presented based on the optimization algorithms and engineering judgment. This serves as a very good comparison to machine-learning.

In [30], the author and his colleagues have shown how this design method could be adapted to better reflect the numerical simulations performed with ANSYS [23] in [37].

Spreading angle

Firstly, the results of materially nonlinear analysis (MNA) at 5% plastic strain, which is assumed to simulate the design yielding resistance, were used to obtain the effective width, i.e., the spreading angle set to 2.5:1 in codes. Two methods were used:

1. Division of numerical resistance by web thickness and yield strength

$$b_{eff,w} = \frac{F_{Rk,MNA}}{t_{wc} \cdot f_{y,wc}} \tag{2.5}$$

2. Distribution of stresses normal to the transverse load at the web directly below the rounding. Only the positive stresses were taken into account in the equivalent area averaging; see Fig. 2.3.

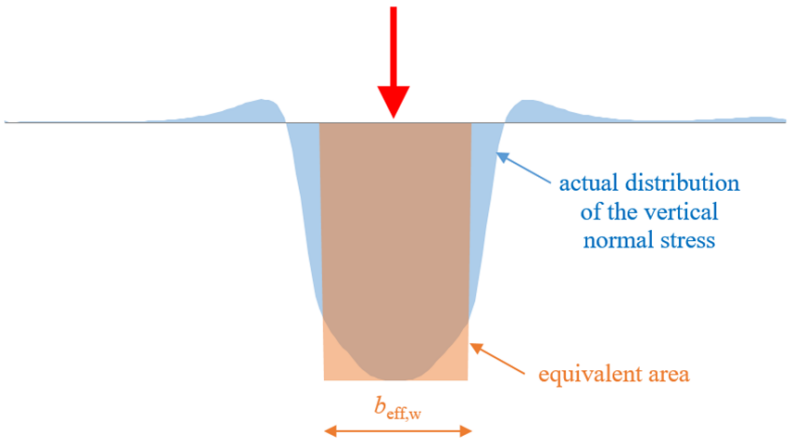


Figure 2.3: *Determination of the effective width $b_{eff,w}$ using normal stress σ_z [30]*

The comparison of effective widths for IPE cross-sections determined using the current Eurocode approach with the spreading angle 2.5:1 (EC3), division of numerical resistance ($F_{Rk,MNA}$), and the distribution of normal stress (σ_z) is shown in Fig. 2.4.

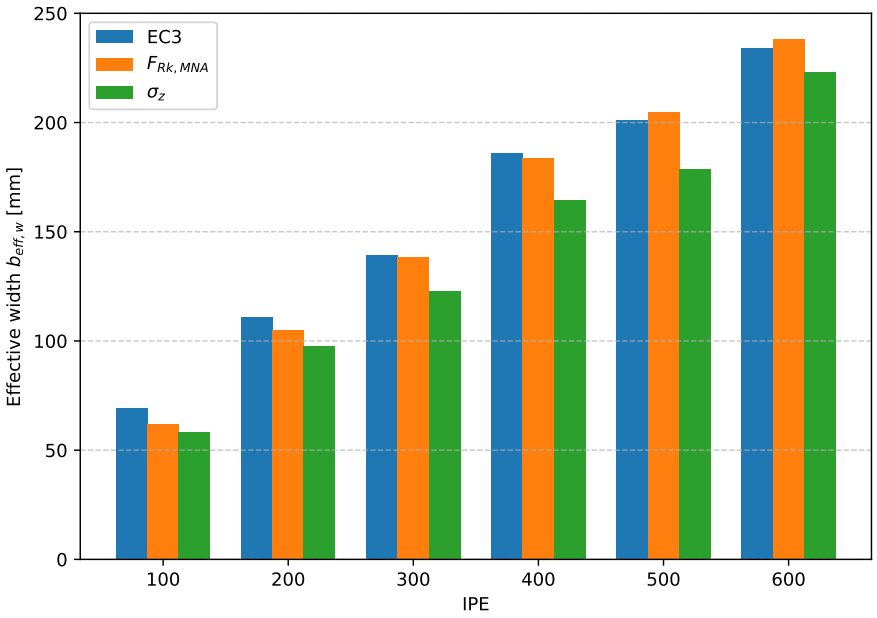


Figure 2.4: *Effective width determined by Eurocode and numerical methods*

It can be seen that the approach using σ_z is the safest, which is probably caused by a more complicated stress distribution including also shear stresses and σ_z cannot be directly compared to yield stress. Fig. 2.5 shows the reliability graphs. These contain number of specimens n, bias b (see Eq. (2.32)), and variation coefficient V_δ (see Eq. 2.36) for quick first order reliability method (FORM) assessment. $F_{Rk,MNA}$ approach can be directly compared to EC approach where the

spreading angle 2.5:1 through the flange and radius shows good estimation with IPE 300 and larger cross-sections and for small cross-sections (IPE 100 and IPE 200) is rather optimistic. In the subsequent calculations, the spreading angle of 2.5:1 is retained. Future work could focus on the source of this issue and perhaps a separate spreading angle through the flange and through the radius.

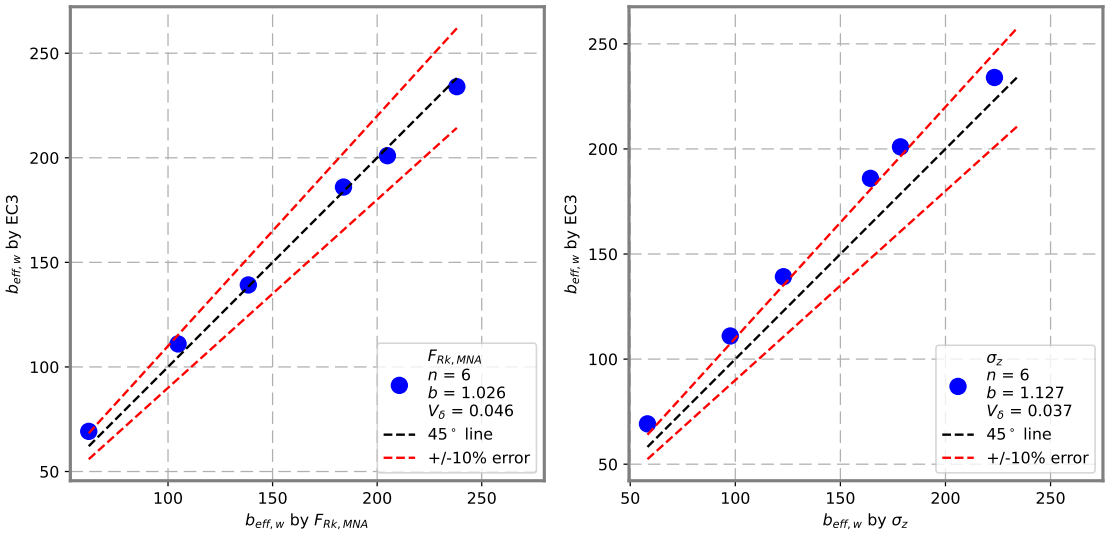


Figure 2.5: *Reliability graphs of the effective width determination by EC3 in comparison to division of numerical resistance $F_{Rk,MNA}$ (left) and distribution of normal stress σ_z (right) [30]*

The resistance determined by MNA decreases as the distance towards the unstiffened end is getting closer. This could be expressed by decreasing the effective width as calculated by Eq. (2.5) by ratio r_e:

$$r_e = \frac{b_{eff,w,e}}{b_{eff,w}} \tag{2.6}$$

where:

- $b_{eff,w,e}$ – effective width for load near the unstiffened end
- $b_{eff,w}$ – effective width for load unaffected by the unstiffened end

From the numerical models, it can be observed that the yield resistance F_{Rk} is not affected by the unstiffened end if the load is further from the edge than $x = e/h = 0.45$, i.e., the ratio $r_e = 1.0$ for $x \geq 0.45$. For the cases closer than that, it seems that a parabola fits well with the datapoints; see Fig. 2.6. Due to this, the coefficient of variation V_δ is rather high, 7.2 %, but for higher loads, the differences are low. It can be assumed that the fitted parabola simulating the effect of unstiffened end proximity is in good agreement with the numerical data.

$$r_e = -2.1 \cdot (x - 0.45)^2 + 1 \tag{2.7}$$

The reliability graph in Fig. 2.6 shows the most significant deviations at low loads, meaning for cross-sections IPE 100 and IPE 200. The difference originates primarily from the assumed distribution angle 2.5:1, which diverges by 10 % and 5 % for IPE 100 and IPE 200, respectively.

Degree of fixity of web to flange transition

The critical force for the plate of infinite length supported at two opposing edges was derived by Timoshenko [38]:

$$F_{cr} = k_{cr} \cdot \frac{\pi^2 \cdot E \cdot t_{wc}^3}{12 \cdot (1 - \nu^2) \cdot d_{wc}} \tag{2.8}$$

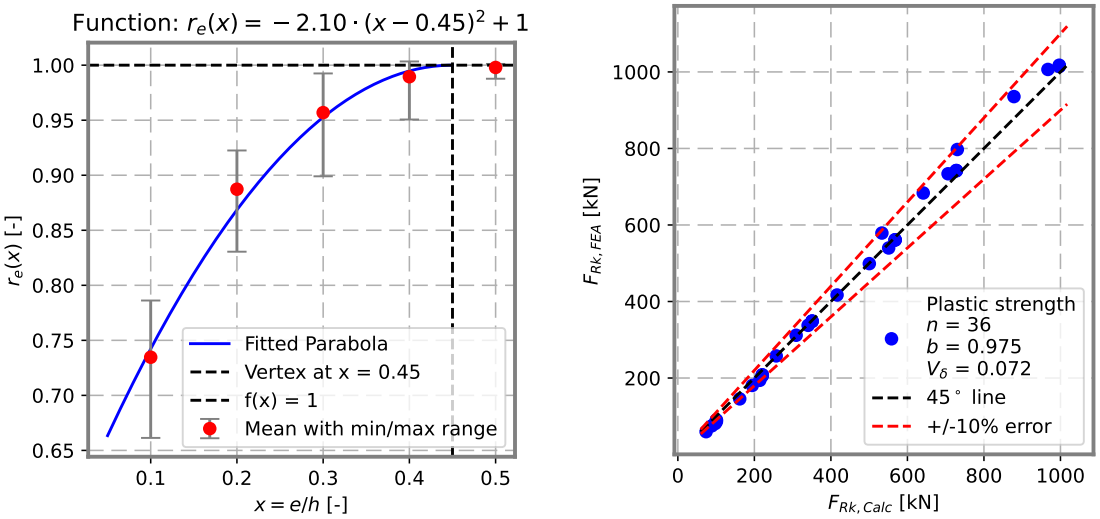


Figure 2.6: *Reduction of effective width in relation to end distance; reliability graph of F_{Rk} estimated by fitted formula [30]*

$$k_{cr} = \begin{cases} 4/\pi, & \text{for simply supported edges} \\ 14/\pi, & \text{for fixed edges} \end{cases} \tag{2.9}$$

where:

- $E = 210000$ MPa – steel modulus of elasticity
- $\nu = 0.3$ – Poisson coefficient
- $d_{wc} = h - 2 \cdot t_{fc} - 2 \cdot r$ – web height between radii (for hot-rolled) or edges of fillet welds (for welded)
- k_{cr} – factor for the degree of fixity of web plate into flange plates [35, 39]

The values of F_{cr} are available from LBA, and since all other parameters are known, k_{cr} can be calculated directly. For the general case where the unstiffened end is sufficiently far away and influences neither the yielding nor the buckling strengths of the web in transverse compression (see $1.0 \cdot h$ in Fig. 2.7), the degree of fixity is on average 3.619, which is closer to the fixed condition ($14/\pi = 4.456$) than to the pinned condition assumed in the code [28] ($4/\pi = 1.273$); see Fig. 2.7 and Eq. (2.9).

For the cases of load near the unstiffened end, the critical force F_{cr} calculated by LBA is decreasing. This can be explained by the fact that the web plate is not infinitely long but its end is getting closer. The buckling behavior is thus more appropriate for the plate supported on three sides rather than all four sides. The degree of fixity of the web to the flanges likely remains the same, but for simplicity, the same factor k_{cr} is used to account for this decrease in the critical force F_{cr}. Very conveniently, the values of k_{cr} show small variation across all IPE cross-sections; the average coefficient of variation is 0.070.

The averages of k_{cr} across all the IPE sections seem to follow a bilinear trend when plotted against relative distance towards the end $x = e/h$, where e is the distance between the load application and the unstiffened end and h is the cross-section depth. By minimizing mean squared error, the following bilinear function was constructed:

$$k_{\mathrm{cr}}(x) = \min\left(3.6,\ 5.7 \cdot x\right) = \min\left(3.6,\ 5.7 \cdot \frac{e}{h}\right) \tag{2.10}$$

Fig. 2.8 proves that the critical force calculated by the proposed formulas closely fit the numerical solution by the LBA.

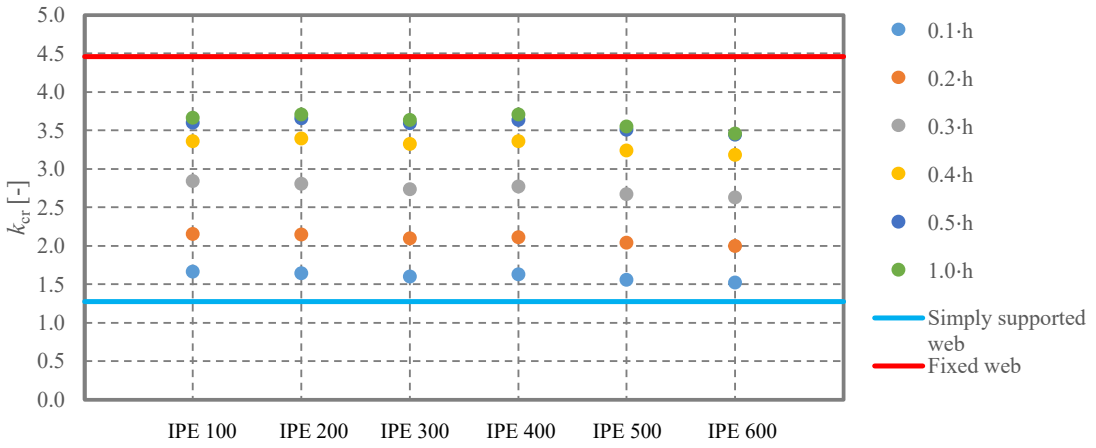


Figure 2.7: *The degree of fixity k_{cr} for varying distance to the unstiffened end [30]*

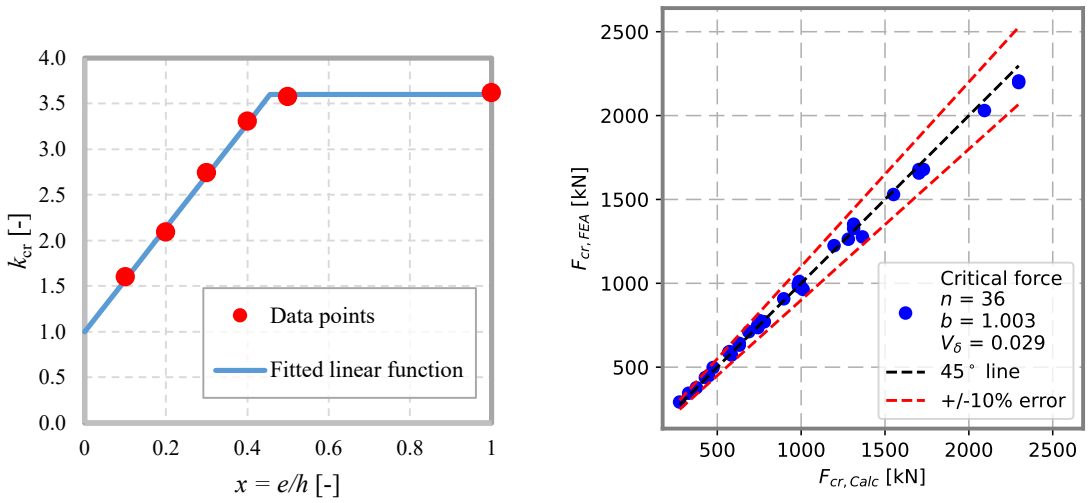


Figure 2.8: *Fitted bilinear curve for k_{cr} as a function of x and comparison of critical forces determined by analytical calculation and by LBA [30]*

Relative slenderness

The code equation (see Eq. (2.4)) for relative slenderness can be rewritten according to EN 1993-1-5, Annex B [32] to:

$$\bar{\lambda}_p = \sqrt{\frac{\alpha_{ult,k}}{\alpha_{cr}}} = \sqrt{\frac{F_{c,wc,Rd}}{F_{cr}}} \tag{2.11}$$

Finite element analysis directly provides the $\alpha_{ult,k}$ or $F_{c,wc,Rd}$ by MNA with failure criterium such as $\varepsilon_{pl} = 5\%$ and α_{cr} or $F_{cr} = F \cdot \alpha_{cr}$ by LBA, where F is the applied load. By the assumption of spreading angle 2.5:1 and above-mentioned equations, the compressive resistance and critical force may also be calculated. The comparison is in Fig. 2.9 showing good approximation with $V_\delta = 4.8\%$ and higher deviations at low values of relative slenderness, where the difference is not extremely important.

Buckling curves

Buckling curves were the most challenging to determine due to the limited amount of data points for various relative distances of the compressive force to the unstiffened end x. Especially for the basic case of $x = 1$, the relative slenderness $\bar{\lambda}$ did not fall below 0.7. More significant reduction

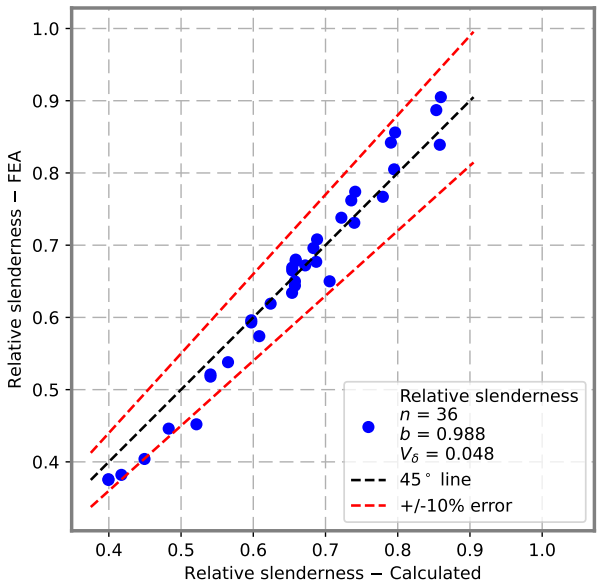


Figure 2.9: *Relative slenderness calculation [30]*

due to buckling occurred only for load close to the unstiffened end; see Fig. 2.10. This makes any optimization difficult and unreliable; therefore, engineering judgment was employed extensively.

Buckling curves in the form specified in EN 1993-1-5, Annex B [32] were chosen:

$$\rho = \frac{1}{\phi_p + \sqrt{\phi_p^2 - \bar{\lambda}_p}} \tag{2.12}$$

$$\phi_p = \frac{1}{2} \cdot \left[1 + \alpha_p \cdot (\bar{\lambda}_p - \bar{\lambda}_{p0}) + \bar{\lambda}_p\right] \tag{2.13}$$

By optimization algorithms, using the Differential Evolution function in `scipy.optimize` package in Python [40], the parameters of α_p and $\bar{\lambda}_{p0}$ were sought. The datapoints were the reduction factor ρ, calculated as the ratio of GMNIA to MNA resistances, and relative slenderness calculated by Eq. (2.11) from MNA and LBA. The optimization algorithm was able to minimize the error, but the optimized buckling curves made little sense; see Fig. 2.10. The imperfection factors α_p were wildly different, which should not happen for the same hot-rolled sections.

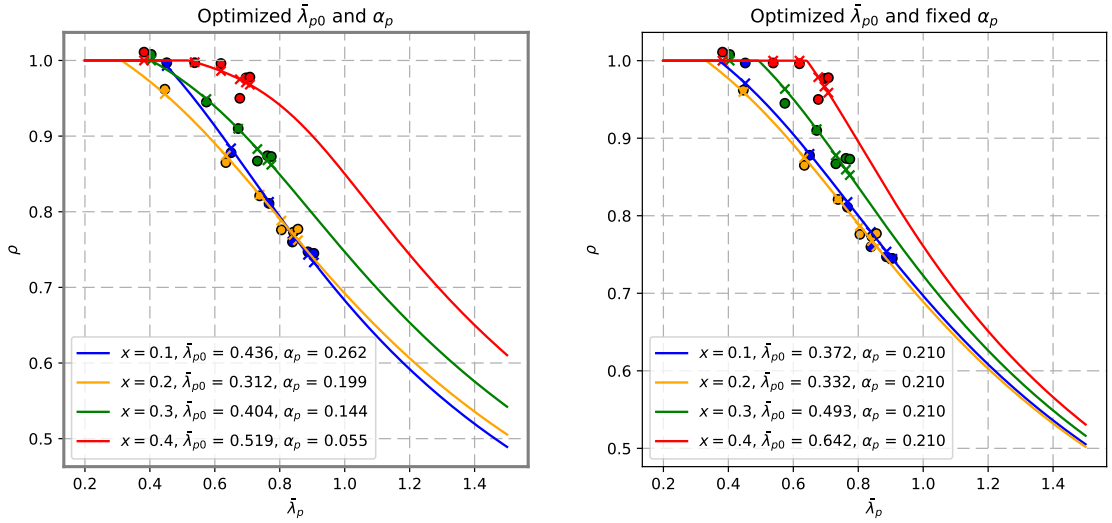


Figure 2.10: *Optimized buckling curves [30]*

The second step was to fix the imperfection factor α_p to 0.21, which seems reasonable for hot-rolled IPE sections and is close to the average for distances $x = 0.1$, 0.2, and 0.3. It was assumed that the distance $x = 0.4$ consists of too few points, the reduction ρ does not go below 0.9, and the imperfection factor for the red curve is meaningless. The error was slightly larger, but still within acceptable limits.

However, it was assumed that as the distance x is decreasing, the plate edge is getting closer and the buckling reduction should be more prominent. This is true for most cases, but strangely, the $\bar{\lambda}_{p0} = 0.372$ for $x = 0.1$ and $\bar{\lambda}_{p0} = 0.332$ (smaller!) for larger distance $x = 0.2$. This was assumed as a numerical error, and the final buckling curves were chosen with $\alpha_p = 0.21$ and $\bar{\lambda}_{p0}$ according to Eq. (2.14), although the error increased for $x = 0.1$ and 0.2; see Fig. 2.11.

$$\lambda_{p0} = \min\left(0.673,\ 0.223 + 0.9 \cdot x\right) = \min\left(0.673,\ 0.223 + 0.9 \cdot \frac{e}{h}\right) \tag{2.14}$$

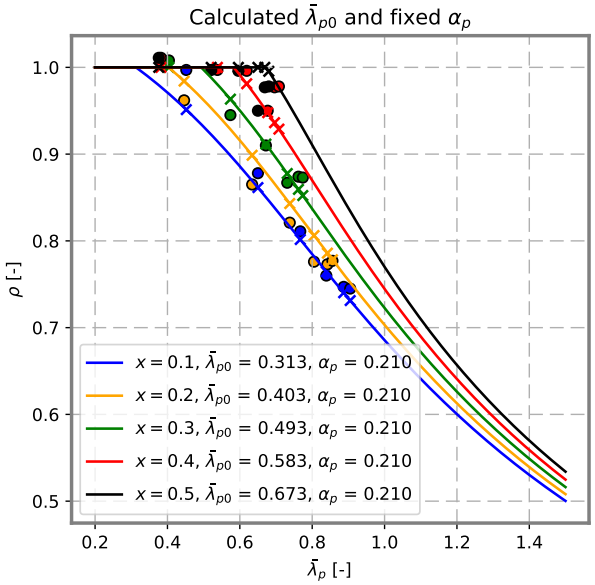


Figure 2.11: *Buckling curves according to EN 1993-1-5, Annex B [32] with calculated parameters [30]*

Buckling resistance for the whole set

The buckling resistance using above-mentioned equations were compared to the whole set of FEA models made in ANSYS [37] as well as the experiments performed at Brno University of Technology in 2023 [30]. Note that the numerical models utilize nominal geometrical and material properties, design equivalent geometric imperfections are applied, and therefore they may be considered as numerical design calculations proving design resistances [5]. For the experiments, measured geometry, yield strength, and modulus of elasticity was used in the above-mentioned formulas.

In total, 10 experiments and 153 numerical design calculations are compared; see Fig. 2.12. Several issues are found with the proposed approach:

- Welded sections have much higher resistances, which indicates the spreading angle 1:2.5 is not constant through the flange and through the fillet radius or fillet weld.
- The reduction in EN 1993-1-8 – Eq. (A.25) [28] for compressive force in the column is insufficient. The calculations are to the unsafe side.

Solving these issues is a topic for future research.

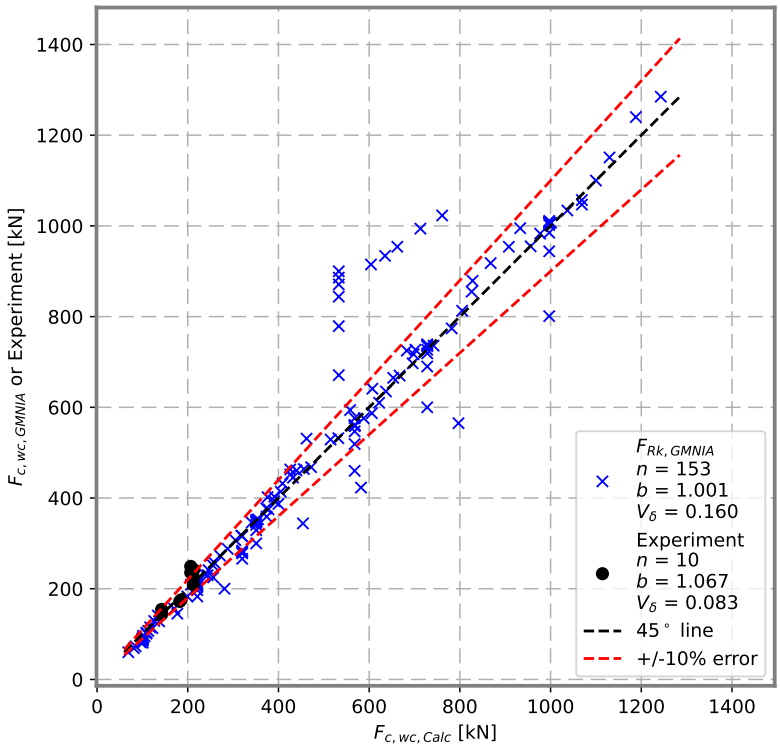


Figure 2.12: *Reliability graph of the proposed method to all ANSYS numerical design calculations [37] and BUT experiments*

2.1.2 Net section failure in tension

Net section failure of plates loaded in tension is considered as the most dangerous for finding a reliable plastic strain limit for numerical design calculations. Therefore, the approaches in design codes are briefly summarized in this section.

EN 1993-1-1 [4] requires verification of members loaded in tension in Clause 6.2.3 by the following equations:

$$\frac{N_{Ed}}{N_{t,Rd}} \leq 1.0 \tag{2.15}$$

$$N_{t,Rd} = \min\left\{N_{pl,Rd} = \frac{A \cdot f_y}{\gamma_{M0}},\ N_{u,Rd} = \frac{0.9 \cdot A_{net} \cdot f_u}{\gamma_{M2}}\right\} \tag{2.16}$$

FprEN 1993-1-1 [12] in Clause 8.2.3 modifies the net section failure as follows:

$$N_{u,Rd} = \frac{k \cdot A_{net} \cdot f_u}{\gamma_{M2}} \tag{2.17}$$

where:

- $k = 1.0$ for sections with smooth holes (i.e. holes without notches), for example holes fabricated by drilling or water jet cutting;
- $k = 0.9$ for sections with rough holes (i.e. holes with notches), for example holes fabricated by punching or flame cutting;
- $k = 0.9$ for structures subjected to fatigue.

So for standard cases, the net section resistance is 10% higher than in the first generation of Eurocodes.

Note that both generations of Eurocode consider only one possibility of shear lag effect – in an angle cross-section connected by only one flange.

AISC 360-22 [18] in Section D.2 determines the tensile strength as the minimum of tensile yielding (corresponding to gross section yielding) and tensile rupture (corresponding to net section failure):

$$P_n = \min\{\phi_{t,y} \cdot A \cdot f_y,\, \phi_{t,u} \cdot A_{net} \cdot f_u\} \tag{2.18}$$

where:

- A – gross section area
- A_{net} – effective net area specified in Section D.3
- f_y – yield strength
- f_u – ultimate strength
- $\phi_{t,y} = 0.9$ – resistance factor for tensile yielding
- $\phi_{t,u} = 0.75$ – resistance factor for tensile rupture

Interestingly, AISC 360 covers in Table D3.1 (see Fig. 3.3) a large variety of cases where shear lag effect decreases the net section area.

Note that the ratio between safety assigned to gross section yielding and net section failure is different according to each code.

$$\begin{cases} \frac{A \cdot f_y}{A_{net} \cdot f_u} \cdot \frac{0.9}{0.75} & \text{for EN 1993-1-1,} \\ \frac{A \cdot f_y}{A_{net} \cdot f_u} \cdot \frac{1.25}{0.9} & \text{for FprEN 1993-1-1,} \\ \frac{A \cdot f_y}{A_{net} \cdot f_u} \cdot \frac{1.25}{1.0} & \text{for AISC 360.} \end{cases} \tag{2.19}$$

Provided that the ratio of f_u/f_y remains constant, with increasing weakening, net section failure begins to govern first according to the current EN 1993-1-1, then according to the second generation FprEN 1993-1-1, and lastly according to AISC 360.

2.2 Experiments on web in transverse compression

This section describes the physical experiments performed at BUT and elsewhere. The simplest setup is to apply the load from both sides against an I-section; see Fig. 2.13. Applying load from one side is complicated, because the web panel in shear or the member in bending may fail first. The geometry for such testing is therefore limited.

The overview of experimental research is given in Tab. 2.1. Most tests are for double-sided concentrated load [30, 24, 44, 35]. Note that the four specimens included in [24] are the same as

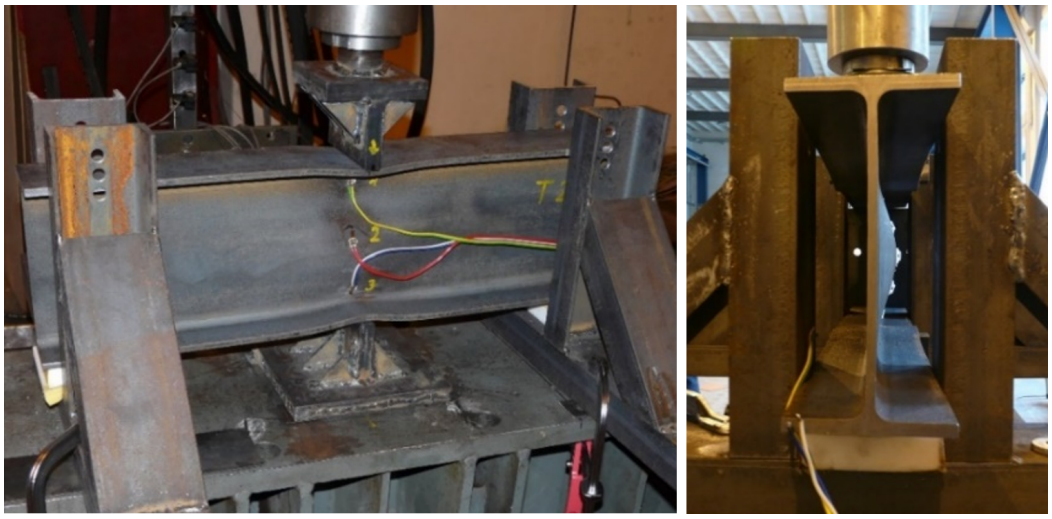


Figure 2.13: *Experiments performed at BUT – test setup (left) and web buckling after significant yielding (right) [24]*

Table 2.1: *Overview of experimental programs on column web in compression*

Paper	**No. of experiments**	**Measured geometry**	**Measured material**
Balázs et al. (2025) [30]	10	Yes	Yes
Bougoffa et al. (2022) [41]	2	No	Yes
Bougoffa et al. (2021) [42]	4	No	Yes
Rodilla and Kowalkowski (2021) [43]	7	No	No
De Mita et al. (2008) [44]	12	No	Yes
Kuhlmann and Kühnemund (2001) [45]	16	Yes	Yes
Bose (1998) [46]	7	No	Yes
Aribert et al. (1992) [35]	33	No	f_y only
Total	**91**		

in [30]. More challenging and well-documented tests were performed at the University of Stuttgart [45], where also the axial load was applied to the HEA 240 and HEB 240 sections. In the USA research, the load was applied also from one side only (specimens labeled SC) [43]. Bose [46] tested complete beam-to-column single-sided joints with end-plate connections that failed in the column web in transverse compression. The resistance of this component must be recalculated from the bending moment and the assumed lever arm.

As can be seen, the researchers rarely provide all the data that are necessary to properly analyze the reliability of design methods. The following data are required:

- Exact explanation of loading and boundary conditions
- Nominal geometry
- Nominal material properties
- Measured geometry – Note that the differences may be high, e.g., some specimens tested by Kuhlmann and Kühnemund [45] have the web thickness 8.25 mm instead of the nominal 7.5 mm (10 % difference). This may be critical especially for buckling where the thickness is to the power of three for the determination of moment of inertia.
- Measured material properties – Measured material properties are crucial, because the variation may be high. Ideally, the full stress-strain curve of coupon tests is provided for numerical simulation. Analytical methods and numerical design calculations may require only yield strength.
- Load-deformation curve – The stiffness of the web panel in transverse compression is relatively high and it is also important to subtract the stiffness of the test rig, as was done by De Mitta [44] and Balázs [24, 30]. The full load-deformation curve is preferable, because the stiffness may be derived and the load resistance may be determined in a different way than simply the peak load – e.g. by load at specified accepted deformation, specific stiffness reduction or specific secant stiffness.

Subsequently, mean material and geometrical properties and the coefficient of variation of these properties can be found in FprEN 1993-1-1 – Annex E [12].

The necessary results are the load-deformation curves. The transverse deformation in the direction of the load must be measured and post-processed. The stiffness of webs in transverse compression may be relatively high and the stiffness of the testing rig is relevant. An additional experiment with a solid steel block should be tested in the testing rig and this stiffness subtracted from the load-vertical deformation curves.

Lateral deformations of the web should also be measured at several levels. Every experiment of the web in transverse compression ends with a buckled web (for rolled sections after significant yielding). Buckling may be a perfect bow shape but typically it is unsymmetrical and at least five layers of lateral displacements are needed to sufficiently capture the buckling shape.

Important research subjects are the stresses and strains below the transverse load. Both Eurocode [1] and AISC Specification [18] assume the spreading angle 2.5:1 through the flange and fillet radius; see Fig. 2.2. This is a significant simplification – the spreading angle is probably different through the flange and the fillet radius (or fillet weld of welded sections) and it may change after yielding. Observing stresses and strains is nowadays possible by digital image correlation (DIC), however, no experiment with such testing equipment has been performed yet.

The research into the component column web panel in transverse compression continues. Further 16 specimens of IPE 400 made of S 355 steel grade were recently tested at BUT. The results have not yet been processed and published. Another experimental campaign is in progress at the University of Coimbra with and without axial load in the member.

2.3 Finite element analysis

The use of IDEA StatiCa for dataset generation offers both advantages and limitations. Recently, the developers significantly enhanced the parametric design capabilities and introduced a REST API [47]. With tools such as the Python client, it becomes feasible—and not overly time-consuming—to generate large datasets.

However, several drawbacks need to be acknowledged. IDEA StatiCa models are based on shell elements with a relatively coarse mesh. A particularly critical issue is the neglect of the fillet radius between the flange and the web, which plays an essential role in the load transfer mechanism. Current code formulas assume the same spreading angle through both the flange and the fillet radius, whereas IDEA StatiCa simplifies this geometry. Another important limitation is the absence of imperfection definitions and GMNIA analysis in IDEA StatiCa Connection. In cases where buckling governs the behavior, results must either be post-processed using the general method in EN 1993-1-5, Annex B.1 [32], or excluded from the evaluation. Furthermore, the material model is fixed to linear elastic – perfectly plastic with a nominal plateau slope for numerical stability, $E_{sh} = E/1000$.

Despite these shortcomings in the finite element modeling, IDEA StatiCa was chosen as a practical tool for dataset generation.

For the case of column web panel in transverse compression, Pešek [37] showed that IDEA StatiCa with the mesh size of 12 elements per web height h does not significantly diverge from predictions by ANSYS [23] with solid elements and fine mesh; see Tab. 2.2.

Table 2.2: *Summarization of resistance ratios (MNA, GMNIA) and critical loads (LBA) [37]*

Parameter	Analysis: IDEA StatiCa/ANSYS		
	MNA	GMNIA	LBA
Minimum	0.83	0.83	0.68
Maximum	1.15	1.40	1.55
Average value	0.98	0.98	1.06
Standard deviation	0.08	0.08	0.21

The case of tensile rupture of plates is very simple regarding modeling. The main limitations of IDEA StatiCa design model is the simplicity of the material model and the use of shell elements neglecting the through-thickness contraction. Most notably, the ultimate strength, which governs the experimental resistance, is not a variable taken into account in the model. As a result, the coefficient of variation of experimental to numerical resistances is high when plates with multiple ultimate to yield strength ratios (f_u/f_y) are grouped together.

2.3.1 Mesh sensitivity

The first step after creating the geometry and assigning the boundary conditions and loading is the mesh sensitivity study. That consists of the following steps:

1. Select a coarse mesh and calculate your model. Write down the load resistance, F, and a number of elements, N.

2. Select a slightly finer mesh and calculate your model. Write down the load resistance and the number of elements.

3. Repeat step 2 until you have at least five load resistances. The load resistances should not jump wildly and should decrease with increasing number of elements. The mode of failure should not change.

4. Create a graph with $1/N$ at x axis and F at y axis.

5. Extrapolate the curve created by your analyses until zero point on x axis. Use linear extrapolation with beam elements, quadratic extrapolation with shell elements, and cubic extrapolation with solid elements.

6. Note down the ideal load resistance where the extrapolated curve intersects with zero. This is the resistance of an idealized model with an infinite number of finite elements.

7. Meshes whose load resistance does not differ more than 5 % from the ideal load resistance are qualified. Typically, use the coarsest qualified mesh.

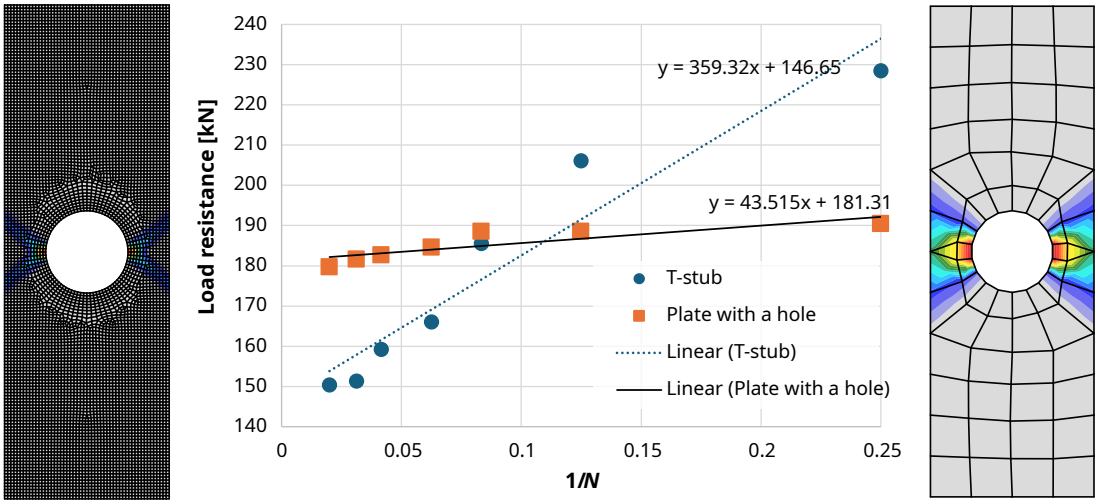


Figure 2.14: *Mesh sensitivity for simple models of plate with a hole and a T-stub failing in mode 1*

The mesh sensitivity study is shown in Fig. 2.14 for two cases: Plate with a large hole and a T-stub with four M16 bolts failing in mode 1. IDEA StatiCa Connection with shell finite elements is used. The number of elements at the plate edge is changed from 4, 8, 12, 16, 24, 32, and up to 50. This number of elements at the plate edge is used in the graph rather than the total number of elements, so the linear trend of mesh sensitivity is expected. The idealized load resistances are 181 kN and 147 kN for the T-stub and plate, respectively. Accepting the 5 % error qualifies meshes with resistances below 190 kN and 154 kN for T-stub and plate, respectively. The load resistance of the T-stub is not very dependent on the mesh and all the tested meshes could be used. On the other hand, the resistance of the plate with a hole decreases rapidly with increasing mesh density. Only meshes with 50 and 32 elements at the plate edge should be used.

This procedure is simple and does not require an extremely fine mesh to determine the "correct" load-resistance. For more details, refer to [48]. On the other hand, note that such a result may be too conservative if there are some simplifications that should be considered. For example, shell models connected node-to-node do not have the added stiffness of thicker plates. When the T-stub flange bends, the plastic hinge appears below the web, which is unrealistic and decreases the T-stub tensile resistance. Furthermore, the area where plastic strain appears must be significant and not infinitesimal. The mesh that provides the load resistance closest to reality is not always the finest.

Mesh sensitivity comes hand-in-hand with failure criteria. The load resistance determined by the plastic strain limit is increasing with decreasing mesh size. This is clearly seen in Fig. 2.15, where meshes that provide the same moment-rotation response show a failure determined by 5% plastic strain at different locations. Note that the bending moment resistance is not very affected, but rotation, ϕ_i, is very different. A practical recommendation is to use the coarsest mesh within

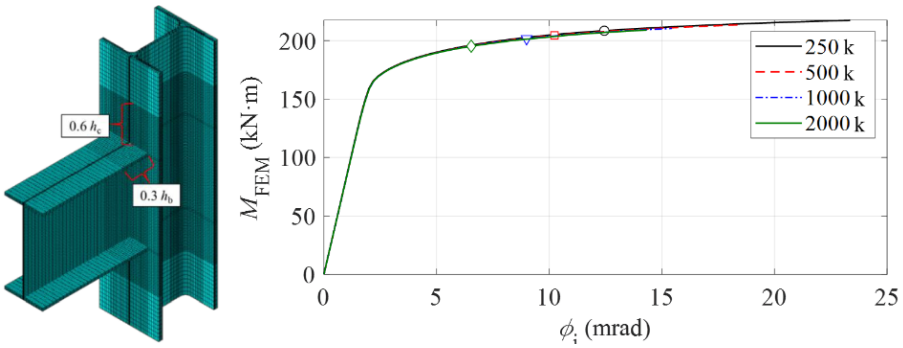


Figure 2.15: *Decreasing load resistance with decreasing mesh size [49]*

the 5% error to the idealized infinite mesh. Another option is to use strain averaging across multiple finite elements as suggested by L. da Silva in [50]. However, the determination of a volume, over which strain averaging should be performed is as complicated as the determination of the plastic strain limit.

2.3.2 Failure criteria

Failure criterion C1 regarding the maximum attained load is tackled by the use of equivalent initial imperfections in geometrically nonlinear analysis with imperfections (GNIA) or also including material nonlinearities (GMNIA) [5]. The reliability of initial imperfection amplitude for example for column buckling curves was worked out by F. Walport [51]. In conclusion, the nominal value or fifth percentile should be used for material and geometrical properties, and with the recommended amplitude of initial imperfections the GMNIA provides reliable load resistance. Cases of shell buckling may be much more complicated and require testing several imperfection shapes including the shapes not originating from the eigenvalue analysis but from the consideration of manufacturing process. However, this is not important in connection design and will not be worked out in this thesis.

A crucial parameter in connection design is the equivalent plastic strain limit – **failure criterion C2**. The current accepted version of FprEN 1993-1-14 [5] in Cl. 8.1.5 states:

If there are no other values given, a recommended value for the maximum equivalent plastic strain of the material may be assumed by $0.33 \cdot \varepsilon_u$ *for uniaxial and* $0.15 \cdot \varepsilon_u$ *for biaxial stress-strain state. The value of* ε_u *for all steels may be taken from Formula (5.5). In case of triaxial stress-strain state or multiaxiality, strain limits should be based on advanced damage theory.*

Currently, the rules across various Eurocode 3 parts are inconsistent and unstructured; see Tab. 2.3. In principle, the strain limits for beam finite elements should be the lowest and may be higher for shell or solid elements. However, this is not the case. Furthermore, it is unclear what rules should be followed when designing e.g. plated girders with some connection elements – there will always be at least welds. Is it the EN 1993-1-5 or EN 1993-1-8 rule? The generic rule in EN 1993-1-14 distinguishes between uniaxial, biaxial, and multiaxial. In practice, considering the Poisson effect, engineers may consider even axial loading to cause a triaxial stress state. As a result, to be safe, $0.15 \cdot \varepsilon_u$ will be used. Furthermore, strain limits for numerical design calculation, as the other two rules, $0.33 \cdot \varepsilon_u$ and $0.15 \cdot \varepsilon_u$, are meant, based on advanced damage theory are meaningless. In nearly every case, the strain limits based on advanced damage theory via numerical simulation will reach much higher values. Therefore, the strain limit for a triaxial stress state will be higher than for a biaxial or even an uniaxial if the clause is followed to the letter. The intention was to warn against edge cases of extreme triaxial stress states that are not possible if traditional detailing rules are followed, but for which numerical tools could be used.

Considering the expectation that engineers will mostly lean towards the safer $0.15 \cdot \varepsilon_u$, the generic rule seems to be too low based on the current practice and experience. Further issue is the unclear distinction between numerical design calculation (directly provides the design load resistance) and numerical simulation (has to be subjected to reliability analysis to find the design load resistance). Yet another problem is the missing link to element type – beam, shell, or solid. It could be argued that beam elements are tackled in Annex C, so these limits apply to shell and solid elements both, but this is not clearly stated. Next, it is unclear, what failure mode the criterion

Table 2.3: *Strain limits according to Eurocode 3 parts (A. Taras, personal communication, Sep. 20, 2025)*

EC3 Part	**Rule**	**S235**	**S355**	**S460**
EN 1993-1-14 (C2, generic)	Max eq. plastic strain = $0.33 \cdot \varepsilon_u$ (uniaxial) / $0.15 \cdot \varepsilon_u$ (bi- or triaxial), with $\varepsilon_u = 0.6 \cdot (1 - f_y/f_u)$.	6.88% / **3.13%**	4.84% / **2.20%**	2.93% / **1.33%**
EN 1993-1-14 – Annex C (beams)	$\varepsilon_{Ed} \leq \varepsilon_{csm}$ (CSM). With $\Omega = 15$, the cap $\approx 15 \cdot \varepsilon_y$	1.68%	2.54%	3.29%
EN 1993-1-5 (plated girders) & EN 1993-1-3 (cold-formed)	Max acceptable plastic strain $\varepsilon_{mpp} = 5\%$	5%	5%	5%
EN 1993-1-6 (shells)	$\varepsilon_{p,eq,Ed} \leq a_{p,eq} \cdot (0.04 - f_{yd}/40\,000)$, with $a_{p,eq} = 2$ and $f_{yd} = f_y$ (N/mm^2).	6.83%	6.23%	5.70%
EN 1993-1-8 (joints)	No explicit % limit; in practice use EN 1993-1-14 (C2)	—	—	—

C2 attempts to capture. Is it to limit the deformation at bolt holes or should it allow for ultimate bearing resistance; see FprEN 1993-1-8 – Tab. 5.9 [28]? Or is it the limit deformation of 3% used for the determination of hollow section joints; see CIDECT Design Guides [52]?

As can be seen, the determination of a reliable plastic strain limit for steel structures is still an unfinished work and an amendment for the modification of this Clause 8.1.5 was accepted in September 2025.

This topic is the primary goal of the project Inter Excellence, LUAUS23114, a collaboration between the Brno University of Technology with the Czech Technical University in Prague and the University of Tennessee, Knoxville. The procedure and investigation is moved into a standalone Chapter 3.

In the following investigation of machine learning, the outcome of this research was used, which is to limit the equivalent plastic strain limit for all situations to:

$$\varepsilon_u = \max\{0.06, 0.6 \cdot (1 - f_y/f_u)\} \tag{2.20}$$

$$\varepsilon_{lim} = 0.25 \cdot \varepsilon_u \tag{2.21}$$

2.4 Data generation

Machine learning requires data, preferably a large amount of data. These can be generated using numerical simulations or numerical design calculations. Numerical simulations will likely have less scatter to the experimental data and will require a safety factor γ_{FE}. The quality of data directly correlates to the quality of machine-learning prediction, because machine learning is able to correlate very closely to the training data; see Fig. 1.1.

It is very tempting to use most-commonly approach to distribute data perfectly regularly using nested cycles. Also, such an approach is better for human understanding and plotting of dependencies. However, such structured sampling introduces patterns that may bias machine learning models or lead to overfitting localized regions of the input space. Machine learning works better with randomly distributed data.

Randomly distributed input samples (e.g., using Latin Hypercube Sampling) provide better coverage of the multidimensional input space. This reduces correlation among features and avoids alignment with the grid structure of the model, ultimately improving generalization [53]. Moreover, random sampling better reflects the natural variability and uncertainty present in real-world data. A comparison between regularly and randomly distributed datasets is shown in Fig. 2.16.

There may be exceptions to the general approach of using independent, continuous variables in machine learning models. For instance, when investigating the resistance of rolled steel cross-

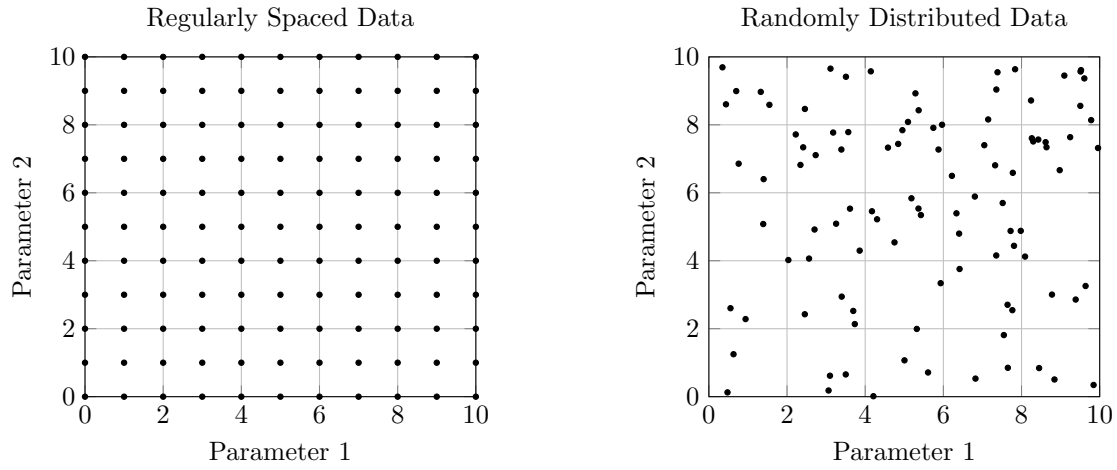


Figure 2.16: *Comparison of regularly spaced and randomly distributed data*

sections, the geometrical dimensions – such as flange thickness, web thickness, overall depth, and width – are typically standardized and manufactured in discrete steps according to product norms (e.g., EN 10365:2017 [54]). These dimensions are interdependent; for example, it is unlikely to encounter a flange thickness of 10 mm paired with a web thickness of 20 mm. Therefore, it is often more appropriate to directly use the real cross-sectional properties of rolled profiles. In this context, the profile designation (e.g., IPE 100) can be treated as a categorical input (random variable), while its geometrical properties serve as numerical features for model training.

A similar consideration applies to steel grades. Their nominal mechanical properties – such as yield strength and ultimate tensile strength – are predefined according to standards (e.g., EN 10025-2:2019 [55]), and the ratio of ultimate to yield strength is partially dependent on the yield strength itself.

The range of validity of the model is a critical factor. In general, expanding the range of input variables increases the amount of data required for accurate training. Moreover, the author has observed that enlarging the validity domain can reduce prediction accuracy, even when the training dataset is proportionally expanded. This phenomenon may be attributed to the increased complexity and variability introduced by the broader input space, which challenges the model's generalization capability.

In this context, it is crucial to accept that the neural network will not be able to predict every case, and to make a reasonable compromise based on investigation of current design practice. In the field of connection design, the largest dataset is the Connection Library of IDEA StatiCa [56]. After proper organization, the dataset could be investigated and the range of validity covering a desirable percentage of cases, e.g. 80 %, could be selected.

It is also important to note that design customs vary across the regions. For example, in the USA, the commonly-used wide flange sections have a thicker web than the European counterparts. Also, the common design utilizes rigid base plates and end plates that do not allow prying actions, i.e., the base plates and end plates are typically thicker.

2.5 Neural network

This section outlines the methodology adopted for developing a data-driven prediction model using supervised machine learning. The approach leverages artificial neural networks (ANNs), implemented in PyTorch [57], and encompasses standard procedures such as feature scaling, data splitting, network training, and performance evaluation.

Supervised learning is a paradigm where a model learns a mapping from input features $\mathbf{X}$ to a target variable $\mathbf{y}$ using labeled data [58]. In this case, the objective is to predict the characteristic resistance of components of structural joints based on multiple input parameters derived from numerical simulations or experiments.

ANNs have their origins in the 1940s with the early models of McCulloch and Pitts [59], further developed by Hebb [60] and later by Minsky and Papert [61]. Initially conceived as simplified representations of biological neurons, ANNs have evolved into powerful tools for solving

complex, non-linear problems. In recent years, their application in architecture, engineering, and construction has grown significantly, driven by improved hardware (notably GPUs), algorithmic advancements, and accessible software frameworks.

A single artificial neuron combines input features x_n with trainable weights W_n, adds a bias b_{NN}, and applies a non-linear activation function:

$$y(x) = a \cdot \left(\sum W_n \cdot x_n + b_{NN} \right) \tag{2.22}$$

Networks consist of multiple interconnected neurons organized in layers, allowing the model to learn hierarchical feature representations. During training, weights and biases are adjusted to minimize prediction error, quantified by a loss function such as mean squared error (MSE) or mean absolute error (MAE). This optimization typically relies on gradient-based algorithms such as backpropagation and stochastic gradient descent.

Data preprocessing is critical for effective training. Normalization or standardization ensures that all features contribute equally by removing differences in scale and distribution; see Tab. 2.4. The outputs for training should be also scaled, so that large values do not cause mathematical issues during the training process.

The output of a neural network may subsequently be passed through another activation function (e.g., sigmoid for binary classification or softmax for multiclass classification), depending on the task. In this study, no activation function was used for the output.

2.5.1 Neural network architecture

In this work, deep neural networks (DNNs) are trained to predict the non-linear resistance of component column web panel in transverse compression and tensile rupture of weakened plates. The model architecture and hyperparameters follow prior studies by Andreas Müller [62], [63], [64] and following investigations in the scope of the project Machine Learning-based Design Optimization of Steel Connections (MADESCO) in a cooperation of IDEA StatiCa, ETH Zürich and CTU Prague. In the following, the effects of the input and output scalers, activation functions, optimizers, learning rate, and the width and depth of the neural network are investigated. As a dataset, a knee welded joint loaded by a combination of bending moment and shear force was chosen. This dataset contains over 800 thousand analyzed models, but for this exercise, only 5 % of the available data was used. The graphs were done primarily by Tamer Çakır at ETH. The DNN architecture, in terms of the number of neurons and layers, was varied in each of the following investigations. The differences between wider DNN (with more neurons) and leaner DNN (fewer neurons in one layer) are not great. The wider DNN typically achieves better error in the form of R^2 than the leaner DNN. On the other hand, it should be noted that the more neurons, the longer the training takes.

Overall, the variation in parameters (scalers, optimizers, learning rate, activation functions) shows that the correct selection of DNN architecture is very important. Some parameters perform much worse than others. These comparisons indicate that the DNN architecture must be tested for a particular problem, and various options should be considered.

2.5.2 Input and output scaler

Prior to model training, all input features may be subjected to dimensional scaling to cross-section depth of e.g. 200 mm using similitude theory [65], [66]. For example, IPE 300 with cross-section depth of 300 mm is scaled by $L = 200/300 = 2/3$. Then, features are scaled as follows:

- Lengths $\sim L^1$
- Areas and forces $\sim L^2$
- Volumes, masses, and bending moments $\sim L^3$
- Constants, stresses, and strains $\sim L^0$ (invariant under scaling)
- Stiffness (force/displacement) $\sim L^1$

This is called a Similitude scaling, and it utilizes the physical similarity of the model and the same stress distribution as in the scaled model. This allows drastically reducing the number of cases for training and potentially increasing the range of validity. Training data are generated by

physical testing, numerical simulations or numerical design calculations, which are the most time- and resource-intensive. When using physical experiments, some experimental variations may get lost and it can be beneficial to add scaled data to the original dataset. When using numerical simulations or numerical design calculations, similitude theory should be perfectly valid and the original dataset may be replaced by the scaled dataset.

Next, all input features should be scaled according to one of the following methods. This preprocessing step ensures that all features contribute equally to the learning process, avoids dominance by features with larger magnitude, and enhances convergence during optimization [58].

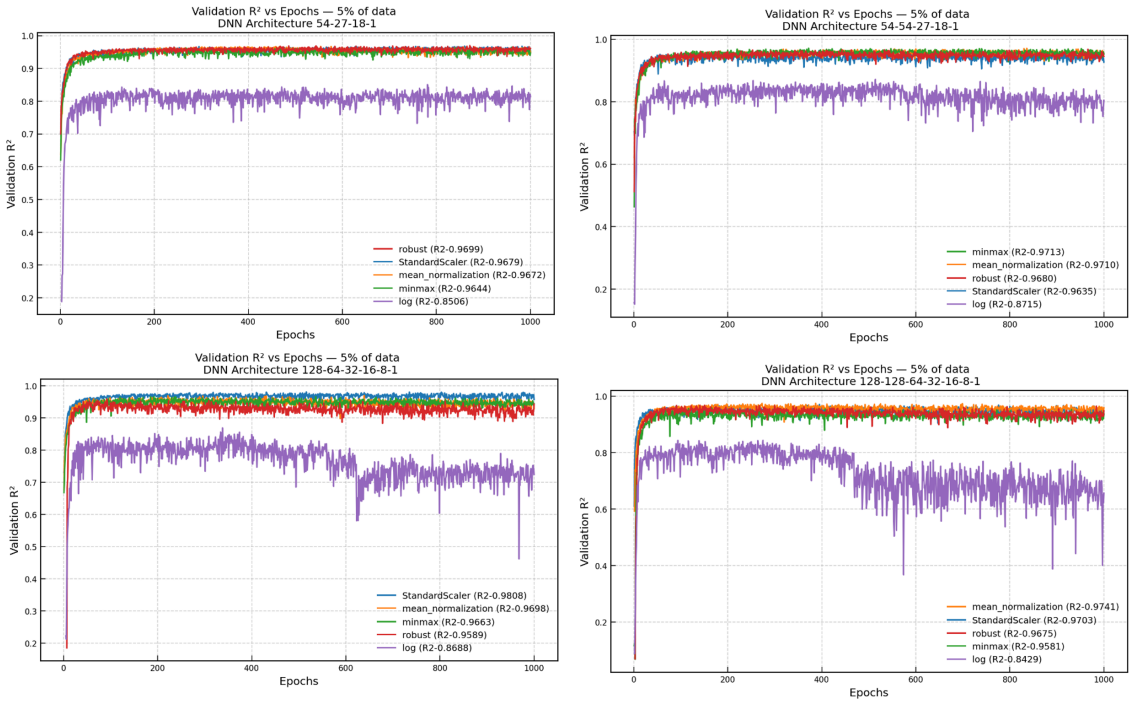


Figure 2.17: *Comparison of input and output scalers for different widths and depths of neural network*

Table 2.4: *Overview of common data scaling and transformation methods*

Method	**Formula**
Standardization	$x'_i = \dfrac{x_i - \mu}{\sigma}, \quad \mu = \dfrac{1}{n}\sum_{j=1}^{n} x_j, \quad \sigma = \sqrt{\dfrac{1}{n}\sum_{j=1}^{n}(x_j - \mu)^2}$
Mean Normalization	$x'_i = \dfrac{x_i - \mu}{x_{\max} - x_{\min}}, \quad \mu = \dfrac{1}{n}\sum_{j=1}^{n} x_j$
Min–Max Scaling	$x'_i = \dfrac{x_i - x_{\min}}{x_{\max} - x_{\min}}$
Robust Scaling	$x'_i = \dfrac{x_i - \text{median}(x)}{Q_3 - Q_1}, \quad Q_1 = 25^{\text{th}}$ percentile, $Q_3 = 75^{\text{th}}$ percentile
Logarithmic Transform	$x'_i = \log(x_i + \varepsilon), \quad \varepsilon > 0$

Fig. 2.17 shows a comparison of the different scalers; see Tab. 2.4. In this case, the outputs were normalized to the bending and shear resistance of the connected beam. Note that the results may vary significantly when the outputs, such as load resistance, are not normalized. The performance of all four tested DNNs is very poor with the log scaler, where both inputs and outputs are transformed into a logarithmic scale. The performance of the other scalers is relatively similar, with the best choice varying depending on the DNN width and depth.

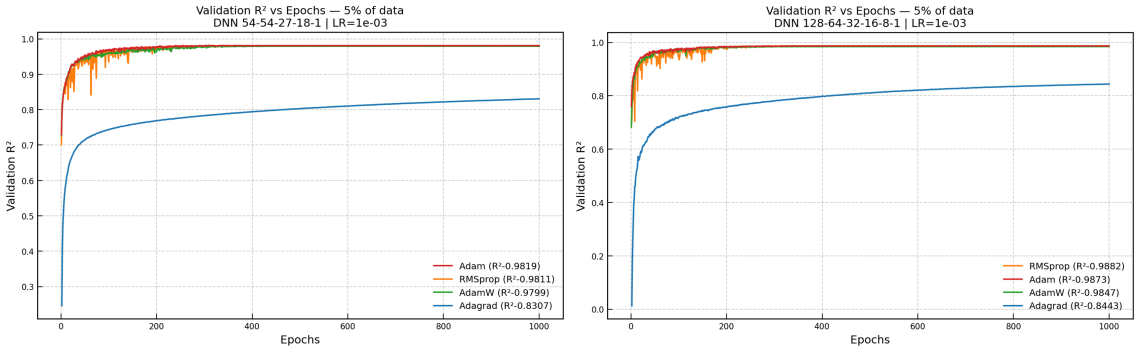


Figure 2.18: *Comparison of different optimizers*

2.5.3 Optimizer

Fig. 2.18 shows a preliminary comparison of the optimizers. A thorough evaluation would require tuning additional hyperparameters for each selected method. Adagrad clearly performs worst in this comparison, while RMSProp achieves the highest R^2. However, RMSProp exhibits significant fluctuations during the early stages of training, whereas Adam and AdamW remain much smoother. The overview of optimizers used historically and in current machine-learning tasks is in Tab. 2.5.

Table 2.5: *Overview of selected gradient-based optimizers*[a]

Optimizer	**Update rule / principle**	**Remarks**
SGD	$w_t = w_{t-1} - \eta \nabla L_t$	Simple stochastic gradient descent; sensitive to learning rate; may oscillate in narrow valleys.
SGD + Momentum	$v_t = \beta v_{t-1} + \nabla L_t$ $w_t = w_{t-1} - \eta v_t$	Accelerates convergence in relevant directions; dampens oscillations; introduces momentum hyperparameter β.
Adagrad [67]	$G_{t,i} = G_{t-1,i} + (\nabla_i L_t)^2$ $w_{t,i} = w_{t-1,i} - \frac{\eta}{\sqrt{G_{t,i}} + \epsilon} \nabla_i L_t$	Effective for sparse data; learning rates decrease monotonically, may become too small.
RMSProp [68]	$v_{t,i} = \beta v_{t-1,i} + (1-\beta)(\nabla_i L_t)^2$ $w_{t,i} = w_{t-1,i} - \frac{\eta}{\sqrt{v_{t,i}} + \epsilon} \nabla_i L_t$	Mitigates Adagrad's diminishing learning rate; works well in non-stationary settings; requires decay tuning.
Adam [69]	$m_t = \beta_1 m_{t-1} + (1-\beta_1)\nabla L_t$ $v_t = \beta_2 v_{t-1} + (1-\beta_2)(\nabla L_t)^2$ $\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t}$ $w_t = w_{t-1} - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t$	Combines momentum and adaptive scaling; fast and robust; may generalize worse than SGD with momentum.
AdamW [70]	Same as Adam, but with decoupled weight decay (regularization applied separately from the gradient step).	Improves generalization; widely used in large-scale models such as Transformers.

[a] Notation: w_t – parameter vector at iteration t; ∇L_t – gradient of the loss at iteration t; $\nabla_i L_t$ – gradient w.r.t. parameter w_i; η – learning rate; ϵ – small constant to prevent division by zero; $G_{t,i}$ – accumulated squared gradients (Adagrad); $v_{t,i}$ – exponential moving average of squared gradients (RMSProp, Adam); m_t – exponential moving average of gradients (Adam); $\hat{m}_t, \hat{v}_t$ – bias-corrected estimates (Adam); β, β_1, β_2 – decay rates for moving averages; v_t – momentum term in SGD + momentum.

2.5.4 Learning rate

Fig. 2.19 shows the DNN performance for varying learning rates. Very high learning rates may overshoot the global minima and fail to converge to a good solution, while very low learning

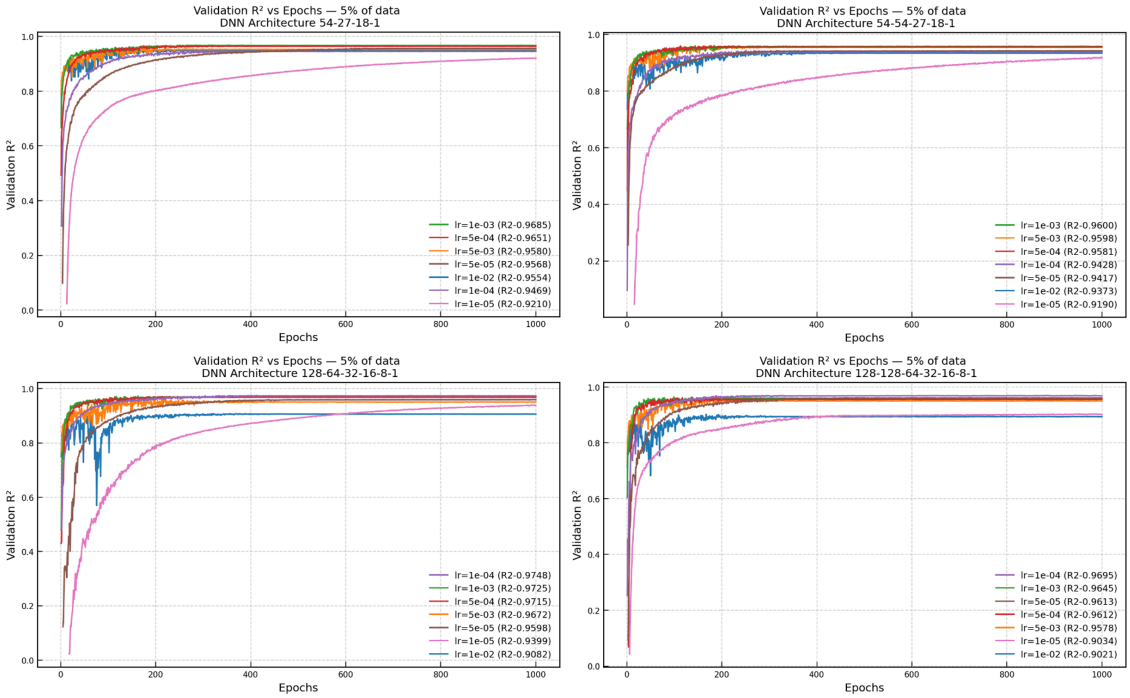


Figure 2.19: *Comparison of learning rates*

rates converge too slowly. The results indicate that particularly large learning rates should be avoided. Interestingly, Adam should dynamically adjust the learning rate based on estimates of first and second moments of the gradients. It slows down the learning rate when nearing a minimum, which helps not overshooting the minimum and reducing bouncing around the minimum. However, an additional (manual) adaptive learning rate that decreases as the loss function becomes smaller provides the best results. PyTorch provides a built-in scheduler for this purpose, called `ReduceLROnPlateau`.

2.5.5 Activation function

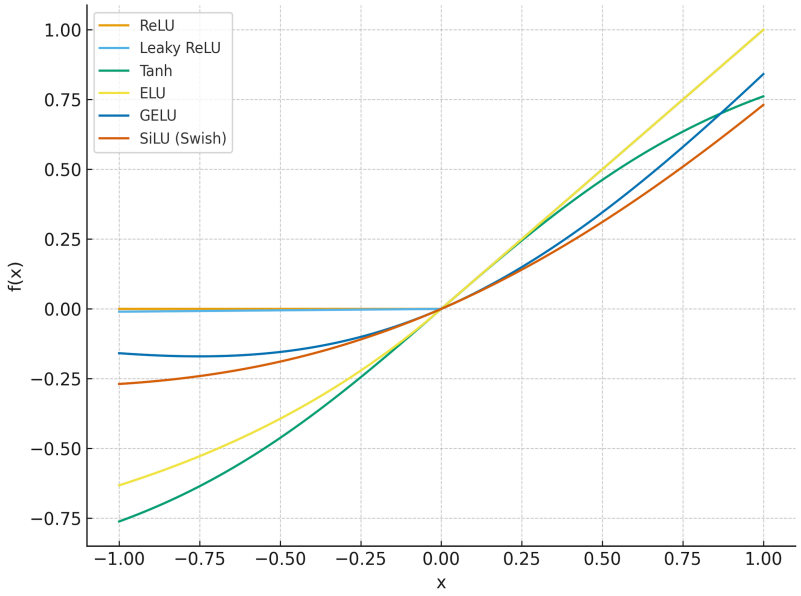


Figure 2.20: *Activation functions in the range* $[-1, 1]$

Fig. 2.21 compares the model performance across different activation functions. Among these,

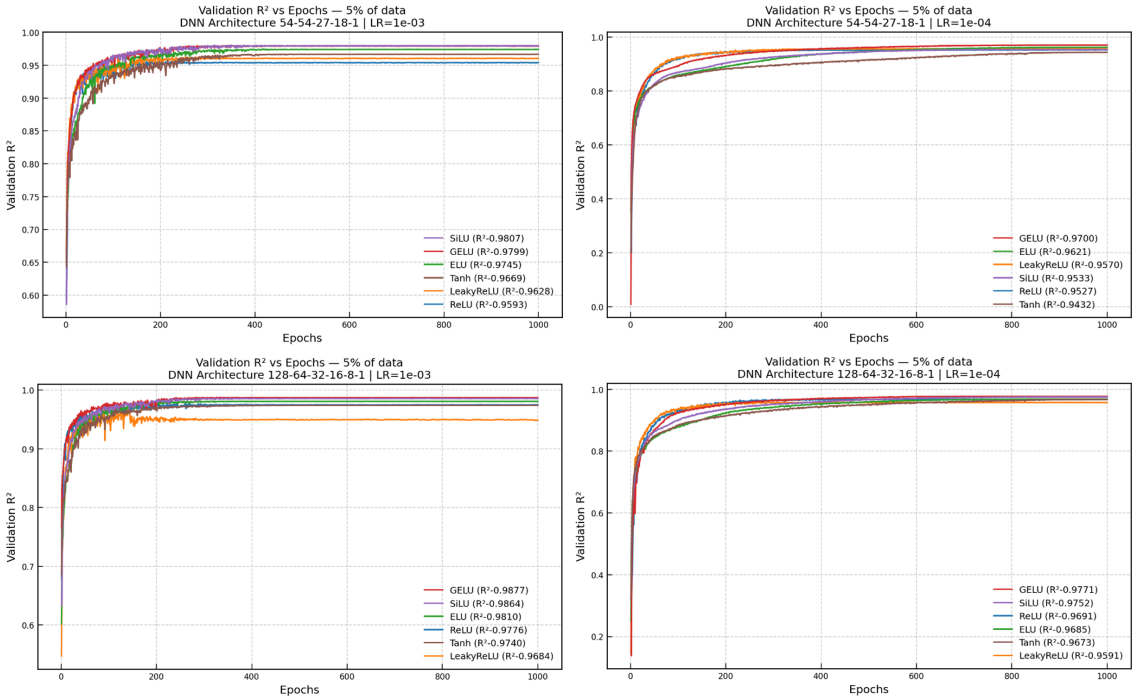


Figure 2.21: *Comparisons of activation functions*

the ReLU function is currently the most widely used due to its simplicity and computational efficiency. An overview of commonly implemented activation functions available in frameworks such as PyTorch is provided in Tab. 2.6 while Fig. 2.20 illustrates their behavior over the input domain $[-1, 1]$. Interestingly, different activation functions converge to a stable error R^2.The GELU function achieves the highest performance with a final error of $R^2 \approx 0.99$ whereas some functions plateau below $R^2 \approx 0.96$. This difference is substantial in terms of prediction quality. A more comprehensive investigation would require evaluating combinations of activation functions across hidden layers and output layers to fully understand their impact.

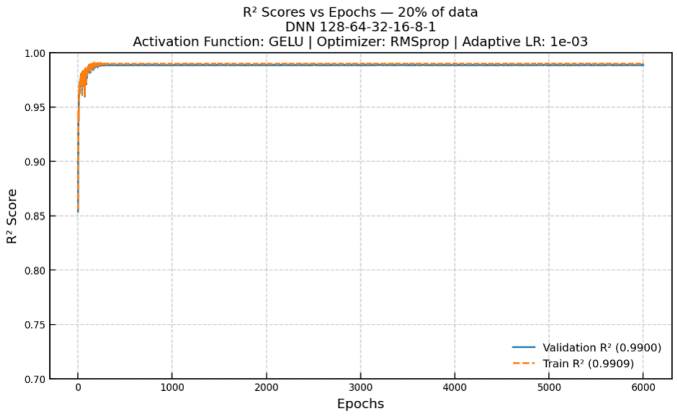


Figure 2.22: *Optimized model*

The optimized model shown in Fig. 2.22 is able to quickly achieve very high prediction quality with a final error $R^2 \approx 0.99$. In this optimized model, RMSProp optimizer, GELU activation function, adaptive learning rate using `ReduceLROnPlateau` in-built scheduler and a wide neural network with the initial width of 128 neurons progressively decreasing across 5 hidden layers to one output. The optimization considers only the resulting error R^2 without any regard for the speed of learning process. That is because the calculation of dataset takes days or even months while the learning process is typically in minutes.

Table 2.6: *Overview of Activation Functions*

Activation Function	**Formula**	**Range**	**Key Characteristics**
ReLU [71]	$f(x) = \max(0, x)$	$[0, \infty)$	Simple, computationally efficient; can suffer from the "dying ReLU" problem.
Leaky ReLU	$f(x) = \max(\alpha x, x)$	$(-\infty, \infty)$	Allows a small, non-zero gradient when the unit is inactive and not saturating.
Tanh	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$(-1, 1)$	Smooth, zero-centered output; suffers from vanishing gradients for large inputs.
ELU	$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha(e^x - 1) & \text{if } x \leq 0 \end{cases}$	$(-\alpha, \infty)$	Smooth and differentiable; aims to push mean activations closer to zero, mitigating vanishing gradients.
GELU [72]	$f(x) = x \cdot \Phi(x)$	$(-\infty, \infty)$	Smooth, non-monotonic; approximates identity for large positive inputs and zero for large negative inputs.
SiLU (Swish)	$f(x) = x \cdot \sigma(x)$	$(-\infty, \infty)$	Self-gated activation function; combines properties of sigmoid and linear functions.

[a] Notation: α – scaling factor for negative inputs (Leaky ReLU, ELU); $\Phi(x)$ – cumulative distribution function of the standard normal distribution; $\sigma(x)$ – sigmoid function.

2.5.6 Loss function

Minimizing the loss function encourages the model to make predictions close to the actual target values. The most commonly used loss function for regression problems is the Mean Squared Error (MSE):

$$L_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{2.23}$$

The disadvantage is that MSE penalizes large errors more severely. The prediction tends to lean towards the outputs with high magnitudes. This skews the results when the dataset contains members or connecting elements of different sizes. For example, the results with IPE 600 cross-section are more relevant than the results with IPE 100 cross-section.

There are several approaches to mitigate this issue. The outputs, e.g., component load resistance, can be transformed into a logarithmic scale. Furthermore, different loss functions may be used:

- Mean Absolute Error (MEA)

$$L_{\text{MEA}} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{2.24}$$

- Huber loss [73]

$$L_\delta(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{if } |y - \hat{y}| \leq \delta \\ \delta\left(|y - \hat{y}| - \frac{1}{2}\delta\right) & \text{otherwise} \end{cases} \tag{2.25}$$

- Relative mean squared error (RMSE)

$$L_{\text{RMSE}} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{(y_i - \hat{y}_i)}{y_i} \right)^2 \tag{2.26}$$

- Log-Cosh

$$L_{\text{Log-Cosh}} = \frac{1}{N}\sum_{i=1}^{N} \log\left(\cosh\left(y_i - \hat{y}_i\right)\right) \tag{2.27}$$

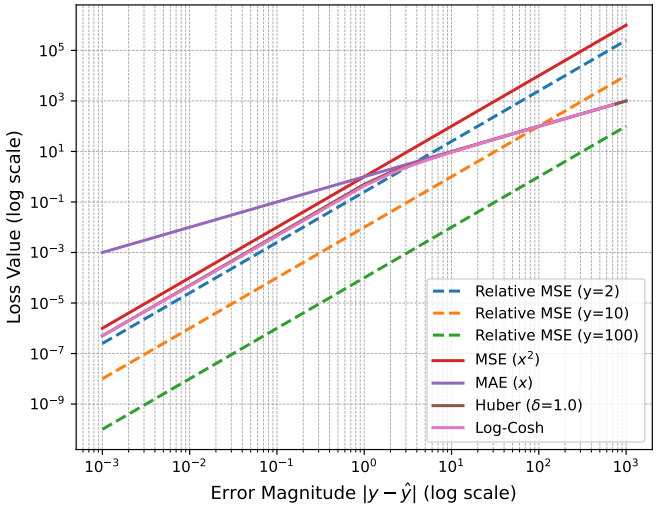


Figure 2.23: *The overview of different loss functions and their output with increasing error*

The comparison of loss values in relation to error magnitude is shown in Fig. 2.23. For effective optimization, larger errors must be penalized more heavily. MSE increases quadratically with error, i.e., in log-log scale is with the steady slope. MAE increases only linearly, i.e., the slope in log-log scale is smaller than for MSE. Huber loss function provides a transition from MSE to MAE at error δ, blending quadratic and linear behavior around this point. The Log-Cosh loss function behaves similarly to the Huber loss with $\delta = 1$, producing nearly the same curve with only slight differences near the transition region.

Note that using RMSE loss function prevents use of some types of scalers (Mean normalization, Standardization, and Robust scaling) for output. The output values y_i and $\hat{y}_i$ cannot be equal to zero or close to zero. This would produce a mathematical error.

2.5.7 Final model for web in compression

In the investigation of column web in transverse compression, these parameters of neural network as a feedforward architecture were adopted:

- 8 input features that include geometric and material properties, 128 neurons, 64 neurons, 32 neurons, 16 neurons and 1 output layer – plastic resistance $F_{pl,Rd}$ or $\alpha_c r$. The model was implemented using `torch.nn.Sequential`.
- Similirity scaler is applied – all inputs and outputs are scaled to the reference beam depth $h = 200$ mm.
- `StandardScaler` from *scikit-learn* [74] for the inputs.
- Custom scaler for the output so that the values are between 0 and 10.

$$y_{scaled} = \frac{y}{y_{max}} \cdot 10 \tag{2.28}$$

- Adam optimizer, because it is more advanced and does not suffer from the initial fluctuations. Its performance scored similarly as the first, RMSProp.
- Learning rate with and initial value of 0.01 and the use of scheduler with the factor of 0.9 (learning rate is reduced only by 10 % loss function is not decreasing for 20 epochs) and minimum learning rate of 0.0001.

- To avoid unnecessary training cycles and prevent overfitting, early stopping was applied. Training was halted if the coefficient of variation (CoV) on the validation set failed to improve over 2 000 consecutive epochs. A minimum of 3 000 epochs was enforced to allow sufficient training. A maximum of 10 000 epochs are used.
- GELU activation functions in hidden layers and no activation function at the output layer.
- Loss function implemented as Relative Mean Squared Error. The purpose of this loss function is for each data point, no matter the size of member, to add the same weight into the training process.

To evaluate the generalization capability of the model, the dataset is randomly split into training and testing subsets. The training set is used to fit the model, while the testing set provides an unbiased evaluation of prediction accuracy. An 80:20 split ratio was employed using `train_test_split` from `scikit-learn` [74]. The split is utilized to:

- Prevent overfitting to the training data,
- Allow estimation of model performance on unseen data,
- Simulate real-world deployment conditions.

The following metrics were used to assess model performance:

- Bias: Measures systematic deviation of predictions from true values; see Eq. (2.32).
- Coefficient of Variation (CoV): Measures relative variability in predictions. CoV is defined as:

$$\text{CoV} = \frac{\text{std}(y/\hat{y})}{\text{mean}(y/\hat{y})} \tag{2.29}$$

- Relative Error: Used to identify outliers with poor prediction accuracy.

Predicted vs. true values were visualized using scatter plots, including $\pm 10\%$ error bounds. Matplotlib [75] was used for plotting. Any outliers with relative error $> 10\%$) were extracted and analyzed for patterns or data quality issues.

The final trained model was serialized and saved as a `.pth` file using `torch.save`, enabling future inference without retraining. Note that only weights and biases of the model are not enough. Input and output scaling parameters and DNN options must also be preserved to maintain prediction consistency.

2.6 Reliability

Eurocode EN 1990 [6] allows two types of reliability verification:

- Semi-probabilistic approach using partial safety factors
- Reliability-based design and assessment

2.6.1 Semi-probabilistic approach

This chapter provides step-by-step guidance for the reliability assessment according to Eurocode standards and is mostly built upon SAFEBRICTILE deliverable D1.1 [15], which was not published.

First, it is necessary to split the reliability of actions and resistances, which greatly simplifies the task. Eurocode assumes factors for resistance $\alpha_R = 0.8$ and for loads $\alpha_E = 0.7$. The probability that the actual resistance R is smaller than the design resistance R_d is then:

$$P\left(R \leq R_d\right) = \Phi\left(-\alpha_R \cdot \beta\right) = \Phi\left(-0.8 \cdot 3.8\right) = \Phi\left(-3.04\right) \tag{2.30}$$

In other words, 1183 samples should fail out of a million. Note that the target reliability is the same for all types of failure modes; member yielding or bolt or weld fracture, with possibly higher consequences, keep the same reliability target. Also, all reliability classes (RC 1, RC 2,

RC 3) are defined by the same resistance reliability target $\alpha_R \cdot \beta = 3.04$. The reliability differentiation is achieved by either through increases of the load-sided partial factors or through increased supervision and fabrication inspection levels [15].

Then a set of experiments should be specified, and experimental resistance $r_{e,i}$ determined. This may seem like an easy task, but the load resistance is often an ambiguous term that needs to be clarified. For the example of tensile rupture, the load resistance is assumed to be at maximum load.

Then the numerical design model is created for each experiment and the numerical resistances $r_{t,i}$ are obtained. The geometrical and material properties are set as measured in the experiment. The numerical model should be subjected to a validation and verification process as described in FprEN 1993-1-14:2024 – Chapter 7 [5].

The next step is to determine the basic variables that affect the results. These variables are, for example, the plate thickness or the material yield strength. The mean values and coefficient of variation for the basis for calibration of partial factors are now codified in FprEN 1993-1-1:2017 – Table E.1 [6]. Interestingly, the variability in dimensional properties does not match the manufacturing tolerances, e.g. in EN 10029:2010 [7], but it outlines the assumptions, upon which the reliability of Eurocodes is built. Next, the mean value of bias b should be calculated:

$$b = \sum_{i=1}^{n} r_{e,i} r_{t,i} / \sum_{i=1}^{n} r_{t,i}^2 \tag{2.31}$$

Where n is the number of specimens. There is a problem with this bias when dataset contains varying sizes of members or connection elements, e.g., bolts or welds. This bias is affected disproportionately more by large members and elements. A more appropriate formula is the simple mean:

$$b = \frac{1}{n} \sum_{i=1}^{n} \frac{r_{e,i}}{r_{t,i}} \tag{2.32}$$

FprEN 1993-1-14 – Annex A [5] states that b should be in the range of $0.8 < b < 1.25$. SAFEBRICTILE [14] requires even more restrictive range of $0.85 < b < 1.15$. However, for numerical design calculation models with bilinear material model and small plastic strain limit, it is expected to reach high values of b, e.g. 1.3. Such higher values should be expected when neglecting strain-hardening and should not be the reason to terminate the reliability assessment due to inaccuracy.

Then the coefficient of variation of the error V_δ should be calculated by the following equations:

$$\delta_i = \frac{r_{e,i}}{b \cdot r_{t,i}} \tag{2.33}$$

$$\Delta_i = \ln(\delta_i) \tag{2.34}$$

$$s_\Delta^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(\Delta_i - \bar{\Delta}\right)^2 \tag{2.35}$$

$$V_\delta = \sqrt{\exp(s_\Delta^2) - 1} \tag{2.36}$$

Where $\bar{\Delta}$ is the average of Δ_i. The next step is to calculate the coefficient of variation V_{rt} for the basic input variables. For that, the geometrical and material properties should be changed in these numerical models to mean values according to Table E.1 [12]. The load resistances of these numerical models are $r_{t,i}(\underline{X_m})$. Furthermore, the numerical models with slightly changed mean values should be calculated. For example, if the basic variables are plate thickness and yield strength, a set of numerical models with changed plate thickness (load resistances $r_{t,i}(\Delta t)$) and another set of numerical models with changed yield strength should be calculated (load resistances $r_{t,i}(\Delta f_y)$). The change should be small but meaningful, and it should not completely change the failure mode. The authors suggest using values close to the standard deviation of the basic variable σ.

$$V^2_{r,t,i,\Delta t} = \left[\frac{\left(r_{t,i}\left(\underline{X_m}\right) - r_{t,i}\left(\Delta t\right)\right)}{\Delta t} \cdot \sigma\left(t\right) \right]^2 \tag{2.37}$$

$$V^2_{r,t,i,\Delta f_y} = \left[\frac{\left(r_{t,i}\left(\underline{X_m}\right) - r_{t,i}\left(\Delta f_y\right)\right)}{\Delta f_y} \cdot \sigma\left(f_y\right) \right]^2 \tag{2.38}$$

$$V^2_{rt,i} = \left(V^2_{rt,i,\Delta t} + V^2_{rt,i,\Delta f_y}\right) / r^2_{t,i}(\underline{X_m}) \tag{2.39}$$

$$V_{rt} = \sum_{i=1}^{n} \sqrt{V^2_{rt,i}} \tag{2.40}$$

It is probable that many numerical models will have more basic variables. In that case, the formulas should be expanded accordingly. The values of V_{rt} depend mainly on the relationship of the result to the basic variables. If the relationship is linear, such as for tensile rupture, the expected values of V_{rt} are about 5%. For phenomena, where relationship is quadratic or even on a power of three, such as plate bending or buckling, the values of V_{rt} will be higher. It is difficult to establish V_{rt} specifically for every experimental set, and V_{rt} may be set based on previous experience; e.g. first generation of Eurocode was built mainly with $V_{rt} = 7\%$.

Finally, the two errors, V_δ and V_{rt} are added together to the coefficient of variation:

$$V_r = \sqrt{V^2_\delta + V^2_{rt}} \tag{2.41}$$

Now variation coefficients Q can be calculated either for individual $V_{rt,i}$ for each specimen, or global V_{rt}:

$$Q_{rt,i} = \sqrt{\ln\left(V^2_{rt,i} + 1\right)}, \text{ or for global } V_{rt}: \quad Q_{rt} = \sqrt{\ln\left(V^2_{rt} + 1\right)} \tag{2.42}$$

$$Q_\delta = \sqrt{\ln\left(V^2_\delta + 1\right)} \tag{2.43}$$

$$Q_i = \sqrt{\ln\left(V^2_\delta + V^2_{rt,i} + 1\right)}, \text{ or for global } V_{rt}: \quad Q = \sqrt{\ln(V_r + 1)} \tag{2.44}$$

The design resistance is calculated according to the number of specimens n (for global V_{rt}, replace $Q_{rt,i}$ for Q_{rt} and Q_i for Q):

$$r_{d,i} = \begin{cases} b \cdot r_{t,i}\left(\underline{X_m}\right) \exp\left(-k_{d,\infty} \cdot \left(\dfrac{Q^2_{rt,i}}{Q_i}\right) - k_{d,n} \cdot \left(\dfrac{Q^2_\delta}{Q_i}\right) - 0.5 \cdot Q^2_i\right), & \text{for } n \le 100 \\ b \cdot r_{t,i}\left(\underline{X_m}\right) \exp\left(-k_{d,\infty} \cdot Q_i - 0.5 \cdot Q^2_i\right), & \text{for } n > 100 \end{cases} \tag{2.45}$$

Where $k_{d,n}$ and $k_{d,\infty}$ are from EN 1990 – Table D2 [6], and FprEN 1993-1-14 – A3(4) [5] suggests that V_x unknown should be used. Often, the number of specimens is higher than 30 and interpolation between 30 and infinity should be made. Authors suggest treating 100 specimens as infinity, i.e., $k_{d,100} = k_{d,\infty} = 3.04$ and interpolate between 30 and 100.

Finally, the partial safety factor can be calculated for each model and then the average:

$$\gamma_{M,i} = r_{nom,i} / r_{d,i} \tag{2.46}$$

$$\gamma_M = \frac{1}{n} \sum_{i=1}^{n} \gamma_{M,i} \tag{2.47}$$

This partial safety factor should be below an acceptable limit, otherwise, it should be applied to each numerical design calculation using these model settings.

SAFEBRICTILE [15] recommends acceptable levels of γ_M according to current Eurocode practice based on coefficient of variation V_r; see Table 2.7. These acceptance limits are within the target reliability according to Equation (1). In case γ_M is below acceptance limit, it does not need to be applied.

Furthermore, coefficient of variation V_δ may be reduced by two options:

Table 2.7: *Recommended acceptance limits of γ_M [15]*

Range of V_r	**Acceptance limit for γ_M**
$0.00 < V_\mathrm{r} < 0.04$	1.03
$0.04 \leq V_\mathrm{r} < 0.20$	$1.03 + 0.75 \cdot (V_\mathrm{r} - 0.04)$
$V_\mathrm{r} \geq 0.20$	1.15

- Sample subdivision: The set of specimens may be divided into groups, for which one model or model settings is more appropriate than the other. For example, some phenomena may be well captured only by GMNIA. Another appropriate subdivision may be into mild steel and high-strength steel. Note that with decreasing number of specimens in the set, $k_{d,n}$ increases.

- Tail approximation. This is useful for cases where many results of numerical model are very safe compared to experiments, which leads to high coefficient of variation V_δ. Such very safe cases may be disregarded in the reliability assessment, which leads to decrease in b value but also decrease in V_δ. In practice, the samples may be disregarded from the safest (smallest $r_{t,i}/r_{e,i}$) until partial safety factor γ_M decreases. This takes into account the fact that the distribution is not normal, but rather log-normal and the tail-end may be ignored. The value of $k_{d,n}$ remains unchanged.

This procedure can be used for reliability assessment of analytical formulas, such as those included in Eurocodes, numerical design calculations, numerical simulations, or machine-learning predictions.

2.6.2 Direct reliability approach

A direct reliability approach based on crude Monte Carlo simulation requires a huge number of simulations. For structures classified as RC2, with a target reliability index of $\beta = 3.8$, this corresponds to approximately one million simulations [76].

$$N \approx \frac{1}{V_\mathrm{r}^2 \cdot P_\mathrm{f}} \tag{2.48}$$

Where V_r is coefficient of variation and P_f is probability of failure.

In practice, analytical equations may be used for this number of simulations, but finite element analyses are not feasible. A practical alternative is the use of machine-learned predictions. It should be noted, however, that these predictions carry their own errors. Provided there is no systematic error in the range of predictions and the means values approximately coincide, this should not pose a significant issue, and a smaller, more conservative estimate of load resistance can be achieved; see 2.24.

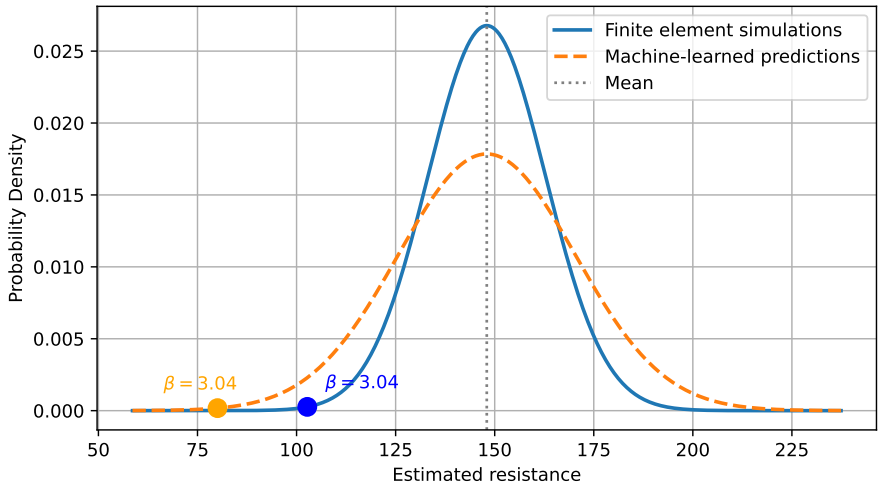


Figure 2.24: *Prediction by FEA and ML with the same mean*

However, if the means do not coincide and there is a systematic error in the ML predictions, the result may be dangerous despite the variation coefficient, V_r, being larger. Therefore, it is important to create at least a single finite element simulation with mean values to confirm the means do coincide with an acceptable error. Typical accepted error is 5 %.

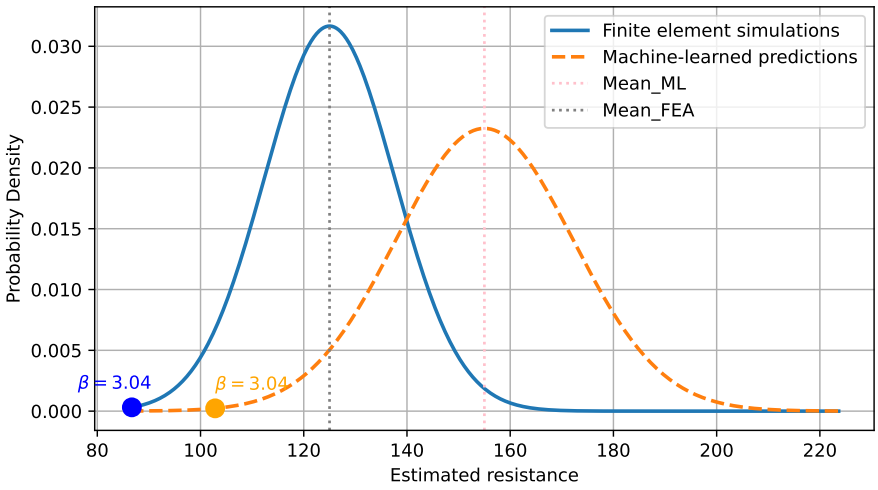


Figure 2.25: *Prediction by FEA and ML with different means*

Note that the outliers in machine learned predictions are common, in probabilistic simulations, millions of predictions with slight changes in inputs are carried out. The chance of outliers consistently pushing the mean is therefore much lower than an occurrence of a single outlier.

An interesting paper opening the doors for a direct reliability approach was recently published by Ljubinković et al. [77]. The primary benefit of the direct reliability approach lies in its application to existing structures. Using a Bayesian update, input parameters can be specified more realistically, and the remaining service life—typically shorter than that of a new structure—can be explicitly considered. Other approaches together with comparisons with DNN predictions for direct reliability assessment are described e.g. in [78, 79].

The direct reliability approach is not within the scope of this thesis.

Chapter 3

Determination of reliable strain limit

3.1 Explanation of concept

The author, together with several colleagues, aims to find a safe **failure criterion C2**, which typically governs the connection resistance. The concept of this approach was created by František Wald, Martin Vild, Jaromír Kabeláč, and Kirill Golubiatnikov and is discussed within Eurocode working groups. Most of the tedious work of numerical modeling was done by Kirill Golubiatnikov.

It must be assumed that the equivalent plastic strain limit is relevant only for steel plates unaffected by, for example, the heat-affected zone near the weld. The resistances of welds or bolts should be determined in a different way.

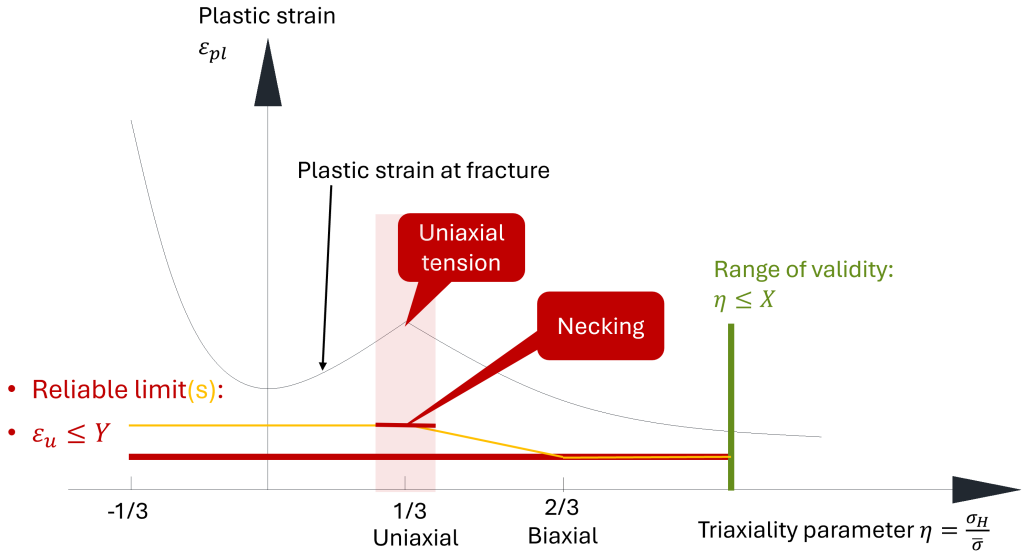


Figure 3.1: *Assumption: Plastic strain at fracture depends on triaxility; plastic strain at failure might depend too*

In the latest discussions in the development of Fpr EN 1993-1-14 [5], it is argued that the strain limit at fracture is dependent on the triaxility parameter, which is the ratio of hydrostatic stress to the von Mises equivalent stress.

There is a lot of research focusing on the effect of triaxiality parameter and lode angle on the strain at fracture, e.g. [80, 81, 82, 83]. The strain at fracture represents the upper bound for the strain at failure (at design resistance). However, this does not guarantee that the strain at design failure actually depends on triaxiality; see Fig. 3.1. The strain at fracture is typically much higher than the strain at design failure.

The strain at failure could be simplistically estimated using the safety factor γ_{M2}, which was crudely evaluated during the 1980s and 1990s with the development of the first generation of

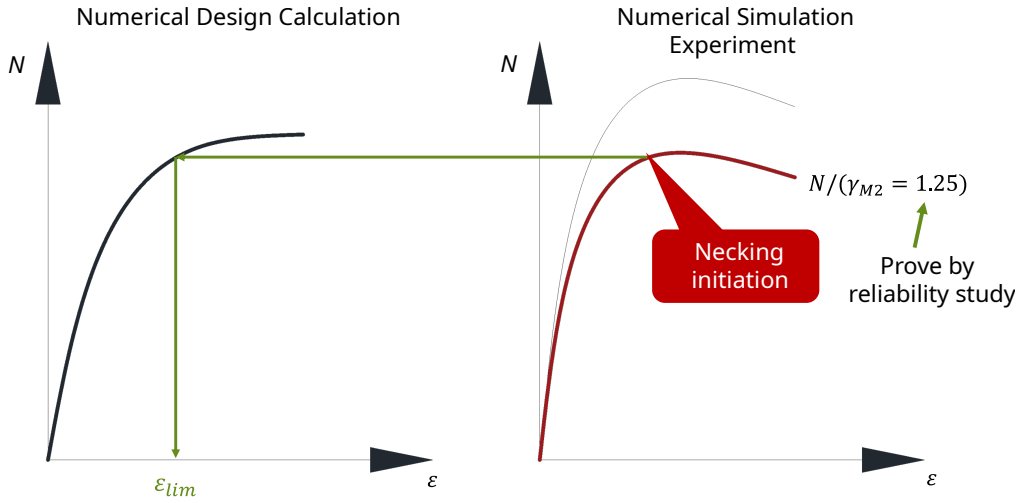


Figure 3.2: *Assumption: Safety factor can be applied to numerical simulation*

Eurocodes, but later refined and confirmed as 1.23 [84, 85, 86]. Therefore, from the literature it seems this simple approach might be sufficient; see Fig. 3.2. However, a confirmation is needed.

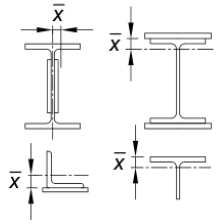
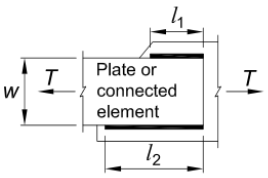
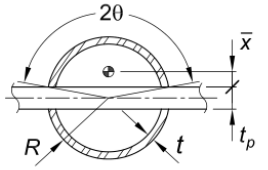
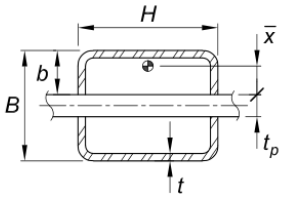
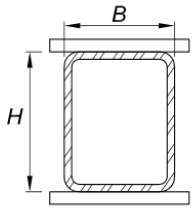
The criterion C2 is dominant especially for cases of tensile rupture (net-section failure). A large variety of physical tests were collected and modeled in IDEA StatiCa Connection to determine the ability of the numerical model with rough mesh to capture tensile rupture since this failure mode is checked in a fundamentally different way than the components [87]. However, in the large series of 529 specimens, only very few failed at a lower load than $A_g \cdot f_y$, gross section area multiplied by measured yield strength. The shear lag factors in Tab. D3.1 in AISC Specification [18] would not be needed if the resistance (safety) factor for tensile rupture were not required to be higher. In fact, in the Eurocodes [4, 1], these factors are missing, and only provisions for angles bolted by one leg (Cl. 3.10.3) or for unsymmetrical angles welded by the shorter leg (Cl. 4.13) are included [1].

Therefore, the plate specimens with holes and notches were selected as the most critical elements of steel structural joints. They are not only critical for a relatively large triaxiality factor, but also for large manufacturing tolerances – for the plates of a nominal thickness of 5 mm, the real thickness can be as small as 4.4 mm (12% reduction!) [88]. The real thicknesses were measured in the SAFEBRICTILE project and the distribution was normalized in FprEN 1993-1-1 – Annex E: Basis for calibration of safety factors [12]. Furthermore, the tolerances for bolt hole misalignment and ovalisation are normalized in EN 1090-2 [11]. Nearly every variable is known; the only issue is the unknown distribution of bolt hole tolerances and the effect of production, i.e., hole roughness. Still, a reasonable distribution can be derived using an uncertainty factor – a distribution of experimental tensile resistance to tensile resistance determined by the numerical simulation.

This research group believes that the net-section resistance is the most critical for the value of equivalent plastic strain limit for numerical design calculation, and when such limit is found, it is generally applicable to all numerical design calculations. To find this limit, a large set of plate specimens with stress concentrations induced by holes or notches was tested at Brno University of Technology and Czech Technical University in Prague, and numerical models with nominal geometry and also edge-case imperfections [11] were created.

The workflow is as follows:

1. Gather data from literature and perform experiments
2. Create numerical simulations with good alignment to the experiments
3. Approximate an analytical formula with good alignment to the numerical simulations
4. Create a large set of Monte Carlo simulations using analytical formulas with the distributions of material properties and geometry within the codified tolerances
5. Find the design resistance using reliability approach in EN 1990 [6]

Case	Description of Element	Shear Lag Factor, U	Examples
1	All tension members where the tension load is transmitted directly to each of the cross-sectional elements by fasteners or welds (except as in Cases 4, 5, and 6).	$U = 1.0$	–
2	All tension members, except HSS, where the tension load is transmitted to some but not all of the cross-sectional elements by fasteners or by longitudinal welds in combination with transverse welds. Alternatively, Case 7 is permitted for W, M, S, and HP shapes and Case 8 is permitted for angles.	$U = 1 - \frac{\bar{x}}{l}$	
3	All tension members where the tension load is transmitted only by transverse welds to some but not all of the cross-sectional elements.	$U = 1.0$ and A_n = area of the directly connected elements	–
4[a]	Plates, angles, channels with welds at heels, tees, and W-shapes with connected elements, where the tension load is transmitted by longitudinal welds only. See Case 2 for definition of $\bar{x}$.	$U = \frac{3l^2}{3l^2 + w^2}\left(1 - \frac{\bar{x}}{l}\right)$	
5	Round and rectangular HSS with single concentric gusset through slots in the HSS.	$\bar{x} = \frac{R\sin\theta}{\theta} - \frac{1}{2}t_p$ θ in rad $U = \left[1 + \left(\frac{\bar{x}}{l}\right)^{3.2}\right]^{-10}$	
		$\bar{x} = b - \frac{2b^2 + tH - 2t^2}{2H + 4b - 4t}$ $U = 1 - \frac{\bar{x}}{l}$	
6	Rectangular HSS with two side gusset plates.	$U = \frac{BU_B + HU_H}{H + B}$ $U_B = \frac{3l^2}{3l^2 + B^2}$ $U_H = \frac{3l^2}{3l^2 + H^2}$	

B = overall width of rectangular HSS member, measured 90° to the plane of the connection, in. (mm); D = outside diameter of round HSS, in. (mm); H = overall height of rectangular HSS member, measured in the plane of the connection, in. (mm); d = depth of section, in. (mm); for tees, d = depth of the section from which the tee was cut, in. (mm); l = length of connection, in. (mm); w = width of plate, in. (mm); $\bar{x}$ = eccentricity of connection, in. (mm).

[a] $l = \frac{l_1 + l_2}{2}$, where l_1 and l_2 shall not be less than 4 times the weld size.

Figure 3.3: *Table D3.1: Shear lag factors for connections to tension members from AISC Specification [18]*

6. Create numerical design calculation models with nominal material properties and geometry

7. Find a reliable plastic strain limit for numerical design calculation; see Fig. 3.4

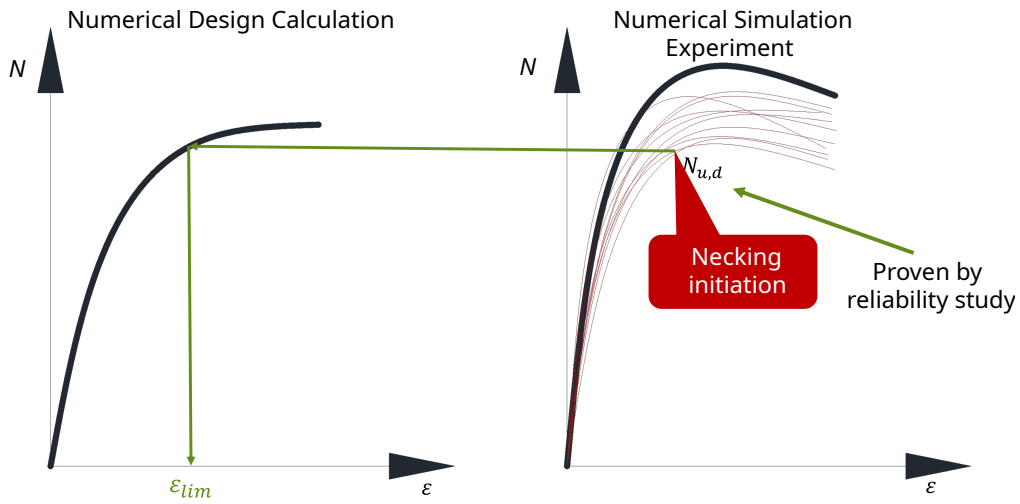


Figure 3.4: *Damage initiation at numerical simulation with probability of failure* $P_f(-\beta \cdot \alpha_R)$ *corresponds to reliable load resistance*

Reliable resistance by numerical simulation means the damage initiation (reaching ultimate strain ε_u at any location) of the simulation where only 0.1183 % of simulations provide the lower load at the damage initiation. Such a simulation is at the design resistance level with a probability of failure:

$$P_f(-\beta \cdot \alpha_R) = P_f(-3.8 \cdot 0.8) = 1.183 \cdot 10^{-3} \tag{3.1}$$

Consequently, the maximum plastic strain at the numerical design calculation equal to this reliable load resistance is the design plastic strain limit.

Based on the maximum possible weakening in EN 1993-1-8 – Tab. 3.3 [1] (see Fig. 3.5) and knowledge of fatigue, the most critical variants were selected; see Fig. 3.6. The stress concentration factor at elastic stage is determined from numerical simulations or for simpler cases according to Murakami [89]:

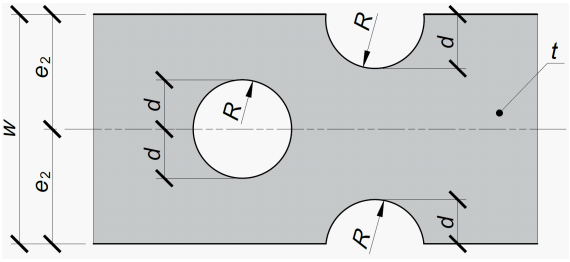


Figure 3.5: *Maximum weakening in structural plates according to EN 1993-1-8 [1]*

$$\mathrm{SCF}_{\mathrm{init}} = \frac{\sigma_{\max}}{\sigma_{\mathrm{nom}}} = 1 + 2\sqrt{\frac{h}{R}} \tag{3.2}$$

where:

- $\sigma_{\max}$ – maximum normal stress near the stress concentrator
- σ_{nom} – uniformly distributed normal stress far from the stress concentrator
- h – width of weakening
- R – radius of weakening

The initial triaxility parameter η_{init} is determined from finite element analysis; see Fig. 3.7.

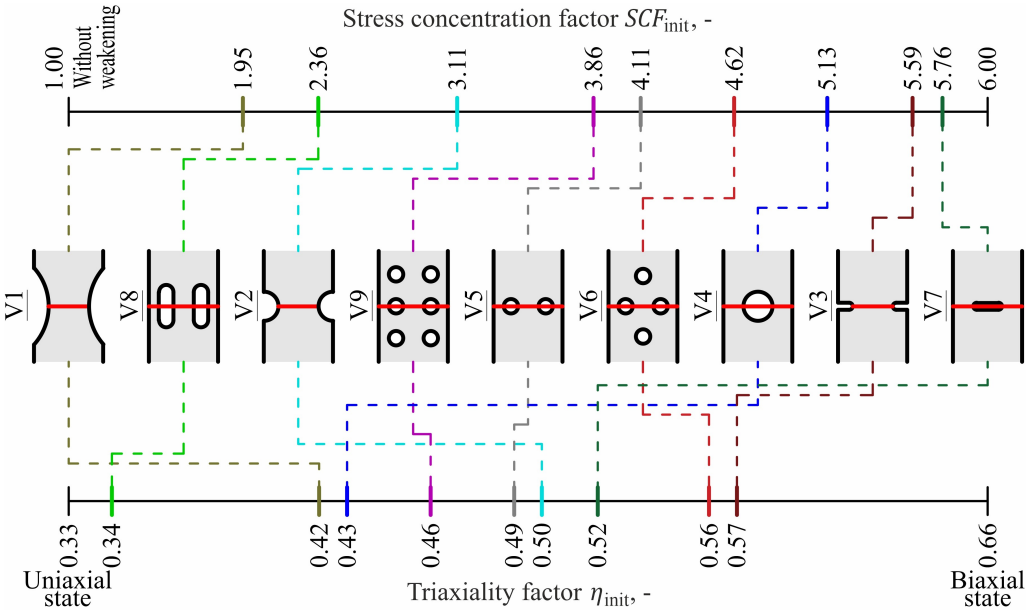


Figure 3.6: *Selected geometries with stress concentration factor and initial triaxiality parameter*

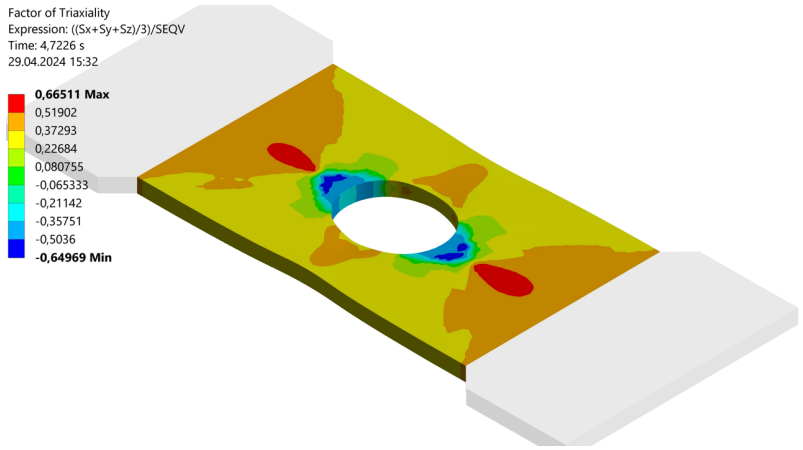


Figure 3.7: *Triaxiality parameter determined by ANSYS at elastic stage*

3.2 Experimental investigation

3.2.1 Brno University of Technology

The first set of experiments was performed using Labortech 6.1000.1 testing machine at AdMaS research center at the Faculty of Civil Engineering, Brno University of Technology in July 2024. A total of 18 specimens were tested. The plates were laser-cut and holes drilled. Three geometries (see Fig. 3.8) and two steel grades (S235 and S355; see Tab. 3.1) varied; each geometry of a steel grade was tested three times; see Tab. 3.2. The nominal dimensions of length × width × thickness were 500 × 100 × 6 mm, respectively. The thickness of each specimen was measured at multiple locations and the average calculated. Notably, specimens made of S235 were thinner, at the edge of tolerance limits. Only the force and displacement over the original gauge length of 220 mm, with bolt holes in the middle, were measured. The speed of loading was 2.8 mm/min.

Table 3.1: *Material properties of used steel from tensile coupon tests [10]*

Steel grade	E [MPa]	f_y [MPa]	f_u [MPa]	Strain at f_u [%]	Strain at fracture [%]
S235	209067	279	405	18.7	33.9
S355	202167	383	537	14.9	26.6

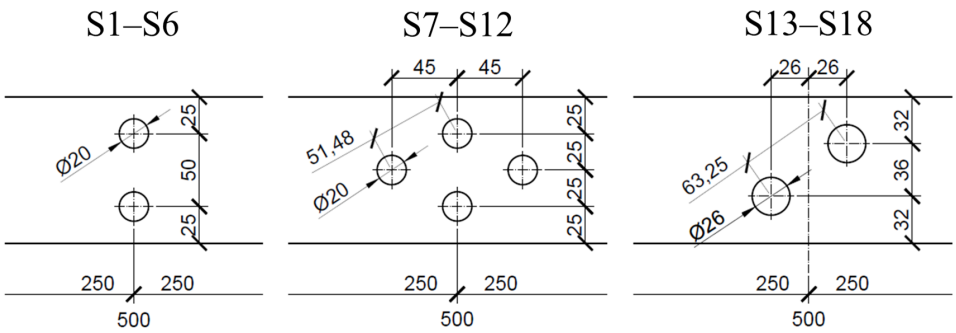


Figure 3.8: *Geometry of three specimen types for BUT experimental set; hole layout 2×d20 – type V5 (left), 4×d20 – type V6 (middle), 2×d26 – type V11 (right)*

Table 3.2: *Geometric and material properties of specimens*

Specimen	Holes	Steel grade	Thickness [mm]	d_0 [mm]	A_{net} [a] [mm^2]	A_{g} [mm^2]
S1	2×d20	S235	5.50	20.00	330.0	550.0
S2	2×d20	S235	5.49	20.00	329.3	548.8
S3	2×d20	S235	5.54	20.00	332.3	553.8
S4	2×d20	S355	5.95	20.10	355.8	595.0
S5	2×d20	S355	5.95	20.23	354.3	595.0
S6	2×d20	S355	5.83	20.15	347.8	582.5
S7	4×d20	S235	5.76	20.05	345.2	576.3
S8	4×d20	S235	5.88	20.08	351.6	587.5
S9	4×d20	S235	5.81	20.10	347.6	581.3
S10	4×d20	S355	6.26	20.15	373.9	626.3
S11	4×d20	S355	6.19	19.83	373.4	618.8
S12	4×d20	S355	6.06	20.10	362.5	606.3
S13	2×d26	S235	5.83	26.10	387.8	582.5
S14	2×d26	S235	5.91	26.00	394.8	591.3
S15	2×d26	S235	5.88	26.03	392.0	587.5
S16	2×d26	S355	6.26	26.00	418.2	626.3
S17	2×d26	S355	6.25	26.03	417.0	625.0
S18	2×d26	S355	6.13	26.08	408.1	612.5

[a] Calculated according to EN 1993-1-1, Cl. 6.2.2.2 [4].

The photographs of failed specimens are in Fig 3.9. Specimens S7–S12 with four bolt holes seemingly failed in the same way as specimens S1–S6 with two bolt holes, but the deformation at failure was considerably larger. There was a significant strain achieved near the single hole before the tensile rupture at two bolt holes occurs. The load-deformation curves are plotted in Fig 3.10. The results are very consistent. Note that the specimens with shear plane (S13–S18) achieved the peak load just before tensile rupture, but the shear plane ultimately failed at very large deformations.

Eurocode reliability EN 1993-1-1 [4] provides guidance on the calculation of the net area in Cl. 6.2.2.2 and tensile rupture resistance in Cl. 6.2.3. In the conducted experiments, tensile rupture always governed over tensile yielding. To maintain consistent reliability, the ratio of experimental resistance to calculated resistance should be constant, but this ratio changes dramatically in Tab. 3.3. Results of specimen triplets are averaged, and measured material properties are used in calculations. N_t stands for simple multiplication of A_{net} and f_u, Interestingly, specimens S1–S6 reached even greater resistances by 10%, indicating that the normal stress exceeded the nominal ultimate strength. This can be explained by stress biaxiality. Due to stress flow and Poisson effect,

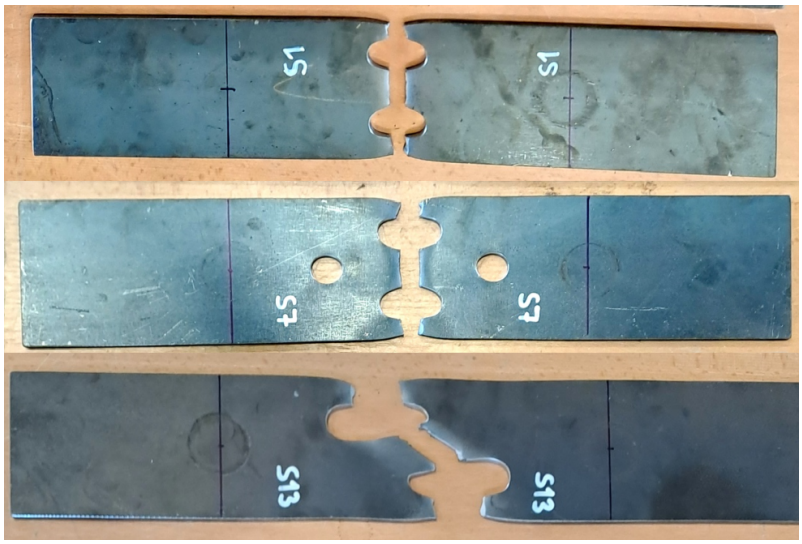


Figure 3.9: *Photographs of ruptured plates with centered and staggered bot holes*

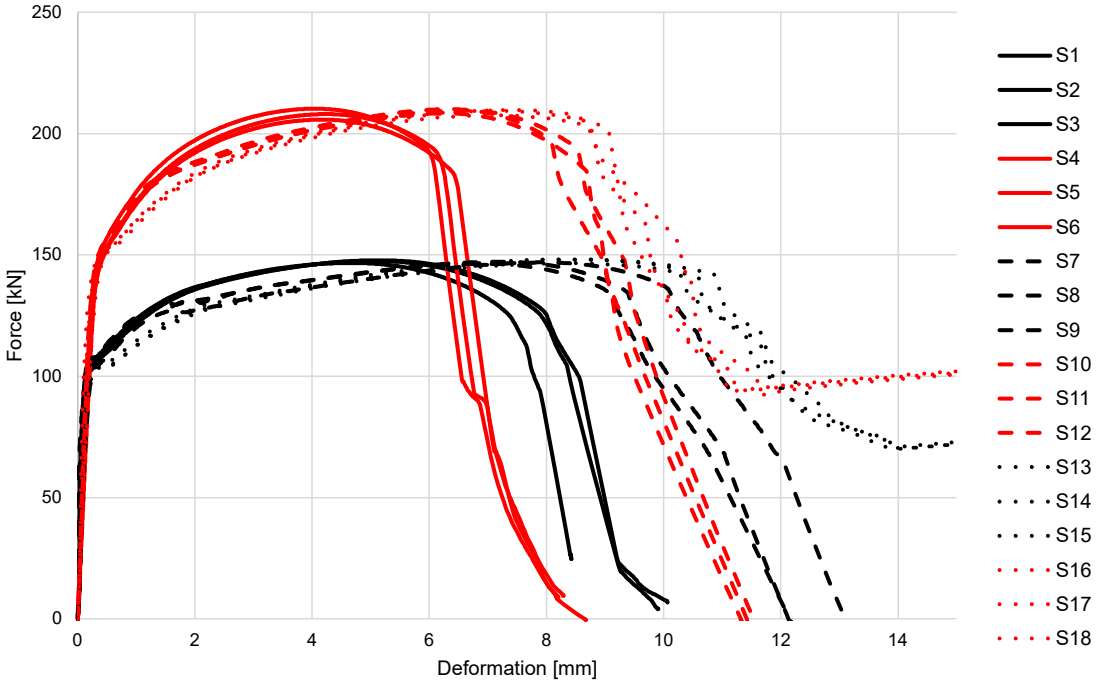


Figure 3.10: *Results of specimens S1–S18 tested at Brno University of Technology*

the steel at bolt holes is in biaxial tension and the normal strength is increased according to von Mises theorem. For specimens S7–S12, the effect was similar but lessened by the occurrence of another hole, which disrupted the stress flow.

$N_{t,Rd}$ is calculated according to Cl. 6.2.3. The formulas are highly conservative for specimens S1–S12, overestimating the resistance on average by 50%. The overestimation for specimens S13–S18 is significantly lower, only 30%. Clearly, the reliability for only tensile rupture and combined tensile and shear rupture is different. These findings are supported by another research [86]. The reason for very high safety margin was an assumption of fatigue fracture at the bolt hole. However, this is inconsistent; the assumption is not kept for combined tensile and shear rupture. A revision of tensile rupture calculation would be beneficial, but the formulas stay the same also in the next generation of Eurocode [12]. On the other hand, the reliability of both steel grades, S235 and S355, seem very consistent.

Table 3.3: *Comparison of analytical, design, and experimental tensile resistances*

Specimen	S1–S3	S4–S6	S7–S9	S10–S12	S13–S15	S16–S18
A_{net} [mm^2]	330.5	352.6	348.1	369.9	391.6	414.4
A_{g} [mm^2]	550.8	590.8	581.7	617.1	587.1	621.3
N_{t} [kN]	133.9	189.4	141.0	198.7	158.6	222.6
$N_{\text{t,Rd}}$ [kN]	96.4	136.3	101.5	143.0	114.2	160.2
$N_{\text{y,Rd}}$ [kN]	153.7	226.3	162.3	236.3	163.8	237.9
$N_{\text{t,exp}}$ [kN]	147.3	208.0	146.8	209.3	147.6	208.9
$N_{\text{y,exp}}$ [kN]	102.3	132.1	101.9	138.3	103.7	134.5
$N_{\text{t,exp}}/N_{\text{t}}$ [%]	110.1	109.9	104.1	105.4	93.1	93.9
$N_{\text{t,exp}}/N_{\text{t,Rd}}$ [%]	152.9	152.6	144.6	146.4	129.3	130.4

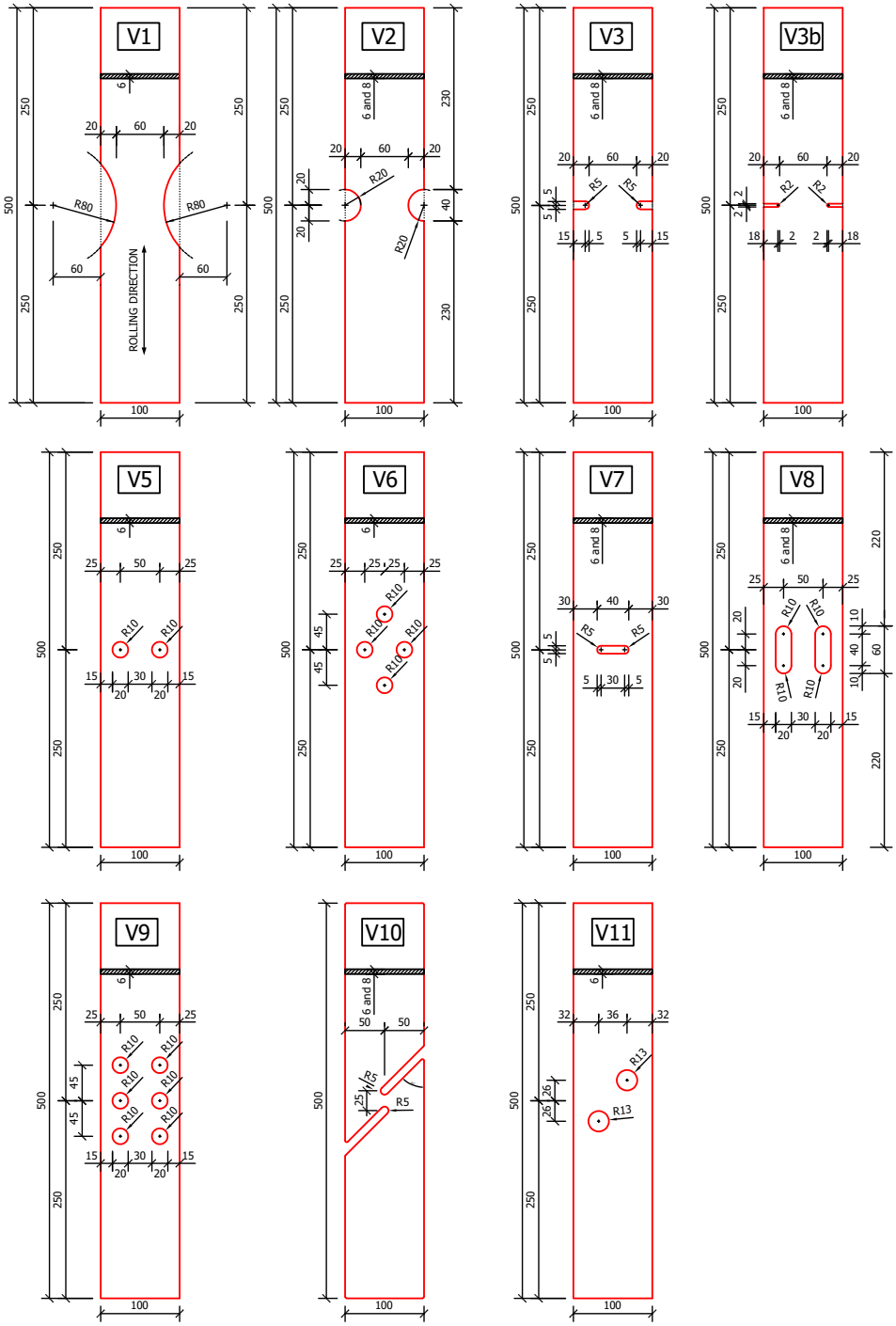


Figure 3.11: *Geometries of the second set of specimens tested at BUT*

The second set of experiments was performed again at the AdMaS center at BUT using the same test machine. The experiments proceeded in the same manner as the first set, only the loading rate was decreased to 1 mm per minute. The measurements and testing were performed by Amina Hajdarević, Michal Štrba, Martin Vild, Mykola Lastovetskyi, Petr Daněk, and Pavel Schmid.

This is a very large set of experiments containing 11 geometries, two thicknesses (see Fig. 3.11), and four different nominal materials: S 235 (Note that the delivered specimens consisted of two distinct materials, labeled based on visual surface texture: S for smooth and R for rough. Tensile coupon tests were conducted for both materials.), S 355, S 460, and S 700.

Every type with the same geometry, type, and thickness was manufactured and tested three times. In total, 204 specimens were tested. However, due to some issues with the test machine, several specimens had to be discarded.

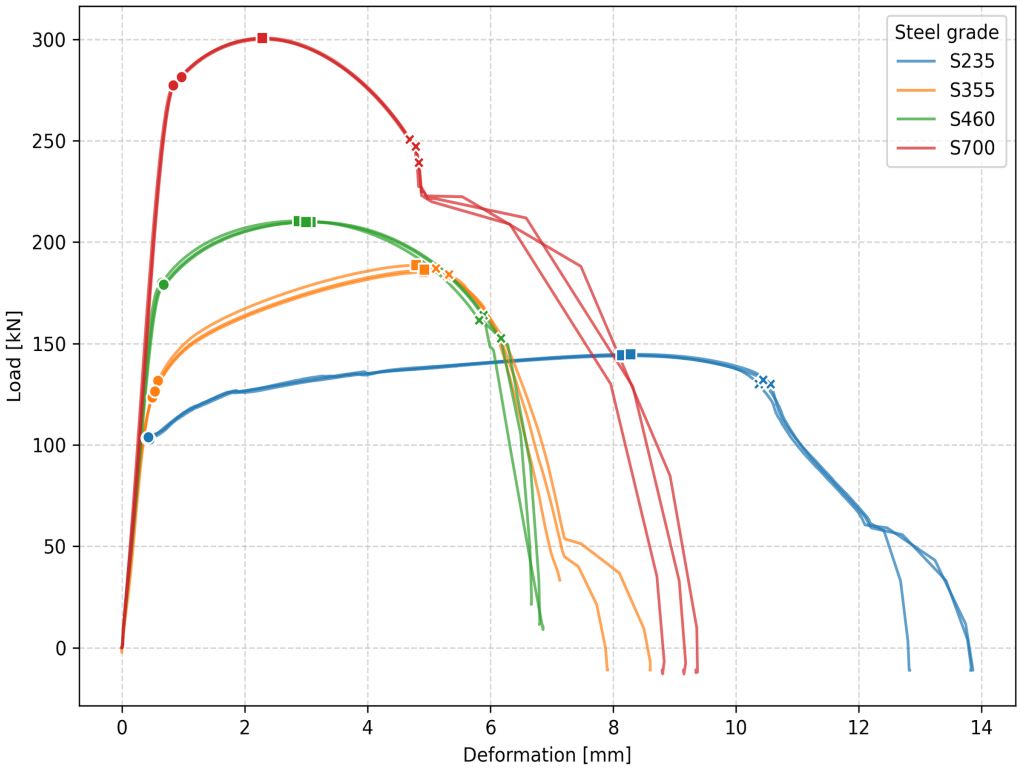


Figure 3.12: *Load-deformation curves for specimen V6 and nominal thickness of 6 mm for all steel grades*

Examples of load-deformation curves for geometry type V6 with four bolt holes are shown in Fig. 3.12. The yield point is marked by circle. It was determined by fitting the initial stiffness, shifting it by 0.0005×gauge length equal to 220 mm and finding the intersection with load-deformation curve. The peak point is marked by square. Ultimate strength was likely achieved at slightly smaller load, but it cannot be experimentally recognized. The fracture is marked by a cross, and was determined by the gradient of -50 kN/mm, but for many cases manually overridden by the engineering judgment and alignment to the DIC measurements. Nevertheless, the real fracture initiation will be found later with the usage of numerical simulations. Fracture initiates where the numerical simulation with the true stress-strain material model begins to diverge – the experiment starts descending faster than the numerical simulation.

The data can be visualized in the form of histograms. All geometry types and thicknesses are grouped together. The average engineering stress (calculated as the peak load divided by original measured area) at the peak load is shown in Fig. 3.13. Note that some specimens failed in pure tension, others in pure shear (V10) and others in combined tension and shear. The average peak stress is greatly affected by this. S 700 steel grade achieved by far the highest average stresses; the differences between other steel grades are not so apparent. Both S 235 and S 355 were very likely much stronger than their nominal strengths.

The deformations at the ultimate load are shown in Fig. 3.14 and are very relevant to the

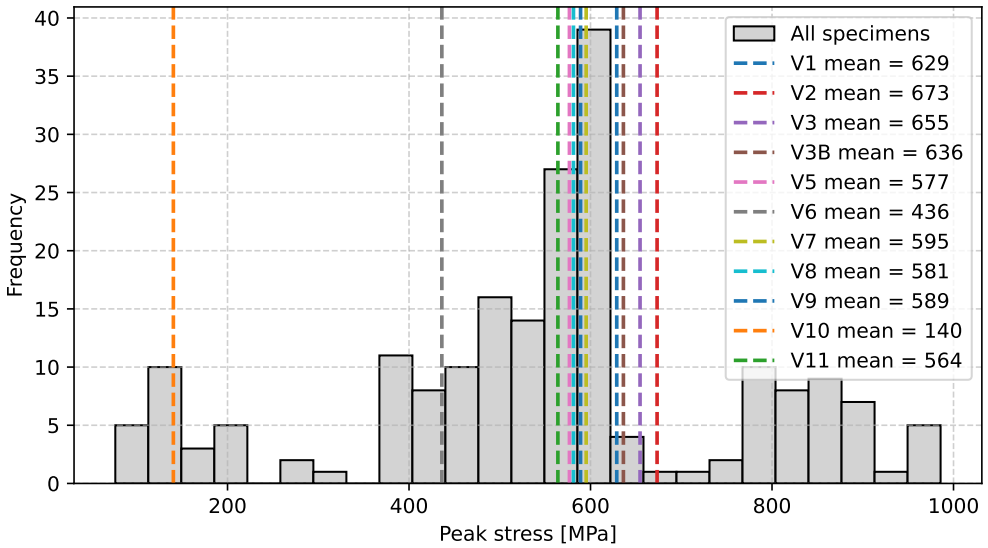


Figure 3.13: *Histogram of peak stresses achieved for all specimens*

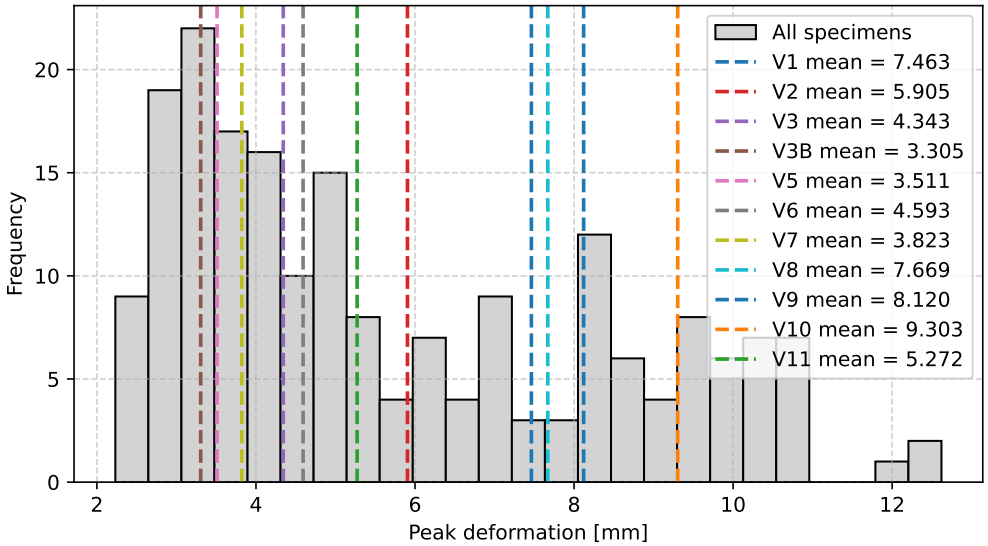


Figure 3.14: *Histogram of deformations achieved at peak load*

research of plastic strain limit in numerical design calculation. The deformations were measured at the gauge length of 220 mm by LVDT, therefore some very small portion of deformation is attributed to the elastic deformation of plates. The smallest deformations at peak load were generally achieved by plates made of S 700 and by geometries with the highest stress concentration factor (geometry V3b); see Tab. 3.4.

Deformation at fracture determined by a gradient or manually measured at 220 mm gauge length is shown in a histogram in Fig. 3.15. The order of geometry types by fracture deformation is similar to that of deformations at the peak load. The exception is geometry type V10 loaded by pure shear, where all the specimens failed very shortly after reaching the peak load.

Most of the specimen types and materials were measured by digital image correlation by Kirill Golubiatnikov and Vojtěch Stančík. Several outputs are in Fig. 3.16 and 3.17.

From DIC measurements, several preliminary findings could be drawn:

- The peak strains in the plastic region were in the center of the weakened specimen and the fracture originated there for all steel grades for the geometries with low stress concentration factor, i.e. type V1 and V8.

Table 3.4: *Mean values of deformation at peak load for individual geometry types*

Geometry	Mean deformation [mm]
V1	7.463
V2	5.905
V3	4.343
V3b	3.305
V5	3.511
V6	4.593
V7	3.823
V8	7.669
V9	8.120
V10	9.303
V11	5.272

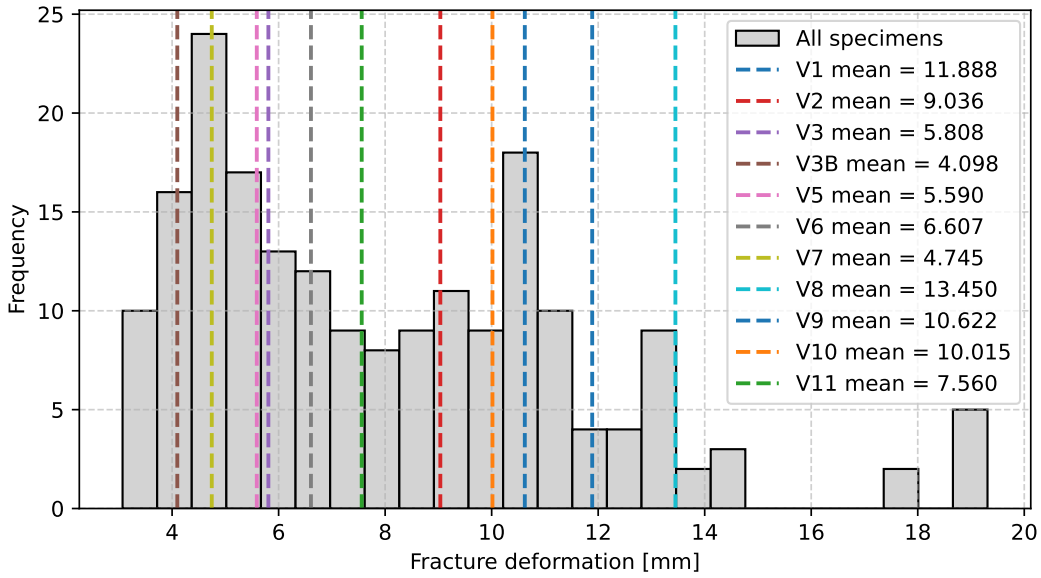


Figure 3.15: *Deformation at fracture (significant load gradient drop)*

- The peak strains and fracture origin were located at the weakening edge for higher SCF_{init}, i.e. geometry types V2, V3, V3B.
- For V5 (two bolt holes) and V6 (four staggered holes), the strain peaks and fracture origins were located at the inner edge of the bolt holes, but interestingly, this switched to the outer edge of the bolt holes for V9 (six bolt holes).
- The stress concentration factor appears to govern the maximum strains measured by DIC immediately before fracture. The smallest strains occurred for geometry type V3B (SCF_{init} = 9.944; see Eq. (3.2)) around 0.1. Note that such weakening geometry is outside the scope of structural engineering detailing rules where the minimum notch radius should be 5 mm (EN 1090-2 – Cl. 6.7 [11]). The highest strains were reached for V1 (around 0.3) and for V8 (up to 0.5).
- The difference between steel grades is noticeable but not great. For several geometries, the strain at fracture initiation as measured by DIC was actually lower for S 355 than for S 700. Typically, the strains at fracture initiation were for S 700 lower by about 30 % than for other steel grades.

The fracture area was captured by an electronic microscope by Karel Vařeka, a Ph.D. student at CEITEC at BUT. A comparison of micrographs with the same magnification of steel grades

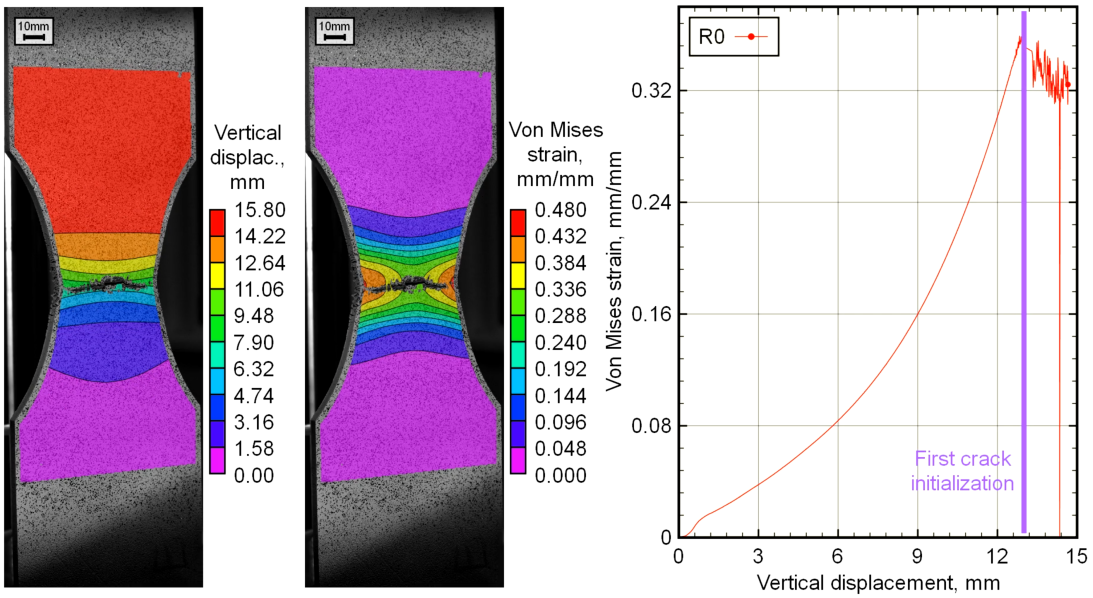


Figure 3.16: *Specimen V1 made of S235 steel grade captured by DIC*

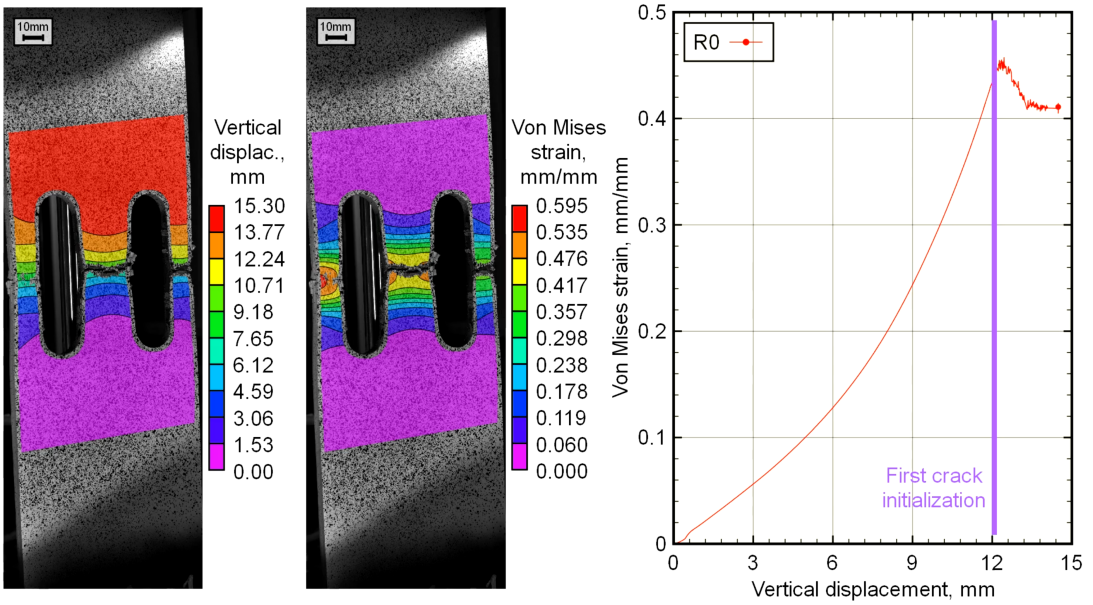


Figure 3.17: *Specimen V8 made of S460 steel grade captured by DIC*

S 355 and S 700 is shown in Fig. 3.18, 3.19, and 3.20. From these images, the following observations can be made:

- S 355 (left):
  - At small and medium magnifications, the fracture surface is irregular with multiple indentations and uneven topography. No clear layering is visible; the surface appears more homogeneous.
  - The fracture surface at large magnification is rough and irregular. Numerous dimples and microvoid coalescence are visible, which indicates a ductile fracture morphology.
- S 700 (right):

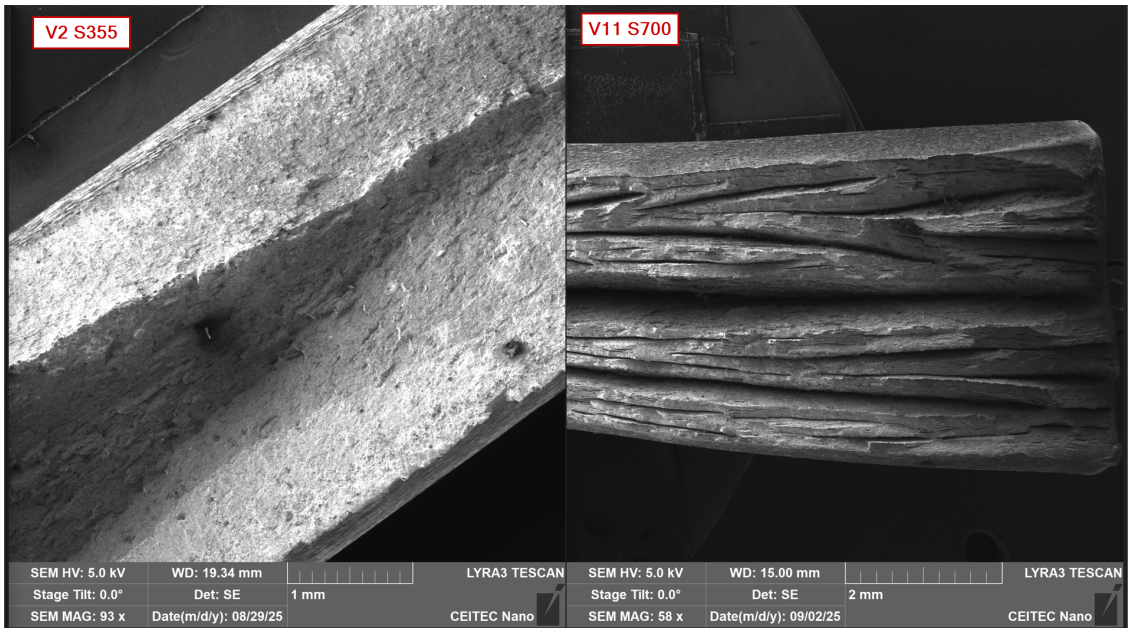


Figure 3.18: *Small magnification of fracture surfaces: Significant delamination is observed at each S 700 specimen*

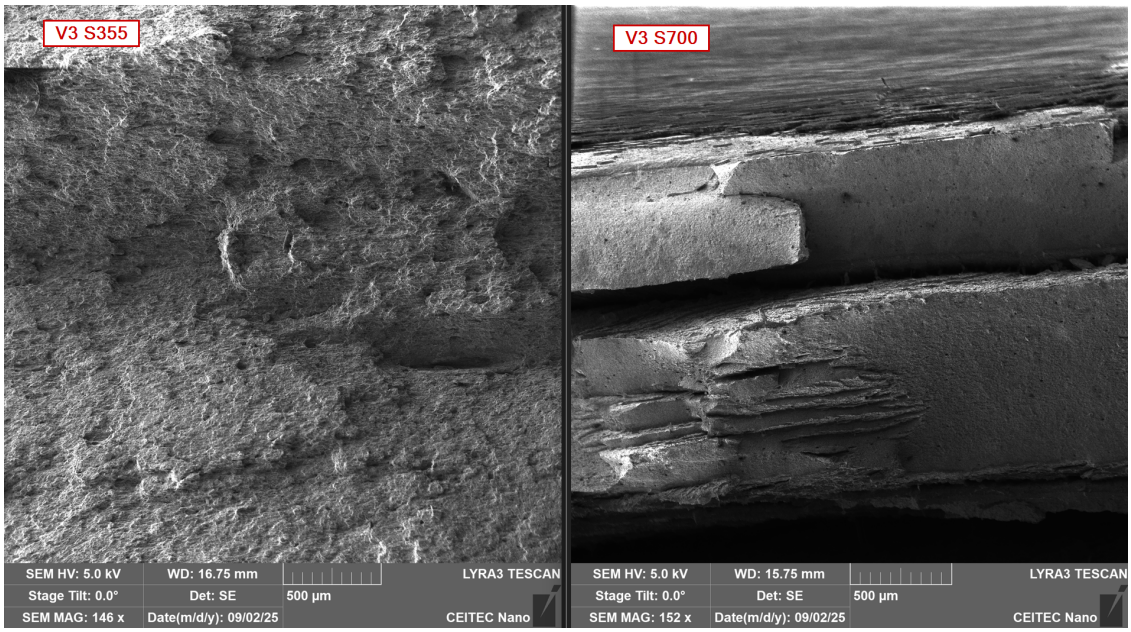


Figure 3.19: *Medium magnification of fracture surfaces*

- At small and medium magnifications, the cross-sectional view is with distinct layers. There are visible fracture separations between layers, suggesting a more complex mechanism, involving delamination. The structure looks more organized compared to S 355.
- The fracture surface at large magnification is smoother and more layered. There are fewer dimples and voids compared to S355, which suggests a less ductile (more brittle) fracture morphology.

This theme is repeated in all cases. Steel grade S 700 shows a distinctly different fracture pattern with delamination, while mild steel (S 235 to S 460) fractures with large uniform surfaces at an angle about 45° from the longitudinal axis. These surfaces are switched suddenly several times; see Fig. 3.18.

The experimental measurements, DIC data, and photographs will be subjected to further inves-

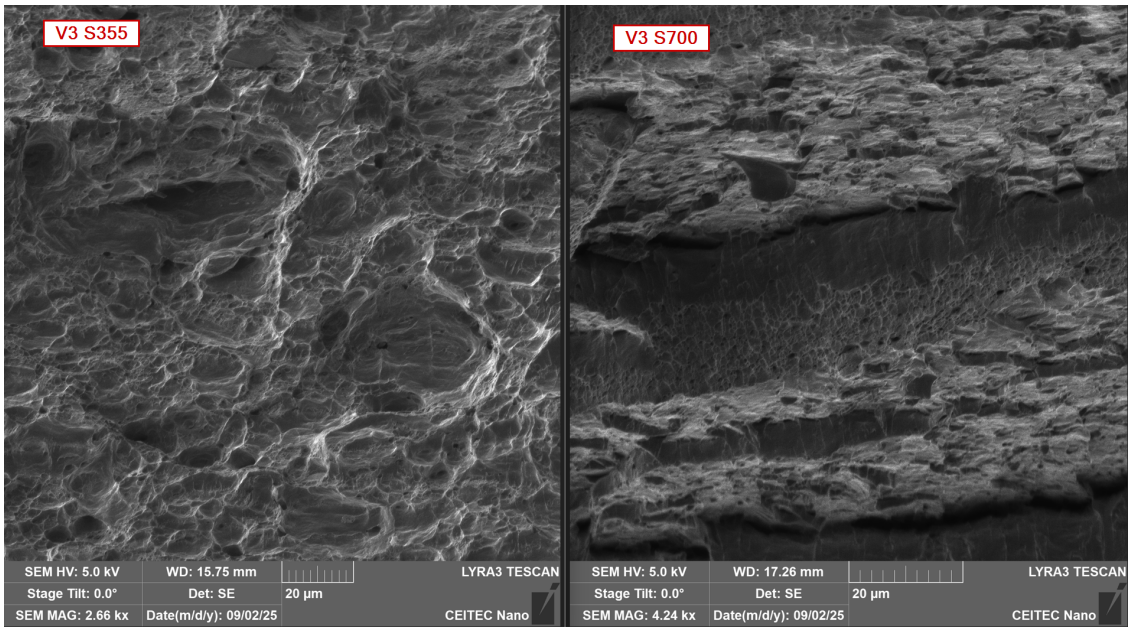


Figure 3.20: *Large magnification of fracture surfaces*

tigations and numerical simulations with the aim of determining the design strain limit for failure in numerical design calculation as well as parameters for the model fracture, i.e., the fracture initiation and descending branch of the load-deformation diagram for numerical simulation. The data will be uploaded to the Zenodo database once a research paper is published.

3.2.2 Czech Technical University in Prague

Another set of experiments was performed at the Czech Technical University in Prague (CTU). These experiments were designed to find the load resistance when the geometrical tolerances are at their limits. The plate thickness was machined to the limits in EN 10029 [88] and the bolt holes were drilled at the extreme out-of-center and out-of-shape tolerances in EN 1090-2 [11]. The experiments are described in detailed in a paper by Golubiatnikov [90] and another in review [91]. The basic geometry of experiments was very similar to those of geometry types V1–9 shown in Fig. 3.11. Shimadzu 300 kN testing machine performed the test with the controlled deformation speed at 0.26 mm/min; see Fig. 3.21.

All the performed experiments together with the well-described experiments from the literature are critical for validation of numerical simulations and determination of numerical-analytical model and uncertainty factor for Monte Carlo simulations.

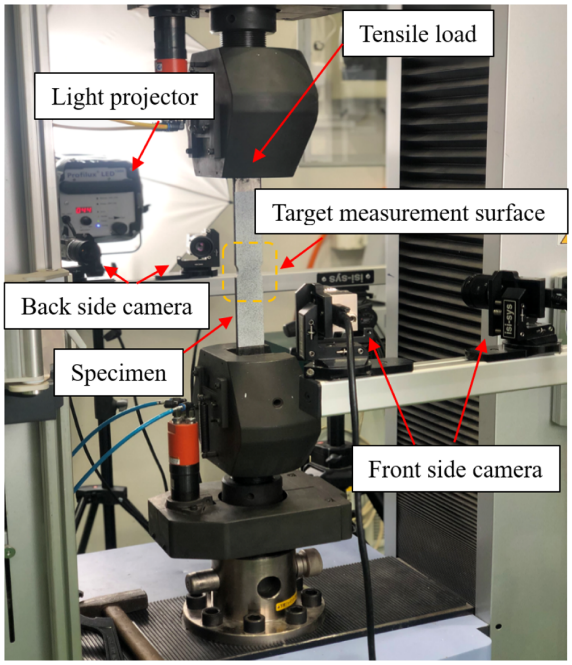


Figure 3.21: *Tensile test at CTU Prague [90]*

3.3 Numerical models

Numerical simulations were created by Kirill Golubiatnikov in ANSYS [23] at CTU Prague. The models are described in detail in publications [90, 91]. Three model types were created: (i) damage model able to simulate the descending branch of load-deformation diagram, i.e. steel fracture, (ii) solid model that is valid only up to necking, i.e. reaching f_u in the most stressed element. GMNIA was used in both models, although geometric nonlinearities and imperfections did not affect the results noticeably. And finally (iii) numerical design calculations using nominal material models and nominal geometry.

3.3.1 Damage model

Damage model utilizes explicit dynamic solver and 4-node tetrahedron elements. Damage model is meshed relatively coarsely. The goal was finding parameters that simulate experimental load-deformation curve of notched tensile test specimens and stress and strain distributions captured by DIC; see Fig. 3.22.

3.3.2 Solid model

Solid model is finely meshed by 20-node quadratic cubic elements. Because the damage criteria are missing, the model is valid only until the first element reaches necking, ultimate strength f_u. For tensile coupon test [10] with uniaxial loading, this moment is coincidental with the peak in the load-displacement diagram. For specimens with weakenings, such as notches or bolt holes, f_u is reached sooner, at the load ≈3 % lower, depending on the weakening geometry. Assuming this is the peak load is therefore a safe assumption. Fig. 3.23 shows the comparison of experimental load-deformation curve with the numerical with real material stress-strain diagram (Real MC) and four-point stress-strain diagram with measured yield and ultimate strengths (Artificial MC). The load resistance determined by Artificial MC is at most 7 % lower, while the deviation in displacements is greater.

Solid models were used for numerical simulations and determination of the effects of geometric tolerances and varying material curves with four-point material model [7].

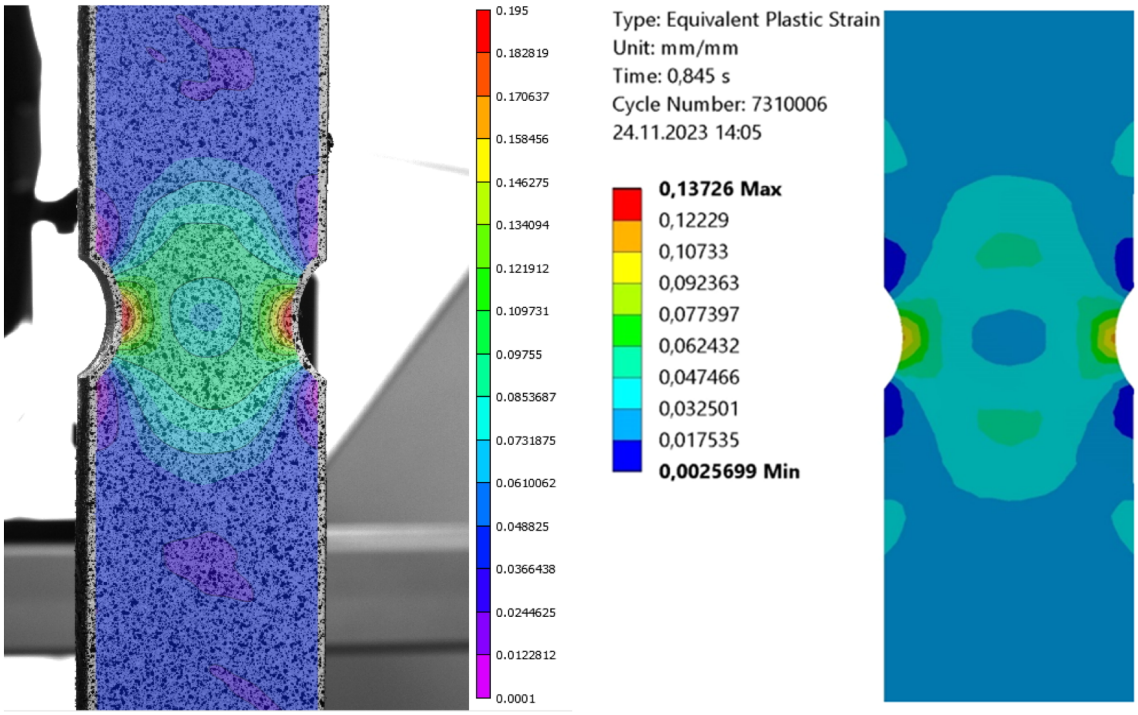


Figure 3.22: *Validation of numerical simulation – load at the lower yield point [90]*

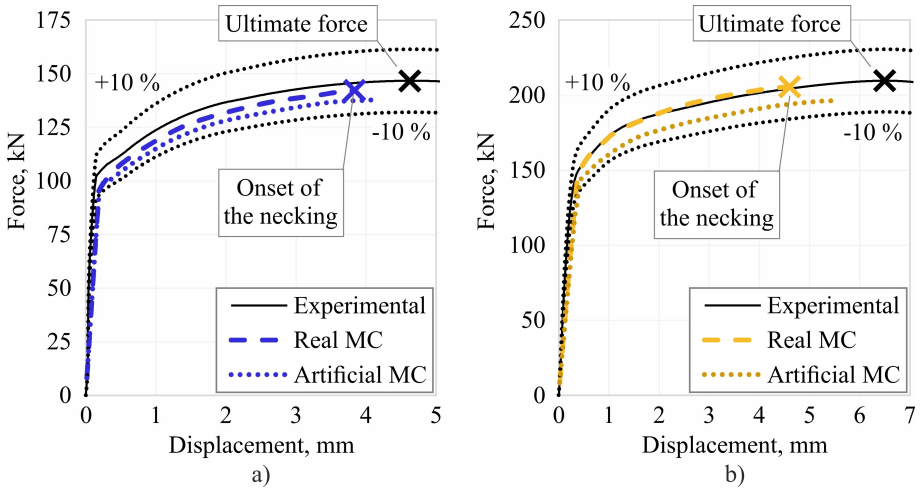


Figure 3.23: *Comparison of real and four-point material curve for geometry types V2 (left) and V3 (right) [91]*

3.3.3 Shell model

Shell models were used for numerical design calculations. That means nominal geometry and nominal material properties were utilized. The effect of mesh densities, small or large deflections theories, and element types were investigated. As shown in Fig. 2.14, net section failure is not very dependent on mesh density, nor is it particularly dependent on geometrical nonlinearities or element types. FprEN 1993-1-14 [5] allows different material diagrams; see Fig. 1.2. It was shown that the four-point material model provides the highest load resistance [90]. In Fig. 3.4, shell models with four-point material model are utilized to create the left graph for each geometry type.

3.4 Numerical-analytical model

A single numerical simulation took two and a half hour. Numerical-analytical model (curve-fitted formulas based on numerical simulations) is therefore needed to run millions of simulations for

Monte Carlo reliability analysis. Solid models described in Section 3.3.2 were created and analyzed in large quantities.

The numerical-analytical model is assumed to be a function of material M, geometry G, and uncertainty U:

$$N_{\mathrm{Rd}} = A_{\mathrm{net}} \cdot f_{\mathrm{u}}/\gamma_{\mathrm{M2}} = f_{N_{\mathrm{Rd}}}(M, G, U) \tag{3.3}$$

3.4.1 Material factor

Eurocode EN 1993-1-1 in Cl. 3.2.2 [4] defines the material, which may be used for structural purposes. There are three limits:

- Ratio of ultimate to yield strength $f_u/f_y \geq 1.10$.
- Elongation at failure must be not less than 15 %.
- Ultimate strain $\varepsilon_u \geq 15 \cdot \varepsilon_y$, where ε_y is the yield strain ($\varepsilon_y = f_y/E$).

It is assumed that any structural steel must conform to these limits and otherwise, the distribution in FprEN 1993-1-1 – Annex E [12] is used; see Fig. 3.24.

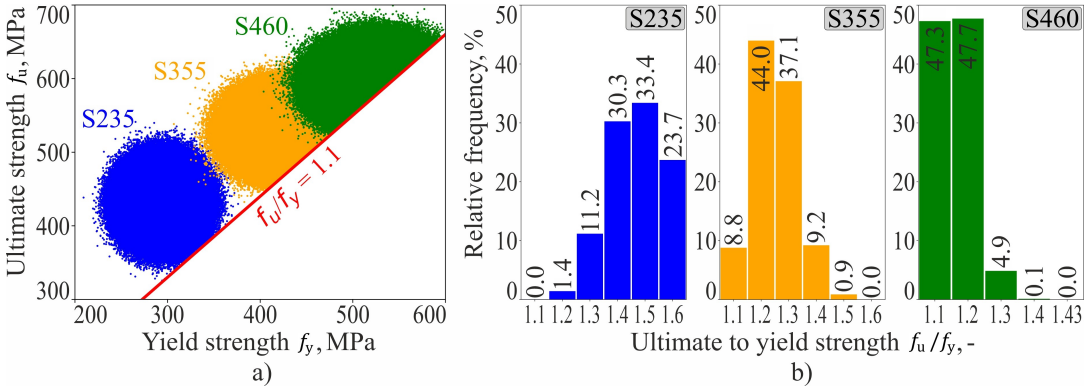


Figure 3.24: *Material properties of European structural steel – real (left) and simplified (right) distribution [92]*

A set of numerical simulations was performed for three steel grades (S 235, S 355, and S 460) and six groups of f_u/f_y ratios: f_u/f_y = 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, except for steel grade S 460 with maximum ratio of f_u/f_y = 1.43 for all geometry types V1–V9. Used material models are shown in Fig. 3.25. They are represented in the Monte Carlo simulations by relative frequencies shown on the right in Fig. 3.24. The resistances of numerical simulations are further modified by geometry reduction factor G (typically reducing the resistance) and uncertainty factor U (typically increasing the resistance) into Monte Carlo simulations.

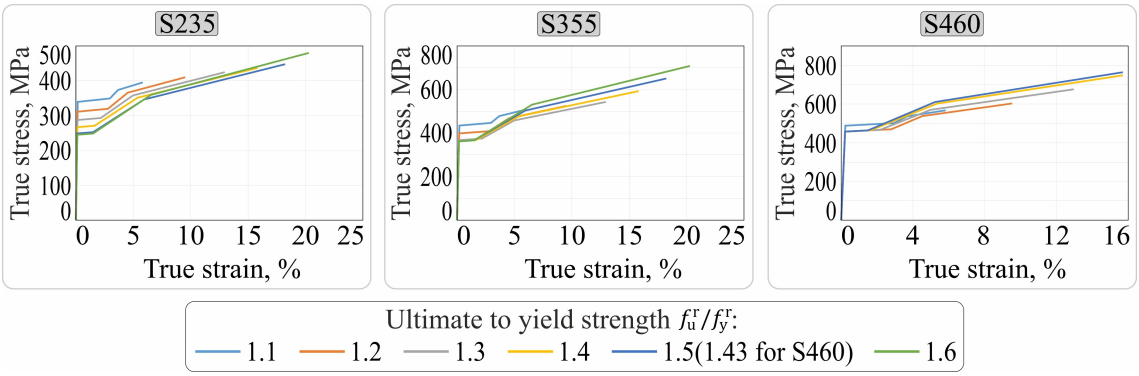


Figure 3.25: *Four-point material models used for numerical simulations*

3.4.2 Geometry reduction factor

Geometry reduction factor G takes into account the effect of manufacturing tolerances, such as hole misalignment or variability in plate thickness. In total, 198 numerical simulations were performed. Three thicknesses (4.40 mm, 4.55 mm, and 5.00 mm), nine geometry types with and without geometric imperfections at the edge of tolerances [11] (in total 33 geometries), and two material models were varied. The Geometry reduction factor was determined as a lower-bound linear function of the plate thickness taking into account any combination of geometrical imperfections due to tolerances:

$$G = p_1 \cdot \frac{t_{real}}{t_{nom}} + p_2 \tag{3.4}$$

where t_{real} is the real thickness, t_{nom} is the nominal thickness, and p_1 and p_2 are parameters in Tab. 3.5.

Table 3.5: *Values of parameters p_1 and p_2 for geometry types V1–V9 [91]*

Variant	V1	V2	V3	V4	V5	V6	V7	V8	V9
p_1	0.915	0.915	0.939	0.950	0.962	0.944	0.874	0.963	1.041
p_2	-0.005	0.004	-0.053	0.019	-0.021	-0.024	0.023	-0.005	-0.090

3.4.3 Uncertainty factor

Uncertainty factor U simulates the production-related factors, such as hole drilling/punching/-cutting methods and counter weighs the conservatism of modeling approach. Uncertainty factor was determined by comparing the numerical simulations with the above-described modeling procedures to the experimental research [93, 90, 85, 94, 95, 96, 97] comprising 178 static tensile tests. The uncertainty factor here therefore coincides with the bias b. Fig. 3.26 shows the distribution of experimental r_e to numerical-analytical r_t resistance ratio. The histogram of r_e/r_t ratios was approximated by normal distribution with mean $\mu_U = 1.115$ and standard deviation $\sigma_U = 3.69\,\%$ with the minimum value of $U = 1.00$ and maximum value of $U = 1.20$.

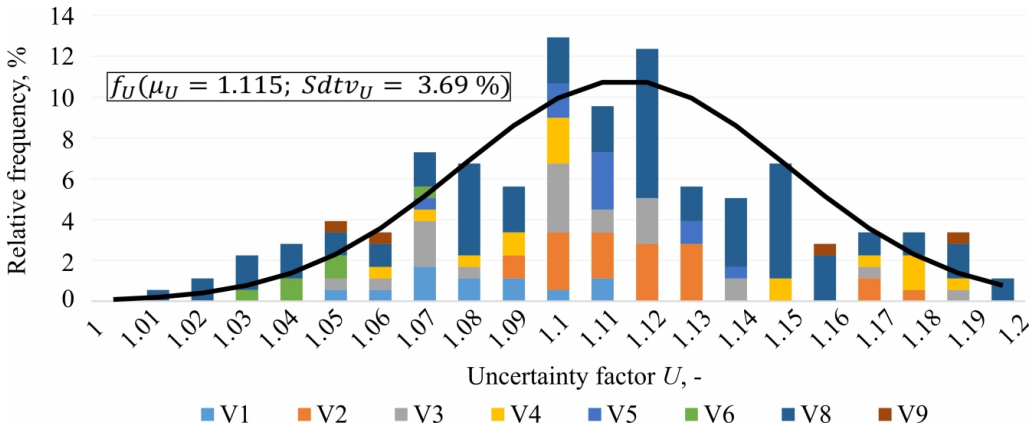


Figure 3.26: *Distribution of experimental to numerical-analytical resistance ratios [91]*

3.5 Monte Carlo simulations

3.5.1 Design resistance

Monte Carlo simulations were generated using a Python script, where input sets of parameters – geometry type, ultimate-to-yield strength ratios, real thicknesses, uncertainty factors – were generated independently as pseudorandom parameters (parameters are independent of each other). The resistance was calculated by numerical-analytical method described above.

The required sample size was determined through a sensitivity study of the design net cross-section resistance $N_{\mathrm{R,d}}$ for geometry V3 of S235 steel grade, which exhibited the highest local stress peak. Input sets were generated for several sample sizes, and the lowest 0.1184 % of resistances were discarded. The smallest remaining value was taken as the design resistance $N_{R,d}$. Each case was repeated ten times per sample size, and standard deviations from the mean were evaluated. A target maximum deviation of 0.01 was achieved with approximately 2.8 million samples. Three million samples were used for each geometry in simulations.

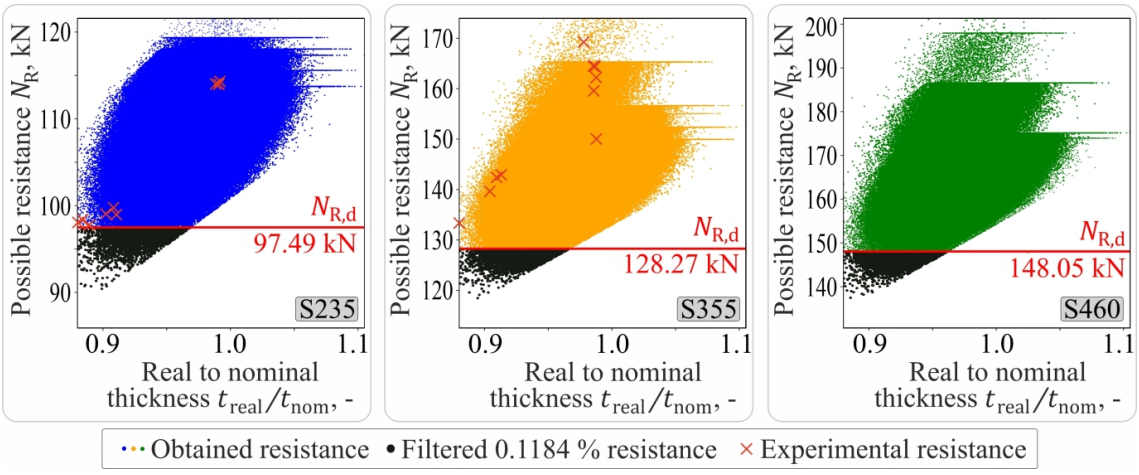


Figure 3.27: *Determination of design resistances for varying steel grades, geometry type V3 [92]*

Fig 3.27 shows an example of the generated resistances for geometry V3 across all considered steel grades. Colored dots represent the simulated results, while red cross marks experimental resistances of specimens with different permitted deviations; see Section 3.2.2. A reminder that the experimental specimens were fabricated in three thicknesses: 5 mm, 4.55 mm, and 4.4 mm – the lowest allowed deviation from the nominal 5 mm thickness.

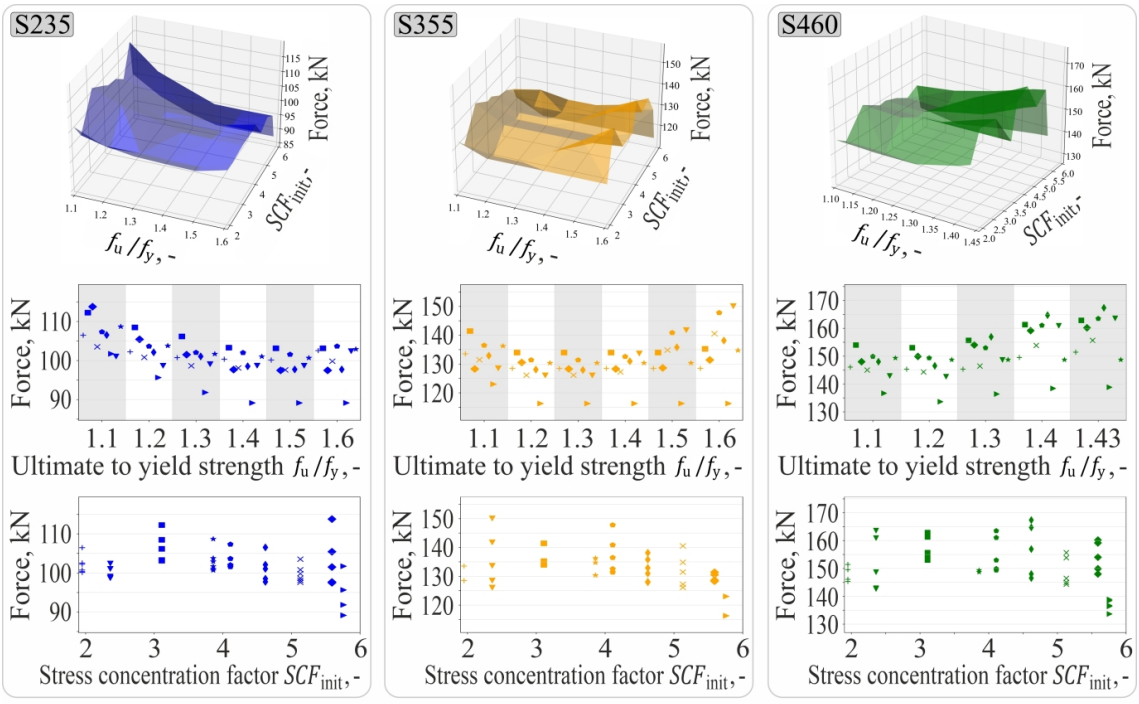


Figure 3.28: *Design resistances $N_{R,d}$ for geometry types V1–9 [92]*

Fig. 3.28 presents the resulting design resistances $N_{R,d}$ for all geometry types categorized according to steel grades, initial stress concentration factors SCF_{init}, and ultimate-to-yield strength

ratios f_u/f_y. Geometry type V7 (the highest $SCF_{init} = 5.76$, above-average triaxiality factor $\eta_{init} = 0.52$; see Fig. 3.6) shows consistently the lowest design resistance of all geometry types. Otherwise, no clear correlation with stress concentration or triaxiality was found.

3.5.2 Partial safety factor γ_{M2}

The same procedure as for design resistance was used for the partial safety factor γ_{M2}. They were determined as the ratio of nominal resistance $A_{net} \cdot f_u$ to the real possible resistance $N_R = f(M, G, U)$. The nominal resistances are calculated from the material curves in Fig. 3.25. The highest 0.1184 % of partial safety factors were discarded.

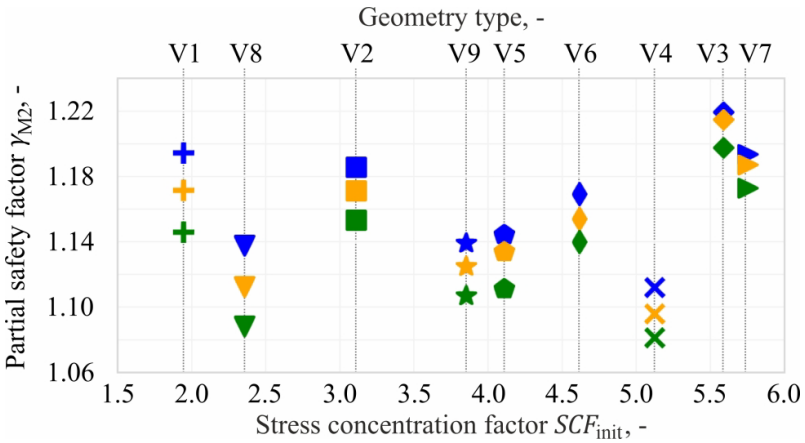


Figure 3.29: *Partial safety factors for net section failure for each steel grade and geometry type*

The maximum partial safety factor $\gamma_{M2} = 1.22$ for geometry type V3. This confirms the Eurocode $\gamma_{M2} = 1.25$ and is consistent with recommended $\gamma_{M2} = 1.23$ from the literature [85, 86].

3.6 Plastic strain limit for numerical design calculation

Numerical design calculations should directly provide design resistance when nominal geometry and nominal material properties are used as an input. Different solid elements (20-node and 8-node), shell elements (8-node and 4-node) and mesh densities ($1.0 \times 1.0t_n$, $0.5 \times 0.5t_n$, $0.25 \times 0.25t_n$, $0.2 \times 0.2t_n$, where t_n is the nominal thickness) were tested. Material nonlinear analysis was performed (small deflection theory).

Section 3.5.1 found the reliable design resistance from a representative set of numerical-analytical simulations (see the right side of Fig. 3.4), these numerical design calculations allow finding the reliable plastic strain limit ε_{lim} (see the left side of Fig. 3.4). Note that in some cases, the numerical design calculation did not reach the reliable design resistance $N_{R,d}$. In such instances, a conversion based on the deformation energy ratio [98] was applied; see steel grade S235 in Fig. 3.30. Such cases are not critical, i.e. causing minimal plastic strain limit, and will not be further discussed here.

Numerical design calculation with the reliable plastic strain limit ε_{lim} provides at most the design resistance $N_{R,d}$ determined by Monte Carlo simulations. The reliable plastic strain limit is suggested in a form of multiplication of ultimate strain ε_u; see Eq. (2.20). The ultimate strains are $\varepsilon_u = 20.8$ %, 16.5 %, and 8.9 % for steel grades S 235, S 355, and S 460, respectively. A reduction factor $\gamma_X = \varepsilon_{lim}/\varepsilon_u$ is sought.

Fig. 3.30 and 3.31 show that the element type nor the mesh density are relevant for the reliable reduction factor γ_X; see also Fig. 2.14. Surprisingly, the geometry types with low initial stress concentration factor determined the smallest γ_X. This is due to high thickness tolerances and critical material models with $f_u/f_y = 1.1$ ratios that cause low design resistance $N_{R,d}$, and at the same time the strains in numerical design calculation being evenly distributed across the net section.

The recommended plastic strain limit for numerical design calculation is:

$$\varepsilon_{lim} = \gamma_X \cdot \varepsilon_u = 0.25 \cdot \varepsilon_u \tag{3.5}$$

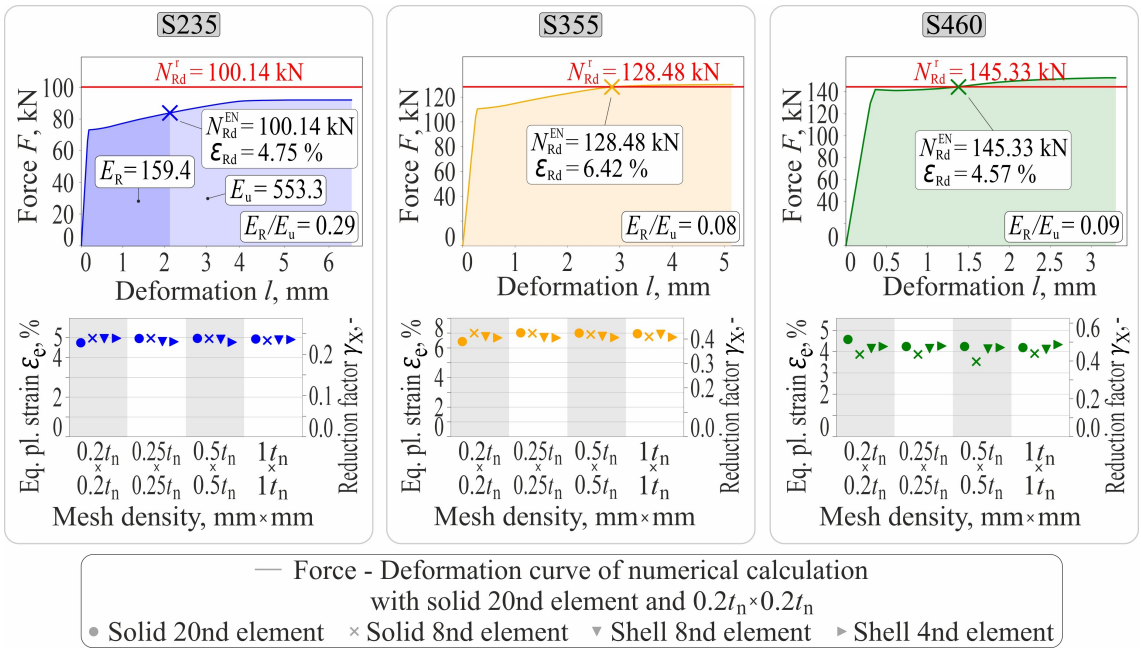


Figure 3.30: *Plastic strain limit and reduction factor γ_X for geometry type V1*

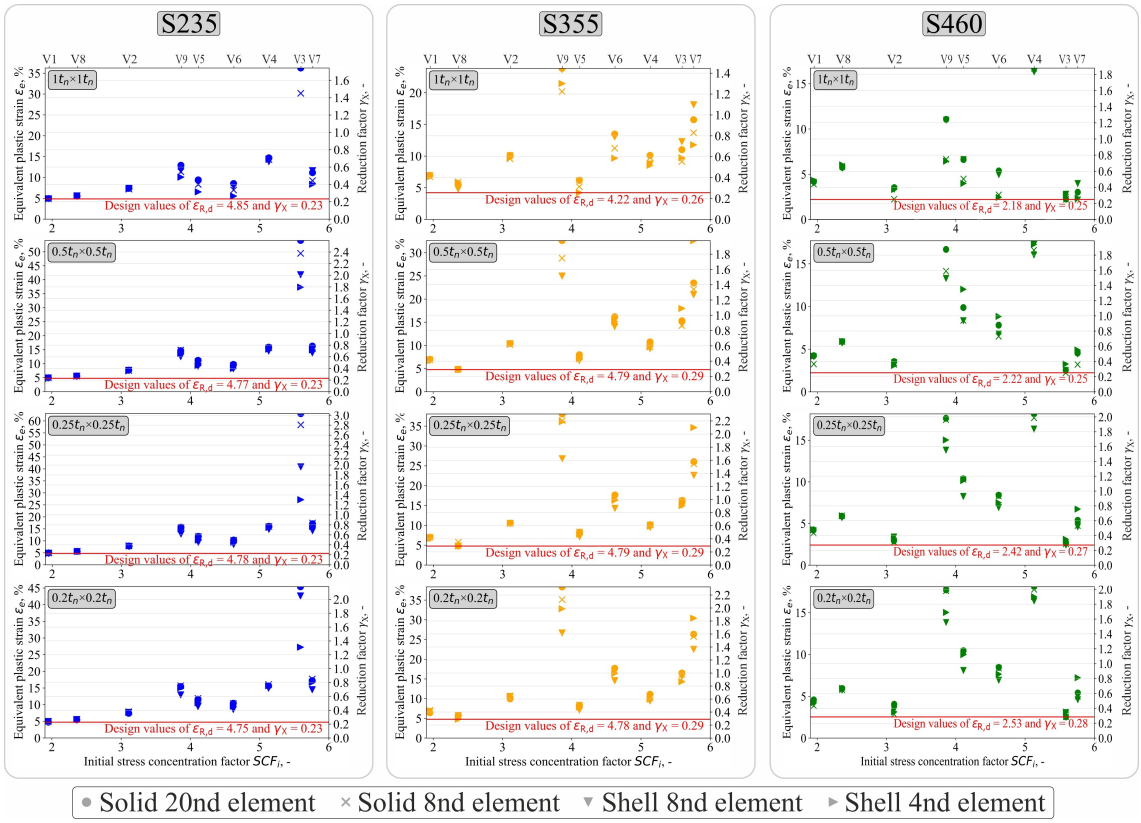


Figure 3.31: *Plastic strain limits ε_{lim} and reduction factors γ_X for all geometries, element types, and mesh densities*

The numerical design calculations with this plastic strain limit provide in the vast majority of samples smaller design resistance than the reliable design resistance $N_{\mathrm{R,d}}$. The experimental load resistances of specimens failing in net section [90, 93, 85, 94, 96, 97] were plotted against the numerical design calculations with the proposed equivalent plastic strain limit according to Eq. (3.5) in Fig. 3.32.

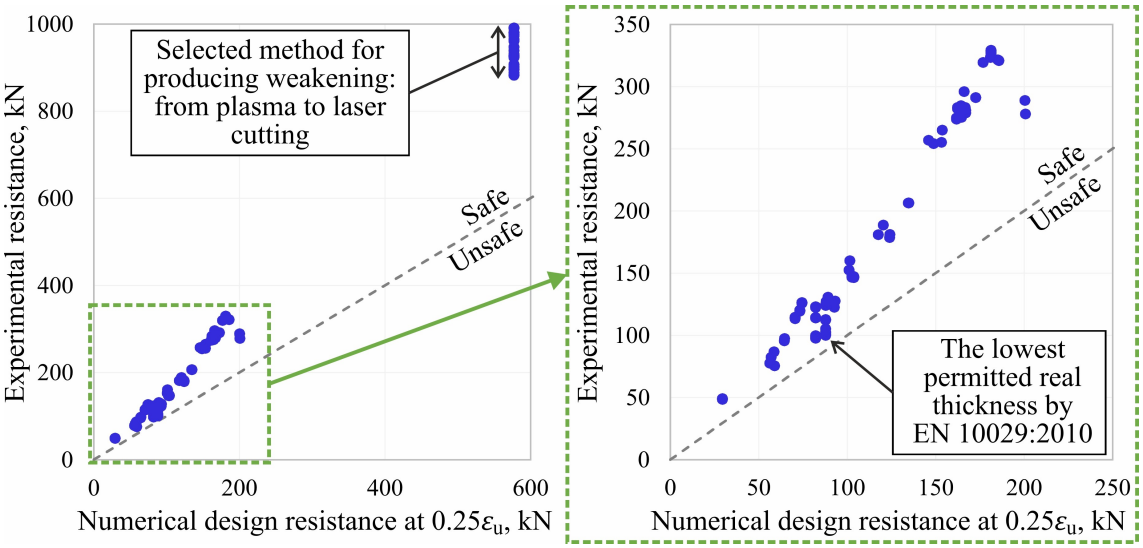


Figure 3.32: *Reliability graph: comparison of numerical design calculations with the proposed plastic strain limit ε_{lim} with the experiments*

This proposed equivalent plastic strain limit must precede the values of fracture strain for structural steel from the available literature, e.g. [99]. This requirement is satisfied.

3.7 Outlook

The proposed equivalent plastic strain limit is simple to use and provides similar values to 5 % for steel grades S 235, S 275, and S 355. The equivalent plastic strain limit will gradually decrease for higher steel grades – decreasing $f_{u,nom}/f_{y,nom}$. The author and colleagues from CTU will attempt to push this proposal into FprEN 1993-1-14 [5] and FprEN 1993-1-8 [28].

In next months, the BUT experiments will be included in the experimental database of net section failures and numerical design calculation models will be analyzed. The reliability of numerical design calculations with the proposed plastic strain limit for all specimens will be determined using the procedure in Section 2.6.

Furthermore, numerical simulations will be built in Abaqus [100] for the BUT specimens to find the plastic strain at fracture initiation. This plastic strain must be higher than the suggested plastic strain limit for the numerical design calculation and will illustrate the safety of the proposal.

Chapter 4

Application

In this chapter, the complete workflow is illustrated using the example of a column web panel in transverse compression. The current design methods, along with possible improvements, were presented in Section 2.1, followed by experimental research by the author and others in Section 2.2. The numerical model is described, and the process of a dataset creation is shown. The performance of the neural network with various settings is investigated. Finally, the reliabilities of code formulas, the new proposed method, and machine-learned predictions are evaluated.

4.1 Numerical model

A numerical model in IDEA StatiCa Connection was used because through API, thousands of models can be generated and solved within a reasonable time. The trade-off is the model precision.

The in-house-made four-node shell finite elements [101] are used for all plates. The rolled I-section is made of three plates, where the nodes of flange are directly connected to the nodes of the web. The number of elements dividing the flange along the cross-section therefore has to be even. The meshing follows the default settings of IDEA StatiCa Connection. The biggest member cross-section size h is divided between 12 elements; see Fig. 4.1. The biggest side of plates is divided into 20 elements. The stiffeners are exempt from this rule for plates and instead follow the meshing rule for cross-sections. The minimum element size is set to 8 mm and maximum size to 50 mm. The software attempts to create elements as close to a square shape as possible.

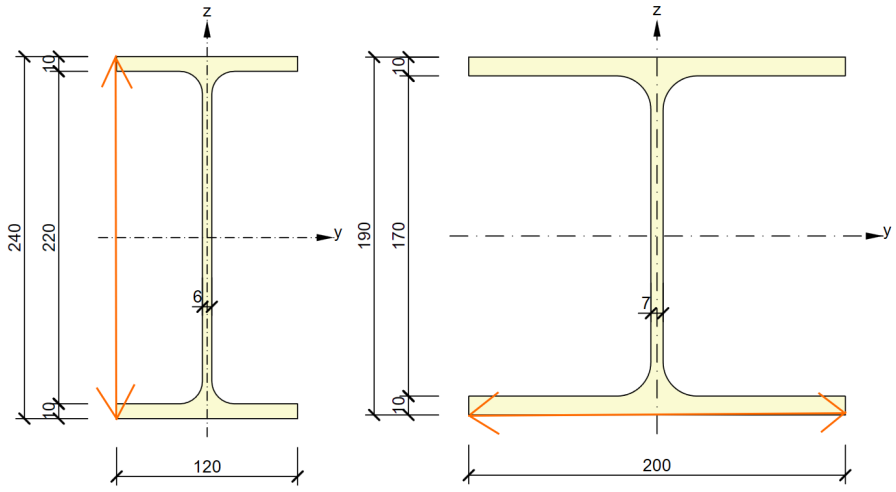


Figure 4.1: *Biggest member cross-section size h is height or width, whichever is bigger*

The model in IDEA StatiCa is created by a series of manufacturing operations; see Fig. 4.2. The top and bottom blocks are solid steel with a cross-section of 200×200 and a length of $1.25 \times h$ of shell elements plus $4 \times h$ as condensed beam element. These blocks are intended to be as rigid as possible. Between the blocks and the I-section, spreading plates are placed with the cross-section

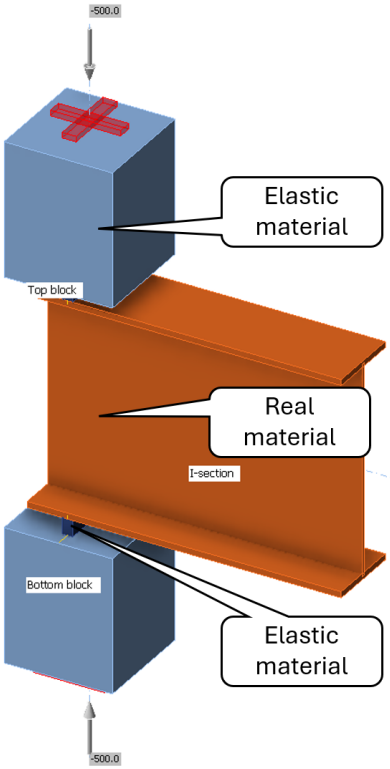


Figure 4.2: *Model in IDEA StatiCa Connection*

height of 30 mm and width equal to the cross-section width of the I-section. The thickness of these plates is one of the examined variables. Both blocks and the spreading plates are modeled with elastic material ($f_y = 4\ 000$ MPa).

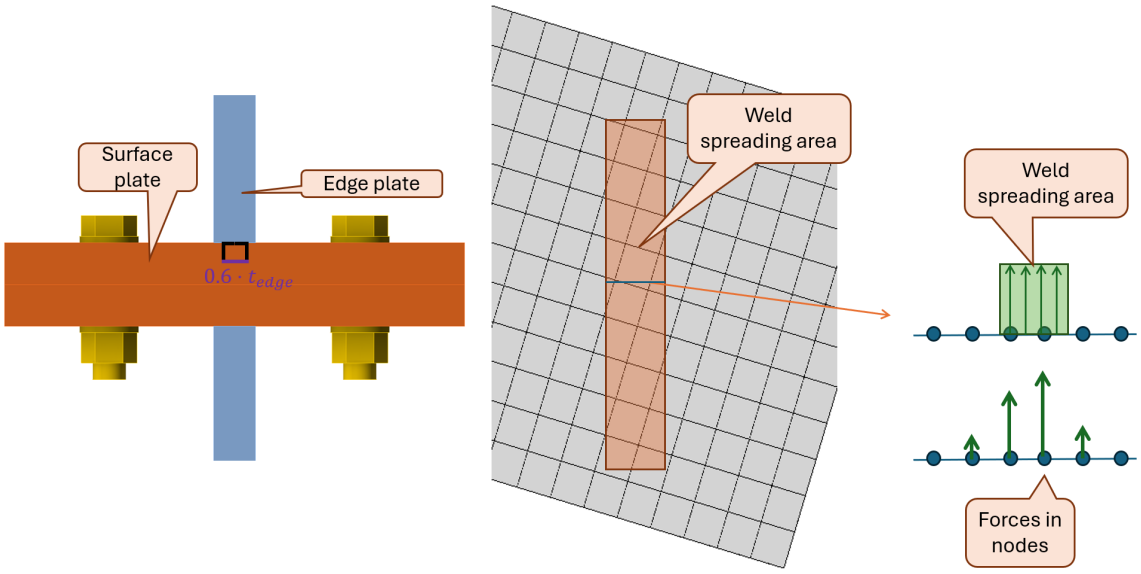


Figure 4.3: *The model of butt welds in IDEA StatiCa Connection*

The transition between the blocks, spreading plate, and I-section is modeled using butt welds, represented as links between the nodes with weights proportional to the weld spreading area, calculated as $0.6 \cdot t_p$, where t_p is the edge plate thickness; see Fig. 4.3.

For each model, materially nonlinear analysis (MNA) and linear buckling analysis (LBA) is solved; see Fig. 4.4. The failure criteria in this dataset do not follow the IDEA StatiCa defaults of 5% plastic strain limit for structural steel and 1% for high-strength steel. Instead, the conclusions

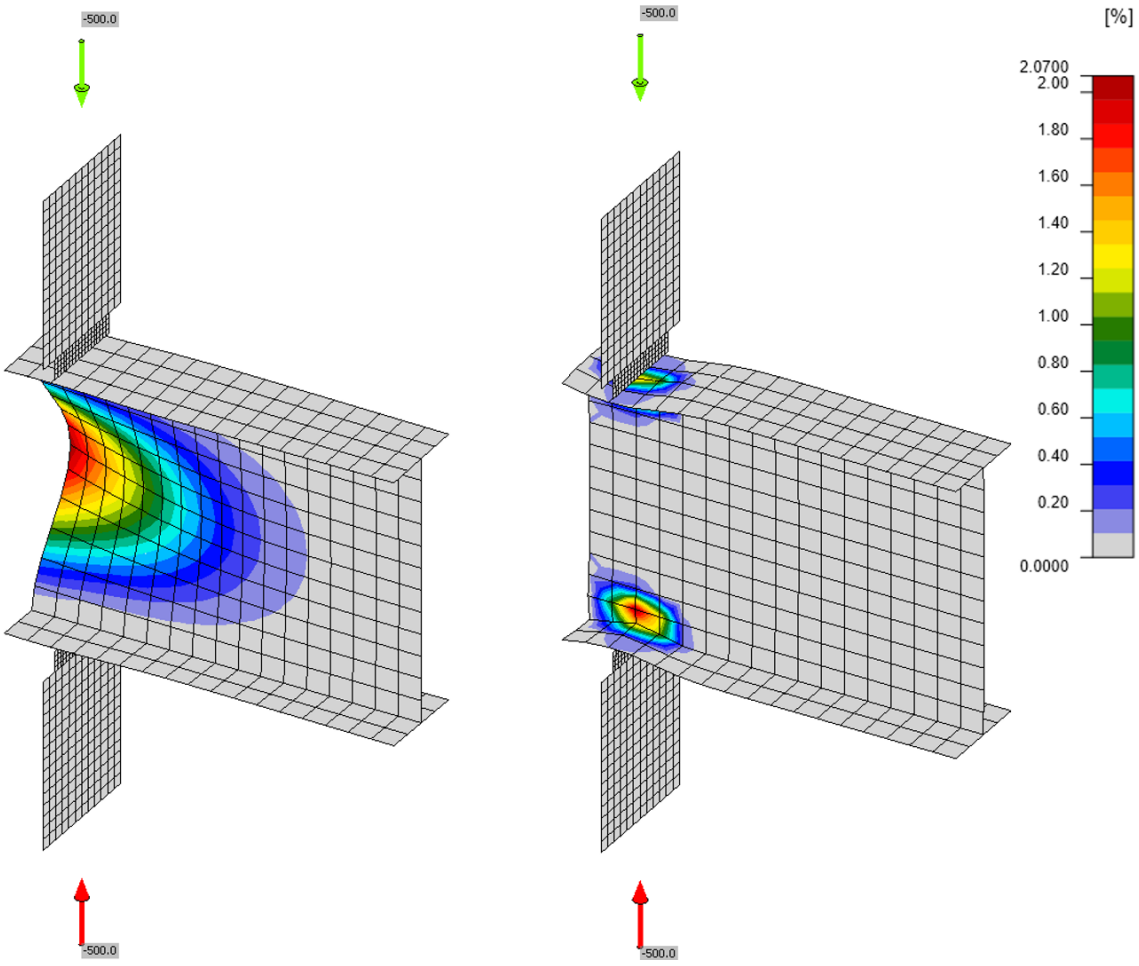


Figure 4.4: *Buckling shape (left) and plastic strain (right)*

of current research into plastic strain limit shown in Chapter 3 are used, i.e., the plastic strain limit is calculated according to Eq. (2.21).

The Stop at limit strain function is utilized, which means the load between the converging step (plastic strain is below the set limit) and the diverging step (plastic strain is above the set limit) is halved eight times. The achieved plastic strain at the last converged step is below but may still be quite far away from the set plastic strain limit, e.g. $\varepsilon_{pl} = 3.52\ \% < \varepsilon_{lim} = 4.13\%$. Furthermore, it should be noted that the scale in the software's 3D scene shows values averaged in nodes, which are typically even smaller. The values presented in the Check tables show values averaged in elements that are more suitable for strain check.

Mesh sensitivity has already been observed by Pesek [37]. IDEA StatiCa model of the web panel in transverse compression is strongly dependent on mesh size; see Fig. 4.5 for an example of IPE 300 cross-section made of S 355 steel grade.

Web height was divided into 6, 8, 12, 16, and 20 elements. The highest resistance is achieved with the coarsest mesh and the lowest resistance with the finest mesh. The theoretical plastic resistance at infinite mesh determined by MNA is $F_{pl,Rd} = 224$ kN. Accepting the 5 % error, $1.05 \cdot 224 = 235$ kN is the qualified plastic resistance. However, even the finest tested mesh with 20 elements per cross-section height reaches 262 kN. The theoretical critical force at infinite mesh determined by LBA is $F_{cr} = \alpha_{cr} \cdot F_{pl,Rd} = 459$ kN. Accepting the 5 % error, $1.05 \cdot 459 = 482$ kN is the qualified critical force. Again, even the finest tested mesh with 20 elements per cross-section height reaches 501 kN.

The default value of 12 elements per cross-section height causes an error of 33 % in plastic resistance and 12 % in critical force compared to the infinite mesh. On the other hand, the model is simplified, the fillet radius is not modeled, nor does the fillet weld add any stiffness to the force distribution. These simplifications decrease the resistance against transverse compression. The

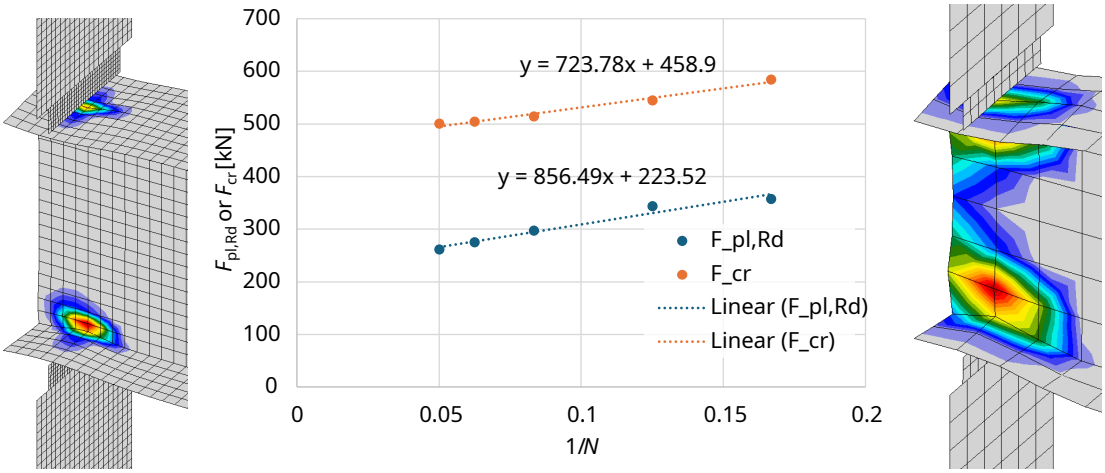


Figure 4.5: *Mesh sensitivity for a web panel in transverse compression model – plastic resistance $F_{pl,Rd}$ and critical force F_{cr}*

default value of 12 elements per cross-section height was found to be a good balance for hot-rolled cross-section and this value is used in the dataset creation.

4.2 Dataset creation

The dataset is created as follows:

- The whole range of IPE, HEA, HEB (European), and W (USA) sections is used totalling 432 cross-sections. This approach was chosen to maintain the realistic ratios of cross-section dimensions; see Fig. 4.6.

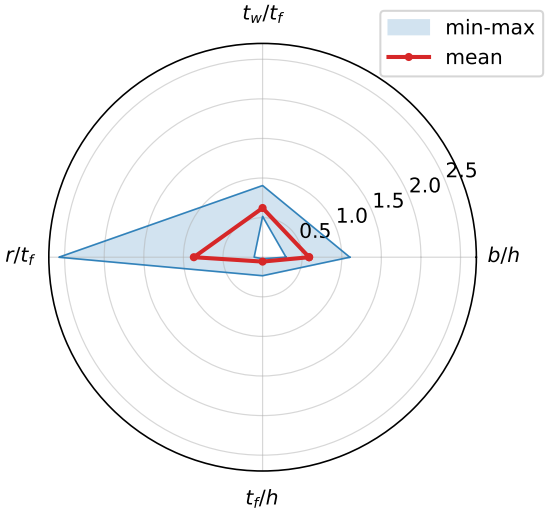


Figure 4.6: *Histogram of cross-section geometry ratios*

- European steel grades were chosen in this study; see Tab. 4.1.

- Relative distance from the unstiffened end $x = e/h$ is generated in a log-spaced grid between 0.05 and 1. Log-spaced grid is chosen because it was observed that the load resistance changes rapidly with a small change near the unstiffened end but is stabilized with increasing distance x. Therefore, it is beneficial to have a denser population of data points in the region of large load resistance gradients – with small x near the unstiffened end.

Table 4.1: *Steel grades and corresponding mechanical properties*

Steel Grade	f_y [MPa]	f_u [MPa]	ε_u	ε_{lim}
S 235	235	360	20.8%	5.21%
S 275	275	430	21.6%	5.41%
S 355	355	490	16.5%	4.13%
S 450	440	550	12.0%	3.00%
S 420 MH/MLH	420	520	11.5%	2.88%
S 460 MH/MLH	460	540	8.9%	2.22%
S 620 Q/QL/QL1	620	700	6.9%	1.71%
S 690 Q/QL/QL1	690	770	6.2%	1.56%
S 500 MC	500	550	6.0%	1.50%
S 550 MC	550	600	6.0%	1.50%
S 650 MC	600	650	6.0%	1.50%
S 700 MC (tle8 mm)	700	750	6.0%	1.50%
S 700 MC (t>8 mm)	680	750	6.0%	1.50%

- Relative spreading plate thickness $a = t_p/t_f$ is uniformly randomly distributed between $a \in [1,\ 5]$. In this case, it is beneficial to have all values evenly populated.

The dataset was selected to directly provide the design load resistance when using nominal material and geometric properties. It should be noted that the choice of properties would differ if the objective were to generate millions of simulations for a Monte Carlo analysis aimed at a reliability-based design approach. In such a case, it would be beneficial to treat yield strength and ultimate strength as random variables interlinked with each other, e.g. $f_u \in [1.1\, f_y,\ 1.6\, f_y]$, because a uniform distribution is preferable to a step-wise distribution of feature values.

The whole dataset is shown in histograms in Fig. 4.7. The variations in beam sizes can be seen. The small number of datapoints in the regions of very thick flanges ($t_f > 80$ mm) and very thick webs ($t_w > 55$ mm) will likely cause issues with predictions of neural network. It is useful to filter out such datapoints and reduce the range of validity. The relative end distance x also has a decreasing trend, but that is by design and should not cause any troubles. The outputs – plastic resistance $F_{pl,Rd}$ and especially F_{cr} may reach extremely high values, which cannot directly be used as targets for training a neural network, as such models work best with targets normalized to ranges [-1, 1] and of comparable order of magnitude. Therefore, some form of normalization needs to be devised for these variables. It is strictly needed to scale these outputs down and most likely also necessary to filter out the extremely high values. A viable alternative is using α_{cr} – the multiplier of $F_{pl,Rd}$ to obtain critical force F_{cr}. This value is basically already scaled by $F_{pl,Rd}$. Another option is to scale $F_{pl,Rd}$ against some cross-sectional property, e.g. the web thickness multiplied by the flange thickness.

The Python code utilizing IDEA StatiCa Connection API [47] that was used to model and calculate the dataset is published at a Github repository [102] along with the generated dataset. In total, 9 241 models were solved and the load resistance F_{pl} at strain limit as determined by MNA, critical force F_{cr} and α_{cr} as determined by LBA were noted.

The main limitations of this dataset are as follows:

- Buckling resistance is not provided and must be further calculated by analytical equations from the results of MNA and LBA
- Shell models are utilized, the effect of the fillet between the web and the flange is totally neglected.

Range of validity is determined by the dataset. Each feature is divided into 30 bins, and bins containing fewer than 5 samples, along with all subsequent bins, are removed. Sparsely populated regions of the feature space would likely lead to poor predictions and degrade the performance of the training process. The remaining dataset defines the range of validity shown in Tab. 4.2.

The range of validity of geometrical features may be exceeded provided the ratios plotted in Fig. 4.6 are maintained if the Similitude scaler is utilized.

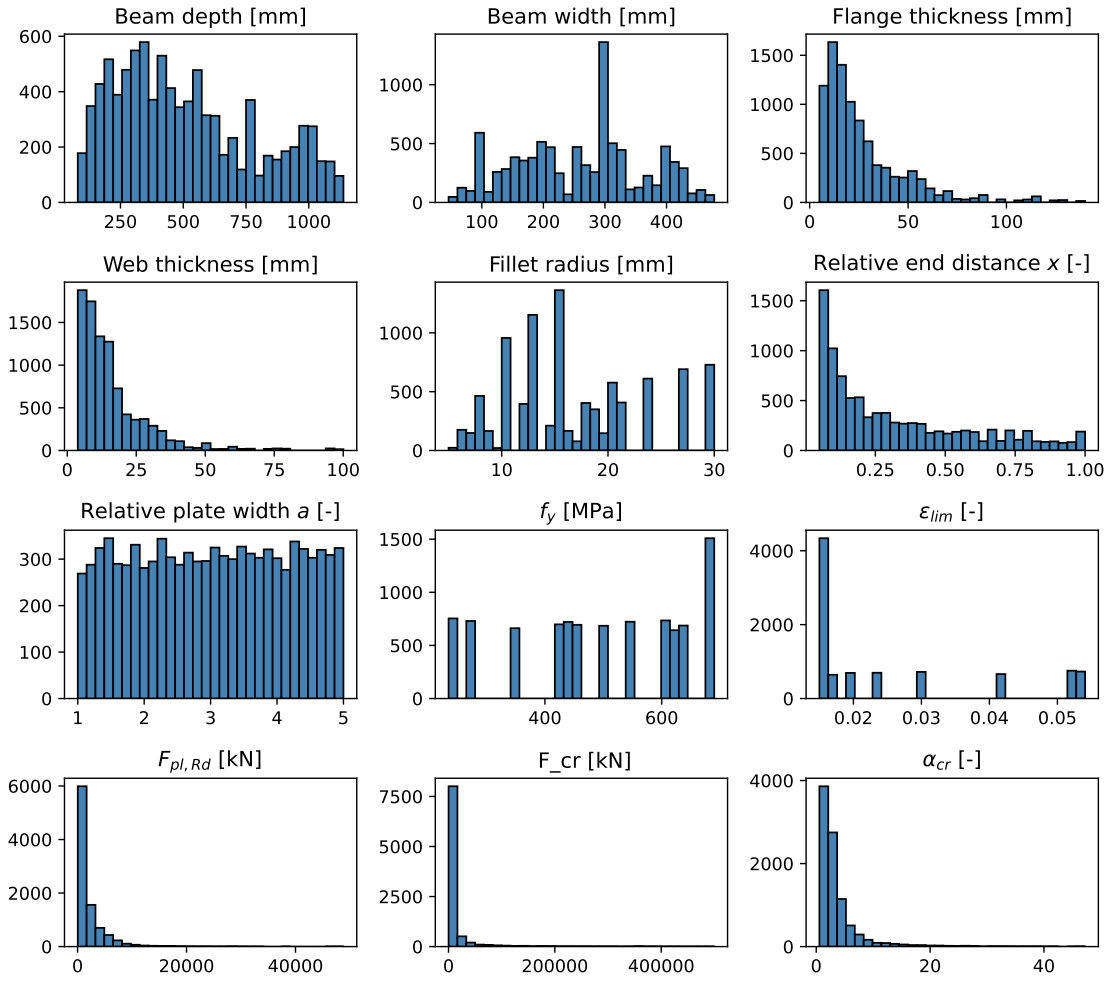


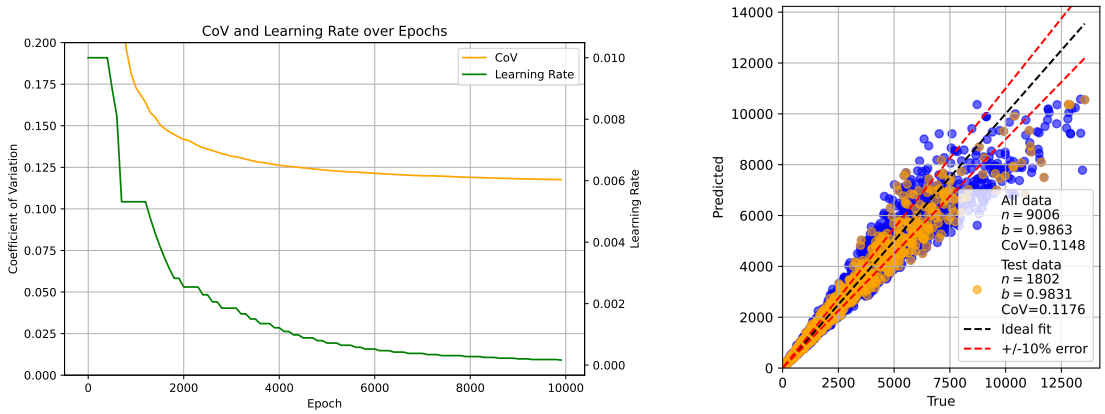
Figure 4.7: *Histogram of all features with outputs*

Table 4.2: *Minimum and maximum values of input and output features after filtering*

Feature	Minimum	Maximum
Depth h [mm]	80.0	1137.9
Width b [mm]	46.0	447.0
Flange thickness t_f [mm]	4.9	89.9
Web thickness t_w [mm]	3.8	51.3
Fillet radius r_w [mm]	5.0	30.0
Relative end distance x	0.05	1.0
Relative plate width a	1.0	5.0
Yield strength f_y	235	690
Plastic strain limit ε_{lim}	0.015	0.054
Load resistance $F_{\text{pl,Rd}}$ [kN]	34	13549

4.3 Dataset scaling

Dataset scaling may be difficult to grasp only by engineering judgment. How could the predictions be better when only different units (e.g. MN instead of kN for the output of load resistance) are used? The neural networks are numerical optimizers; therefore, all inputs should have the same scale and outputs should be in a reasonable range. Large values lead to large gradients and therefore unstable training; small values lead to slow training. The benchmark case without any scaler is shown in Fig. 4.8. The predictions of this model are terrible clearly indicating that scalers are necessary and greatly improve model performance.



(a) *CoV and learning rate progress* (b) *Reliability graph*

Figure 4.8: *Model performance without any scaler*

Similitude scaling is explained in Section 2.5.2. Its usage is shown in the following graph in Fig. 4.9 where cross-sectional dimensions are varying and the outputs in the form of yield resistance and critical force can be predicted by scaling. The dimensions of cross-section – beam depth h, beam width b, flange thickness t_f, and web thickness t_w are all multiplied by a scaling factor. The plastic resistance $F_{pl,Rd}$ and critical force F_{cr} are multiplied by the scaling factor on the power of two.

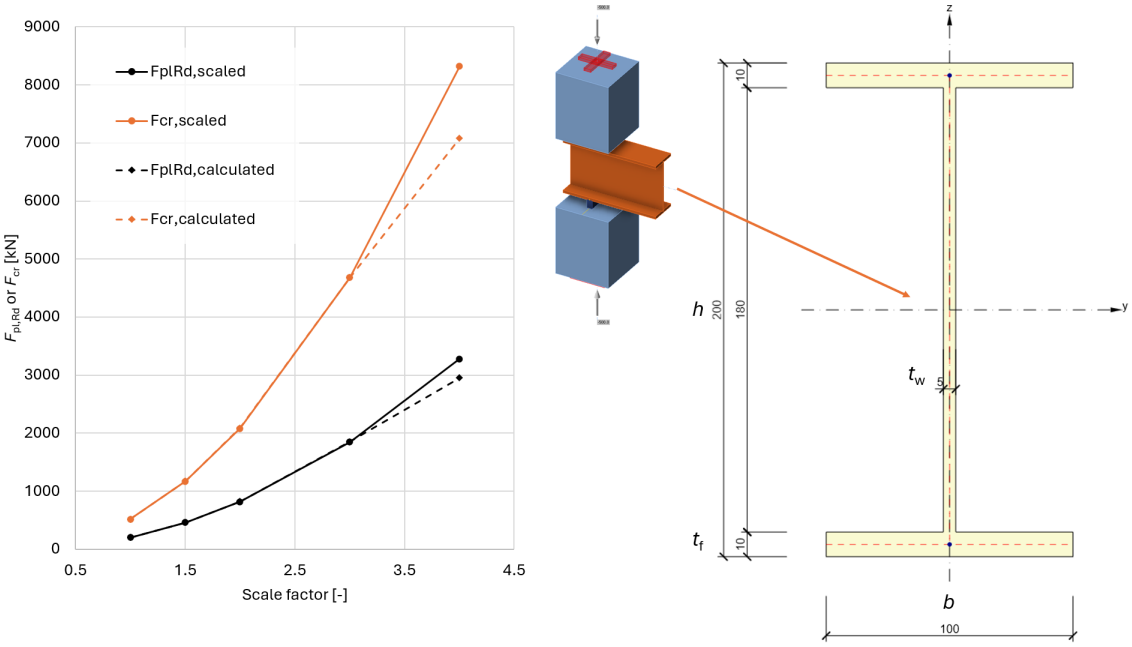


Figure 4.9: *Scaling dimensions perfectly proportionally allows prediction of plastic resistance and critical force*

It can be observed that for the factor up to 3, the scaling prediction fits the finite element analysis nearly perfectly. However, for larger factor, the prediction and calculation diverge. The reasons why similitude scaling may not work well are as follows:

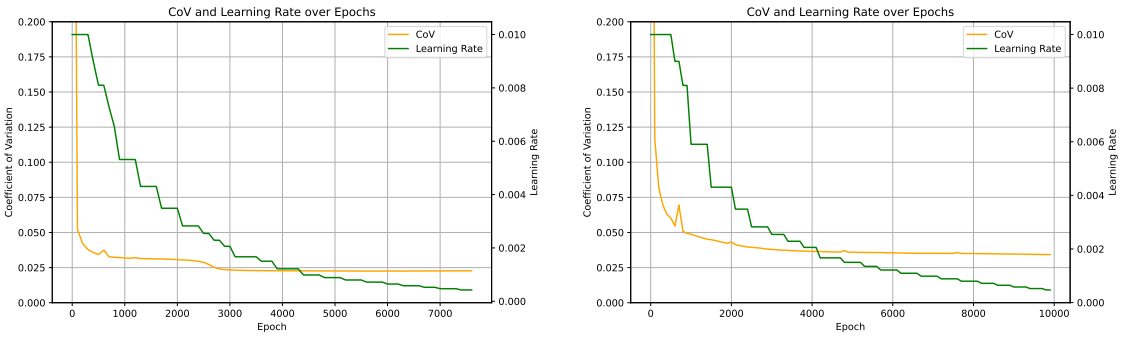
- Yield strength of some grades is thickness-dependent – thicker plates may be assigned a lower yield strength. For some standards, such as EN 10025 [55], the yield strength for thickness above 40 mm is reduced. For other standards, such as EN 10149 [103], the yield strength is reduced for thickness above 8 mm [12].
- Not all elements in the model are scaled – in the model for the column web in transverse

compression (see Fig. 4.2), the top and bottom stubs are kept constant.

- Meshing rules include the minimum and maximum finite element size. Very small sections ($h < 12 \cdot 8 = 96$ mm) consist of smaller number of finite elements. Very large sections ($h > 12 \cdot 50 = 600$ mm) consist of more finite elements. This is what causes the divergence in the example in Fig. 4.9.

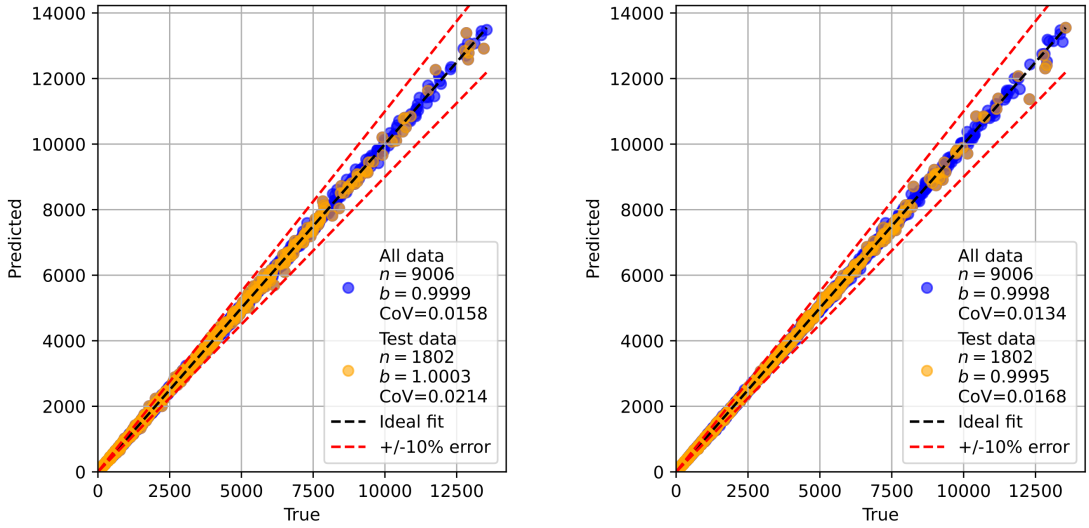
Inputs are further scaled using `StandardScaler` from *scikit-learn* [74].

Outputs are scaled by dividing each value by the maximum output and multiplying by 10. This ensures that the maximum output is 10 and the minimum remains greater than 0. This may seem as a useless step but the increase in model performance can be clearly seen when the output scaler is applied. Using the output scaler leads to $b = 0.9992$, CoV $= 0.0228$, Minimal ratio $= 0.843$, and Maximal ratio $= 1.111$. Compare to not using output scaler: $b = 0.9994$, CoV $= 0.0342$, Minimal ratio $= 0.727$, Maximal ratio $= 1.166$. The convergence is also much slower; see Fig. 4.10.



(a) *With output scaling* (b) *Without output scaling*

Figure 4.10: *Comparison of learning rate and CoV over epochs with and without output scaling*



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.11: *Comparison of the Similitude scaler*

The Similitude scaler actually makes the predictions worse, as seen in Fig. 4.11, where input and output scalers are both applied with and without the Similitude scaler. This is no surprise for the reasons stated above – the dataset contains both very small and very large members and the meshing difference and yield strength reduction cause issues. If this was taken into account in the dataset creation by removing the minimum and maximum element sizes, this issue would be

mitigated. This is the tradeoff for an increased range of validity. If these issues were dealt with, the similitude scaler might actually prove useful for datasets with a larger number of parameters, as it effectively reduces the dimensionality of the input domain by one. However, in the dataset considered here, it didn't seem necessary.

4.4 Dataset size

A key question is the minimum dataset size required to achieve sufficiently accurate predictions. Figure 4.12 illustrates the effect of varying the training set size by adjusting the train/validation split; for example, 20 % is used for training, 80 % for testing, then 30 % and 70 %, and so on. Across all training sizes, the model without the Similitude scaler (but using standard input scaling and custom output scaling) consistently outperformed the model with the Similitude scaler, both in terms of CoV and maximum prediction error.

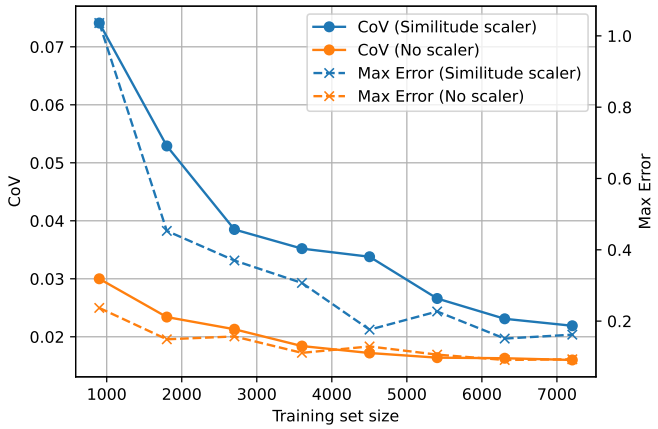


Figure 4.12: *CoV and maximum error of DNN with and without the Similitude scaler for varying training set size*

Increasing the number of cases always improves the prediction quality, but it converges to a limit value, and adding additional samples is less and less effective.

4.4.1 Active learning

Urben, in his master's thesis [104] prepared within the MADESCO project investigated **active learning** (AL) in the analysis of steel connections using DNNs. AL means strategically increasing the dataset in locations of unsatisfactory model performance. The main issue is recognizing the areas in the multidimensional space of bad predictions without the knowledge of ground truth (calculated models).

Several learning functions were implemented and evaluated:

- Random Sampling: Serves as a baseline, selecting data points without any heuristic.
- Error-Based Acquisition: Prioritizes samples with high relative prediction error, effective when ground truth is available.
- Latin Hypercube Sampling (LHS): Ensures uniform coverage of the input space, independent of model predictions.
- Cluster-Based Sampling: Uses unsupervised clustering to select representative samples from distinct regions.
- Uncertainty-Based Acquisition: Employs Monte Carlo Dropout (MCD) to estimate epistemic uncertainty, guiding sample selection without requiring ground truth.

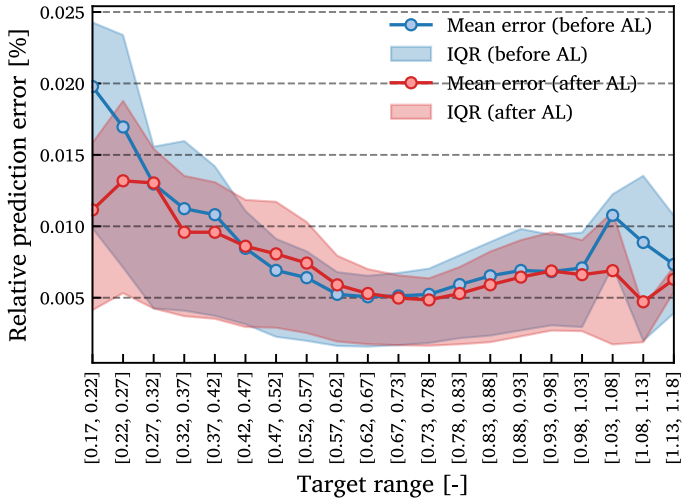


Figure 4.13: *Active learning – adding more training data into regions of bad prediction accuracy [104]*

Empirical evaluations revealed that error-based and uncertainty-based strategies yielded the most sustainable improvements in model performance. However, error-based acquisition is not applicable when ground truth is unavailable, which limits its practical use.

Two complementary AL frameworks were developed:

- Local Active Learning (LAL): Targets existing designs with poor predictive performance. For each design, new load combinations are selected using a multi-objective function that balances exploration (distance from existing loads) and exploitation (proximity to high-uncertainty regions). These combinations are evaluated via FEM and added to the training set.
- Global Active Learning (GAL): Explores the broader design space by generating novel connection designs. A feasible hypercubemap (FHCM) is constructed to partition the design space into feasible and infeasible regions. Samples are drawn from feasible hypercubes, clipped to match standardized profiles, and evaluated using the DNN and MCD uncertainty. High-uncertainty samples are selected for FEM evaluation and dataset augmentation.

Active learning, supported by uncertainty quantification and automated FEM integration, enables efficient and targeted dataset expansion for structural engineering applications. The proposed frameworks demonstrate the feasibility of autonomous, data-driven model refinement, laying the groundwork for intelligent design tools in steel connection engineering. The areas of bad predictions can be recognized and further samples added to such areas; see Fig. 4.13.

However, the full loop that would assign the parameters of CBFEM model has not yet been developed. Active learning remains a promising strategy to reduce the number of time-consuming and costly FEM analyses while achieving the target accuracy, e.g. CoV < 0.02.

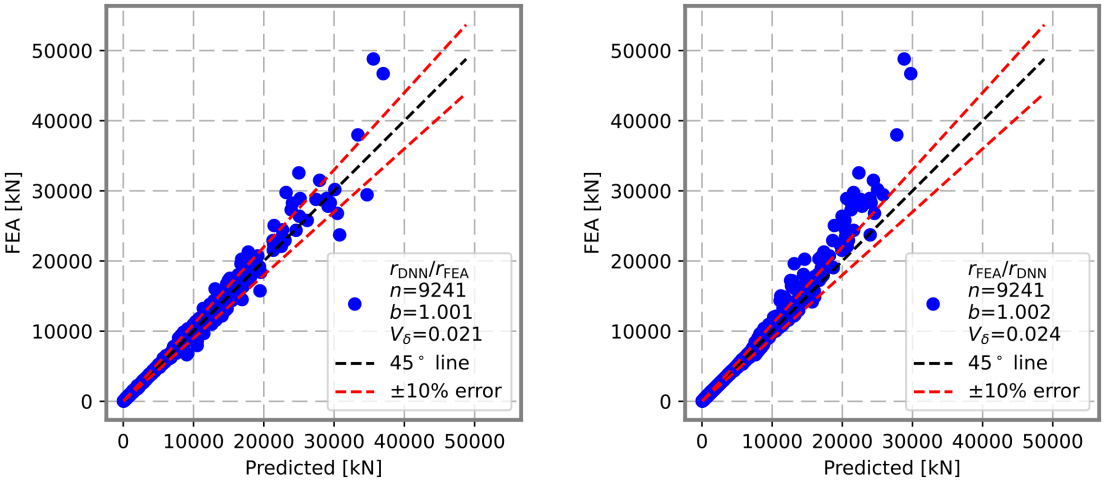
4.5 Predictions of different datasets

The predictions of the current dataset is amazingly precise; the DNN provides nearly the same results as IDEA StatiCa Connection numerical design calculation that was used for the data generation. In the following section, the quality of predictions towards different datasets is investigated. Two evaluation models are used:

- Without the Similitude scaler – performed better for all known data
- With the Similitude scaler – anticipated to also remain reliable outside the geometric range of validity

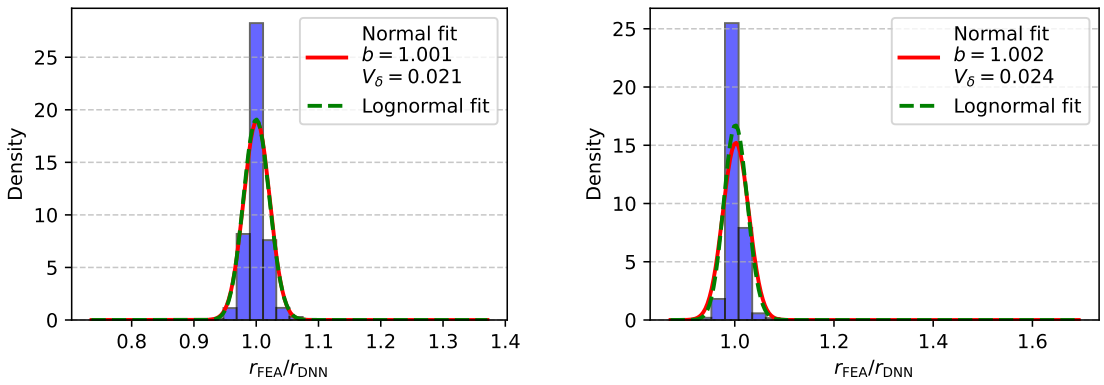
4.5.1 Full dataset without filtering

The aim of this section is to quantify how well the DNN can predict also the unknown data outside the range of validity.



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.14: *Reliability graphs of full dataset – plastic resistance* $F_{pl,Rd}$



(a) *With Similitude scaling* (b) *Without Similitude scaling*

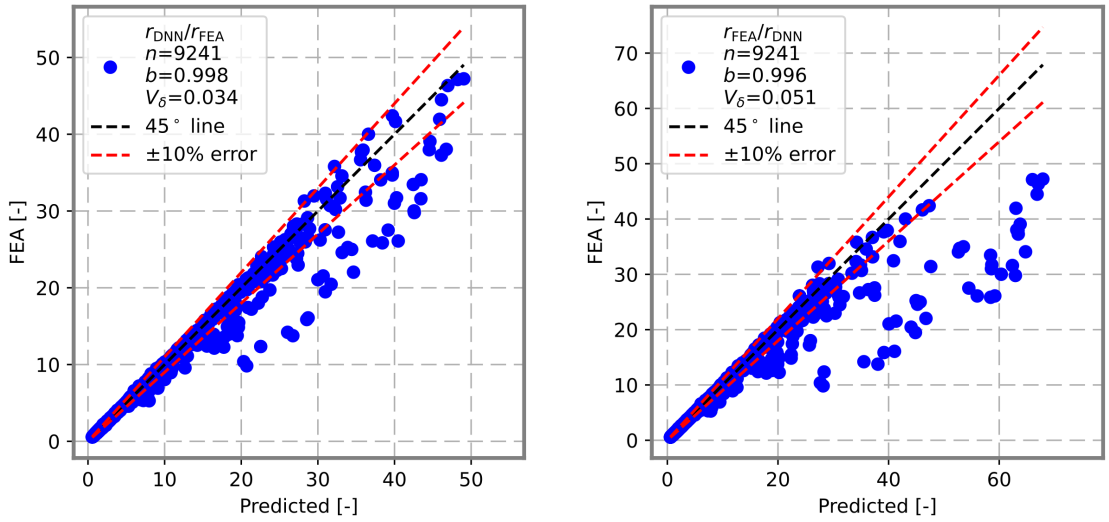
Figure 4.15: *Histogram of plastic resistance* $F_{pl,Rd}$ *predictions*

Fig. 4.14 and 4.15 show that the Similitude scaler may indeed help predicting the samples outside the range of validity (for cut-offs, see Tab. 4.2). However, the unknown ratios, e.g. too thick web to beam depth ratio, remain an issue. Without the Similitude scaler, the DNN tends to undervalue the load resistance $F_{pl,Rd}$ for jumbo cross-sections. By including all data, the Similitude scaler becomes more appropriate, with superior variation coefficient V_δ; see Eq. (2.36).

For buckling analysis, the same procedure is done for the predictions of critical buckling factor $\alpha_{cr} = F_{cr}/F_{pl,Rd}$. Predicting α_{cr} instead of F_{cr} is convenient, because the buckling factor is already normalized and the range of values is much smaller; see Fig. 4.7. In the Similitude scaler, it must be remembered that the output in the form of coefficient α_{cr} is no longer scaled.

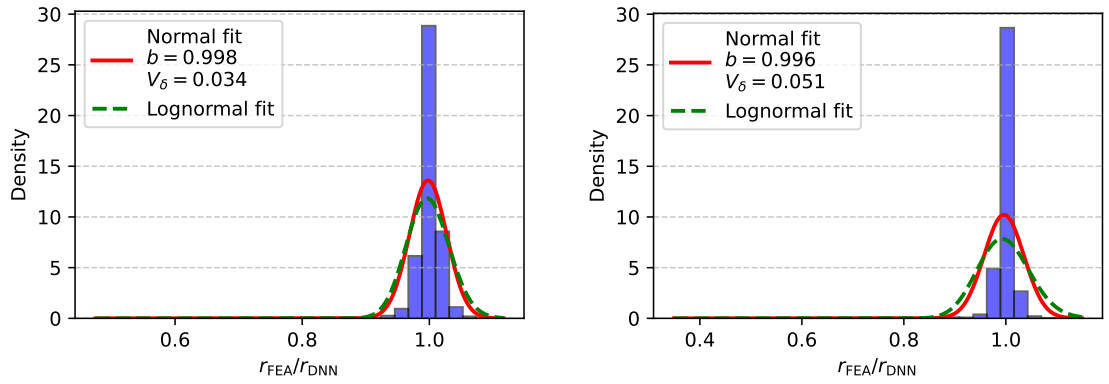
The predictions of the buckling factor are much worse. Both predictions with and without the Similitude scaler overestimate the buckling factor α_{cr} for unknown data, despite the fact that the predictions for training and testing dataset with the data filtering were excellent; see Tab. 4.3.

Overall, the Similitude scaler allows predicting datapoints outside the range of validity better. However, the results were still too far away from the ground truth – finite element analysis (FEA). Predicting outside the range of validity cannot be recommended, although future research may show that going outside the geometry limits while keeping the geometry ratios may be viable with the Similitude scaler.



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.16: *Reliability graphs of full dataset – buckling factor* α_{cr}



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.17: *Histogram of buckling factors* α_{cr} *predictions*

Table 4.3: *Bias, CoV, and the worst predictions for scalers with and without Similitude scaling for filtered dataset*

Scaler	Bias	CoV	Min ratio	Max ratio
Similitude	0.9993	0.0194	0.895	1.120
No similitude	0.9997	0.0159	0.926	1.169

4.5.2 IDEA StatiCa dataset created by nested for loops

The second dataset should not cause any troubles for the DNN. The parameters are within the range of validity created by the same software that was used for models used for training, IDEA StatiCa Connection. In this dataset, six cross-sections are used: IPE 100, IPE 200, ..., IPE 600 and relative distance to the unstiffened end x varies between 0.05 and 1. Otherwise, relative plate width $a = 1$, S 355 steel grade, $\varepsilon_{lim} = 0.05$ are kept constant.

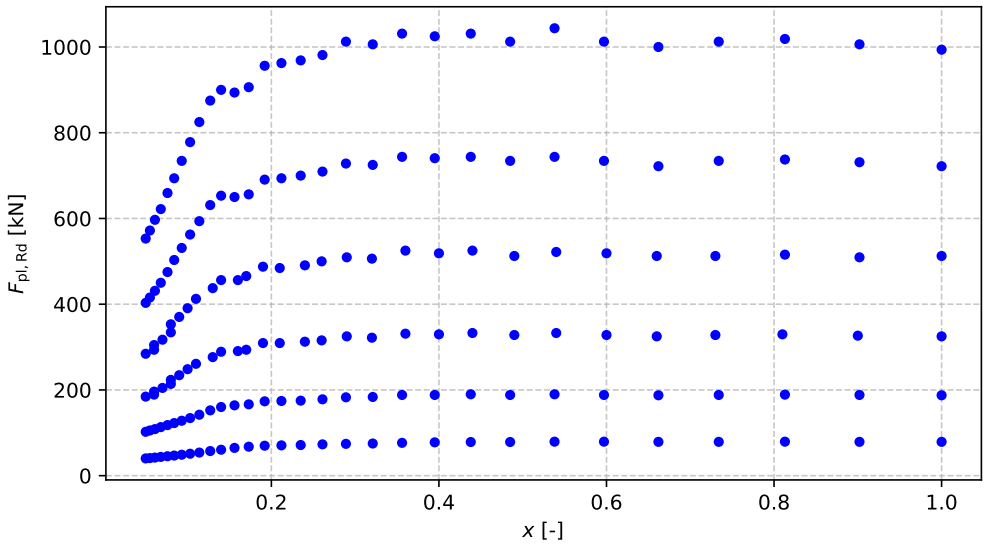
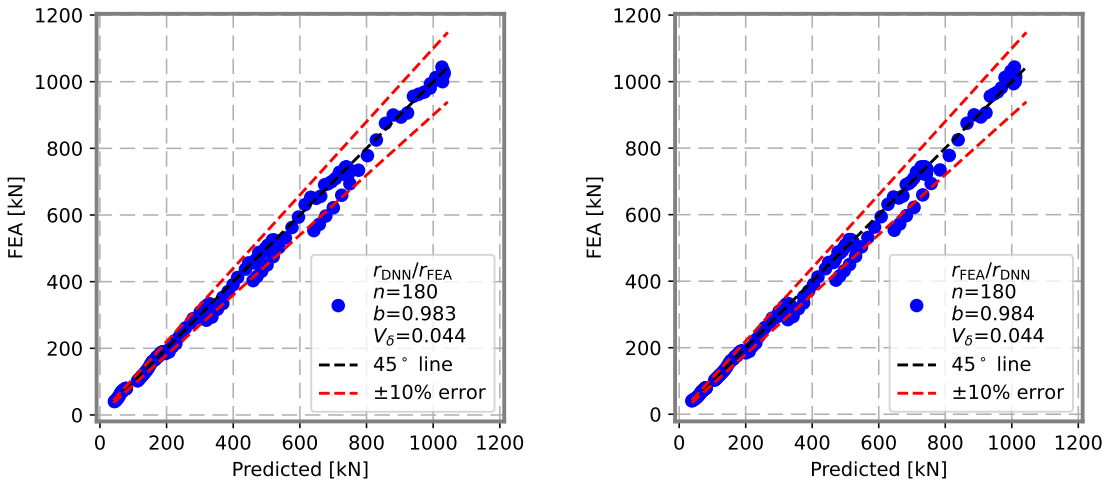


Figure 4.18: *Dataset created by nested for loops*

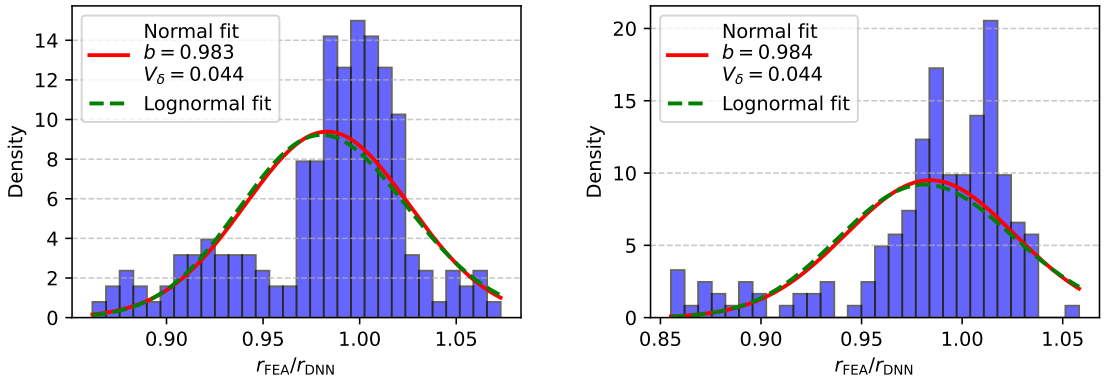


(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.19: *Reliability graphs of dataset created by for loops – plastic resistance by MNA,* $F_{pl,Rd}$

As stated above, such dataset is suitable for human perception and it allows noticing interesting remarks about the numerical design calculation in IDEA StatiCa:

- By varying relative distance x, the plastic resistance slightly jumps up and down although it is stabilized for $x > 0.4$; see Fig. 4.18. This is caused by the number of finite element nodes being active in direct transfer of transverse force; see Fig. 4.3.
- By keeping the same mesh, this fluctuation propagates throughout all cross-section sizes.



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.20: *Histogram of dataset created by for cycles*

The DNN trained on this dataset is actually able to capture this numerical error. That is not desirable and is a sign of overfitting.

- Interestingly, this numerical error increases with increasing cross-section size. The relative errors calculated as $F_{\min}/F_{\max}$ in the range $x = [0.4, 1.0]$ are 99.01%, 98.77%, 97.65%, 97.02%, 97.06%, 95.21% for IPE 100 to IPE 600, respectively.
- Because this numerical error reaches up to 5%, achieving higher levels of accuracy by DNN prediction does not seem feasible. This finding is relevant also for other datasets created by IDEA StatiCa Connection.

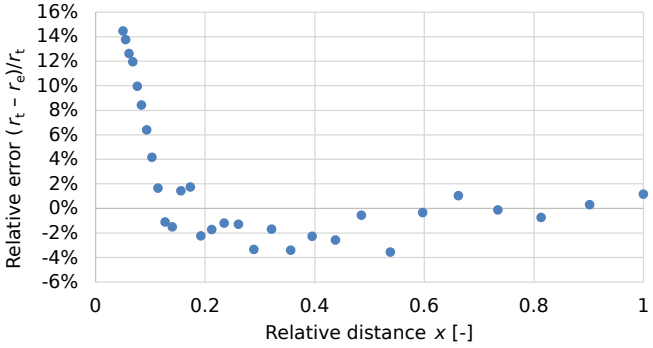


Figure 4.21: *The plot of relative prediction error to relative distance for IPE 600*

Against expectations, Fig. 4.19 and 4.20 show that the DNN has troubles to perfectly imitate the behavior of dataset. Fig. 4.19 shows clear bands of results. The DNN overpredicts the plastic resistance for small relative distance x. The relative errors are the smallest for IPE 200 and IPE 300, but increase up to 14 % for IPE 400, 500, and 600. The relative error of prediction without the Similitude scaler plotted against relative distance x for IPE 600 is shown in Fig. 4.21. There is not much of a difference between the usage of the Similitude scaler.

The training set was intentionally created with high density of samples with small relative distance x exactly to prevent this issue; see Fig. 4.7. It is an unpleasant surprise that the predictions are not capturing the load decrease near the unstiffened end for some cross-sections well enough. On the other hand, the variation coefficient $V_\delta = 0.044$ for both DNNs is small.

4.5.3 Dataset created in ANSYS software

The aim of this section is to investigate the differences between the predictions trained on a crude numerical design calculation in IDEA StatiCa compared to the finely-meshed solid models created

in ANSYS by Ondřej Pešek [37]. Note that the same ANSYS models were used in papers [25, 26, 24, 30] and their responses have been validated on physical experiments.

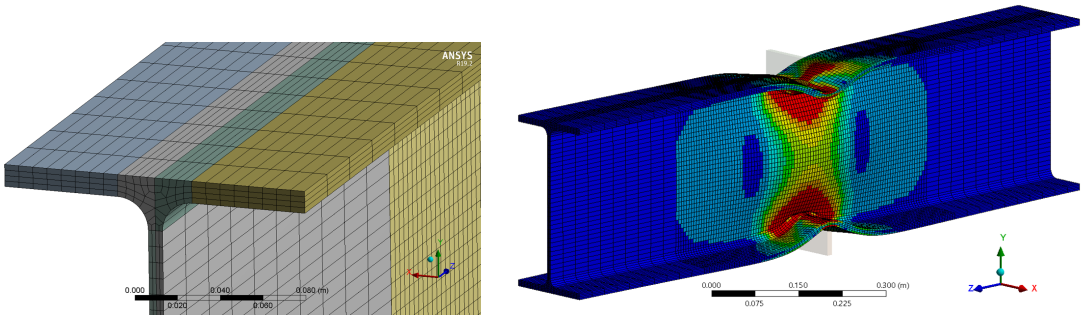


Figure 4.22: *Web and flange is divided between 4 solid elements across their thickness; beam length is equal to* $4 \cdot h$ *[37]*

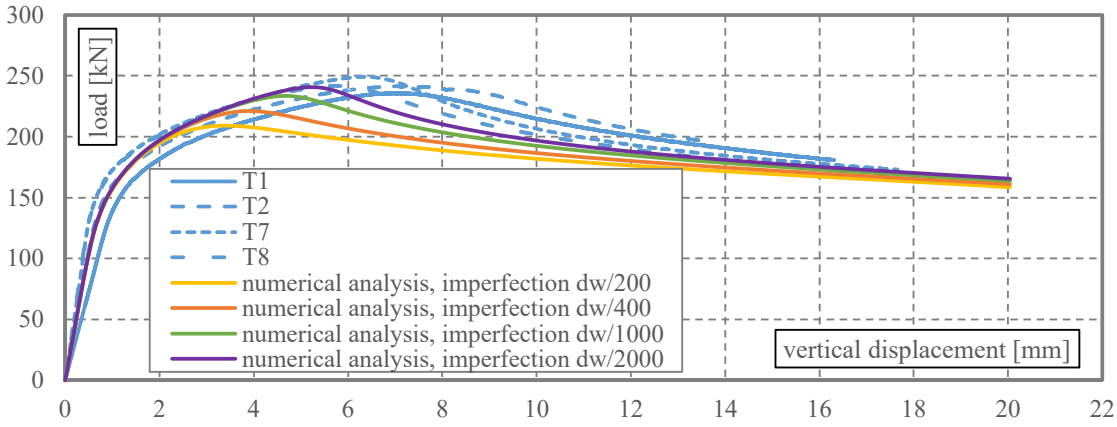


Figure 4.23: *Validation of ANSYS model – transverse load-displacement curves for IPE 200 [24]*

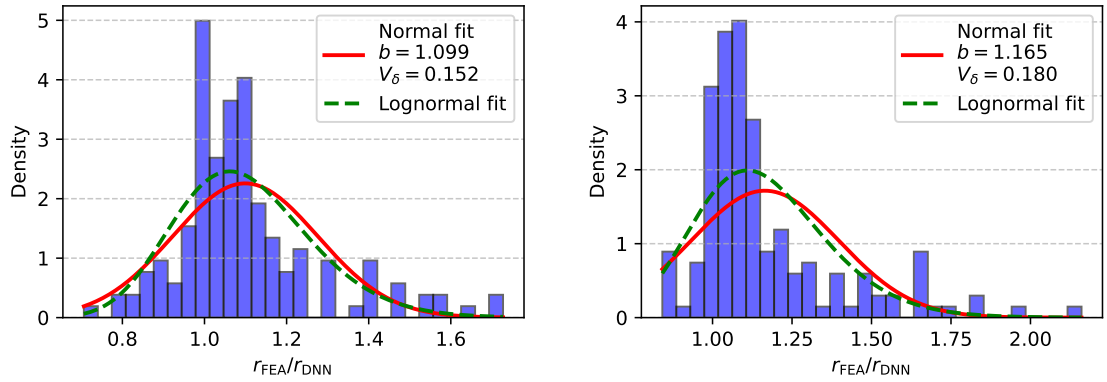
The mesh and von Mises stress for the base model of IPE 400 are shown in Fig. 4.22. The model is validated by using the measured true stress-strain curve and low geometric imperfections. The basic case of a web in transverse compression is validated on BUT tests described in [24]; see Fig. 4.23. Models with axial force are validated on tests by Kuhlmann [45] and with load near the unstiffened end on other BUT tests [30]. The models with the end plate have not been validated yet, but due to their simplicity, no issues are expected.

In the following comparisons, ANSYS models utilize bilinear stress-strain curve with insignificant strain-hardening of $E/1000$ and plastic strain limit of ε_{lim} as suggested by EN 1993-1-5 [32]. Equivalent geometric imperfections equal to $d_w/200$ as suggested by FprEN 1993-1-14 [5], where d_w is the depth of straight portion of column web (clear depth).

The dataset created in ANSYS brings two additional parameters for several models: (i) axial force in the member in the range $n = N_{Ed}/N_{pl,Rd} \in [0.1, 0.9]$, (ii) end plate with the thickness in the range $t_{ep}/t_f \in [0.5, 2]$. The presence of an axial force cannot be captured by the trained model. The end plate thickness was transformed into the relative plate width a using the Eurocode assumption, where the distribution slope is 1:1, i.e. $a_{mod} = a + 2 * t_{ep}$. Furthermore, two welded sections are introduced; one is small and thick-webbed, which may not be an issue, but the other is with the depth of $h = 550$ mm and web thickness of $t_w = 10$ mm. Such high web slenderness is unusual among the hot-rolled sections.

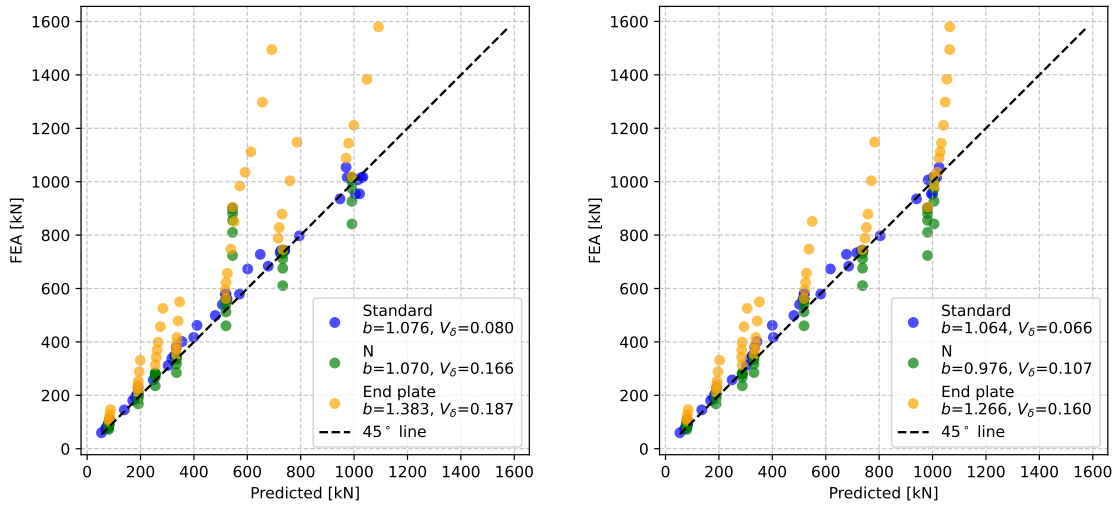
Also note that for the DNN predictions, relative distance x was set to 1 (the highest value in the range of validity), although in ANSYS models, it was $x = 2$. It is assumed that any models with the distance over $x > 1$ provide the same load resistance as $x = 1$.

Plastic resistance $F_{pl,Rd}$ by MNA is investigated first. The histograms in Fig. 4.24 show high scatter, which requires further investigation. Understandably, the DNN cannot accurately predict the additional parameters.



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.24: *Histogram of dataset created by validated ANSYS models – $F_{pl,Rd}$*



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.25: *Reliability graphs of dataset created by validated ANSYS models – plastic resistance by MNA, $F_{pl,Rd}$*

Fig. 4.25 shows reliability graph in several batches:

- Standard: Contains IPE and welded sections without any additional parameters.
- N: Contains the same cross-sections but loaded also by axial force $n = N_{Ed}/N_{pl,Rd} = \{0.1, 0.3, 0.5, 0.7, 0.9\}$
- End plate: Contains the same cross-sections but with the transverse load acting through an end plate with the relative thickness of $t_{ep}/t_f = \{0.5, 0.75, 1.0, 1.5, 2.0\}$

The Standard batch shows good alignment of DNN predictions towards ANSYS models. IDEA StatiCa models appear to be well calibrated. The exception is the welded section with high web slenderness ($r_e = 902$ kN). The DNN with the Similitude scaler grossly underestimated its resistance ($r_t = 545$ kN), whereas DNN without the Similitude scaler overestimated it ($r_e = 981$ kN). This point disrupts the variation coefficient V_δ for Similitude scaler.

The N batch shows that the predicted values are constant, equal to the standard case of $n = 0$. The load resistance determined by ANSYS progressively decreases with increasing axial load ratio n. Therefore, many points fall to the unsafe side, below 45° line. Again, the welded section disrupts DNN results with the Similitude scaler, pushing bias b even above one.

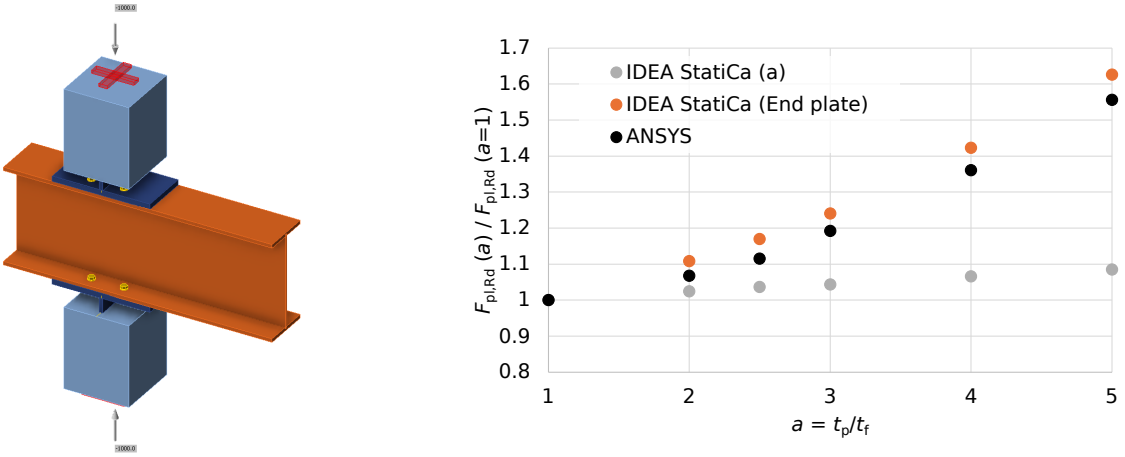
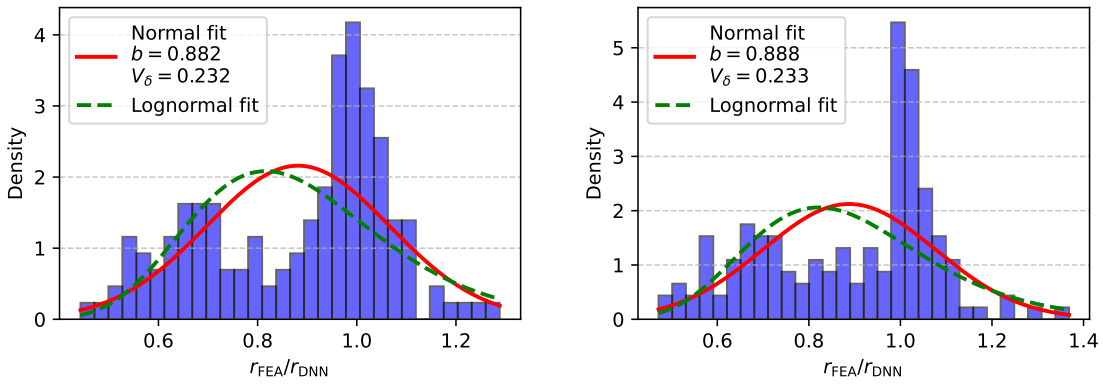


Figure 4.26: *Load resistances by IDEA StatiCa and ANSYS for varying plate width a and added end plate thickness*

The End plate batch is the most interesting. The DNN predictions are extremely safe. There may be two reasons for this discrepancy; either the spreading angle 1:1 is not correct, or IDEA StatiCa Connection does not predict well high values of relative plate thickness a. This can be quickly verified by several calculations with IPE 300 cross-section in IDEA StatiCa Connection; see Fig. 4.26. The ratio of IDEA StatiCa resistance $F_{pl,Rd}(a = 5)$ is only 8 % higher than $F_{pl,Rd}(a = 1)$. On the other hand, ANSYS with end plate $t_{ep}/t_f = 2$ carries 56 % more.

An additional IDEA StatiCa model with an added end plate was created and plotted into the same graph. The shell model with end plate follows the same trend as ANSYS model.

Critical buckling factor α_{cr} by LBA is shown next.



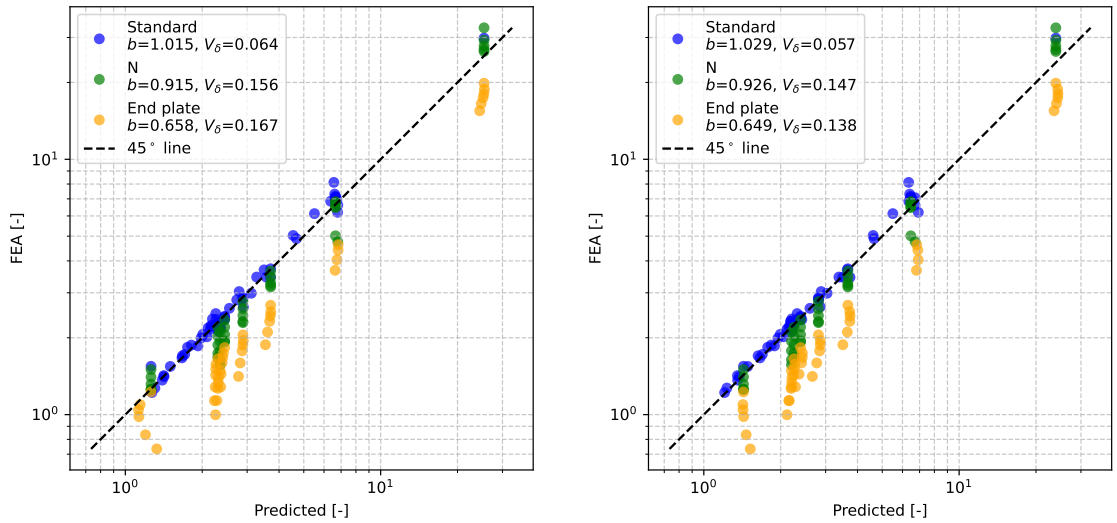
(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.27: *Histogram of dataset created by validated ANSYS models – α_{cr}*

Fig. 4.27 shows the histograms of $\alpha_{cr,FEA}$ to $\alpha_{cr,DNN}$ ratios. The bias b is smaller than one, meaning DNN predicts on average higher critical factor than determined by ANSYS, which is unsafe. Furthermore, the scatter is very large with $V_\delta = 0.233$. Note that $\alpha_{cr} = F_{cr}/F_{pl,Rd}$, and $F_{pl,Rd}$ is already not determined correctly and this error propagates also into α_{cr}.

Fig. 4.28 utilizes the same division into batches as Fig. 4.25. Standard batch shows good agreement of DNN predictions with ANSYS models. Batches N and End plate are not correctly captured.

From the above-mentioned graphs, it can be seen that: (i) End plate thickness cannot be simulated by increased bearing plate thickness a in the IDEA StatiCa Connection model and therefore in trained DNNs, (b) additional parameter of normal force significantly disrupts the alignment of predictions to the ground truth (ANSYS model) despite Eurocode assumption that



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.28: *Reliability graphs of dataset created by validated ANSYS models – buckling factor α_{cr} (log scale)*

up to $n = 0.7$, no reduction is necessary, and (iii) unknown welded cross-section may cause serious issues, although all the geometric parameters are within the range of validity; see Tab. 4.2.

Buckling resistance $F_{b,Rd}$ can be calculated using geometrically and materially nonlinear analysis with imperfections (GMNIA) or estimated using General method in EN 1993-1-1 – Cl. 6.3.4 [4] or EN 1993-1-5 – Annex B [32]. Web buckling resistance was determined in ANSYS by GMNIA, while only predictions of $F_{pl,Rd}$ and α_{cr} were trained on MNA and LBA in IDEA StatiCa Connection.

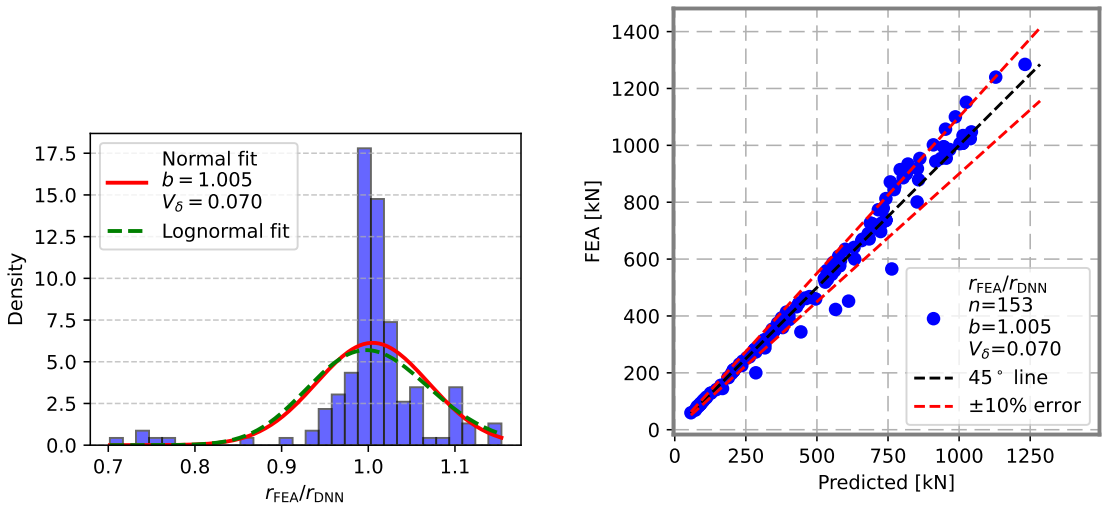
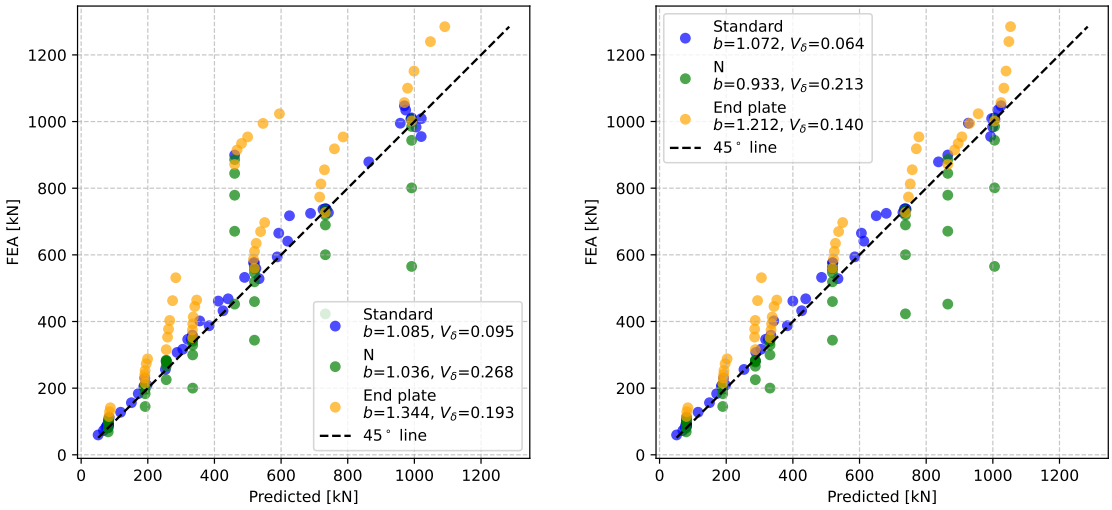


Figure 4.29: *The comparison of General method (r_t) and GMNIA (r_e) using only ANSYS data*

At first, the test is performed to determine whether the General method using Eurocode buckling curves and the new proposal described in Section 2.1.1 is feasible; see Fig. 4.29. In some cases, the General method provides nearly the same buckling resistance factor, ρ – for standard cases including load near unstiffened end. However, the buckling resistance factor decreases with increasing axial load; for $n = 0.9$, the drop is down to 70 % and the General method does not follow the decreasing trend nearly as much – only down to ∼92 %. Although, it should be mentioned

that for $n = 0.7$, the drop is not nearly as dramatic, and for $n \leq 0.5$, it is not noticeable. The end plate thickness is predicted better, but on the other hand, smaller buckling resistance is predicted. In sum, the General method in combination with the Eurocode buckling curve for the web panel in transverse compression and new proposed method may be used with the exception of extreme axial forces ($n \geq 0.7$).



(a) *With Similitude scaling* (b) *Without Similitude scaling*

Figure 4.30: *Reliability graphs of dataset created by validated ANSYS models – GMNIA calculation in ANSYS vs General method using predicted $F_{pl,Rd}$ and α_{cr}*

Fig. 4.30 shows the comparison of buckling resistance $F_{b,Rd}$ determined directly by GMNIA in ANSYS and the General method using the predicted plastic resistance $F_{pl,Rd}$ and buckling factor α_{cr} by DNN. The results for Standard cases and cases with low axial force have good alignment except for the welded section.

It is concerning that such a poor prediction is obtained for a relatively standard welded cross-section. The welded cross-section has a large web slenderness – ratio of $(h - 2 \cdot t_f)/t_w$. However, all the individual dimensions and this ratio was included in the training set – the range of validity of this ratio is $(h - 2 \cdot t_f)/t_w \in [3.19, 62.52]$. What is apparently outside the range of validity is the fillet radius.

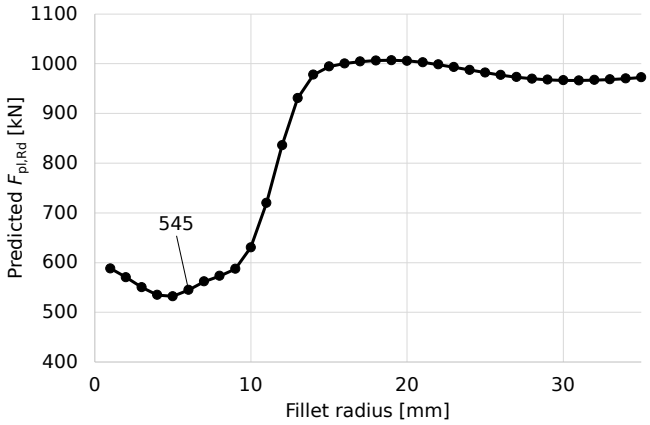


Figure 4.31: *Predictions of plastic resistance by DNN with the Similitude scaler for varying fillet radius*

Fig. 4.31 shows the predictions of the same welded cross-section with only one varying geometric parameter – fillet radius r_w. As shown earlier, the calculations of IDEA StatiCa remain constant

with varying r_w. Also, the SHAP graphs (see Section 4.7) show that r_w does not affect the predictions nearly at all. However, going outside the range of validity with the scaled $r_{w,scaled} = 6/550 \cdot 200 = 2.18$ mm < 3.48 mm (minimum $r_{w,scaled}$ in the training dataset) by the Similitude scaler, the predictions are wildly modified.

4.6 Reliability

This section presents the reliability of the current approach of the second generation Eurocode [28], the reliability of the proposed model by Balázs et al. [30] with an increased degree of fixity between the flange and web, better corresponding to the numerical models and also with a possibility to provide design resistance close to the unstiffened end (see Section 2.1.1), and finally the reliability of predictions made by the neural network described in the previous sections of this chapter. The procedure for reliability assessment is described in Section 2.6.

4.6.1 Eurocode reliability

The experimental set of 91 specimens has to be reduced – specimens with load applied close to the unstiffened end tested by Balázs [30] are removed. A total of 85 specimens remain to be calculated according to Eurocode procedures. The ratios of experimental results to Eurocode predictions using measured material and geometrical (if available) properties are shown in Fig. 4.32 and 4.33. No experiment failed at a lower than calculated load with measured properties. It can be seen that some specimens failed at extremely large ratios, demonstrating the Eurocode prediction is highly conservative. Taking into account all tests, the bias calculated as the average of ratios is $b = 1.462$, the coefficient of variability of prediction errors is rather high, $V_\delta = 0.199$, and by assuming variation of the basic variables $V_{rt} = 0.07$, as was used in Eurocode development, the coefficient of variation $V_r = 0.211$ is obtained. Theoretically, V_{rt} could be calculated more precisely, but with high V_δ this does not make any sense. The resulting safety factor is $\gamma_M = 1.128$, which is below the acceptance limit 1.15.

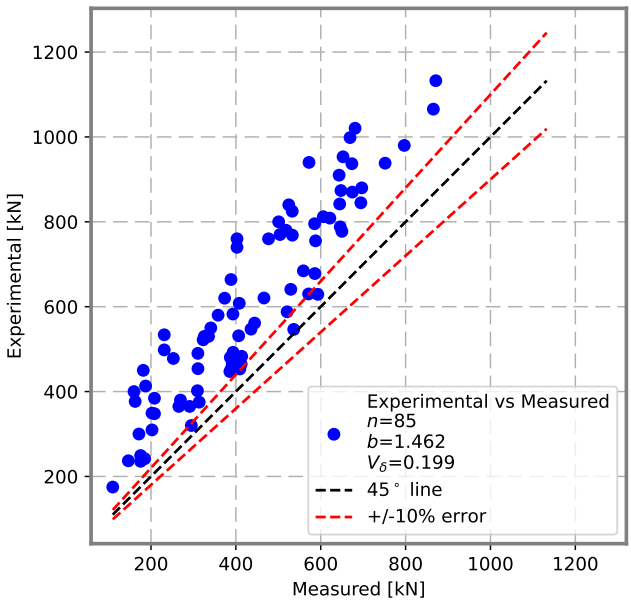


Figure 4.32: *Reliability graph of all specimens – comparison of experiments to calculations with measured properties*

The histogram shows a log-normal rather than the normal distribution assumed in Eurocode reliability calculations. The tail-fitting should be applied. In Fig. 4.34, where the specimens with the highest Experimental/Measured ratios, r_e/r_e, are deleted one-by-one and γ_M is calculated for each step, the best γ_M is achieved at 43 removed specimens.

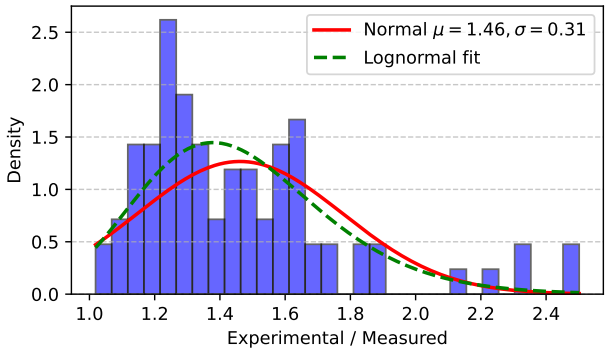


Figure 4.33: *Experimental resistance at peak load vs. Eurocode calculation [28]*

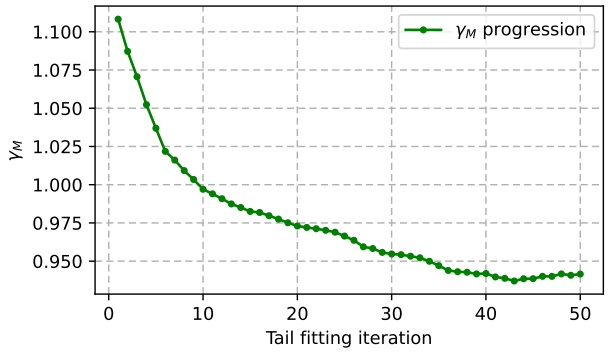


Figure 4.34: *Tail fitting with safety factor evaluated for each iteration*

The resulting parameters of reliability assessment for the dataset of 44 removed specimens with the highest r_e/r_e are as follows: $b = 1.228$, $V_\delta = 0.071$, $V_r = 0.100$, $\gamma_M = 0.937$, and the acceptance limit is 1.075. That means Eurocode formulas are well-calibrated even with a sufficient safety margin. However, half of the specimens removed by tail-fitting fail at a much higher load in the experiment than in the EC calculation. The formulas may clearly be improved.

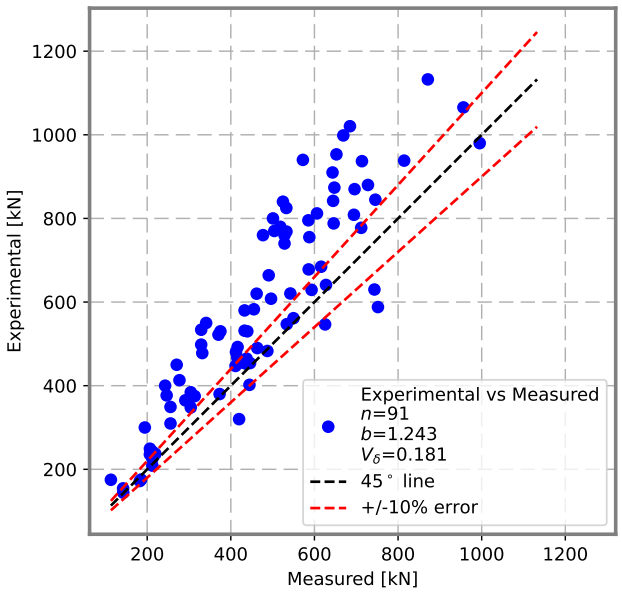


Figure 4.35: *Reliability graph of all specimens calculated by new proposal*

4.6.2 Reliability of new method

The new method proposed in Balázs [30] and shown in the second part of Section 2.1.1 allows the calculation of all 91 specimens including those with concentrated load near the unstiffened end. The reliability graph in Fig. 4.35 and the histogram in Fig. 4.36 show that the coefficient of variation is smaller. On the other hand, multiple specimens failed with a lower experimental resistance than determined by calculation, which indicates that the safety factor $\gamma_M > 1.0$ must be applied.

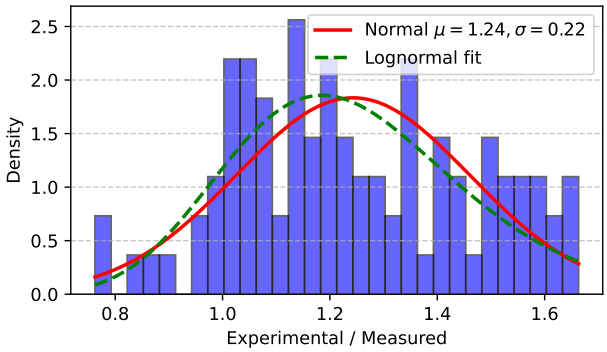


Figure 4.36: *Experimental resistance at peak load vs. calculation according to the new proposal [30]*

Taking into account all specimens results into the following parameters of reliability assessment: $b = 1.243$ (significantly lower than EC), $V_\delta = 0.181$ (slightly lower than EC), $V_r = 0.194$, $\gamma_M = 1.226$, and the acceptance limit is 1.15.

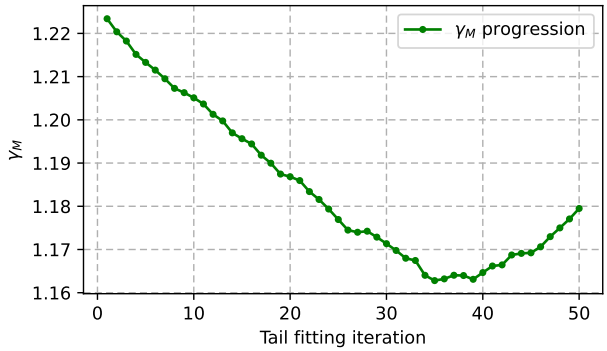


Figure 4.37: *Tail fitting with safety factor evaluated for each iteration*

The histogram in Fig. 4.36 again indicates that tail-fitting may be beneficial. According to Fig. 4.37, removing 35 specimens results in the smallest safety factor. The resulting parameters of reliability assessment for the dataset of 35 removed specimens with the highest r_e/r_e are as follows: $b = 1.098$, $V_\delta = 0.119$, $V_r = 0.138$, $\gamma_M = 1.163$, and the acceptance limit is 1.104. The safety factor is above the acceptance limit and the safety factor has to be applied. It can be calculated as the ratio of γ_M and the acceptance limit, i.e., the final $\gamma_M = 1.163/1.104 = 1.053$.

However, some safety should be added for the potential compressive force in the column, which is typical. The Eurocode formulation decreases the resistance of the column web panel in transverse compression component only for loads at 70 % of the column plastic axial resistance and higher. The experiments by Kuhlmann and Kühnemund [45] suggest that the component resistance decreases even at smaller axial loads. Therefore, a higher safety factor, about $\gamma_M \approx 1.15$ should be applied for this new procedure.

Alternatively, a safer assumption of resistance reduction for axial force in the column may be adopted, e.g. by Corman [105].

4.6.3 Reliability of neural network

Reliability of predictions by DNN with and without the Similitude scaler is shown on a set of all experiments except for tests by Bougoffa [42, 41] that were welded with large patch loading (feature a) outside the range of validity; experimental dataset comprises 85 specimens. In both DNNs, only one specimen failed at a slightly lower load in the experiment than the DNN predictions with measured material and geometrical properties – Delft 5.1 [35] but only by 2% and 1% with and without the Similitude scaler, respectively.

DNN was trained to predict plastic resistance, $F_{pl,R}$ and buckling factor α_{cr}. The final buckling resistance, $F_{b,R}$, which can be compared to the experimental resistance, must be calculated by the following procedure. The relative slenderness is calculated:

$$\bar{\lambda}_p = \sqrt{1/\alpha_{cr}} \tag{4.1}$$

For the load sufficiently far away from the unstiffened end ($x > 0.5$), the standard Eurocode [28] equation for the buckling reduction coefficient according to Eq.(2.3) is used. For the load near the unstiffened end, the new approach is used, i.e., Eq. (2.13) followed by Eq. (2.12) where λ_{p0} is calculated using Eq. 2.14.

Finally, the buckling resistance of the web in transverse compression is calculated:

$$F_{b,R} = \rho \cdot F_{pl,R} \tag{4.2}$$

where ρ is calculated according to Eq. (2.12) for $x \leq 0.5$ and according to Eq. (2.3) otherwise. Three buckling resistances are calculated as in previous reliability assessments:

1. Using measured material properties and measured geometric properties (if available)
2. Using nominal material and geometrical properties
3. Using mean material and geometrical properties

Similitude scaler provides very conservative predictions, with bias $b = 1.565$ and relatively large scatter $V_\delta = 0.221$; see Fig. 4.38. When considering all specimens, the partial safety factor $\gamma_M = 1.100$ is below the acceptance limit of 1.15.

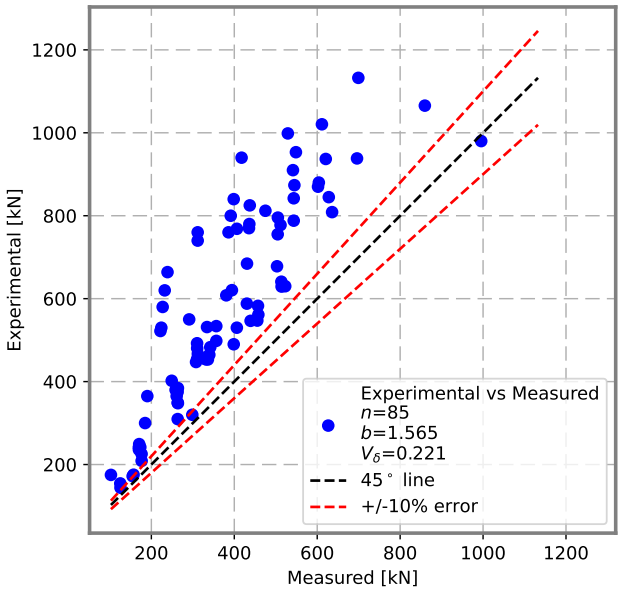


Figure 4.38: *Reliability graph of all specimens predicted by DNN with the Similitude scaler*

Fig. 4.39 shows again lognormal rather than normal distribution, and therefore the tail fitting is used. Fig. 4.40 shows that the best result is achieved after removing 26 specimens with the highest r_e/r_t ratio. The final bias $b = 1.368$ and variation coefficient $V_\delta = 0.112$. Again, assuming

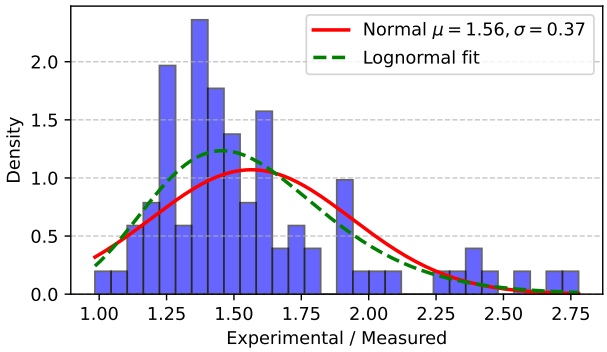


Figure 4.39: *Experimental resistance at peak load vs. prediction by DNN with the Similitude scaler*

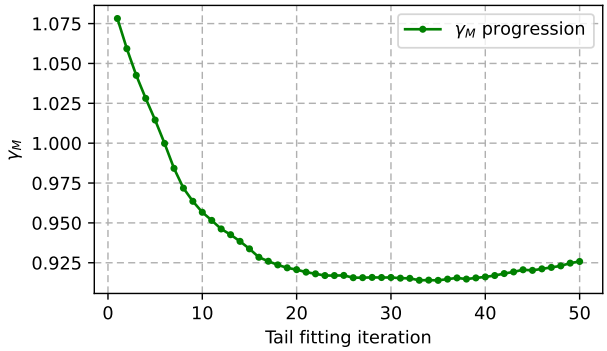


Figure 4.40: *Tail fitting with safety factor evaluated for each iteration*

$V_r t = 0.07$ results in $V_r = 0.132$ and the partial safety factor $\gamma_M = 0.916$ below the acceptance limit of 1.099.

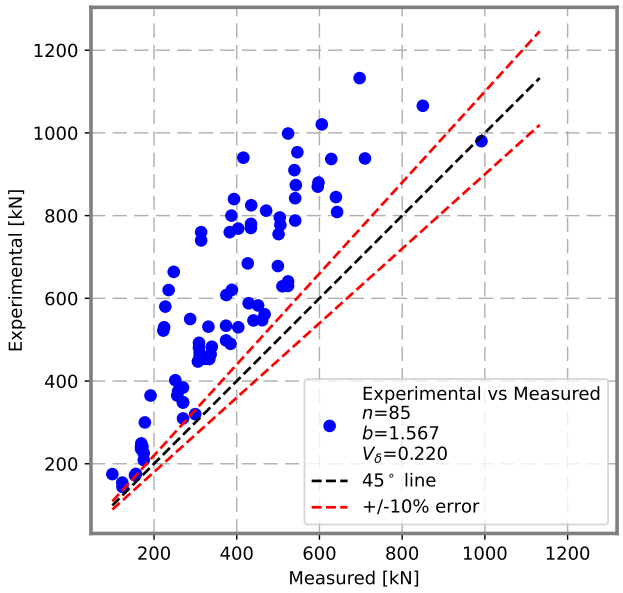


Figure 4.41: *Reliability graph of all specimens predicted by DNN without the Similitude scaler*

Without Similitude scaler , the DNN predictions are very similar. Bias $b = 1.567$ and large scatter $V_\delta = 0.220$; see Fig. 4.41. Assuming all specimens, the partial safety factor $\gamma_M = 1.093$ is within the acceptance limit of 1.15.

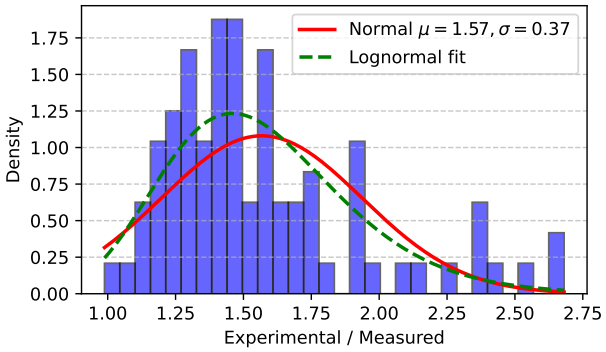


Figure 4.42: *Experimental resistance at peak load vs. prediction by DNN without the Similitude scaler*

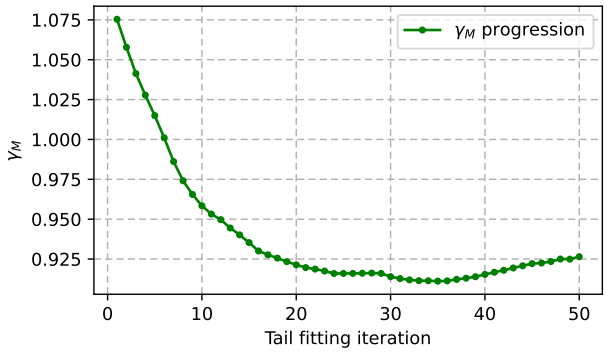


Figure 4.43: *Tail fitting with safety factor evaluated for each iteration*

The histogram in Fig. 4.42 shows lognormal distribution, so again, tail fitting is utilized. Fig. 4.43 shows that removing 35 specimens yields the best partial safety factor. The final bias $b = 1.330$ and variation coefficient $V_\delta = 0.100$, $V_r = 0.122$; partial safety factor $\gamma_M = 0.911$ is below the acceptance limit of 1.092.

Both with and without the Similitude scaler, the predictions are extremely conservative for specimens with end plates. The unsuitability of the model was shown already in ANSYS comparison; see Fig.4.25. These specimens are removed by tail fitting.

4.6.4 Comparison of methods

Both DNNs provide nearly the same results, comparable to the reliability of Eurocode equations, but both failed to improve the predictions. The DNN provides the most conservative results of all methods and unfortunately also the largest scatter to the experimental results; see Tab. 4.4.

Table 4.4: *Comparison of statistical parameters before and after tail fitting*

	No tail fitting				After tail fitting			
Model	Spec.	b	V_δ	$\gamma_M \lessgtr \gamma_{M,\mathrm{lim}}$	Spec.	b	V_δ	$\gamma_M \lessgtr \gamma_{M,\mathrm{lim}}$
EC	85	1.462	0.199	$1.128 < 1.15$	42	1.228	0.071	$0.937 < 1.075$
New proposal	91	1.243	0.181	$1.226 > 1.15$	56	1.098	0.119	$1.163 > 1.104$
DNN Simil.	85	1.565	0.221	$1.100 < 1.15$	59	1.368	0.112	$0.916 < 1.099$
DNN No Simil.	85	1.567	0.220	$1.093 < 1.15$	50	1.330	0.100	$0.911 < 1.092$

A closer look at individual specimens is necessary to find specific issues of each approach; see Tab. 4.5. Eurocode predictions are very conservative for the slender specimens tested by Bougoffa [42, 41]. Such high relative slenderness is rare in the dataset. This indicates that the degree of fixity between the web and flange is really higher than as assumed in Eurocode.

Many specimens in [35] are badly predicted by DNNs – with end plates, which confirms Fig. 4.26, but also others. Strangely, the calculations by Aribert according to Eurocode in some cases differ markedly from the calculations using the reported data, suggesting potential discrepancies in the Aribert database.

The tests by Bose [46] are the full beam-column end plate connections and it is likely that the governing failure mode as reported by the author – column web in compression – may be affected by web panel in shear or other effects.

Table 4.5: *Ratio r_e/r_t for experimental dataset – All approaches*

Specimen	Section (Note)	EC	New Proposal	DNN SIM	DNN No SIM
[30] – IPE200-T1	IPE200	1.343	1.142	1.386	1.397
[30] – IPE200-T2	IPE200	1.352	1.160	1.422	1.432
[30] – IPE200-T3	IPE200 ($x = 0.5$)		0.984	1.183	1.192
[30] – IPE200-T4	IPE200 ($x = 0.5$)		1.062	1.278	1.287
[30] – IPE200-T5	IPE200 ($x = 0.125$)		1.085	1.235	1.253
[30] – IPE200-T6	IPE200 ($x = 0.125$)		1.009	1.148	1.165
[30] – IPE200-T7	IPE200	1.420	1.207	1.461	1.473
[30] – IPE200-T8	IPE200	1.305	1.134	1.395	1.405
[30] – IPE200-T9	IPE200 ($x = 0.25$)		0.946	1.117	1.124
[30] – IPE200-T10	IPE200 ($x = 0.25$)		0.943	1.111	1.117
[41] – POS_508	W508x200x10x6	2.316	1.528		
[41] – POS_370	W369x200x10x6	2.206	1.490		
[42] – US_1	IPE300	1.892	1.441		
[42] – US_2	W400x180x12.3x5.5	1.616	1.079		
[42] – US_3	W395x200x10x6	2.475	1.663		
[42] – US_4	W528x200x10x6	2.501	1.647		
[35] – INSA L1	HEB140	1.253	1.253	1.398	1.425
[35] – INSA L2	HEB200	1.528	1.528	1.768	1.773
[35] – INSA L3	HEB260	1.290	1.250	1.445	1.460
[35] – INSA L4	HEB140	1.196	1.196	1.424	1.453
[35] – INSA L5	HEB200	1.505	1.505	1.788	1.795
[35] – INSA L6	HEB200	1.548	1.548	1.886	1.896
[35] – INSA L7	HEB260	1.263	1.209	1.457	1.473
[35] – INSA N1	HEB160	1.613	1.613	1.889	1.913
[35] – INSA T1	HEB200+EP10	1.593	1.593	1.968	1.984
[35] – INSA T2	HEB200+EP15	1.597	1.597	2.044	2.066
[35] – INSA T3	HEB200+EP20	1.601	1.601	2.107	2.135
[35] – INSA T4	HEB200+EP30	1.642	1.642	2.251	2.259
[35] – INSA M1	IPE140	1.592	1.546	1.707	1.763
[35] – INSA M2	HEA260	1.490	1.225	1.597	1.622
[35] – INSA M3	IPE220	1.748	1.543	1.622	1.687
[35] – INSA M4	IPE360	1.579	1.206	1.305	1.313
[35] – S.T. Delft 1.1	IPE240	1.410	1.018	1.470	1.454
[35] – S.T. Delft 2.1	IPE240	1.085	0.762	1.071	1.068
[35] – S.T. Delft 3.1	HEA240	1.167	0.991	1.412	1.421
[35] – S.T. Delft 4.1	HEA300	1.102	0.847	1.205	1.202
[35] – S.T. Delft 5.1	HEA500	1.230	0.985	0.984	0.988
[35] – INSA MH1	HEA140	1.374	1.219	1.919	1.905
[35] – INSA MH2	HEA160	1.628	1.413	2.365	2.361
[35] – INSA MH3	HEA160	1.617	1.409	2.352	2.346
[35] – INSA MH4	HEA200	1.887	1.438	2.438	2.421
[35] – INSA MH5	HEA200	1.838	1.400	2.374	2.357
[35] – INSA MH6	HEA200A	1.299	0.904	1.614	1.598
[35] – INSA MH7	HEA300A	1.129	0.782	1.366	1.371
[35] – INSA MH8	IPE240	1.461	1.018	1.366	1.409
[35] – INSA MH9	IPE360A	1.578	1.058	1.229	1.272

Specimen	Section (Note)	EC	New Proposal	DNN SIM	DNN No SIM
[35] – INSA MH10	HEA160+EP10	1.619	1.340	2.552	2.551
[35] – INSA MH11	HEA160+EP15	1.660	1.343	2.670	2.635
[35] – INSA MH12	HEA160+EP20	1.709	1.353	2.779	2.681
[45] – A1	HEA240($n = 0.67$)	1.126	1.061	1.368	1.379
[45] – A2	HEA240($n = 0.6$)	1.107	1.042	1.348	1.359
[45] – A3	HEA240	1.307	1.228	1.590	1.603
[45] – A4	HEA240($n = 0.11$)	1.242	1.165	1.548	1.557
[45] – A5	HEA240($n = 0.23$)	1.159	1.087	1.453	1.461
[45] – A6	HEA240($n = 0.34$)	1.193	1.124	1.497	1.505
[45] – A7	HEA240($n = 0.49$)	1.135	1.081	1.466	1.471
[45] – A8	HEA240	1.251	1.183	1.587	1.595
[45] – B1	HEB240	1.285	1.285	1.496	1.507
[45] – B2	HEB240($n = 0.47$)	1.157	1.157	1.348	1.357
[45] – B3	HEB240($n = 0.66$)	1.061	1.061	1.224	1.232
[45] – B4	HEB240($n = 0.12$)	1.414	1.414	1.680	1.687
[45] – B5	HEB240($n = 0.26$)	1.349	1.349	1.603	1.609
[45] – B6	HEB240($n = 0.38$)	1.306	1.306	1.550	1.556
[45] – B7	HEB240($n = 0.61$)	1.220	1.220	1.450	1.456
[45] – B8	HEB240	1.461	1.461	1.737	1.744
[43] – W10x19 SC-NS	W10x19	1.548	1.225	1.189	1.160
[43] – W10x19 DC-NS	W10x19	1.739	1.376	1.335	1.302
[43] – W10x39 SC-NS	W10x39	1.494	1.290	1.284	1.300
[43] – W12x26 SC-NS	W12x26	1.684	1.152	1.327	1.298
[43] – W12x26 DC-NS	W12x26	1.860	1.273	1.467	1.434
[43] – W16x31 SC-NS	W16x31	2.171	1.525	1.408	1.342
[43] – W16x31 DC-NS	W16x31	2.326	1.634	1.508	1.437
[46] – Test 1	254x254 UC 73 (EP12)	1.256	1.025	1.201	1.185
[46] – Test 2	254x254 UC 89 (EP12)	1.301	1.165	1.272	1.258
[46] – Test 3	254x254 UC 73 (EP15)	1.264	1.021	1.226	1.201
[46] – Test 4	254x254 UC 73 (EP15)	1.210	1.021	1.248	1.223
[46] – Test 5	254x254 UC 89 (EP15)	1.248	1.152	1.348	1.321
[46] – Test 6	254x254 UC 89 (EP15)	1.391	1.313	1.510	1.491
[46] – Test 7	254x254 UC 89 (EP20)	1.216	1.133	1.345	1.320
[44] – HE200A	HEA200	1.340	1.340	1.709	1.724
[44] – HE200B	HEB200	1.442	1.442	1.891	1.904
[44] – HE220A	HEA220	1.167	1.045	1.373	1.389
[44] – HE220B	HEB220	1.359	1.359	1.575	1.578
[44] – HE240A	HEA240	1.019	0.874	1.244	1.241
[44] – HE240B	HEB240	1.300	1.300	1.620	1.625
[44] – HE260A	HEA260	1.331	1.144	1.574	1.598
[44] – HE260B	HEB260	1.492	1.492	1.887	1.905
[44] – HE280A	HEA280	1.224	1.111	1.589	1.604
[44] – HE280B	HEB280	1.499	1.490	1.669	1.685
[44] – HE300A	HEA300	1.196	1.092	1.520	1.538
[44] – HE300B	HEB300	1.231	1.114	1.240	1.254

4.7 Explainability of DNN predictions

The predictions of DNNs are often considered a "black box". The output is generated by a mathematical model in which each neuron is defined by its weight and bias, while the entire neural network is characterized by its architecture – namely width, depth, and activation functions. In principle, any prediction can be decomposed into the contributions of individual neurons. However, practically, this is impossible to grasp for human perception.

To address this limitation, model-agnostic explainability methods have been developed. In this study, the SHapley Additive exPlanations (SHAP) framework [106] is employed. SHAP is based on Shapley values from cooperative game theory, which provide a theoretically consistent measure of the contribution of each feature to a given prediction. For every input sample, SHAP assigns additive importance values to all input features, such that their sum explains the deviation of the prediction from the mean model output.

Several types of SHAP visualizations are adopted to analyze the trained DNN without the Similitude scaler:

- **Summary plots**: show the distribution of SHAP values across all samples, indicating both global feature importance and the direction of influence (positive or negative contribution).
- **Bar plots**: rank the features by their mean absolute SHAP value, providing a straightforward measure of global importance without considering the sign.
- **Waterfall plots**: decompose a single prediction into baseline value and successive contributions of features, thereby explaining why the model produced the given output for an individual specimen.

Through these visualizations, SHAP allows the identification of the most influential geometric and material parameters and clarifies whether their increase tends to strengthen or weaken the predicted structural response. In this way, the DNN predictions can be interpreted in engineering terms rather than remaining opaque numerical results.

Note that scaling input parameters has both positive and negative outcome. Scaling allows directly comparing percentage changes in input values and is valuable when input parameters are of very different magnitude, e.g., beam depth h and web thickness t_w. On the other hand, scaled inputs may be hard to understand – their scale is no longer e.g. in [mm] units.

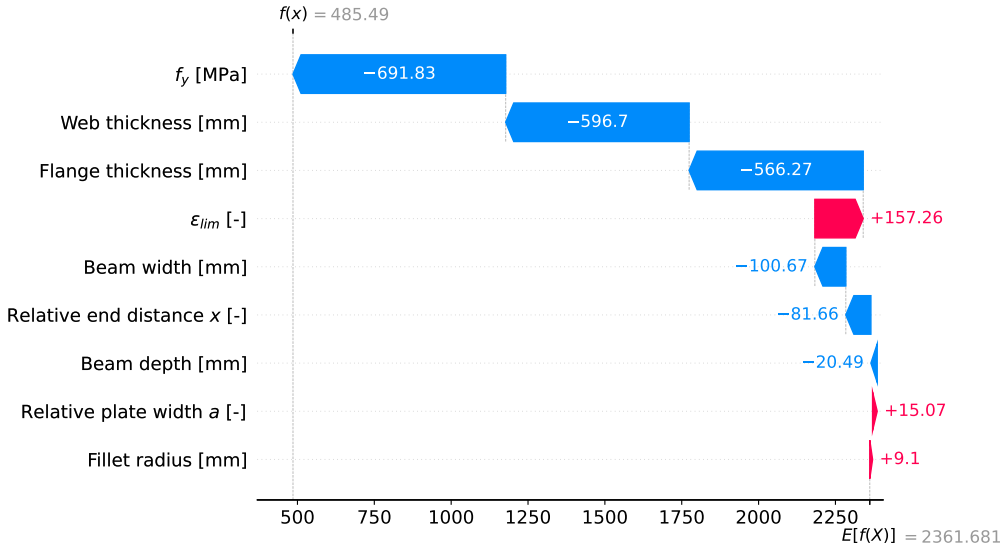


Figure 4.44: *SHAP waterfall plot for IPE 400 sample*

An example of one particular specimen is shown in Fig. 4.44. This is IPE 400 ($h = 400$ mm, $b = 180$ mm, $t_f = 13.5$ mm, $t_w = 8.6$ mm), steel grade S 355 ($f_y = 355$ MPa, $\varepsilon_{lim} = 4.13\%$), relative distance to unstiffened end $x = 0.739$, relative plate thickness $a = 1.68$, calculated plastic resistance $F_{pl,Rd} = 504.6$ kN. The mean predicted output is $E[f(X)] = 2361.681$ kN. The waterfall plot shows that the fillet radius has the smallest impact on the predicted output – plastic resistance $F_{pl,Rd}$. This is expected, since in IDEA StatiCa calculations used for training, fillet radius has no impact at all. This feature could be totally omitted for this training set. The second smallest impact has relative plate width a. It was observed in Fig. 4.26 that IDEA StatiCa is not very sensitive to this parameter. Strangely, the value of $a = 1.68$, which is below the average of $a = 3$ increases the resistance. The next feature that changes the mean resistance only lightly is the beam depth h, likely because $h = 400$ mm is close to the dataset mean. Next feature is the relative end distance $x = 0.739$, which decreases the mean output by 81.66 kN, which again comes as a surprise. ε_{lim} is above average, so it increases the resistance by 157 kN. More influential features follow. Flange thickness for IPE cross-section is low, thus reducing the resistance by 566 kN; the same applies for web thickness that reduces the resistance by another 597 kN. The most impactful is the yield strength, reducing the resistance by further 692 kN to final prediction of 485 kN.

For the entire dataset, a bar plot was constructed; see Fig. 4.45. It presents the average impact of individual features on the output – plastic resistance $F_{pl,Rd}$ – can be observed. As expected, the

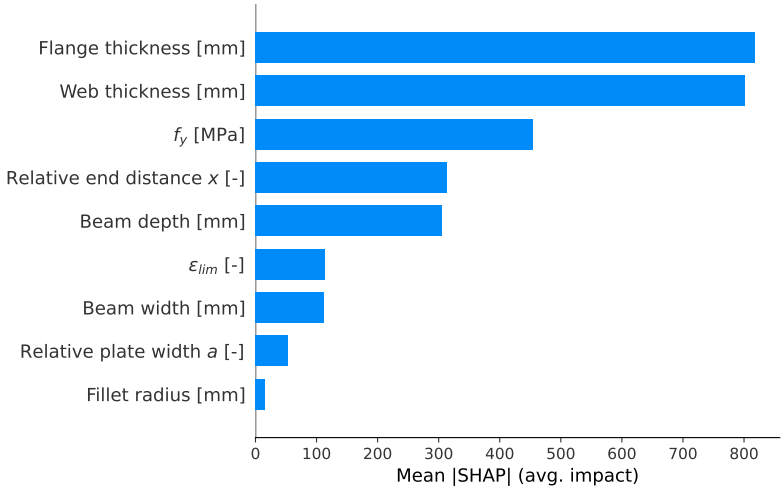


Figure 4.45: *SHAP bar plot shows the mean importance of each feature*

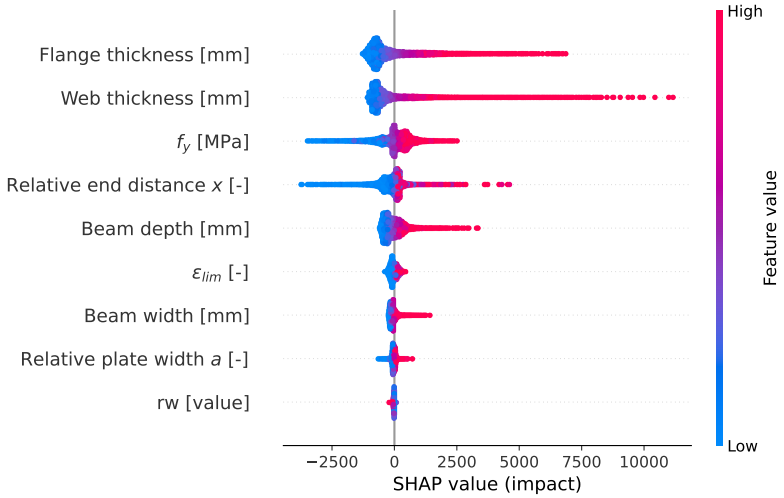


Figure 4.46: *Summary plot with dot plot (beeswarm plot)*

three most influential features are flange thickness, web thickness and yield strength. Note that the feature relative plate width is scaled according to flange thickness ($a = t_p/t_f$), which may increase the SHAP of flange thickness feature. Also, as expected, the fillet radius does not influence the IDEA StatiCa calculations used in the DNN training, and thus does not affect DNN prediction of the plastic resistance. This is caused by the simplicity of the model rather than the reality.This limitation arises from the simplicity of the training model rather than from structural behaviour.

A summary plot using the default dot representation is shown in Fig. 4.46. In this plot, each sample in the dataset is represented by a single dot, and the color indicates the feature value (e.g., low to high). The horizontal position of each dot corresponds to the SHAP value, i.e., the impact of that feature on the model output. This visualization provides an additional perspective on feature importance: not only can the magnitude of influence of each feature be assessed, but the direction of its effect on the prediction can also be observed. Specifically, features with dots predominantly on the positive side tend to increase the output, whereas features with dots on the negative side tend to decrease it. Furthermore, the color gradient allows identifying whether high or low feature values are associated with positive or negative contributions. Overall, the summary plot conveys both global feature importance and local feature effects across the dataset.

A more detailed investigation of individual features can be performed using SHAP scatter plots, which visualize the relationship between feature values and their contribution to the model output. In Fig. 4.47, the SHAP scatter plot for flange thickness shows a near-linear increase of SHAP values (corresponding to the predicted plastic resistance) with increasing flange thickness.

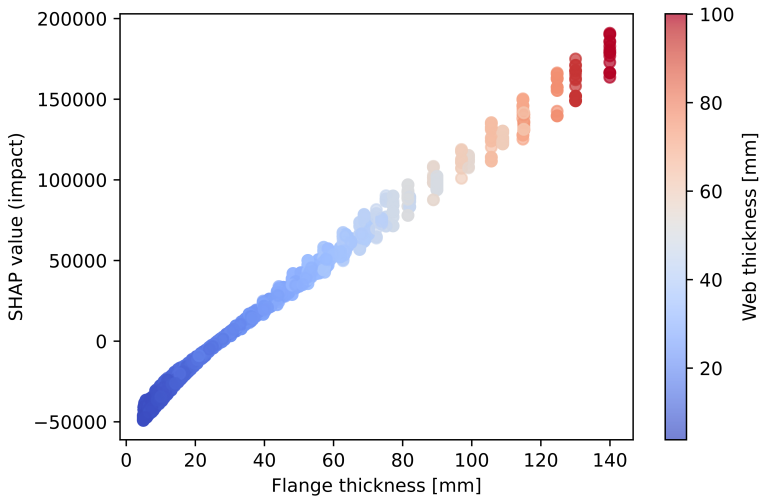


Figure 4.47: *SHAP scatter plot – SHAP impact of flange thickness and web thickness (Note that SHAP values are in a different scale in this graph)*

This trend suggests that thicker flanges generally lead to higher plastic resistance. The observed linearity is likely influenced by correlations with other geometric parameters in the dataset, such as web thickness, beam width, and overall depth, which tend to increase proportionally with flange thickness in the sampled beams.

Such scatter plots are particularly useful to identify non-linear effects, thresholds, or saturation points in feature contributions, as well as potential interactions with other features. For instance, deviations from the linear trend may indicate regions where other parameters, material properties, or joint detailing start to dominate the resistance behavior. This enables more targeted interpretation of the model's predictions and highlights areas where design optimization or additional experimental validation may be needed.

Chapter 5

Conclusion and future outlook

In this thesis, a framework for developing a reliable machine-learned model was outlined. It covered all stages of the process, and potential challenges were always highlighted. It is expected that only numerical design calculations are able to provide enough data samples for the successful creation of machine-learned model; therefore, big attention is given to the process of modeling, verifying, and especially failure criteria, which is still unresolved within the development of the second generation of Eurocode. An example of a column web panel in transverse compression component that aligns with the latest research focus of the author is selected. Analytical methods were described, experimental results collected, and a large-scale dataset of numerical models was generated with IDEA StatiCa Connection. The process of verifying numerical design calculations was described with special attention on failure criterion C2 – a plastic strain limit. The procedure for reliability assessment using the current state-of-the-art was described and the reliability of four different methods was assessed using the experimental dataset.

Machine learning is able to predict the training dataset with near-perfect accuracy, provided that a sufficient amount of data (several thousand samples) is available. The reliability of the machine-learned predictions is basically the same as that of samples used for training. There is a big potential in machine learning used in engineering practice for the following reasons:

- The precision of predictions is unparalleled. The coefficient of variation less than 2% is achievable.
- The speed of predictions is extremely fast, comparable to any calculation using code formulas. Machine-learned predictions may be used for a direct reliability approach using Monte Carlo simulations.
- The reliability of machine-learning predictions within a certain range of validity may be assessed in the same way as the reliability of any other approach.

On the other hand, there are several critical issues that anyone using neural networks must keep in mind.

- The range of validity must be strictly kept. This applies not only to the boundaries of individual features but also to the ratios of feature values; for example, not only must the thickness of the web and flange lie within the range of validity, but also the ratio of web to flange thicknesses.
- The machine-learned model, if perfectly applied, copies the training dataset. If the training dataset has any shortcomings, they directly propagate into the machine-learned model.

Dataset for training was created using IDEA StatiCa Connection, where thousands of models can be effortlessly calculated using Python programming language and IDEA StatiCa API. More than 9 000 models simulating the behavior of the column web in transverse compression component were calculated. The dataset comprises all hot-rolled I- or H-sections used in Europe and USA (IPE, HEA, HEB, and W), main steel grades used in Europe (S 235, S 275, S 355, S 450, S 420 MH/MLH, S 460 MH/MLH, S 620 Q/QL/QL1, S 690 Q/QL/QL1, S 500 MC, S 550 MC, S 600 MC,

S 650 MC, S 700 MC), varying distance to the unstiffened end, and varying width of plate inducing the transverse compression. IDEA StatiCa provides good approximation of load resistance for the standard cases, however, there were issues in the chosen strategy. The width of plate inducing the transverse compression has very little effect on the load resistance and wide plate cannot be used to simulate patch loading or loading through the end plate.

For these reasons, the machine-learned predictions failed to improve the code formulas for web panel in transverse compression. However, after expanding the dataset by cases with end plate and with an axial force in the column, the variation coefficient of machine-learned predictions will become lower than code calculations.

The geometrically and materially nonlinear analysis with imperfections, which would directly predict buckling resistance comparable to the experimental resistance, is only available in IDEA StatiCa Member without API. IDEA StatiCa Connection provides only materially nonlinear analysis and linear buckling analysis. The general method was utilized to estimate buckling resistance using the combination of both analyses.

There are possibilities for creating the dataset with higher-quality models:

- Similitude scaler could be used for enlarging the training and testing dataset rather than being used in the input and output scaling. In principle, for every calculated numerical model, another hundred geometrically similar models could be created.
- Machine learning could also be involved in the creation of a dataset:
  - It was shown that the General method works well for some cases and badly in others. Machine learning could be utilized to predict the results of computationally demanding GMNIA from the results of MNA, LBA and other parameters, such as geometry and material.
  - Assume Abaqus or ANSYS models created with solid elements. There is very likely a correlation between models with coarse and fine mesh. Again, machine learning could predict the results of numerical models with mesh conforming to the mesh sensitivity conditions from the simple and quickly analyzed models with a coarse mesh. This way, models could be calculated in minutes instead of several hours.
  - Much as likely, there is a correlation between IDEA StatiCa shell models and detailed models finely meshed with solid elements. Approximating correctly the results of detailed models by machine learning could speed up the process of dataset creation even more – seconds per model.

The machine-learned predictions achieve almost perfect alignment with the dataset used for training and testing – coefficient of variation equal to 0.0158 and 0.0134 for two different scaling strategies. This precision is only possible by **optimizing the neural network architecture**. Within the thesis, the effects of width (number of neurons in a layer) and depth (number of layers) were investigated together with learning rate, activation functions, optimizers, loss functions, and input and output scalers. The dataset size even for this simple model with only a few features (varying parameters) must be large – several thousand models. Each added model increases the prediction precision; however, this benefit diminishes with growing numbers. The coefficient of variation asymptotically approaches the numerical error of the finite element model. Active learning strategies were presented, but not applied.

The performance of machine-learned predictions was tested on several other datasets – (i) full dataset including large cross-sections that were filtered out with the aim of observing model performance outside the range of validity, (ii) dataset created in a traditional manner using nested for cycles, which is suitable for human perception but may cause overfitting of neural networks, and (iii) dataset created in another software – ANSYS where the shortcomings of IDEA StatiCa shell models were identified.

Building the numerical models with higher accuracy is an ideal task for Ph.D. students. These models should overcome the limitations of the numerical design calculations used in this thesis – missing geometrically nonlinear analysis with imperfections, neglected fillet radius, and insufficient end plate simulation. A sufficiently large database may serve as a training dataset for machine learning the foundations for improvements of current analytical models. Methods to enlarge the database by e.g. Similitude scaling or physically-informed methods could be investigated as well

as active learning. Such advancements may provide sufficiently large datasets for the creation of reliable machine learning models of structural components with much smaller deviations from numerical simulations and experiments.

Reliability assessment in accordance with current Eurocodes [6], based on the SAFEBRICTILE project [15], was presented as a step-by-step procedure. This procedure can be applied to any method – analytical, numerical simulations, numerical design calculations, or machine-learned predictions.

Reliability is a central topic of this thesis. It was evaluated for a Eurocode component, a new approach aimed at reducing Eurocode conservatism by increasing the degree of fixity between the web and the flange, and particularly for machine-learned predictions with and without the Similitude scaler. Achieving this is only possible through the collection of a large number of experimental results from the literature. Unfortunately, this is a challenging task, as many publications omit crucial information. There is room for improvement in verifying reported data and recovering missing information, such as measured geometrical properties.

The reliability assessment procedure is suitable for sets of similar experiments involving similar materials. However, when multiple steel grades or different types of experiments are considered, the limitations of the method become apparent. For instance, a material model with a high plastic strain limit and significant strain hardening exhibits low variability between experimental and numerical resistances. Such a model may appear highly reliable, allowing the application of a small partial safety factor γ_M. Conversely, if the model is made objectively safer – by reducing the plastic strain limit or using a bilinear material model that neglects strain hardening – the variability increases substantially, particularly when different steel grades with varying yield-to-ultimate strength ratios are involved. Paradoxically, this more conservative model may show lower reliability than the model closely matching experimental data. Clearly, this is inconsistent, highlighting the need for improvements in the reliability assessment procedure.

Data availability is more important today than ever. Collecting published data enables confident use of material models or assessment of manufacturing tolerances. Recent examples include the formulation of a characteristic material model by Yun and Gardner [7], distributions of material properties and thicknesses of plates and hot-rolled sections from the SAFEBRICTILE project [14], and bolt tests compiled by Ding and Elkadhy [107]. Published experimental data are essential for validating numerical models, which is of paramount importance. Just as machine learning models have a limited range of validity, modeling choices may perform well for one failure mode but poorly for another. Validation against a wide variety of experiments builds confidence and allows the development of sufficiently general solutions.

The work by Ding and Elkadhy [107] in compiling bolt experiments highlights common issues with insufficient reporting in research papers and technical reports. Firstly, researchers often measure different and incompatible displacements. Secondly, published values are frequently post-processed without clear explanation and presented only in graphs, making it difficult – or sometimes impossible – to extract key parameters such as initial stiffness. Out of potentially hundreds or thousands of bolt tests, only about 60 were found to be reported with sufficient detail.

A potential solution is the development of unified databases of standardized tests, such as tensile coupon tests [10] or bolt tensile tests [108], which could gradually expand to include additional test types, such as connection component tests and end-plate connection tests. Such databases would encourage researchers to measure and report the necessary data while also providing guidance and shared experience for early-career researchers and laboratory personnel.

Fields in which data sharing and proper documentation are standard—such as software engineering—demonstrate much higher productivity than civil engineering. Time is wasted on the literature reviews again and again.

Deep neural networks are often considered black boxes, but in reality they are optimizers whose behavior can be interpreted. In this thesis, SHAP plots [106] were employed to facilitate this interpretation. These plots visualize the deviation of each prediction from the mean and allow assessment of the contribution of each input feature (parameter). Tools like SHAP enable a deeper understanding of neural network behavior.

A critical aspect of using neural networks correctly is ensuring predictions remain within the range of validity. For machine-learned predictions to be safely adopted by the engineering com-

munity, engineers must be informed whether the input parameters fall within this valid range. Predictions outside the range of validity should not be provided, as they are likely to be unreliable or misleading.

Neural networks are suitable for significant improvements in design formulas. The reliability of design resistances can be maintained while increasing the load resistance. Thus, significant savings in material are possible.

Bibliography

[1] *EN 1993-1-8:2005 – Eurocode 3: Design of steel structures – Part 1-8: Design of joints.* CEN. Brussels, 2005.

[2] K. Weynand et al. *Proposal for implementation of the component method for tubular joints in EN 1993-1-8.* Tech. rep. CIDECT, Oct. 2015.

[3] *Eurocode No.3 – Design of Steel Structures – Part 1 - General Rules and Rules for Buildings, Background Documentation D.04 – Statistical analysis of strength functions for welded H-section joint with respect fo available experimental data.* Tech. rep. Eurocode 3 Editorial Group, Nov. 1989.

[4] *EN 1993-1-1:2005 – Eurocode 3: Design of steel structures – Part 1-1: General rules and rules for buildings.* CEN. Brussels, 2005.

[5] *FprEN 1993-1-14:2024 – Eurocode 3: Design of steel structures – Part 1-14: Design assisted by finite element analysis.* CEN. Brussels, 2024.

[6] *EN 1990 – Eurocode – Basis of structural design.* CEN. Brussels, 2002.

[7] Xiang Yun and Leroy Gardner. "Stress-strain curves for hot-rolled steels". In: *Journal of Constructional Steel Research* 133 (2017), pp. 36–46. ISSN: 0143-974X. DOI: https://doi.org/10.1016/j.jcsr.2017.01.024. URL: https://www.sciencedirect.com/science/article/pii/S0143974X16305181.

[8] Adam J. Sadowski, J. Michael Rotter, and Thomas Ummenhofer. "12.13: On recent characterisations of the post-yield properties of structural carbon steels". In: *ce/papers* 1.2-3 (2017), pp. 3577–3583. DOI: https://doi.org/10.1002/cepa.413. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cepa.413. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cepa.413.

[9] Leroy Gardner and Xiang Yun. "Description of stress-strain curves for cold-formed steels". In: *Construction and Building Materials* 189 (2018), pp. 527–538. ISSN: 0950-0618. DOI: https://doi.org/10.1016/j.conbuildmat.2018.08.195. URL: https://www.sciencedirect.com/science/article/pii/S0950061818321317.

[10] *EN ISO 6892-1:2019 – Metallic materials – Tensile testing – Part 1: Method of test at room temperature.* CEN. Brussels, 2019.

[11] *EN 1090 – Execution of steel structures and aluminium structures – Part2: Technical requirements for steel structures.* CEN. Brussels, 2008.

[12] *FprEN 1993-1-1:2024 – Eurocode 3: Design of steel structures – Part 1-1: General rules and rules for buildings.* CEN. Brussels, 2024.

[13] F.S.K. Bijlaard, G. Sedlacek, and J.W.B. Stark. *Procedure for the determinantion of design resistance from tests.* Tech. rep. TNO Institute for Building Materials and Structures, June 1988.

[14] L. da Silva et al. *Standardization of Safety Assessment Procedures across Brittle to Ductile Failure Modes (SAFEBRICTILE).* Tech. rep. Research Fund for Coal and Steel, June 2016.

[15] A. Taras et al. *Standardization of Safety Assessment Procedures across Brittle to Ductile Failure Modes (SAFEBRICTILE) – Deliverable D1.1: Guideline for Safety Assessment of Design Rules for Steel Structures in Line with EN 1990.* Tech. rep. Research Fund for Coal and Steel, June 2016.

[16] *PyTorch Documentation.* PyTorch Foundation. 2025. URL: https://docs.pytorch.org/docs/stable/index.html.

[17] OpenAI. *ChatGPT*. Large language model, GPT-5, accessed on 2025-08-29. 2025. URL: https://chat.openai.com/.

[18] American Institute of Steel Construction. *Specification for Structural Steel Buildings (ANSI/AISC 360-22)*. Chicago, IL: AISC, 2022. URL: https://www.aisc.org/products/publication/standards/aisc-360/specification-for-structural-steel-buildings-ansiaisc-360-16-download2/.

[19] František Wald et al. *Component-Based Finite Element Design of Steel Connections*. Czech Technical University in Prague, 2021. ISBN: 9788001068625. URL: https://payhip.com/b/p0Hr.

[20] Mark D. Denavit et al. *Steel Connection Design by Inelastic Analysis*. Hoboken, NJ: Wiley, 2024. ISBN: 9781394222162. URL: https://www.wiley.com/en-us/Steel+Connection+Design+by+Inelastic+Analysis-p-9781394222155.

[21] IDEA StatiCa. *IDEA StatiCa Support Center – Verifications*. https://www.ideastatica.com/support-center-verifications?product=steel. Accessed on 2025-08-29. 2025.

[22] IDEA StatiCa. *Web in Transverse Compression*. 2024. URL: https://www.ideastatica.com/support-center/web-in-transverse-compression (visited on 06/19/2025).

[23] ANSYS, Inc. *ANSYS Mechanical User's Guide*. Release 2024 R1. ANSYS, Inc. Canonsburg, PA, USA, 2024. URL: https://www.ansys.com (visited on 06/19/2025).

[24] Ivan Balázs et al. "Numerical and experimental verification of column web in transverse compression". In: *Thin-Walled Structures* 203 (2024). Cited by: 0; All Open Access, Hybrid Gold Open Access. DOI: 10.1016/j.tws.2024.112267.

[25] Ivan Balázs et al. "Geometrically and materially nonlinear analysis of selected aspects influencing resistance of column webs in transverse compression". In: vol. 2928. 1. Cited by: 0. 2023. DOI: 10.1063/5.0170547.

[26] Ivan Balázs et al. "Resistance of Column Webs Subjected to Transverse Compression at the Unstiffened End". In: vol. 2950. 1. Cited by: 0. 2023. DOI: 10.1063/5.0180866.

[27] Jorge Conde et al. "Welded beam-to-column steel joints: Assessment of European design rules". In: *Thin-Walled Structures* 199 (2024). Cited by: 5; All Open Access, Green Open Access, Hybrid Gold Open Access. DOI: 10.1016/j.tws.2024.111844.

[28] *EN 1993-1-8:2024 – Eurocode 3: Design of steel structures – Part 1-8: Design of joints*. CEN. Brussels, 2024.

[29] V. H. Cochrane. "Rules for Rivet-Hole Deduction in Tension Members". In: *Engineering News-Record* 89.20 (1922), pp. 847–848.

[30] Ivan Balázs et al. "Experimental and Numerical Analysis of Column Web Subjected to Transverse Compression at the Unstiffened End". Manuscript submitted for publication. 2025.

[31] B. Johansson et al. *Commentary and worked examples to EN 1993-1-5 "Plated structural elements"*. Luxembourg: Office for Official Publications of the European Communities, 2007.

[32] *EN 1993-1-5:2006 – Eurocode 3: Design of steel structures – Part 1-5: Plated structural elements*. CEN. Brussels, 2006.

[33] Harald Unterweger and Andreas Taras. "Steifenlose Krafteinleitung bei Biegeträgern". In: *Stahlbau* 84.6 (2015), pp. 435–448. DOI: https://doi.org/10.1002/stab.201510279. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/stab.201510279. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/stab.201510279.

[34] Tudor Golea, Jean-Pierre Jaspart, and Jean-François Demonceau. "Characterisation of the Behaviour of Beam-to-Column Steel Joints up to Failure". In: *Advanced Steel Construction* 18.3 (2022), pp. 679–686. DOI: 10.18057/IJASC.2022.18.3.5. URL: https://www.ascjournal.com/index.php?option=com_content&view=article&id=608:vol18no3-5&catid=153:vol18no3&Itemid=538.

[35] J. M. Aribert, A. Lachal, and M. Moheissen. "Modelling and experimental investigation of plastic resistance and local buckling of H or I steel sections submitted to concentrated or partially distributed loading". In: *Contact Loading and Local Effects in Thin-Walled Plated and Shell Structures*. IUTAM Symposium Prague. Berlin: Springer-Verlag, 1992, pp. 101–110.

[36] George Winter. "Strength of Thin Steel Compression Flanges". In: *Transactions of the American Society of Civil Engineers* 112.1 (1947), pp. 527–554. DOI: 10.1061/TACEAT.0006092. eprint: https://ascelibrary.org/doi/pdf/10.1061/TACEAT.0006092. URL: https://ascelibrary.org/doi/abs/10.1061/TACEAT.0006092.

[37] Ondřej Pešek, Ivan Balázs, and Martin Horáček. *Verification of numerical model of I-beam in IDEA StatiCa Steel*. Tech. rep. Institute of Metal and Timber Structures, Faculty of Civil Engineering, Brno University of Technology, Oct. 2022.

[38] S. P. Timoshenko and J. M. Gere. *Theory of Elastic Stability*. 2nd. New York: McGraw-Hill, 1985.

[39] Jean-Pierre Jaspart, Adrien Corman, and Jean-François Demonceau. "Mechanical Properties of the Component "Column Web in Compression" in Steel Beam-to-column Joints". In: *ce/papers* 5.4 (2022), pp. 242–250. DOI: https://doi.org/10.1002/cepa.1752. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cepa.1752. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cepa.1752.

[40] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17.3 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2. URL: https://doi.org/10.1038/s41592-019-0686-2.

[41] Sabra Bougoffa et al. "Full Length Transverse Stiffener Under Compression". In: *ce/papers* 5.4 (2022), pp. 967–973. DOI: https://doi.org/10.1002/cepa.1841. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cepa.1841. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cepa.1841.

[42] Sabra Bougoffa et al. "Compression zone with partial stiffeners in beam-to-column steel connections". In: *Engineering Structures* 229 (2021), p. 111674. ISSN: 0141-0296. DOI: https://doi.org/10.1016/j.engstruct.2020.111674. URL: https://www.sciencedirect.com/science/article/pii/S0141029620342759.

[43] Javier Alvarez Rodilla and Keith Kowalkowski. "Determination of Capacities of Eccentric Stiffeners Part 1: Experimental Studies". In: *Engineering Journal* 58.2 (2021), pp. 79–98. DOI: 10.62913/engj.v58i2.1176. URL: https://doi.org/10.62913/engj.v58i2.1176.

[44] L. De Mita, V. Piluso, and G. Rizzano. "Theoretical and Experimental Analysis of Column Web in Compression". In: *The Open Construction and Building Technology Journal* 2 (2008), pp. 294–303. ISSN: 1874-8368. DOI: 10.2174/1874836800802010294. URL: https://doi.org/10.2174/1874836800802010294.

[45] Ulrike Kuhlmann and Frank Kühnemund. *Proposal of a new design resistance of the joint component "column web in compression"*. Universität Stuttgart, Institut für Konstruktion und Entwurf, 2001.

[46] B. Bose. "Design Resistance of Unstiffened Column Web Subject to Transverse Compression in Beam-to-Column Joints". In: *Journal of Constructional Steel Research* 45.1 (1998), pp. 1–15. ISSN: 0143-974X. DOI: 10.1016/S0143-974X(97)00065-5. URL: https://doi.org/10.1016/S0143-974X(97)00065-5.

[47] IDEA StatiCa. *IDEA StatiCa Developer Site: API Documentation and Resources*. Accessed on 2025-08-29. 2025. URL: https://developer.ideastatica.com/.

[48] *CEN/TR 1993-1-141:2024 – Eurocode 3: Design of steel structures – Part 1-141: Background and Explanations on EN 1993-1-14 Design assisted by finite element analysis*. CEN. Brussels, 2024.

[49] Luis Simões da Silva et al. *Moment Resistance of Strong-Axis Open-Section Beam-to-Column Welded Steel Joints*. Tech. rep. ISISE UC – Institute for Sustainability and Innovation in Structural Engineering, June 2024.

[50] Luís Simões da Silva et al. "FE-based assessment of the resistance of steel joints using strain averaging". In: *Thin-Walled Structures* (2025), p. 114038. ISSN: 0263-8231. DOI: https://doi.org/10.1016/j.tws.2025.114038. URL: https://www.sciencedirect.com/science/article/pii/S0263823125011279.

[51] F. Walport, L. Gardner, and D.A. Nethercot. "Equivalent bow imperfections for use in design by second order inelastic analysis". In: *Structures* 26 (2020), pp. 670–685. ISSN: 2352-0124. DOI: https://doi.org/10.1016/j.istruc.2020.03.065. URL: https://www.sciencedirect.com/science/article/pii/S2352012420301478.

[52] Jaap Wardenier et al. *Design Guide for Circular Hollow Section (CHS) Joints under Predominantly Static Loading. Construction with Hollow Steel Sections, Guide No. 1.* Ed. by Comité International pour le Développement et l'Étude de la Construction Tubulaire (CIDECT). 2nd ed. Stuttgart, Germany: LSS Verlag, 2008. ISBN: 978-3-938817-03-2.

[53] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd. Springer, 2009. ISBN: 978-0-387-84858-7.

[54] European Committee for Standardization. *EN 10365:2017 - Hot rolled steel channels, I and H sections – Dimensions and tolerances.* CEN, Brussels. 2017.

[55] European Committee for Standardization. *EN 10025-2:2019 - Hot rolled products of structural steels – Part 2: Technical delivery conditions for non-alloy structural steels.* CEN, Brussels. 2019.

[56] *IDEA StatiCa Connection Library.* https://connectionlibrary.ideastatica.com/. Accessed: 2025-08-06; Cloud-based database of over 1 000 000 real-world steel connection designs. IDEA StatiCa, 2025.

[57] Adam Paszke et al. "PyTorch: An imperative style, high-performance deep learning library". In: *Advances in Neural Information Processing Systems* 32 (2019).

[58] Christopher M. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006.

[59] Warren S. McCulloch and Walter Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity". In: *The Bulletin of Mathematical Biophysics* 5 (1943), pp. 115–133. DOI: 10.1007/BF02478259.

[60] Donald O. Hebb. *The Organization of Behavior: A Neuropsychological Theory.* New York: Wiley, 1949.

[61] Marvin Minsky and Seymour A. Papert. *Perceptrons: An Introduction to Computational Geometry.* Reissue of the 1988 Expanded Edition with a new foreword by Léon Bottou. Cambridge, MA: MIT Press, 2017. ISBN: 9780262534772.

[62] A. Müller and A. Taras. "Prediction of the local buckling strength and load-displacement behaviour of SHS and RHS members using Deep Neural Networks (DNN)–Introduction to the Deep Neural Network Direct Stiffness Method (DNN-DSM)". In: *Steel Construction* 15.S1 (2022). Special Issue, pp. 78–90.

[63] A. Müller, A. Taras, and M. A. Kraus. "Scientific Machine and Deep Learning Investigations of the Local Buckling Behaviour of Hollow Sections". In: *ce/papers* 5.4 (2022), pp. 1034–1042. DOI: 10.1002/cepa.1848. URL: https://doi.org/10.1002/cepa.1848.

[64] A. Müller and A. Taras. "Prediction of the deformation and local buckling behavior of structural systems using the Deep Neural Network Direct Stiffness Method (DNN-DSM)". In: *Proceedings of the Annual Stability Conference, Structural Stability Research Council.* Online. Charlotte, North Carolina, 2023, p. 18.

[65] E. Buckingham. "On physically similar systems; illustrations of the use of dimensional equations". In: *Physical Review* 4.4 (1914), pp. 345–376. DOI: 10.1103/PhysRev.4.345.

[66] Zdenek P. Bazant and Jordi Planas. *Fracture and Size Effect in Concrete and Other Quasibrittle Materials.* Boca Raton, FL: CRC Press, 1998. ISBN: 084938284X.

[67] John Duchi, Elad Hazan, and Yoram Singer. "Adaptive subgradient methods for online learning and stochastic optimization". In: *Journal of Machine Learning Research* 12.7 (2011), pp. 2121–2159.

[68] Geoffrey Hinton. *Lecture 6e rmsprop: Divide the gradient by a running average of its recent magnitude.* Coursera: Neural Networks for Machine Learning. Lecture slides. 2012.

[69] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization.* https://arxiv.org/pdf/1412.6980. arXiv preprint arXiv:1412.6980. 2014.

[70] Ilya Loshchilov and Frank Hutter. "Decoupled Weight Decay Regularization". In: *International Conference on Learning Representations (ICLR).* 2019.

[71] Vinod Nair and Geoffrey E Hinton. "Rectified linear units improve restricted Boltzmann machines". In: *Proceedings of the 27th International Conference on Machine Learning (ICML)*. 2010, pp. 807–814.

[72] Dan Hendrycks and Kevin Gimpel. "Gaussian Error Linear Units (GELUs)". In: *arXiv preprint arXiv:1606.08415* (2016).

[73] Peter J. Huber. "Robust estimation of a location parameter". In: *Annals of Mathematical Statistics* 35.1 (1964), pp. 73–101. DOI: 10.1214/aoms/1177703732.

[74] Fabian Pedregosa et al. "Scikit-learn: Machine learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[75] John D Hunter. "Matplotlib: A 2D graphics environment". In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95.

[76] Robert E. Melchers and André T. Beck. "Integration and Simulation Methods". In: *Structural Reliability Analysis and Prediction*. John Wiley and Sons, Ltd, 2017. Chap. 3, pp. 63–93. ISBN: 9781119266105. DOI: 10.1002/9781119266105.ch3.

[77] Filip Ljubinković Dr et al. "Towards Stochastic Characterization of Welded Beam-to-Column Steel Joints Using Artificial Neural Networks". In: *Structural Engineering International* 0.0 (2025), pp. 1–17. DOI: 10.1080/10168664.2025.2526010. eprint: https://doi.org/10.1080/10168664.2025.2526010. URL: https://doi.org/10.1080/10168664.2025.2526010.

[78] David Lehký, Lukáš Novák, and Drahomír Novák. "Surrogate Modeling for Stochastic Assessment of Engineering Structures". In: *Machine Learning, Optimization, and Data Science*. Ed. by Giuseppe Nicosia et al. Cham: Springer Nature Switzerland, 2023, pp. 388–401. ISBN: 978-3-031-25891-6.

[79] David Lehký and Martina Šomodíková. "Reliability calculation of time-consuming problems using a small-sample artificial neural network-based response surface method". In: *Neural Computing and Applications* 28.6 (2017), pp. 1249–1263. DOI: 10.1007/s00521-016-2485-3. URL: https://link.springer.com/article/10.1007/s00521-016-2485-3.

[80] European Commission et al. *Modern plastic design for steel structures*. Publications Office, 2010. DOI: 10.2777/91747.

[81] Yingbin Bao and Tomasz Wierzbicki. "On fracture locus in the equivalent strain and stress triaxiality space". In: *International Journal of Mechanical Sciences* 46.1 (2004), pp. 81–98. ISSN: 0020-7403. DOI: https://doi.org/10.1016/j.ijmecsci.2004.02.006. URL: https://www.sciencedirect.com/science/article/pii/S0020740304000360.

[82] Marvin Nahrmann and Anton Matzenmiller. "Modelling of nonlocal damage and failure in ductile steel sheets under multiaxial loading". In: *International Journal of Solids and Structures* 232 (2021), p. 111166. DOI: 10.1016/j.ijsolstr.2021.111166. URL: https://doi.org/10.1016/j.ijsolstr.2021.111166.

[83] Helen Bartsch et al. "Damage mechanics based failure prediction for structures under seismic action". In: *Journal of Constructional Steel Research* 173 (2020), p. 106264. ISSN: 0143-974X. DOI: https://doi.org/10.1016/j.jcsr.2020.106264. URL: https://www.sciencedirect.com/science/article/pii/S0143974X20308166.

[84] Franc Sinur and Darko Beg. "Reliability analysis of net cross-section resistance with accidental eccentricity of holes". In: *International Journal of Steel Structures* 9.2 (2009), pp. 153–160. DOI: 10.1007/BF03249490. URL: https://doi.org/10.1007/BF03249490.

[85] H.H. Snijder, R.W.A. Dekker, and P.A. Teeuwen. "13.01: Net cross-section failure of steel plates at bolt holes: Numerical work and statistical assessment of design rules". In: *ce/papers* 1.2-3 (2017), pp. 3679–3688. DOI: https://doi.org/10.1002/cepa.424. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cepa.424. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cepa.424.

[86] Hubertus H. Snijder et al. "Safety Assessment for Capacity Design of Bolted Steel Connections in Tension". In: *ce/papers* 5.4 (2022), pp. 282–289. DOI: https://doi.org/10.1002/cepa.1757. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cepa.1757. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cepa.1757.

[87] Kayla Truman-Jarrell, Avery Burnham, and Mark D. Denavit. *Structural Steel Connection Design for Tensile Rupture by Advanced Inelastic Analysis*. Final Report. Submitted to IDEA StatiCa. Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, Oct. 2022. URL: https://assets-us-01.kc-usercontent.com/1ca05609-4ad1-009e-bc40-2e1230b16a75/da1b274c-1434-4626-a6c1-4088f045cec9/Structural%20steel%20connection%20design%20for%20tensile%20rupture.pdf.

[88] *EN 10029: Hot-rolled steel plates 3 mm thick or above – Tolerances on dimensions and shape*. Brussels: European Committee for Standardization (CEN), 2010.

[89] Yukitaka Murakami. *Theory of Elasticity and Stress Concentration*. Hoboken: Wiley, 2017. ISBN: 9781119295136.

[90] Kirill Golubiatnikov et al. "Determination of design plastic strain limit for notched structural steel tensile specimens". In: *Alexandria Engineering Journal* 121 (2025), pp. 38–52. ISSN: 1110-0168. DOI: https://doi.org/10.1016/j.aej.2025.02.075. URL: https://www.sciencedirect.com/science/article/pii/S1110016825002510.

[91] K. Golubiatnikov, M. Vild, and F. Wald. "A Numerical-Analytical Prediction of the Net Cross-Section Resistance of Weakened Tensile Plates". In: *SSRN Electronic Journal* (2025). Preprint. DOI: 10.2139/ssrn.5491159. URL: https://ssrn.com/abstract=5491159.

[92] K. Golubiatnikov, M. Vild, and F. Wald. "Design Net Cross-Section Resistances for Numerical Design Analyses of Weakened Tensile Plates with Real Material Properties". In: *SSRN Electronic Journal* (2025). Preprint. DOI: 10.2139/ssrn.5499750. URL: https://ssrn.com/abstract=5499750.

[93] Pawel J. Romanowicz, Bogdan Szybinski, and Mateusz Wygoda. "Application of DIC Method in the Analysis of Stress Concentration and Plastic Zone Development Problems". In: *Materials* 13.16 (2020). ISSN: 1996-1944. DOI: 10.3390/ma13163460. URL: https://www.mdpi.com/1996-1944/13/16/3460.

[94] Hendrik Baarssen et al. "An experimental investigation on the net cross-section failure of damaged plates containing holes". In: *Procedia Structural Integrity* 41 (2022). 2nd Mediterranean Conference on Fracture and Structural Integrity, pp. 183–191. ISSN: 2452-3216. DOI: https://doi.org/10.1016/j.prostr.2022.05.020. URL: https://www.sciencedirect.com/science/article/pii/S2452321622004784.

[95] Bernatowska, Edyta and Sleczka, Lucjan. "Net section fracture assessment of steel bolted joints with shear lag effect". In: *MATEC Web Conf.* 262 (2019). DOI: 10.1051/matecconf/201926209002. URL: https://doi.org/10.1051/matecconf/201926209002.

[96] Justin D. Brown et al. *Evaluation of Influence of Hole Making Upon the Performance of Structural Steel Plates and Connections*. Technical Report FHWA/TX-07/0-4624-1. Center for Transportation Research, The University of Texas at Austin, Jan. 2007. URL: https://library.ctr.utexas.edu/ctr-publications/0-4624-1.pdf.

[97] Nestor I. Prado, Julián Carrillo, and Claudia Retamoso. "Effective net area of failure in steel plates subjected to tension". In: *Revista de la Construcción. Journal of Construction* 23.3 (Jan. 2025), pp. 608–622. DOI: 10.7764/RDLC.23.3.608. URL: https://revistadelaconstruccion.uc.cl/index.php/RDLC/article/view/62829.

[98] Kirill Golubiatnikov and František Wald. "Conversion between the numerical simulation and the calculation using the Deformation Energy Ratio approach". In: *MethodsX* 14 (2025), p. 103221. ISSN: 2215-0161. DOI: https://doi.org/10.1016/j.mex.2025.103221. URL: https://www.sciencedirect.com/science/article/pii/S2215016125000688.

[99] Xue-Wei Zhang et al. "Effects of the stress state on plastic deformation and ductile failure: Experiment and numerical simulation using a newly designed tension-shear specimen". In: *Fatigue & Fracture of Engineering Materials & Structures* 42.9 (2019), pp. 2079–2092. DOI: https://doi.org/10.1111/ffe.13084. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ffe.13084. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/ffe.13084.

[100] Dassault Systèmes Simulia Corp. *Abaqus 2024: Unified FEA Software*. User's Guide and Theory Manual. Dassault Systèmes Simulia Corp. Johnston, RI, USA, 2024. URL: https://www.3ds.com/products-services/simulia/products/abaqus/.

[101] IDEA StatiCa. *Shell Elements – Benchmark Studies*. https://preview.ideastatica.com/support-center/shell-elements-benchmark-studies. Accessed: 2025-09-16.

[102] Martin Vild. *MartinVild/Machine-Learning: Steel Joint Prediction using Neural Networks*. Version 1.0. Oct. 2025. DOI: 10.5281/zenodo.17359933. URL: https://doi.org/10.5281/zenodo.17359933.

[103] British Standards Institution (BSI). *EN 10149-1:2013 — Hot rolled flat products made of high yield strength steels for cold forming. Part 1: General technical delivery conditions*. British Standard implementing the European Standard EN 10149-1 (approved 17 August 2013). 2013.

[104] Benjamin Andrea Luca Urben. "Use of Machine Learning in the Design and Analysis of Steel Connections". Chair of Structural Mechanics and Monitoring. Master's Thesis. Zürich, Switzerland: ETH Zürich, Institute of Structural Engineering (IBK), June 2025. URL: https://github.com/benurben08/Master_Thesis.

[105] Adrien Corman, Jean-François Demonceau, and Jean-Pierre Jaspart. "Analytical model for the panel zone resistance in welded steel beam-to-column joints". In: *Journal of Constructional Steel Research* 189 (2022), p. 107099. ISSN: 0143-974X. DOI: https://doi.org/10.1016/j.jcsr.2021.107099. URL: https://www.sciencedirect.com/science/article/pii/S0143974X21005824.

[106] Scott M Lundberg and Su-In Lee. "A Unified Approach to Interpreting Model Predictions". In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon et al. Vol. 30. Curran Associates, Inc., 2017. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

[107] Zizhou Ding and Ahmed Elkady. "Generalized spring model for steel bolts in tension considering uncertainty and loading speed". In: *Journal of Constructional Steel Research* 231 (2025), p. 109574. ISSN: 0143-974X. DOI: https://doi.org/10.1016/j.jcsr.2025.109574. URL: https://www.sciencedirect.com/science/article/pii/S0143974X25002524.

[108] *Mechanical properties of fasteners made of carbon steel and alloy steel — Part 1: Bolts, screws and studs with specified property classes — Coarse thread and fine pitch thread*. International Standard. Geneva, Switzerland: International Organization for Standardization, 2013.

List of Figures

List of Tables