Doctoral theses at NTNU, 2020:120

Ivanna Baturynska

Machine Learning for Quality Assurance in Polymer Powder Bed Fusion Additive Manufacturing

Mapping build layout design to quality of AM products as a part of an intelligent system for quality assurance

Doctoral thesis

Ivanna Baturynska

ISBN 978-82-326-4588-6 (printed ver.)
ISBN 978-82-326-4589-3 (electronic ver.)
ISSN 1503-8181

NTNU
Norwegian University of Science and Technology

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Engineering
Department of Manufacturing and Civil Engineering

Doctoral theses at NTNU, 2020:120

NTNU
Norwegian University of Science and Technology

NTNU

Ivanna Baturynska

Machine Learning for Quality Assurance in Polymer Powder Bed Fusion Additive Manufacturing

Mapping build layout design to quality of AM products as a part of an intelligent system for quality assurance

Thesis for the Degree of Philosophiae Doctor

Trondheim, April 2020

Norwegian University of Science and Technology
Faculty of Engineering
Department of Manufacturing and Civil Engineering

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Engineering
Department of Manufacturing and Civil Engineering

ISBN 978-82-326-4588-6 (printed ver.)
ISBN 978-82-326-4589-3 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2020:120

Printed by NTNU Grafisk senter

to my parents, Liubov and Valerii Ostapchuk

Abstract

The latest developments in the field of additive manufacturing (AM) have led to the wider use of this technology for the production of end-user products. As a result, more stringent requirements are set to the quality of parts produced with AM technology. In order to satisfy the needs of industries and to make additive manufacturing more attractive technology, a more in-depth understanding of the factors that cause variations is needed.

Since each AM process works with a limited number of materials and requires optimization of different process parameters, each AM category defines a variety of materials and parameters that can be optimized. There are seven different categories of AM, while a polymer powder bed fusion process is analysed in this work. Typically, the main components of the AM process, which are material, process, and CAD models(s), are investigated isolated from each other, or by limited combinations. However, a comprehensive mathematical description of any AM process should be based on all related components. Therefore, a new approach towards quality assurance in AM is needed.

This thesis aims at describing two perspectives: (1) gaining more knowledge about the effect of build layout design on the quality of the produced parts; (2) and how the obtained knowledge can be transformed into a user-friendly decision support tool to minimize variations in polymer powder bed fusion process prior to fabrication. The STL model characteristics, nine coordinates, three orientation angles, building platform utilization, platform density, material, run number, and build height are the parameters used to describe the build layout design.

As a result, experimental work on the EOS P395 AM system has been conducted, and 1526 standard dogbones specimens have been produced.

In order to collect data from these specimens in a systematic way, a new data acquisition and analysis approach has been proposed, which includes data registration rules for stages before, during, and after the AM process. With the help of model-based system engineering, a number of calculation modules have been grouped into an intelligent system for quality assurance. Each calculation module consists of a set of functional requirements and a number of selected predictive models trained from the collected data using machine learning algorithms. Since the quality of machine learning models can be affected by the quality of used data, the data analysis pipeline has been proposed as a process of data preparation and has been developed based on the data science foundations. Data cleaning, data integration, data normalization, and feature selection have been used as the main data preprocessing steps. The resulting predictive models are:

1. the Random Forest machine learning model with an accuracy of 99.16% for the estimation of geometric deviations.
2. the Random Forest machine learning model with accuracy 47.14% (RMSE = 0.026) for estimation of compensation ratios in x, y, and z axes for each object individually.
3. the Random Forest machine learning model with accuracy 66.24% for estimation of tensile modulus, tensile strength, and strain at break.

This work also describes an attempt that was made to develop the predictive model for optimization of part location in the build chamber with respect to mechanical properties. However, this model requires further work.

Acknowledgements

Big thanks to my main supervisor, Kristian Martinsen, for his support and motivation during these four years. I would like to thank my co-supervisors, Knut Sørby and Erik Andreassen, for their valuable feedback and critical thinking.

I would like to thank the Norwegian Research Council and the MKRAM project for funding my work. I wish to express my gratitude to the colleaguesn at the SINTEF Materials and Chemistry, for helping me to conduct the experiments, especially Marius Johansen and Erik Andreassen.

Big thanks to my colleagues at the Department of Manufacturing and Civil Engineering for providing a friendly environment and interesting discussions. Thanks to Pål Erik Endrerud, Kenneth Kalvåg, and Tor Erik Nicolaisen for their time and support in all my laboratory work at Additive Manufacturing laboratory "AddLab". Thanks to Torbjørn Skogsrød and Iver Eugen Jensen for exceptional leadership. Thanks to Torbjørn Leirmo for exciting conversations. Thanks to Chunhong Luo for her top-notch administrative support.

Thanks to all my friends for being there in hard and good times, especially Anastasiia Moldavska, Christoffer Hallstensen, Alla Marchenko, Andrii Shalaginov, Marina Shalaginova, Khrystyna Solomchak, Vasilis Gkioulos, Ambika Shrestha, Sushil Chitrakar, Eivind Johansen, and Therese Havnsund.

Big thanks to my parents for their support and teaching me never give up.

Tremendous thank to my beloved Oleksandr Semeniuta for his patience and enormous support throughout this journey.

Contents

List of Tables

List of Figures

List of Abbreviations

ABR	AdaBoost regressor
AM	Additive manufacturing
AMF	Additive Manufacturing file format
ANN	Artificial neural network
ANOVA	Analysis of variance
ASCII	American Standard Code for Information Interchange
ASTM	ASTM International (formerly known as American Society for Testing and Materials)
BP	Back-propagation
BPNN	Back-propagation neural network
CAD	Computer-aided design
CMM	Coordinate measuring machine
CSV	comma-separated values (file format)
DIN	Deutsches Institut für Normung (German for *German national organization for standardization*)
DOE	Design of Experiment
EBM	Electron Beam Melting (AM process)
FDM	Fused Deposition Modeling (AM process)
FEA	finite Element Analysis
FEM	Finite Element Method
FFF	Fusion Filament Fabrication (AM process)
FS	Feature Selection
GA	Genetic algorithms
GBR	Gradient boosting regressor
HA	hydroxyapatite (material)
HP	The Hewlett-Packard Company
ISO	The International Organization for Standardization

KDE	Kernel density estimation (function)
LAPS	Large area projection sintering
MAE	Mean Absolute Error
MGGP	Multi-gene genetic programming algorithm
MKRAM	Material Knowledge for Robust Additive Manufacturing
ML	Machine learning
MLP	Multilayer perceptron (ANN)
MOPSO	Multi-objective particle swarm optimizer
MSE	Mean Squared Error
NN	Neural Network
NRC	Norwegian Research Council
NSGA-II	Non-dominated sorting genetic algorithm
NTNU	Norwegian University of Science and Technology
PA	Polyamide
PBF	Powder Bed Fusion
PCA	Principal component analysis
PCS	Polar coordinates system
PPBF	Polymer Powder Bed Fusion
PSO	Particle swarm optimization
PSW	The process software (controls the EOS AM system)
RF	Random forest method
RMSE	Root mean square error
RP	Rapid Prototyping
RQ	Research question
SDG	Sustainable Development Goals
SE	System Engineering
SLM	Selective Laser Melting
SLS	Selective Laser Sintering
SPADE	System engineering framework stays for Stakeholders, Problem, Alternatives, Decisions and Evaluation
STL	Stereolithography (file format)
SVM	Support Vector Machine
SVR	Support Vector Regressor
USC	University of Southern California
WCS	World Coordinate System

Chapter 1

Introduction

1.1 Introduction

Additive Manufacturing[1] has been developed and improved significantly during the past a few decades. Patents expiration has opened up new opportunities for both researchers and industry to obtain a deeper understanding of AM processes. However, a complete mathematical description of the AM processes is yet a research gap.

Any AM process consists of three main components; machine (usually characterized by process parameters), material, and CAD model(s). The development of an envisioned comprehensive mathematical description of a process should be based on all related aspects of the components mentioned above. The formulation of an analytical form is infeasible at this time. At the current state, a more in-depth understanding of components and their interrelations should be acquired first. While the CAD model can be the same for any AM process, the choice of material and machine is dependent on each other. In other words, if a metal material needs to be used, limited types of AM machines can be of interest. Conversely, a choice of AM machine also sets constraints for a material choice.

Seven major types of AM process categories are presented in ISO/ASTM 52900:2015 [62], which are binder jetting, direct energy deposition, material

[1] AM is a process of joining materials, usually layer upon layer, in order to produce parts from 3D models [62]

extrusion, material jetting, powder bed fusion, sheet lamination, and vat photopolymerization. Each AM process category differs by material, and the type of energy used to fuse the material. For instance, powder bed fusion processes[2], also known as SLS, SLM, and EBM, fuse metal or polymer powder with laser or electron beam. In contrast, material extrusion, also known as Fused Deposition Modeling (FDM) or Fusion Filament Fabrication (FFF), selectively dispenses a plastic fillament[3] through a nozzle [62]. Thus, different AM process categories have different parameters that are used to describe the process. Additionally, metals and polymers have different material parameters. Therefore, a mathematical description of each AM process should be done individually, depending on the used material and process.

Usually, a set of developed models that describe the physical phenomena of a process is used for numerical simulation to detect possible defects in order to avoid them beforehand. Process and material models are developed independently from each other due to the inherent complexity of the task. Smith et al. [102] report that modeling of AM process "is exceedingly difficult due to the highly localized and drastic evolution that often occurs over the course of the manufacture time of each component." Since this research is a part of the Norwegian knowledge-based project, MKRAM[4], and powder bed fusion processes are predefined by the project, this AM process category is of interest in the thesis.

Modeling of the powder bed fusion process is often tackled from two different angles. On the one hand, machine-related process parameters are optimized to improve the quality of fabricated parts. On the other hand, the packing of a build chamber[5] by optimizing part's position and orientation is considered as another machine-related group of parameters, which will be referred to as build chamber parameters.

Optimization of process parameters such as laser power, scan spacing, scan speed, beam compensation, contouring and hatching, hatch length, the inertia of scanning mirror, scan direction, bed temperature, and layer thickness

[2]PBF is an AM process in which thermal energy is selectively applied to fuse regions of a powder bed [62]

[3]Usually, thermoplastic or composite material

[4]Material Knowledge for Robust Additive Manufacturing [1]

[5]Build chamber is "enclosed location within the additive manufacturing system where the parts are fabricated" [62]

has been already presented in the literature [18, 21, 22, 30, 33, 33, 74, 90, 97, 101]. While part orientation, position, build chamber packing density, and strategy have also been reported as important parameters [22, 59, 119], only a few have attempted to investigate a combination of build chamber parameters with machine process parameters [97].

In addition to process parameters, influence of material parameters is also extensively described in the literature. Powder morphology, viscosity and particle size are defined as important material characteristics used to predict quality of the parts with a focus on mechanical properties [37, 39, 40, 50, 51, 77]. Besides, combination of the material, process and build chamber parameters (mainly part orientation) have been intensively investigated and used in Finite Element Analysis [16, 18, 22, 26, 36, 37, 44, 59, 64, 68, 69, 77, 91, 92, 93, 113, 114, 116, 119].

Zhu et al. [124] have presented three main mechanisms of error generation in AM processes, which are the mathematical geometry approximation error (conversion from CAD to STL model), error due to process parameters, and material-related error (thermal shrinkage and material distortion). In addition, Calignano [20] has reported that a combination of part position, orientation, and a number of mesh triangles for the powder bed fusion AM process have an impact on dimensional accuracy. The importance of part position within a build chamber was also highlighted in [86] due to its effect on the cooling time (which is essential for mechanical properties).

Therefore, there is a need to investigate how the synergy of STL model parameters and build chamber parameters, namely part orientation and position, affect the quality of fabricated parts. In this work, geometric accuracy and mechanical properties are used as the two main aspects of quality measures. The integration of information about the build chamber and STL model parameters is referred to as a build layout design.

In addition to the need for a better understanding of how a build layout design affects the geometry and mechanical properties, there is a need for predictive models to control and manage possible deviations. This task is important since AM has found an increased use for fabrication of functional parts [102, 118], especially for electronics, automotive, medical, and aerospace industries [102, 105, 117]. Thus, requirements for quality have become more stringent.

On the one hand, mathematical modeling, in combination with finite element analysis, is widely used to model a sintering process. Still, developed models are simplified and restricted to a specific set of parameters [16]. Thus, simulated results are not within the required tolerances. Machine learning (ML) methods, on the other hand, show promising results when modeling tasks are complex and have a large number of both linear and nonlinear interrelated parameters. Machine learning techniques can find hidden patterns in the data, and then, the accumulated information in the form of trained ML models can be used to solve new unknown problems.

Even though there is a wide choice of machine learning techniques, Artificial Neural Network (ANN) and Genetic Algorithm (GA) are the most used machine learning methods in AM. For example, several studies [46, 80, 90] reported that geometric deviation could be minimized via prediction of scaling ratio or through optimization of process parameters by applying ANN machine learning techniques. Rong-Ji et al. [90] and Shen et al. [99] used ANN to investigate the effect of process parameters on the density of produced parts and to develop predictive models.

In the recent publications, machine learning techniques have been used for prediction of shrinkage behavior, surface quality, optimization of process parameters, real-time process monitoring using both electric signals and images of each layer [32, 81, 83, 88, 94, 100]. This is only a short list of available literature, but it shows that application of machine learning techniques is justified for getting understanding of variations in AM.

Additionally, the application of machine learning is especially effective for data analysis collected from process monitoring by using different types of sensors [10, 17, 78, 93]. The process monitoring systems, available nowadays, are typically used for quality assurance, but only at the stage of fabrication. On the one hand, such monitoring systems contribute to quality improvements. On the other hand, if part orientation or position has not been optimized, such monitoring systems are not able to compensate for the possible deviations via tuning only process parameters of the AM machine. Therefore, there is a need for a decision support tool that can be used at a stage of the build layout design by operators.

The work underlying in this thesis combines two perspectives: (1) gaining knowledge about how a build layout design influences the quality of the produced parts; (2) and how the obtained knowledge can be transformed into

a user-friendly decision support tool to minimize variations in AM prior to the fabrication. One of the main goals of this work is focused on evaluating which machine learning techniques are the best for the development of models that will become a part of an intelligent system for quality assurance in the powder bed fusion machine.

1.2 The MKRAM project

The work underlying in this thesis is a part of the Norwegian project MKRAM (Material Knowledge for Robust Additive Manufacturing) funded by the Norwegian Research Council. The main goal of this project has been set to gain more in-depth knowledge about material behavior, and how material variation influences mechanical properties for the powder bed fusion AM process, which would result in material models for selected materials. The project investigated plastic and metallic materials separately, and this work is a part of the polymer research focus.

While the main interest was to develop material models for their further use in finite element analysis (FEA), this work shows another perspective on how different build layout designs influence material properties and geometric accuracy. How part's position, part's orientation, and other build layout related features can be used for the prediction of material properties and geometric variations is also presented in this work.

1.3 Research questions

This thesis aims at answering the following research questions:

RQ1. How does the build layout design affect product quality in polymer powder bed fusion systems?

This research question aims at understanding how the quality of parts produced by a polymer powder bed fusion system is affected by different build layout designs. Part orientation, part positioning, and STL model characteristics are chosen as the main parameters to describe build layout design. Quality of the produced parts is described in terms of geometric accuracy and mechanical properties, namely tensile modulus, tensile strength, and strain at break.

RQ2. How can machine learning techniques improve dimensional accuracy

of AM?

The second research question aims at addressing the issues of insufficient geometric accuracy of additively manufactured parts. Many different factors may have an impact on dimensional accuracy. While traditional mathematical descriptions of isolated issues may be found in the literature, there is a need for ML techniques that will allow using historical data to find hidden patterns in the data. Therefore, different machine learning methods are applied, and their performance is compared to select the most suitable method for the task.

RQ3. How can application of ML techniques contribute to control and management of mechanical properties of AM built products?

Similarly to RQ2, this research question aims at addressing the issues of inconsistent mechanical properties of additively manufactured parts. The prediction of mechanical properties for a specific position of a part is an important task. Therefore, machine learning techniques are applied to develop predictive models and determine which parameters are significant for this task.

RQ4. How to utilize a build chamber of powder bed fusion machines in a more sustainable way?

In the powder bed fusion system, if more parts can be fit in one building cycle, the cost of each part will be reduced, and material usage will be more efficient. Therefore, this research question aims at addressing how the developed models can help to utilize more space in a build chamber while still maintaining the desired quality.

1.4 Contributions of the thesis

The thesis describes the following contributions:

- Experimental investigation of how a build layout design affects product quality in polymer powder bed fusion system, EOS P395, has been done.

- Data from 1526 standard specimens for tensile testing has been collected through the experimental work, which was not available from before, and can be used by other researchers in the future as a starting point for new models' developments.
- A new data acquisition and analysis approach has been developed, which includes data registration rules for the stages before, during, and after the AM process.
- Design of calculation modules and functional requirements of an intelligent system for quality assurance for powder bed fusion machines has been proposed.
- Machine learning models have been developed for the prediction of dimensional deviations based on the part position, orientation, STL model characteristics, and build layout design. A prototype of a module of the intelligent system for quality assurance has been developed based on the developed models.
- Machine learning models are developed for prediction of scaling ratio individually for each part in the build layout design. A prototype of a module of the intelligent system for quality assurance has been developed based on the developed models.
- Machine learning models are proposed for mechanical properties, namely tensile modulus, tensile strength and strain at break, based on the part position, orientation, STL model characteristics, and build layout design. A prototype of a module of the intelligent system for quality assurance has been developed based on the developed models.
- Optimized part placement in the build chamber with the help of newly developed machine learning models has been proposed. A prototype of a module of the intelligent system for quality assurance has been developed based on the developed models.

1.5 Thesis outline

Due to the nature of the research, the rest of the thesis is structured, as shown in Fig. 1.1. The theoretical background is divided into two chapters. Chapter 2 aims at describing theoretical background about AM, and a current state-of-the-art regarding issues raised in the research questions, while

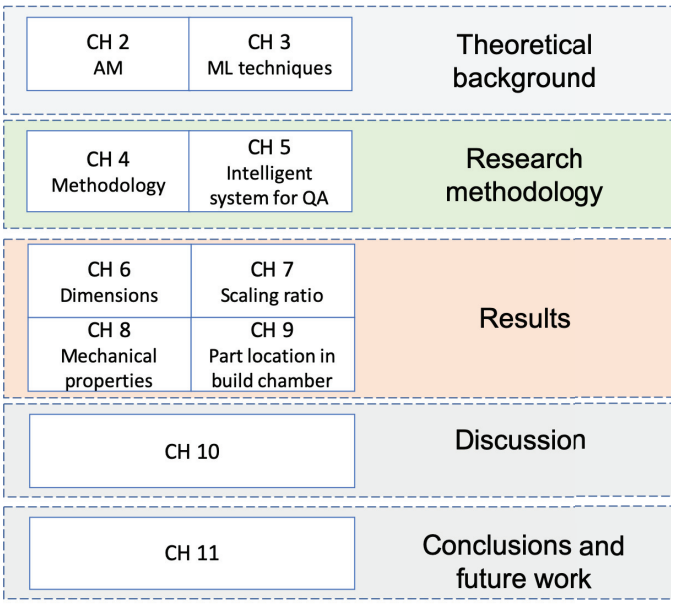


Figure 1.1: Thesis outline

Chapter 3 is devoted to the description of foundations of data science techniques that are used for data analysis.

The research philosophy, design of experiments, and mechanical testing are described in the details in Chapter 4. Then the description of how system engineering was applied to design an intelligent system for quality assurance for powder bed fusion machines is introduced in Chapter 5. Additionally, Chapter 5 also describes the reasons behind the structure of the representation of the results.

The results are presented as separate modules that correspond to the design of the intelligent system for quality assurance of PPBF. Thus, Chapter 6 is devoted to (1) the analysis of measured length, width and thickness, (2) to the investigation of correlation between chosen parameters (build layout design and STL model properties) and dimensions, and (3) to the description of the developed machine learning models for prediction of the geometric deviations with the respect to the orientation and location of the part within a build chamber.

Chapter 7 aims at addressing similar issues as Chapter 6, but in terms of the

compensation ratios in x, y, and z axes individually for each part. This step is necessary since it helps to simplify a process of compensating geometric deviations. Chapter 8 is devoted to the analysis of the measured mechanical properties, namely tensile modulus, tensile strength, and strain at break. The correlations between chosen parameters and mechanical properties are also described in this chapter, and the developed machine learning models for the estimation of mechanical properties are presented at the end of the chapter.

Chapter 9 discusses the opportunities related to the optimization of part placement for the desired mechanical properties, and aims at providing machine learning models for the defined task. Discussion about the obtained results with respect to the formulated research questions is provided in Chapter 10. The conclusions and ideas for future work are presented in Chapter 11.

Chapter 2

Theoretical background and state-of-the-art

Description of the theoretical background is started with Section 2.1 focused on the foundations of AM in terms of powder bed fusion system, material, and STL model characteristics. Section 2.2 provides a brief overview of the current progress in the state-of-the-art regarding geometric deviations and mechanical properties in additive manufacturing.

2.1 Foundations about additive manufacturing technology

This section aims at describing the main principles of the powder bed fusion additive manufacturing process. The first section starts with a description of a powder bed fusion system in detail. The second section is about polymer powder and its characteristics. The last section describes the STL model and characteristics that are important for powder bed fusion systems.

2.1.1 Polymer powder bed fusion system

The original patent that describes the main principle of the powder bed fusion process has a name ”Molding process”, and was issued in 1981 [53]. The main idea was to join molding and casting processes into one process, allowing them to produce 3-dimensional (3D) objects faster. It was proposed to use simultaneously two materials, powdered dry sand to form a mold (in

the patent referred to as "mold material") and dry mixture of plaster of paris and sand (referred to as "casting material" in the patent).

Inventors described that materials are filled in a matrix that is divided into square cells (openings in Fig. 2.1). The matrix is inside of a container and always stays at the fixed position. A movable table also called a frame, is placed under the matrix at the first layer to ensure that material remains in the cells. Each cell is filled with the corresponding material: either casting or mold material. When all cells are filled, the table is lowered by a defined step so that the matrix can be filled again. When all layers are finished, water is loaded into the container for a couple of hours to create a chemical reaction between the plaster of paris and water, which allows the material to fuse. As the last step, the mold material is removed by water under high pressure, resulting in the completed 3D object.

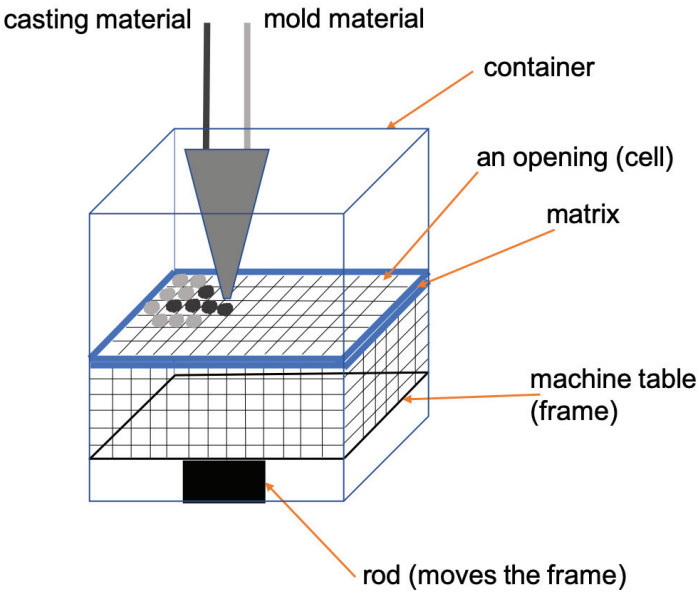


Figure 2.1: Schematic representation of the idea of the molding process from a patent issued in 1981 (adopted from [53])

Housholder [53] described many variations of the abovementioned molding process, and only one of them reminds of the powder bed fusion process as it is known nowadays. The author presented that only one type of powder

material can be used if a laser beam is used as a source of energy to melt the material. In general, this description is the same as known nowadays as powder bed fusion.

According to the ISO/ASTM52900-15 standard [62], powder bed fusion constitutes an "additive manufacturing process in which thermal energy selectively fuses regions of a powder bed". The schematic representation of the powder bed fusion machine is shown in Fig. 2.2. The machine table (powder bed) is first lowered to a one layer thickness. Then the powder is distributed on the bed, and thermal energy is applied to selected regions on the surface to fuse the material[1]. All these steps are repeated in a cycled manner, where the total number of cycles is equal to the number of layers that are predefined beforehand. A number of layers depends on the layer thickness and the size of the object to be produced. For example, if layer thickness is chosen to be 0.12 mm with the height of the object being 75 mm, the number of layers equals 625.

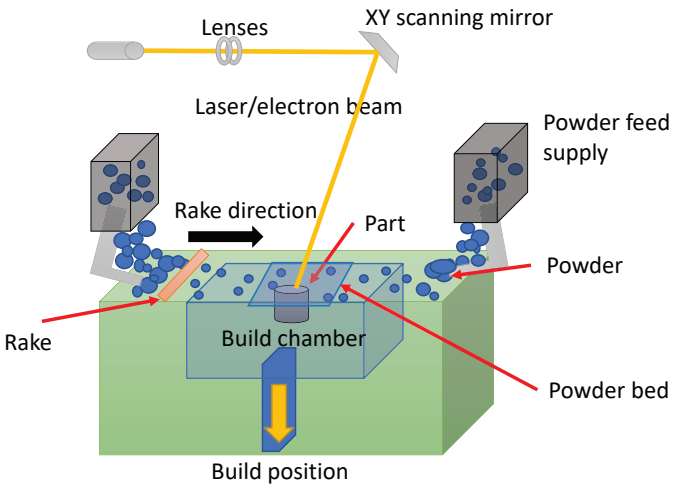


Figure 2.2: Schematic representation of powder bed fusion machine

In polymer powder bed fusion processes, a pulsed laser beam is the most

[1]Metallic, ceramic, composite, and polymer are types of material that can be used for fabrication by powder bed fusion additive manufacturing process.

widely used source of energy [65]. In order to melt or sinter material in selected regions, the laser beam needs to be projected on these regions. The direction of the beam is realized with galvanometer-controlled mirrors [30]. The scanning strategy defines the way how the laser beam scans the surface of the material.

As shown in Fig. 2.3, laser beam movement from one side of the part contour to another side can be defined as a scan line. The direction of the scan lines at each layer determines the scanning strategy. Fig. 2.3 illustrates schematic representation of the *'sorted'* scanning strategy. The thin blue line, in Fig. 2.3, represents the contour of a part at one layer. The thick blue line, shown on PSW example, is the actual contour of a part defined by its CAD model (the laser beam follows a thin line since it is a center of the path line for contour solidification), and green lines/arrows correspond to scan lines. Usually, scan lines are referred to as hatch lines, and the distance between them is called hatch distance (d in Fig. 2.3). According to EOS, the default hatch distance between hatch lines is $d = 0.3$ mm. In addition to the hatch distance, laser power, scan speed, beam offset, and hatch pattern are the main controllable parameters in a scanning strategy. Laser beam offset is used to compensate for the increase in the size of the contour due to curing zones, which are material specific. According to EOS, usually, beam offset is set to 0.33 mm.

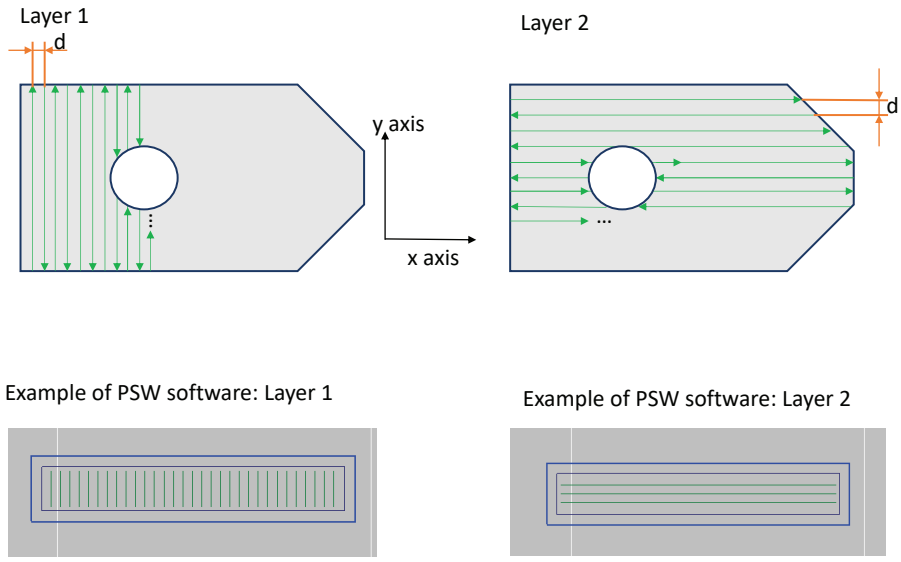


Figure 2.3: Schematic representation of the sorted scanning strategy, including an example from the PSW software by EOS

It is important to mention that the distribution of hatch lines at one layer is dependent on the chosen hatch distance. In other words, one layer can be considered as a 2-dimensional coordinate system consisting of x and y axes (Fig. 2.3). The scan line at each layer starts from the origin point $(0, 0)$ on the "world" coordinate system of the AM machine. Thus, scan lines along y-direction will be distributed as $y = \{0, d, 2d, 3d, ..\}$ for layer 1, and scan lines along x-direction will be distributed as $x = \{0, d, 2d, 3d, ..\}$ for layer 2. So d is a step for hatch line positioning at each layer, and therefore, even and odd layers will always consist of the same number of hatch lines, respectively. According to Pavan et al. [86], this information is important because energy density is considered as a factor that influences both geometric accuracy and mechanical properties.

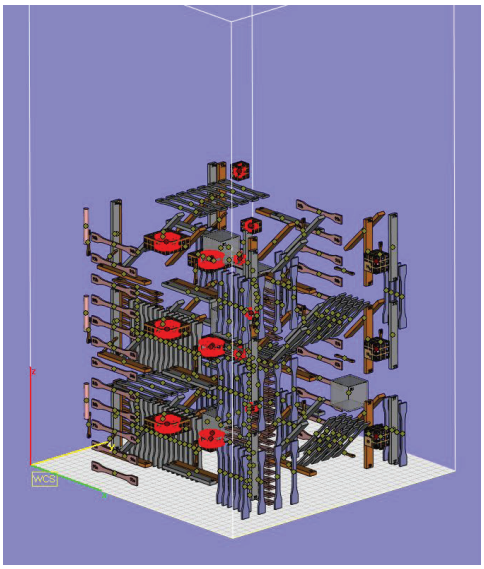


Figure 2.4: An example of a build layout in Magics 20.0 software

In addition to the scanning strategy, build layout[2] is another factor that has an effect on temperature distribution in the build chamber (an example is shown in Fig.2.4) [97]. Build packing strategy describes how parts should be oriented and positioned in the build chamber for effective use of the build volume and the material. Part position (or part placement) in the build chamber is defined in terms of rotation and translation respect to the build chamber global coordinate system (World Coordinate System, WCS).

Since Magics 20.0 software provides information about part placement in terms of central, minimal, and maximal coordinates (see Fig.2.5), it is important to define what these coordinates mean. Minimal coordinate corresponds to a point on the part that is placed closest to the origin of WCS of the build chamber. In contrast, the maximal coordinate describes the position on the part that is farthest from the origin of WCS.

Fig. 2.5 illustrates how minimal and maximal coordinates are defined for a part that is placed in the build chamber in two different orientations. For

[2]Build layout is a result of a build packing strategy, in which optimization of parts' placement and orientation is performed to maximize utilization of the build volume and material usage without negative impact on the quality.

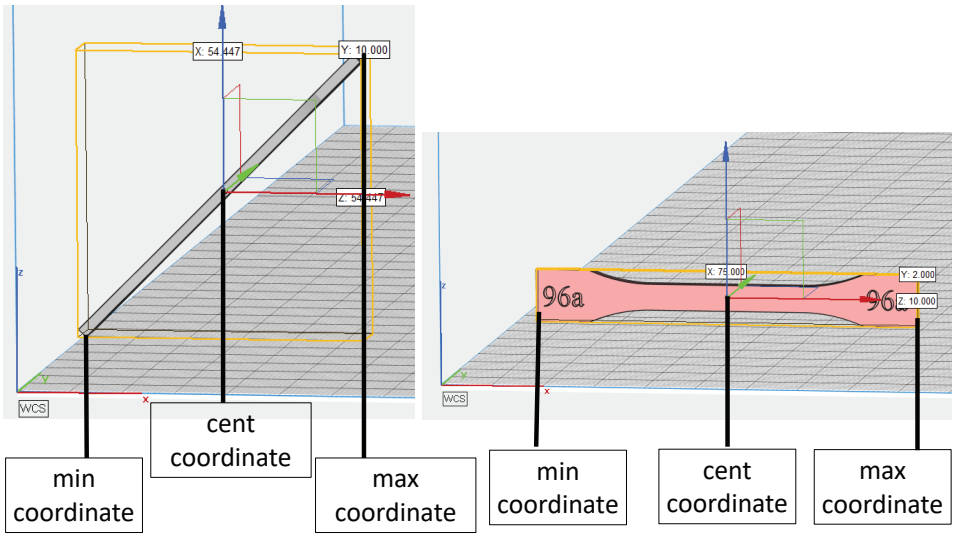


Figure 2.5: Example of part placement description in the build chamber through maximal, central and minimal coordinates for Angle (left hand side) and XZY (right hand side) orientations using Magics 20.0

the part in XZY orientation, the difference between maximal and minimal coordinates corresponds to its dimensional features. However, a distance between maximal and minimal coordinates for parts in Angle orientation will not correspond to the value of a dimensional feature. In other words, part placement coordinates are typically used to describe the part placement in the build chamber, but not to define dimensional features of parts.

2.1.2 An introduction to mechanical properties

A polymer is a material used in the experimental work underlying this thesis. It is well known that material is an essential factor influencing the quality of additively manufactured parts. Therefore, even though the material is not directly included in the list of investigated parameters in this research, a basic overview of material and related quality characteristics need to be presented.

Gedde [47] defined polymer as "a substance composed of molecules characterized by the multiple repetitions of one or more species of atoms or groups of atoms (constitutional repeating units)." PA2200 is a powder based on the material polyamide-12 (PA12), which constitutes of the n - number of the

constitutional repeating unit shown in Fig. 2.6.

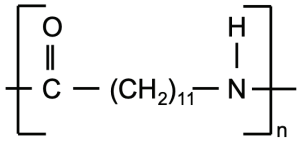


Figure 2.6: The polyamide-12 constitutional repeating unit (adopted from [47])

Investigation of the polyamide behavior includes analysis of powder morphology, particle size distribution, particle density, flowability properties, and powder materials physical properties [111]. Researchers look into the interconnection between material characteristics and AM parameters (e.g., energy density, layer thickness, laser power). However, the main driving force why these interconnections are investigated is due to the aim of revealing their correlation with the final part quality. Depending on the application areas of the produced parts, final quality can be described in terms of part density, surface quality, geometric accuracy, internal build flaws, and mechanical properties [111].

Mechanical properties can be described differently with respect to the chosen mechanical testing techniques. In this work, tensile testing is conducted to evaluate the strength and stiffness of the produced specimens (standard dogbones described in more detail in Chapter 4). A tensile test is a technique used to define the relationship between stress and strain of the material. Different material characteristics can be determined from a stress-strain diagram. In this work, the mechanical properties of PA 2200 will be evaluated and analyzed with respect to Young's modulus, tensile strength, and elongation at break.

Young's modulus, also known as tensile modulus or modulus of elasticity, measures material resistance to elastic deformation under the applied load. According to ISO 527-1:2019 standard [61], tensile modulus is a " slope of the stress/strain curve $\sigma(\epsilon)$ in the interval between the two strains $\epsilon_1 = 0.05\%$ and $\epsilon_2 = 0.25\%$", and is expressed in megapascals (MPa).

Tensile strength, also known as maximum stress, is referred to as strength in ISO 527-1:2019 standard [61] and is defined as "stress at the first local maximum observed during a tensile test." Similar to the tensile modulus,

strength is also expressed in megapascals (MPa).

The elongation at break, also known as the strain at break, is the maximum relative deformation of the specimen before it breaks. The strain in uniaxial tension, defined in eq. 2.1, is the ratio of the change in gauge length upon loading (ΔL) to the original gauge length (L) [61].

$$\epsilon = \frac{\Delta L}{L} * 100\% \tag{2.1}$$

A more detailed overview of the parameters that influence investigated mechanical properties is described in Section 2.2.2.

2.1.3 Highlights of STL model characteristics and their importance

In order to fabricate a 3-dimensional (3D) CAD model by additive manufacturing, it should be first converted into STereoLithography (STL) model (saved in 'STL' file format). This file format was developed in 1987 by 3D Systems. STL is considered as the most used CAD file format in AM [9]. Conversion of CAD model into an STL model is a process of tessellation of CAD model into a set of triangular facets, which represents a surface feature of a geometry [9, 58].

Tessellation of CAD model into STL format frequently ends with errors, which lead to gaps and holes in data structures [106]. Usually, these types of error create open loops in cross-sections, and such cross-sections will not be possible to produce with AM technology. Due to these issues and a need for more information being stored, there is a broad discussion in AM on using more advanced geometry file formats, namely AMF and 3MF. Additionally, more advanced users make attempts to use CAD models directly for AM fabrication. However, STL still stays as the most used geometry file format, especially for powder bed fusion AM processes.

There are two types of STL files; binary STL and ASCII STL format (examples are shown in Fig. 2.7). Even though ASCII STL is easier to read, the binary STL is mostly used in practice. The main reason for that is because, for a large part, the ASCII STL file becomes too large and inconvenient to use. Therefore, in this work, the binary STL file is used and referenced to as STL file format.

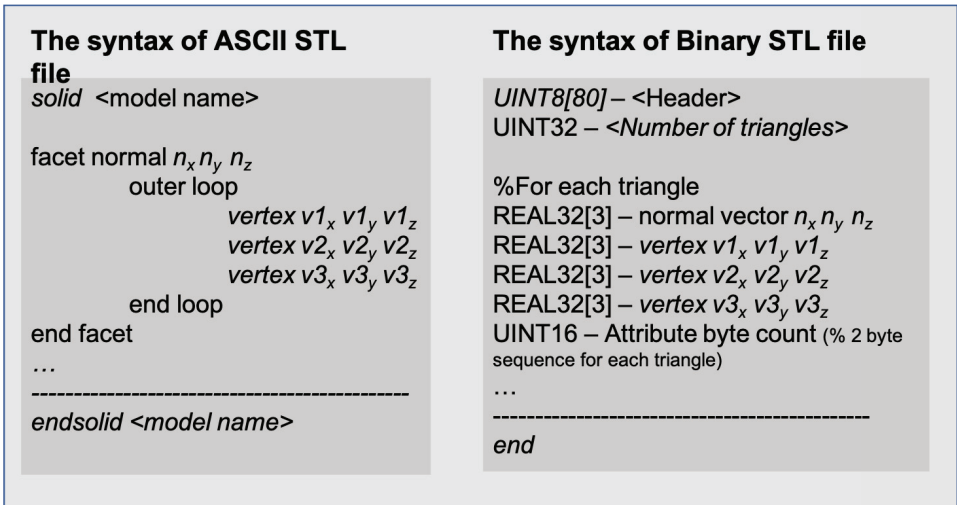


Figure 2.7: Representations of ASCII STL format and Binary STL file format

STL file is a set of oriented triangles (facets), which can be numerically represented as (n_x^i, n_y^i, n_z^i) coordinates of outward normal of i^{th} facet, and $(v_{kx}^i, v_{ky}^i, v_{kz}^i)$ coordinates of k^{th} vertex of the i^{ith} triangle, while $i = 1, .., N$ is a number of mesh triangles [106]. Graphical representation of vertices and normal vector is shown in Fig. 2.8. The normal vector needs to be perpendicular to the facet and should point outwards with respect to 3D model.

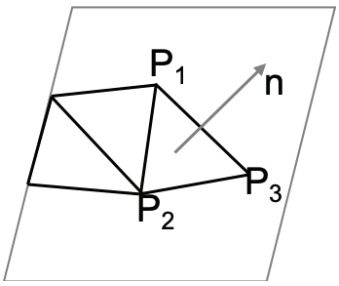


Figure 2.8: An example of a facet in STL model with the corresponding normal vector and the verteces (e.g. $P1 = (v1_x, v1_y, v1_z)$)

All triangular facets in the STL file must comply with the following two rules [58]:

1. **Vertex rule**. Each triangle must share two vertices with its adjacent triangles.

2. **Facet orientation rule**. The orientation of a facet must be described to satisfy two conditions; (1) direction of the normal vector is outward, and (2) the order of listing vertices is counterclockwise when looking at the object from outside.

Often, many CAD systems could generate STL files that are incorrect since they do not comply with the two rules described above. Typically, the main reason for that is a very complex design of the CAD model. The incorrect direction of the normals, overlapping facets, and cracks are common the STL file errors [58]. Therefore, it is recommended to use STL file repairing programs before 3D printing them. In some cases, not all errors can be corrected, which means that either there is a need for model redesign or using algorithms that allow repairing contour at 2D level, formed as slices from the STL model rather than repairing the STL file in general [58].

Since generated STL file needs to be sliced into layers of 2D contours [9, 106], contour generation is an important part of this process. Fig. 2.9 illustrates a schematic representation of how slicing of the mesh model is performed for contour generation by a generic algorithm [76].

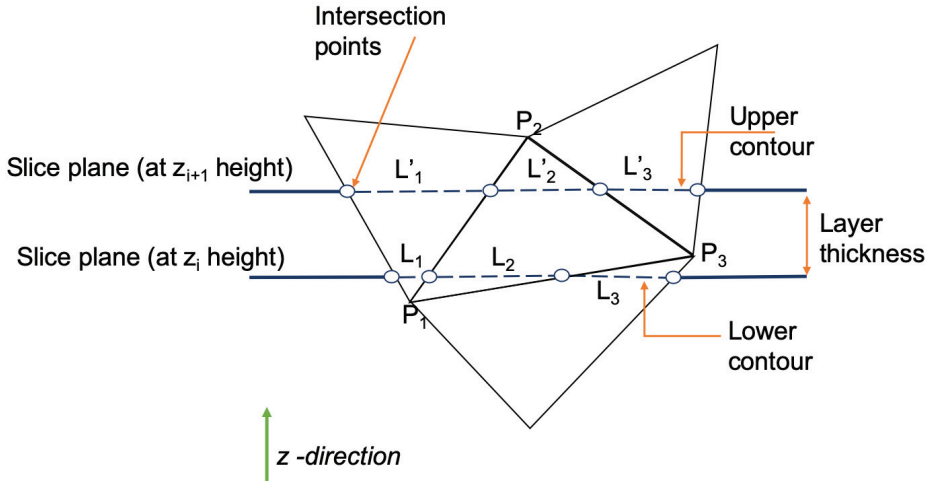


Figure 2.9: Using horisontal planes for contour generation (adopted from [76])

If to look at the simplified process of contour generalization, one can consider slicing of the mesh model with horizontal planes (parallel to XY plane), incremented along the z-axis (shown in Fig. 2.9). Intersections of a plane and triangle edges create intersection points (vertices). Created intersection points within one slice plane are connected into a set of lines. Lower contour can be described as set of L_1, L_2 and L_3 line segments, and line segments $L_1^{'}, L_2^{'}$ and $L_3^{'}$ define the upper contour. Distance between the upper and lower contours is equal to layer thickness. Thus, two generated contours can be defined as a set of segments $\{L_1, L_2, L_3\}$ and $\{L_1^{'}, L_2^{'}, L_3^{'}\}$.

Therefore, the STL file format is an important factor in additive manufacturing. Accuracy of 2D contours may be affected by changing the number of triangles, and thus resulting in a smaller or larger volumetric error, which leads to geometric deviations. Volumetric error determines numerically difference between generated contours of the STL model and designed a CAD model with respect to the part's orientation [76].

2.2 A brief overview of the current state-of-the-art

A brief overview of the current state-of-the-art is presented as three different subsections. In the first subsection, an overview of the recent studies regarding geometric variations in AM is presented. The second subsection describes the current state-of-the-art of the studies about mechanical properties in additive manufacturing. The last subsection illustrates an overview of how machine learning techniques are used for additive manufacturing technologies.

2.2.1 Geometric variations in additive manufacturing

Compensation of shrinkage effect based on the STL model

In the series of research publications [54, 55, 56, 57, 73, 124] the issues of geometric deviations related to the shape deformations at the stage of processing STL file are raised and addressed. Zhu et al. [124] highlighted that due to the layer-wise nature of additive manufacturing technology, there exist in-plane and out-plane deviations of the product shape from the designed shape. The in-plane deviations occur as a result of an error effect inside each layer, while out-plane deviations are due to errors between layers.

Zhu et al. [124] have adopted a multi-task Gaussian process machine learning algorithm to model the local variations, namely in-plane variations. A general model for prediction of in-plane (x-y plane) geometric deviations for various shapes (cylindrical [57], polygon [55] and prismatic shapes [56]) has been described. The main idea of Luan and Huang [73] was to decouple the complexity of geometric shapes from the deformation modeling and thus to develop a predictive learning strategy of freeform objects.

The optimal compensation policy for 2D and 3D shapes has been proposed by Huang [54], where minimum area deviation and minimum volume deviation are used as quality measures. While deviation compensation for 2D shapes is addressed via transformation of shape deformations into deviation profiles by using a polar coordinates system, the 3D shapes deformations have been addressed as a unified formulation of in-plane and out-plane deformations by using a spherical coordinate system. Both of the provided annotations are built on the previously published works [55, 56, 73] and show great potential for addressing issues of dimensional inaccuracies in additive manufacturing.

In the most recent research activities, application of machine learning techniques for predictive modeling of dimensional inaccuracies of freeform 3D shapes based on the STL model [31]. Decker and Huang [31] have used a random forest machine learning method to predict the physical position of a vertex in the print bed. The orientation and curvature of a surface are significant for defining surface slope properties with possible errors like the stair-step effect.

Usually, the stair-step effect is tackled from another perspective, where researchers focus on the development of new slicing algorithms that will help to minimize geometric deviation [9, 35, 76, 109]. Since larger CAD models with more complex designs result in the increasing size of the STL model, and thus, resulting in the decreased efficiency of existing slicing algorithms. Therefore, Ding et al. [35] proposed a multi-direction slicing algorithm for 3D shapes with respect to part's orientation. In order to be able to use the proposed algorithm, AM machines need to be able to conduct a deposition path trajectory along multiple directions.

For instance, error compensation by correcting STL files has been evaluated on fused filament fabrication [109]. Two types of compensation errors have been defined: translational error individually for x, y and z axes, and

rotational error in y and z axes. However, dimensional inaccuracy in the z-direction was not significantly improved due to layer quantization, which means that geometric variations along the z-axis are less controllable.

Hao et al. [52] have developed a method for decomposing and reconstructing the STL model into several meaningful sub-models, which are easier and more efficient to fabricate. The authors have validated the proposed curvature-based partitioning algorithm on the large-scale STL model, which improved the geometric accuracy. However, the usage of the feature plane as the cut plane leads to incomplete cutting corners [52].

Even though these solutions have been validated on simpler additive manufacturing processes, like stereolithography and material extrusion, there is no discussion on how the developed models can be utilized for powder bed fusion machines, or how optimization of process parameters can contribute to the minimization of geometric inaccuracies.

Shrinkage effect compensation by optimization of process parameters

Even though compensation of geometric deviations based on STL model improvements have shown great potential, Cheng et al. [23] highlighted that including process parameters is also an important step towards minimization of geometric inaccuracies. Therefore, the authors have extended the proposed in-plane compensation model from Huang et al. [57] by including information about process parameters of material extrusion additive manufacturing technology. The proposed framework, namely Gaussian Process and Kernel Smoothing scheme, consists of the first part, based on in-plane shape deviation with respect to STL file and process parameters, and the second part, smoothing the new compensated contours of the 2D shape.

Often researchers connect shrinkage effect with geometric deviations. Thus, Negi and Sharma [81] reported that dimensional accuracy could be improved as a result of minimizing shrinkage effect by optimizing part bed temperature, scan speed, and scan spacing for polymer powder bed fusion machine. The main algorithm utilizes an artificial neural network. The authors discussed the relationship between shrinkage effect and polymer crystallization rate, which is a significant factor for defining a tendency for part's curling or other geometric deviations, and is dependent on the temperature inside the build chamber.

Optimization of laser power, scan spacing, bed temperature, and hatch length of polymer powder bed fusion process (3D Systems) was performed to predict shrinkage effect [101]. The authors used Analysis of Variance (ANOVA) based on the quadratic model for shrinkage, and their results showed that all parameters are significant, with scan spacing being the most significant parameter. When scan spacing is too narrow, the overlapping zones of hatch lines lead to a higher temperature in such places and an additional sintering-cooling sequence.

However, if scan spacing is too large, poor packing of the particle will take place, and thus layers can have a tendency of curling and clinging to the powder distribution rake (roller) [101]. Similarly, powder bed temperature also influences the process of the powder recrystallization process. Lower temperature leads to lower recrystallization and smaller shrinkage effect, while higher powder bed temperature provides better density and strength for fabricated parts, but larger dimensional inaccuracy [101, 119].

Caulfield et al. [22] investigated thickness and width in terms of their dependence on energy density and build orientation. However, authors haven't included the length of the part and "Y" orientation that could provide more information for a better understanding of the sintering process. As a result, Caulfield et al. [22] have documented that the "role of the build orientation and parts dimensions may be more complicated than the influence of energy density." Therefore, more investigation of dimensional accuracy interaction with build properties is needed.

Yang et al. [119] proposed a set of models for optimization of shrinkage ratio for part placement in X, Y, or Z directions. Models are built based on the results from Taguchi and Analysis of Variance and supported their optimization with experimental testing of models. The results obtained from these models led to decreased shrinkage error from 84% to 64% comparing with the results of other commercially available methods, and thus improving the accuracy of part dimensions significantly [119] .

Senthilkumaran et al. [97] investigated the influence of different building strategies on the shrinkage effect. Authors studied the effects of beam compensation, contouring, and hatching, the inertia of scanning mirror, scan direction, and compensation of positioning errors on shrinkage effect. Additionally, the impact of part's orientation has been evaluated with respect to geometry deviations per unit length. Senthilkumaran et al. [97] observed

that shrinkage effect of parts oriented in XYZ orientation is higher than in YXZ direction ($(0.4 - 0.6)\%$ and $(-0.4 - 0.35)\%$, respectively).

Later, Senthilkumaran et al. [98] introduced a new model for shrinkage compensation based on the results and gained knowledge from the previous study. This model was developed for compensation of shrinkage "at every layer and at every hatch length, unlike a uniform compensation scheme applied to the entire part" [98]. Results were compared with the compensation suggested by the machine manufacturer, and improvements of dimensional accuracy approximately by 55-62% were observed for the newly developed compensation scheme.

Similarly, Senthilkumaran et al. [98] proposed a model for compensating shrinkage behavior with respect to a part position and exposure strategies (such as scan length and scan distance). The shrinkage effect was reported as highly non-uniform, especially along the z-axis due to lower scan length. While in-plane shrinkage compensation is easier to achieve by using developed models along x and y directions separately, the shrinkage effect along the z-axis is affected by both part position, and distance between hatchlines [98]. As a result, the authors highlighted the complexity of such a model and haven't developed model for shrinkage compensation along the z-axis.

Another research work [65] shows that optimization of scanning strategy is one of the solutions for minimization of shrinkage effect, and thus geometric deviations. Jhabvala et al. [65] stated that scan patterns, including scanning orientation, hatch spacing, and beam offset, are among the most significant parameters. Moreover, these parameters play an important role in the thermal evolution, and deformations related to the temperature distribution in the build chamber [123]. The authors proposed a thermal model for layer-wised simulation of the powder bed fusion process as a means for geometry improvements via scanning parameters' optimization [123].

Investigation of the relationship between the shrinkage and layer thickness, hatch spacing, laser power, scanning speed, the temperature of a working environment, interval time, and scanning mode have been presented byWang et al. [113]. Delgado et al. [33] also evaluated the significance of the effects of process parameters on dimensional error, surface roughness, and mechanical properties for metal powder bed fusion systems. The authors also reported that research on dimensional accuracy for two metal materials is limited

comparing with surface roughness and mechanical properties.

As shown above, the compensation of geometric deviations with the help of optimizing STL models and AM process parameters are actively investigated nowadays. The proposed solutions show great potential, while at the same time, there is a number of research gaps yet to be addressed. First of all, there is a need for more in-depth knowledge on how the STL model parameters affect the geometric deviations. While optimization of AM process parameters and correction of STL files are performed isolated from each other, the combination of these two aspects is yet not well studied. Besides, the current state-of-the-art shows that there is a lack of understanding of how dimensional inaccuracies vary within the build chamber of the powder bed fusion process. Most of the presented studies describe the geometric deviations based on the limited number of produced objects. Moreover, the usage of larger space in the build chamber could contribute to a different deviation profile.

2.2.2 Mechanical properties in additive manufacturing

The current state-of-the-art [14, 16, 18, 22, 26, 36, 37, 44, 59, 68, 69, 77, 91, 92, 93, 103, 104, 113, 114, 116, 119] describes the importance of part orientation, powder morphology, and machine process parameters as a means towards the control and management of variation in polymer powder bed fusion system.

Laser power, scan speed, hatch distance, scan strategy, melting temperature, and powder bed temperature are among the most investigated AM machine process parameters [36, 44, 68, 91, 113, 114]. There is a number of studies [41, 49, 50, 79] which report that laser power, scan speed, hatch distance, and layer thickness can be used to define the line energy and how their variation may influence mechanical properties of the part. In addition to energy applied to solidify polyamide, Mielicki et al. [77] have also reported the importance of layer thickness and powder distribution in each layer. Powder distribution is dependent on the size of the particles and powder viscosity.

Furthermore, Drummer et al. [37] and Gümüs et al. [51] studied how the size of the particles of the polymer powder and its viscosity influence mechanical properties. While Drummer et al. [37] investigated the degradation behavior of PA12 based on the analysis of phase transition temperature and melt

viscosity for both virgin and aged powder, Gümüs et al. [51] have reported that the size and morphology of the particles could lead to the creation of pores, gaps, or/and voids in the fabricated parts. Flodberg et al. [40] have reported that porosity is an important factor leading to the degradation of mechanical properties. Two different types of polyamide-12 was evaluated, namely, Duraform ProX Nylon PA12 and PA12 reinforced by carbon fibers. It was found the PA12 reinforced with carbon fibers have better mechanical properties and lower porosity including smaller sizes of pores, while Duraform ProX Nylon PA12 has larger porosity level (including larger sizes of pores) and lower mechanical properties. Besides, Flodberg et al. [40] reported that tensile strength and tensile modulus were less affected by the part's orientation, while strain at breaks has significant variation depending on the part's orientation.

Stichel et al. [103] have performed a Round Robin procedure on six different polymer powder bed fusion machines with the best set of process parameters for each machine, and corresponding choice of the material, namely 12 Duraform PA (3D Systems) and PA2200 (EOS). In a series of publications [103, 104], researchers have reported on "the minor reproducibility regarding the mechanical and dimensional properties." While microstructural origin in mechanical properties with a significant focus on the importance of porosity and residual particles and cores is described in [103], the backtracing of pore morphology to the process parameters were presented in [104].

The highest ductility (elongation at break) and ultimate tensile strength have been observed when a high melt efficiency without coplanar residual particle core arrangements was achieved [103]. The pore morphology was found to be less significant for mechanical properties, but it is reflected in the anisotropic mechanical response depending on the samples' orientation in the build chamber. Stichel et al. [103] have highlighted that porosity and residual particles and core are not linked and can vary independently depending on the machine configurations.

The backtracing of porosity, pore density, pore shape, and pore arrangments of laser-sintered polyamide-12 samples to the process parameters have been described in [104] as a following up article of [103]. The laser energy input temperature was found to be the most critical process parameter that may help to reduce porosity and improve mechanical properties. The combination of 100 μm layer thickness, the area energy density value around 4

J/cm^2, and process temperature close to the melting point are recommended to achieve microstructure with lower porosity, and thus higher mechanical properties [104].

The mechanical properties of samples produced with three different polymer powder bed fusion machines, namely EOS Formiga P100, multijet fusion by HP, and large area projection sintering (LAPS), were compared by Craft et al. [27]. The importance of temperature in terms of lower porosity and higher elongation at break has also been reported [27], and these results are in line with the findings described in [103, 104].

In another study, monitoring and control of a powder morphology by using machine vision in combination with Design of Experiment (DoE) were proposed by Flores Ituarte et al. [41]. In order to control powder flowability, the authors described how machine process parameters need to be varied (e.g., line energy, change the speed of the recoater, adjust layer thickness).

The influence of part orientation on mechanical properties has already been described in detail by [22, 74, 108]. It is important to note that there is a difference between what is reported in the literature and what is provided by EOS datasheets. While [22, 74, 108] have reported that tensile modulus, elongation at break, and maximal stress are affected by the part orientation, EOS reports in their datasheets for PA12-Balanced process parameters group that Tensile Modulus is the same in all orientations. Ituarte et al. [64] have also reported that part orientation has the most significant influence on mechanical properties, among other investigated parameters based on the Taguchi DOE. Caulfield et al. [22] reported that the thermal distribution in the build chamber also has an impact on the mechanical properties of the fabricated parts.

In contrast, Faes et al. [39] reported that mechanical properties are isotropic, and part orientation is not a significant factor. The authors assume that such observation could be connected to the age of the used AM machine since newer AM machines have highly optimized parameter settings, namely "scanning patterns, laser beam settings, improvements to the preheating homogeneity of the powder bed, and also continuous improvement in the sintering characteristics of the material itself" [39]. However, part location within the build chamber was found to be an important factor for tensile modulus and elongation at break.

Tasch et al. [107] investigated the dependence of mechanical properties on the thickness of polymer powder bed fusion produced samples. The thickness was varied from 0.6 mm to 2 mm, and mechanical properties decreased significantly for thickness lower than 0.8 mm. In addition, Tasch et al. [107] highlighted that the results are non-homogeneous, AM system-specific, with anisotropic material properties that depend on position and wall thickness of the produced parts.

The impact of various process parameters on mechanical properties and investigation of material properties with a focus on new material developments are well presented in the current state-of-the-art. Even though researchers have highlighted the importance of energy distribution within the build chamber as a significant factor that affects mechanical properties, the impact of build chamber packing strategies on temperature distribution in the build chamber is not investigated in depth. Several studies evaluated the effect of the parts' position and orientation. Still, more effort towards mapping build packing strategies to mechanical properties needs to be put in the future.

2.2.3 Machine learning in additive manufacturing

Considering latest progress in the fields of AM and ML, Razvi et al. [89] reported in their overview that there are four main ML application areas in AM, namely (1) design, (2) process and performance optimization, (3) in-situ process monitoring and control, and (4) inspection, testing and validation.

The in-situ process monitoring and control was highlighted as one of the most widely presented machine learning applications within additive manufacturing technologies. The data collected from various sensors and cameras are used for mapping process performance to quality metrics [89, 115, 120, 122]. The main interest is set to defining deviations during fabrication and adjusting process parameters in real-time. Since even small changes in process parameters may introduce a larger deviation, this task requires more in-depth knowledge about all possible factors affecting product quality.

The optimization of process parameters in order to improve product performance (or in other words, quality metrics such as surface roughness, dimensional accuracy, strength, strain at break, etc.) is also widely investigated in AM. For example, Artificial Neural Network (NN), mainly

backpropagation NN, and the Genetic Algorithm are the most used machine learning methods for process modeling, optimization, and prediction of process parameters.

Vosniakos et al. [112] have proposed a neural network model to be used as a cost function of a genetic algorithm to optimize volumetric accuracy of shape approximation and build time but for powder bed fusion. At the same time, Rong-Ji et al. [90] attempted to determine the best process parameters to fabricate parts with a higher level of accuracy. Authors focused on such parameters of SLS as the layer thickness, hatch spacing, scanning speed, scanning mode, laser power, interval time, and work surrounding temperature. To obtain optimum process parameters listed above, Rong-Ji et al. [90] applied combination of genetic algorithm and backpropagation (BP) NN algorithm. In this study, results from BPNN were used as input parameters for fitness function in GA. A genetic algorithm was used as a method to determine optimal process parameters based on minimum shrinkage ratio [90].

Padhye and Deb [85] tested and evaluated different methods for multi-objective optimization and multi-criteria decision making. The authors borrowed already described by other scientists two multi-objective evolutionary algorithms, namely non-dominated sorting genetic algorithm (NSGA-II) and multi-objective particle swarm optimizer (MOPSO). These algorithms were tested on 16 different 3D CAD models considering surface roughness and build time as the main parameters that should be minimized [85]. Their results showed that some geometries do not have one best solution, while for other shapes, NSGA-II found a better optimization solution. However, the authors mentioned that this work hadn't been experimentally verified.

When it comes to the task of compensation of dimensional inaccuracies prior to fabrication, the application of conventional mathematical models is a difficult task for estimation of the shrinkage effect [113]. The main reasons are the complexity of the AM processes and the presence of nonlinear and multivariable relationships between dimensional properties and process parameters. Development of a neural network model based on the series of experiments to describe relationships between process parameters and shrinkage effect have been already proposed in [97, 101, 113]. Several studies [24, 31, 110, 124] have also addressed how different ML techniques can be used for estimating dimensional deviations by comparing 3D point clouds

of nominal CAD model and model created as a result of scanning printed product. However, the proposed solution is limited to specific shapes and are not universal.

When it comes to mechanical properties, the porosity (density) estimation with the help of ML techniques is one of the most presented quality metrics in the literature [25, 45, 99, 114]. For instance, Wang et al. [114] and Shen et al. [99] used ANN to develop predictive models for prediction of density for the powder bed fusion process based on laser power, scan speed, scan spacing, and layer thickness parameters.

Garg et al. [45] investigated the prediction of open porosity for SLS fabricated parts from self-made powder as a mix of hydroxyapatite (HA) and polyamide (PA). The authors applied a multi-gene genetic programming algorithm (MGGP) and ensemble-MGGP (EN-MGGP). Results showed that layer thickness, laser power, and laser scan speed has a significant impact on open porosity, and EN-MGGP is better than the classical MGGP algorithm.

Other mechanical parameters are usually modelled with a help of traditional mathematical models that have a number of constraints. Baturynska [15] has made an attempt to estimate tensile modulus, tensile strength, and strain at break through mapping part orientation, part location within a build chamber and STL model properties to the mechanical properties. The four different machine learning methods have been evaluated and compared with linear models, and gradient boosting regressor has outperformed all investigated methods.

As a result, the application of machine learning techniques in additive manufacturing has shown that process monitoring is one of the most promising areas due to the availability of large datasets. At the same time, artificial neural networks, and genetic algorithms are among the most common machine learning techniques applied in other AM areas such as estimation of density, mechanical properties, or dimensional deviations. However, the current state-of-the-art shows that more experimental data needs to be collected in order to benefit from ML techniques. The application of machine learning methods needs to be performed using well-defined methodologies, which are not available yet. Finally, the quality of the collected data should also be controlled either by using existing data science techniques or by the development of new ones.

Chapter 3

Data science as a means towards robust AM

This chapter describes the theoretical foundations of data science techniques that have been used at different stages of data analysis. The first section is focused on the explanation of feature selection methods used to evaluate the correlation between investigated parameters and outcome. The second section describes the theoretical background of the machine learning techniques used to predict the quality characteristics of additively manufactured parts. The third section describes the process of how the model's hyperparameters are optimized and evaluated.

3.1 Feature selection

In data science, features are defined as a number of parameters that are used to predicted desired output. Feature selection filtering methods are used to identify nonsignificant features to exclude them from a model. This step allows for minimizing risks for the model's overfitting and improving generalization capabilities. The selected features from this step are used for training machine learning models by Multi-Layer Perceptron (MLP) ANN method. However, other ML techniques used in this work have a feature selection step incorporated inside their algorithms, and therefore, this step is not required for them.

Pearson correlation test, Spearsman's Rho correlation test, and Mutual In-

formation methods are used to analyze correlation independently between each feature and each target.

Pearson correlation test measures the linear correlation between two datasets, and requires that their distribution is close to normal distribution [3]. The correlation coefficient varies from -1 to 1 and implies no correlation when it is equal to 0. For feature x and a target y, negative coefficient values correspond to a decrease of y when x increases and otherwise.

Spearman's Rho correlation test, unlike the Pearson correlation test, is a nonparametric measure of the relationship between two variables, which are not required to be normally distributed [4]. Spearsman's correlation test is also often referred to as a non-linear test, with its correlation coefficient varying between -1 and 1, and having the same meaning as for the Pearson correlation coefficient.

Mutual Information filtering method measures a dependency between two random variables [5]. This coefficient has non-negative values, and variables are independent only if Mutual Information rank is equal to 0, otherwise higher value means higher dependency. Unlike Pearson and Spearsman's correlation tests, the Mutual Information test does not have a boundary of 1. The concept of mutual information is linked to the Information theory through a combination of entropy estimation of a random variable from k-nearest neighbor distances [67]. While correlation tests have an assumption about the variable's distribution, in the mutual information method, it is assumed that the variable is not categorical.

3.2 Machine Learning Techniques

The first step of an application of any machine learning method constitutes the splitting of the main dataset into training and testing subsets. This is a necessary step since the training subset is used to train models, while a testing subset is used to evaluate models' performance and generalization ability. Depending on the number of available data points, the ratio of training/testing subsets varies. Since, in this work, five-fold cross-validation is conducted at the training stage, the train/test split ratio is set to 85%/15%, respectively. The normalized data is split data with the help of train_test_split Scikit-Learn function.

3.2.1 Artificial Neural Network

Artificial Neural Networks constitute a class of machine learning models that allows defining a complicated non-linear relationship between inputs and outputs. The core idea behind ANNs is in constructing a complex model as a network of processing functions and learning the parameters of these functions using backpropagation. The latter constitutes a method for computation of gradients of a cost function with respect to functions' parameters by propagating the error back through the network architecture and applying the chain rule for differentiation.

Multilayer Perceptron (MLP) is the classical neural network model, based on a sequence of fully connected layers of neurons, where the linear layer-to-layer mapping is activated with a non-linear function.

MLP using backpropagation

A feed-forward multi-layer perceptron using backpropagation is one of the machine learning techniques. This method can be applied for modeling of complex tasks, where more conventional mathematical modeling is difficult or unsuitable. A performance of MLP neural network can be described based on its operational unit, the perceptron. The perceptron takes a set of features as an input vector. Typically, it is represented as a vector $\mathbf{x} \in \mathbb{R}^n$ where n is a number of features. A set of features and the corresponding outputs $y \in \mathbb{R}$ should be preliminary collected. Thus, an algorithm will map the input values to output as a function $f : \mathbb{R}^n \to \mathbb{R}$. The function f is evaluated based on the sum of weighted inputs and bias factors.

The most common MLP is a three-layer neural network that uses different layers for processing information sequentially. These layers are an input layer, a hidden layer, and an output layer, which are schematically represented in Fig. 3.1. Each hidden unit approximates an input layer to the output layer using the activation function ($a(\cdot)$):

$$h_j = a(\sum_{i=1}^{n} x_i w_{ji} + b_j) \tag{3.1}$$

where h_j is the output of jth hidden unit, n is a number of inputs, w_{ji} is a weight (connection link) for ith neuron, and b_j is a bias.

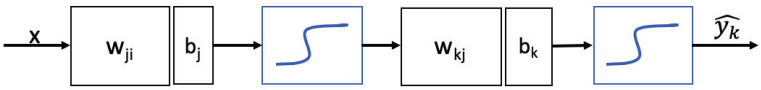


Figure 3.1: Schematical representation of three-layer feed-forward backpropagation MLP

The approximated output is calculated by using the output from equation 3.1 as an input:

$$\hat{y_k} = f(\sum_{j=1}^{N} h_j w_{kj} + b_k) \tag{3.2}$$

where $\hat{y_k}$ is an approximated value of the kth output unit, N is a number of neurons in a hidden layer, and b_k is a bias.

Optimization of weights needs to be performed to minimize the difference (e) between the observed and approximated outcomes:

$$e = \underset{w_{ji},b_j,w_{kj},b_k}{\arg\min} \ (\frac{1}{2}\sum_{k=1}^{m}(y_k - \hat{y_k})^2) \tag{3.3}$$

where y_k refers to the observed outcome, $\hat{y_k}$ is the approximated outcome, and m is the number of outcomes.

3.2.2 Decision Tree Regressor

A decision tree is also one of the machine learning techniques used in this thesis. Typically, this method is used for classification tasks, but there is a possibility of applying it for a regression task. Opposite to utilizing an artificial neural network as a black-box, a decision tree method is an open and easy to understand method.

For a given training vector $\mathbf{x} \in \mathbb{R}^n$ (where n is a number of features) and a training label $\mathbf{y} \in \mathbb{R}^l$ ($i = 1, 2, ..., l$ where l is a number of labels) the regression tree algorithm recursively partitions the features domain into smaller regions (separate classes). It is important to choose correct metrics for best data split and determining when a tree node should become a terminal.

Since, in this work, a decision tree algorithm is used for a regression task, the target constitutes a continuous value. For a node m, which represents a region R_m with N_m observations, Mean Squared Error (MSE) or Mean Absolute Error (MAE) are possible regression criteria to minimize impurity function $H(\cdot)$ as for determining locations for future data splits. Minimization of an error can be done by using mean values of the terminal nodes for MSE [87]:

$$H(X_m) = \frac{1}{N_m} \sum_{i=1}^{N_m} (y_i - \tilde{y}_m)^2 \tag{3.4}$$

or by using MAE:

$$H(X_m) = \frac{1}{N_m} \sum_{i=1}^{N_m} |y_i - \tilde{y}_m| \tag{3.5}$$

where X_m is training data in node m, and $\tilde{y}_m$ is a mean of all targets in node m.

However, when it comes to the analysis of the big amount of data, this method has issues with scalability, stability, and robustness [11, 66]. Another issue that should be addressed is an increase of the complexity when large data samples are used. The total number of nodes, total number of leaves, tree depth, and the number of attributes are hyperparameters that can be controlled in order to minimize the complexity of the decision tree [66]. Since these issues not always can be addressed, ensembles of decision trees are used instead, which are more robust.

3.2.3 Gradient Boosting Regressor

Gradient boosting regression (GBR) machine learning method can be described as an ensemble of decision trees (see Fig. 3.2). Instead of building one tree, this method predicts the desired outcome based on the additive regression model that uses decision trees as weak learners [121]. Sequential fitting of a parameterized function (base learner) to current "pseudo"-residuals is done at each iteration by optimizing regression loss (e.g., least squares, absolute error) [43]. Friedman [43] describes "pseudo" residuals as minimization of the gradient of a loss function with respect to values of the regression model at each training data point for the current step.

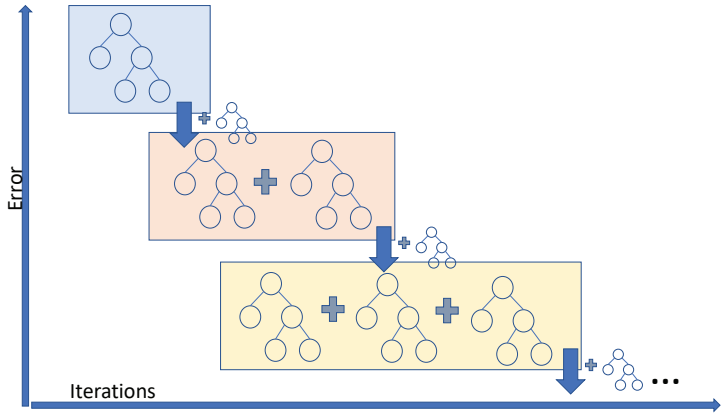


Figure 3.2: Schematical representation of gradient boosting regression in regards to algorithm iterations

The introduction of randomization in the process of training data set selection allows to improve accuracy and reduce the possibility for overfitting. This way of compiling a decision tree allows minimizing the errors at each following step, and therefore boosting regressor is considered as a more reliable and robust method comparing to the classic decision tree regressor.

3.2.4 Adaptive Boost Regressor

AdaBoost Regressor (ABR, short for Adaptive Boost) is another ensemble machine learning method. It works similarly to the Gradient Boost regressor, with the only difference being the way weak learners are created at each iteration. AdaBoost changes the sample distribution at each iteration by varying the weights of each feature (the ones with the biggest error will have the highest weights).

3.2.5 Random Forest

Random Forest (RF) is another ensemble machine learning method. Similarly to GBR and ABR, it also uses a decision tree as the elementary model. However, in the random forest, successive trees are chosen independently from the previous results (opposite to AdaBoost and Gradient Boosting regressors where results of previous trees are used for constructing the new ones), and each tree is constructed by using bootstrapped[1] data samples

[1]Bootstrap sample is a smaller sample that is resampled from the larger data set.

[71]. At the last step, a prediction is made based on the simple majority vote. Randomness in the random forest comes from two levels; datasets are randomly resampled into smaller subsets used to train decision trees, and a number of features are randomly chosen at node splitting step.

The random forest method is not sensitive to either the number of trees in a forest or to the number of features at each node, and it is robust against overfitting [71]. The main steps in the random forest method are the following[71]:

1. Generating n - bootstrapped samples from a training dataset (drawn randomly and can share some of the data points).

2. Generating a forest by training regression trees for each bootstrap sample for a randomly chosen number of features for node splitting.

3. The average value of obtained predictions from each tree is provided as the overall output of the forest.

4. A part of training data, which has not been used for training, is used to evaluate predictors' performance, typically called an out-of-bag dataset. The Mean Squared Error is used as a metric for performance evaluation.

3.3 Model optimization and evaluation

3.3.1 Model optimization with Grid Search

Every machine learning algorithm consists of two types of model parameters; model parameters and hyperparameters. While model parameters are tuned while training a model, hyperparameters need to be passed by an ML engineer. Since there is a large number of possible combinations of hyperparameters, trying each of them manually is a time-consuming process. Therefore, a grid search Scikit-learn module can be used to address this issue.

Grid search entails training the chosen ML model for every combination of hyperparameters from the given set. This process also includes the cross-validation step and accuracy evaluation of regression estimators.

An example of a code listing of the grid search for multi-layer perceptron is presented below:

```
from sklearn.neural_network import MLPRegressor
from sklearn.model_selection import GridSearchCV
parameters = {'hidden_layer_sizes':[19, 22, 25, 27, 35, 52],
       'activation':('relu', 'logistic'),
       'solver': ('lbfgs', 'sgd')}
model = MLPRegressor()
clf = GridSearchCV(model, parameters, cv=5)
clf.fit(X_train, y_train)
```

3.3.2 Model evaluation with 5-fold cross-validation

The five-fold cross-validation is a process when training data is randomly split into k folds, in our case $k = 5$, and then the model is trained on the $k - 1$ folds, while one is left for testing (an example is illustrated on Fig. 3.3). This procedure is repeated k times. The last step of cross-validation is the final evaluation of the model performance, which uses testing data (separated from the original dataset before starting cross-validation) and evaluates the accuracy of the model.

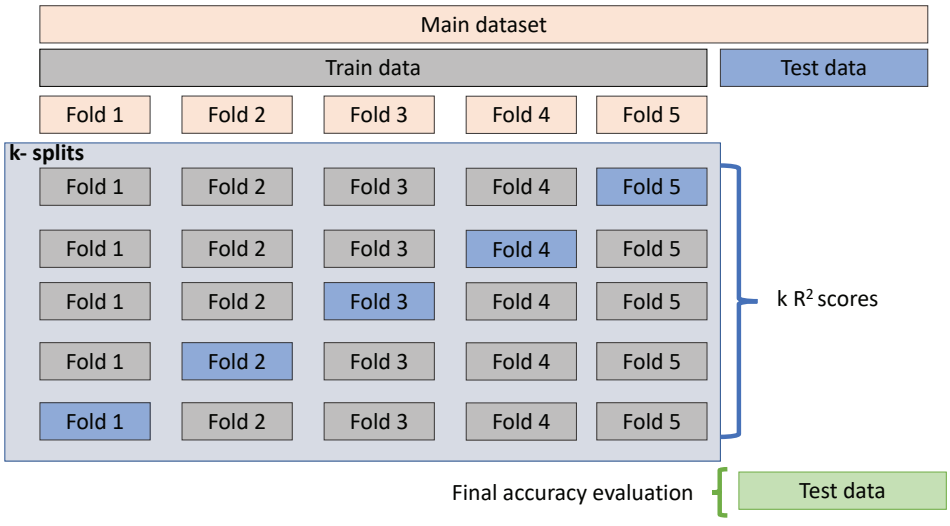


Figure 3.3: Description of 5-fold cross-validation

Chapter 4

Methodology

This chapter describes the research philosophy, design, and methods used to define the problem, design a system, collect and analyze data. Since data is an essential resource for this work, the process of collecting data is underlined in a more detailed way so that anyone could repeat the experimental work.

4.1 Research Philosophy

Research design is always shaped by a discipline area of the researcher, his/her beliefs, academic environment, and the researcher's previous experience [28]. According to Creswell [28], there are four main philosophical paradigms that describe "a basic set of beliefs that guide action", which are positivism, pragmatism, realism, and interpretivism.

While realism and interpretivism are research philosophies mostly related to qualitative studies, this work is about quantitative research design. Therefore, a research design is mainly shaped by the combination of the other two philosophical paradigms.

Typically, additive manufacturing is considered as a technology where a set of variables influence the outcome. For example, Goodridge and Ziegelmeier [49] have investigated how laser power, scan speed, hatch distance, layer thickness influence the mechanical properties of fabricated parts. Another example is an investigation of how scan spacing and direction, laser power, bed temperature, hatch length, and contouring influence a geometric accur-

acy of 3D printed parts [21, 33, 97, 98, 101, 119]. An approach presented in these reports is based on the experimental studies, where a small dataset is used to evaluate the hypothesis where investigated parameters have an impact on the outcome. Such an approach is known as a positivism (or postpositivism) philosophical paradigm. The researchers determins causes that influence the outcome, usually via experiments, by reducing a problem into small discrete tasks. Therefore, one of the main tasks in positivism is empirical observations and measurements that are used to verify a theory.

Similarly, the first experiment in this work has been designed by testing a hypothesis that if one controls material properties, uses the same build layout design and process parameters, the quality of AM produced parts will be similar. This hypothesis arose from a study reported by [91], where researchers have controlled the material and received good repeatability between the runs that is usually a challenge for additive manufacturing processes.

After this hypothesis has been proven to be true, the next challenge that was of interest has come from the observations from other published works. Often a limited number of samples is produced for a better understanding of AM processes. However, in the real-world, packing the whole build chamber with many parts is of interest due to the reduction of the costs. Therefore, the first research question has been formulated as a result of the first observations. The **RQ1** is about investigation how changes in build layout design can influence the quality of AM produced parts.

While the first part of experiments was inspired by the positivism philosophical paradigm, the second set of experiments is shaped by the pragmatism paradigm. Even more, when it comes to methods used for data analysis, the previously used methods like linear regression and ANOVA are not suitable for this research problem [14]. Thus, all available research methods need to be applied to understand the problem and derive knowledge about the problem (as in pragmatism philosophical paradigm) [28].

Since it is not possible to describe the impact of changes in build layout design on the quality of AM produced parts as a linear function, more sophisticated methods are required. Additionally, the complexity of parameter sets that have an impact on the quality of AM parts is emphasized in [48, 60, 93, 94, 95]. Typically, there is a set of unknown factors that influence quality. Therefore, methods capable of dealing with these factors are

of interest. Applying machine learning techniques that can identify different patterns in the data is one of the solutions, which is widely used in computer science, computer vision, and traditional manufacturing [70, 75, 82, 96].

Therefore, coming back to the positivism paradigm, another two research questions have been formulated. **RQ2** states that it is of interest to investigate how machine learning methods can help to understand polymer powder bed fusion process and mathematically describe its behavior with a focus on geometric deviations. The **RQ3** raises a similar question but with application to the mechanical properties.

So, as can be seen, there is a synergy of positivism and pragmatism philosophical paradigms that influenced the way research has been designed. Additive manufacturing is still a new technology is many unknown factors that influence an outcome. Therefore, the researcher needs to be flexible and react based on the results at each stage of the research path.

The design of experiments in such a way has shown an interesting observation, which was outside the scope of this study. Typically, it is assumed that fabricating parts in the center of the build chamber will always lead to the best quality of a part. However, results obtained in this research study have shown that this assumption can be questioned because strong mechanical properties have been observed not only in the center part of the build chamber, and more factors influence the mechanical properties.

Therefore, an understanding of how to pack parts in the build chamber before producing them can allow making AM even more sustainable by reducing material waste, energy consumption, and obtaining desired quality. These challenges are formulated as **RQ4**.

4.2 Limitations of the work

By the reason that this research has a limited timeframe, there is a need to set boundaries for the research.

This study is limited to the polymer powder bed fusion AM process with the main material PA 2200. The only one machine will be used in this work, which is located at NTNU Gjovik campus. Testing results on other machines or other types of AM process is out of the scope of this work. Similarly, polymer material, which is used in this laboratory, is investigated in this study, and other types of polymers, composites, or metals are out of

the scope.

Even though additive manufacturing is known for its flexibility of producing complex shapes parts, this work is limited only to the investigation of one type of specimen used for mechanical testing. First of all, additive manufacturing is still considered an expensive manufacturing process. Therefore, data collection for the investigation of both geometry and mechanical properties needs to be combined. This was the first reason why specimens for mechanical testing is chosen as the main object for study. The second reason is that there is no systematic approach or standard for the evaluation of geometric deviations and mechanical properties of complex shapes. Therefore, complex designs are out of the scope of this work.

The predictive models developed in this study are not final models ready to be used in manufacturing. These models are considered as prototypes that allow establishing a systematic process of data collection, preprocessing and analysis, identifying which and how machine learning techniques can be used.

4.3 Research methods

As was already described, this research is quantitative and is based on experimental work. The knowledge is derived from the results of experimental work. Therefore, the research methods are chosen with a correspondence to the research questions and philosophical paradigms that influenced the research design. The design of experiment, 3D printing, measurements of dimensions, conditioning, mechanical testing, and data analysis are considered as main research activities in this work. Fig.4.1 illustrates which hardware and software have been used in each activity.

However, the list of research methods used in this work is even longer. Since, in addition to listed activities, it also includes a literature review, application of system engineering, and process of validation and verification of the results. Each of the research methods is described in the details below.

4.3.1 Literature review

As can be seen in Figure 4.2, a literature review has been performed throughout the whole Ph.D. study. The main motivation behind this decision was to follow the current state-of-the-art and being able to go deeper into the

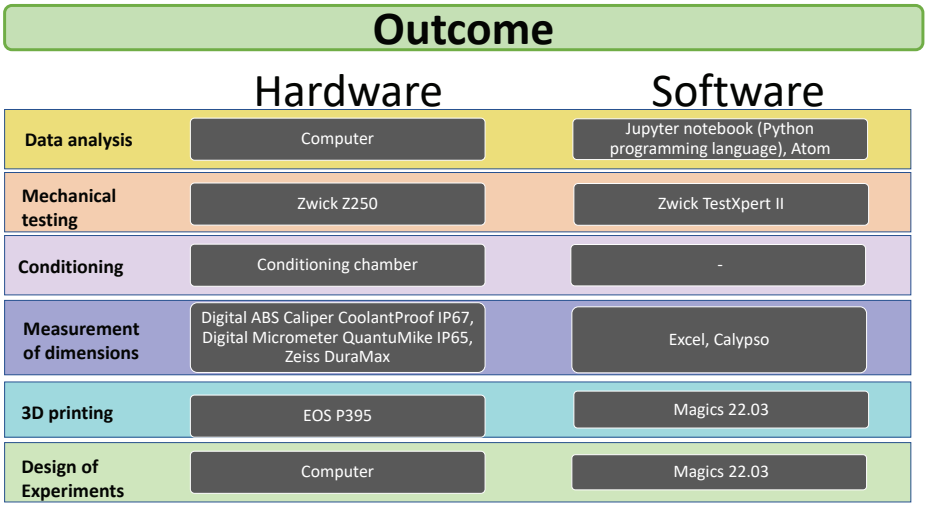


Figure 4.1: Correspondence of the used hardware and software with respect to the research activities

topic. Since researchers started to pay more attention to AM and the application of machine learning during the past couple of years, the number of published articles has been increased exponentially. Therefore, continuously following the developments and new findings were necessary and beneficial for this research.

According to the requirements and needs of the project, the first stage of the literature review was performed in the first year of Ph.D. study and was focused on the material models for FEA of polymer powder bed fusion systems. However, one of the main outcomes of this literature review was an understanding that material models for PA 12 include many simplifications, and cannot be used for finite element analysis. Therefore, a more extensive literature review was needed for proposing another solution on how geometric accuracy and mechanical properties can be analyzed with a larger number of parameters.

Thus, the literature review conducted throughout the whole Ph.D. study was organized in the following way. The main keywords that have been used in a search for literature can be divided into several groups: 'material', 'process name', 'type of modeling', and 'other'.

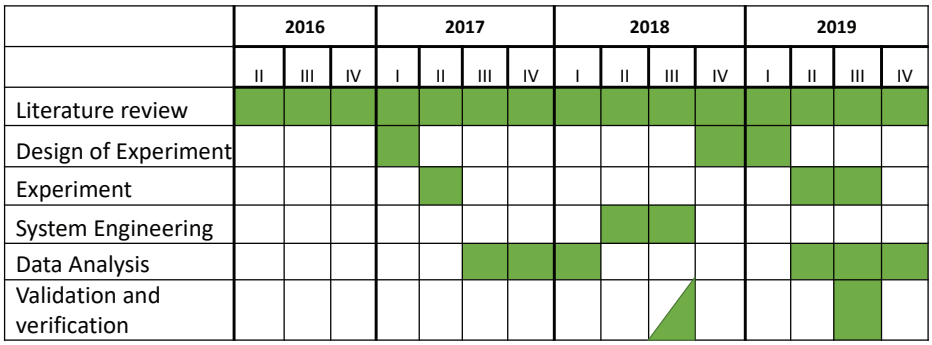


Figure 4.2: Grant chart about PhD research path

Thus, the 'material' group consisted of the following terms: *polymer, polyamide 12, PA12, nylon, nylon 12, nylon 6, polyamide 6, PA6, PA11, plastic, plastic powder, polymer powder, polymers.*

The 'process name' group included such terms:*additive manufacturing, additive manufacturing technology, AM, AM process, SLS, selective laser sintering, rapid prototyping, layered manufacturing, 3D printing, 3D-printing, 3d printing, 3D printer, 3D-printer, 3D-printed, powder bed fusion, powder bed, powder bed fusion system, polymer powder bed fusion system, additively-manufactured, SLM, selective laser melting.*

The 'type of modeling' group in the first year of Ph.D. study (second and third quarters on Fig. 4.2) was mainly focused on the terms like *mathematical modeling, material model, material models, finite element analysis, FEA, finite element method, FEM, Taguchi method, ANOVA, analysis of variance, design of experiments, S/N analysis.* From the fourth quarter of 2016 year (see Fig. 4.2), the keywords of this group have been changed to *regression modeling, design of experiments, DOE, data analysis, intelligent data analysis, machine learning, artificial intelligence, artificial neural network, ANN, feature selection, intelligent methods, genetic algorithm, GA, support vector machines, SVM, Bayesian methods.*

The 'other' keyword group was expending every year based on the topic of interest. For example, when a more in-depth understanding of STL file was

needed, *stl model, STL models, CAD model, STL model features, STL model properties, number of mesh triangles, mesh triangles, contouring, contouring failures, triangular mesh* have been used in combination with other keyword groups. Similarly, any search for literature (including patents) have been performed by mixing keywords from different groups.

4.3.2 Design of experiments

Based on the first results of the literature review, the **RQ1** has been defined, and the first experiment has been designed. The sequential design of experiment[1] was used in this work due to two main reasons. First of all, practical experiments in additive manufacturing usually have been limited to a small number of samples, which have also led to the limitations of understanding what can be expected in the results when a larger number of samples are produced in one run. Second of all, the results of each experimental work may influence the design of the next one, and therefore sequential DOE is the best choice in this research study.

4.3.3 Experimental work

Experimental work was performed in two stages. The first one was executed in the 2017 year, and the second stage was conducted in the 2019 year. As Fig. 4.2 shows, an extensive literature review has been done before the first experiment was designed. This experiment was designed by using Rüsenberg et al. [91] article as a reference work because it could help to get a better understanding of how control of material properties can influence the resulting quality.

After the first experiment, collected data was analyzed based on the defined **RQ1**, and different modeling techniques were tested and evaluated. The analysis of the collected data stimulated the definition of the **RQ2** and **RQ3**, and the design of experimental work in 2019 has also been done with taking into account observed weak sides of the first experiment. More details about experimental work are described in Section 4.4.

4.3.4 Model-based system engineering

System engineering is a field of study that provides different techniques to design complex systems. In this work, a combination of the experiments and analysis of the collected data is considered as a complex system. The term

[1]The second experiment builds on the results from the first experiment [28]

"system" in this work is regarded as "any organized assembly of resources and procedures united and regulated by interaction or interdependence to accomplish a set of specific functions"[72]. Since any system consists of components and interrelations between the components, the understanding of how the collected data can be used efficiently for the analysis of dimensions and mechanical properties is needed. Therefore, model-based system engineering is applied in this work to define a roadmap for data analysis and to determine a systematic way of presenting results. The proposed roadmap is represented as an intelligent system for quality assurance in additive manufacturing, and more details on how it was developed are presented in a separate Chapter 5.

4.4 Experimental work

In this work, each experiment consists of the specific steps that need to be done before the experiment can be executed. Similarly to any additive manufacturing process, a polymer powder bed fusion system requires a 3-dimensional (3D) model of an object to be produced. Therefore, computer-aided design (CAD) is converted into an STL file format that is readable by AM machine. STL models are shallow triangular meshes that are first placed in the build chamber of a machine and then sliced into layers. A build layout design can be defined as a set of STL models placed in the same build chamber by following specific rules for the chosen material and AM machine.

The main steps conducted during all experimental work are illustrated on Fig.4.3 and Fig.4.4. Preparation of the material mix (mix of virgin and used powder with defined ratio - typically recommended by AM machine producer as 50/50 % ratio) and designing a build layout are two first steps that should be done before the experiment can be conducted. Then based on the material properties, AM machine process parameters are selected with a consideration that working chamber temperature is dependent on the powder melting point. After parameters are defined, the fabrication job is executed, and the build chamber needs to be cooled down at least for the same time as the printing process has taken place. The next step is getting produced part out from the build cake and cleaning parts from the powder.

While in metal additive manufacturing, typically, the next step would be post-processing of the parts; in polymer additive manufacturing, one can usually skip this step. Therefore, the next stage is measuring the dimensions

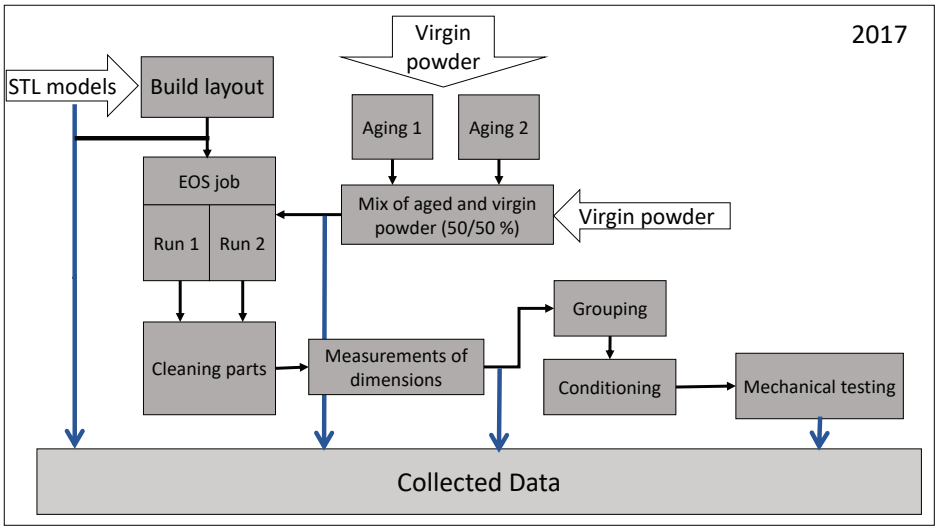


Figure 4.3: Experiments 2017

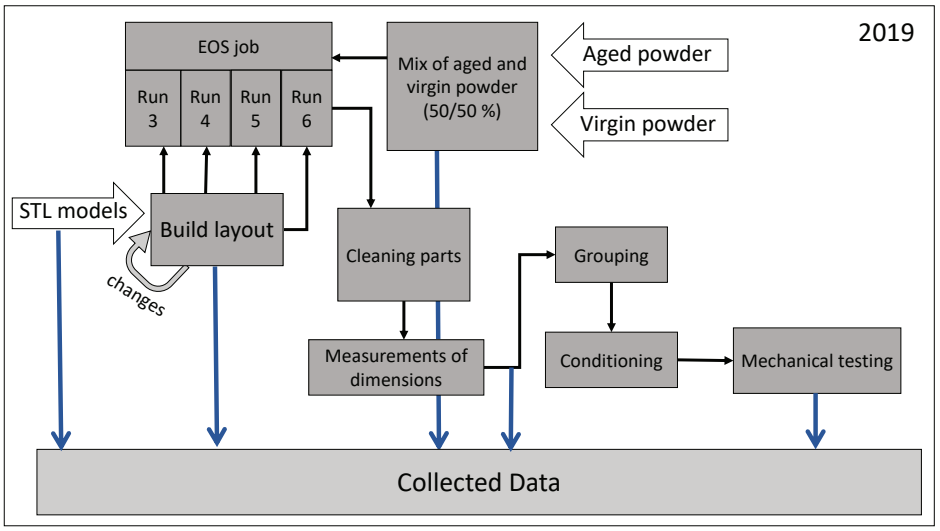


Figure 4.4: Experiments 2019

Table 4.1: Material and process parameters used in the experiments

Parameters	Value
Virgin/aged PA2200 powder ratio, %	50/50
EOS P395 system settings	Balance
AM system warm up time, min	120
AM system cooling down time, min	240
Working chamber temperature, °C	180.5
Removal chamber temperature, °C	130.0

of produced parts.

Since in this work, part quality is considered as both geometric (dimensional) accuracy and mechanical properties, destructive testing - tensile test - is executed at the end of the experiment. Following the DIN EN ISO 527-1 standard, produced samples have been conditioned first, and only then the tensile test has been performed. Additionally, it is important to mention that experimental work conducted in 2017 is different from the experimental work performed in the 2019 year due to changes in a build layout design and material properties.

As can be seen in Fig.4.3, only two runs were produced at the first stage of experimental work. While at the second stage of the experimental work, which is illustrated in Fig.4.4, four runs were made, and three out of four runs had changes in the build layout design.

4.4.1 EOS P395 polymer powder bed fusion system

An EOS P395 polymer powder bed fusion system has been used in the experiments performed to collect data. Machine process parameters have been used the same for all six runs, and more details about known material and process parameters are depicted in Table 4.1.

PA 2200 was used in all runs with virgin/aged powder ratio of 50/50 %. While in the first two runs, the material was kept the same, starting from Run 3, polymer powder has been used with the same virgin/aged ratio but with different material properties.

In order to be able to control material properties and keep them constant in the first two runs, an idea of controlling material properties via self-aging

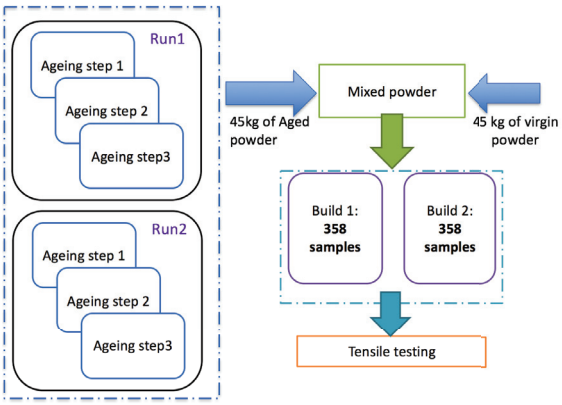


Figure 4.5: Schematic representation of the polyamide powder self-aging steps

powder was presented by Rüsenberg et al. [91] and applied in this work. The process of self-aging PA12 is shown in Fig. 4.5. The powder was self-aged through 3 cycles of running EOS P395 without laser deposition.

4.4.2 Build layout design

Build layout design for runs 1-3

Even though Rüsenberg et al. [91] was used as a reference for the design of the first and second runs, strategy for placement and orientation of specimens was chosen to be different. I assumed that the build layout should be designed similarly to real manufacturing conditions.

Based on this assumption, the maximum number of parts is chosen to be the main criterion for the design of the build layout. It means that the parts are placed as close to each other as possible, and the minimum distance between the specimens is set to *5 mm* based on the recommendations from Magics 20.0 (and Magics 22.03) software. Additional attention was paid to the specimens placed in the same orientation for verification and validation of the results. In other words, more than five specimens in the same orientation were placed close to each other for better control of potential variations related to the position of the specimen.

Description of the build layout for runs 1-3 is combined because these runs have used the same layout. So, in total, 358 specimens were produced in one run (or 1074 specimens for the three runs combined). However, as a part

of this Ph.D., data have been collected only from 217 (or 651 in total for runs 1-3) specimens of type ISO 527-2 1BA for mechanical testing. This is done for a possibility of comparison results obtained for geometric accuracy and mechanical properties, and evaluating whether the same dataset can be used for modeling different quality metrics.

An example of a specimen with a dimensional description is shown in Fig.4.6. A final build layout design (filtered to only investigated specimens) used for these three runs is illustrated in Fig.4.7(a).

All investigated specimens have been placed in four different orientations (see Fig. 4.6), and names of the orientations have been defined according to the ISO/ASTM 52921:2013 [63] standard:

- Group 1. XYZ -oriented parts (XY on Fig.4.6)
- Group 2. XZY -oriented parts (XZ on Fig.4.6)
- Group 3. ZYX -oriented parts (Z on Fig.4.6)
- Group 4. Angle-oriented parts (Angle on Fig.4.6)

The Angle-oriented specimens are parts (the word ”parts” is and will be used as a synonym) oriented at 45° between x and z axes.

Since the requirement to fit as many specimens as possible has been defined during the design of experiment, the number of specimens in each orientation differs. Thus, each run consisted of 65 parts placed in XYZ orientation, 24 parts in XZY orientation, 84 parts in ZYX orientation, and 44 in Angle orientation.

To identify parts and be able to connect the dimensional measurements, tensile testing results with part's position, every specimen has its label, which is placed on two sides of the part. This led to variations in the STL model properties, which are number of mesh triangles, surface, and volume, for each part within the build layout. Although, it is important to mention that there is no variation in STL model properties between Run 1, Run 2, and Run 3 due to the usage of the same build layout.

In addition to named characteristics of a build layout design, platform volume utilization, current nesting density, build height, and coordinates

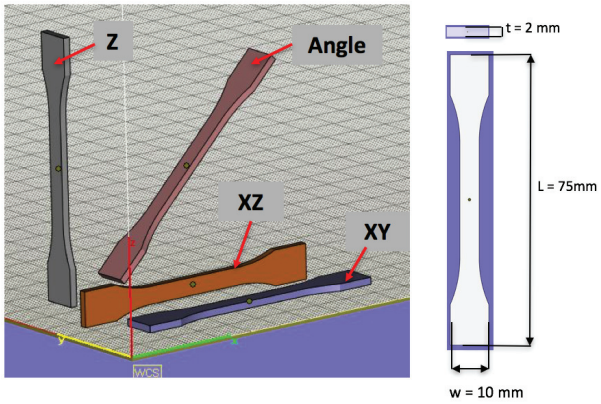


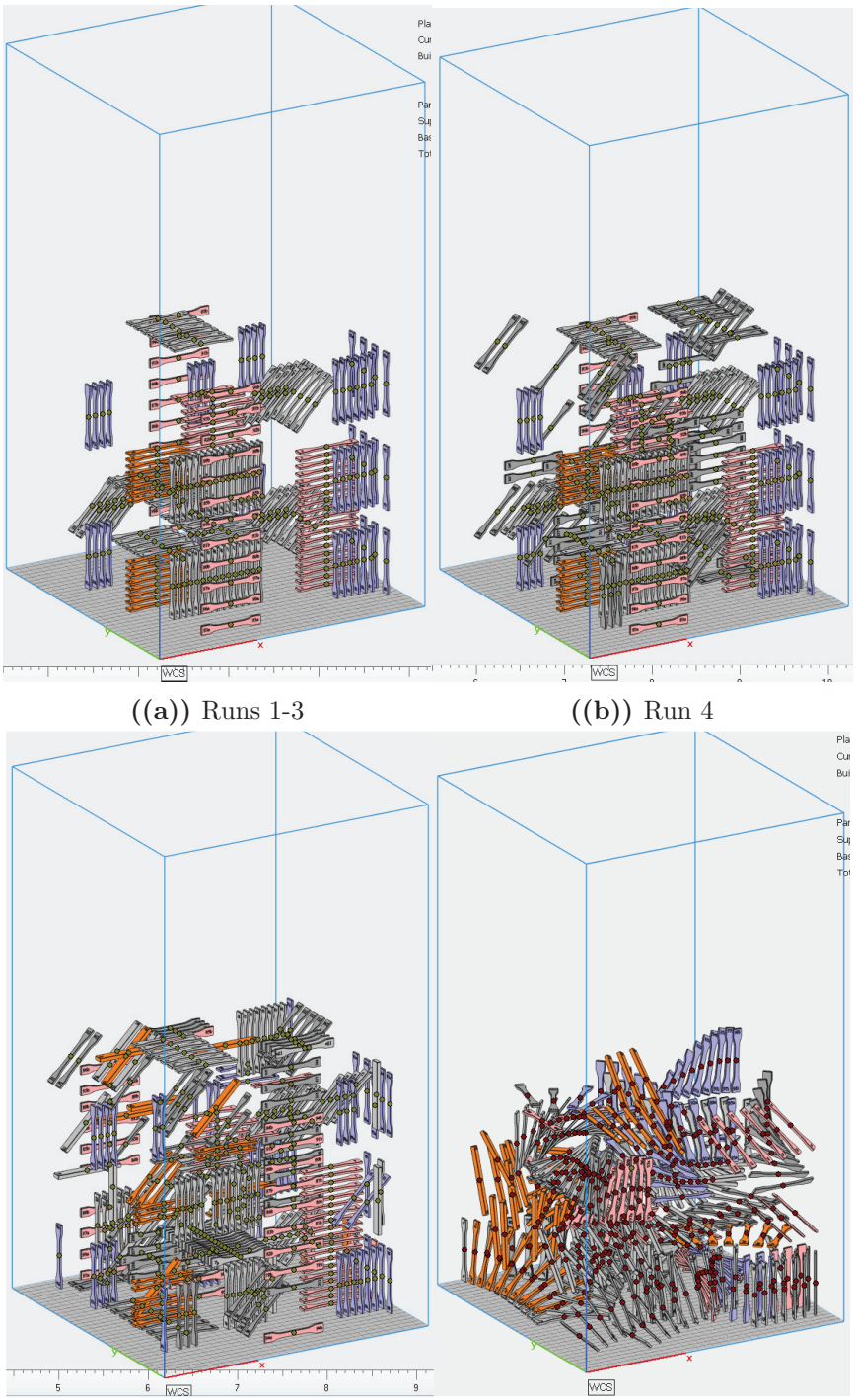
Figure 4.6: (Schematic visualization of parts' orientation and dimensional features (where t – thickness, w – width and L – length)

that describe the part position in a build chamber will also be considered as important features in this work. The resulting build height for this build layout is 314.75 mm. Platform volume utilization is 0.75%, and current nesting density is 1.48%.

Build layout design for Run 4

In a build layout for Run 4, the majority of the parts, which are not used for analysis, have been taken out of the build layout design, and on their place, new 75 specimens of type ISO 527-2 1BA for mechanical testing have been fitted in build layout. Among these new 75 specimens, 14 are in XYZ orientation, 26 in XZY orientation, 3 in ZYX orientation, and 32 in Angle orientation. The main focus was set to produce more specimens in XZY and Angle orientation because, in previous build layout, their number was significantly smaller than for the other two orientations. Thus, the number of specimens in the new build layout in each orientation group has increased to 79 (XYZ - 65 previously), 50 (XZY-24), 87 (ZYX-84), and 72 (Angle-44). The final build layout for Run 4 is illustrated on Fig.4.7(b).

The resulting build height is also 314.75 mm. Platform volume utilization is 0.76%, and current nesting density is 1.51%.



((a)) Runs 1-3 ((b)) Run 4

((c)) Run 5 ((d)) Run 6

Figure 4.7: Build layout design in Magics 22.03: home view

Build layout design for Run 5

A redesign of the build layout for Run 4 into a build layout for Run 5 has been conducted in three main steps. The first step was focused on:

- changing the orientation of specimens in XYZ orientation into XZY, ZYX, and Angle orientations but keeping the same positions;
- changing the orientation of specimens in XZY into XYZ, ZYX, and Angle orientations but keeping the same positions;
- changing the orientation of specimens in ZYX into XYZ, XZY, and Angle orientations but keeping the same positions;
- changing the orientation of specimens in Angle into XYZ, XZY, and ZYX orientations but keeping the same positions.

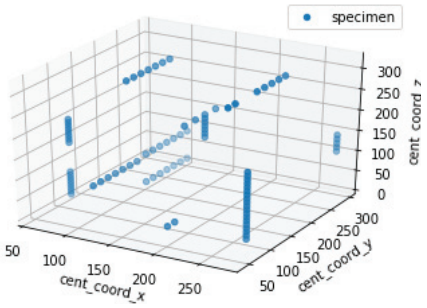
However, by the reason that temperature distribution is dependent on the build layout design and nesting density of the build chamber [13, 19], a limited number of specimens can be rotated and moved. Thus, not all of the listed conditions were possible to fulfill, especially for such cases when orientation needs to be changed from XYZ/XZY to Angle/ZYX orientation, or another way around.

Therefore, the next step was to change the position of specimens that have not been rotated. This step is mandatory due to overlaps created by rotating specimens, and because of the need for having a 7 mm distance between specimens, so they do not sinter together.

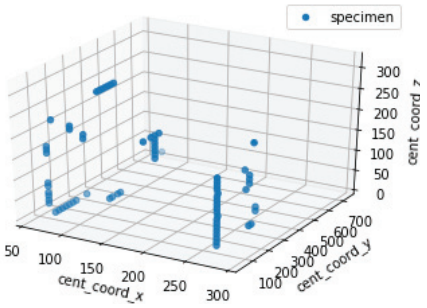
In the third step, the position of rotated specimens that are too close to other specimens was changed. In other words, specimens, which interfere with neighboring specimens (had less than 7 mm distance in x, y, or z axes), need to be moved to an empty area in a build chamber.

The comparison of orientation groups for Run 4 and Run 5 is shown in Fig.4.8.

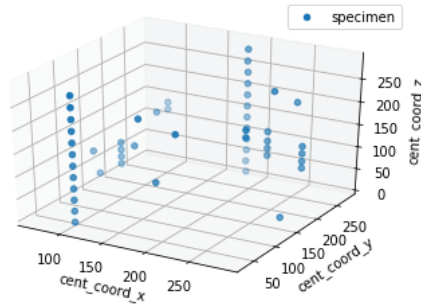
The resulting build is shown on Fig.4.7(c), and its height is 314.75 mm. Platform volume utilization is smaller than for the other runs and is equal to 0.56%, and current nesting density is 1.10%.



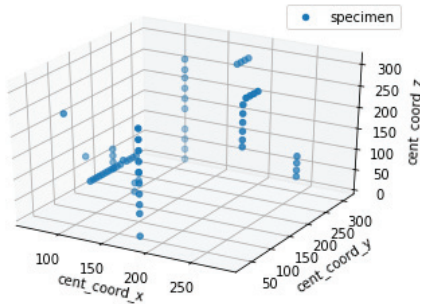
((a)) XYZ for Run 4



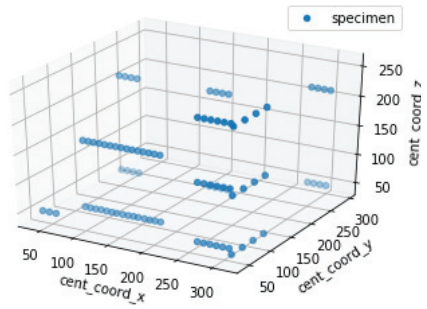
((b)) XYZ for Run 5



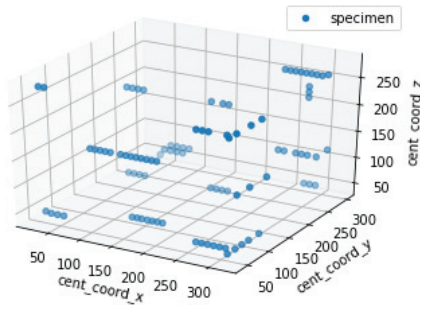
((c)) XZY for Run 4



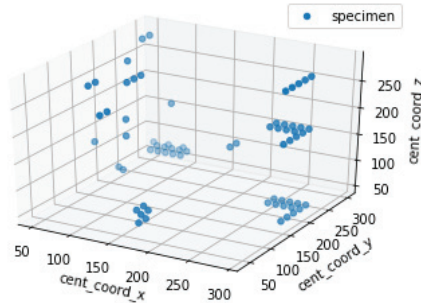
((d)) XZY for Run 5



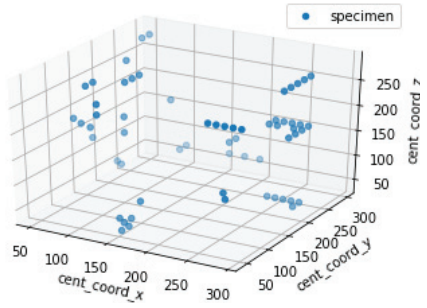
((e)) ZYX for Run 4



((f)) ZYX for Run 5



((g)) Angle for Run 4



((h)) angle for Run 5

Figure 4.8: Orientation-based comparison of changes conducted on the build layout Run 4 to create a new build layout for Run 5

Build layout design for Run 6

A build layout design for Run 6, shown in Fig. 4.7(d), was done by Torbjorn Leirmo (Ph.D. candidate in Additive Manufacturing at NTNU) by utilizing the same specimens produced in runs 4-5. However, dimensional measurements and mechanical testing were done by Ivanna Baturynska.

The main goal for this run was set to use a build layout designed by another person because it will help to evaluate how sensitive developed models could be for an external user. Besides, using data collected from this run will also be beneficial for models' generalization in the future, while weaknesses could be identified for further improvements.

This build layout design consists of 8 different families. Each family has a number of groups of specimens rotated around y or z axes by 5- and 10 -degrees intervals. The starting positions for different groups are XYZ and YXZ orientations. The distance between specimens in all axes was set to 7 mm by following recommendations from Materialize software and EOS (producer of used AM machine system).

The resulting build height was reduced to 299.45 mm, which is less than in other build layouts. Platform volume utilization is the same as for runs 1-3 and is equal to 0.75%, and current nesting density is 1.56%.

4.4.3 Measurement of dimensions

In the first experimental work, data was collected from two identical runs and was used to evaluate the dimensional accuracy of the produced specimens. Length value was measured using a Digital ABS Caliper Coolant-Proof IP67 with an accuracy of $\pm 0.02mm$. Width and thickness were measured using a Digital Micrometer QuantuMike IP65 with an accuracy of $\pm 1\mu m$.

In the second stage of experimental work performed in 2019, data was collected with the help of the Zeiss DuraMax coordinate measurement machine (CMM) with an accuracy of $\pm(2.7 + L/250)\mu m$ (where L - is measured value) [2]. A stylus assembled with two probes with a diameter of $0.5mm$ has been used for measurements based on the expert's recommendation from the Zeiss company. Due to the size of measured specimens and weight of the part (light material), a special fixturing was designed to provide stable fixation of the specimen while dimensions are measured, and it is shown in

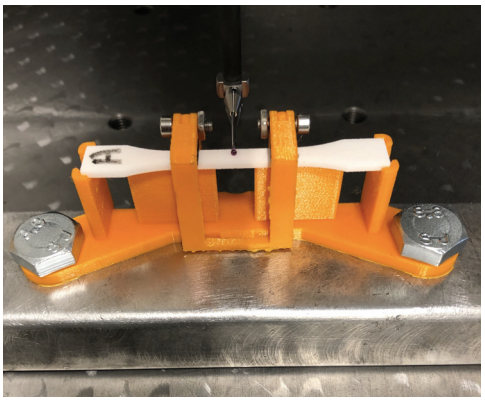


Figure 4.9: Specially designed fixturing for measurements with Zeiss coordinate measurement machine

Fig.4.9.

In addition, the final value of each dimensional feature (see Fig. 4.6) was calculated as a mean of three repeated measurements to minimize a measurement error.

While measurements of width and thickness in the first two runs have been conducted without registration of a measured side, the width and thickness from runs 3-6 have been measured in three different locations, as it is shown in Fig.4.10. This change in the locations was based on the observation from data analysis of runs 1-2, and discussions with an expert from EOS company. The expert has described that due to temperature distribution in the powder bed fusion system, variations in the z-axis are complicated and non-linear. Therefore, it is of interest to look at the geometric variations along a part, especially those that are printed in ZYX and Angle orientations, and measuring specimens in three locations (Side A, side B, and side C) will allow conducting this analysis.

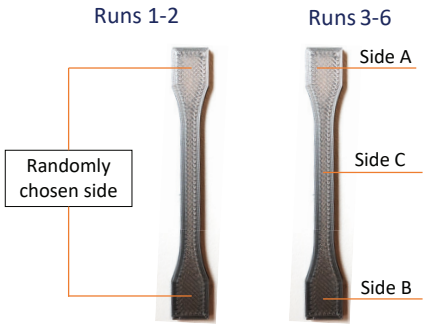


Figure 4.10: Description of differences between measurement areas from runs 1-2 and runs 3-6

4.4.4 Mechanical testing

Tensile testing was performed according to ISO 527-1 [61]. PA12 absorbs some moisture from the atmosphere, and this affects the mechanical properties. Hence, the specimens were conditioned to a moisture content corresponding to saturation at the standard condition 50% relative humidity and 23 °. This conditioning was implemented by storing the specimens for 7 days in a climate chamber at 62% relative humidity and 70 °, which is accelerated conditioning that results in the same moisture content as the standard condition mentioned above [ISO 1110]. After this, the specimens were kept 1-2 days in a climate chamber at the standard condition before testing. The Zwick Z250 machine was loaded with 2.5 kN cell. The specimens were mounted in wedge grips with grip to grip distance set to 55 mm, and an initial gauge length of the extensometer of 25 mm.

4.5 Data preparation and analysis

Python programming language is used in this work for prototyping and development of the modules, which can be used later as a core of the intelligent system for quality assurance. This programming language has been chosen because it is simple, easy to use, and supports multiple programming paradigms. An extensive choice of standard open-source libraries allows ex-

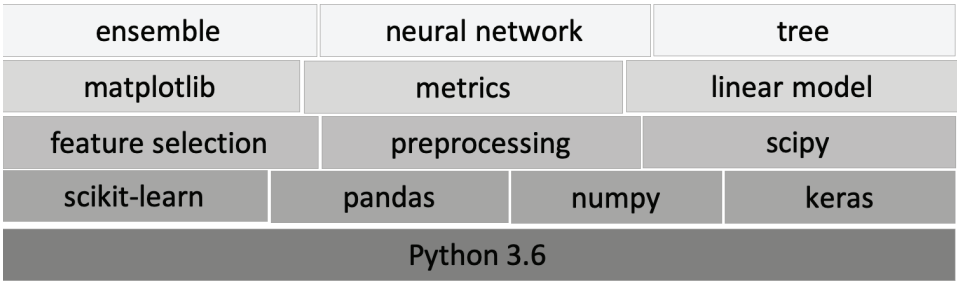


Figure 4.11: Python stack for the codebase of the data anlysis

ecuting complex algorithms in a simple "one-line" code manner. It thus provides a possibility of focusing on the application area rather than programming all algorithms from scratch.

Python stack used for the codebase of the data analysis is presented in Fig.4.11. The latest version of Python 3.6 is used for the data analysis in this work. Analysis of the collected data is conducted in this work through three main steps:

- **Data preprocessing** (data preparation) is a necessary step towards quality assurance of the collected data because raw data is vulnerable to noise, outliers, missing values, or grammatic errors. Transforming data into correct data classes, cleaning data from the outliers, and empty values (like NaN or null), grouping data with respect to necessary analysis are conducted at this stage of the research. Such libraries, like `pandas, scipy, numpy, and scikit-learn`, are used for tasks listed above. The `preprocessing` is a scikit-learn module used for data normalization and data split into training, testing, and validation datasets. Data normalization is needed when, for example, values of different features vary from 1 to 5000, and it will help to avoid overweighting some of the features only due to their large values rather than relevance rank.
- **Data visualization** is performed throughout all process of data analysis. For a better understanding of data and its quality, graphical visualization is used for illustrating different data distributions, and therefore, may also be of use for detecting data outliers. At the stage of data processing, visualization of the model's performance as a func-

tion of accuracy and training epochs have also been used for the evaluation of models. The `matplotlib` library has been used for data visualization in combination with scikit-learn `metrics` module.

- By **data processing**, in this work, is considered a process of models' development. Linear modeling is conducted by using scikit-learn `linear model` module, and more complex machine learning methods are applied with a help of `neural network` , `ensemble (AdaBoost Regressor, Random Forrest` and `Gradient Boosting Regressor`), and `tree (Decision Trees for regression)` modules. While multilayer perceptron (MLP) neural network has been chosen as ML technique in this study since this method is already widely applied in AM, the other ML techniques have been chosen based on the two major requirements. The techniques have to be suitable for the regression task, and they should provide additional information on features (parameters used to map build layout design to quality metrics) importance. The second requirement was introduced in this work because it is of interest which of the investigated parameters are significant for predictive models.

All the abovementioned steps are performed for analysis of both geometric accuracy and mechanical properties, and more details are presented below.

4.5.1 Data cleaning

In this work, several mistakes have been found in the collected data. For instance, feature's names spelled differently (e.g., 'Orientation' and 'orientation' are considered as two different feature names), punctuation signs were in a wrong order, and specimen indices were spelled inconsistently (e.g., 'a' in one table and 'a.' in another table).

In addition to that, in Run 6, one sample was lost during cleaning samples from the powder after the printing process, and 20 samples have not been mechanically tested due to the transportation equipment-related delays. Data corresponding to one lost sample was removed from Run 6 for all data analysis steps, while data corresponding to 20 not tested samples are removed from the feature list used for analysis related to mechanical properties. In total, 1526 samples were used for measuring dimensions, and 1506 samples were mechanically tested. A list of all features and targets are depicted on Tables 4.2 and 4.3.

Table 4.2: Description of the features (also refered to as inputs)

Features/Input	Short name
central coordinate x	cent_coord_x
central coordinate y	cent_coord_y
central coordinate z	cent_coord_z
max coordinate x	max_coord_x
max coordinate y	max_coord_y
max coordinate z	max_coord_z
min coordinate x	min_coord_x
min coordinate y	min_coord_y
min coordinate z	min_coord_z
volume	volume
surface	surface
number of mesh triangles	num_mesh_triang
number of mesh points	num_mesh_triang
orientation angle by x axis	reor_a
orientation angle by y axis	reor_b
orientation angle by z axis	reor_c
build hight	bh
platform volume utilization	pvu
current nesting density	cnd

Table 4.3: Description of the targets (also refered to as outputs)

Target/outcome
length
thickness
width
Tensile modulus
Nominal stress
Elongation at break
Scaling ratio - length
Scaling ratio - thickness
Scaling ratio - width
Placement - cent coord x
Placement - cent coord y
Placement - cent coord z

More details that will help to understand target data are described in more detail in Chapter 6 and Chapter 8, for dimensions and mechanical properties, respectively.

4.5.2 Data integration

Data integration is an important part of any data analysis process. Since, in this study, six different experiments have been performed with different controlled variables at each stage, the run-based data integration process is illustrated in Fig. 4.12.

The data collected from Run 1 and Run 2 are joined into one dataset (Column A) because these two experiments were characterized by the same build layout, material, and machine process parameters.

In order to understand whether the change in material has an impact on feature importance, data collected from Run 3 is joined with Column A into a new dataset described as Column B. The next experiment was conducted with a similar material as Run 3, but additional 75 specimens have been fitted into the same build layout as used in the previous runs. The joined data is presented as Column C.

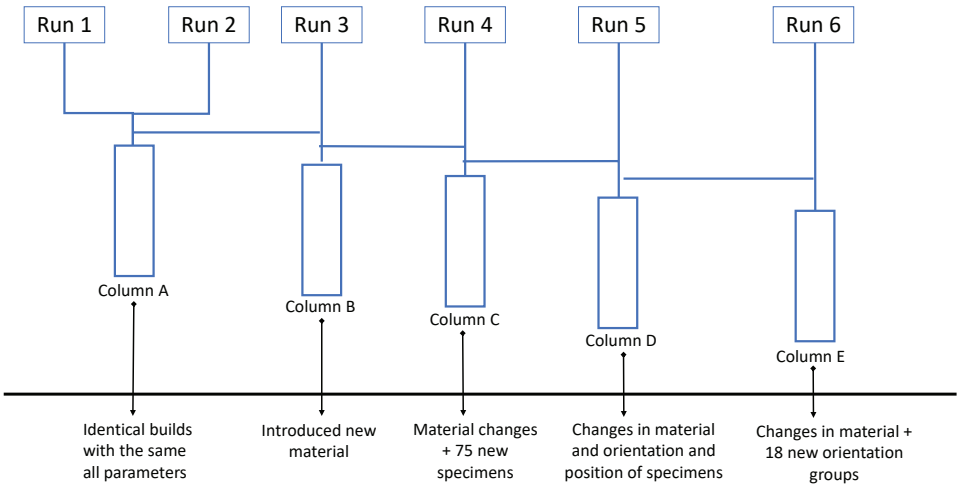


Figure 4.12: Description of data structure used for feature selection and analysis of the results

Column D is a description of datasets that consists of data from runs 1-5, where in the last run a small variation in the material is introduced, but the main difference constitutes changes of orientations and positions of specimens in the build chamber within the same orientation groups as in the previous runs. However, Run 6 introduces an extreme change in the build layout with more than 18 orientation groups being defined and produced.

Besides, files extracted from different machines required additional formating. The integration of data from these files was conducted into two files, one containing features, and the other contains target values. The file that contains features values is the same for both dimensions and mechanical properties, while the target values are saved separately for dimensions and mechanical properties.

4.5.3 Data normalization

The application of machine learning requires the normalization of features in the input data set. ML estimators perform poorly if a feature has variance with orders of magnitude larger compared with other features. Therefore this step is conducted after data integration. Parameters' values have to be scaled to zero mean and unit variance. As such, feature (input) data are

scaled to zero mean and unit variance using StandardScaler (Scikit-learn module), while target values are not changed.

The prepared data is used to develop intelligent predictive models with the help of MLP, Random Forest, Gradient Boosting Regressor, and AdaBoost regressor machine learning techniques. The next step is the validation and verification of the results, which are described below.

4.6 Validation and verification of the results

Since the results of this research are both predictive models and prototypes of software modules for the intelligent system for quality assurance, the developed models and modules need to be validated and verified. The process of the validation of the models is directed towards the predictive models, while the design of an intelligent system, which will incorporate the proposed predictive models as modules, is considered as a verification process.

Validation of the obtained results is performed at different levels. First of all, models developed on the data collected from the first experiments conducted in the 2017 year (Run 1 and Run 2) are evaluated with the help of Run 3. Run 3 consists of a number of specimens that have been rescaled with respect to the prediction of geometric deviations - this step can be called as compensation of deviations.

The next level of the models' validation is set to the integration of the five-fold cross-validation into the process of training models. In such a way, issues with memorizing data by models (overfitting) can be overcome. However, the generalization of the models is still under the question.

Therefore, the third level of models' validation is based on using a specific testing dataset, which is not used in the previous two validation levels, for evaluation of model performance on them. This step can help to show us whether models are sensitive to new unseen data. The prediction accuracy of the proposed models is used as the main metric for performance evaluation and for making a final decision on which model should be used or further developed.

Chapter 5

System engineering as a means towards an intelligent system for quality assurance

This chapter describes how model-based system engineering is used to design an intelligent system for quality assurance in additive manufacturing. The main goal is to understand which components in a system should be developed, which interconnections exist, and how to provide a successful solution at the end of the research.

Typically, four steps illustrated in Fig. 5.1 are used as a guideline for the development of a technical system. However, in this work, only the first step is highlighted, namely *requirement analysis*, since the development of the whole system is a time-consuming task. However, a brief analysis of the stakeholders' needs and description of the functional behavior analysis, architectural synthesis, and validation and verification stages are presented in Appendix B.

In this chapter, the main attention is paid to the description of the components of the proposed system and the corresponding functional requirements.

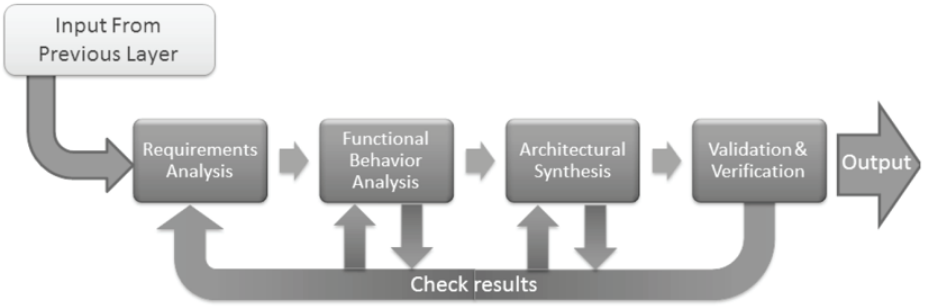


Figure 5.1: Process of working across the domains in SE [72]

5.1 Requirements analysis

Defining requirements is a process that first and foremost involves stakeholders and their needs. Therefore, the requirements to the system are identified through the analysis of the MKRAM project description, stating what each participant needs and expects to obtain at the end of the project. However, for a better understanding of how all their needs can be included in one system, the top-down model-based system engineering principle is used. According to this approach, a system is analyzed in terms of layers, starting with a general description and then going into more details with every following layer.

5.1.1 Layer 1. General understanding of a system

In the first layer, the general explanation of system components and interrelationships are visualized, see Fig. 5.2. There are three main components, namely customers, an intelligent system (a tool for quality assurance), and a production floor (additive manufacturing machine). A customer and the intelligent system should be able to communicate, with the intelligent system being able to understand what the customer wants and providing some solution, which is then sent to the additive manufacturing machine. As such, the intelligent system should also be able to communicate with the production floor.

In more details **functional requirements** to the system shown in Fig.5.2 are listed below:

- Accept request from the customer

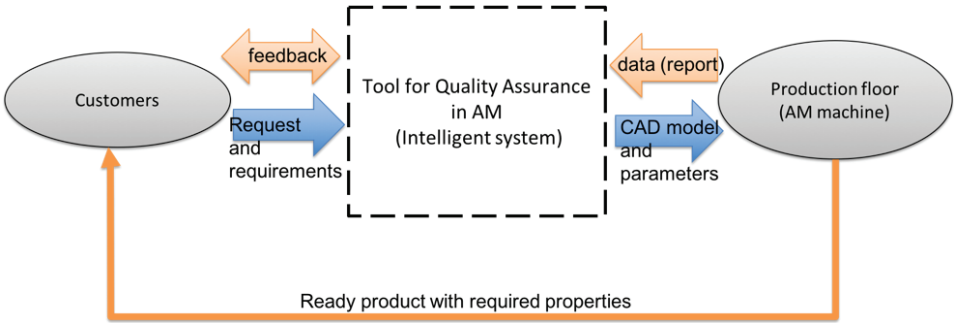


Figure 5.2: First layer of a system for quality assurance in additive manufacturing

- Understand request from the customer
- Analyze request from the customer (CAD model)
- React to request from the customer (send feedback, optimization/prediction)
- Save optimized results
- Send the results to the AM machine
- Accept report after part is produced (parameters and quality check)
- Accept feedback from the customer about product quality.

5.1.2 Layer 2. Description of main components of the intelligent system

Since a general understanding of what intelligent system should do is already described, the next step is to look into more details of the intelligent system (a tool for quality assurance in additive manufacturing in Fig. 5.3). Based on the previously defined functional requirements, to provide communication between customer and system, an interface for the intelligent system should be created. A request should be analyzed using the calculation core, with the results being saved in the database. In addition, the calculation core should have access to the required data, and thus, it should be connected to the database. Results from the calculation core should also be passed to the additive manufacturing machine.

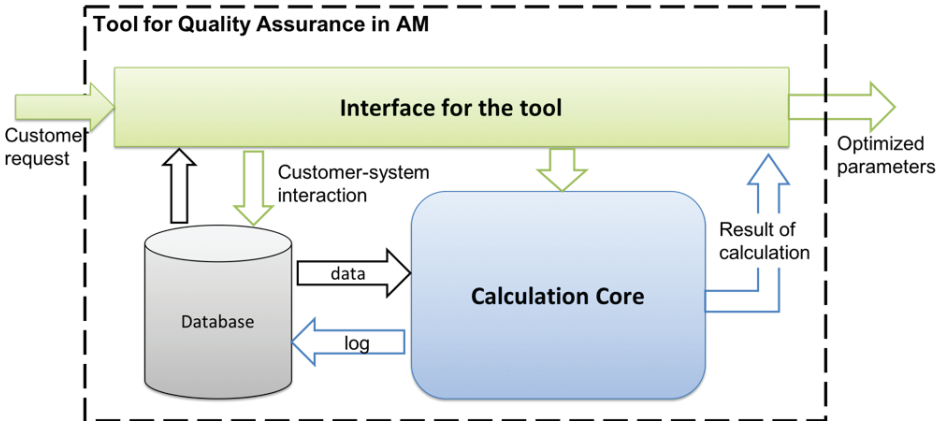


Figure 5.3: Second layer of system for quality assurance in additive manufacturing

In more details, **functional requirements** to interface shown on Fig. 5.3 are listed below:

- Should have clear links to operations and be easy to navigate
- Should have a field to enter a request
- Should have a button to upload a CAD model
- Should have a field to enter the type of AM machine
- User should be able to choose a type of material
- User should be able to choose a type of a problem

5.1.3 Layer 3. Description of the database component of the intelligent system

To define the requirements for the next component of the intelligent system, namely the database, more information about the input is provided in Figure 5.4.

Since the database is not the main priority of this research due to time limitation and is out of the scope of this study, just general **functional requirements to the database** (data storage) are proposed below:

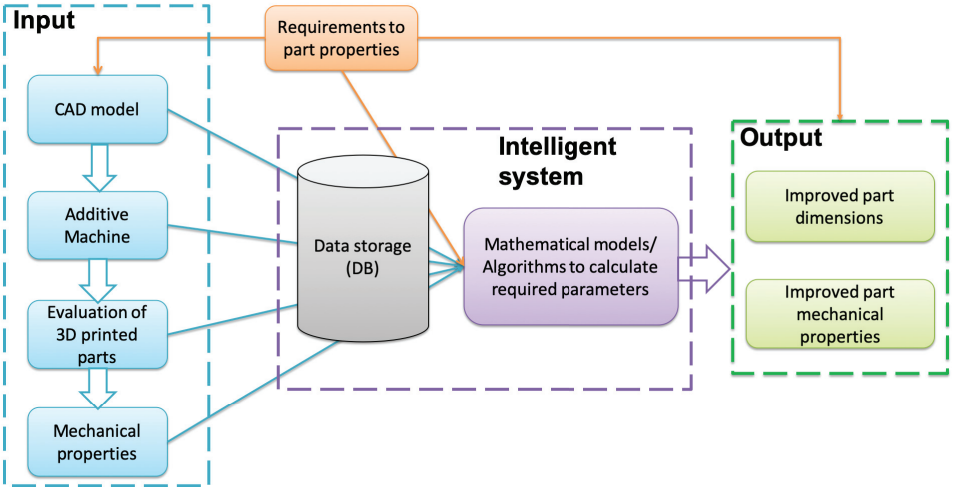


Figure 5.4: Third layer of system for quality assurance in additive manufacturing

- Should be able to retrieve data sets based on the queries
- Should store data in a structured way
- Should be flexible (to make changes in the future easily)
- Should work fast (as its operation will influence the calculation speed)

5.1.4 Layers 4-5. Description of the calculation core of the intelligent system

The main focus of this research is set on the development of the calculation core, as it will process and analyze requests from the customers. In order to address the stakeholders' needs, the calculation core, which could be considered as a set of mathematical models and algorithms, should be able to perform classification and prediction tasks (Fig. 5.5).

The classification module is important for this system because its result can predict which of the part may be produced with defects. Fig. 5.6 illustrates three classification groups, namely dimensional properties, mechanical properties, and their combination. Therefore, the following **functional requirements to the Classification component** should be included in the intelligent system:

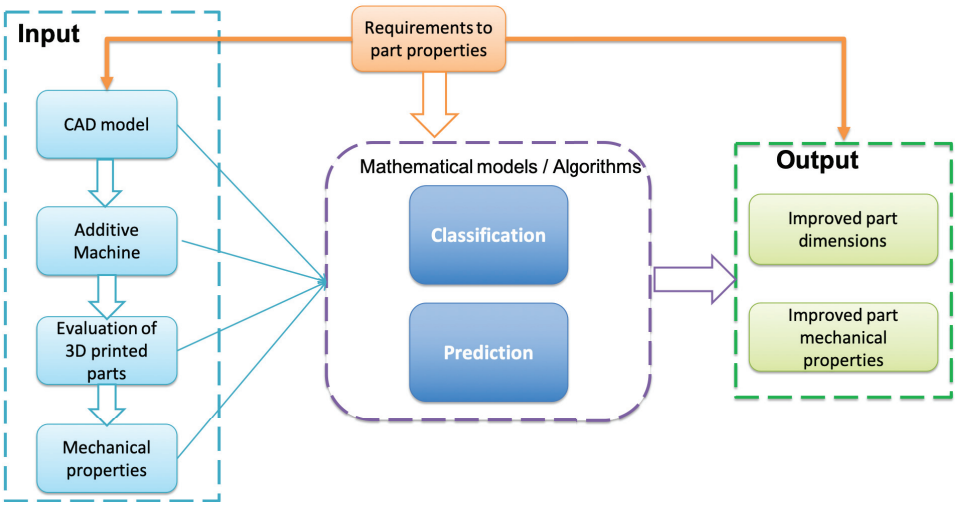


Figure 5.5: Fourth layer of system for quality assurance in additive manufacturing

- Classify parts that are within the tolerance range for dimensions
- Classify parts that are within the tolerance range for mechanical properties)
- Classify good thickness in regards to mechanical properties

Prediction component of the system, shown in Fig. 5.6, consists of 12 modules. They correspond to **functional requirements to the Prediction component**, which are the following:

P1 - Predict dimensions

P2 - Predict compensation ratio in x, y, and z axes individually for each part

P3 - Predict part position/placement (x, y, and z coordinates)

P4 - Predict the number of mesh triangles/surface/volume (CAD model properties)

P5 - Predict which standard set of parameters to use (TopQuality, Performance, Balance, Speed, TopSpeed)

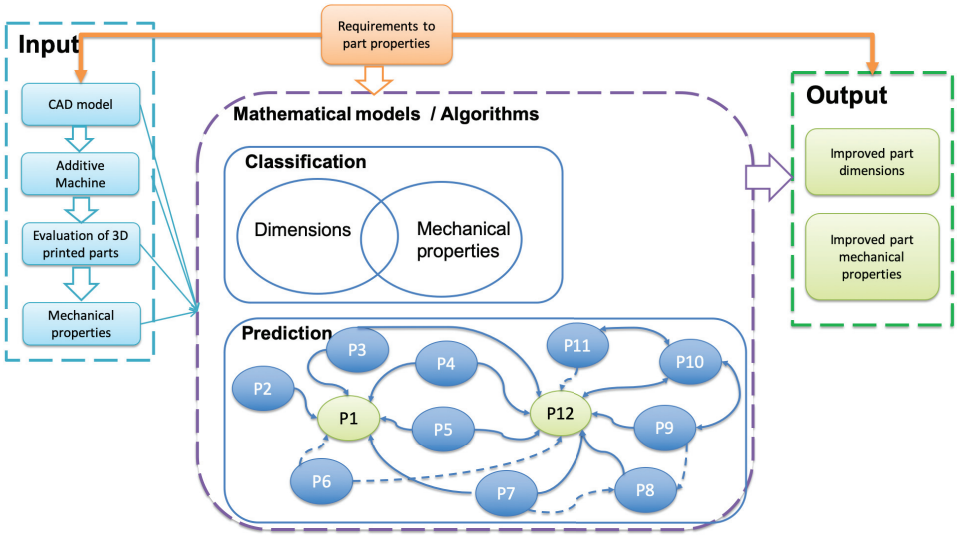


Figure 5.6: Fifth layer of system for quality assurance in additive manufacturing

P6 - Predict humidity level and temperature in the laboratory where the AM machine is placed

P7 - Predict layer thickness (process parameter)

P8 - Predict porosity level

P9 - Predict preheating temperature fo the AM machine (process parameter)

P10 - Predict virgin/used powder ratio (material parameter)

P11 - Predict powder bed temperature (process parameter)

P12 - Predict mechanical properties

In addition, connections between the modules are also shown in Fig. 5.6, where modules correspond to the parameters that need to be predicted or optimized. P1 and P12 modules are directly connected to the output, and therefore they are highlighted in the same color as output. Connections with solid lines mean that it is known from the previous studies that parameters are related to each other, while connections with dashed lines need to be investigated.

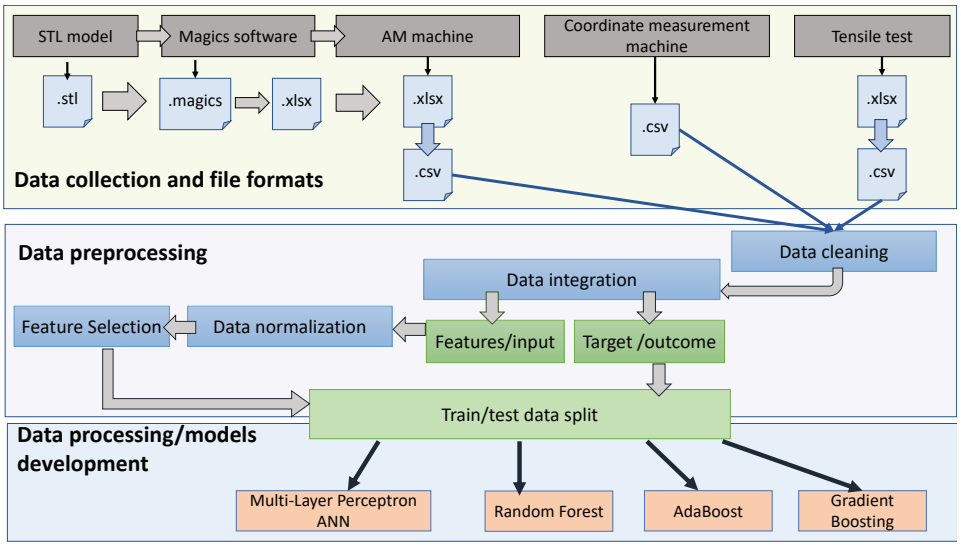


Figure 5.7: Representation of data pipeline conducted individually for each module; starting with the used equipment description

Due to the time limits of this work, only prototypes of modules P1, P2, P3, and P12 are developed based on the collected data. Fig. 5.7 shows a data pipeline that is used as a guideline for developing and describing prototypes for each module individually. Three main stages can be highlighted; data collection and file formats, data preprocessing, and data processing/models' development.

At the first stage, a typical process of how the data has been collected from the equipment is shown (more details are described in Chapter 4), and which file formats are extracted from each equipment and transformed into the CSV file format.

At the second stage, data preprocessing is started from cleaning data in each $*.csv$ file and then was passed through other important steps of data preprocessing. The last stage is the development of the models, which are a core part of the corresponding modules. The theoretical foundation of the data science techniques illustrated in Fig. 5.7 is described in Chapter 3. Selected modules for prototyping in this work are described in the corresponding Chapters 6 - 9.

Chapter 6

Module P1: Prediction of dimensional deviations

The results of this Ph.D. study are presented in four different chapters. This chapter describes how predictive models for estimation of dimensional deviations are developed. The description of the collected data is presented in the first section. The selection of significant features based on the filtering methods is presented in the second section. The description and comparison of intelligent predictive models are provided in the third section of this chapter.

6.1 Data exploration

In order to develop predictive models based on the collected data, there is a need for data understanding and preprocessing. This process allows preparing good quality data that doesn't contain empty cells or/and meaningless information. Therefore, this section aims at describing the collected data through graphical visualization of data distributions. Additional data explanations such as various tables and figures have been used for comparison of different runs.

Since additive manufacturing is a relatively new technology, there is a limited number of international standards that are ready to be applied for quality evaluation and assurance in the manufacturing. Therefore, using standards developed for traditional manufacturing processes could be con-

Table 6.1: Possible deviations (in *mm*) from DIN 16742:2013 [34] tolerance standard for injection molding for relevant values

Tolerance classes	0.5 – 3 mm	6 – 30 mm	30 –120 mm
fine	±0.05	±0.1	±0.15
medium	±0.1	±0.2	±0.3
coarse	±0.2	±0.5	±0.6
very coarse	—	±1	±1.5

sidered as a solution until the required standards for AM processes are developed.

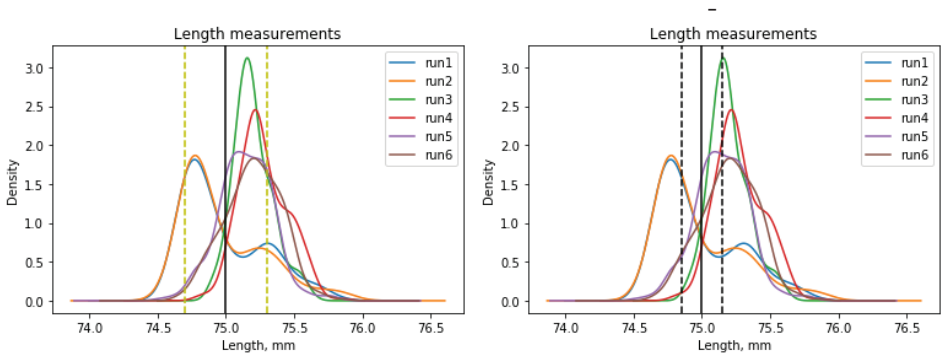
Results of the polymer powder bed fusion process are often compared with results from the injection molding process. Since the latter traditional manufacturing process has well-defined international standards for quality inspection and assurance, the DIN 16742:2013 [34] tolerance standard for injection molding will be used for comparison in this work. This standard has been chosen because it is used by researchers investigating the injection molding process [8, 29, 38].

According to DIN 16742:2013 [34], there are four main tolerance classes that are illustrated on Table 6.1. Depending on the application area, different tolerance classes are of interest. In this work, it is assumed that future applications could be the products with tight geometric tolerances. Therefore, the fine tolerance class is used as the main class in this work, while the medium class is also used for comparison.

6.1.1 Description of length measurements

The specimens produced in six experimental runs were designed with the length of 75 mm. All measured values (1526 specimens in total for six runs) are illustrated in Fig.6.1 as distributions based on kernel density estimation (KDE) for each run separately including a comparison of medium and fine tolerance classes.

As can be seen from Fig.6.1, the difference between experiments performed in 2017 and 2019 is present. While comparing two tolerance classes, Table 6.2 shows that even for medium tolerance class, there is a number of specimens that are outside of tolerance range. In cases when the tolerance



((a)) Medium tolerance class **((b))** Fine tolerance class

Figure 6.1: KDE distributions of length measurements for all runs with a reference to DIN-17642-2013-10 standard for injection molding tolerances

Table 6.2: Number of specimens outside the tolerance range (length measurements) based on medium (M) and fine (F) tolerance classes according to DIN 16742:2013 [34] tolerance standard for injection molding

Measured dimension	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6
Length (M)	67	63	50	114	65	95
Length (F)	163	173	127	224	162	194

range is tighter, like the fine tolerance class, most of the specimens will not meet requirements set to the length value. Since it was already defined that the fine tolerance class is chosen as the main class for quality evaluation, the upcoming analysis of length measurements will be done by comparing results with this tolerance class.

Additionally, it is important to take a look at how the collected data is distributed since it will provide an overview of run-to-run variations, and a comparison of the obtained results will be easier to conduct. Thus, Fig.6.2(a) shows that the measured length from Run 1 and Run 2 has multimodal distributions with two peaks, where the first (”larger”) peak corresponds to the length value and is smaller than the nominal value (75 mm). The second (”smaller”) peak corresponds to a value that is larger than the nominal one.

Two possible factors may lead to the bimodal distribution, namely the

parts orientation or their location in the build chamber. The more detailed orientation-based comparison of the results will be presented hereafter in the section.

The results from runs 3-6 are distributed normally (see Fig.6.2(b)), but still, most of the measured values are outside of the fine tolerance class and are larger than nominal values. Additionally, one can observe that variation for the results obtained from runs 1-2 is larger than the variation observed for runs 3-4.

The main difference between the runs (except for runs 1 and 2 because they are identical with the same material, process parameters, build layout design and produced within one week) are material [1] and build layout design. For example, Fig.6.2(c) shows the difference between experiments in 2017 and the first experiment (Run 3) in 2019 when the main difference is the material mix. Since Run 3 reflects how a change of the material influences the quality, the material used in 2019 may have had better properties than the material used in 2017 because the variation of the measured length has decreased.

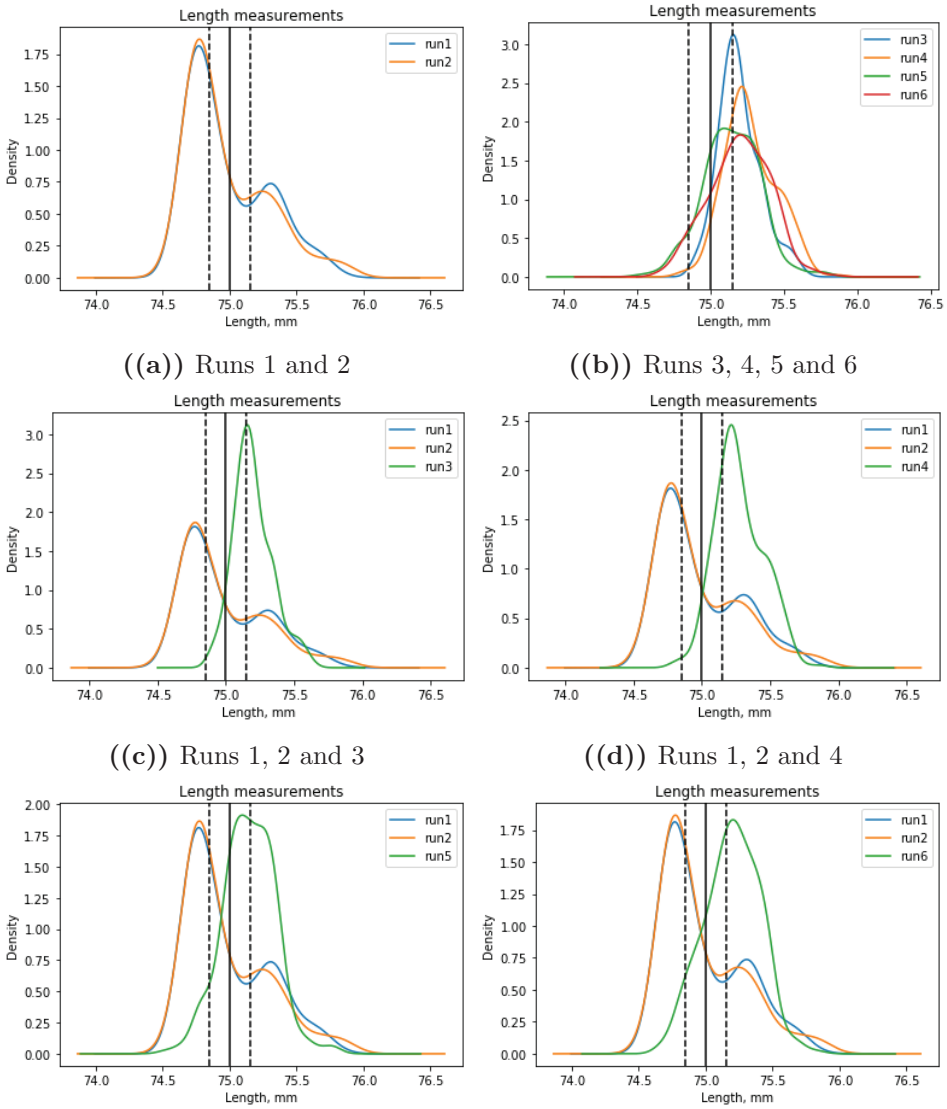
When all parameters have been controlled and kept the same (Fig.6.2(a)) data distributions are similar, and results show good repeatability. However, results obtained by using the same build layout design and process parameters but different material show different data distributions. Similar data distributions were observed when small changes to build design have been done (see Fig.6.2(d), Fig.6.2(e), and Fig.6.2(f)).

The change in the material is an important factor in powder bed fusion systems, and it has already been reported by researchers in previous studies. By summarizing the observed results, one can see that the first group of experiments (year 2017) resulted in mainly smaller specimens, while the second group of experiments (year 2019) has resulted in larger specimens.

Orientation-to-orientation comparison of length measurements

As presented earlier, multimodal distributions observed for runs 1-2 could be a result of orientation groups and/or their locations in the build chamber. In order to understand how different orientations contribute to the data distributions in general, Fig. 6.3 is illustrated below.

[1] Material is PA12 from different batches of mixed virgin powder and regularly aged powder



((a)) Runs 1 and 2

((b)) Runs 3, 4, 5 and 6

((c)) Runs 1, 2 and 3

((d)) Runs 1, 2 and 4

((e)) Runs 1, 2 and 5

((f)) Runs 1, 2 and 6

Figure 6.2: Kernel Density Estimation distributions of length measurements for all runs with a reference to DIN-17642-2013-10 tolerance standard for injection molding

While in Fig. 6.1, multi-modal distribution was observed for runs 1-2, Fig. 6.3 shows that Run 3 also has multimodal distribution for XZY and ZYX orientations, while Run 5 has multimodal distributions in ZYX and Angle orientations. Additionally, the similarity of the distributions for runs 1-2 is observed between ZYX and partly Angle orientations. However, the first peak in Fig. 6.1 mainly correspond to the joint distributions of XYZ and XZY orientation, while the second peak corresponds to the results for ZYX and Angle orientations.

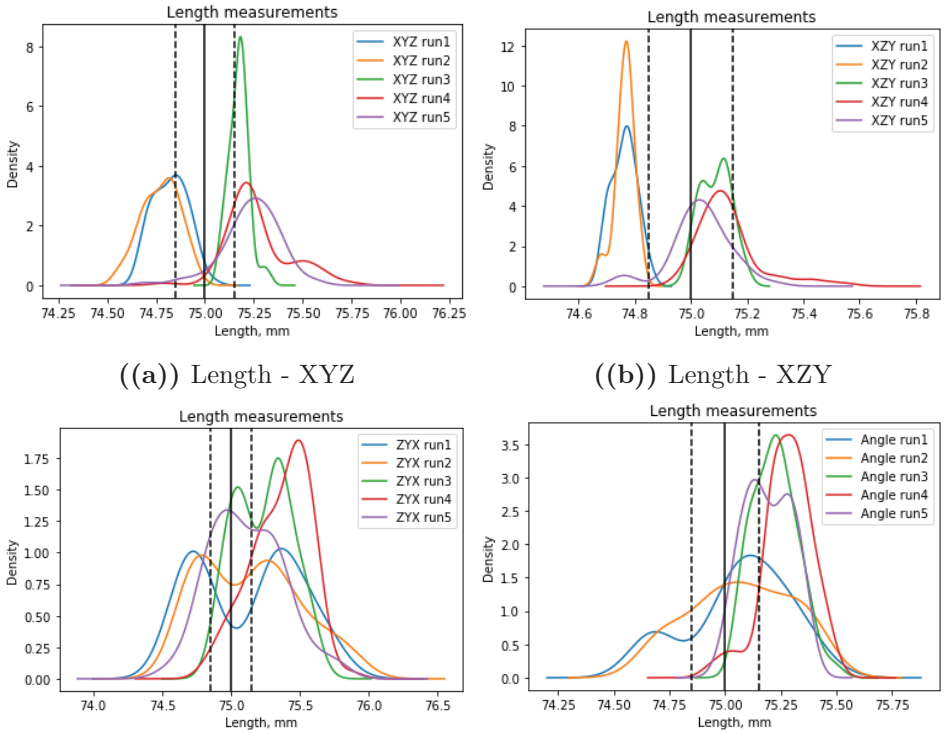
In other words, samples fabricated in different orientations tend to have different variations within one run. The XYZ and XZY orientations provide roughly normally distributed results, while ZYX and Angle orientation are more alike, and there is a need to understand which factors can lead to the bimodal distributions for these orientations.

Since in ZYX orientation length value corresponds to the dimension that is measured along the z-axis, it is assumed that parts' location in the build chamber can be one of the reasons leading to the observed results. Therefore, Fig. 6.4 shows distributions of the measured length values along the z-axis in the build chamber only for the runs and orientations that have multimodal distribution in Fig. 6.3.

As can be seen, there is a difference between variations related to the part location along the z-axis in the build chamber. For example, specimens are larger on the last layers of the build chamber when $z>200$ (bright green color on Fig. 6.4), while specimens are smaller or close to the desired values in the first layers of the build chamber where $z<75$ in the build chamber. In other words, the temperature distribution and cooling time are different for the described regions if to look at them as a function of time. While the first layers have already started a process of cooling down, the last layers haven't been produced yet. Therefore, these factors are important for the observed deviations along the z-axis, and they need to be taken into account in the future experiments.

6.1.2 Description of width measurements

While length measurements were done on the same areas of the specimen for all runs, the width and thickness measurements have additional values for runs 3-6. As described in Chapter 4, each specimen has two sides with width designed to be 10 mm and a narrow central part, which is intended



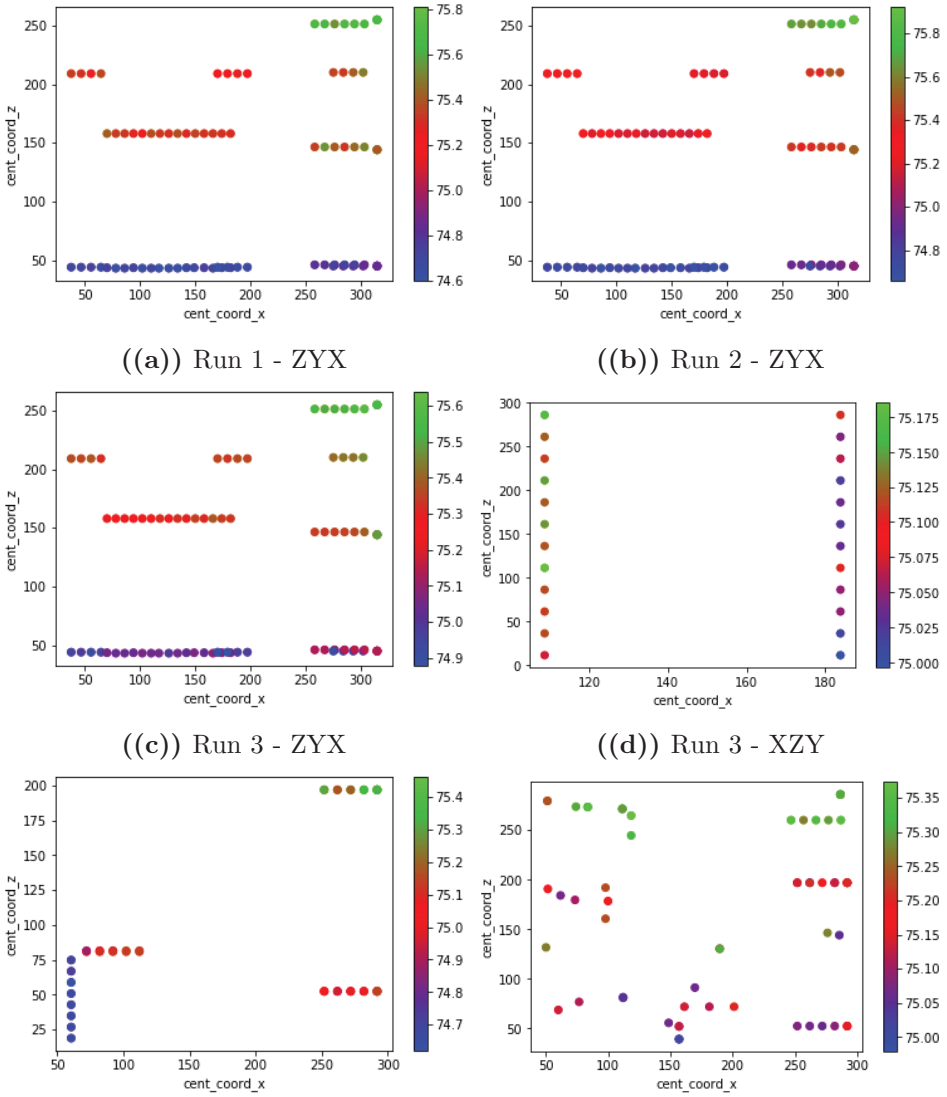
((a)) Length - XYZ

((b)) Length - XZY

((c)) Length - ZYX

((d)) Length - Angle

Figure 6.3: Orientation-based illustration of length measurements via Kernel Density Estimation distributions with reference to DIN-17642-2013-10 tolerance standard for injection molding



((a)) Run 1 - ZYX

((b)) Run 2 - ZYX

((c)) Run 3 - ZYX

((d)) Run 3 - XZY

((e)) Run 1 - Angle

((f)) Run 5 - Angle

Figure 6.4: Distribution of measured length values in x and z axes of the build chamber

Table 6.3: Number of specimens outside the tolerance range (width measurements) based on medium (M) and fine (F) tolerance classes according to DIN 16742:2013 [34] tolerance standard for injection molding

Measured dimension	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6
Width Side A (M)	79	82	5	29	13	15
Width Side A (F)	174	179	62	109	91	71
Width Side B (M)	–	–	5	30	14	14
Width Side B (F)	–	–	62	107	89	87
Width Center (M)	–	–	34	69	85	78
Width Center (F)	–	–	116	164	157	156

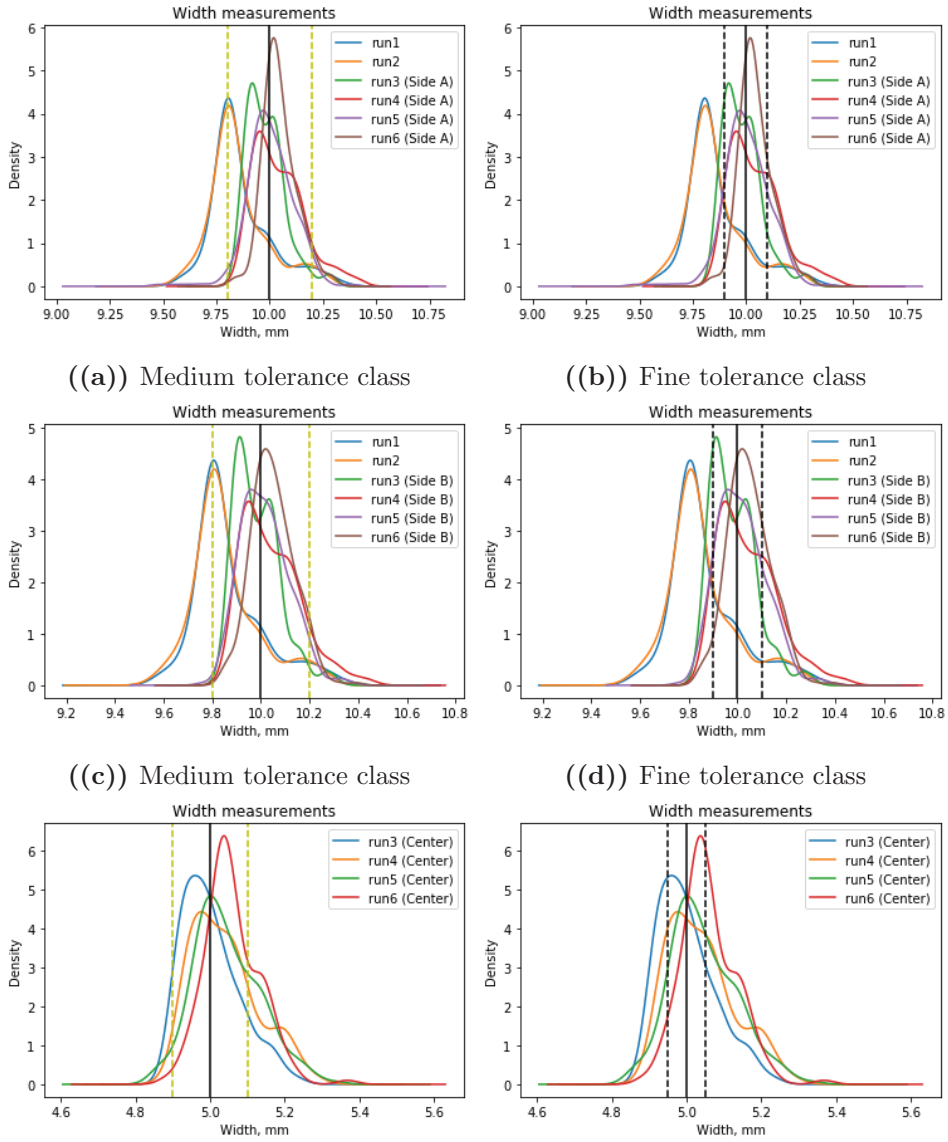
to be 5 mm. During the experiments from the 2017, a single side [2] has been measured (10 mm), and it was not registered which side it was. In the process of data analysis after Run 1 and Run 2, the author understood that the missing information was important. Therefore this issue has been addressed in all other measurements since 2019.

Thus, Fig.6.5 illustrates measurements for side A, side B, and center (side C) of the specimen for runs 3-6 separately, while runs 1-2 are compared with both sides A and B. Similarly to how it was done for length measurement, medium and fine tolerance classes are shown in Fig.6.5. Additionally, the number of parts outside the tolerance range is calculated using the medium and fine tolerance classes as a reference, and the results are summarized in Table 6.3.

In the case of medium tolerance class, most of the width measurements for runs 3-6 are within the tolerance range, while approximately 80 out of 217 specimens from Run 1 and Run 2 are out of the tolerance range (See Fig.6.5(a) and Fig.6.5(c)). However, when it comes to the fine tolerance class, approximately 170 out of 217 width measurements from runs 1 and 2 are out of tolerance range (see Table 6.3 and Fig.6.6).

Another important issue is the repeatability of mesurements on two sides of one sample. Results in Table 6.3 and Fig.6.6 illustrates that KDE distributions for Side A and Side B are similar at each run, but Run 3 and Run 4 have multimodal distributions (see Fig.6.6(a) and Fig.6.6(c)). Therefore,

[2]The side was randomly chosen



((a)) Medium tolerance class

((b)) Fine tolerance class

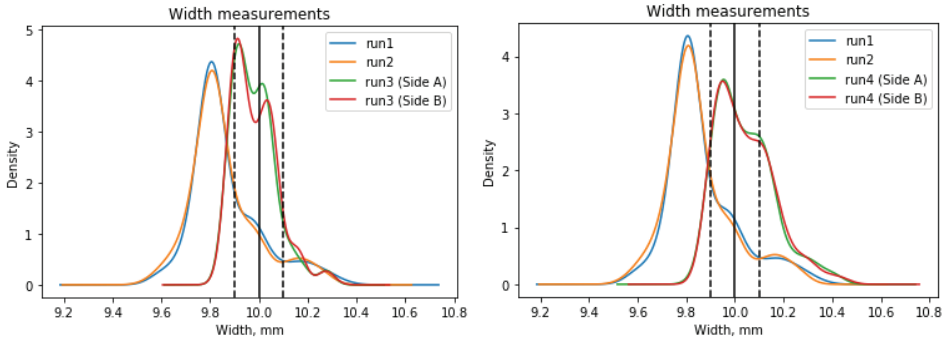
((c)) Medium tolerance class

((d)) Fine tolerance class

((e)) Medium tolerance class

((f)) Fine tolerance class

Figure 6.5: Kernel Density Estimation distributions of width measurements for all runs with reference to DIN-17642-2013-10 standard for injection molding tolerances



((a)) Runs 1 and 2

((b)) Runs 1, 2 and 4

((c)) Runs 1, 2 and 5

((d)) Runs 1, 2 and 6

Figure 6.6: Kernel Density Estimation distributions of width measurements for all runs with reference to DIN-17642-2013-10 tolerance standard for injection molding

there is a need for orientation-based analysis.

Moreover, Run 6 also has a different number of defects for Side A and Side B when the fine tolerance class is used (see Table 6.3). One of the reasons that leads to different variations on the sides could be a non-even temperature distribution within the build chamber, which could be caused by build layout design. The distribution of the samples within the build chamber or specific locations in the build chamber could affect the physical process. For instance, some specimens can cool down faster than others (in the corners of the build chamber).

Similarly to the results for length measurements, change of material batches may have the most significant impact on the deviations between results for experiments from 2017 and 2019. Additionally, it is also observed that changes in the build layout could lead to a change from multimodal distribution to Gaussian distribution for Run 5 and Run 6 (See Fig. 6.6(c) and Fig. 6.6(d)). However, these assumptions need to be analyzed in more detail with corresponding statistical techniques, which will be presented in the following chapter.

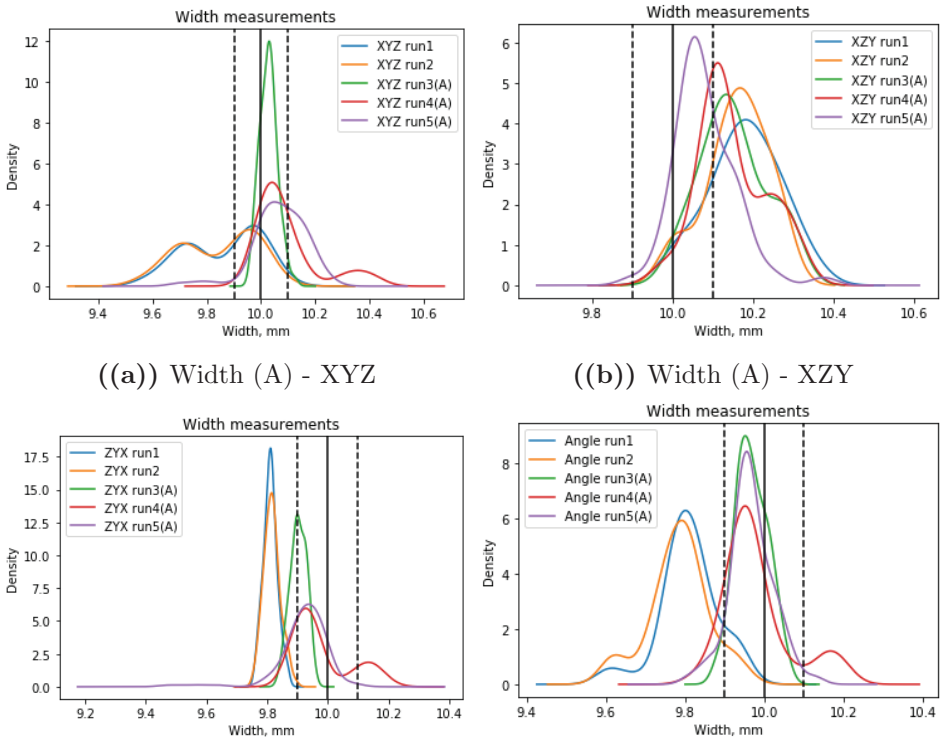
Orientation-to-orientation comparison of width measuements

Orientation-to-orientation analysis of width is also performed separately for side A, B, and C (center of the specimen) for runs 1-5. However, as can be seen from Fig. 6.7 and Fig. 6.8, distributions within the same orientation groups look similar for width measured on sides A and B. Thus, it can be assumed that factors, impacting on the variations of measured values are the same.

By comparing the distributions from runs 1-2 and Run 3, one can see that even though results for Run 2 repeats the result from Run 1, the distributions are multimodal, and width variation is larger than for other runs. Although, width measured for Run 3 has a normal distribution that is inside the tolerance range, and material is the only factor that was varied.

The results for runs 4-5 have similar behavior as for Run 3, while the material is similar between runs 3-5, the build layout has been changed in the runs 4-5 what led to larger variation. However, the results are still better than for the runs 1-2.

One of the reasons why this phenomenon is observed is related to the ma-



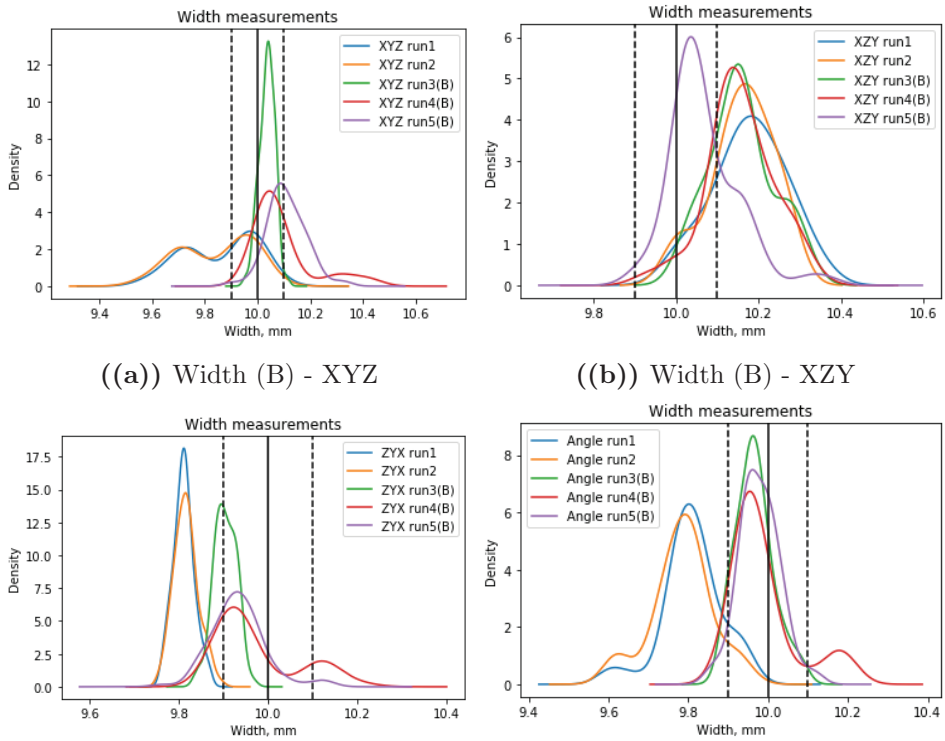
((a)) Width (A) - XYZ

((b)) Width (A) - XZY

((c)) Width (A) - ZYX

((d)) Width (A) - Angle

Figure 6.7: Orientation-based illustration of width (side A) measurements via Kernel Density Estimation distributions with reference to DIN-17642-2013-10 tolerance standard for injection molding



((a)) Width (B) - XYZ

((b)) Width (B) - XZY

((c)) Width (B) - ZYX

((d)) Width (B) - Angle

Figure 6.8: Orientation-based illustration of width (side B) measurements via Kernel Density Estimation distributions with reference to DIN-17642-2013-10 tolerance standard for injection molding

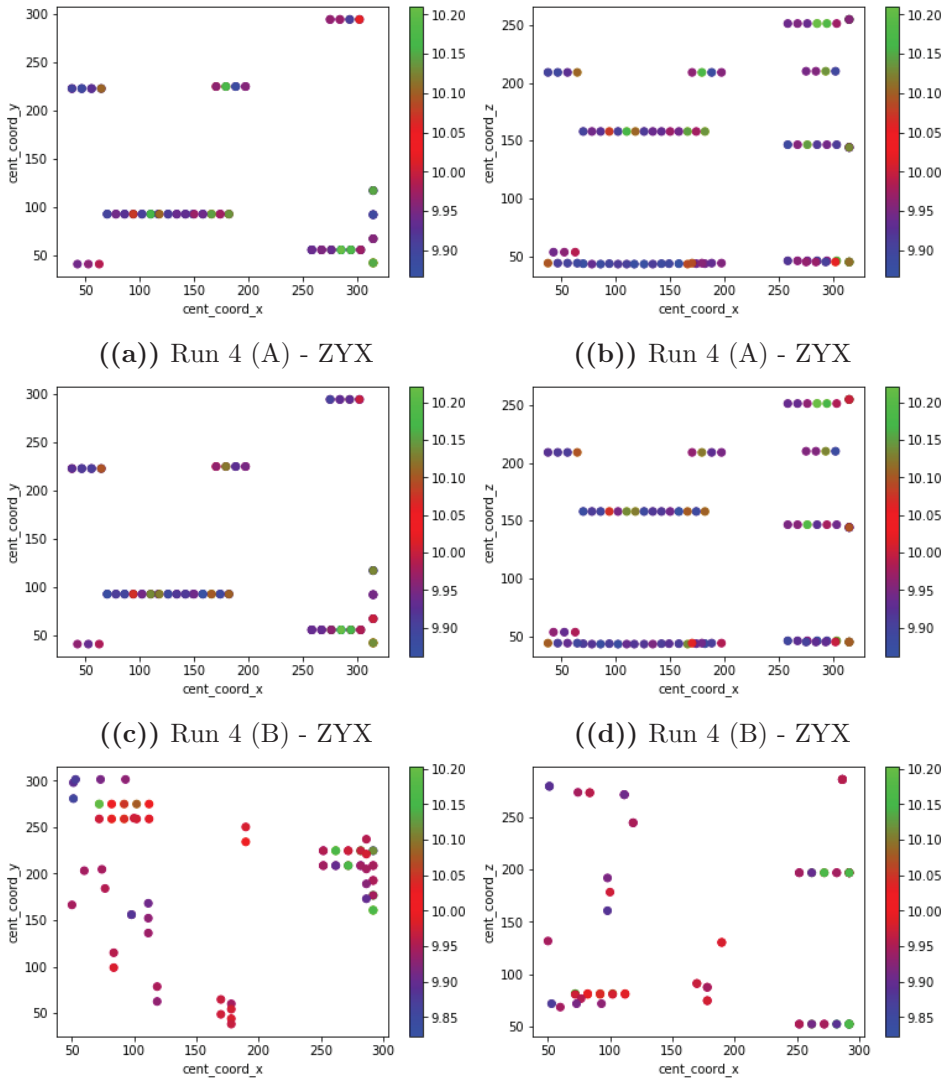
terial properties, but there is another not controlled factor that may have an impact on the quality. The powder bed fusion machine has a scheduled maintenance, in which the machine is recalibrated. Since the first two runs were conducted in 2017, while other runs in 2019, the powder bed fusion machine had a number of scheduled maintenance in two years. Therefore, in addition to the material variation, maintenance of the powder bed fusion system could have influenced the quality of the fabricated specimens.

At the same time, the results obtained in XZY, ZYX, and Angle orientation groups look similar for all runs. For instance, Fig. 6.7(b) shows that width in XZY orientations for all runs is larger than designed, and all distributions have a large variation. In ZYX orientation, distributions are narrow, meaning small variation, and are normally distributed except for Run 4. Fig. 6.7(c) shows that width values for Run 4 has multimodal distribution with two peaks, which is similar to the data distribution for Run 4 observed for Angle orientation (see Fig. 6.7(d)).

It is assumed that multimodal distribution is observed for ZYX, and Angle orientation groups of Run 4 are related to the parts' location in the build chamber. Since results for width measured on side A and side B are similar, analysis of the impact of specimens' location on the measured values will be illustrated only for one side (side A). However, after a comparison of the results for Run 4 measured on side A and side B for ZYX orientation, Fig. 6.9(a)- 6.9(b) and Fig. 6.9(c) -6.9(d) show similarities between two runs, but still, in the build chamber, there are randomly distributed specimens that have different values in the same position. Additionally, it can be seen that in the right side of the build chamber the specimens are larger than in other positions of the build chamber.

It can be summarized that there is a small difference between measurements on side A and side B in specific locations. This difference may be present due to random error or temperature distribution related to the size of powder particles, and how the material is distributed on the build plate at each layer. A similar observation is made for the Angle orientation of Run 4, where the part location is related to the multimodal distribution. Moreover, the same side of the build chamber as for other measurements, including length, provides larger specimens than it was designed.

When it comes to the orientation-to-orientation comparison of the measured width on side C (in the central part of the specimen), this measure-



((a)) Run 4 (A) - ZYX

((b)) Run 4 (A) - ZYX

((c)) Run 4 (B) - ZYX

((d)) Run 4 (B) - ZYX

((e)) Run 4 (A) - Angle

((f)) Run 4(A) - Angle

Figure 6.9: Distribution of measured length values for Run 4 in (x,y) and (x,z) axes of the build chamber

ment hasn't been done for experiments from 2017. Thus the influence of the material change is assumed to be the same for other width measurements. However, Fig. 6.16 shows that distributions in XYZ, ZYX, and Angle orientations are normal, and Run 3 has the smallest variations in ZYX orientation compared to the other two runs.

While multimodal distributions have been observed for Run 4 in ZYX and Angle orientations for both side A and side B, different behavior is noted for the central part of the specimen. One of the reasons that may lead to the obtained results could be a temperature distribution in the build chamber that is influenced by the density of part placement (build layout design). This assumption is hard to investigate based on the experiments performed but should be considered as a task for future work.

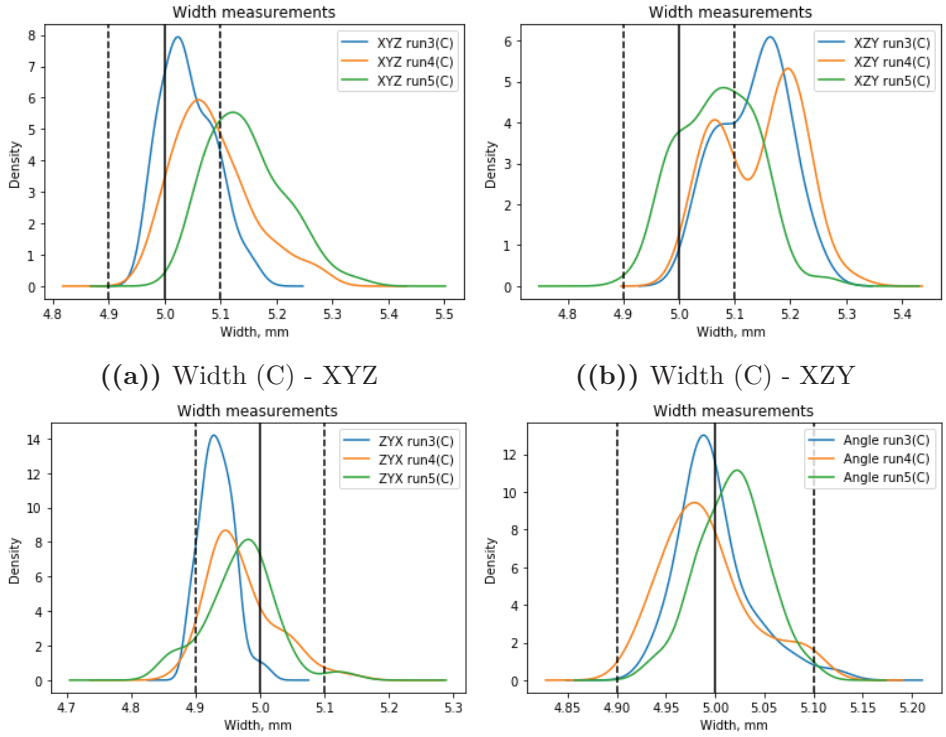
The multimodal distribution is observed in XZY orientation for Run 3, which is similar to the data distribution for length and Run 4. Since the previous analysis has shown that usually, multimodal distribution is present, in most cases, due to locations of the specimens in the build chamber, a similar analysis should be conducted for Run 3 and Run 4 in XZY orientation.

However, Fig. 6.11 shows that variations in measured width on side C are random, and there is no clear pattern or specific area in the build chamber that corresponds to larger or smaller specimens. Thus, there either should be some additional factors that may lead to the deviations, which are not known to the author, or this is a random error that needs to be defined and controlled in some way.

6.1.3 Analysis of thickness measurements

Even though thickness was measured in the same way as width, its value should be the same (2 mm) across the whole specimen. Similarly to the width measurements, thickness in Run 1 and Run 2 has been measured only on one side that was randomly chosen. The obtained results are compared to the measurements in different location on the specimen, which are sides A, B and C. Thus, Fig.6.12 illustrates comparison of the results for thickness measured on sides A, B and C (center of the specimen) for runs 3-6 with measurements for Run 1 and Run 2.

While thickness measurements are very similar on Side A for all runs, measured value in a central part of specimens has the highest variation from run to run. However, when it comes to the tolerance ranges defined by the



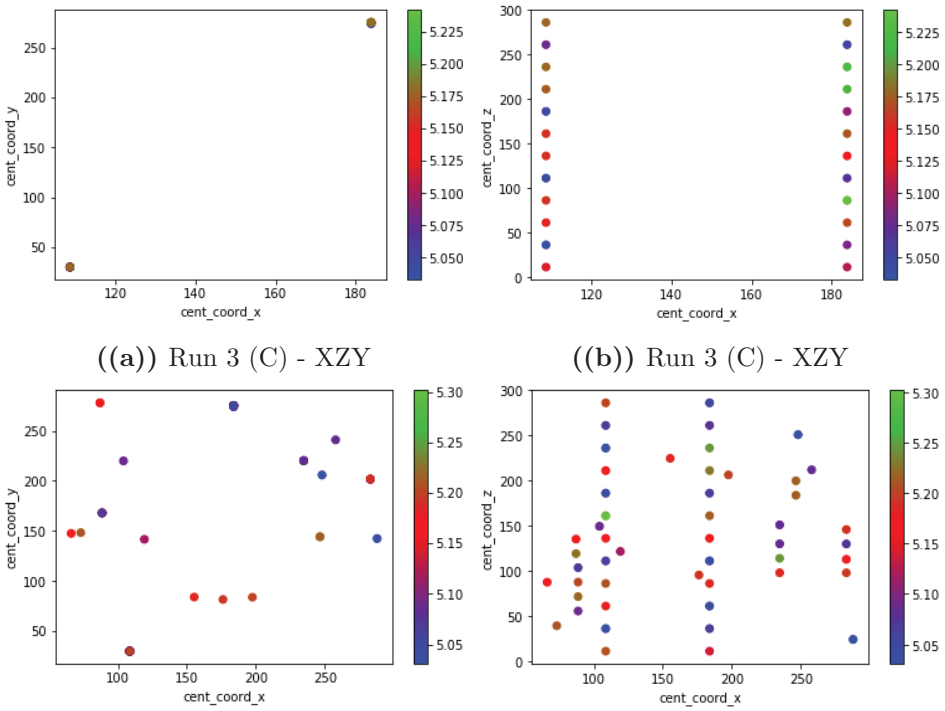
((a)) Width (C) - XYZ

((b)) Width (C) - XZY

((c)) Width (C) - ZYX

((d)) Width (C) - Angle

Figure 6.10: Orientation-based illustration of width (side C) measurements via Kernel Density Estimation distributions with reference to DIN-17642-2013-10 tolerance standard for injection molding



((a)) Run 3 (C) - XZY

((b)) Run 3 (C) - XZY

((c)) Run 4 (C) - XZY

((d)) Run 4 (C) - XZY

Figure 6.11: Distribution of measured width (side C) values in the build chamber along x, y and z axes

Table 6.4: Comparison of the number of specimens outside the tolerance range (thickness measurements) based on medium (M) and fine (F) tolerance classes according to DIN 16742:2013 [34] tolerance standard for injection molding

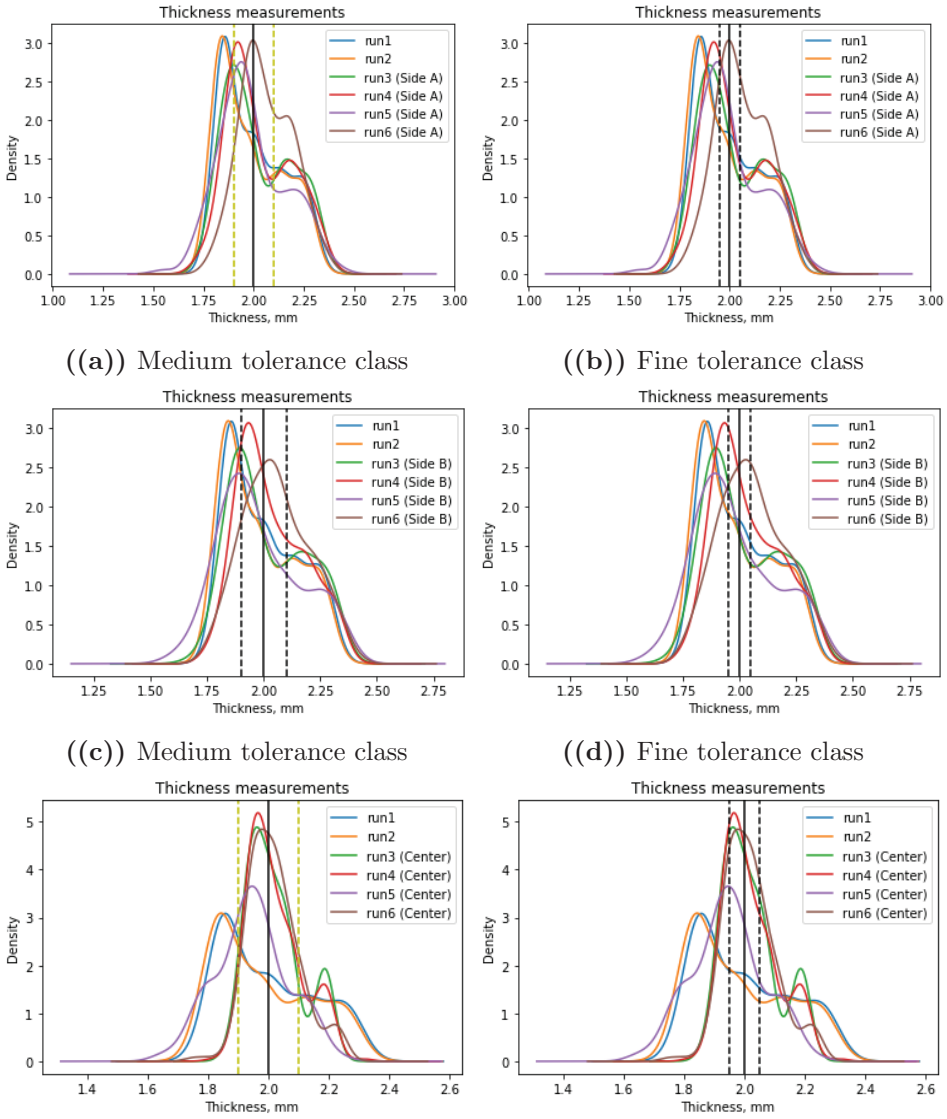
Measured dimension	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6
Thickness Side A (M)	162	157	147	162	164	138
Thickness Side A (F)	170	174	180	230	222	200
Thickness Side B (M)	–	–	146	146	185	148
Thickness Side B (F)	–	–	187	225	248	213
Thickness Center (M)	–	–	44	57	134	62
Thickness Center (F)	–	–	123	158	208	152

standard for injection molding [34], measurements from the central area of specimens are mainly within the medium tolerance class (see Fig.6.12(e)). In contrast, most of the measurements on sides A and B are out of both the medium and fine tolerance classes (see Table 6.4).

As the distributions for side A and Side B have similar shapes, frequency, and width, results from each run are visualized separately in Fig. 6.13. Both data distributions for runs 1-2 and Run 3 are multimodal. A change of material in Run 3 has led to the smaller range of variations in the thickness in the central part, but similar distribution for measurements conducted on sides for runs 1-3 (see Fig.6.13(a)). Then the question that arises from such results is why variation on sides of the specimen is different from the variation in the central area of the specimen within one run (given the same material).

Since width measurements have similar behavior, but its value is similar on the sides (10 mm), and smaller in the central area (5 mm), the concentration of energy applied to this area could be different. As a result, narrower regions of the part may be characterized by a faster colling process, thus leading to a different shrinkage ratio for all dimensions affected by this area (in this case, it is width and thickness of a specimen).

Similar behavior can be seen for Run 4 (shown in Fig.6.13(b)), where additional 75 specimens were fitted into an existing build layout design, and Run 5, where build layout design was changed (see Fig.6.13(d)). However, shape, frequency, and width (variation) of data distribution for the central



((a)) Medium tolerance class

((b)) Fine tolerance class

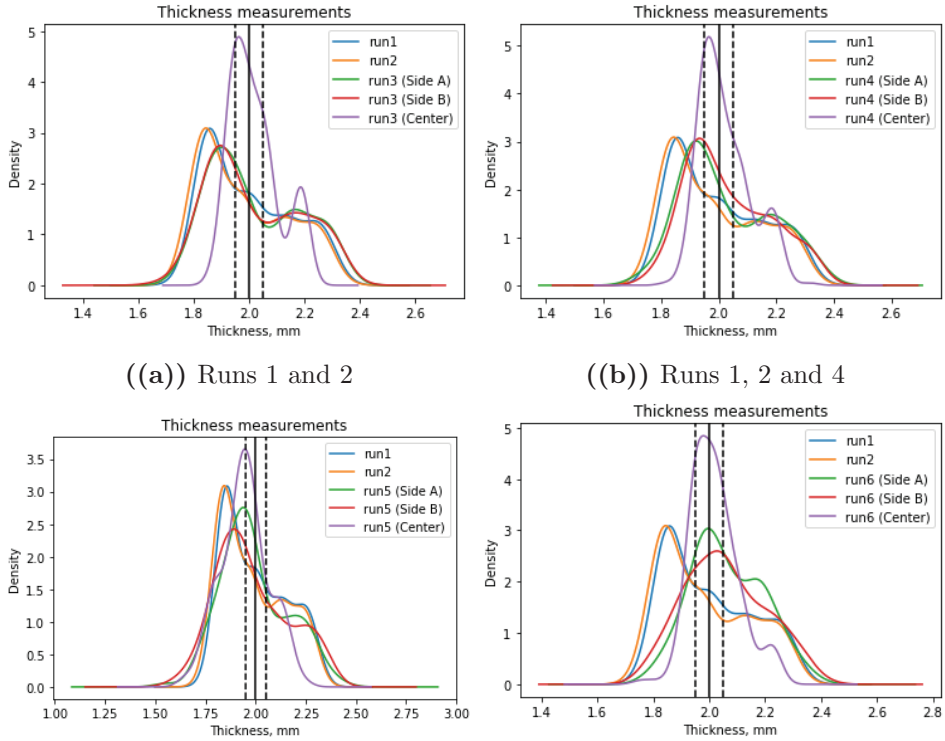
((c)) Medium tolerance class

((d)) Fine tolerance class

((e)) Medium tolerance class

((f)) Fine tolerance class

Figure 6.12: Kernel Density Estimation distributions of width measurements for all runs with reference to DIN-17642-2013-10 standard for injection molding tolerances



((a)) Runs 1 and 2

((b)) Runs 1, 2 and 4

((c)) Runs 1, 2 and 5

((d)) Runs 1, 2 and 6

Figure 6.13: Kernel Density Estimation distributions of thickness measurements for all runs with reference to DIN-17642-2013-10 tolerance standard for injection molding

area for Run 5 is more similar to the data distribution on sides. Therefore a change in the build layout is one of the possible reasons for the observed results. In other words, the build layout design can influence the distribution of the energy at each layer, and thus having an impact on the cooling time, which results in a similar shrinkage ratio at different locations within one sample.

Similarly, Fig. 6.13(d) shows that changing build layout and introducing more orientations than in the previous runs leads to a smaller variation, and multimodal distributions have the second peak that is much smaller than for other runs. In order to better understand the nature of the observed multimodal distributions for all measurements, orientation-to-orientation analysis is described below.

Orientation-to-orientation comparison of thickness measurements

One of the assumptions that different orientations lead to multimodal data distribution has been made for thickness measurements. However, Fig. 6.14 - 6.16 show that multimodal distributions are also observed for XYZ, XZY and Angle orientations. In XYZ orientation, the distributions for runs 1-3 for side A and side B are similar. At the same time, the distribution for Run 5 is normal on side A and multimodal on side B. The assumption behind this phenomenon is similar as for the width and length measurements, namely that in addition to the orientations, part location in the build chamber also affects the dimensional variation.

In other words, as it was already presented in other publications [36], the temperature distribution in the build chamber is not uniform. Thus packaging of the build platform (strategy for part placement in the build chamber) has an impact on how much energy is applied at each layer to the specific area of the build chamber. Therefore, part location can provide more information that is necessary to understand the described phenomenon.

In addition to the part location and orientation, material importance is also reflected in the results for thickness measurements. For example, multimodal distribution in XZY orientation is observed for the experiments (runs 1-2) conducted in 2017, but after changing only material, the thickness has a smaller deviation from the nominal value and is normally distributed for Run 3. Even after adding new specimens to the build layout, the data distribution is very similar. Only when changing the build layout by chan-

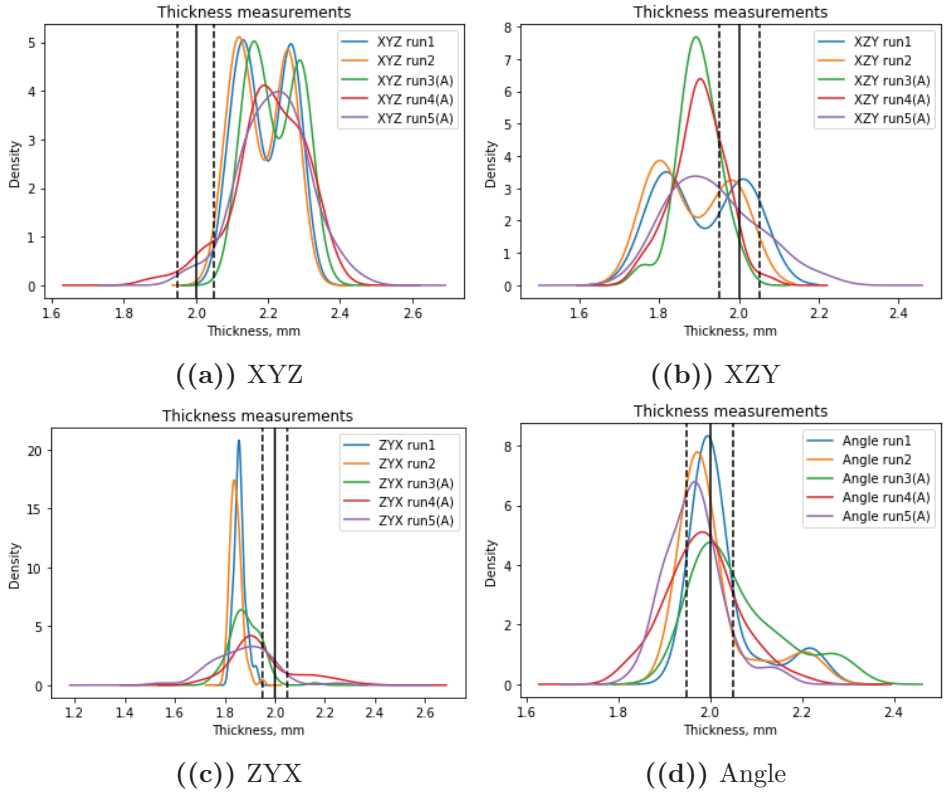
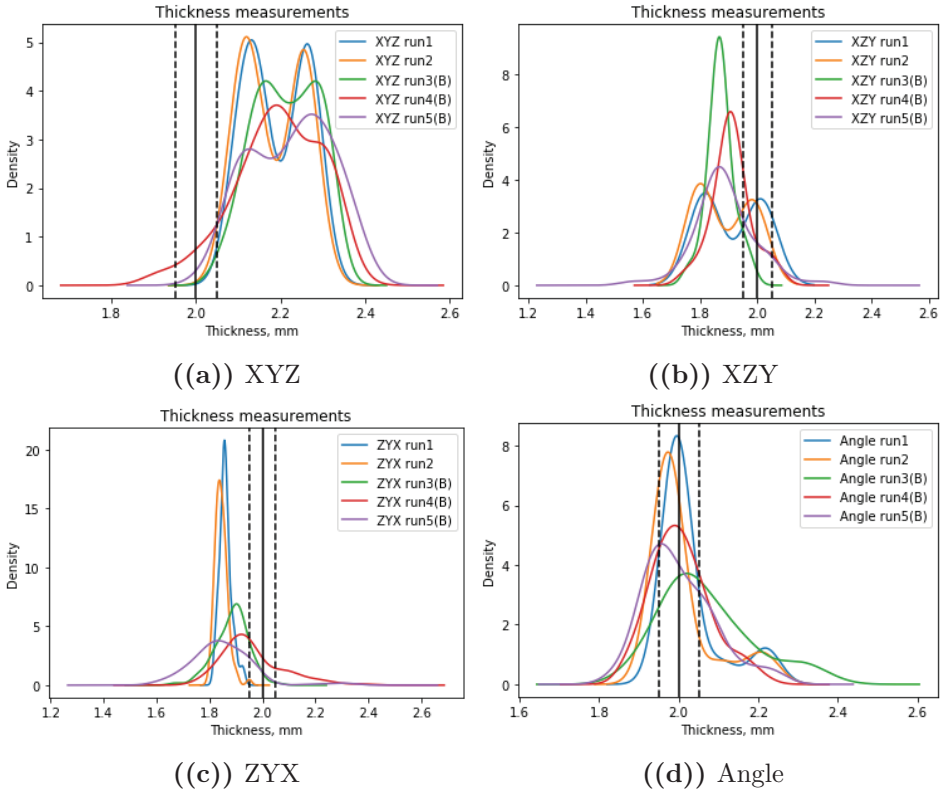


Figure 6.14: Orientation-based illustration of thickness (side A) measurements via Kernel Density Estimation distributions with reference to DIN-17642-2013-10 tolerance standard for injection molding

ging the specimens' orientation and location, the variation has increased to a large degree, as it is shown in Fig. 6.14(b).

Another new observation can be made for Angle orientation for Side A, which is shown in Fig. 6.14(d). The distribution of the measured thickness for all runs is similar in all aspects, namely distribution width, shape, and frequency. Therefore, one can see that for thickness measurement, specimens fabricated in different orientations and locations are less affected by the changes in the experiments.

The results of thickness measurements on side B is similar to the observed results for Side A. Fig. 6.14 - 6.15 also show that specimens fabricated



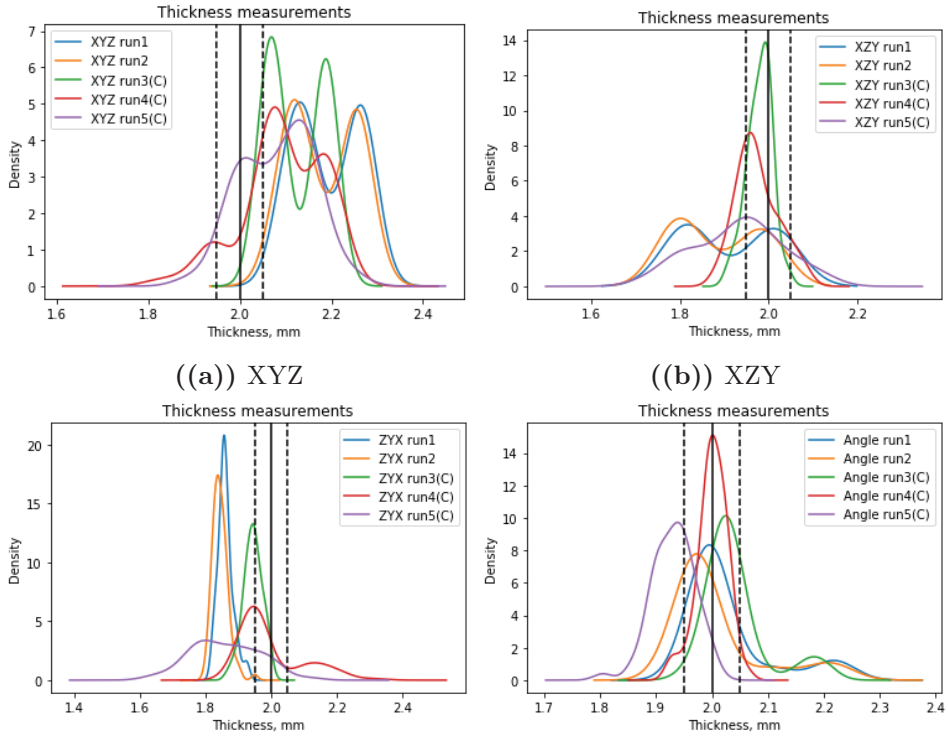
((a)) XYZ ((b)) XZY

((c)) ZYX ((d)) Angle

Figure 6.15: Orientation-based illustration of thickness (side B) measurements via Kernel Density Estimation distributions with reference to DIN-17642-2013-10 tolerance standard for injection molding

in XYZ orientation are larger than designed. Most of the measurements are smaller than designed in XZY and ZYX orientations, while in Angle orientation value is both smaller and larger than it was designed. However, thickness measured in the central area (side C) of the specimen has a similar shape of the deviations, but their width and frequency are different.

For instance, thickness measurements have multimodal distributions in XYZ orientation for all runs. Even more, as can be seen from Fig. 6.16(a), distribution for Run 4 has three peaks that have not been observed before for either length, width, or thickness (sides A and B) measurements. An additional assumption, which hasn't been made before, could lead to the observed variations; positions of the specimens in one orientation, for instance,



((a)) XYZ ((b)) XZY

((c)) ZYX ((d)) Angle

Figure 6.16: Orientation-based illustration of thickness (side C) measurements via Kernel Density Estimation distributions with reference to DIN-17642-2013-10 tolerance standard for injection molding

XYZ, may be affected by specimens from another orientation groups. Since the powder bed fusion process solidifies material in a layered manner, it is also important to investigate the scanning strategy that defines the way energy is distributed at each layer.

Thus, it can be seen that powder bed fusion technology is a complex process in which dimensional variations could be caused by different factors that are hard to capture in a small number of experiments. Therefore, all highlighted issues need to be further investigated in future experiments.

Analysis of Run 6 as a special case study

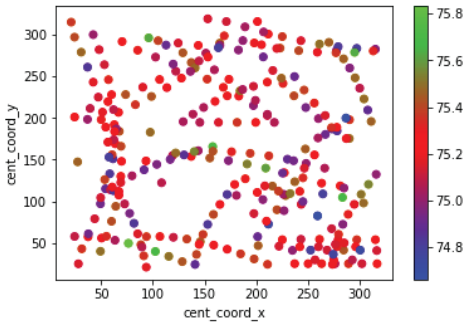
Analysis of Run 6 is a special case due to the number and type orientation groups applied to the current build layout design. A build chamber is a 3-dimensional space, which is described through the world coordinate systemin $\mathbb{R}^3$ (x, y and z axes). Therefore, length measurements are illustrated in Fig. 6.17(a) shows the distribution of specimens in 2-dimensional space with x and y used as axes; in Fig. 6.17(b), which illustrates the distribution of specimens in the build chamber in x and z axes, and Fig. 6.17(c), which shows specimens distribution along y and z axes.

For such type of the build layout design, visualization of length distribution shows that there are no specific areas in the build chamber where specimens tend to shrink or expand more. Both shrinkage and expansion of length are randomly distributed in the build chamber. Even more, most of the parts are larger than it was designed.

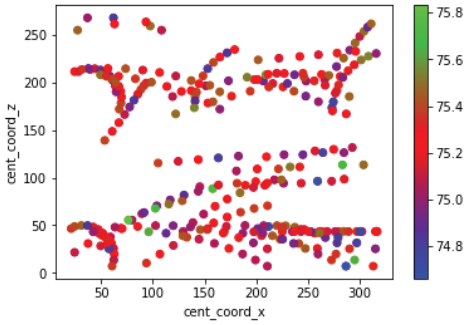
Width measurements are shown on Fig. 6.18 only on X/Y and X/Z axes in order to be able to compare visually different sides measured for this dimensional property. However, how width measurements are distributed in Y/Z axes has been performed and compared with results for X/Z axes beforehand, and the most informative 2D space has been chosen.

Fig. 6.18 shows that that results for side A and side B are similar, and the variation range is also similar. Only some of the points differ, but no pattern is observed that can detect in which location samples tend to shrink/expand more than in the other. There is also no difference between the center of the build layout and the sides. Along with that, the variation range for width measured on side C resembles the ones for sides A and B.

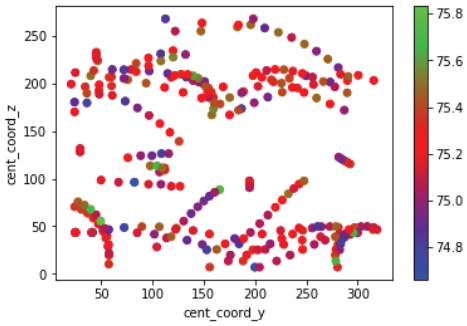
Since the current nesting density is larger than for other builds, it may be



((a)) X/Y axes

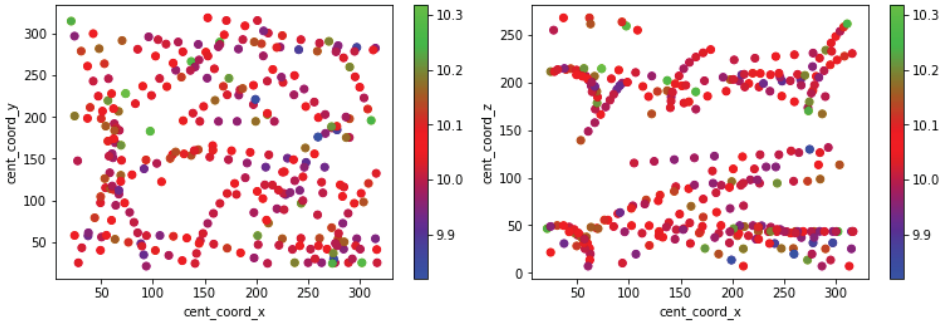


((b)) X/Z axes



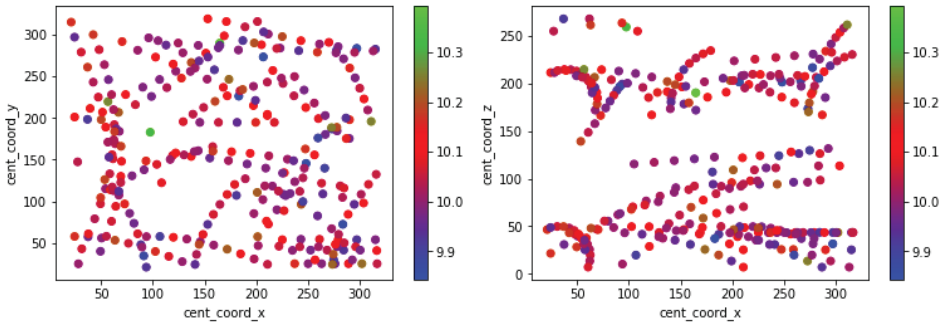
((c)) Y/Z axes

Figure 6.17: Distribution of measured length values for Run 6



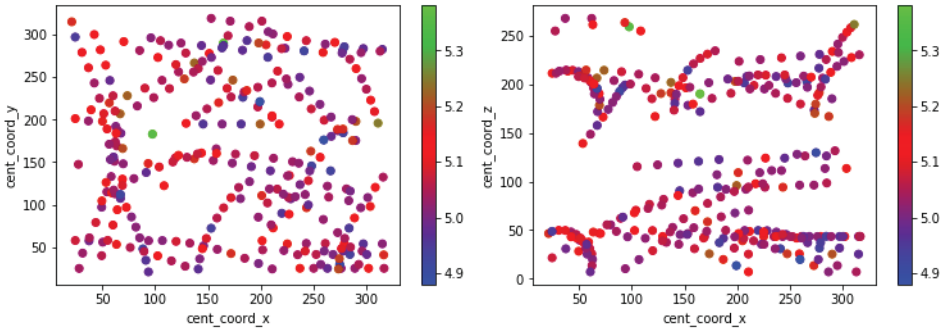
((a)) Side A - X/Y axes

((b)) Side A - X/Z axes



((c)) Side B - X/Y axes

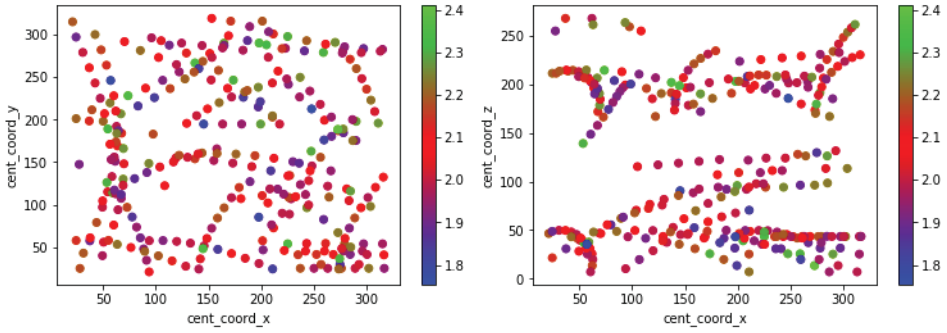
((d)) Side B - X/Z axes



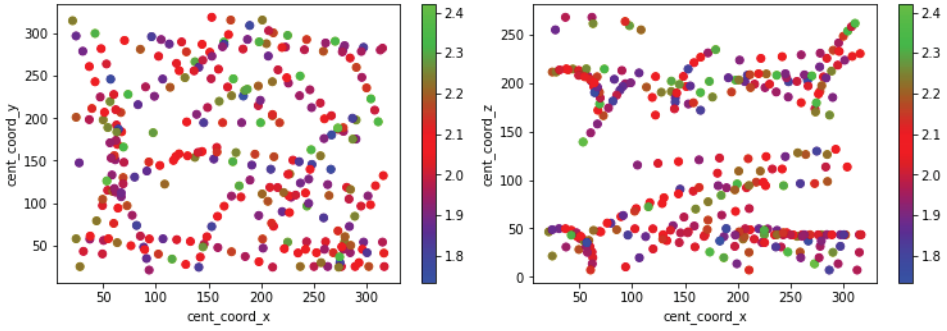
((e)) Side C - X/Y axes

((f)) Side C - X/Z axes

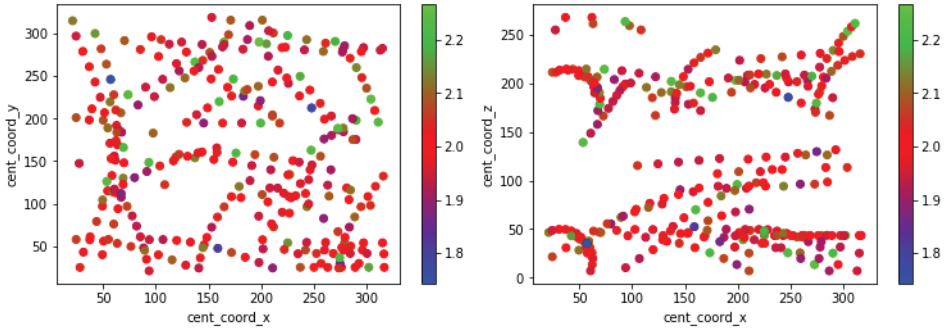
Figure 6.18: Distribution of measured width values for Run 6



((a)) Side A - X/Y axes ((b)) Side A - X/Z axes



((c)) Side B - X/Y axes ((d)) Side B - X/Z axes



((e)) Side C - X/Y axes ((f)) Side C - X/Z axes

Figure 6.19: Distribution of measured thickness values for Run 6

assumed that temperature within the build layout has been affected and has led to a more uniform cooling process. However, this assumption needs to be investigated in more detail with the help of additional thermal sensors.

Thickness measurements for Run 6 are analyzed similarly to the width, and results are shown in Fig. 6.19. The measurements on side A and side B are very similar, and have the same variation range, while thickness measured in the central area of the specimen has a smaller variation range. The same observations have been made for thickness measurements for other runs. Therefore, one can see a dependency between width and thickness. For larger width areas, thickness has larger variations, while for smaller width areas, thickness deviation has a smaller range. In other words, more energy is directed to the area, more expansion (phenomenon opposite to the shrinkage) is observed for this area in all directions.

6.2 Feature selection

Typically development of models based on machine learning techniques has challenges related to overfitting and long training time. Many different methods can be used to overcome these challenges. Feature selection is one of these techniques. The selection of the relevant features helps to reduce model overfitting, shorter training time, and improve prediction accuracy by decreasing noise that might come from irrelevant features.

Therefore, this section is focused on the selection of significant features for length, width, and thickness. The selected features will be used for the development of predictive models.

6.2.1 Feature selection for length dimensional property

Feature selection for Run 1 and Run 2

Table 6.5 shows a comparison of three different methods used for understanding what investigated features are significant and should be used for prediction of length. While the Mutual Information method and Spearman's rank correlation coefficient test (Spearman's Rho method) include analysis of non-linear correlations, the Pearson correlation test is a measure of only linear relationships between two parameters.

In order to understand which method should be used for further analysis, a quick check on the predictive model's performance has been done. Multi-

layer perceptron (MLP) has been trained on the same data, but with features selected by each method, and prediction accuracy obtained from the five-fold cross-validation test is used as a metric to choose a method for feature selection for other runs.

As can be seen from Table 6.5, the model has higher prediction accuracy with features selected based on the methods that take into account non-linear relationships, namely Mutual Info and Spearman's methods. Additionally, neglecting non-significant features has led to an increase of prediction accuracy from 64.44% to 91.63 % (models without and including feature selection, respectively).

Therefore, it can be summarized that features having a correlation rank higher than 0.1 are considered significant for prediction of length. In Table 6.5 features are listed in an order according to their importance rank based on Spearsman's correlation rank method. All features except for run number (*Run* on Table 6.5), orientation angles by y and z axes (*reor_b and reor_c*, respectively), material, build height, platform volume utilization, current nesting density and build layout design have been considered as the features that have significant correlation with length based on the collected data from runs 1-2.

The non-significant features do not correlate with the length for the Runs 1-2 since all of them were constant and not varied. Therefore, results from the correlation tests correspond to the physical aspects of the experiments.

Feature selection for all runs based on Spearman's rho ranking method

The first column A on Table 6.6 is taken from Table 6.5, and it shows how strong are correlations between investigated features and length for the case when build layout design, material, and machine process parameters are controlled and kept the same.

Column B on Table 6.6 shows feature selection results for data merged from runs 1-3. Run 3 was performed with the same build layout design and machine process parameters but with another batch of material. Therefore, it can be seen that the correlation between material and length has significantly increased. However, the model's prediction accuracy has decreased compared with the results from column A. Additionally, the significance level for all features has changed.

Table 6.5: Evaluation of feature selection methods based on the correlation ranks for each feature and length for the collected data from Run 1 and Run 2

Features	Mutual info	Pearson	Spearman's
max_coord_z	0.755	0.589	0.544
cent_coord_z	0.798	0.474	0.503
min_coord_x	0.582	0.459	0.406
min_coord_z	0.754	0.341	0.385
cent_coord_x	0.531	0.390	0.360
max_coord_x	0.537	0.302	0.274
reor_a	0.126	-0.255	-0.253
mesh_triang	0.189	0.302	0.202
mesh_points	0.191	0.302	0.202
surface	0.330	0.264	0.152
volume	0.339	-0.242	-0.138
max_coord_y	0.552	0.004	0.107
min_coord_y	0.541	-0.003	0.104
cent_coord_y	0.530	0.0004	0.103
reor_b	0.456	0.261	0.078
Run	0.0002	-0.009	-0.009
reor_c	0	NaN	NaN
material	0	NaN	NaN
bh	0	NaN	NaN
pvu	0	NaN	NaN
cnd	0	0	NaN
build_layout	0	NaN	NaN
MLP accuracy	83.23%	83.27%	91.63%

It is important to mention that material feature in this work is a numeric representation of the batches of virgin material (the material was assigned to 1 for Runs 1-2, while Run 3 has material annotation of 2, Run 4 as 2.1, Run 5 as 2.2, and Run 6 as 2.3). The numeric annotation of 2.1, 2.2, and 2.3 means that material batch is the same, but it was used on different days. Thus it was assumed that material properties might have a small variation due to climate conditions in the laboratory. In the future, these values should be changed to reflect the real material properties.

Column C in Table 6.6 corresponds to the feature selection results for runs 1-4, where 75 new specimens are added to the existed build layout in Run 4. This change has been reflected in the importance of the *Run, build layout, current nesting density (cnd), and platform volume utilization (pvu)* features. Only two features, namely build height and orientation by z-axis (reor_c in Table 6.6), have no correlation with the length, and therefore they are not considered as significant during the development of MLP models. All other features have been considered significant.

Column D corresponds to the feature selection results for runs 1-5, where the number of specimens were kept the same as for Run 4, but orientation and position have been changed for some of the specimens. Similar to the results described for runs 1-4 (column C), material properties have been uncontrolled, and this also has an impact on the results. Although, the batch for virgin material was the same, the main difference between material from Run 4 is related to the day when a mix of used and virgin powder was used, and thus, it is expected that Polyamide viscosity could have changed.

Even though the significance level for the *Run* feature has decreased, it is still among the features with the strongest correlation with length dimensional property. One of the reasons why these results are observed is related to changes conducted in the build layout design, and it is reflected on other features that describe different aspects of build layout design, namely pvu, cnd, and build_layout. The list of non-significant features is the same as for Column C since these features haven't been varied.

Results for feature selection for runs 1-6 show similar results to previous runs (Column E on Table 6.6. The main difference of Run 6 from other is the introduction of new orientation groups and completely different build layout design with different build height. Therefore, all features have correlation larger than 0, and there is no need for model comparison with and without feature selection results.

The prediction accuracy for the Column E has decreased significantly because in addition to unknown material properties, introduction of more than 18 orientation groups has led to an increase of noise and uncertanties in the data. In order to improve model performance there is a need for either repeating the last run one more time or collecting additional information about material and AM machine parameters.

Table 6.6: Feature selection for length dimensional property based on the collected data from all runs using Spearman's rho ranking method (where A is for runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6, and FS stays for feature selection)

Features	A	B	C	D	E
material	NaN	0.405	0.482	0.309	0.276
Run	-0.009	0.193	0.401	0.260	0.243
max_coord_z	0.544	0.400	0.323	0.279	0.227
build_layout	NaN	NaN	0.383	0.230	0.222
cent_coord_z	0.503	0.360	0.298	0.248	0.198
cnd	0	0	0.383	0.184	0.196
pvu	NaN	NaN	0.383	0.184	0.170
min_coord_z	0.385	0.275	0.248	0.209	0.164
surface	0.152	0.158	0.188	0.185	0.164
min_coord_x	0.406	0.313	0.218	0.155	0.132
volume	-0.138	-0.152	-0.159	-0.140	-0.117
bh	NaN	NaN	NaN	NaN	-0.105
cent_coord_x	0.359	0.273	0.192	0.129	0.103
mesh_triang	0.202	0.171	0.121	0.107	0.086
mesh_points	0.202	0.171	0.121	0.108	0.086
max_coord_x	0.274	0.207	0.156	0.098	0.072
reor_a	-0.253	-0.169	-0.058	-0.073	-0.068
cent_coord_y	0.103	0.048	0.039	0.057	0.046
max_coord_y	0.107	0.051	0.040	0.058	0.039
min_coord_y	0.104	0.050	0.042	0.057	0.027
reor_c	NaN	NaN	NaN	NaN	0.021
reor_b	0.078	0.097	0.026	0.004	-0.002
MLP accuracy	91.63%	73.60%	65.71%	71.33%	44.53%
MLP accuracy without FS	64.44%	65.72%	64.01%	68.93%	44.53%

Another interesting observation can be made for correlation ranks of coordinates within the same axis. For example, the correlation between maximal coordinate x and length is different from the correlation between central coordinate x and length. In this case, central coordinates have a larger correlation with length, while minimal coordinate x has an even larger cor-

relation with length. Similar behavior is observed for max, min, and central coordinates x and y. This result repeats from run to run. By the reason that actual measurements of length are done on either minimal or maximal coordinates of x/y/z-axis, this observation can have a meaning related to the physical process of additive manufacturing.

However, at this moment, it is not possible to compare obtained results with findings from other researchers because similar research hasn't been done or published. Therefore, there is a need for new experiments that can allow evaluating whether similar results are present for other powder bed AM systems.

6.2.2 Feature selection for width dimensional property

Feature selection for width dimensional property is conducted in the same way as for length, but with only one difference. Since width and thickness are measured differently for runs 1-2 and runs 3-6, feature selection is performed separately for each measurement area (Side A, Side B and Side C on Fig.4.10) for runs 3-6, and results for each side individually is combined with the data from runs 1-2.

Feature selection for Run 1 and Run 2

Selected features for the collected data from Run 1 and Run 2 are shown in Table 6.7, and Spearman's Rho test shows the best performance, among all the investigated methods. All methods have a similar set of features that have some correlation with width. Therefore, in order to be able to compare the methods' performance, the features that have absolute correlation rank higher than 0.03 have been considered as significant.

On the one hand, the Pearson correlation test could also be used for feature selection due its relatively high prediction accuracy of 78.52%, but it considers a linear correlation between two features. The Spearsman's correlation test, on the other hand, is chosen by the reason that it also takes into account the non-linear correlation between two parameters.

Table 6.7: Feature selection for width based on the collected data from Run 1 and Run 2

Features	Mutual info	Pearson	Spearman's rho
max_coord_z	0.674	0.177	0.187
cent_coord_z	0.626	0.245	0.190
min_coord_x	0.836	-0.299	-0.301
min_coord_z	0.649	0.301	0.200
cent_coord_x	0.761	-0.255	-0.268
max_coord_x	0.899	-0.197	-0.285
reor_a	0.306	0.766	0.539
mesh_triang	0.325	0.174	0.135
mesh_points	0.324	0.174	0.135
surface	0.484	0.265	0.180
volume	0.443	-0.299	-0.189
max_coord_y	0.777	0.165	0.247
min_coord_y	0.763	0.189	0.252
cent_coord_y	0.764	0.177	0.251
reor_b	0.466	-0.176	-0.064
Run	0	-0.038	-0.029
reor_c	0	NaN	NaN
material	0	NaN	NaN
bh	0	NaN	NaN
pvu	0	NaN	NaN
cnd	0	0	NaN
build_layout	0	NaN	NaN
MLP accuracy	74.55%	78.52%	80.76%

Comparison of feature selection for width side A and side B considering all runs based on Spearman's rho ranking method

Tables 6.8 - 6.9 show the results of feature selection based on the Spearsman's correlation test. The correlation ranks and prediction accuracies are similar for both sides and they are also similar for each column. Therefore, their analysis can be done together.

Table 6.8: Feature selection for width (Side A) dimensional property based on the collected data from all runs based on Spearsman's method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	**A**	**B**	**C**	**D**	**E**
max_coord_z	0.187	0.120	0.074	0.071	0.044
cent_coord_z	0.190	0.142	0.098	0.084	0.057
min_coord_x	-0.301	-0.229	-0.169	-0.127	-0.104
min_coord_z	0.200	0.169	0.119	0.093	0.070
cent_coord_x	-0.268	-0.190	-0.135	-0.107	-0.093
max_coord_x	-0.285	-0.182	-0.116	-0.098	-0.087
reor_a	0.539	0.445	0.350	0.251	0.169
mesh_triang	0.135	0.109	0.005	-0.015	-0.030
mesh_points	0.135	0.109	0.005	-0.014	-0.030
surface	0.180	0.154	0.151	0.174	0.172
volume	-0.189	-0.162	-0.117	-0.102	-0.101
max_coord_y	0.247	0.111	0.064	0.079	0.114
min_coord_y	0.252	0.116	0.069	0.080	0.061
cent_coord_y	0.251	0.115	0.068	0.078	0.088
reor_b	-0.064	-0.081	-0.109	-0.090	-0.120
Run	-0.029	0.232	0.482	0.434	0.474
reor_c	NaN	NaN	NaN	NaN	0.054
material	NaN	0.505	0.584	0.490	0.506
bh	NaN	NaN	NaN	NaN	-0.291
pvu	NaN	NaN	0.459	0.070	0.069
cnd	NaN	NaN	0.459	0.070	0.257
build_layout	NaN	NaN	0.459	0.410	0.459
MLP accuracy	80.76%	72.53%	59.43%	59.28%	56.06%
MLP accuracy without FS	80.07%	68.88%	63.81%	54.56%	56.06%

Table 6.9: Feature selection for width (Side B) dimensional property based on the collected data from all runs based on Spearsman's method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	A	B	C	D	E
max_coord_z	0.187	0.115	0.082	0.088	0.077
cent_coord_z	0.190	0.139	0.106	0.100	0.090
min_coord_x	-0.301	-0.229	-0.162	-0.124	-0.099
min_coord_z	0.200	0.166	0.125	0.107	0.105
cent_coord_x	-0.268	-0.190	-0.129	-0.101	-0.083
max_coord_x	-0.285	-0.176	-0.107	-0.086	-0.068
reor_a	0.539	0.442	0.345	0.274	0.199
mesh_triang	0.135	0.111	0.010	-0.003	-0.022
mesh_points	0.135	0.111	0.010	-0.002	-0.022
surface	0.180	0.147	0.155	0.194	0.184
volume	-0.189	-0.154	-0.119	-0.110	0.104
max_coord_y	0.247	0.107	0.064	0.066	0.116
min_coord_y	0.252	0.112	0.068	0.067	0.065
cent_coord_y	0.251	0.111	0.067	0.065	0.091
reor_b	-0.064	-0.087	-0.115	-0.103	-0.132
Run	-0.029	0.233	0.482	0.457	0.478
reor_c	NaN	NaN	NaN	NaN	0.058
material	NaN	0.507	0.585	0.512	0.511
bh	NaN	NaN	NaN	NaN	-0.276
pvu	NaN	NaN	0.458	0.047	0.046
cnd	NaN	NaN	0.458	0.047	0.230
build_layout	NaN	NaN	0.458	0.433	0.462
MLP accuracy	80.76%	73.68%	59.85%	60.06%	59.50%
MLP accuracy without FS	80.07%	67.60%	65.02%	56.12%	59.50%

As can be seen from Tables 6.8 - 6.9, material feature has the highest correlation rank starting from the Run 3, when the new material was introduced, and stays approximately at the same level throughout runs 3-6 for both sides. Similar behavior is observed for build layout feature. Column C describes the data from runs 1-4, and the build layout is changed for the first time for Run 4 by integrating 75 new specimens in the previous build

layout. This change has resulted in different values of build layout, current nesting density, and platform volume utilization features. Therefore, the significance of the listed features has increased. However, depending on the variations introduced at each new run, the correlation between platform volume utilization, build height, and current nesting density also varies.

Another important observation can be made regarding the correlation between part location, part orientation and width. When material and build layout design were kept the same, orientation around the x-axis was the most significant feature, followed by part's location in the build chamber and STL model properties. However, the significance of these features decreases when new variations are introduced. This phenomenon could mean that each variation in the material and build layout design leads to different values of width at the same location and with the same orientation.

Additionally, the prediction accuracy of 56% informs that the proposed features describe variations that take place as a result of the AM process only partly. This statement is supported by the finding from other researchers [12, 50], where they report that scanning strategy, powder morphology, layer thickness, and applied energy are other types of the parameters that influence the dimensional quality of fabricated parts.

Feature selection for width (Side C) considering all runs based on the Spearman's rho ranking method

The results for the correlation test between features listed in Table 6.10 and width measured in the central area of the specimen (side C) is completely different from the presented earlier results for width. The correlation between part's location and orientation decreases for each run, starting from the Run 3. Additionally, a high correlation coefficient can be observed for material feature, especially in Column B, where new material is used. The prediction accuracy reaches 99.92%, and stays very high for both five-fold cross-validated testing data (MLP accuracy in Table 6.10) and for models without feature selection. Thus, width on side C is correlated with the material, and build layout-related features that describe build layout in general rather than a detailed position of each sample individually. For instance, Fig. 6.18 illustrates an example where the measured width value is similar for different orientations and locations. Therefore, the results presented in the analysis of measured width and correlation test are similar. However, it is not clear why the correlation test results for side A and side B differ

from the ones for side C.

Table 6.10: Feature selection for width (Side C) dimensional property based on the collected data from all runs based on Spearsman's method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	A	B	C	D	E
max_coord_z	0.187	0.085	0.008	-0.010	-0.002
cent_coord_z	0.190	0.086	0.004	-0.023	-0.002
min_coord_x	-0.301	-0.140	-0.054	-0.020	-0.011
min_coord_z	0.200	0.095	0.004	-0.032	0.005
cent_coord_x	-0.268	-0.124	-0.051	-0.025	-0.009
max_coord_x	-0.285	-0.131	-0.067	-0.051	-0.022
reor_a	0.539	0.243	0.069	-0.016	0.003
mesh_triang	0.135	0.067	0.044	0.048	0.029
mesh_points	0.135	0.067	0.044	0.047	0.028
surface	0.180	0.090	0.024	0.013	0.017
volume	-0.189	-0.094	-0.047	-0.051	-0.021
max_coord_y	0.247	0.100	0.026	0.014	0.040
min_coord_y	0.252	0.103	0.025	0.018	0.040
cent_coord_y	0.251	0.102	0.026	0.016	0.040
reor_b	-0.064	-0.027	0.031	0.031	0.040
Run	-0.029	-0.422	-0.592	-0.550	-0.463
reor_c	NaN	NaN	NaN	NaN	-0.020
material	NaN	-0.816	-0.772	-0.657	-0.536
bh	NaN	NaN	NaN	NaN	0.138
pvu	NaN	NaN	-0.467	-0.023	-0.022
cnd	NaN	NaN	-0.467	-0.023	-0.114
build_layout	NaN	NaN	-0.467	-0.464	-0.399
MLP accuracy	80.76%	99.92%	99.91%	99.88%	99.85%
MLP accuracy without FS	80.07%	99.93%	99.90%	99.82%	99.85%

Table 6.11: Feature selection for thickness based on the collected data from Run 1 and Run2

Parameters	Mutual info	Pearson	Spearman's rho
max_coord_z	0.830	-0.035	-0.048
cent_coord_z	0.836	0.132	0.123
min_coord_x	1.023	-0.336	-0.252
min_coord_z	0.827	0.288	0.242
cent_coord_x	0.877	-0.208	-0.125
max_coord_x	0.973	-0.062	0.046
reor_a	0.126	-0.211	-0.230
mesh_triang	0.184	-0.309	-0.312
mesh_points	0.185	-0.309	-0.312
surface	0.276	-0.505	-0.503
volume	0.286	0.499	0.487
max_coord_y	0.919	0.167	0.300
min_coord_y	0.909	0.159	0.300
cent_coord_y	0.924	0.164	0.300
reor_b	0.795	-0.590	-0.566
Run	0.029	-0.052	-0.098
reor_c	0	NaN	NaN
material	0	NaN	NaN
bh	0.009	NaN	NaN
pvu	0	NaN	NaN
cnd	0	0	NaN
build_layout	0	NaN	NaN
MLP accuracy	95.44%	95.78%	96.33%

6.2.3 Feature selection for thickness dimensional property

The feature selection for thickness dimensional property is performed in the same ways as for length and width. Similarly, the thickness was measured at three different locations on the specimen (side A, side B, and side C), and therefore, the results for each side are presented separately.

Feature selection for Run 1 and Run 2

All three methods rank the significance of the investigated features a bit differently. For instance, in Table 6.11, one can see that the Mutual Information test highlights a strong correlation between thickness and minimal, maximal, and central coordinates in x, y, and z axes. In contrast, Pearson and Spearsman's correlation tests show a weak correlation between maximal coordinates in x and z axes and thickness. Therefore, the prediction accuracy of MLP models that are used for methods evaluation defines which method should be further used.

An additional adjustment needs to be made for Pearson and Spearsman's correlation tests. Since both methods have defined the same feature sets, an additional condition is created to separate those methods. Thus, features that have an absolute correlation coefficient higher than 0.09 are considered significant, and MLP models are evaluated for the sets of significant features. Spearsman's correlation test outperformed other methods and is used for feature selection for all runs.

Comparison of feature selection for thickness on Side A, side B and side C considering all runs based on Spearman's rho ranking method

Tables 6.12 - 6.14 show the results of correlation test for thickness measured on side A, side B and side C, and the results are similar for all sides. Therefore their analysis can be done simultaneously.

The results for the correlation test for thickness are different from the reported results for width and length. The main difference is in the significance of such features as material, build layout, and Run number. While the material is one of the most significant features for width, thickness measured at all locations has the strongest correlation with coordinates in the y-axis and orientation angle around the y-axis, but still, the correlation coefficient is weak compared with the value of correlation rank for the material feature for width.

Another important observation is related to the decrease in prediction accuracy when a new dataset is added to the analysis. Thus, when build layout and material properties are kept the same, the prediction accuracy accounts to 96%. However, as can be seen from Tables 6.12 - 6.14, changing to a new material has led to decrease in prediction accuracy from 96% to 55-65% for

side A and B, which are affected more than measurements on Side C. With the introduction of more variations in the experiment, correlation between the investigated features and thickness decreases. Similarly, the prediction accuracy of models also decreases.

Therefore, there should be some external factors that influence dimensional quality. The assumptions made for length and width dimensional properties are also applicable to the thickness, and they are described above (e.g., real material properties, maintenance of the powder bed system, and others).

Table 6.12: Feature selection for thickness dimensional property based on the collected data from all runs based on Spearman's rho ranking method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	A	B	C	D	E
max_coord_z	-0.048	0.001	0.016	0.028	0.005
cent_coord_z	0.123	0.119	0.093	0.075	0.054
min_coord_x	-0.252	-0.196	-0.160	-0.133	-0.142
min_coord_z	0.242	0.201	0.145	0.106	0.100
cent_coord_x	-0.125	-0.113	-0.103	-0.093	-0.104
max_coord_x	0.046	0	-0.021	-0.038	-0.054
reor_a	-0.230	-0.125	-0.114	-0.096	-0.068
mesh_triang	-0.312	-0.161	-0.153	-0.120	-0.125
mesh_points	-0.312	-0.161	-0.153	-0.121	-0.126
surface	-0.503	-0.277	-0.150	-0.070	-0.039
volume	0.487	0.266	0.159	0.072	0.080
max_coord_y	0.300	0.192	0.176	0.180	0.197
min_coord_y	0.300	0.194	0.175	0.181	0.155
cent_coord_y	0.300	0.193	0.176	0.179	0.178
reor_b	-0.566	-0.377	-0.328	-0.295	-0.275
Run	-0.098	-0.005	0.084	0.021	0.132
reor_c	NaN	NaN	NaN	NaN	0.062
material	NaN	0.111	0.129	0.045	0.147
bh	NaN	NaN	NaN	NaN	-0.174
pvu	NaN	NaN	0.096	0.072	0.064
cnd	NaN	NaN	0.096	0.072	0.169
build_layout	NaN	NaN	0.096	0.022	0.136
MLP accuracy	96.33%	66.28%	45.88%	41.86%	25.44%
MLP accuracy without FS	96.22%	65.97%	46.23%	42.34%	25.44%

Table 6.13: Feature selection for thickness dimensional property based on the collected data from all runs based on Spearman's rho ranking method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	**A**	**B**	**C**	**D**	**E**
max_coord_z	-0.048	0.009	0.016	0.021	0.009
cent_coord_z	0.123	0.135	0.098	0.075	0.064
min_coord_x	-0.252	-0.186	-0.172	-0.135	-0.149
min_coord_z	0.242	0.220	0.152	0.107	0.115
cent_coord_x	-0.125	-0.098	-0.110	-0.093	-0.108
max_coord_x	0.046	0.022	-0.019	-0.029	-0.050
reor_a	-0.230	-0.118	-0.116	-0.111	-0.057
mesh_triang	-0.312	-0.160	-0.167	-0.125	-0.133
mesh_points	-0.312	-0.160	-0.167	-0.125	-0.133
surface	-0.503	-0.284	-0.175	-0.095	-0.067
volume	0.487	0.272	0.183	0.087	0.097
max_coord_y	0.300	0.192	0.178	0.172	0.208
min_coord_y	0.300	0.194	0.177	0.173	0.164
cent_coord_y	0.300	0.195	0.178	0.171	0.188
reor_b	-0.566	-0.390	-0.356	-0.317	-0.293
Run	-0.098	-0.009	0.104	-0.001	0,099
reor_c	NaN	NaN	NaN	NaN	0.035
material	NaN	0.101	0.146	0.020	0.112
bh	NaN	NaN	NaN	NaN	-0.142
pvu	NaN	NaN	0.122	0.123	0.106
cnd	NaN	NaN	0.122	0.123	0.177
build_layout	NaN	NaN	0.122	0	0.102
MLP accuracy	96.33%	64.17%	46.18%	46.28%	31.29%
MLP accuracy without FS	96.22%	56.11%	45.76%	38.40%	31.29%

Table 6.14: Feature selection for thickness dimensional property based on the collected data from all runs based on Spearman's rho ranking method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	A	B	C	D	E
max_coord_z	-0.048	-0.039	-0.012	0.006	-0.004
cent_coord_z	0.123	0.099	0.090	0.075	0.061
min_coord_x	-0.252	-0.223	-0.201	-0.150	-0.150
min_coord_z	0.242	0.199	0.167	0.125	0.126
cent_coord_x	-0.125	-0.130	-0.125	-0.097	-0.098
max_coord_x	0.046	0.004	-0.016	-0.022	-0.030
reor_a	-0.230	-0.147	-0.086	-0.089	-0.042
mesh_triang	-0.312	-0.212	0.189	-0.133	-0.137
mesh_points	-0.312	-0.212	0.189	-0.134	-0.139
surface	-0.503	-0.356	-0.235	-0.166	-0.123
volume	0.487	0.343	0.249	0.153	0.157
max_coord_y	0.300	0.176	0.150	0.142	0.184
min_coord_y	0.300	0.177	0.151	0.146	0.154
cent_coord_y	0.300	0.177	0.151	0.145	0.172
reor_b	-0.566	-0.423	-0.383	-0.331	-0.308
Run	-0.098	0.022	0.104	-0.082	0.020
reor_c	NaN	NaN	NaN	NaN	0.046
material	NaN	0.111	0.159	-0.050	0.040
bh	NaN	NaN	NaN	NaN	-0.109
pvu	NaN	NaN	0.107	0.216	0.196
cnd	NaN	NaN	0.107	0.216	0.215
build_layout	NaN	NaN	0.107	-0.096	0.015
MLP accuracy	96.33%	85.09%	72.20%	61.34%	50.10%
MLP accuracy without FS	96.22%	84.40%	68.65%	60.20%	50.10%

6.3 Intelligent predictive models

While distribution-based analysis of measured dimensional properties is important to understand both run-to-run and orientation-to-orientation variations, the selection of significant features is valuable as a source of more

information about the correlation between dimensions and different build layout parameters. There is a need for predictive models that can be used for the estimation of possible geometric deviations in order to be able to compensate them before the AM fabrication.

As presented in Chapter 2.2, currently, researchers propose different models for compensating deviations through scaling ratios in x, y, and z axes for the whole build. However, the analysis of the collected data in this study shows that variation for length and width, especially along the z-axis, is different at different locations in the build chamber. Therefore, in this chapter predictive models for estimating geometric deviations for length, width and thickness individually for each specimen are proposed below, while Chapter 7 proposes predictive models for compensation ratios in x, y, and z axes separately for each part based on its location, orientation, and STL model properties.

In order to develop models with the best performance, grid search is used to investigate all possible combinations from the listed hyperparameters. Each method has its hyperparameters that need to be tuned. The learning rate for MLP is not optimized, but it is mentioned as one of the hyperparameters since MLP internally defines which learning rate should be used, and it is an important hyperparameter for the model.

The set of hyperparameters used in the grid search is the same for all mechanical properties. Thus, for MLP models, the following set of hyperparameters have been evaluated:

- hidden layer sizes: [19, 22, 25, 27, 35, 52],
- activation function:('relu', 'logistic'),
- solver function: ('lbfgs', 'sgd').

For Random Forest models, the following set of hyperparameters have been used for grid search:

- number of estimators: [50, 80, 100, 150, 200, 300, 500],
- maximum features :('auto', 'sqrt', 'log2').

The set of hyperparameters used for grid search for AdaBoost Regressor is the following:

- learning rate: [0.01, 0.001, 0.05, 0.1],
- loss function: ('exponential', ' square').

While a number of estimators for AdaBoost is tuned internally by the algorithm, the set of hyperparameters investigated for Gradient Boost Regressor is similar to the AdaBoost, but types of loss function are different due to the used algorithms, and are the following:

- number of estimators: [80, 100, 150, 200],
- learning rate: [0.01, 0.001, 0.05, 0.1],
- loss function: ('ls', 'lad', 'huber').

Depending on the number of hyperparameters in the lists, the model training time is different. For example, with the lists of hyperparameters proposed for MLP models, the training of a model including grid search and five-fold cross-validation has taken approximately 30 minutes, while Random Forest needed about 15 minutes to train a model and find the best hyperparameters of the model. Even though Gradient Boost Regressor has a large number of combinations of hyperparameters, similarly to MLP, the time need for grid search, and training a model with cross-validation was similar to the Random Forest, and hasn't exceeded 20 minutes.

6.3.1 Length dimensional property

The final combination of hyperparameters for each model, which are the result of the grid search, are presented in Table 6.15. In order to avoid overfitting and evaluate generalization capability, in addition to using the 85/15% ratio for training and testing datasets, an additional five-fold cross-validation on training dataset is used. It is also important to mention that the Multilayer Perceptron model is trained using all features based on the results of feature selection. While MLP during feature selection had prediction accuracy 44.53%, after optimizing hyperparameters, its prediction

accuracy has increased to 50.23% for cross-validation and 49.52% for test accuracy.

Even though MLP has outperformed ensemble methods such as AdaBoost and Gradient Boost Regressor, the Random Forest model has the best performance with test accuracy 53.57%. However, this prediction accuracy is still low, and such a model cannot be used yet, and requires improvements and developments in the future. Material description in this work should be changed to powder viscosity or another material-related property.

Table 6.15: Results of optimization of predictive models for length dimensional property

Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	relu	-	-	-
hidden layer	35	-	-	-
solver	lbfgs	-	-	-
learning rate	0.0001	-	0.1	0.05
loss	-	-	exponential	huber
n_estimators	-	500	50	100
max_features	-	auto	-	-
Accuracy	54.61%	55.14%	51.90%	55.41%
Cross-validation	50.23%	55.09%	51.52%	55.49%
Test accuracy	49.52%	53.57%	49.24%	46.81%

6.3.2 Width dimensional property

The predictive models for width are developed separately for each side, similarly to how feature selection was conducted. Table 6.16 shows results of the grid search with the best combination of hyperparameters for each model for width measured on side A. Table 6.17 shows comparison of predictive models for width measured on side B, and Table 6.18 presents the developed predictive models for width measured in central area of a specimen. Similarly to the feature selection for width, models' prediction accuracies for sides A and B are similar, while the prediction accuracies for side C is higher than 99.88%.

Table 6.16: Results of optimization of predictive models for width (side A) dimensional property

Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	relu	-	-	-
hidden layer	35	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.001	0.1
loss	-	-	exponential	ls
n_estimators	-	150	50	150
max_features	-	sqrt	-	-
Accuracy	56.08%	60.81%	51.84%	59.06%
Cross-validation	55.89%	60.25%	52.38%	59.02%
Test accuracy	50.52%	59.33%	49.16%	52.33%

Table 6.17: Results of optimization of predictive models for width (side B) dimensional property

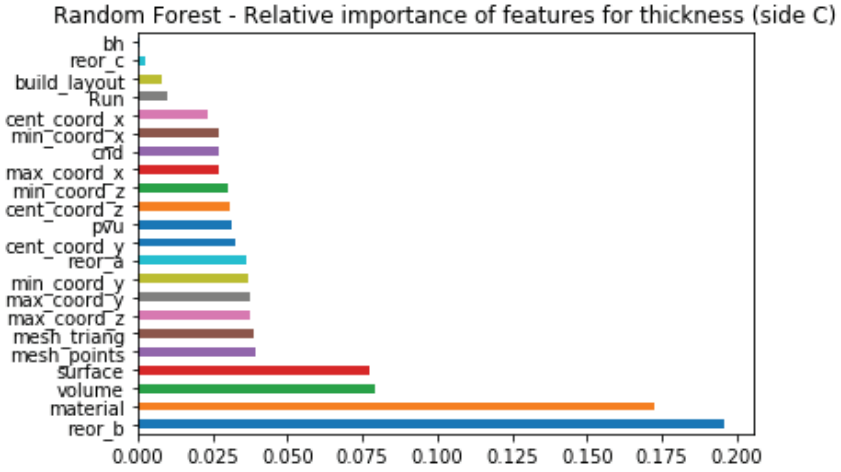
Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	relu	-	-	-
hidden layer	27	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.01	0.1
loss	-	-	square	ls
n_estimators	-	80	50	200
max_features	-	sqrt	-	-
Accuracy	58.44%	64.14%	53.60%	62.35%
Cross-validation	57.35%	64.14%	52.93%	61.80%
Test accuracy	51.70%	59.16%	49.17%	55.81%

It is important to mention that the developed models for width measured on

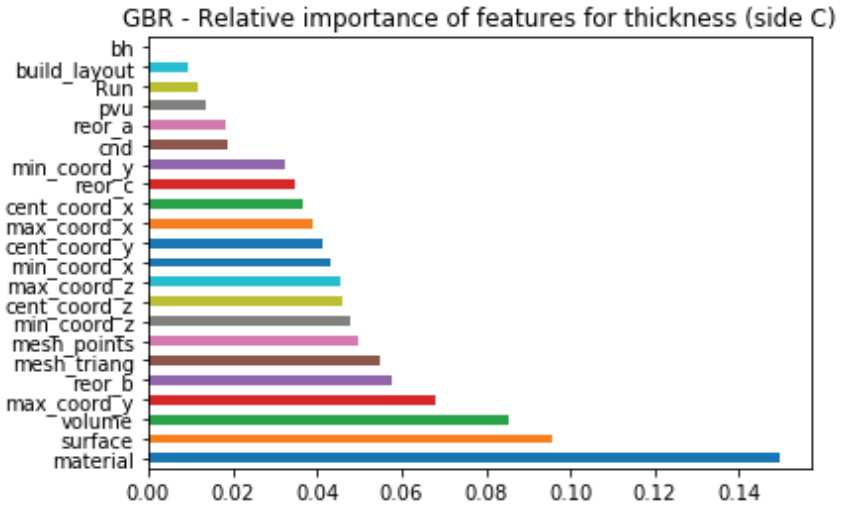
side A and side B have different architectures and hyperparameters. This means that different datasets can be described with different combinations of hyperparameters. Therefore, it is important to evaluate different combinations during the models' development, and grid search allows doing this process in an organised automated manner. Another benefit of using the grid search is improved prediction accuracy of the model due to the optimized architecture of the model.

Similarly to the models for length, Random Forest models have outperformed their counterparts, but accuracy of ca. 59% is still too low to be used in the production. Therefore, similar actions that were described for predictive models for length need to be done in the future.

All the optimized predictive models for width on side C have high prediction accuracy. By the reason that different types of model evaluation techniques are used (training, cross-validation, and testing), the model overfitting should not be considered as an issue. Although, if another type of geometry is introduced, there is a high risk that model performance will decrease to a large extent. Additionally, it is also interesting to compare the relative importance of features for two models with the same performances. Fig. 6.20 shows that even though two models have the same performance, the relative importance of features varies between the methods.



((a)) Random Forest



((b)) Gradient Boost Regressor

Figure 6.20: Relative importance of features for predictive models for width (side C)

Table 6.18: Results of optimization of predictive models for width (side C) dimensional property

Model's parameter	MLP FS=No FS	Random Forest	AdaBoost	GBR
activation	relu	-	-	-
hidden layer	52	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.1	0.1
loss	-	-	exponential	ls
n_estimators	-	300	50	200
max_features	-	sqrt	-	-
Accuracy	99.89%	99.90%	99.87%	99.90%
Cross-validation	99.87%	99.90%	99.87%	99.90%
Test accuracy	99.85%	99.88%	99.84%	99.88%

Any machine learning model proposed for width (side C) can be used in the future, but due to the reason that Random Forest is among the most robust models and requires less time for training, it is recommended to choose this model among the other proposed.

6.3.3 Thickness dimensional property

The developed predictive models for thickness have similar results as it was described in the section about feature selection. Table 6.19 shows the results of grid search for thickness measured on side A. Table 6.20 demonstrates the developed models for thickness measured on side B, while Table 6.21 presents the predictive models for thickness measured on side C. The models from four proposed machine learning methods are optimized and evaluated based on three types of prediction accuracies.

Table 6.19: Results of optimization of predictive models for thickness (side A) dimensional property

Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	logistic	-	-	-
hidden layer	25	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.05	0.1
loss	-	-	exponential	ls
n_estimators	-	250	50	200
max_features	-	sqrt	-	-
Accuracy	42.96%	42.76%	23.24%	40.42%
Cross-validation	38.62%	42.59%	23.65%	39.59%
Test accuracy	14.52%	36.02%	19.72%	24.26%

Similarly to low prediction accuracy observed at the stage of feature selection, the models have prediction accuracy 34-36% for side A and B, and 49% for side C even after defining the best (among proposed) combinations of the hyperparameters. Even though Random Forest outperformed the other machine learning methods for thickness measurements, the model's architecture for each side is different. Therefore, it could be assumed that there is a difference in datasets that affect the choice of hyperparameters of the model.

Table 6.20: Results of optimization of predictive models for thickness (side B) dimensional property

Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	logistic	-	-	-
hidden layer	19	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.05	0.1
loss	-	-	exponential	huber
n_estimators	-	200	50	200
max_features	-	auto	-	-
Accuracy	38.39%	43.19%	22.06%	38.95%
Cross-validation	34.84%	42.87%	21.70%	38.65%
Test accuracy	15.09%	34.24%	18.23%	28.21%

Table 6.21: Results of optimization of predictive models for thickness (side C) dimensional property

Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	relu	-	-	-
hidden layer	27	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.1	0.1
loss	-	-	exponential	huber
n_estimators	-	300	50	200
max_features	-	auto	-	-
Accuracy	51.75%	51.20%	36.63%	48.76%
Cross-validation	41.50%	51.09%	36.46%	49.10%
Test accuracy	41.36%	49.88%	34.95%	32.10%

Since most of the proposed models for all dimensional properties have low

prediction accuracy, and only material and build layout features have been varied from run to run, it is assumed that considering each dimension isolated from each other could result in the loss of some information. Besides, length, thickness, and width measurements physically describe one sample and could be correlated with each other. Therefore, in the next section, multi-output learning is applied for prediction of length, width, and thickness simultaneously.

6.3.4 Geometric deviations as a vector of length, width, and thickness

Resulting prediction of each dimensional property is presented above, and as it can be seen, the accuracy is not high enough. Since one of the possible reasons for that is not enough numerical information about factors that have an impact on the output, the multi-output learning is of interest. The multi-output learning is used when the prediction of two or more outputs needs to be done simultaneously, i.e. as a vector.

Among the investigated machine learning techniques, MLP and Random Forest allow predicting more than one output simultaneously. Therefore, these ML methods have been used to predict length, width, and thickness simultaneously, and their results are compared in Table 6.22.

Table 6.22: Optimized hyperparameters of predictive models for length, width and thickness simultaneously

Model's hyperparameters	**MLP (FS=No FS)**	**Random Forest**
activation	relu	-
hidden layer	52	-
solver	lbfgs	-
learning rate	0.001	-
n_estimators	-	500
max_features	-	auto
Accuracy	99.28%	99.329%
Cross-validation	99.26%	99.328%
Test accuracy	98.64%	99.16%

The grid search optimization is used to determine the hyperparameters of the predictive models. Grid search automatically tries all possible combinations from the sets of hyperparameters provided by the user, and the model with the combination yielding the best performance is provided for the user (see Table 6.22).

The sets of hyperparameters used in the grid search optimization for the MLP neural network are the following:

- size of hidden layer: [19, 22, 25, 27, 35, 52],
- activation function: ('relu', 'logistic'),
- solver function: ('lbfgs', 'sgd').

The set of hyperparameters used in the grid search optimization for the Random Forest method is the following:

- number of estimators: [80, 100, 150, 200, 250, 300, 500],
- maximum features: ('auto', 'sqrt', 'log2').

Additionally, the model's performance is verified through five-fold cross-validation. The scores from cross-validation for MLP neural network and random forest models are shown in Table 6.23.

Table 6.23: Comparison of five-fold cross-validation scores for MLP and Random Forest machine learning methods

ML method	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Multilayer Perceptron	99.34%	99.30%	99.43%	99.17%	99.08%
Random Forest	99.34%	99.31%	99.47%	99.24%	99.28%

As can be seen from Tables 6.22 - 6.23, both methods have similar prediction accuracy, but Random Forest is more stable to datasets variations. The difference between prediction accuracies at each fold for Random Forest is smaller than for the Multilayer Perceptron model. Additionally, the Random Forest model has a better performance than the MLP model on new unseen data (see test accuracy in Table 6.22).

Moreover, comparison of predicted values by MLP and Random Forest with observed (measured) values is shown on Fig. 6.21 - 6.23. Even though 229 specimens have been used for testing models performance, only 17 randomly chosen are shown on the Fig. 6.21 - 6.23.

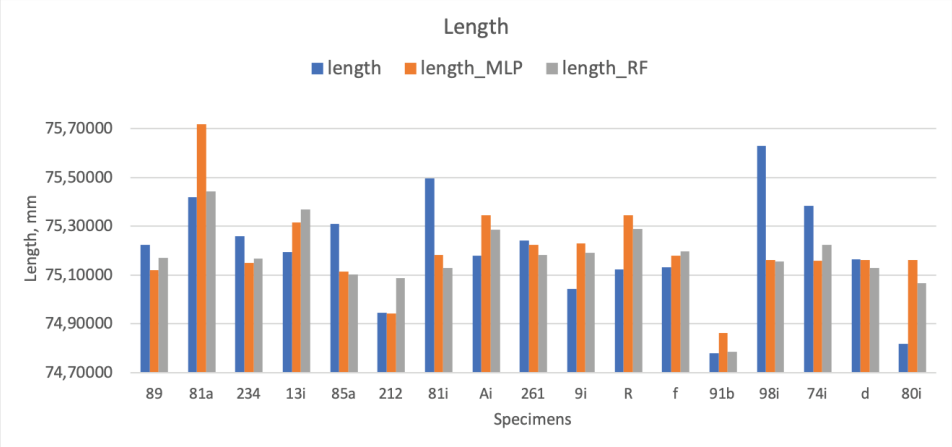


Figure 6.21: Comparison of observed with predicted length values

As can be seen from Fig 6.21, both models have very similar results, but for the specimen with index 81a, the MLP model has a large deviation comparing to other samples. However, for Random Forest, such behavior is not observed.

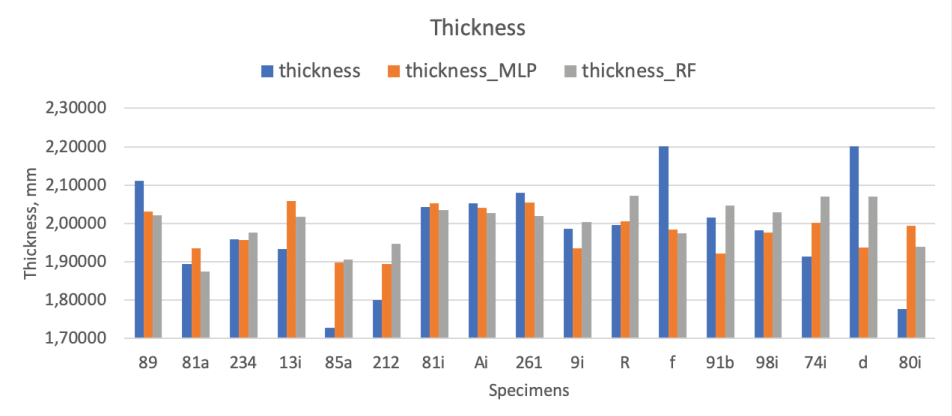


Figure 6.22: Comparison of observed with predicted thickness values

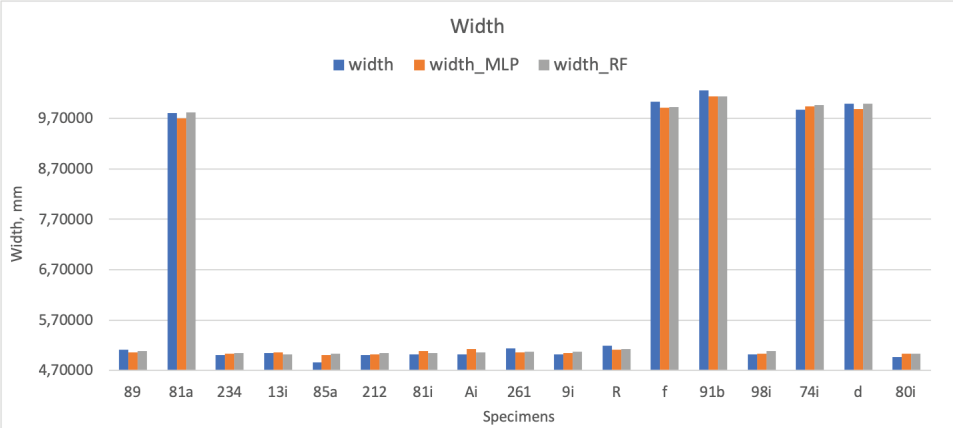


Figure 6.23: Comparison of observed with predicted width values

Similarly, Random Forest also results in a better prediction accuracy for thickness and width values. When comparing the difference between the predicted and observed values for length, both models, as can be seen from Fig. 6.22 have a small difference between the observed and predicted thickness values. The predicted values for width has the smallest deviation from the observed width values.

Thus, it can be seen that both methods can be used for prediction of geometric deviations, but the Random Forest model is both more robust and more accurate than the MLP model.

Additionally, it is important to highlight that considering dimensional properties as a vector has resulted in better model performance. This is caused by the reason that datapoints representation of dimensions has moved from $\mathbb{R}$ to $\mathbb{R}^3$. There is a possibility that data points in $\mathbb{R}^3$ have a stronger correlation with other features that are represented as some patterns that are captured by machine learning models.

Chapter 7

Module P2: Prediction of compensation ratio

Even though in Chapter 6 geometric deviations are predicted for each specimen separately, the obtained results need further transformation for compensating the deviations. Therefore, this chapter aims at describing the developed models for estimation of compensation ratios in x, y, and z axes for each part separately.

7.1 Preliminaries

A compensation ratio for each dimension is calculated as follows:

$$cr_{i,j} = \frac{C_{i,j}^{CAD}}{C_{i,j}^{mea}} \tag{7.1}$$

where cr represents compensation ratio, $i \in \{1, 2, 3\}$ corresponds for length, width and thickness, $j \in \{1, 2, .., k\}$ describes index of a part, and k stays for the number of parts, $C_{i,j}^{CAD}$ represents designed (CAD) dimensional value, and $C_{i,j}^{mea}$ describes a measured dimensional value.

In order to use the defined compensation ratio in practice, the operator needs to multiply dimensions from the CAD model by corresponding compensation ratio. In software used to design a build layout, this option is available for each specimen and thus can be performed as the last step of the build

layout design process. However, in the future, one of the tasks is to expand the proposed modules by an additional service for automated compensation of the predicted deviations performed without human interaction. At this moment, due to time limitations, this task is out of the scope of the current work.

The feature list is the same as in Chapter 6, and compensation ratios for length, width, and thickness are considered to be the outputs. By following the proposed data pipeline in Fig. 5.7 from Chapter 5, the data was already preprocessed, and only feature selection needs to be conducted as the target is different.

7.2 Feature selection

Mutual information, Pearson and Spearman's correlation test are applied only to the final dataset containing a joined data from all six experiments, and comparison of the obtained results is shown on Tables 7.1, 7.2, and 7.3 for compensation ratios for length, width and thickness respectively.

The correlation ranks for Spearsman's method described on Table 7.1 are similar to the correlation ranks obtained for length in Chapter 6. However, such features as material, Run, build layout, build height, platform volume utilization, current nesting density, and maximal, minimal, and central coordinates z have higher correlation rank with the compensation ratio for length than with measured length values. All other correlation ranks are similar in both cases.

Besides, such features as number of mesh triangles, mesh points, and orientation angles by x and z axes have similar correlation ranks for different correlation tests. The correlation ranks are still weak, but it can be seen that the prediction models still can find patterns in the data despite the weak correlation (prediction accuracy of Random Forest model for dimensions as a vector is ca. 99.20%).

Therefore, the main interest in the feature selection process is to determine whether there are features with zero correlation. As can be seen from Table 7.1, all features have the correlation coefficient greater than 0, and therefore all of them will be used for model training.

Table 7.1: Comparison of feature selection methods based on the correlation ranks for each feature and compensation ratio for length for the collected data from all runs

Features	Mutual info	Pearson	Spearman's
material	0.330	-0.277	-0.396
cent_coord_x	0.172	-0.088	-0.108
min_coord_x	0.169	-0.109	-0.130
max_coord_x	0.119	-0.066	-0.082
cent_coord_y	0.120	-0.038	-0.034
min_coord_y	0.115	-0.033	-0.027
max_coord_y	0.102	-0.046	-0.042
cent_coord_z	0.151	-0.308	-0.281
min_coord_z	0.159	-0.273	-0.234
max_coord_z	0.155	-0.332	-0.317
mesh_triang	0.030	-0.069	-0.076
mesh_points	0.032	-0.069	-0.076
surface	0.033	-0.168	-0.154
volume	0.045	0.104	-0.043
reor_a	0.076	0.053	0.073
reor_b	0.124	0.020	-0.019
reor_c	0.029	-0.027	-0.018
Run	0.180	-0.246	-0.271
build_layout	0.209	-0.226	-0.210
bh	0.045	0.114	0.114
pvu	0.122	-0.170	-0.004
cnd	0.211	-0.201	-0.032

The correlation between features and width (side C, in this case, represents combination of 10 and 5 mm) has similar results as shown in Chapter 6 and Table 7.2. The material feature has the strongest correlation based on all filtering methods. However, Pearson and Spearsman's correlation tests provide more similar correlation ranks between each other than with the Mutual information method. All features have correlation ranks greater than 0, and therefore, all should be considered during modeling of the compensation ratio for width.

Table 7.2: Comparison of feature selection methods based on the correlation ranks for each feature and width compensation ratio for the collected data from all runs

Features	**Mutual info**	**Pearson**	**Spearman's**
material	0.339	-0.572	-0.554
cent_coord_x	0.175	0.080	0.077
min_coord_x	0.163	0.091	0.085
max_coord_x	0.200	0.073	0.066
cent_coord_y	0.208	-0.107	-0.117
min_coord_y	0.213	-0.077	-0.093
max_coord_y	0.206	-0.136	-0.146
cent_coord_z	0.172	-0.063	-0.079
min_coord_z	0.208	-0.074	-0.092
max_coord_z	0.186	-0.060	-0.063
mesh_triang	0.069	0.053	0.063
mesh_points	0.067	0.053	0.063
surface	0.085	-0.202	-0.183
volume	0.073	0.095	-0.051
reor_a	0.103	-0.160	-0.137
reor_b	0.132	0.126	0.136
reor_c	0.052	-0.059	-0.029
Run	0.219	-0.543	-0.501
build_layout	0.207	-0.533	-0.483
bh	0.062	0.311	0.281
pvu	0.077	0.026	0.196
cnd	0.204	-0.204	0.124

The correlation ranks shown in Table 7.3 for thickness are also similar to the ones presented in Chapter 6, and all correlation coefficients are greater than 0. Therefore, all features need to be considered during the development of the models for compensation ratio for thickness. Since correlation tests' results are similar for the compensation ratio and actual dimensional properties (from Chapter 6), it means that the compensation ratio is derived correctly, and correlation can be described in the same way as for dimensional properties in Chapter 6.

Table 7.3: Comparison of feature selection methods based on the correlation ranks for each feature and thickness compensation ratio for the collected data from all runs

Features	Mutual info	Pearson	Spearman's
material	0.315	-0.040	-0.049
cent_coord_x	0.164	0.098	0.096
min_coord_x	0.183	0.150	0.150
max_coord_x	0.159	0.030	0.037
cent_coord_y	0.199	-0.172	-0.147
min_coord_y	0.201	0.154	-0.130
max_coord_y	0.184	-0.184	-0.161
cent_coord_z	0.231	-0.061	-0.070
min_coord_z	0.174	-0.125	-0.132
max_coord_z	0.237	0.004	-0.005
mesh_triang	0.026	0.137	0.133
mesh_points	0.026	0.139	0.135
surface	0.026	0.123	0.132
volume	0.089	-0.157	-0.013
reor_a	0.100	0.042	0.088
reor_b	0.207	0.308	0.307
reor_c	0.036	-0.046	-0.052
Run	0.214	-0.020	0.004
build_layout	0.171	-0.015	0.012
bh	0.047	0.109	0.098
pvu	0.120	-0.196	-0.224
cnd	0.170	-0.215	-0.231

Based on the obtained results from feature selection, all of the listed features are used for the development of predictive models for compensation ratio.

7.3 Predictive models for compensation ratio

The process of model development describe earlier is also applied for this task. First models are developed separately for compensation ratio that corresponds only to one-dimensional properties, and then as a $\mathbb{R}^3$ vector. The obtained results for individual models are compared with the model proposed for compensation ratios for the dimensions in $\mathbb{R}^3$.

In order to develop models with the best performance, grid search is also used to define the best combinations of the model hyperparameters from the listed ones. The sets of hyperparameters for each method differ from the hyperparameters presented in Chapter 6, and therefore, updated sets of hyperparameters for each ML method is described below. As such, the sets of hyperparameters used in the grid search for MLP models is the following:

- hidden layer sizes: [22, 25, 27, 35, 52, 55],
- activation function:('relu', 'logistic'),
- solver function: ('lbfgs', 'sgd').

For Random Forest models, the following set of hyperparameters have been used for grid search:

- number of estimators: [150, 200, 250, 300, 500, 600],
- maximum features :('auto', 'sqrt', 'log2').

The set of hyperparameters used for grid search for AdaBoost Regressor is the following:

- learning rate: [0.01, 0.001, 0.05, 0.1],
- loss function: ('exponential', ' square').

The hyperparameters of the Gradient Boost Regressor model are defined from the following set:

- number of estimators: [80, 100, 150, 200],
- learning rate: [0.01, 0.001, 0.05, 0.1],
- loss function: ('ls', 'lad', 'huber').

The evaluation of the model's performance is done at three different levels. First of all, all data is divided into two datasets, namely training and testing, with a ratio of 85/15 %, respectively. During the training process, the training accuracy is calculated for each model (accuracy is a coefficient of determination - R^2). Second of all, five-fold cross-validation on the training dataset is performed as the second stage of the model's performance evaluation. Finally, the model with the highest five-fold cross-validation result is evaluated on the testing dataset, and testing accuracy represents the obtained results. The main goal is to define a model, which has either the most similar training/cross-validation/testing accuracies or the one with the highest testing accuracy.

7.3.1 Predictive models for compensation ratio for length

As a result of grid search and cross-validation, Table 7.4 shows the final models' description in terms of the combination of hyperparameters chosen by grid search for compensation ratio (length). Among all investigated machine learning methods, the Gradient Boost Regressor has the highest testing accuracy (65.87%), while the Random Forest model has higher training and cross-validation accuracies. However, as it was already mentioned earlier, the preference should be given to the method with the higher testing accuracy.

7.3.2 Predictive models for compensation ratio for width

Among all models presented in Table 7.5, the Random Forest model has the best performance, and the accuracy level, on average, is about 58.5%. However, this value is lower than for the compensation ratio of length.

Table 7.4: Results of optimization of predictive models for length compensation ratio dimensional property

Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	relu	-	-	-
hidden layer	22	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.01	0.1
loss	-	-	exponential	huber
n_estimators	-	600	50	80
max_features	-	auto	-	-
Accuracy	22.66%	66.21%	56.86%	63.65%
Cross-validation	24.03%	66.18%	55.84%	63.28%
Test accuracy	30.14%	61.84%	54.77%	65.87%

Table 7.5: Results of optimization of predictive models for compensation ratio for width dimensional property

Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	relu	-	-	-
hidden layer	52	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.01	0.1
loss	-	-	exponential	huber
n_estimators	-	500	50	150
max_features	-	sqrt	-	-
Accuracy	45.47%	58.91%	45.38%	55.36%
Cross-validation	44.17%	58.33%	45.11%	54.88%
Test accuracy	40.21%	58.60%	43.26%	56.73%

7.3.3 Predictive models for compensation ratio for thickness

According to the results illustrated in Table 7.6, the Random Forest model for compensation ratio of thickness has the highest training, cross-validation, and testing accuracies. The prediction accuracy is similar to the one observed for the compensation ratio of the width, though 3% lower.

However, comparing with the predictive models for thickness described in Chapter 6, the predictive model for compensation ratio of thickness has higher prediction accuracy, while for other dimensional features, the opposite is true. This observation means that redefining tasks differently can also help to develop models with better performance by using the same data.

Table 7.6: Results of optimization of predictive models for compensation ratio for thickness dimensional property

Model's hyperparameter	**MLP FS=No FS**	**Random Forest**	**AdaBoost**	**GBR**
activation	relu	-	-	-
hidden layer	22	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.1	0.1
loss	-	-	exponential	ls
n_estimators	-	600	50	150
max_features	-	auto	-	-
Accuracy	50.49%	55.74%	39.06%	52.96%
Cross-validation	48.98%	55.64%	39.35%	53.66%
Test accuracy	53.28%	55.82%	38.44%	52.06%

7.3.4 Predictive models for compensation ratios in $\mathbb{R}^3$

By the reason that in Chapter 6 the model developed for dimensions in $\mathbb{R}^3$ has improved prediction accuracy from 40-60% for length and thickness to 99.16%, it is assumed that similar behavior could be observed the for compensation ratio.

Therefore, multilayer perceptron and random forest are chosen for this task

since they can provide a model with multi-output prediction. Table 7.7 shows the selected combinations of parameters for each model as a result of grid search and five-fold cross-validation. Even though in terms of training and cross-validation accuracies, the random forest has outperformed MLP, their testing accuracy is similar. In order to understand which model is better, the predicted values are compared with the nominal values.

Fig. 7.1, 7.2 and 7.3 illustrates the nominal and predicted values of compensation ratio of length, width and thickness respectively. The comparison is made by using 17 randomly selected specimens from the testing dataset. Through visual comparison of predicted values by MLP and Random Forest, the one can observe that the latter one performs better than MLP.

However, an additional metric needs to be introduced because the difference between the predicted and nominal values is small (even though the prediction accuracy is not as high as for dimensional deviations). Therefore, the mean squared error (MSE) is used as an additional metric of models' performance. One can be observe that the value of MSE is also similar for both models (see Table 7.7).

Table 7.7: Optimized hyperparameters of predictive models for compensation ratio of length, width and thickness simultaneously

Model's hyperparameters	**MLP (FS=No FS)**	**Random Forest**
activation	relu	-
hidden layer	25	-
solver	lbfgs	-
learning rate	0.001	-
n_estimators	-	200
max_features	-	auto
Accuracy	51.86%	57.66%
Cross-validation	50.35%	57.72%
Test accuracy	47.17%	47.14%
MSE	0.0006938	0.0006941

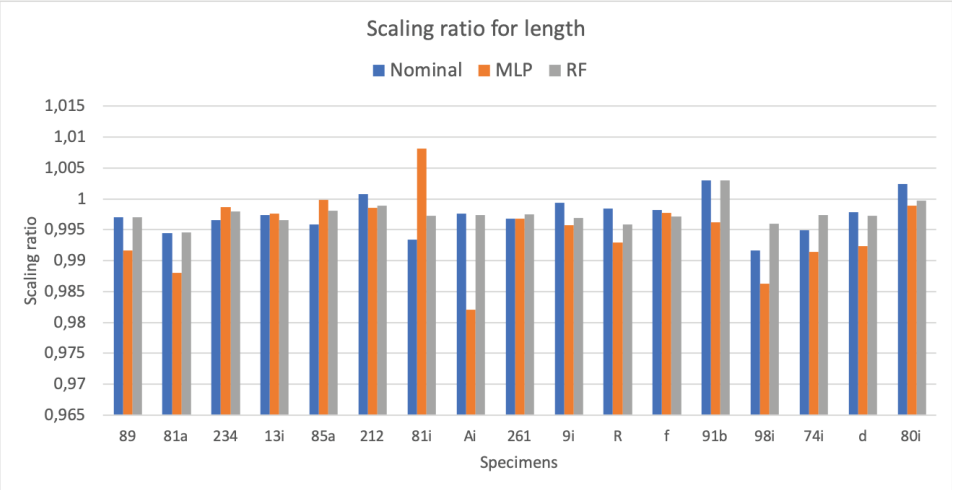


Figure 7.1: Comparison of observed with predicted length compensation ratio values

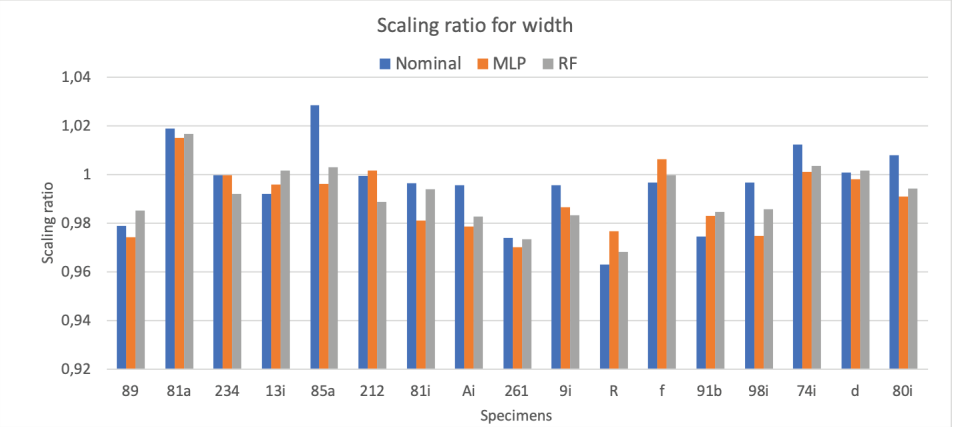


Figure 7.2: Comparison of observed with predicted width compensation ratio values

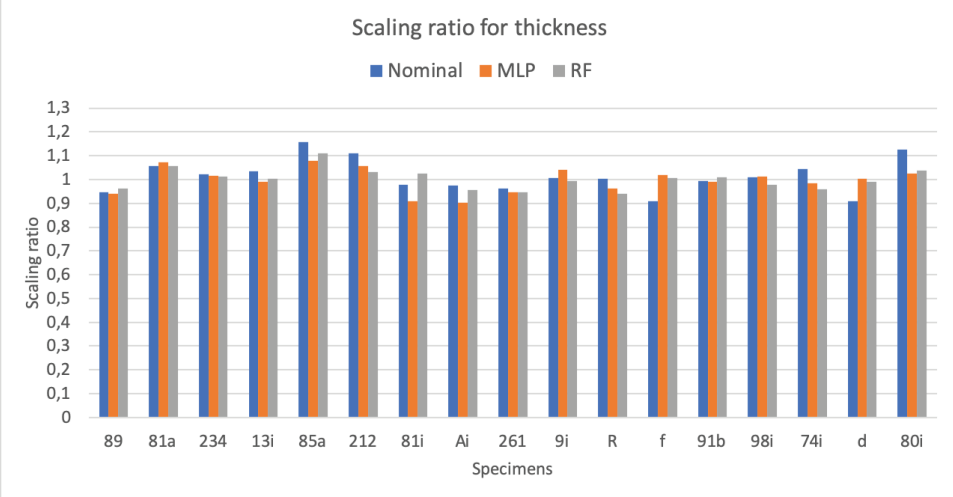


Figure 7.3: Comparison of observed with predicted thickness compensation ratio values

Since the predicted values and nominal values, especially for the random Forest model, are very close and MSE is equal to 0.00069, it is assumed that models predicting a scalar compensation ratio have higher prediction accuracies than the model for compensation ratio in $\mathbb{R}^3$.

The results show that models predicting a scalar compensation ratio have smaller mean squared errors, and therefore should be used in the future. Additionally, it is important to mention that relatively small prediction accuracy does not fully describe the correct model performance in the case when low numerical values need to be predicted.

Chapter 8

Module P12: Prediction of mechanical properties

This chapter describes how the predictive models for estimation of mechanical properties were developed. The chapter consists of three main sections. The first section describes the collected data. The second section provides information about the selection of significant features. The chosen features are used for the development of predictive models. The obtained models are described and compared in the last section.

8.1 Data exploration

It is generally adviced to use the central part of the build chamber in order to produce products with good mechanical properties. Such underutilization of build space leads to the increased price per product, longer time-to-market for larger batches, and increased amount of degradation and waste of material. Therefore, the analysis of the mechanical properties in different areas of the build chamber is investigated in this section.

According to Caulfield et al. [22], mechanical properties depend on the part orientation. Therefore, the description of the mechanical properties should be done separately for each orientation. The four orientations, which are XYZ, XZY, ZYX, and Angle (45° between x and z axes), have been analyzed for runs 1-5, and analysis of Run 6 has been done separately due to the difference in the built layout design (refer to Chapter 4 for a detailed

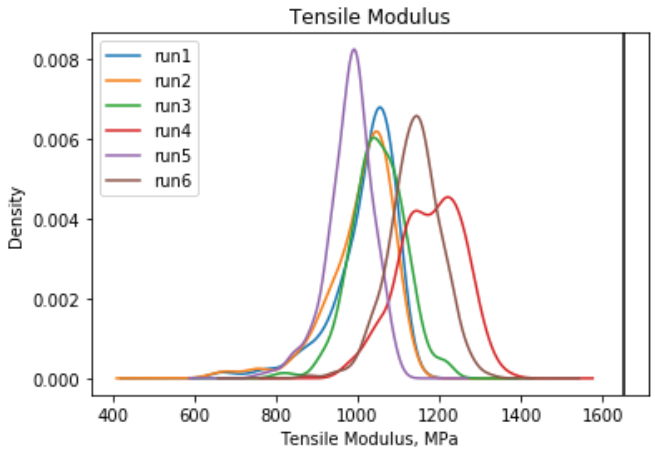


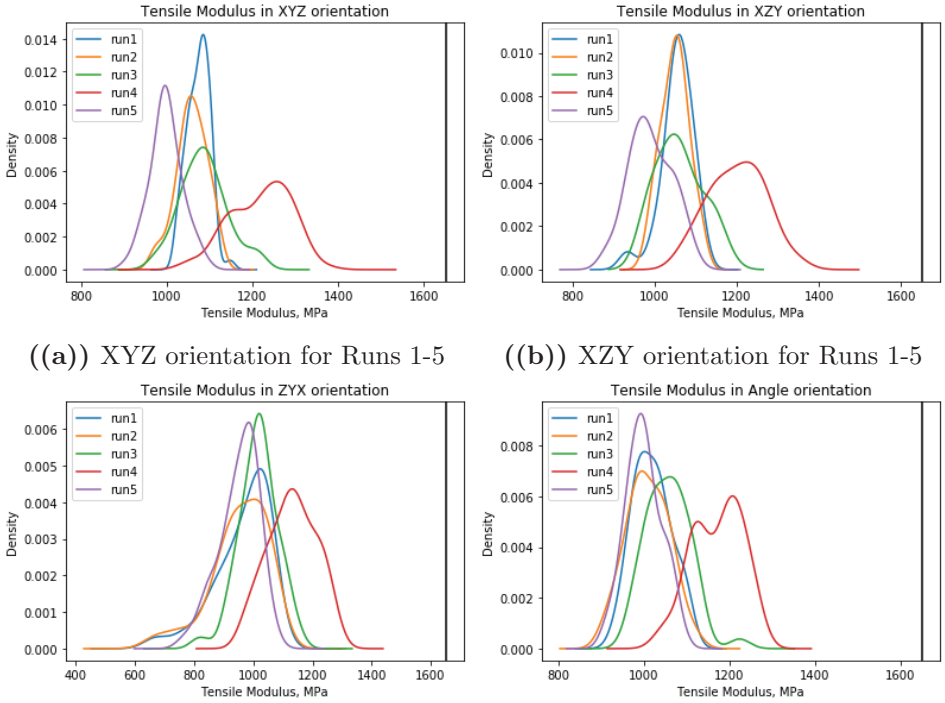
Figure 8.1: Distribution of tensile modulus for different runs based on kernel distribution estimation (The straight line (1650 MPa) corresponds to the value from EOS Balanced datasheet)

description).

8.1.1 Analysis of tensile modulus

The result of kernel density estimation of tensile modulus for all runs is illustrated in Fig.8.1. All data is distributed normally except for the results obtained from Run 4, which has multimodal distribution. This result could be caused by the different build orientation, or due to the specimens' placement variations in the build chamber. However, Fig.8.2 shows that tensile modulus distributions for different orientations have the same trend as shown in Fig.8.1. The only difference is observed for the Run 4, where data is distributed normally in XZY and ZYX orientations. In contrast, distributions for XYZ and Angle orientations resemble multimodal distributions, and the one for Angle orientation looks similar to the distribution for Run 4 shown in Fig.8.1. This phenomenon is described with more details later in this section.

Another concern related to the material properties that are usually reported for polyamide is anisotropy. Anisotropic behavior of PA12 has been reported earlier by [51, 59]. The results illustrated in Fig.8.2 do not support this statement for tensile modulus, which is similar to the datasheets provided by EOS for balanced machine settings for PA12 material (EOS is both a



((a)) XYZ orientation for Runs 1-5

((b)) XZY orientation for Runs 1-5

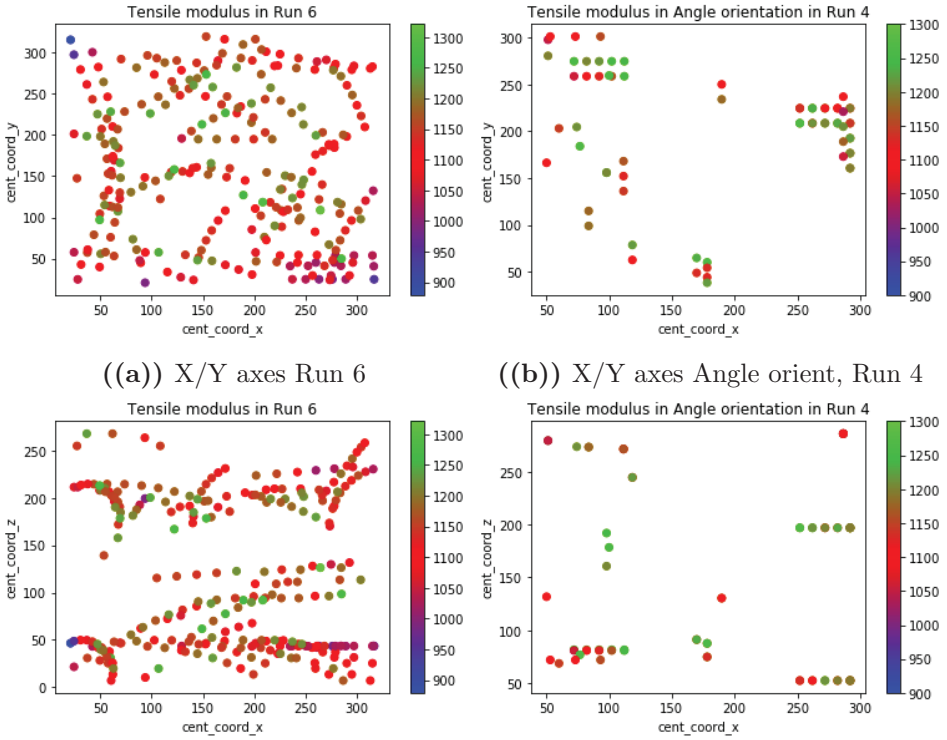
((c)) ZYX orientation for Runs 1-5

((d)) Angle orientation for Runs 1-5

Figure 8.2: Distribution of tensile modulus for different orientations based on kernel distribution estimation (The straight line (1650 MPa) corresponds to the value from EOS Balanced datasheet)

supplier of the material and producer of AM machine, which have been used in the experiments). However, none of the runs have a maximum or modal tensile modulus that corresponds to 1650 MPa, which is the nominal value from the EOS datasheet. For example, maximum values for Run 4 and Run 6 are 1372 MPA and 1320 MPa, correspondingly.

Orientation-based description of the data collected from Run 6 is not compared with other five runs because it consits of more than four orientations shown in Fig.8.2. Due to a large number of orientation groups, distribution of tensile modulus for Run 6 is illustrated in Fig.8.3 as a color map. Fig.8.3(a) and Fig.8.3(b) show how measured tensile modulus for all specimens in Run 6 and Run 4 (Angle orientation) are distibuted in x and y axes, while distribution of measured tensie modulus along z axis is illustrated in



((a)) X/Y axes Run 6

((b)) X/Y axes Angle orient, Run 4

((c)) X/Z axes Run 6

((d)) X/Z axes Angle orient. in Run 4

Figure 8.3: Distribution of the tensile modulus in the build chamber with regards to the coordinates for Run 6 and Angle orientation in Run 4

Fig.8.3(c) (for Run6) and Fig.8.3(d) (for Angle orientation Run 4).

One can observe in Fig.8.3 that the tensile modulus values for Run 6 range between ca. 900 MPa and 1320 MPa, while the majority of values are in the range of 1000-1200 MPa. The weakest specimens are positioned in the corners, but it is challenging to see some patterns related to the 18 orientation groups. Additionally, strong specimens can also be observed close to the borderlines (e.g., coordinates for strong specimens in x and y axes are (50, 200) or (280, 50)) of the chamber for both Run 6 and Run 4.

As it was mentioned earlier, tensile modulus distribution for Run 4 has two peaks, and Angle orientation has similar data distribution. In order to try to understand what could be a reason for this behavior, a color map that

corresponds to the part positioning in the build chamber is shown in Fig.8.3. There is no linear dependence between the position and tensile modulus. In some locations, both relatively strong and weak specimens are observed. Although, data in Angle orientation for Run 4 could be clustered in two groups along the x-axis. However, it is only one of the assumptions, and a more thorough statistical analysis is needed to get a better understanding of how the part placement correlates to the tensile modulus.

8.1.2 Analysis of tensile strength

Similarly to the tensile modulus, distributions of tensile strength for all runs are illustrated in Fig.8.4. In addition to the distributions based on the kernel density estimation, values in XYZ (solid line) and ZYX (dashed line) orientations from EOS datasheets for balanced parameters are also shown in Fig.8.4. All of the values are lower than the nominal values, and distributions could be described as skewed (Runs 4, 5, and 6) or multimodal (Runs 1, 2, and 3).

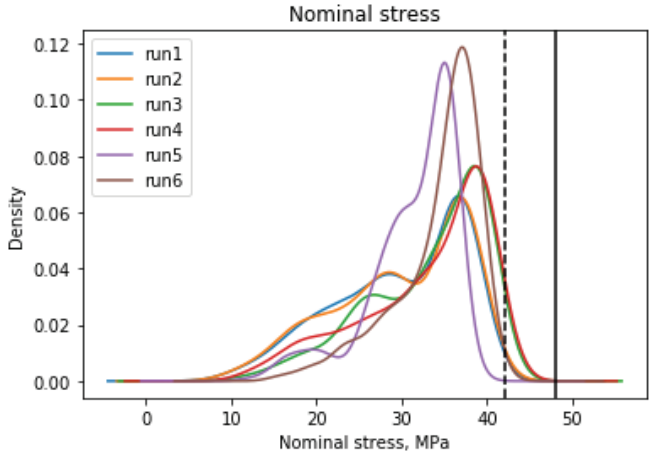


Figure 8.4: Distribution of tensile strength for different orientations based on kernel distribution estimation. The straight solid (48 MPa for XYZ orientation) and dashed lines (42 MPa for ZYX orientation) correspond to the value from EOS Balanced datasheet

Therefore, there is a need to describe data distributions by looking at orientations separately. Similarly to the description of the tensile modulus results,

tensile strength for Run 6 is visualized in a separate Fig.8.6 by the reason that this run has more than 18 different orientations. Orientation-based KDE distributions for other runs are illustrated on Fig.8.5.

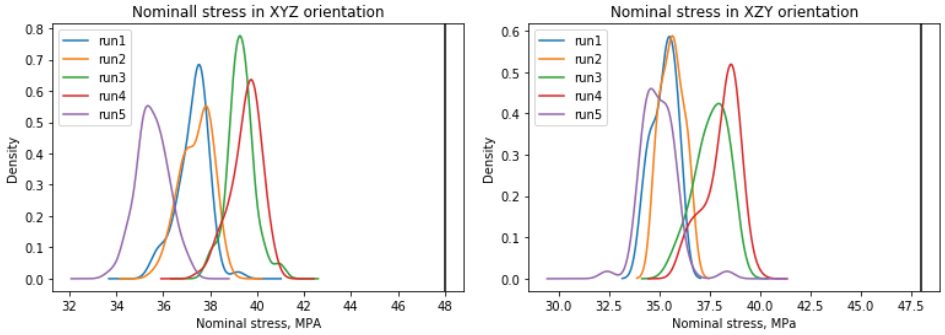
For XYZ orientation, all data distributions look similarly, the main difference is in the modal tensile strength values for each run. As it can be seen in Fig.8.5(a), Run 1 and Run 2 nearly repeat each other, while Run 4 has the lowest tensile strength values. Run 3 has the narrowest distribution, and Run 4 is a bit wider than Run 3 and has the highest modal value. However, all results are less than the nominal one from EOS datasheets. Such observed variation could be caused by changes in the materials since good repeatability is observed for the first two runs where the same material was used.

In terms of the values, data in XZY orientation is distributed in a similar range (ca. 34-41 MPa for XYZ orientation and ca. 33-40 MPa for XZY orientation). However, distributions for different runs are placed tighter and could be described as two groups. The first group consists of Runs 1, 2, and 5, which almost repeat each other. Since there is a difference in used material and built layout design, such repeatability is an interesting phenomenon. However, more analysis is needed to explain the observed behavior. Run 3 and Run 4 could be seen as the second group, where Run 4 has the highest modal tensile strength among all runs.

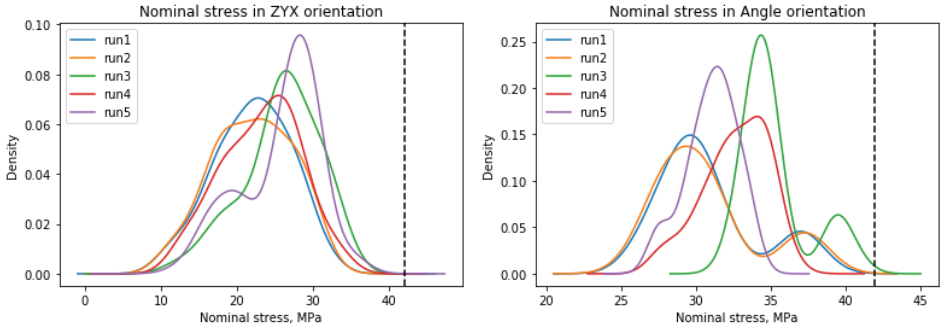
The difference between measured tensile strength values for XYZ and ZYX orientations is present for both in the EOS datasheet for balance parameters and Fig.8.5. This phenomenon is called anisotropic behavior and was already mentioned as an important issue for PA12, and it is observed for the tensile strength results. While tensile strength in ZYX orientation is in the range of ca. 10-35 MPa, the nominal value is 42 MPa (see Fig.8.5(b)). The results for Angle orientation are analyzed similarly to ZYX orientation, but the value range is similar to XZY orientation. The distribution is wider and resembles ZYX orientation (see Fig.8.5(d)).

When it comes to the evaluation of Run 6 with a large number of orientation groups, measured values are shown in Fig.8.6 as a distribution of values with respect to the position in the build chamber.

As can be seen from Fig8.6, orientations have an impact on the tensile strength, since some of the values in the center of the build are weaker than

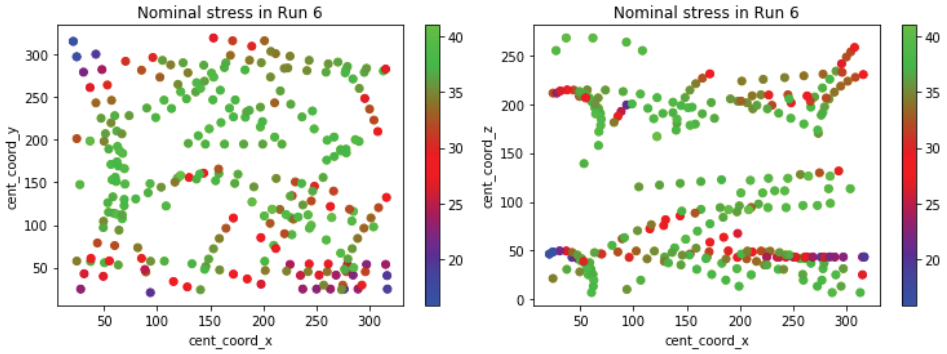


((a)) XYZ orientation for Runs 1-5 ((b)) XZY orientation for Runs 1-5



((c)) ZYX orientation for Runs 1-5 ((d)) Angle orientation for Runs 1-5

Figure 8.5: Kernel density estimation distributions of tensile strength for all runs. The straight solid (48 MPa for XYZ orientation) and dashed lines (42 MPa for ZYX orientation) correspond to the value from EOS Balanced datasheet



((a)) X/Y axes Run 6 ((b)) X/Z axes Run 6

Figure 8.6: Distribution of the tensile strength in the build chamber with regards to the coordinates for Run 6

their neighbors. Although a number of weaker specimens in the center of the build chamber is less than a number of weaker (and the weakest) specimens located in the corners of the build chamber. Thus both orientation and positioning constitute important factors that influence tensile strength values.

8.1.3 Analysis of strain at break

Strain at break, comparing with the two other mechanical properties, has more similar results to the ones provided by EOS and is illustrated in Figure 8.7. According to the EOS datasheet, strain at break is expected to be 18% for XYZ orientation and 4 % for ZYX orientation. The modal value of strain at break for different orientations differs from run to run. However, strain at break for Run 4 has the nearest modal value to the nominal value from the EOS datasheet (see Fig.8.8(a)), while results for other runs are alike and all higher than the respective nominal values.

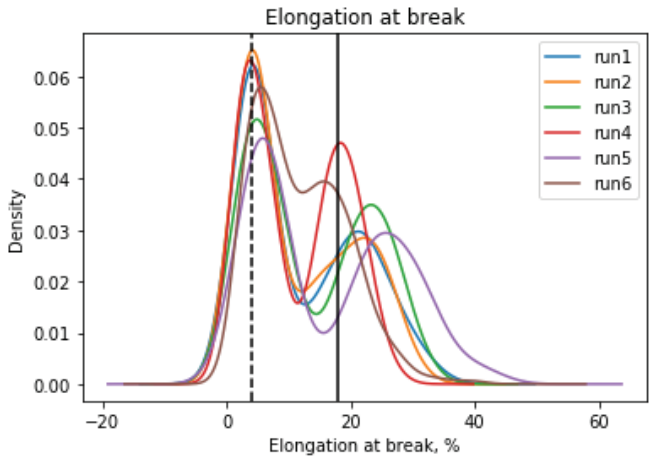
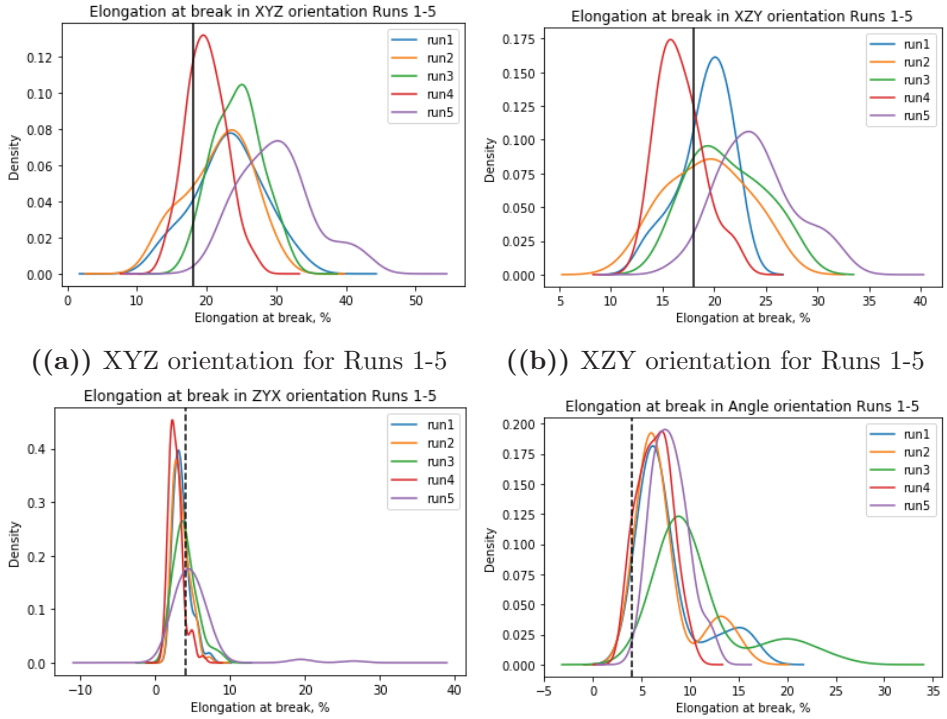


Figure 8.7: Distribution of strain at break for different orientations based on kernel distribution estimation. The straight solid (18 % for XYZ orientation) and dashed lines (4% for ZYX orientation) correspond to the value from EOS Balanced datasheet

Results for ZYX orientation are opposite to the ones described above for XYZ orientation. As it can be seen from Fig.8.8(c), the modal values of measured strain at break almost for all runs is around nominal value (4 %) from the EOS datasheet. However, Run 1 and Run 2 have modal values ca. 3.5 % and 2.8 % for Run 4, while Run 3 and Run 5 have the modal values of ca. 4 % and 5%.

Since there is no information from the EOS for other orientations, the observed values for XZY and Angle orientations cannot be compared with the nominal ones. However, it is expected the XZY orientation should have similar results to XYZ orientation due to the dominating x-axis, and thus slicing of the specimens is performed perpendicularly to the direction of force load in a tensile test. Similarly, ZYX and Angle orientations are sliced a way to facilitates crack generation under a tensile test.

As it can be seen from Fig.8.8(b), two runs in XZY orientation, have modal values around the nominal value (18 %) for XYZ orientation, where Run 4 has smaller modal value and Run 1 has a larger value than the nominal one. Similarly, the results observed in Angle orientation are more similar to the nominal value for ZYX orientation (4 %), but all of them have modal values



((a)) XYZ orientation for Runs 1-5

((b)) XZY orientation for Runs 1-5

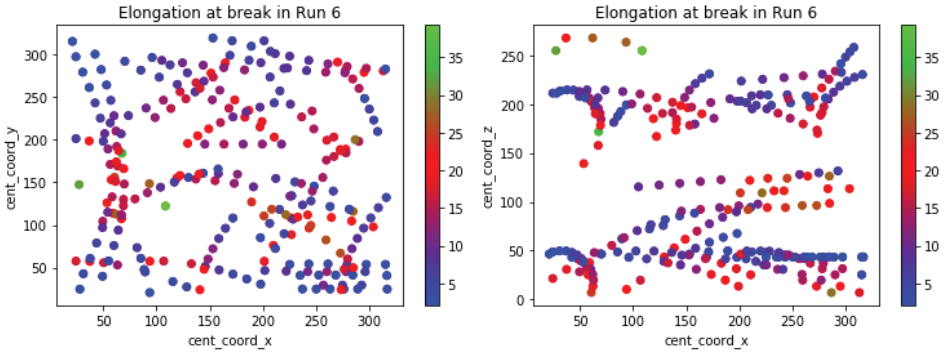
((c)) ZYX orientation for Runs 1-5

((d)) Angle orientation for Runs 1-5

Figure 8.8: Kernel Density Estimation distributions of strain at break for all runs. The straight solid (18 % for XYZ orientation) and dashed lines (4% for ZYX orientation) correspond to the value from EOS Balanced datasheet

greater than expected. Additionally, data distributions also are similar from run to run.

Strain at break values from Run 6 can also be described as a multimodal distribution, but the second peak is not as well defined as for other runs. The main reason for this observation is the number of orientation groups. Fig.8.9 supports this assumption, and it can be seen that orientations have a small impact on the measured values compared with how specimens' positioning influence the outcome.



((a)) X/Y axes Run 6 ((b)) X/Z axes Run 6

Figure 8.9: Distribution of strain at break in the build chamber with regards to the coordinates for Run 6

8.2 Feature selection

Similarly to the chapter about geometric deviations, feature selection for mechanical properties should also be made before developing intelligent models. Since tensile modulus, tensile strength and strain at break are of interest in this work, they will be analyzed separately.

Since there are more than one feature selection methods, data collected from Run 1 and Run 2 is used for defining a selection method that will be used for the analysis of data from all other runs. Pearson correlation test, Mutual information correlation test, and Spearsman's correlation rank methods are compared between each other, and the method that has led to the highest prediction accuracy of MLP model is used for feature selection for other runs as well. The process is similar to the one described in Chapter 6.

8.2.1 Feature selection for tensile modulus

The list of investigated features is the same as used for dimensional properties, and the interest is to develop a model based on Run, specimen's position, orientation, information about the STL model, and build layout design. The MLP prediction accuracy for tensile modulus of runs 1-2 without feature selection is 53.41%, while as it can be seen from Table 8.1, it is possible to improve prediction accuracy up to 80.18% by dropping the irrelevant features.

The *Run, reor_c, reor_a , mesh_triang, mesh_points, build height (bh), platform volume utilization (pvu), material, and current nesting density* features have been considered as irrelevant (when coefficient rank is lower than 0.1) based on the correlation rank values from the Mutual information test. According to the correlation coefficient provided by the Pearson correlation test, *Run, build height (bh), platform volume utilization (pvu), material, build layout, reor_c, current nesting density, and min_coord_y* features have been considered as irrelevant. The Spearsman's rho methods have identified *reor_c, build height (bh), platform volume utilization (pvu), material, current nesting density, min_coord_z, min_, max_, and cent_coord_y* features as irrelevant for prediction of tensile modulus for runs 1-2.

Since the Mutual Information method has the highest prediction accuracy, the feature selection for other runs is performed using this method. Additionally, the development of models based on the data collected from runs 1-2 will be done considering the selected features for techniques that do not include feature selection (e.g., Multilayer Perceptron Neural Network).

Table 8.1: Feature selection for tensile modulus based on the collected data from Run 1 and Run2

Features	Mutual Info	Pearson	Spearman's
min_coord_x	0.675	-0.623	-0.542
cent_coord_x	0.544	-0.548	-0.465
max_coord_x	0.465	-0.445	-0.339
max_coord_y	0.444	0.132	0.044
cent_coord_y	0.435	0.134	0.046
min_coord_y	0.432	0.138	0.046
reor_b	0.300	-0.420	-0.303
max_coord_z	0.297	-0.192	-0.215
cent_coord_z	0.235	-0.068	-0.066
min_coord_z	0.208	0.055	0.081
volume	0.176	0.197	0.204
surface	0.123	-0.215	-0.221
mesh_triang	0.026	-0.192	-0.192
mesh_points	0.025	-0.192	-0.192
reor_a	0.023	0.173	0.180
Run	0.0	-0.071	-0.093
reor_c	0	NaN	NaN
material	0	NaN	NaN
bh	0	NaN	NaN
pvu	0	NaN	NaN
cnd	0	0	NaN
build_layout	0	NaN	NaN
MLP accuracy	80.18%	53.27%	78.76%

However, an interesting observation can be made by looking at the changes of the correlation ranks when data from every run is merged with the data from the preceding runs. Table 8.2 shows that more data brings more uncertainty, and thus, the relevance of the investigated features decreases after each run. Column B corresponds to the dataset that includes data collected from run 1-3. While all parameters, including built layout, have been kept the same, there are some external factors that influence the outcome (tensile modulus). A list of the irrelevant features varies between the runs.

Another interesting observation can be made for the *Run* feature. This feature in the first three runs is considered as irrelevant, but after adding data from Run 4, its relevance significantly increases. This can be explained by looking at the activities connected to the build layout design. In Run 4, 75 new specimens have been added to the build layout design, while the number of specimens and their placement were kept the same for runs 1-3. This change could have had an impact on the mechanical properties as the specimens were packed more tightly, thus leading to the different temperature distribution within the build chamber.

A similar observation is made for current nesting density (cnd), which numerically describes how dense the parts are packed in the build chamber. The material, build height (bh) and build layout design features also vary when data from runs 4-6 are used for the analysis. The material feature is included not as a real material property, but as a numeric annotation on the batches (considering virgin/used ratio) of material being used. Therefore, this feature doesn't fully represent the real material factor, but it shows that material is an important factor that should be defined and used as a feature for predictive models in the future. It should be determined by material properties corresponding to the used material (e.g., viscosity fo polymers). Futhermore, it is assumed that changing from a numeric annotation to a physical measurement of viscosity could lead to a positive impact on the predictive models.

The resulting list of selected features for each column is different. All features that have correlation coefficient greater than 0 have been considered as significant. For the resulting dataset, which corresponds to Column E, all features are significant because the correlation level is much lower comparing to the Column A. Every change in the experiment contributes to the variation in the outcome. Therefore, predictive models for tensile modulus are developed using all the listed features. However, some of the machine learning techniques have a feature selection as a part of the algorithm. Thus, it would be interesting to compare the significance of the features obtained by such methods.

Table 8.2: Feature selection for tensile modulus mechanical property based on the collected data from all runs based on the Mutual information method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	A	B	C	D	E
min_coord_x	0.675	0.303	0.246	0.131	0.109
cent_coord_x	0.544	0.240	0.197	0.103	0.097
max_coord_x	0.465	0.198	0.171	0.109	0.114
max_coord_y	0.444	0.257	0.199	0.101	0.094
cent_coord_y	0.435	0.244	0.197	0.100	0.094
min_coord_y	0.432	0.251	0.199	0.118	0.109
reor_b	0.300	0.144	0.098	0.059	0.094
max_coord_z	0.297	0.161	0.123	0.046	0.035
cent_coord_z	0.235	0.143	0.117	0.047	0.052
min_coord_z	0.208	0.080	0.085	0.053	0.061
volume	0.176	0.061	0.047	0.045	0.042
surface	0.123	0	0.023	0.028	0.031
mesh_triang	0.026	0	0.026	0.030	0.031
mesh_points	0.025	0	0.025	0.029	0.030
reor_a	0.023	0.009	0.015	0.006	0.064
Run	0.0	0.006	0.344	0.406	0.435
bh	0	0	0.098	0.058	0.093
reor_c	0	0	0	0	0.072
pvu	0	0	0	0	0.071
material	0	0	0.015	0.006	0.063
build_layout	0	0	0.048	0.044	0.042
cnd	0	0	0.026	0.027	0.031
MLP accuracy	80.18%	47.28%	61.27%	57.77%	62.81%
MLP accuracy without FS	53.41%	41.51%	60.16%	62.01%	62.81%

8.2.2 Feature selection for tensile strength mechanical property

Feature selection for tensile strength mechanical property is performed in the same way as for the tensile modulus. As it can be seen from Table 8.3, MLP prediction accuracy for tensile strength without feature selection is

84.89%, and it has increased up to 99.02% with the help of Mutual information method. The lower threshold of the correlation coefficient for selecting features has been set to 0.05. Thus all features with the correlation coefficient that exceeds the chosen threshold are considered as relevant for tensile strength mechanical property.

Table 8.3: Feature selection for tensile strength based on the collected data from Run 1 and Run2

Features	**Mutual info**	**Pearson**	**Spearman's**
min_coord_x	1.155	-0.551	-0.472
max_coord_x	1.048	-0.245	-0.139
cent_coord_x	0.992	-0.411	-0.341
max_coord_y	0.941	0.153	0.091
min_coord_y	0.925	0.160	0.094
cent_coord_y	0.915	0.157	0.093
reor_b	0.881	-0.687	-0.575
min_coord_z	0.845	0.244	0.223
max_coord_z	0.811	-0.120	-0.161
cent_coord_z	0.811	0.066	0.050
volume	0.538	0.245	0.345
surface	0.465	-0.269	-0.368
mesh_triang	0.336	-0.165	-0.230
mesh_points	0.336	-0.165	-0.230
reor_a	0.282	0.261	0.140
Run	0.015	0.006	0.02
reor_c	0	0	0
material	0	NaN	NaN
bh	0	NaN	NaN
pvu	0	NaN	NaN
cnd	0	0	NaN
build_layout	0	NaN	NaN
MLP accuracy	99.02%	98.13%	91.69%

Mutual Information test is chosen as a filtering method for other runs based on the prediction accuracy obtained with features selected by this method. Similarly to the tensile modulus, the position of the specimens in the build

chamber has a large correlation with the tensile strength. By the reason that Mutual Information doesn't have an upper threshold for correlation rank, the minimal and maximal x coordinates have the largest correlation with the tensile strength in runs 1-2, and it is greater than 1. However, all correlation tests report that *reor_c, material, build height, platform volume utilization, current nesting density, and build layout design* features do not have any correlation with the tensile strength in the first two runs. This phenomenon is observed because all of the listed parameters are the same for runs 1-2.

When comparing how variations at each run contribute to the correlation between investigated features and tensile strength, one can see that the correlation coefficient also decreases from run to run, similarly to the results for tensile modulus. As it can be seen from Table 8.4, by changing the number of specimens (Column C), and specimens' location and orientation (Columns D and E), correlation between *cent_coord_y, max_coord_x and tensile strength* varies to a large degree, namely from strong correlation to no correlation).

Thus, in the first three runs, when the build layout design was the same, and even though the material has been changed to a new batch of polyamide, a correlation between all coordinates and tensile strength was strong. However, after variation of specimens' placement and orientation was introduced, the correlation level has changed, and it has decreased to a large degree.

One of the main assumptions why this behavior is observed is connected to the number of runs with the same build layout. In other words, Runs 4-6 haven't been repeated more than once, and each of them was characterized by a unique build layout design. The correlation between the *build_layout* feature supports this assumption. When the build layout was the same for the first three runs, the correlation between this feature and tensile strength was equal to 0, but after the build layout was changed, the correlation coefficient has increased. Similar behavior is also observed for *build height (bh), Run, reor_c, material and current nesting density (cnd)* features (see Table 8.4).

Along with that, it is also important to take a look at the feature that defines the material. As stated before, this feature is a numeric representation of the changes in the material. Therefore, it could be a reason why in Column

B (when we introduce new material batch), this feature is not correlated with tensile strength.

Table 8.4: Feature selection for tensile strength mechanical property based on the collected data from all runs based on the Mutual information method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	**A**	**B**	**C**	**D**	**E**
min_coord_x	1.155	1.159	0.268	0.159	0.119
max_coord_x	1.048	1.048	0.064	0.168	0.161
cent_coord_x	0.992	0.993	0.103	0.200	0.192
max_coord_y	0.941	0.939	0.198	0.275	0.252
min_coord_y	0.925	0.923	0.256	0.152	0.134
cent_coord_y	0.915	0.913	0	0	0.027
reor_b	0.881	0.880	0.265	0.172	0.152
min_coord_z	0.845	0.831	0.227	0.128	0.100
max_coord_z	0.811	0.809	0.264	0.155	0.089
cent_coord_z	0.811	0.807	0.064	0.167	0.161
volume	0.538	0.536	0.147	0.100	0.067
surface	0.465	0.463	0.134	0.101	0.066
mesh_triang	0.336	0.337	0.277	0.172	0.150
mesh_points	0.336	0.335	0.178	0.119	0.081
reor_a	0.282	0.282	0.269	0.172	0.098
Run	0.015	0.014	0.064	0.167	0.161
reor_c	0	0	0	0.201	0.129
bh	0	0	0.221	0.127	0.136
pvu	0	0	0	0	0.0
material	0	0	0.047	0.029	0.032
build_layout	0	0	0.1478	0.072	0.062
cnd	0	0	0.134	0.086	0.053
MLP accuracy	99.02%	52.07%	43.27%	36.41%	33.30%
MLP accuracy without FS	84.89%	60.09%	37.08%	33.88%	26.06%

Only one feature is considered as irrelevant for all datasets, which is *pvu* (platform volume utilization), and it is recommended to neglect this feature

when developing machine learning models for tensile strength if the method doesn't have feature selection included in the algorithm.

8.2.3 Feature selection for strain at break mechanical property

Comparing the results of different correlation tests, as can be seen from Table 8.5, the Spearsman's Rho method is among the best feature selection techniques for strain at break mechanical property. This choice is based on the improvements of prediction accuracy of MLP model from 93.93% (without feature selection) to 96.38% by neglecting all features with correlation rank lower than 0.1, which are *max_coord_x, Run, reor_c, max_coord_z, material, build height(bh), platform volume utilization (pvu), current nesting density (cnd), and build_layout.*

However, the Mutual information method provides different correlation ranks for *max_coord_x and max_coord_z* features, which are much higher than the other features. Additionally, prediction accuracies for all feature selection methods are relatively high and close to each other. Therefore, since two of three applied methods have provided low correlation rank for *max_coord_x and max_coord_z* features, it is decided to neglect them and use Spearsman's test to analyze data from all runs.

Strain at break mechanical property, comparing with tensile modulus and tensile strength, has a weaker correlation with specimens' position and orientation. This observation is based on the feature selection results shown in Table 8.6. While orientation along the y-axis (reor_b in Table 8.6) has a much stronger correlation with the other two mechanical properties, it has still the highest correlation with strain at break. However, orientation along the z-axis is either not correlated (runs 1-5) or has a close to 0 correlation rank (Run 6). Even though the correlation is close to 0 when it was neglected, the accuracy of the MLP model has decreased. Similar results are observed for build height (bh in Table 8.6) feature.

Features related to the general description of build layout design, such as *platform volume utilization (pvu), build_layout, and current nesting density (cnd)*, have similar correlation with strain at break, and are greater than 0 starting from Run 4 (when more parts are added to the build layout used for runs 1-3). Similarly, an increase in the correlation between the material feature and strain at break is observed for Run 3 (Column B in Table 8.6),

and this is related to the fact that new batch of material was used in this run.

Table 8.5: Feature selection for strain at break based on the collected data from Run 1 and Run2

Features	**Mutual info**	**Pearson**	**Spearman's Rho**
reor_b	0.897	-0.522	-0.598
min_coord_x	0.941	-0.325	-0.370
min_coord_z	0.773	0.352	0.341
surface	0.337	-0.345	-0.337
volume	0.363	0.315	0.311
reor_a	0.141	0.293	0.277
cent_coord_x	0.790	-0.182	-0.228
cent_coord_z	0.774	0.171	0.190
mesh_triang	0.221	-0.193	-0.186
mesh_points	0.223	-0.193	-0.186
min_coord_y	0.844	0.044	0.118
cent_coord_y	0.842	0.040	0.117
max_coord_y	0.856	0.035	0.112
max_coord_x	0.942	-0.020	-0.023
Run	0.0	-0.028	-0.027
max_coord_z	0.816	-0.023	-0.018
reor_c	0	NaN	NaN
material	0	NaN	NaN
bh	0	NaN	NaN
pvu	0	NaN	NaN
cnd	0	0	NaN
build_layout	0	NaN	NaN
MLP accuracy	96.30%	96.15%	96.38%

One can see that strain at break is more correlated with the material feature than the tensile modulus and tensile strength. This correlation could have been even stronger if real material properties have been used as a material feature. An additional observation can be made for the relationship between specimens' location and strain at break. In Column E, where data from all runs is joined, minimal coordinates along x and z axes have a stronger correl-

ation with strain at break than other coordinates. This could be related to the temperature distribution inside the build chamber. As it was described in Chapter 2, scanning strategy is a part of the process of how energy is applied to each layer and each specimen, especially when it comes to the hatch lines and layer distributions. However, more research work should be done to investigate how hatch lines distributions in connection with the part location contribute to variations in mechanical properties.

Table 8.6: Feature selection for strain at break mechanical property based on the collected data from all runs based on the Spearman's Rho method, (where A is for Runs 1-2, B is for runs 1-3, C is for runs 1-4, D is for runs 1-5, and E is for runs 1-6)

Features	A	B	C	D	E
reor_b	-0.598	-0.522	-0.255	-0.181	-0.159
min_coord_x	-0.370	-0.325	-0.171	-0.124	-0.116
min_coord_z	0.341	0.352	0.145	0.142	0.125
surface	-0.337	-0.345	-0.156	-0.076	-0.071
volume	0.311	0.315	0.136	0.091	0.068
reor_a	0.277	0.293	0.117	0.118	0.087
cent_coord_x	-0.228	-0.182	-0.099	-0.066	-0.068
cent_coord_z	0.190	0.171	0.070	0.086	0.074
mesh_triang	-0.186	-0.193	-0.120	-0.092	-0.080
mesh_points	-0.186	-0.193	-0.120	-0.092	-0.081
min_coord_y	0.118	0.044	0.032	0.010	0.005
cent_coord_y	0.117	0.040	0.032	0.012	0.005
max_coord_y	0.112	0.036	0.030	0.012	0.007
max_coord_x	-0.023	-0.020	-0.003	0.007	-0.004
Run	-0.027	-0.028	-0.051	0.121	0.073
max_coord_z	-0.018	-0.023	-0.025	0.012	0.012
reor_c	NaN	NaN	NaN	NaN	-0.007
bh	NaN	NaN	NaN	NaN	0.010
pvu	NaN	NaN	-0.075	-0.192	-0.169
material	NaN	0.101	-0.021	0.138	0.084
build_layout	NaN	NaN	-0.075	0.120	0.070
cnd	0	0	-0.075	-0.192	-0.127
MLP accuracy	96.38%	55.87%	45.24%	25.30%	22.31%
MLP accuracy without FS	93.93%	55.18%	46.28%	22.57%	22.31%

Feature selection is only one of the techniques used in machine learning for overcoming overfitting problems and shortening the training time. While one can argue why there is a need for feature selection because the prediction accuracy is so low that none of the models can be used for the prediction of mechanical properties, there exist other ways of improving the predictive

models. For example, using other machine learning methods that are more complex is one of the options. Additionally, there are machine learning methods that have feature selection included in the algorithms. Therefore, in the next section, different ML methods for developing intelligent predictive models will be compared and described in detail.

8.3 Intelligent predictive models

Similarly to the described models for geometric deviations in Chapter 6, predictive models for mechanical properties are also developed by applying the same four machine learning methods, namely Multilayer Perceptron, Random Forest, AdaBoost Regressor, and Gradient Boosting Regressor.

Results of feature selection for each mechanical property are incorporated in the development of MLP models. Since all features have a correlation coefficient greater than 0 for tensile modulus and strain at break, there is no need for comparing models' performance with and without feature selection. However, the platform volume utilization feature does not correlate with tensile strength, and therefore MLP models with and without feature selection are compared for tensile strength.

The process of developing predictive models for mechanical properties is the same as the one used for geometric deviations. However, a set of the hyperparameters that were tuned (as a part of the grid search) is described below as a reminder.

The set of hyperparameters used in the grid search is the same for all mechanical properties. Thus, for MLP models, the following set of hyperparameters have been evaluated:

- hidden layer sizes: [19, 22, 25, 27, 35],
- activation function:('relu', 'logistic'),
- solver function: ('lbfgs', 'sgd').

For Random Forest models, the following set of hyperparameters have been used for grid search:

- number of estimators: [50, 80, 100, 150, 200, 300],

- maximum features :('auto', 'sqrt', 'log2').

The set of hyperparameters used for grid search for AdaBoost Regressor is the following:

- learning rate: [0.01, 0.001, 0.05, 0.1],
- loss function: ('exponential',' square').

The number of estimators for AdaBoost is tuned internally by the algorithm. The set of hyperparameters investigated for Gradient Boost Regressor is similar to the AdaBoost, but types of loss function are different due to the used algorithms and are the following:

- number of estimators: [80, 100, 150],
- learning rate: [0.01, 0.001, 0.05, 0.1],
- loss function: ('ls', 'lad', 'huber').

8.3.1 Tensile modulus mechanical property

The chosen combinations of hyperparameters for each ML models are shown on Table 8.7. Prediction accuracy for all models is within the range from 64 to 70 %. However, the model's ability to generalization is an important factor when the choice should be made. Therefore, a test of the accuracy on an unseen testing dataset has been done to evaluate generalization ability. Additionally, five-fold cross-validation was used to evaluate the stability of the models. In other words, the evaluation of how sensitive the model is for different sets of data is done similarly, as presented in Chapter 6.

By comparing training, testing, and cross-validation accuracy, one can see that Random forest and Gradient Boost Regressor outperformed the other two models, and their results are alike. However, the prediction of newly introduced data is very important for the long term use. Therefore, it is recommended to choose the Random Forest model among all ML models. However, the accuracy of ca. 67% is still too low to be able to use the developed models for similar shapes. Therefore, there is a need for further improvements of the models.

Table 8.7: Results of optimization of predictive models for tensile modulus mechanical property

Model's hyperparameter	MLP FS = No FS	Random Forest	AdaBoost	GBR
activation	relu	-	-	-
hidden layer	35	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.01	0.1
loss	-	-	exponential	huber
n_estimators	-	200	50	150
max_features	-	auto	-	-
Accuracy	66.03%	69.93%	64.44%	69.93%
Cross-validation	65.72%	64.04%	61.58%	69.62%
Test accuracy	56.74%	66.84%	58.01%	64.30%

There are two ways how the improvement can be achieved. First of all, more features need to be defined, and real material property should be included as a feature. Second of all, experiments should be repeated by using the same build layouts as in runs 4-6, and machine-related process parameters should also be introduced and varied.

8.3.2 Tensile strength mechanical property

Even though the prediction accuracy of models developed to predict tensile modulus is too low to use the models, the prediction accuracy of ML models for tensile strength is even lower. The highest prediction accuracy (Accuracy in Table 8.8) is observed to be 37.4 % for Gradient Boost Regressor, while AdaBoost has the highest testing accuracy of 28.88 %.

The low accuracy can be explained in the way that using only information about build layout design with real material and process parameters is not enough for accurate predictive models for tensile strength. There is a need for more experimental work that will include new features about the material and machine parameters.

Table 8.8: Results of optimization of predictive models for tensile strength mechanical property

Model's hyperparameter	MLP No FS	MLP FS	Random Forest	AdaBoost	GBR
activation	relu	relu	-	-	-
hidden layer	19	25	-	-	-
solver	lbfgs	lbfgs	-	-	-
learning rate	0.001	0.001	-	0.001	0.05
loss	-	-	-	square	ls
n_estimators	-	-	100	50	80
max_features	-	-	auto	-	-
Accuracy	35.87%	34.15%	32.69%	35.38%	37.40%
Cross validation	23.40%	31.71%	32.12%	35.18%	37.13%
Test accuracy	11.69%	22.74%	23.93%	28.88%	27.05%

An additional observation can be made that another type of machine learning showed a better performance for tensile modulus. Such observation is important because it shows that there is more than only one machine learning method that can be effective for any predictive task. Whenever new features or data are used to retrain the models, different ML methods should be used as conducted in this work.

8.3.3 Strain at break mechanical property

Comparing with the other two mechanical properties, the strain at break has more similar results to tensile strength. As it can be seen from Table 8.9, the prediction accuracies are low for all models. Even though the AdaBoost model has outperformed other models, it cannot be used in its current form.

Similar actions need to be made in order to improve the models' performance, and this result shows that using information about build layout is meaningful but cannot be used only by itself. From the perspective of how polymer powder bed fusion AM process works, all research performed by other researchers supports the statement about a need for more information about material and machine.

Table 8.9: Results of optimization of predictive models for strain at break mechanical property

Model's hyperparameter	MLP FS = No FS	Random Forest	AdaBoost	GBR
activation	relu	-	-	-
hidden layer	19	-	-	-
solver	lbfgs	-	-	-
learning rate	0.001	-	0.001	0.05
loss	-	-	exponential	huber
n_estimators	-	150	50	150
max_features	-	auto	-	-
Accuracy	18.70%	22.28%	23.32%	23.25%
Cross-validation	23.07%	22.20%	22.89%	22.21%
Test accuracy	13.31%	16.08%	20.57%	17.51%

8.3.4 Mechanical properties as a vector of tensile modulus, tensile strength, and strain at break

By the reason that predictive models for mechanical properties have low prediction accuracy, which is very similar to the results observed for dimensional variations described in Chapter 6, it is important to evaluate the prediction of mechanical properties simultaneously.

Similarly, the grid search technique was used to find a combination of the model's hyperparameter that results in the best accuracy. However, the results, which are shown in Table 8.10, are not as promising as it was observed for dimensional variations. Random Forest model has outperformed the MLP model by almost 10% for testing accuracy and 13% for cross-validation. The model is similar to the Random Forest of tensile modulus, but it is less sensitive to different datasets due to similar results for training accuracy and cross-validation accuracy.

Table 8.10: Optimized hyperparameters of predictive models for tensile modulus, tensile strength, and strain at break simultaneously

Model's hyperparameters	MLP (FS=No FS)	Random Forest
activation	relu	-
hidden layer	27	-
solver	lbfgs	-
learning rate	0.001	-
n_estimators	-	200
max_features	-	auto
Accuracy	65.69%	69.43%
Cross-validation	56.77%	69.40%
Test accuracy	56.62%	66.24%

For better understanding of how predicted values differ from measured values, Fig. 8.10 - 8.12 illustrate prediction results for 17 randomly chosen specimens. In total 226 specimens have been used for testing model's performance.

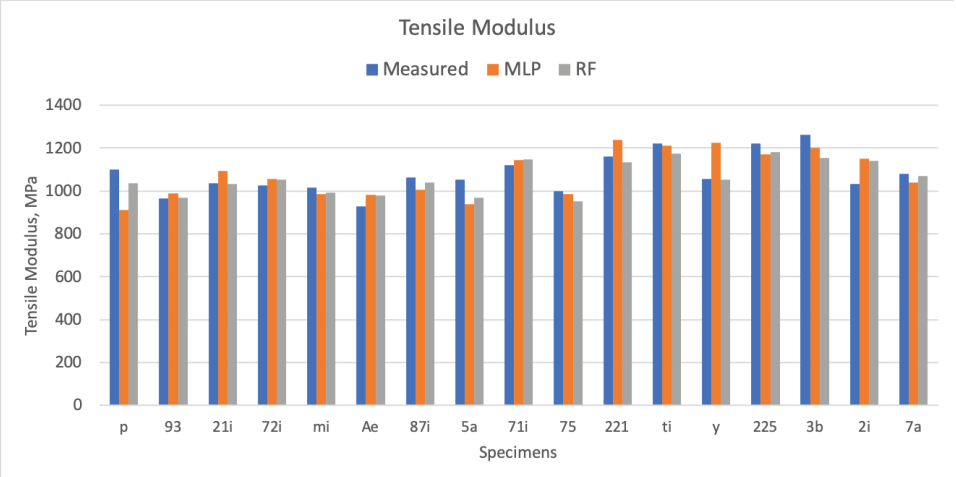


Figure 8.10: Comparison of observed and predicted tensile modulus values

As can be seen from Fig. 8.10, the prediction error for tensile modulus is the smallest comparing with two other mechanical properties. For some specimens, the predicted values are very close to the measured (e.g., specimens 93, 71i, or 7a).

The prediction of tensile strength has a higher prediction error comparing with tensile modulus. However, there is a number of specimens (93, 87i, or 221 on Fig. 8.11) for which at least one of the ML methods results in the predicted values very close to the measured ones. Therefore, by predicting mechanical properties simultaneously, the prediction accuracy for the tensile strength is improved, and these predictions could help to understand in which range the value of tensile strength can be in a specific location in the build chamber.

Similarly to the observed results for individual prediction of mechanical properties, Fig. 8.12 shows that stain at break has the largest deviations between predicted and measured values. Even though the Random Forest model outperformed the MLP model, there is still a large prediction error (e.g., specimens 75 or p) that doesn't allow using the proposed model for strain at break.

Therefore, the obtained results can be used to get a better understanding of the possible outcome based on a location and orientation of the specimens, but only for tensile modulus and tensile strength mechanical properties.

In general, all models need to be further improved by introducing information about the material and AM machine. More experimental work is necessary to collect the required data. Still, the presented results show that information about build layout design is important and should be used in the predictive models for mechanical properties.

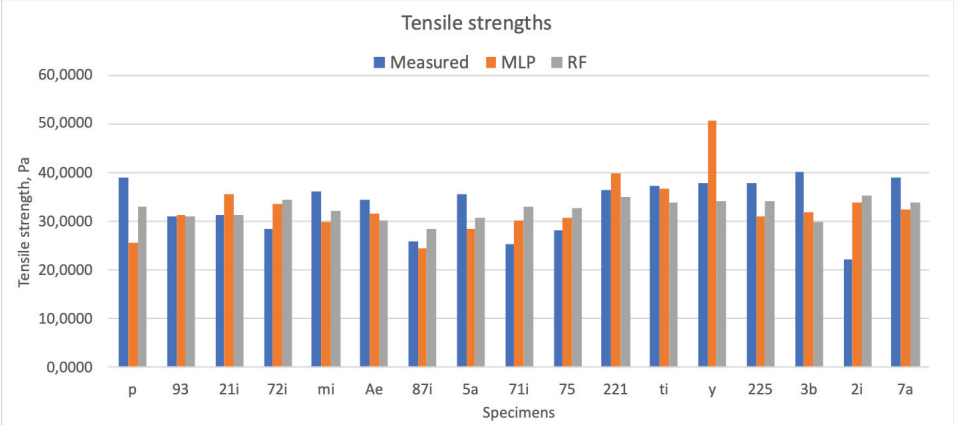


Figure 8.11: Comparison of observed and predicted tensile strength values

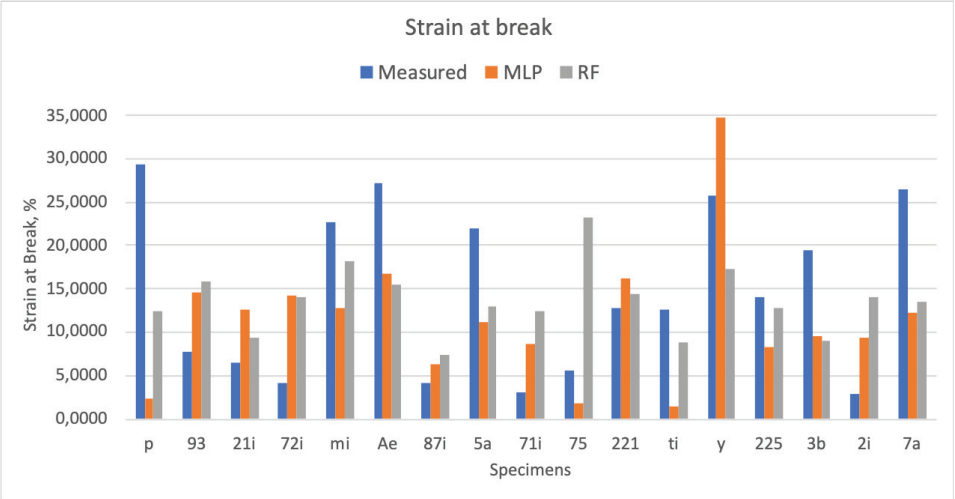


Figure 8.12: Comparison of observed and predicted strain at break values

Chapter 9

Module P3: Optimization of part placement in a build chamber

This chapter aims at describing how predictive models for estimation of the part placement in the build chamber are developed. The data preparation, including feature selection, is presented in Section 9.2. The predictive models and their comparison are presented in Section 9.3.

9.1 Preliminaries

Typically, optimization of part location within a build chamber in powder bed fusion processes is set to minimize build height, fabrication time, and cost [80]. In the literature, this task is typically reffered to as parts' packing. Most of the published works describe orientation as one of the significant issues with respect to the surface quality, dimensional accuracy, and build space utilization.

However, to be able to utilize build chamber space even more efficiently, there is a need for a decision-making tool that will help to optimize part location within the build chamber with respect to the size and mechanical properties of the product. Such a tool, for instance, could allow producing parts with different requirements for the mechanical properties within one build chamber.

Therefore, an algorithm for optimization of part location based on the requirements for mechanical properties is described in this work. Since different solutions for optimization of part orientation have already been proposed in various studies, it is assumed that part orientation was chosen by using one of the existing solutions.

The same data, which was used in Chapters 6, 8 and 7, is also used for developing models and evaluating the data analysis pipeline proposed in Chapter 3. As a reminder, the algorithm consists of three main steps; (1) data collection, (2) data preprocessing, and (3) data processing -models developing. The first step has already been conducted and described in the previous chapter, while the other two steps need to be performed for the defined task.

9.2 Data preprocessing

Data preprocessing is also important for this task because the list of features (inputs) and targets (outputs) is changed. Therefore, data integration followed by data normalization and feature selection needs to be performed again. Besides, it is important to remind that, usually before data integration is done, there is a need for data cleaning. In this work, data has already been cleaned, and the results are described in Chapter 4.

In order to evaluate whether it is possible to estimate part placement based on the information available before fabrication, mechanical properties, designed dimensions, material property, run number, build layout design number, predefined orientation, and STL model characteristics are used as inputs. Central coordinates in x, y, and z axes in $\mathbb{R}^3$ vector are considered as the output. Besides, it is proposed to develop individual models for tensile modulus, tensile strength, and strain at break. Therefore, data was reintegrated into three files that consist of information about input features, and one file for central coordinates (target). Data normalization is done in the same way as described in Chapter 4. The next step is feature selection for machine learning models.

9.2.1 Feature selection

Even though the output is handled as a vector of central coordinates x, y, and z, the feature selection needs to be conducted separately for each parameter since filtering methods work only for 1/1 (one feature and one

output parameter) arrays. The results of application of Mutual information, Pearson and Spearsman's correlation tests are compared in Tables 9.1, 9.2 and 9.2 for central coordinates x, y and z, respectively.

The feature selection process is not used to determine a cause-effect relationship, but it is used to filter insignificant features. As it can be seen from Tables 9.1, 9.2 and 9.2, in general there is a weak correlation between all investigated features and central coordinates. In contrast to Pearson and Spearsman's correlation tests, Mutual information defines relatively stronger correlations, especially for central coordinate x.

Table 9.1: Comparison of the feature selection methods based on the correlation ranks for listed features and central coordinate x

Features	**Mutual info**	**Pearson**	**Spearman's**
tensile modulus	0.113	-0.137	-0.118
Run	0.279	-0.028	-0.038
build_layout	0.317	-0.031	-0.039
strain at break	0.108	-0.078	-0.079
length	0.163	0.121	0.104
material	0.278	-0.021	-0.038
mesh_points	0.146	0.003	-0.012
mesh_triang	0.150	0.002	-0.013
reor_a	0.126	0.014	0.007
reor_b	0.273	-0.042	-0.056
reor_c	0.094	0.020	0.0169
tensile strength	0.125	-0.134	-0.110
surface	0.228	-0.026	-0.032
thickness	0.146	-0.113	-0.108
volume	0.270	-0.024	0.023
width	0.412	0.018	-0.028

According to the Mutual Information test results, width, build layout design, material, Run, volume, surface, and orientation by y-axis have the strongest correlation with all central coordinates. However, it is important to remember that none of the investigated features influence part position within the build chamber. Therefore, the correlations are not strong, and there is a high possibility that the provided information does not allow predicting the

part placement within the build chamber. Since the prediction of only one of the coordinates is not of interest in this work because the build chamber is a 3-dimensional space, the part location must be described as a vector of x, y, and z coordinates. Thus, based on the results in Chapter 6 and Chapter 7, one can assume that even though the correlation between listed features and coordinates is weak, optimization of part location within the build chamber could be done by using predictive models.

Table 9.2: Comparison of the feature selection methods based on the correlation ranks for listed features and central coordinate y

Features	**Mutual info**	**Pearson**	**Spearman's**
tensile modulus	0.089	0.075	0.051
Run	0.286	0.051	0.050
build_layout	0.294	0.055	0.052
strain at break	0.159	0.077	0.083
length	0.091	0.029	0.032
material	0.263	0.045	0.050
mesh_points	0.148	-0.035	-0.027
mesh_triang	0.153	-0.035	-0.028
reor_a	0.101	-0.081	-0.081
reor_b	0.231	-0.130	-0.125
reor_c	0.051	0.001	0.007
tensile strength	0.194	0.115	0.075
surface	0.233	-0.027	-0.027
thickness	0.164	0.120	0.142
volume	0.239	0.001	0.018
width	0.399	-0.050	0.008

Table 9.3: Comparison of the feature selection methods based on the correlation ranks for listed features and central coordinate z

Features	Mutual info	Pearson	Spearman's
tensile modulus	0.104	-0.030	-0.033
Run	0.273	0.011	0.002
build_layout	0.264	0.003	0.002
strain at break	0.143	0.065	0.072
length	0.113	0.197	0.198
material	0.247	0.026	0.002
mesh_points	0.249	0.041	0.041
mesh_triang	0.250	0.040	0.040
reor_a	0.153	-0.013	-0.021
reor_b	0.239	0.007	0.005
reor_c	0.063	-0.015	-0.025
tensile strength	0.217	0.029	0.024
surface	0.378	0.026	0.061
thickness	0.164	0.075	0.063
volume	0.342	-0.033	-0.076
width	0.329	-0.037	-0.010

Similar behavior was observed in Chapter 6 when correlations between features and dimensional properties were weak, but the predictive models for dimensions as a vector have the accuracy of 99.20%. Therefore, features with the correlation ratio greater than 0 are considered as significant for model development.

9.3 Predictive models for part placement as a vector of x, y and z coordinates

In previous chapters, ML models have been developed individually for each target. For this task, in contrast, the development of multi-output models is of interest because part location within a build chamber is always described by the set of coordinates in $\mathbb{R}^3$. This requirement sets constraints on the type of ML techniques that can be used for multi-output models. As a result, MLP and Random Forest ML techniques are used for this task.

The development of predictive models is performed individually for tensile

modulus, tensile strength, and strain at break. As the first step, the data is split into training and testing datasets with a ratio of 85/15 %, respectively. The training dataset consists of 1280 data points, while the testing dataset has 226 data points. The model development is conducted by using the training dataset, and five-fold cross-validation is also performed on the training dataset. The evaluation of the proposed models is also done on the testing dataset.

The model architecture has been chosen based on the results of a grid search, similarly as it was done in other chapters. The set of the hyperparameters that were used to tune machine learning models' hyperparameters (as a part of grid search) is the same for all models. Thus, for MLP models, the following set of hyperparameters have been evaluated:

- hidden layer sizes: [11, 13, 15, 17, 19, 21, 25, (8,27), 27, (12, 15), 32, 52], where tuple corresponds to the case of more than one hidden layer,
- activation function: ('relu', 'logistic'),
- solver function: ('lbfgs', 'sgd').

For Random Forest models, the following set of hyperparameters have been used for grid search:

- number of estimators: [30, 50, 80, 100, 150, 200, 300, 500, 600],
- maximum features: ('auto', 'sqrt', 'log2').

Resulting architectures of the developed MLP and Random Forest models, and comparison of the models are described in Tables 9.4, 9.5, and 9.6.

In the first case, when part location should be optimized for the specific tensile modulus, the prediction accuracies of MLP and Random Forest models are lower than 20% (see Table 9.4). Considering that data used for this task is collected from the parts of the same shape, size, and relatively similar mechanical properties, defining clear patterns between one point in 3D space and hyperparameters listed above, is a complicated task. The more historical data and a broader range of objects with different properties are

needed for evaluation of whether the proposed machine learning techniques are suitable for this task.

Table 9.4: Part placement: Optimized hyperparameters of predictive models with respect to tensile modulus

Model's hyperparameters	MLP (FS=No FS)	Random Forest
activation	relu	-
hidden layer	25	-
solver	lbfgs	-
learning rate	0.001	-
n_estimators	-	500
max_features	-	auto
Accuracy	22.44%	20.25%
Cross-validation	9.30%	19.90%
Test accuracy	-0.028%	14.23%

Table 9.5: Part placement: Optimized hyperparameters of predictive models with respect to tensile strength

Model's hyperparameters	MLP (FS=No FS)	Random Forest
activation	relu	-
hidden layer	(8,27)	-
solver	lbfgs	-
learning rate	0.001	-
n_estimators	-	500
max_features	-	auto
Accuracy	13.28%	21.70%
Cross-validation	11.08%	21.42%
Test accuracy	0.002%	18.12%

Table 9.6: Part placement: Optimized hyperparameters of predictive models with respect to strain at break

Model's hyperparameters	**MLP (FS=No FS)**	**Random Forest**
activation	relu	-
hidden layer	(12, 15)	-
solver	lbfgs	-
learning rate	0.001	-
n_estimators	-	600
max_features	-	sqrt
Accuracy	19.45%	20.55%
Cross-validation	11.76%	20.32%
Test accuracy	-0.134%	16.88%

The performance of predictive models with respect to tensile strength and strain at break is similar to the one described for tensile modulus. Even though the Random Forest method has outperformed the MLP, the models' performance is unsatisfactory as the prediction accuracies are in the range of 14-20%. This means that the current dataset and investigated features do not provide meaningful information that can be used for the development of predictive models for optimization of part placement based on the mechanical properties. Therefore, this task needs to be investigated in more detail, and other techniques should be used for optimizing part placement for the required mechanical properties.

Chapter 10

Discussion

The discussion chapter aims at discussing the research questions that have been previously formulated based on the obtained results and the current state of the art.

10.1 Discussing RQ1.

How does the build layout design affect product quality in polymer powder bed fusion systems?

The results of this research have shown that build layout design, in terms of part location, orientation, STL model properties, and packing density, is an important factor that could be used for quality management in additive manufacturing. However, usually, researchers mainly focus on the importance of material and process parameters when discussion about quality management of fabricated parts arises.

Considering the results from correlation tests for orientation and dimensional properties, orientation has the largest correlation with thickness and width, while correlation with length is close to 0. In contrast, Ituarte et al. [64], Pavan et al. [86] reported that orientation is an important factor related to the geometric deviations. The relative importance of the parameters for development of machine learning models also shows that orientation is an important factor (see Fig. 6.20(a) in Chapter 6).

In addition to the importance of orientation, location within a build cham-

ber is found as an important factor for both mechanical and dimensional properties. These findings are in line with the results reported by Senthilkumaran et al. [97]. In addition to the importance of part location, number of mesh triangles has also been previously reported by Calignano [20]. The authors presented how different combinations of part position and number of mesh triangles have resulted in variations of geometric inaccuracy, and similar results have been obtained in this study as well.

For mechanical properties, the researchers [39, 40] highlighted that orientation of a part in combination with the powder morphology, layer thickness [50] and scanning strategy [12, 81] have a large impact on the quality of parts. The results for the mechanical properties obtained in this study are not fully in line with this statement. For example, Tables 8.2, 8.4, and 8.6 show that tensile modulus and strain at break have weak correlation with orientation, while tensile strength have significantly stronger correlation with orientation parameters. These results are partly in line with the results reported by Faes et al. [39], where authors report that part orientation is not significant for mechanical properties. However, at the same time, the opposite results are presented by [40, 50].

However, an interesting observation can be made for tensile modulus and tensile strength. Even though both mechanical properties have low correlation with all of the investigated features, predictive models for tensile modulus describe 70% of the variation, while only 39 % of variation can be described for tensile strength. Strain at break has more similar results to the tensile strength, and when new material is introduced, the correlation between other parameters decreases significantly. This means that material is an important factor, but not the only one influencing the quality of the parts. When the density of the build packing is changed, and the parts' position and orientation are varied, the correlation also becomes weaker.

An interesting observation can be made by looking at the performance of predictive models for estimating geometric variations and mechanical properties as a vector (i.e. together). The performance for tensile strength and strain break becomes more similar to the tensile modulus, which could be a result of the presence of the correlation between the mechanical properties. Similar results are observed for dimensional properties as well. This correlation also helps to define more clear patterns in the investigated parameters.

It is important to remember that one-to-one correlation analysis only shows

the relationship between two parameters, while machine learning methods are able to define a more complex linear and non-linear correlation between two or more parameters. In general, the results show that part positioning, part orientation, packing density, and material have an impact on the quality of the AM produced parts. This means that not only typically defined process parameters and material should be of interest to the operators, but more attention should be paid to the build layout design.

10.2 Discussing RQ2.

How can machine learning techniques improve dimensional accuracy of AM?

In this work, the application of machine learning techniques to develop predictive models for estimation of geometric deviations is described in Chapters 6 and 7. The proposed models are able to provide information about the possible deviations individually for each object, depending on its orientation and location within the build chamber. It is worth noting that this is not a first attempt to find a solution for compensating geometric deviations, other researchers [22, 23, 46, 101, 114] have also made attempts to develop the mathematical models for prediction of shrinkage effect. While the estimation of deviations along x and y axes in the build chamber is more successful, the estimation of shrinkage effect along the z-axis is a more complex task, which is not fully addressed yet.

The shrinkage effect in the powder bed fusion machines is present due to the thermal processes in these machines. Depending on the laser energy input and material properties, parts either shrink or are produced larger than they are designed. While the results of this study show that temperature within the build chamber is also dependent on the packing density, parts location, and orientation (see Chapter 6), this information should be considered as a whole rather than isolated entities. As a result, every change of the location or orientation of one part influences the dimensional accuracy of the adjacent parts. The task of mathematically describing a dynamic system becomes even more complicated when, in addition to the build layout design, process and material parameters should also be a part of the system. Therefore, the application of machine learning techniques has a benefit over traditional mathematical modeling in terms of being able to define a relationship between a large number of parameters affecting the final results.

On the one hand, machine learning serves as a powerful tool that could help to adjust predictive models to rapid changes. On the other hand, there is a need for a large amount of data to make machine learning models more effective. The data could be acquired on an everyday basis while using AM machines, but there is a need for a systematic way of data collection. In other words, there is a need for standardized reporting systems so that these reports could be transformed into inputs to ML routines.

Traditionally, the design of experiments is a systematic technique used for data collection and analysis. Even though the structured DoE is beneficial for the optimization of individual manufacturing parameters of AM systems and allows developing a systematic experimental approach to simultaneously analyzing multiple production requirements [64], the structured DoE comes with drawbacks. For example, Flores Ituarte et al. [41] have reported that the "presented DoE was incapable of replicating this phenomenon, as the range energy density was limited to a narrower window due to excessive geometrical distortion", which means that the investigation of different combinations of AM-related parameters depends on the understanding of the process. In the case where some of the process conditions are not satisfied, the data from the experiment can be unrialable or unfit for statistical methods. Therefore, from this perspective, the application of the machine learning methods can help to avoid the existing issues by using the available data for data analysis and modeling.

Another benefit of using machine learning techniques over classical mathematical models lies in the ability to define the relationship between parameters automatically. For example, Chapter 7 describes how 22 parameters are used as an input for machine learning models. While correlation tests for one-to-one parameters' analysis show weak correlations, the machine learning models have prediction accuracy of 99.20%. There are several reasons for that. First of all, evaluation of all possible combinations of interaction effect for 13 parameters is a time-consuming task, and the machine learning models are able to do it in reasonable time.

However, the prediction of the compensation ratios for each object individually leads to another issue. If 200 objects need to be produced by the PBF machine, and each of the objects needs to be scaled in x, y, and z axes prior to the fabrication, the 600 compensations (3 compensations per object) need to be manually applied by an operator. This challenge is not yet addressed

in this study, but with the help of numpy-STL module, which is a new module available for the python programming language, the application of compensation ratios can be implemented as an algorithm in the future. This module allows to read the STL file, make changes to it, and then save it as an updated STL file. Since the developed models are also programmed by using python programming language, the numpy-STL module will allow for incorporating predicted compensation ratios automatically without human interaction.

10.3 Discussing RQ3.

How can application of ML techniques contribute to control and management of mechanical properties of AM built products?

Typically, mathematical models for mechanical properties of polyamide material in powder bed fusion systems are used for failure analysis using the finite element method. While this approach provides information about the possible places for crack generation, it doesn't provide any knowledge on where in the build chamber should the part be placed and whether two or more parts placed in the build chamber may have an effect on the quality of the adjacent parts. Therefore, in addition to the traditional approaches towards failure modeling, in this work, machine learning techniques are used for estimation of tensile modulus, tensile strength, and strain at break.

Discussion regarding the usage of machine learning for compensation of geometric variations shows that data collected from six experiments can be used for different types of modeling. The same data was also used for developing predictive models for mechanical properties with no changes made. Usually, one can see that DoE is an important technique used for systematic analysis and development of linear models. Still, there is a high risk of losing some of the data points due to errors in the AM process, and thus resulting in difficulties for further analysis [41]. Machine learning techniques work better with unstructured data and allows defining relationship that is hard to determine with classic methods, as shown in Chapter 8.

A large number of machine learning methods provide flexibility in improving the model's performance when more data is available. However, for mechanical properties presented in this work, the destructive tensile test is required to collect the data. This method becomes expensive and time-consuming when a large number of samples need to be tested. Additionally, the mod-

els are yet immature and cannot be trusted with validation, meaning that every time a part is produced, it has to be tested to evaluate the real values. Therefore, there is a concern about how to use the proposed models in the future. However, the benefit of using machine learning techniques lies in the possibility of faster adapting to changes, especially related to the introduction of new materials. The presented procedure for the experiments could be repeated a number of times. As a result, machine learning methods will be capable of redefining relationships when a new variation in the data is introduced.

While one can argue that the proposed predictive models developed based on a single machine, the same material and the same type of the shape have too many limitations, it is important to remember that transfer learning could be considered as one of the solutions to address these limitations. Transfer learning allows reusing previously collected data contained in the related tasks to solve new but similar problems more effectively [124]. However, not always reuse the same data for a different task is possible. For example, data acquired in this work provides good results for the tasks of predicting geometric deviation, compensation ratios, and mechanical properties, while optimization of part placement based on the mechanical properties is not possible yet (Chapter 9). Even though the idea is important, more research needs to be conducted in the future to address this issue.

10.4 Discussing RQ4.

How to utilize a build chamber of powder bed fusion machines in a more sustainable way?

Among 17 Sustainable Development Goals (SDGs) presented by the United Nations, there are two SDGs, namely ”Goal 9: Industry, Innovation and Infrastructure” [7] and ”Goal 12: Responsible Consumption and Production”[6], which could be addressed by the results of this work. Each sustainable development goal is comprised of a number of targets, and the results of this work could be considered as a small contribution to several targets.

In the SDG about industry, innovation, and infrastructure, several targets state that there is a need to enhance scientific research, upgrade the technological capabilities of industrial sectors, and ”upgrade infrastructure and retrofit industries to make them sustainable, with increased resource-use efficiency and greater adoption of clean and environmentally sound techno-

logies and industrial processes" [7]. In addition, a substantial reduction of the waste generated through prevention, reduction, recycling, and reuse is also stated in SDG as an important target [6].

Even though researchers report that AM technology could be considered as a more sustainable than traditional manufacturing due to reuse of the material, localized recycling, and decrease of the material waste [42], a lack of the in-depth knowledge about AM processes still results in a high material waste caused by "trial-and-error" approach, which is used to optimize machine, material and process parameters for the high-quality product. In other words, the efficient use of the AM processes is highly dependent on the operators' knowledge and experience. Therefore, to use AM technology even more efficiently, there is a need for decision support tools that will allow fabricating products with the "first-time-right" quality.

Different roads may lead to the "first-time-right" quality for the AM produced parts, which might result in a more sustainable AM production. However, this work has focused on the investigation of the build chamber utilization and how different build layouts influence the quality of the parts. Typically, it is recommended to use a center of the build chamber to obtain the best quality, while sides of the build chamber should be avoided due to temperature behaviors in the build chamber. However, if all space in a build chamber can be utilized, this will allow decreasing price per product, increase material usage, and increase time-to-market. Along with that, by being able to estimate dimensional and mechanical properties prior to the fabrication, geometric deviations could be compensated with the help of solutions described in Chapters 6 and 7, while estimation of mechanical properties, described in Chapter 8, could help to produce only parts that have satisfactory estimated properties.

On the one hand, the proposed solutions can contribute to a more efficient way of using powder bed technologies. On the other hand, several limitations need to be highlighted. All the proposed solutions are developed only for a single powder bed fusion machine, and cannot yet be used for other machines. Similarly, only one type of material and shape have been used to develop predictive models, which brings limitations to the type of objects that can be analyzed by the already existing models. While the idea and algorithms are important and should be used in the future, still more experimental work on different machines, materials, and objects is required for

the idea presented in this work to be successful.

Chapter 11

Conclusions and future work

This chapter consists of two sections. The 1^{st}section concludes the work underlying this thesis. The 2^{nd} section describes the ideas that are important to address in future work.

11.1 Conclusions

In this Ph.D. work, the new approach towards quality assurance in additive manufacturing is presented. The main effort has been set to investigate how different components of build layout design influence dimensional and mechanical properties of the fabricated parts. The main question that has been addressed in this work is how to get the first time right quaily of AM produced parts.

The thesis consisted of two parts. The first part described the contribution made in the design of experiments, with the novel data acquisition process being proposed. Producing a large number of labeled standard specimens have provided important insights towards the better understanding of how geometric deviations and mechanical properties differ based on the location within the build chamber. One of the findings highlight that mechanical properties are strong on the sides of the build chamber and not only in the central part of the build chamber (as postulated by the AM producer). The volume and number of mesh triangles were found to be important parameters, providing additional information about the object being fabricated. Along with that, orientation and coordinates that describe a part location

within the build chamber are also important parameters for the estimation of geometric variations and mechanical properties. The design of experimental work contributes to a better understanding of which data should be collected at each run of the AM machine, and could be used as an additive manufacturing machine calibration tool.

The second part of this work has focused on the development of intelligent predictive models for quality assurance in AM based on the empirically collected data. The model-based system engineering has been applied to design an intelligent system for quality assurance for the powder bed fusion AM process. The designed system should consist of 12 modules, where each corresponds to a specific prediction task. The connection between the modules has been developed and used as guidelines in the data analysis process. Based on the results from model-based system engineering, the predictive models have been developed as four separate modules. The data analysis pipeline has been proposed as a process of data preparation and has been developed based on the data science foundations. The collected data always need to go through several steps before being used in machine learning. Data cleaning, data integration, data normalization, and feature selection are the main data preprocessing steps.

The STL model characteristics, nine coordinates, three orientation angles, building platform utilization, platform density, material, run number, and build height are the parameters used to describe the build layout design. The listed parameters are used as input to train machine learning models. In total, 22 parameters were used to develop models, 1521 samples have been prepared, and used to train machine learning models, and 17 different predictive models have been proposed. As a result:

1. The predictive model with the accuracy of 99.16% for geometric deviations has been developed by using the Random Forest machine learning method. This model is a part of module P1 in the designed intelligent system for quality assurance.

2. The intelligent predictive model with accuracy 47.14% (RMSE = 0.026) was developed for predicting compensation ratios in x, y, and z axes for each object individually. The MLP and Random Forest have shown similar performance, and both models can be used in the future. This model is a part of module P2 in the designed intelligent system for

quality assurance.

3. The intelligent predictive model with accuracy 66.24% for estimating tensile modulus, tensile strength, and strain at break has been proposed by using the Random Forest machine learning method. This model is a part of module P12 in the designed intelligent system for quality assurance.

4. An attempt was made to develop the predictive model for optimization of part placement in the build chamber with respect to mechanical properties. This model is designed as a part of module P3 in the intelligent system for quality assurance. However, this model requires more work.

All developed models have limitations that have been presented earlier and should be addressed in the future. The next section presented the perspectives on future work.

11.2 Future work

With respect to the findings and defined limitations of this work, future work should be focused on the following aspects:

- The more experimental work should be done with a focus on the material and process parameters. Different scanning strategies and AM process settings need to be investigated and included in the proposed predictive models.
- The aspects of transfer learning should be investigated in more detail in order to generate a systematic approach for being able to use the proposed models on different powder bed fusion machines.
- The proposed models in this work are limited to a single shape being produced, and therefore, there is a need to develop a novel process for feature extraction from different shapes. The main interest should be set at answering thequestion of how the estimation of geometric deviations can be conducted for new, previously unseen objects.
- The modules presented as a part of the intelligent system for quality assurance, which are not addressed in this work, should be developed in the future.

- The universal database for data storage needs to be designed and implemented in the future. This element is important since it will help to make data analysis more automated and easier to perform.
- The material should be presented in the models with the respective properties.

Since this work was limited by a single type of shape, the investigation of shape distributions through the utilization of shape functions with respect to the different geometric attributes has already started in collaboration with the University of Southern California (USC, Los Angeles). The initial idea was to use shape distributions as a means for feature extraction to represent different shapes in the machine learning models presented in this work. The original concept was presented by Osada et al. [84] in the field of study of pattern recognition. The authors have used shape distribution based on the Euclidean distances between randomly sampled points on the surface of a 3D mesh model in order to evaluate the similarity between different objects and classify to which category they correspond. While this idea could be used as a technique for feature extraction, there are other ways how shape distributions could be used in AM. Quick quality inspection of additively manufactured products with a focus on deviation profiles is currently under investigation in collaboration with the researchers at USC.

List of publications

1. Baturynska, I., Martinsen K. (2020). Prediction of geometry deviations in additive manufactured parts: Comparison of linear regression with machine learning algorithms. Journal of Intelligent Manufacturing. Springer (Accepted)

2. Baturynska, I. (2019). Application of Machine Learning Techniques to Predict the Mechanical Properties of Polyamide 2200 (PA12) in Additive Manufacturing. Applied Sciences, 9(6), 1060.

3. Andreassen, E., Baturynska, I., Johansen, M., Andersen, M. (2019). Mechanical properties of polyamide 12 parts made by powder bed fusion (additive manufacturing). Nordic Polymer Days 2019, Trondheim.

4. Baturynska, I., Semeniuta, O., Wang, K. (2018, September). Application of machine learning methods to improve dimensional accuracy in additive manufacturing. In International Workshop of Advanced Manufacturing and Automation (pp. 245-252). Springer, Singapore.

5. Baturynska, I. (2018). Statistical analysis of dimensional accuracy in additive manufacturing considering STL model properties. The International Journal of Advanced Manufacturing Technology, 97(5-8), 2835-2849.

6. Baturynska, I., Semeniuta, O., Martinsen, K. (2018). Optimization of process parameters for powder bed fusion additive manufacturing by combination of machine learning and finite element method: A conceptual framework. Procedia CIRP, 67, 227-232.

Bibliography

[1] Mkram - material knowledge for robust additive manufacturing. https://www.sintef.no/en/projects/mkram-material-knowledge-for-robust-additive-manuf/. Accessed: 2019-09-05.

[2] The right measuring equipment for the shopfloor. https://www.zeiss.com/metrology/products/systems/coordinate-measuring-machines/production-cmms/duramax.html. Accessed: 2019-08-21.

[3] Pearson correlation test. https://docs.scipy.org/doc/scipy-0.15.1/reference/generated/scipy.stats.pearsonr.html, . Accessed: 2019-09-13.

[4] Spearsman correlation test. https://docs.scipy.org/doc/scipy-0.16.1/reference/generated/scipy.stats.spearmanr.html, . Accessed: 2019-09-13.

[5] Mutual information - feature selection filtering method. https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.mutual_info_regression.html. Accessed: 2019-09-16.

[6] Envision2030 goal 12: Responsible consumption and production. https://www.un.org/development/desa/disabilities/envision2030-goal12.html, . Accessed: 2019-11-15.

[7] Envision2030 goal 9: Industry, innovation and infrastructure. https://www.un.org/development/desa/disabilities/envision2030-goal9.html, . Accessed: 2019-11-15.

[8] Metin Aclan, Rauf Oezden, and Saliba Danho. Investigation of the geometric change of plastic articles made by polyamide. In *MATEC Web of Conferences*, volume 178, page 04003. EDP Sciences, 2018.

[9] FA Adnan, FRM Romlay, and M Shafiq. Real-time slicing algorithm for stereolithography (stl) cad model applied in additive manufacturing industry. In *IOP Conference Series: Materials Science and Engineering*, volume 342, page 012016. IOP Publishing, 2018.

[10] SO Akande, KW Dalgarno, J Munguia, and J Pallari. Assessment of tests for use in process and quality control systems for selective laser sintering of polyamide powders. *Journal of Materials Processing Technology*, 229:549–561, 2016.

[11] Tomàs Aluja-Banet and Eduard Nafria. Stability and scalability in decision trees. *Computational Statistics*, 18(3):505–520, 2003.

[12] Yiğit M Arısoy, Luis E Criales, Tuğrul Özel, Brandon Lane, Shawn Moylan, and Alkan Donmez. Influence of scan strategy and process parameters on microstructure and its optimization in additively manufactured nickel alloy 625 via laser powder bed fusion. *The International Journal of Advanced Manufacturing Technology*, 90(5-8):1393–1417, 2017.

[13] Alexander Arndt, Heike Hackbusch, and Reiner Anderl. An algorithm-based method for process-specific three-dimensional nesting for additive manufacturing processes. In *26th Solid Freeform Fabrication Symposium*, volume 10, page 2015, 2015.

[14] Ivanna Baturynska. Statistical analysis of dimensional accuracy in additive manufacturing considering stl model properties. *The International Journal of Advanced Manufacturing Technology*, pages 1–15, 2018.

[15] Ivanna Baturynska. Application of machine learning techniques to predict the mechanical properties of polyamide 2200 (pa12) in additive manufacturing. *Applied Sciences*, 9(6):1060, 2019.

[16] Ivanna Baturynska, Oleksandr Semeniuta, and Kristian Martinsen. Optimization of process parameters for powder bed fusion additive manufacturing by combination of machine learning and finite element method: A conceptual framework. *Procedia CIRP*, 67C:227–232, 2018.

[17] Sebastian Berumen, Florian Bechmann, Stefan Lindner, Jean-Pierre Kruth, and Tom Craeghs. Quality control of laser-and powder bed-based additive manufacturing (am) technologies. *Physics procedia*, 5: 617–622, 2010.

[18] Tomaz Brajlih, Bogdan Valentan, Joze Balic, and Igor Drstvensek. Speed and accuracy evaluation of additive manufacturing machines. *Rapid prototyping journal*, 17(1):64–75, 2011.

[19] Gabriel Bugeda Miguel Cervera and Guillermo Lombera. Numerical prediction of temperature and density distributions in selective laser sintering processes. *Rapid Prototyping Journal*, 5(1):21–26, 1999.

[20] F Calignano. Investigation of the accuracy and roughness in the laser powder bed fusion process. *Virtual and Physical Prototyping*, 13(2): 97–104, 2018.

[21] G Casalino, SL Campanelli, N Contuzzi, and AD Ludovico. Experimental investigation and statistical optimisation of the selective laser melting process of a maraging steel. *Optics & Laser Technology*, 65: 151–158, 2015.

[22] Brian Caulfield, PE McHugh, and S Lohfeld. Dependence of mechanical properties of polyamide components on build parameters in the sls process. *Journal of Materials Processing Technology*, 182(1):477–488, 2007.

[23] Longwei Cheng, Andi Wang, and Fugee Tsung. A prediction and compensation scheme for in-plane shape deviation of additive manufacturing with information on process parameters. *IISE Transactions*, 50(5):394–406, 2018.

[24] Sushmit Chowdhury and Sam Anand. Artificial neural network based geometric compensation for thermal deformation in additive manufacturing processes. In *ASME 2016 11th International Manufacturing Science and Engineering Conference*. American Society of Mechanical Engineers Digital Collection, 2016.

[25] Cai Cong-Zhong, Pei Jun-Fang, Wen Yu-Feng, Zhu Xing-Jian, and Xiao Ting-Ting. Density prediction of selective laser sintering parts based on support vector regression. *Acta Physica Sinica*, 58:S1–S19, 2009.

[26] Johannes A Coy, Clara-Maria Kuball, Daniel B Roppenecker, and Tim C Lueth. Flexural modulus of lasersintered pa 2200. In *ASME 2013 International Mechanical Engineering Congress and Exposition*, pages V02AT02A011–V02AT02A011. American Society of Mechanical Engineers, 2013.

[27] Garrett Craft, Justin Nussbaum, Nathan Crane, and JP Harmon. Impact of extended sintering times on mechanical properties in pa-12 parts produced by powderbed fusion processes. *Additive Manufacturing*, 22:800–806, 2018.

[28] John W Creswell. Research design: Qualitative and mixed methods approaches. *London and Thousand Oaks: Sage Publications*, 2009.

[29] Ademir Linhares de Oliveira and Gustavo Daniel Donatelli. Historical measurement data reuse and similarity analysis for dimensional production tolerancing of injected plastic parts. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 39(10):4161–4175, 2017.

[30] Carl R Deckard, Joseph J Beaman, and James F Darrah. Method for selective laser sintering with layerwise cross-scanning, October 13 1992. US Patent 5,155,324.

[31] Nathan Decker and Qiang Huang. Geometric accuracy prediction for additive manufacturing through machine learning of triangular mesh data.

[32] Brian L DeCost, Harshvardhan Jain, Anthony D Rollett, and Elizabeth A Holm. Computer vision and machine learning for autonomous characterization of am powder feedstocks. *Jom*, 69(3):456–465, 2017.

[33] Jordi Delgado, Joaquim Ciurana, and Ciro A Rodríguez. Influence of process parameters on part quality and mechanical properties for dmls and slm with iron-based materials. *The International Journal of Advanced Manufacturing Technology*, 60(5-8):601–610, 2012.

[34] DIN 16742:2013. Plastics mouldings: tolerances and acceptance conditions. Standard, German Institute for Standardization, 2013.

[35] Donghong Ding, Zengxi Pan, Dominic Cuiuri, Huijun Li, and Stephen van Duin. Advanced design for additive manufacturing: 3d slicing and 2d path planning. *New Trends in 3D Printing*, 2016.

[36] Lin Dong, Nicolas Barth, JPM Correia, and Saïd Ahzi. Modeling and numerical simulation of selective laser sintering. In *Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), 2016 17th International Conference on*, pages 1–4. IEEE, 2016.

[37] Dietmar Drummer, Katrin Wudy, and Maximilian Drexler. Modelling of the aging behavior of polyamide 12 powder during laser melting process. In *AIP Conference Proceedings*, volume 1664, page 160007. AIP Publishing, 2015.

[38] Tobias Eifler, Boorla S Murthy, and Thomas J Howard. Toward meaningful manufacturing variation data in design–feature based description of variation in manufacturing processes. *Procedia CIRP*, 43:190–195, 2016.

[39] Matthias Faes, Yueqi Wang, Pascal Lava, and David Moens. Variability, heterogeneity, and anisotropy in the quasi-static response of laser sintered pa12 components. *Strain*, 53(2):e12219, 2017.

[40] Göran Flodberg, Henrik Pettersson, and Li Yang. Pore analysis and mechanical performance of selective laser sintered objects. *Additive Manufacturing*, 24:307–315, 2018.

[41] Iñigo Flores Ituarte, Olli Wiikinkoski, and Anton Jansson. Additive manufacturing of polypropylene: A screening design of experiment using laser-based powder bed fusion. *Polymers*, 10(12):1293, 2018.

[42] Simon Ford and Mélanie Despeisse. Additive manufacturing and sustainability: an exploratory study of the advantages and challenges. *Journal of Cleaner Production*, 137:1573–1587, 2016.

[43] Jerome H Friedman. Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4):367–378, 2002.

[44] Mario Ganci, Wei Zhu, Gianluca Buffa, Livan Fratini, Song Bo, and Chunze Yan. A macroscale fem-based approach for selective laser

sintering of thermoplastics. *The International Journal of Advanced Manufacturing Technology*, 91(9-12):3169–3180, 2017.

[45] A Garg, Jasmine Siu Lee Lam, and MM Savalani. A new computational intelligence approach in formulation of functional relationship of open porosity of the additive manufacturing process. *The International Journal of Advanced Manufacturing Technology*, 80(1-4):555–565, 2015.

[46] Akhil Garg and K Tai. An ensemble approach of machine learning in evaluation of mechanical property of the rapid prototyping fabricated prototype. In *Applied mechanics and materials*, volume 575, pages 493–496. Trans Tech Publ, 2014.

[47] ULF Gedde. *Polymer physics*. Springer Science & Business Media, 2001.

[48] Christian Gobert, Edward W Reutzel, Jan Petrich, Abdalla R Nassar, and Shashi Phoha. Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging. *Additive Manufacturing*, 21: 517–528, 2018.

[49] R Goodridge and S Ziegelmeier. Powder bed fusion of polymers. In *Laser Additive Manufacturing*, pages 181–204. Elsevier, 2017.

[50] RD Goodridge, CJ Tuck, and RJM Hague. Laser sintering of polyamides and other polymers. *Progress in Materials science*, 57(2):229–267, 2012.

[51] Serap Gümüs, Juergen M Lackner, Şeyda Polat, Wolfgang Kraschitzer, Hermann Hanning, Alperen Bayram, Mesut Kaya, Metin Çallı, and Attila Alkan. Failure behavior of pa12 based sls lattice structure with macro-porosity. In *MATEC Web of Conferences*, volume 188, page 03007. EDP Sciences, 2018.

[52] Jingbin Hao, Liang Fang, and Robert E Williams. An efficient curvature-based partitioning of large-scale stl models. *Rapid Prototyping Journal*, 17(2):116–127, 2011.

[53] Ross F Housholder. Molding process, January 27 1981. US Patent 4,247,508.

[54] Qiang Huang. An analytical foundation for optimal compensation of three-dimensional shape deformation in additive manufacturing. *Journal of Manufacturing Science and Engineering*, 138(6):061010, 2016.

[55] Qiang Huang, Hadis Nouri, Kai Xu, Yong Chen, Sobambo Sosina, and Tirthankar Dasgupta. Predictive modeling of geometric deviations of 3d printed products-a unified modeling approach for cylindrical and polygon shapes. In *2014 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 25–30. IEEE, 2014.

[56] Qiang Huang, Hadis Nouri, Kai Xu, Yong Chen, Sobambo Sosina, and Tirthankar Dasgupta. Statistical predictive modeling and compensation of geometric deviations of three-dimensional printed products. *Journal of Manufacturing Science and Engineering*, 136(6):061008, 2014.

[57] Qiang Huang, Jizhe Zhang, Arman Sabbaghi, and Tirthankar Dasgupta. Optimal offline compensation of shape shrinkage for three-dimensional printing processes. *Iie transactions*, 47(5):431–441, 2015.

[58] S-H Huang, L-C Zhang, and M Han. An effective error-tolerance slicing algorithm for stl files. *The International Journal of Advanced Manufacturing Technology*, 20(5):363–367, 2002.

[59] Sung-Min Hur, Kyung-Hyun Choi, Seok-Hee Lee, and Pok-Keun Chang. Determination of fabricating orientation and packing in sls process. *Journal of Materials Processing Technology*, 112(2):236–243, 2001.

[60] Farhad Imani, Aniruddha Gaikwad, Mohammad Montazeri, Prahalada Rao, Hui Yang, and Edward Reutzel. Layerwise in-process quality monitoring in laser powder bed fusion. In *ASME 2018 13th International Manufacturing Science and Engineering Conference*, pages V001T01A038–V001T01A038. American Society of Mechanical Engineers, 2018.

[61] ISO 527-1: 2019. Plastics - Deformation of tensile properties - Part 1: General principles. Standard, ISO 2019, 2019.

[62] ISO/ASTM 52900:2015(E). Standard terminology for additive manufacturing - General Principles - Terminology. Standard, ISO/ASTM International, 2015.

[63] ISO/ASTM 52921:2013(E). Standard terminology for additive manufacturing: Coordinate systems and test methodologies. Standard, ISO/ASTM International, 2013.

[64] Iñigo Flores Ituarte, Eric Coatanea, Mika Salmi, Jukka Tuomi, and Jouni Partanen. Additive manufacturing in production: a study case applying technical requirements. *Physics Procedia*, 78:357–366, 2015.

[65] Jamasp Jhabvala, Eric Boillat, Thibaud Antignac, and Rémy Glardon. On the effect of scanning strategies in the selective laser melting process. *Virtual and physical prototyping*, 5(2):99–109, 2010.

[66] Sotiris B Kotsiantis. Decision trees: a recent overview. *Artificial Intelligence Review*, 39(4):261–283, 2013.

[67] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical review E*, 69(6):066138, 2004.

[68] T Kumaresan, R Gandhinathan, M Ramu, M Ananthasubramanian, and K Banu Pradheepa. Design, analysis and fabrication of polyamide/hydroxyapatite porous structured scaffold using selective laser sintering method for bio-medical applications. *Journal of mechanical science and technology*, 30(11):5305–5312, 2016.

[69] Pil-Ho Lee, Haseung Chung, Sang Won Lee, Jeongkon Yoo, and Jeonghan Ko. Dimensional accuracy in additive manufacturing processes. In *ASME Proceedings*, page V001T04A045, 2014.

[70] Liangzhi Li, Kaoru Ota, and Mianxiong Dong. Deep learning for smart industry: Efficient manufacture inspection system with fog computing. *IEEE Transactions on Industrial Informatics*, 2018.

[71] Andy Liaw, Matthew Wiener, et al. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.

[72] David Long and Zane Scott. *A primer for model-based systems engineering*. Lulu. com, 2011.

[73] He Luan and Qiang Huang. Predictive modeling of in-plane geometric deviation for 3d printed freeform products. In *2015 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 912–917. IEEE, 2015.

[74] Candice Majewski and Neil Hopkinson. Effect of section thickness and build orientation on tensile properties and material characteristics of laser sintered nylon-12 parts. *Rapid Prototyping Journal*, 17(3):176–180, 2011.

[75] Kristian Martinsen, Jonathan Downey, and Ivanna Baturynska. Human-machine interface for artificial neural network based machine tool process monitoring. *Procedia CIRP*, 41:933–938, 2016.

[76] SH Masood, W Rattanawong, and P Iovenitti. A generic algorithm for a best part orientation system for complex parts in rapid prototyping. *Journal of materials processing technology*, 139(1-3):110–116, 2003.

[77] C Mielicki, B Gronhoff, and J Wortberg. Effects of laser sintering processing time and temperature on changes in polyamide 12 powder particle size, shape and distribution. In *AIP Conference Proceedings*, volume 1593, pages 728–731. AIP, 2014.

[78] Jorge Mireles, Shakerur Ridwan, Philip A Morton, Alejandro Hinojos, and Ryan B Wicker. Analysis and correction of defects within parts fabricated using powder bed fusion technology. *Surface Topography: Metrology and Properties*, 3(3):034002, 2015.

[79] Kadephi Vuyolwethu Mjali and Annelize Botes. The influence of the concept of ?line energy? on the mechanical properties of laser formed commercially pure grade 2 titanium alloy plates. *Procedia Manufacturing*, 26:267–275, 2018.

[80] J Munguía, J Ciurana, and C Riba. Neural-network-based model for build-time estimation in selective laser sintering. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 223(8):995–1003, 2009.

[81] Sushant Negi and Rajesh Kumar Sharma. Study on shrinkage behaviour of laser sintered pa 3200gf specimens using rsm and ann. *Rapid Prototyping Journal*, 22(4):645–659, 2016.

[82] Olga Ogorodnyk, Ole Vidar Lyngstad, Mats Larsen, Kesheng Wang, and Kristian Martinsen. Application of machine learning methods for prediction of parts quality in thermoplastics injection molding. In *International Workshop of Advanced Manufacturing and Automation*, pages 237–244. Springer, 2018.

[83] Ikenna A Okaro, Sarini Jayasinghe, Chris Sutcliffe, Kate Black, Paolo Paoletti, and Peter L Green. Automatic fault detection for laser powder-bed fusion using semi-supervised machine learning. *Additive Manufacturing*, 27:42–53, 2019.

[84] Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Shape distributions. *ACM Transactions on Graphics (TOG)*, 21(4):807–832, 2002.

[85] Nikhil Padhye and Kalyanmoy Deb. Multi-objective optimisation and multi-criteria decision making in sls using evolutionary approaches. *Rapid Prototyping Journal*, 17(6):458–478, 2011.

[86] M Pavan, M Faes, D Strobbe, B Van Hooreweder, T Craeghs, D Moens, and W Dewulf. On the influence of inter-layer time and energy density on selected critical-to-quality properties of pa12 parts produced via laser sintering. *Polymer testing*, 61:386–395, 2017.

[87] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[88] Harsh Rana and Nicholas Benoit. Applying machine learning for real time optimization of powder bed manufacturing. 2019.

[89] Sayyeda Saadia Razvi, Shaw Feng, Anantha Narayanan, Yung-Tsun Tina Lee, and Paul Witherell. A review of machine learning applications in additive manufacturing. In *ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection, 2019.

[90] Wang Rong-Ji, Li Xin-hua, Wu Qing-ding, and Wang Lingling. Optimizing process parameters for selective laser sintering based on

neural network and genetic algorithm. *The International Journal of Advanced Manufacturing Technology*, 42(11-12):1035–1042, 2009.

[91] Stefan Rüsenberg, Stefan Josupeit, and Hans-Joachim Schmid. A method to characterize the quality of a polymer laser sintering process. *Advances in Mechanical Engineering*, 6:185374, 2014.

[92] MM Savalani, L Hao, PM Dickens, Y Zhang, KE Tanner, and RA Harris. The effects and interactions of fabrication parameters on the properties of selective laser sintered hydroxyapatite polyamide composite biomaterials. *Rapid Prototyping Journal*, 18(1):16–27, 2012.

[93] Manfred Schmid and Gideon Levy. Quality management and estimation of quality costs for additive manufacturing with sls. In *Fraunhofer Direct Digital Manufacturing Conference 2012*. ETH-Zürich, 2012.

[94] Luke Scime and Jack Beuth. Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm. *Additive Manufacturing*, 19:114–126, 2018.

[95] Luke Scime and Jack Beuth. A multi-scale convolutional neural network for autonomous anomaly detection and classification in a laser powder bed fusion additive manufacturing process. *Additive Manufacturing*, 24:273–286, 2018.

[96] Oleksandr Semeniuta, Sebastian Dransfeld, and Petter Falkman. Vision-based robotic system for picking and inspection of small automotive components. In *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 549–554. IEEE, 2016.

[97] K Senthilkumaran, Pulak M Pandey, and PVM Rao. Influence of building strategies on the accuracy of parts in selective laser sintering. *Materials & Design*, 30(8):2946–2954, 2009.

[98] K Senthilkumaran, Pulak M Pandey, and PVM Rao. New model for shrinkage compensation in selective laser sintering. *Virtual and Physical Prototyping*, 4(2):49–62, 2009.

[99] Xianfeng Shen, Jin Yao, Yang Wang, and Jialin Yang. Density prediction of selective laser sintering parts based on artificial neural network. In *International Symposium on Neural Networks*, pages 832–840. Springer, 2004.

[100] Sergey A Shevchik, Giulio Giulio Masinelli, Christoph Kenel, Christian Leinenbach, and Kilian Wasmer. Deep learning for in situ and real-time quality monitoring in additive manufacturing using acoustic emission. *IEEE Transactions on Industrial Informatics*, 2019.

[101] Sharanjit Singh, Vishal S Sharma, and Anish Sachdeva. Optimization and analysis of shrinkage in selective laser sintered polyamide parts. *Materials and Manufacturing Processes*, 27(6):707–714, 2012.

[102] Jacob Smith, Wei Xiong, Wentao Yan, Stephen Lin, Puikei Cheng, Orion L Kafka, Gregory J Wagner, Jian Cao, and Wing Kam Liu. Linking process, structure, property, and performance for metal-based additive manufacturing: computational approaches with experimental support. *Computational Mechanics*, 57(4):583–610, 2016.

[103] Thomas Stichel, Thomas Frick, Tobias Laumer, Felix Tenner, Tino Hausotte, Marion Merklein, and Michael Schmidt. A round robin study for selective laser sintering of polyamide 12: Microstructural origin of the mechanical properties. *Optics & Laser Technology*, 89: 31–40, 2017.

[104] Thomas Stichel, Thomas Frick, Tobias Laumer, Felix Tenner, Tino Hausotte, Marion Merklein, and Michael Schmidt. A round robin study for selective laser sintering of polymers: Back tracing of the pore morphology to the process parameters. *Journal of Materials Processing Technology*, 252:537–545, 2018.

[105] Stoyan Stoyanov and Chris Bailey. Machine learning for additive manufacturing of electronics. In *2017 40th International Spring Seminar on Electronics Technology (ISSE)*, pages 1–6. IEEE, 2017.

[106] M Szilvśi-Nagy and GY Matyasi. Analysis of stl files. *Mathematical and Computer Modelling*, 38(7-9):945–960, 2003.

[107] David Tasch, Anna Mad, Robert Stadlbauer, and Martin Schagerl. Thickness dependency of mechanical properties of laser-sintered polyamide lightweight structures. *Additive Manufacturing*, 23:25–33, 2018.

[108] Mohammad Taufik and Prashant K Jain. Role of build orientation in layered manufacturing: a review. *International Journal of Manufacturing Technology and Management*, 27(1-3):47–73, 2013.

[109] Kun Tong, Sanjay Joshi, and E Amine Lehtihet. Error compensation for fused deposition modeling (fdm) machine by correcting slice files. *Rapid Prototyping Journal*, 14(1):4–14, 2008.

[110] M Samie Tootooni, Ashley Dsouza, Ryan Donovan, Prahalad K Rao, Zhenyu James Kong, and Peter Borgesen. Classifying the dimensional variation in additive manufactured parts from laser-scanned three-dimensional point cloud data using machine learning approaches. *Journal of Manufacturing Science and Engineering*, 139(9):091005, 2017.

[111] Silvia Vock, Burghardt Klöden, Alexander Kirchner, Thomas Weißgärber, and Bernd Kieback. Powders for powder bed fusion: a review. *Progress in Additive Manufacturing*, pages 1–15, 2019.

[112] GC Vosniakos, T Maroulis, and D Pantelis. A method for optimizing process parameters in layer-based rapid prototyping. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 221(8):1329–1340, 2007.

[113] Rong-Ji Wang, Lingling Wang, Lihua Zhao, and Zijian Liu. Influence of process parameters on part shrinkage in sls. *The International Journal of Advanced Manufacturing Technology*, 33(5-6):498–504, 2007.

[114] Rong-Ji Wang, Jianbing Li, Fenghua Wang, Xinhua Li, and Qingding Wu. Ann model for the prediction of density in selective laser sintering. *International Journal of Manufacturing Research*, 4(3):362–373, 2009.

[115] K Wasmer, C Kenel, C Leinenbach, and SA Shevchik. In situ and real-time monitoring of powder-bed am by combining acoustic emission and artificial intelligence. In *International Conference on Additive Manufacturing in Products and Applications*, pages 200–209. Springer, 2017.

[116] Andreas Wegner and Gerd Witt. Understanding the decisive thermal processes in laser sintering of polyamide 12. In *AIP Conference Proceedings*, volume 1664, page 160004. AIP Publishing, 2015.

[117] Terry Wohlers. *Wohlers report 2016*. Wohlers Associates, Inc, 2016.

[118] Terry Wohlers. *Wohlers report 2019*. Wohlers Associates, Inc, 2019.

[119] H-J Yang, P-J Hwang, and S-H Lee. A study on shrinkage compensation of the sls process by using the taguchi method. *International Journal of Machine Tools and Manufacture*, 42(11):1203–1212, 2002.

[120] Dongsen Ye, Geok Soon Hong, Yingjie Zhang, Kunpeng Zhu, and Jerry Ying Hsi Fuh. Defect detection in selective laser melting technology by acoustic signals with deep belief networks. *The International Journal of Advanced Manufacturing Technology*, 96(5-8):2791–2801, 2018.

[121] Jerry Ye, Jyh-Herng Chow, Jiang Chen, and Zhaohui Zheng. Stochastic gradient boosted distributed decision trees. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 2061–2064. ACM, 2009.

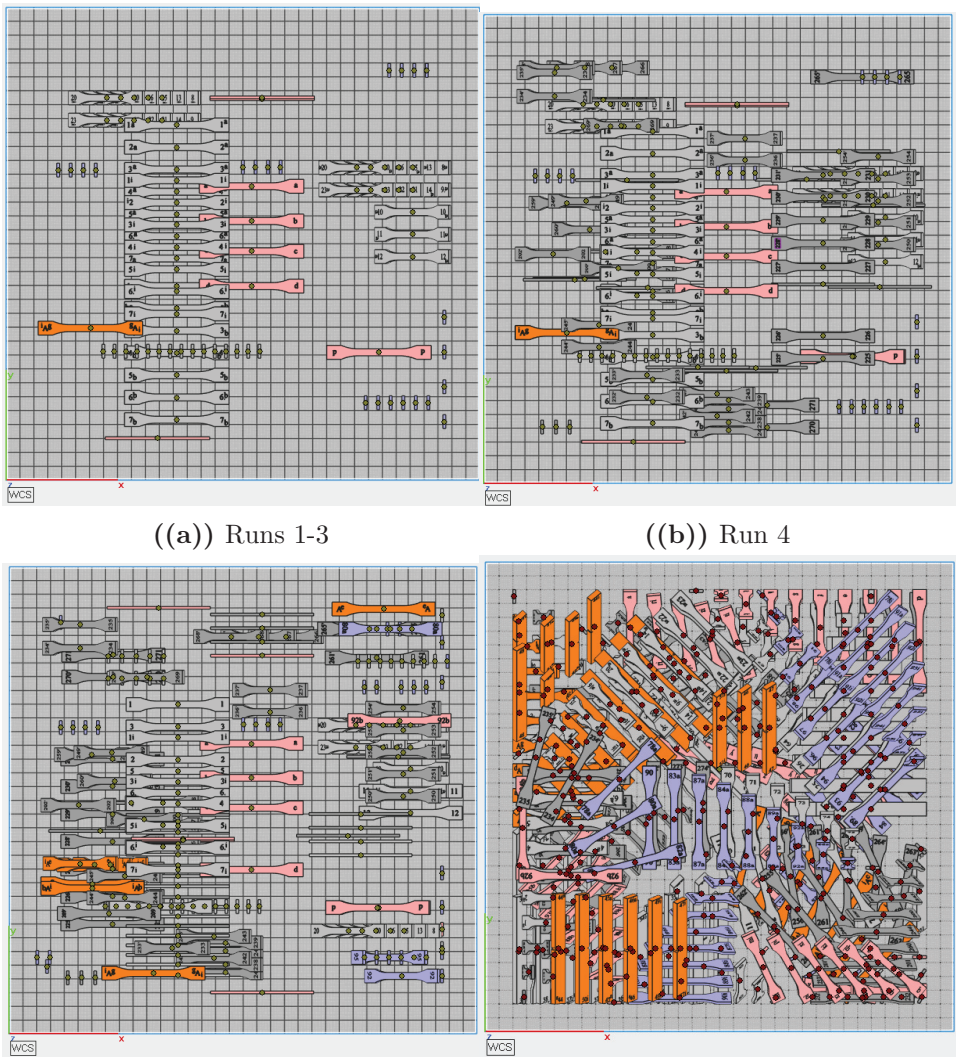
[122] Bodi Yuan, Gabriel M Guss, Aaron C Wilson, Stefan P Hau-Riege, Phillip J DePond, Sara McMains, Manyalibo J Matthews, and Brian Giera. Machine-learning-based monitoring of laser powder bed fusion. *Advanced Materials Technologies*, 3(12):1800136, 2018.

[123] K Zeng, N Patil, H Gu, H Gong, D Pal, T Starr, BE Stucker, et al. Layer by layer validation of geometrical accuracy in additive manufacturing processes. In *Proceedings of the Solid Freeform Fabrication Symposium, Austin, TX, Aug*, pages 12–14, 2013.

[124] Zuowei Zhu, Nabil Anwer, Qiang Huang, and Luc Mathieu. Machine learning in tolerancing for additive manufacturing. *CIRP Annals*, 67 (1):157–160, 2018.

Appendix A

Build layout design details from Magics 22.03



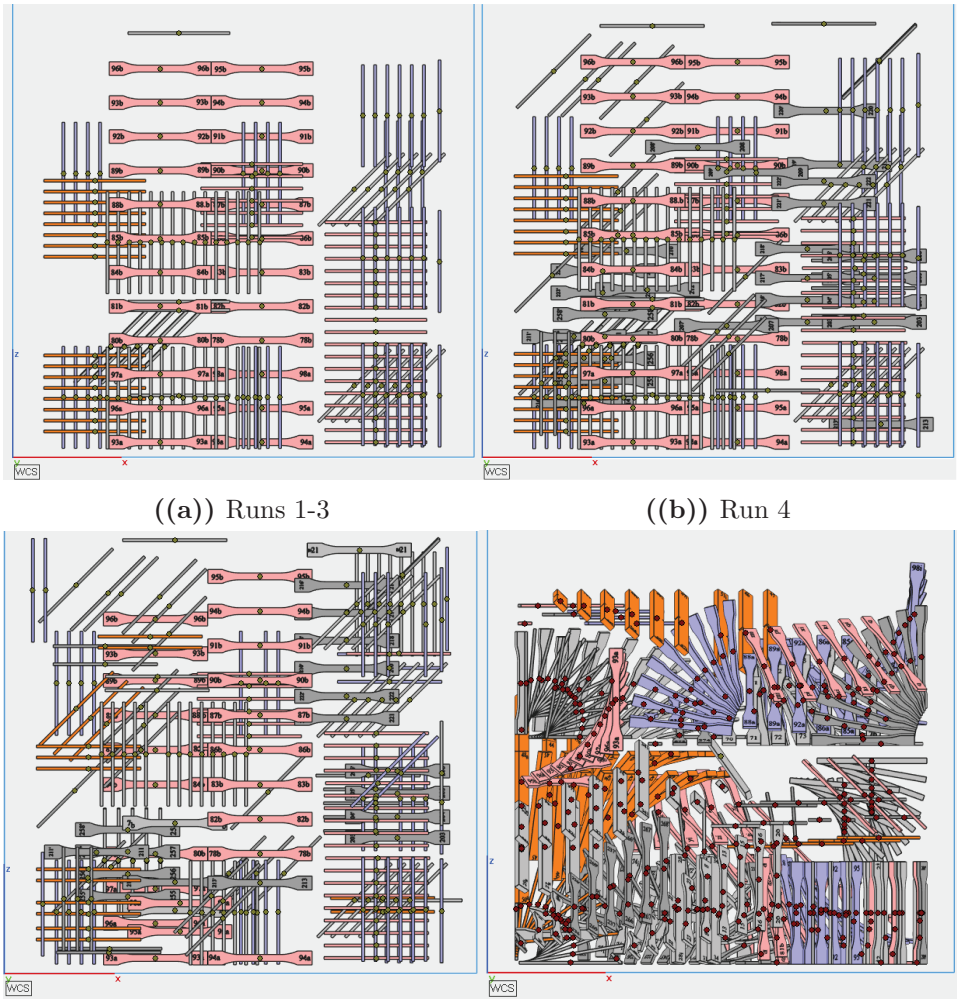
((a)) Runs 1-3

((b)) Run 4

((c)) Run 5

((d)) Run 6

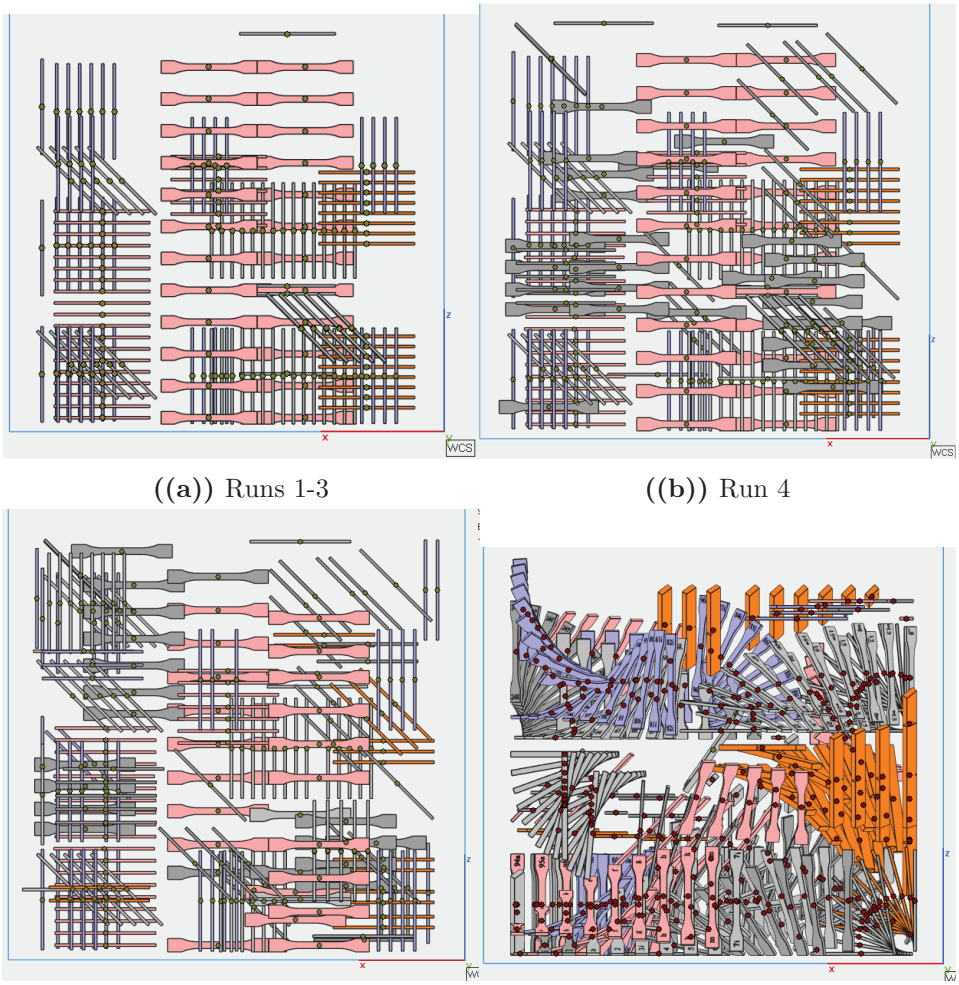
Figure A.1: Build layout design in Magics 22.03: top view



((a)) Runs 1-3 ((b)) Run 4

((c)) Run 5 ((d)) Run 6

Figure A.2: Build layout design in Magics 22.03: front view



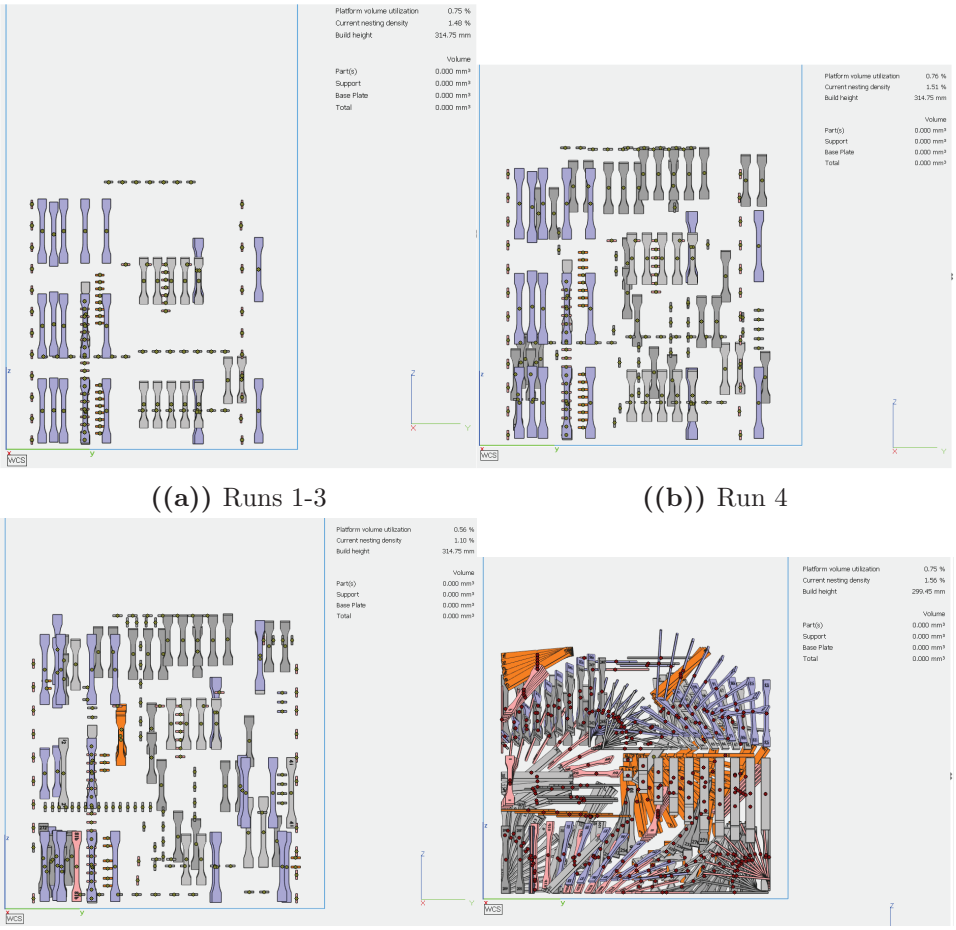
((a)) Runs 1-3

((b)) Run 4

((c)) Run 5

((d)) Run 6

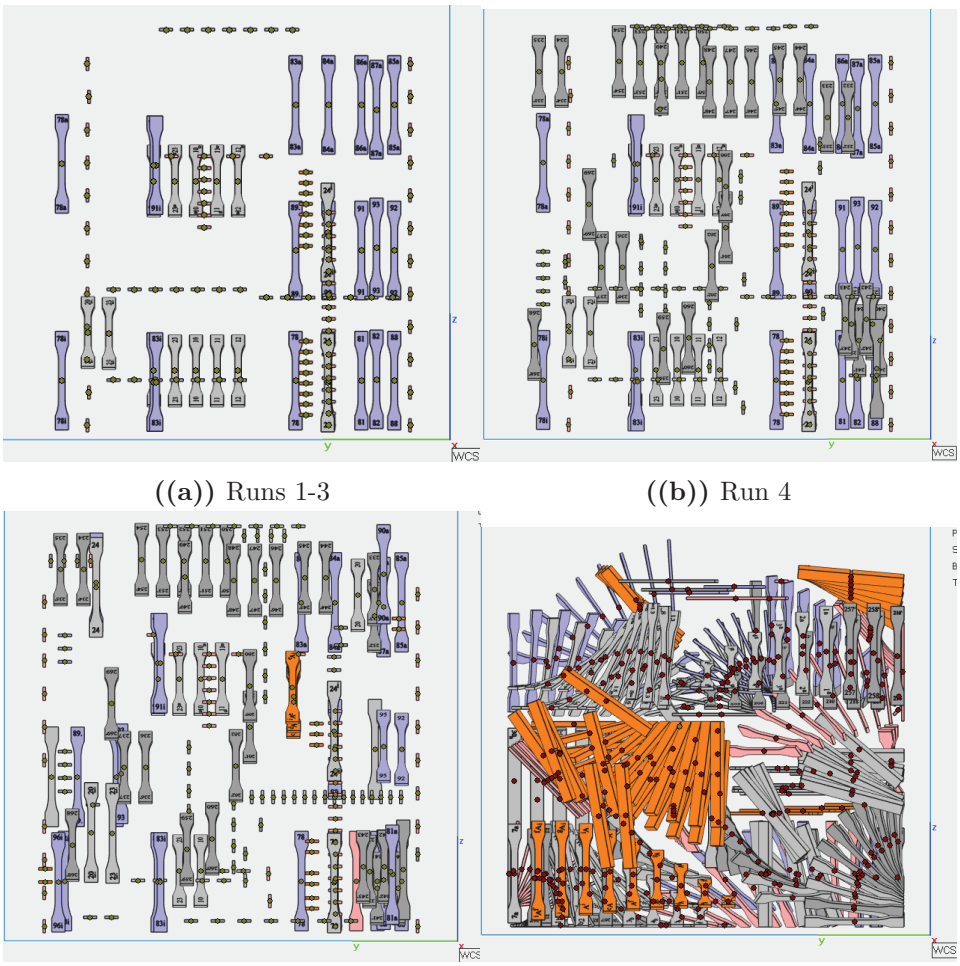
Figure A.3: Build layout design in Magics 22.03: back view



((a)) Runs 1-3 **((b))** Run 4

((c)) Run 5 **((d))** Run 6

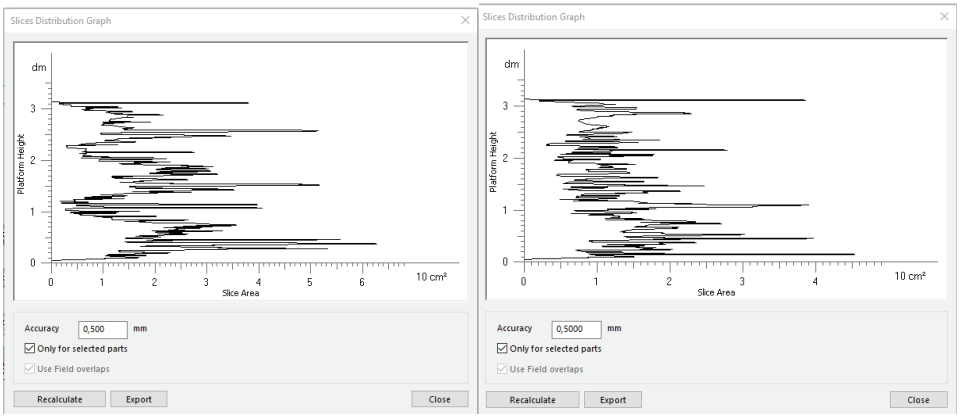
Figure A.4: Build layout design in Magics 22.03: right view



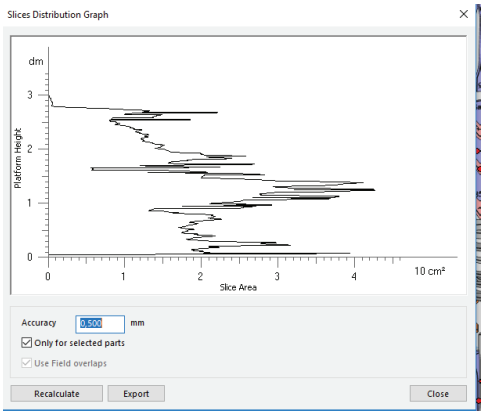
((a)) Runs 1-3 ((b)) Run 4

((c)) Run 5 ((d)) Run 6

Figure A.5: Build layout design in Magics 22.03: left view



((a)) Runs 1-3 ((b)) Run 4



((c)) Run 5

Figure A.6: Slice distribution for each build layout design in Magics 22.03

Appendix B

Application of model-based system engineering to design a technical system

B.1 Identification of Stakeholders and their needs

According to SPADE, stakeholders and their needs should be identify first. Stakeholders needs will define expectation and requirements to the system. If this step is not performed at the beginning of system development, it may lead to negative consequences later.

Since this PhD is a part of MKRAM project, all participants are stakeholders of the system designed by using System Engineering principles. Additionally, MKRAM project funding is provided by the Norwegian Research Council (NRC), and therefore, NRC is also a stakeholder of the system that is under development (Fig.B.1).

As it was described earlier, Additive Manufacturing is relatively new technology and lack of standards is one of the reasons why industry does not adopt this technology as a part of a production line. Another important issues are lack of understanding of the process, how to control it, how to design new shapes, how evaluate quality of nonstandard shapes and presence of inconsistency every time in the results. Therefore, results of this research can be used as a contribution to development of new standards, and therefore systematic approach is very important.

Hospitals are also in stakeholders list because they can benefit from the results of developed system. Nowadays, human bones are already fabricated with additive manufacturing, and one of the main requirements is dimensional accuracy with specific mechanical properties.

Visualization of relationships between system that is under development - Intelligent System for Quality Assurance in Additive Manufacturing - and stakeholders is presented on Fig. B.2. It is expected that requirements and data as an input are collected from the participants of MKRAM project, namely, NTNU, SINTEF and manufacturing companies. In return, knowledge about process behavior and how to control variations in quality, systematically organized data that is collected from the practical experiments, and results from data analysis will be provided to NTNU, SINTEF and manufacturing companies.

Since material properties have an impact on the quality of 3D printed parts, supplier of the material for additive manufacturing has also impact on the product and should be mentioned in the context of intelligent system. When tool for quality assurance is developed and adopted, customers will be a part

Stakeholders	Level of interest	Needs
NTNU	*Primary*	• new knowledge • process understanding • data storage for use in the future • repeatable results • new funding • student evaluation
SINTEF	*Primary*	• new knowledge • samples for mechanical testing • systematic approach to research • repeatability of results • process understanding
Manufacturing companies (GKN Aerospace, Konsberg Automotive, Nammo Raufoss, OM BE Plast, Sandvik Teenes)	*Primary*	• process understanding • adoption of additive manufacturing • high quality • minimize cost • positive contribution to environment • guidelines • consistency
The Norwegian Research Council	*Primary*	• new knowledge • sustainability aspects • share knowledge with academia • adoption of additive manufacturing
Standards/Regulations	*Secondary*	• systematic research (contribution to development of new standards)
Hospitals	*Secondary*	• high quality products • dimensional accuracy • fast design and product delivery (may save human life)
Natural environment, community and society	*Secondary*	• minimize negative impact on the environment

Figure B.1: Stakeholders and their needs

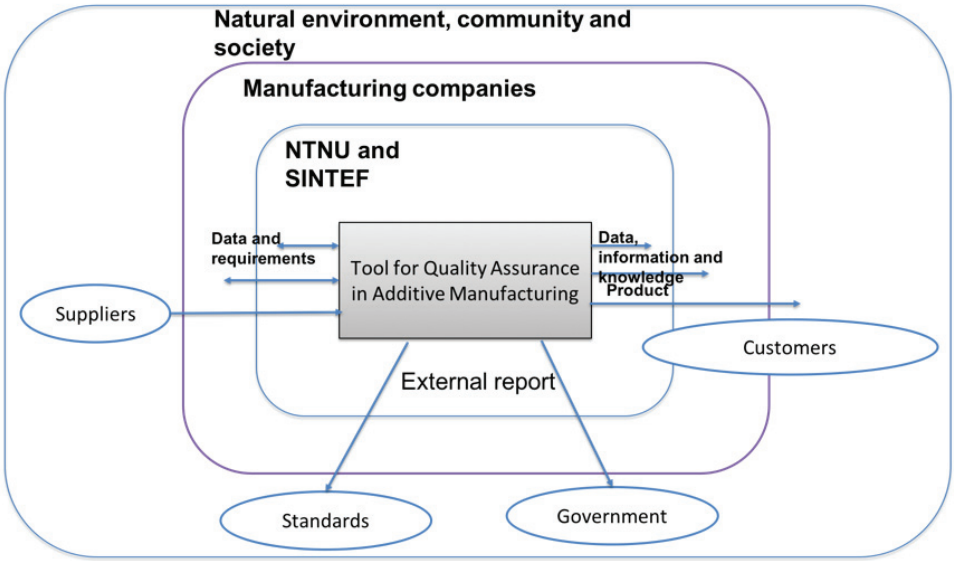


Figure B.2: Visualization of context model for Intelligent System for quality Assurance in Additive Manufacturing

of the system because they will be the users of the product fabricated with AM process. Hospitals are included in the list of customers, however they may also contribute to the requirements.

All results of this research will be available to the community and academia around the world via open source publications. Additional external reports will be a valuable contribution for development of new standards and policies (government).

B.2 Functional behavior analysis

System behavior is described in this work by using a concept of functional flow diagram that should be read from the left to the right [**?**]. General overview of how systems functions is shown on Fig.B.3.

The first step in the functional flow diagram shown on Fig.B.3 is when user makes first interaction with the system. Depending on the choice of the user, system can either do some changes to the data or do some sort of calculation. After this, either of the results should be accepted by the user, and user should be forward to the next stage.

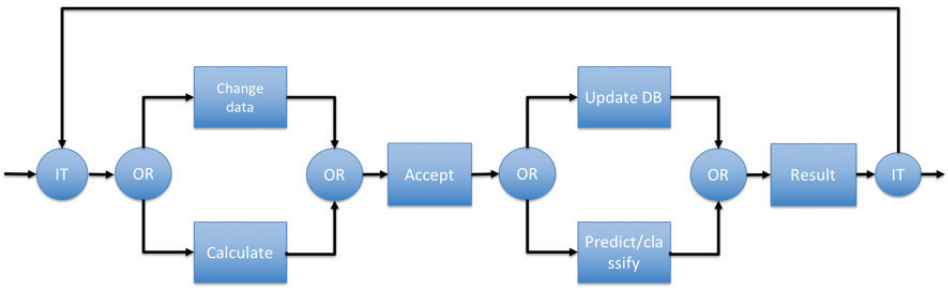


Figure B.3: Functional flow block diagram for Intelligent system for quality assurance in additive manufacturing: general overview

The next stage is dependent on the first choice of the user, if it was to change a data, then at this stage database should be updated. Otherwise, system should make a prediction or classification. Results should be shown to the user and if user is interested, (s)he will be redirected to the first page where (s)he started interaction with the system. All steps will be repeated again as many times as it is needed.

B.3 Architectural synthesis

Architectural synthesis is about development and application of an effective partitioning strategy for physical components with reference to behavioral model. Since, there is a limit in time, this part of model-based system engineering is not covered in this work. However, proposed visualization of the components illustrated on Fig.5.2 - Fig.5.6 should be used as guidelines.

B.4 Validation and verification

Validation and verification is very important step while designing system. It should be performed at each stage and each layer when model-based system engineering is applied. In this work, visualization of the components helped to verify that all functional requirements are described for every component of the system. Validation of functional requirements were also done by analysis of MKRAM project description and published articles that are relevant to the current research. However, this step should be continuously performed throughout the future system modification at each stage and each layer of intelligent system design.