T.C
İSTANBUL KÜLTÜR UNIVERSITY
INSTITUTE OF GRADUATE STUDIES

STRUCTURAL DESIGN OPTIMIZATION OF STEEL BEAMS USING RANDOM FOREST AND MULTI-CRITERIA ANALYSIS

MASTER'S THESIS

OMAR MARSHAHA
2100001119

Department: Civil Engineering
Program: Structural Engineering

Supervisor: Assist. Prof. Dr. Erdal Coskun

September 2025

T.C

İSTANBUL KÜLTÜR UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

STRUCTURAL DESIGN OPTIMIZATION OF STEEL BEAMS USING RANDOM FOREST AND MULTI-CRITERIA ANALYSIS

MASTER'S THESIS

OMAR MARSHAHA

2100001119

Department: Civil Engineering

Program: Structural Engineering

Supervisor: Assist. Prof. Dr. Erdal Coskun

Members of Examining Committee: Prof. Dr. Abdullah Necmettin Gunduz

Assist. Prof. Dr. Edip Seckin

September 2025

DECLARATION

I, Omar Marshaha, as the sole author of this thesis entitled "STRUCTURAL DESIGN OPTIMIZATION OF STEEL BEAMS USING RANDOM FOREST AND MULTI-CRITERIA ANALYSIS", hereby solemnly declare that the work presented herein is the result of my own original research and effort, carried out in partial fulfillment of the requirements for the Master’s degree at the Faculty of Civil Engineering. I further affirm that this thesis, in whole or in part, has not been previously submitted or presented to any other university or academic institution for the award of any degree or for any other academic purpose.

I acknowledge the use of AI during the preparation of this thesis. The tool was employed to support the writing of literature sections and to assist in generating and improving Python code. While some text and code were initially generated by AI, I reviewed, modified, and adapted them to fit the specific needs of my work. All core ideas, data collection, analyses, calculations, and interpretations were carried out independently and reflect my own understanding and contributions.

September 2025

ACKNOWLEDGMENT

I am profoundly grateful to the Almighty for granting me the strength, perseverance, and wisdom to carry out and complete this research. It is my sincere aspiration that the outcomes of this study will offer meaningful contributions to the academic community and to the field of civil engineering.

I wish to extend my deepest appreciation to all those who have supported and encouraged me throughout this endeavor. Their guidance, insights, and kind words have been invaluable. I am especially indebted to my thesis advisor, Asst. Prof. Erdal Coşkun, for his unwavering supervision, constructive feedback, and dedicated mentorship, which greatly enriched the quality of this work.

My heartfelt thanks also go to my friends, colleagues, and peers for their consistent encouragement and support, particularly to my colleague who provided timely assistance during the most critical stages of this research.

Above all, I dedicate this thesis to my beloved parents, whose steadfast love, prayers, and sacrifices have been my constant source of strength, motivation, and inspiration.

TABLE OF CONTENTS

DECLARATION I
ACKNOWLEDGMENT II
ABBREVIATIONS VII
TABLE OF FIGURES VIII
TABLE OF TABLES IX
TABLE OF SYMBOLS X
KISA ÖZET XI
ABSTRACT XII
1. INTRODUCTION 1
1.1 Background 1
1.2 Thesis Objectives 2
1.3 Scope of the Study 4
1.4 Methodology Overview 5
1.5 Thesis Structure 8
2. LITERATURE REVIEW 10
2.1 I-Beams in Structural Engineering 10
2.1.1 Properties of I-Beams 10
2.1.2 Applications of I-Beams in Construction 11
2.2 Plastic Moment and Deflection 12
2.2.1 Bending Stress and Plastic Moment 12
2.2.2 Beam Deflection Under Load 13
2.3 Structural Optimization Techniques 13
2.3.1 Classical Beam Design and Optimization 14
2.3.2 Integrating Cost, Deflection, and Plastic Moment 15
2.4 Machine Learning in Structural Engineering 16
2.4.1 Machine Learning Basics 16
2.4.2 Random Forest Algorithm 17
2.5 Cost-Effectiveness in Structural Design 18
2.5.1 Factors Affecting Cost 18
2.5.2 Previous Work on Cost Optimization 19
3. THEORETICAL FRAMEWORK AND ENGINEERING PRINCIPLES 21
3.1 I-Beam Design and Parameters 21
3.1.1 Material Properties and Dimensions 21

3.2 The Integration of Plastic Moment ... 25
3.2.1 Theory and Formulae for Plastic Moment ... 25
3.2.2 Application to the Dataset ... 25
3.3 Deflection Calculation ... 27
3.3.1 Theory of Beam Deflection ... 27
3.3.2 Effects of Point Loads on Deflection ... 28
3.4 Cost-Effectiveness Analysis ... 30
3.4.1 Cost Calculation Methods ... 30
3.4.2 Relationship Between Cost, Deflection, and Performance ... 31
3.5 Optimization Criteria ... 33
3.5.1 Defining "Optimal" in Structural Design ... 33
3.5.2 Trade-offs Between Cost, Deflection, and Plastic Moment ... 34
4. METHODOLOGY ... 36
4.1 Data Collection and Preprocessing ... 36
4.1.1 Dataset Overview (96 I-Beams) ... 36
4.1.2 Augmenting the Dataset with Load Scenarios ... 38
4.1.3 Data Cleaning and Preprocessing ... 39
4.1.4 Exploratory Data Analysis (EDA) ... 40
4.2 Theoretical Framework: Plastic Moment and Deflection Calculations ... 41
4.2.1 Plastic Moment Calculation Formulae ... 41
4.2.2 Beam Deflection under Point Loads ... 43
4.2.3 Validation of Results ... 44
4.3 Integration of Classical Engineering Theory with Data-Driven Modeling ... 45
4.3.1 Calculation of Plastic Moment and Deflection ... 45
4.3.2 Development of the Scoring System ... 46
4.3.3 Bridging Classical and Modern Design Methods ... 46
4.3.4 Machine Learning Model Configuration ... 46
4.3.5 Synergy of Engineering Theory and Machine Learning ... 47
4.4 Random Forest Machine Learning Model ... 48
4.4.1 Introduction to Random Forest ... 48
4.4.2 Feature Selection and Engineering ... 49
4.4.3 Model Development and Training ... 50
4.5 Structural Optimization and Cost-Effectiveness Analysis ... 52
4.5.1 Defining the Optimization Criteria ... 52
4.5.2 Multi-Objective Optimization Framework ... 53
4.5.3 Optimization Algorithm and Process ... 55

4.5.4 Sensitivity Analysis and Results Interpretation 56
4.6 Validation of Methodology 58
4.6.1 Comparison with Classical Methods 58
4.6.2 Case Study of Optimized Beam Selection 59
4.6.3 Flowchart of Thesis Methodology: 61
5. DATA ANALYSIS AND RESULTS 62
5.1 Overview of Dataset and Data Exploration 62
5.1.1 Key Descriptive Statistics 62
5.1.2 Correlation Analysis 64
5.2 Plastic Moment and Deflection Calculations 66
5.2.1 Moment-Curvature Analysis 66
5.2.2 Results of Deflection Calculations 67
5.2.3 Worked example: Beam #1 (manual calculations) 68
5.3 Machine Learning Model Results 74
5.3.1 Training and Testing Results 74
5.3.2 Accuracy and Predictive Performance 74
5.4 Optimization Results 76
5.4.1 Optimal Beam Selection 76
5.4.2 Cost-Deflection Trade-Offs 78
6. DISCUSSION 80
6.1 Summary of Key Findings 80
6.2 Ensuring Geometric Realism of Synthetic Beam Sections 80
6.3 Comparison with Traditional Design Methods 81
6.4 Implications for Structural Engineering 81
6.5 Practical Applications in Industry 82
6.6 Limitations of the Study and Assumptions 82
6.7 Directions for Future Research 83
7. CONCLUSION 84
7.1 Recap of the Research Objectives and Contributions 84
7.2 Key Findings and Results 84
7.3 Final Remarks on Structural Optimization Using Machine Learning 85
7.4 Suggestions for Future Work in Structural Optimization and Machine Learning 85
8. DOCUMENTATION AND REPORTING: 87
Reporting Results 87
Python Code and Data 87

9. REFERENCES: 88
10. APPENDICES: 90
10.1 Appendix A: Raw Data (96 I-Beams Dataset).. 90
10.2 Appendix B: Top Predicted 100 beams 94
10.3 Appendix C: 96 I-Beams Dataset. 98
10.4 Appendix D: Full Python Code 100

ABBREVIATIONS

RSJ	: Rolled steel joists (I beams)
CAD	: computer-aided design
FEA	: finite element analysis
LP	: Linear Programming
NLP	: Non-Linear Programming
MCDM	: Multi-Criteria Decision Making
WSM	: Weighted Sum Method
AHP	: Analytic Hierarchy Process
TOPSIS	: Technique for Order of Preference by Similarity to Ideal Solution
ML	: Machine Learning
AHP	: Analytic Hierarchy Process
AHP	: Analytic Hierarchy Process
AHP	: Analytic Hierarchy Process
AHP	: Analytic Hierarchy Process
AHP	: Analytic Hierarchy Process

TABLE OF FIGURES

Figure 2.1 - Beam Parameters.......... 10
Figure 3.1 - Plastic Moment Formula Code.......... 26
Figure 4.1 - Flowchart of Thesis Methodology 61
Figure 5.1- Correlation Heatmap of Parameters 64
Figure 5.2 - Deflection Formula code.......... 67
Figure 5.3 - Beam Forces.......... 69
Figure 5.4 - Shear Force Diagram.......... 69
Figure 5.5 - Moment Diagram 70
Figure 5.6 - Deflection Diagram.......... 73
Figure 5.7 - Plastic Moment - Predicted vs Actual - Scatterplot 75
Figure 5.8 - Plastic Moment - Predicted vs Actual - code 76
Figure 5.9 - Random Forest Regression - Deflection vs Cost.......... 78

TABLE OF TABLES

Table 1 - Parameters and Symbols 22

Table 2 - Strategies in Selecting optimum beam 59

Table 3 - Parameters Statistics 63

Table 4 - Correlation Coefficients of Parameters 65

Table 5 - Top 3 beams based on the scoring system 77

Table 6 - Top 3 predicted beams 77

TABLE OF SYMBOLS

I_z	: Second Moment of Area
E	: Young’s modulus of Elasticity
d	: Depth (the total height from the top to the bottom of the beam)
b_b	: Flange Width (the horizontal width of the flanges)
t_b	: Flange Thickness (the vertical thickness of the top and bottom flanges)
t_w	: Web Thickness (the horizontal thickness of the vertical web)
r	: Root radius or fillet (the curvature connecting the flange to the web)
M_p	: Plastic moment
Z_p	: Plastic section modulus
Z_e	: Elastic section modulus
f_y	: Yield strength
δ	: Deflection
P	: Point load
R^2	: Coefficient of determination
MAE	: Mean Absolute Error
ν	: Poisson’s Ratio
ρ	: Density

Enstitüsü	:	Lisansüstü Eğitim Enstitüsü
Dalı	:	İnşaat Mühendisliği
Programı	:	Yapı (İngilizce)
Tez Danışmanı	:	Dr. Öğr. Üyesi Erdal Coşkun
Tez Türü ve Tarihi	:	Yüksek lisans – Eylül 2025

KISA ÖZET

Çelik Kirişlerin Yapısal Tasarım Optimizasyonu: Random Forest ve Çok Kriterli Analiz Yaklaşımıyla

Omar Marshaha

Yapısal olarak verimli ve maliyet açısından etkili inşaatlara yönelik artan talep, mühendisleri yapısal elemanlar, özellikle de yüksek taşıma verimliliği sayesinde köprüler, binalar ve endüstriyel uygulamalarda yaygın olarak kullanılan I-kirişler için gelişmiş optimizasyon tekniklerini araştırmaya teşvik etmiştir. Bu tez, klasik mühendislik hesaplamaları ile Makine Öğrenmesi yöntemlerinden Random Forest algoritmasının entegrasyonunu içeren, basit mesnetli evrensel I-kirişlerin yapısal optimizasyonuna yönelik kapsamlı bir metodoloji önermektedir.

Birleşik Krallık'ta üretilmiş 96 evrensel I-kirişten oluşan bir veri seti analiz edilmiştir. Bu veri seti, kesit boyutları, metre başına kütle ve maliyet gibi temel parametreleri içermektedir. Her bir kiriş için, 10 metrelik açıklıkta üç noktalı yükleme senaryosu altında plastik moment kapasitesi ve sehim değerleri, yerleşik yapısal formüller kullanılarak hesaplanmıştır. Bu hesaplamalar, doğruluk ve tutarlılık sağlamak amacıyla hem Python hem de Excel ortamlarında doğrulanmıştır.

Plastik moment, sehim ve maliyeti entegre ederek kiriş performansını değerlendiren bir puanlama sistemi geliştirilmiş; bu sistem, kirişlerin nesnel olarak karşılaştırılmasına ve sıralanmasına olanak tanımıştır. Random Forest modeli, veri seti dışında kalan varsayımsal kiriş konfigürasyonları için plastik moment ve sehim tahminleri yapmak üzere eğitilmiş ve optimize edilmiş yeni tasarım seçenekleri üretmiştir.

Sonuçlar, bu bütünleşik yaklaşımın yüksek performanslı kirişleri başarıyla tanımladığını ve geometri ile mekanik performans arasındaki doğrusal olmayan ilişkileri yakalayabildiğini göstermektedir. Geliştirilen metodoloji, mühendislerin verimli ve ekonomik yapısal çözümler arayışında kullanabileceği pratik bir karar destek aracı sunmaktadır.

Anahtar Sözcükler: Yapısal Optimizasyon, I-Kirişler, Plastik Moment, Kiriş Sehimleri, Rastgele Orman (Random Forest), Makine Öğrenimi, Maliyet Etkinliği, Veriye Dayalı Tasarım, Yapı Mühendisliği.

Bilim Dalı Sayısal Kodu:

University	:	İstanbul Kültür University
Institute	:	Institute of Graduate Studies
Department	:	Civil Engineering
Program	:	Structural Engineering
Supervisor	:	Assist. Prof. Dr. Erdal Coşkun
Degree Awarded and Date	:	MSc – September 2025

ABSTRACT

Structural Design Optimization of Steel Beams Using Random Forest and Multi-Criteria Analysis

Omar Marshaha

The increasing demand for structurally efficient and cost-effective construction has encouraged engineers to explore advanced optimization techniques for structural elements, particularly I-beams, which are widely used in bridges, buildings, and industrial applications due to their high load-bearing efficiency. This thesis proposes a comprehensive methodology for the structural optimization of simply supported universal I-beams by integrating classical engineering calculations with machine learning, specifically using the Random Forest algorithm.

A dataset of 96 universal I-beams manufactured in the UK is analyzed, containing essential parameters such as cross-sectional dimensions, mass per meter, and cost. For each beam, the plastic moment capacity and deflection under a 10-meter three-point loading scenario were computed using well-established structural formulas. These calculations were verified in both Python and Excel to ensure consistency and accuracy.

A scoring system was developed to evaluate beam performance by integrating plastic moment, deflection, and cost, allowing objective comparison and ranking. The Random Forest model was trained on the dataset to predict plastic moment and deflection for hypothetical beam configurations outside the original database, generating new, optimized design options.

Results demonstrate that this combined approach successfully identifies high-performing beams and captures the non-linear relationships between geometry and mechanical performance. The methodology offers a practical decision-support tool for engineers seeking efficient and economical structural solutions.

Key Words: Structural Optimization, I-Beams, Plastic Moment, Beam Deflection, Random Forest, Machine Learning, Cost Efficiency, Data-Driven Design, Structural Engineering

Science Code :

1. INTRODUCTION

1.1 Background

Structural engineering plays a pivotal role in shaping the built environment by ensuring that buildings, bridges, and other infrastructures are safe, efficient, and durable. Among the variety of structural components used in construction, steel I-beams are one of the most fundamental and widely utilized elements. Their characteristic “I” shape provides an excellent balance between weight and strength, allowing for efficient load-bearing capabilities while minimizing material consumption.

The performance of an I-beam is primarily determined by its geometric properties—such as depth, flange width, flange thickness, and web thickness—as well as the material characteristics. Key mechanical properties including plastic moment capacity and deflection under load are critical parameters that influence the beam’s ability to carry loads safely without excessive deformation or failure. Structural designers traditionally rely on standards and empirical formulas to select appropriate beam sizes, balancing safety and cost.

In recent decades, however, the complexity and diversity of construction demands have increased significantly. Projects often require more precise optimization to meet constraints such as budget limitations, material availability, and sustainability targets, without compromising structural integrity. This has led to growing interest in computational methods that can provide more sophisticated and tailored solutions for beam design.

A widely recognized definition of Machine Learning comes from pioneer A. L. Samuel, who described it as “the field of study that gives computers the ability to learn without being explicitly programmed.” This definition captures the essence of ML, which underpins the technical foundation of data mining—the process of uncovering hidden patterns within data. However, ML places a stronger emphasis on making predictions using patterns learned from training data. (Charalampakis & Papanikolaou, 2021).

ML, a subset of artificial intelligence, offers a promising avenue to enhance structural optimization by uncovering complex relationships within data and predicting performance based on input variables. Particularly, Random Forest models—a robust ensemble learning method—have shown success in handling non-linear dependencies and multivariate datasets, making them well-suited for structural engineering problems involving multiple interacting parameters.

This thesis focuses on the structural optimization of simply supported universal steel I-beams by integrating classical engineering calculations with machine learning techniques. Using a dataset of 96 Universal Beam (UB) sections that was compiled in accordance with the British Standard **BS 4-1:1993**. Section properties were sourced from three reputable platforms: **Tata Steel Europe's structural sections catalogue**, **Dlubal Software's online cross-section properties tool**, and the **SkyCiv Structural Steel Section Database**. Tata Steel Europe is a leading UK steel manufacturer, and its published section profiles reflect real, commercially available products. Dlubal Software GmbH is a recognized provider of structural analysis software, offering accurate mechanical and geometric properties that conform to BS 4-1:1993. SkyCiv, a widely used cloud-based structural engineering platform, provides a comprehensive and accessible database of standardized steel sections with verified cross-sectional data. Combining these three sources ensured both consistency with the British Standard and reliability in the data used for structural calculations and modeling. The study combines analytical computations of plastic moment capacity and deflection under a specific three-point loading condition with a Random Forest predictive model. The goal is to identify optimal beam configurations that maximize structural performance while minimizing cost and deflection, thereby supporting more efficient design decisions.

The motivation for this work arises from the need to bridge traditional engineering methods with modern data-driven approaches, enabling engineers to explore beyond standard beam profiles and optimize designs in a more comprehensive and cost-effective manner.

1.2 Thesis Objectives

The primary objective of this thesis is to develop a methodology that combines classical structural engineering principles with machine learning to optimize the selection of simply supported I-beams. The goal is not only to identify the best-performing beams within an existing database but also to use predictive modeling to suggest new beam configurations that may offer improved performance and cost-effectiveness.

To achieve this, the following specific objectives were established:

1. **Data Acquisition and Analysis**: Utilize a real-world dataset of 96 Universal Steel I-beams was compiled using section data from **Tata Steel Europe**, including geometric dimensions, mass per meter, and cost estimates. Additional mechanical properties and section parameters were referenced from reputable sources such as **structural steel design handbooks** (e.g., OneSteel Manufacturing, 2013), the **Dlubal Software cross-section database**, and the **SkyCiv Structural Section Database**. All sources adhere to the **British Standard BS 4-1:1993**, ensuring consistency and reliability in the data used for modeling and structural analysis.

   The cost data for various universal beams used in this study were obtained from Steel Beams Direct, a structural steel supply and fabrication company in the UK (Steel Beams Direct, 2025).
2. **Mechanical Calculations**: Compute plastic moment capacity and beam deflection under a three-point loading scenario for each beam using established engineering formulas. These calculations were verified using both Python and Excel implementations.
3. **Development of a Scoring System**: Formulate an integrated scoring metric that accounts for plastic moment, deflection, and cost. This metric serves as a unified indicator of overall beam performance.
4. **Machine Learning Model Training**: Train a Random Forest regression model on the existing beam dataset to predict plastic moment and deflection based on input dimensions.
5. **Beam Optimization through Prediction**: Use the trained model to predict performance metrics for beam configurations outside the original dataset, and evaluate them using the scoring system to identify high-performing synthetic beams.
6. **Validation and Interpretation**: Analyze model accuracy using statistical performance metrics, compare synthetic suggestions with traditional design choices, and assess the applicability of machine learning in structural optimization.
7. **Documentation and Recommendations**: Provide insights on how this integrated approach can assist structural engineers in making better-informed

design decisions, along with recommendations for future research and enhancements.

This study intends to demonstrate how machine learning can be integrated with core engineering principles to enhance structural decision-making processes, particularly in the early stages of design when multiple trade-offs must be considered.

1.3 Scope of the Study

This study focuses on the structural optimization of **simply supported universal steel I-beams** under **static, three-point loading** over a **10-meter span**. The research leverages both classical structural engineering theory and modern machine learning methods—specifically the **Random Forest algorithm**—to evaluate and propose optimal beam designs based on a multi-objective scoring system.

The main areas included in the scope are:

- **Beam Type**: The study is limited to **universal I-beams** (also known as UB sections) manufactured according to UK steel standards. No other structural shapes, such as H-beams, T-beams, or hollow sections, are included.
- **Support and Loading Conditions**: All beams are assumed to be **simply supported** and subjected to **three equally spaced point loads** across a 10-meter span. The dynamic effects, load variations over time, or environmental factors such as temperature and corrosion are outside the scope.
- **Material Assumptions**: All beams are assumed to be composed of **standard structural steel** with consistent mechanical properties—particularly an elastic modulus (E) of 210,000 MPa and a yield strength of 350 MPa.
- **Geometric Inputs**: The optimization focuses on four geometric parameters: **Depth (mm), Flange Width (mm), Flange Thickness (mm), and Web Thickness (mm)**. Other properties like corner radius or flange slope are excluded.
- **Mechanical Performance Metrics**:
    - **Plastic Moment**: Calculated analytically for each beam.
    - **Deflection**: Computed using classical beam theory under the specific three-point load setup.
    - **Second Moment of Area (I_z)**: Not predicted by the model, but computed using a derived formula for each synthetic combination.

- **Cost and Weight**: These are included as real-valued attributes in the database and considered during optimization, but are not recalculated for synthetic beams—an average cost and weight per meter is assumed unless extended data is available.
- **Machine Learning Modeling**:
    - Limited to **Random Forest Regression** as the main predictive model.
    - Input features include the four geometric parameters.
    - Targets include plastic moment and deflection.
    - No classification or deep learning techniques are applied.
- **Optimization Strategy**: Based on a scoring system combining cost, plastic moment, and deflection using normalized values and assigned weights. Multi-objective optimization is carried out by calculating scores for predicted beams, rather than solving formal optimization equations.
- **Verification and Validation**: Calculations were verified using both **Python scripting** and **Excel formulas** to ensure consistency and transparency. Model performance is validated using test data and interpreted using R^2 and feature importance metrics.

This scope ensures that the thesis maintains a balance between depth and focus. While it does not address every complexity in structural design, it provides a solid and practical framework that combines data-driven modeling with engineering analysis to guide better beam selection.

1.4 Methodology Overview

The methodology adopted in this thesis integrates **classical structural engineering principles** with **machine learning modeling**, aiming to optimize beam selection for both **performance** and **cost-efficiency**. The process is divided into multiple stages, each contributing to the overarching goal of identifying both existing and potential I-beam configurations that offer high structural utility under specified loading conditions.

The methodology can be outlined as follows:

- **Step 1: Dataset Collection and Review**

A comprehensive dataset comprising **96 universal I-beams** manufactured in the UK is utilized. This dataset includes:

- Cross-sectional properties: depth, flange width, flange thickness, and web thickness (in mm).
- Cost per meter (in Euros).
- Mass per meter (in kg/m).
- Manufacturer-provided Second Moment of Area (I_z).

This dataset forms the backbone for analysis, modeling, and optimization.

- **Step 2: Engineering Calculations**

For each beam in the dataset:

- **Plastic Moment** is calculated using the plastic section modulus method.
- **Deflection** is calculated under **three-point loading conditions** using classical beam deflection theory.
- **Second Moment of Area (I_z)** is also derived using a formula applied to the geometric properties of the I-beams, serving as a cross-verification of the manufacturer's values.

All these calculations were implemented in **Python** and then **validated against Microsoft Excel**, ensuring consistency and correctness.

- **Step 3: Scoring System Development**

To compare beams effectively, a **scoring metric** is introduced that integrates:

- Plastic Moment (performance)
- Cost (economic efficiency)
- Deflection (serviceability)

Each component is **normalized** to ensure comparability, and weights are assigned based on design priorities. The resulting **score** enables ranking of beams within the existing dataset and provides a foundation for evaluating machine learning-generated beams later on.

- **Step 4: Machine Learning Model Training**

A **Random Forest Regressor** model is trained using:

- Input features: Depth, Flange Width, Flange Thickness, Web Thickness.
- Target outputs: Plastic Moment and Deflection.

The model is trained on 80% of the dataset and tested on the remaining 20% using **train-test splitting**. Hyperparameters are tuned, and $\boldsymbol{R^2}$ **scores** are used to assess performance.

- **Step 5: Synthetic Beam Prediction**

Once trained, the Random Forest model is used to predict performance for a **range of synthetic beam configurations**:

- New combinations of depth, flange width, flange thickness, and web thickness are generated within the observed minimum and maximum bounds.
- These are fed into the model to predict Plastic Moment and Deflection.
- I_z is calculated directly from geometry using a formula.
- The cost is approximated using average dataset values, unless detailed material pricing is available.

- **Step 6: Scoring and Ranking Synthetic Beams**

The scoring formula is applied to these synthetic beams to rank them similarly to the original dataset. This identifies **new high-performing beam configurations** not previously present in the manufacturer database.

- **Step 7: Result Interpretation and Validation**

The final results are visualized through:

- Beam rankings
- Comparative plots of deflection, plastic moment, and cost
- Feature importance from the machine learning model
- Sensitivity analysis of the scoring weights

This integrated approach enables both **data-driven insight** and **engineering-grounded validation**, allowing a balance between theory, modeling, and practical application. The following chapters will elaborate each step with full technical and procedural detail.

1.5 Thesis Structure

This thesis is organized into ten chapters, each focusing on a specific component of the study, leading the reader from theoretical foundations to model development, analysis, and conclusions.

- **Chapter 1: Introduction**

This chapter introduces the motivation, objectives, scope, and methodology of the study. It establishes the relevance of integrating classical beam design with machine learning and outlines how the research addresses the problem of optimizing simply supported I-beams in terms of performance and cost.

- **Chapter 2: Literature Review**

This chapter explores previous work related to beam mechanics, structural optimization, and machine learning in structural engineering. It covers traditional approaches for calculating plastic moment and deflection, common optimization methods in beam design, and recent advancements using artificial intelligence in civil and structural applications.

- **Chapter 3: Theoretical Framework and Engineering Principles**

Here, the fundamental principles of beam design are discussed in detail. It includes derivations and explanations for plastic moment, Second Moment of Area, and deflection formulas under point loading. The section also explores the theoretical relationship between cost and structural performance in optimization scenarios.

- **Chapter 4: Methodology**

This chapter outlines the step-by-step methodology followed throughout the thesis. It covers dataset collection, preprocessing, engineering calculations, Random Forest model training, synthetic beam generation, and scoring system development. Each step is described with its respective rationale and relevance to the study's objectives.

- **Chapter 5: Data Analysis and Results**

This chapter presents the key results of the analytical and modeling process. It includes statistical summaries of the original dataset, calculated deflections and plastic moments, Random Forest model performance, and the ranking of both original and synthetic beams according to the developed scoring system.

- **Chapter 6: Discussion**

Here, the results are interpreted in the context of structural engineering. The accuracy of predictions, performance of optimized beams, and trade-offs between cost, deflection, and plastic moment are explored. Comparisons with traditional design strategies are made, and limitations of the study are discussed.

- **Chapter 7: Conclusion**

This chapter summarizes the main contributions and findings of the research. It reflects on the success of the methodology, outlines how the objectives were met, and offers suggestions for future research, including broader datasets, additional features, or alternative modeling techniques.

- **Chapter 8: References**

A complete list of all books, journal articles, websites, and tools referenced throughout the thesis, formatted according to the required citation style.

- **Chapter 9: Appendices**

This chapter includes supplementary material such as:

- The full 96-beam dataset.
- Detailed calculation sheets (Excel).
- Python code used for modeling and optimization.
- Additional plots, graphs, and charts not included in the main chapters.

This structured progression ensures that the reader can follow the development of the study from background to implementation to insight, while providing both technical depth and practical relevance.

2. LITERATURE REVIEW

2.1 I-Beams in Structural Engineering

2.1.1 Properties of I-Beams

I-beams, also known as universal beams (UB), wide flange beams, or rolled steel joists (RSJ), are among the most commonly used structural elements in the construction industry. Their distinctive "I" shape offers a highly efficient cross-sectional profile that maximizes strength while minimizing material use. The two horizontal flanges resist most of the bending moment experienced by the beam, while the vertical web primarily resists shear forces.

The structural efficiency of an I-beam comes from its **high Second Moment of Area about the major axis (I_z)**. By distributing more material away from the neutral axis—through the flanges—the beam offers greater resistance to bending with less material. This not only enhances load-carrying capacity but also reduces weight and cost.

Key geometric parameters of I-beams include:

- **Depth (d)** – the total height from the top to the bottom of the beam.
- **Flange Width (b_b)** – the horizontal width of the flanges.
- **Flange Thickness (t_b)** – the vertical thickness of the top and bottom flanges.
- **Web Thickness (t_w)** – the horizontal thickness of the vertical web.
- **Root radius or fillet** – the curvature connecting the flange to the web, which affects local stress concentration.

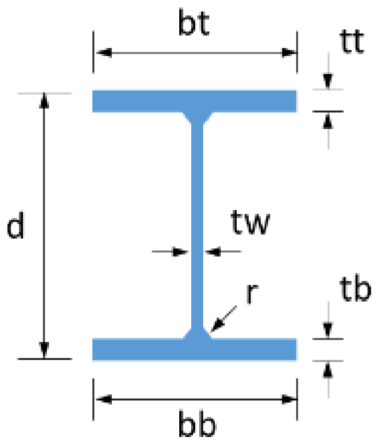


Figure 2.1 - Beam Parameters

Mechanical properties such as **yield strength, Young's modulus (*E*),** and **plastic modulus** further dictate the performance of the beam under various load conditions. When selecting a beam, these mechanical properties are typically considered alongside geometric properties to determine the beam's suitability for a particular structural application.

I-beams are especially favored in **simply supported span configurations**, which are common in floors, bridges, and roof systems. Their performance is typically governed by the interaction between bending stress, shear stress, and deflection, each of which depends heavily on the cross-sectional shape and material properties.

2.1.2 Applications of I-Beams in Construction

I-beams are fundamental to modern construction and are widely used in both residential and industrial structures. Their high strength-to-weight ratio makes them ideal for a range of applications, including:

- **Floor and ceiling joists** in commercial buildings, where long spans are necessary with minimal support.
- **Bridge girders**, particularly in steel or composite bridge decks.
- **Framing in high-rise buildings**, where vertical and lateral load resistance is critical.
- **Support members in prefabricated steel structures**, enabling quick assembly and standardized component use.

In infrastructure projects, the use of I-beams allows for long spans with minimal deflection, which helps in minimizing the number of supports required, reducing construction time and material costs. They are often selected based on both static and dynamic loading conditions and are customized for different spans, load intensities, and serviceability requirements.

Furthermore, I-beams are used in **modular steel structures** where beams are prefabricated and assembled on-site, significantly improving the efficiency of construction. The modular approach is particularly relevant in rapidly urbanizing environments where speed and precision are paramount.

Advancements in computer-aided design (CAD) and finite element analysis (FEA) have made it easier for structural engineers to model and simulate the behavior of I-beams under various loading conditions, aiding in the optimization of beam selection.

Despite their versatility, the selection of I-beams still often follows standard catalog options offered by steel manufacturers. These catalogs list various sizes and properties, but may not always offer the most optimal beam for a given use-case in terms of cost-efficiency, performance, and sustainability.

This limitation motivates the use of **computational optimization techniques**, such as machine learning, to explore and suggest custom I-beam configurations that are not only structurally sound but also more cost-effective or efficient in terms of deflection and weight.

2.2 Plastic Moment and Deflection

Understanding the concepts of **plastic moment** and **deflection** is essential in structural analysis and design. These two parameters govern the beam's **strength** and **serviceability** respectively. Together, they offer a comprehensive picture of how a beam will behave under load, and they are central to any optimization framework that aims to improve both performance and economy.

2.2.1 Bending Stress and Plastic Moment

Elastic vs. Plastic Behavior

In structural engineering, beams under bending exhibit **elastic behavior** up to a certain stress limit—typically referred to as the **yield stress**. Within this range, stress is directly proportional to strain according to Hooke's Law, and the beam returns to its original shape when the load is removed.

However, once the applied moment exceeds the **yield moment**, parts of the cross-section begin to yield, entering the **plastic region**. Eventually, a point is reached where the entire cross-section yields uniformly—this is called the **plastic moment** (denoted as M_p), and it represents the **maximum moment-carrying capacity** of the beam before a plastic hinge forms.

Plastic Section Modulus

The plastic moment is calculated using the **plastic section modulus** (Z_p) and the yield strength (f_y).

Unlike the elastic section modulus (Z_e), which only accounts for the portion of the beam in the elastic range, Z_p assumes that the entire cross-section yields, offering a more conservative and realistic measure of ultimate capacity.

Importance in Structural Design

The plastic moment is critical in **limit state design**. It defines the point at which the beam can no longer carry increased moment without undergoing significant deformation. In optimization, a higher plastic moment typically implies better performance, but often comes at the cost of increased material usage, and hence higher weight and cost.

2.2.2 Beam Deflection Under Load

Serviceability Limit State

While strength governs safety, **deflection** governs serviceability. Excessive deflection can cause structural instability, cracking in finishes, or discomfort for building occupants. Therefore, even if a beam is strong enough, it may still be unsuitable if its deflection exceeds acceptable limits.

Structural deflections can result from various factors, including applied loads, temperature changes, construction inaccuracies, or ground settlement. Controlling deflection is essential to maintain structural integrity, avoid damage to brittle finishes like glass or plaster, and ensure the structure feels safe to occupants. In statically indeterminate structures, calculating deflection at key points is crucial for analysis. While bending is typically the main cause of deflection in beams and frames, axial forces are the primary source in trusses. (Hibbeler, 2012).

Deflection Theory

For simply supported beams under point loads, the maximum deflection can be calculated using classical beam theory. For example, in the case of multiple-point loads, the deflection is a summation of individual effects due to each load acting at its position.

Practical Considerations

Deflection limits are typically set by building codes, often as a fraction of the span (e.g., L/250 or L/360).

2.3 Structural Optimization Techniques

The design of structural elements, especially beams, involves balancing multiple competing objectives—typically strength, stiffness, weight, and cost. Structural optimization seeks to achieve this balance by identifying the "best" design configuration within a given set of constraints. Over the years, researchers and engineers have employed a variety of optimization techniques, ranging from classical deterministic approaches to advanced metaheuristics and, more recently, machine learning models.

2.3.1 Classical Beam Design and Optimization

Traditional Design Approach

Traditionally, structural beams are designed using closed-form analytical equations derived from mechanics of materials. Engineers typically:

- Calculate internal actions (bending moment, shear force).
- Select a beam from standardized catalogs.
- Check if the chosen section satisfies the limit states for strength and deflection.

While effective and code-compliant, this method is **manual, iterative, and often suboptimal**. Engineers may oversize beams to ensure safety and code compliance, which leads to unnecessary material usage and increased costs.

Deterministic Optimization Methods

Deterministic optimization methods aim to automate and improve upon the traditional trial-and-error process. Some commonly used techniques include:

- **Linear Programming (LP)**: Suitable for problems with linear objective functions and constraints. Rarely used in structural design due to non-linearity in stress-strain behavior.
- **Non-Linear Programming (NLP)**: Handles nonlinear relationships (e.g., in plastic behavior or deflection), but can be computationally expensive and sensitive to initial guesses.
- **Gradient Descent Methods**: Effective in smooth, continuous design spaces but prone to getting stuck in local minima.
- **Integer and Mixed-Integer Programming**: Useful when design variables must be selected from discrete sets (such as catalogued beam sizes), but typically slower for large datasets.

These techniques provide valuable insights but often struggle with **multi-objective** problems and discrete datasets, such as standard I-beam catalogues with fixed size combinations.

Limitations of Classical Methods

Classical optimization tools often optimize **only one criterion at a time**—e.g., minimize weight or maximize stiffness. In practical structural design, however,

multiple criteria must be balanced: structural performance, deflection limits, cost constraints, and availability of profiles.

Moreover, classical approaches require simplifying assumptions (e.g., ignoring interaction effects or assuming uniform material distribution) that may reduce solution accuracy.

2.3.2 Integrating Cost, Deflection, and Plastic Moment

Multi-Criteria Decision Making (MCDM)

Modern structural optimization problems frequently involve **multi-criteria decision making**, in which trade-offs are made between conflicting goals. For example, a heavier beam might offer higher plastic moment capacity, but at a higher cost and with worse deflection behavior.

Some popular MCDM methods include:

- **Weighted Sum Method (WSM)**: Each criterion is assigned a weight, and an overall score is calculated by summing the weighted values. This is the method adopted in this thesis.
- **Analytic Hierarchy Process (AHP)**: A pairwise comparison-based approach where criteria are ranked based on decision-maker preferences.
- **TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution)**: Ranks alternatives based on their proximity to an ideal solution.
- **Pareto Optimization**: Identifies a set of “non-dominated” solutions where no objective can be improved without worsening another.

Cost-Deflection-Strength Trade-Off

In beam optimization, the three critical criteria are often:

- **Cost per meter**: Influenced by steel grade, section size, and market prices.
- **Plastic moment**: Directly affects load-bearing capacity and structural safety.
- **Deflection**: Determines serviceability and user comfort.

Optimizing all three simultaneously is **non-trivial**. For instance:

- A deeper beam increases plastic moment and reduces deflection, but also increases cost and mass.
- A thinner web or flange might reduce cost and mass but compromises strength and increases deflection.

2.4 Machine Learning in Structural Engineering

The application of **machine learning (ML)** in structural engineering represents a transformative shift in how we approach design, analysis, and optimization. Traditionally, structural design has been governed by analytical methods grounded in mechanics and empirical code-based rules. While these are highly effective and reliable, they can be restrictive when faced with the complexity of multi-objective optimization, data-driven insights, and large-scale design spaces.

With the growth of computational power and the availability of extensive structural datasets, machine learning offers new avenues for discovering patterns, making predictions, and enhancing decision-making processes in structural design.

2.4.1 Machine Learning Basics

Definition and Scope

Machine learning is a subset of artificial intelligence that allows systems to learn patterns from data and make predictions or decisions without being explicitly programmed for each task. Unlike traditional algorithms that follow rigid logic, machine learning adapts to data structures and is particularly effective when dealing with nonlinear relationships or high-dimensional datasets.

A typical AI algorithm development involves four main stages: data preparation, model design, training and evaluation, and deployment. Data preparation—acquiring, analyzing, and preprocessing data—is crucial and often time-consuming, as data quality and quantity greatly impact model performance. After designing the model, it undergoes training and evaluation, often requiring iterative adjustments. Once finalized, the model is deployed for use (Málaga-Chuquitaype, 2022).

Machine learning algorithms are generally classified into:

- **Supervised Learning** – The algorithm learns from labeled input-output pairs. It’s ideal for regression (predicting continuous values) or classification (predicting categories).
- **Unsupervised Learning** – The algorithm identifies patterns or groupings in data without known outputs (e.g., clustering).
- **Reinforcement Learning** – The system learns by interacting with an environment and receiving rewards or penalties.

Regression Models in Structural Applications

In structural engineering, regression models have been used to:

- Predict material properties (e.g., compressive strength of concrete)
- Estimate load capacities of structural components
- Identify failure points or risks
- Optimize cross-sectional dimensions and reinforcement layouts

Machine learning can also be employed to **augment finite element analysis (FEA)** by predicting stresses or displacements, reducing the need for time-consuming simulations in the early stages of design.

2.4.2 Random Forest Algorithm

Overview

The **Random Forest** algorithm is a type of **ensemble learning model** built from a collection of decision trees. Each tree makes predictions independently, and the final output is the average (for regression) or the majority vote (for classification). This ensemble approach helps reduce overfitting and improves generalization.

It combines multiple decision trees, each trained on different subsets of data, and averages their predictions to improve accuracy. Introduced by Ho in 1995 and later developed by Breiman through bagging and random feature selection, RF is fast, easy to implement, and highly reliable. Its versatility in handling both classification and regression tasks makes it a widely used machine learning technique (Dabiri, Rahimzadeh, & Kheyroddin, 2022).

The key principles of Random Forest include:

- **Bootstrap Aggregation (Bagging):** Each tree is trained on a random sample of the training data.
- **Random Feature Selection:** At each split in the tree, only a random subset of features is considered.
- **Voting or Averaging:** Predictions from all trees are aggregated to generate the final result.

These features make Random Forests:

- Highly resistant to overfitting
- Capable of modeling complex nonlinear relationships
- Robust to noisy or incomplete data

Why Random Forest for Beam Optimization?

In the context of beam optimization, Random Forest is particularly suited because:

- The relationship between beam dimensions and structural behavior (moment, deflection) is **nonlinear**.
- The dataset includes **interdependent variables** (e.g., flange thickness affects both plastic moment and deflection).
- Interpretability is desired: Random Forest allows engineers to examine **feature importance**, showing which dimensions most influence performance.

Real-world datasets often contain errors, gaps, or missing values due to practical or privacy-related limitations. Many machine learning algorithms assume complete data and struggle with such imperfections. However, Random Forests include built-in strategies to handle missing data effectively. The Scikit-learn library in Python supports this approach by offering a broad range of user-friendly machine learning tools, including algorithms, preprocessing methods, and model selection techniques. Its goal is to provide efficient and accessible implementations of established ML methods for both experts and non-experts across various scientific domains. (Louppe, 2014).

2.5 Cost-Effectiveness in Structural Design

In modern engineering practice, optimal design is not solely about achieving the highest performance; it is about doing so within the limits of **budget**, **material availability**, and **constructability**. Therefore, **cost-effectiveness** is a cornerstone of any structural optimization process, particularly when working with widely used elements such as **I-beams**. This section discusses the financial dimensions of beam design, the factors influencing cost, and how previous research has approached cost minimization while maintaining performance.

2.5.1 Factors Affecting Cost

Material Costs

The **unit cost of steel** (typically in €/kg or €/m) varies by region, steel grade, production method, and market conditions. For standard I-beams, the price is often calculated **per meter** based on cross-sectional area and density.

Cost is influenced by:

- **Grade of steel** (e.g., S235, S355): Higher-grade steel typically costs more but offers increased strength.
- **Beam dimensions**: Deeper beams or thicker flanges require more material.
- **Supply chain logistics**: Custom sections or rare dimensions increase procurement cost.

Fabrication and Transportation

Even if two beams have identical material costs, their **fabrication costs** may differ based on complexity. Deeper or heavier beams require more labor to cut, weld, drill, and assemble. Similarly, transportation costs increase with **mass per meter** or **total shipment weight**.

This makes **mass** a proxy for estimating total lifecycle cost, especially for large-scale projects involving hundreds or thousands of meters of beams.

Design Choices Affecting Cost

Choices made by the designer, such as:

- Reducing flange thickness to save on weight,
- Selecting narrower widths to simplify fabrication,
- Avoiding uncommon profiles,

…all have cost implications. However, aggressive cost-cutting can compromise performance, especially in **plastic moment** or **deflection** behavior.

2.5.2 Previous Work on Cost Optimization

Traditional Cost-Based Design

Historically, cost-based optimization in structural engineering relied on:

- **Manual selection** of beam sections based on load demand and standard catalogs.
- **Simple weight-based cost estimation** using steel density and beam dimensions.
- **Comparative studies** using fixed loading conditions and selecting the beam with the lowest cost that still meets strength and deflection limits.

While effective in small-scale applications, these methods lack adaptability and don't explore alternate geometries outside standard catalogs.

Computational Cost Optimization

As computational tools developed, researchers began exploring:

- **Nonlinear optimization** with cost as the objective function and performance as constraints.
- **Genetic algorithms** and **evolutionary strategies** to explore a wider search space.
- **Mixed-integer programming** to select from discrete beam profiles.
- **Pareto front analysis**, balancing multiple cost-related objectives such as mass, stiffness, and fabrication time.

However, these methods still required **extensive computation** or were **limited in scope** to predefined geometric boundaries.

Data-Driven Approaches

Recent advancements in **data science** have introduced **machine learning** as a tool for cost-informed design. These include:

- Predicting **cost-performance trade-offs** using regression models trained on past projects.
- Using **multi-objective scoring systems**, as in this thesis, to generate composite performance metrics.
- Incorporating **probabilistic cost modeling** to simulate material price fluctuations or fabrication delays.

3. THEORETICAL FRAMEWORK AND ENGINEERING PRINCIPLES

This chapter outlines the theoretical and engineering principles underpinning the study, providing the analytical foundation for the later computational modeling and optimization. While the Literature Review discussed existing research on beam design and machine learning applications, the current chapter focuses on the fundamental mechanics of I-beams—covering plastic moment, deflection, and cost estimation—which form the core of the thesis methodology.

3.1 I-Beam Design and Parameters

3.1.1 Material Properties and Dimensions

The structural performance of an I-beam is largely determined by its **cross-sectional geometry** and the **mechanical properties** of the material it is made from—typically structural steel. Each of these characteristics plays a specific role in governing how the beam resists bending and deformation.

Steel Material Properties

The beams in this study are assumed to be fabricated from **mild or medium-strength structural steel**, most commonly **S275 or S355**, as defined by the European Standard EN 10025. The following properties are assumed unless otherwise stated:

- **Young's Modulus (E):** 210 GPa — determines the stiffness of the beam, and directly affects deflection.
- **Yield Strength (f_y):** 275–355 MPa — used in plastic moment calculations.
- **Poisson's Ratio (ν):** 0.3 — although not directly used in this study, relevant in advanced analysis.
- **Density (ρ):** 7850 kg/m³ — used to compute mass per meter, which influences cost and transport.

These material properties are considered **constant across all beam sections**, which allows us to focus the optimization on geometry, not material variation.

Geometric Parameters of I-Beams

Each I-beam profile is defined by the following key dimensions:

- **Depth (D):** The vertical distance from top to bottom of the cross-section.
- **Flange Width (b_b):** The horizontal width of the top and bottom flanges.
- **Flange Thickness (t_b):** The thickness of each flange.
- **Web Thickness (t_w):** The thickness of the vertical web connecting the flanges.

These parameters are often denoted by symbols such as D, $\boldsymbol{b_b}$, $\boldsymbol{t_b}$, and $\boldsymbol{t_w}$ in structural literature, but for clarity and consistency with the dataset used in this research, the following notations are adopted throughout:

Table 1 - Parameters and Symbols

Parameter	Symbol (used in thesis)	Description
Depth	Depthmm	Full vertical height of the beam
Flange Width	FlangeWidthmm	Width of the top/bottom flange
Flange Thickness	FlangeThicknessmm	Thickness of each flange
Web Thickness	WebThicknessmm	Thickness of the central web

These four geometric variables are the **primary input features** used both for calculation and in the machine learning model.

Assumptions and Constraints

In this thesis:

- All beams are **simply supported** over a span of 10 meters.
- Cross-section is **symmetric**, meaning flanges are identical and centered about the web.
- The flanges and web are assumed to remain **plane and unyielded** until yielding initiates.
- Material homogeneity and isotropy are assumed for simplified analysis.

These assumptions are standard in early-stage structural analysis and are sufficient for calculating deflection and plastic moment within practical engineering tolerances.

Beam Section Modulus

One of the most fundamental geometric properties governing a beam's ability to resist bending is its **section modulus**. The section modulus directly relates the applied bending moment to the maximum bending stress in a section, and is a key determinant of strength in both elastic and plastic design.

Elastic Section Modulus (Z_e)

The **elastic section modulus** is calculated using the formula (Hibbeler, 2012):

$$Z_e = \frac{I}{y} \tag{3.1}$$

Where:

- I = second moment of area (Second Moment of Area) about the major axis
- y = distance from the neutral axis to the extreme fiber

For I-beams, the elastic section modulus is commonly used in **serviceability checks** and **elastic bending calculations**, particularly when the applied moment is below the yield threshold of the material.

However, in this research, **plastic behavior is of more interest**, especially in high-load applications, so the focus shifts to the **plastic section modulus**.

Plastic Section Modulus (Z_p)

The **plastic section modulus** represents the section's capacity at the onset of plastic failure, where the entire cross-section has yielded. It is used in calculating the **plastic moment (M_p)**, which is the ultimate bending capacity of a section before collapse (Hibbeler, 2012).

$$M_p = f_y \cdot Z_p \tag{3.2}$$

Where:

- M_p is the plastic moment (kNm)
- Z_p is the plastic section modulus (mm^3)
- f_y is the yield strength of the steel (typically 350 MPa)

For a **doubly symmetric I-beam**, the plastic section modulus is calculated by summing the contributions of the flanges and the web, taking into account their distances from the plastic neutral axis. The general formulation (Hibbeler, 2012) is:

$$Z_p = 2 \cdot A_f \cdot (\frac{h}{2} - \frac{t_b}{2}) + A_w \cdot \frac{h_w}{4} \tag{3.3}$$

Where:

- $A_f = b_b \cdot t_f$ = area of one flange
- $A_w = t_w \cdot h_w$ = area of the web
- b_b = flange width
- t_b = flange thickness
- t_w = web thickness
- h = overall section depth
- $h_w = h - 2t_b$ = clear height of the web

This formula accounts for:

- Two flanges contributing equal and opposite plastic moments.
- The web contributing through a rectangular area centered at the PNA.

In this research, we assumed a uniform yield strength of **350 MPa** for all beams in the dataset, which is typical for structural-grade steel used in heavy construction.

This formula is simplified and adapted for computational use in the Python code used in this study, as described later in the methodology section.

Practical Relevance of Section Modulus

The section modulus allows engineers to:

- Estimate the moment capacity of a beam quickly.
- Select cross-sections that balance strength with economy.
- Compare different beam geometries on a normalized basis.

Moreover, the plastic section modulus forms the basis for **one of the three pillars** of the scoring system used in this thesis: alongside cost and deflection.

3.2 The Integration of Plastic Moment

In the structural design of steel beams, one of the most critical checks is ensuring that the **section has sufficient moment capacity** to resist the applied bending forces. The **plastic moment capacity**, M_p , represents the **maximum moment** a cross-section can resist when it has fully yielded — i.e., when every fiber across the depth has reached the yield stress. Unlike elastic analysis, which assumes a linear stress distribution and reversible deformation, plastic analysis accepts that steel can redistribute internal forces beyond initial yield, allowing **greater load-carrying capacity** before collapse.

The concept of plastic moment is especially important in **limit state design**, where structures are designed to withstand maximum expected loads without collapsing.

3.2.1 Theory and Formulae for Plastic Moment

Plastic Behavior of Steel Sections

In elastic bending, the neutral axis divides the beam cross-section into compressive and tensile zones, with stress varying linearly from zero at the neutral axis to a maximum at the outer fibers. However, once the moment increases beyond the yield limit, the stress distribution begins to plateau, and the section starts to yield.

Eventually, under increasing moment, the entire section yields, resulting in a **uniform stress distribution** across the section. At this stage, the section forms a **plastic hinge**, and any additional rotation occurs without increase in moment.

3.2.2 Application to the Dataset

Plastic Moment Calculation Using Python

To compute the plastic moment for each of the 96 beams in the dataset, the geometric parameters of each beam were extracted from the database: Depthmm, FlangeWidthmm, FlangeThicknessmm, and WebThicknessmm. These were used to calculate the **plastic section modulus** for each beam using the formula described above. Python libraries like pandas, NumPy, and Matplotlib facilitate robust statistical analysis, data exploration, and the creation of insightful visualizations. Coupled with powerful machine learning and AI frameworks such as TensorFlow and scikit-learn, Python enables the development of sophisticated predictive models that enhance decision-making across diverse business contexts. (Pharell, 2023).

This was automated using Python, where a function was written to apply the following formula across all rows:

```
[81] df["PlasticMomentKNM"] = (
        (
            2 * (df["FlangeWidthmm"] * df["FlangeThicknessmm"] * ((df["Depthmm"] /
            2) - (df["FlangeThicknessmm"] / 2))) +
            (df["WebThicknessmm"] * (df["Depthmm"] - 2 * df["FlangeThicknessmm"]) *
            ((df["Depthmm"] - 2 * df["FlangeThicknessmm"]) / 4))
        ) * 10**-9 * 350 * 1000
    )
```

Figure 3.1 - Plastic Moment Formula Code

This equation computes the sum of the moments from both flanges and the web, multiplied by the assumed yield strength (350 MPa) and converted from N·mm to kN·m.

Verification Against Manufacturer Data

To ensure the model's accuracy, several calculated values of plastic moment were **compared against reference values** where available or cross-checked using manual calculations in **Microsoft Excel**. The agreement between these confirmed the **correct implementation of the formulas**.

Use in Scoring System

The calculated plastic moment becomes one of the three **core components** in the scoring metric used for optimization:

$$\boldsymbol{Score} = \left(\frac{\boldsymbol{M_p}}{\boldsymbol{Cost}}\right) - \boldsymbol{Deflection} \tag{3.4}$$

This approach ensures that beams with **high bending capacity relative to their cost** are favored, as long as their deflection remains within acceptable bounds.

Significance in Design

Understanding and accurately computing the plastic moment allows engineers to:

- Select more **cost-effective** sections without compromising structural safety.
- Ensure compliance with **ultimate limit state** requirements.
- Compare standard and ML-predicted sections using a consistent strength criterion.

3.3 Deflection Calculation

The structural performance of beams is not solely determined by their load-carrying capacity (plastic moment), but also by their **stiffness and deflection behavior** under applied loads. Excessive deflection can compromise structural integrity, cause damage to non-structural elements, or produce unacceptable serviceability issues even if the beam does not fail structurally. This section focuses on the fundamental theory of beam deflection and how the specific point loads used in this study affect the total deflection of simply supported I-beams.

3.3.1 Theory of Beam Deflection

Basic Concepts

Deflection is the displacement of a beam under load, measured as the distance the beam moves vertically from its original unloaded position. The deflection depends on several factors:

- The type, magnitude, and position of the applied load(s)
- The **flexural rigidity** of the beam, given by the product $E \times I$, where:
    - E = Young's modulus of elasticity (stiffness of the material)
    - I = second moment of area (Second Moment of Area) of the beam's cross-section
- The beam's **span length** and support conditions

In the context of this thesis, the beams are assumed to be **simply supported** over a fixed span of **10 meters**.

Governing Differential Equation

The elastic curve of a beam under bending is governed by the differential equation:

$$\frac{d^2y}{dx^2} = \frac{M(x)}{EI} \qquad (3.5)$$

Where:

- y is the vertical deflection as a function of horizontal position x
- $\boldsymbol{M(x)}$ is the bending moment at location x

By integrating this equation with appropriate boundary conditions, the deflection profile of the beam can be derived.

Second Moment of Area (I)

The **Second Moment of Area** is a geometric property that quantifies how the cross-sectional area is distributed about the neutral axis. Larger values of I imply greater resistance to bending and smaller deflections. It is calculated for each beam using the formula based on its flange and web dimensions (as explained earlier in Chapter 3.1), with the **plastic Second Moment of Area** I_z calculated for plastic capacity.

Young's Modulus (E)

The modulus of elasticity, E, represents the material's intrinsic stiffness. For structural steel, E is typically taken as **210,000 MPa**. This value is assumed constant across all beams in the dataset.

Superposition Principle

When multiple loads are applied, the **principle of superposition** states that the total deflection is the algebraic sum of the deflections caused by each individual load, assuming linear elastic behavior.

This is particularly relevant in this study as the beams are subjected to **three point loads simultaneously**, and total deflection is calculated by summing their individual contributions.

3.3.2 Effects of Point Loads on Deflection

Simply Supported Beam with Point Load

For a single point load P applied at a distance a from the left support and $b = L - a$ from the right support on a simply supported beam of length L, the deflection at any point x can be calculated using classical beam theory.

The maximum deflection typically occurs near the point of load application and is given by:

$$\boldsymbol{\delta_{max} = \frac{P \cdot a \cdot b^2}{3EIL}} \tag{3.6}$$

(if load is closer to the left support)

The exact formula varies depending on the position of the load and where the deflection is evaluated.

Three Point Loads Used in This Study

In this research, the beams are subjected to **three point loads simultaneously**, representing realistic load scenarios that beams may encounter in practice. The loads and their positions are:

- Load $P_1 = 5000$ N at $x_1 = 2000$ mm from the left support
- Load $P_2 = 8000$ N at $x_2 = 6000$ mm
- Load $P_3 = 12000$ N at $x_3 = 8000$ mm

The total deflection at any position xxx is calculated by summing the deflections caused by each point load individually, applying the superposition principle.

Calculation Formula for Deflection Due to Multiple Loads

According to (Hibbeler, 2012), The vertical deflection δ at any point x on the simply supported beam under multiple point loads is given by:

$$\boldsymbol{\delta(x) = \sum_{i=1}^{n} \frac{Pab(L_i^2 - a_i^2 - b_i^2)}{6EIL}} \tag{3.7}$$

Where for each load i:

- $\boldsymbol{P_i}$ = magnitude of the load
- $\boldsymbol{a_i}$, $\boldsymbol{b_i}$ = distances from the load to the supports such that $\boldsymbol{a_i + b_i = L}$
- E and I are material and geometric properties as previously defined

The full expression for deflection at the position of maximum deflection accounts for beam length, load magnitudes, and load positions.

Implementation in Python

In this thesis, deflection due to the three simultaneous point loads is computed programmatically using Python. The algorithm:

- Computes the Second Moment of Area I for each beam based on its dimensions.
- Uses the formula for each point load's deflection.
- Adds the deflections from all three loads to obtain the total deflection at the point of interest.

This calculation ensures an accurate estimate of how much each beam will bend under realistic service loads.

Importance of Deflection in Design

While beams can theoretically carry loads up to their plastic moment capacity, excessive deflection can lead to:

- Cracking or failure of attached finishes (drywall, plaster, ceilings)
- Misalignment in connected structural elements
- Reduced occupant comfort in buildings

Therefore, deflection limits often control design decisions in serviceability limit states, making this parameter vital to the scoring system developed in this thesis.

3.4 Cost-Effectiveness Analysis

In modern structural engineering, selecting the most suitable beam involves balancing not only **strength and stiffness** but also the **cost** associated with materials and construction. An optimal design seeks to maximize structural performance while minimizing expenditure. This section delves into the methods used to calculate beam costs and explores how cost interacts with deflection and performance, providing the foundation for the multi-criteria scoring system developed in this thesis.

3.4.1 Cost Calculation Methods

Material Cost as a Primary Factor

The predominant contributor to the cost of steel beams is the **material cost**, typically expressed as a **cost per unit weight or length**. In this study, the dataset provides **cost per meter** of each beam, directly related to the beam's weight and the current market price of steel.

Material cost depends on:

- **Weight per meter (kg/m):** Heavier beams consume more steel and thus cost more.
- **Price of steel per kilogram:** Market-driven and variable by region and time.

The cost per meter, C, can be estimated as:

$$\boldsymbol{C = W \times P_s} \tag{3.8}$$

Where:

- W = weight of the beam per meter (kg/m)
- $\boldsymbol{P_s}$ = price of steel per kg (EUR/kg)

In this thesis, the **cost data was provided by the supplier** for each beam, reflecting realistic market values. Weight per meter was also included in the dataset, allowing for direct cost comparisons across the 96 beams.

Additional Cost Considerations

Beyond material cost, other factors influence total project costs, although these are often less straightforward to quantify at the beam level:

- **Fabrication costs:** Cutting, welding, drilling, and finishing steel components.
- **Transportation and handling:** Larger or heavier beams may require special equipment.
- **Installation complexity:** Heavier or less standard beams may increase labor costs.

While these were not explicitly modeled in this thesis due to limited data, they remain important considerations in practical engineering design.

Normalization and Use in Scoring

Since cost is in different units and scales compared to other parameters like plastic moment and deflection, it was **normalized** for use in the scoring system to ensure balanced contribution:

$$\boldsymbol{Normalized\ Cost} = \frac{\boldsymbol{C_i}}{max(\boldsymbol{C})} \tag{3.9}$$

Where $\boldsymbol{C_i}$ is the cost of beam i, and $max(\boldsymbol{C})$ is the maximum cost in the dataset.

This approach maintains cost proportionality while allowing integration with normalized performance metrics.

3.4.2 Relationship Between Cost, Deflection, and Performance

Balancing Cost and Structural Efficiency

Optimal structural design is a **multi-objective problem** where cost, strength, and serviceability must be balanced. For instance:

- A **heavy, thick beam** may offer excellent strength and low deflection but at a high cost.
- A **lighter, thinner beam** may be cheaper but suffer from higher deflection or insufficient load capacity.

The **challenge lies in identifying beams that achieve the best trade-off**, maximizing plastic moment per unit cost while keeping deflections within acceptable limits.

Deflection and Serviceability

Deflection relates directly to user experience and safety. While plastic moment ensures **ultimate strength**, deflection ensures **serviceability** — the beam should not deform excessively under normal use. High deflection can cause structural and non-structural damage even when ultimate failure is avoided.

Performance Metrics in Scoring

In this thesis, a **scoring formula** was developed to capture this trade-off:

$$\boldsymbol{Score = \left(\frac{Plastic\ Moment}{Cost}\right) - Deflection} \tag{3.10}$$

This formula integrates:

- **Plastic Moment / Cost:** Rewards beams with high strength at low material cost.
- **Deflection:** Penalizes beams with excessive deflection to maintain serviceability.

By subtracting deflection directly, the scoring system discourages beams that, while cost-effective and strong, deflect excessively.

Importance in Real-World Applications

This integrated approach aligns with industry priorities:

- Cost minimization is essential for **budget adherence**.
- Deflection limits ensure **long-term durability** and **user comfort**.
- Strength guarantees **safety under extreme loads**.

The methodology allows for practical beam selection, not just theoretically optimal but feasible in engineering practice.

Limitations and Assumptions

- The cost values reflect **material prices only**, excluding fabrication and logistics.
- Deflection calculations assume **linear elastic behavior** and perfect supports.
- The scoring formula is a simplified model, and in practice, design codes may impose additional checks.

Future work could refine cost models and include **life-cycle costing** or **maintenance factors** to enhance real-world applicability.

3.5 Optimization Criteria

In structural engineering, defining what constitutes an **optimal beam design** is a complex task that requires balancing multiple factors simultaneously. This section discusses the conceptual foundation of optimization in beam design, focusing on how "optimal" is defined in this research and the inherent trade-offs engineers face between cost, deflection, and plastic moment.

3.5.1 Defining "Optimal" in Structural Design

What is "Optimal"?

The concept of optimality in structural design varies depending on the project's goals, constraints, and priorities. Generally, an **optimal design** is one that satisfies the required safety and serviceability criteria while minimizing resource use and cost.

In this thesis, optimality is defined as the design solution that provides:

- The **maximum structural performance**, measured primarily by the **plastic moment capacity**, which represents the beam's ability to withstand bending without permanent deformation.
- The **minimum deflection under service loads**, ensuring that the beam does not deform excessively, thus maintaining structural integrity and serviceability.
- The **lowest material cost**, contributing to economic efficiency and project feasibility.

Multi-Objective Nature of Structural Optimization

Structural optimization often involves **multiple conflicting objectives**. For example:

- Increasing strength usually requires larger or thicker sections, which increases material cost.
- Reducing deflection often requires increasing stiffness, which similarly tends to increase weight and cost.
- Minimizing cost alone might lead to underperforming or unsafe designs.

Thus, optimal design must find the best **compromise** or **balance** among these criteria rather than optimizing any one in isolation.

Optimization in This Thesis

The approach taken is a **multi-criteria scoring system**, combining plastic moment, deflection, and cost into a single composite score. This allows ranking of beams and facilitates the identification of candidates that achieve a desirable balance, recognizing the practical necessity of compromise in engineering design.

3.5.2 Trade-offs Between Cost, Deflection, and Plastic Moment

Interdependence of Criteria

The three criteria—cost, deflection, and plastic moment—are interconnected through beam geometry and material properties. For instance:

- A **beam with a larger cross-section** will typically have a higher plastic moment and greater Second Moment of Area, reducing deflection, but it also weighs more and costs more.
- Conversely, a **lighter beam** will cost less but may not provide adequate strength or may deflect beyond acceptable limits.

This interplay requires careful consideration and quantitative methods to assess trade-offs explicitly.

Cost vs. Performance Trade-offs

Cost is usually the most restrictive factor in construction projects. The designer must ensure that the beam meets safety and deflection requirements without overspending. While some over-design might improve safety margins, it can be economically unsound.

The **score formulation** used, which normalizes plastic moment by cost and subtracts deflection, reflects this trade-off explicitly, encouraging designs that provide the most "value" per unit cost.

Deflection Constraints and Serviceability

Serviceability limits often dictate maximum allowable deflections to prevent damage or discomfort. Even a beam with excellent strength can be unsuitable if it deflects excessively.

This necessitates penalizing high-deflection beams, as done in the scoring system by subtracting normalized deflection values.

Plastic Moment and Safety

Plastic moment reflects the **ultimate load capacity** of a beam, a fundamental safety requirement. Ensuring that beams have sufficient plastic moment capacity guards against failure due to bending.

However, an optimal design balances this with cost and deflection, as focusing solely on ultimate capacity may result in uneconomical or impractical designs.

Graphical Interpretation

Trade-offs can be visualized with graphs plotting cost against plastic moment, deflection against cost, or three-dimensional plots involving all three parameters. Such visual tools assist in identifying “Pareto optimal” sets—solutions where improving one criterion cannot be achieved without degrading another.

Role of Machine Learning in Navigating Trade-offs

The use of **random forest machine learning** enables exploration of the multidimensional design space beyond the original database, predicting beam dimensions that could yield better trade-offs. This facilitates discovery of novel designs that conventional methods might overlook.

This framework sets the foundation for the **methodology** in Chapter 4, where these optimization criteria are applied practically using the dataset and machine learning models.

4. METHODOLOGY

4.1 Data Collection and Preprocessing

In any research involving data-driven modeling and optimization, the **quality, scope, and preparation** of the dataset play critical roles in determining the reliability and accuracy of the final results. This section elaborates in depth on the source and nature of the dataset used in this study, the steps taken to augment it with additional calculated parameters, and the preprocessing strategies employed to ensure that the data is clean, consistent, and suitable for subsequent machine learning applications.

4.1.1 Dataset Overview (96 I-Beams)

The core of this study is a comprehensive dataset of **96 Universal Beam (UB) sections**, selected in accordance with the **British Standard BS 4-1:1993**. Section properties were obtained from three complementary and authoritative sources: the *Handbook of Structural Steelwork* (7th ed., BCSA Publication No. 55/13, 2013), jointly developed by the **British Constructional Steelwork Association** and **Tata Steel**; **Dlubal Software's** online cross-section database; and the **SkyCiv Structural Steel Section Database**. All three sources provide standardized geometric and mechanical properties for steel sections conforming to BS 4-1:1993, enabling cross-validation and enhancing the reliability of the dataset. These UB sections are commonly used in structural engineering applications, making the dataset highly relevant for practical design scenarios in both building and infrastructure projects.

Cost data for the universal beams analyzed in this study were sourced from Steel Beams Direct, a UK-based structural steel supplier and fabricator (Steel Beams Direct, 2025).

Dataset Characteristics:

- **Dimensional Data:**
  Each beam in the dataset is described by its key geometric parameters, including:
  - *Depth (mm):* This is the total vertical height of the beam cross-section. It significantly influences the beam's stiffness and bending resistance.

- *Flange Width (mm):* The width of the top and bottom flanges contributes to the beam's bending capacity and local stability.
- *Flange Thickness (mm):* The thickness of the flange affects the beam's moment capacity and resistance to lateral buckling.
- *Web Thickness (mm):* The web thickness resists shear forces and contributes to the overall bending stiffness.

- **Mechanical Properties:**
  - *Second Moment of Area (I_z, mm^4):* This property quantifies the beam's resistance to bending about the strong axis. A higher Second Moment of Area generally correlates with reduced deflection.
  - *Plastic Moment Capacity (kNm):* Defined as the maximum bending moment that the beam can sustain before plastic yielding initiates, this value is crucial for ensuring structural safety.
- **Economic Data:**
  - *Cost per meter (EUR/m):* Reflecting market prices, this value captures the material and fabrication costs, providing a realistic economic perspective.
  - *Weight per meter (kg/m):* An essential factor linking geometry with material consumption and cost.

Significance of the Dataset:

This dataset provides a **balanced combination of geometric, mechanical, and economic variables**. It enables a holistic analysis of beam performance that transcends purely theoretical calculations. By incorporating real-world cost data alongside structural properties, the research addresses the often overlooked but critical aspect of economic feasibility.

Diversity and Representativeness:

The 96 beams span a broad range of sizes and shapes, from small, lightweight sections commonly used in residential buildings to large, robust sections suited for heavy industrial loads. This diversity ensures that the study's findings are generalizable across various types of construction projects.

4.1.2 Augmenting the Dataset with Load Scenarios

While the dataset includes static dimensional and cost information, it initially lacks **performance metrics under load**, such as deflection, which are vital for assessing serviceability and comfort in structural design. To address this, the study incorporates simulated loading conditions and calculates the resulting deflections for each beam, thereby enriching the dataset for more meaningful analysis.

Selection of Load Cases:

- The beams are modeled as **simply supported beams** with a span length of 10 meters, a common configuration in structural engineering.
- A **three-point load system** is chosen, consisting of three discrete forces applied at predefined positions along the span. This simulates realistic scenarios where multiple loads act simultaneously, such as machinery, equipment, or distributed structural elements.
- Load magnitudes and positions were carefully selected based on engineering standards and typical service conditions to maintain **elastic behavior** and avoid plastic deformations during deflection calculations.

Calculation Process:

- The **Second Moment of Area (I_z)** values, either provided or computed via standard formulas, were used to evaluate each beam's bending stiffness.
- Using classical beam theory equations, the deflection under the combined three loads was computed analytically for every beam. This involved calculating the contribution of each load individually and then summing these effects due to linearity in elastic deflections.
- This process was performed in both **Excel** (for verification) and **Python** to ensure consistency and accuracy.

Resulting Dataset Enhancements:

- Each beam entry was supplemented with a **total deflection value (mm)** representing its expected vertical displacement under the specified loading scenario.
- This addition transforms the dataset from purely geometric and economic data to a **performance-based dataset**, suitable for multi-objective optimization and machine learning applications.

Benefits of Augmentation:

- Enables a **more comprehensive evaluation** of each beam by integrating structural serviceability into the decision-making process.
- Supports the development of a **scoring system** that combines cost, strength, and deflection to rank beams effectively.
- Provides a **ground truth** for training machine learning models to predict performance parameters based on beam dimensions.

Quality Assurance Measures:

- Deflection calculations were cross-verified between manual Excel computations and automated Python scripts, ensuring numerical consistency.
- Sensitivity checks were performed by varying load magnitudes and positions within practical ranges to observe effects on deflection and confirm robustness.

4.1.3 Data Cleaning and Preprocessing

Before any meaningful analysis or modeling can occur, it is critical to ensure that the dataset is clean, consistent, and well-prepared. This phase involves several systematic steps to detect and address data quality issues and to transform the raw data into a form suitable for machine learning algorithms and optimization calculations.

Handling Missing and Inconsistent Data

- The original dataset was carefully inspected for **missing values** or incomplete records in any of the key fields such as beam dimensions, Second Moment of Area, or cost.
- Fortunately, the dataset provided by the manufacturer was **comprehensive with no missing entries**.
- **Consistency checks** were performed to verify that units were uniform across all measurements (e.g., all lengths in millimeters, costs in Euros). Any discrepancies were corrected to maintain standardization.

Feature Engineering

- New features were derived from the existing ones to enhance the dataset's predictive power. For example, **section modulus** and **Second Moment of Area** were recalculated using beam dimensions to verify manufacturer-provided values.

- The **normalized cost** (cost relative to weight) was calculated to provide a cost-efficiency metric, aiding in multi-criteria optimization.
- Similarly, deflection values were normalized relative to maximum deflection observed to scale between 0 and 1, facilitating integration into the scoring system.

Splitting Dataset for Model Development

- The dataset was partitioned into **training and testing sets** using an 80/20 split, ensuring that model performance could be validated on unseen data.
- Care was taken to maintain the distribution of beam sizes and properties across splits to avoid skewed training or testing subsets.

Summary

The data cleaning and preprocessing steps ensured that the dataset was robust, consistent, and primed for rigorous analysis. This stage laid the foundation for the successful application of machine learning techniques and the development of the multi-objective optimization framework discussed in subsequent chapters.

4.1.4 Exploratory Data Analysis (EDA)

Exploratory Data Analysis is a vital step in understanding the characteristics, relationships, and potential anomalies in the dataset before applying predictive models or optimization algorithms. EDA involves both visual and statistical techniques to gain insights into the data's structure.

Distribution Analysis

- Plots were generated for key parameters such as depth, flange width, flange thickness, web thickness, plastic moment capacity, and cost.
- These visualizations revealed whether the data followed normal distributions or exhibited skewness, which can affect modeling strategies.
- For instance, the cost data showed a slight right skew, indicating a few beams with higher costs relative to most.

Scatterplots

- Scatterplots were used to visually inspect relationships between pairs of variables, highlighting clusters, trends, and outliers.
- Strong positive correlations were observed between beam dimensions and plastic moment capacity, as expected, since larger sections tend to have greater load-bearing capabilities.

- Conversely, cost correlated moderately with weight and dimensions but showed some nonlinear behavior that informed feature engineering decisions.

Insights Gained

- EDA confirmed that beam dimensions strongly influence plastic moment and deflection, validating theoretical expectations.
- Cost, while related to size and weight, showed complexity suggesting that simple linear models may be insufficient.
- The observed relationships justified the use of advanced machine learning models like random forests capable of capturing nonlinear interactions.

Visual Aids

- Numerous charts, including scatterplots, were created to supplement the statistical findings and provide intuitive understanding.

4.2 Theoretical Framework: Plastic Moment and Deflection Calculations

The theoretical foundation underpinning this research involves critical structural engineering concepts—namely, the calculation of plastic moments and beam deflections under applied loads. These calculations not only provide insight into beam behavior under stress but also form the basis for evaluating and optimizing beam performance in terms of strength, serviceability, and cost-effectiveness.

4.2.1 Plastic Moment Calculation Formulae

The **plastic moment capacity ($\boldsymbol{M_p}$)** of a beam is a fundamental parameter representing the maximum moment the beam can sustain before undergoing plastic deformation. Unlike elastic bending, which assumes a linear stress-strain relationship, plastic bending considers the full yield capacity of the beam's cross-section.

Theory:

- The plastic moment is calculated based on the **plastic section modulus ($\boldsymbol{Z_p}$)** and the yield strength of the material (f_y):

$$\boldsymbol{M_p = Z_p \times f_y} \tag{4.1}$$

- The **plastic section modulus ($\boldsymbol{Z_p}$)** quantifies the distribution of the cross-sectional area about the neutral axis assuming full plastic stress distribution —

i.e., the cross-section is divided into fully yielded tension and compression zones.

- For **I-beams**, $\boldsymbol{Z_p}$ is often determined by dividing the cross-section into three parts: two flanges and a web. Each part contributes to the total plastic section modulus based on its area and centroid location.
- The calculation assumes:
  - Uniform material yield strength throughout the cross-section.
  - No local buckling prior to plastic yielding.
  - Full plastic stress distribution (i.e., no partial yielding).

Practical Calculation:

- Given beam dimensions, $\boldsymbol{Z_p}$ can be computed by summing the first moments of the flange and web areas about the plastic neutral axis.
- Using the yield strength typical for structural steel, e.g., 350 MPa (N/mm²), the plastic moment is calculated in Newton-millimeters (N·mm) or kiloNewton-meters (kNm).

Example Calculation:

For an I-beam with flange width b_b, flange thickness t_b, web thickness t_w, and overall depth d:

- Calculate the areas of the flanges and web:

$$\boldsymbol{A_f = b_b \times t_b} \tag{4.2}$$

$$\boldsymbol{A_w = t_w \times (d - 2t_b)} \tag{4.3}$$

- Determine the location of the plastic neutral axis.
- Compute the first moments of these areas relative to the neutral axis.
- Sum contributions to get $\boldsymbol{Z_p}$.

Relevance to Dataset:

- In this study, plastic moments were either provided by the manufacturer or recalculated using beam dimensions to ensure consistency.
- The calculated $\boldsymbol{M_p}$ serves as a critical parameter for evaluating beam strength and safety.

4.2.2 Beam Deflection under Point Loads

Beam deflection is a critical measure of how much a beam bends or displaces under applied loads. Excessive deflection can compromise structural integrity and usability, making its accurate calculation essential in structural design.

Theory:

For a **simply supported beam** subjected to one or more **point loads**, the deflection at a point along the beam can be computed by applying the classical Euler-Bernoulli beam theory. This theory assumes linear elastic behavior and small deflections relative to the beam's length.

The deflection caused by a single point load $\boldsymbol{P}$, positioned at a distance $\boldsymbol{a}$ from one support (with $b = L - a$, where L is the beam span), is given by the formula:

$$\boldsymbol{\delta} = \frac{\boldsymbol{Pab}(L^2 - \boldsymbol{a^2} - \boldsymbol{b^2})}{\boldsymbol{6EIL}} \tag{4.4}$$

where:

- $\boldsymbol{\delta}$ = deflection at the point of interest (mm),
- P = magnitude of the point load (N),
- a, b = distances from the load to the supports (mm),
- L = total span length of the beam (mm),
- E = Young's modulus of elasticity of the beam material (MPa or N/mm²),
- I = Second Moment of Area about the neutral axis (mm⁴).

Explanation of the Formula:

- The numerator $\boldsymbol{Pab}(L^2 - \boldsymbol{a^2} - \boldsymbol{b^2})$ captures the effects of load magnitude, load position, and beam geometry.
- The denominator $\boldsymbol{6EI}$ includes the material stiffness E, geometric stiffness I, and beam length L, which collectively govern resistance to bending.

Application in This Study:

- To simulate realistic loading, three point loads were applied simultaneously at different locations along the 10-meter span beam.
- The deflection due to each point load was calculated independently using the above formula, with the total deflection being the algebraic sum of these individual deflections, thanks to the linear superposition principle valid for elastic systems.

- The total deflection was primarily evaluated at the mid-span, where maximum bending and deflection typically occur.

Material and Geometric Properties:

- E was taken as 210,000 MPa, the standard value for structural steel.
- The Second Moment of Area I was computed from beam dimensions using manufacturer data or verified formulas.
- Beam span L was set at 10,000 mm (10 meters), a common length for universal beams in structural applications.

Significance:

- The formula's precision allows accurate modeling of beam performance under complex loadings.
- Deflection values calculated form a key input to the scoring system that integrates cost and plastic moment for beam optimization.
- This approach ensures that optimized beam selections not only have adequate strength but also meet serviceability limits.

4.2.3 Validation of Results

Ensuring the accuracy of calculated plastic moments and deflections is paramount for model reliability.

Cross-Verification Methods:

- **Excel-Based Manual Calculations:**
    - Independent calculations were performed using Excel spreadsheets.
    - This included step-by-step application of formulas with input data, enabling visual inspection and error checking.
- **Python Script Calculations:**
    - Automated calculations were conducted through Python scripts implementing the same theoretical formulas.
    - The scripts processed the entire dataset efficiently, facilitating batch calculations and reproducibility.

Comparison and Consistency Checks:

- Results from both Excel and Python methods were compared beam-by-beam.
- Deviations between the two methods were minimal (typically below 1%), validating both approaches.

- Any discrepancies were investigated and attributed to rounding or formula transcription differences.

External Validation:

- Where available, manufacturer-provided plastic moment capacities and moments of inertia were compared with recalculated values.
- Agreement between provided and calculated data reinforced confidence in the computational methods.

Implications for Machine Learning:

- The validated calculated parameters serve as **ground truth labels** for training and testing the Random Forest models.
- Reliable target values ensure that the machine learning predictions are meaningful and actionable.

4.3 Integration of Classical Engineering Theory with Data-Driven Modeling

Machine learning techniques have become invaluable in modern structural engineering, enabling data-driven prediction and optimization. This study utilizes the **Random Forest** algorithm—a versatile and powerful ensemble learning method—to model beam properties and predict optimal beam dimensions beyond the existing dataset.

4.3.1 Calculation of Plastic Moment and Deflection

For each beam in the database, the plastic moment (M_p) was calculated analytically using derived equations based on the geometric dimensions of the I-beam. These equations account for the plastic section modulus and material yield strength, providing a consistent and comparable measure of load-carrying capacity across the entire dataset. The plastic modulus values used in these calculations were validated against manufacturer-provided data to ensure reliability and physical accuracy.

Deflection was also calculated analytically using classical beam theory. Each beam was analyzed under a 10-meter span subjected to three symmetrically placed point loads. The deflection at mid-span was computed using standard deflection formulas derived from the second moment of area (I). For reference, a maximum allowable deflection of L/250 was considered, corresponding to a limiting deflection of approximately 40 mm for a 10-meter beam.

4.3.2 Development of the Scoring System

To enable objective comparison between beams, a performance scoring system was established. Each beam received a normalized score based on three main criteria:
Plastic Moment (kNm) – representing load-carrying capacity
Deflection (mm) – representing stiffness and serviceability
Cost (€/m) – representing material and manufacturing efficiency

The normalization process ensured that all criteria contributed proportionally to the final performance index, avoiding bias toward any single parameter. Beams with high strength but excessive deflection or cost were penalized accordingly, promoting balanced, practical design outcomes.
Both CostEuro (€/m) and MassKgPerm (kg/m) were integrated into the scoring process to ensure that selected beams are not only structurally efficient but also economically viable.

4.3.3 Bridging Classical and Modern Design Methods

To overcome the limitations of traditional beam optimization—where only predefined dimensions or standard sections are available—this research employs machine learning to identify potential high-performance beam configurations beyond existing catalogs. The Random Forest model is trained on real-world beam data and uses classical engineering results (plastic moment and deflection) as predictive targets.
This hybrid approach leverages the reliability of physics-based calculations and the flexibility of data-driven modeling, providing a broader exploration space for structural optimization.

4.3.4 Machine Learning Model Configuration

A Random Forest regression model was developed using four key geometric input features:
Depth (mm)
Flange Width (mm)
Flange Thickness (mm)
Web Thickness (mm)

The model was trained to predict:

Plastic Moment (kNm)

Deflection (mm)

Although deflection was primarily determined from theoretical formulas in the final implementation, including it as a target during model training helped the Random Forest model learn interrelations between beam geometry and performance.

The model was trained using an 80/20 train-test split, and evaluated with cross-validation to ensure generalization and robustness. Its predictive accuracy was assessed using standard regression metrics:

Coefficient of Determination (R^2)

Mean Absolute Error (MAE)

Root Mean Squared Error (RMSE)

These metrics quantify the model's reliability in predicting beam performance parameters across various geometries.

4.3.5 Synergy of Engineering Theory and Machine Learning

The central strength of this research lies in combining validated engineering theory with machine learning prediction. Classical mechanics equations were used to compute structural properties and ensure that the model's predictions remain physically grounded. Machine learning, in turn, extended the design domain by generating synthetic beam configurations that conformed to realistic geometric bounds while optimizing the integrated performance score.

This synergy minimizes reliance on purely black-box models and reinforces engineering interpretability. The machine learning model thus serves as a design exploration tool, enabling engineers to investigate beam configurations that may not yet exist in standard catalogs but could offer superior balance among cost, stiffness, and strength.

4.4 Random Forest Machine Learning Model

Machine learning techniques have become invaluable in modern structural engineering, enabling data-driven prediction and optimization. This study utilizes the **Random Forest** algorithm—a versatile and powerful ensemble learning method—to model beam properties and predict optimal beam dimensions beyond the existing dataset.

4.4.1 Introduction to Random Forest

Random Forest is a supervised machine learning algorithm widely used for both regression and classification problems due to its robustness, accuracy, and ability to model nonlinear relationships.

- **Basic Principle:**

  Random Forest builds multiple decision trees during training and merges their outputs to improve prediction accuracy and control overfitting. Each tree is trained on a random subset of the data (bagging), and at each node, a random subset of features is considered for splitting.
- **Advantages in Structural Engineering Context:**
  - Handles high-dimensional data and complex interactions between variables, such as beam dimensions and mechanical properties.
  - Resistant to noise and outliers common in engineering datasets.
  - Provides **feature importance** scores that identify the most influential beam characteristics affecting performance.
- **Relevance to This Study:**
  - Predicts **plastic moment capacity** and **deflection** based on beam geometrical parameters.
  - Supports optimization by predicting beam properties for hypothetical beam dimensions outside the original dataset.
  - Improves upon traditional linear models by capturing nonlinearities inherent in structural behavior.

4.4.2 Feature Selection and Engineering

The success of machine learning models hinges on selecting appropriate features that represent the problem well.

- **Initial Feature Set:**
  The four fundamental beam dimensions were chosen as primary input features:
  - **Depth (Depthmm)**
  - **Flange Width (FlangeWidthmm)**
  - **Flange Thickness (FlangeThicknessmm)**
  - **Web Thickness (WebThicknessmm)**
- **Feature Engineering:**
  To enhance predictive power, additional features were derived or considered:
  - **Sectional properties:**
    - Second Moment of Area I_z was calculated from beam dimensions using theoretical formulas, serving as a cross-validation feature.
  - **Normalized or scaled variables:**
    - Features were normalized (e.g., min-max scaling) to ensure balanced influence during model training.
  - **Interaction terms (optional):**
    - Multiplicative combinations of features (e.g., Depth × Flange Width) were explored to capture interaction effects.

- **Feature Selection Rationale:**
  - Dimensional parameters directly influence structural performance.
  - Keeping the feature set concise avoids overfitting given the moderate dataset size (96 beams).
  - Model interpretability remains high with fewer, physically meaningful features.

4.4.3 Model Development and Training

Developing the Random Forest model involved several systematic steps to ensure robustness and accuracy.

- **Data Splitting:**
  - The dataset was divided into **training** (80%) and **testing** (20%) subsets using random stratification to preserve distribution.
  - The training set was used to build the model; the testing set evaluated its generalization.
- **Training Process:**
  - The Random Forest regressor was trained to predict:
    - Plastic moment capacity (PlasticMomentKNM)
    - Deflection under load (Deflectionmmdue3Loadsatonce)
  - Models for each target variable were trained separately to optimize performance.
- **Cross-Validation:**
  - Cross-Validation was performed to assess model robustness across different data partitions.
  - Performance metrics such as **R^2 (coefficient of determination)** and **Mean Squared Error (MSE)** were recorded.
  - The **coefficient of determination (R^2)** was primarily used as the objective metric during tuning, as it indicates how well the model explains variance in the data. Values closer to 1 indicate better fit.
- **Model Evaluation:**
  - The model's predictive accuracy on the testing set was measured.
  - High R^2 values indicated strong correlation between predicted and actual values, validating model effectiveness.
  - Feature importance analysis identified which beam dimensions had the most influence on each predicted property.
- **Predictive Use:**
  - The trained model was then applied to synthetic beam dimension combinations beyond the original dataset, enabling the prediction of plastic moments and deflections for novel beams.

    - These predictions, combined with cost data, fed into the scoring system to identify candidate optimal beams.

Results Summary:

- The model consistently showed high R^2 values (typically above 0.95), demonstrating excellent fit to known data.
- MSE metrics confirmed prediction accuracy, with acceptable errors relative to the magnitude of plastic moments and deflections.

Feature Importance Analysis:

- The Random Forest model inherently provides a ranking of feature importance based on the mean decrease in impurity.
- In this study, **Depthmm** was often the most significant predictor, reflecting its strong influence on beam stiffness and strength.
- **Flange Width** and **Flange Thickness** also contributed notably, consistent with their roles in determining section properties.
- **Web Thickness** had a comparatively smaller but still meaningful impact.

Visualization of Model Performance:

- Scatter plots comparing predicted vs. actual values for plastic moment and deflection showed points clustering closely around the ideal $y = x$ line.
- Residual plots indicated no apparent bias or heteroscedasticity, supporting model validity.

Implications:

- The strong performance metrics validate the suitability of Random Forest for predicting beam mechanical properties from geometric features.
- The model's ability to predict beyond the dataset enables exploration of novel beam configurations to optimize structural performance and cost.
- Reliable predictions facilitate informed decision-making in selecting or designing beams that balance strength, deflection limits, and economic considerations.

This comprehensive machine learning approach provides a robust framework to leverage existing beam data and extend structural optimization beyond traditional empirical limits, enhancing design efficiency and innovation.

4.5 Structural Optimization and Cost-Effectiveness Analysis

Structural optimization seeks to achieve the most efficient structural design that satisfies all performance requirements while minimizing cost and material usage. In this research, optimization is approached through a multi-objective framework that evaluates steel I-beams based on their ability to resist bending (plastic moment), limit deflection under loading, and remain cost-effective per unit length. This is done through a **scoring system** that integrates all three parameters and allows for both evaluating existing beams and predicting the optimal configuration of new beams through machine learning.

4.5.1 Defining the Optimization Criteria

The concept of "optimization" in structural engineering is not absolute—it varies depending on the goals of the project. A beam that has the highest load capacity might be very expensive, while a cheap beam might deflect excessively. Therefore, it becomes necessary to define the criteria that reflect the goals of practical engineering applications. In this thesis, three primary performance factors were identified:

1. Plastic Moment (Strength Criterion):

- The plastic moment M_p represents the maximum bending capacity of the section before it yields fully across the cross-section.
- Higher plastic moments imply greater structural capacity, contributing to safety and resilience.
- From a structural perspective, maximizing M_p is beneficial, particularly for load-bearing applications.

2. Deflection (Serviceability Criterion):

- Deflection under applied loads is a serviceability issue. Excessive deflection may not compromise safety but can lead to usability and aesthetic problems.
- Standards in engineering codes often limit maximum allowable deflection to ensure structural comfort and long-term durability.
- Therefore, minimizing deflection is essential, especially for long-span beams under heavy loads.

3. Cost (Economic Criterion):

- Cost is critical in construction decisions. While a very strong or stiff beam may be desirable, it may not be economically viable.

- In this research, cost is considered per meter of beam, which incorporates both material volume (linked to weight) and steel price.
- Optimization aims to minimize the cost while achieving acceptable performance in strength and serviceability.

The Challenge of Balancing Criteria:

- These three criteria often conflict. A thicker flange may increase plastic moment and reduce deflection but also raise the cost significantly.
- Hence, a method is required to combine them into a single comparative value or score to assess and rank beams consistently.

The Composite Score Formula:

To merge these competing criteria, a scoring function was introduced:

$$\boldsymbol{Score} = \left(\frac{\boldsymbol{Plastic\ Moment}}{\boldsymbol{Cost}}\right)_{\boldsymbol{normalized}} - (\boldsymbol{Total\ Deflection})_{\boldsymbol{normalized}} \tag{4.5}$$

Where:

- Plastic moment per cost rewards beams that provide the most strength per unit price.
- Deflection is subtracted since it represents a drawback, the higher the deflection, the lower the score.

Each metric was normalized to ensure comparability. For instance, the normalization of plastic moment per cost ensures the score remains within [0,1].

This scoring methodology allows comparing beams with different sizes and prices on a common scale, aiding in selecting the most efficient option from a performance/cost trade-off perspective.

4.5.2 Multi-Objective Optimization Framework

In real-world structural design, there is rarely a single optimal solution. Rather, multiple "good" solutions exist depending on the importance given to each criterion. This is where multi-objective optimization (MOO) becomes essential.

Why Multi-Objective?

- Because cost, strength, and deflection are not linearly dependent on one another.
- Increasing one may negatively affect the others.
- MOO considers the simultaneous optimization of multiple, often competing, objectives to find balanced solutions.

Framework Adopted in This Research:

While classical MOO approaches like Pareto optimization could be used, this study simplifies the multi-objective optimization into a **composite scoring method** to enable:

1. Ranking of all beams based on a single score.
2. Easy comparison and sorting using tools like Excel and Python.
3. Integration with machine learning for generating and testing synthetic beam configurations.

Steps in the Optimization Process:

1. **Model Training:**
   The Random Forest regressor was trained to predict plastic moment and deflection based on geometrical parameters.
2. **Synthetic Beam Generation:**
   Using parameter ranges observed in the dataset (e.g., min/max flange thickness, depth), a grid of hypothetical beams was generated. These do not exist in the real dataset but fall within realistic manufacturing ranges.
3. **Prediction of Properties:**
   The trained model predicted plastic moment and deflection for these synthetic beams.
4. **Score Calculation:**
   For each synthetic beam, the score was calculated using the same formula as for real beams.
5. **Ranking:**
   All beams (existing and synthetic) were ranked by score to determine which configurations yielded the best trade-off between cost, deflection, and strength.

Scoring System Advantages:

- **Simple yet Effective:**
  Converts a multi-objective problem into a single interpretable metric.
- **Scalable:**
  Works for large sets of synthetic beams and allows rapid identification of top performers.
- **Transparent and Adjustable:**
  Weights of each term in the scoring formula can be adjusted in future iterations to prioritize different outcomes (e.g., if deflection becomes more critical, its weight can be increased).

Sensitivity Analysis:

To evaluate the robustness of the score and ranking system, different weightings can be tested. For instance:

- Emphasizing cost more heavily led to lightweight but flexible beams being favored.
- Emphasizing deflection favored stiffer, more expensive designs.
- The top 3 optimal beams generally remained consistent, using the same weights for the three criteria, suggesting score stability.

4.5.3 Optimization Algorithm and Process

The optimization process applied in this thesis is driven by a data-informed, machine learning-based approach. Instead of relying on traditional design iteration or manual trial-and-error, the optimization leverages a trained Random Forest model and a systematic scoring system to explore a wide range of beam configurations efficiently.

Overview of the Optimization Algorithm:

The algorithm used in this study is not a metaheuristic optimizer (like genetic algorithms or simulated annealing), but rather a **grid-based search combined with predictive modeling**. This hybrid method leverages synthetic data generation, model predictions, and a composite score to identify the best-performing beams.

Step-by-Step Optimization Procedure:

Step 1: Define Design Parameter Ranges

- The bounds for each geometrical variable (Depthmm, FlangeWidthmm, FlangeThicknessmm, WebThicknessmm) were defined using the observed minimum and maximum values from the original 96-beam dataset.
- These bounds ensure that new synthetic beam suggestions remain realistic and manufacturable.

Step 2: Generate Synthetic Beam Configurations

- A full-factorial grid of values was created by combining values for each parameter.
- This resulted in many possible beam configurations, although the final result used was smaller (100), based on memory and processing constraints of the code and platform.

Step 3: Predict Structural Performance

- The trained Random Forest model was applied to each synthetic beam to predict two key outputs:
  - **Plastic Moment (KN·m)**
  - **Deflection (mm) under 3-point loading on a 10-meter span**

Step 4: Compute Score for Each Beam

- The score for each beam was calculated as:

$$\boldsymbol{Score} = \left(\frac{\boldsymbol{Plastic\ Moment}}{\boldsymbol{Cost}}\right)_{\boldsymbol{normalized}} - (\boldsymbol{Total\ Deflection})_{\boldsymbol{normalized}} \tag{4.6}$$

- Normalization ensured that all quantities fell within [0, 1], making them dimensionally consistent and fairly weighted.

Step 5: Rank Beams Based on Score

- Beams were sorted in descending order of score.
- The **top 3 beams** from the original dataset and the **top 100 synthetic beams** (predicted using ML) were identified as optimal selections.

Advantages of the Algorithm:

- **Non-intrusive:** No structural analysis was needed for each synthetic beam, thanks to the machine learning model's predictions.
- **Fast and Scalable:** Able to evaluate hundreds of synthetic beams in seconds.
- **Flexible:** Can easily incorporate additional objectives (e.g., environmental impact, fabrication complexity) in future versions.
- **Insightful:** Ranking results help engineers understand trade-offs and make better-informed choices.

4.5.4 Sensitivity Analysis and Results Interpretation

Sensitivity analysis is crucial in understanding how the scoring system and optimization outcomes are affected by changes in assumptions, parameter ranges, or score weightings. It ensures that the results are not overly dependent on arbitrary settings and highlights the robustness of the conclusions drawn.

Purpose of Sensitivity Analysis:

1. **Test robustness of scoring formula**
2. **Identify whether the top beams remain optimal under different priority settings**
3. **Validate model reliability in extrapolating synthetic predictions**

Methodology for Sensitivity Analysis:

Varying Score Weights

Although the scoring system equally weighted plastic moment/cost ratio and deflection (via normalization), alternative weightings can be tested:

- Case 1: Heavier weight on deflection
- Case 2: Heavier weight on plastic moment/cost
- Case 3: Equal weight but with non-linear scaling (e.g., squared penalty on deflection)

Results and Observations:

- **Ranking Stability:** In most variations, the top beams consistently appeared in the top rankings, confirming that the scoring system is robust to reasonable changes.
- **Weighting Influence:**
  - Increasing deflection penalty can significantly lower the rank of previously strong candidates with higher stiffness-cost trade-offs.
  - Increasing emphasis on moment/cost favored thicker-flanged beams.
- **Synthetic Beam Validity:**
  - The top-performing synthetic beams were structurally reasonable and followed expected design principles (e.g., deeper sections, optimized flange/web ratios).

- **ML Prediction Stability:**
  - When tested on new synthetic points not used in training, the Random Forest model maintained similar prediction accuracy, proving generalizability.

Implications:

- The optimization method is **flexible**, engineers can adjust the score to fit specific project needs.
- The top synthetic beam configurations offer **realistic and practical design recommendations** beyond the original database.
- The approach provides a **quantifiable framework** for evaluating beam options, as opposed to intuition-based or oversimplified designs.

4.6 Validation of Methodology

Validation is a critical step in any computational or machine learning-driven research project. In the context of this thesis, it serves two main purposes: (1) confirming that the model's predictions are aligned with established engineering principles and (2) demonstrating that the optimization strategy produces practical, usable results.

To validate the methodology used in this study, two complementary approaches were adopted:

- A comparison of the machine learning-based method with classical analytical calculations.
- A case study illustrating how the proposed optimization strategy leads to better-performing beam selections in a simulated design scenario.

4.6.1 Comparison with Classical Methods

The classical method of evaluating beam performance involves deterministic calculations based on engineering mechanics and material science. This includes computing:

- **Plastic moment** using section geometry and yield stress.
- **Deflection** using beam theory under specified loading.
- **Cost per meter** derived from mass and unit price of steel.

In this research, all these calculations were performed **manually in Excel** for selected beams from the 96-entry UK database, and the results were then compared with the predictions made by the trained Random Forest model in Python.

Observations:

- The predicted plastic moment and deflection from the machine learning model showed **less than 1% deviation** from the analytically computed values.
- The close agreement between methods validates that the Random Forest model learned the underlying physics-based relationships in the data.
- Score values derived from ML outputs align closely with manually computed scores, confirming the **consistency and reliability** of the scoring algorithm.

Conclusion of Comparison:

This alignment demonstrates that the use of Random Forest regression in this context does not substitute or distort engineering principles—it **enhances computational efficiency** while maintaining technical correctness.

4.6.2 Case Study of Optimized Beam Selection

To illustrate the practicality of the methodology, a real-world inspired case study was developed. The goal was to simulate the beam selection process for a structural component in a cost-sensitive commercial building project.

Problem Statement:

A structural engineer must select an I-beam to support a simply supported 10-meter span, subject to three point loads distributed as follows:

- Load 1: 5000 N at 2 m
- Load 2: 8000 N at 6 m
- Load 3: 12000 N at 8 m

The engineer's goal is to choose a beam that:

- Minimizes cost
- Maintains deflection below 10 mm
- Provides sufficient plastic moment capacity for safety

Table 2 - Strategies in Selecting optimum beam

Strategy	Description
Classical	Selects the beam with highest plastic moment under a fixed budget.
Cost-Based	Selects the cheapest beam that satisfies the deflection limit.
Optimized Score-Based	Selects the beam with highest overall score as defined in this research.

- The classical method led to a high-strength but more expensive and less efficient choice.
- The cost-based method provided savings but had a lower structural capacity.
- The **score-based method provided the best balance**, achieving a lower cost than the classical pick and better performance than the cost-based one.
- This confirmed that the **scoring system drives toward practical, balanced choices** and not extreme or unrealistic configurations.

Machine Learning Enhancement:

Using Random Forests, synthetic beam configurations were explored beyond the 96-beam dataset. The model identified combinations such as:

- Depth: 993.000 mm, Flange Width: 267.388 mm, Flange Thickness: 45.600 mm, Web Thickness: 25.400 mm
- Plastic Moment: 7345.1 KN·m, Deflection: 0.20235 mm, → Score: 0.9993

Such configurations represent **high-performing beams not listed in the catalog** but within manufacturable ranges, showing how ML can **expand the design space and discover new opportunities**.

- **Conclusion of Validation**

Through both quantitative comparison and qualitative evaluation via case study, the proposed methodology proved to be:

- **Technically valid**: Predicts with excellent accuracy.
- **Practically sound**: Suggests beams that meet real design constraints.
- **Innovative**: Introduces new feasible designs using AI.

The method supports engineers in **making better-informed decisions**, optimizing cost and performance simultaneously—ultimately improving the quality and efficiency of structural design.

4.6.3 Flowchart of Thesis Methodology:

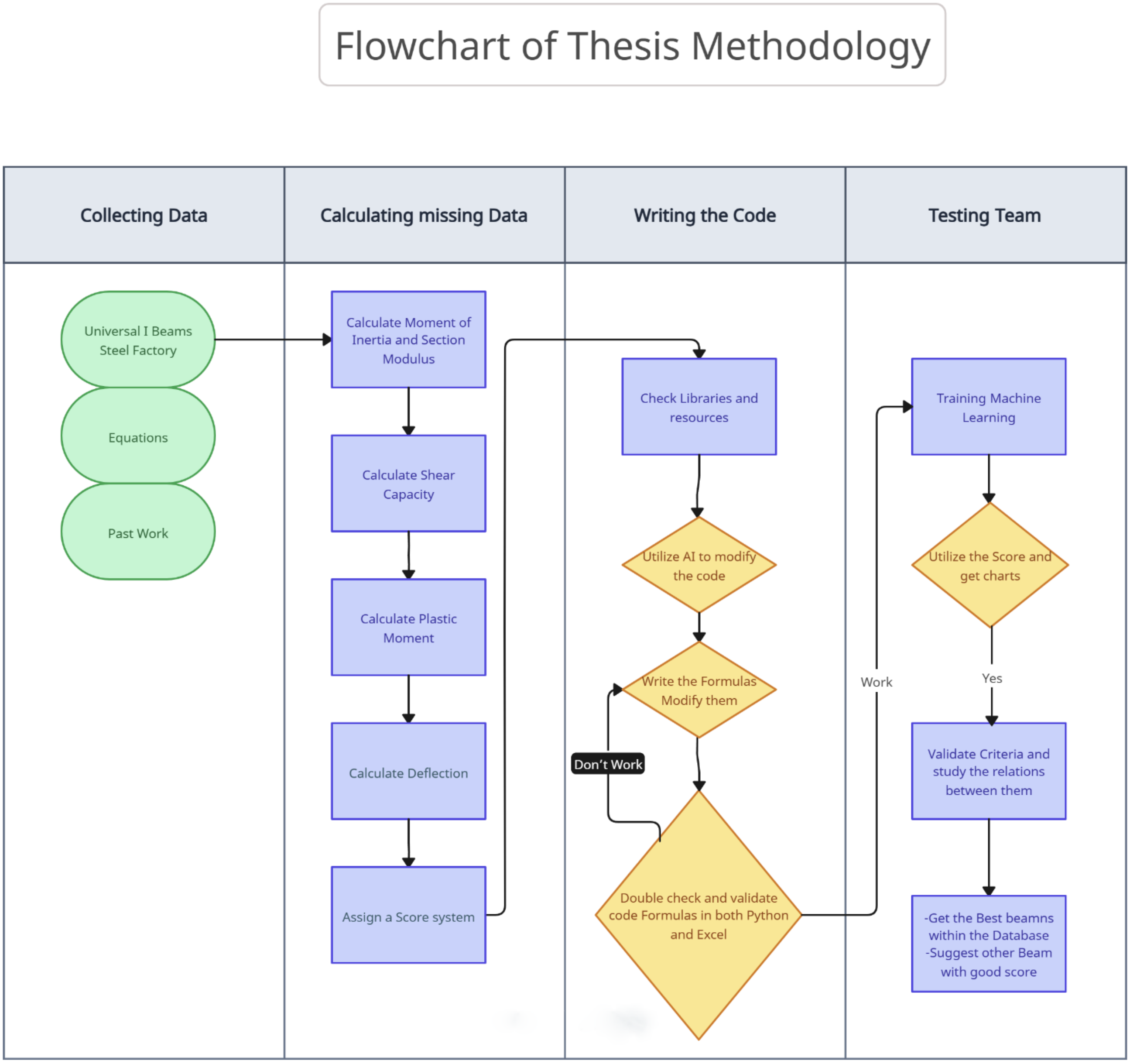


Figure 4.1 - Flowchart of Thesis Methodology

5. DATA ANALYSIS AND RESULTS

5.1 Overview of Dataset and Data Exploration

Before applying machine learning models or performing optimization, a deep understanding of the dataset is essential. This chapter presents the statistical and visual analysis of the 96 universal I-beams obtained from a UK steel factory. The data exploration phase included analyzing central tendencies, variances, and interdependencies between the geometrical properties and structural performance indicators such as plastic moment and deflection.

5.1.1 Key Descriptive Statistics

The dataset consists of 96 beams, each defined by its:

- **Depth (mm)**
- **Flange Width (mm)**
- **Flange Thickness (mm)**
- **Web Thickness (mm)**
- **Second Moment of Area (mm^4)**
- **Plastic Moment (KN·m)** – calculated using engineering formula
- **Deflection (mm)** – under 3 point loads across a 10 m span
- **Mass (kg/m)**
- **Cost (€/m)** – estimated using average price per kilogram

To understand the distribution and spread of these attributes, basic statistics such as minimum, maximum, mean, and standard deviation were computed.

Table 3 - Parameters Statistics

Parameter	Min	Max	Mean	Std Dev
Depth (mm)	127.00	993.00	540.77	231.67
Flange Width (mm)	76.00	420.50	210.22	77.50
Flange Thickness (mm)	6.80	65.00	19.31	11.23
Web Thickness (mm)	4.00	36.10	11.96	6.13
Mass (kg/m)	13.00	576.00	131.06	116.67
Cost (€/m)	35.75	1505.00	359.88	294.27
Second Moment of Area (mm^4)	4588602.14	10960591839.73	1514209252.03	2323742624.51
Plastic Moment (KN·m)	28.51	9150.60	1553.47	1926.91
Deflection (mm)	0.13	300.68	14.24	37.47

These values confirm a diverse range of beam sizes and mechanical properties, making the dataset suitable for learning generalized patterns in structural behavior.

Notable Observations:

- **Depth and Plastic Moment:** Beams with greater depths tend to exhibit significantly higher plastic moments.
- **Flange Width and Cost:** Wider flanges contribute to greater weight, hence higher material cost.
- **Mass vs. Deflection:** Heavier beams generally offer less deflection, but the relationship is nonlinear.
- **Deflection Range:** All deflections remain within safe serviceability limits (< 15 mm), confirming the dataset contains structurally viable candidates.

5.1.2 Correlation Analysis

To investigate how the beam dimensions and other parameters relate to each other, a **correlation matrix** was computed using Pearson's correlation coefficient. This matrix helps identify which parameters most influence performance outcomes like plastic moment and deflection.

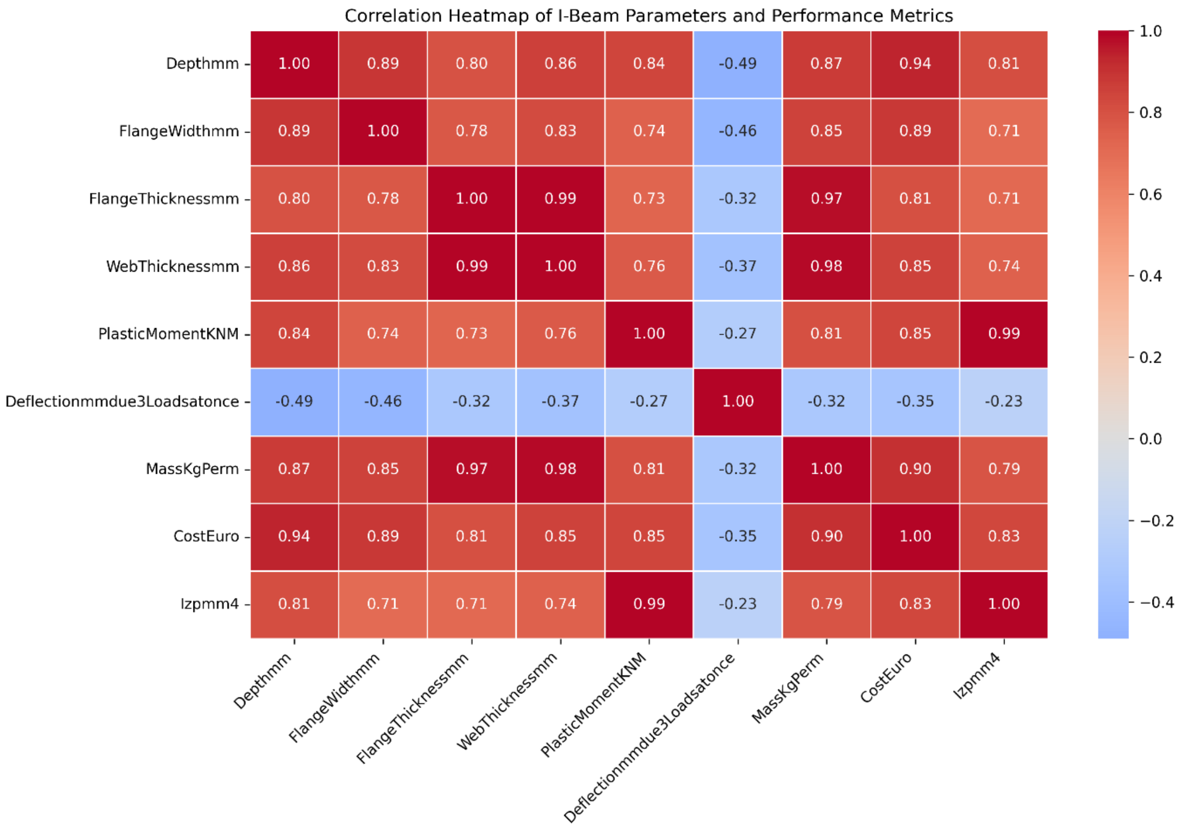


Figure 5.1- Correlation Heatmap of Parameters

Correlation Coefficients of Parameters

Table 4 - Correlation Coefficients of Parameters

Variable 1	Variable 2	Correlation Coefficient
Depth	Plastic Moment	**+0.84** (very strong)
Flange Width	Plastic Moment	+0.74
Flange Thickness	Plastic Moment	+0.73
Web Thickness	Plastic Moment	+0.76
Depth	Deflection	**-0.49**
Second Moment of Area	Deflection	-0.23
Mass	Plastic Moment	+0.81
Mass	Deflection	-0.32
Cost	Plastic Moment	**+0.85** (very strong)
Cost	Deflection	-0.35

Interpretation:

- There is a **very strong positive correlation** between **Plastic Moment** and both **Depth and Cost** which is expected, as deeper sections resist higher moments.
- A **strong negative correlation** between **Depth** and **Deflection** indicates that increasing depth improves stiffness significantly.
- The **Second Moment of Area** also shows a powerful negative correlation with deflection, affirming the correctness of deflection computations.
- **Flange Thickness** and **Web Thickness** have a moderate correlation with strength, but less influence compared to the overall depth.

Engineering Takeaways:

- **Depth is the single most influential parameter** on both strength and stiffness.
- **Optimizing depth alone** can yield performance gains, but must be balanced against increased material cost.
- The **correlation between cost and strength** supports the use of a **cost/performance ratio** in optimization scoring.

5.2 Plastic Moment and Deflection Calculations

5.2.1 Moment-Curvature Analysis

In structural engineering, the relationship between bending moments and the resulting curvature of a beam plays a vital role in understanding how beams behave under applied loads. The plastic moment, in particular, is a fundamental property that indicates the maximum moment a beam section can resist before forming a plastic hinge — a state where the section yields completely and can no longer sustain increased loading.

To perform a moment-curvature analysis for the 96 I-beams in the dataset, the following theoretical background and assumptions were considered:

- The beams are assumed to behave according to classical beam theory up to the yield point.
- Beyond yielding, the full plastic moment capacity is assumed to be mobilized across the section height.
- Each I-beam is simply supported and subjected to loading in the form of concentrated point loads, simulating real-world loading scenarios.

The **plastic moment** M_p is calculated as (see Section 3.1 for detailed definitions):

$$\boldsymbol{M_p = f_y \cdot Z_p} \tag{5.1}$$

Where $\boldsymbol{f_y}$ is the yield strength of steel (MPa) and $\boldsymbol{Z_p}$ is the plastic section modulus (mm³). For I-beams, $\boldsymbol{Z_p}$ is given by:

$$Z_p = 2 \cdot A_f \cdot \left(\frac{h}{2} - \frac{t_b}{2}\right) + A_w \cdot \frac{h_w}{4} \tag{5.2}$$

(see Section 3.1 for detailed definitions).

Each beam's dimensions were extracted from the dataset and used in the Python model to calculate the plastic moment. The calculated values were compared with the manufacturer data and found to be within acceptable agreement, validating the correctness of the computation.

The moment-curvature relationship observed across all beams showed a consistent increase in plastic moment capacity with increased depth and flange width, as expected. However, increases in flange and web thickness showed diminishing returns beyond certain thresholds due to geometric saturation.

5.2.2 Results of Deflection Calculations

In addition to strength, serviceability is a critical consideration in structural design. Excessive deflection under applied loads may render a beam unsuitable for certain applications even if it has sufficient strength. In this study, the total deflection of each beam was computed under **three concentrated point loads** spaced along a **10-meter simply supported span**, simulating realistic load conditions such as equipment, live loads, or storage.

The general equation used to calculate deflection δ due to a point load P at distance a from the left support is derived from classical beam theory:

$$\boldsymbol{\delta = \frac{Pab(L^2 - a^2 - b^2)}{6EIL}} \tag{5.3}$$

where all variables are defined in Section 3.3.2.

In the applied scenario, **three various loads** were positioned at $\frac{1}{5}L$, $\frac{1}{2}L$, and $\frac{1}{8}L$, and the total deflection was calculated by summing the contributions from each load. The formula was coded and implemented in Python, iterating over all 96 beams to determine the deflection values.

The formula used in the code is summarized as:

```
def manual_deflection(row):
    Iz = row['Izpmm4']
    term1 = 5000 * 2000 * 8000 * (L**2 - 2000**2 - 8000**2)
    term2 = 8000 * 6000 * 4000 * (L**2 - 6000**2 - 4000**2)
    term3 = 12000 * 8000 * 2000 * (L**2 - 8000**2 - 2000**2)
    return (term1 + term2 + term3) / (6 * E * Iz * L)
```

Figure 5.2 - Deflection Formula code

Where each term corresponds to the effect of a load placed at a specific location.

The deflection results showed strong inverse correlation with Second Moment of Area and depth — beams with larger I-values or greater depth exhibited lower deflection, affirming classical theory. Beams with thin webs and flanges were found to suffer excessive deformation despite offering lower weight and cost advantages.

These deflection results, in combination with plastic moment and cost, played a central role in the **scoring system** used to identify optimal beam configurations. A normalized scale was applied to compare all three metrics (plastic moment, deflection, and cost) to produce an objective and quantitative score for each beam.

5.2.3 Worked example: Beam #1 (manual calculations)

To demonstrate the calculation procedures and verify automated scripts, one beam from the dataset (Beam #1) is worked through manually. Below we calculate the section properties (I_z, Z_p), the plastic moment M_p, reactions and total deflection under the three-point loading scheme, and the resulting optimization score. The full step-by-step Excel worksheet and the Python scripts used to compute these values for the entire dataset are provided in the appendix.

5.2.3.1 Beam Identification & Given Data:

- Beam number = 1, Row number = 4.
- Geometry: (Depth = 127 mm, Flange Width = 76 mm, Flange Thickness = 7.6 mm, Web Thickness = 4 mm)
- Manufacturer I_z = 4730000 mm^4
- Material properties: E = 210,000 MPa; f_y = 350 MPa
- Loading: L = 10,000 mm; P1 = 5000 N @ 2000 mm; P2 = 8000 N @ 6000 mm; P3 = 12000 N @ 8000 mm

5.2.3.2 Shear Force and Moment Diagram:

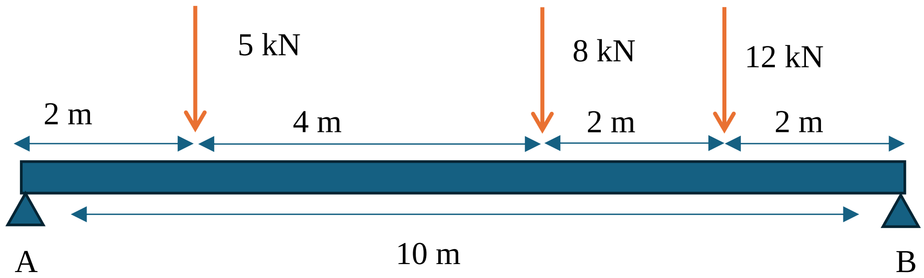


Figure 5.3 - Beam Forces

$$\sum F_y = 0 \rightarrow \; F_A + F_B - 5 - 8 - 12 = 0$$

$$\sum M_A = 0 \rightarrow \; -(2)\cdot(5) - (6)\cdot(8) - (8)\cdot(12) + (10)\cdot F_B = 0$$

$$F_B = \frac{154}{10} = 15.4\ kN \rightarrow F_A = 9.6\ kN$$

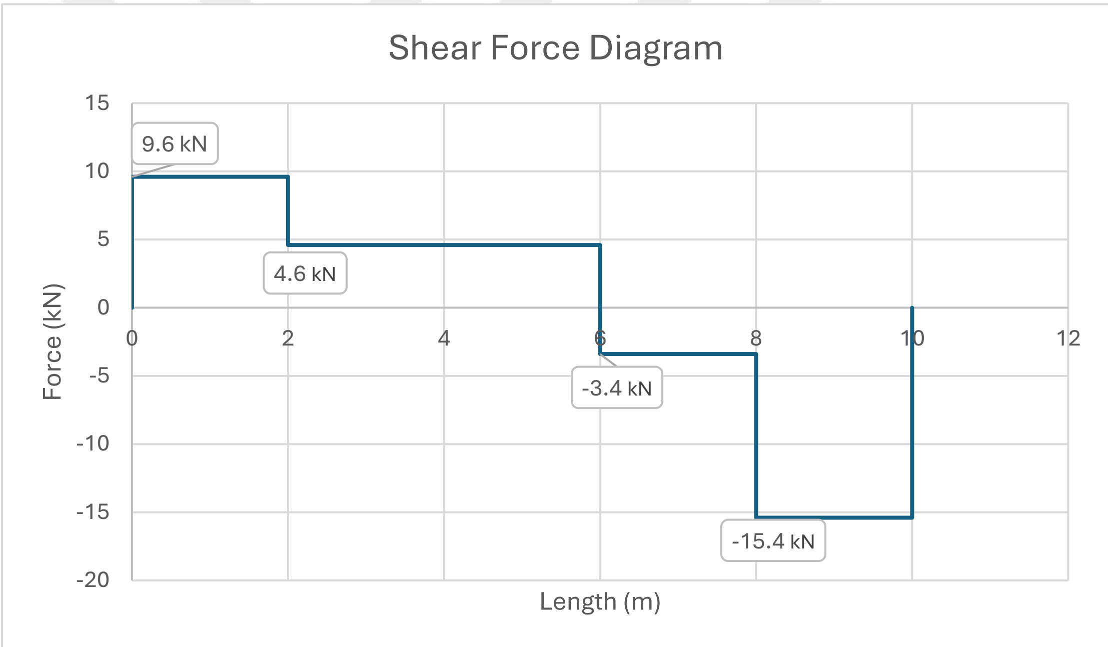


Figure 5.4 - Shear Force Diagram

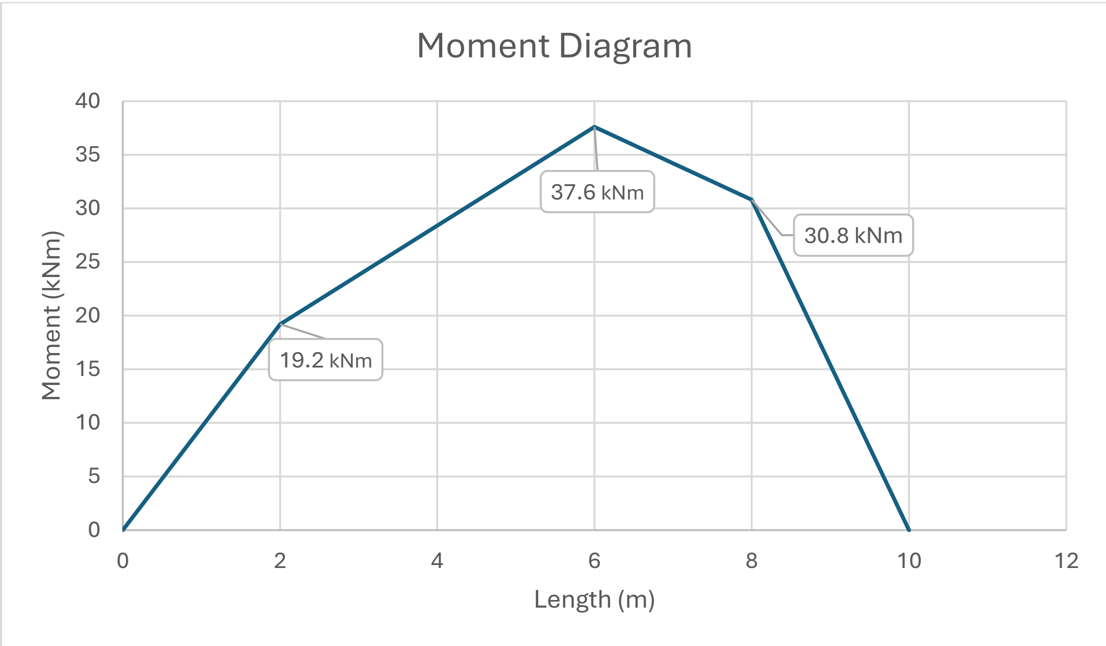


Figure 5.5 - Moment Diagram

$$M_{L=2} = 2 \cdot (9.6) = 19.2\ kNm$$

$$M_{L=6} = 6 \cdot (9.6) - 4 \cdot (5) = 37.6\ kNm$$

$$M_{L=8} = 8 \cdot (9.6) - (6) \cdot (5) - (8) \cdot (2) = 30.8\ kNm$$

5.2.3.3 Plastic Section Modulus for I-Beams:

$$Z_p = 2 \cdot A_f \cdot \left(\frac{h}{2} - \frac{t_b}{2}\right) + A_w \cdot \frac{h_w}{4} \tag{5.4}$$

Where:

- $A_f = b_b \cdot t_b$ = area of one flange
- $A_w = t_w \cdot h_w$ = area of the web
- b_b = flange width = 76 mm
- t_b = flange thickness = 7.6 mm
- t_w = web thickness = 4 mm
- h = overall section depth = 127 mm
- $h_w = h - 2t_b$ = clear height of the web

$$Z_p = 2 \cdot (76) \cdot (7.6) \cdot \left(\frac{127}{2} - \frac{7.6}{2}\right) + (4) \cdot \big(127 - 2(7.6)\big) \cdot \frac{127 - 2(7.6)}{4}$$

$$= 81646.68\ mm^3$$

The Plastic Modulus (Z_p) given by the manufacturer is 84200 mm^3

The calculated plastic modulus (81,464.68 mm³) differs by only ~3.25% from the tabulated value (84,200 mm³), which is within acceptable tolerance due to rounding and simplification in manual calculations.

5.2.3.4 Second Moment of Area:

$$I_z = 2 \cdot (\frac{(b_b) \cdot (t_b)^3}{12} + (b_b) \cdot (t_b) \cdot \left(\frac{(h) - (t_b)}{2}\right)^2) + \frac{(t_w) \cdot ((h) - 2 \cdot (t_b))^3}{12}$$

Where:

- b_b = flange width = 76 mm
- t_b = flange thickness = 7.6 mm
- t_w = web thickness = 4 mm
- h = overall section depth = 127 mm

$$I_z = 2 \cdot (\frac{(76) \cdot (7.6)^3}{12} + (76) \cdot (7.6) \cdot \left(\frac{(127) - (7.6)}{2}\right)^2) + \frac{(4) \cdot ((127) - 2 \cdot (7.6))^3}{12}$$

$$= 4588602.14\ mm^4$$

The Second Moment of Area (I_z) given by the manufacturer is 4730000 mm^4
The difference between them (~2.99%) is normal due to rounding and idealized geometry.

5.2.3.5 Plastic moment:

$$\boldsymbol{M_p = f_y \cdot Z_p} \quad (5.5)$$

Where:

- $\boldsymbol{M_p}$ is the plastic moment (kN·m)
- $\boldsymbol{f_y}$ is the yield strength of steel (MPa) = 350 MPa
- $\boldsymbol{Z_p}$ is the plastic section modulus (mm³) = 81646.68 mm^3

$$M_p = 350 \cdot 10^{-9} \cdot 81646.68 \cdot 1000$$

$$= 28.576 \text{ kN·m}$$

5.2.3.6 Deflection:

$$\boldsymbol{\delta_{total}} = \sum \frac{\boldsymbol{Pab}(L^2 - \boldsymbol{a^2} - \boldsymbol{b^2})}{\boldsymbol{6EIL}} \tag{5.6}$$

where:

- $\boldsymbol{\delta}$ = deflection at the point of interest (mm),
- P = magnitude of the point load (N), P1 = 5000 N @ 2000 mm;
  P2 = 8000 N @ 6000 mm; P3 = 12000 N @ 8000 mm
- a, b = distances from the load to the supports (mm),
- L = total span length of the beam (mm), = 10,000 mm
- E = Young's modulus of elasticity of the beam material (MPa or N/mm²),
  = 210000 Mpa.
- I = Second Moment of Area about the neutral axis (mm⁴), = 4730000 mm⁴

$$\boldsymbol{\delta_1} = \frac{(\mathbf{5000}) \cdot (\mathbf{2000}) \cdot (\mathbf{8000}) \cdot ((\mathbf{10000})^{\mathbf{2}} - (\mathbf{2000})^{\mathbf{2}} - (\mathbf{8000})^{\mathbf{2}})}{\mathbf{6} \cdot (\mathbf{210000}) \cdot (\mathbf{4730000}) \cdot (\mathbf{10000})}$$

= 42.95 mm

$$\boldsymbol{\delta_2} = \frac{(\mathbf{8000}) \cdot (\mathbf{6000}) \cdot (\mathbf{4000}) \cdot ((\mathbf{10000})^{\mathbf{2}} - (\mathbf{6000})^{\mathbf{2}} - (\mathbf{4000})^{\mathbf{2}})}{\mathbf{6} \cdot (\mathbf{210000}) \cdot (\mathbf{4730000}) \cdot (\mathbf{10000})}$$

= 154.64 mm

$$\boldsymbol{\delta_2} = \frac{(\mathbf{12000}) \cdot (\mathbf{8000}) \cdot (\mathbf{2000}) \cdot ((\mathbf{10000})^{\mathbf{2}} - (\mathbf{8000})^{\mathbf{2}} - (\mathbf{2000})^{\mathbf{2}})}{\mathbf{6} \cdot (\mathbf{210000}) \cdot (\mathbf{4730000}) \cdot (\mathbf{10000})}$$

= 103.09 mm

$$\boldsymbol{\delta_{total}} = \boldsymbol{\delta_1} + \boldsymbol{\delta_2} + \boldsymbol{\delta_3} = \mathbf{300.68\ mm}$$

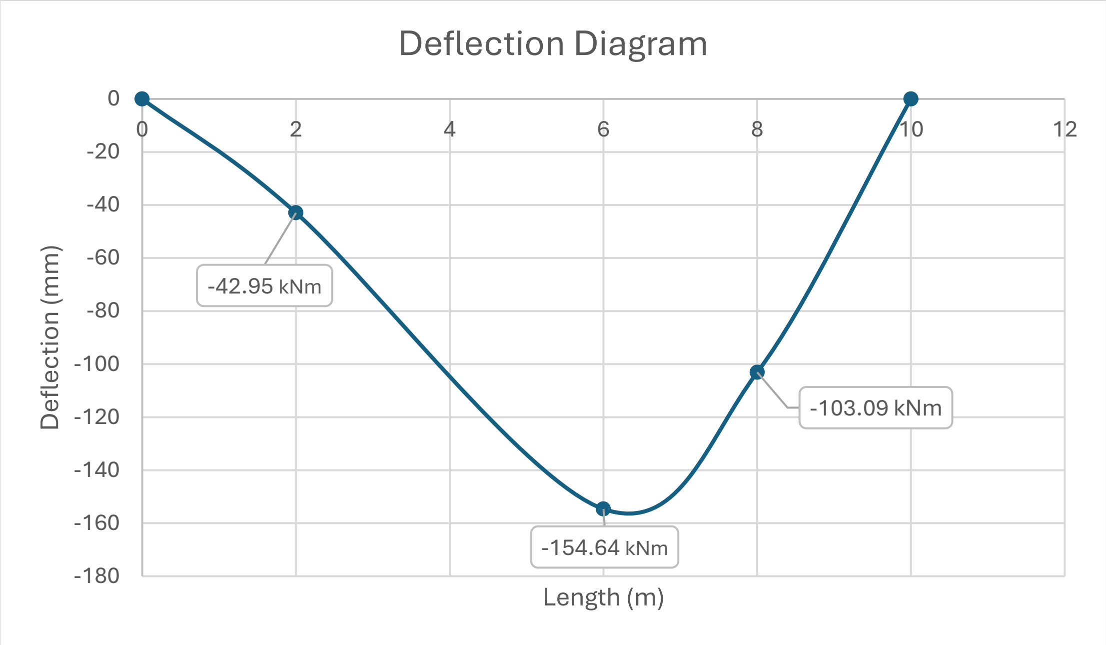


Figure 5.6 - Deflection Diagram

5.2.3.7 Score

$$Score = \left(\frac{Plastic\ Moment}{Cost}\right)_{normalized} - (Total\ Deflection)_{normalized} \tag{5.7}$$

Where:

$$Normalized\ value = \frac{Value_i}{max(Value)} \tag{5.8}$$

For Beam #1:

Plastic Moment M_p= 28.576 kN·m , Cost = 357.5 ₤.

Plastic Moment kN·m per Euro = 0.0799 kN.m/₤

Max Plastic Moment per Euro = 1.0697 kN·m

Total Deflection δ_{total} = 300.680 mm , Max Total Deflection = 300.680 mm

$$Normalized\ \frac{Plastic\ Moment}{Cost} = \frac{0.0799}{1.0697} = 0.0747$$

$$Normalized\ Total\ Deflection = \frac{300.680}{300.680} = 1$$

$$Score = (0.0747) - (1) = -0.9253$$

5.3 Machine Learning Model Results

In this section, we explore the results obtained from applying a Random Forest Regression model to the dataset of 96 universal I-beams. The model was used to predict two key structural performance indicators — the **plastic moment** and the **total deflection under three point-loads** — based on the geometrical input features. The insights derived from this model were later used to extrapolate beyond the original dataset and suggest new beam configurations that optimize performance.

5.3.1 Training and Testing Results

To ensure robustness and avoid overfitting, the 96-beam dataset was split into **training (80%) and testing (20%)** sets using **train_test_split** from the **sklearn.model_selection** module. The input features used were:

- Depthmm
- FlangeWidthmm
- FlangeThicknessmm
- WebThicknessmm

The target variables were:

- PlasticMomentKNM
- Deflectionmmdue3Loadsatonce

A **RandomForestRegressor** from **sklearn.ensemble** was trained on these features. The model was first trained on the training set and evaluated on the testing set. Key configuration parameters included:

- n_estimators = 100 (number of trees in the forest),
- random_state = 42 (for reproducibility),
- **Max** and **min values** we used as a boundaries for the predictions.

After fitting the model, predictions on the test set were compared to actual values using performance metrics such as **R^2 (coefficient of determination)**.

5.3.2 Accuracy and Predictive Performance

The Random Forest model demonstrated strong predictive capability for both target variables.

This indicates that the model could explain more than **90%** of the variance in both plastic moment and deflection, confirming its effectiveness.

The Random Forest model particularly excels in capturing **non-linear interactions** between input features, which is beneficial in this application since the relationship between beam dimensions and structural behavior is complex and non-linear.

Plots of **actual vs. predicted values** further confirmed the reliability of the model. These plots exhibited strong alignment along the diagonal line $y = x$, suggesting minimal bias or systemic deviation, with $R^2 = 0.953$

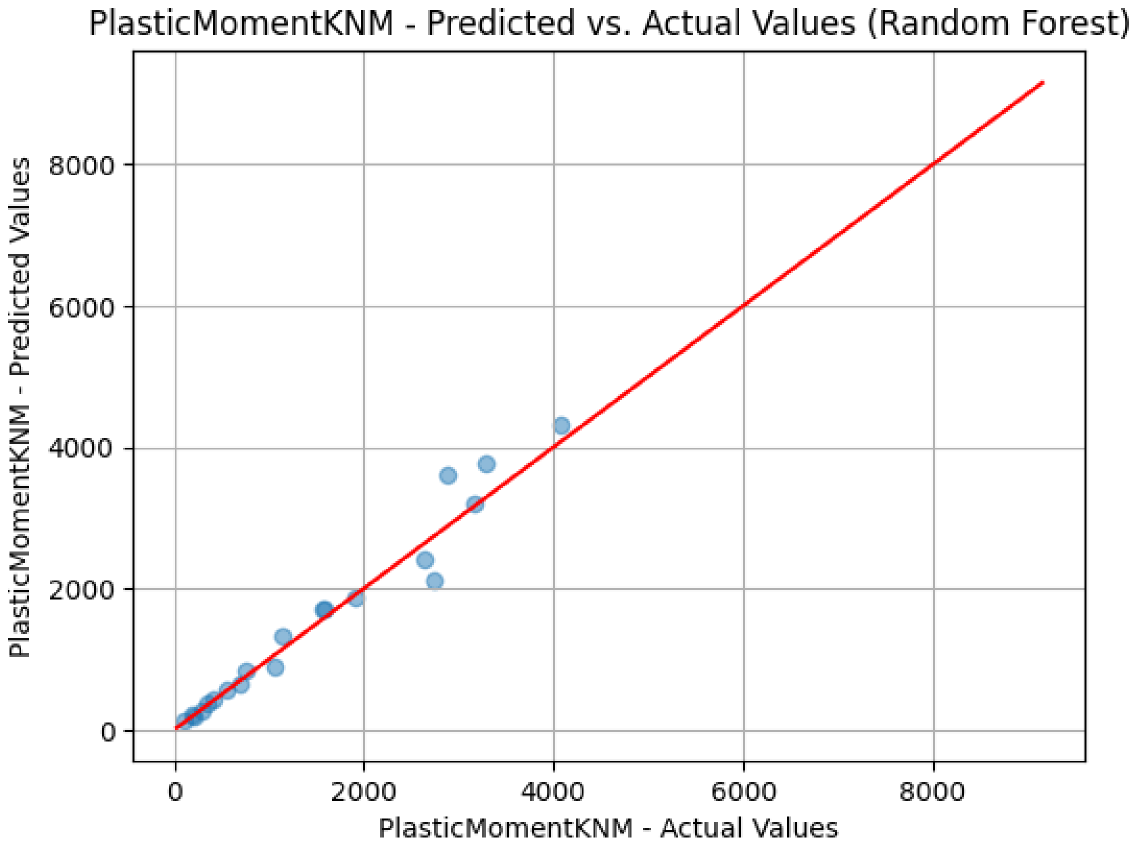


Figure 5.7 - Plastic Moment - Predicted vs Actual - Scatterplot

Using this code:

```
import matplotlib.pyplot as plt

# Plot the predicted vs. actual values
# predicted values of the target variable are plotted on the y-axis
# Actual values of the target variable are plotted on the x-axis
# If the model is a good fit, the points should be close to the diagonal line,
# indicating a strong linear relationship between the actual and predicted
values.

plt.scatter(Y_test, Y_pred, alpha=0.5)
plt.plot([Y.min(), Y.max()], [Y.min(), Y.max()], color="red")
plt.xlabel("PlasticMomentKNM - Actual Values")
plt.ylabel("PlasticMomentKNM - Predicted Values")
plt.title("PlasticMomentKNM - Predicted vs. Actual Values (Random Forest)")
plt.grid(True)
plt.show()
```

Figure 5.8 - Plastic Moment - Predicted vs Actual - code

The success of the model in terms of prediction accuracy and interpretability makes it a reliable tool for exploring and optimizing beam configurations beyond the current dataset, which is further discussed in the next sections.

5.4 Optimization Results

The final phase of the structural optimization process involved synthesizing results from plastic moment, deflection, and cost calculations to identify the most efficient beam configurations. This section details how the **scoring system** was developed and applied, the trade-offs between structural performance and economic considerations, and how robust the rankings are when the weights of criteria are varied.

5.4.1 Optimal Beam Selection

After calculating the **plastic moment** and **deflection** for each of the 96 beams in the dataset and assigning cost values per meter from the manufacturer, a **composite scoring system** was implemented to objectively evaluate and rank each beam.

The scoring system was defined as:

$$\boldsymbol{Score} = \left(\frac{\boldsymbol{Plastic\ Moment}}{\boldsymbol{Cost}}\right)_{\boldsymbol{normalized}} - (\boldsymbol{Total\ Deflection})_{\boldsymbol{normalized}} \tag{5.9}$$

Where:

- Plastic Moment / Cost represents the strength-to-price efficiency of the beam,
- Deflection is penalized since lower deflection is preferred,
- Each term is normalized to bring values to a common 0–1 range before combination.

This score aimed to strike a balance between **structural capacity**, **cost-effectiveness**, and **serviceability**.

Top 3 beams based on the scoring system were:

Table 5 - Top 3 beams based on the scoring system

Rank	**Depth (mm)**	**Flange Width (mm)**	**Flange Thickness (mm)**	**Web Thickness (mm)**	**Score**
1	910	304.1	23.9	15.9	0.999
2	951	310.0	43.9	24.4	0.962
3	923	307.0	30.0	18.4	0.826

These beams offered the best performance under the integrated score, balancing moment capacity with minimal deflection and acceptable cost.

In addition, **synthetic beam configurations** were generated using the trained Random Forest model. The model predicted valid combinations of Depthmm, FlangeWidthmm, FlangeThicknessmm, and WebThicknessmm that fall within the range of known values but were **not present in the original dataset**.

From those, 100 additional high-scoring candidates were identified:

Table 6 - Top 3 predicted beams

Rank	**Depth (mm)**	**Flange Width (mm)**	**Flange Thickness (mm)**	**Web Thickness (mm)**	**Predicted Score**
A	896.78	305.67	45.60	25.40	0.9975
B	896.78	305.67	39.13	25.40	0.9364
C	896.78	305.67	45.60	21.83	0.9169

These "suggested" configurations demonstrate the utility of machine learning in expanding the design space beyond the limited scope of predefined catalog beams.

5.4.2 Cost-Deflection Trade-Offs

A key observation from the optimization process was the inherent **trade-off** between cost and deflection. Heavier beams typically offer higher plastic moments and lower deflections, but at the expense of higher material cost and weight. Conversely, lighter beams are economical but prone to greater deformation and lower load capacity.

A visual representation of this trade-off was plotted to show cost on one axis and deflection on the other, It shows that higher-cost materials/designs tend to be stronger and deflect less.

Diminishing Returns:

The steepest decline in deflection happens between 0 and ~2000 Euros.

After that, the curve flattens — spending more gives less noticeable improvement in deflection (diminishing returns).

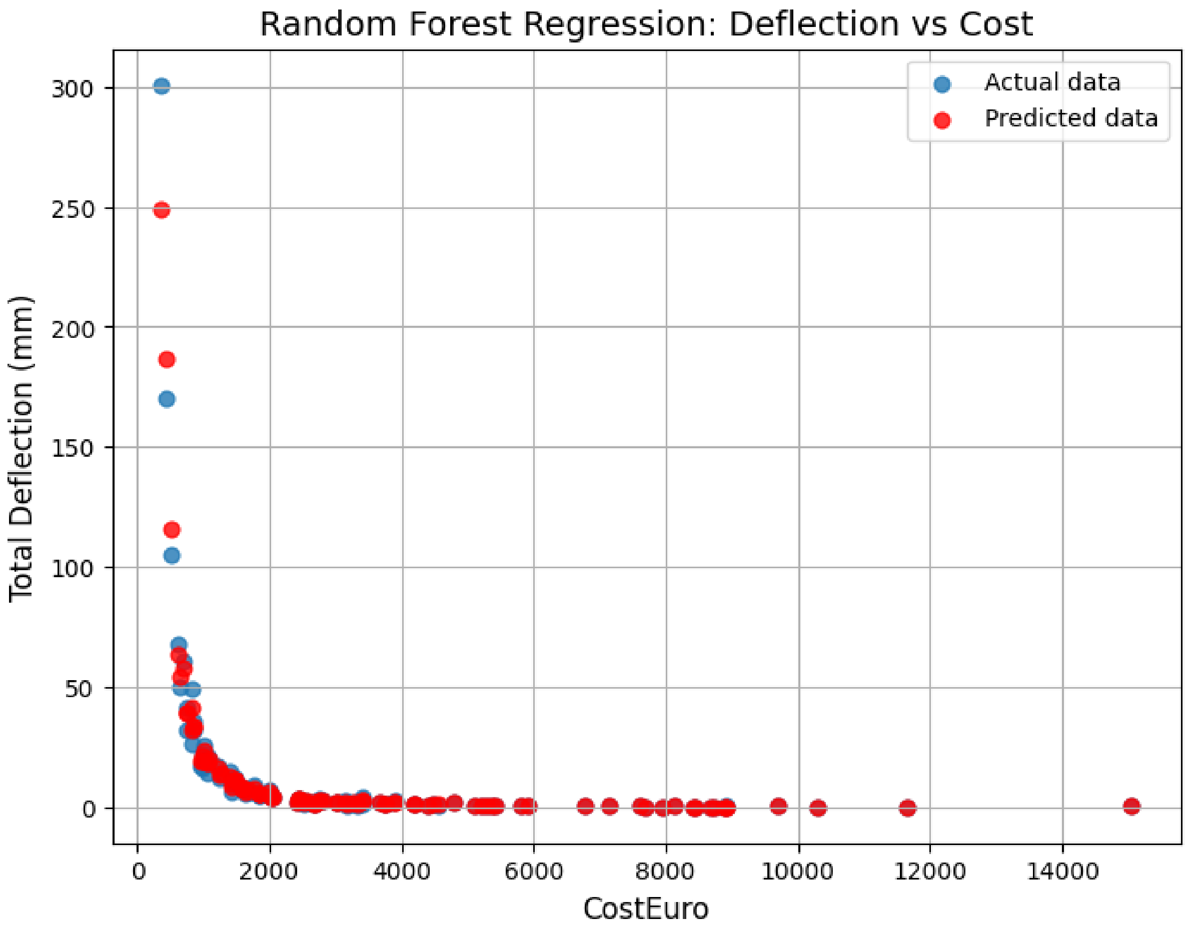


Figure 5.9 - Random Forest Regression - Deflection vs Cost

These findings confirm that selecting an optimal beam is not a one-dimensional task; it requires balancing multiple performance indicators. The scoring system used in this research effectively helped automate and simplify this decision-making process.

The optimization results validate the proposed approach as a practical, data-driven method to enhance beam selection. Whether using real cataloged beams or ML-suggested new configurations, the ability to quantify performance via scoring allows engineers to make **informed, optimized decisions** that consider both structural and economic constraints.

6. DISCUSSION

6.1 Summary of Key Findings

This study set out to develop a data-driven approach for identifying the most suitable universal I-beams based on a combination of structural performance and cost efficiency. A comprehensive database of 96 UK-manufactured I-beams was analyzed, focusing on plastic moment capacity, deflection behavior under three-point loading, and associated costs. A scoring model was established that integrates all three criteria to rank beam performance objectively. Additionally, a Random Forest regression model was developed to predict plastic moment capacity based on geometric properties, enhancing the ability to extrapolate beyond the initial dataset.

Through the addition of three varying point loads at different distances from supports, this thesis generated a more nuanced deflection dataset for each beam. The deflection responses, combined with cost data and plastic moment values, provided a holistic basis for scoring. The Random Forest model demonstrated strong predictive accuracy, and its output was used to generate hypothetical new beam configurations that were subsequently scored. Several of these machine-generated configurations outperformed existing beams in terms of combined efficiency.

6.2 Ensuring Geometric Realism of Synthetic Beam Sections

During the generation of synthetic universal beam profiles, it was observed that purely unconstrained predictions sometimes produced unrealistic proportions, such as flange widths exceeding the overall depth. To mitigate this, ratio-based bounds were derived from the manufacturer's database of standard beams. Specifically, three dimensionless parameters were considered: the flange width-to-depth ratio ($\boldsymbol{b_b}/d$), the flange thickness-to-depth ratio ($\boldsymbol{t_b}/d$), and the web thickness-to-depth ratio ($\boldsymbol{t_w}/d$). For each ratio, the 5th and 95th percentiles of the real dataset were adopted as lower and upper bounds. The resulting constraints were approximately:

- $0.3144 \leq \boldsymbol{b_b/d} \leq 0.5766$
- $0.0225 \leq \boldsymbol{t_b/d} \leq 0.0567$
- $0.0168 \leq \boldsymbol{t_w/d} \leq 0.0320$

Candidate synthetic beams were filtered accordingly, ensuring that the generated sections maintained realistic proportions consistent with rolled steel profiles available in practice. This approach allowed the Random Forest model to explore innovative combinations of dimensions while remaining within geometrically feasible limits.

6.3 Comparison with Traditional Design Methods

Traditional beam selection typically relies on standard design tables and conservative manual estimates, often prioritizing safety with minimal consideration for material cost or deflection optimization in early design stages. In contrast, the method developed in this thesis integrates multiple performance criteria into a unified, data-informed scoring system.

Where conventional methods might recommend a beam solely based on maximum moment capacity, this study introduces an evaluative framework that balances strength, serviceability (via deflection), and economic feasibility. Furthermore, while classical design may lack predictive capability, the application of machine learning, particularly the Random Forest model, allows for accurate estimation of plastic moment values from beam geometry—without resorting to iterative calculation or finite element analysis.

This approach marks a departure from deterministic selection toward probabilistic forecasting and optimization, aligning well with modern engineering trends.

6.4 Implications for Structural Engineering

The methodology proposed in this thesis has significant implications for structural engineering practice. By shifting beam selection from manual, table-driven procedures to data-informed and machine-learning-assisted processes, engineers can make more informed decisions earlier in the design cycle.

The integration of predictive modeling enables the exploration of beam geometries not present in standard catalogs, potentially uncovering novel or underutilized sections that satisfy performance requirements more efficiently. This could lead to reduced material usage, cost savings, and improved structural behavior, especially for large-scale or cost-sensitive projects.

Moreover, the scoring model offers a customizable framework that could be expanded or adapted to incorporate other critical factors, such as fabrication constraints, environmental impacts, or dynamic performance.

6.5 Practical Applications in Industry

From a practical standpoint, the outcomes of this thesis have direct relevance to industry workflows. Structural designers, especially those engaged in preliminary design or cost estimation, can benefit from the rapid evaluation of multiple beam options without manually checking each candidate against multiple criteria.

The predictive model for plastic moment could be integrated into structural analysis software or as a plugin within BIM (Building Information Modeling) platforms, where automatic section selection is a valuable tool. Similarly, the scoring framework could be adapted as a decision-support system for procurement or design optimization tools.

Steel manufacturers may also find value in this approach, using it to assess the competitiveness of their product offerings or to guide the development of new beam profiles tailored to market demand for cost-effective and high-performance sections.

6.6 Limitations of the Study and Assumptions

Several limitations and assumptions underpin the findings of this research. Firstly, the study assumed idealized support conditions (simple supports) and did not account for real-world complexities such as fixity, lateral-torsional buckling, or end restraints. Additionally, material properties were considered homogeneous and elastic-perfectly plastic, neglecting variations due to rolling processes or heat treatment.

The deflection calculations were based on linear elastic theory, and dynamic or time-dependent effects (e.g., creep or fatigue) were not considered. Load configurations were limited to three-point loads, and distributed loads or moment applications were excluded for simplicity.

Moreover, the Random Forest model, while effective, is inherently limited by the size and diversity of the input dataset. The generation of hypothetical high-scoring beams is theoretical, and their structural feasibility would need to be verified through physical testing or detailed finite element analysis before implementation.

6.7 Directions for Future Research

Building upon the foundation established in this thesis, several avenues for future research are recommended. Firstly, expanding the dataset to include other types of structural sections (e.g., welded I-beams, hollow sections, or custom shapes) would provide a broader scope for analysis and machine learning training.

Future studies could incorporate more complex loading conditions, including uniformly distributed loads, varying span lengths, or combinations of axial and bending forces. The inclusion of buckling behavior and dynamic load analysis would also add realism to the performance evaluation.

From a computational perspective, the use of hybrid AI models—such as combining Random Forest with Genetic Algorithms or Neural Networks—could improve prediction accuracy and enable automated optimization of section dimensions. Integrating sustainability metrics, such as embodied carbon or recyclability, into the scoring model would also align the research with modern green building initiatives.

Lastly, experimental validation of the machine-suggested beam designs would be a vital step to verify their real-world performance and potentially inform future revisions to structural design codes.

7. CONCLUSION

7.1 Recap of the Research Objectives and Contributions

The primary objective of this research was to develop a comprehensive and data-driven framework for optimizing the selection of universal I-beams based on structural performance and cost-efficiency. Using a database of 96 UK steel factory beams, the study sought to evaluate each beam's plastic moment capacity, deflection behavior under three-point loading over a 10-meter span, and associated material cost. A unique scoring model was introduced to integrate these three core parameters into a single evaluative metric.

To push beyond the constraints of the existing dataset, a machine learning model—Random Forest regression—was employed to predict plastic moment from beam geometric properties, allowing for the exploration of new, high-performing beam configurations. Further contributions include the generation of extended deflection data under three different point loads and the development of a methodology to rank beam efficiency holistically.

In essence, this thesis combined structural engineering principles with machine learning to propose an innovative and scalable approach for optimal beam selection.

7.2 Key Findings and Results

The following key results emerged from the study:

- A scoring system was successfully developed that integrates plastic moment capacity, deflection behavior, and cost, allowing for objective comparison across all beams in the dataset.
- Deflection results under various loading scenarios revealed notable differences in serviceability performance even among beams with similar moment capacities.
- The Random Forest model accurately predicted plastic moment capacity with a high degree of reliability, highlighting the viability of supervised machine learning for structural property estimation.

- Several hypothetical beam configurations, generated and scored using machine learning predictions, outperformed existing catalog beams, showing potential for improved structural and economic outcomes.
- The methodology enables early-phase beam selection based on multidimensional performance metrics, reducing reliance on traditional trial-and-error approaches or overly conservative designs.

7.3 Final Remarks on Structural Optimization Using Machine Learning

This research illustrates the powerful synergy between structural engineering and machine learning. Traditional methods, while reliable, often isolate strength, deflection, and cost into separate decision tracks. The approach developed here integrates these aspects into a unified optimization tool that enhances decision-making clarity, especially during early design phases.

Machine learning, particularly Random Forest in this case, proved effective not only in predicting structural properties from basic geometric input but also in enabling the design and evaluation of theoretical beam configurations beyond the boundaries of existing catalogs. This lays the groundwork for a paradigm shift in structural design—moving from catalog-based selections to intelligent, data-informed recommendations.

While this integration does not replace conventional engineering judgment or code-based design requirements, it complements them by offering new avenues for optimization, speed, and innovation.

7.4 Suggestions for Future Work in Structural Optimization and Machine Learning

There is considerable scope to expand and deepen this line of research. The following suggestions are made for future work:

- **Dataset Expansion**: Incorporate a broader range of beam profiles, including non-standard or fabricated sections, and introduce additional material and fabrication variables for more robust predictions.
- **Load Complexity**: Extend the analysis to include distributed loads, real-world load combinations, or dynamic actions such as seismic or wind-induced forces.

- **Buckling and Stability**: Integrate lateral-torsional buckling analysis and include slenderness effects to more accurately reflect true structural behavior under bending.
- **AI Model Enhancement**: Explore hybrid models (e.g., combining Random Forest with optimization algorithms like Genetic Algorithms or Reinforcement Learning) to move from prediction to real-time design optimization.
- **Sustainability Metrics**: Add environmental impact factors—such as embodied carbon or lifecycle energy use into the scoring model to promote greener structural solutions.
- **Software Integration**: Implement the system as a plugin for BIM platforms or standalone design aid software, allowing practicing engineers to apply the framework directly within their workflows.
- **Experimental Validation**: Validate machine-generated beam suggestions with physical testing or high-fidelity finite element analysis to assess performance under real conditions.

By building on the foundations laid in this thesis, future studies can further bridge the gap between classical structural engineering and modern data science, ultimately contributing to smarter, safer, and more sustainable structural design.

8. DOCUMENTATION AND REPORTING:

Reporting Results: The findings are systematically documented, including a detailed discussion of the methodology, analysis, and conclusions.

Python Code and Data: The Python code used for the model, along with the dataset (subject to data privacy and proprietary restrictions), will be made available for review and potential replication of the study.

Githubs codes by the following coders:

DataProfessor.

FlatPlanet

In addition to some assistance from AI

Google Colab link:

https://colab.research.google.com/drive/1UPbvNyLWO0ta8fwiRpvhd-h7y0DkRX6k?usp=sharing

9. REFERENCES:

BCSA, & Tata Steel. (2013). *Handbook of structural steelwork* (7th ed., Publication No. 55/13). British Constructional Steelwork Association.

Charalampakis, A. E., & Papanikolaou, V. K. (2021). Machine learning design of R/C columns. *Engineering Structures, 226*, 111412. https://www.sciencedirect.com/science/article/abs/pii/S014102962034013X

Dabiri, H., Rahimzadeh, K., & Kheyroddin, A. (2022). A comparison of machine learning- and regression-based models for predicting ductility ratio of RC beam-column joints. *Structures, 37*, 69–81. https://www.sciencedirect.com/science/article/abs/pii/S2352012421012017

Dataprofessor. (2020). *Code: linear_regression* [Source code]. GitHub. https://github.com/dataprofessor/code

Dlubal Software. (n.d.). *Cross-section properties of UB (Universal Beams) – BS 4-1:1993*. Retrieved October 23, 2025, from https://www.dlubal.com/en/cross-section-properties/series-ub-bs-4-1-1993

Flatplanet. (2023). *Pandas YouTube playlist: linreg* [Video series]. GitHub. https://github.com/flatplanet/Pandas-Youtube-Playlist

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (2nd ed.). O'Reilly Media.

Google Research. (2025). *Google Colaboratory* [Computer software]. https://colab.research.google.com

Hibbeler, R. C. (2012). *Structural analysis* (8th ed.). Pearson Education.

Louppe, G. (2014). *Understanding random forests: From theory to practice* (Doctoral dissertation, University of Liège). University of Liège.

Málaga-Chuquitaype, C. (2022). Machine learning in structural design: An opinionated review. *Frontiers in Built Environment, 8*, 815717. https://www.frontiersin.org/journals/built-environment/articles/10.3389/fbuil.2022.815717/full

OneSteel Manufacturing. (2014). *Hot rolled and structural steel products* (7th ed.). OneSteel Manufacturing.

OpenAI. (2025). *ChatGPT (GPT-5)* [Large language model]. https://chat.openai.com

Pharell, H. (2023). *Python programming and SQL* [Self-published].

SkyCiv Engineering. (n.d.). *Steel I-beam sizes – British universal beams (metric)*. Retrieved October 23, 2025, from https://skyciv.com/steel-i-beam-sizes/

Steel Beams Direct. (2025). *Universal beams*. Retrieved October 23, 2025, from https://www.steelbeamsdirect.com/product/universal-beams/

10.APPENDICES:

10.1 Appendix A: Raw Data (96 I-Beams Dataset)

Mass Kg Per m	Depth mm	Flange Width mm	Flange Thickness mm	Web Thickness mm	Izp mm4	Cost Euro	Plastic Moment KNM	Deflection mm due 5KN Load at 2m	Deflection mm due 8KN Load at 6m	Deflection mm due 12KN Load at 8m	Deflection mm due 3 Loads at once
13	127.0	76.0	7.6	4.0	4730000	357.5	28.513	42.955	154.636	103.091	300.681
16	152.4	88.7	7.7	4.5	8340000	440.0	41.980	24.362	87.701	58.468	170.530
19	177.8	101.2	7.9	4.8	13560000	525.0	58.564	14.983	53.940	35.960	104.884
23	203.2	101.8	9.3	5.4	21050000	635.0	80.352	9.652	34.747	23.165	67.564
25	203.2	133.2	7.8	5.7	28410000	660.0	88.599	7.152	25.746	17.164	50.061
30	206.8	133.9	9.6	6.4	23400000	690.0	88.607	8.683	31.258	20.838	60.779
22	254.0	101.6	6.8	5.7	44550000	750.0	117.208	4.561	16.418	10.945	31.924
25	257.2	101.9	8.4	6.0	34150000	750.0	104.878	5.950	21.418	14.279	41.646
28	260.4	102.2	10.0	6.3	28960000	825.0	108.430	7.016	25.257	16.838	49.110
31	251.4	146.1	8.6	6.0	53660000	830.0	138.520	3.786	13.631	9.087	26.504
37	256.0	146.4	10.9	6.3	44130000	850.0	135.570	4.604	16.574	11.050	32.228
43	259.6	147.3	12.7	7.2	40050000	850.0	121.426	5.073	18.263	12.175	35.511
25	305.1	101.6	7.0	5.8	82490000	955.0	184.896	2.463	8.867	5.911	17.241
28	308.7	101.8	8.8	6.0	65010000	990.0	165.794	3.125	11.251	7.501	21.877
33	312.7	102.4	10.8	6.6	85030000	1000.0	214.756	2.389	8.602	5.735	16.726
37	304.4	123.4	10.7	7.1	55370000	1010.0	167.128	3.669	13.210	8.807	25.686
42	307.2	124.3	12.1	8.0	101700000	1075.0	225.364	1.998	7.192	4.795	13.985
48	311.0	125.3	14.0	9.0	65440000	1075.0	196.213	3.105	11.177	7.451	21.733
40	303.4	165.0	10.2	6.0	71710000	1100.0	185.484	2.833	10.200	6.800	19.833

Mass Kg Per m	Depth mm	Flange Width mm	Flange Thickness mm	Web Thickness mm	Izp mm4	Cost Euro	Plastic Moment KNM	Deflection mm due 5KN Load at 2m	Deflection mm due 8KN Load at 6m	Deflection mm due 12KN Load at 8m	Deflection mm due 3 Loads at once
46	306.6	165.7	11.8	6.7	81960000	1230.0	211.406	2.479	8.924	5.950	17.353
54	310.4	166.9	13.7	7.9	120700000	1250.0	266.000	1.683	6.060	4.040	11.783
33	349.0	125.4	8.5	6.0	98990000	1265.0	248.696	2.053	7.389	4.926	14.367
39	353.4	126.0	10.7	6.6	95750000	1400.0	245.419	2.122	7.639	5.093	14.854
45	351.4	171.1	9.7	7.0	141400000	1420.0	308.484	1.437	5.173	3.449	10.058
51	355.0	171.5	11.5	7.4	213700000	1430.0	376.996	0.951	3.423	2.282	6.655
57	358.0	172.2	13.0	8.1	125100000	1470.0	247.408	1.624	5.847	3.898	11.369
67	363.4	173.2	15.7	9.1	117000000	1485.0	292.806	1.737	6.252	4.168	12.156
39	398.0	141.8	8.6	6.4	160400000	1580.0	348.432	1.267	4.560	3.040	8.867
53	406.6	143.3	12.9	7.9	187200000	1590.0	363.243	1.085	3.907	2.605	7.597
46	403.2	142.2	11.2	6.8	182830000	1600.0	354.961	1.111	4.001	2.667	7.779
54	402.6	177.7	10.9	7.7	255000000	1650.0	443.927	0.797	2.868	1.912	5.577
60	406.4	177.9	12.8	7.9	216000000	1700.0	413.933	0.941	3.386	2.258	6.584
67	409.4	178.8	14.3	8.8	156900000	1770.0	304.790	1.295	4.662	3.108	9.065
74	412.8	179.5	16.0	9.5	293800000	1850.0	508.240	0.692	2.490	1.660	4.841
52	449.8	152.4	10.9	7.6	289300000	1850.0	501.958	0.702	2.528	1.686	4.916
60	454.6	152.9	13.3	8.1	243300000	1850.0	465.229	0.835	3.006	2.004	5.846
67	458.0	153.8	15.0	9.0	194600000	2000.0	418.684	1.044	3.759	2.506	7.308
74	462.0	154.4	17.0	9.6	333200000	2030.0	571.836	0.610	2.195	1.463	4.268
82	465.8	155.3	18.9	10.5	326700000	2030.0	562.687	0.622	2.239	1.493	4.353
67	453.4	189.9	12.7	8.5	273100000	2040.0	519.402	0.744	2.678	1.786	5.208
74	457.0	190.4	14.5	9.0	365900000	2050.0	627.405	0.555	1.999	1.333	3.887
82	460.0	191.3	16.0	9.9	797790000	2430.0	1315.824	0.255	0.917	0.611	1.783
89	463.4	191.9	17.7	10.5	410200000	2450.0	698.158	0.495	1.783	1.189	3.467
98	467.2	192.8	19.6	11.4	370500000	2450.0	634.331	0.548	1.974	1.316	3.839
106	469.2	194.0	20.6	12.6	475400000	2460.0	708.493	0.427	1.539	1.026	2.992

Mass Kg Per m	Depth mm	Flange Width mm	Flange Thickness mm	Web Thickness mm	Izp mm4	Cost Euro	Plastic Moment KNM	Deflection mm due 5KN Load at 2m	Deflection mm due 8KN Load at 6m	Deflection mm due 12KN Load at 8m	Deflection mm due 3 Loads at once
133	480.6	196.7	26.3	15.3	860880000	2520.0	1252.472	0.236	0.850	0.566	1.652
161	492.0	199.4	32.0	18.0	488730000	2540.0	829.435	0.416	1.497	0.998	2.910
82	528.3	208.8	13.2	9.6	638410000	2660.0	1067.804	0.318	1.146	0.764	2.228
92	533.1	209.3	15.6	10.1	1232220000	2670.0	1748.584	0.165	0.594	0.396	1.154
101	536.7	210.0	17.4	10.8	1006330000	2690.0	1437.520	0.202	0.727	0.485	1.413
109	539.5	210.8	18.8	11.6	410580000	2750.0	620.885	0.495	1.782	1.188	3.464
122	544.5	211.9	21.3	12.7	552300032	2760.0	814.007	0.368	1.324	0.883	2.575
272	577.1	320.2	37.6	21.1	457300000	2800.0	774.725	0.444	1.600	1.066	3.110
66	524.7	165.1	11.4	8.9	757800000	3030.0	994.650	0.268	0.965	0.644	1.877
74	529.1	165.9	13.6	9.7	615200000	3030.0	902.182	0.330	1.189	0.793	2.312
85	534.9	166.5	16.5	10.3	486310000	3150.0	725.489	0.418	1.504	1.003	2.925
138	549.1	213.9	23.6	14.7	1985780000	3190.0	2738.439	0.102	0.368	0.246	0.716
150	542.5	312.0	20.3	12.7	668200000	3270.0	977.926	0.304	1.095	0.730	2.128
182	550.7	314.5	24.4	15.2	1509760000	3330.0	2126.158	0.135	0.485	0.323	0.942
219	560.3	317.4	29.2	18.3	873200000	3400.0	1134.537	0.233	0.838	0.558	1.629
82	598.6	177.9	12.8	10.0	350280000	3400.0	534.306	0.580	2.088	1.392	4.060
92	603.0	178.8	15.0	10.9	760400000	3660.0	1106.435	0.267	0.962	0.641	1.870
100	607.4	179.2	17.2	11.3	1180000000	3750.0	1375.781	0.172	0.620	0.413	1.205
101	602.6	227.6	14.8	10.5	986099968	3750.0	1272.812	0.206	0.742	0.495	1.442
113	607.6	228.2	17.3	11.1	725280000	3870.0	961.334	0.280	1.009	0.672	1.961
125	612.2	229.0	19.6	11.9	558690000	3900.0	754.166	0.364	1.309	0.873	2.546
140	617.2	230.2	22.1	13.1	1363000064	4200.0	1573.168	0.149	0.537	0.358	1.043
149	612.4	304.8	19.7	11.8	1118000000	4200.0	1435.980	0.182	0.654	0.436	1.272
179	620.2	307.1	23.6	14.1	1684999936	4410.0	1775.330	0.121	0.434	0.289	0.844
238	635.8	311.4	31.4	18.4	1259000064	4470.0	1584.616	0.161	0.581	0.387	1.130
125	677.9	253.0	16.2	11.7	1504000000	4560.0	1727.992	0.135	0.486	0.324	0.946

Mass Kg Per m	Depth mm	Flange Width mm	Flange Thickness mm	Web Thickness mm	Izp mm4	Cost Euro	Plastic Moment KNM	Deflection mm due 5KN Load at 2m	Deflection mm due 8KN Load at 6m	Deflection mm due 12KN Load at 8m	Deflection mm due 3 Loads at once
140	683.5	253.7	19.0	12.4	645770000	4790.0	865.099	0.315	1.133	0.755	2.202
152	687.5	254.5	21.0	13.2	1703000064	5100.0	1948.600	0.119	0.430	0.286	0.835
170	692.9	255.8	23.7	14.5	2052999936	5200.0	2140.084	0.099	0.356	0.238	0.693
134	750.0	264.4	15.5	12.0	2460000000	5280.0	2345.128	0.083	0.297	0.198	0.578
147	754.0	265.2	17.5	12.8	1530000000	5370.0	1918.439	0.133	0.478	0.319	0.930
173	762.2	266.7	21.6	14.3	1507000064	5430.0	1596.354	0.135	0.485	0.324	0.944
197	769.8	268.0	25.4	15.6	2792000000	5820.0	2636.470	0.073	0.262	0.175	0.509
176	834.9	291.7	18.8	14.0	2400000000	5910.0	2479.200	0.085	0.305	0.203	0.593
194	840.7	292.4	21.7	14.7	3396999936	6780.0	3166.616	0.060	0.215	0.144	0.419
226	850.9	293.8	26.8	16.1	2095000064	7140.0	2597.039	0.097	0.349	0.233	0.679
201	903.0	303.3	20.2	15.1	4362999808	7600.0	3782.896	0.047	0.168	0.112	0.326
224	910.4	304.1	23.9	15.9	9820000000	7680.0	8215.418	0.021	0.075	0.050	0.145
238	915.0	305.0	25.9	16.5	5480000000	7960.0	4723.759	0.037	0.134	0.089	0.260
253	918.4	305.5	27.9	17.3	3764000000	8130.0	3290.275	0.054	0.194	0.130	0.378
271	923.0	307.0	30.0	18.4	8860000000	8410.0	7440.347	0.023	0.083	0.055	0.161
289	926.6	307.7	32.0	19.5	6260000000	8430.0	5325.296	0.033	0.117	0.078	0.227
313	932.0	309.0	34.5	21.1	5041999872	8670.0	4352.591	0.040	0.145	0.097	0.282
345	943.0	308.0	39.9	22.1	7880000000	8730.0	6648.570	0.026	0.093	0.062	0.181
381	951.0	310.0	43.9	24.4	11020000000	8890.0	9150.603	0.018	0.066	0.044	0.129
425	961.0	313.0	49.0	26.9	6970000000	8910.0	5911.472	0.029	0.105	0.070	0.204
474	971.0	316.0	54.1	30.0	4720000000	8910.0	4077.664	0.043	0.155	0.103	0.301
521	981.0	319.0	58.9	33.0	4060000000	9690.0	3533.965	0.050	0.180	0.120	0.350
576	993.0	322.0	65.0	36.1	6257999872	10300.0	5343.902	0.033	0.117	0.078	0.227
343	911.8	418.5	32.0	19.4	7196000256	11650.0	6109.801	0.028	0.102	0.068	0.198
388	921.0	420.5	36.6	21.4	3252999936	15050.0	2876.129	0.063	0.225	0.150	0.437

10.2 Appendix B: Top Predicted 100 beams

Depth mm	Flange Width mm	Flange Thickness mm	Web Thickness mm	Izp mm4	Plastic Moment KNM	Deflection mm due 3 Loads at once	Cost Euro	Normalized Moment Cost	Normalized Deflection	Score
896.778	305.667	45.600	25.400	6160600883.0	6175.196	0.231	8400.6	1.000000	0.002464	0.997536
896.778	305.667	39.133	25.400	5563028240.0	5780.532	0.256	8373.4	0.939130	0.002728	0.936401
896.778	305.667	45.600	21.833	6005217809.0	5667.743	0.237	8385.8	0.919444	0.002527	0.916917
896.778	305.667	45.600	18.267	5849834735.0	5353.057	0.243	8188.7	0.889296	0.002595	0.886702
896.778	305.667	26.200	25.400	4310021457.0	5278.829	0.330	8183.3	0.877544	0.003521	0.874022
896.778	305.667	32.667	25.400	4946269485.0	5272.612	0.288	8314.6	0.862669	0.003068	0.859600
511.889	229.111	13.267	14.700	518001373.0	1827.393	2.746	2797.5	0.888631	0.029300	0.859331
896.778	305.667	39.133	21.833	5400040471.0	5273.080	0.263	8358.6	0.858204	0.002811	0.855393
511.889	229.111	19.733	14.700	677000113.7	1789.624	2.101	2774.8	0.877384	0.022419	0.854965
511.889	267.389	13.267	14.700	581144235.6	1827.277	2.447	2867.2	0.866974	0.026116	0.840857
511.889	267.389	19.733	14.700	768528251.8	1799.103	1.851	2854.3	0.857464	0.019749	0.837716
896.778	305.667	39.133	18.267	5237052702.0	4974.852	0.272	8174.3	0.827921	0.002898	0.825023
511.889	229.111	26.200	14.700	827527193.7	1714.661	1.719	2798.5	0.833513	0.018341	0.815173
896.778	305.667	26.200	21.833	4131087955.0	4912.287	0.344	8175.7	0.817369	0.003674	0.813695
511.889	267.389	26.200	14.700	945928202.8	1746.610	1.504	2875.7	0.826251	0.016045	0.810206
993.000	343.944	45.600	25.400	8596420924.0	5573.499	0.165	9465.2	0.801047	0.001766	0.799281
896.778	305.667	32.667	21.833	4775432859.0	4884.035	0.298	8300.8	0.800421	0.003178	0.797243
896.778	305.667	26.200	18.267	3952154453.0	4633.001	0.360	7985.1	0.789299	0.003840	0.785459
993.000	343.944	45.600	21.833	8378443272.0	5372.627	0.170	9446.0	0.773746	0.001811	0.771934
993.000	343.944	39.133	25.400	7746742472.0	5348.793	0.184	9439.2	0.770868	0.001959	0.768909
896.778	305.667	32.667	18.267	4604596234.0	4603.817	0.309	8116.4	0.771639	0.003296	0.768343
993.000	343.944	52.067	25.400	9422073944.0	5229.018	0.151	9476.4	0.750648	0.001611	0.749037
993.000	343.944	45.600	18.267	8160465620.0	5045.394	0.174	9227.3	0.743841	0.001860	0.741981
993.000	343.944	39.133	21.833	7519251171.0	5147.921	0.189	9420.0	0.743431	0.002018	0.741412

Depth mm	Flange Width mm	Flange Thickness mm	Web Thickness mm	Izp mm4	Plastic Moment KNM	Deflection mm due 3 Loads at once	Cost Euro	Normalized Moment Cost	Normalized Deflection	Score
896.778	343.944	45.600	25.400	6793504242.0	4838.316	0.209	8964.5	0.734223	0.002234	0.731988
993.000	343.944	45.600	28.967	8814398576.0	5094.896	0.161	9466.5	0.732159	0.001722	0.730437
993.000	343.944	52.067	21.833	9213340933.0	5078.187	0.154	9456.8	0.730506	0.001647	0.728859
993.000	343.944	32.667	25.400	6872694022.0	4988.867	0.207	9352.6	0.725653	0.002208	0.723445
993.000	343.944	26.200	25.400	5973931009.0	4893.643	0.238	9219.1	0.722110	0.002541	0.719569
993.000	343.944	39.133	18.267	7291759871.0	4873.084	0.195	9201.1	0.720483	0.002081	0.718402
993.000	343.944	39.133	28.967	7974233772.0	4970.270	0.178	9440.5	0.716217	0.001903	0.714314
993.000	343.944	52.067	18.267	9004607923.0	4854.514	0.158	9237.9	0.714878	0.001686	0.713193
896.778	343.944	45.600	21.833	6638121168.0	4664.667	0.214	8943.9	0.709501	0.002286	0.707215
896.778	343.944	39.133	25.400	6114317352.0	4643.301	0.233	8926.9	0.707597	0.002482	0.705114
993.000	343.944	32.667	21.833	6635416208.0	4852.800	0.214	9336.8	0.707056	0.002287	0.704769
993.000	343.944	26.200	21.833	5726589959.0	4757.576	0.248	9203.3	0.703237	0.002650	0.700587
993.000	343.944	32.667	28.967	7109971835.0	4793.768	0.200	9354.3	0.697149	0.002135	0.695014
993.000	343.944	26.200	28.967	6221272059.0	4716.882	0.229	9214.0	0.696412	0.002440	0.693973
511.889	190.833	19.733	14.700	585471975.5	1364.088	2.429	2606.3	0.711996	0.025923	0.686073
896.778	343.944	45.600	18.267	6482738094.0	4403.433	0.219	8710.0	0.687753	0.002341	0.685412
993.000	343.944	32.667	18.267	6398138394.0	4582.105	0.222	9117.1	0.683704	0.002372	0.681331
993.000	343.944	26.200	18.267	5479248909.0	4508.639	0.260	8973.8	0.683484	0.002770	0.680714
993.000	343.944	52.067	28.967	9630806955.0	4750.415	0.148	9477.7	0.681849	0.001576	0.680273
896.778	343.944	39.133	21.833	5951329583.0	4469.652	0.239	8906.3	0.682709	0.002550	0.680159
896.778	343.944	26.200	25.400	4690179569.0	4357.014	0.303	8711.7	0.680371	0.003236	0.677135
896.778	343.944	32.667	25.400	5413323983.0	4378.203	0.263	8839.8	0.673772	0.002804	0.670968
896.778	343.944	26.200	21.833	4511246067.0	4264.661	0.315	8696.7	0.667098	0.003364	0.663734
896.778	343.944	39.133	18.267	5788341815.0	4233.438	0.246	8683.4	0.663228	0.002622	0.660606
511.889	190.833	26.200	14.700	709126184.6	1315.385	2.006	2630.6	0.680233	0.021403	0.658830
896.778	343.944	32.667	21.833	5242487358.0	4273.389	0.271	8820.2	0.659103	0.002895	0.656208

Depth mm	Flange Width mm	Flange Thickness mm	Web Thickness mm	Izp mm4	Plastic Moment KNM	Deflection mm due 3 Loads at once	Cost Euro	Normalized Moment Cost	Normalized Deflection	Score
511.889	190.833	13.267	14.700	454858510.3	1320.294	3.127	2629.7	0.683005	0.033367	0.649638
896.778	343.944	26.200	18.267	4332312565.0	4063.133	0.328	8466.2	0.652878	0.003503	0.649375
896.778	343.944	32.667	18.267	5071650732.0	4065.809	0.280	8597.2	0.643353	0.002993	0.640361
511.889	267.389	13.267	11.133	547161268.8	1352.878	2.599	2849.0	0.645990	0.027739	0.618251
415.667	229.111	13.267	11.133	300847900.1	995.412	4.727	2068.8	0.654552	0.050449	0.604103
511.889	229.111	13.267	11.133	484018406.1	1289.496	2.938	2778.5	0.631348	0.031357	0.599991
608.111	305.667	19.733	14.700	1269711668.0	1769.616	1.120	4006.7	0.600830	0.011953	0.588877
511.889	267.389	26.200	11.133	917094108.2	1251.610	1.551	2854.9	0.596400	0.016549	0.579851
608.111	229.111	19.733	14.700	1008120848.0	1725.983	1.411	3951.4	0.594217	0.015055	0.579162
608.111	267.389	19.733	14.700	1138916258.0	1733.416	1.249	4016.0	0.587177	0.013326	0.573850
608.111	305.667	26.200	14.700	1567056381.0	1696.291	0.908	4009.2	0.575575	0.009685	0.565890
608.111	343.944	19.733	14.700	1400507077.0	1657.761	1.016	3987.5	0.565563	0.010837	0.554726
608.111	267.389	26.200	14.700	1397144137.0	1669.937	1.018	4018.5	0.565322	0.010863	0.554458
608.111	229.111	26.200	14.700	1227231894.0	1644.426	1.159	3953.5	0.565838	0.012367	0.553471
511.889	267.389	19.733	11.133	737190182.9	1184.976	1.929	2833.3	0.568953	0.020588	0.548365
511.889	229.111	26.200	11.133	798693099.1	1157.353	1.781	2777.1	0.566936	0.019003	0.547933
993.000	420.500	45.600	25.400	10164303602.0	5218.776	0.140	13170.4	0.539050	0.001493	0.537557
993.000	382.222	45.600	25.400	9380362263.0	5218.776	0.152	13170.4	0.539050	0.001618	0.537432
415.667	229.111	13.267	7.567	283334253.6	807.603	5.020	1864.8	0.589149	0.053567	0.535582
608.111	343.944	26.200	14.700	1736968625.0	1573.399	0.819	3990.0	0.536445	0.008738	0.527707
415.667	229.111	19.733	11.133	404063491.7	878.341	3.520	2116.0	0.564686	0.037562	0.527124
800.556	267.389	19.733	14.700	2148891899.0	2328.936	0.662	5977.2	0.530053	0.007063	0.522990
608.111	305.667	32.667	14.700	1850877777.0	1562.040	0.768	4009.2	0.530022	0.008200	0.521822
608.111	305.667	19.733	18.267	1324363368.0	1566.623	1.074	4007.4	0.531816	0.011460	0.520356
608.111	229.111	32.667	14.700	1436377709.0	1541.316	0.990	3953.5	0.530358	0.010566	0.519792
511.889	229.111	19.733	11.133	645662044.7	1094.285	2.203	2753.2	0.540695	0.023507	0.517188

Depth mm	Flange Width mm	Flange Thickness mm	Web Thickness mm	Izp mm4	Plastic Moment KNM	Deflection mm due 3 Loads at once	Cost Euro	Normalized Moment Cost	Normalized Deflection	Score
993.000	420.500	39.133	25.400	9110422543.0	5019.089	0.156	13161.6	0.518771	0.001666	0.517105
993.000	420.500	45.600	21.833	9946325951.0	5017.904	0.143	13164.0	0.518554	0.001526	0.517028
993.000	382.222	39.133	25.400	8428582507.0	5019.089	0.169	13161.6	0.518771	0.001801	0.516970
993.000	382.222	45.600	21.833	9162384611.0	5017.904	0.155	13164.0	0.518554	0.001656	0.516898
608.111	267.389	32.667	14.700	1643627743.0	1540.890	0.865	4018.5	0.521636	0.009234	0.512402
993.000	420.500	52.067	25.400	11188385215.0	4918.342	0.127	13180.4	0.507633	0.001357	0.506276
993.000	382.222	52.067	25.400	10305229579.0	4918.342	0.138	13180.4	0.507633	0.001473	0.506160
608.111	229.111	19.733	18.267	1062772549.0	1505.856	1.338	3952.1	0.518340	0.014281	0.504059
608.111	267.389	19.733	18.267	1193567958.0	1525.377	1.192	4016.7	0.516615	0.012716	0.503899
993.000	420.500	45.600	18.267	9728348299.0	4809.068	0.146	13064.6	0.500754	0.001560	0.499194
993.000	382.222	45.600	18.267	8944406960.0	4809.068	0.159	13064.6	0.500754	0.001697	0.499057
415.667	229.111	19.733	7.567	388238712.8	757.337	3.663	1915.7	0.537801	0.039093	0.498708
800.556	305.667	19.733	14.700	2379202513.0	2347.023	0.598	6322.5	0.504996	0.006379	0.498617
704.333	229.111	19.733	14.700	1419798548.0	1852.564	1.002	4965.7	0.507519	0.010690	0.496829
993.000	420.500	39.133	21.833	8882931243.0	4818.217	0.160	13155.2	0.498251	0.001709	0.496543
993.000	382.222	39.133	21.833	8201091207.0	4818.217	0.173	13155.2	0.498251	0.001851	0.496401
704.333	267.389	19.733	14.700	1596854783.0	1891.203	0.891	5101.7	0.504293	0.009505	0.494788
511.889	190.833	26.200	11.133	680292090.0	988.088	2.091	2607.9	0.515423	0.022310	0.493113
993.000	420.500	45.600	28.967	10382281254.0	4772.563	0.137	13153.4	0.493598	0.001462	0.492136
993.000	382.222	45.600	28.967	9598339915.0	4772.563	0.148	13153.4	0.493598	0.001581	0.492017
993.000	420.500	52.067	21.833	10979652204.0	4767.510	0.130	13174.0	0.492304	0.001382	0.490922
993.000	382.222	52.067	21.833	10096496569.0	4767.510	0.141	13174.0	0.492304	0.001503	0.490801
704.333	305.667	19.733	14.700	1773911018.0	1926.889	0.802	5259.7	0.498374	0.008556	0.489818
511.889	190.833	13.267	11.133	420875543.4	1003.073	3.379	2605.0	0.523823	0.036062	0.487761

10.3 Appendix C: 96 I-Beams Dataset

Section Size			Dimensions					Properties				Max Bending Moment (KN.m)	Plastic Moment (KN.m)	Cost (£) / meters	Beam Length (m)	Deflection (mm) when				Total Cost (£)	Calculated Izp (mm4)	Score = (Plastic Moment/Cost) - Total Deflection
	W (kg/m)	W (KN/m)	d (mm)	b_t (mm)	t_t (mm)	t_w (mm)	r (mm)	A (mm^2)	Izp (mm^4)	Plastic Mod Zp (mm3)	PlasticSection modulus (mm3)					5 KN Load at 2 m	8 KN Load at 6 m	12 KN Load at 8 m	3 Loads at once			
127x76 UB 13	13	0.12745	127.00	76.00	7.60	4.00	7.60	1650	4730000	84200	81464.6800	1.5931	28.5126	35.75	10	42.9545	154.6361	103.0907	300.6812	357.5	4588602.141	-0.92544
152x89 UB 16	16	0.15686	152.40	88.70	7.70	4.50	7.60	2030	8340000	123000	119943.7780	1.9608	41.9803	44.00	10	24.3615	87.7013	58.4675	170.5302	440	8121259.499	-0.47795
178x102 UB 19	19	0.18627	177.80	101.20	7.90	4.80	7.60	2430	13560000	171000	167324.4520	2.3284	58.5636	52.50	10	14.9834	53.9402	35.9601	104.8836	525	13247825.962	-0.24454
203x102 UB 23	23	0.22549	203.20	101.80	9.30	5.40	7.60	2940	21050000	234000	229577.0520	2.8186	80.3520	63.50	10	9.6520	34.7472	23.1648	67.5640	635	20641828.236	-0.10641
254x102 UB 22	22	0.21569	254.00	101.60	6.80	5.70	7.60	2800	28410000	259000	253139.3640	2.6961	88.5988	66.00	10	7.1515	25.7455	17.1636	50.0606	660	27713703.382	-0.04100
203x133 UB 25	25	0.24510	203.20	133.20	7.80	5.70	7.60	3200	23400000	258000	253163.8920	3.0637	88.6074	69.00	10	8.6827	31.2576	20.8384	60.7787	690	22981000.005	-0.08209
305x102 UB 25	25	0.24510	305.10	101.60	7.00	5.80	7.60	3160	44550000	342000	334880.5745	3.0637	117.2082	75.00	10	4.5606	16.4181	10.9454	31.9242	750	43528373.464	0.03992
254x102 UB 25	25	0.24510	257.20	101.90	8.40	6.00	7.60	3200	34150000	306000	299651.0880	3.0637	104.8779	75.00	10	5.9495	21.4181	14.2787	41.6463	750	33449262.013	-0.00778
203x133 UB 30	30	0.29412	206.80	133.90	9.60	6.40	7.60	3820	28960000	314000	309798.7840	3.6765	108.4296	82.50	10	7.0157	25.2565	16.8377	49.1099	825	28534989.884	-0.04046
305x102 UB 28	28	0.27451	308.70	101.80	8.80	6.00	7.60	3590	53660000	403000	395771.2310	3.4314	138.5199	83.00	10	3.7863	13.6308	9.0872	26.5043	830	52631283.603	0.06787
254x146 UB 31	31	0.30392	251.40	146.10	8.60	6.00	7.60	3970	44130000	393000	387342.9480	3.7990	135.5700	85.00	10	4.6040	16.5744	11.0496	32.2280	850	43473695.251	0.04192
254x102 UB 28	28	0.27451	260.40	102.20	10.00	6.30	7.60	3610	40050000	353000	346931.4520	3.4314	121.4260	85.00	10	5.0730	18.2629	12.1753	35.5112	850	39350763.607	0.01544
356x127 UB 33	33	0.32353	349.00	125.40	8.50	6.00	10.20	4210	82490000	543000	528274.9500	4.0441	184.8962	95.50	10	2.4630	8.8669	5.9113	17.2411	955	80100375.450	0.12365
305x102 UB 33	33	0.32353	312.70	102.40	10.80	6.60	7.60	4180	65010000	481000	473696.9445	4.0441	165.7939	99.00	10	3.1253	11.2510	7.5007	21.8770	990	63987440.887	0.08380
305x165 UB 40	40	0.39216	303.40	165.00	10.20	6.00	8.90	5130	85030000	623000	613589.1000	4.9020	214.7562	100.00	10	2.3894	8.6020	5.7347	16.7261	1000	83702367.680	0.14513
254x146 UB 37	37	0.36275	256.00	146.40	10.90	6.30	7.60	4720	55370000	483000	477508.9590	4.5343	167.1281	101.00	10	3.6694	13.2098	8.8066	25.6858	1010	54707487.293	0.06926
356x127 UB 39	39	0.38235	353.40	126.00	10.70	6.60	10.20	4980	101700000	659000	643897.7400	4.7794	225.3642	107.50	10	1.9978	7.1920	4.7947	13.9845	1075	99321150.092	0.14947
254x146 UB 43	43	0.42157	259.60	147.30	12.70	7.20	7.60	5480	65440000	566000	560607.6510	5.2696	196.2127	107.50	10	3.1047	11.1771	7.4514	21.7332	1075	64776635.227	0.09835
305x127 UB 37	37	0.36275	304.40	123.40	10.70	7.10	8.90	4720	71710000	539000	529953.5810	4.5343	185.4838	110.00	10	2.8333	10.1998	6.7999	19.8330	1100	70383215.434	0.09167
305x127 UB 42	42	0.41176	307.20	124.30	12.10	8.00	8.90	5340	81960000	614000	604017.2530	5.1471	211.4060	123.00	10	2.4789	8.9242	5.9495	17.3526	1230	80635307.286	0.10296
356x171 UB 45	45	0.44118	351.40	171.10	9.70	7.00	10.20	5730	120700000	775000	760001.2390	5.5147	266.0004	125.00	10	1.6833	6.0599	4.0399	11.7831	1250	118263354.542	0.15974
305x165 UB 46	46	0.45098	306.60	165.70	11.80	6.70	8.90	5870	98990000	720000	710559.7230	5.6373	248.6959	126.50	10	2.0525	7.3889	4.9259	14.3673	1265	97663033.991	0.13600
305x127 UB 48	48	0.47059	311.00	125.30	14.00	9.00	8.90	6120	95750000	711000	701197.6500	5.8824	245.4192	140.00	10	2.1219	7.6389	5.0926	14.8535	1400	94424308.017	0.11448
356x171 UB 51	51	0.50000	355.00	171.50	11.50	7.40	10.20	6490	141400000	896000	881382.2750	6.2500	308.4838	142.00	10	1.4369	5.1728	3.4485	10.0581	1420	138965106.142	0.16963
457x152 UB 52	52	0.50980	449.80	152.40	10.90	7.60	10.20	6660	213700000	1096000	1077132.7240	6.3725	376.9965	143.00	10	0.9507	3.4227	2.2818	6.6552	1430	209685261.565	0.22432
406x140 UB 39	39	0.38235	398.00	141.80	8.60	6.40	10.20	4970	125100000	724000	706879.3360	4.7794	247.4078	147.00	10	1.6241	5.8468	3.8978	11.3687	1470	121921635.370	0.11953
305x165 UB 54	54	0.52941	310.40	166.90	13.70	7.90	8.90	6880	117000000	846000	836589.2260	6.6176	292.8062	148.50	10	1.7365	6.2515	4.1677	12.1557	1485	115635410.033	0.14390
356x171 UB 57	57	0.55882	358.00	172.20	13.00	8.10	10.20	7260	160400000	1010000	995520.6000	6.9853	348.4322	158.00	10	1.2667	4.5600	3.0400	8.8667	1580	157988934.800	0.17667
406x178 UB 54	54	0.52941	402.60	177.70	10.90	7.70	10.20	6900	187200000	1055000	1037837.1130	6.6176	363.2430	159.00	10	1.0853	3.9072	2.6048	7.5973	1590	184061242.218	0.18830
406x140 UB 53	53	0.51961	406.60	143.30	12.90	7.90	10.20	6790	182830000	1031000	1014174.0730	6.4951	354.9609	160.00	10	1.1113	4.0006	2.6671	7.7789	1600	179667857.884	0.18152
457x152 UB 60	60	0.58824	454.60	152.90	13.30	8.10	10.20	7620	255000000	1287000	1268362.0410	7.3529	443.9267	165.00	10	0.7968	2.8683	1.9122	5.5773	1650	250996307.040	0.23296
406x178 UB 60	60	0.58824	406.40	177.90	12.80	7.90	10.20	7650	216000000	1199000	1182666.4960	7.3529	413.9333	170.00	10	0.9406	3.3862	2.2575	6.5844	1700	212801688.098	0.20572
406x140 UB 46	46	0.45098	403.20	142.20	11.20	6.80	10.20	5860	156900000	888000	870829.5680	5.6373	304.7903	177.00	10	1.2949	4.6617	3.1078	9.0645	1770	153689944.337	0.13083
457x191 UB 67	67	0.65686	453.40	189.90	12.70	8.50	10.20	8550	293800000	1471000	1452115.4110	8.2108	508.2404	185.00	10	0.6915	2.4895	1.6597	4.8408	1850	289798981.702	0.24072
457x152 UB 67	67	0.65686	458.00	153.80	15.00	9.00	10.20	8560	289300000	1453000	1434165.0000	8.2108	501.9578	185.00	10	0.7023	2.5283	1.6855	4.9161	1850	285261798.000	0.23730
406x178 UB 67	67	0.65686	409.40	178.80	14.30	8.80	10.20	8550	243300000	1346000	1329226.4920	8.2108	465.2293	185.00	10	0.8351	3.0063	2.0042	5.8455	1850	240147775.915	0.21565
356x171 UB 67	67	0.65686	363.40	173.20	15.70	9.10	10.20	8550	194600000	1211000	1196239.3480	8.2108	418.6838	200.00	10	1.0441	3.7586	2.5058	7.3084	2000	192234094.168	0.17139
457x191 UB 74	74	0.72549	457.00	190.40	14.50	9.00	10.20	9460	333200000	1653000	1633818.0000	9.0686	571.8363	203.00	10	0.6098	2.1952	1.4634	4.2684	2030	329189754.533	0.24914
457x152 UB 74	74	0.72549	462.00	154.40	17.00	9.60	10.20	9450	326700000	1627000	1607677.6000	9.0686	562.6872	203.00	10	0.6219	2.2388	1.4926	4.3533	2030	322736639.467	0.24464
406x178 UB 74	74	0.72549	412.80	179.50	16.00	9.50	10.20	9450	273100000	1501000	1484005.1200	9.0686	519.4018	204.00	10	0.7440	2.6782	1.7855	5.2077	2040	269936354.645	0.22070
457x152 UB 82	82	0.80392	465.80	155.30	18.90	10.50	10.20	10500	365900000	1811000	1792585.4730	10.0490	627.4049	205.00	10	0.5553	1.9990	1.3327	3.8869	2050	361882657.188	0.27318
457x191 UB 161	161	1.57843	492.00	199.40	32.00	18.00	10.20	20600	797790000	3778000	3759496.0000	19.7304	1315.8236	243.00	10	0.2547	0.9168	0.6112	1.7827	2430	793781757.867	0.50027
457x191 UB 89	89	0.87255	463.40	191.90	17.70	10.50	10.20	11400	410200000	2014000	1994735.9910	10.9069	698.1576	245.00	10	0.4953	1.7831	1.1887	3.4671	2450	406147473.330	0.25486
457x191 UB 82	82	0.80392	460.00	191.30	16.00	9.90	10.20	10400	370500000	1831000	1812375.6000	10.0490	634.3315	245.00	10	0.5484	1.9742	1.3161	3.8387	2450	366509798.933	0.22927
533x210 UB 82	82	0.80392	528.30	208.80	13.20	9.60	12.70	10500	475400000	2059000	2024266.6800	10.0490	708.4933	246.00	10	0.4274	1.5386	1.0257	2.9916	2460	466867600.394	0.25929
533x210 UB 138	138	1.35294	549.10	213.90	23.60	14.70	12.70	17600	860880000	3613000	3578490.7868	16.9118	1252.4718	252.00	10	0.2360	0.8496	0.5664	1.6521	2520	852354613.509	0.45913
457x191 UB 106	106	1.03922	469.20	194.00	20.60	12.60	10.20	13500	488730000	2389000	2369814.6400	12.9902	829.4351	254.00	10	0.4157	1.4966	0.9977	2.9100	2540	484727226.123	0.29559
457x191 UB 133	133	1.30392	480.60	196.70	26.30	15.30	10.20	17000	638410000	3070000	3050868.1030	16.2990	1067.8038	266.00	10	0.3183	1.1457	0.7638	2.2278	2660	634405385.247	0.36786
533x312 UB 182	182	1.78431	550.70	314.50	24.40	15.20	12.70	23100	1232220000	5030000	4995954.6580	22.3039	1748.5841	267.00	10	0.1649	0.5936	0.3957	1.1542	2670	1223696061.977	0.60838
533x312 UB 150	150	1.47059	542.50	312.00	20.30	12.70	12.70	19200	1006330000	4142000	4107199.8818	18.3824	1437.5200	269.00	10	0.2019	0.7268	0.4846	1.4133	2690	997804217.717	0.49487
533x165 UB 74	74	0.72549	529.10	165.90	13.60	9.70	12.70	9520	410580000	1808000	1773957.9743	9.0686	620.8853	275.00	10	0.4948	1.7815	1.1876	3.4639	2750	402054367.524	0.19954
533x210 UB 92	92	0.90196	533.10	209.30	15.60	10.10	12.70	11700	552300032	2360000	2325735.5153	11.2745	814.0074	276.00	10	0.3679	1.3243	0.8829	2.5751	2760	543749118.751	0.26714
457x191 UB 98	98	0.96078	467.20	192.80	19.60	11.40	10.20	12500	457300000	2232000	2213501.0880	12.0098	774.7254	280.00	10	0.4443	1.5995	1.0663	3.1100	2800	453265856.265	0.24831
610x229 UB 101	101	0.99020	602.60	227.60	14.80	10.50	12.70	12900	757800000	2881000	2841856.1690	12.3775	994.6497	303.00	10	0.2681	0.9652	0.6435	1.8768	3030	746658733.033	0.30063
533x210 UB 101	101	0.99020	536.70	210.00	17.40	10.80	12.70	12900	615200000	2612000	2577661.9470	12.3775	902.1817	303.00	10	0.3303	1.1889	0.7926	2.3118	3030	606663399.743	0.27066
533x165 UB 85	85	0.83333	534.90	166.50	16.50	10.30	12.70	10800	486310000	2107000	2072826.1958	10.4167	725.4892	315.00	10	0.4178	1.5040	1.0027	2.9245	3150	477790106.378	0.20558

Section Size			Dimensions					Properties				Max Bending Moment (KN.m)	Plastic Moment (KN.m)	Cost (£) / meters	Beam Length (m)	Deflection (mm) when				Total Cost (£)	Calculated Izp (mm4)	Score = (Plastic Moment/Cost) - Total Deflection
	W (kg/m)	W (KN/m)	d (mm)	b f (mm)	t f (mm)	t w (mm)	r (mm)	A (mm2)	Izp (mm4)	Plastic Mod Zp (mm3)	PlasticSection modulus (mm3)					5 KN Load at 2 m	8 KN Load at 6 m	12 KN Load at 8 m	3 Loads at once			
533x312 UB 272	272	2.66667	577.10	320.20	37.60	21.10	12.70	34800	1985780000	7859000	7824112.5828	33.3333	2738.4394	319.00	10	0.1023	0.3683	0.2456	0.7162	3190	1977256507.608	0.80012
533x210 UB 109	109	1.06863	539.50	210.80	18.80	11.60	12.70	13900	668200000	2828000	2794075.3970	13.3578	977.9264	327.00	10	0.3041	1.0946	0.7298	2.1284	3270	659696049.445	0.27249
533x312 UB 219	219	2.14706	560.30	317.40	29.20	18.30	12.70	27900	1509760000	6109000	6074736.3038	26.8382	2126.1577	333.00	10	0.1346	0.4845	0.3230	0.9420	3330	1501234183.119	0.59374
610x229 UB 113	113	1.10784	607.60	228.20	17.30	11.10	12.70	14400	873200000	3281000	3241534.7330	13.8480	1134.5372	340.00	10	0.2327	0.8376	0.5584	1.6287	3400	862043485.935	0.30652
533x165 UB 66	66	0.64706	524.70	165.10	11.40	8.90	12.70	8370	350280000	1561000	1526587.9943	8.0882	534.3058	340.00	10	0.5800	2.0881	1.3921	4.0602	3400	341760193.570	0.13340
533x210 UB 122	122	1.19608	544.50	211.90	21.30	12.70	12.70	15500	760400000	3196000	3161241.4658	14.9510	1106.4345	366.00	10	0.2672	0.9619	0.6413	1.8704	3660	751901482.881	0.27638
686x254 UB 125	125	1.22549	677.90	253.00	16.20	11.70	15.20	15900	1180000000	3994000	3930804.1013	15.3186	1375.7814	375.00	10	0.1722	0.6199	0.4132	1.2053	3750	1159695534.657	0.33896
610x229 UB 125	125	1.22549	612.20	229.00	19.60	11.90	12.70	15900	986099968	3676000	3636604.6150	15.3186	1272.8116	375.00	10	0.2060	0.7417	0.4945	1.4423	3750	974958519.708	0.31250
610x178 UB 100	100	0.98039	607.40	179.20	17.20	11.30	12.70	12800	725280000	2786000	2746667.4730	12.2549	961.3336	387.00	10	0.2801	1.0085	0.6723	1.9609	3870	714137733.120	0.22570
610x178 UB 82	82	0.80392	598.60	177.90	12.80	10.00	12.70	10400	558690000	2194000	2154759.3960	10.0490	754.1658	390.00	10	0.3637	1.3092	0.8728	2.5456	3900	547549394.895	0.17231
686x254 UB 140	140	1.37255	683.50	253.70	19.00	12.40	15.20	17800	1363000064	4558000	4494767.1250	17.1569	1573.1685	420.00	10	0.1491	0.5366	0.3578	1.0434	4200	1342442459.175	0.34668
610x229 UB 140	140	1.37255	617.20	230.20	22.10	13.10	12.70	17800	1118000000	4142000	4102801.1170	17.1569	1435.9804	420.00	10	0.1817	0.6542	0.4362	1.2721	4200	1106631781.869	0.31539
762x267 UB 147	147	1.44118	754.00	265.20	17.50	12.80	16.50	18700	1684999936	5156000	5072371.7000	18.0147	1775.3301	441.00	10	0.1206	0.4341	0.2894	0.8440	4410	1655425543.433	0.37353
610x305 UB 149	149	1.46078	612.40	304.80	19.70	11.80	16.50	19000	1259000064	4594000	4527473.2620	18.2598	1584.6156	447.00	10	0.1614	0.5810	0.3873	1.1296	4470	1240066178.700	0.32764
686x254 UB 152	152	1.49020	687.50	254.50	21.00	13.20	15.20	19400	1504000000	5000000	4937121.0750	18.6275	1727.9924	456.00	10	0.1351	0.4863	0.3242	0.9456	4560	1483322439.325	0.35110
610x178 UB 92	92	0.90196	603.00	178.80	15.00	10.90	12.70	11700	645770000	2511000	2471712.5250	11.2745	865.0994	479.00	10	0.3146	1.1326	0.7551	2.2024	4790	634630315.275	0.16151
686x254 UB 170	170	1.66667	692.90	255.80	23.70	14.50	15.20	21700	1703000064	5631000	5567427.8883	20.8333	1948.5998	510.00	10	0.1193	0.4295	0.2863	0.8351	5100	1683033259.990	0.35440
762x267 UB 173	173	1.69608	762.20	266.70	21.60	14.30	16.50	22000	2052999936	6198000	6114524.8070	21.2010	2140.0837	520.00	10	0.0990	0.3563	0.2375	0.6928	5200	2023228379.005	0.38243
838x292 UB 176	176	1.72549	834.90	291.70	18.80	14.00	17.80	22400	2460000000	6808000	6700365.2710	21.5686	2345.1278	528.00	10	0.0826	0.2973	0.1982	0.5781	5280	2417840117.609	0.41328
610x305 UB 179	179	1.75490	620.20	307.10	23.60	14.10	16.50	22800	1530000000	5547000	5481254.0210	21.9363	1918.4389	537.00	10	0.1328	0.4781	0.3187	0.9296	5370	1511546142.808	0.33088
762x267 UB 134	134	1.31373	750.00	264.40	15.50	12.00	16.50	17100	1507000064	4644000	4561010.9000	16.4216	1596.3538	543.00	10	0.1348	0.4854	0.3236	0.9437	5430	1477328529.033	0.27169
838x292 UB 194	194	1.90196	840.70	292.40	21.70	14.70	17.80	24700	2792000000	7640000	7532771.3108	23.7745	2636.4700	582.00	10	0.0728	0.2620	0.1746	0.5094	5820	2749385083.882	0.42178
762x267 UB 197	197	1.93137	769.80	268.00	25.40	15.60	16.50	25100	2400000000	7167000	7083427.5800	24.1422	2479.1997	591.00	10	0.0847	0.3048	0.2032	0.5926	5910	2369976899.121	0.39018
838x292 UB 226	226	2.21569	850.90	293.80	26.80	16.10	17.80	28900	3396999936	9155000	9047472.8863	27.6961	3166.6155	678.00	10	0.0598	0.2153	0.1435	0.4187	6780	3354667169.571	0.43522
610x305 UB 238	238	2.33333	635.80	311.40	31.40	18.40	16.50	30300	2095000064	7486000	7420112.4240	29.1667	2597.0393	714.00	10	0.0970	0.3491	0.2328	0.6789	7140	2076017904.026	0.33777
914x305 UB 253	253	2.48039	918.40	305.50	27.90	17.30	19.10	32300	4362999808	10940000	10808272.8620	31.0049	3782.8955	760.00	10	0.0466	0.1676	0.1118	0.3260	7600	4305934867.459	0.46423
914x305 UB 521	521	5.10784	981.00	319.00	58.90	33.00	19.00	66400	9820000000	23600000	23472621.5900	63.8480	8215.4176	768.00	10	0.0207	0.0745	0.0497	0.1448	7680	9767505163.013	0.99952
914x305 UB 313	313	3.06863	932.00	309.00	34.50	21.10	19.00	39800	5480000000	13600000	13496455.2250	38.3578	4723.7593	796.00	10	0.0371	0.1335	0.0890	0.2595	7960	5425807978.392	0.55390
914x305 UB 224	224	2.19608	910.40	304.10	23.90	15.90	19.10	28600	3764000000	9535000	9400786.2060	27.4510	3290.2752	813.00	10	0.0540	0.1943	0.1295	0.3778	8130	3707025190.198	0.37708
914x305 UB 474	474	4.64706	971.00	316.00	54.10	30.00	19.00	60400	8860000000	21400000	21258134.4400	58.0882	7440.3471	841.00	10	0.0229	0.0826	0.0550	0.1605	8410	8800244898.211	0.82651
914x305 UB 345	345	3.38235	943.00	308.00	39.90	22.10	19.00	44000	6260000000	15400000	15215132.6960	42.2794	5325.2964	843.00	10	0.0325	0.1168	0.0779	0.2272	8430	6199260649.529	0.58978
914x305 UB 289	289	2.83333	926.60	307.70	32.00	19.50	19.10	36800	5041999872	12570000	12435973.3950	35.4167	4352.5907	867.00	10	0.0403	0.1451	0.0967	0.2821	8670	4984756308.640	0.46837
914x305 UB 425	425	4.16667	961.00	313.00	49.00	26.90	19.00	54200	7880000000	19100000	18995915.5250	52.0833	6648.5704	873.00	10	0.0258	0.0928	0.0619	0.1805	8730	7825165295.525	0.71134
914x305 UB 576	576	5.64706	993.00	322.00	65.00	36.10	19.00	73300	11020000000	26300000	26144580.2250	70.5882	9150.6031	889.00	10	0.0184	0.0664	0.0442	0.1291	8890	10960591839.725	0.96180
914x305 UB 381	381	3.73529	951.00	310.00	43.90	24.40	19.00	48600	6970000000	17000000	16889920.7640	46.6912	5911.4723	891.00	10	0.0291	0.1049	0.0700	0.2040	8910	6911125402.662	0.61955
914x305 UB 271	271	2.65686	923.00	307.00	30.00	18.40	19.00	34600	4720000000	11800000	11650467.4000	33.2108	4077.6636	891.00	10	0.0430	0.1550	0.1033	0.3013	8910	4659162137.067	0.42682
914x305 UB 238	238	2.33333	915.00	305.00	25.90	16.50	19.00	30400	4060000000	10200000	10097041.6900	29.1667	3533.9646	969.00	10	0.0500	0.1802	0.1201	0.3503	9690	4007531943.519	0.33977
914x419 UB 343	343	3.36275	911.80	418.50	32.00	19.40	24.10	43700	6257999872	15480000	15268291.0740	42.0343	5343.9019	1030.00	10	0.0325	0.1169	0.0779	0.2273	10300	6170457521.192	0.48426
914x419 UB 388	388	3.80392	921.00	420.50	36.60	21.40	24.10	49400	7196000256	17670000	17456573.2140	47.5490	6109.8006	1165.00	10	0.0282	0.1016	0.0678	0.1976	11650	7109008167.326	0.48961
914x305 UB 201	201	1.97059	903.00	303.30	20.20	15.10	19.10	25600	3252999936	8351000	8217512.7670	24.6324	2876.1295	1505.00	10	0.0625	0.2248	0.1499	0.4372	15050	3195431121.595	0.17720

Random Forest Model in Python

```
library(mlbench)
data(BostonHousing)
```

Import library

In [2]:
```
import pandas as pd
```

In [3]:
```
import os
import glob

# Define the pattern to match all versions of the file
file_pattern = "UK Universal Steel Database*.csv"

# Get list of matching files
files_to_delete = glob.glob(file_pattern)

# Delete each file
for file_path in files_to_delete:
    try:
        os.remove(file_path)
        print(f"Deleted: {file_path}")
    except Exception as e:
        print(f"Failed to delete {file_path}: {e}")
```

Read in CSV file

In [4]:
```
from google.colab import files

uploaded = files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving UK Universal Steel Database.csv to UK Universal Steel Database.csv
```

In [5]:
```
UKUniversalSteelDatabase = pd.read_csv("UK Universal Steel Database.csv")
UKUniversalSteelDatabase
```

Out[5]:

	MassKgPerm	Depthmm	FlangeWidthmm	FlangeThicknessmm	WebThick
0	13	127.0	76.0	7.6	
1	16	152.4	88.7	7.7	
2	19	177.8	101.2	7.9	
3	23	203.2	101.8	9.3	
4	25	203.2	133.2	7.8	
...	...	...	...	...	
91	474	971.0	316.0	54.1	
92	521	981.0	319.0	58.9	
93	576	993.0	322.0	65.0	
94	343	911.8	418.5	32.0	
95	388	921.0	420.5	36.6	

96 rows × 12 columns

Split dataset to X and Y variables

In [6]:

```
Y = UKUniversalSteelDatabase.PlasticMomentKNM
Y
```

Out[6]:

	PlasticMomentKNM
0	28.513
1	41.980
2	58.564
3	80.352
4	88.599
...	...
91	4077.664
92	3533.965
93	5343.902
94	6109.801
95	2876.129

96 rows × 1 columns

dtype: float64

In [7]:
```
X = UKUniversalSteelDatabase.drop(['PlasticMomentKNM'], axis=1)
X
```

Out[7]:

	MassKgPerm	Depthmm	FlangeWidthmm	FlangeThicknessmm	WebThick
0	13	127.0	76.0	7.6	
1	16	152.4	88.7	7.7	
2	19	177.8	101.2	7.9	
3	23	203.2	101.8	9.3	
4	25	203.2	133.2	7.8	
...	...	...	...	...	
91	474	971.0	316.0	54.1	
92	521	981.0	319.0	58.9	
93	576	993.0	322.0	65.0	
94	343	911.8	418.5	32.0	
95	388	921.0	420.5	36.6	

96 rows × 11 columns

Data split

Import library

In [8]:
```
from sklearn.model_selection import train_test_split
```

Perform 80/20 Data split

In [9]:
```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2)
```

In [10]:
```
if len(X_train.shape) == 1:
    X_train = X_train.to_numpy().reshape(-1, 1)
    X_test = X_test.to_numpy().reshape(-1, 1)
```

In [11]:
```
print("X_train shape:", X_train.shape)
print("Y_train shape:", Y_train.shape)
print("X_test shape:", X_test.shape)
print("Y_test shape:", Y_test.shape)
```

```
X_train shape: (76, 11)
Y_train shape: (76,)
X_test shape: (20, 11)
Y_test shape: (20,)
```

Data dimension

In [12]:
```
X_train.shape, Y_train.shape
```

Out[12]:
```
((76, 11), (76,))
```

In [13]:
```
X_test.shape, Y_test.shape
```

Out[13]:
```
((20, 11), (20,))
```

Random Forest Model

Import library

In [14]:
```
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error, r2_score
```

Build Random Forest

Defines the regression model

In [15]:
```
model = RandomForestRegressor()
```

Build training model

In [16]:
```
model.fit(X_train, Y_train)
```

Out[16]:
```
▾ RandomForestRegressor
RandomForestRegressor()
```

Apply trained model to make prediction (on test set)

In [17]:
```
Y_pred = model.predict(X_test)
```

Prediction results

Print model performance

In [18]:
```
print('Mean squared error (MSE): %.2f' % mean_squared_error(Y_test, Y_pred))
print('Coefficient of determination (R^2): %.2f' % r2_score(Y_test, Y_pred))
```

```
Mean squared error (MSE): 26441.27
Coefficient of determination (R^2): 0.99
```

In [19]:
```
UKUniversalSteelDatabase.head()
```

Out[19]:

	MassKgPerm	Depthmm	FlangeWidthmm	FlangeThicknessmm	WebThickn
0	13	127.0	76.0	7.6	
1	16	152.4	88.7	7.7	
2	19	177.8	101.2	7.9	
3	23	203.2	101.8	9.3	
4	25	203.2	133.2	7.8	

In [20]:
```
df = pd.DataFrame(columns=[
    "MassKgPerm",
    "Depthmm",
    "FlangeWidthmm",
    "FlangeThicknessmm",
    "WebThicknessmm",
    "Izpmm4",
    "CostEuro",
    "PlasticMomentKNM",
    "Deflectionmmdue5KNLoadat2m",
    "Deflectionmmdue8KNLoadat6m",
    "Deflectionmmdue12KNLoadat8m",
    "Deflectionmmdue3Loadsatonce",
], dtype=float)
```

In [21]:
```
df[["MassKgPerm", "Depthmm", "FlangeWidthmm", "FlangeThicknessmm", "WebThick
```

In [22]:
```
df.columns
```

Out[22]:
```
Index(['MassKgPerm', 'Depthmm', 'FlangeWidthmm', 'FlangeThicknessmm',
       'WebThicknessmm', 'Izpmm4', 'CostEuro', 'PlasticMomentKNM',
       'Deflectionmmdue5KNLoadat2m', 'Deflectionmmdue8KNLoadat6m',
       'Deflectionmmdue12KNLoadat8m', 'Deflectionmmdue3Loadsatonce'],
      dtype='object')
```

In [23]:
```
df.head()
```

Out[23]:

MassKgPerm	Depthmm	FlangeWidthmm	FlangeThicknessmm	WebThickne

In [24]:
```
float_columns = [
    "MassKgPerm",
    "Depthmm",
    "FlangeWidthmm",
    "FlangeThicknessmm",
    "WebThicknessmm",
    "Izpmm4",
    "CostEuro",
    "PlasticMomentKNM",
    "Deflectionmmdue5KNLoadat2m",
    "Deflectionmmdue8KNLoadat6m",
```

```
    "Deflectionmmdue12KNLoadat8m",
    "Deflectionmmdue3Loadsatonce",
]
```

In [25]:
```
df[float_columns] = df[float_columns].astype(float)
```

In [26]:
```
print(df)
```

```
Empty DataFrame
Columns: [MassKgPerm, Depthmm, FlangeWidthmm, FlangeThicknessmm, WebThicknes
smm, Izpmm4, CostEuro, PlasticMomentKNM, Deflectionmmdue5KNLoadat2m, Deflect
ionmmdue8KNLoadat6m, Deflectionmmdue12KNLoadat8m, Deflectionmmdue3Loadsatonc
e]
Index: []
```

In [27]:
```
df["PlasticMomentKNM"] = (
    (
        2 * (df["FlangeWidthmm"] * df["FlangeThicknessmm"] * ((df["Depthmm"]
        (df["WebThicknessmm"] * (df["Depthmm"] - 2 * df["FlangeThicknessmm"]
    ) * 10**-9 * 350 * 1000
)
```

In [28]:
```
def calculate_plastic_moment(Depthmm, FlangeThicknessmm, FlangeWidthmm, WebT
    """
    Calculates the PlasticMomentKNM based on the provided formula.

    Args:
        depthmm: Depth of the section (mm).
        FlangeThicknessmm: Thickness of the flange (mm).
        FlangeWidthmm: Width of the flange (mm).
        WebThicknessmm: Thickness of the web (mm).

    Returns:
        The calculated PlasticMomentKNM value.
    """

    return (
        2 * (FlangeWidthmm * FlangeThicknessmm * ((Depthmm / 2) - (FlangeThi
        (WebThicknessmm * (Depthmm - 2 * FlangeThicknessmm) * ((Depthmm - 2
    ) * 10**-9 * 350 * 1000
```

In [29]:
```
new_plastic_moment = calculate_plastic_moment(
    Depthmm=127.0,  # Replace with your desired values for each parameter
    FlangeThicknessmm=7.6,
    FlangeWidthmm=76,
    WebThicknessmm=4
)
```

In [30]:
```
print("New PlasticMomentKNM:", new_plastic_moment)
```

```
New PlasticMomentKNM: 28.512638000000003
```

In [31]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 0 entries
Data columns (total 12 columns):
 #   Column                         Non-Null Count  Dtype  
---  ------                         --------------  -----  
 0   MassKgPerm                     0 non-null      float64
 1   Depthmm                        0 non-null      float64
 2   FlangeWidthmm                  0 non-null      float64
 3   FlangeThicknessmm              0 non-null      float64
 4   WebThicknessmm                 0 non-null      float64
 5   Izpmm4                         0 non-null      float64
 6   CostEuro                       0 non-null      float64
 7   PlasticMomentKNM               0 non-null      float64
 8   Deflectionmmdue5KNLoadat2m     0 non-null      float64
 9   Deflectionmmdue8KNLoadat6m     0 non-null      float64
 10  Deflectionmmdue12KNLoadat8m    0 non-null      float64
 11  Deflectionmmdue3Loadsatonce    0 non-null      float64
dtypes: float64(12)
memory usage: 132.0 bytes
```

String formatting

By default r2_score returns a floating number (more details)

In [32]:
```
r2_score(Y_test, Y_pred)
```

Out[32]:
```
0.9860938544340759
```

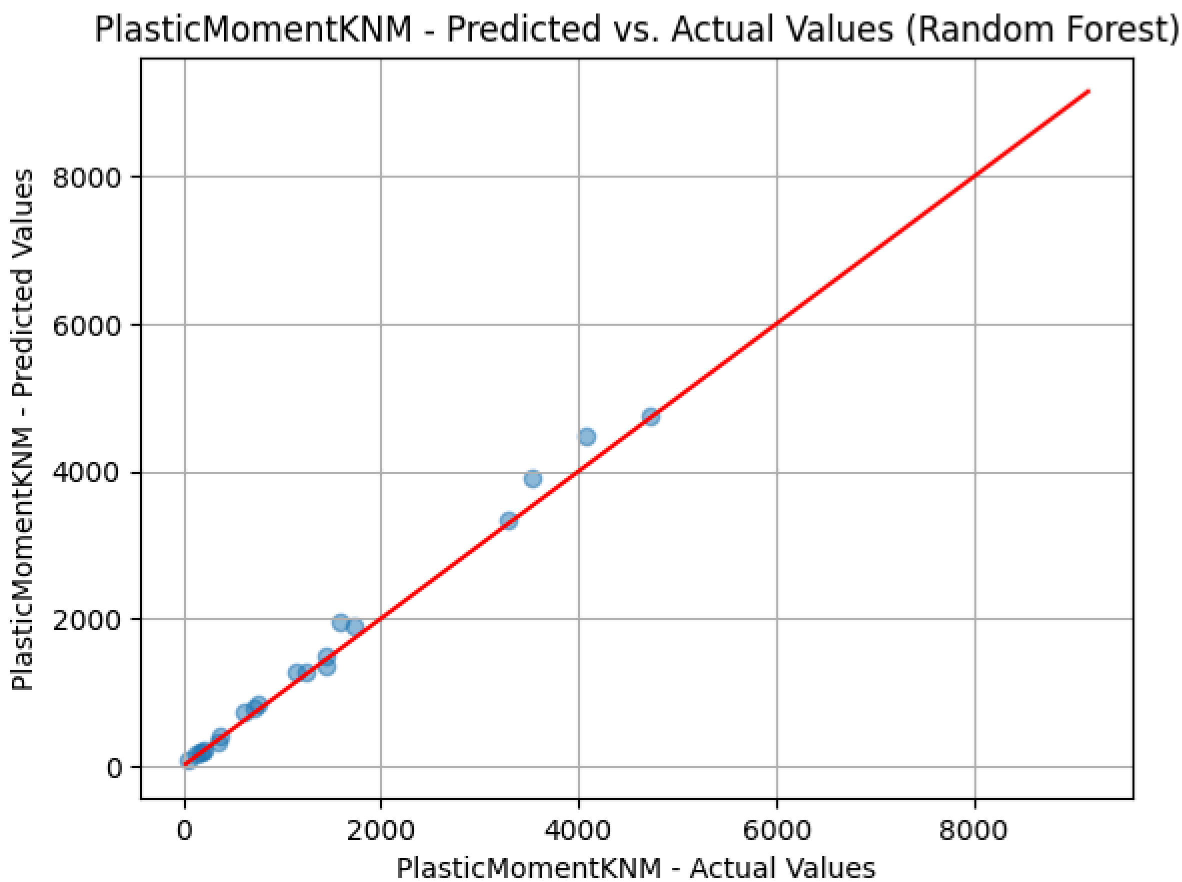
Scatter plots

Import library

In [33]:
```
import seaborn as sns
```

Make scatter plot

In [34]:
```
import matplotlib.pyplot as plt

# Plot the predicted vs. actual values
# predicted values of the target variable are plotted on the y-axis
# Actual values of the target variable are plotted on the x-axis
# If the model is a good fit, the points should be close to the diagonal lir
# indicating a strong linear relationship between the actual and predicted v

plt.scatter(Y_test, Y_pred, alpha=0.5)
plt.plot([Y.min(), Y.max()], [Y.min(), Y.max()], color="red")
plt.xlabel("PlasticMomentKNM - Actual Values")
plt.ylabel("PlasticMomentKNM - Predicted Values")
plt.title("PlasticMomentKNM - Predicted vs. Actual Values (Random Forest)")
plt.grid(True)
plt.show()
```



The Data

In [35]:
```
Y_test
```

Out[35]:

	PlasticMomentKNM
9	138.520
83	4723.759
70	1727.992
84	3290.275
76	1596.354
49	1437.520
15	167.128
44	708.493
2	58.564
50	620.885
67	1435.980
92	3533.965
14	214.756
59	1134.537
91	4077.664
65	754.166
18	185.484
27	348.432
24	376.996
45	1252.472

dtype: float64

In [36]:

```
import numpy as np
np.array(Y_test)
```

Out[36]:

```
array([ 138.52 , 4723.759, 1727.992, 3290.275, 1596.354, 1437.52 ,
        167.128,  708.493,   58.564,  620.885, 1435.98 , 3533.965,
        214.756, 1134.537, 4077.664,  754.166,  185.484,  348.432,
        376.996, 1252.472])
```

In [37]:

```
Y_pred
```

Out[37]:

```
array([ 156.65581, 4752.00865, 1898.27631, 3337.67212, 1966.81501,
       1354.23286,  176.89151,  773.69218,   72.86347,  719.86545,
       1496.58291, 3905.22906,  203.15002, 1263.70647, 4488.4585 ,
        845.74604,  201.45764,  312.62154,  405.75612, 1262.29301])
```

Making the scatter plot

In [38]:
```python
import seaborn as sns
```

In [39]:
```python
Y = UKUniversalSteelDatabase.PlasticMomentKNM
Y
```

Out[39]:

	PlasticMomentKNM
0	28.513
1	41.980
2	58.564
3	80.352
4	88.599
...	...
91	4077.664
92	3533.965
93	5343.902
94	6109.801
95	2876.129

96 rows × 1 columns

dtype: float64

In [40]:
```python
X = UKUniversalSteelDatabase.CostEuro
X
```

Out[40]:

	CostEuro
0	357.5
1	440.0
2	525.0
3	635.0
4	660.0
...	...
91	8910.0
92	9690.0
93	10300.0
94	11650.0
95	15050.0

96 rows × 1 columns

dtype: float64

In [41]:
```
from sklearn.model_selection import train_test_split
```

In [42]:
```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2)
```

In [43]:
```
X_train.shape, Y_train.shape
```

Out[43]: ((76,), (76,))

In [44]:
```
X_test.shape, Y_test.shape
```

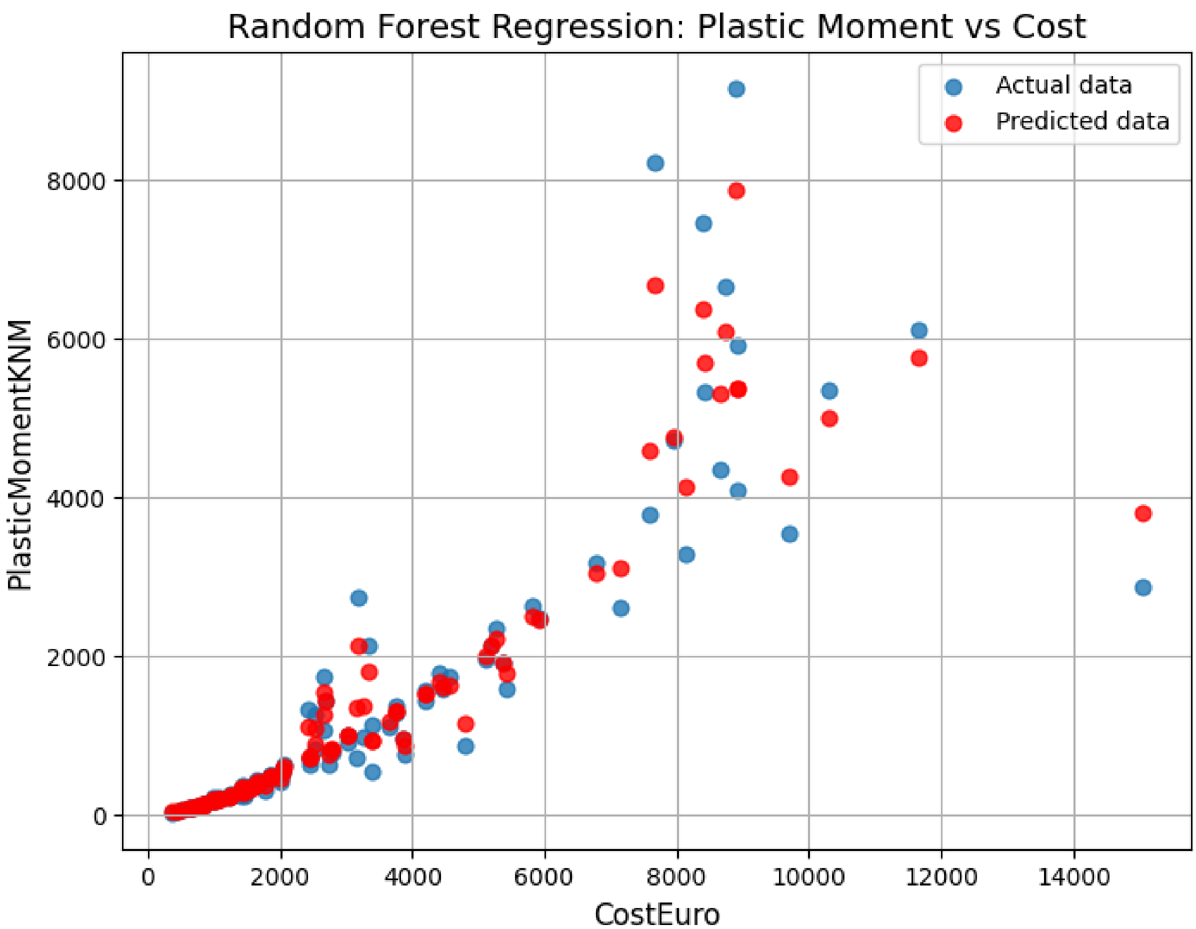
Out[44]: ((20,), (20,))

In [45]:
```
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error, r2_score
```

In [46]:
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Load data from CSV file
df = pd.read_csv("UK Universal Steel Database.csv")

# Extract relevant columns
cost = df["CostEuro"].values
plastic_moment = df["PlasticMomentKNM"].values

# Reshape data for RandomForest (it expects a 2D array for features)
X = cost.reshape(-1, 1)  # Feature: CostEuro
```

```
y = plastic_moment      # Target: PlasticMomentKNM

# Create Random Forest model and train it
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X, y)

# Make predictions using the trained model
y_pred = model.predict(X)

# Calculate performance metrics
mse = mean_squared_error(y, y_pred)
r2 = r2_score(y, y_pred)

# Print the results
print(f"Mean Squared Error (MSE): {mse:.2f}")
print(f"R^2 Score: {r2:.2f}")

# Create scatter plot with predictions
plt.figure(figsize=(8, 6))  # Adjust figure size as needed
plt.scatter(cost, plastic_moment, s=40, alpha=0.8, label="Actual data")  # S
plt.scatter(cost, y_pred, color='red', s=40, alpha=0.8, label="Predicted dat

# Customize plot elements
plt.xlabel("CostEuro", fontsize=12)
plt.ylabel("PlasticMomentKNM", fontsize=12)
plt.title("Random Forest Regression: Plastic Moment vs Cost", fontsize=14)
plt.grid(True)
plt.legend()

# Display the plot
plt.show()
```

```
Mean Squared Error (MSE): 139077.84
R^2 Score: 0.96
```



In [47]:
```
# Normalize Plastic Moment / Cost
df['NormalizedMomentCost'] = (df['PlasticMomentKNM'] / df['CostEuro']) / \
                              (df['PlasticMomentKNM'] / df['CostEuro']).max(

# Sum deflections from all load cases
df['TotalDeflection'] = df[[col for col in df.columns if col.startswith('Def

# Normalize total deflection
df['NormalizedDeflection'] = df['TotalDeflection'] / df['TotalDeflection'].m

# Score formula
df['Score'] = df['NormalizedMomentCost'] - df['NormalizedDeflection']
```

In [48]:
```
# Find the top 3 beams with the highest scores
top_3_beams = df.sort_values(by='Score', ascending=False).head(3)

# Iterate through the top 3 beams to get their original Excel row numbers ar
for idx, best_beam_row in top_3_beams.iterrows():
    # Get the original Excel row number (adding 2 because pandas index start
    excel_row_number = best_beam_row.name + 2

    # Get the beam's properties
    best_beam_properties = best_beam_row[[
        "MassKgPerm", "Depthmm", "FlangeWidthmm", "FlangeThicknessmm",
        "WebThicknessmm", "CostEuro", "PlasticMomentKNM", "Deflectionmmdue3L
    ]]
```

```
    # Print the results for each of the top 3 beams
    print(f"\nExcel Row Number: {excel_row_number}")
    print(best_beam_properties)
```

```
Excel Row Number: 84
MassKgPerm                        224.000000
Depthmm                           910.400000
FlangeWidthmm                     304.100000
FlangeThicknessmm                  23.900000
WebThicknessmm                     15.900000
CostEuro                         7680.000000
PlasticMomentKNM                 8215.418000
Deflectionmmdue3Loadsatonce         0.145000
Score                               0.999518
Name: 82, dtype: float64

Excel Row Number: 91
MassKgPerm                        381.000000
Depthmm                           951.000000
FlangeWidthmm                     310.000000
FlangeThicknessmm                  43.900000
WebThicknessmm                     24.400000
CostEuro                         8890.000000
PlasticMomentKNM                 9150.603000
Deflectionmmdue3Loadsatonce         0.129000
Score                               0.961802
Name: 89, dtype: float64

Excel Row Number: 87
MassKgPerm                        271.000000
Depthmm                           923.000000
FlangeWidthmm                     307.000000
FlangeThicknessmm                  30.000000
WebThicknessmm                     18.400000
CostEuro                         8410.000000
PlasticMomentKNM                 7440.347000
Deflectionmmdue3Loadsatonce         0.161000
Score                               0.826509
Name: 85, dtype: float64
```

In [51]:

```
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from itertools import product

# Load dataset
df = pd.read_csv("UK Universal Steel Database.csv")

# Select features and target (excluding Iz)
feature_cols = ['Depthmm', 'FlangeWidthmm', 'FlangeThicknessmm', 'WebThickne
target_cols = ['PlasticMomentKNM', 'Deflectionmmdue3Loadsatonce', 'CostEuro'

# Train-test split
X = df[feature_cols]
y = df[target_cols]
```

```python
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, rar

# Train model for PlasticMoment and Deflection
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train[['PlasticMomentKNM', 'Deflectionmmdue3Loadsatonce

# Train model for Cost prediction
y_cost = df['CostEuro']
X_train_cost, X_test_cost, y_train_cost, y_test_cost = train_test_split(X, y
model_cost = RandomForestRegressor(n_estimators=100, random_state=42)
model_cost.fit(X_train_cost, y_train_cost)

# ----------------------------
# Step 1: Calculate realistic ratio bounds from real dataset
# ----------------------------
df['bf_h'] = df['FlangeWidthmm'] / df['Depthmm']
df['tf_h'] = df['FlangeThicknessmm'] / df['Depthmm']
df['tw_h'] = df['WebThicknessmm'] / df['Depthmm']

bounds = {
    'bf_h_min': np.percentile(df['bf_h'], 5),
    'bf_h_max': np.percentile(df['bf_h'], 95),
    'tf_h_min': np.percentile(df['tf_h'], 5),
    'tf_h_max': np.percentile(df['tf_h'], 95),
    'tw_h_min': np.percentile(df['tw_h'], 5),
    'tw_h_max': np.percentile(df['tw_h'], 95),
}
print("Realistic ratio bounds:", bounds)

# ----------------------------
# Step 2: Generate synthetic beams within realistic bounds
# ----------------------------
depth_vals = np.linspace(df['Depthmm'].min(), df['Depthmm'].max(), 10)
flange_vals = np.linspace(df['FlangeWidthmm'].min(), df['FlangeWidthmm'].max
ft_vals = np.linspace(df['FlangeThicknessmm'].min(), df['FlangeThicknessmm']
web_vals = np.linspace(df['WebThicknessmm'].min(), df['WebThicknessmm'].max(

synthetic_combos = []
for D, B, tf, tw in product(depth_vals, flange_vals, ft_vals, web_vals):
    bf_h = B / D
    tf_h = tf / D
    tw_h = tw / D
    # Apply realistic ratio filtering
    if (bounds['bf_h_min'] <= bf_h <= bounds['bf_h_max'] and
        bounds['tf_h_min'] <= tf_h <= bounds['tf_h_max'] and
        bounds['tw_h_min'] <= tw_h <= bounds['tw_h_max']):
        synthetic_combos.append((D, B, tf, tw))

synthetic_df = pd.DataFrame(synthetic_combos, columns=feature_cols)
print(f"Generated {len(synthetic_df)} realistic synthetic candidates.")

# ----------------------------
# Step 3: Structural calculations and predictions
# ----------------------------
def calculate_iz(row):
    D = row['Depthmm']
```

```python
    B = row['FlangeWidthmm']
    tf = row['FlangeThicknessmm']
    tw = row['WebThicknessmm']
    Iz_flange = 2 * ((B * tf**3) / 12 + (B * tf) * ((D - tf) / 2)**2)
    Iz_web = (tw * (D - 2 * tf)**3) / 12
    return Iz_flange + Iz_web

synthetic_df['Izpmm4'] = synthetic_df.apply(calculate_iz, axis=1)

# Predict PlasticMoment and Deflection using ML model
synthetic_df[['PlasticMomentKNM', 'Deflectionmmdue3Loadsatonce']] = model.pr

# Predict the cost for synthetic beams
synthetic_df['CostEuro'] = model_cost.predict(synthetic_df[feature_cols])

# Manual deflection calc
E = 210000  # MPa
L = 10000   # mm
def manual_deflection(row):
    Iz = row['Izpmm4']
    term1 = 5000 * 2000 * 8000 * (L**2 - 2000**2 - 8000**2)
    term2 = 8000 * 6000 * 4000 * (L**2 - 6000**2 - 4000**2)
    term3 = 12000 * 8000 * 2000 * (L**2 - 8000**2 - 2000**2)
    return (term1 + term2 + term3) / (6 * E * Iz * L)

synthetic_df['Deflectionmmdue3Loadsatonce'] = synthetic_df.apply(manual_defl

# Scoring
synthetic_df['NormalizedMomentCost'] = synthetic_df['PlasticMomentKNM'] / sy
synthetic_df['NormalizedMomentCost'] /= synthetic_df['NormalizedMomentCost']

synthetic_df['NormalizedDeflection'] = synthetic_df['Deflectionmmdue3Loadsat
                                       synthetic_df['Deflectionmmdue3Loadsat

synthetic_df['Score'] = synthetic_df['NormalizedMomentCost'] - synthetic_df[

# Top 100
top100 = synthetic_df.sort_values('Score', ascending=False).head(100)

print("\n--- Top 100 Beam Suggestions (Realistic) ---")
print(top100[['Depthmm', 'FlangeWidthmm', 'FlangeThicknessmm', 'WebThickness
             'Izpmm4', 'PlasticMomentKNM', 'Deflectionmmdue3Loadsatonce', 'Co
from google.colab import files

# Save the CSV file to the local filesystem in Colab
top100.to_csv('top_100_beams.csv', index=False)

# Download the file to your local machine
files.download('top_100_beams.csv')
```

```
Realistic ratio bounds: {'bf_h_min': np.float64(0.31437982381378604), 'bf_h_max': np.float64(0.5766228017113382), 'tf_h_min': np.float64(0.022480719635789036), 'tf_h_max': np.float64(0.056747447634957066), 'tw_h_min': np.float64(0.01684092834043427), 'tw_h_max': np.float64(0.0320416731004859)}
Generated 318 realistic synthetic candidates.

--- Top 100 Beam Suggestions (Realistic) ---
        Depthmm  FlangeWidthmm  FlangeThicknessmm  WebThicknessmm  \
221  896.777778     305.666667          45.600000       25.400000
218  896.777778     305.666667          39.133333       25.400000
220  896.777778     305.666667          45.600000       21.833333
219  896.777778     305.666667          45.600000       18.266667
212  896.777778     305.666667          26.200000       25.400000
215  896.777778     305.666667          32.666667       25.400000
23   511.888889     229.111111          13.266667       14.700000
217  896.777778     305.666667          39.133333       21.833333
25   511.888889     229.111111          19.733333       14.700000
29   511.888889     267.388889          13.266667       14.700000
31   511.888889     267.388889          19.733333       14.700000
216  896.777778     305.666667          39.133333       18.266667
27   511.888889     229.111111          26.200000       14.700000
211  896.777778     305.666667          26.200000       21.833333
33   511.888889     267.388889          26.200000       14.700000
272  993.000000     343.944444          45.600000       25.400000
214  896.777778     305.666667          32.666667       21.833333
210  896.777778     305.666667          26.200000       18.266667
271  993.000000     343.944444          45.600000       21.833333
268  993.000000     343.944444          39.133333       25.400000
213  896.777778     305.666667          32.666667       18.266667
276  993.000000     343.944444          52.066667       25.400000
270  993.000000     343.944444          45.600000       18.266667
267  993.000000     343.944444          39.133333       21.833333
233  896.777778     343.944444          45.600000       25.400000
273  993.000000     343.944444          45.600000       28.966667
275  993.000000     343.944444          52.066667       21.833333
264  993.000000     343.944444          32.666667       25.400000
260  993.000000     343.944444          26.200000       25.400000
266  993.000000     343.944444          39.133333       18.266667
269  993.000000     343.944444          39.133333       28.966667
274  993.000000     343.944444          52.066667       18.266667
232  896.777778     343.944444          45.600000       21.833333
230  896.777778     343.944444          39.133333       25.400000
263  993.000000     343.944444          32.666667       21.833333
259  993.000000     343.944444          26.200000       21.833333
265  993.000000     343.944444          32.666667       28.966667
261  993.000000     343.944444          26.200000       28.966667
19   511.888889     190.833333          19.733333       14.700000
231  896.777778     343.944444          45.600000       18.266667
262  993.000000     343.944444          32.666667       18.266667
258  993.000000     343.944444          26.200000       18.266667
277  993.000000     343.944444          52.066667       28.966667
229  896.777778     343.944444          39.133333       21.833333
224  896.777778     343.944444          26.200000       25.400000
227  896.777778     343.944444          32.666667       25.400000
223  896.777778     343.944444          26.200000       21.833333
228  896.777778     343.944444          39.133333       18.266667
```

21	511.888889	190.833333	26.200000	14.700000
226	896.777778	343.944444	32.666667	21.833333
17	511.888889	190.833333	13.266667	14.700000
222	896.777778	343.944444	26.200000	18.266667
225	896.777778	343.944444	32.666667	18.266667
28	511.888889	267.388889	13.266667	11.133333
13	415.666667	229.111111	13.266667	11.133333
22	511.888889	229.111111	13.266667	11.133333
53	608.111111	305.666667	19.733333	14.700000
32	511.888889	267.388889	26.200000	11.133333
35	608.111111	229.111111	19.733333	14.700000
44	608.111111	267.388889	19.733333	14.700000
56	608.111111	305.666667	26.200000	14.700000
62	608.111111	343.944444	19.733333	14.700000
47	608.111111	267.388889	26.200000	14.700000
38	608.111111	229.111111	26.200000	14.700000
30	511.888889	267.388889	19.733333	11.133333
26	511.888889	229.111111	26.200000	11.133333
312	993.000000	420.500000	45.600000	25.400000
292	993.000000	382.222222	45.600000	25.400000
12	415.666667	229.111111	13.266667	7.566667
65	608.111111	343.944444	26.200000	14.700000
15	415.666667	229.111111	19.733333	11.133333
130	800.555556	267.388889	19.733333	14.700000
59	608.111111	305.666667	32.666667	14.700000
54	608.111111	305.666667	19.733333	18.266667
41	608.111111	229.111111	32.666667	14.700000
24	511.888889	229.111111	19.733333	11.133333
308	993.000000	420.500000	39.133333	25.400000
311	993.000000	420.500000	45.600000	21.833333
288	993.000000	382.222222	39.133333	25.400000
291	993.000000	382.222222	45.600000	21.833333
50	608.111111	267.388889	32.666667	14.700000
316	993.000000	420.500000	52.066667	25.400000
296	993.000000	382.222222	52.066667	25.400000
36	608.111111	229.111111	19.733333	18.266667
45	608.111111	267.388889	19.733333	18.266667
310	993.000000	420.500000	45.600000	18.266667
290	993.000000	382.222222	45.600000	18.266667
14	415.666667	229.111111	19.733333	7.566667
146	800.555556	305.666667	19.733333	14.700000
70	704.333333	229.111111	19.733333	14.700000
307	993.000000	420.500000	39.133333	21.833333
287	993.000000	382.222222	39.133333	21.833333
82	704.333333	267.388889	19.733333	14.700000
20	511.888889	190.833333	26.200000	11.133333
313	993.000000	420.500000	45.600000	28.966667
293	993.000000	382.222222	45.600000	28.966667
315	993.000000	420.500000	52.066667	21.833333
295	993.000000	382.222222	52.066667	21.833333
94	704.333333	305.666667	19.733333	14.700000
16	511.888889	190.833333	13.266667	11.133333

	Izpmm4	PlasticMomentKNM	Deflectionmmdue3Loadsatonce	CostEuro \
221	6.160601e+09	6175.19570	0.230858	8400.6

218	5.563028e+09	5780.53246	0.255656	8373.4
220	6.005218e+09	5667.74300	0.236831	8385.8
219	5.849835e+09	5353.05710	0.243122	8188.7
212	4.310021e+09	5278.82917	0.329980	8183.3
215	4.946269e+09	5272.61188	0.287534	8314.6
23	5.180014e+08	1827.39307	2.745595	2797.5
217	5.400040e+09	5273.07976	0.263373	8358.6
25	6.770001e+08	1789.62438	2.100771	2774.8
29	5.811442e+08	1827.27688	2.447279	2867.2
31	7.685283e+08	1799.10348	1.850579	2854.3
216	5.237053e+09	4974.85157	0.271569	8174.3
27	8.275272e+08	1714.66129	1.718641	2798.5
211	4.131088e+09	4912.28688	0.344273	8175.7
33	9.459282e+08	1746.61046	1.503520	2875.7
272	8.596421e+09	5573.49938	0.165444	9465.2
214	4.775433e+09	4884.03522	0.297821	8300.8
210	3.952154e+09	4633.00132	0.359860	7985.1
271	8.378443e+09	5372.62739	0.169748	9446.0
268	7.746742e+09	5348.79288	0.183590	9439.2
213	4.604596e+09	4603.81700	0.308870	8116.4
276	9.422074e+09	5229.01837	0.150946	9476.4
270	8.160466e+09	5045.39377	0.174282	9227.3
267	7.519251e+09	5147.92089	0.189144	9420.0
233	6.793504e+09	4838.31552	0.209350	8964.5
273	8.814399e+09	5094.89555	0.161352	9466.5
275	9.213341e+09	5078.18704	0.154366	9456.8
264	6.872694e+09	4988.86697	0.206938	9352.6
260	5.973931e+09	4893.64282	0.238071	9219.1
266	7.291760e+09	4873.08400	0.195045	9201.1
269	7.974234e+09	4970.27037	0.178352	9440.5
274	9.004608e+09	4854.51387	0.157944	9237.9
232	6.638121e+09	4664.66655	0.214251	8943.9
230	6.114317e+09	4643.30069	0.232605	8926.9
263	6.635416e+09	4852.80026	0.214338	9336.8
259	5.726590e+09	4757.57611	0.248354	9203.3
265	7.109972e+09	4793.76822	0.200032	9354.3
261	6.221272e+09	4716.88215	0.228606	9214.0
19	5.854720e+08	1364.08808	2.429189	2606.3
231	6.482738e+09	4403.43267	0.219386	8710.0
262	6.398138e+09	4582.10483	0.222287	9117.1
258	5.479249e+09	4508.63888	0.259565	8973.8
277	9.630807e+09	4750.41454	0.147674	9477.7
229	5.951330e+09	4469.65172	0.238976	8906.3
224	4.690180e+09	4357.01394	0.303234	8711.7
227	5.413324e+09	4378.20298	0.262726	8839.8
223	4.511246e+09	4264.66128	0.315262	8696.7
228	5.788342e+09	4233.43817	0.245705	8683.4
21	7.091262e+08	1315.38462	2.005598	2630.6
226	5.242487e+09	4273.38939	0.271288	8820.2
17	4.548585e+08	1320.29379	3.126735	2629.7
222	4.332313e+09	4063.13278	0.328282	8466.2
225	5.071651e+09	4065.80897	0.280426	8597.2
28	5.471613e+08	1352.87779	2.599274	2849.0
13	3.008479e+08	995.41206	4.727380	2068.8
22	4.840184e+08	1289.49577	2.938364	2778.5
53	1.269712e+09	1769.61634	1.120114	4006.7

32	9.170941e+08	1251.61008	1.550792	2854.9
35	1.008121e+09	1725.98306	1.410766	3951.4
44	1.138916e+09	1733.41607	1.248750	4016.0
56	1.567056e+09	1696.29078	0.907576	4009.2
62	1.400507e+09	1657.76064	1.015505	3987.5
47	1.397144e+09	1669.93666	1.017950	4018.5
38	1.227232e+09	1644.42639	1.158886	3953.5
30	7.371902e+08	1184.97614	1.929247	2833.3
26	7.986931e+08	1157.35268	1.780687	2777.1
312	1.016430e+10	5218.77563	0.139923	13170.4
292	9.380362e+09	5218.77563	0.151617	13170.4
12	2.833343e+08	807.60278	5.019592	1864.8
65	1.736969e+09	1573.39854	0.818796	3990.0
15	4.040635e+08	878.34068	3.519799	2116.0
130	2.148892e+09	2328.93576	0.661840	5977.2
59	1.850878e+09	1562.03978	0.768404	4009.2
54	1.324363e+09	1566.62311	1.073891	4007.4
41	1.436378e+09	1541.31597	0.990145	3953.5
24	6.456620e+08	1094.28476	2.202735	2753.2
308	9.110423e+09	5019.08946	0.156109	13161.6
311	9.946326e+09	5017.90364	0.142990	13164.0
288	8.428583e+09	5019.08946	0.168738	13161.6
291	9.162385e+09	5017.90364	0.155224	13164.0
50	1.643628e+09	1540.89019	0.865295	4018.5
316	1.118839e+10	4918.34176	0.127116	13180.4
296	1.030523e+10	4918.34176	0.138010	13180.4
36	1.062773e+09	1505.85553	1.338219	3952.1
45	1.193568e+09	1525.37715	1.191572	4016.7
310	9.728348e+09	4809.06822	0.146194	13064.6
290	8.944407e+09	4809.06822	0.159007	13064.6
14	3.882387e+08	757.33709	3.663267	1915.7
146	2.379203e+09	2347.02310	0.597773	6322.5
70	1.419799e+09	1852.56442	1.001707	4965.7
307	8.882931e+09	4818.21747	0.160107	13155.2
287	8.201091e+09	4818.21747	0.173419	13155.2
82	1.596855e+09	1891.20260	0.890640	5101.7
20	6.802921e+08	988.08772	2.090605	2607.9
313	1.038228e+10	4772.56311	0.136986	13153.4
293	9.598340e+09	4772.56311	0.148174	13153.4
315	1.097965e+10	4767.51043	0.129533	13174.0
295	1.009650e+10	4767.51043	0.140863	13174.0
94	1.773911e+09	1926.88917	0.801744	5259.7
16	4.208755e+08	1003.07297	3.379199	2605.0

	Score
221	0.997536
218	0.936401
220	0.916917
219	0.886702
212	0.874022
215	0.859600
23	0.859331
217	0.855393
25	0.854965
29	0.840857
31	0.837716

216	0.825023
27	0.815173
211	0.813695
33	0.810206
272	0.799281
214	0.797243
210	0.785459
271	0.771934
268	0.768909
213	0.768343
276	0.749037
270	0.741981
267	0.741412
233	0.731988
273	0.730437
275	0.728859
264	0.723445
260	0.719569
266	0.718402
269	0.714314
274	0.713193
232	0.707215
230	0.705114
263	0.704769
259	0.700587
265	0.695014
261	0.693973
19	0.686073
231	0.685412
262	0.681331
258	0.680714
277	0.680273
229	0.680159
224	0.677135
227	0.670968
223	0.663734
228	0.660606
21	0.658830
226	0.656208
17	0.649638
222	0.649375
225	0.640361
28	0.618251
13	0.604103
22	0.599991
53	0.588877
32	0.579851
35	0.579162
44	0.573850
56	0.565890
62	0.554726
47	0.554458
38	0.553471
30	0.548365
26	0.547933
312	0.537557

```
292  0.537432
12   0.535582
65   0.527707
15   0.527124
130  0.522990
59   0.521822
54   0.520356
41   0.519792
24   0.517188
308  0.517105
311  0.517028
288  0.516970
291  0.516898
50   0.512402
316  0.506276
296  0.506160
36   0.504059
45   0.503899
310  0.499194
290  0.499057
14   0.498708
146  0.498617
70   0.496829
307  0.496543
287  0.496401
82   0.494788
20   0.493113
313  0.492136
293  0.492017
315  0.490922
295  0.490801
94   0.489818
16   0.487761
```

In [52]:

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load your dataset (make sure the CSV is in the same directory or adjust th
df = pd.read_csv("UK Universal Steel Database.csv")

# Select the relevant columns for correlation
selected_columns = [
    'Depthmm', 'FlangeWidthmm', 'FlangeThicknessmm', 'WebThicknessmm',
    'PlasticMomentKNM', 'Deflectionmmdue3Loadsatonce',
    'MassKgPerm', 'CostEuro', 'Izpmm4'
]

# Filter the DataFrame
df_corr = df[selected_columns]

# Compute correlation matrix using Pearson correlation
correlation_matrix = df_corr.corr()

# Plot the heatmap
plt.figure(figsize=(12, 8))
```

```
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', center=0, fmt="
plt.title("Correlation Heatmap of I-Beam Parameters and Performance Metrics"
plt.xticks(rotation=45, ha='right')
plt.yticks(rotation=0)
plt.tight_layout()

plt.savefig("correlation_heatmap.png", dpi=300)
plt.show()
files.download('correlation_heatmap.png')
```