Forging Preform Design Optimization for Structural Magnesium Components

by

Tharindu Kodippili

A thesis

presented to the University of Waterloo

in fulfillment of

the thesis requirement for the degree of

Doctor of Philosophy

in

Mechanical and Mechatronics Engineering

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	**Dr. Kamran Behdinan** Professor, Department of Mechanical and Industrial Engineering University of Toronto, Toronto, ON
Supervisor	**Dr. Stephan Lambert** Professor, Department of Mechanical and Mechatronics Engineering University of Waterloo, Waterloo, ON
Co-Supervisor	**Dr. Arash Arami** Professor, Department of Mechanical and Mechatronics Engineering University of Waterloo, Waterloo, ON
Internal Member	**Dr. Marry Wells** Professor, Department of Mechanical and Mechatronics Engineering University of Waterloo, Waterloo, ON
Internal Member	**Dr. Hamid Jahed** Professor, Department of Mechanical and Mechatronics Engineering University of Waterloo, Waterloo, ON
Internal Member	**Dr. Alex Wong** Professor, Department of Systems Design Engineering University of Waterloo, Waterloo, ON

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contribution included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The research detailed in this thesis was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) through the Strategic Partnership Grant (SPG) with contributions from Multimatic Technical Centre, and CanmetMATERIALS, Natural Resources Canada. The contents presented in Chapter 3 are taken from a manuscript that is currently being prepared for publication. The candidate is the primary author of the manuscript and was responsible for developing the new material models, conducting forging simulations, and analyzing results. The contents presented in Chapter 4 are taken from two manuscripts that were published. The candidate is the primary author of both manuscripts. The individual contributions from all of the co-authors are as follows:

Prof. Stephan Lambert and Prof. Arash Arami: provided both technical and editorial feedback for the manuscripts, as well as supervised and mentored all work which was completed in this thesis.

Dr. Seyed Behzad Behravesh: provided both technical and editorial feedback in the preparation of the manuscript content presented in Chapter 3, as well as the machining schematic Figure of the Mg alloy billet.

Dr. Andrew Gryguć: provided both technical and editorial feedback in the preparation of the manuscript content presented in Chapter 3.

Dr. Paresh Prakash: conducted small-sample uniaxial compression tests to provide flow stress curves, and provided both technical and editorial feedback in the preparation of the manuscript content presented in Chapter 3.

Mr. Alex Duquette: Liaised between Multimatic Technical Centre and the University of Waterloo to aid with the casting and forging tool fabrication and facilitated component-level testing, which helped identify the defects presented in Chapter 3.

Dr. Fateh Fazeli: provided feedback on torsion sample design and conducted torsion tests, the results of which are presented in Chapter 3.

Dr. Bruce Williams and Dr. Jonathan McKinley: Liaised between CanmetMATERIALS and the University of Waterloo to aid with cast-forging of the various Magnesium alloy components presented in this work.

Mr. Lucian Blaga: For conducting the laboratory-scale forging experiments at CanmetMATERIALS forging facility.

Mr. Erfan Azqadan: performed mechanical tests on forged I-beam samples and provided the Quasi-static tensile strain-stress curve figure, and yield and ultimate tensile strengths and fracture strain table for the cast-billet and optimized preform forging samples, in addition to the analysis of mechanical test data presented in Chapter 4.

Mr. Jared Uramowski: provided flow stress curves of the as-cast AZ80 alloy, which were used to conduct all optimization simulations presented in this work.

Abstract

An increasing emphasis on emissions reduction and improved fuel economy is leading to a broader utilization of magnesium (Mg) alloys in vehicle light-weighting applications. Mg alloys are the lightest structural metals and hold great promise in automotive applications owing to their high strength, stiffness-to-weight ratio, castability, machinability, and damping. Mg alloy components are typically used in non-load-bearing applications and are produced by casting processes, which are cost-effective methods for producing components with intricate geometry. However, as-cast components can exhibit poor mechanical properties due to porosity and microstructure inhomogeneity. On the other hand, forged components exhibit superior mechanical properties compared to their as-cast counterpart but have been predominantly limited to high-cost sports and military applications due to the poor formability of the material. In addition, a workpiece may be subjected to bending and pre-forming before forging, which can be resource-intensive and result in significant material waste. While both forging and casting methods are suitable for large-scale production of components, typically, only forged components exhibit the adequate mechanical properties that are required for structural applications in vehicles. To leverage the benefits of both casting and forging, a novel hybrid manufacturing technique is introduced to sequentially combine casting and forging steps to produce high-strength Mg alloy structural components that can be both intricate in shape and cost-effective to manufacture. In this novel approach, the intermediate workpiece (or preform) is cast and then forged into the desired shape. The current research is part of a larger advanced manufacturing and lightweight materials research project (the SPG project), with a primary objective of cast-forging an industrial-scale front lower control arm (FLCA) for the 2013 Ford Fusion vehicle using an AZ80 Mg alloy.

The focus of this thesis is on forging preform design optimization for effectively engineering material distribution within forging dies to induce the desired levels of strain throughout the forged component while minimizing material waste and fully filling the die. Preform design optimization is computationally intensive, demands manual computer-aided design (CAD) modelling efforts, and places considerable reliance on engineering judgment and experience. In addition, the use of disjointed CAD and Finite Element Method (FEM) software makes it difficult to effectively incorporate FEM simulation responses to inform design updates. The contributions of this thesis include (i) a set of phenomenological material models (both anisotropic and isotropic models) for use in FEM simulations to predict the deformation behaviour of AZ80 alloy during hot forging; (ii) a global design optimization method using a data-driven multi-objective optimization framework for optimizing three-dimensional forging preform designs; and (iii) a novel local design optimization method using a topology-based optimization framework to iteratively and automatically update three-dimensional forging preform designs.

Acknowledgements

Foremost, I would like to express my sincere gratitude to Profs. Stephan Lambert, Arash Arami, Hamid Jahed, and Mary Wells for providing me with the opportunity to work on this project. To my supervisors, Stephan Lambert, and Arash Arami, I am deeply grateful for the guidance and support you've both given me throughout my academic journey. Your expertise, invaluable feedback, and mentorship have been instrumental in shaping my research acumen and enhancing the quality of this work.

To my research colleagues, Paresh Prakash, Andrew Gryguć, Erfan Azqadan, Jared Uramowski, Behzad Behravesh, Massimo Diciano, Talal Paracha, Gou Yu and Alex Strong from University of Waterloo and beyond, thank you all for the invaluable contributions you have made to this project. It was a privilege to work alongside you, and I am truly appreciative of the valuable experience and knowledge that you have all shared with me.

I would also like to thank Alex Duquette from Multimatic Technical Centre, for playing a pivotal role in coordinating tool manufacturing, component-level testing and more. I am also truly appreciative of the words of encouragement you have given me in the past, which has served as constant source of inspiration throughout my academic journey.

To Bruce Williams Jonathan, Jonathan McKinley from CanmetMATERIALS, thank you for all the help you have provided me throughout my academic journey, from tool design to forge modelling, and for the conducting various laboratory-scale forging experiments at the forging facility.

To my lab members, Erfan Azqadan, Mohammad Shushtari, Iara Santelices, Sahand Ajami, Julia Goyal, Hannah Dinovitzer, Sola Weng, Paul Wolfe, Milad Nazarahari, and Rezvan Nasiri, I am deeply grateful for the numerous stimulating conversations and the laughs we have shared. The inquisitive and joyful atmosphere that you all helped foster within our lab group has been a very memorable part of my academic journey.

Furthermore, I would like to acknowledge the support of Natural Sciences and Engineering Research Council of Canada (NSERC) through the Strategic Partnership Grant (SPG) under STPGP 521551 - 18 grant throughout this work.

Finally, I would like to sincerely thank my parents for their encouragement and support throughout this journey.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANN — Artificial Neural Network

BESO — Bi-Direction Evolutionary Structural Optimization

CAD — Computer-Aided Design

DOE — Design of Experiments

DRX — Dynamic Recrystallization

EA — Evolutionary Algorithm

FEM — Finite Element Method

FFD — Free-Form Deformation

FLCA — Front Lower Control Arm

GA — Genetic Algorithm

HCP — Hexagonal Close-Packed

HCR — High Cooling Rate

HPDC — High-Pressure Die Casting

RBF — Radial Basis Function

SBD — Simulation-Based Design

SBDO — Simulation-Based Design Optimization

SDK — Software Development Kit

SPEAII — Strength-Pareto Evolutionary Algorithm 2

KPI — Key Performance Indicator

LCR — Low Cooling Rate

LHS — Latin Hypercube Sampling with Multi-dimensional Uniformity

LHSMDU — Latin Hypercube Sampling with Multi-Dimensional Uniformity

MIGA — Multi-Island Genetic Algorithm

MSE — Mean-Squared-Error

MAE — Mean Absolute Error

NSGA-II — Non-Dominating Sorting Genetic Algorithm-II

PMC — Permanent Mould Casting

RSM — Response Surface Models

YS — Yield Strength

1. Introduction

1.1 Overview

In an effort to improve vehicle performance and efficiency, the automotive industry is relying more on lightweight materials for manufacturing structural components. Magnesium (Mg) alloys are among the lightest commercially available structural metals. They hold significant promise in this area due to their weight-saving potential and favourable material properties compared to conventional materials used in the industry. However, they remain underutilized in structural automotive applications due to challenges associated with manufacturing, processing, assembly, in-service performance, and cost [1]. Much of their use has been limited to producing non-load-bearing components using casting processes [2]. While casting offers greater design flexibility, cast products can exhibit coarse and non-uniform microstructure and porosity defects. In addition, large grains and large intermetallic particles often present in cast microstructure contribute to subpar quasi-static and fatigue strengths [3]. These drawbacks limit casting as a primary method for manufacturing structurally-demanding, load-bearing components.

On the other hand, hot forged Mg alloy components exhibit superior mechanical properties compared to their cast counterpart as the strain energy imposed during forging refines the microstructure of the as-cast or wrought workpiece into a more uniform and fine-grained structure. Typically, hot forging of as-cast or wrought billets requires one or more pre-forming steps to splay material throughout the die before the final forging step. Naturally, the extra required tooling increases the cost associated with multi-step forging.

This research introduces a novel manufacturing process, cast-forging, which combines casting and forging as a sequence of manufacturing steps. Complex-shaped forging workpieces (preforms) can be produced via casting (eliminating the cost associated with fabricating pre-forming dies). Then, the preform is subjected to a hot forging step to induce plastic deformation of the as-cast structure to produce a forged component with improved strength and ductility.

The current work builds on the foundational knowledge acquired from an NSERC Automotive Partnership Canada (APC) project, where multidisciplinary research teams collaborated to manufacture an industrial-scale Mg alloy front lower control arm (FLCA) for a 2013 Ford Fusion vehicle. Following a multi-step forging process, a Mg alloy preform was forged into the desired shape of a control arm. The sequence of manufacturing steps included a heated billet bending operation followed by a heated radial flattening operation (pre-forming). Lastly, a closed-die hot forging operation to force the preform material into the die impression (see Appendix A.1). The APC research project provided critical insights into the hot deformation behaviour of as-cast and wrought AZ80 Mg alloys. A thorough study was conducted to understand how various process

parameters, such as temperature and strain rate of deformation, and material-specific parameters, such as initial texture, precipitate content, and material processing history, influence microstructure and texture evolution throughout the deformation process. The FLCAs forged during the APC project revealed a critical aspect of the manufacturing process that needed improvement—the preform design—based on observations of flow-related defects on FLCA forgings. Following the completion of the APC project, an NSERC Strategic Partnership Grant (SPG) project was initiated to continue the research and development of a cost-effective cast-forging process, which involved casting an AZ80 alloy preform and then hot forging it to produce FLCAs. The SPG project was composed of three major tasks: (i) to establish a link between structure-process-properties in cast-forgings of Mg alloys; (ii) to develop an AI-based optimization algorithm to optimize the design of the cast preform shape to minimize material use and maximize post-forging quasi-static and cyclic properties; and (iii) to cast-forge an exemplar complex-shaped automotive suspension component (FLCA) out of Mg with at least 30% weight saving compared to the Aluminum benchmark [4].

As part of the broader research effort to develop a cost-effective cast-forging process for producing Mg alloy FLCAs, this thesis work aims to develop suitable forging preform design optimization methods to conceive optimal preform designs that will produce components with superior mechanical properties while minimizing material waste (in the form of flash) and achieving complete die fill. To this end, the proposed research aims to achieve the following objectives:

1. Systematically develop phenomenological material models (both anisotropic and isotropic models) for use in finite element method (FEM) simulations to predict the deformation behaviour of AZ80 alloy during hot forging.

2. Develop a global design optimization method using a data-driven multi-objective optimization framework to evolve preform designs.

3. Develop a local design optimization method using a novel topology-based optimization framework, which fully automates modelling efforts at each incremental topology update step.

1.2 Thesis Structure

The subsequent chapter of this thesis, Chapter 2, aims to provide a review of critical concepts relevant to Mg alloy forging with sub-sections focusing on the casting and forging of the AZ80 alloy, past work on forging simulations, material data collection procedures, and the broad topics of design optimization and machine learning. Chapter 3 highlights the simulation results of Gleeble® compression, "Coin", I-beam, and FLCA forgings. The key contributions of this section are a set of phenomenological material models for predicting the anisotropic deformation behaviour of wrought AZ80 alloy and the isotropic deformation behaviour of as-cast AZ80 alloy. Chapter 4 introduces a data-driven multi-objective optimization framework that was used to obtain an optimal preform shape for forging an AZ80 alloy structural I-beam component—a key contribution that helped fulfill objective (ii) of the SPG project. Chapter 5 introduces a novel topology-based optimization framework and demonstrates its functionality on a toy problem to distribute material within a volumetric segment of the FLCA forging die. The key contribution of this chapter is this framework and the algorithms that drive fully automated three-dimensional shape optimization. Chapter 6 summarizes the findings and conclusions of each chapter. Chapter 7 presents a theoretical formulation of a generative design-based expert system as part of future work, which extends the functionality of the proposed topology-based optimization framework by integrating state-of-the-art physics and agent-based machine learning techniques. Lastly, the list of references and appendices are presented.

2. Literature Review

2.1 Magnesium Alloys

Magnesium (Mg) alloys are the lightest of all structural metals, with a density of 1.74 g/cm^3. It is approximately two-thirds the density of aluminum (Al) and one-fourth that of steel. They are also less dense than most glass-fibre-reinforced polymers used in automotive applications. They have a similar density to carbon fibre composites, which are relatively more expensive and challenging to recycle [1], [5]–[7]. Mg has a hexagonal close-packed (HCP) crystal structure—unlike Al or iron, which have a cubic crystal structure—and is more difficult to deform at room temperature (RT). For structural use in automotive applications, Mg is alloyed with additional metals to improve its strength, formability, fluidity, and corrosion resistance [6]. The use of Mg alloys in the automotive industry dates back to the mid-1930s when Volkswagen first used them to produce an engine block for the Beetle. Since then, more use cases, including power train, chassis, and body structural applications, have been demonstrated [2]. In most of these use cases, Mg alloy components were cast, leveraging the excellent fluidity property of Mg, making it possible to produce very complex thin-walled shapes. However, despite the increasing trend in Mg alloy utilization in the automotive market (based on 2022 US automotive market data [8]), low formability, low corrosion resistance, and limited overall design and manufacturing knowledge render them underutilized compared to other structural metals [9].

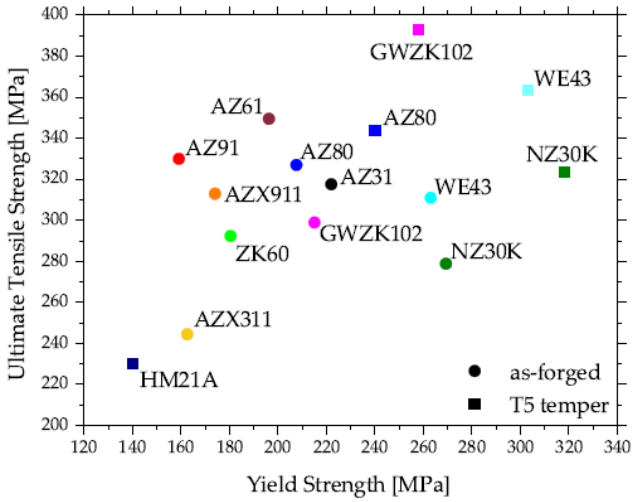


Figure 2-1: Tensile properties of various Mg alloys at RT and T5 aged condition (cooled and artificially aged only) are shown (reproduced from [5]).

As part of a broader research effort to cast-forge a fatigue-critical automotive component, the AZ80 alloy of the AZ family (Mg-Al-Zn) was selected as the material of interest as it is a versatile alloy that can be cast and forged (see Figure 2-1). Mg alloy families are represented by two letters representing the major elements and numbers showing the alloy concentration as per ASTM

Standards. Alloying elements are added to improve the mechanical properties of the metal, with aluminum being the most common alloying element that is added to improve strength, corrosion resistance, and hardness. Increasing the Al content increases the fluidity of the metal, thereby improving its castability and tensile strength; however, it reduces the material's ductility. Zinc is combined with Al to improve the RT strength of the alloy. It also increases the fluidity of the alloy. However, increasing Zinc content decreases the ductility, increases micro-porosity, and reduces corrosion resistance. Manganese is the third most common alloying element that is added, increasing yield strength and improving corrosion resistance [6], [10].

2.1.1 Casting AZ80 Alloys

The AZ series of alloy has been widely used for automotive applications and is commonly produced by gravity, low-pressure casting, high-pressure die casting (HPDC), squeeze casting, and sand or permanent mould casting process [2], [5]. The cast microstructure of commercial AZ alloys is primarily composed of α-Mg matrix and non-equilibrium eutectic, comprising of $Mg_{17}Al_{12}$ (a hard, brittle phase) and a super-saturated aluminum rich α-Mg solid solution phase [11]. The morphology of primary α-Mg, the morphology and volume fraction of the eutectic phase, and porosity defects play a critical role in the evolution of the as-cast microstructure during forging and, ultimately, on the mechanical properties of the forging.

As part of objective (i) of the SPG project, colleagues Azqadan and Uramowski investigated the influence of casting process parameters, such as cooling rate and degassing (by adding a Hexachloroethane degassing agent), on an AZ80 alloy wedge cast sample to examine the microstructure of as-cast samples and their deformation behaviour [12], [13]. Their work aimed to identify casting process parameters that would produce the optimal starting cast microstructure for forging. They noted that the microstructure of samples subjected to a high cooling rate (HCR at ~25 ºC/s) during solidification showed, on average, smaller grain sizes of 119 μm $\pm$8 μm. In comparison, samples subjected to a low cooling rate (LCR at ~3.5 ºC/s) during solidification showed average grain sizes of 201 μm $\pm$25 μm. They state that the α-Mg exhibits a fine dendritic morphology at higher cooling rates. Consequently, the eutectic $Mg_{17}Al_{12}$ that solidifies in the interdendritic regions (at the later stages of solidification) is smaller and shows either a fully divorced or partially divorced morphology. A fully divorced morphology with fine and discrete β-intermetallic particles was observed at relatively higher cooling rates, and a coarse network of β-intermetallic particles was observed at relatively lower cooling rates. At even lower cooling rates, a more cellular or globular α-Mg structure was observed, with the formation of a more continuous eutectic and precipitation of a lamellar $Mg_{17}Al_{12}$ from super-saturated α-Mg (see Figure 2-2). They noted that the addition of the degassing agent reduced gas porosity but did not change the volume fraction of eutectic $Mg_{17}Al_{12}$ in both HCR and LCR samples in comparison to their as-cast counterparts; similar grain sizes were observed in both as-cast and de-gassed HCR samples.

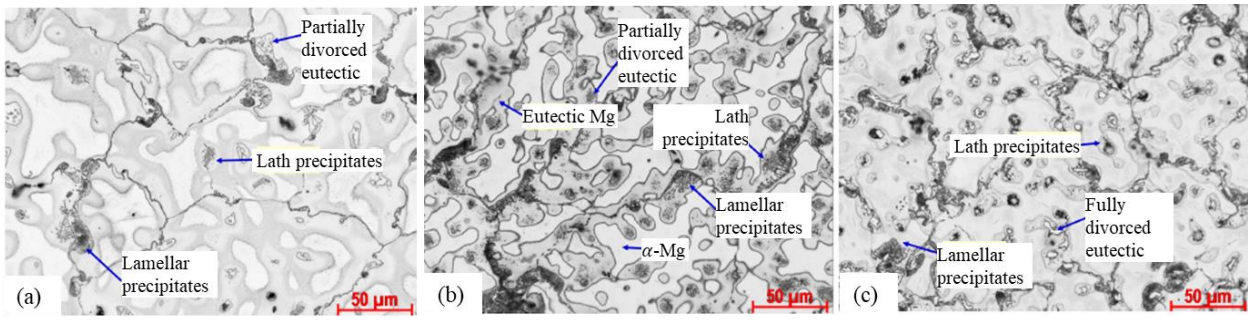


Figure 2-2: (a) Microstructure of AZ80 alloy wedge cast sample (without degassing) at LCR; (b) microstructure of the cast sample (without degassing) at HCR; (c) microstructure of the cast sample (with degassing) at HCR—the morphology difference compared to sample (b) is a result of the higher cooling rate due to faster melt pour rate (reproduced from [13]).

The deformation behaviour of HCR and LCR wedge cast samples were compared by evaluating their compressive stress-strain behaviour. HCR samples with smaller grain sizes showed higher peak compressive stresses compared to LCR samples. No significant difference in compressive stress was observed between as-cast and degassed HCR and LCR samples. The forgeability of degassed HCR and LCR samples was evaluated based on their propensity to crack during compression over a range of temperatures (250 ºC – 350 ºC) and compressive strain rates ($0.01s^{-1}$ - $1s^{-1}$). Based on these experimental results, they claim that porosity is the most significant contributing factor influencing the propensity for cracking (see Figure 2-3). They recommended that to obtain a good starting cast microstructure, forging preforms should be cast at a high cooling rate to benefit from a refined grain structure and degassing to lessen the susceptibility towards cracking, providing a wider forging process window in terms of deformation temperature and strain rate.

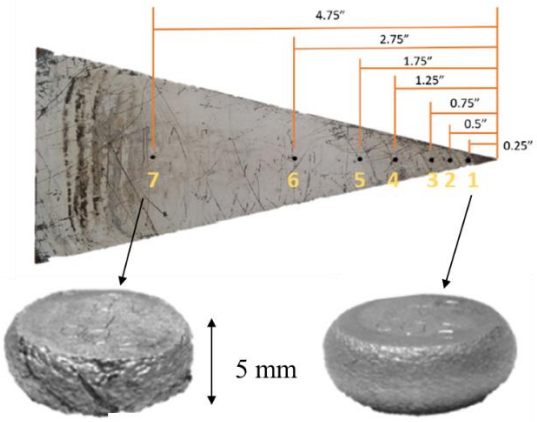


Figure 2-3: A degassed AZ80 LCR sample with visible surface cracks (left) and a degassed AZ80 HCR sample without visible surface cracks (right), after compression at a temperature of 300 ºC and a strain rate of $1s^{-1}$ to strain of 1.0; LCR sample extraction location (7), and HCR sample extraction location (1) from the wedge cast is shown (reproduced from [13]).

In the SPG project, AZ80 alloy preforms were cast using permanent mould casting (PMC), which is a process that uses reusable moulds made from steel. The main steps in the PMC process included preparing the mould, melting the metal, melt processing to remove sources of defects and refine the melt, pouring the metal into the mould, and removing the casting from the mould upon solidification (a detailed description on the casting process can be found in [14]). Considering the advantages of PMC, which include high production rates, dimensional control, and improved mechanical properties [9], [14], it was chosen as the primary method for producing the starting AZ80 alloy preforms used in die forging.

2.1.2 Forging AZ80 Alloys

Forged AZ80 alloys have increased strength and ductility compared to their as-cast counterparts, which is desirable for structural applications. In closed-die forging, a preform is deformed plastically due to compressive forces acting along the forging axis. Simultaneously, lateral material movement is constrained due to friction through a narrowing gap at the flash land, forcing material into the die cavities. From a microstructural standpoint, during deformation, crystal planes slide in relation to one another, causing an increase in the system's potential energy. As a result, the material undergoes grain refinement, thereby improving its material properties as competing mechanisms of recovery, recrystallization, and grain growth ensue to drive the system back toward equilibrium [15]. These restorative mechanisms are heavily influenced by forging process parameters such as temperature, strain rate, strain, and homogenization treatment.

In past work, Prakash et al. investigated the effects of temperature on the deformation behaviour of as-cast AZ80 alloy showed that at a forging temperature of 300 ºC, the microstructure is bimodal with a significant presence of $Mg_{17}Al_{12}$ precipitates [16]. At forging temperatures greater than 350 ºC, the bimodality of the microstructure was shown to decrease due to the dissolution of $Mg_{17}Al_{12}$ precipitates in the α-Mg matrix. At 400 ºC, the microstructure was shown to be relatively homogenous with no observable $Mg_{17}Al_{12}$ precipitates. The presence of $Mg_{17}Al_{12}$ precipitates strongly influences the material's dynamic recrystallization (DRX) behaviour during deformation—the mechanism by which new strain-free grains are formed [17], [18]. DRX can enhance the material's workability, and depending on the deformation condition, DRXed grains may be randomly oriented, attenuating texture development and anisotropy in the material.

Based on microstructure observations of compressed wedge cast samples (at strain rates of $0.01s^{-1}$ and $0.1s^{-1}$, at a temperature of 300 ºC), Uramowski noted that only a small fraction of DRX grains were visible at an average strain of 0.15, while an increasingly higher fraction of DRX grains were visible with continued compression to strains upward of 0.4 [13]. In addition, finer DRX grain sizes were observed at the highest tested strain rate of $1s^{-1}$.

A forged microstructure with fine grains can ultimately enhance the yield strength (YS) of the forged alloy based on contributions from several strengthening mechanisms, such as dislocation strengthening, precipitation strengthening, solid solution strengthening, and grain boundary strengthening. In contrast, forged components with larger grain sizes have a reduced YS [19], [20]. Based on these findings, Uramowski recommended that forging preforms should be degassed and cast at high cooling rates to benefit from a fine-grained cast microstructure that is less susceptible to cracking during forging.

The effects of homogenization were also investigated by subjecting wedge cast samples to a homogenization treatment at 420 °C for 24 hours. The homogenization treatment was found to dissolve the $Mg_{17}Al_{12}$ intermetallic compounds, leading to a higher volume fraction of DRXed grains during deformation and a more uniform microstructure. In general, a decrease in peak compressive stress was observed in the homogenized samples, suggesting that a homogenization treatment may be beneficial to reducing the overall forging load during material deformation.

In past work by Gruguć, the upper bound forging temperature of the AZ80 alloy was determined to be around 427 °C (based on the incipient melting temperature of the $Mg_{17}Al_{12}$ phase). Forging the alloy at a temperature beyond this upper-bound limit was expected to cause shrinkage, cold shut defects and segregation within the microstructure. The lower-bound forging temperature limit was determined to be around 250 °C, although the exact lower-bound limit varies depending on the forging strain rate. This limit should be determined based on defect manifestation and severity (based on macroscopic and microscopic crack observations at the edges of gleeble compression samples where edge stretch is pronounced). The flow behaviour of the AZ80 alloy is highly dependent on the strain rate. Higher strain rates result in higher flow stresses, which will result in higher forging loads. Therefore, the upper limit of the forging strain rate is a hard physical constraint that limits forging to a specific domain of the processing window, based on the forging equipment being utilized. Ultimately, Gruguć recommended that the optimal forging condition is at the lowest temperature and the highest viable forging strain rate that would result in a defect-free forging outcome (for more information, refer to [3]).

2.2 Forging Simulations

A convenient and cost-effective method for evaluating forging outcome involves using FEM simulations to predict the evolution of material during deformation. Past work on simulating component-level die forging of Mg often employed phenomenological models of plasticity that treated the metal as an isotropic body. Examples include work by Ju et al., where they carried out uniaxial compression tests from ambient to 420 ◦C at varying strain rates (0.001, 0.01 and 0.1 s−1) to obtain flow stress data to simulate forging of an AZ80 Mg alloy automotive wheel geometry; the simulation results were used to identify the stress and strain distribution throughout the component, as well as the energy propagation paths during the forging process [21]. Similarly,

Yuan et al. simulated isothermal forging of an extruded AZ80 Mg wheel at 330 °C and a rate of 1 mm/sec, where they found a significant difference in the resulting forged properties based on the location within the component, with the highest strength and elongation observed in the web region with uniform flow lines parallel with the direction of loading [22]. Chen et al. simulated isothermal forging of a complex upper receiver component from an AZ80 Mg alloy to identify optimal preform shape, processing parameters, and the forging procedure to forge a component with a refined and homogenous microstructure with good mechanical properties [23]. In general, these aforementioned works employ forging simulations treating the AZ80 alloy material as an isotropic body to predict forging outcome.

However, anisotropy is a major phenomenon that needs to be accounted for during forging, especially in wrought Mg alloys. A simplified approach for determining the anisotropic properties of wrought Mg alloys was proposed by Kobold et al. [24]. They extracted uniaxial flow stress data by conducting small-sample compression tests of extruded AZ80 alloy samples under multiple loading directions at 300, 350 and 400 °C with deformation rates ranging from 0.4 to 2.3s^{-1}, along with shear flow stress data corresponding to the same processing conditions. Using the uniaxial and shear yield data, a constant set of Hill's anisotropic coefficients (based on a simplified assumption that anisotropic properties remain constant throughout deformation) was used to define anisotropic material models of the alloy. Then, they simulated the small-sample compressions in orthogonal loading directions and concluded that simplified material models with a constant set of Hill's anisotropic yield coefficients can be used in industrial applications to predict forging outcome, striking a favourable trade-off between reliable computation and computational time. Based on their work, Yu [25] and Paracha [26] developed AZ31, ZK60, and AZ80 Mg alloy material models to simulate the forging of wrought Mg alloys during the APC research project using DEFORM®-3D [27].

2.3 Material Data

During the SPG project, a new set of wrought and as-cast AZ80 alloy material models were developed for various processing conditions. The material models were defined using three inputs: tabular flow stress data as a function of strain rate and temperature; a yield criterion (a Hill's anisotropic yield criterion for simulating wrought alloy deformation or a von Mises yield criterion for simulating as-cast alloy deformation); and a hardening rule (an isotropic hardening model was used in this work, making a simplification assumption neglecting the Bauschinger effect observed in Mg alloys).

Table 2-1: Composition of the wrought and as-cast material, as per the ASTM B91-12 standard.

Material	Al %	Zn %	Mn %	Mg %
AZ80	7.8-9.8	0.2-0.8	0.12-0.5	Bal.

The wrought material used in this work was a commercially available AZ80 alloy extruded (wrought) billet. The material was received from Luxfer MEL Technologies in the as-fabricated condition. The cast material used in this work was melted from commercial grade AZ80 cast billets obtained from Mag Specialties (see Table 2-1). All lab-scale casting and forging operations were conducted at CanmetMATERIALS (Hamilton, Canada). The methodology for extracting uniaxial and shear flow stress data and determining the friction coefficient between the workpiece and tool are described in the subsequent subsections.

2.3.1 Uniaxial Hot Compression Tests

To characterize the uniaxial stress-strain state of material during deformation, uniaxial flow stress curves of the material for a range of deformation conditions need to be provided as tabular inputs to the DEFORM®-3D [27]. In the present research, lab-scale uniaxial compression tests were conducted on cylindrical samples measuring ø10 × 15 mm, extracted from the as-received extruded billet (ø63.5 × 1 m) at a radial distance of 22 mm and in two orientations: (i) parallel and (ii) perpendicular to the prior extrusion axis (henceforth referred to as ED and TD samples, respectively, and illustrated in Figure 2-4(a)). It was necessary to obtain flow stress curves in these two orientations since the (wrought) material exhibits different flow behaviour during compression along the parallel and perpendicular directions with respect to the prior extrusion axis. The yield strengths measured under uniaxial compression along these two orientations were used to calculate Hill's yield function coefficients used in this work to define anisotropic AZ80 alloy material models [24], [28]. To define the isotropic material models, uniaxial flow stress curves were obtained from compression tests of cylindrical samples measuring ø10 × 15 mm, extracted from the locations of the wedge cast as shown in Figure 2-4(b).

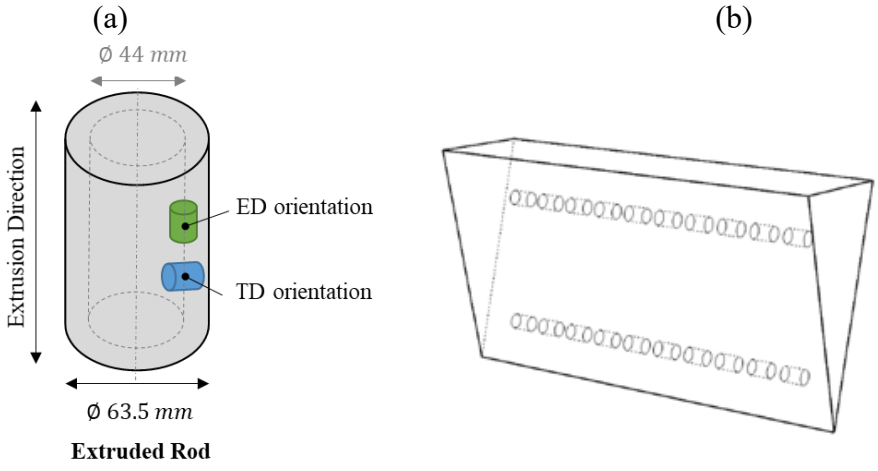


Figure 2-4: (a) Cylindrical Gleeble compression sample extraction orientations of the as-received extruded billet; (b) cylindrical Gleeble compression sample extraction orientation of the as-cast wedge (reproduced from [13]).

The uniaxial compression tests were conducted by Prakash and Uramowski at the University of Waterloo using a Gleeble® 3500 thermal mechanical simulator at temperatures of 300 °C and 400 °C, and constant true strain rates ranging from 0.001 s^{-1} to 1 s^{-1}. The deformation conditions were selected based on past work and the literature on comparable AZ alloys for high formability and an enhanced occurrence of DRX during the deformation [13], [16]. Before compressing the samples, a graphite-based lubricant was applied to their flat ends to reduce friction with the anvils. The samples were heated via electric resistive heating to the compression test temperature at a rate of 5 °C·s^{-1} and held for 60 seconds before compression to allow for the temperature to homogenize throughout the sample. The temperature was controlled during the tests based on active feedback from thermocouple wires welded at the mid-height level of the sample's surface. The samples were deformed to a true strain of 1.0 and immediately quenched in water to preserve the as-deformed microstructure. The load-stroke data obtained from the Gleeble® was converted to true stress and strain data and smoothed using a polynomial fit. The flow stress data were then corrected for friction using a procedure outlined in [15].

With wrought AZ80 alloy, material anisotropy was evident in TD samples, resulting in an elliptical cross-section perpendicular to the compression axis. On the other hand, ED samples deformed symmetrically, with a resultant circular cross-section (see Figure 2-5). The deformation behaviour of the as-cast material was similar in both ED and TD directions due to the intial random texture orientation of the as-cast microstructure. The uniaxial stress-strain curves corresponding to the extruded alloy are shown in Figure 2-6 (at both 300 °C and 400 °C), and the uniaxial stress-strain curves corresponding to the as-cast alloy are shown in Figure 2-7 (at 300 °C).

Figure 2-5: Wrought samples compressed along the ED and TD directions (as viewed from the compression direction) at 400°C at a strain rate of 0.1 s-1 to a strain of 1.0 (wrought AZ80 alloy).

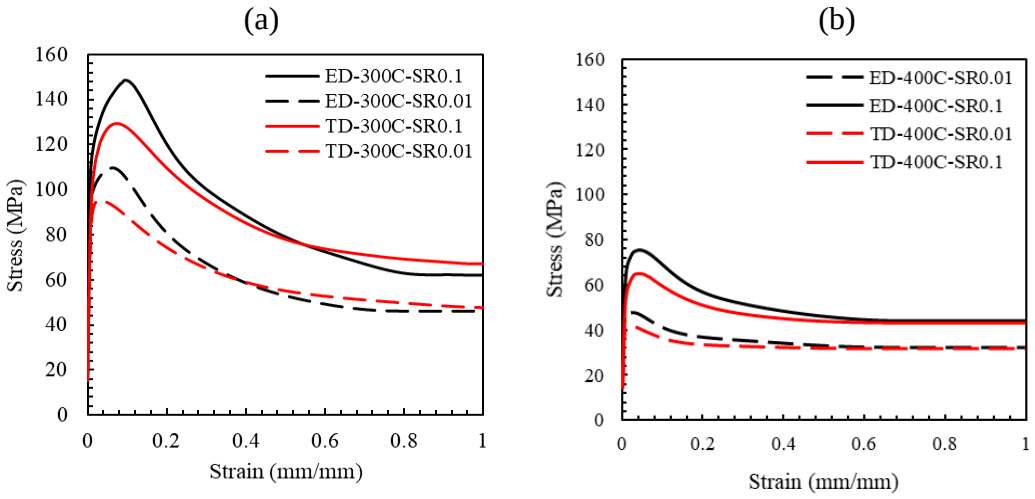


Figure 2-6: (a) Extruded AZ80 Mg alloy normal ED and TD stress-strain curves at 300C at strain rates of 0.01 s-1, and 0.1s-1; (b) ED and TD stress-strain curves at 400C at strain rates of 0.01 s-1, and 0.1s-1.

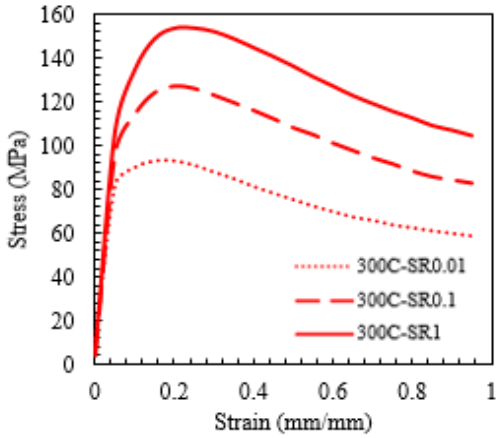


Figure 2-7: (a) As-cast AZ80 Mg alloy normal stress-strain curves at 300C at strain rates of 0.01 s^{-1} and 0.1s^{-1} (the cast billets from which the sample was extracted were cooled at a rate of 3.5 ºC/s).

2.3.2 Hot Torsion Tests

Hot torsion tests were conducted to extract shear yield strength data to derive Hill's anisotropy coefficients for the wrought alloy using a BÄHR STD 811 hot torsion plastometer at CanmetMATERIALS. The hot torsion plastometer required a minimum sample length of 170 mm, and since the diameter of the starting billet was smaller than this value, it was not possible to extract TD oriented torsion samples with a length of 170 mm from the as-received material. A workaround to this problem was developed by machining gauge sections (cylindrical in shape) measuring 30 mm in length and 10 mm in diameter along ED and TD from the as-received billet and fitting (clearance fit) them to grips made using a W1 drill rod, rendering the total length of assembled sample 170 mm. For comparison, unibody (a fully machined sample) and assembled

sample designs of torsion specimens are schematically illustrated in Figure 2-8(a). The degree to which an assembled sample replicates the torsion behaviour of a corresponding unibody sample was verified by comparing the torque-rotation angle data of the unibody and assembled ED samples at 300 °C at an equivalent strain rate of 0.01 s^{-1}. The resultant data plotted in Figure 2-8(b) shows a similar trend in torque curves of the assembled and unibody torsion samples. The slightly higher torque value observed in the assembled sample is likely due to the higher stiffness of the steel grip sections where the degree of twist was measured. The shear yield strength data used in this study (see Figure 2-9) was obtained based on torsion tests that were conducted using the assembled samples along the ED and TD orientations at 300 °C and 400 °C, at equivalent strain rates of 0.01/s and 0.1/s, to a rotation angle of 10 radians (which corresponds to an equivalent strain of 1.2 at the gauge surface). Torque, rotation angle, shear stress, and true strain data were directly obtained from the plastometer. Two tests were conducted for each deformation condition/orientation to check for data repeatability.

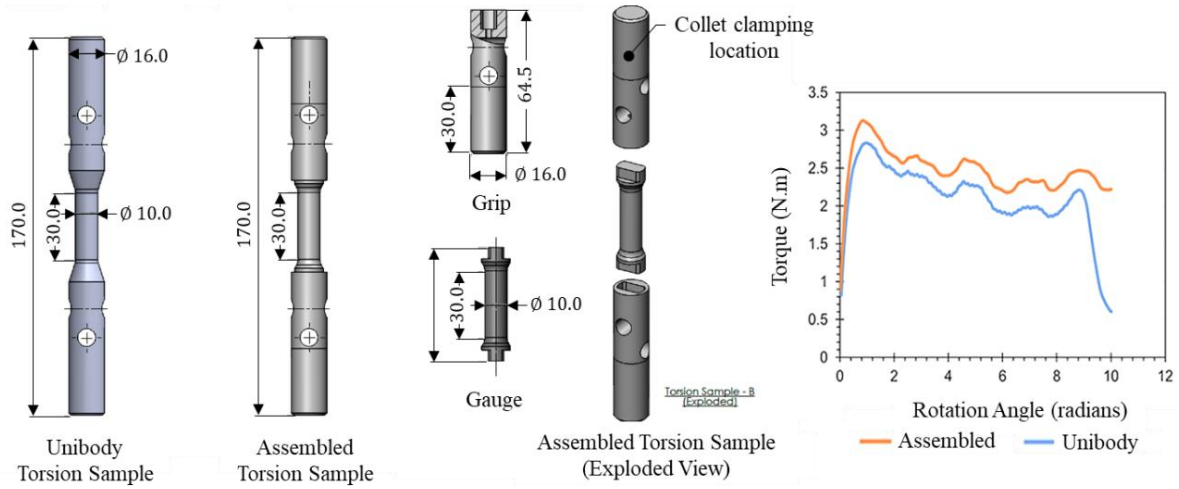


Figure 2-8: Unibody and assembled torsion samples used in the torsion tests (all dimensions are in mm), and Torque–rotation angle data from torsion tests on the unibody and assembled ED oriented samples at 300 °C and an equivalent strain rate of 0.01 s^{-1}.

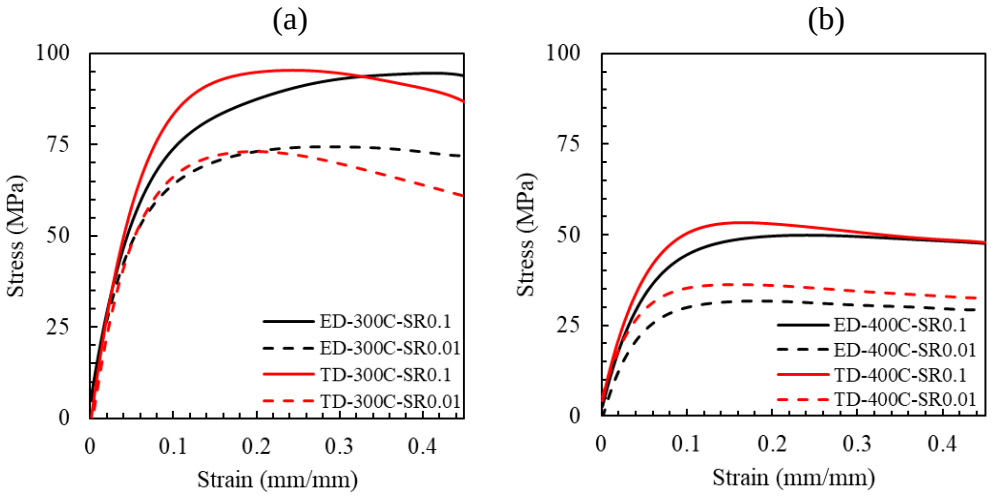


Figure 2-9: (a) Extruded AZ80 Mg alloy shear ED and TD stress-strain curves at 300C at strain rates of, 0.01 s^{-1}, and 0.1s^{-1}; (b) ED and TD stress-strain curves at 400C at strain rates of 0.01 s^{-1}, and 0.1s^{-1}

2.3.3 Ring Compression Tests

The interfacial friction coefficient is an important parameter that needs to be provided as input to the FEM model. In past work, the experimental values of friction coefficients were determined using ring compression (RC) tests that were carried out on cylindrical shell samples (with the outer diameter, inner diameter, and height in the ratio of 6:3:2). Yu reported on RC tests performed at CanmetMATERIALS on samples machined from an extruded AZ31 alloy rod [25]. The samples had an outer diameter of 88.9 mm, an inner diameter of 44.45 mm, and a height of 29.64 mm (see Figure 2-10(a)). Two lubricants, graphite and boron nitride, were tested. Prior to compression, the flat ends of the sample that made contact with the anvils were coated with the desired lubricant. The tests were conducted at 400 °C by compressing the samples to a 50% reduction in height at forging rates of 0.17 mm/s and 6.67 mm/s. The tests were repeated once to validate the data. Subsequently, the RC test was also modelled numerically using a range of friction coefficients (0 to 1, in increments of 0.1) as input. The percentage decrease in the inner diameter of the sample as a function of the percentage reduction in its height from the simulations was plotted to be compared with experimental data to determine the experimental friction coefficients. The test results indicated that the graphite lubricant results in a lower friction coefficient (of about 0.2) compared to boron nitride (see Figure 2-10(b)). Ring compression tests at a slower forging rate of 0.17 mm/s showed similar results. The graphite lubricant was used for open-die and semi-closed die forgings in this study, where a value of 0.2 was used as the friction coefficient in both 300 °C and 400 °C forging simulations. Later, a graphite-polymer lubricant was used during closed-die-forging, where a value of 0.08 was used for the friction coefficient in simulations (determined by CanmetMATERIALS).

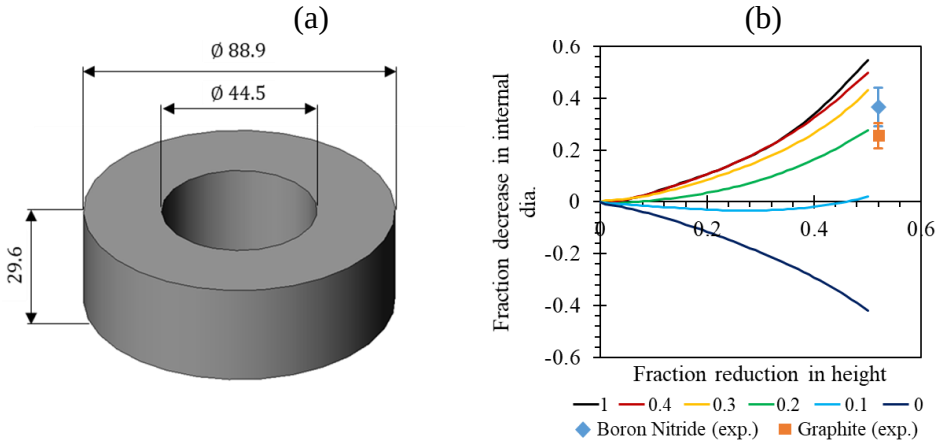


Figure 2-10: Calibration of coefficient of friction at 400 °C and 6.67 mm/s. Solid lines represent calibration curves at various friction coefficients. Diamond and square dots represent experiments.

2.4 Design Optimization

Engineering design is an iterative process in a product development workflow. The conventional engineering design process can be divided into a sequence of five phases to establish (i) requirements and specifications, (ii) a conceptual design, (iii) a preliminary design, (iv) a detailed design, and (v) a final design. Following this sequence, engineers make design decisions at each phase based on accrued knowledge, which can be relatively scant during the initial product design phase, especially during an innovative product development process. Paradoxically, the design decisions that are made during the initial phases can have the most significant influence on product development. This phenomenon is called the Design Paradox, where design freedom decreases with increasing knowledge of the problem space [29].

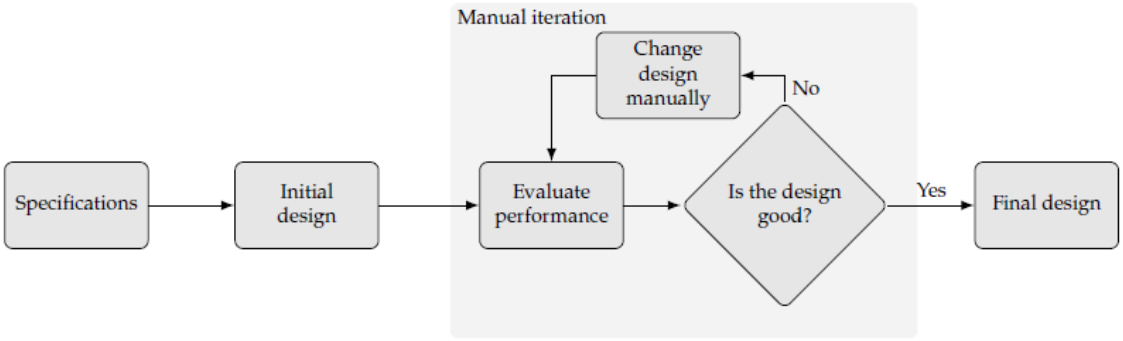


Figure 2-11: Conventional (manual) design iteration process (reproduced from [29]).

Design optimization methodologies can mitigate some of these challenges by leveraging numerical simulations and predictive models to explore design permutations, effectively expanding the space of design possibilities and enabling significant early-stage knowledge acquisition. A design can be

optimized manually in an iterative loop, where the initial design is analyzed and modified based on the evaluation of its performance (see Figure 2-11). Alternatively, the design optimization can be implemented algorithmically according to a formal problem formulation based on design variables that characterize design features, an objective(s) to be minimized, and constraints (see Figure 2-12). While design optimization is seldom a fully automated end-to-end process, still requiring human intervention at particular stages of the processes to determine specifications or to formulate the design problem, it confers several advantages over a conventional design process as designs with good performance can be obtained within a shorter time span, resulting in a reduction of cost of the overall design process; and uncertainty in design performance can be reduced more quickly.

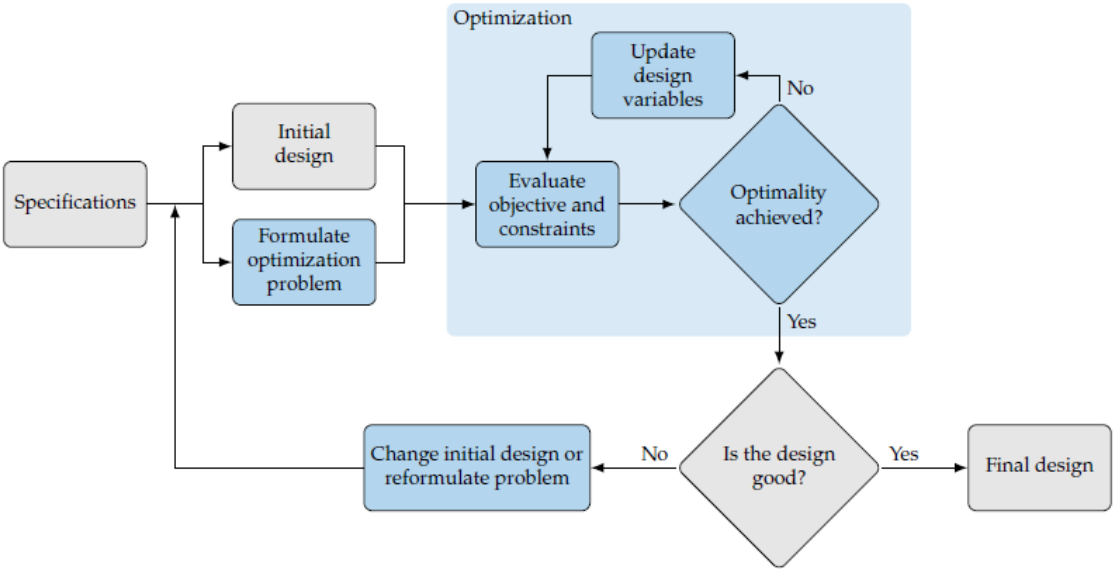


Figure 2-12: Design optimization process (reproduced from [29]).

To formulate an optimization problem, design variables, objectives, and constraints need to be established. Design variables are represented in a column vector:

$$x = [x_1, x_2, \dots, x_n], \tag{1}$$

where x denotes a different design and n defines the dimensionality of the design space. Design variables can be discrete, continuous, or mixed. The objective function F is a scalar function that measures design performance and is either minimized or maximized during optimal design search. The search space of most optimization problems is confined to a feasible region in design space which is defined using constraints. Constraints are characterized as equality constraints denoted by $h(x) = 0$, which restricts the function to a fixed value or inequality constraints, denoted by $g(x) \leq 0$, which limits the function to equal to or less than a particular value. A single-objective optimization problem based on these three aspects can be formulated according to the following general statement:

$$\begin{aligned} &\text{minimize} && F(x) \\ &\text{by varying} && a \le x_i \le b, \quad i = 1, \dots n_x \\ &\text{subject to} && g_j(x) \le 0, \quad j = 1, \dots n_g \\ & && h_k(x) = 0, \quad k = 1, \dots n_h \end{aligned} \tag{2}$$

In some optimization problems, exploring trade-offs between competing objectives is necessary. For such cases, a multi-objective optimization problem formulation can be used to quantify the trade-offs to obtain a "family" of designs for making better-informed design decisions. A multi-objective optimization problem can be formulated by modifying the single-objective statement to:

$$\text{minimize} \quad F(x) = \begin{bmatrix} F_1(x) \\ F_2(x) \\ \vdots \\ F_n(x) \end{bmatrix}, \text{where } n \ge 2 \tag{3}$$

With multiple objectives, the concept of Pareto optimality is used to resolve conflicting objectives. The concept is illustrated in Figure 2-13, which shows three designs, A, B, and C, measured in terms of two objectives, F_1 and F_2. When comparing design A with B, A can be seen as "dominating" design B based on both objective values. A design is said to dominate another if its objective values are superior to that of the other design. Comparing designs A and C, the objective value of F_1 corresponding to design A is superior to that of design B, but the objective value F_2 corresponding to design B is inferior. In this case, neither design dominates the other. A non-dominated design is not dominated by any other design point, in which case it is considered Pareto-optimal—and a set of non-dominated design points form a vector of functions referred to as the Pareto front, which describes the trade-off sensitivity between designs.

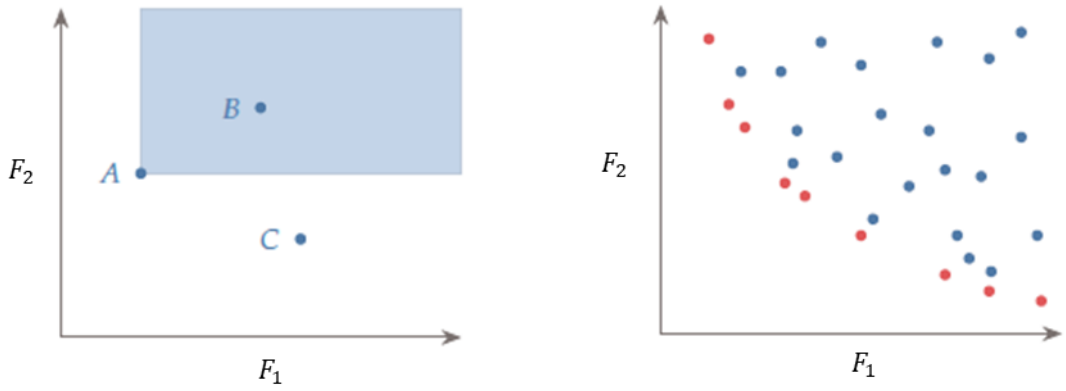


Figure 2-13:(a) Shows three design points, A, B and C plotted against two objectives F_1 and F_2; the shaded region highlights the region in objective space that is dominated by point A; (b) shows a plot of design solutions in objective space, where red and blue colours differentiate between non-dominated and dominated solutions, respectively (reproduced from [29]).

Simulation-based design optimization (SBDO) problems are often considered non-convex and non-linear as they involve large design spaces and complex, non-linear system interactions. Convexity is related to the concept of modality, where an objective or constraint function can either be unimodal, having a single minimum, or multi-modal, having multiple minima. A function is said to be convex if a line connecting any two points on the function lies above the function, else it is considered non-convex. At a minimum, optimization algorithms require models to provide derivative-free or zeroth-order information to compute objective and constraint values for a given set of design variables, in which case the optimization algorithm is called a gradient-free algorithm. On the other hand, if the optimization algorithm relies on gradient information of both the objectives and constraints, i.e., first-order information, then it is referred to as a gradient-based algorithm. Both gradient-based and gradient-free optimization strategies have been applied in the past in forging preform shape optimization problems.

In a gradient-based approach, the derivative of the objective function is defined with respect to shape variables and calculated using finite element method (FEM) simulations, and mathematical principals guide the iterative search process. Gao and Grandhi employed an approach based on sensitivity analysis to optimize the microstructure of a turbine disk while meeting die-fill constraints [30]. Using a similar approach, Zhao et al. implemented a multi-objective optimization of a preform tool design to produce forgings with improved deformation uniformity [31]. While gradient-based optimization approaches have been successfully applied in the past, there are considerable difficulties with establishing sensitivity functions and integrating them with numerical simulations. Also, the availability of objective function gradients is strictly dependent on whether they are accessible through forging simulation software, which significantly impedes the incorporation of gradient-based methods in practical preform optimization applications.

More recent work in forging preform shape optimization looks to gradient-free methods to drive shape optimization. One such method is Bi-direction Evolutionary Structural Optimization (BESO), a structural optimization method. The BESO approach follows a process where elements from the preform boundary surface are added or deleted iteratively according to a criterion based on hydrostatic stress and strain components in the loading direction to distribute material effectively within the design envelope. BESO performs a local search in shape space based on heuristic principles to guide the iterative search process. A local search involves taking small steps toward the optimum, starting from a single point. Conversely, in a global search, the whole design space is searched in an attempt to find the global optimum. Typically, in structural optimization algorithms, a seed geometry is predefined to confine the search space of designs. BESO methods evaluate objective and constraint functions based on the performance responses obtained directly from numerical models, making them entirely simulation-based optimization approaches suitable if a good seed geometry can be defined and if the number of forging simulations that can be run is limited. An example of a BESO application includes work by Shao et al., where they optimized a

three-dimensional preform for an aerofoil forging operation for which they achieved improved material flow and deformation uniformity, eliminated high local strains, achieved a complete die-filling, and reduced material waste [32]. Structural optimization is a local search method that attempts to find a local optimum in high-dimensional shape space. Alternatively, shapes can be parametrized to perform a global search in a lower-dimensional parameter space (i.e., design space). Gradient-free hybrid approaches using surrogate models have been commonly used for such optimization tasks. Thiyagarajan and Grandhi developed a reduced-order modelling technique for three-dimensional shape optimization where primitive shapes were linearly combined to obtain resultant preform shapes [33]. This technique was applied to generate models based on a design of experiments (DOE). During the offline phase of the optimization process, three-dimensional models were sampled using a Latin Hypercube Sampling (LHS) method. These were then subjected to FEM simulations to generate data to construct a response surface model (RSM) that describes forging quality in terms of strain variance and underfill. For creating design alternatives, the reduced-order modelling technique they used offers an elegant solution for combining a set of basis shapes to generate resultant shapes. In a similar approach, Chan et al. introduced an integrated FEM and Artificial Neural Network (ANN) methodology to approximate the design performance of a metal forming tool as a function of billet design parameters [34]. They used FEM data to train ANNs to construct performance surface graphs that can later be used to identify optimal billet designs based on maximum deformation load and maximum von Mises effective stresses on the tool. In a more recent study, Shao et al. optimized the airfoil preform shape to enhance material flow during deformation using an evolutionary optimization approach known as Multi-Island Genetic Algorithm (MIGA) with an integrated Radial Basis Function (RBF) model [35]. An integrated FEM and ANN methodology offers a way to make coarse evaluations of design performances without having to rely on simulations continuously. Non-gradient-based evolutionary multi-optimization techniques such as MIGA [36], Non-dominating Sorting Genetic Algorithm-II (NSGA-II) [37], or Strenth-Pareto Evolutionary Algorithm 2 (SPEAII) [38] are proven multi-objective optimization algorithms that can be used in conjunction with ANNs to improve the efficiency of the optimization process.

2.4.1 CAD Modelling Techniques

Oftentimes in gradient-free shape optimization methods, the dimensionality of the design space is defined by the number of design variables that are used to parameterize a shape. In such cases, it is advantageous to reduce the number of design variables as much as possible while still maintaining the shape variety—to avoid a phenomenon known as the curse of dimensionality. The curse of dimensionality arises when the dimensionality of the data increases, and the availability of data becomes sparse, i.e., the distance between data points becomes large, making it challenging to find meaningful patterns or relationships in the data.

Common parameterization techniques that are used in shape optimization include radial basis function (RBF) morphing, free-form deformation (FFD), and parametric modelling. Both RBF and FFD are a class of CAD-free modelling techniques involving surface morphing of mesh geometry. In FFD parametrization, the spatial coordinates defining a geometry are embedded in a volume spline that can be modified based on its control points. FFD relies on an algebraic scheme to modify the underlying geometry mesh based on the perturbation of control points, which has the added benefit of enabling the calculation of volume control point derivatives with respect to shape variables of the underlying geometry mesh [39]. The downside of this approach is that the entire underlying geometry mesh is modified when volume control points are perturbed (see Figure 2-15). As a workaround, multiple FFD volumes that only influence particular regions of the geometry can be defined. However, this increases the complexity of the algorithm as it would necessitate modifications to the algorithm to resolve topological inconsistencies between adjacent FFD volumes [40]. In addition, it is not trivial to incorporate design envelopes that constrain the morphable space.

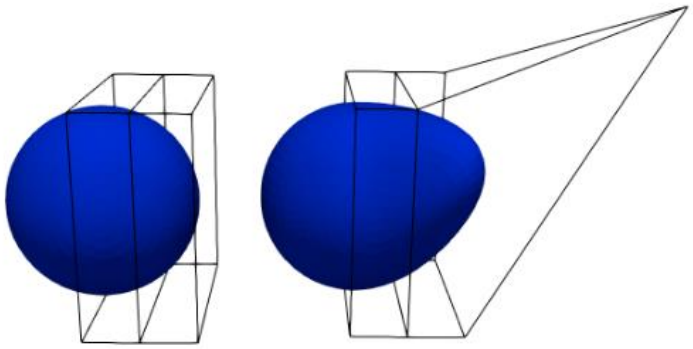


Figure 2-14: FFD-based parameterization of an aircraft wing (reproduced from [41]).

RBF morphing uses radial basis kernel functions that are placed at boundary points to morph the surface of underlying geometry mesh based on boundary point displacements. The morphing problem can be viewed as an interpolation problem where a function $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ exactly interpolates the prescribed boundary displacements and smoothly interpolates these displacements onto the underlying mesh (see Figure 2-15). The RBF morphing function formulation is based on a linear combination of weighted kernel functions that are located at boundary points and a linear polynomial to guarantee precision. While smooth morphing of the underlying mesh makes RBF morphing a powerful technique, it can be computationally expensive as the computation of weight coefficients requires a very large system of equations to be solved, as each mesh vertex is factored into the computation. Also, similar to FFD, implementing RBF mesh morphing while enforcing design envelope constraints is not trivial [42].

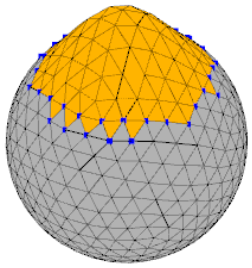


Figure 2-15: An implementation of the RBF morphing algorithm on a spherical mesh.

Most Computer-Aided Design (CAD) models used in design engineering workflows use conventional modelling techniques such as solid or surface modelling, where primitive three-dimensional geometry or curve geometry is modified incrementally using Boolean operations or transformation operations such as extrusion, revolution, and sweeping to reflect design intent. These conventional CAD-based modelling techniques are relatively straightforward. However, they demand time and attention, making it difficult to generate design variations in a design optimization loop efficiently. Alternatively, under a parametric CAD-based modelling paradigm, feature-based parametric CAD systems that can automatically reflect design changes based on numerical inputs can be developed. These systems enable rapid design generation without the need for involved remodelling efforts and by virtue of their construction, enable design constraints to be directly embedded in the model. Feature-based parametric CAD models rely on a feature tree to encode procedural generation logic that governs the generative behaviour of curve and surface geometry, giving the freedom to explicitly define the range of unique possibilities of design variation that can be generated [43], [44]. Parametric CAD systems are propagation-based systems [45] that compute new design solutions whenever input variables are modified, triggering the automatic reconstruction of a new CAD model [46].

In the present work, sketches were defined in Rhinoceros® [47], and feature trees (operational graphs) were defined in Grasshopper (a graphical algorithm editor that is integrated with Rhinoceros® [47]) to carry out the procedural generation logic. Grasshopper provides an easy-to-use design environment that is accessible to non-programmers, making it an ideal environment for prototyping and developing procedural generation logic by connecting various components (i.e., nodes in the operational graph) that operate on numerical and geometric data. Additionally, Rhinoceros® [47] provides access to Rhinocommon, a powerful software development kit (SDK) that is cross-platform and can be used with various programming languages such as C# and Python (see Figure 2-16).

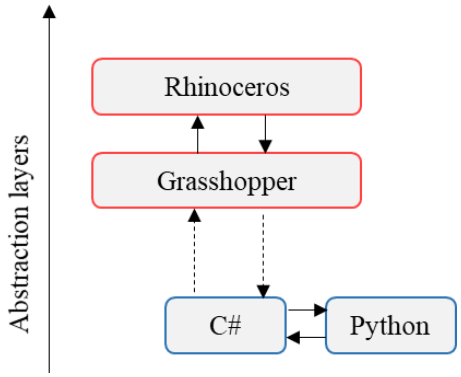


Figure 2-16: A high-level flow chart of information between the different code libraries. Rhino is used as the user interface, taking commands and displaying results (adapted from [48]).

2.4.2 Multi-objective Evolutionary Algorithm

Evolutionary optimization algorithms such as MIGA, NSGA2 and SPEA2 are multi-objective optimization algorithms. These algorithms extend the application of the generic single-objective genetic algorithm (GA) to multi-objective problems. A GA is a stochastic metaheuristic for search and optimization problems that mimic the evolutionary process described in the Darwinian theory of natural selection, where the fittest members of a population survive environmental pressures and so have a greater chance of reproducing and propagating genetic information to the subsequent generation [49], [50]. Multi-objective optimization algorithms such as SPEA2 simultaneously optimize multiple objective functions for a given design vector, such that optimization converges to a set of Pareto-optimal solutions—a set of non-dominated solutions in objective space that define a boundary beyond which any further improvement of any one objective will worsen at least one other objective. The SPEA2 algorithm was used in this research to iteratively solve for non-dominated solutions.

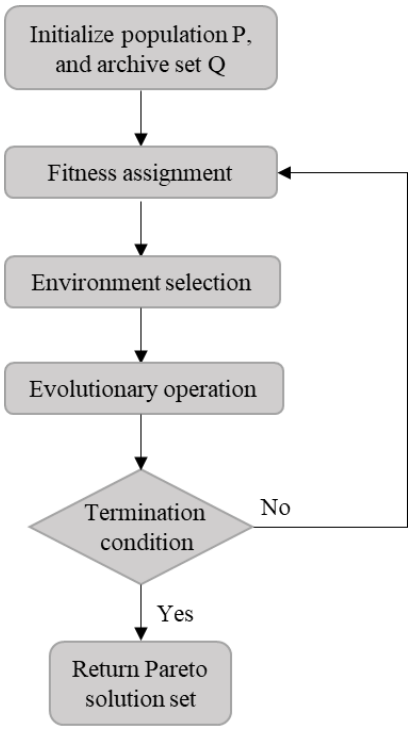


Figure 2-17: SPEA2 flow diagram.

The SPEA2 algorithm comprises of five main steps (see Figure 2-17, and for a more detailed description of the algorithm and equations, refer to [51]):

1. Initialization of the population: a population of designs, P_t of size, N is initialized with a random subset of design vectors referred to as chromosomes (subscript t refers to the generation number). A design vector can be represented as a binary vector where design parameters or features are coded as a string of bits (i.e., a string of 1s and 0s) or as a vector of integers or continuous variables. From an algorithmic standpoint, the population can be represented as an $(N \times M)$ two-dimensional array, with N representing the number of individuals in the population and M representing the characteristic genes or features of the chromosome. In addition, in SPEA2, an empty archive set, A_t of size $\overline{N}$ is also initialized. The archive set is used to store and update non-dominated solutions that are discovered during the optimization process that ultimately represent the Pareto-optimal set of solutions.

2. Fitness evaluation: unlike in GA, in SPEA2, the overall fitness values of individuals in the population are calculated based on multiple objective functions and the domination order of solutions to determine which solutions are to be pushed into the archive set. To calculate the overall fitness, the strength $S(i)$ of each member is first calculated to determine the number of solutions in the population that each member dominates. For every solution $i \in$

P_t, the strength $S(i)$ is calculated as the number of solutions j that i dominates in the population (Equation 4):

$$S(i) = |\{j | j \in P_t + A_t \wedge j \succ i\}| \tag{4}$$

The strength is used to calculate raw fitness, $R(i)$, of individuals (Equation 5). The raw fitness is calculated as the summation of strengths of solutions j, dominating solution i. Thus, a non-dominated solution would have a raw fitness value, $R(i) = 0$. A non-zero raw fitness value is indicative of a solution being dominated by solutions in one or more objective dimensions.

$$R(i) = \sum_{j \in P_t + A_t, i > j} S(j) \tag{5}$$

The raw fitness and density information of a solution is considered in the selection mechanism for copying solutions to the archive set. As a final step towards calculating fitness, additional density information is considered to distinguish between solutions which may have identical raw fitness values. In the SPEA2 algorithm, a *k*-th nearest neighbour method is applied to assign a density value to a solution as a decreasing function of the distance to the *k*-th nearest solution in the population (Equation 6). $D(i)$ captures the density of a solution.

$$D(i) = \frac{1}{\sigma_i^k + 2} \tag{6}$$

where the *k*-th element in the sorted list (sorted according to increasing order), σ_i^k is the distance to the *k*-th nearest point, and $k = (N + \bar{N})^{1/2}$. Finally, the fitness of the solution can be calculated based on raw fitness and density to determine Pareto dominance (Equation 7).

$$F(i) = R(i) + D(i) \tag{7}$$

The set of solutions that have a fitness value $F(i) < 1$ are non-dominated solutions and are copied to the archive set. To maintain an archive set of the specified size, a distance-based (i.e., Euclidean distance in objective space) truncation method is used to maintain the diversity of solutions.

3. Selection: the selection process is implemented to identify the individuals with the highest fitness scores that will reproduce to create the subsequent generation (based on an elitist selection scheme). This is a crucial step, as the selection strategy directly influences the convergence rate. Many selection strategies exist, such as Fitness Proportionate Selection, which subdivides into Roulette Wheel Selection (RWS) and Stochastic Universal Sampling (SUS) [52]. In the case of an RWS strategy, selection pressure in the form of a probabilistic

condition is imposed to ensure individuals in the population with higher finesses are selected more often for reproduction.

4. Reproduction: during reproduction, a new generation of 'offspring' chromosomes is produced based on the exchange of genetic information between two parent chromosomes of the previous generation (gene cross-over) and gene mutation. From an algorithmic standpoint, cross-over operations involve swapping a sequence of bits between two arrays, while a mutation operation would involve switching a binary state of a bit at a random location on an 'offspring' chromosome based on a mutation probability.

5. Termination criteria: if a satisfactory solution is obtained given the convergence criteria, the GA search loop is terminated. If not, the search process is executed once more until the convergence criteria are met.

2.5 Machine Learning

At a minimum, evolutionary optimization algorithms require models to provide zeroth-order information to perform evolutionary computations. Data-driven models such as artificial neural networks (ANNs) are considered universal approximators well suited for such tasks because they can discern complex non-linear patterns in data [52]. This subsection aims to provide a foundational knowledge base on key concepts relating to artificial neural networks. An artificial neural network consists of multiple interconnected neurons that learn complex patterns in an observed dataset by learning an approximation function to map input data to output data. In this context, a neuron is a mathematical function constituting the basic building block of a neural network. A neuron transforms a linear combination of input signals and a bias term by a non-linear activation function to produce an output. This computation can be expressed according to the following equation:

$$y = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right), \tag{8}$$

where x_i is an input signal, w_i is a learnable weight parameter, b is a bias term, and σ is a non-linear activation function. The input signal is a feature vector, which is an ordered list of numerical values that characterize the properties of a data sample. The learnable weight parameters are numerical values that are tuned during the training or optimization process, along with the bias term. Lastly, the activation function introduces non-linearity to the output of the neuron [53].

The activation function introduces non-linearity to the neural computations, which enables the learning of complex patterns in the data. Non-linear activation functions produce non-linear

decision boundaries, which enable a sequence of layers in the network to build increasingly more complex representations of the input data. The choice of activation function is problem-specific and is based on several factors, such as:

- Network behaviour: whether the network is attempting to capture linear or non-linear observations in the data. Linear activation functions can also be used as activation functions. However, this will only enable the network to capture linear trends in the data. On the other hand, non-linear activation functions can capture parabolic and tangential trends as well.

- Range: the output range of the activation function. Depending on the learning task, it may be desirable to map the summed inputs to a neuron to a particular range such as [0,1], [-1,1], or [0, inf] (commonly used activation functions that perform such mappings include the sigmoid (Sigmoid), hyperbolic tangent (Tanh), or rectified linear unit (ReLU) activation functions, respectively). The choice of activation function has a direct impact on learning stability in gradient-based learning approaches.

- Derivative: the learning speed of a network is influenced by the derivatives of the selected activation function.

A feed-forward network is a basic artificial neural network that recruits multiple neurons to form a fully connected graph where information is passed sequentially through layers (a cluster of neurons that form a multivariate function): (i) an input layer that accepts a feature vector representation of data samples, (ii) hidden layer(s) that increase the degree of freedom or complexity of the network, and (iii) an output layer. The computation performed by an MLP with a single hidden layer can be expressed in the following matrix form equation:

$$y = \sigma_2(W_2\sigma_1(W_1x + b_1) + b_2), \tag{9}$$

where W_i is the weight matrix, and b_i is the bias vector of the *i*-th layer. In fact, according to the Universal Approximation Theory proposed by Cybenko [54], an ANN with a single hidden layer (with enough hidden neurons) and non-linear activation functions can, in theory, be a universal approximator of any non-linear relationship on a compact subset $\mathbb{R}^n$.

The learnable weight parameters, W_i are optimized during training. The most commonly used optimization algorithm for training ANNs is the backpropagation algorithm [55], which is the de facto choice for finding the minima of typically highly non-convex cost function landscapes of neural networks. The cost function, f in this context, can be mean-squared-error (MSE), mean-absolute-error (MAE), Cross entropy (CE), etc., which measures the distance in metric space between the network prediction $\hat{y}^i$, and the ground truth, y^i. The derivative of this error with

respect to every weight parameter in the network (or bias) is then backpropagated through the network to update the current state of network parameters according to the following update rule:

$$\begin{aligned} W_i &:= W_i - \eta \nabla f, \\ b_i &:= b_i - \eta \nabla f, \end{aligned} \tag{10}$$

where ∇f is the derivative of the error with respect to network parameters, and η is the step size or learning rate. Over several cycles (or epochs), the iterative process of forward and backward information passes updates the state of the network parameters such that the deviation between the prediction and the ground truth is minimized.

3. Simulation of Forging Processes

3.1 Introduction

This section describes the development of forging FEM models of the AZ80 alloy for various industrial processing conditions using DEFORM®-3D [27]. The present research builds on the foundational studies of the past [24]–[26], employing a phenomenological model to characterize the hot deformation behaviour of the wrought and as-cast forms of the alloy under isothermal deformation conditions. New material models were developed in this work to predict isotropic and anisotropic deformation of the AZ80 alloy materials procured from Luxfer MEL Technologies (extruded billet) and Mag Specialties (as-cast billets). The forging models were developed systematically, progressing from lab-scale uniaxial compression tests of small cylindrical samples (see Section 2.3) to forgings of intermediate size and complexity using semi-closed-die or closed-die forging. At each step, simulations and corresponding lab-scale forgings were compared for validation. The approach presented here can be extended to other Mg alloys and forging operations involving complex component geometries.

3.2 Methods

3.2.1 Semi-closed Die Forging

Semi-closed-die forging (henceforth referred to as SF or "Coin" forging) experiments were conducted on cylindrical samples measuring ø 20 mm x 45 mm, extracted at a radial distance of 40 mm from the as-received extruded billets, with their longitudinal axis parallel to the prior extrusion axis. Forgings were conducted at CanmetMATERIALS using a 110-ton hydraulic press. The die geometry and the sample orientation within the die are presented in Figure 3-1(a), and images of forged coin samples are presented in Figure 3-1(b). As illustrated in Figure 3-1(a), the upper die has a flat surface, while the lower die has a T-shaped cross-section with a flat web section and a central rib with an inclined step feature (designed by Yu [25]). The rib and web were designed to have a thickness of 3 mm and an inner corner radius of 10 mm to produce coin forgings with an underfill in the rib so that it could be measured to evaluate the prediction accuracy of anisotropic flow behaviour. The sample was oriented within the press such that the radial direction was aligned with the direction of the press stroke (the forging direction was aligned with the TD direction of the billet). The billet and tooling were heated separately to 400°C. The billet was held at 400 °C for 1 hour prior to forging to ensure all thermal gradients had decayed. Forging experiments were carried out at average ram speeds of 0.04 mm/s and 0.4 mm/s under isothermal conditions and then air-cooled.

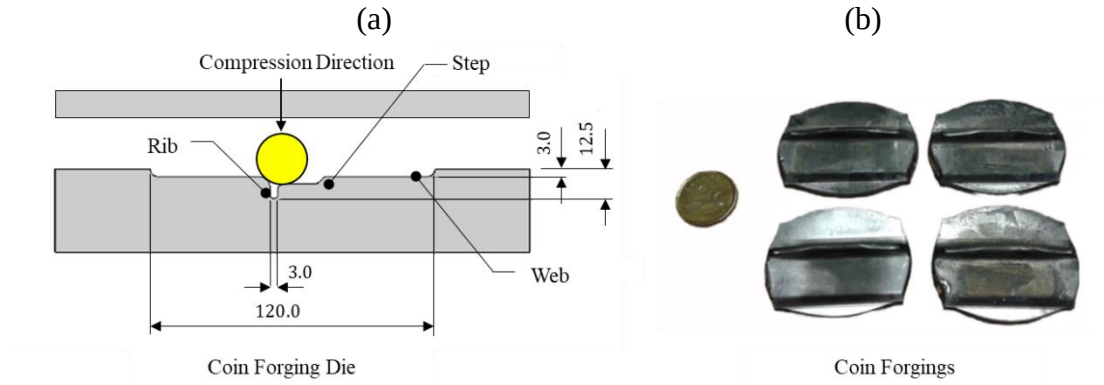


Figure 3-1: (a) Tooling and sample used in the semi-closed-die forging experiments (all dimensions are in mm); (b) images of as-forged "Coin" samples

3.2.2 Closed-Die Forging

Closed-die forgings were conducted using both wrought and as-cast forms of the AZ80 alloy. Anisotropic material models were used to simulate wrought material deformation, while isotropic material models were used to simulate as-cast material deformation. Although texture development is expected during deformation due to the realignment of grains into preferred orientations (in both forms of the alloy), the consequent change in material anisotropy was not modelled.

Cast-forging experiments of I-beams were conducted to validate the isotropic material models used in preform optimization tasks. The I-beam geometry was selected as it consists of an H-section, a standard feature common in structural components. It was expected that due to multi-directional metal flow and varying levels of deformation, distinct regions of the forging would exhibit varying levels of effective plastic strain and diverse microstructure developments, suitable for the material characterization task as per objective (i) of the SPG project. The AZ80 alloy cylindrical billets (ø 63.5mm x 70mm) were cast using a PMC process, which involved purging the casting mould with a CO_2 + 0.5% SF_6 protective gas prior to moulding (refer to [15] for information on tool design). Then, the molten metal was degassed with a Hexachloroethane degassing agent at a concentration of 100g/60kg prior to pouring. The castings were cooled at various rates (6, 4, 3.5, and 1.5 ºC/s). The material model of the as-cast AZ80 alloy discussed in the remainder of this work was defined using stress-strain curves corresponding to a starting material that was cooled at a rate of 3.5°C/s (see Figure 3-2).

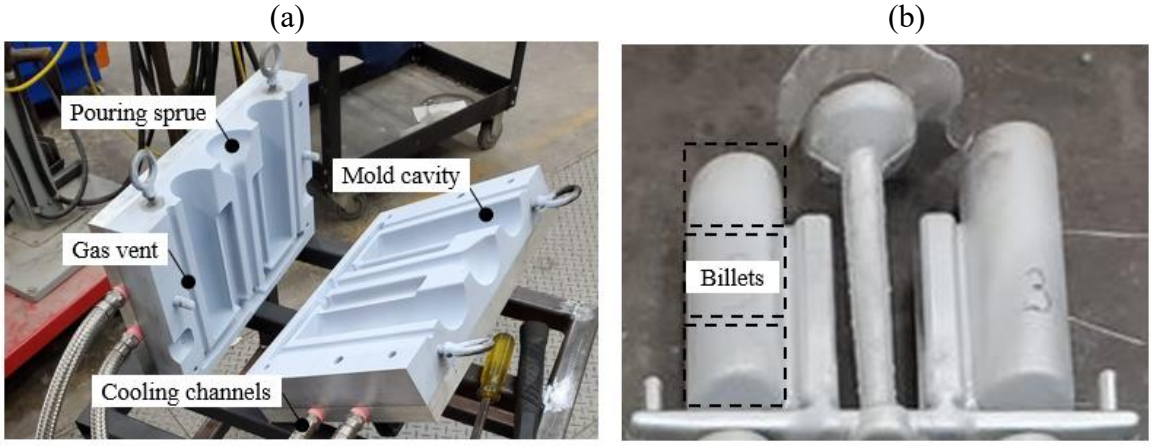


Figure 3-2: (a) PMC mould; (b) as-cast cylindrical billets (courtesy of CanmetMATERIALS)

An hour prior to forging, the cast cylindrical billets were heated in a furnace to elevated temperature ranges between 250–350 ºC and held at that temperature (at the forging temperature of the process) to ensure all thermal gradients had decayed. I-beam forgings discussed in this work were produced under isothermal forging conditions at 300 ºC using an Interlaken 500-ton press (see Figure 3-3). During lab-scale I-beam forging, a constant downward press velocity of approximately 5.6 mm/s was recorded, where a deflection of approximately 45 mm was observed in the insulation laminates beneath the top and bottom dies. To account for this deflection, a press velocity of 2.5 mm/s was used in forging simulations.

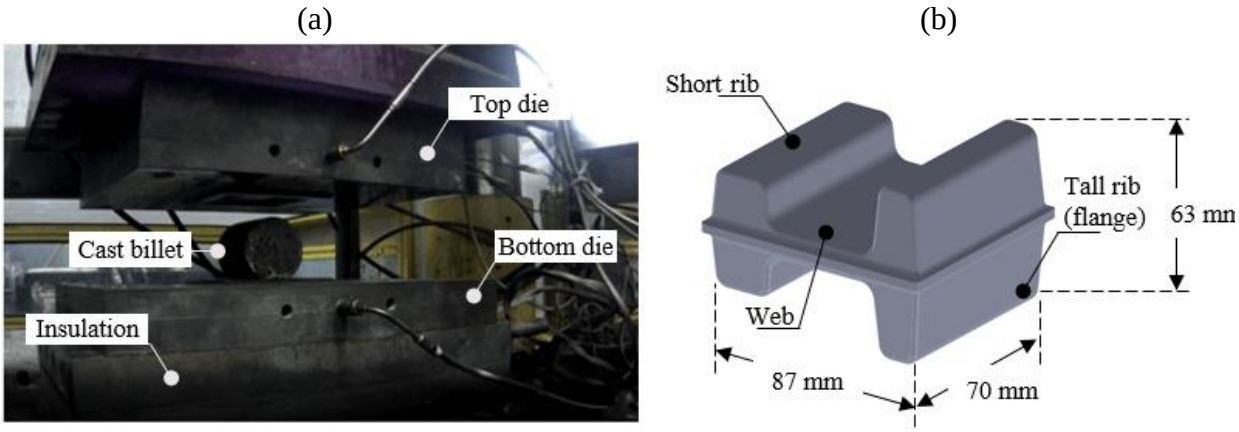


Figure 3-3: (a) I-beam forging tool used in closed-die forging experiments (courtesy of CanmetMATERIALS); (b) the I-beam geometry

Closed-die forgings of the front lower control arms (FLCA) were carried out to develop the knowledgebase and methodology for manufacturing high-performance, lightweight Magnesium alloy suspension components as SPG project objective (iii). This chapter presents wrought AZ80 Alloy isothermal forging of FLCAs at 300ºC under a multi-step forging process. The steps involved bending an extruded AZ80 billet (ø 63.5mm x 780mm) and then flattening it at an elevated temperature using a pair of anvils to obtain the preform for the subsequent forging step.

Figure 3-4 shows the forging tool and the preform shape (refer to [56] for information on tool design). The forging trials were conducted at CanmetMATERIALS, utilizing a Macrodyne 1500 Ton press. The preform and tooling were heated separately to 300°C. The preform was held at 300 °C for 1 hour prior to forging to ensure all thermal gradients had decayed. Forging experiments were carried out according to a variable ram speed schedule where the initial speed of 4mm/s was stepped down to 0.4 mm/s and then to 0.04 mm/s towards the end of the forging cycle (a variable press cycle was used to deform the material at a slower strain rate towards the end, to lessen the forging load in order to fill the die).

(a) (b)

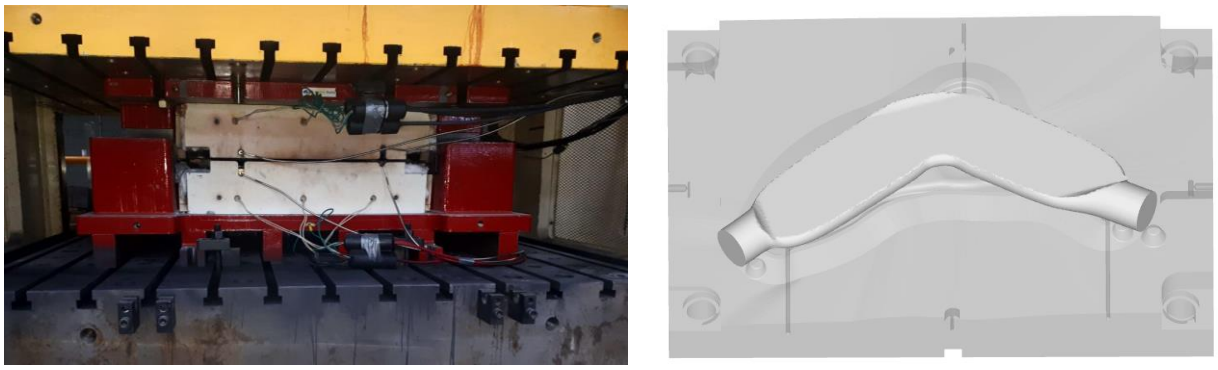


Figure 3-4: (a) The control arm forging tool used in closed-die forging experiments; (b) the wrought AZ80 alloy preform used to forge the control arm—obtained following bending and flattening operations

3.3 Finite Element Model Formulation

3.3.1 Viscoplastic Material Formulation

DEFORM®-3D [27] was used to simulate the forging process based on a rigid-viscoplastic FEM formulation and updated Lagrangian approach. The rigid-viscoplastic formulation describes the material response during the deformation according to the following governing equations [57], [58]:

Equilibrium equation:

$$\sigma_{ij,j} = 0 \tag{11}$$

Compatibility and incompressibility condition:

$$\dot{\epsilon}_{ij} = \frac{1}{2}\left(v_{i,j} + v_{j,i}\right) \tag{12}$$

Constitutive relation:

$$\sigma'_{ij} = \frac{2\bar{\sigma}}{3\dot{\bar{\epsilon}}}\dot{\epsilon}_{ij}, \qquad \bar{\sigma} = \sqrt{\frac{3}{2}\sigma'_{ij}\sigma'_{ij}}\ , \qquad \dot{\bar{\epsilon}} = \sqrt{\frac{2}{3}\dot{\epsilon}_{ij}\dot{\epsilon}_{ij}} \tag{13}$$

Boundary conditions:

$$\sigma_{ij}n_j = \bar{F}_i \text{ on } S_F, \qquad v_i = v_i^* \text{ on } S_u \tag{14}$$

where $\sigma_{ij}, \dot{\epsilon}_{ij}$, and v_i are stress, strain rate, and velocity components, respectively. $\sigma'_{ij} = \sigma_{ij} - \delta_{ij}\sigma_{kk}/3$ defines the deviatoric stress components, where δ_{ij} is Kronecker delta. $\bar{F}_i$ denotes the force per unit area (traction) on the surface S_F, and n_j is the unit normal vector. v_i^* denotes deformation velocity over the boundary surface S_u. The equilibrium equation is solved according to the variational principle:

$$\int_V \bar{\sigma}\delta\dot{\bar{\epsilon}}dV + K\int_V \dot{\epsilon}_v\delta\dot{\epsilon}_v dV - \int_S F_i\delta v_i dS = 0 \tag{15}$$

where $\bar{\sigma}$, $\dot{\bar{\epsilon}}$, and $\dot{\epsilon}_v$ are the effective stress, effective strain rate, and volumetric strain rate, respectively. *K* is a positive constant that penalizes volume change. The variational form presented in Equation 15 is converted into a set of non-linear algebraic equations and solved using the direct iteration method and conjugate gradient solver in DEFORM®-3D [27].

3.3.2 Hill's Yield Criterion

During plastic deformation, anisotropy develops due to the re-orientation of the crystal grains. Increasing strain leads to a gradual development of preferred crystal orientation, leading to the development of material anisotropy [59]. Preferred crystal orientation is not the only cause of anisotropy. Lattice defects such as laminar inclusions and cavities can also lead to the development of material anisotropy [60]. Simulation of lattice defects would require the use of physics-based constitutive models, which can be difficult to implement in commercial FEM software. Therefore, phenomenological constitutive models were used to characterize the material deformation behaviour in this work, where the anisotropic material models were defined using a Hill's (1948) anisotropic yield criterion.

The anisotropic yield criterion proposed by R. Hill [28] builds on von Mises criterion of yielding:

$$f(\sigma_{ij}) = Constant \tag{16}$$

where f is the plastic potential (i.e., a function of stress tensor components characterizing the state of the material), which describes the macroscopic and non-uniform behaviour of anisotropic metals during plastic deformation. The assumption that the sole cause of anisotropy is preferred grain orientation underpins the formulation of this criterion. The strain-increment tensor $d\epsilon_{ij}$ that is used in conjunction with the yield criterion is given below:

$$d\epsilon_{ij} = \frac{\partial f}{\partial \sigma_{ij}} d\lambda \tag{17}$$

where $d\lambda$ is a positive scalar factor. Hill proposed a quadratic anisotropic yield law, a plastic potential f based on the assumptions that yielding is not influenced by hydrostatic stress and that tensile and compressive yield stresses are equivalent in each principal direction. The Hill's (1948) yield criterion has the following form:

$$2f \equiv F(\sigma_{22} - \sigma_{33})^2 + G(\sigma_{33} - \sigma_{11})^2 + H(\sigma_{11} - \sigma_{22})^2 + 2L{\sigma_{23}}^2 + 2M{\sigma_{31}}^2 + 2N{\sigma_{12}}^2 \tag{18}$$

where F,G,H,L,M, N are constants that characterize the current state of anisotropy of the material. In this work, the six independent anisotropic coefficients in equation 18 were determined experimentally using yield stresses that act along the principal axes of anisotropy. Coefficients F, G, and H were determined using the uniaxial yield stresses $\bar{\sigma}_{11}$, $\bar{\sigma}_{22}$, and $\bar{\sigma}_{33}$ obtained from the uniaxial compression tests of the extruded alloy samples. Coefficients L, M, and N were determined using the shear yield stresses $\bar{\sigma}_{12}$, $\bar{\sigma}_{23}$, and $\bar{\sigma}_{31}$, obtained from the torsion tests of the extruded alloy samples. The Hill's anisotropic coefficients were normalized into dimensionless constants according to Equation 19, prior to being assigned in the DEFORM®-3D material model.

$$\left.\begin{aligned} F &= \frac{\sigma_m^2}{2}\left(\frac{1}{\bar{\sigma}_{22}^2} + \frac{1}{\bar{\sigma}_{33}^2} - \frac{1}{\bar{\sigma}_{11}^2}\right) \\ G &= \frac{\sigma_m^2}{2}\left(\frac{1}{\bar{\sigma}_{33}^2} + \frac{1}{\bar{\sigma}_{11}^2} - \frac{1}{\bar{\sigma}_{22}^2}\right) \\ H &= \frac{\sigma_m^2}{2}\left(\frac{1}{\bar{\sigma}_{11}^2} + \frac{1}{\bar{\sigma}_{22}^2} - \frac{1}{\bar{\sigma}_{33}^2}\right) \\ L = \frac{\sigma_m^2}{2\bar{\sigma}_{23}^2}\ ; M &= \frac{\sigma_m^2}{2\bar{\sigma}_{31}^2}\ ;\ N = \frac{\sigma_m^2}{2\bar{\sigma}_{12}^2} \end{aligned}\right\} \tag{19}$$

where the squared average value of the stress components σ_m^2 was expressed as,

$$\sigma_m^2 = \frac{1}{3}(\bar{\sigma}_{11}^2 + \bar{\sigma}_{22}^2 + \bar{\sigma}_{33}^2) \tag{20}$$

In reality, the principal axes of anisotropy vary throughout a continuum and also from its initial configuration during deformation. While the use of Hill's yield criterion with a constant set of coefficients does not account for this, i.e., for the changes in anisotropy due to increasing grain distortion and rotation of axes of anisotropy, it is used in this work to enable reasonably accurate computation of anisotropic material flow.

3.3.3 Derivation of Hill's Anisotropic Coefficients

Hill's yield function coefficients (F,G,H,L,M, and N) were calculated using uniaxial and shear yield stresses obtained from the extruded samples in the ED and TD directions. In the simulation environment, the principal axes of anisotropy (i.e., unit normal vectors $e_{11}, e_{22},$ and e_{33} of the material model) were aligned with the X, Y, and Z axes of the coordinate system (see Figure 3-5). In this work, two distinct sets of Hill's coefficients were defined to simulate ED and TD compression of the extruded alloy since the compression direction was fixed along the X-axis (see Figure 3-6).

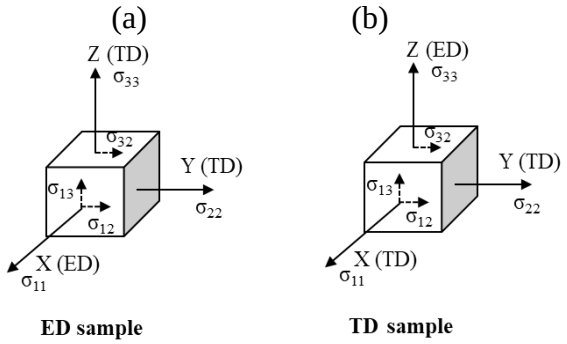


Figure 3-5: Cauchy stress tenor orientations that are assumed in the material models used in ED sample compression (a) and TD sample compression (b) along the X-axis

To fully characterize the state of anisotropy, six independent yield stresses, $\bar{\sigma}_{11}, \bar{\sigma}_{22}, \bar{\sigma}_{33}, \bar{\sigma}_{12}, \bar{\sigma}_{13},$ and $\bar{\sigma}_{23}$ are needed. Each of these yield stress values can be defined as the average value of stresses ranging from the yield point to some logarithmic strain value, φ_T or φ_L (see Equation 21) which is calculated based on true measurements $D_0, D_T,$ and D_L of the compressed sample footprint (see Figure 3-6). Using the experimental measurements of material extension along the major and minor axes of the footprint, it is possible to obtain ratios of yield stress values that provide a good approximation of the overall anisotropic flow behaviour (for the uniaxial loading conditions). The underlying assumption here is that the principal stress and strain-increment directions coincide with the principal axes of anisotropy and that the directions of the principal axes remain invariant throughout the deformation. Kobold et. al., [24] showed in their work that yield stress values acquired in this way lead to good approximations of anisotropic flow behaviour. This work applies the same procedure to obtain Hill's coefficients. Then, the Hill's coefficients were calibrated by adjusting the yield stress ratios to minimize the error between the simulated and experimental

footprint dimensions. Table 3-1 and Table 3-2 provide the final sets of yield stresses that were used to calculate Hill's coefficients used in forging simulations at 300 and 400 °C at strain rats of $0.01s^{-1}$ and $0.1s^{-1}$.

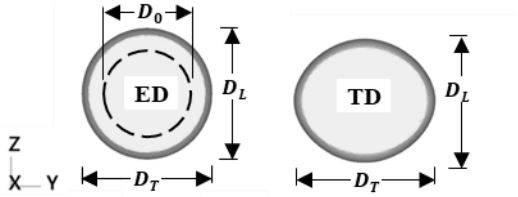


Figure 3-6: Compressed samples along the ED and TD directions (as viewed from the compression direction) at 400°C at a strain rate of 0.1 s^{-1} to a strain of 1.0. The ED oriented sample shows a circular profile, while the TD oriented sample shows an elliptic profile due to material flow anisotropy. D_T measures the material expansion along the major axis, while D_L measures the material expansion along the minor axis, and D_0 refers to the initial diameter of the sample.

$$\varphi_L = ln\left(\frac{D_L}{D_0}\right) ; \varphi_T = ln\left(\frac{D_T}{D_0}\right) \tag{21}$$

Table 3-1: ED compression – Normal and shear yield stresses along different orthogonal directions

Temp. (°C)	**Strain rate (1/s)**	**Yield Stress (MPa)**						
		$\bar{\sigma}_{11}$	$\bar{\sigma}_{22}$	$\bar{\sigma}_{33}$	$\bar{\sigma}_{12}$	$\bar{\sigma}_{13}$	$\bar{\sigma}_{23}$	σ_{m}
400	0.01	37	34	34	29	29	20	35
	0.1	75	64	64	46	46	37	68
300	0.01	79	73	73	45	45	46	75
	0.1	113	104	104	67	67	56	107

Table 3-2: TD compression - Normal and shear yield stresses along different orthogonal directions

Temp. (°C)	**Strain rate (1/s)**	**Yield Stress (MPa)**						
		$\bar{\sigma}_{11}$	$\bar{\sigma}_{22}$	$\bar{\sigma}_{33}$	$\bar{\sigma}_{12}$	$\bar{\sigma}_{13}$	$\bar{\sigma}_{23}$	σ_{m}
400	0.01	42	42	48	20	29	29	44
	0.1	65	71	78	37	45	45	68
300	0.01	63	69	82	46	45	45	72
	0.1	90	98	117	56	67	67	102

The yield stress values shown in Table 3-1 and Table 3-2 were used with Equations 19 and 20 to derive Hill's coefficients to describe the deformation behaviour of wrought alloys along the ED and TD directions, respectively. The final set of Hill's coefficients are shown in Table 3-3 and Table 3-4.

Table 3-3: Hill's coefficients used in ED compression simulations

Temp. (°C)	Strain rate (1/s)	F	G	H	L	M	N
400°C	0.01	0.61	0.45	0.45	1.63	0.75	0.75
	0.1	0.72	0.41	0.41	1.70	1.11	1.11
300°C	0.01	0.61	0.45	0.45	1.30	1.39	1.39
	0.1	0.61	0.45	0.45	1.83	1.27	1.27

Table 3-4: Hill's coefficients used in TD compression simulations

Temp. (°C)	Strain rate (1/s)	F	G	H	L	M	N
400°C	0.01	0.29	0.55	0.66	1.55	1.43	2.42
	0.1	0.29	0.47	0.63	1.14	1.14	1.69
300°C	0.01	0.28	0.49	0.81	1.28	1.22	1.22
	0.1	0.28	0.48	0.81	1.15	1.15	1.66

3.4 Results and Discussion

The DEFORM®-3D [27] pre-processor was used to set up Gleeble®, "Coin", I-beam, and FLCA forging simulations. The die and workpiece geometries for all simulations were generated using SOLIDWORKS® [61] and imported into the pre-processor environment. The workpiece was considered rigid viscoplastic and meshed using tetrahedral elements with an average element size of 2 mm. The top and bottom dies were defined as rigid bodies. Isothermal forging conditions were assumed during deformation (in the experimental setup, cartridge heaters were embedded into the forging tool to maintain a fixed temperature condition), and both the workpiece and forging dies were prescribed a constant uniform temperature. To prevent volume loss of the workpiece during remeshing, target volume compensation was enabled. To account for the frictional forces that arise during contact between the workpiece and die, a coulomb friction function was used with a friction coefficient of 0.2 or 0.08, based on the use of either a graphite or graphite/polymer lubricant. Material properties for isothermal forging conditions at 300°C and 400°C were defined using corresponding uniaxial flow curves at strain rates of 0.01 s^{-1}, 0.1 s^{-1}, and 1.0 s^{-1}, and the derived Hills coefficients were used to define the anisotropic material models (the choice of coefficients were determined based on the forging temperature and the average deformation strain rate). Other input parameters used during the simulation setup are provided in Table 3-5, along with the computation time required to run each simulation on an i7-7820X CPU.

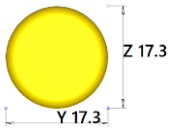
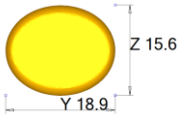
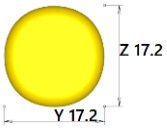
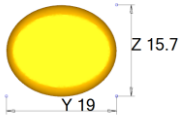
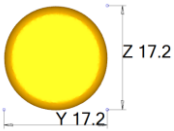
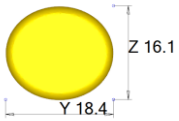
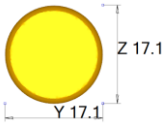
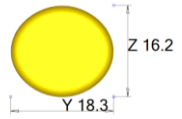
Table 3-5: DEFORM®-3D model input parameters and computation time

Input Parameters				
	Gleeble®	**Coin**	**I-beam**	**Control Arm**
Remeshing Algorithm	Global	Global	Global	Global
Remeshing Parameter	0.7	0.7	0.7	0.7
Average Element Size	2 mm	2 mm	2 mm	2 mm
Number of Elements	50k	150k	150k	300k
Step Definition	0.1 mm/step	0.1 mm/step	0.1 mm/step	0.1 mm/step
Target Volume	Active in FEM & meshing	Active in FEM & meshing	Active in FEM & meshing	Active in FEM & meshing
Die-Workpiece Friction Coefficient	0.2	0.2	0.08	0.08
Die Velocity	Strain rates: $0.1s^{-1}$, $0.01s^{-1}$	Average die speed: 0.4 mm/s (SF14), 0.04 mm/s (SF6)	Average die speed: 2.5 mm/s	Variable
Forging Temperature	300°C/400°C	400°C	300°C	300°C
Material Model	Anisotropic	Anisotropic	Isotropic	Anisotropic
Internal Heat Generation	Not modelled	Not modelled	Not modelled	Not modelled
Die-Workpiece Heat Generation	Not modelled	Not modelled	Not modelled	Not modelled
Computation Time				
Total simulation time	15 Mins	2.5 Hrs	3 Hrs	7 Hrs

3.4.1 Gleeble Compression Simulations

Gleeble® samples extracted from the extruded alloy were compressed in ED and TD directions at processing conditions of 300 °C and 400 °C, at strain rates of 0.01 s^{-1} and 0.1 s^{-1}. The flow stress data obtained from the lab-scale compression tests were provided as tabular inputs to the DEFORM®-3D material model. Distinct material models were defined based on the processing conditions, and each material model was assigned a unique set of Hill's coefficients from Table 3-3 or Table 3-4. Gleeble® compression simulations were used to validate the anisotropic material models and calibrate coefficients. Table 3-6 gives a comparison of simulation and experimental results. The major and minor axes dimensions of compressed samples were compared along with compression loads. In general, simulation predictions in TD and ED compression conditions show a good degree of dimensional accuracy relative to the true measurements of the major and minor axes dimensions with around 1-2% error.

Table 3-6: Major and minor axes dimensions of simulations are shown along with measured dimensions of compressed samples

Temp (°C)	Strain Rate (0.1/s)								Strain Rate (0.01/s)							
	ED				TD				ED				TD			
300																
Dim (mm)	Maj	Min	e^{Maj}	e^{Min}	Maj	Min	e^{Maj}	e^{Min}	Maj	Min	e^{Maj}	e^{Min}	Maj	Min	e^{Maj}	e^{Min}
	17.3	17.3	0%	0%	19.3	15.7	2%	1%	17.3	17.3	1%	1%	19.0	15.8	0%	1%
Load (kN)	Sim	Exp	e		Sim	Exp	e		Sim	Exp	e		Sim	Exp	e	
	16.9	14.3	18%		15.6	15.2	3%		10.7	10.7	0%		10.5	10.9	4%	
400																
Dim (mm)	Maj	Min	e^{Maj}	e^{Min}	Maj	Min	e^{Maj}	e^{Min}	Maj	Min	e^{Maj}	e^{Min}	Maj	Min	e^{Maj}	e^{Min}
	17.3	17.3	1%	1%	18.5	16.2	1%	1%	17.2	17.2	1%	1%	18.3	16.2	0%	0%
Load (kN)	Sim	Exp	e		Sim	Exp	e		Sim	Exp	e		Sim	Exp	e	
	10.6	10.4	2%		10.3	9.9	4%		8.9	7.4	14%		6.6	7.2	9%	

Plastic deformation occurs by dislocation slip or twinning in HCP metals along certain crystallographic planes and directions as a result of shear strains [62]. With continued deformation, an accumulation of strain history causes different slip systems to activate, which leads to changes in the strain path. Changes in strain path induce rotation in crystal grains towards a preferred orientation, resulting in the development of texture and anisotropy [63]. Ideally, with the change in the principal directions of anisotropy, Hill's coefficients should be updated to account for the evolving state of the material at each simulation step. However, the results in Table 3-6 show that

the cumulative effect of anisotropy can be approximated reasonably well using a fixed set of Hill's coefficients that are calibrated to minimize the error between simulated and experimental major and minor axes footprint dimensions.

3.4.2 Coin Forging Simulations

This section presents the experimental results and simulation predictions of forgings with a slightly more complex sample geometry, that of the "Coin" forging shown in Figure 3-7. These forgings were carried out at 400ºC, at average forging speeds of 0.04 mm/s and 0.4 mm/s. The corresponding samples are henceforth referred to as SF6 and SF14, respectively. The "Coin" forging simulation study aimed to investigate how well anisotropic deformation under multi-axial loading can be predicted using a constant set of Hill's coefficients.

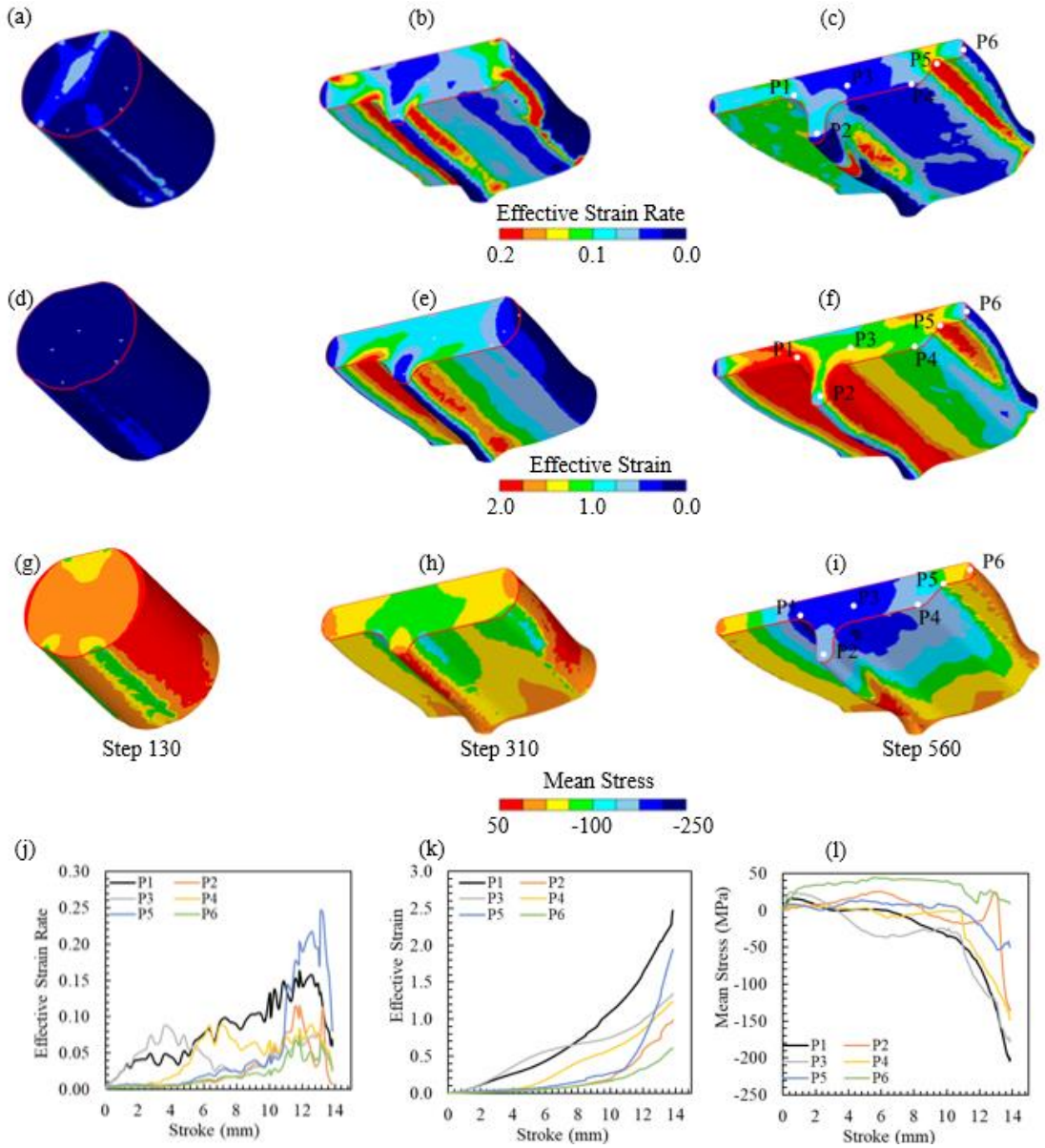


Figure 3-7: (a)-(c) effective strain rate evolution at intermediate steps of the forging simulation SF14 (average forging speeds of 0.4 mm/s at 400ºC); (d)-(f) effective strain evolution at intermediate steps; (g)-(i) mean stress evolution at intermediate steps; (j) effective strain rate evolution of points P1-P6; (k) effective strain evolution of points P1-P6; (l) mean stress evolution of points P1-P6.

Figure 3-7 shows the intermediate stages of plastic deformation during Coin forging, along with effective strain, effective strain rate, and mean stress plots corresponding to material points P1-P6 that were tracked throughout the deformation. Referring to Figure 3-7(a)-(c), high strain rates can be observed along the length of the web-to-rib transition region, as well as along the length of the step-to-web transition region (refer to Figure 3-1). A sudden change in effective strain rate, effective strain and mean stress can be seen in Figure 3-7(j),(k), and (l) at around a stroke distance of 11 mm in the region surrounding material point P5, which is situated along the length of the step at the bottom die. It is likely that the material in this region experiences a notable change in

microstructure and preferred texture orientation and intensity as a result of multi-axial loading and changes in strain path (see Figure 3-7(d)-(f)). Under these conditions, the material may undergo rapid strain hardening or softening, which would have a great degree of influence on the deformation behaviour. This change in material deformation behaviour can also be seen in the load-stroke plots in Figure 3-8, which show a divergence between predicted and observed forging loads in both SF14 and SF6 conditions at around a stroke distance of 11mm.

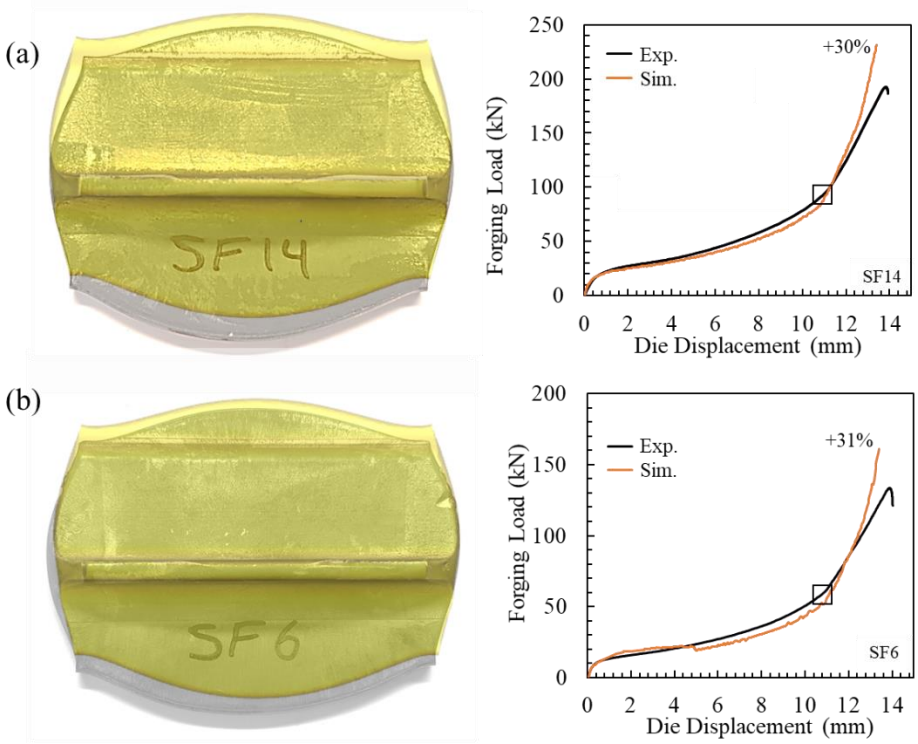


Figure 3-8: Simulation and experimental forging shape profile overlays (the two images are aligned using the rib feature as a reference) and corresponding forging load vs. die displacement plots; (a) SF14: average forging speeds of 0.4 mm/s at 400ºC; (b) SF6: average forging speeds of 0.04 mm/s at 400ºC.

In the simulation corresponding to the SF14 forging, the predicted and experimental forging loads were 250 kN and 190 kN (a 30% error in load prediction). The predicted contour shape of the SF6 forging is in better agreement with the experiment. The predicted and experimental loads were 170kN and 130 kN (a load prediction error of 31%). The difference between the predicted contour shape and the experiment is most pronounced in the SF14 forging (see Figure 3-8). Table 3-7 shows the measured values of the indicated indices.

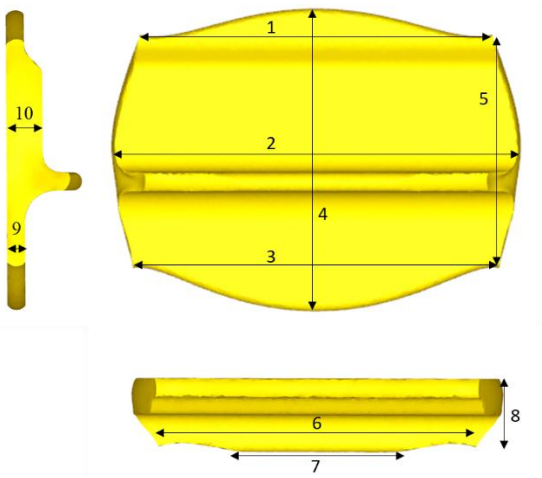


Figure 3-9: Coin forging measurement indices.

Table 3-7: Simulation and experimental Coin forging measurements.

Coin (AZ80)		Avg. Speed (mm/s)	Measurements									
			1	2	3	4	5	6	7	8	9	10
SF14	Exp.	0. 4	56.2	64.4	55.6	48.9	35.5	49.6	14.3	11.3	3.4	6.5
	Sim.		55.6	64.1	57.4	47.3	37.1	56.3	19.3	11.7	3.4	6.4
	e		1%	1%	3%	3%	4%	12%	26%	3%	0%	1%
SF6	Exp.	0.04	53.1	62.6	55.2	49.9	35.9	50.1	15.4	11.2	3.5	6.6
	Sim.		55.4	64.6	57.5	47.5	36.7	53.9	20.2	12.2	3.4	6.4
	e%		4%	3%	4%	5%	2%	7%	24%	9%	0%	1%

Uncertainty in the shear flow curve data undoubtedly has an impact on the simulation results. In addition, the true response of the material under multi-axial loading is not captured in the flow stress data acquired from uniaxial compression tests. This is expected to be the most prominent factor which contributes to errors in load and flow prediction. Ideally, to account for the re-orientation of the principal axes of anisotropy, the stress components in Equation 19 should be transformed at each simulation step to reflect the yield stress magnitude in the direction of the rotated principal axes of anisotropy. A deeper discussion on the revolution of principal axes of anisotropy during deformation can be found in [63]. Also, referring to Figure 3-7(g)-(i), the mean stress, which indicates the overall state of stress of a mesh element, can be seen varying throughout the forging, with red and orange contours indicating elements under a tensile stress state. It is important to note here that the tension-compression asymmetry in wrought Mg alloys was not taken into consideration in this work, resulting in inadequate predictions of asymmetric yielding. An extension of the Hill's yield criterion, which accounts for the tension-compression asymmetry, can be found in [64]. In addition, a combination of isotropic and kinematic hardening rules may be adopted to formulate the evolution of the yield surface to more accurately characterize the

hardening behaviour of Mg during deformation (refer to [65] for more information). Non-uniform friction conditions can also contribute to flow prediction error, being influenced by lubrication, forging pressure, surface topography, tool-workpiece interface temperature, deformation velocity, and material properties [66]. While a constant friction condition was assumed in the simulations, in practice, friction conditions are likely to vary during deformation with the tool-workpiece interface temperature and contact pressure. The tool-workpiece interface temperature has a direct influence on the lubricant viscosity, while contact pressure has an influence on surface expansion. If these conditions vary greatly during deformation, uneven frictional stresses and uneven spreading of lubricant can be expected [67],[68].

3.4.3 I-Beam Forging Simulation

This section presents the lab-scale and simulation results of an I-beam forging using an as-cast billet forged at 300 ºC at an average forging speed of 2.5 mm/s. The forging outcome of the lab-scale sample ICF11 was compared with the corresponding forging simulation to validate the isotropic material model, which was used in all subsequent preform design optimization tasks discussed in this thesis.

Figure 3-10: I-beam simulation and laboratory-scale forging flash comparisons.

As shown in Figure 3-10, the simulated flash profile does not reflect that of the lab-scale forging. This was expected since the material model used in the simulation does not account for the evolution of anisotropy, especially at the stage where the material is expanding into the flash land. In terms of forging load, it is interesting to note that between press cycle times of 5 and 15 seconds, there is a noticeable under-prediction of load in the forging simulation. The higher loads observed in the lab-scale forging suggest of complex microstructural and (or) friction-related conditions, which are not captured in the simulation. Despite this, after 15 seconds, when the material front has expanded into the flash, forging loads begin to track one another closely until 18 seconds into the press cycle. Beyond this point, there is a reduction in forging speed in the lab-scale experiment,

which is not reflected in the forging simulation (see Figure 3-11(a)). In general, higher effective plastic strains were observed at the flash exit and at the web-rib transition surfaces (see Figure 3-11(b)). The strain rates that were observed during forging (within the impression) ranged between roughly 0 and $1s^{-1}$ (see Figure 3-11(c)); the material response to loading at strain rates within this range is linearly interpolated by the FEM solver based on the flow stress data (corresponding to strain rates of $0.01s^{-1}$, $0.1s^{-1}$, and $1s^{-1}$) used in the material model.

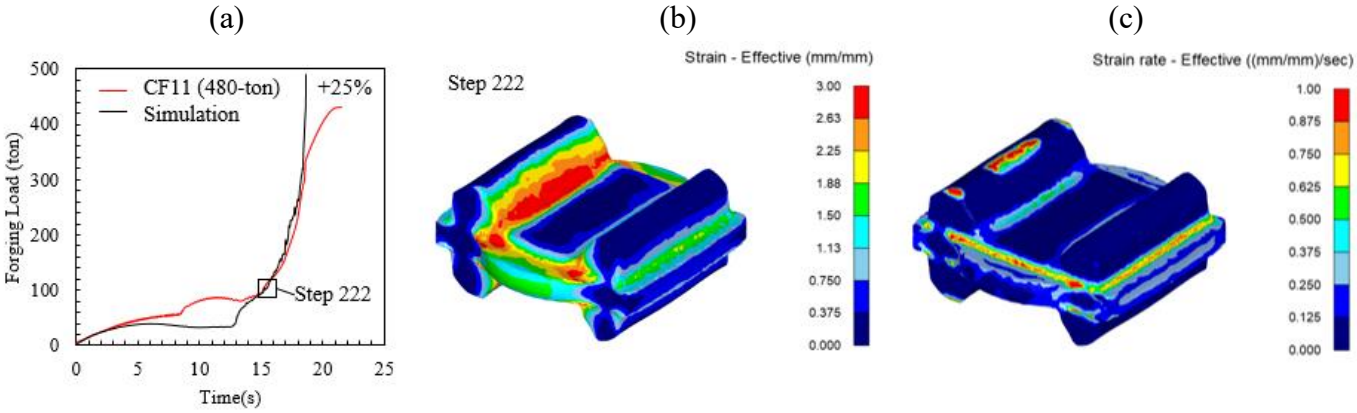


Figure 3-11: (a) I-beam simulation vs. laboratory-scale forging loads and (b) effective strain and (c) effective strain rate ~18s into the forging cycle are shown.

3.4.4 Control Arm Forging Simulation

This section presents forging results of the AZ80 alloy control arm (FLCA) at 300ºC under a variable die speed schedule starting at 4mm/s and stepping down to 0.04 mm/s and then 0.004 mm/s towards the end of the forging (see Figure 3-12). Note that the lab-scale forging results discussed in this section correspond to the forged component, CA48. It was forged under a multi-step forging process where a wrought AZ80 alloy billet was pre-formed following a series of bending and flattening operations and then hot forged (see Appendix A.1). The input parameters to the FEM model are presented in Table 3-5. The purpose of this forging simulation study was to use the validated anisotropic material model to assess the forgeability of the control arm, and to investigate the formation of forging defects that were observed in lab-scale forgings, which ultimately motivated the exploration of cast-forging as an alternate manufacturing method for producing structural Mg components. The material model was defined using flow stresses of the material in the TD direction, along with a fixed set of Hill's coefficients corresponding to a strain rate of 0.1/s (in general, strain rates observed during forging ranged between 0 and 0.1 s^{-1}).

Figure 3-12: (a) Control arm forging simulation; (b) laboratory-scale forging of CA48 with a 2 mm underfill at the bushing in the indicated region.

Figure 3-12(a) shows the simulated forging outcome, and Figure 3-12(b) shows the corresponding lab-scale forging of CA48. In the lab-scale forging, the control arm was near-fully formed except at the bushing, where there was about a 2 mm underfill (as a result of reaching the max press load capacity). Comparing the simulation and experiment, the flash contour predicted by the simulation is noticeably different. This was expected, considering the material that is forced into the flash is under tension, and the Hill's coefficients were calibrated to predict material anisotropy under compression. Additionally, the use of an isotropic hardening rule, assumptions of no internal heat generation, unchanged friction conditions throughout deformation, and the use of a constant set of anisotropic coefficients throughout the forging cycle all contribute to the deviation in metal flow that is observed in the simulation.

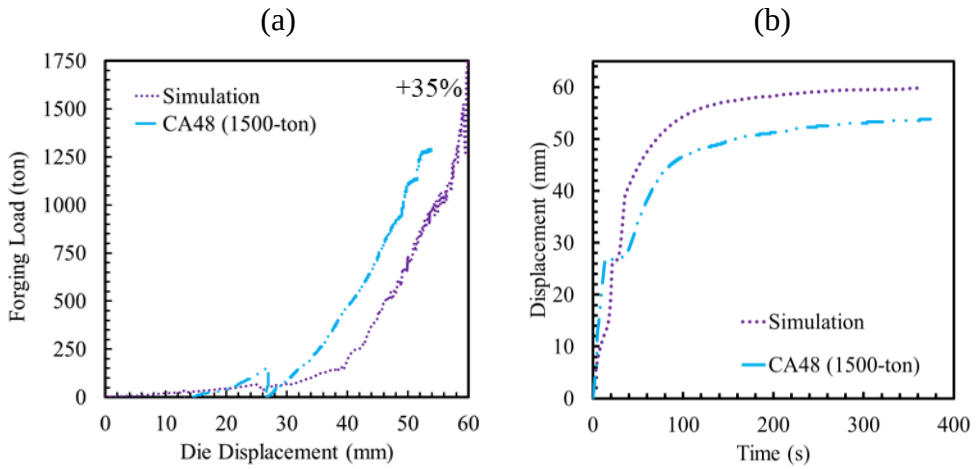


Figure 3-13: (a) Forging load vs die displacement plot, and (b) die displacement vs time plot comparing the simulation with experiment CA48.

In terms of peak forging load, the simulated load prediction was ~1750 Tons, while the recorded load was ~1300 Tons, which amounts to +35% over prediction in forging load. The overall trend between the simulation and experiment press load and displacement curves are agreeable (see

Figure 3-13). However, there is a noticeable offset of about 10 mm between the simulated and experiment die displacements, which is likely due to the deflection of insulation laminates. The changes in ram speed from 4 mm/s to 0.4 mm/s and 0.04 mm/s can be noticed at a die displacement of around 25 mm and 45 mm. The stepped press cycle schedule was adopted to ensure that the dies were able to reach the intended die shut height. By stepping down the ram speed, the workpiece is subjected to lower deformation strain rates, which reduces the material flow stress.

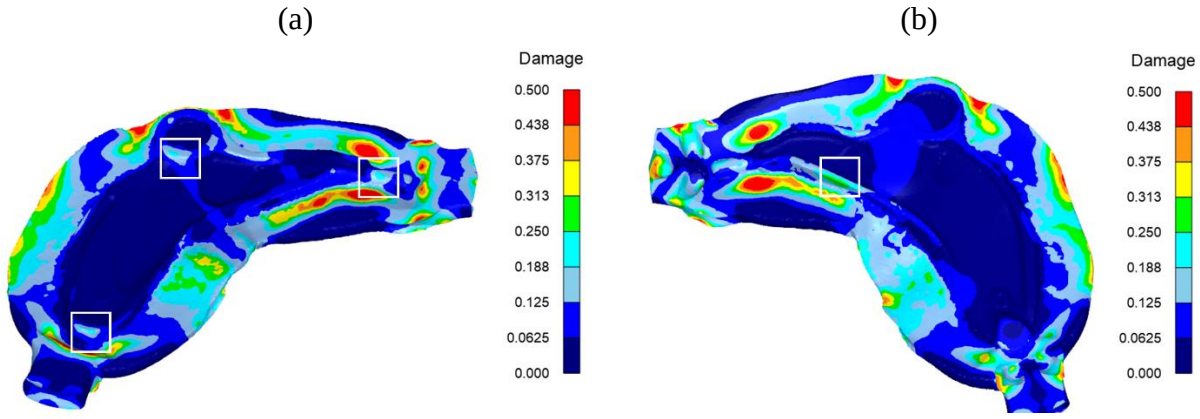


Figure 3-14: Damage distribution (Normalized Cockroft & Latham criterion) of the control arm as viewed from (a) top view and (b) bottom view.

The normalized Cockroft & Latham (C-L) damage criterion was used to visualize the damage distribution shown in Figure 3-14(a) and (b), highlighting the potential sites for crack initiation. The C-L damage criterion is expressed according to the following equation [69]:

$$\int_0^{\varepsilon_i(t)} \frac{\sigma_{max}}{\sigma_i} d\varepsilon_i(t) \geq C \tag{19}$$

where σ_{max} is the maximum principal stress, σ_i is the effective stress, ε_i is the effective strain, and C is the critical damage value. The critical value C is dependent on the material temperature, strain rate and the stress state which makes the precise determination of cracking difficult [69]. However, past researchers like Xue et al. have reported the critical damage value for AZ80 to be between 0.26 and 0.46 at 400ºC for strain rates of 0.001 s^{-1} to 0.1 s^{-1} [70]. Based on this, regions in Figure 3-14 exhibiting damage values greater than 0.25 have been identified as potential sites for crack initiation.

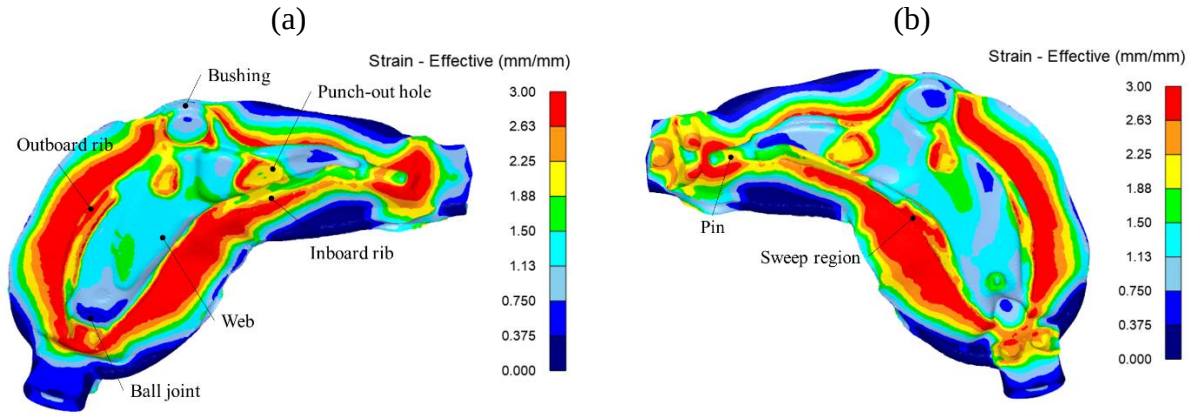


Figure 3-15: Effective strain distribution of the control arm as viewed from (a) top view and (b) bottom view.

The distribution of effective strain shown in Figure 3-15 is fairly uniform along the web region with a strain value of 1.5. Non-uniform effective strains beyond a value of 2.0 can be seen in both the outboard and inboard ribs, the bushing, and pin regions. The lowest effective strains were observed at the ball joint and near the web-to-pin transition region. Non-uniform strain distribution is more common throughout the forging in regions with sharp topological variations (i.e., at the bushing, the outboard and inboard rib, and sweep regions), which can result in microstructural and mechanical property variations. Additionally, it can lead to the development of residual stresses within the forging, which may cause distortions or cracking [71].

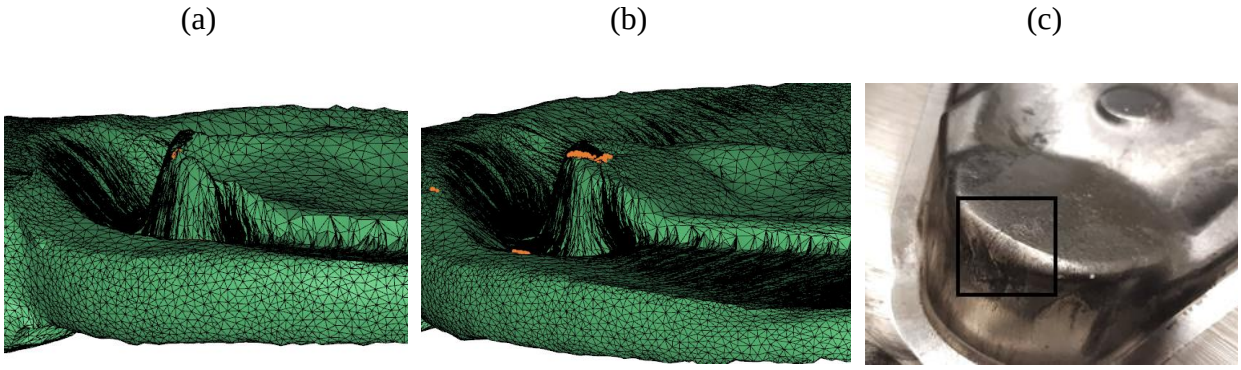


Figure 3-16: Flownet visualization of metal flow in the ball joint region are shown in (a) and (b); forging quality at the location of the fold site is shown in (c).

In this case, the preform geometry was not optimized. As a result, poor flow behaviour was observed in certain regions of the forging via simulation, and defects in the lab-scale forging. Using the "flownet" visualization feature in DEFORM®-3D [27], folding was observed during the formation of the ball joint (see Figure 3-16(a) and (b)). At the onset of forging, the top die digs into the thick volume of material at the end of the preform, causing uneven filling of the ball joint

cavity as the material is forced up the cavity wall until it bottoms out and curls back onto itself, creating a fold site. This fold sight was detected by the software and indicated by the orange dots as shown in Figure 3-16(b). A similar pattern of formation can be observed at the sweep region, where the material front can be observed squeezing into the inboard rib and folding back on itself after it has bottomed out in the cavity. The micrographs of the cross-section at the sweep region reveal poorly fused "fold-lines" and large graphite particles that were entrained in the leading metal front, which compromised the structural integrity of the forging (see Figure 3-17).

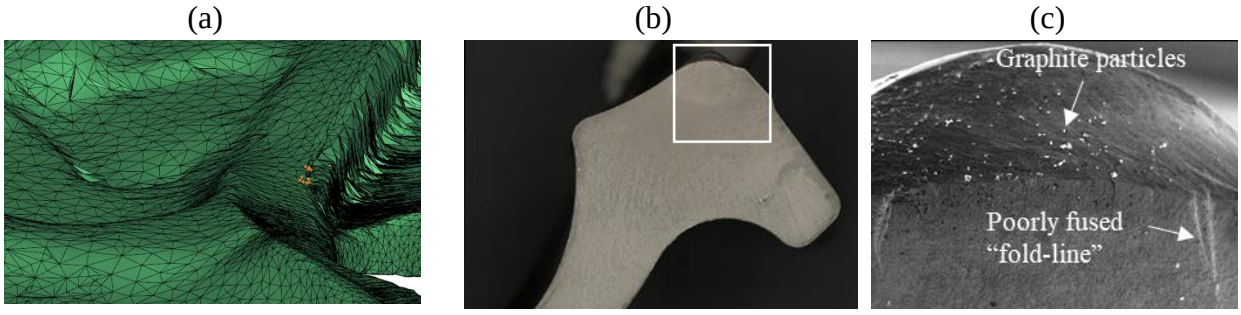


Figure 3-17: (a) Flownet visualization of metal flow in the sweep region; (b) cross-section taken at the sweep region (reproduced from [72]); (c) a magnified view of the cross-section revealing forging defects (reproduced from [72]).

Fold sites were also ubiquitous at the bushing and pin regions. The relative velocity vector plots in Figure 3-17(a) and Figure 3-18(a) reveal the formation of these regions. At the bushing, the metal flow can be seen separating while flowing into and around the punch-out hole. The sharp topologic features in this region induce variations in the local forging direction, further disrupting the metal flow. In Figure 3-19(a) and (b), poor formation of the pin can also be observed as material is seen folding on to itself.

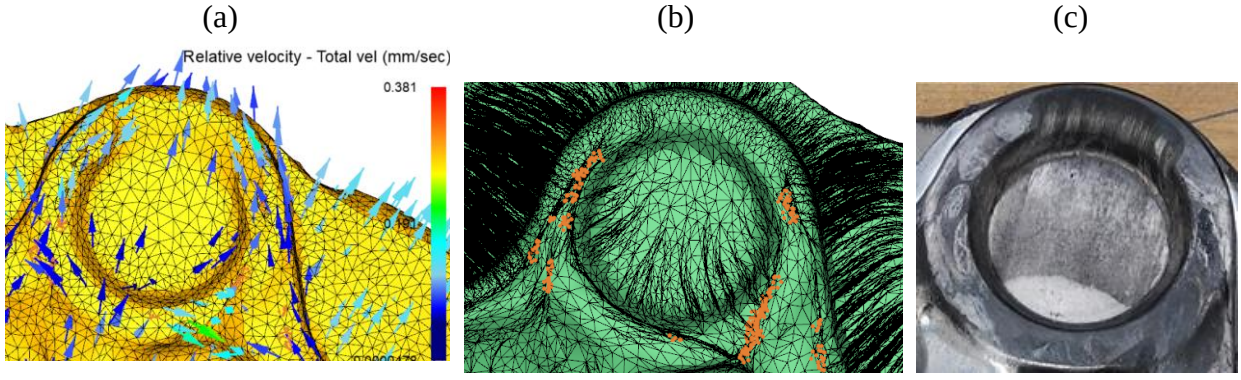


Figure 3-18: (a) Relative velocity vector plot and (b) flownet visualizations of the bushing at intermediate forging steps; (c) shows an under fill at the edge of the bushing.

(a) (b)

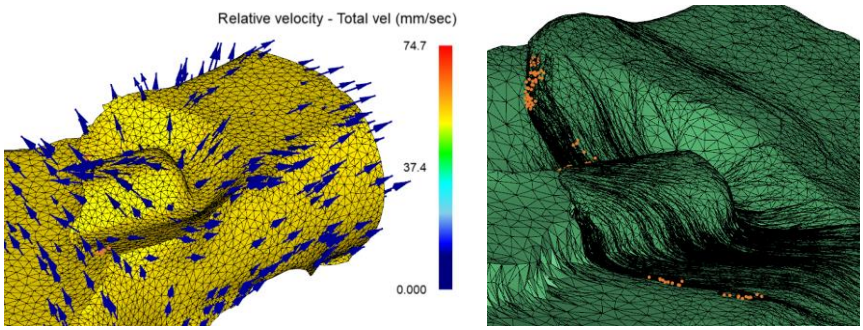


Figure 3-19: (a) Relative velocity vector plot and (b) flownet visualization of pin formation at an intermediate forging step.

3.5 Conclusions

In conclusion, the key contributions of this chapter are the phenomenological material models that were developed for simulating the anisotropic deformation behaviour of extruded AZ80 Mg alloy at processing conditions of 300°C and 400°C, at equivalent strain rates of 0.01^{-1} and $0.1s^{-1}$, as well as the isotropic material model of the as-cast AZ80 Mg alloy at processing conditions of 300°C, at equivalent strain rates of 0.01^{-1}, $0.1s^{-1}$ and $1s^{-1}$. The Hill's (1948) quadratic anisotropic yield criterion was used to predict anisotropic material deformation, where the six independent coefficients of the yield function were obtained using uniaxial and shear yield stress data corresponding to the specific processing conditions. Material models with calibrated Hill's coefficients were validated systematically, starting with simulations of uniaxial compression tests of small cylindrical samples (open-die forging) in both extrusion (ED) and radial (TD) directions. Good dimensional agreement between lab-scale compression samples and corresponding simulations indicated that using a constant set of Hill's coefficients, anisotropic metal flow can be reliably predicted under uniaxial loading conditions (<2% error in major and minor axes dimensions was observed in simulations). It should be emphasized that the material models used in this work should only be used as a reference for simulating deformation conditions under 300°C and 400°C, up to a max strain rate of $1s^{-1}$.

Next, "Coin" forgings of intermediate size and complexity (semi-closed-die forging) were simulated at a processing temperature of 400°C at an average speed of 0.4 mm/s (SF14) and 0.04 mm/s (SF6). The forging behaviour of the "Coin" samples (under both SF14 and SF6 forging conditions) was predicted very accurately until the final steps of the simulations, at which point the load-stroke curves diverged, resulting in an over-prediction of forging loads by +30% and +31% in the simulations corresponding to the SF14 and SF6 processing conditions, respectively. In terms of dimensional accuracy, the simulated "Coin" forging shape, under the SF6 processing

condition, showed good agreement with the corresponding experimental forging shape. A lower degree of dimensional accuracy was observed under the SF14 processing condition. This highlights the fact that under multi-axial loading conditions, using a constant set of Hill's coefficients will result in a less accurate (but acceptable) forging outcome prediction. This claim was reinforced based on the multi-step closed-die forging outcome of the FLCA. Good agreement between the simulated control arm forging and its respective lab-scale forging CA48 reveals that these models can provide a reasonably accurate prediction of forging outcome within a reasonable amount of time. To obtain better simulation predictions of anisotropic deformation behaviour, the Hill's anisotropic yield criterion should be modified to account for yield strength asymmetry between tension and compression; yield stresses used to compute Hill's coefficients should be updated frequently to reflect the dynamic revolution of the principal axes of anisotropy; a more suitable hardening rule should be used; and thermal effects should be considered.

The forging simulation of the I-beam was carried out using an isotropic material model, as the starting microstructure of the cast material is comprised of randomly oriented grains. However, the material develops anisotropic characteristics during deformation, as evidenced by the lab-scale I-beam forging footprint, which reveals non-uniform material stretching in the flash. The evolution of anisotropy is not captured in the isotropic material model, so this contributes to the deviations between the simulated and lab-scale forging flash profile shapes. Nevertheless, as this thesis mainly focuses on optimizing cast preform geometries, the isotropic material model was used in the remainder of this work.

4. Data-Driven Multi-Objective Design Optimization Framework

This chapter describes the development of a data-driven multi-objective framework for preform design optimization. A good preform shape is crucial for mitigating forging defects and material waste, such as those observed in the FLCA forging (CA48). Therefore, it is vital to distribute material effectively when designing forging preforms to (i) prevent disruptive metal flow, (ii) minimize material loss in the form of flash, (iii) minimize tool wear by reducing the degree of metal movement within the die, and (iv) adequately deform material to achieve desired mechanical properties. This framework was developed mainly to achieve preform design objectives (ii) and (iv), and it was applied to optimize a cast-preform shape for the AZ80 alloy I-beam forging operation.

The contents of this chapter have been published[1,2] in Material Communication Today (June 2022) [73] and in the International Journal of Advance Manufacturing Technology (October 2023) [74]. Modifications have been made to integrate their contents with the larger thesis documents. A version of this work was also presented at the 12th International Conference on Magnesium Alloys and their Applications (June 2021) [75].

4.1 Introduction

Three-dimensional preform shape optimization is challenging. Typically, a design would be modified manually in an iterative loop, where implementing design changes to the starting geometry would demand CAD modelling efforts and reliance on expert knowledge to interpret the simulation results to make effective design modifications. Therefore, in the past, researchers have explored techniques to automate simulation-based design (SBD) using optimization algorithms (SBDO). Primarily, SBDO consists of three main elements: (i) the simulation software, (ii) an optimization algorithm, and (iii) CAD software [76]. In preform design optimization, researchers have relied on surrogate models, such as response surface models (RSM) or artificial neural networks (ANNs), to obtain reduced-order approximations of forging response. These models were then used with global derivative-free optimization algorithms, such as evolutionary algorithms (EAs), to evolve preform shapes.

1 Tharindu Kodippili, Stephan Lambert, Arash Arami, "Data-driven prediction of forging outcome: Effect of preform shape on plastic strain in a magnesium alloy forging," Materials Today Communications, Vol 31, 103210 (June 2022).
2 Tharindu Kodippili, Erfan Azqadan, Stephan Lambert, Arash Arami, Hamid Jahed " Multi-objective Optimization of a Cast-preform Shape for a Magnesium Alloy Forging Application" Journal of Advance Manufacturing Technology (July 2023).

Hybrid FEM and ANN methodologies have been successfully employed in the past to reduce the computational burden of simulations. ANNs are universal approximators, which makes them ideal for approximating complex nonlinear functions from simulation data. Additionally, they can be easily integrated with EAs, such as genetic algorithms (GAs), to drive the optimization process. Derivative-free optimization algorithms such as GAs are preferred in large deformation shape optimization problems since FEM computations are expensive and noisy, making it challenging to formulate analytical expressions for objective functions that are parameterized using shape variables. While AI-based EA frameworks have been used in the past in preform design optimization problems, the forging geometries that were considered have either been two-dimensional or very simple three-dimensional geometries. Additionally, there is a noticeable absence of discussion on effective shape parameterization techniques for such problems. The data-driven multi-objective design optimization framework discussed in this chapter was developed to fill this gap. It relies on a reduced-order modelling technique for generating versatile parametric CAD models and data-driven models (feedforward neural networks) for predicting forging quality in spatially varying regions of interest within the die.

4.2 Methods

The proposed data-driven multi-objective design optimization framework consists of three sub-systems: a parametric CAD model for creating preform CAD, feedforward neural networks for evaluating forging outcome of preforms, and a multi-objective genetic optimization algorithm (SPEAII) for searching the design space of the parametric CAD model for viable solutions. The general approach for applying this framework for shape optimization is as follows:

1. Set up a design of experiments (DOE) based on Latin Hypercube Sampling with multi-dimensional uniformity (LHSMDU) to sample the design space and create a preform CAD model database.

2. Run forging simulations of preforms and process them to reduce the large information content output to key performance indicators (KPIs) of the forging outcome to curate a dataset of design and corresponding performance space variables.

3. Use the dataset to train data-driven models to evaluate the KPIs of forging outcomes.

4. Once the offline training of data-driven models is complete, i.e., the models have learned accurate design-to-performance space mapping functions, use the models in conjunction with the multi-objective optimization algorithm to evolve a Pareto-optimal set of preform designs.

4.2.1 Parametric Modelling

Parametric CAD models are ideal for design tasks that require the creation of multiple design variations. Parametric modelling allows designers to capture design intent using features and constraints which can be modified, enabling the generation of many variants of a core design. The design features of a parametric CAD model are embedded in an operational graph (a feature tree) in Grasshopper. An operational graph is a network of interconnected nodes with input and output terminals that maintain geometric information, i.e., parent/child associations between geometric features of the model. The nodes on this graph directly operate on geometric and non-geometric data [44], [77], [78]. Weight parameters are non-geometric inputs to specific nodes in the graph that scale or transform geometric data. By modifying the weight parameters of the operational graph, rapid design alterations can be made to the core parametric CAD model design. This section introduces two parametric CAD models, 1 and 2. Both models were hand-engineered to minimize the dimensionality of the design space, i.e., use as few input parameters as possible to define shape features while attempting to maximize the variability of shapes that can be generated. The two models were used for dedicated purposes: the first was used to develop foundational knowledge in parametric modelling and training data-driven models for purely predicting forging outcome of I-beam-preforms, while the second was used in the design optimization framework to evolve a Pareto-optimal set of I-beam-preforms. Both models were developed in Grasshopper, a visual programming environment that runs on Rhinoceros® [47].

Parametric CAD Model-1

Parametric CAD model-1 was developed by deliberately engineering a design space using a reduced-order modelling technique to minimize the number of input parameters required to control the geometric features of a design. It takes as input seven weight parameters, $\vec{a} \in [0.0, 1.0]^7$ to carry out a sequence of geometric transformations on a set of immutable[3] basis curves, φ^1, φ^2, …, φ^5. The sequence of geometric transformation operations, i.e., the procedural generation logic, was defined using native Grasshopper components. The native components were connected in series to form the operational graph that manipulates geometric information as it propagates downstream once provided with weight parameter inputs. Essentially, the operation graph maps weight parameter inputs—the coordinates of a point in design space—to a three-dimensional shape in Euclidean space. The general procedure used in this work for constructing parametric CAD models is described below:

1. Draw immutable geometry, a set of basis curves $\varphi^1, \varphi^2, \ldots, \varphi^k$ that define the shape extremes based on design intent and boundary constraints conveyed using the cross-sectional curves of the impression die.

[3] Immutable data refers to data that cannot be modified or deleted once it has been created

2. Apply a reduced-order modelling technique (RBT) [33], where basis curves containing large information content (i.e., m equidistant points on a curve) are linearly combined according to the weight parameters contained in a design vector $\vec{a}$, to interpolate the resultant cross-sectional curve, φ^R.

3. Loft together resultant cross-sectional curves, i.e., interpolate a surface between curves, or alternatively, sweep cross-sectional curves along a guide curve(s) (henceforth referred to as a rail) to maintain precise control over the generated three-dimensional closed surface.

For parametric CAD model-1, this entailed drawing basis curves with a general H-section profile to introduce extra material to the die's short and tall rib regions to promote a greater degree of material deformation along the compression axis (along the y-axis shown in Figure 4-1(a)). This model relies on the weight parameters to produce two cross-sectional curves Φ^1 and Φ^2, with three weight parameters being used per cross-sectional curve (see Figure 4-1(b)). Crosssectional curve Φ^2 is offset in the normal direction to the sketch plane of curve Φ^1 according to the seventh input weight parameter. Crosssectional curve Φ^2 is mirrored about the same sketch plane to produce cross-sectional curve Φ^3. Then, the resultant cross-sectional curves Φ^1, Φ^2, and Φ^3 are lofted together to form a closed-surface model of the preform (an operation that interpolates a surface between curves).

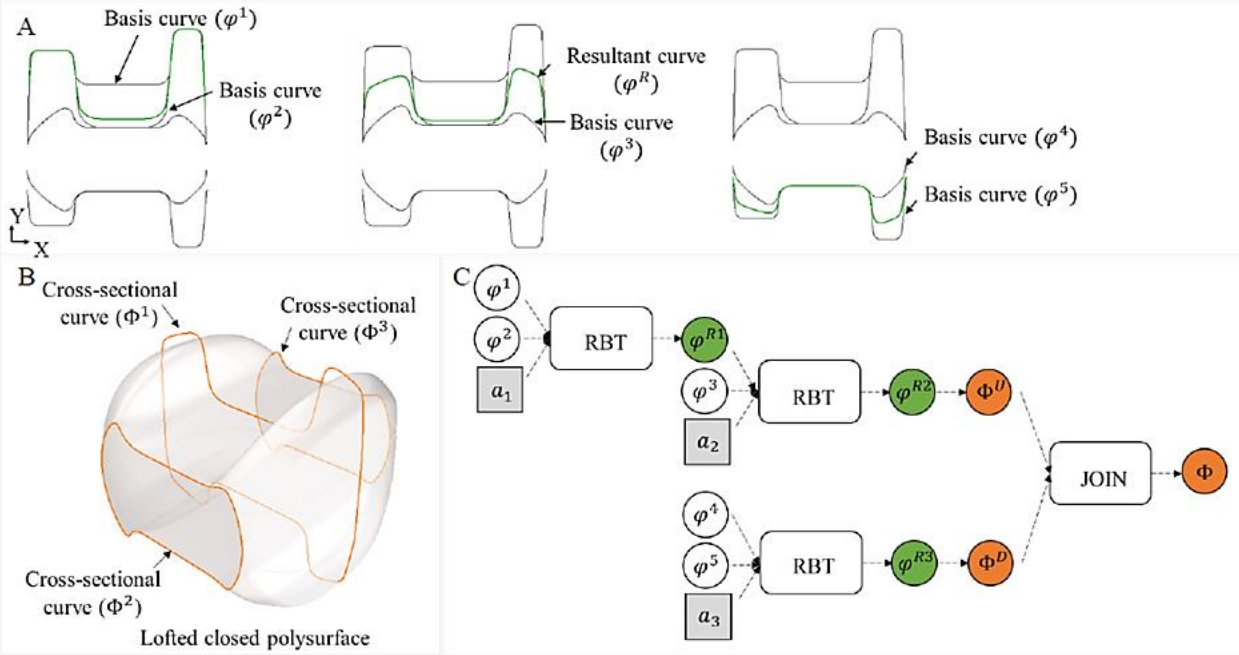


Figure 4-1: A) The set of five basis curves (black) used for creating cross-sectional curves; B) Lofted closed-surface of the model's wireframe; C) The reduced basis technique function (RBT) node takes three inputs: two basis curves (black) and a weight parameter, and outputs a resultant curve (green).

Figure 4-1(c) illustrates the reduced-order modelling technique used to generate cross-sectional curve Φ, which is encoded in its operational graph. Starting with basis curves φ^1 and φ^2, they are linearly combined according to weight parameter a_1 to produce a resultant curve φ^{R1} (Equation 22). Weight parameter a_1 essentially controls the web thickess of the preform (see Figure 4-2).

$$\varphi^{R1} = a_1\varphi^2 + (1 - a_1)\varphi^1 \tag{22}$$

Then, the resultant curve φ^{R1} and basis curve φ^3 are combined according to weight parameter a_2 to produce resultant curve Φ^U (the upper segment of cross-sectional curve Φ). By combining resultant curve φ^{R1} and φ^3, rib height and curvature of the upper section of the preform can be controlled (Equation 23).

$$\Phi^U = \varphi^{R2} = a_2\varphi^{R1} + (1 - a_2)\varphi^3 \tag{23}$$

Similarly, the lower segment of the cross-sectional curve Φ^D is generated using basis curves φ^4 and φ^5, and weight parameter a_3. By combining these two basis curves, the rib height and curvature of the lower section of the preform can be controlled (Equation 24).

$$\Phi^D = \varphi^{R3} = a_3\varphi^5 + (1 - a_3)\varphi^4 \tag{24}$$

Finally, curves Φ^U and Φ^D are joined to form a cross-sectional curve, Φ.

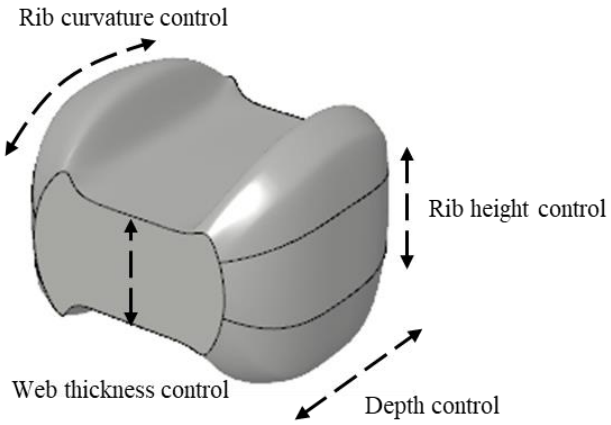


Figure 4-2: Preform shape features that can be controlled by tuning weight parameters (parametric CAD model-1)

Parametric CAD Model-2

In the second version of the parametric CAD model, improvements were made to the procedural generation logic based on the observation that parametric CAD model-1 performs many-to-one mappings between design points and preform shapes, where the spatial proximity between

uniformly sampled design points in a DOE was not preserved between three-dimensional shapes in Euclidean space. This was problematic since it reduced the data efficiency[4] of the ANN training dataset.

Parametric CAD model-2 addressed this issue. This model generated preforms with H-sections using an 8-dimensional input vector, $\vec{a} \in [0.0, 1.0]^8$. Eight weight parameters were required to control the geometric features of generated CAD models: web height and curvature, rib height and curvature, depth, and the distance between the bottom surface of the web and die surface (see Figure 4-3). The particular features of the core design and the allowable range of modifications were determined based on engineering judgment: 1) to control height-wise and depth-wise stretching of the general H-section to increase/decrease the overall volume of material that is compressed along the loading direction (affecting the overall degree of deformation), and 2) to control the distribution of material in the web and ribs by controlling the surface curvature to reduce excess material that gets pushed into the flash during deformation. The reduced-order modelling technique was also applied in this model; basis curves were drawn to increase shape variability and to guarantee a one-to-one mapping between design points and preform shapes.

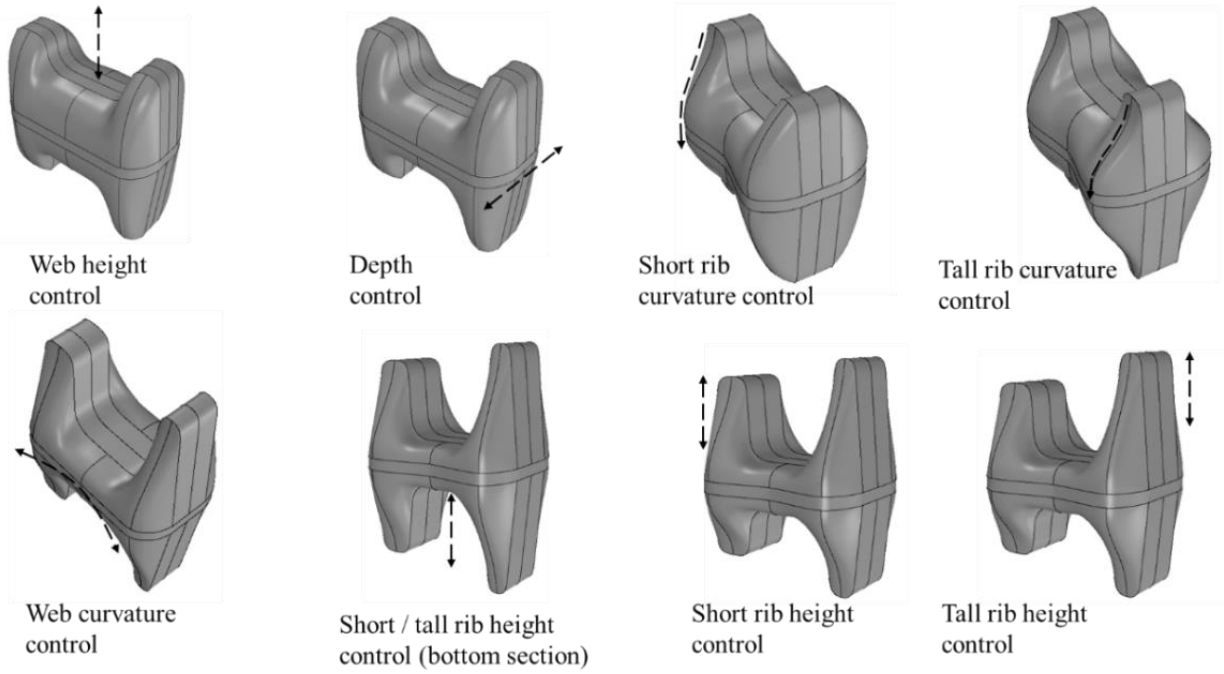


Figure 4-3: Geometric features of the preform shape that can be controlled by tuning weight parameters

The models presented in this section were used to generate CAD for two experiments: parametric CAD model-1 was used for preliminary work on data-driven prediction of forging outcome, and

[4] In the context of training neural networks, data efficiency refers to the ability of a neural network to achieve high performance with a limited amount of training data

parametric CAD model-2 was used in the data-driven multi-objective optimization framework to train its data-driven models.

4.2.2 Design of Experiments

In addition to having a good generative system that was able to output a large variation of preform designs, an effective method for evaluating designs was also required. For this purpose, feedforward models were trained on a subset of FEM simulations to evaluate forging outcomes. These models were trained to map input parameters, i.e., design space coordinate vectors to KPI values in performance space. A DOE was set up using a Latin hypercube sampling (LHS) technique with multi-dimensional uniformity (MDU) to select the subset of preform shapes for FEM simulation. In a traditional LHS technique, input variables' univariate cumulative distribution functions (CDFs) are stratified into non-overlapping, equiprobable intervals, followed by random sampling of a point from within each stratum to form a Latin hypercube [79]. LHSMDU extends this idea of univariate uniformity to a multivariate condition and so has the effect of spreading out sampling point coordinates in a quasi-random manner across the entire input space [80].

In the preliminary study, which investigated the effects of preform shape on forging outcome, LHSMDU was used to sample 300 design points. These were then fed to parametric CAD model-1 to generate the corresponding CAD models for simulation (Figure 4-4). DEFORM®-3D [27] was used to run low-fidelity forging simulations, requiring a total simulation run time of about 50 hours, approximately 10 minutes per simulation (the preform geometry was meshed with approximately 30k elements), running on 4 cores of an i7-7820X CPU @ 3.6 GHz.

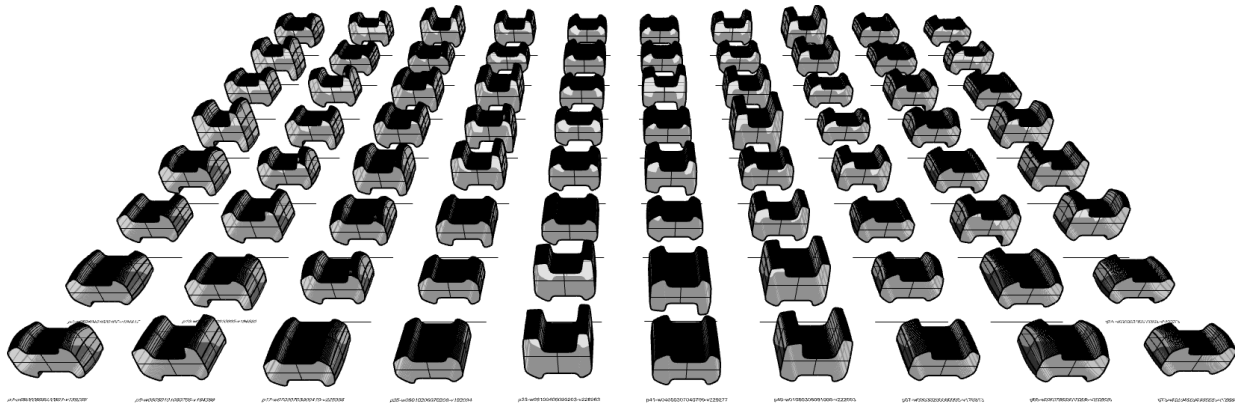


Figure 4-4: Preform designs that were sampled for simulation; 100 samples generated using parametric CAD model-1 are shown.

In subsequent work, to evolve a Pareto-optimal set of candidate I-beam preform designs, the design space of parametric CAD model-2 was uniformly sampled using LHSMDU to obtain 250 design points for FEM simulation. The total simulation time in this case was about 83 hours,

requiring approximately 20 minutes per simulation (the preform geometry was meshed with approximately 50k elements). To minimize user effort, the data generation process was automated, where a general simulation template was defined using the die geometry, material, and processing conditions (forging speed, temperature, and friction coefficient) described in Section 3.4. An automation script then called this template to iteratively import a new preform mesh geometry, generate a simulation database file, and then submit it to the numerical solver.

4.3 Data-Driven Prediction of Forging Outcome

4.3.1 Training Feedforward Networks

Using CAD models that were generated using parametric CAD model-1, 300 forging simulations were conducted (the material and process parameters of the simulation template is defined according to Table 3-5). Then, each simulation's effective plastic strain and mesh data (node and element connectivity lists) were post-processed using an automation script. Effective plastic strain is a state variable that can be extracted from a simulation to describe the degree of deformation, which can be correlated with microstructural changes. This script looped through the raw data files, reduced mesh data to a point cloud representation, and associated each point (mesh element centroid) with the corresponding element effective plastic strain value. Next, this large information content was reduced by culling points outside of the predefined regions of interest (see regions of interest: grid indices 26, 63, 79, and 116 in Figure 4-5) and by further reducing the effective plastic strain values in each region to an average effective plastic strain value, ϵ_{avg}^{p}. Then, the design vector, which consists of the input weight parameter values and the ϵ_{avg}^{p} values of each region were concatenated to form a data sample—300 such samples were used in the feedforward network training dataset.

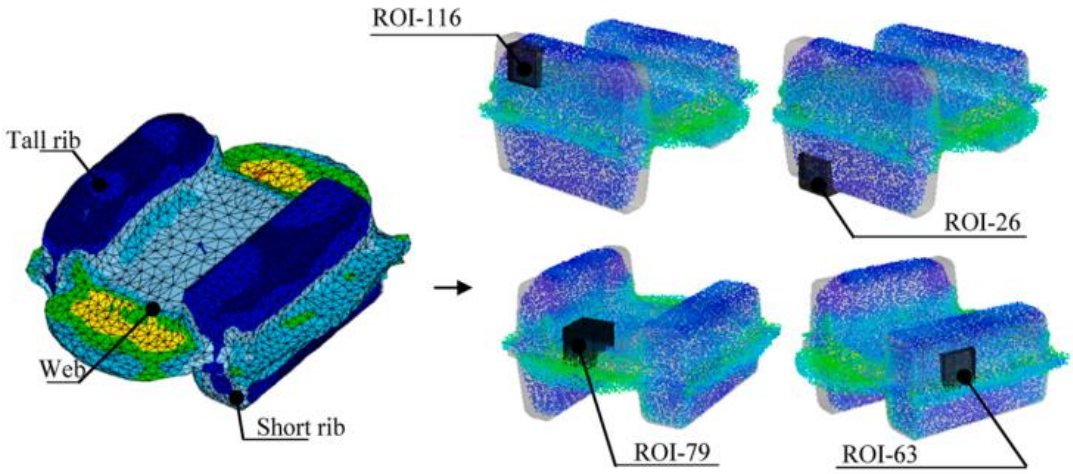


Figure 4-5: Point cloud representation of FEM mesh-elements. Region of interest (ROI): 26, 63, 79, and 116 are shown.

Four feedforward network models were trained to predict the average effective plastic strain in regions 26, 63, 79, and 116. Each model had seven input neurons that received a design vector with weight parameter values characterizing preform shape and a single output neuron which returned a prediction of average effective plastic strain for the respective region of interest. 10-fold cross-validation was used to evaluate the predictive performance of the models developed for each region of interest [81]. The dataset was randomly partitioned into ten equal-sized folds. Nine out of the ten folds were used for training the model, and the left-out fold was used for testing. The left-out fold was rotated so that all ten folds were tested.

Each model had two hidden layers with three and two neurons in the first and second hidden layers, respectively, and an output layer with a single neuron. A hyperbolic tangent activation function was used in the hidden layers to introduce nonlinearity to the model. The model's initialization weights and biases were drawn from a uniform distribution, $U[-1,1]$. The Huber loss function (see Equation 25) was used to evaluate prediction error. Huber loss behaves quadratically for small error residuals and linearly for large residuals [82]. Given the target output distribution of the training dataset, Huber loss proved to be a more robust loss function in comparison to mean-squared-error loss as it was less sensitive to the outliers in the dataset.

$$\text{Huber Loss} = \begin{cases} \frac{1}{2}x^2, & |x| \leq \alpha \\ \alpha\left(|x| - \frac{1}{2}\alpha\right), & |x| > \alpha \end{cases} \tag{25}$$

where the alpha parameter, α, indicates the transition between a quadratic error and a linear error (an alpha value of 0.5 was used). The Adam optimization algorithm was used for training the feedforward networks [83]. The learning rate was set to 0.001. The exponential decay rate for the first moment, β_1, and second moment β_2, were set to 0.9, and 0.999, respectively. Lastly, a batch size of four was used, and the models were trained for 5000 training epochs.

4.3.2 Results and Discussion

Cross-validation results were summarized to evaluate the predictive performance of the models. Mean absolute error (MAE) was used to assess the error between model predictions and target outputs. The coefficient of determination (R^2) was used to determine how well the models fit the data. The lowest MAE was recorded for model predictions in region 26, and the highest in region 79 (see Table 4-1). Model predictions of average effective plastic strain response in region 26 showed the best fit to data with an R^2 value of 0.93±0.04 (mean±standard deviation).

Conversely, model response predictions for region 79 showed a less accurate fit to the data with an R^2 value of 0.84±0.08. In general, model predictions for regions 26, 63, 79 and 116 showed a strong positive correlation, indicating that the trained models can effectively predict strain response in regions of interest for any preform that can be generated using this parametric CAD model.

Table 4-1: 10-Fold cross-validation statistical results summary

ROI	R^2	R^2_{MAX}	MAE	MAE_{MIN}
26	0.93±0.04	0.97	0.02±0.003	0.01
63	0.88±0.06	0.93	0.04±0.004	0.03
79	0.84±0.08	0.93	0.05±0.004	0.04
116	0.91±0.04	0.95	0.04±0.002	0.01

Composite residual plots based on cross-validation model ensemble predictions were also generated to observe patterns in the residual distributions for average effective plastic strain, see Figure 4-6. The residual is defined as the difference between the target and predicted values. By analyzing the residuals of each model, model bias and the relative magnitude of individual errors were identified. The residuals (see Figure 4-6 A,D,G,J) that lie below the horizontal zero-line indicate an overestimation of strain response, whereas residuals that lie above the zero-line indicate an underestimation of response. Along with the residual plots, preforms and their respective forged state corresponding to the lowest (see Figure 4-6: C,F,I,L) and highest residuals (see Figure 4-6: B,E,H,K) were visualized to draw correlations between the geometric features of a preform and the predicted response. For regions 26, 63, and 116, the marginal histograms of strain response (plotted on the x-axis) revealed a right-skewed distribution, where response prediction of preforms with the smallest residuals was clustered around the distribution mean. The marginal histograms of residuals for all four plots were normally distributed (plotted along the y-axis) and were centred at zero, implying little to no model bias. The response histograms, which revealed a right-skewed distribution (Figure 4-6: A, D, and J), show a low sample count at the tail-end. This skewed distribution indicated an under-representation of shapes in the dataset which exhibited high effective plastic strains after deformation. These under-represented shapes tend to have more material allocated in the flanges and undergo a greater deformation during forging, resulting in higher effective plastic strains.

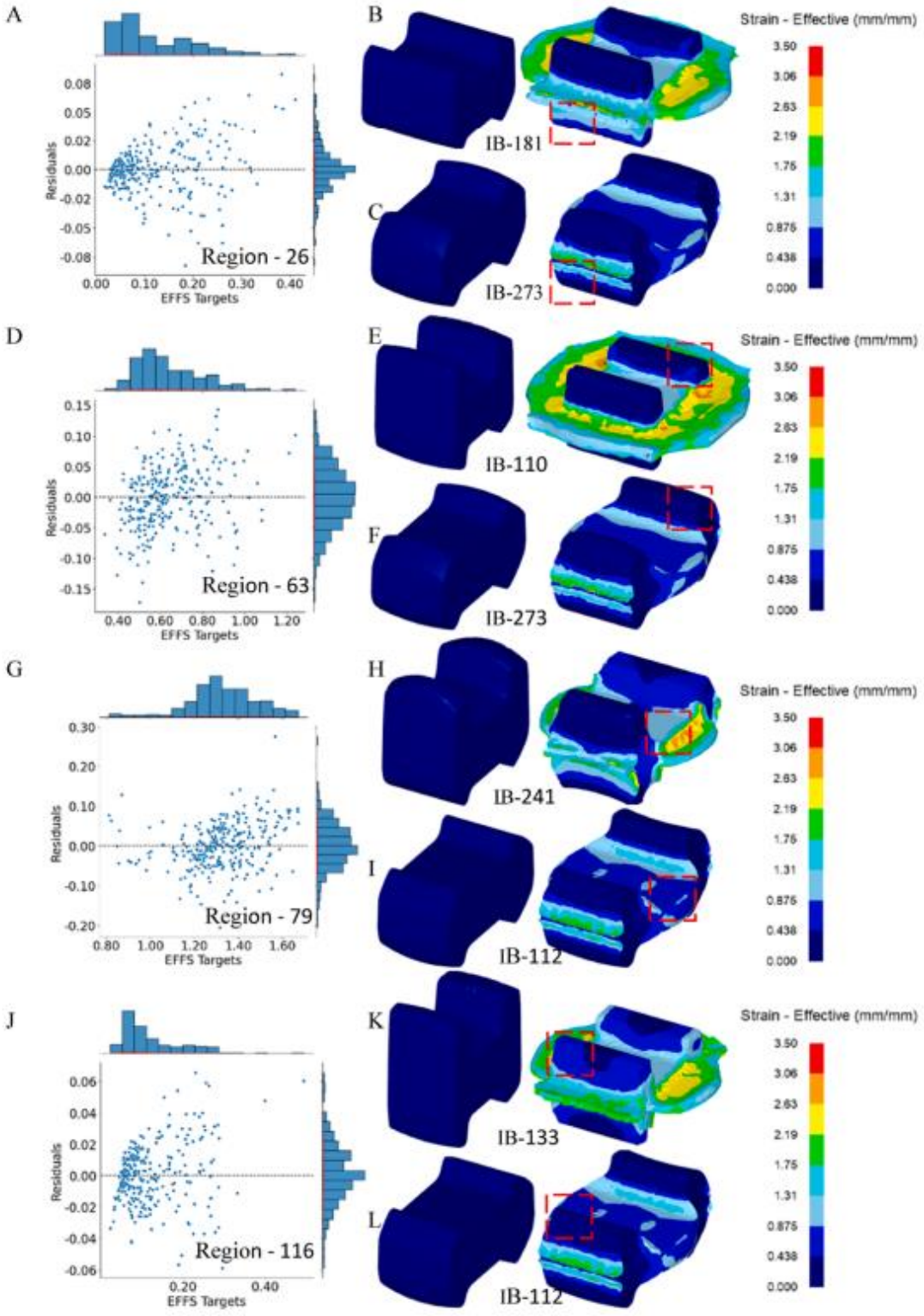


Figure 4-6: Region 26: A) composite residual plot; B) preform IB-181 with a strain of 0.41; C) preform IB-273 with a strain of 0.02. Region 63: D) composite residual plot; E) preform IB-110 with a strain of 1.24; F) preform IB-273 with strain of 0.33. Region 79: G) composite residual plot; H) preform IB-241 with a strain of 1.67; I) preform IB-112 with a strain of 0.81. Region 116: J) composite residual plot; K) preform IB-133 with a strain of 0.49; L) preform IB-112 with a strain of 0.02.

High residuals indicate that the feedforward models were less adept at predicting strain response for these preforms. Conversely, smaller residuals were observed in preforms with shorter flanges, suggesting a larger account of such shapes in the dataset. This observation can be explained as a consequence of the many-to-one mapping dilemma that was previously discussed. Nevertheless, due to a sufficiently large dataset size used to train the feedforward models, they still learned relatively accurate mappings between the design space and their respective performance spaces of average effective plastic strains for the different regions of interest. In general, the trained models predicted forging responses to within ±8% of the ground truth.

4.4 Cast-Preform Shape Optimization for I-Beam Forging

Next, a data-driven multi-objective design optimization was developed. Since it relied on parametric CAD model-2 for shape generation, a new set of feedforward models were trained to evaluate the performance of preform designs, and a multi-objective evolutionary algorithm was used to carry out a design space search. This framework was used to evolve a Pareto-optimal set of candidate preform designs, from which a single design was selected for the PMC process to produce an AZ80 alloy preform for I-beam forging.

A new set of feedforward networks were trained on 250 simulations, which were post-processed to extract effective plastic strain and raw mesh data. Fewer simulations were run to train the new set of feedforward models since the new parametric CAD model that was developed guarantees a one-to-one mapping between design and shape space, improving data efficiency during training. As before, the mesh data corresponding to the final step of each simulation was reduced to a collection of element centroids with effective plastic strain attributes. This point cloud was then further reduced to get an average effective plastic strain (ϵ_{avg}^{p}), this time in 21 regions of interest (see Figure 4-7). The material waste produced for each preform was evaluated as a percentage of the flash volume to the total volume of the preform. The die fill volume percentage was also calculated for each simulation. In total, 23 feedforward networks were trained to predict ϵ_{avg}^{p} in 21 regions of interest, and overall flash and die fill volume percentages. The feedforward networks (8-4-4-1) consisted of two hidden layers with four neurons, each with hyperbolic tangent (tanh) activation functions and a single output neuron with a linear activation function. Each network accepted an 8-dimensional design vector and returned a scalar value quantifying the performance of the corresponding shape.

The dataset was randomly partitioned into training, validation and test sets (with a partition ratio of 80-10-10). Each model was trained for 5000 training epochs with early stopping (using the validation set). As before, the Adam optimization algorithm was used with a learning rate of 0.001. The exponential decay rate for the first moment, β_1, and second moment β_2, were set to 0.9, and 0.999, respectively; and a batch size of two was used. Model performance was evaluated using the test set (see to Appendix B.1)

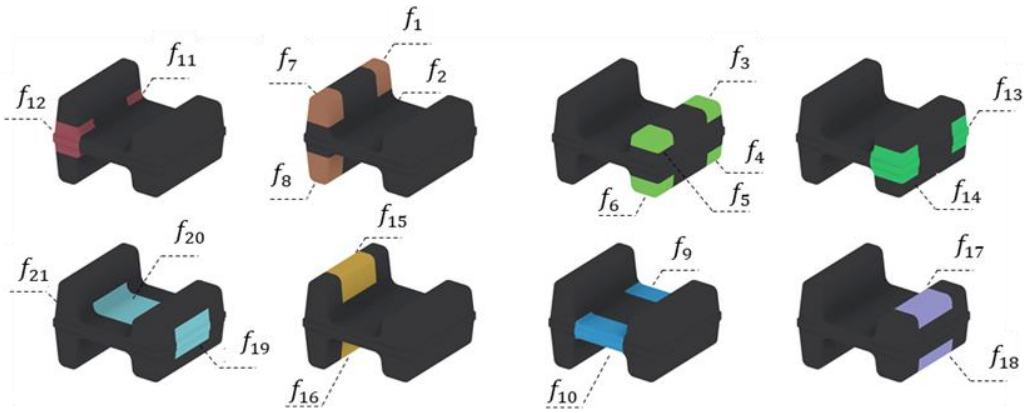


Figure 4-7: The set of dedicated feedforward networks (f_i) used for evaluating the average effective plastic strain (ϵ_{avg}^{p}) in regions of interest within the I-beam, $f_i: X \rightarrow Y,\ i = 1, \ldots, 23$, ($f_{22}$ *and* f_{23} networks evaluate flash and fill volume percentages)

In general, model predictions on test data showed very high coefficients of determination (R^2) ranging between 0.9 and 1.0, showing a good model fit to the target data—validating the modifications that were made to the procedural generation logic of parametric CAD model-2, which ensured a one-to-one mapping between design points and preform shapes.

4.4.1 Multi-Objective Optimization

After training feedforward models to learn accurate mapping functions between design and performance spaces, they were connected to a multi-objective optimization algorithm to evaluate the performance (or fitness) of preform shapes that were evolved by the algorithm—eliminating further reliance on simulations. The Strength Pareto Evolutionary Algorithm 2 (SPEAII) was chosen as the optimization algorithm in this work to optimize several objective functions simultaneously given a design vector. The SPEAII algorithm converges to a set of Pareto-optimal solutions, which is a set of non-dominated solutions in objective space that define a boundary beyond which any further improvement of any one objective will worsen at least one other objective [84]. This algorithm was used to search the design space of parametric CAD model-2 for a set of Pareto-optimal I-beam preform designs. The algorithm iteratively solves for the non-dominated set of solutions according to the following steps (refer to Section 2.2.2 for a detailed description):

1. Initialize the algorithm with a random starting population of input vectors and an empty archive set.

2. Calculate the fitness values of individuals in the population and archive sets based on objective functions.

3. Add non-dominated solutions in the population and archive set of the current generation to the archive set of the subsequent generation based on the fitness of an individual solution (Pareto-dominance) and a nearest neighbour-based archive truncation method to maintain diversity (diversity in shapes of the non-dominated solutions)

4. Apply genetic operations of recombination and mutation to the mating pool to obtain the starting population of the subsequent generation according to:

    a. A roulette wheel selection scheme to draw parent chromosomes from the population (i.e., two design vectors) and multi-point cross-over to combine genetic information of parent chromosomes by randomly selecting and swapping values between element indices of the two design vectors.

    b. Mutation operations on each offspring based on a mutation probability, where an offspring chromosome was mutated by randomly selecting an index along its design vector and randomly assigning a value between 0.0 and 1.0.

5. Terminate the optimization search if the loop iteration count exceeds the predefined termination criteria.

6. If the termination criteria were not met, apply the selection scheme once more to select a mating pool consisting of individuals from the current population to continue the optimization cycle from step 2.

The multi-objective I-beam preform design optimization problem was defined according to the following expressions:

$$\begin{aligned} \boldsymbol{max} \quad & F(X) = (\, F_1(X),\, F_2(X),\, F_3(X))^T \\ \boldsymbol{s.t} \quad & 0.95 < \, G(X) \\ & 0.0 \, \leq x_i \, \leq 1.0, \quad X = (x_1,\, x_2, \ldots,\, x_8)^T \end{aligned} \tag{26}$$

where $F_i(X)$ and $G(X)$ were objective and constraint functions, respectively. Here, the outputs of these functions $F_i(X)$ and $G(X)$ were evaluated using the trained feedforward networks. The problem was defined to search the design space for preform shapes with high ϵ_{avg}^{p} and low variance of ϵ_{avg}^{p} between sub-regions in the rib, and a low flash percentage and a die fill percentage greater than 95%. To this end, $F_1(x)$ calculated an objective value based on the geometric mean of ϵ_{avg}^{p} in regions 12 and 14 (see Equation 27). $F_2(x)$ calculated an objective value based on strain variance, $\sigma^2(\cdot)$, between regions 12 and 21 in the short rib, and regions 14 and 19 in the tall rib (Equation 28). $F_3(x)$ calculated an objective value based on the flash percentage (Equation 29),

and constraint function $G(x)$ based on fill percentage (Equation 30). As per step (2) in the SPEAII algorithm, the three objective functions and the constraint function were used to evaluate the fitness of individuals. Note that lowercase "f" refers to feedforward network function, "X" is the design vector input, and uppercase "F" refers to the objective value function.

$$F_1(X) = \sqrt{f_{12}(X) \cdot f_{14}(X)} \tag{27}$$

$$F_2(X) = 1 - \sigma^2(f_{12}(X), f_{21}(X)) - \sigma^2(f_{14}(X), f_{19}(X)) \tag{28}$$

$$F_3(X) = 1 - f_{22}(X) \tag{29}$$

$$G(X) = f_{23}(X) \tag{30}$$

The optimization parameters of the SPEA2 algorithm used in this study were as follows: population size=50, archive size=30, number of generations=50, crossover probability=0.55, and mutation probability=0.05. The hyperparameters of the SPEA2 algorithm were determined based on a grid search, and the values which led to the best optimization results are reported here. The termination criteria for the algorithm was triggered after reaching the maximum number of generations.

4.4.2 Results and Discussion

Optimization Results

Following the optimization, three solutions, 2, 14, and 19 were identified from the set of preforms in the archive set (the set of dominant solutions evolved during optimization, which consists of diverse solutions that are measured in terms of their objective values) after loop termination. Among the three solutions, solution 2 had the highest mean ϵ_{avg}^{p} objective value, followed by solutions 19 and 14. In terms of the flash and ϵ_{avg}^{p} variance objectives, solution 14 ranked above solution 19, followed by solution 2. Figure 4-8 shows the Pareto-optimal solutions that were obtained in objective space. Figure 4-8(a) shows a projection of the Pareto front onto objective dimensions F_1 and F_3 (mean ϵ_{avg}^{p} vs. flash objectives), and Figure 4-8(b) shows a projection on to dimensions F_2 and F_3 (ϵ_{avg}^{p}variance vs. flash objectives). In general, both solutions 14 and 19 were similar in shape, i.e., in web thickness, the bow-shaped rib feature, and preform depth. Considering the proximity of solutions to one another in objective space, this was to be expected. In contrast, solution 2 was further away from solutions 14 and 19, and the difference in geometric features was noticeable. Solution 2 had a thicker web, tall ribs, and a bow shape with a greater inward inflection. From a die-fill standpoint, all three solutions fully filled the die impression. This was also expected since the constraint function explicitly confined the search space of solutions to those with an expected fill volume greater than 95%. In general, solution 14 showed a better overall performance based on objective values (see Table 4-2). Weight parameters of the three solutions were passed to the generative system to obtain CAD models of preforms

shown in Figure 4-9. Volume measurements of solutions 2, 14, and 19 were $302.9(10^3)\ mm^3$, $200.9(10^3)\ mm^3$ and $210.4(10^3)\ mm^3$, respectively, while the impression volume was $193.1(10^3)\ mm^3$. The flash volume percentage of the preform was obtained by subtracting the impression volume from the total preform volume and dividing it by the impression volume. Flash volume percentages of solutions 2, 14, and 19 were approximately 56%, 4%, and 9%, respectively. Disregarding solution 2, considering the amount of material waste, high-fidelity forging simulations of solutions 14 and 19 were run. Both simulations revealed comparable plastic strain responses due to the likeness in the general shape and metal flow behaviour. From these two solutions, solution 14 was selected for a final detailed modification (to ensure good placement in the impression die) prior to being used to create the impression of the casting mould design.

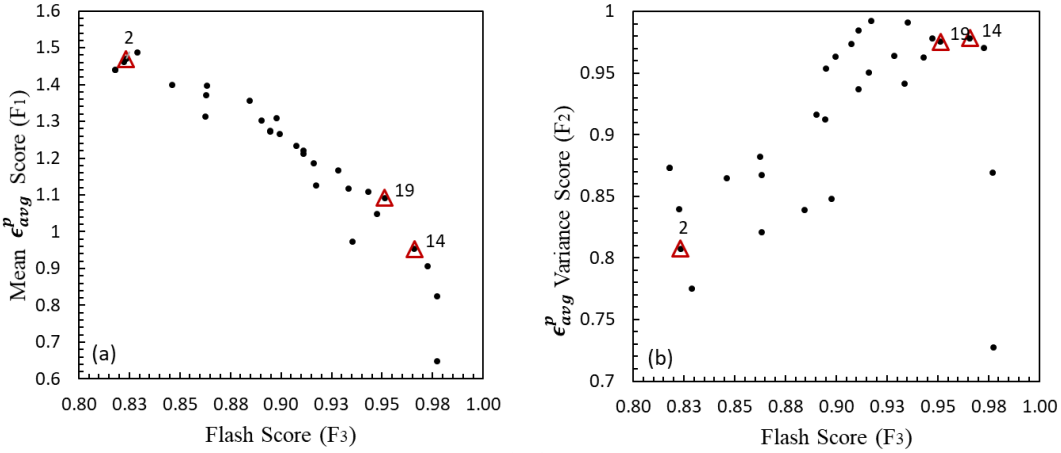


Figure 4-8: (a) Projection of the Pareto front of non-dominated solutions onto F_1 and F_3 (mean plastic strain vs. flash objectives);(b) Projection of the Pareto front onto F_1 and F_2 (strain variance vs. flash objectives)

Table 4-2: Objective function scores of flash, mean ϵ^p_{avg} and variance ϵ^p_{avg} of solutions 2, 14 and 19

Solution	Variance ϵ^p_{avg} $F_1(x)$	Mean ϵ^p_{avg} $F_2(x)$	Flash $F_3(x)$
2	0.807	1.469	0.823
14	0.965	0.952	0.978
19	0.951	1.091	0.975

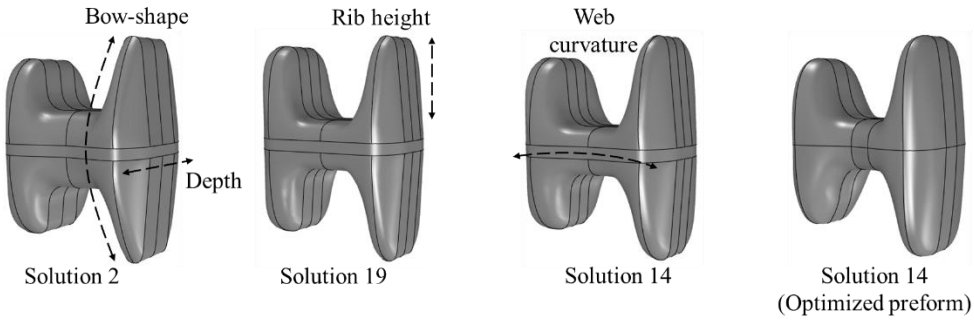


Figure 4-9: Pareto-optimal solutions 2, 19, and 14 obtained after 50 generations of preform shape evolution using SPEA2 multi-objective optimization algorithm

Referring to the forging simulation results in Figure 4-10, axisymmetric metal flow can be observed in the cast-preform and cast-billet as a result of shape symmetry in the top and bottom impressions. As a result, the effective plastic strain contour pattern can be seen mirrored at each step of the simulation about the centerline of the web. At the onset of forging of the cast-billet, the material between the web sections is compressed along the longitudinal direction, forcing the surrounding material to bulge out into the rib cavities, producing the resultant strain path (see Figure 4-10(a)). Continued deformation forces metal into the ribs, and the grain flow starts to follow the contours of the die. At this point, a high degree of plastic deformation (~2.0) can be observed at the web-to-rib fillet transition as divergent normal forces displace material longitudinally at the fillet transition point closest to the web and laterally at the transition point closest to the rib (see Figure 4-10(b)). In AZ80 alloy forging, a good effective plastic strain value was considered to be a value that was greater than 0.4, which was shown to mark the initiation of DRX in the alloy during uniaxial compression testing [11]. Towards the end of forging, a high degree of plastic strain variance can be observed in the ribs as severely deformed material from the web region has partially migrated into the ribs (see Figure 4-10(c)). Continued deformation leads to material flow into the flash land and gutter. The fully forged state of the billet—throughout its cross-section—revealed plastic strains ranging between 1.25~2.0 in the web, 0.25~2.0 in the short rib, and 0~2.0 in the tall rib. A lack of internal pressure to force metal flow further into the ribs at the final stages of forging leads to flow stagnation within the ribs, creating dead metal zones with minimal plastic deformation (0.0~0.25). In contrast, at the onset of deformation of the cast-preform, the metal flows along the longitudinal direction in both short and long flanges simultaneously, and plastically deforming along the height of the ribs (see Figure 4-10(e)). Continued deformation leads to contact in the fillet transition zones and further material compression within the ribs (see Figure 4-10(f)). However, small regions of dead metal zones can still be observed in the preform forging at the midway point along the length of the ribs (see Figure 4-10(g)). In general, throughout the displayed cross-section, effective plastic strains ranged between 1.0~1.25 in the web, 0.75~1.0 in the short rib, and 0.5~0.75 in the tall rib. A comparison of Figure 4-10(c) and (g) reveals a more uniform degree of material deformation in the short and tall rib of the preform forging and a smaller dead metal zone compared to the cast-billet forging. Figure 4-10 (d) and (h) show the outcome of experimental forgings of the cast-billet

and the optimized preform, respectively. To quantify the generated flash volume, it was trimmed from both forgings and weighed. The trimmed flash of the cast-billet forging weighed 42.6g, and the optimized preform forging, 20.8g. The measured flash-to-impression volume percentage was 12% and 6% for the cast-billet and preform forgings, while the simulation-based predictions were approximately 15% and 7% for the cast-billet and cast-preform forging, respectively. Since anisotropy-related flow behaviour which develops during forging was not explicitly captured in simulations, a discrepancy in the predicted flash profiles between simulations and experiments can be observed, especially at the end of the stroke when the effect of anisotropy is most pronounced, in both forgings.

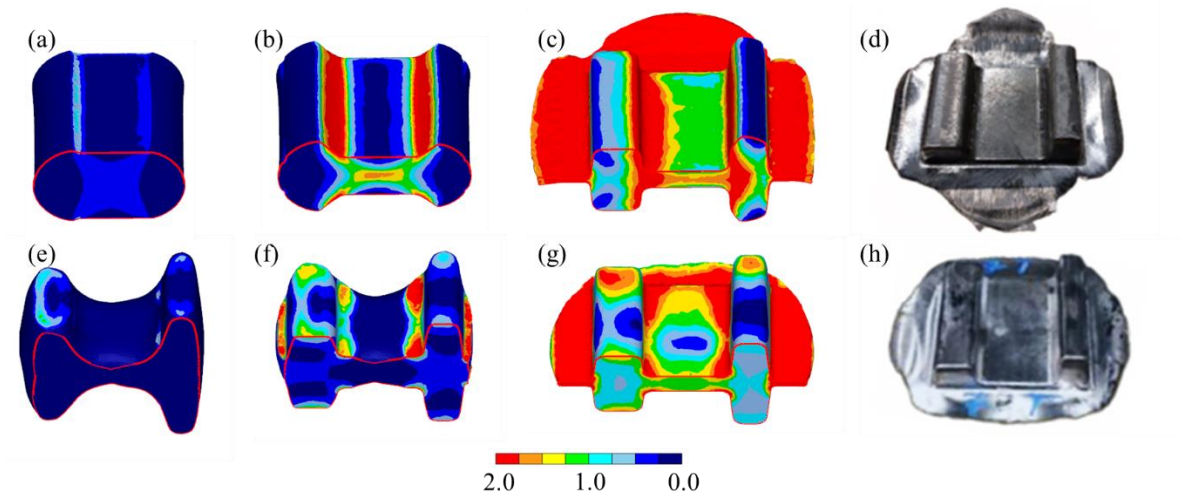


Figure 4-10: (a)-(c) intermediate forging steps of the cast-billet forging simulation; (d) forging outcome of the cast-billet; (e)-(g) intermediate forging steps of the cast-preform forging simulation; (h) forging outcome of the cast-preform; effective plastic strain is visualized

Experimental Results

The I-beam forgings were produced sequentially by first casting the preforms using a PMC process and then hot forging them at an elevated temperature of 300 ºC. The molten metal for casting both the baseline cylindrical billet and the optimized preform was prepared using a pre-alloyed AZ80 billet from Mag Specialties. In addition to preform geometry, the forging process and material parameters also play a decisive role in the hot deformation behaviour and, consequently, the mechanical properties of the forging (for a more detailed discussion on material and process parameters, refer to Section 2.1). During the forging step, as-cast preforms were hot forged under isothermal conditions at 300℃ (see Figure 4-11), reducing casting defects, closing pores, breaking down the primary phases of the as-cast microstructure, and promoting grain refinement. Both casting and forging operations were conducted at CanmetMATERIALS.

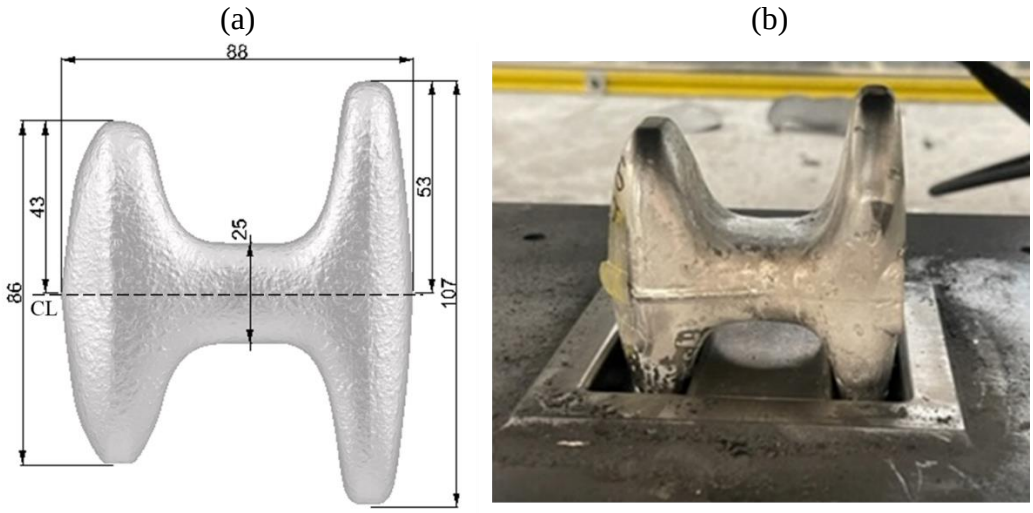


Figure 4-11: (a) PMC mould; (b) as-cast cylindrical billets; (c) optimized I-Beam preform (courtesy of CanmetMATERIALS)

Azqadan determined the mechanical properties of both forgings using quasi-static tensile and stress-controlled cyclic tests. Figure 4-12(a) shows the engineering stress-strain curves corresponding to samples extracted from the tall rib, short rib, and web regions of the cast billet and optimized preform forgings. While both forgings showed comparable yield and ultimate tensile strengths, the preform forging showed a significant increase in fracture strain, suggesting improved toughness. The yield, ultimate tensile strength, and fracture strain values are provided in Table 4-3. Stress-controlled cyclic tests were also performed at two different stress amplitudes for several samples extracted from all three locations. As shown in Figure 4-12(b), samples extracted from the preform forgings failed at higher cyclic lives compared to their cast-billet counterparts. This improvement stems from a more uniform distribution of the local effective strain and an adequate level of strain energy being imparted onto the preform during forging to trigger dynamic recrystallization. For a more complete account of experimental results correlating microstructural features with mechanical properties, refer to [85].

Table 4-3: Yield and ultimate tensile strengths and fracture strain

	Sample	Yield Strength (MPa)	Ultimate Tensile Strength (MPa)	Fracture Strain (%)
Short rib	Cast-billet forging	192	258	2.75
	Optimized preform forging	206	305	7.89
Web region	Cast-billet forging	220	302	4.28
	Optimized preform forging	222	302	5.71
Tall rib	Cast-billet forging	166	260	5.5
	Optimized preform forging	174	246	1.9

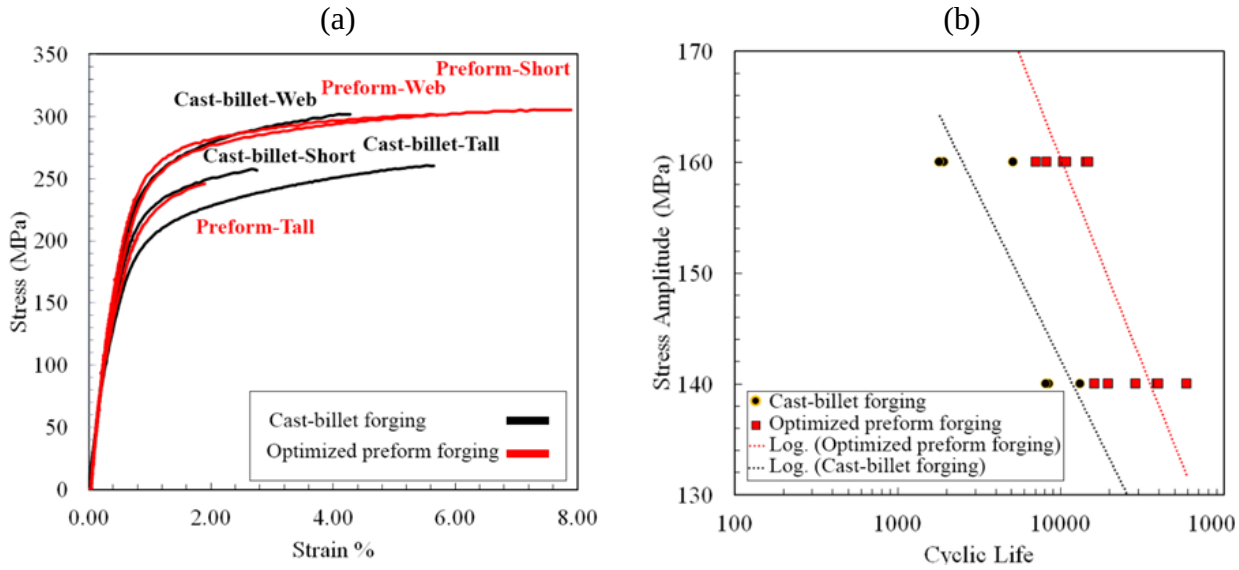


Figure 4-12: (a) Quasi-static tensile strain-stress curves for the cast-billet and optimized preform forging samples extracted from different locations; (b) S-N curves for the cast-billet and optimized preform forging samples extracted from different locations tested at 140 and 160 MPa stress amplitudes

4.5 Conclusions

In this chapter, two parametric CAD models were introduced, each with the specific task of generating CAD models for training feedforward networks for predicting forging outcome in spatially varying regions within the die. Parametric CAD model-1 was used to generate input data for training feedforward models to predict the average effective plastic strain in four regions of interest within the forging die. Design points were sampled from a 7-dimensional design space (with each dimension ranging between 0 and 1, inclusive) using LSHMDU. The elements of the coordinate vector of a design point are the weight parameter inputs to the parametric CAD model, which triggers the procedural generation logic embedded within the model's operational graph to generate closed-surface geometries of preforms. In total, 300 design points were sampled to create a dataset of 300 preform CAD models. DEFORM®-3D [27] was used to run low-fidelity forging simulations of the CAD models, requiring a total simulation run time of about 50 hours, approximately 10 minutes per simulation (the preform geometry was meshed with approximately 30K elements). Then, the simulations were post-processed to curate a dataset, which was used to train feedforward models to predict forging outcome (average effective plastic strain) in spatial varying regions of interest. In general, the trained models predicted forging responses to within ±8% of the ground truth.

The following insights regarding the construction of parametric CAD models were drawn from this preliminary work:

- When drawing immutable (basis) curve geometry to be linearly combined in downstream operations, it is important to ensure that shape similarity between basis curves is as low as possible to ensure a wide range of geometries can be generated.

- To prevent a many-to-one mapping scenario between design and shape spaces, any two basis curves should only be combined once.

A data-driven multi-objective design optimization framework was developed following this preliminary work to optimize the preform shape for an AZ80 alloy I-beam cast-forging process. It consisted of three components: parametric CAD model-2 for shape generation, a new set of feedforward networks for forging outcome prediction, and an SPEAII algorithm to drive multi-objective optimization. Using this framework, an optimal I-beam-like preform shape was obtained and manufactured in a sequence of cast-forging steps. The forging outcome of the cast preform was compared to that of a cast cylindrical billet using simulations and lab-scale experiments. In general, forging simulations showed good agreement with the experiments in terms of the predicted flash volume—the measured flash-to-impression volume percentage was 12% and 6% for the cast-billet and preform forgings, while the simulation-based predictions were approximately 15% and 7% for the cast-billet and cast-preform. Quasi-static tensile and stress-controlled cyclic tests were conducted to evaluate mechanical properties. Compared to the cast-billet forging, on average, the optimized preform forging exhibited a 4.3%, 4.2%, and 51.6% improvement in yield, ultimate tensile strength, and fracture strain, respectively. Comparable yield and ultimate tensile strengths were observed in both forgings, and a significant increase in fracture strain was observed in the preform forging, suggesting improved toughness, fulfilling objective (ii) of the SPG project.

The advantages of using this framework are described below:

- The ability to explore multiple design solutions within a large design space. This approach is suitable during an initial product design phase where there is little accrued knowledge regarding particular designs' performance advantages and disadvantages.

- While a substantial number of simulations need to be run to curate a training dataset, user involvement in this process can be near-fully eliminated by using a data pipeline to automate simulation assembly, submittal to the FEM solver, and post-processing (as implemented in this work).

The limitations of this framework are described below:

- A parametric CAD model is only able to explore designs within the confines of its design space; the size of its design space is fixed, and the model cannot extrapolate designs that lie beyond the bounds of its design space.

- To output geometries with greater shape variation, i.e., shapes from drastically different topological spaces from the one containing its core design, a more complex set of procedural generation logic would need to be encoded in the operational graph. Herein lies a major issue: at a certain point, it becomes increasingly more challenging to hand-engineer a design space that is both diverse— including designs from varying topological spaces— and defined using as few design parameters as possible.

- The data-driven models used in this framework do not leverage an obvious relational inductive bias, the spatial locality in strain response, i.e., the similarity in effective plastic strain between neighbouring regions of interest. Leveraging this inductive bias may significantly improve data efficiency, making it possible to train neural networks with fewer data samples. Alternate neural network architectures like GNNs [86], [87] might be better suited in this case over simple feedforward neural networks or even non-parametric data modelling approaches using Gaussian Processes [88].

5. Topology-Based Optimization Framework

Due to computational limitations and SPG project timeline constraints, the evolutionary design optimization framework introduced in the previous chapter was not used to optimize the preform design of the control arm (FLCA) forging. Instead, a manual iterative design process that mirrored a Bi-directional Evolutionary Structural Optimization (BESO) workflow was followed to fulfill objective (iii) of the SPG project (refer to Appendix C.1 for more information). The knowledge gained from this work was used to develop the fully automated topology-based optimization framework presented in this chapter.

5.1 Introduction

When developing a parametric CAD model that can effectively parameterize the shape features of a complex geometry such as FLCA, more weight parameters are required. The complexity of the geometry is loosely defined here based on the planarity of the centerline or spine of the geometry. As an example, the centerline of the I-beam geometry is planar, a straight line, whereas the centerline of the control arm geometry is non-planar, with a curvature that varies in all three dimensions along its length. Typically, more parameters will be required to characterize a complex preform shape. Using more parameters increases the dimensionality of the parametric CAD model, which translates to needing more training data to train data-driven models—the curse of dimensionality.

An alternative approach to shape optimization is presented in this chapter, in a fully automated topology-based optimization framework with a similar workflow to a BESO algorithm. In a traditional BESO algorithm, the starting preform geometry needs to be modelled manually between each iterative cycle to convert a voxelized mesh[5] (a form that enables voxels elements to be added and removed from a mesh [32]) to a surface mesh representation, which used in the FEM model. While algorithms such as Marching Cubes[6] [89] and Laplacian smoothing[7] [90] can be used to automate the conversion between the two mesh representations, it is still necessary to make manual adjustments to the surface model to ensure that there is no preform-die interference. The presented topology-based optimization framework eliminates manual modelling efforts during the optimization cycle by using computational algorithms to directly operate on the underlying curve geometry (the wireframe) that make up a closed-surface model of a preform. In addition, this framework runs on a single design environment (currently in the Rhinoceros® [47] workspace) to

[5] A voxel mesh is an object that is divided into a grid of voxels (cubes), with each voxel representing a small volume of the object.
[6] An algorithm for converting voxel data to surface meshes
[7] Laplacian smoothing is an algorithm used to smooth a polygonal mesh by updating the postions of its vertices based on local position information of neighboring vertices.

automatically post-process DEFORM®-3D [27] simulation data, and to regenerat it in the workspace, providing an integrated display of both the surface model and the simulation mesh at any step during deformation.

The contribution of this chapter is this framework, which includes the procedural generation algorithms that drive automated surface modelling and the optimization algorithm that interprets simulation response data to formulate the directives to guide surface modelling. A description of how these algorithms operate is presented in this chapter, along with a demonstration of its functionality on a toy problem—to iteratively distribute material of a starting preform shape within a volumetric segment of the FLCA forging die. The code libraries of the framework were developed using C# and the RhinoCommon[8] SDK.

5.2 Methods

The topology-based optimization framework runs the procedural generation algorithm and the optimization algorithm with the DEFORM®-3D [27] FEM solver (in the loop) to iteratively modify the topology of a staring preform shape according to the following steps (see Figure 5-1):

1. The user passes immutable geometry and weight parameters as inputs to the procedural generation algorithm to generate the starting surface mesh model of the preform.

2. The mesh model is imported into a forging simulation template to generate a simulation database file, which is then submitted to the FEM solver.

3. Once the simulation has completed running, simulation mesh and state-variable data are post-processed by the optimization algorithm, making it possible to visualize all deformed mesh states from the Rhinoceros® [47] workspace, as well as a strain tracked (or backtrack) mesh, which indicates the regions on the preform mesh where material needs to be added or removed (effective plastic strain is the only state-variable information backtracked in the current version of the algorithm).

4. Backtracked state-variable information is evaluated based on a heuristic, which is defined in a state-action lookup table to prescribe point displacement values as actions to direct the procedural generation algorithm on how to modify the underlying wireframe geometry of the surface model. Once the topology update is made, the algorithm restarts from step 2.

[8] A software development kit (SDK) for Rhino, providing libraries that enable development with Rhino geometry.

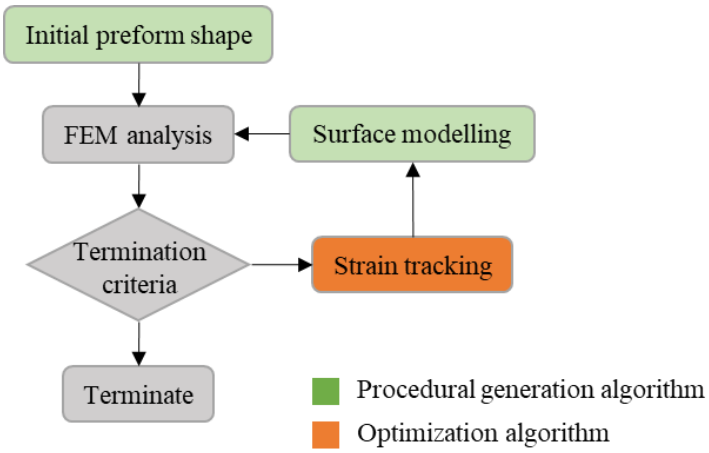


Figure 5-1: A flow diagram of the fully automated topology-based optimization framework; initial preform shape generation and surface modelling are implemented using the procedural generation algorithm; the optimization algorithm carries out strain tracking.

5.2.1 Procedural Geometry Generation Algorithm

At a high level, the procedural generation algorithm constructs a surface mesh by utilizing a series of 2D cross-sectional curves stationed along 3D curves, referred to as rails. The surface mesh is generated by sweeping[9] cross-sectional curves along the rails, ensuring that it does not interfere with the die so that the generated preform can be securely placed on the die. Additionally, the algorithm forces the endpoints of the cross-sectional curves to coincide with the rails to guarantee a valid closed-surface mesh model output. The procedural generation algorithm generates the initial geometry and iteratively updates this geometry based on inputs from the optimization algorithm. Geometric inputs to initiate the algorithm include the *ImpressionMesh, CentreRail, OutboardRail,* and *InboardRail*, and the numerical inputs include *FrameCount,* and *WeightParameters* (see Figure 5-2). The algorithm executes eight procedural generation logic steps. Upon initialization of the optimization loop, the initial preform geometry is generated by executing steps 1 to 5. Then, at every iteration, steps 5 to 8 are executed to update the topology of the preform geometry of the previous step.

Step 1 involves generating a number of cross-sectional planes along the centre rail at equidistant points, with their respective Z and X coordinate axes pointing in perpendicular directions to the tangent vector of the centre rail at each plane-curve intersection point (at the plane's origin), based on the number of frames. All curve geometry generated in subsequent steps is drawn on these planes (the *ReferenceFrames*). Then, for each reference frame, closed curves of the impression mesh boundary are extracted and separated into top and bottom curve segments (*TopImpressionCurve.Boundary* and *BotImpressionCurve.Boundary*). Additional curve trimming

[9] In surface modelling, a sweep operation is used to create a surface by sweeping a 2D profile (cross-sectional curve) along a 3D path (rail).

operations are performed on the boundary curves to ensure that the curve geometry does not extend out of the impression volume and into the flash (see Figure 5-2).

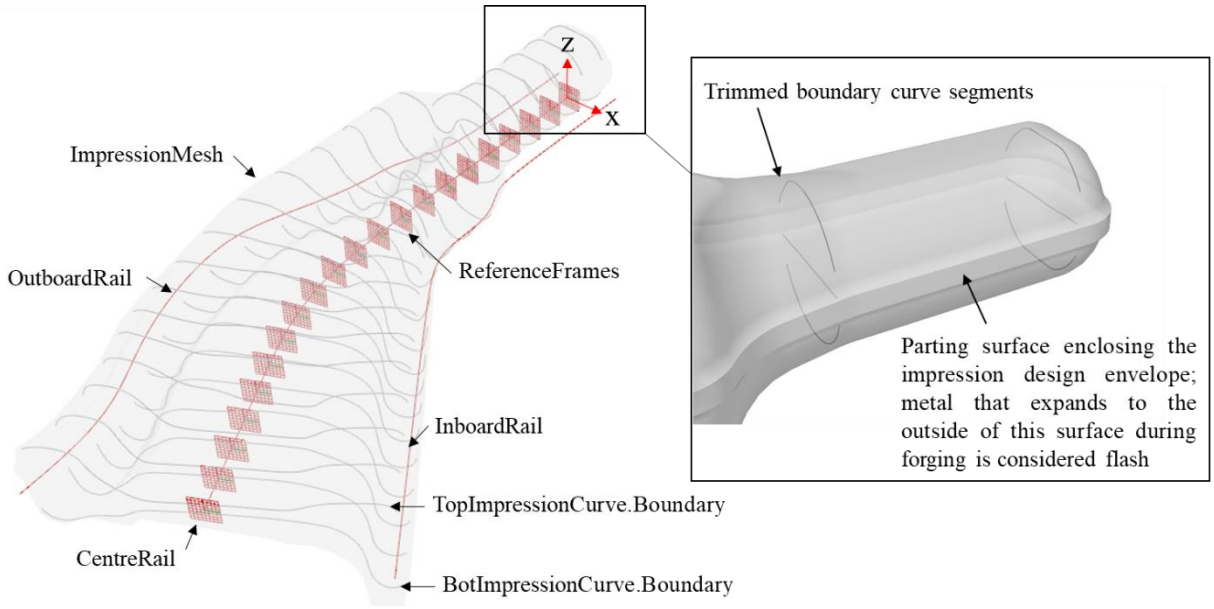


Figure 5-2: Curve and mesh geometry inputs that are used to extract boundary curves are shown.

Step 2 involves iterating through the boundary curve geometry of each reference frame to create a new set of curves (basis curves), which are copies of the top and bottom boundary curves that are internally offset by a small amount. The basis curves are the primary set of construction curves used for creating cross-sectional curves for the preform. Potential preform-die interference issues are mitigated by offsetting these curves from the impression mesh surface (note the impression mesh is used to create the negative of the forging die). Additionally, a parting curve (another primary construction curve) is created, which passes through the mid-point between the top and bottom boundary curves and along the local X-axis of the reference frame. The parting curves split the internal area equally and are used to establish the allowable minimum and maximum web width and height of each cross-section—to establish a design envelope (see Figure 5-3).

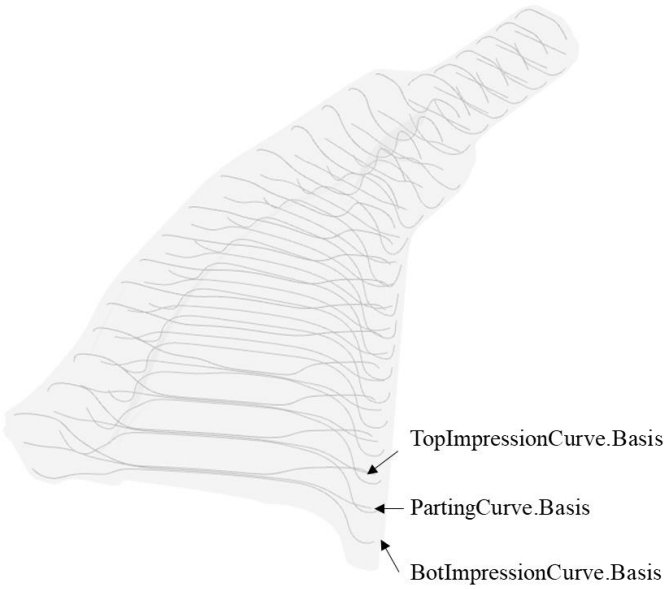


Figure 5-3: The basis curve geometry is shown.

Step 3 involves generating a preliminary set of rails for sweep operations. In total, three weight parameters, $\vec{a} \in [0.0, 1.0]^3$, are used to generate the starting preform shape. Weight parameter a_1, is used to construct the rails that run along the length of the preform geometry (*Rail.A1* and *Rail.T1)* (see Figure 5-4). The rails are created by shooting rays in the positive and negative global Z-direction from equidistant points along the centre rail. Then, non-uniform rational B-spline (NURBS[10]) curves are created at each ray-mesh intersection point; the NURBS curve construction function performs a weighted interpolation between the intersection points to generate smooth curves (see Equation 31 and Figure 5-5).

$$C(t) = \frac{\sum_{i=1}^{n} N_{i,p}(t) w_t P_i}{\sum_{i=1}^{n} N_{i,p}(t) w_t}, \tag{31}$$

where $C(t)$ is the point on the NURBS curve at parameter t, $N_{i,p}(t)$ is the i-th B-spline basis function of degree p, w_i is the weight associated with the i-th control point, and P_i is the i-th control point [91]. Then, based on the value of weight parameter a_1, the top rail (*Rail.A1*) is translated along the global Z-direction, which has the effect of increasing or decreasing the overall web thickness of the generated preform geometry.

[10] A Non-Uniform Rational B-Spline (NURBS) curve is a mathematical representation of a 2D or 3D curve using four paramters: the degree, control points, knots, and an evaluation rule [78]

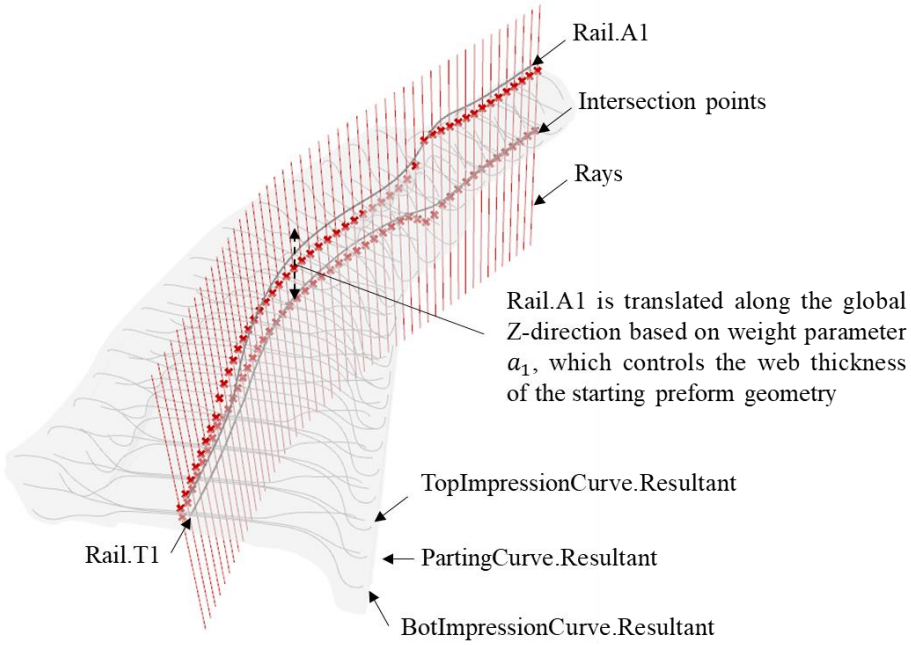


Figure 5-4: Construction geometry used to obtain Rail.A1 and Rail.T1 are shown.

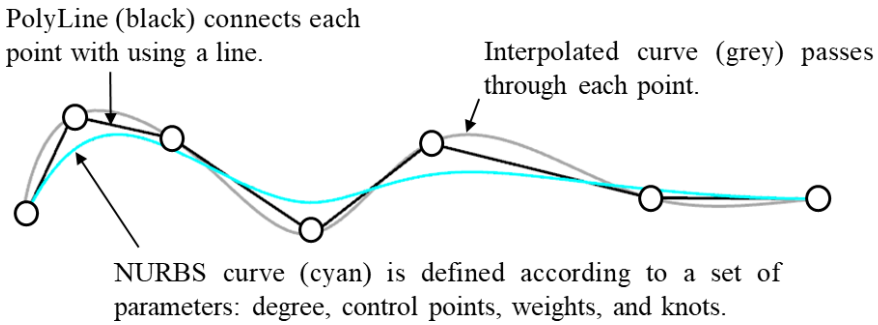


Figure 5-5: The difference between a polyline (black), an interpolated curve (gray), and a NURBS curve (cyan) is shown.

Step 4 involves constructing a second set of curve (resultant curve) geometry for each reference frame, used for creating the cross-sectional curves of the wireframe model. Resultant curves are copies of basis curves that are forced to guarantee coincidence with top and bottom rails. Newly formed intersection points between *Rail.T1* and *Rail.A1* are called ReferencePoint.T1 and *ReferencePoint.A1*, respectively (see Figure 5-6). Then, the remaining weight parameters, a_2 and a_3, are used to define the locations of points, *ReferencePoint.C1*, and *ReferencePoint.S1* along the parting curve domain (between the minimum and maximum allowable range highlighted in orange in Figure 5-6).

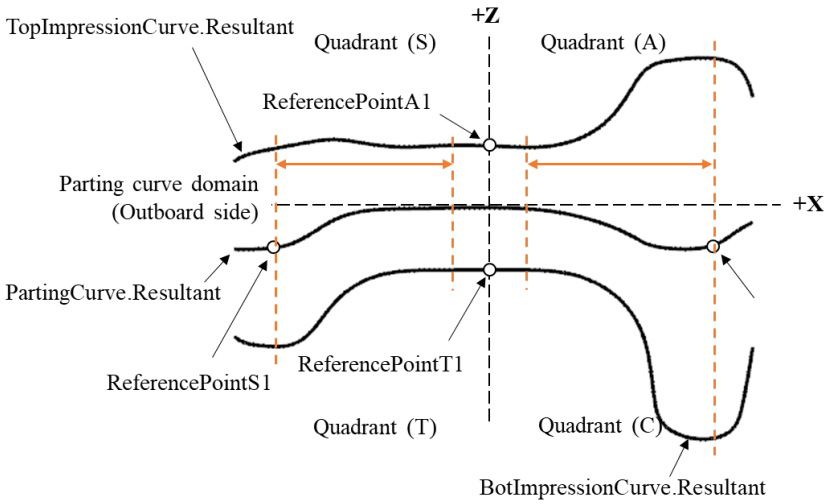


Figure 5-6: Curve geometry of a cross-section that is projected onto the ZX-plane, illustrating quadrants C, A, S, and T; quadrants C and A lie on the inboard side, and quadrants S and T lie on the outboard side; the permitted range along the parting curve where reference points can be placed is indicated using orange hidden lines.

Step 5 involves translating control points based on their displacement value inputs. In total, 7 displacement values are used to update the coordinates of 7 control points (the control points are translated with respect to the reference points that were created in step 4). The start and end points of curve segments of the underlying wireframe model intersect the control points. When the optimization loop is initialized, all point displacement values are set to zero to generate the reference (or starting) preform design. Each quadrant of a cross-sectional consists of three control points (see Figure 5-7). An additional point, *CentrePoint.A2*, is used to independently control the web height of each cross-section by translating the (resultant) top impression curve (shown in cyan in Figure 5-7) along the Z-direction. The other two centre points control the web thickness by morphing the (resultant) impression curve. The rib points are translated horizontally along the impression curve profile to either increase or decrease the web width of each cross-section. Lastly, parting points are dynamically calculated based on the locations of rib points.

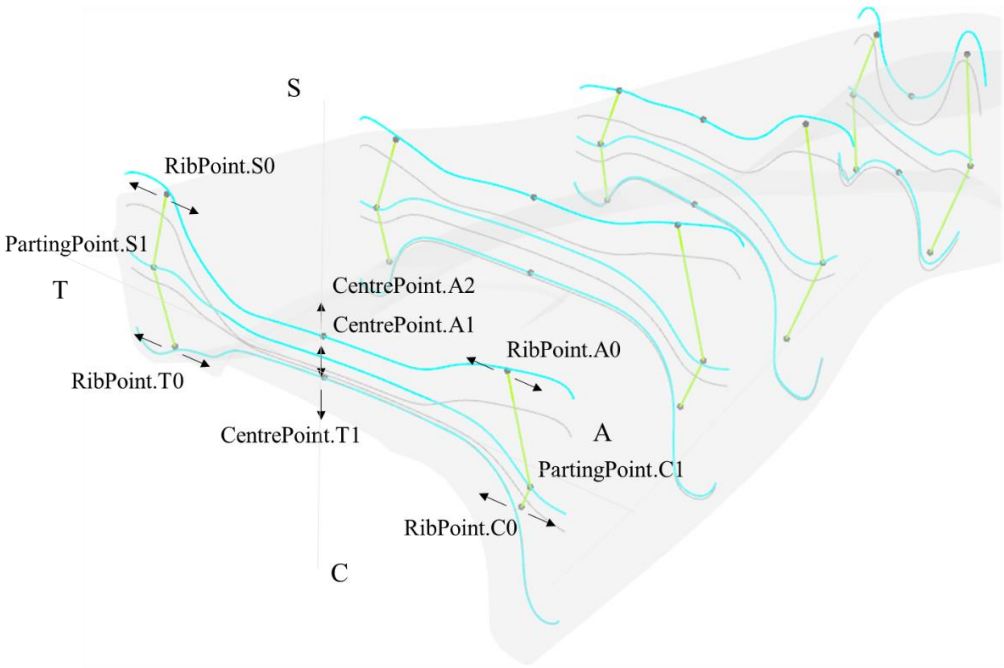


Figure 5-7: General point displacement directions along a cross-section; control point quadrant designation and point number based on a counter-clockwise ordering convention are shown.

Step 6 involves re-constructing the top and bottom rails, *Rail.A1* and *Rail.T1*, to ensure the rails pass through their respective control points. Rails, *Rail.C0*, *Rail.C1*, *Rail.S0*, *Rail.S1*, and *Rail.T0* are constructed using their respective control points using the NURBS curves construction function (see Figure 5-8).

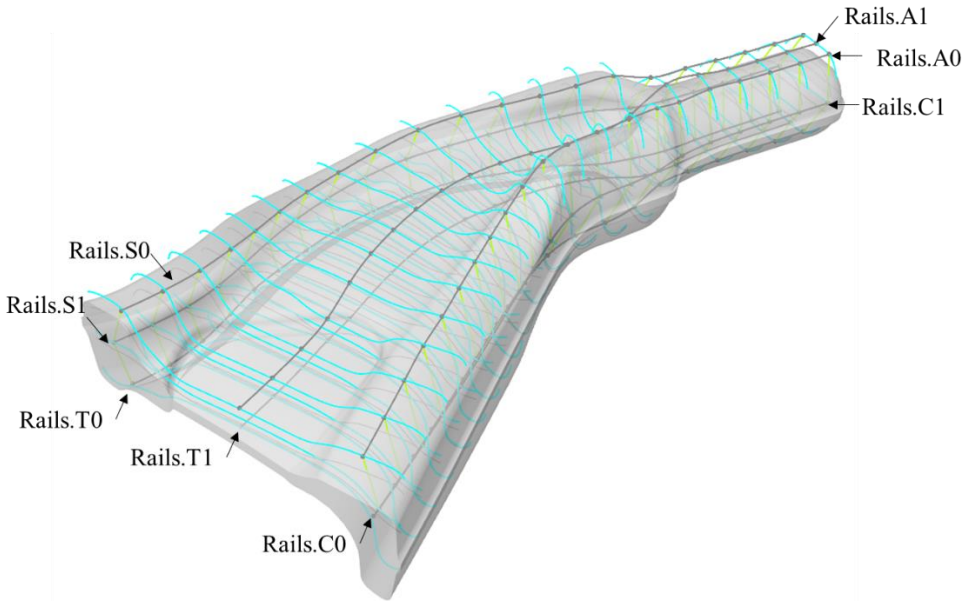


Figure 5-8: Rail curves and the intersection points are shown.

Step 7 involves creating curve segment (quadrant curve) geometry. Quadrant curve geometry consists of two curve segments, a left-side curve (*LeftSegment*) and a right-side curve

(*RightSegment*) (see Figure 5-9). First, a right curve segment is created, connecting the rib point to the parting point. Then, it is modified based on whether or not it intersects the impression curve (see the curve modification made based on the Boolean expression, *isRightSegmentIntersecting* in Figure 5-9). If it does intersect the impression curve, the portion of the curve segment that extends outside of the impression curve (outside the design envelop) is trimmed, and the impression curve segment between the rib point and the intersection point is extracted and merged with the trimmed right curve segment (see Figure 5-9(b)).

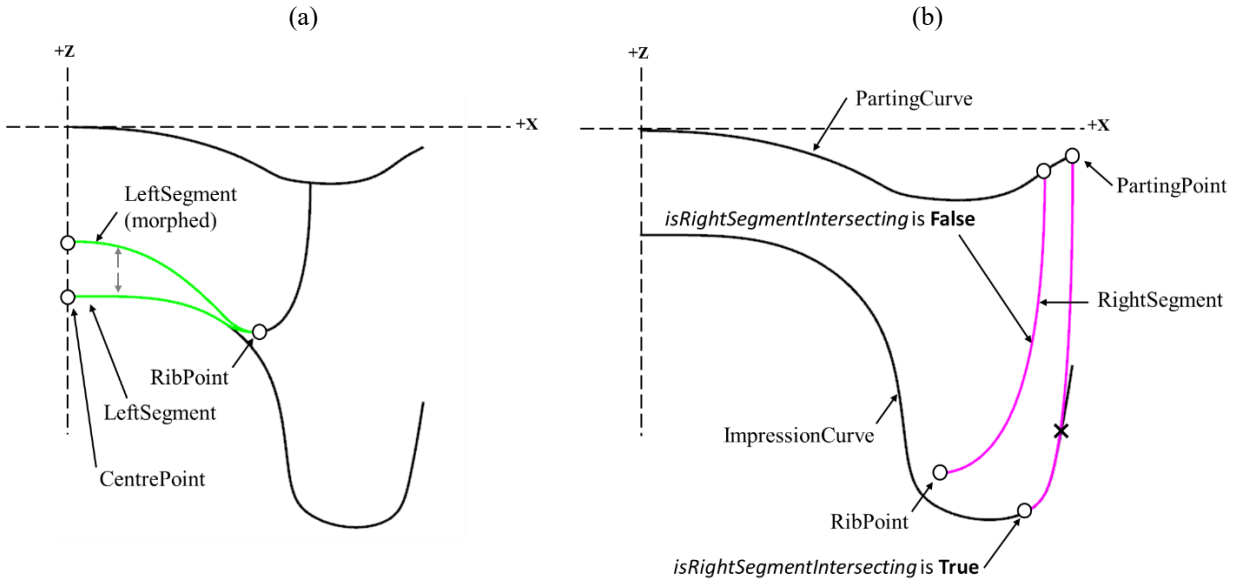


Figure 5-9: (a) LeftSegment in its morphed and original configurations are shown; (b) two RightSegment curves are shown, one generated with and another without a secondary curve modification step.

Next, the left curve segment is created, connecting the centre point to the rib point. One of four procedural generation logic sequences is executed to construct the left curve segment based on the combined expression of two Boolean values (refer to Figure 5-10, which shows the combination of Boolean values *isExtLinePointIntersecting*, and *isRibPointOnImpressionCurve* and the specific construction logic sequences that are triggered). The Boolean value, *isExtLinePointIntersecting*, conveys information on whether or not a horizontal line drawn from the rib point towards the origin of the frame intersects the impression curve. The Boolean value, *isRibPointOnImpressionCurve*, conveys information on whether the rib point lies on the impression curve. When constructing the right and left curve segments, it is necessary to evaluate the respective test conditions to execute the correct set of procedural generation logic sequences, to guarantee that the generated curve segments lie within the design envelope and pass through the respective control points. Lastly, both the right and left curve segments are merged to form a unified curve, the *QuadrantCurve* (see Figure 5-11).

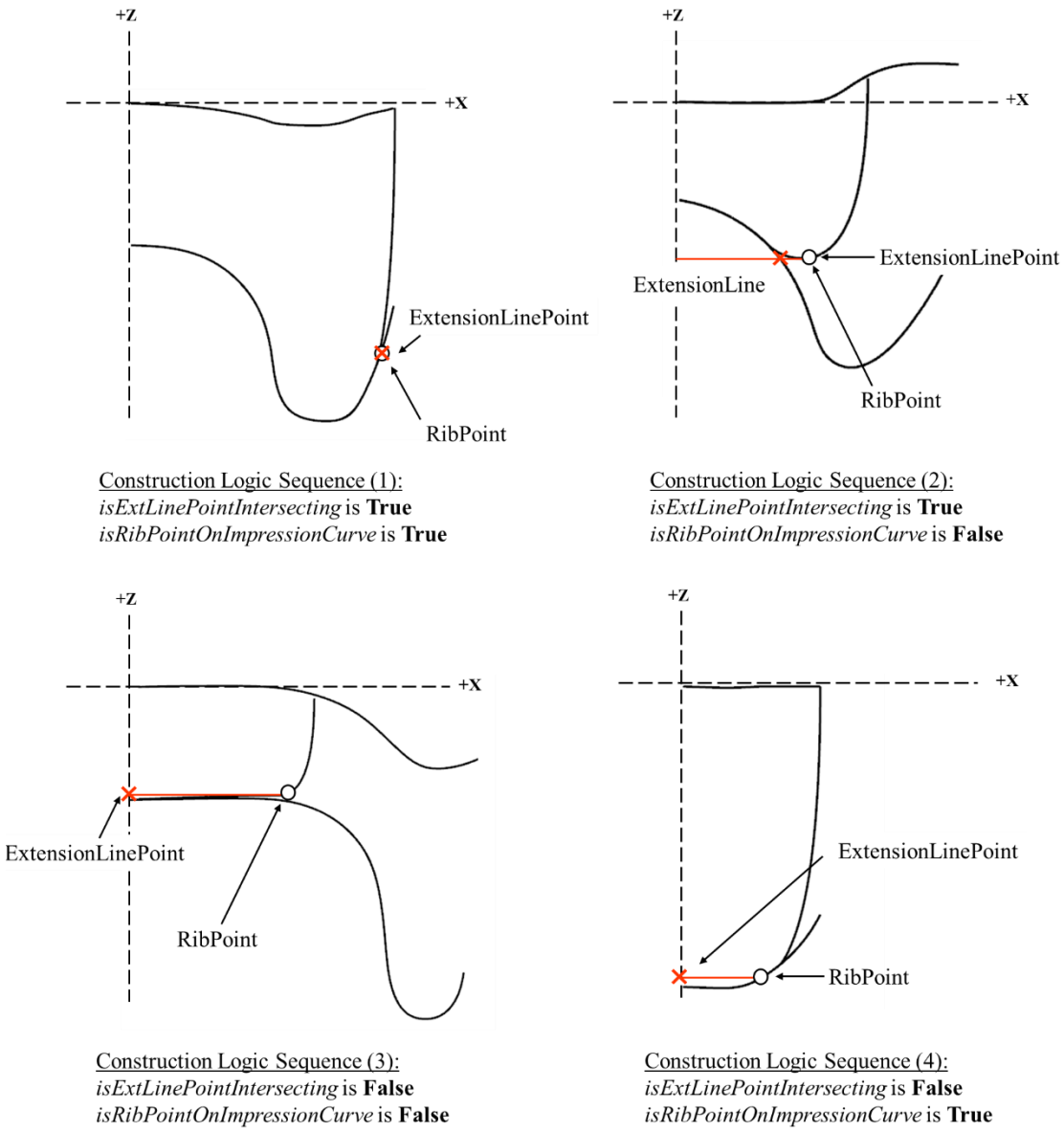


Figure 5-10: One of four construction logic sequences is executed based on the tested Boolean Conditions.

Finally, in step 8, the surface model of the preform is constructed. The rails and quadrant curves are sorted in a counter-clockwise orientation and used in sweep operations to generate surface panels, which are then joined together to form a closed-surface geometry and a surface mesh geometry of the preform (see Figure 5-11 and Figure 5-12).

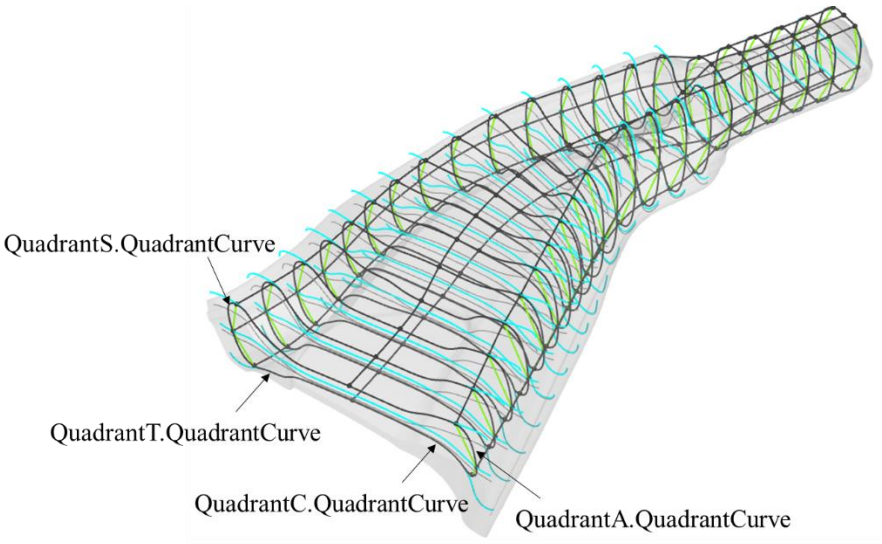


Figure 5-11: The curve geometry that is used to sweep surfaces is shown.

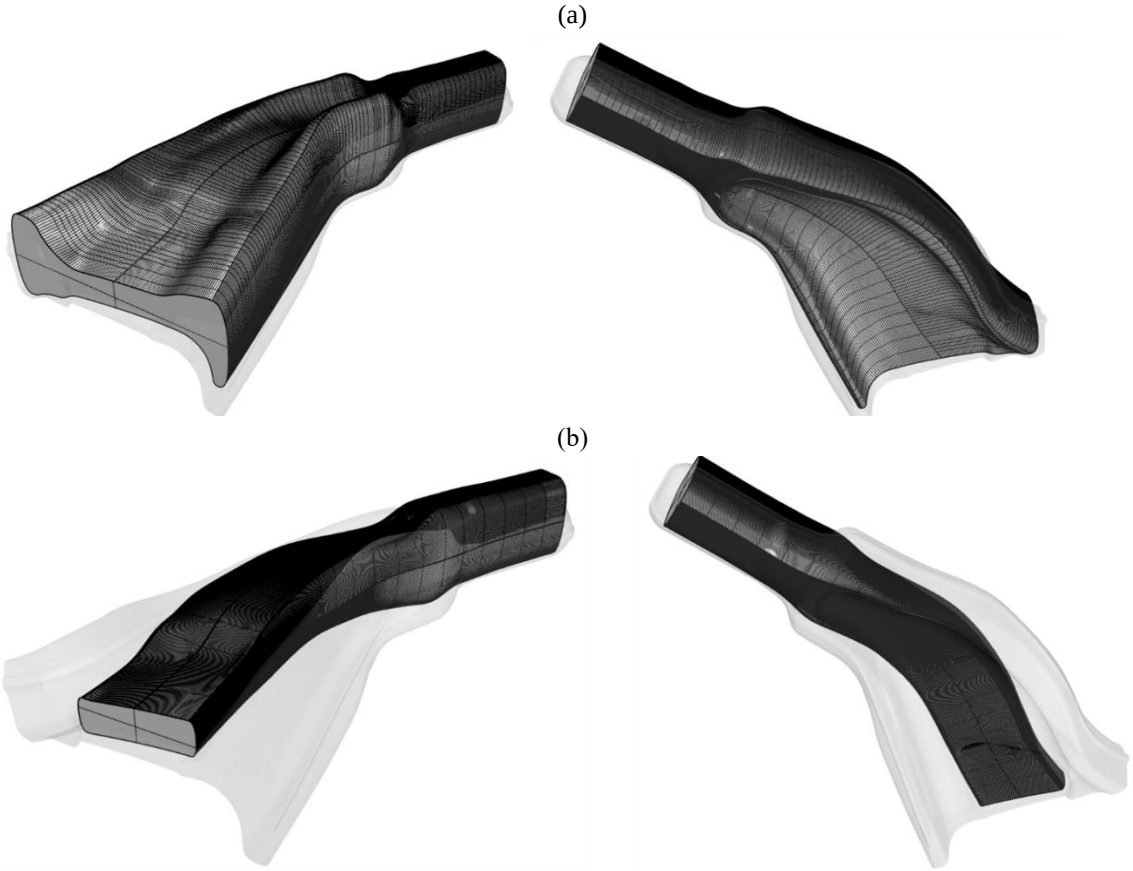


Figure 5-12: (a) Closed-surface preform model generated using weight parameter values: 0.5, 0.8, and 0.8; (b) Closed-surface preform model generated using weight parameter values: 0.3, 0.1, and 0.1.

5.2.2 Optimization Algorithm

The optimization algorithm runs a sequence of post-processing operations to extract mesh and state variable information (only effective plastic strain is extracted in this version of the algorithm) from forging simulations, for each step of the simulation. This information is then used to calculate the state of each preform cross-section and the set actions (point displacement values) to signal the procedural generation algorithm on how to displace the control points of each cross-section.

Step 1, in this algorithm, involves processing simulation output text files—the Deform3D® [27] post-processor was configured to output three text files containing mesh vertex, element connectivity, and strain information of every forging simulation step. The large corpus of text data is processed to output structured numerical data arrays. Next, the mesh element data is reduced to element centroids to generate point clouds with effective plastic strain point attributes. Remeshing[11] steps are also identified in this step to support strain tracking (or backtracking)—mapping the strain field of the last point cloud to the first point cloud. The strain field between dissimilar point clouds (before and after a remeshing steps) is mapped from the leading cloud to the trailing cloud based on a *k*-nearest neighbour interpolation approach by assigning the average strain value of *k* closest points in the leading cloud to a target point in the trailing cloud; this operation is carried out for each point in the trailing cloud. Once the strain field of the last point cloud has been mapped back to the first point cloud, it is imposed onto the vertices of the starting mesh, making it possible to identify the regions where material needs to be added or removed—based on the simple optimization heuristic which is defined in a lookup-table (see Figure 5-13).

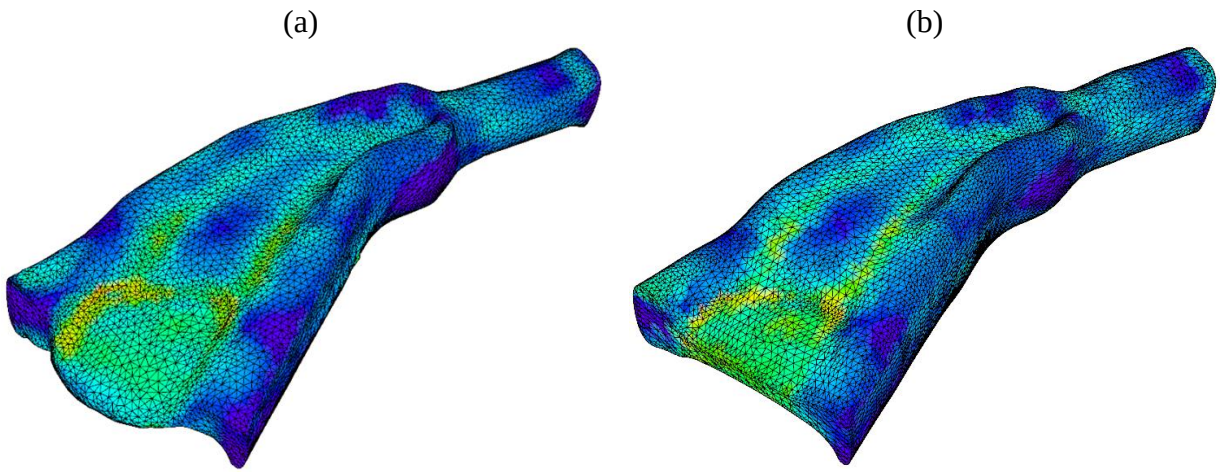


Figure 5-13: (a) Shows the mesh state at the last step of the forging (open-die forging) simulation where metal is pushed out in both lateral and longitudinal directions; (b) shows the strain field corresponding to the last simulation step imposed on the starting preform mesh.

[11] Remeshing is crucial in large deformation simulations to maintain mesh quality so that material deformation behavior is accurately represented

In total, nine states are defined in the lookup table based on a combination of strain values at the centre, parting, and rib control points of a quadrant (see Figure 5-14). Based on the state of a cross-section, the actions involve either prescribing a positive point displacement value of a fixed step size to add material, prescribing a negative point displacement value of a fixed step size to remove material, or prescribing a point displacement value of zero.

Quadrant - C - State - 1									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4	x	x	x		Add	x	x	x
	0.4 < ε < 0.7					Nominal			
	ε > 0.7					Remove			

Quadrant - C - State - 2									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4	x	x			Add	x	x	
	0.4 < ε < 0.7			x		Nominal			x
	ε > 0.7					Remove			

Quadrant - C - State - 3									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4	x	x			Add	x		x
	0.4 < ε < 0.7					Nominal			
	ε > 0.7			x		Remove		x	

Quadrant - C - State - 4									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4			x		Add		x	
	0.4 < ε < 0.7	x	x			Nominal	x		x
	ε > 0.7					Remove			

Quadrant - C - State - 5									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4					Add			
	0.4 < ε < 0.7	x	x	x		Nominal	x	x	x
	ε > 0.7					Remove			

Quadrant - C - State - 6									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4					Add			x
	0.4 < ε < 0.7	x	x			Nominal	x		
	ε > 0.7			x		Remove		x	

Quadrant - C - State - 7									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4			x		Add			x
	0.4 < ε < 0.7					Nominal		x	
	ε > 0.7	x	x			Remove	x		

Quadrant - C - State - 8									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4					Add			
	0.4 < ε < 0.7			x		Nominal		x	x
	ε > 0.7	x	x			Remove	x		

Quadrant - C - State - 9									
Strain Value		State			Translate		Action		
		CentrePoint	RibPoint	PartingPoint			CentrePoint	RibPoint	TopCurvePoint
	ε < 0.4					Add			
	0.4 < ε < 0.7					Nominal			
	ε > 0.7	x	x	x		Remove	x	x	x

Figure 5-14: The state-action lookup table encapsulated in the optimization heuristic is shown; the state of each control point is quantified based on the average effective plastic strain, and the action is prescribed based on its corresponding state.

A total of 12 strain responses are calculated per cross-section based on the average strain of mesh vertices within spherical regions centred at each control point. Control points that are shared between quadrants are displaced based on the average value of the prescribed point displacement values. Referring to Figure 5-15, the state of a quadrant is assessed based on whether the average strain value corresponding to a control point is below 0.4, between 0.4 and 0.7, or greater than 0.7. A minimum strain value of 0.4 was used here as it was noted to be the point beyond which DRX effects were significantly more visible in wedge-cast compression samples [13]. Based on the observed state, the action response is as follows: if the average strain value is below 0.4, add material to that region; else, if the average strain value is greater than 0.4 and less than 0.7, neither add nor remove material; if the average strain value is greater than 0.7, then remove material (the choice of the upper-bound strain value of 0.7 was determined arbitrarily). After the state of each cross-section has been evaluated, the corresponding actions conveying point displacement values are sent as inputs to the procedural generation algorithm—to modify the locations of control points, which triggers a reconstruction of quadrant curves and rails that make up the wireframe and the surface model.

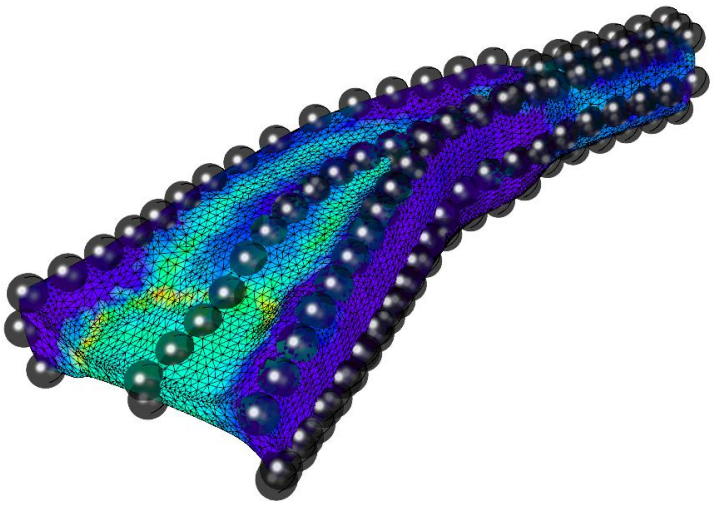


Figure 5-15: Spherical regions centred about construction points are generated to evaluate the average effective strain within the surrounding region to evaluate the “state” of a cross-section

Figure 5-16 illustrates the cross-sectional curve shape changes that occur given point displacements. Figure 5-16(a) shows the original curve shape of an arbitrary preform cross-section; the shape of this section is determined based on the set of input weight parameters and the input geometry that are provided upon initialization of the optimization loop. Figure 5-16(b)-(f) shows the possible curve shape changes given point displacements.

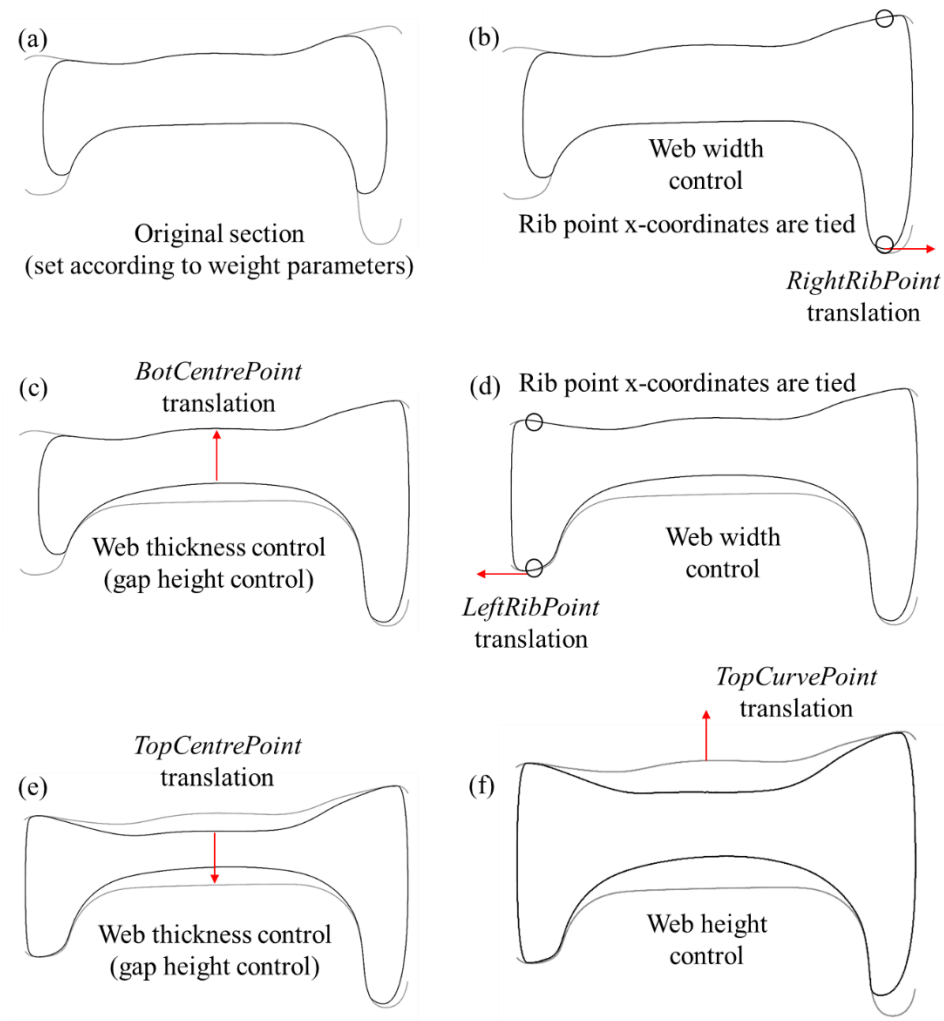


Figure 5-16: (a) Illustrates a cross-sectional curve geometry with zero point displacement values; (b)-(f) illustrates how the shape of the original curve is modified when point displacements are used to translate the control points.

5.3 Results and Discussion

A toy preform shape optimization problem was set up to validate the topology optimization framework. This section illustrates the topology modifications made to the starting preform design based solely on strain response information from simulations. Note that flash and underfill responses were not used in this version of the algorithm. The optimization loop was run for 14 design iterations. No strict termination criteria were set in this problem, and the iteration loop was terminated at an arbitrary point that was deemed sufficient to display the effectiveness of the underlying algorithms based on the improvement in the strain field of forging simulations. Another important fact to note is that a fixed point displacement step size of 2 mm was used from iterations 1 to 10 and increased to 4 mm from iterations 10 to 14 (to assess the quality of topology change when larger displacement values are used). The simulation model template was defined using an isotropic material model, using the flow stress curves of the as-cast AZ80 alloy deformation at 300

ºC. The preform geometries were meshed using 100K elements, while all other simulation parameters were prescribed according to Table 3-5.

The design optimization process begins with the starting preform design at iteration 1, which was automatically generated given a set of weight parameters and input geometry. Inspecting Figure 5-17(a), the initial geometry shows an effective plastic strain response greater than 1.0 at the punch-out region and in a localized region of the outboard rib—roughly all other regions except those surrounding the central web exhibit low strain values less than 0.25. From iteration 1 to iteration 8, there is a steady increase in overall effective plastic strain throughout the forging, which is indicative of the preform being subjected to higher amounts of deformation. This indicates that the optimization algorithm prescribes control point displacements that increase the overall web height and thickness of preform cross-sections while proportionately reducing the web width to prevent excessive flash formation. There is a noticeable underfill in the rib near the punch-out region. This is partly due to the fact that this toy problem is an open-die problem; metal flow in the normal directions to the front and rear-most cross-sectional planes are not constrained. Therefore, there is a lack of inward pressure, which would otherwise force the metal into the rib. However, a more significant fact is that this simple optimization heuristic does not condition action responses based on underfill or flash volume assessments, which would provide additional information that can be used to define a better optimization heuristic. Considering the overall shape of preform designs, iterations 1 to 8 show a smooth evolution of the topology, while design iterations 12 and 14 show pronounced topological variations along the length of its non-planar centerline (spine). The reason behind these observations can be attributed to the fixed displacement step size, which was set to a value of 2 mm from iterations 1 to 10 and 4 mm from iterations 10 to 14. The rough topology observed in design iteration 12 highlights several issues with the existing procedural generation algorithm, the first being related to the smoothness of NURBS curves, which form the generated rails. The smoothness of NURBS curves can deteriorate when the control points are displaced using a fixed step size, allowing for large relative distances in the longitudinal (Z-direction) between neighbouring control points. This displacement upsets the aspect ratio between the lateral (XY-plane) and longitudinal point spacings, which affects the smoothness of the NURBS curves that can be created using the control points (see Figure 5-17(d)). Reducing the number of cross-sections is a simple solution to counteract this effect; however, it will also reduce the level of detail of the preform geometry since there would be fewer cross-sectional curves parameterizing the volumetric shape of the preform. A better solution would be to include an additional curve smoothing step involving a curve subdivision operation to discretize the rough curve using fewer equidistant points, thereby improving the aspect ratio by relaxing the lateral spacing between control points. A curve rebuild operation can then follow this step to rebuild the NURBS curve using fewer control points. Alternatively, a function can be formulated to factor in the effective plastic strain and the minimum distance to the die surface to obtain continuous point displacement values, instead of fixed values. This would also be an effective method to resolve rough topology updates between design iterations (as seen in between Figure 5-17 (c) and (e)).

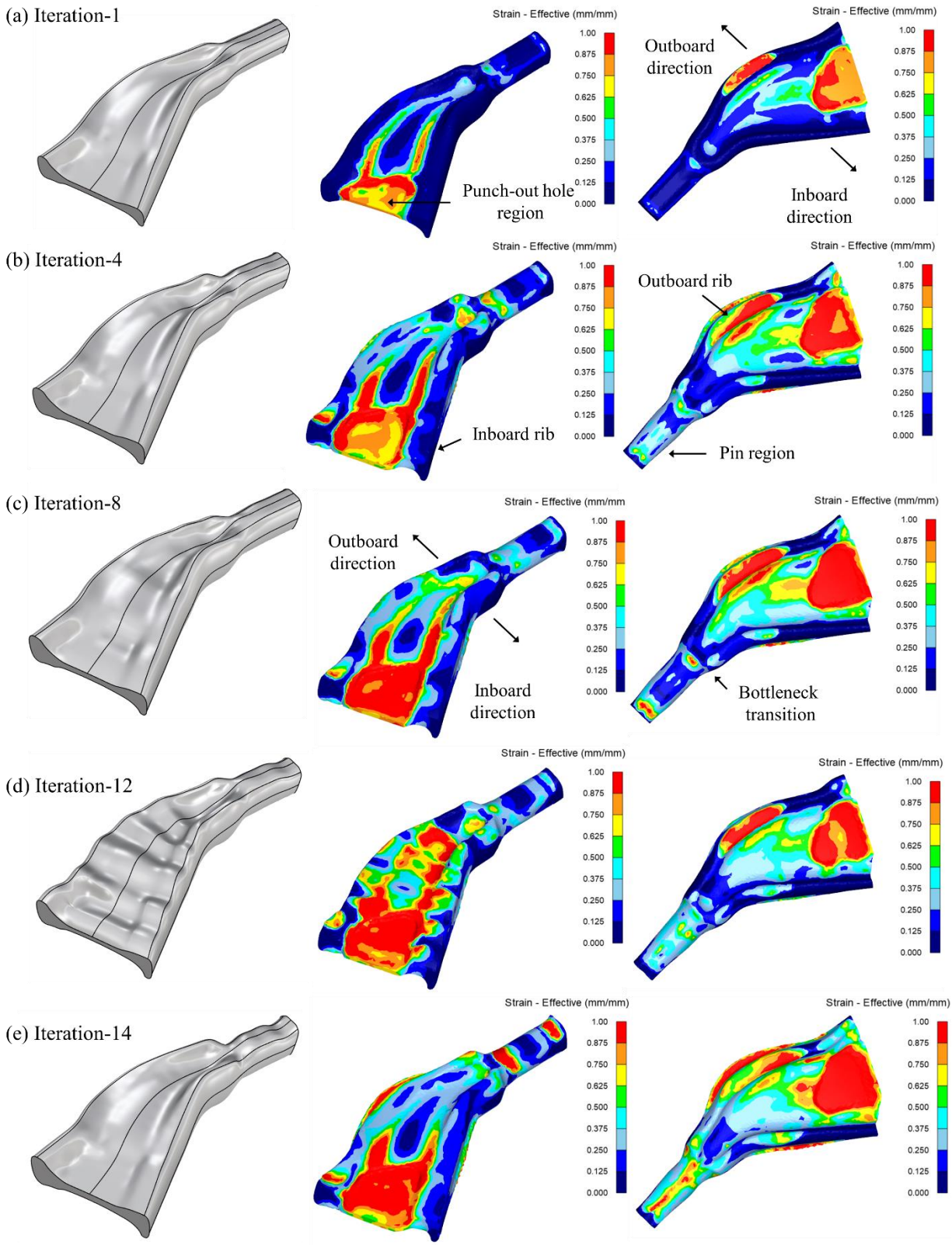


Figure 5-17: Preform shape design iterations are shown along with strain response; metal flow in the directions parallel to the front and rear-most section normals are not shown

A noticeable increase in the effective plastic strain response can be observed in design iteration 14 compared to design iteration 1. The absence of coloured mesh elements in Figure 5-18(a), corresponding to design iteration 14, indicates that a larger surface area on the forging exceeds the effective plastic strain cut-off limit of 0.4. Also, when comparing the minimum distance state variable, which shows the gap distance between the forging and the impression die surface at the last simulations step, the improvement in the material distribution between design iterations 1 and 14 is evident. While the optimization heuristic used in this problem to manipulate the shape of cross-sections is relatively simple, the state-action responses that are defined in the lookup table appear to be effectively guiding topology updates to the starting preform design in a direction that would lead to a favourable forging outcome, with little flash formation and an increased degree of deformation throughout most of the forging. However, there is underfill at the inboard rib near the punch-out hole and at the bottleneck transition near the pin that is not filled adequately, even after 14 design iterations. This outcome was somewhat expected since no direct signalling mechanism indicates the amount of underfill or flash formation in this version of the optimization algorithm. Lastly, since no manual modelling efforts were required to implement shape modifications, the total optimization time was around 7 hours, accounting for only the time required to run 14 simulations (with each simulation taking approximately 30 mins to complete).

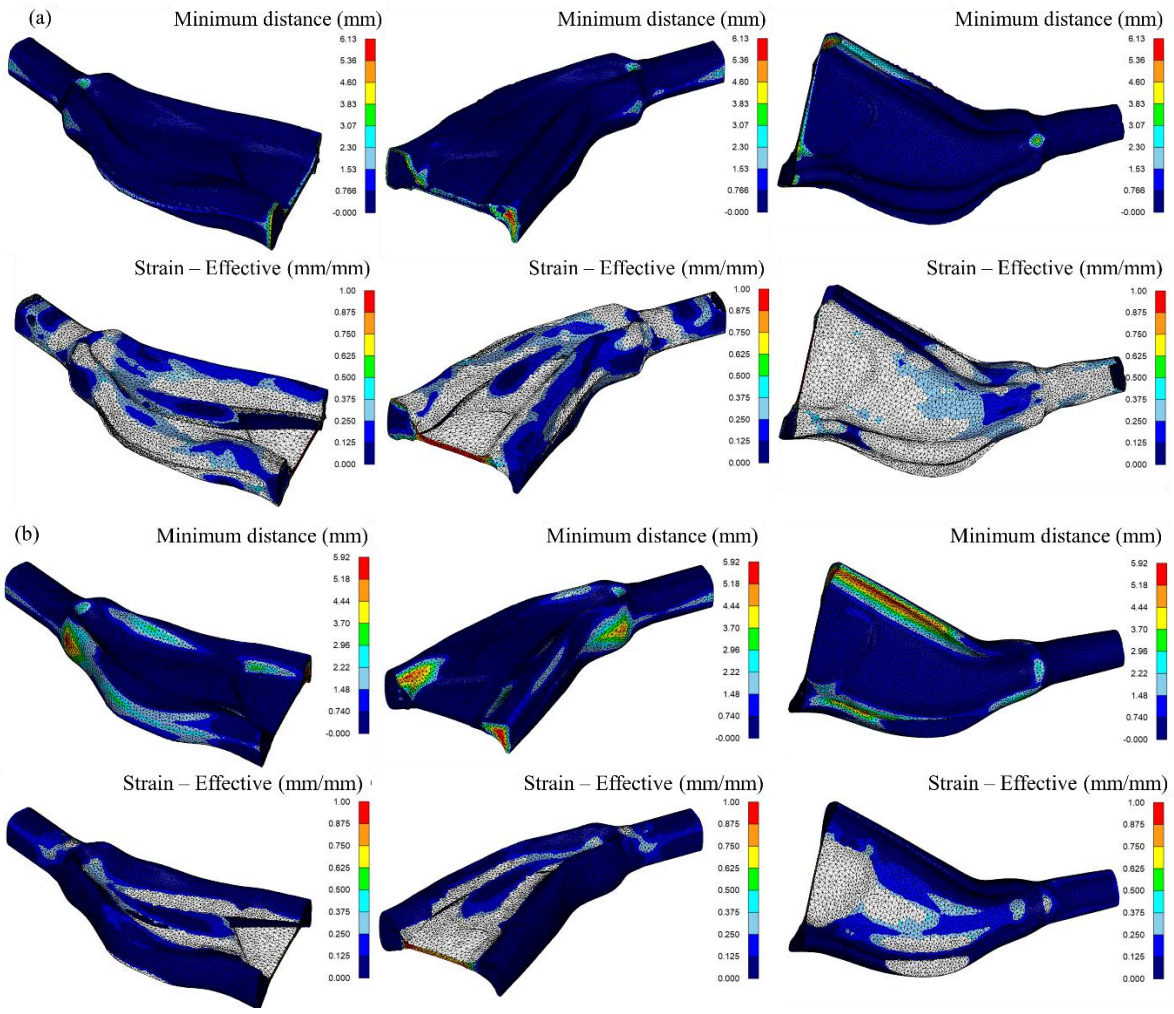


Figure 5-18: (a) Forging outcome of design iteration 14; (b) forging outcome of design iteration 1; the maximum effective plastic strain colour cut-off limit is 0.4.

5.4 Conclusions

This chapter introduces a fully automated topology-based preform design optimization framework to carry out local shape optimization, i.e., the starting preform shape is modified incrementally by taking small steps to update its topology. It comprises of two main algorithms: a procedural generation algorithm and an optimization algorithm. The procedural generation algorithm generates the starting preform geometry based on a set of immutable geometric and numerical inputs and iteratively modifies its topology based on inputs from the optimization algorithm. The framework's functionality is demonstrated on a toy problem to optimize the material distribution of a preform design within a volumetric segment of the control arm forging die. In total, 14 design iterations were carried out. A noticeable improvement in the effective plastic strain was observed between design iterations 1 and 14, with iteration 14 exhibiting a larger surface area fraction of effective plastic strains greater than 0.4 (marking the onset of DRX for the

particular forging condition). A noticeable improvement in the minimum distance between the forging and impression die surfaces was also observed. The total optimization time was around 7 hours—to run 14 simulations, each taking approximately 30 mins to complete.

The advantages of this optimization framework are listed below:

- Unlike in voxel-based topology optimization methods, this framework operates directly on the underlying curve geometry (the wireframe of a surface model), making it possible to directly control the topology of the surface mesh of the preform.

- The framework integrates simulation information into the design environment by post-processing simulation data and displaying it directly in the CAD modelling workspace.

- The procedural generation algorithm provides control over defining the number of cross-sections that are used to create the starting preform shape, giving the flexibility to generate either very smooth organic preform shapes (by using a few numbers of cross-sections), or shapes that are very similar to the die impression (by using a high number of cross-sections).

- The procedural generation algorithm can be used in any shape optimization problem to fully automate geometry generation and modification, reducing modelling time significantly (only the impression mesh and rail curves need to be modelled as they are required inputs).

- Generated preform CAD models are guaranteed to avoid any preform-die interference issues.

The current limitations of this framework are listed below:

- The simple optimization heuristic does not consider flash volume or die-fill information to calculate the action responses.

- The subroutine used in the procedural generation algorithm for rail generation does not currently apply a supplemental curve smoothing operation, which is necessary to prevent drastic topology changes between contiguous cross-sections when large fixed step sizes are used.

6. Conclusions and Recommendations

The research detailed in this thesis was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) through the Strategic Partnership Grant (SPG) with contributions from Multimatic Technical Centre and CanmetMATERIALS, Natural Resources Canada. This research was part of a larger project (the SPG project) to develop a cost-effective hybrid manufacturing process, which involved sequentially casting (an AZ80 alloy preform) and forging to produce a lightweight complex-shaped automotive suspension front lower control arm (FLCA) for the 2013 Ford Fusion vehicle. The work presented in this thesis contributed directly or indirectly to the SPG project's overarching objectives. The primary contributions of this thesis are listed below:

1. A set of phenomenological material models for use in hot forging FEM simulations to predict the flow behaviour of wrought and as-cast AZ80 Mg alloy under industrial processing conditions.

2. A global design optimization method using a data-driven multi-objective optimization framework, which was applied to obtain a preform design for an AZ80 Mg alloy structural I-beam forging process.

3. A fully automated local design optimization method using a topology-based optimization framework, which was applied to a toy problem with a complex volumetric shape to demonstrate the functionality of the framework and its constitutive algorithms.

Details on the first contribution are presented in Chapter 3. Phenomenological material models were developed to simulate the anisotropic deformation behaviour of wrought AZ80 alloy at processing conditions of 300°C and 400°C, at equivalent strain rates of 0.01^{-1} and $0.1s^{-1}$. For predicting wrought alloy deformation, a Hill's (1948) quadratic anisotropic yield criterion was used. The material models were calibrated and validated systematically, starting with uniaxial compression simulations of Gleeble® samples in both ED and TD directions. "Coin" forgings of intermediate size and complexity (semi-closed die forging) were then simulated. Next, isotropic deformation of as-cast AZ80 alloy billet during I-beam forging (using flow stress curves obtained from wedge-cast sample material tests) was simulated. Lastly, the anisotropic deformation of wrought AZ80 alloy preform during FLCA forging was simulated and then compared with the lab-scale FLCA forging (CA48). The contents and recommendations pertaining to this chapter are summarized below:

- Good dimensional agreement between lab-scale Gleeble® compression samples and corresponding simulations was observed, suggesting the anisotropic metal flow of wrought

AZ80 alloy can be reliably predicted under uniaxial loading conditions (<2% error in major and minor axes dimensions was observed in simulations), using Hill's anisotropic coefficients. Significant forging load over-predictions (max. 18%) were observed for two processing conditions, which was likely due to inaccuracies in shear flow data.

- "Coin" sample forging was predicted very accurately until the final steps of the simulations, at which point the load-stroke curves diverged, resulting in an over-prediction of forging loads by +30% and +31% in the simulations corresponding to laboratory-scale forgings, SF14 (forged at a temperature of 400°C and an average speed of 0.4 mm/s) and SF6 (forged at a temperature of 400°C and an average speed of 0.04 mm/s). Good dimensional accuracy was observed between the simulated coin forging shape and SF6; a lower degree of dimensional accuracy was observed between the simulated coin forging shape and SF14. "Coin" forgings revealed that material models that are defined using a fixed set of Hill's coefficients are less adept at predicting flow behaviour under multi-axial loading conditions.

- The general deformation behaviour of the as-cast billet I-beam forging was simulated to generate a simulation template for use in Chapter 4 during the application of the data-driven multi-objective optimization framework.

- The forging behaviour of an FLCA wrought preform was simulated. Despite making simplification assumptions (i.e., isothermal forging conditions, uniform coefficient of friction, an isotropic hardening rule, and a constant set of Hill's coefficients), good agreement between the simulation and lab-scale forging CA48 was observed.

- It is recommended that the evolution of material anisotropy during deformation should be considered to further improve the prediction accuracy of forging simulations. In addition, coupled thermo-mechanical FEM models should be developed for industrial applications.

Details on the second contribution are presented in Chapter 4. This chapter highlights a global design optimization method. A data-driven multi-objective optimization framework was developed to evolve an optimal preform design for an I-beam forging operation. It was then cast using an AZ80 alloy and then forged to produce the component geometry. The proposed framework consists of three components: a parametric CAD model for shape generation, data-driven models for shape evaluation, and an SPEAII algorithm to drive a multi-objective optimization search over the design space of the parametric CAD model. The contents and recommendations pertaining to this chapter are summarized below:

- A reduced-order modelling technique was used to define the procedural generation logic for a parametric CAD model, capable of generating preform shapes with varying H-sections, given a set of weight parameters.

- An automation pipeline was developed to process large batches of simulation files to assemble and submit jobs to the simulation solver and to post-process results to eliminate the tedium and the chance of error. By automating the data engineering process, over 250 simulations were run to generate training data, which was then used to train 23 feedforward neural networks for predicting forging outcomes in 21 spatially varying regions of the forging, with the remaining two used to predict flash volume and die fill percentages.

- The preform shape was optimized according to the following objectives and constraints: maximize average effective plastic strain and minimize average strain variance in select regions of interest while minimizing flash percentage and achieving a die fill percentage greater than 95%.

- The multi-objective optimization algorithm (SPEAII) converged to a Pareto-optimal set of preform shapes. From this set, a single solution was selected and refined (to ensure secure placement on the forging die) for cast-forging AZ80 alloy I-beams. A multi-objective algorithm is better suited over a single objective algorithm (for this application) as it evolves a diverse set of non-dominated solutions that capture the trade-offs between different objectives.

- DEFORM®-3D [27] was used to run low-fidelity forging simulations, requiring a total simulation run time of about 50 hours, approximately 10 minutes per simulation (the preform geometry was meshed with approximately 30K elements).

- In general, forging outcome predictions showed good agreement with the laboratory-scale forgings. The optimized preform produced about half the material waste (~6% flash) compared to the cast-billet forging while fully filling the die.

- Quasi-static tensile and stress-controlled cyclic tests were conducted to evaluate the mechanical properties of the cast-billet and cast-preform forgings. On average, preform forgings exhibited a 4.3%, 4.2%, and 51.6% improvement in yield, ultimate tensile strength, and fracture strain, respectively. The significant increase in fracture strain suggested improved toughness.

- It is recommended that this global optimization framework be used during the initial design conception phase to broadly explore a design space to accrue knowledge regarding the trade-offs between different designs.

- For direct process-property (such as hardness) estimation of forging outcome, effective plastic strain information should be incorporated along with process parameters such as casting cooling rate and forging temperature. This would enable hardness (which can also be correlated with material properties such as tensile strength, resistance to deformation and wear properties) to be correlated with effective plastic strain.

Details on the third contribution are presented in Chapter 5. A fully automated topology-based optimization framework (a local shape optimization method) was developed to incrementally improve the topology of a starting preform shape. In complex-shaped preform design optimization problems, geometric modelling is time-consuming and places considerable reliance on engineering experience and judgment, while simulations can be computationally intensive. In addition, there is a lack of integration between CAD and FEM software packages which further reduces the efficiency of the optimization cycle since it is challenging to use simulation response information to update geometry effectively. This framework alleviates most of these pain points in the optimization cycle, and it was used on a toy optimization problem to improve the material distribution of a preform within a volumetric segment of the FLCA forging die. The contents and recommendations pertaining to this chapter are summarized below:

- The topology optimization framework relies on two algorithms to modify the topology of a starting preform shape automatically. The first algorithm executes a set of procedural generation logic to automatically generate the starting preform based on a set of immutable geometric and numerical inputs; and the optimization algorithm encapsulates a state-action lookup table that directs the topology update steps of the procedural generation algorithm.

- Unlike conventional voxel-based topology optimization methods, this framework directly manipulates the underlying curve geometry that makes up the surface model of the preform, while preventing any preform-die interference issues.

- This framework demonstrated its effectiveness in a complex-shaped design optimization task. The total time required to optimize the preform shape was around 7 hours, accounting for only the time required to run 14 simulations. The optimized shape displayed higher effective plastic strains (>0.4) and better die fill with little to no flash, relative to its starting shape. This framework is more computationally efficient than existing methods since it demands zero modelling effort from the user.

- It is recommended that this local optimization framework be used to incrementally improve the design of a candidate solution once it has been established based on engineering judgment.

7. Future Work

A theoretical formulation of a generative design system architecture that integrates state-of-the-art machine learning techniques with the procedural generation algorithm described in Chapter 5 is discussed in this section to expand on in future work. The limitation of the topology-based optimization framework is the high cost of computation, which can limit the number of designs that can be explored. A generative design system that uses a physics-based AI backbone for performance evaluation with an agent-based optimization approach—replacing the optimization heuristic—can drive design changes in an end-to-end manner, eliminating this limitation and improving the iterative topology optimization process (see Figure 7-1). Additionally, multiple instances of such a generative design system can be run concurrently to refine multiple candidate geometries simultaneously. This can be highly advantageous in industrial applications as it confers the demonstrated benefits of the procedural generation algorithm, which fully eliminates human involvement in geometric modelling and the promise of extremely fast non-linear deformation evaluation. The practical applications of such a system can extend far beyond forging preform design optimization to other manufacturing processes, such as blow forming (for parison[12] design), injection moulding (for mould design), etc.—as a general solution for optimizing any design process involving large deformation analysis.

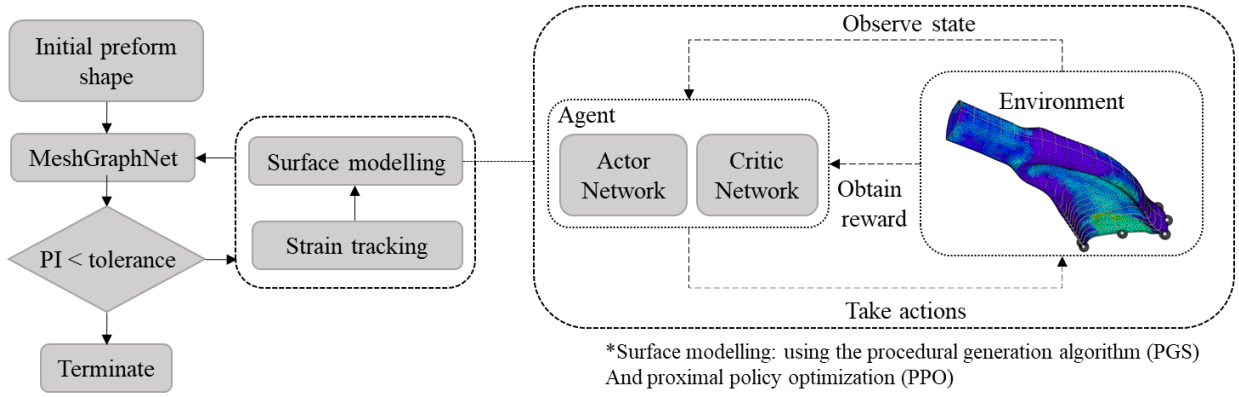


Figure 7-1: The generative design system architecture.

The physics-based AI backbone would be a MeshGraphNet (introduced by Pfaff et al.[87]), a graph neural network for learning mesh-based simulations (see Appendix D.1). The authors of the original work claim that these networks can run 1-2 orders of magnitude faster than the simulations on which they were trained. Additionally, the networks are said to learn mesh resolution independent dynamics, allowing them to generalize well beyond the size of the training data. For example, during training, even if a dataset with a relatively smaller mesh size was used, the

[12] In the blow molding process, the starting geometry is a thermoplastic molten tube called a "parison" [95]

network would still be able to predict the system dynamics of a larger mesh size or more complex boundary conditions. This promise of faster simulation times with an acceptable level of accuracy for coarse shape optimization problems makes the use of MeshGraphNet highly attractive. MeshGraphNets are trained to learn a one-step simulator, i.e., given a mesh M^t at time t, a graph neural network with an Encode-Process-Decode architecture (MeshGraphNet) is tasked with learning the dynamic quantities, such as vertex velocity, acceleration, stress, etc., of mesh M^{t+1} at time $t+1$. Once trained, the MeshGraphNet can simulate or "roll out" the deformation by chaining together a series of one-step deformation predictions (Figure 7-2).

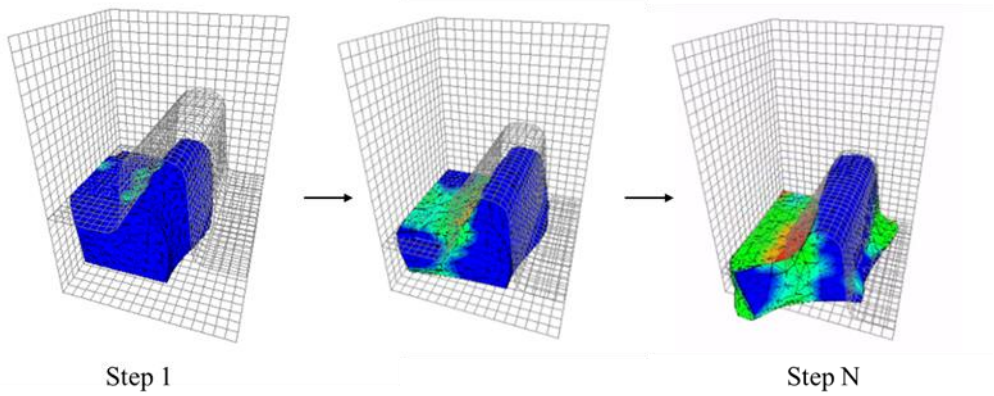


Figure 7-2: A simulation "rollout" of a single data sample (each sample consists of ~30 steps) from ~6000 sample dataset (curated as part of preliminary work in this research to explore the application of MeshGraphNets).

Agent-based optimization is an approach that uses agent(s), an autonomous entity that can interact with the environment, which in the context of the generative system would be the design environment within which the editable CAD and simulation response data are visualized. The role of the agent would be to direct surface modelling, to translate control points and modify CAD geometry. The agent would be trained using reinforcement learning to build an internal world model through self-supervised learning—based on interactions with the design environment—so that is can learn effective policies to modify cross-sections based on the forging response.

Reinforcement learning is a branch of machine learning concerned with optimal decision-making in a complex environment where at some time step $t \in N$, an agent observes the current state s_t of its environment, takes an action a_t and receives a reward signal $r_t \in R$ from the environment. During training, the agent interacts with the environment by taking a series of actions according to a policy π. A collection of these interactive experiences is referred to as a trajectory, i.e., a trajectory of states, actions, and rewards $\tau = (s_0, a_0, r_0, s_1, a_1, r_1 \ldots)$ of the given policy π. After collecting trajectories, the agent's neural networks are updated so that future actions would lead to a higher cumulative discounted reward (refer to Appendix B.3 for a description of proximal policy optimization (PPO), a specific type of RL algorithm that was explored in this work as shape optimization technique).

Proximal policy optimization (PPO) (refer to Appendix D.2) can be used to frame the RL problem where an agent with an actor-critic network is tasked with learning an optimal decision-making policy π for interacting with the design environment that executes the procedural generation algorithm (refer to Chapter 5). PPO is a reinforcement learning algorithm that aims to improve the training stability of the policy by avoiding large policy updates. It is a policy gradient method that learns from online data (through direct interaction with the design environment). The most common implementation of PPO is via the Actor-Critic Model, which uses two deep neural networks, one to take actions (Actor network) and the other to supply rewards (Critic network). The Actor model learns what action to take under a particular observed state of the environment, i.e., the average strain values attributed of control points, and the Critic model learns to evaluate if the action taken—the degree of displacement of a control point—is conducive to improving the preform topology.

References

[1] C. Moosbrugger, “Engineering Properties of Magnesium Alloys,” *ASM Int.*, no. M, pp. 1–12, 2017, [Online]. Available: http://doi.wiley.com/10.1002/9780470905098.ch1.

[2] M. K. Kulekci, “Magnesium and its alloys applications in automotive industry,” *Int. J. Adv. Manuf. Technol.*, vol. 39, no. 9–10, pp. 851–865, 2008, doi: 10.1007/s00170-007-1279-2.

[3] A. Gryguć, “Fatigue of Forged AZ80 Magnesium Alloy,” 2019.

[4] H. Jahed, “Jahed Mechanical Engineering and Mechatronics Strategic Partnership Grants for Projects Advanced Manufacturing / Lightweight Materials and Technologies Cost-effective lightweight complex-shaped structural components made of magnesium : A hybrid manufactur,” *SPG-P Propos. 0618*, 2020.

[5] N. P. Papenberg, S. Gneiger, I. Weißensteiner, P. J. Uggowitzer, and S. Pogatscher, “Mg-alloys for forging applications-A review,” *Materials (Basel).*, vol. 13, no. 4, pp. 1–61, 2020, doi: 10.3390/ma13040985.

[6] C. A. Sager, I. A. Yakubtsov, W. D. MacDonald, S. Shook, B. J. Diak, and M. Niewczas, “Physical metallurgy of Mg AZ80 alloys for forging applications,” *Magnes. Technol.*, no. January 2009, pp. 405–410, 2009.

[7] J. Zhang, V. S. Chevali, H. Wang, and C. Wang, “Current status of carbon fibre and carbon fibre composites recycling,” *Compos. Part B*, vol. 193, no. December 2019, p. 108053, 2020, doi: 10.1016/j.compositesb.2020.108053.

[8] DAWNBREAKER, “Market Research Study U.S. Forging Industry,” pp. 1–43, 2022.

[9] A. A. Luo, “Magnesium casting technology for structural applications,” *J. Magnes. Alloy.*, vol. 1, no. 1, pp. 2–22, 2013, doi: 10.1016/j.jma.2013.02.002.

[10] B. Callegari, T. N. Lima, and R. S. Coelho, “The Influence of Alloying Elements on the Microstructure and Properties of Al-Si-Based Casting Alloys: A Review,” *Metals (Basel).*, vol. 13, no. 7, p. 1174, 2023, doi: 10.3390/met13071174.

[11] P. Prakash *et al.*, “Microstructure and Texture Evolution During Hot Compression of Cast and Extruded AZ80 Magnesium Alloy,” 2019.

[12] E. Azqadan *et al.*, “The effect of cooling rate and degassing on microstructure and mechanical properties of cast AZ80 magnesium alloy,” *Mater. Sci. Eng. A*, vol. 844, no. January, 2022, doi: 10.1016/j.msea.2022.143176.

[13] J. Uramowski, “Influence of As-cast Microstructure on Forging Behaviour of Magnesium Alloy by,” 2021.

[14] ASM International, “Metal Handbook 9th edition Vol.15: Casting,” no. USA, ASM International, pp. 1–2002, 1988, [Online]. Available: http://s1.iran-mavad.com/ASM hanbooks/Vol_15_casting_iran-mavad.com.pdf.

[15] C. A. L. Bailer-jones, D. J. C. Mackay, T. J. Sabin, and P. J. Withers, “Static and Dynamic Modelling of Materials Forging,” *Aust. J. Intell. Inf. Process. Syst.*, vol. 5, no. 1, pp. 10–16, 1998.

[16] P. Prakash, D. Toscano, S. K. Shaha, M. A. Wells, H. Jahed, and B. W. Williams, “Effect of temperature on the hot deformation behavior of AZ80 magnesium alloy,” *Mater. Sci. Eng. A*, vol. 794, no. Ref 11, pp. 13–16, 2020, doi: 10.1016/j.msea.2020.139923.

[17] B. Shi, C. Yang, Y. Peng, F. Zhang, and F. Pan, “Anisotropy of wrought magnesium alloys: A focused overview,” *J. Magnes. Alloy.*, vol. 10, no. 6, pp. 1476–1510, 2022, doi:

10.1016/j.jma.2022.03.006.
[18] Y. Li, P. Hou, Z. Wu, Z. Feng, Y. Ren, and H. Choo, “Dynamic recrystallization of a wrought magnesium alloy: Grain size and texture maps and their application for mechanical behavior predictions,” *Mater. Des.*, vol. 202, 2021, doi: 10.1016/j.matdes.2021.109562.
[19] Z. jin Li *et al.*, “Microstructure evolution and deformation mechanism of AZ80 alloy during die forging,” *Mater. Sci. Eng. A*, vol. 869, no. September 2022, 2023, doi: 10.1016/j.msea.2023.144789.
[20] H. Y. Wang *et al.*, “Achieving high strength and high ductility in magnesium alloy using hard-plate rolling (HPR) process,” *Sci. Rep.*, vol. 5, no. 5988, pp. 1–9, 2015, doi: 10.1038/srep17100.
[21] F. Ju, Z. Xia, B. J. Diak, O. A. Ojo, W. D. MacDonald, and M. Niewczas, “Modeling and simulation of Mg AZ80 alloy forging behaviour,” *SAE Tech. Pap.*, vol. 2008, no. 724, 2008, doi: 10.4271/2008-01-0214.
[22] L. Yuan, Z. Zhao, W. Shi, F. Xu, and D. Shan, “Isothermal forming of the large-size AZ80A magnesium alloy forging with high mechanical properties,” *Int. J. Adv. Manuf. Technol.*, vol. 78, no. 9–12, pp. 2037–2047, 2015, doi: 10.1007/s00170-014-6780-9.
[23] Q. Chen *et al.*, “Isothermal closed-die forming process of magnesium alloy upper receiver: numerical simulation and experiments,” *Int. J. Adv. Manuf. Technol.*, vol. 102, no. 1–4, pp. 685–694, 2019, doi: 10.1007/s00170-018-03209-5.
[24] D. Kobold, G. Gantar, and T. Pepelnjak, “Finite element analysis of magnesium AZ80 wrought alloy anisotropic behaviour during warm forging,” *Mechanika*, vol. 18, no. 3, pp. 251–258, 2012, doi: 10.5755/j01.mech.18.3.1876.
[25] G. Yu, “Forging Specimen Design for Magnesium Alloys,” *(Master’s Thesis), Retrieved from Univ. Waterloo Database*, 2016.
[26] T. Paracha, “Modelling of the Forging Process for a Magnesium Alloy Automotive Control Arm,” 2018.
[27] Scientific Forming Technologies Corporation, “DEFORM®.” Columbus, OH, [Online]. Available: www.deform.com.
[28] B. R. Producers and N. Bureau, “A theory of the yielding and plastic flow of anisotropic metals,” *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.*, vol. 193, no. 1033, pp. 281–297, 1948, doi: 10.1098/rspa.1948.0045.
[29] A. Martins, J. R. R. A. and Ning, *Engineering Design Optimization*. Cambridge University Press, 2022.
[30] Z. Gao and R. V. Grandhi, “Microstructure optimization in design of forging processes,” *Int. J. Mach. Tools Manuf.*, vol. 40, no. 5, pp. 691–711, 2000, doi: 10.1016/S0890-6955(99)00083-8.
[31] G. Zhao, E. Wright, and R. V. Grandhi, “Sensitivity analysis based preform die shape design for net-shape forging,” *Int. J. Mach. Tools Manuf.*, vol. 37, no. 9, pp. 1251–1271, 1997, doi: 10.1016/S0890-6955(96)00087-9.
[32] Y. Shao, B. Lu, H. Ou, and J. Chen, “A new approach of preform design for forging of 3D blade based on evolutionary structural optimization,” *Struct. Multidiscip. Optim.*, vol. 51, no. 1, pp. 199–211, 2015, doi: 10.1007/s00158-014-1110-2.
[33] N. Thiyagarajan and R. V. Grandhi, “3D preform shape optimization in forging using reduced basis techniques,” *Eng. Optim.*, vol. 37, no. 8, pp. 797–811, 2005, doi: 10.1080/03052150500340579.

[34] W. L. Chan, M. W. Fu, and J. Lu, “An integrated FEM and ANN methodology for metal-formed product design,” *Eng. Appl. Artif. Intell.*, vol. 21, no. 8, pp. 1170–1181, 2008, doi: 10.1016/j.engappai.2008.04.001.
[35] Y. Shao, H. Ou, P. Guo, and H. Yang, “Shape optimization of preform tools in forging of aerofoil using a metamodel-assisted multi-island genetic algorithm,” *J. Chinese Inst. Eng. Trans. Chinese Inst. Eng. A*, vol. 42, no. 4, pp. 297–308, 2019, doi: 10.1080/02533839.2019.1584734.
[36] J.-J. Zhang, L.-W. Xu, and R.-Z. Gao, “Multi-island Genetic Algorithm Opetimization of Suspension System,” *TELKOMNIKA Indones. J. Electr. Eng.*, vol. 10, no. 7, pp. 1685–1691, 2012, doi: 10.11591/telkomnika.v10i7.1563.
[37] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002, doi: 10.1109/4235.996017.
[38] X. Liu and D. Zhang, “An improved SPEA2 algorithm with local search for multi-objective investment decision-making,” *Appl. Sci.*, vol. 9, no. 8, 2019, doi: 10.3390/app9081675.
[39] T. W. Sederberg and S. R. Parry, “Free-form deformation of solid geometric models,” *Proc. 13th Annu. Conf. Comput. Graph. Interact. Tech. SIGGRAPH 1986*, vol. 20, no. 4, pp. 151–160, 1986, doi: 10.1145/15922.15903.
[40] J. A. Samareh, “Geometry and grid/mesh generation issues for CFD and CSM shape optimization,” *Optim. Eng.*, vol. 6, no. 1, pp. 21–32, 2005, doi: 10.1023/B:OPTE.0000048535.08259.a8.
[41] B. Hoppenstedt *et al.*, “Dimensionality reduction and subspace clustering in mixed reality for condition monitoring of high-dimensional production data,” *Sensors (Switzerland)*, vol. 19, no. 18, 2019, doi: 10.3390/s19183903.
[42] D. Sieger, S. Menzel, and M. Botsch, “RBF morphing techniques for simulation-based design optimization,” *Eng. Comput.*, vol. 30, no. 2, pp. 161–174, 2014, doi: 10.1007/s00366-013-0330-1.
[43] A. Schulz, J. Xu, B. Zhu, C. Zheng, E. Grinspun, and W. Matusik, “Interactive design space exploration and optimization for CAD models,” *ACM Trans. Graph.*, vol. 36, no. 4, 2017, doi: 10.1145/3072959.3073688.
[44] J. D. Camba, M. Contero, and P. Company, “Parametric CAD modeling: An analysis of strategies for design reusability,” *CAD Comput. Aided Des.*, vol. 74, pp. 18–31, 2016, doi: 10.1016/j.cad.2016.01.003.
[45] A. Eltaweel and Y. SU, “Parametric design and daylighting: A literature review,” *Renew. Sustain. Energy Rev.*, vol. 73, no. February, pp. 1086–1103, 2017, doi: 10.1016/j.rser.2017.02.011.
[46] K. H. Chang and C. Chen, “3D shape engineering and design parameterization,” *Comput. Aided. Des. Appl.*, vol. 8, no. 5, pp. 681–692, 2011, doi: 10.3722/cadaps.2011.681-692.
[47] R. McNeel, “Rhinoceros 3D, Version 6.0. Robert McNeel & Associates.” Seattle, WA, 2010.
[48] C. Hoff, “Structural design optimisation by organisation of stiffness,” 2016.
[49] R. T. F. A. King, K. Deb, and H. C. S. Rughooputh, “Comparison of NSGA-II and SPEA2 on the Multiobjective Environmental/Economic Dispatch Problem,” *Univ. Mauritius Res. J.*, vol. 16, no. 1, pp. 485–511, 2010.
[50] S. W. Kielarova and S. Sansri, “Shape optimization in product design using interactive

genetic algorithm integrated with multi-objective optimization," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10053 LNAI, pp. 76–86, 2016, doi: 10.1007/978-3-319-49397-8_7.
[51] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm," *Evol. Methods Des. Optim. Control with Appl. to Ind. Probl.*, pp. 95–100, 2001, doi: 10.1.1.28.7571.
[52] S. L. Brunton and J. N. Kutz, "Data Driven Science & Engineering - Machine Learning, Dynamical Systems, and Control," p. 572, 2017, [Online]. Available: databook.uw.edu.
[53] T. Gulikers, "An Integrated Machine Learning and Finite Element Analysis Framework, Applied to Composite Substructures including Damage," 2018, [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3A615f2151-bcae-4e78-a2cb-3f1891a28275.
[54] C. Debao, "Degree of approximation by superpositions of a sigmoidal function," *Approx. Theory its Appl.*, vol. 9, no. 3, pp. 17–28, 1993, doi: 10.1007/BF02836480.
[55] D. E. Rumelhart and G. E. Hintont, "Learning Representations by Back-Propagating Errors," *Cogn. Model.*, no. 2, pp. 3–6, 2019, doi: 10.7551/mitpress/1888.003.0013.
[56] T. A. Kodippili, "Hot Forging Tool Design for a Magnesium Alloy Front Lower Control Arm," 2018, [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/14032.
[57] T. Altan and G. Ngaile, *Cold and Hot Forging*. 2005.
[58] S. I. Oh, W. T. Wu, J. P. Tang, and A. Vedhanayagam, "Capabilities and applications of FEM code deform: the perspective of the developer," *J. Mater. Process. Tech.*, vol. 27, no. 1–3, pp. 25–42, 1991, doi: 10.1016/0924-0136(91)90042-D.
[59] T. Yamashita, "Analysis of Anisotropic Material," University of Ohio, 1996.
[60] R. B. Colby, "Ph.D.thesis. Equivalent plastic strain for the Hill's yield criterion under general three-dimensionalloading," pp. 1–45, 2013.
[61] D. S. BIOVIA, "SOLIDWORKS." 2020.
[62] D. W. A. Rees, *BASIC ENGINEERING PLASTICITY: An Introduction with Engineering and Manufacturing Applications*, vol. 4, no. 1. Elsevier Ltd., 2006.
[63] K. I. Y.Izawa, R.Tanaka, "Revolution of Principal Axes of Plastic Anisotropy Developed during Deformation Process," *JSME Int. J.*, 2001, doi: https://doi.org/10.1299/jsmea.45.245.
[64] L. Chen, W. Wen, and H. Cui, "Generalization of Hill's yield criterion to tension-compression asymmetry materials," *Sci. China Technol. Sci.*, vol. 56, no. 1, pp. 89–97, 2013, doi: 10.1007/s11431-012-5037-9.
[65] S. B. Behravesh, "Fatigue Characterization and Cyclic Plasticity Modeling of Magnesium Spot-Welds," p. 264, 2013, [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/7651%5Cnhttp://hdl.handle.net/10012/7651.
[66] Margarita Tsarkouskaya, "A Formulation on Friction in Forging," 2020.
[67] Ö. N. Cora, M. Akkök, and H. Darendeliler, "Modelling of variable friction in cold forging," *Proc. Inst. Mech. Eng. Part J J. Eng. Tribol.*, vol. 222, no. 7, pp. 899–903, 2008, doi: 10.1243/13506501JET419.
[68] P. Groche, P. Kramer, S. Zang, and V. Rezanov, "Prediction of the evolution of the surface roughness in dependence of the lubrication system for cold forming processes," *Tribol. Lett.*, vol. 59, no. 1, pp. 1–9, 2015, doi: 10.1007/s11249-015-0535-0.
[69] A. Gontarz, Z. Pater, K. Drozdowski, A. Tofil, and J. Tomczak, "Fem Analysis of the Forging Process of Hub Part From Az80 Magnesium Alloy," *Proc. 10th World Congr.*

Comput. Mech., no. May 2014, pp. 456–466, 2014, doi: 10.5151/meceng-wccm2012-16805.
[70] Y. Xue, Z. M. Zhang, and Y. J. Wu, "Study on critical damage factor and the constitutive model including dynamic recrystallization softening of AZ80 magnesium alloy," *Sci. Sinter.*, vol. 45, no. 2, pp. 199–208, 2013, doi: 10.2298/SOS1302199X.
[71] M. L. Thorat and R. R. Ligade, "Review on Defects in Hot Forging Process-Investigation," *Int. J. Sci. Dev. Res.*, vol. 3, no. 4, pp. 34–42, 2018.
[72] A. Gryguć, "Private communications." 2023.
[73] T. Kodippili, S. Lambert, and A. Arami, "Data-driven prediction of forging outcome: Effect of preform shape on plastic strain in a magnesium alloy forging," *Mater. Today Commun.*, vol. 31, 2022, doi: 10.1016/j.mtcomm.2022.103210.
[74] T. Kodippili, E. Azqadan, H. Jahed, S. Lambert, and A. Arami, "Multi-objective optimization of a cast-preform shape for a magnesium alloy forging application," *Int. J. Adv. Manuf. Technol.*, pp. 3221–3232, 2023, doi: 10.1007/s00170-023-12478-8.
[75] T. Kodippili, "Multi-objective Optimization of a Cast-preform Shape for a Magnesium Alloy Forging Application," pp. 1–18, 2023.
[76] D. D'Agostino, A. Serani, M. Diez, and E. F. Campana, "Deep autoencoder for off-line design-space dimensionality reduction in shape optimization," *AIAA/ASCE/AHS/ASC Struct. Struct. Dyn. Mater. Conf. 2018*, no. May, 2018, doi: 10.2514/6.2018-1648.
[77] R. Stouffs and W. T. Chang, "Representational programming for design analysis," *EG-ICE 2010 - 17th Int. Work. Intell. Comput. Eng.*, no. June 2010, 2019.
[78] M. Khabazi, *Algorithmic Modelling with Grasshopper (Rhino Plug-in)*. 2009.
[79] Z. Z. Liu, W. Li, and M. Yang, "Two General Extension Algorithms of Latin Hypercube Sampling," *Math. Probl. Eng.*, vol. 2015, 2015, doi: 10.1155/2015/450492.
[80] J. L. Deutsch and C. V. Deutsch, "Latin hypercube sampling with multidimensional uniformity," *J. Stat. Plan. Inference*, vol. 142, no. 3, pp. 763–772, 2012, doi: 10.1016/j.jspi.2011.09.016.
[81] M. Agarwal, "Cross-validated Structure Selection for Neural Networks," vol. 20, no. 2, pp. 175–186, 1996.
[82] A. B. Owen, "A robust hybrid of lasso and ridge regression," no. October, pp. 59–71, 2007, doi: 10.1090/conm/443/08555.
[83] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.
[84] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Struct. Multidiscip. Optim.*, vol. 26, no. 6, pp. 369–395, 2004, doi: 10.1007/s00158-003-0368-6.
[85] E. Azqadan, J. Uramowski, M. A. Wells, A. Arami, and H. Jahed, "Hardness variation in cast-forging process of AZ80 magnesium alloys and its data-driven prediction," *Mater. Today Commun.*, vol. 36, no. May, p. 106833, 2023, doi: 10.1016/j.mtcomm.2023.106833.
[86] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 1, pp. 4–24, 2021, doi: 10.1109/TNNLS.2020.2978386.
[87] T. Pfaff, M. Fortunato, A. Sanchez-Gonzalez, and P. W. Battaglia, "Learning Mesh-Based Simulation with Graph Networks," pp. 1–18, 2020, [Online]. Available: http://arxiv.org/abs/2010.03409.

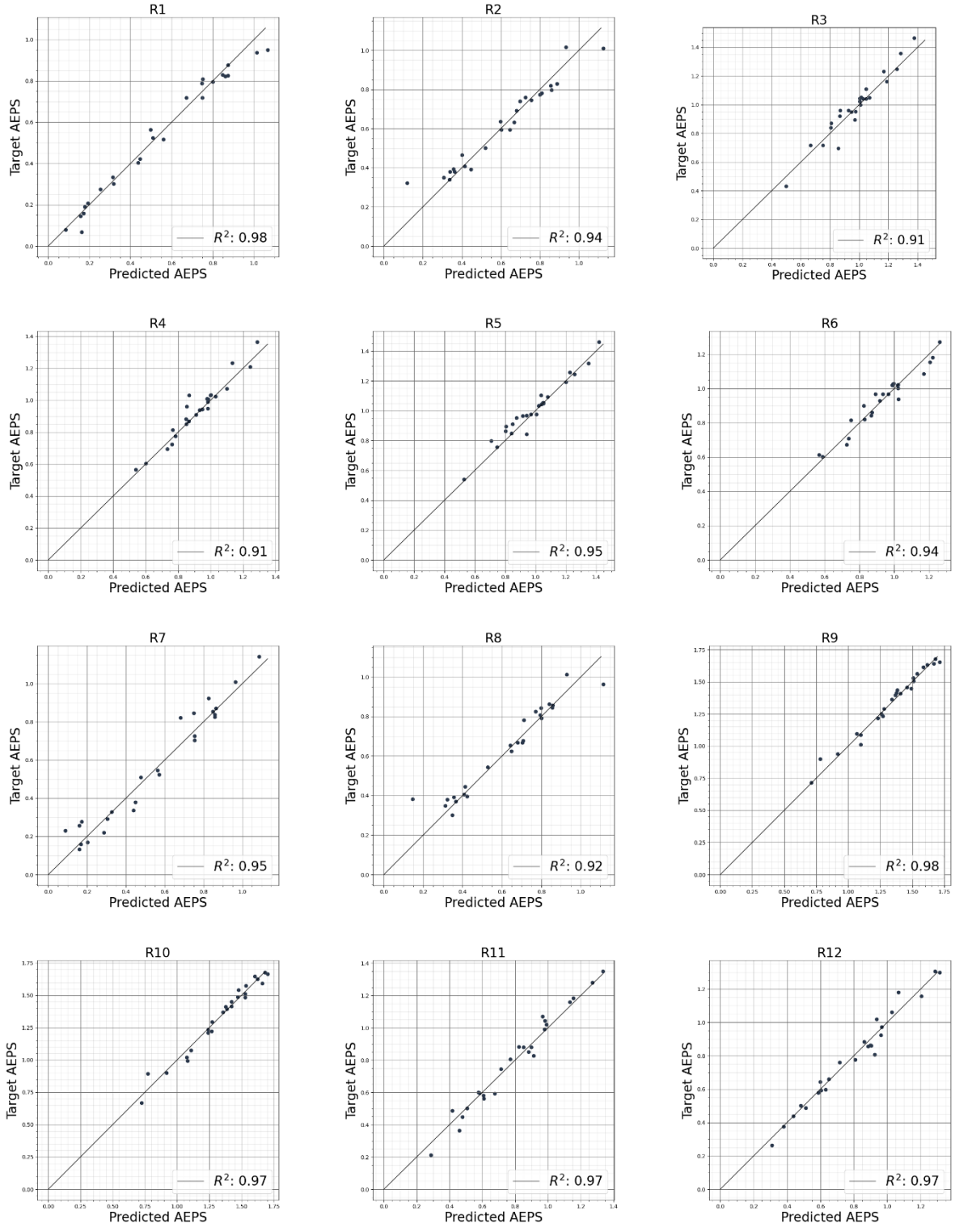
[88] J. Wang, “An Intuitive Tutorial to Gaussian Processes Regression.”
[89] W. E. Lorensen and H. E. Cline, “Marching Cubes: A High Resolution 3D Surface Construction Algorithm,” vol. 21, no. 4, pp. 163–169, 1987, doi: 10.1145/37401.37422.
[90] A. Bacciaglia, A. Ceruti, and A. Liverani, “Surface smoothing for topological optimized 3D models,” *Struct. Multidiscip. Optim.*, vol. 64, no. 6, pp. 3453–3472, 2021, doi: 10.1007/s00158-021-03027-6.
[91] D. F.Rogers, *An Introduction to NURBS with Historical Perspective*. Morgan Kaufmann Publishers, 2001.
[92] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal Policy Optimization Algorithms,” pp. 1–12, 2017, [Online]. Available: http://arxiv.org/abs/1707.06347.
[93] X. Hui, H. Wang, W. Li, J. Bai, F. Qin, and G. He, “Multi-object aerodynamic design optimization using deep reinforcement learning,” *AIP Adv.*, vol. 11, no. 8, 2021, doi: 10.1063/5.0058088.
[94] J. Viquerat, J. Rabault, A. Kuhnle, H. Ghraieb, A. Larcher, and E. Hachem, “Direct shape optimization through deep reinforcement learning,” *J. Comput. Phys.*, vol. 428, 2021, doi: 10.1016/j.jcp.2020.110080.
[95] PDL Handbook Series, *Applied Plastics Engineering Handbook*. 2017.

Appendix A.1: Wrought AZ80 Alloy Forged FLCA

Figure A-1: Cast Al (benchmark) control arm for the 2013 Ford Fusion vehicle and the forged Mg front lower control arm.

Figure A-2: The AZ80 alloy front lower control arm that was forged during the APC project; the multiple pre-forming steps are shown in: (a) bent Mg alloy billet; (b) flattened preform shape; (c) as-forged FLCA; (d) FLCA with flash trimmed and surfaces machined.

Appendix B.1: Feedforward Network Training Results



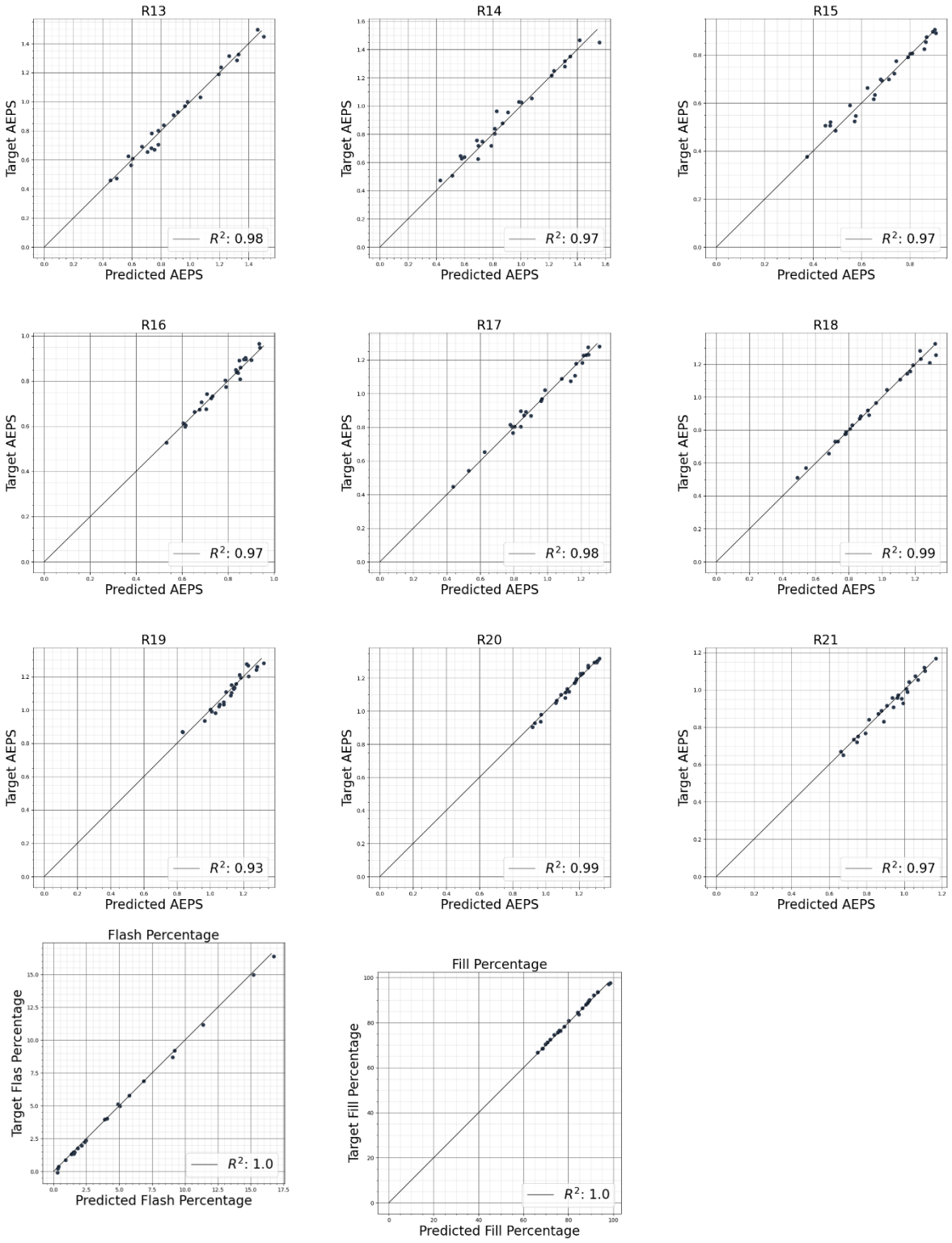


Figure B-1: The prediction accuracy of average effective plastic strain (AEPS), flash and fill percentages on test data of the 23 feedforward networks that were trained for the evolutionary design optimization framework are shown.

Appendix C.1: Manual Design Iteration of the Control Arm Preform for Cast-Forging

This section provides a description of the manual iterative design process that was carried out to optimize the preform design for an AZ80 alloy structural automotive front lower control arm (FLCA), which was cast-forged. The design optimization process was involved iteratively modifying a starting preform shape by manually modelling surface topology changes according to the BESO flow diagram shown in Figure C-1.

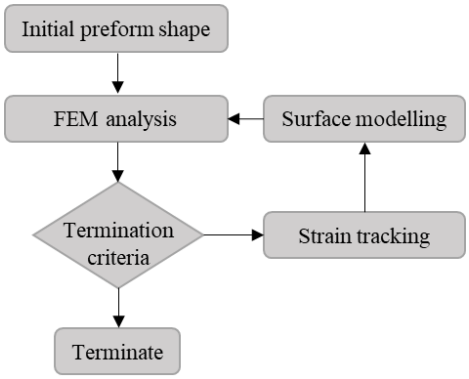


Figure C-1: Design iteration flow diagram (the proposed topology optimization algorithm automates the surface modelling step in this flow diagram)

The FEM simulations were set up using the as-cast AZ80 alloy isotropic material model (based on compression stress-strain response curves corresponding to specimens that were cooled at a rate of 3.5°C/s). Initially, the preform was meshed with a coarse mesh size of approximately 100K elements, which was then increased to a finer mesh size of approximately 300K elements during the last two design iteration cycles. The simulations were conducted under an isothermal forging condition assumption, under a variable die speed schedule starting at 4mm/s and stepping down to 0.04 mm/s and then 0.004 mm/s towards the end of the forging. After the simulation analysis step in the design iteration loop, the forging outcome was evaluated to determine whether KPIs indicate a satisfactory outcome, i.e., the forging has fully filled the die, produced minimal flash, and achieved good metal flow and effective plastic strains greater than 0.4 in structurally critical regions such as the sweep region. The sweep region was of particular interest as the onset of crack initiation was observed first in this region during laboratory-scale durability trials of control arms forged during the APC project (refer to section 3.4.4). If the termination criteria were not satisfied (based on a qualitative assessment of forging quality), strain tracking was performed. Strain tracking (or strain backtracking) is the process of imposing the strain field of the final mesh onto the starting mesh in the first simulation step to identify regions of the preform where material needs to be added or removed (see Figure C-2(a)). Finally, this step was followed by surface modelling to make the necessary adjustments to the surface topology (see Figure C-2(b)).

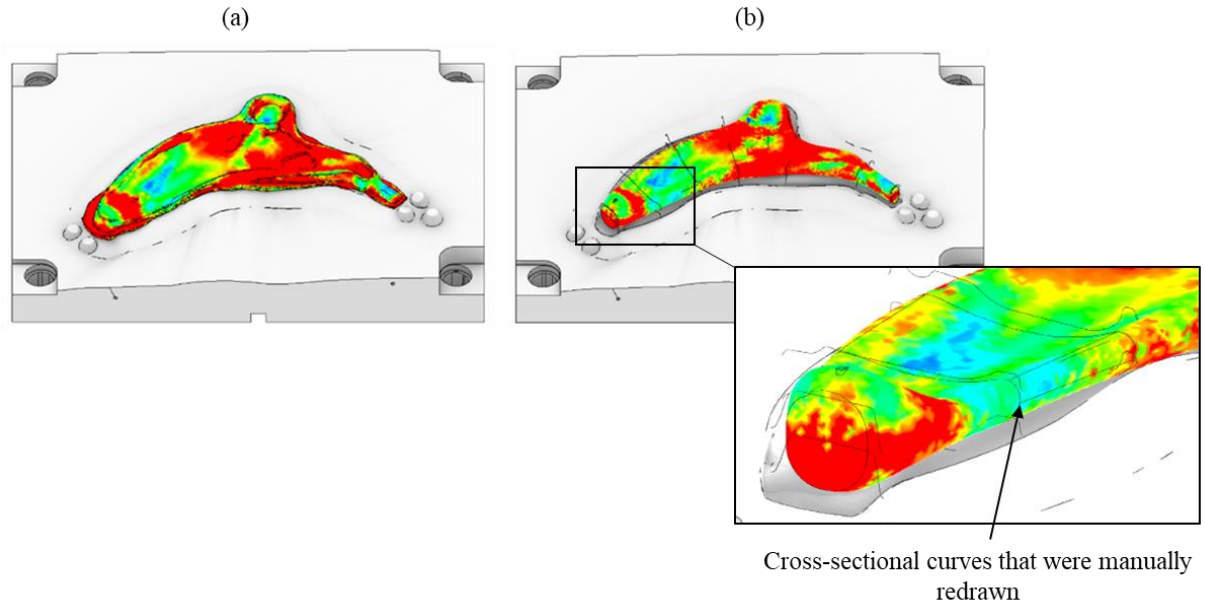


Figure C-2: (a) The strain field and the mesh state in the last simulation step is shown; (b) the strain field of the last simulation step is backtracked and imposed on to the mesh in the first simulation step; high degree of plastic deformation (indicated by high effective strain values greater than 1.0 are shown in red), and regions of inadequate deformation (indicated by low effective strain values less than 0.4 in regions are shown in blue and green.

In total, 11 iterative cycles were carried out to obtain cast-preform iteration 11 (see design CP11 in Figure C-3), which was cast using a PMC process and forged under temperature conditions ranging from 250 °C to 350 °C. The design optimization cycle was initiated with a preform shape with a close semblance to the control arm shape, based on engineering judgement in an effort to minimize surface shear by minimizing the degree of lateral metal flow during deformation. The final preform design is shown in Figure C-4, and the forging outcome of CP11 is shown in Figure C-4 alongside the forging outcome of a control arm CA48 (refer to section 3.4.4). In addition to fully filling the die, the CP11 forging exhibited better flow behavior in the critical sweep region compared to CA48. Also, an evident reduction in the flash can be observed in the CP11 forging compared to the flash produced in the multi-step flattening and forging operation that was involved in producing CA48.

Figure C-3: (a) Control arm cast-preform design iteration 11 (CP11); (b) cast preform (courtesy of CanmetMATERIALS, Lucian Blaga).

In total 11 design iterations were carried out. The average computational time for a control arm forging simulation with a mesh size of 100K elements was around 3 hours and 7 hours with a mesh size of 300K elements. Roughly an additional 2 hours were dedicated towards surface modelling to make the necessary geometric modifications. In total, around 62 hours were dedicated to refining the initial preform geometry: 27 hours to run the first nine course-mesh simulations, 14 hours to run the last two fine-mesh simulations, and approximately 22 hours to modify preform geometry.

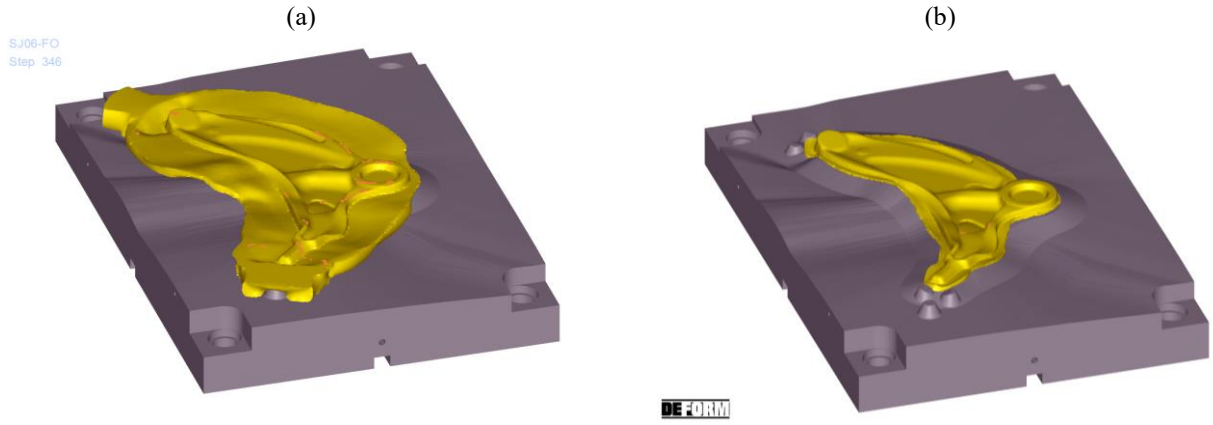


Figure C-4: (a) Multi-step flattening and forging outcome of CA48; (b) cast-forging outcome of CP11

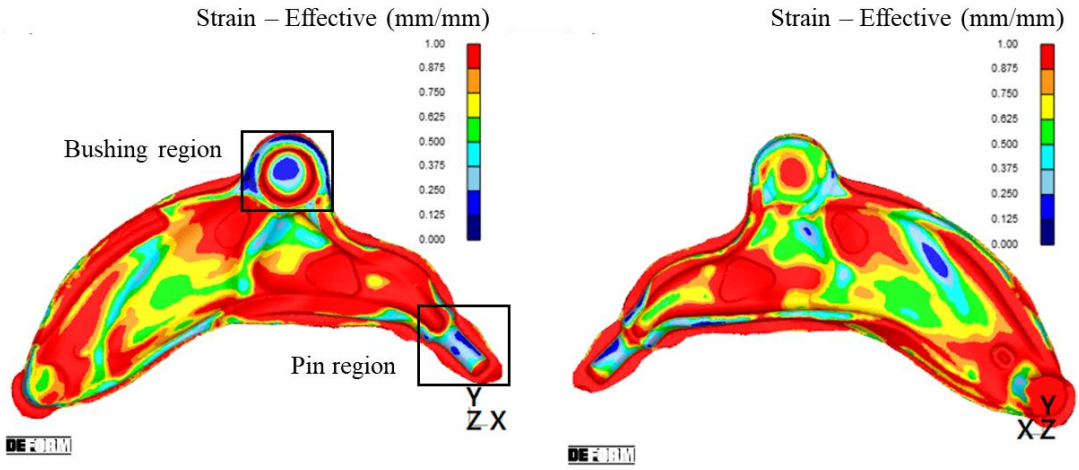


Figure C-5: CP11 forging at 300 ºC (effective plastic strain is shown).

While preform CP11 drastically reduced material waste (less than 5% flash waste relative to the +30% flash waste that was produced during multi-step forging of CA48, the effective strain was visibly lower in certain regions throughout the component, such as in the pin region and the bushing (see Figure C-5). At forging temperature of 300 ºC, the required forging load to reach full die shut height exceeds the press capacity, leaving underfill in the bushing and rib regions of the forging (see Figure C-6); the control arm was fully forged at a temperature of 350 ºC.

Figure C-6: Control arm forging outcome using cast-preform iteration 11 (CP11), forged under 300 ºC temperature condition (courtesy of CanmetMATERIALS, Bruce Williams)

Appendix D.1: Graph-Based Neural Networks

Three different modules are used in the model: an encoder, a processor, and a decoder. The encoder constructs the input graph, representing the system's state comprising the preform mesh and the rigid body die meshes. Meshes are essentially graphs with nodes and edges. Dual space message passing is implemented in the MeshGraphNets to enable the model to reason the nodal interactions to compute the internal dynamics that govern the deformation of mesh manifolds. By encoding relative positions between mesh nodes as edge attributes–interactions between pairs of nodes based on spatial proximity—external dynamics such as self-collision and contact can computed. By providing relative position features as edge attributes, the model is entirely local and translation equivariant (equivariance is a form of symmetry for functions), making it data efficient and capable of generalizing to larger problems. Additionally, node features such as velocity, node type (to differentiate between preform and dies), and effective plastic strain can be embedded in the graph nodes to inject inductive bias or prior knowledge into the model.

As a first step in learning a one-step simulator MeshGraphnet model, the graph's raw node and edge features are encoded into a latent representation of a fixed size. The encoded graph data is then processed using the processor by performing several rounds of message passing, which pools together messages received from both mesh space and world space edges, enabling the model to learn the physical interactions between objects. Processing consists of message passing, aggregation, and updating operations. The graph node and edge embeddings are updated via message passing during the processing step.

Lastly, the decoder maps the updated states to acceleration vectors. This is the network's final output, which is a prediction of the acceleration of the nodes in the current time step. The acceleration vectors are then used to update the next mesh state to “roll out” model predictions.

Appendix D.2: Proximal Policy Optimization

Reinforcement learning is a branch of machine learning concerned with optimal decision-making in a complex environment where, at some time step $t \in N$, an agent observes the current state s_t of its environment, takes an action a_t and receives a reward signal $r_t \in R$ from the environment. During training, the agent interacts with the environment by taking a series of actions according to a policy π. A collection of these interactive experiences is referred to as a trajectory, i.e., a trajectory of states, actions, and rewards $\tau = (s_0, a_0, r_0, s_1, a_1, r_1 \ldots)$ of the given policy π. After collecting trajectories the agent's neural networks are updated so that future actions would lead to a higher cumulative discounted reward. Proximal policy optimization (PPO) is a type of reinforcement learning algorithm. PPO is a policy gradient method that is suited for continuous control tasks. PPO learns directly by interacting with the environment, meaning it uses a batch of experiences to perform a gradient update of the agent's network parameters. General policy optimization methods define the policy gradient loss as the expectation over the policy actions log multiplied by an advantage function estimate [92].

$$L^{PG}(\theta) = \hat{\mathrm{E}}_t\left[log\pi_\theta(a_t|s_t)\widehat{\mathrm{A}}_t\right] \tag{32}$$

where π_θ is the policy network (actor network) that takes as input observations s_t from the environment and outputs log probabilities of actions a_t; and $\widehat{\mathrm{A}}_t$ is an estimate of the relative value of the selected action given the current state. The advantage is computed after the episode sequence is collected from the environment, using the discounted sum of rewards and a value function $V(\cdot)$. The value function (learned by a value network or critic network v_w) provides an estimate of the discounted return based on the current state and future onward states to estimate the final return for the episode from the current state. During training, the critic network is frequently updated using observations s_t collected from the environment and outputs a noisy estimate of the discounted sum of rewards from this point onwards [92].

$$\begin{gathered}\widehat{\mathrm{A}}_t = \delta_t + (\gamma\lambda)\delta_{t+1} + \cdots + (\gamma\lambda)^{T-t+1}\delta_{T-1} \\ \text{where } \delta = r_t + \gamma V(s_{t+1}) - V(s_t)\end{gathered} \tag{33}$$

The objective function that is used in the PPO algorithm has the following form [92]:

$$L^{CLIP}(\theta) = \hat{\mathrm{E}}_t\left[\min\left(r_t(\theta)\widehat{\mathrm{A}}_t\right), clip(r_t(\theta), 1-\epsilon, 1+\epsilon)\widehat{\mathrm{A}}_t\right] \tag{34}$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\mathrm{Old}}}(a_t|s_t)}$ is used to ensure that the updated policy remains within a trust region centered around the old policy, and a clipping operation, to restrict moving r_t outside the

predefined trust-region interval $[1-\epsilon, 1+\epsilon]$. Pseudocode 1 provides the pseudocode for vanilla PPO.

Pseudocode 1: PPO, Actor-Critic Style

for $iteration = 1,2, \dots N$ **do**
  for $actor = 1,2, \dots M$ **do**
    Run policy $\pi_{\theta_{old}}$ in the environment for T time steps
    Compute advantage estimates $\hat{A}_1, \dots, \hat{A}_T$
  end for
  Optimize surrogate L given θ, with K epochs and mini-batch size $s \leq MT$
  $\theta_{old} \leftarrow \theta$
end for

The agent's actor π_θ and critic V_w networks have the same architecture, except at the output layer. While the actor-network takes as input, a point cloud representation of the current state, and outputs an array of action log probabilities, the critic network outputs a Q-value. The critic network is tasked with learning to evaluate whether the action taken by the actor-network led the environment to transition into a better state by outputting a Q-value (a value that represents the long-term desirability of the state-action, in other words, the expected cumulative reward given the current state). When the design environment signals the server to train the agent, the actor-critic network parameter update is carried out according to Pseudocode 2 for a predefined number of network updates I. At the start of the loop, the sequence of states S, rewards R, and critic output Q-values V are recalled from the agent's memory. Then, advantage $\hat{A}$ is calculated for each time step. Finally, the recalled sequence of experience is partitioned into J number of batches of size N, and is used to update the parameters of both the actor and critic networks.

Pseudocode 2: Train Agent

Initialize: Critic network V_w with initial parameters w
Initialize: Actor network π_θ with initial parameters θ
Input: Number of network updates I
for $i = 1,2, \dots I$ **do**
  $S, R, V =$ **Agent.memory.Recall( )**
  for $t = 1,2, \dots T-1$ steps per epoch **do**
    for $k = t, \dots T-1$ steps per epoch **do**
      $\hat{A}[t] \leftarrow (\gamma\lambda)^{k-1}(R[k] + \gamma V[k+1] - V[k])$
    end for
  end for
  for $j = 1,2, \dots J$ batches **do**
    $L(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(\min\left(\frac{\pi_\theta(S[j])}{\pi_{\theta_{Old}}(S[j])}\hat{A}[j]\right), clip\left(\frac{\pi_\theta(S[j])}{\pi_{\theta_{Old}}(S[j])}, 1-\epsilon, 1+\epsilon\right)\hat{A}[j]\right)$
    $L(w) = \frac{1}{N}\sum_{i=1}^{N}\left(\left[\hat{A}[j] + R[j]\right] - V_w[S[j]]\right)^2$
    $L = L(\theta) + \frac{1}{2}L(w)$
    $\theta \leftarrow \theta + \alpha\nabla L$

16: $w \leftarrow w + \alpha \nabla L$
17: **end for**
18: **end for**

An application of PPO in shape optimization includes work by Hui et al., who proposed using PPO to carry out a multi-objective aerodynamic design optimization where they use computational fluid dynamics (CFD) simulations to evaluate the aerodynamic performance of different airfoils to learn a policy to generate optimal airfoil shapes in a single step [93]. In this work, the authors train a PPO agent to learn how to maximize the lift-to-drag ratio of an airfoil cross-section while restricting its thickness. The state description they provided to the learning agent was the geometric coordinates of a baseline airfoil (RAE2822), which the agent then adjusted by translating the control points of a free-form deformation (FFD) mapping frame. In FFD, the target geometry is embedded in a mapping frame where modifications can be made to the shape of the geometry indirectly by controlling the mapping frame. Their algorithm tasks the agent with iteratively modifying the vertical locations of control points of the FFD frame. This generates a new airfoil geometry, and the environment transitions into a new state, and a reward is provided to the agent from the environment. Their multi-objective PPO algorithm was compared with an evolutionary genetic optimization algorithm, NSGA-II, to evolve a Pareto set of airfoils; the PPO algorithm was shown to generate Pareto solutions within 15% of the required by the NSGA-II algorithm. Another application of PPO in shape optimization was demonstrated by Viquerat et. al. [94], who used it for an airfoil shape generation problem where they applied a "degenerate" version of deep reinforcement learning (DRL) to generate 2D airfoil shapes by controlling points on Bezier curves. They also use PPO with CFD simulations to drive the optimization process. The "degenerate" version of DRL applied in this learning problem consisted of a learning episode with a single time step where the agent's network made a one-shot attempt to generate an optimal shape. Their work applied DRL as a direct non-linear optimizer where the agent learns from indirect supervision of the reward signal.