Doctoral theses at NTNU, 2021:71

Olga Ogorodnyk

Towards Intelligent Process Control for Thermoplastics Injection Molding

Machine Learning for prediction of quality of injection molded parts

Doctoral thesis

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

NTNU
Norwegian University of
Science and Technology

Olga Ogorodnyk

Towards Intelligent Process Control for Thermoplastics Injection Molding

Machine Learning for prediction of quality of injection molded parts

Thesis for the Degree of Philosophiae Doctor

Gjøvik, March 2021

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 978-82-326-5810-7 (printed ver.)
ISBN 978-82-326-5906-7 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:71

Printed by NTNU Grafisk senter

To my mom Rozaliia Ogorodnyk and my dad Ievgen Ogorodnyk

Nobody ever figures out what life is all about, and it doesn't matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough.

– Richard P. Feynman

Abstract

Injection molding is a process that is widely used for production of plastic parts of different shapes, sizes and applications. Some of its advantages are high production rates and efficiency. However, to be able to produce parts of high quality and avoid in-process variations, high controllability and repeatability of the process are necessary. To achieve this, definition of critical process parameters that influence the final part quality and their proper values is necessary. Nowadays when starting production of a new part, these values are selected based on the injection molding machine operator's experience using trial and error. Such method might be highly inefficient, requires a lot of time and results in production of scrap. To avoid this, a new approach to monitoring and control of the injection molding process is necessary.

In this PhD dissertation a framework for development of an intelligent control system for thermoplastics injection molding is created using systems engineering approach and prototypes of the system's modules are developed. Such system would allow to turn the injection molding process data into manufacturing intelligence and allow to avoid trial and error method for selection of the proper process settings.

The dissertation aims to investigate the questions of (1) which injection molding process parameters are the most influential on different aspects of the injection molded parts' quality (dimensions and physical properties); (2) how to create prediction models of dimensional (width and thickness) and physical properties (Young's modulus (tensile modulus), tensile strength and tensile strain at break) of injection molded parts using Machine Learning methods; and (3) how to develop an intelligent control system for thermoplastics injection molding.

To answer the questions four experiments are conducted on "ENGEL insert 130" injection molding machine and 798 standard dogbone specimens are produced (72 of the specimens are with 15 mm thickness, while 726 are with 4 mm thickness). The specimens with 15 mm thickness are produced only during experiment 2. Virgin HDPE material is used in the first two experiments, while recycled material from two different suppliers is used in experiments 3 and 4. The experiments are conducted using designs of experiment (DOE). 8 parameters are included in the first experiment's DOE, 7 in the second one and 6 in the DOE used for both experiments 3 and 4. Machine and process parameters data is logged during the dogbone specimens' production: 41 parameters during experiment 1, 65 parameters during experiment 2 and 52 during experiments 3 and 4. To collect the data, a systematic data acquisition approach is proposed in the dissertation. It includes guidelines for collection of the data during and after the production process.

After the data collection is completed, data cleaning, integration, normalization and feature selection are performed in order to obtain high-quality datasets and as the main data preprocessing steps. The datasets are then used to develop machine learning models for prediction of dimensions and physical properties of the dogbone focus parts. The resulting prediction models are:

1. A Random Forest predictive model with $R^2 = 0.95$, RMSE = 0.05 and correlation coefficient = 0.94 for dimensional properties (width and thickness).

2. A Random Forest predictive model with $R^2 = 0.92$, RMSE = 28.32 and correlation coefficient = 0.97 for mechanical properties (tensile modulus, tensile strength, tensile strain at break).
3. A model for prediction of dimensional deviations developed using the Random Forest machine learning method with $R^2 = 0.92$, RMSE = 0.06 and correlation coefficient of 0.94.

The framework for development of an intelligent control system for thermoplastics injection molding is proposed based on the results obtained during the experimental work and analysis of the gathered data, as well as using the model-based systems engineering approach. The modules used for data acquisition, data preprocessing and the machine learning models creation are prototypes of the intelligent system's modules. Such system allows to gather the injection molding process data of interest, select the most influential process parameters using feature selection machine learning methods, as well as develop the quality prediction models. The models can be then used to select the proper process parameter values without use of the resource consuming trial and error method.

Acknowledgements

PhD is a long and lonely path, however, there are a lot of people that made it less lonely and possible for me to make it.

Big thanks to my main supervisor, Kristian Martinsen, for all the support, understanding and for giving me freedom in shaping my research throughout all my PhD years. I would also like to thank my informal co-supervisors from SINTEF Manufacturing AS: Ole Vidar Lyngstad and Mats Larsen. I have received a tremendous amount of help and knowledge from you with laboratory and equipment arrangements, as well as experimental work. Thanks to the Norwegian Research Council and the MegaMould project for funding my research work.

I would like to thank my colleagues/ friends at the Department of Manufacturing and Civil Engineering for providing a friendly working environment and interesting discussions, especially Nina Tvenge, Jo Sterten, Ivanna Baturynska and Oleksandr Seminiuta. Thanks to Torbjørn Skogsrød and Iver Eugen Jensen for the great leadership assistance. Thank you to Tor Erik Nicolaisen for timely and professional laboratory management. Thanks to Chunhong Luo and Mary Hailemariam for the top-notch administrative support.

Thanks to all my friends, especially Nataliia Synko, Daria Ganchukova, Oleksandra Yakovlieva, Malin Victoria Hagebakken, Anders Hagebakken, Radina Stoykova, Anastasiia Moldavska, Christoffer Vargtass Hallstensen, Gina Karoline Dalen, Christina Marie Mitcheltree, Juan Victor Abreu-Peralta, Paul Kengfai Wan, Andrii Shalaginov, Marina Shalaginova, Sri Sudha Vijay Keshav Kolla, Nicolas Ghanbarpour, Anneilen Ytterstad and Irina Bjerkvold.

Thank you to Sergii Banin for always being there by my side in good and bad times.

Thank you to my family, my brother Oleksandr Vovk, my nephew Ievgen Vovk and my sister-in-law Larysa Vovk. Most of all thank you to my mom Rozaliia Ogorodnyk and my dad Ievgen Ogorodnyk for all the love, care, encouragement, for believing in me and helping me to get all the way here. Papa, I miss you every day.

Olga Ogorodnyk, Gjøvik, March 2021

Contents

List of Tables

List of Figures

List of Abbreviations

Adaptive boosting (AdaBoost)

Artificial intelligence (AI)

Artificial neural network (ANN)

Analysis of variance (ANOVA)

Application programing interface (API)

Back-propagation neural network (BPNN)

Computational core (CC)

Correlation-based feature selection (CFS)

Coordinate measurement machine (CMM)

Cyber-physical systems (CPS)

Design of experiment (DOE)

Finite element analysis (FEA)

Finite element method (FEM)

Feature selection (FS)

Genetic algorithm (GA)

Gradient boosting (GBR)

High-density polyethylene (HDPE)

Injection molding (IM)

Injection molding machine (IMM)

Latin hypercube sampling (LHS)

Mean absolute error (MAE)

Model-based systems engineering (MBSE)

Manufacturing Execution Systems (MES)

Machine Learning (ML)

Multilayer Perceptron (MLP)

Mean squared error (MSE)

Non-dominated sorted genetic algorithm (NSGA)

Principal component analysis (PCA)

Particle swarm optimization (PSO)

Pressure-volume-temperature (P–V–T)

Recursive feature elimination (RFE)

Root mean squared error (RMSE)

Response surface methodology (RSM)

Sustainable development goals (SDGs)

Self-organizing map (SOM)

Statistical Process Control (SPC)

Thermoplastics injection molding (TIM)

Chapter 1

Introduction

Study hard what interests you the most in the most undisciplined, irreverent and original manner possible.

– Richard P. Feynman

This chapter gives an introduction to the dissertation. Section 1.1 describes motivation and objectives for the work. Section 1.2 gives a short description of the MegaMould project that this research is a part of. Next, Section 1.3 provides research questions that this thesis is aimed at answering, while Section 1.4 is dedicated to the explanation of scope and boundaries of the PhD project. Section 1.5 outlines the main contributions of this work. Finally, the outline of the dissertation is presented in Section 1.6.

1.1 Motivation and objectives

Nowadays, about 30% of polymeric products are manufactured using injection molding (IM) [1]. It is highly possible that you are surrounded by several injection molded products right now: body of a pen you write with, plastic shell of a computer mouse you used to scroll through a webpage, your lunch box, etc. Injection molding is widely used for mass production of plastic parts of different sizes, geometries, materials and applications. High controllability and repeatability of the process become critically important in this case, as they allow to produce quality products, decrease the scrap rate and energy consumption.

Injection molding includes the following stages: plasticization, injection, packing, cooling and ejection. It might sound simple, but it is a sophisticated process, during which molten polymers undergo significant thermo-mechanical changes that influence quality of the obtained part [1]. During the process, plastic pellets are fed into a heated barrel through the hopper, mixed using a reciprocating screw and then injected into a mold cavity, where the molten material cools and hardens into a shape based on the cavity configuration [2]. To get an injection molded product of a desired quality, it is important to consider several aspects: properties of the chosen material, mold geometry, condition of the injection molding machine and the mold in use, as well as values of the relevant process parameters' [3, 4]. While all the mentioned aspects are equally important for manufacturing of a part free from defects, this PhD dissertation focuses on the latter.

Due to the high complexity of the process, it is often hard to choose optimal parameter values when starting production of a thermoplastic part. The process parameters' selection is often based on the experience of the injection molding machine operators, who use trial-and-error method to define the parameter values [4, 5]. Inability to choose proper settings of process parameters will result in various defects in the obtained product and manufacturing of scrap rather than high quality products [6, 7]. Moreover, it might not always be clear which specifically process parameters need to be tuned, as a different subset of parameters is responsible for producing specific type of a defect and different

scholars argue different process parameters to be the most influential [8]. Therefore, understanding which of the set parameter values are responsible for getting an unsatisfactory output, as well as control of in-process variations are important tasks within the field of thermoplastics injection molding (TIM).

Necessity to have a repeatable and controllable process, definition of the most important process parameters and use of trial-and-error method are challenges relevant not only for the injection molding area. Such industrial processes as machining, additive manufacturing, extrusion, forming also strive to achieve the highest quality of the products and control the production process to the highest extent possible. A significant number of studies stress importance of collection and analysis of data from different manufacturing processes [9] to achieve implementation of self-optimizing machines that would not require use of trial-and-error method to improve the process performance. Industry 4.0, cyber-physical systems (CPS), intelligent control, data-driven methods for prediction and classification are some of the concepts assisting and pushing forward this development. Increasing demands on production quality, system performance and economic feasibility lead to demand of development within such topics as prediction and optimization of production system behavior in real time [10]. However, despite all the ongoing research, Industry 4.0 is still regarded by many as an immature concept [11, 12], while self-optimizing machines remain far from industrial implementation passively listening to commands from operators even when provided parameter settings that are far from optimal [13]. As a result, further research on cutting-edge digital technologies and their application for manufacturing is of high importance [14].

At the same time, the amount of data collected in manufacturing is growing and if its quality is high, it can be further converted into manufacturing intelligence to impose positive effects on all aspects of manufacturing [15, 16]. However, there are several challenges that arise while collecting the manufacturing process data such as: inability to synchronize logging of data from machine built-in and third-party sensors, limitations of manufacturing execution systems (MES) in terms of access to certain parameters data, infrequent sampling rate, lack of suitable communication protocols, etc. [17]. In addition, even when the meaningful data is acquired, a significant number of companies continue using conventional non-data-driven techniques to monitor and control the production processes, thus, devaluing both the data collection efforts and the data itself.

Taking into consideration all the above-mentioned, it is possible to identify some of the important challenges within the field of thermoplastics injection molding:

- Necessity to have controllable and repeatable process to avoid unnecessary in-process and quality variations;
- Application of trial-and-error method when starting production of a new part or tuning the process;
- Definition of the critical process parameters that influence the final part quality;
- Acquisition of the process data and turning this data into the manufacturing intelligence through the corresponding data analysis.

Most of these challenges can be solved by implementation of proper monitoring and control routines and systems. However, when it comes to monitoring and control for the

injection molding, commonly used methods are usually rather simple and are not learning from the data they are processing [9, 13, 18]. One of the most conventional methods for control of the injection molded parts quality is statistical process control (SPC) [19]. However, it is not guaranteed that the locally established SPC routine for an injection molding machine is able to distinguish good and bad parts with high enough success rate [20]. This is due to a reason that SPC, as well as control routines and models created using other conventional algorithms are not being updated based on the real time data but use pre-defined procedures instead. In addition, it might be difficult to know which parameters need to be adjusted to ensure production of the parts with high quality.

Models created with help of data-driven methods, such as Machine Learning (ML), on the other hand, can provide necessary flexibility and robustness for different production scenarios [21]. ML methods have been successfully applied to intrusion detection [22], classification of cardiovascular diseases [23], robotics [24], supply chain demand forecasting [25], manufacturing industry [26] and many more fields. There are also examples of application of the ML techniques to the injection molding, for example, to construct a model of the injection molding process using artificial neural network (ANN) [27], to improve repeatability and product quality through utilization of a model predictive controller combined with an ANN [28] and to investigate the cavity filling process [29].These methods are proven to work well, when there is a large number of parameters that need to be taken into account, have higher precision and shorter model training time in comparison to conventional optimization techniques [30]. At the same time, ML methods are able to give a better performance in comparison to conventional statistical methods and are capable of coping with high level of complexity of the mathematical models [31]. Regular regression models require certain physical and mathematical knowledge about the process to be able to construct a model, while data-driven methods can extract process information from historical data, as well as from online process data [32, 33]. This will allow to increase controllability and repeatability of the overall process, leading to lowering probability of unnecessary in-process variations.

It is important to mention that before applying the ML methods, necessary amount of data needs to be acquired. As suggested by scikit-learn's "Machine Learning Map" [34] (scikit-learn is one of the biggest open source Python programming language libraries for ML methods), it is not recommended to use any of the machine learning methods, if the amount of data points (samples) used to train the model is less than 50. However, some studies attempt to apply ML methods to smaller datasets, as for example, in [35]. This is due to a fact that it is not always easy to get the necessary amount of data.

To solve the above-mentioned challenges an intelligent control system for thermoplastics injection molding needs to be developed. Such system would allow to increase controllability and repeatability of the overall process, decrease need in use of trial and error method and predict quality of the produced parts based on the set of current process parameters. ML methods are the meaningful candidates for application within such system in order to train the regression models for prediction of quality features such as dimensions and physical properties of the produced parts.

The work presented in this thesis describes a framework for development of an intelligent control system for thermoplastics injection molding. The system includes data acquisition and transformation of this data into manufacturing intelligence to minimize variations in produced parts. An important goal of this work is determining which machine learning techniques are best suited for development and training of models for prediction of quality of the injection molding products.

1.2 The MegaMould project

This PhD thesis describes research conducted as part of the Norwegian innovation project *MegaMould* supported by the Norwegian Research Council (project number: 256819). The MegaMould project is focusing on development of innovative processes, tools and methods for injection molding of extra-large thermoplastic components. The goal of the project is to achieve controllable and repeatable process for manufacturing of high-quality products. When talking about quality, the main attention is payed to mechanical properties and dimensions of parts within the strict tolerance specifications. Materials in focus of the MegaMould are virgin and recycled thermoplastics and thermoplastic composites. This PhD work shows how machine learning methods can be applied to the injection molding process data in order to build quality prediction models and achieve repeatable and controllable production process.

The MegaMould project participants are Plasto AS, PipeLife Norge AS and AKVA Group ASA, as well as SINTEF Manufactruing AS and NTNU research institutions. The industrial partners have contributed to the MegaMould project with industrial competence and particular cases drawn from their manufacturing sites. SINTEF Manufacturing AS and NTNU, as R&D partners, were contributing with theoretical and scientific knowledge.

Production of parts used as focus parts within the PhD project and collection of the injection molding process data was conducted at the laboratory at SINTEF Manufacturing AS. A great deal of practical undertakings underlying this work were performed with significant support from members of the Materials Technology and Production Technology departments at SINTEF Manufacturing AS. Collection of the parts quality data was done at NTNU's laboratory facilities.

1.3 Research questions

This dissertation focuses on answering the following research questions:

RQ1. How to select parameters that influence the injection molded part quality using the ML methods?

This research question aims at explaining how machine learning techniques can assist in selecting the most important process parameters influencing quality of the final injection molded part. Injection molding machines (IMMs) operate using a significant number of machine and process parameters that play an important role in the quality of the manufactured product, such as: holding pressure, holding pressure time, barrel temperature, screw speed, etc. However, not all of them are equally important. Therefore,

only some of them need to be included in the models for prediction of the parts quality. Feature selection methods are a subset of the ML methods that can assist in selecting the process parameters that need to be payed special attention to.

RQ2. How can machine learning methods be used for prediction of dimensions of the injection molded parts?

The second research question is focused on addressing the desire of achieving a repeatable process that produces parts with high accuracy within the tolerance specifications of dimensions (width and thickness). It is aimed at providing prediction models trained by the ML methods and based on the collected historical data. Such models provide possibility of finding hidden patterns in the data obtained from production of parts within a particular production system. Unlike traditional mathematical models that focus on generalized description of the process that might not always correspond to a particular case, the prediction models created with help of the ML techniques are based on data collected from a certain production facility and are better describing peculiarities of the process in question.

RQ3. How can machine learning methods be used for prediction of physical properties of the injection molded parts?

Similarly to the RQ2, this research question is related to achieving controllable process that results in production of parts with desired physical properties, namely tensile modulus (Young's modulus), tensile strength and tensile strain at break. The ML methods are applied to train models capable of predicting the physical properties based on the chosen process parameter values.

RQ4. How to create an intelligent control system for thermoplastics injection molding?

The last research question aims at describing how an intelligent control system for thermoplastics injection molding can be created and what is its core functionality. A framework for development of the corresponding system is presented and its boundaries are outlined. A detailed description of different modules of the system is provided. The question also includes explanation of how selection of significant process parameters (RQ1), prediction of models for dimensions and physical properties of final parts (RQ2 and RQ3) contribute to development of such system.

1.4 Scope of the dissertation

Due to the limited timeframe of this research, there is a need to set boundaries of the PhD project.

This study is limited to the injection molding of thermoplastics with the high-density polyethylene (HDPE) as material in focus, namely virgin BorSafe^{TM} HDPE and recycled HDPE from RePro and ContainerService suppliers. All the experiments were conducted on "ENGEL insert 130" vertical injection molding machine with CC300 control unit available at SINTEF Manufacturing AS. Use of other materials and machines is out of scope of this PhD project.

The focus parts used in the project are dogbone specimens of type 1A based on the ISO 527-2:2012 Tensile Properties of Plastics [36] with 4 mm and 15 mm thicknesses. Even though injection molding can be used for production of parts of various geometries, the specimens were chosen due to the molds available in the laboratory. The reason for choosing two different thicknesses is to collect data for parts with regular wall thickness (4 mm) and parts with thick walls (15 mm). Another reason for choosing the specimens as the focus parts is that they fit well for determination of both dimensions and physical properties.

The predictive models developed within the study will include data of only those parameters that can be accessed on the IMM used in the experiments and sensors (if any) installed in the corresponding molds. When producing specimens with the 15 mm thickness, pressure and temperature data will be also logged from multi-sensors installed in the mold. Mold for manufacturing of the 4 mm thick specimens does not include any sensors, therefore no mold data can be acquired.

The developed models cannot be directly used in manufacturing. These models are considered as prototypes used for demonstration and establishment of the necessary routines, such as data collection, preprocessing and analysis. In addition, they play a significant role in understanding which ML techniques can be efficiently used to benefit the controllability of the injection molding process.

This PhD project mentions but does not include development of a database prototype or solution for storing the acquired data and models, as this is out of scope of this dissertation.

1.5 Contributions of the thesis

This dissertation describes the following contributions:

- Data from 435 IMM runs and corresponding 798 dogbone specimens produced has been collected through the experimental work. No similar data has been publicly available before. The produced dataset can be used by other researchers as a starting point for development of predictive models for thermoplastics injection molding.
- A data acquisition, preprocessing and analysis approach has been proposed for collection of data during and after the injection molding of thermoplastic parts (process data and quality data).
- A prototype of a data acquisition system for logging of injection molding machine and mold process data has been developed. Such system can be used by researchers, IMM producers and injection molding companies for collection of the production process near-real-time data.
- Design of a framework and functional requirements for development of an intelligent control system for thermoplastics injection molding is proposed.

- A prototype of the intelligent control system module for determination of the most influential injection molding process parameters has been proposed using the feature selection methods as a subset of ML techniques.
- Machine learning models are developed for prediction of dimensions (width and thickness) of injection molded dogbone specimens with 4 mm and 15 mm thickness produced during four different experiments using virgin or recycled HDPE depending on the experiment number. A prototype of the corresponding intelligent control system module has been developed based on these models.
- Machine learning models are proposed for prediction of physical properties (tensile modulus (Young's modulus), tensile strength and tensile strain at break). A prototype of the intelligent control system has been developed based on these models.

1.6 Thesis outline

The outline of the dissertation is given below:

Chapter 1 presents motivation, challenges and research questions identified within this research work. A short description of the MegaMould project, which this PhD project is a part of, is provided, along with the scope and contributions of the dissertation.

Chapter 2 includes theoretical background on the topics of thermoplastics injection molding, main defects occurring during the production process and state-of-the-art on the topic of monitoring and control for the IM.

Chapter 3 describes foundations of data science and ML for their successful implementation within the area of thermoplastics injection molding.

Chapter 4 describes research philosophy, design of the presented study and methods used to conduct the relevant experiments, data collection and analysis.

Chapter 5 shows how the systems engineering approach was used for designing an intelligent control system for thermoplastics injection molding.

Chapter 6 describes the procedure followed for development of predictive models for width and thickness dimensional variables using general datasets, as well as presents the quality characteristics of the obtained models.

Chapter 7 presents the procedure for development of predictive models for Young's modulus (tensile modulus), tensile strength and tensile strain at break using the general datasets. Quality characteristics (R^2, RMSE, correlation coefficient) of the models are also described.

Chapter 8 describes the development procedure and presents quality scores of predictive models trained on data series datasets for dimensional quality variables.

Chapter 9 shows how predictive models for mechanical properties trained on the corresponding data series datasets were created. The models' quality characteristics are also presented.

Chapter 10 presents models for prediction of dimensional properties deviations and their quality scores.

Chapter 11 discusses the results presented in the PhD thesis and the research questions that were raised in its beginning.

Chapter 12 concludes the dissertation and presents the necessary future steps.

Chapter 2

Theoretical background and literature review

Though alone, he was not lost.

– Jack London, "Love of Life"

This chapter is divided into two parts. Section 2.1 provides theoretical background on the topic of thermoplastics injection molding, main defects that occur during the process and their causes, as well as introduction to material's mechanical properties. Section 2.2 includes a brief literature review on the topic of monitoring and control of the injection molding process. Methodology used for conducting the literature review is described in Chapter 4, Subsection 4.2.1.

2.1 Injection molding process foundations

Injection molding is a cyclic manufacturing process for production of identical parts through injection of molten material into a mold of a chosen shape [37]. The main advantage of this process is its ability to produce high volumes of repetitively fabricated parts of complex geometries and various sizes, from the smallest components to entire body panels of cars [37]. In general, injection molding can be performed with different types of materials including metals (the process is then called die-casting), glasses, elastomers, confections and most commonly thermoplastic or thermosetting polymers. This thesis, in its turn, focuses on thermoplastics injection molding. Thermoplastics are highly suitable for injection molding due to their characteristics, such as ease of recycling and versatility for the wide application range [38].

2.1.1 Thermoplastics injection molding

The history of the injection molding goes back to the "packing machine" invented by the Hyatt brothers, who received a patent in 1872 for the invention of a machine that was able to mold camphor-plasticized cellulose nitrate [39]. The machine used a plunger to inject material into a mold through a heated cylinder. The industry kept on developing over the years, while producing buttons and hair combs. In 1919 Arthur Eichengrün developed the first injection molding press and in 1939 he patented the injection molding of plasticized cellulose acetate [40]. The industry expanded in the 1940s because of the World War II and its demand for inexpensive, mass-produced items. In 1946 James Watson Hendry built the first screw injection molding machine, where the plunger device was substituted by a screw. This machine provided more precise control over the speed of injection and quality of the final parts [41]. The screw injection machine allowed to mix material before injection, leading to possibility of adding colored or recycled material to virgin material and mixing them. Later, several improvements led to development of a reciprocating screw injection molding machine that is widely used nowadays [37] and to significant evolution of the injection molding industry that went all the way from producing combs

and buttons to production of a vast variety of items for multiple industries including automotive, aerospace, construction and packaging [42].

A reciprocating screw injection molding machine includes the following essential components: a hopper, a rotating screw, a heated barrel and a clamping unit containing the mold typically made of two halves. A mold is usually produced by a mold-maker commonly from steel or aluminum and is further machined to precisely fit characteristics of a part to be produced. Molds might have one or multiple cavities of the same or different geometries. Figure 2.1 depicts a sketch of an injection molding machine with a reciprocating screw.

During the injection molding cycle, the following main stages are usually distinguished: plasticization, injection, packing, cooling and ejection [2, 37]. During the **plasticization stage**, plastic pallets are fed into a heated barrel through the hopper. When the pallets have entered the barrel, a reciprocating screw rotates forcing the granules against the walls of the barrel melting, mixing and homogenizing them. The plastic pallets melt due to both conduction from the heating units along the barrel and friction heat created by the screw rotation. The molten material is then moved towards the tip of the screw, while the pressure is being formed against the "closed-off" nozzle. The screw then moves backward to accumulate enough melt at the front of the barrel into a volume known as a shot. A ***shot*** is the material volume enough to fill the mold cavity, compensate for shrinkage and provide a ***cushion*** (usually about 10% of the shot volume, which remains in the barrel and prevents the screw from bottoming out). When enough molten material is accumulated, the screw rotation stops, and the plasticization stage is over.

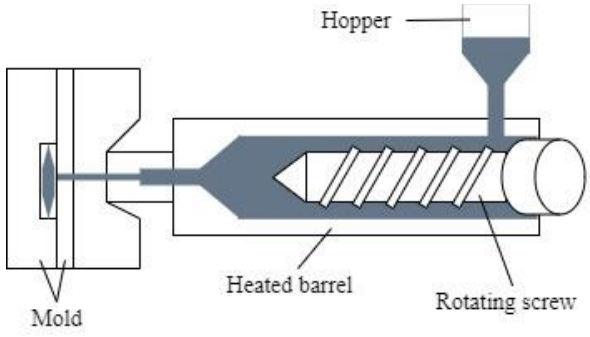


Figure 2.1. A sketch of a reciprocating screw injection molding machine

The next stage is called **injection or filling stage**. Here the clamp unit keeps the mold closed, while the screw moves forward and forces the molten material into the mold cavity at high pressure and velocity. When the polymer melt is injected from the nozzle of the IMM, it flows through the sprue, runner, gate and only then enters the mold cavity. The feed system is a term used for the sprue, runner and gate together. An example of a feed system is shown on Figure 2.2. Runners can be of two types: cold and hot ones. The mold temperature for the cold runner is similar to that in the mold cavity, while the hot runner maintains polymer at the melt temperature.

To prevent pressure spikes during the material injection, the process usually uses a ***switchover point*** (cavity is 95-98% full) to switch from the constant velocity to the constant pressure control. Once the screw switchover position is reached, the packing or holding pressure is applied and the **packing or holding stage** begins**.** When the mold is filled, the screw stays in the forward position or keeps moving with a small displacement to maintain the necessary holding pressure to complete the mold filling and compensate for the material's thermal shrinkage. The material cools down and shrinks allowing a little more material to enter the cavity. The holding pressure is applied until the gate (cavity entrance) solidifies and no more material can enter the cavity.

When the gate has completely frozen, the cavity pressure is being reduced to zero or a low value. The part continues to cool down and solidify, while the screw starts rotating again and moves backwards to start the new plasticization stage. This stage is called **cooling**. After keeping the part in the mold for a sufficient time allowing it to solidify, the mold is opened, and the part is ejected (**ejection stage**). After ejecting the part, the mold closes again, and the new injection molding cycle begins.

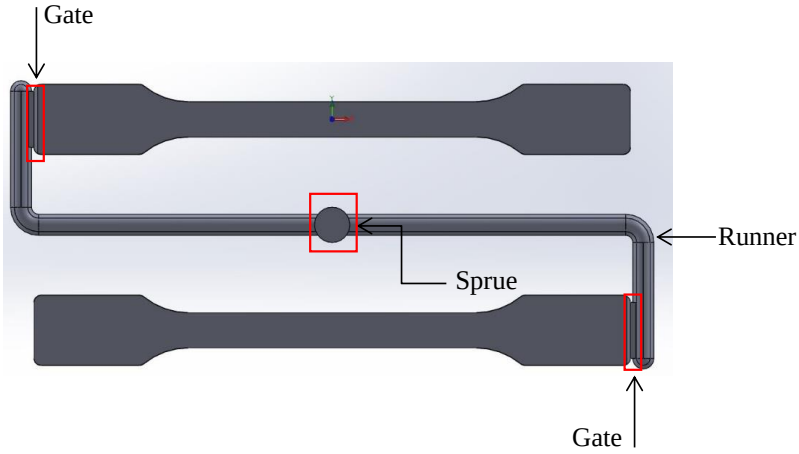


Figure 2.2. A feed system

2.1.2 Injection molding defects and their possible causes

Any manufacturing process has production of high-quality items as its focus. However, Reeves and Bednar [43] state that it is hard to find a universal definition to the term quality and different interpretations of this concept are appropriate depending on the application field. Quality is defined depending on user's needs and neither large nor small variations are good. According to Zheng, Tanner [37], two key problems of the injection molding processing are production of parts with desired (1) dimensional tolerances and surface and (2) mechanical, thermal, optical and other important properties. Therefore, in case of the injection molding process, high-quality part can be defined as a part that has values of the above-mentioned properties that comply with the corresponding tolerance specifications. However, depending on the product's application field, the main focus

might be dedicated to a different property, for example, to visual characteristics, rather than mechanical properties or the other way around. As mentioned above, this PhD dissertation is focusing on dimensions and physical properties of the injection molded parts.

According to a study by Texas Plastic Technologies that was carried out during a 30-years period (from 1963 to 1993), the most common general causes of the defects are machine, mold, material and operator [44]. Figure 2.3 shows distribution of the main defect causes. As it is easy to see, 60% of the defects are caused by improper setting of the machine and process parameters.

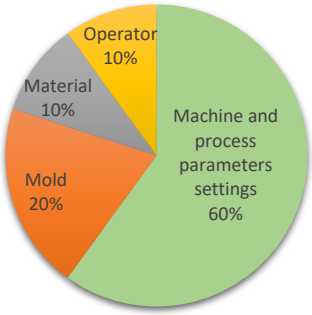


Figure 2.3. Distributions of the main defect causes

Due to the thermo-mechanical changes applied to the material during the process and a significant number of process parameters that influence quality characteristics of the obtained parts, there are several defects that may occur during the injection molding cycle. Table 4.1 shows the list of the possible injection molding defects and their causes in terms of process parameters setting and chosen material. The table does not cover defect causes related to product or mold design flaws, utilization of an unsuitable IMM or operator errors. Table 4.1 is based on the information available on the web-pages of several injection molding companies Toray Plastics [45], Polyplastics [46], GPTmold [47], Ecomolding [48], Bryce [44] and [49].

Some of the defect causes not mentioned in the table might be erratic cycling due to operator's error, poor mold ventilation, not optimal size and positioning of gates, use of machine that is too big or too small for production of a particular product, improper ejection of molded part and a simple need for cleaning of a mold or machine nozzle. In addition, the product itself might not be designed in an optimal way causing defects in the final part. All of these possible causes need to be considered, as well as selection of proper machine and process parameters when troubleshooting the injection molding process.

Table 2.1. Injection molding defects and their causes

Defect name	Description	Causes	Possible countermeasures
Short shot (nonfill)	A molded product is incomplete, because the mold cavity was not fully filled.	1. Material flowability too low, viscosity too high. 2. Material or mold temperature too low. 3. Injection speed and pressure too low. 4. Shot volume too small. 5. Nozzle clogging.	1. Use material with higher flowability and lower viscosity. 2. Increase barrel and/or nozzle temperature. 3. Increase mold temperature. 4. Increase injection speed and/or pressure. 5. Increase shot volume. 6. If the frequent nozzle clogging occurs, increase the mold temperature or shorten the cycle time.
Flash (burrs)	A thin layer of plastic that flows outside the mold cavity, where the two halves of the mold meet.	1. Material flowability too high, viscosity too low. 2. The mold is forced open due to insufficient clamping. 3. Mold temperature too high. 4. Barrel temperature too high. 5. Injection pressure too high. 6. Injection speed too high. 7. Holding pressure or holding pressure time not chosen appropriately. 8. Backpressure or screw speed too high.	1. Use the stiffest material possible without causing nonfill or short shot. 2. Increase clamping force. 3. Decrease mold temperature. 4. Decrease barrel temperature. 5. Decrease injection pressure. 6. Decrease injection speed. 7. Adjust holding pressure or holding pressure time. 8. Decrease backpressure or screw speed.
Voids (bubbles)	The residual voids sealed up in the product.	1. Material is too moist. 2. Changes in material density as material changes from a molten state. 3. The surface of the molded product rapidly loses heat, causing shrinkage concentration in the thickest parts of the product. 4. Injection pressure too low. 5. Injection time too short. 6. Holding pressure too low.	1. Dry material before molding. 2. Increase injection pressure. 3. Increase injection time. 4. Decrease barrel temperature. 5. Decrease mold temperature. Make sure that the mold temperature is uniform. 6. Increase holding pressure.
Sink marks	A defect that mostly appears in the regions of thick features, where in some areas, shrinkage might be higher than in the rest of the product.	1. Shot volume too small. 2. Shrinkage occurring due to thick walls of the product, rapid cooling of the surface and slow cooling of the inside of the part. 3. Melt temperature too high. 4. Mold temperature too high. 5. Holding pressure or holding pressure time too low.	1. Increase shot volume. 2. Increase injection pressure. 3. Increase injection time. 4. Decrease barrel temperature. 5. Decrease mold temperature. Make sure that the mold temperature is uniform.

Defect name	Description	Causes	Possible countermeasures
			6. Increase holding pressure or holding pressure time.
Blisters	A raised area on the part's surface.	1. Cooling time too short. 2. Mold temperature too low. If the part skin is formed too fast, the excessive air cannot leave the part. 3. Backpressure too low. Backpressure is used to push out excessive air before it enters the mold cavity.	1. Increase cooling time. 2. Increase mold temperature. 3. Increase backpressure.
Warpage	Distortion of the original shape of the produced part.	1. The plastic material might cool down and solidify before the mold is packed out due to low injection pressure or too short injection time. 2. Too low mold temperature resulting in a less dense part. Uneven mold temperature. 3. Material is too stiff and doesn't flow fast enough to pack the mold before it cools down and solidifies. 4. Material temperature too high. 5. Cooling time too short.	1. Increase injection pressure or injection time. 2. Increase mold temperature. Make sure that the mold temperature is uniformly distributed. 3. Use material with the highest flowability, viscosity without causing flash. 4. Decrease barrel temperature. 5. Increase cooling time.
Burn marks	Discoloration or dark stains on the molded part.	1. Injection speed or pressure too high. Because of this, molten material is forced into the mold too fast, not letting any air trapped in the runner system or mold cavity out. The air temperature then raises and ignites the surrounding material. 2. Nozzle temperature too high. 3. Barrel temperature too high. 4. Cycle time too long.	1. Decrease injection speed and injection pressure. 2. Decrease nozzle temperature. 3. Decrease barrel temperature. 4. Decrease cycle time.
Flow marks (lines)	"Off-tone" lines or patterns in the molded part mostly around the injection gates.	1. Injection pressure too low. This might result in improperly bonded material, because when it enters the mold cavity, it is not packed enough to form a smooth layer against the mold surface. 2. Mold temperature too low. If the mold temperature is not high enough, the material might not be able to flow far enough before solidifying. 3. Barrel temperature too low. 4. Injection speed too low. 5. Nozzle temperature too low.	1. Increase injection pressure. 2. Increase mold temperature. 3. Increase barrel temperature. 4. Increase injection speed. 5. Increase nozzle temperature.

Defect name	Description	Causes	Possible countermeasures
Shrinkage	Geometric reduction in the size of the part.	1. Material is too stiff. This might not let material to get fully packed in the cavity. 2. Barrel temperature too high. This might result in material absorbing an excessive amount of heat. 3. Mold temperature too high. This might cause improper formation of the part's skin, resulting in additional shrinkage.	1. Choose material with higher flowability. 2. Decrease barrel temperature. 3. Decrease mold temperature.
Weld lines	The material splits and flows in two or more directions, when the flows meet, they form hair-like lines at the meeting points.	1. Material splits and flows in different directions. 2. Material temperature is too low. The material is not fully melted. 3. Mold temperature too low. 4. Injection pressure too low. 5. Injection speed too low.	1. Increase barrel and nozzle temperature. 2. Increase mold temperature. 3. Increase injection pressure. 4. Increase injection speed.
Silver streaks (splay)	Streaks of a silvery white color that appear in the material flow direction.	1. Material is too moist. 2. Mold temperature too low. 3. Injection speed and pressure too high. 4. Material is overheated in the barrel. If the barrel temperature is too high, the material will begin to char or carbonize. Charred or carbonized particles on the surface of the product will result in silver streaks.	1. Dry material before molding. 2. Increase mold temperature. 3. Decrease injection speed and pressure. 4. Decrease barrel and nozzle temperature.
Cracking (finer cracks)	Cracks due to the skin of the molding part not being solid enough upon ejection.	1. Holding pressure or injection speed too high. This might result in molding unnecessary stresses into a product. 2. Material is too moist.	1. Decrease holding pressure or injection speed. 2. Dry material before molding.

2.1.3 Introduction to mechanical properties

High-density polyethylene is a thermoplastic material used in all the experiments underlying this PhD thesis. As mentioned in the previous section, material is one of the important factors influencing quality of an injection molded product. Even though this work does not directly include material in the list of investigated parameters, it is important to give a short overview related to the material and its quality characteristics.

Thermoplastic is a plastic polymer material that becomes moldable at a certain elevation temperature and solidifies after cooling down [50]. Polymer, in its turn, can be defined 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)*" [51]. HDPE is a major polyethylene compound manufactured at low temperatures and pressures [52]. It consists of *n* constitutional repeating units shown in Figure 2.4, where *n* is a number of the units. The absence of branches in the HDPEs structure allows to create a dense material of high strength and moderate stiffness [52].

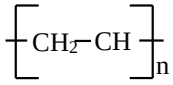


Figure 2.4. The HDPE constitutional repeating unit

Some of HDPE's important attributes are moldability, resistance to corrosion, strength to density ratio and recyclability. Due to the wide use of HDPE in different sectors, a lot of research focuses on meticulous investigation of the material's behavior. These studies include analysis of viscoelastic and viscoplastic behaviors, as well as mechanical behavior under different loading conditions [53]. A high interest for researchers and practitioners is understanding of the effects of injection molding parameters on the final part quality, including the mechanical properties, in order to reveal any underlying relationships.

Mechanical properties can be described in different ways depending on the chosen testing techniques. In this PhD work, tensile stress testing is conducted to investigate the strength and stiffness of focus parts (dogbone specimens described in more details in Chapter 4). Tensile stress testing is used to determine relation between stress and strain of a chosen material. In this thesis, mechanical properties of the virgin BorSafe^{TM} HDPE and recycled HDPE from RePro and ContainerService suppliers will be evaluated with respect to Young's modulus, tensile strength and elongation at break.

Young's modulus, also known as tensile modulus or modulus of elasticity, shows resistance of a material to elastic deformation under the applied load. According to ISO 527-1:2019 standard [54], it is defined as "*slope of the stress/strain curve* $\sigma(\varepsilon)$ *in the interval between the two strains* $\varepsilon_1 = 0{,}05\%$ *and* $\varepsilon_2 = 0{,}25\%$", and is expressed in megapascals (MPa).

Tensile strength, also known as maximum stress, is referred to as strength in ISO 527-1:2019 standard [54] and defined as "*stress at the first local maximum observed during a tensile test*". Tensile strength is also expressed in megapascals (MPa).

Elongation at break, also known as strain at break, is the maximum relative deformation before break. In ISO 527-1:2019 standard [54] it is interpreted as "*strain at the last recorded data point before the stress is reduced to less than or equal to 10% of the strength if the break occurs prior to yielding*". It is the ratio between the changed gauge length (ΔL) and the initial gauge length (L), calculated according to equation 2.1. Elongation at break is expressed as a dimensionless ratio or a percentage (%).

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

More details on the tensile testing conducted within this project are provided in Chapter 4.

2.2 A brief literature review

Injection molding is one of the most commonly used manufacturing processes for mass production of plastic parts. This is due to several reasons: the increased demand on plastic products in daily life, IM's short cycle times, and ability of the IM to manufacture parts of various complex shapes, from different plastic materials and with desirable visual and mechanical properties that meet tight tolerance specifications [55, 56]. "*Quality is a determining factor that affects the productivity and economy of production, especially in the case of mass production*" [57]. However, to achieve high quality of the injection molded parts, it is important to pay attention to the part and mold design, chosen material and process parameters [3, 56, 58].

When the part and the mold have already been designed and the material to be used has been chosen, the only way to improve the product's quality is selection of suitable process parameters' values. It is an important and a rather complicated task, failing which leads to decreased process efficiency and increased production of scrap [7]. Zhang, Mao [59] divide parameters that influence the part's quality into three categories: machine inputs, control trajectories and state variables. Machine inputs are the process parameters that are manually set, control trajectories represent the pressure, speed and temperature curves that the IMM performs, while state variables are related to the conditions of the polymer melt or solidified polymer. In [60] and [61], on the other hand, the values of the injection molding process are divided into machine values, process values and quality values.

The machine and process settings are often obtained based on the experience and knowledge of the IMM operators and a significant amount of trial-and-error is involved in this process [62, 63]. Such approach, however, is extremely inefficient, costly and does not increase competitiveness on the global market, where standards for the injection molded components are very high [28].

As it has already been mentioned, the quality of the injection molded parts can be characterized in terms of dimensional stability, mechanical properties and different features related to the product's appearance [55]. Selection of the faulty process parameters might result in occurrence of various defects, some of the most frequent once

being flash, short shot, sink mark, warpage, shrinkage and flow line [64]. For example, "*low injection pressure, short injection time, and low mold temperature will easily lead to short shot, and low packing pressure and short cooling time will cause warpage*" [65]. Therefore, it becomes crucial to provide high reproducibility of the IM process with the optimal parameter settings.

Industry 4.0, or the fourth industrial revolution, is being accelerated by such factors as the lack of skilled workers, knowledge leakage and population aging [66]. It pushes development of such concepts as cyber-physical systems or CPS that can be defined as *"systems of collaborating computational entities which are in intensive connection with the surrounding physical world and its on-going processes, providing and using, at the same time, data-accessing and data-processing services available on the Internet"* [67]. Implementation of CPS can assist the decision-making process to become more self-centered through application of data-driven and machine learning methods [68, 69]. Therefore implementation of these methods within the traditional manufacturing industries, such as injection molding, is expected to revive them and provide them with the additional competitive advantage [68, 69]. The following subsections provide examples of application of more traditional methods, such as finite element method (FEM), statistical process control (SPC), pressure-volume-temperature (P–V–T) curves, etc., as well as ML methods for monitoring and control of dimensional deviations and mechanical properties of the injection molded parts.

2.2.1 Monitoring and control of dimensional deviations in injection molding

Dimensional consistency is an extremely important quality feature of the injection molded parts, however, due to the nature of polymer materials, all of them shrink after the IM [5]. Amorphous polymers shrink due to the thermal contraction, while semi-crystalline ones due to the volume change during their crystallization [70, 71]. Unfortunately, the shrinkage values for most of the plastic materials are high in relation to the dimensional tolerances, in addition, the shrinkage is not always isotropic in nature and anisotropic shrinkage often leads to warpage. The part's shrinkage varies depending on the geometry and the process settings, and is commonly assumed to be known at the part and mold design stages [58]. However, if errors in the calculation of shrinkage were made, it can limit achievable tolerances, increase the necessary cycle time and lower the overall process efficiency. Therefore, it is important to monitor and control the corresponding machine and process parameters, to avoid excessive shrinkage, warpage and other defects resulting in the dimensional deviations of the molded parts.

Monitoring and control methods can be categorized based on different criteria. One of the ways is to divide them into destructive (those that require installation of additional sensors or other kind of equipment into the mold or the IMM, and involves drilling holes, etc.) and non-destructive (those that do not require such measures). Most of the research focused on the IM process monitoring and control uses destructive methods [7, 72]. However, modern IMMs include a significant number of installed by the IMM's producer sensors and contain a lot of information about the actual process conditions without need to apply the destructive methods [7, 73].

Non-destructive methods

Interactions between the quality characteristics and process parameters are complex and need further investigation for tuning the injection molding process. There are several different methods applied in order to achieve this. One of such methods is finite element method or FEM. It is a deterministic method that allows simulation of the optimal design of the molds and determination of a suitable parameter combination for the production start [74]. However, due to the stochastic nature of the injection molding process, the real production conditions, variations in the material or mold geometry lead to inconsistencies and deviations from the nominal process setting obtained from the FEM. A significant advantage of this method, though, is that it can be used before the real production begins.

In [63] a mathematical model for prediction of warpage is proposed. At first, five of the most influential on occurrence of this defect parameters (melt temperature, coolant temperature, injection time, switchover time and mold temperature) were screened using fractional factorial design of experiments. Next a prediction model was created using central composite design of experiments and FEM. Finally, the model was tested, and the corresponding statistical analysis was carried out using the MoldFlow™ simulation package. "*MoldFlow™ software is a commercial software based on hybrid finite element/finite difference techniques in order to solve pressure, flow, and temperature of the molding process*" [56].

Fu and Ma [75] propose a method for simulation of the early part ejection and possible deformations through integration of MoldFlow™ and Ansys™ simulation software. Similarly to MoldFlow™, Ansys™ is a mechanical FEM software used for simulation of models of structures, electronics or machine components for analysis of strength, toughness, elasticity, temperature distribution, etc. [76]. This method is better at predicting part's dimensions if compared to stand-alone molding simulations. Such simulation can be helpful at the mold design stage to facilitate the part's ejection.

At the same time, in [74] the authors propose a virtual prototyping environment for performing a robust optimization of the injection molding process and prediction of the produced component dimensions. The environment is based on the combination of numerical simulations, response surface methodology (RSM) and stochastic simulations. Combining all of these methods allows to consider stochastic fluctuations of the process and remove some of their time-consuming aspects. RSM is a collection of mathematical and statistical techniques that can be applied when a response of interest is influenced by several parameters. The goal of this methodology is to optimize the response and establish a model that quantifies the relationship between the response and the input variables [74]. This method, however, requires experience for selection of the relevant process variables and selection of the corresponding ranges. In addition, application of this experimentation methodology might be rather time consuming [74]. However, the same way as FEM, RSM can be used before the production start.

Zhao, Zhou [7] propose a non-destructive online method for monitoring of the injection molding process and parameters that influence dimensional deviations that does not involve application of the simulation methods. The technique proposed by the authors, includes collection of signals from electrical sensors installed in the injection molding machine by the machine manufacturer, while ultrasonic monitoring technology [77] is implemented to measure the cavity pressure without drilling any holes in the mold. The

authors also present an algorithm for the injection molding process stage identification using the obtained data. This methodology can assist in optimization of switch-over time and holding time parameters.

Johnston, McCready [78], on the other hand, have designed an auxiliary process controller for online multivariate optimization of the injection molding process. From the control point of view, IM is regarded as a batch process "*such that the resulting dynamics of the polymer melt throughout the mold cavity proceed in an open loop fashion*" [78]. At first, a multivariate model is created using principal component analysis (PCA). The model is able to show relationships between process settings and different part quality characteristics, such as product's dimensions. Second, PCA enables the online modeling of the process through consideration of the streams of the process data. In the end, modeling of the process states dynamics is enabled, allowing a more rapid process convergence.

Destructive methods

The methods used for non-destructive process monitoring and control are also applied within the destructive techniques, however, they are often combined with acquisition of data from additionally installed machine and/ or mold sensors. In [58] the authors propose a button cell type in-mold shrinkage sensor. The sensor's performance was validated and compared to traditional shrinkage prediction methods. Design of experiments (DOE) was used to validate the sensor's performance as a function of holding pressure, melt temperature, cooling time and coolant temperature.

In addition to use of data related to machine and process parameters, cavity pressure is considered an extremely important parameter for achieving a high part quality, since it determines the evolution of the polymer conditions inside the mold [7, 55, 56]. Due to this, Kurt, Kamber [55] have conducted a study of how cavity pressure and temperature influence shrinkage and cyclicity of the focus part. They used three Kistler 615BA piezoelectric pressure transducers in the cavity and one 6190BAG temperature/pressure combined transducer. A Kistler CoMo 2869A injection type aparatus was used to log and analyze the obtained sensor signals. Later, Kurt, Kaynak [56] have investigated relation and influence of molding conditions such as holding pressure, melt temperature, cooling time, as well as cavity pressure, mold temperature and melt temperature on shrinkage and roundness of injection molded parts. They have used three Kistler 6157BA piezoelectric pressure transducers and a 6190BAG temperature-pressure combined transducer in the cavity.

Zhang, Dubay [79] have also analyzed cavity pressure and temperature data. They combined the use of PCA, independent component analysis and model-based predictive control to propose methodology for online control of injection molding process parameters. PCA and independent component analysis were applied to extract and transform cavity pressure and temperature data. The method allows to automatically adjust and vary key process parameters, such as coolant temperature and coolant flow rate to avoid warpage of the produced parts.

Gao, Tang [57] emphasize that even though the initial pressure and temperature of the polymer melt are controlled by the IMM, the actual melt states vary in the mold cavity.

They demonstrate an online product quality monitoring system with the focus on the parts dimensions. Piezoelectric pressure sensors, an in-mold thermocouple, infrared melt pyrometer, a custom-designed multivariate sensor are used to measure melt pressure, temperature, velocity and viscosity within the injection mold. Later, a model based on support vector regression algorithm and the obtained data is developed to monitor and predict the produced parts dimensions.

Another important parameter often used in numerical simulations of the injection molding is Pressure-Volume-Temperature (P–V–T) relationships of polymers or P–V–T curves. They are important in both engineering and polymer physics [80], as they describe relationships between the pressure, density and temperature for a given polymer material [7]. It is constant for different materials, but might not always be accurate enough. Wang, Xie [80] present a novel online testing equipment for measurement of the P–V–T relationships of polymers to predict shrinkage and warpage of produced parts. The equipment is based on the IMM and can be used to obtain the P–V–T data directly using a special testing mold. The authors also propose to create a P–V–T database of polymers to aid further development of software packages for monitoring and control of IM.

At the same time, traditional process parameter optimization procedures for control of the aesthetic defects in the IM often suffer from convergence and stability problems [4]. Therefore, Gao, Zhang [4] propose an injection process parameters optimization model-based procedure, "*utilizing the fact that the feasible parameter domain is usually sandwiched between two opposite defects when a parameter increases from a low level to a high level*". The procedure applies fuzzy reasoning method and is targeting various IM aesthetic defects.

In [81] the authors applied a digital image processing technique to detect shrinkage and flash on the produced injection molded parts and then used the simplex model free optimization algorithm to optimize mold temperature, injection pressure, holding pressure, holding time and cooling time to avoid the above mentioned defects. Park and Nguyen [82], in their turn, use RSM and non-dominated sorting genetic algorithm II to solve a multi-optimization problem for minimization of clamping force and warpage of a plastic car fender. Mold temperature, melt temperature, holding time, holding pressure and cooling time are chosen as the main control parameters of the optimization problem.

Unlike virgin plastic materials, the recycled ones may have properties, such as fluidity and viscosity, that significantly differ from the primary properties of the virgin material [83]. Therefore, Zhang [83] presented a pokayoke system for prediction of flash defect in injection molding targeting the recycled materials use. The system employs accelerometer to measure the injection molding vibration signals, which are then analyzed and used as an input to the logistic regression modeling algorithm.

Monitoring and control of dimensional deviations for the IM, as well as the impact of various process parameters is well presented in the literature. Many scholars use mold temperature and pressure, melt characteristics, holding pressure and time, cooling time and coolant characteristics as parameters included into the control models. However, there are other potentially useful parameter values that can be obtained from the IMM sensors, that are rarely or never used within the developed models. Examples of such

parameters are screw speed, barrel temperature, injection speed and cushion. In addition, some of the simulation methodologies and techniques applied for the monitoring and control of geometrical deviations have significant disadvantages (employed simplifications, big computational efforts, need for highly experienced operators, etc.). Therefore, more efforts need to be put in order to develop a generalized approach to efficient monitoring and control of shrinkage, warpage and other defects influencing the potential dimensional inconsistencies.

2.2.2 Monitoring and control of mechanical properties in injection molding

Monitoring and control of mechanical properties of the manufactured parts is by no means less important than that of their dimensional deviations or aesthetics. "*The structural safety and stability of the products depend on the mechanical performance of the products*" [84]. Shrinkage, warpage, residual stresses and other defects influence the mechanical performance of the parts, and the need for control and optimization of the IM process is extremely high [84]. The residual stresses caused by warpage cause a remarkable influence on the mechanical performance of the produced parts. Therefore, it is essential to consider influence of the molding conditions on the description of the mechanical behavior of the final product [84]. Some of the methods applied for monitoring and control of the mechanical properties are the same as those for optimization of the dimensional consistency (FEM, use of simulation software, RSM, etc.).

Any change in the IM parameters or material might affect the process stability and quality of the serviced products. "*The properties of the plastic materials are directly dependent on the temperature, time and environment*" [8]. Consistent material supply is crucial for production of high quality items and its absence might result in significant problems in the obtained products [8]. Therefore, in [8] the authors demonstrate how near infrared spectroscopy can be used for in-line process monitoring and quality control improvement during the injection molding process. The nozzle of an injection molding machine that was used within the experiments was modified to accommodate two optical fiber probes for monitoring of the injected polymer. The data related to material and color identification, as well as humidity identification was obtained and analyzed both online and off-line for several different materials (polypropylene (PP), polyethylene-terephthalate (PET) and polyamide (PA)).

Kim, Gang [29], in their turn, investigated the behavior of the polymer melt in the mold microchannel cavities during the filling process. It has been described with an analytical model and the cavity pressure-time profiles data. The model shows the relationships between the injection flow rate, peak cavity pressure, mold temperature, melt temperature and the filling length.

In [85] the authors have used DOE and range analysis method to select the best values of injection molding parameters to obtain expected tensile strength of the product. DOE in this case is applied to obtain the necessary parameters data and facilitate the optimization process. It is highlighted that injection pressure, melt temperature, cooling time, holding time, holding pressure, mold temperature and injection speed are some of the most influential parameters affecting the parts tensile strength.

Xu, Zhang [84] concentrated on optimization of process parameters to obtain the desired mechanical properties (von Mises stress) of the vehicle window made of polycarbonate. At first, the FEM is applied. Next a back-propagation neural network (BPNN) is trained to map the nonlinear relationship between the process parameters (mold temperature, melt temperature, injection speed, compression distance, compression force, compression speed, compression waiting time) and mechanical properties. The last step is application of particle swarm optimization (PSO) to the neural network model for obtaining the optimal values of the above-mentioned parameters.

Ozcelik [86] applied Taguchi method to create an experimental plan to identify process parameters (melt temperature, holding pressure, injection pressure) that have a significant influence on mechanical properties (maximum tensile load, extension at break, charpy impact strength (notched)) of the polypropylene specimens with weld line. ANOVA was used to identify the most important parameters, linear models to describe the corresponding dependencies were then created applying the regression analysis. "*The most important parameter affecting the maximum tensile load and the extension at break (for specimen without/with weld line) was injection pressure and melt temperature, and for charpy impact strength (notched) (without/with weld line) was melt temperature and injection pressure, respectively*" [86].

Taguchi method is a design of experiment methodology that aims to use a smaller number of experiments to allow a greater understanding of in-process variation. This is done through conducting experiments from an orthogonal array that should simulate the random environment in which the product is manufactured. The method is capable of converting the quality characteristics into signal-to-noise ratio (SN ratio), providing a corresponding response table showing the optimal production conditions [87]. Using ANOVA, it is also possible to assess importance of each of the factors involved. The main disadvantage of this method though is that Taguchi methodology can find a local optimum, rather than a global one, since its search space is limited to the parameter levels chosen to create the experimental design.

The authors in [88] have conducted a study for minimization and analysis of shrinkage, warpage and von Mises stress for the bone screw parts. A FEM was used to determine the force values and concentration points causing yielding of the screws. An RSM based on the central composite design was carried out to determine effects of the process parameters (coolant temperature, mold temperature, melt temperature, holding time, injection time, and holding pressure) on the focus parts.

In addition to the need for the IM process control, extended use of plastic products in the everyday life and the corresponding environmental concerns create new challenges in the plastics industry, such as reuse, recycle and application of the lightweight materials [89]. As a result, use of recycled plastic materials becomes essential. Fernandez, Muniesa [89] describe a methodology for the rheological testing of polymers during the injection molding process and further optimization of the production process. The method is especially beneficial for defining non-conventional features of the plasticization phase of recycled thermoplastic materials. To achieve this a nozzle of the injection molding machine is equipped with electric band heaters with thermocouples to obtain the molten material temperature data and Kistler 4083A sensors to measure pressure and temperature

of the plastic melt. In addition, the machine plasticization unit is equipped with a wire potentiometer for measurement of the linear displacement of the reciprocating screw and hydraulic pressure sensor. After obtaining the sensors data, the calculation of the material viscosity is performed, and rheological model constants are obtained.

Mehat and Kamaruddin [90], at the same time, investigated the effects of injection molding process parameters on the mechanical properties (stress at yield, flexural modulus) of recycled materials. They have used MoldFlow™ simulation software, Taguchi methodology and variation analysis (ANOVA) to conduct their experiments. The results of the experiments have shown that injection time significantly influences the material flexural modulus, while melt temperature is the most important parameter for obtaining the expected stress at yield. The other investigated parameters include mold temperature, melt temperature, injection time, holding pressure, holding time and cooling time.

The number of studies directly focusing on the monitoring and control of the mechanical properties of the IM products is smaller than of those related to the geometric deviations control. Simultaneously, research concerning the IM parameters that effect different aspects of the part quality is still an ongoing field of research [79]. Nevertheless, there is a number of studies well describing these challenges and their possible solutions. Still, more effort needs to be employed to map relations between various machine and process parameters and their influence on the mechanical properties of the final IM parts.

2.2.3 Machine learning approaches for injection molding

The previous subsections have shown examples of various methodologies applied for monitoring, control, optimization and investigation of different parameters influencing the IM process. However, many of the above-mentioned methods have significant disadvantages that make their application non-feasible. For example, FEM often involves huge computational efforts [91], numerical modeling methods apply simplifications that might not be able to adequately reflect on the non-linear material behavior [5], Taguchi methodology can find only the best set of the specified process parameter level combinations, and not the global optimum, etc. [92]. ML methods, on the other hand, require smaller computational time, are capable of modeling non-linear relationships and give better results when it comes to process modelling and forecasting [31].

In the recent years, collection and analysis of big amounts of production processes data become more common and enable application of the ML methods, as they are able to provide adequate results only if enough statistical data is used to train the models for process control and optimization [79, 93]. At the same time, it is extremely important to choose a proper ML method depending on the application purpose [72]. ANNs are often applied for quality prediction due to their ability of nonlinear mapping between the noisy sets of input and output data [92]. Their utilization is often coupled with self-organizing maps (SOM), genetic algorithms or fuzzy logic [92]. For example, in [92] the authors have trained two different models, the first one used combination of SOM and a back-propagation neural network (BPNN) and the second one only on the BPNN. The models were developed to create a dynamic quality predictor for the injection molding process. The predicted quality characteristics is part weight. Nine parameters (injection stroke

curve, injection velocity curve, pressure curve, injection time, packing pressure, injection velocity, packing time, injection stroke, VP switch position) were included into the first model and six into the second one. Taguchi's parameter design was utilized to optimize the parameters of the neural network. The datasets included 120 and 40 samples of the experimental data correspondingly.

Manjunath and Krishna [94] developed an ANN model for forward and reverse mapping prediction. In the forward mapping process parameters (holding pressure, injection speed, mold temperature, melt temperature) were used as input variables for prediction of dimensional shrinkage of the injection molded parts. In the reverse mapping, it was attempted to predict an appropriate set of process parameters needed to reach certain quality of the part. The model was trained using 1000 data samples randomly generated through previously reported injection molding regression equations.

Kuo, Su [87] have used combination of Taguchi quality method, analysis of variance (ANOVA) and a BPNN to reduce the dimensional deviations of injection molding. The Taguchi quality method was applied to establish the design of experiment, which consisted of 18 combinations of nine process parameters. Next ANOVA was used to assess influence of each individual factor and choose the most influential ones. Finally, five parameters (pre-plasticity amount, injection pressure, injection speed, screw rotation speed, cooling time) were chosen to train a BPNN for prediction of the injection molded parts dimensions. The model was created using 180 data samples. Similarly to [87], in [95] a DOE is established using the Taguchi method, and ANOVA is carried out to define the most important process parameters. 81 data samples were collected and used to train multioutput support vector machine for regression and a BPNN. Afterwards, the multi-objective optimization was performed through the application of the nondominated sorting genetic algorithm. The main quality objectives were haze ratio and peak-to-valley 20 of the plastic optical lens focus part.

In [96] the authors compared performance of ANN and response surface methodology for establishment of the process window for a plastic lens production. The models were created for mold temperature, cooling time and holding time process parameters. The models were able to predict the accuracy of the lens's form. The optimal lens form obtained after using the parameter values proposed by the ANN was better than that received as a result of the response surface methodology model application. However, the form accuracies obtained from both models were quite consistent.

Lotti, Ueki [5] used DOE to define the experiments and ANNs to create models for prediction of shrinkage of the focus parts. The models included four processing parameters, namely: melt and mold temperatures, holding pressure and flow rate. The models were trained using 30 data samples. The ANN has shown better results in comparison to those achieved with the help of the MoldFlow™ software package and the corresponding finite element analysis. Yin, Mao [97] also compared performance of an ANN with the FEM conducted in the MoldFlow™. They applied BPNN for prediction of warpage of an automobile glove compartment cap and optimization of mold temperature, melt temperature, holding pressure, holding time and cooling time. The results obtained from the BPNN model were compared with those from FEM conducted in the MoldFlow™ simulation software. The BPNN is able to successfully predict the possible

warpage occurrence, as well as optimize the desired parameter values, simulation time used by the BPNN is significantly shorter than that required for the FEM.

Nagorny, Pillet [98], on the other hand, collected process data and thermal images of 204 rectangular specimens and extracted 97 scalar statistical features from the obtained data. The features were then used to train geometric dimensions prediction models using support vector regression, random forest, k-nearest neighbors, stochastic gradient descent, bagging decision tree, Ada Boosting, as well as convolutional neural network and long-short term memory network. Neural networks models show the best results among the tested algorithms and corresponding models.

Zhu and Chen [99] describe development of a fuzzy neural network-based in-process prediction system for prediction of flash defect occurrence during injection molding process. The main goal is to create a fuzzy neural network capable of predicting flash when using recycled mixed plastics. Such important process parameters as injection speed, melt temperature and holding pressure are varied within the experiment. The model was trained using 180 data samples collected during a 3x2x2 factorial experimental design.

The authors of [100] propose a reinforcement learning control system for the injection molding process. At first, the relevant process data that includes quality characteristics (warpage measured in three different points) and corresponding parameters settings (holding pressure, holding time, melt temperature, mold temperature) is obtained through the established DOE and a MoldFlow™ simulation software analysis. Secondly, a self-prediction quality model (ANN regression model) is trained using the obtained data. Finally, the reinforcement learning decision model (Markov decision process model) for adjustment of learning process parameters is trained.

Kozjek, Kralj [101] proposed a data mining approach to the identification of complex faults, such as unplanned machine stops during the injection molding production process. Such methods as J48 decision trees, random forest, JRip rules, naïve Bayes and k-nearest neighbors are applied to the industrial data (about 2.2 million cycles on five observed IMMs).

In [91] an ANN is trained using process parameters data for mold temperature, melt temperature, injection time, holding time, holding pressure and cooling time, a target variable is expressed with a quality index that includes warpage deformation, thickness uniformity, etc. The quality index is evaluated using MoldFlow™ Plastic Insight and optimization procedure is performed through the parametric sampling evaluation function. The models are built for three different products (scanner, TV frame, plastic lens), a small-sized DOE consisting 20 data samples is obtained using Latin hypercube sampling method (LHS) for each of the products. This approach is focused on avoiding necessity for the use of FE and computational efforts related to its application.

In [102] two different BPNNs were trained for prediction of part weight and part length. The first BPNN used discrete data from a Box-Behnken DOE for such parameters as injection pressure, holding pressure, injection time and cushion. In total 150 data samples were collected. The second experiment used continuous variable data (complete cycle cavity pressure sensor data). The models' performance was compared to that of SPC,

BPNN has shown better result, falsely accepted and rejected parts rate was decreased with 50% with use of the BPNN approach.

The main idea of SPC is building a stochastic model that allows to control the process based on the measurements of the manufactured parts and avoid over-controlling it. A good fit to the model and a clear connection between the out-of-control signal and actions that help to bring the process in control are essential. At first, historical data of a stable process behavior is collected and the control limits for the expected measurements are calculated. Secondly, the process data is collected, the measurements should fall between the control limits for the process to be under control. The measurements that fall out of the control limits are examined to see if they belong to the same population as the initial data or if not, in that case the control limits need to be recalculated [103]. The corresponding examination and calculation of the new control limits needs to be conducted by the skilled personnel.

Chen, Nguyen [104] apply combination of Taguchi method, response surface methodology and hybrid genetic algorithm and particle swarm optimization (GA-PSO) for optimization of warpage and length quality characteristics. The control parameters are melt temperature, injection speed, holding pressure, holding time and cooling time.

However, according to [105] the above-mentioned application of the ML methods includes some deficiencies, as it is focused on optimization of a single criterion, such as shrinkage or warpage, while a multi-objective optimization needs to be conducted to optimize several criteria at once. Therefore, in [105] a framework for tackling the Pareto optimum of the injection molding process parameters for multi-objective optimization of the final part quality is presented. The studied processing parameters are injection time, melt temperature, holding pressure, holding time, cooling temperature and cooling time. The optimization systems proposed in the study includes two levels. The first level utilizes an improved efficient global optimization algorithm to approximate the nonlinear relations between the process parameters and the quality measures (volumetric shrinkage). In the second stage non-dominated sorting-based genetic algorithm II is used to find a better spread of design solutions.

The authors in [65, 106, 107] also focus on the multi-objective optimization problems. In [106] an intelligent methodology that combines variable complexity methods, non-dominated sorted genetic algorithm (NSGA), BPNN and MoldFlowTM analysis are applied to solve the multi-objective optimization problem of the injection molding parameters, where the optimization objective function is focused on minimization of volumetric shrinkage, cycle time and total volume in the runner system. The sample data is collected using the MoldFlowTM analysis software. In [65] and [107] the authors adopt a multi-objective optimization approach for minimization of warpage and cycle time, and volume shrinkage and clamping force correspondingly. A sequential approximate optimization using radial basis function network (a feed-forward network) is applied to define the pareto-frontier between the cycle time and warpage objective functions in [65] and volume shrinkage and clamping force in [107]. In [65] holding pressure profile, melt temperature, injection time, coolant temperature and cooling time parameter values are optimized, while in [107] injection pressure, holding time, holding pressure, melt

temperature, injection time and mold temperature. The initial sampling points are generated using the LHS in both cases.

Most of the above-mentioned research studies use either simulated or real process data , however, in [108] it is suggested that the transfer learning can assist in overcoming the gap between the real and simulation data. The authors have trained two ANN models for prediction of the parts quality and the results have shown that the model which was created using both simulation and real data had better performance in comparison to the one using only experimental data.

In addition to application of the ML methods to concrete case studies, some of the researchers propose frameworks for development of artificial intelligence or smart manufacturing control systems for the injection molding process [68, 93, 109]. According to the authors, such systems will decrease the need for human interference in the future production systems. However, at the moment "*real time process parameter optimization and control without any human interference is still a distant approach due to complexity of injection molding process and related input and output parameters*" [93].

In [93] the authors propose an artificial intelligence (AI) based control system for consistent product quality in injection molding process. The system includes two pressure and two temperature cavity sensors, which collect process data during the production cycles. The sensor data is then analyzed using ML techniques and under sampling method. If the product failure occurs, the machine operator is notified, and parameter adaptation is performed. The control system is tested with a car door injection molded part. Lee, Ryu [68], on the other hand, propose a systematical framework for a smart injection molding system capable of self-learning know-how and knowledge on characteristics related to product quality and search of the optimal parameter settings. The framework includes detailed description of the systems functions, data flow, procedures for data analysis, decision-making and control of the process based on the real-time data.

In [109], at the same time, an intelligent system based on fuzzy logics is proposed. The system employs adaptive membership functions, which are defined based on the different quality defect behavior curves. The system can be tuned by an operator by means of a set of adaptive regression membership functions. The systems assists in correlation of quality inspection of the manufactured parts with the set of appropriate process parameters. Liau, Lee [110], in their turn describe a framework for implementation of a digital twin for the injection molding process that would allow to combine digital and physical environments for the production process improvement.

Hopmann, Ressmann [28] suggest that application of ANNs provides good control results, however, has several disadvantages: it requires a complex controller setup, large amount of training data and low amount of process knowledge incorporated into the obtained ANN models. Therefore, in their work Hopmann, Ressmann [28] developed a model-based controller that uses online optimization to determine the control inputs. Unlike other types of controllers that apply a well-tuned, but relatively simple algorithms, this controller is based on a physical-based process model, which is able to reproduce the general behavior of the injection molding system. Later, Hopmann, Abel [60] also proposed a control strategy based on the iterative learning control that allows controlling

in-mold cavity pressure with respect to the reference generated by the P–V–T optimization tailored for a specific material.

Even though most of the above-described research studies report good results after application of the various numerical and stochastic modeling techniques, as well as machine learning, fuzzy logics or reinforcement learning, more work needs to be conducted to develop an overall framework for their implementation. Some of the applied methods, such as FEM, RSM, SPC and Taguchi method have disadvantages that make their application non-feasible, due to either significant simplification involved, high computational costs, inability of finding a global optimal solution or a need for constant involvement of skilled personnel. In this case, application of the ML methods for the IM process monitoring and control is a promising area, as these methods are capable of "learning" from the statistical data and mapping complex non-linear relationships. Nevertheless, some of the studies that employ ML methods use very small data sets to train the models, as, for example, in [5] an [91], meaning that more experimental data needs to be obtained. At the same time, it is important to control quality of this data to ensure high quality of the obtained models. Moreover, most of the scholars, except [98, 101], apply the same ML methods, mostly ANN and GA, without attempts to use other machine learning methods that might show better performance depending on the problem that needs to be solved. Some of the scholars use ANOVA [86, 87, 90, 95] to define the most influential parameters based on the collected data, but none of them use feature selection methods (a subset of ML methods) for this purpose. Therefore, development of a general framework and guidelines for implementation of the ML methods for monitoring and control of the IM process is essential to achieve creation of an intelligent system that could be used not for a limited number of case studies, but production of parts with various geometries and materials.

Chapter 3

Machine Learning methods as means for monitoring and control of thermoplastics IM

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.

– Marie Curie

This chapter provides theoretical foundations of ML methods used at different stages of this work to conduct the necessary data analysis. Section 3.1 provides a general overview of the ML approach, while Sections 3.2 – 3.5 include description of different categories of the ML methods used within this work. Section 3.6 sums up advantages and disadvantages of the ML methods application.

3.1 A general overview of the ML approach

Machine Learning (ML) is a sub-field of Computer Science that describes ability of a computer program or a machine to learn from previously collected data (experience) **E** to solve the class of problems (tasks) **T** with performance characteristics **P** with condition of **P** improving over time [111]. According to Fayyad, Piatetsky-Shapiro [112] there are several fundamental tasks that the ML can solve:

- Classification – division of data into the known groups based on their previously defined properties.
- Clustering – division of data into groups based on their common characteristics with no prior knowledge about these groups.
- Regression/ Forecasting – prediction of future or an event based on the available historical data.
- Rules learning/ Associated rules – extraction of common characteristics of the available data and representation of these characteristics in a suitable form.
- Optimization – selection of an optimal solution based on predefined criteria and constraints.

These tasks can also be divided into the groups of supervised and unsupervised learning. Supervised learning includes classification, regression, rules learning and optimization. The main feature of the supervised learning algorithms is that training input and output data samples are provided [113]. At the same time, unsupervised learning is a self-learning technique, where the system has to discover the features of the input population without prior knowledge of groups this data belongs to. One of the tasks of the unsupervised learning is clustering. In this dissertation the regression and optimization tasks are addressed. In general, to solve a task from the above-described classes, ML algorithms follow a routine similar to that shown on Figure 3.1, also described in [114].

As depicted in Figure 3.1 the ML scheme starts with the problem understanding and formulation and further proceeds to the relevant data acquisition. The next step includes

data preprocessing – transformation of raw data (sensor signals, log data) into a format suitable for further processing and analysis. This step might include normalization, interpolation, padding, etc. Further on, the process is divided into two stages:

Training – a process of fitting the model to the available training data (training dataset) in accordance to the specified performance measures that are capable of showing the model's ability to fit the data. It includes two main stages:

- Feature construction and selection – a set of methods for construction and selection of relevant and non-redundant features/ parameters to be used for further model construction. Depending on the problem, sometimes only feature construction or selection is needed. In this dissertation the feature construction step is omitted, as it is not relevant for the task.
- Model selection and training – selection of a suitable ML algorithm and corresponding hyper-parameters depending on the task that needs to be solved and training of a model based on the selected ML method.

Testing – is a second stage of the ML approach. It is a process of model evaluation using new data samples (testing dataset) to predict unknown parameters or classify which group does the sample belong to.

- *Feature evaluation* – evaluation of the new data sample based on the set of the constructed and selected features.
- *Performance of the chosen task and its evaluation* – the model trained based on the training data performs classification, regression, etc., and its performance is evaluated using the relevant performance measures.

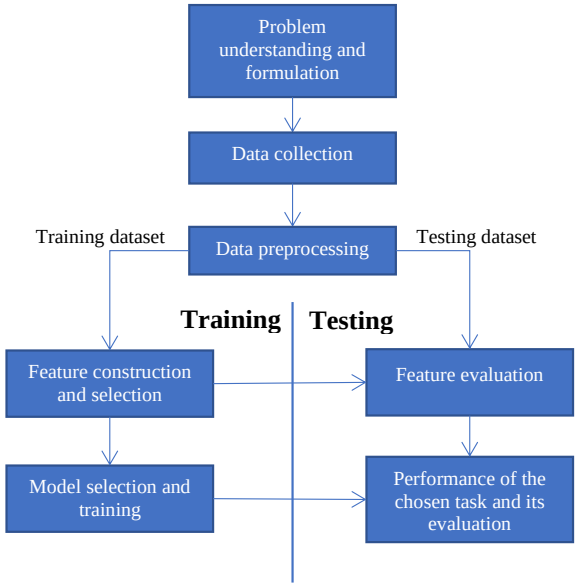


Figure 3.1. A ML algorithm routine

The described above scheme often needs to be supported by the relevant expert knowledge related to selection of a suitable modeling method and tuning of the corresponding hyper-parameters, as well as collection of relevant data that contains potentially interesting information and can result in generation of a useful model. Therefore, it is important to understand that: (i) the problem to be solved needs to be well understood and formulated, (ii) enough of high quality data needs to be acquired prior to application of the ML methods, and (iii) the corresponding ML algorithms capable of solving the formulated task need to be carefully selected and their hyperparameters tuned. Taking into account these three points is a key factor to successful solution of a relevant data analysis task. When one or more of these points are ignored, overfitting or underfitting might occur. Overfitting happens when a model learns details and noise in the training data to such extent that it negatively influences its ability to generalize and perform on the new data. Underfitting, on the other hand, is related to a model that is neither able to model the training data, nor generalize on the new data.

3.2 Data preprocessing

Data preprocessing includes a number of procedures performed on the raw data prior to its further utilization. According to [113] there are three main types of preprocessing operations: outlier removal, data normalization and dealing with missing data.

Outlier removal

An *outlier* can be defined as a data point whose value lies very far from the mean of the random variable under consideration [113]. For a normally distributed normal variable, 95 % of the points lie within the distance of two standard deviations from the mean, while 99 % are within the three standard deviations distance. Points that are significantly different from the mean result in large errors during the model training. This might cause serious effects on the quality and performance of the final model. If the number of outliers is small, they might appear due to erroneous measurements, in such case they are usually discarded. At the same time, if there are many outliers, they might be result of a distribution with long tails. In this case, cost functions that are not very sensitive to the presence of outliers need to be adopted during the model construction. More information on such cost functions can be found in [115].

Missing data

Unfortunately, it often occurs that some of the feature vectors have missing values. This might happen due to several reasons some of which are partial response in surveys or application of the remote sensing technologies, when certain regions are covered by a different number of sensors. Some of the most popular techniques for dealing with missing data are *interpolation* and completing the missing values with (a) zeros (called *padding*); (b) the unconditional mean computed using the available values of the corresponding feature; (c) the conditional mean, in case one has an estimate of the probability distribution of the missing values [113]. In addition, when having the large amounts of data, it is sometimes decided to discard samples that have missing values, however, most of the time such measure is considered as an unnecessary luxury.

At the same time, more sophisticated methods for dealing with the missing values are available. An example of such method is *imputing from a conditional distribution*. The main idea of this method is to substitute the missing values with values with respect to their statistical nature. In this case, a missing value is replaced by a random draw from a corresponding distribution. Another more advanced variation of this method is *multiple imputation* described in [116]. In this case, for each missing value, $m > 1$ samples are generated that are later combined to fulfil certain statistical properties. More information on the methods for dealing with missing data can be found in [117].

Data normalization

In many cases features or parameters that are to be further used to create a model lie within different ranges. As a result, features with larger values might have a bigger influence in the cost function in comparison to those with smaller values. However, this does not necessarily reflect their significance in the model's design [113]. This issue is solved through the application of data normalization. There are different normalization techniques that are to be applied depending on the dataset and problem at hand, here are examples of some of them:

- *Scaling at range* is a technique that limits the feature values to the range of [0, 1] applying the equation 3.1 for N available data points. It is best used when the approximate minimum and maximum values of the feature are known, and data distribution is close to uniform.

$$\overline{x}_i = \frac{x_i - x_{min}}{x_{max} - x_{min}}, i = 1, 2, \dots, N \tag{3.1}$$

- *Feature clipping* is a useful technique in case of presence of extreme outliers. In this case, all the feature values above or below a certain value are set to a selected fixed value.
- *Z-score* is used to receive the normalized feature values with zero mean and unit variance using equation 3.2 for N available data points, where μ is a mean and σ is variance of the original feature. This technique is quite useful in case of small number of outliers that are not as extreme as when the clipping is needed.

$$\overline{x}_i = \frac{x_i - \mu}{\sigma}, i = 1, 2, \dots, N \tag{3.2}$$

- *Logarithmic scaling* is example of a nonlinear normalization method, which helps to deal with feature vectors, where the data is not evenly distributed around its mean with help of the equation 3.3.

$$\overline{x}_i = \log(x_i), i = 1, 2, \dots, N \tag{3.3}$$

3.3 Feature selection

In data science and ML, *a feature* is one of the parameters used to predict the output of interest. Theoretically, the more features are used to construct a model, the better performance it will show. However, practical experience with the ML methods shows that this is not always the case [118]. One of the reasons for this is that presence of big amounts of noisy and redundant data decreases performance of the created model.

Therefore, feature selection needs to be performed to choose the features containing the most information about the predicted response.

Feature selection (FS) can be, therefore, defined as a process of identifying and eliminating as much irrelevant and unnecessary data as possible [118]. This reduces dimensionality of the data at hand and in many cases allows the learning algorithm to perform faster and in a more efficient way. Feature selection takes places for various reasons, but the most popular ones are: (1) reducing the number of features, (2) gaining a better understanding of the features and their relationship to the response variable, (3) enhancing the generalization of the model by reducing overfitting, (4) models simplification for better interpretation by users [119, 120]. In this work feature selection methods are mainly used before application of the ANN MLP and k-Nearest Neighbors methods, since the other ML modeling methods include the feature selection step in their algorithms.

A feature selection method needs to be selected depending on the main aim for its application, as some of them might be good for gaining better understanding of the features-response relationships, but not as good for the dimensionality reduction and vice versa. In general, the feature selection methods are divided into the three main groups: *filter*, *wrapper* and *embedded*.

- *Filter methods* select the important features regardless of the model using a useful descriptive measure to rank them. The main benefits of these methods are their low computation time and ability to avoid overfitting [120]. However, not all these methods take into consideration interactions or correlation between the features. Filter methods used in this work are Pearson correlation test, Spearman's Rho correlation test, RReliefF, and Correlation-based feature selection (CFS).

Pearson correlation and Spearman's Rho correlation are univariate feature selection methods. Such methods examine each feature individually to measure the strength of the relationship between this feature and the response variable. Such methods are usually simple to use and understand and good for getting a better understanding of the dataset.

Pearson correlation determines the level of linear correlation between two variables (a feature and a response variable). This method requires the variables' distributions to be close to normal distribution [121]. The resulting value lies in the range of [-1;1]. Pearson correlation value equal to -1 means perfect negative correlation (as one variable decreases, the other one increases), +1 means perfect positive correlation (as one variable increases, the other increases as well), while 0 indicates no correlation.

Spearman's Rho correlation, on the other hand, is a nonparametric measure of the relationship between two variables, which do not necessarily need to be normally distributed [122]. Spearman's correlation is often referred to as a non-linear correlation test, whose value also lies in the range of [-1;1] and has the same meaning as Pearson correlation.

RReliefF or regressional ReliefF is a feature selection method from the Relief algorithms family. The main idea of the initial RELIEF method [123] is to assess features based on their ability to distinguish between the dataset instances whose values are close

to each other. RELIEF method can be used only for the two-class problems. It searches for two nearest neighbors of a randomly selected instance, one of them from the same class (nearest hit H) and another one from the opposite class (nearest miss M). Later it updates the quality estimates for all the features depending on the values of the randomly selected instance, M and H. The RELIEF algorithm is capable of dealing with both discrete and continuous features. ReliefF method, on the other hand is a modification of the RELIEF for multi-class problems that is also more robust and can work with noisy and incomplete data [120, 124]. RReliefF, at the same time, is the original algorithm's modification that can be applied for regression tasks, when the predicted value is continuous. For this purpose, the algorithm assesses the features ability to conclude that the predicted values of two instances are different [124]. The Relief family algorithms have low bias and can capture the local dependencies that other filter methods miss.

Correlation-based feature selection (CFS) is another filter feature selection algorithm that ranks features depending on a value of a correlation-based heuristic evaluation function [118]. The correlation function is used to determine the feature subsets that have high correlation with the predicted variable and no or low correlation with each other, it is calculated using equation 3.4. The acceptance of a feature depends on its ability to predict the correct response variable value in the areas of the instance space not yet covered by the other features.

$$M_s = \frac{k\overline{r_{cf}}}{\sqrt{k+k(k-1)\overline{r_{ff}}}}, \tag{3.4}$$

where M_s is the heuristic evaluation function of a feature subset S containing k features, $\overline{r_{cf}}$ is the main feature-response correlation ($f \in S$), $\overline{r_{ff}}$ is the average feature-feature correlation.

- *Wrapper methods* evaluate the feature subsets, which allows to consider possible interactions between the variables. Main disadvantages of such methods are an increased overfitting risk and significant computation capacity and time needed. Examples of the wrapper methods are forward selection, backward selection, stepwise selection and Recursive feature elimination (RFE). In this dissertation the RFE algorithm is tested and compared to application of the filter methods.

Recursive feature elimination (RFE) method works by recursively removing the features and building a model including the remaining variables [125]. At first, the estimator is trained on the full set of features and their initial importance characteristic is obtained. Later, the least important features are eliminated, and the procedure recursively repeats until the desired number of features is selected.

- *Embedded methods* try to combine the advantages of both filter and wrapper methods. They perform feature selection as part of the model creation during the algorithm's execution. This allows to avoid some disadvantages of the previous methods, as the selection is done with connection to the model tuning. Examples of such methods are Lasso and Ridge regression, as well as decision tree algorithms.

3.4 Machine learning techniques

As it has been mentioned before, the ML approach includes training and testing stages. During these stages training and testing sets are used correspondingly. They are obtained through splitting the initial dataset used for the model development. Then, the training set is used to train the model and the testing set to test its performance and generalization abilities on the previously unseen data. The size of the training and testing datasets depends on the size of the original dataset. In this work, 70% training and 30% testing ratio is used, while 5-folds cross-validation is performed on the train set.

3.4.1 Artificial Neural Networks

Artificial Neural Networks (ANNs) is a class of machine learning models used for mapping complicated non-linear relationships between the inputs and outputs that can be used to solve various tasks, including classification and regression. The main idea behind the ANNs is construction of a model, which is a network of simple processing units (neurons) connected between each other (presence or absence of a connection between the particular neurons depends on the chosen architecture of the ANN), capable of calculating a weighted sum of their inputs and using an activation function to calculate its output [114]. The ANNs might use different activation functions, such as a binary step function, sigmoid function, rectified linear units function (ReLu), etc. The network is then used to adjust its weights in order to maximize the model's performance and minimize the appropriate cost function of its output. One of the ways to perform this adjustment is application of backpropagation, which is a method for computation of gradients of the cost function and propagating the corresponding error back through the network. This allows to adjust the model's weights depending on the values of the backpropagated error. While ANN is a powerful tool for modeling of various phenomena, its main disadvantage is low interpretability of the obtained model for a human expert. In this dissertation a feed-forward multilayer perceptron with backpropagation is used.

Multilayer Perceptron (MLP) with backpropagation

Multilayer Perceptron (MLP) is one of the classical ANN architectures. Perceptron is a feed-forward neural network with only input and output layers. MLP, on the other hand, includes one or more additional layers that are called hidden layers. These layers are fully connected, and layer-to-layer mapping is activated using a non-linear function. Multilayer perceptron can be used to model any nonlinear function, when other conventional mathematical modeling is difficult or inappropriate [114]. To do this a suitable number of layers and neurons in them need to be selected.

It is possible to explain how an MLP works using example of a perceptron. The perceptron takes a set of features as an input vector $x \in \mathbb{R}^n$, where n is the number of features, and provides an output $y \in \mathbb{R}$, as a result. This way it maps the input values to the output, as a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$. The corresponding f function is calculated based on the sum of weighted inputs and bias factors.

One of the most common MLP architectures is a three-layered network, it includes an input layer, a hidden layer and an output layer. Schematic representation of a three-layered MLP is shown on Figure 3.2, for simplification purposes the bias terms are not

depicted. Each hidden unit approximates the input to the output layer using an activation function $g()$ using the equation 3.5:

$$H_k = g(\sum_{i=1}^{n} x_i w_{ki} + b_k), \tag{3.5}$$

where H_k is an output of the kth hidden unit, n is a number of outputs, w_{ki} – weight for the ith neuron, and b_k is bias.

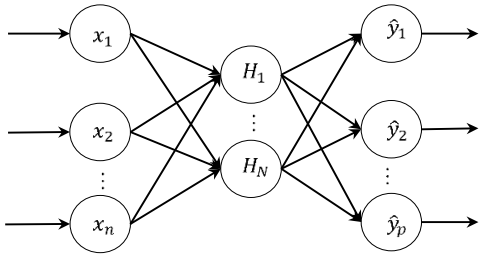


Figure 3.2. Schematic representation of a three-layer multi-layered perceptron

The output of the MLP is calculated using the equation 3.6, where $\hat{y}_m$ is an approximated value of the mth output unit, N is a number of neurons in a hidden layer, while b_m is bias:

$$\hat{y}_p = f(\sum_{j=1}^{N} H_j w_{pj} + b_p) \tag{3.6}$$

To minimize the difference (e) between the observed and predicted values of the output equation 3.7 is used:

$$e = arg\ min(\frac{1}{2}\sum_{m=1}^{M}(y_m - \hat{y}_m)^2), \tag{3.7}$$

where y_m is the observed value of the response variable, $\hat{y}_m$ is the predicted value, and M is the number of samples.

3.4.2 k-Nearest Neighbors

k-nearest neighbors (kNN) is a simple ML algorithm that uses samples similarity to predict values of previously unseen data points [114], it can be used both for classification and regression. When a new example is presented to the algorithm, it searches for k most similar to the new one samples. To assess which samples are closest to the new one, different distance measures are used. The most popular for the continuous target variables is Euclidean distance. To find the value of the variable of interest for the new sample, the mean value of target variables from the set of the nearest neighbors is taken. kNN is sensitive to the chosen distance measure and k – the chosen number of the nearest neighbors. On the one hand, if the noise is present, increasing the number k might average the results and reduce probability of an erroneous prediction. On the other hand, it might increase probability of considering far nearest neighbors, which will decrease the algorithm's prediction quality. The main disadvantages of this method is its high computational complexity for large datasets.

3.4.3 Decision Trees for Regression

Decision trees are another class of the ML modeling techniques that can be used for both classification and regression tasks. Unlike ANNs, decision tree models are easy to interpret and understand. The decision tree "asks" a series of questions to the data, narrowing the possible values until the model is confident enough to make a single prediction. The obtained model has a flowchart-like structure and includes three main components: nodes, branches and leaves [114]. *Nodes* in this case correspond to the features, *branches* to the feature values (conjunctively combined), while *leaves* or *termination nodes* represent the decision value (predicted value of a target variable in case of regression and predicted class for classification). Decision trees can be useful for gaining more understanding of the investigated production process and influence of the process parameters on the response of interest [114].

For a given input vector of features $x \in \mathbb{R}^n$ and an output training vector $y \in \mathbb{R}$, the regression tree algorithm recursively splits the data into subsets that contain instances with similar values. It is important to choose the relevant features using which the data is split. In case of a regression problem, the following metrics can be used: standard deviation, mean squared error (MSE), mean absolute error (MAE), etc. In this dissertation an implementation of scikit-learn decision tree regressor [126] is used, therefore MSE or MAE can be regression criteria for impurity function $H()$ minimization and determination of data splits [127]. Equation 3.8 can be applied for the utilization of MSE and equation 3.9 for MAE:

$$H(X_m) = \frac{1}{N_m}\sum_{i=1}^{N_m}(y_i - \tilde{y}_m)^2 \tag{3.8}$$

$$H(X_m) = \frac{1}{N_m}\sum_{i=1}^{N_m}|y_i - \tilde{y}_m|, \tag{3.9}$$

where X_m is training data in node m, $\tilde{y}_m$ is a mean of all response variable values in the node m.

At the same time, when it comes to application of the decision trees algorithms to big amounts of data, there might occur issues related to scalability, complexity, stability and robustness of this method [128]. This might be partially avoided through tuning of the corresponding hyper-parameters, such as: tree depth, total number of instances in a leaf, total number of nodes or leaves.

3.4.4 Gradient Boosting Regression

To avoid disadvantages of the regular decision tree algorithms, Gradient Boosting (GBR) method can be applied. It is a machine learning technique available both for classification and regression that is based on "*general problem of producing a very accurate prediction rule by combining rough and moderately inaccurate rules-of-thumb*" [129]. Instead of building one decision tree, this technique creates many trees, where each of them is capable of predicting certain set of the training samples well. It is an ensemble algorithm that creates an additive model, where simple models (weak learners) are added one at a time, while the existing decision tree learners remain the same.

Gradient Boosting regression calculates pseudo-residuals (remaining error or the difference between the known target variable value and the predicted value). As a next step it fits a weak learner to these pseudo-residuals and performs a gradient descent procedure to minimize the loss when adding a new learner to the overall model. During this procedure the new decision tree is parametrized, and its parameters are modified. This allows to push the overall model to the desirable target through optimization of a chosen differentiable loss function. In case of regression tasks, it can be an MSE or MAE. Repeating this step over and over again helps to increase the model's prediction quality.

3.4.5 Adaptive Boost Regression

Adaptive Boost regression or AdaBoost (AB) is also an ensemble method partially similar to the Gradient Boosting learning method. It also creates weak learners to enhance the model's predictive abilities. However, unlike Gradient Boosting, it changes the sample distribution by modifying the weights attached to each of the instances [129]. It increases weights of the samples that were wrongly predicted and decreases them for those predicted correctly. As a result, the weak learner added on each new iteration focuses on the "problematic" samples.

3.4.6 Random Forest

Random forest is another ensemble algorithm that uses decision trees as separate learners [130], in this case a collection of multiple learners is called a forest. However, unlike GBR and AdaBoost, it does not add new decision trees based on the results obtained from the previous ones. The random forest is not sensitive to the number of trees in the forest or a number of features in each node, at the same time it is robust to overfitting [131]. The main steps of the algorithm are as follows:

1. Generate n bootstrapped[1] samples from a training dataset. These samples are drawn randomly and some of the samples might repeat.
2. Generate a forest by fitting the decision tree regressors to each of the n bootstrapped samples for a randomly chosen number of features for the node splitting.
3. The average value of all the decision tree predictions is used as the model's output.
4. Evaluate the overall random forest model performance using part of the training data that has not been used (out-of-bag dataset) and one of the metrics for the performance valuation (MSE, MAE, etc.).

3.5 Model optimization and evaluation

3.5.1 Model optimization using Grid Search

To tune hyper-parameters of the models obtained using the different ML methods, different search methods can be utilized, for example, Grid Search or Random Search. Grid Search is a method that exhaustively looks through each combination of the hyper-parameters. At the same time, Random Search uses random combinations of the hyper-

[1] Bootstrapping is a resampling technique that creates smaller data samples from the larger dataset with replacement.

parameter values. As a result, not all the hyper-parameter values are tried, but a number of combinations limited with an iterations number. In case of a large hyperparameter space, Random Search is able to explore a wider range of the searched space.

3.5.2 Model evaluation

Due to development of exclusively regression models in this work, to evaluate the obtained models' performance 5-folds cross-validation, correlation coefficient, root mean squared error (RMSE) and R^2 are used. Cross validation is a statistical method used to estimate skill of various machine learning models [114]. To perform k-folds cross-validation the next steps can be followed:

1. The dataset is randomly shuffled.
2. The dataset is split into k groups. In our case, $k = 5$.
3. For each of the k obtained groups:
   - Remove group from the rest of the data and leave it as a test dataset.
   - Use the remaining groups as a training set.
   - Train the model using the obtained training dataset and evaluate the model using the training set.
   - Obtain the model's evaluation scores.
4. Summarize the evaluation scores of the obtained model.

3.6 Advantages and disadvantages of ML methods

Machine learning is an extremely powerful tool that can be used to solve various problems and often outperform the classic mathematical modeling methods. Some of its main advantages are:

1. **Ability to map complex non-linear relationships.** As it has been discussed, ML is capable of processing large amounts of data and find previously unseen patterns and relationships.
2. **Handling of multi-dimensional data and data of different types.** ML can deal with highly multi-dimensional datasets that include data of different types using features pre-processing and feature selection methods to filter out the irrelevant features and leave only those that have a high impact on the variable of interest.
3. **Process automation.** Proper implementation of the machine learning methods allows to create models capable of learning from the data, making predictions and improving the algorithms and processes on their own.
4. **Continuous improvement.** As more data is obtained and the model is updated based on this data, its prediction capabilities are improving.
5. **Wide application range.** ML methods are applied to solve tasks in various fields, such as: information security [19], medicine [20], robotics [21], manufacturing industry [23], etc.

At the same time, any approach has its disadvantages, and ML is not an exception:

1. **Data acquisition.** ML approach requires significant amounts of high-quality data to be available. This means that corresponding procedures to gather and ensure quality of the data need to be available and implemented for the task in focus.

2. **Time and computation demand.** To create high quality models, enough computation power and time resources need to be allocated.
3. **Models interpretability.** Some of the ML methods create models that are hard to interpret for a human expert (ANN, Random forest, AdaBoost, etc.). Due to this such models can not contribute with explicit process understanding [28].

Chapter 4

Methodology

Although I do not suppose that either of us knows anything really beautiful and good, I am better off than he is – for he knows nothing, and thinks that he knows. I neither know nor think that I know.

– Socrates

This chapter presents the research philosophy used within this study and describes methods that were applied to conduct the relevant experiments, data collection and analysis, so that those interested in this research topic can repeat the experimental work when needed.

4.1 Research philosophy

Research design depends on several factors, such as the field of research, academic environment, as well as the researcher's background [132, 133]. Creswell [132] highlights four classical research paradigms that take part in shaping the research design: *positivism*, *pragmatism*, *realism* and *interpretivism*. These paradigms can be described as an intellectual context to conduct the research, beliefs that guide corresponding actions and a source of assumptions and ideas [132, 133].

- *Positivism* is a philosophy that regards reality as that consisting of discrete events that can be precisely observed, described and further used to produce general rules to forecast events or behaviors with a minimum uncertainty.
- *Pragmatism* "*is concerned with action and change and the interplay between knowledge and action*" [134]. It is an action-oriented paradigm that includes intervention into the world and not just its observation. For a pragmatist, research starts with a practical problem that needs a practical solution for future practice [135]. Within this paradigm, the best research methods are those that help to solve the research question in the most effective way.
- *Interpretivism* is completely opposite to the positivism. Here, the reality is considered as a socially constructed one, meaning that it is affected by the beliefs, opinions and values of those looking at it.
- *Realism*, at the same time, is rather close to positivism, however, takes into account the subjective nature of research and the corresponding conclusions drawn by it. Here, fewer claims are made regarding the knowledge that perfectly describes the object of study.

Interpretivism and realism are paradigms mostly used in qualitative studies, while the one described in this dissertation is a quantitative work based on experiments and the obtained results. Therefore, combination of positivism and pragmatism are used here. More discussion on this is presented in Chapter 11.

Injection molding is a net-shape manufacturing process that usually provides significant time and cost advantages, where, however, quality of the final product is highly dependent on the initial process parameter settings [57]. Therefore, many research studies in this field, as shown in Chapter 2, collect certain amount of data and observe influence of the investigated parameters on different quality aspects of the produced parts to establish relationships between the parameters and the observed variables. Such studies correspond to the positivistic paradigm.

Following this way of thinking and the positivism research paradigm, the first experiment was planned to evaluate a hypothesis related to the influence of various process parameters on the quality features of the focus part. Correspondingly, the RQ1 is related to selection of the most influential process parameters.

RQ2 and RQ3, at the same time, are focused not only on evaluation of the relationships between the quality and the process parameters, but also on the application of the ML methods that have some advantages over the conventional modeling techniques. This point of view is highly correlated with that of the pragmatism, as a more effective method and its application to the question of interest are investigated.

To answer the RQ4 the first three research questions need to be answered and incorporated so that a framework for an intelligent control system for thermoplastics injection molding is proposed. As a result, combination of positivism and pragmatism are used. Synergy of these two philosophical research paradigms allows greater flexibility in the research design of the study. It provides certain freedom for planning of the research path depending on the results obtained on different stages of the presented work.

4.2 Research methods

As mentioned above, the research study described in this dissertation is a quantitative one. It is based on experimental work and the knowledge is derived through analysis of the obtained experimental data. Research methods are chosen according to the research questions of interest and the relevant philosophical research paradigms. Design of experiments, injection molding of the focus parts (experimental work), data collection and analysis are the essential research activities conducted in this work. The performed data collection, in its turn, can be divided into three stages: collection of the IM process data, collection of dimensional data and collection of physical properties data of the obtained parts. The dimensional and mechanical properties data together form quality data. Table 4.1 shows hardware and software solutions used at the different stages of the research study. In addition, literature review, application of model-based systems engineering (MBSE), validation and verification were also performed. A Gant chart depicting the timeline of the main research activities is shown on Figure 4.1.

Table 4.1. Hardware and software used in the research activities

	Hardware	**Software**
Design of Experiments	Computer	modeFrontier [136]
Injection Molding	"ENGEL insert 130" with CC300 control unit	Linux OS
IM process data collection	Computer, RevPi Core3 [137], Kistler 4021B10H1P1 temperature and pressure multi-sensor	Python 3 programming language, PyCharm, Atom

	Hardware	Software
Dimensional data collection	Zeiss DuraMax [138]	Zeiss Calypso
Physical properties data collection	Instron 5960	Bluehill Universal
Data analysis	Computer	Windows OS, Python 3 programming language, PyCharm, Jupyter notebook

4.2.1 Literature review

Literature review has been an important part of this study. As shown on Figure 4.1, it has been conducted all the way along the PhD work in parallel to the other activities. This was necessary to be up to date on the state-of-the-art in the relevant study field. Scopus, Science Direct and Google Scholar databases were used to perform the literature search. The articles were filtered based on the relevance to the topic of interest and publication year. Papers published before the year 2000 were not included in the literature review presented in Chapter 2.

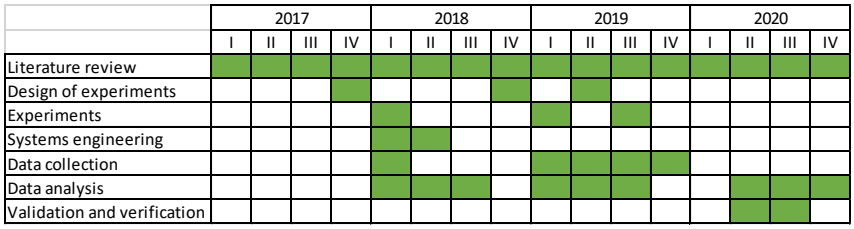


Figure 4.1. PhD research project Gant chart

The first stage of the literature review was conducted in the very beginning of the project in order to get familiar with the current developments in the field of monitoring and control of thermoplastics injection molding. As a result, it became clear that the field of IM still does not utilize process data in full capacity to extract relevant patterns for monitoring and control, and a lot of scholars use such methods as Taguchi method, FEM, RSM and ANOVA to obtain the optimal process parameter settings, while SPC is used to monitor the process. Since those methods have several disadvantages, there was a number of papers where the ML methods were applied and argued to be a better alternative, due to their ability of learning from the data and effectively mapping non-linear relationships. Most of those papers, however, lack a universal approach and were often focused on rather specific cases. Therefore, it became apparent that a general framework for application of ML methods for intelligent control of thermoplastics injection molding process is needed. However, to propose a good framework (RQ4), responses to the RQ1, RQ2 and RQ3 also need to be given.

Later, literature review was conducted again and again to be able to keep up with the current developments in the relevant field. During the literature review different search words were used that can be divided into three groups “process in general”, “modeling, monitoring and control methods” and “other”. Examples of the search words used within each category are presented in Table 4.2. The “process in general” keywords were mostly used to find the relevant research works about the overall description of the injection molding process, important process parameters that influence quality of the produced parts, the most common IM defects, etc. “Modeling, monitoring and control” search words included different queries related to the more conventional (FEM, RCM, Taguchi

method), as well as ML modeling, monitoring and control techniques. The group "other" consists of the key words related to different accompanying concepts necessary for the in-depth understanding of topics, such as data acquisition process, mechanical testing of parts or coordinate measurement.

Table 4.2. Examples of keywords used in the literature review

"process in general"	**"modeling, monitoring and control methods"**	**"other"**
Injection molding, thermoplastics injection molding, injection molding of plastic parts, injection molded parts, injection molding parameters, the most important injection molding parameters, plastic injection molding important parameters, common injection molding defects, prediction of injection molding parts quality, optimization of the injection molding process parameters	Taguchi method, SPC, statistical process control, FEA, finite element analysis, FEM, finite element method, RCM, response surface methodology, ANOVA, analysis of variance, DOE, design of experiments, Latin Hypercube, ML methods, machine learning methods, classification, regression, ANN, artificial neural networks, decision trees, random forest, feature selection methods, correlation-based feature selection	CAD model, coordinate measuring, tensile stress testing, tensile testing, mechanical testing, Young's modulus, tensile modulus, tensile strength, tensile strain at break, data acquisition, data acquisition system, MES, manufacturing execution system

4.2.2 Design of experiments

Based on the first stage of the literature review, the RQ1, RQ2 and RQ4 were formulated and the first experiment was designed. To design the first and the following experiments, Latin hypercube sampling (LHS) [139] was used. Unlike simple random sampling, the LHS stratifies simultaneously on all input dimensions, therefore, leading to a better coverage of the corresponding parameters hyperspace. Since DOE for the experiments in this work included from 6 to 8 different parameters and value range for some of them was quite wide, LHS was chosen as a sampling method capable of covering as much of the hyperspace of the parameter values of interest as possible.

4.2.3 Experimental work

Experimental work was performed during the years 2018 and 2019, as shown on Figure 4.1, and included four experiments. Experiment 1 was planned and conducted during the year 2018, after the first stage of the literature review and formulation of the RQ1, RQ2 and RQ4. In 2019 the experiments 2-4 were performed and the RQ3 was formulated. After completing the injection molding of the focus parts and collection of the corresponding process parameters data within the experiments, dimensional measurements and tensile testing were performed to obtain the quality data.

Originally, the first experiment was planned to answer the RQ1, RQ2 and then RQ4. The RQ3 was added later due to the importance of the parts mechanical properties when defining the final product quality. This, however, did not prevent using data obtained from the first experiment to assist answering the RQ3. More details about the experimental work are given in Section 4.3.

4.2.4 Model-based systems engineering

Systems engineering is an interdisciplinary approach that includes techniques for successful realization of complex systems [140]. A system, in this case, can be defined as *"any organized assembly of resources and procedures united and regulated by*

interaction or interdependence to accomplish a set of specific functions" [141]. In this work the answer to the RQ4 is meant to propose a framework that can assist development of an intelligent control system for thermoplastics injection molding. Such system consists of multiple components and connections between them, where the relevant data is collected and analyzed. Model-based systems engineering is, therefore, applied to improve the data collection and analysis process, present the obtained results in a systematic way and propose a framework for the intelligent system of interest and application of the ML methods within it. More details on the model-based systems engineering application are presented in Chapter 5.

4.3 Experimental work

As mentioned before, the experimental work within this study consisted of four experiments. All of them were conducted on the "ENGEL insert 130" vertical IMM with CC300 control unit. During the experiments dogbone specimens with 4 mm and 15 mm thickness, as depicted in Figure 4.2, were produced. The experiments were performed following different DOEs designed in advance using the Latin hypercube sampling method. The used DOEs are presented in Appendix A of the dissertation. Table 4.3 provides more information on the difference between the experiments. When it comes to the difference in the DOEs, in experiment 1 eight parameters were included (holding pressure, holding time, backpressure, cooling time, injection speed, screw speed, barrel temperature, mold temperature) and 32 parameter combinations, in experiment 2 – six parameters (barrel temperature and mold temperature were excluded) and 24 combinations. In experiments 3 and 4 the same DOE was used that included seven parameters (mold temperature excluded) and 32 parameter combinations.

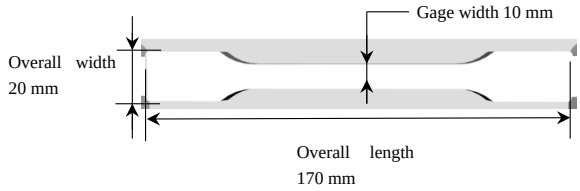


Figure 4.2. Injection molded dogbone specimen

In the first experiment, each combination was launched five times, while in the experiments 2-4 – three times. However, during the experiments 3-4 some combinations were used more times to finish the material in the IMM machine. In the experiment 3: combination number 10 was launched ten times and number 32 – seven times, while in the experiment 4 combination 32 – nine times.

As it is shown in the Table 4.3 there are several parameters that distinguish the experiments. During the experiments 1, 3-4 dogbone specimens with 4 mm thickness were produced, while during the experiment 2 15 mm thick specimens were manufactured. In the experiments 1-2 the virgin HDPE material was utilized, while in the experiments 3-4 recycled HDPE from two different suppliers was used.

Due to a large number of IMM runs within the same experiment and some of the parameter adjustments taking a long time to be applied, not all of the specimens in the experiments 1 and 2 were produced during the same day. In the experiment 1 specimens for the DOE combinations 1-10 were produced on 22.03.2018, 11-21 on 23.03.2018 and 22-32 on 10.04.2018. In the experiment 2, batches 1-2 were produced on the 28.01.2019, 3-10 on the 31.01.2019 and 11-24 on the 08.02.2019. However, within each experiment all the specimens were produced using the same material batch.

Table 4.3. Description of experiments 1-4

Exp. #	DOE parameters	Num. of combin. in DOE	Focus parts	Num. of specimens in one run	Material	Mold sensor data collection	Total num. of param. logged	Total num. of IMM runs	Num. of runs per DOE combin.
1	8	32	4 mm dogbone specimen	2	BorSafe™ virgin HDPE	No	41	160	5
2	6	24	15 mm dogbone specimen	1	BorSafe™ virgin HDPE	Yes	65	72	3
3	7	32	4 mm dogbone specimen	2	Recycled HDPE from ContainerService supplier	No	52	101	3
4	7	32	4 mm dogbone specimen	2	Recycled HDPE from RePro supplier	No	52	102	3

Table 4.3 depicts that during the experiments different number of the corresponding process parameters were logged. In the experiments 1, 3-4 only the parameters data available from the IMM was obtained, while in the experiment 2 corresponding mold pressure and temperature data was also logged from the mold cavity multi-sensors. This data was not obtained in the other experiments due to the sensor availability only in the mold for production of the 15 mm thick dogbone specimens.

After each of the experiments, the produced specimens were measured on the coordinate measuring machine to obtain the dimensional data (width and thickness) and then tensile tested to get the mechanical properties data. Figure 4.3 shows a general flow of the experiments conducted in this work. On the figure large blue arrows indicate the experiment's inputs (DOE and material), gray rectangles are different stages of the experimental work, blue rectangles are outputs of the different stages of the experiment (data and injection molded specimens). Green arrows indicate collected data, while blue arrows connect different stages of the experimental process.

4.3.1 System used as a case

All the experiments were performed using "ENGEL insert 130" vertical IMM with CC300 control unit. Process parameter settings were varied according to the corresponding design of experiments shown in Appendix A. Prior to the start of the injection molding, the used material was dried, and the IMM was warmed up. In addition, several cycles not included into the collected data were performed before the beginning of the experiment and when the process parameter combination was changed. This was done to stabilize the production process with the new parameter values.

To log the injection molding process parameters data an application programing interface (API) developed at SINTEF Manufacturing AS in Python 3 programming language was used. The data was collected each 0.5 s from the closing until the opening of the mold throughout each production cycle. As mentioned before, during the experiment 2 mold pressure and temperature data was also obtained using Kistler 4021B10H1P1 temperature and pressure multi-sensor. More information about the sensor is provided in Table 4.4. The mold for production of 15 mm dogbone specimens has two of the sensors installed, as shown on Figure 4.5 (a), however, data from only one of them was logged as during this experiment one dogbone specimen was produced per cycle. Examples of mold pressure and temperature curves data obtained from the sensors is shown in Figure 4.4.

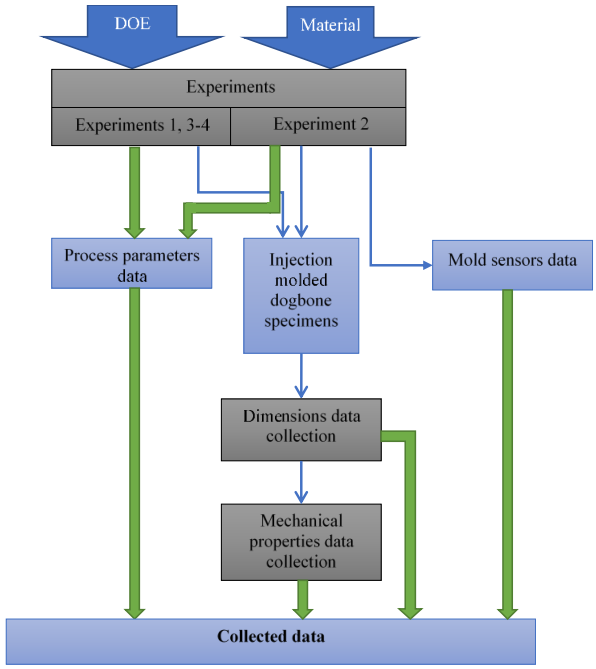
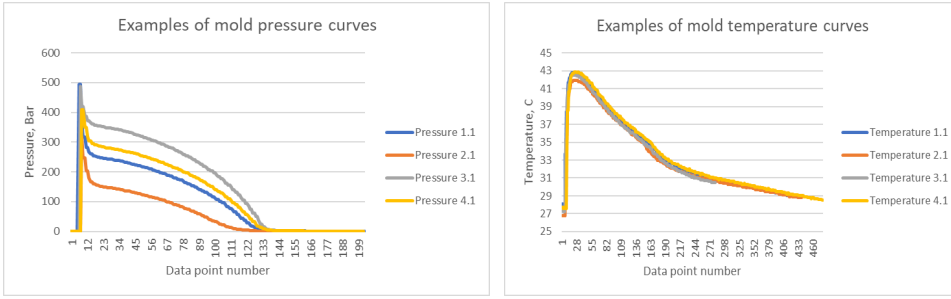


Figure 4.3. A general experimental flow

Table 4.4. Pressure and temperature multi-sensor characteristics

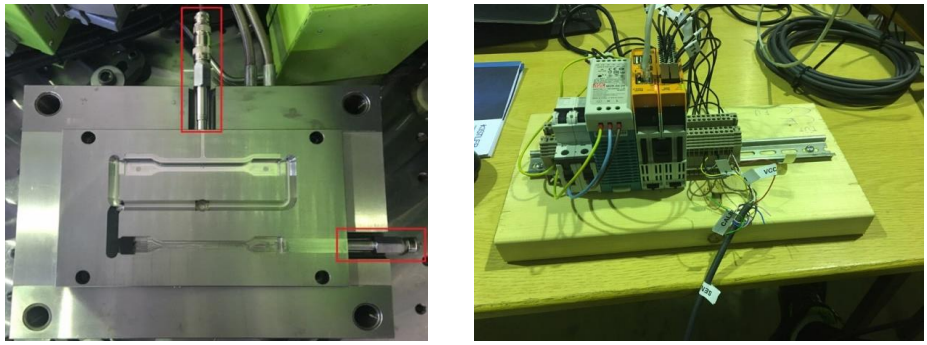
Specifications	**Kistler type 4021B10H1P1**
Model	Measuring chain
Calibration	Calibrated by Kistler
Measuring Range pressure (Bar)	0..1000
Measuring Range temperature (^{0}C)	0..350
Temperature accuracy (^{0}C)	±5
Diameter (mm)	21
Height (mm)	91.5
Natural Frequency (kHz)	>165



(a) (b)

Figure 4.4 (a) Examples of mold pressure curves obtained from the in-mold sensors; (b) Examples of mold temperature curves

To acquire data from the sensors the industrial Raspberry Pi RevPi Core3 and RevPi Analog Input Output (AIO) modules for the digital sampling of the signals were used, shown on Figure 4.5 (b). The modules are connected through PiBridge. Unlike commercial DAQs from manufacturers like NI and HBM, RevPi is an open, modular industrial PC which provides flexibility of the software alternatives [137]. The sampling rate is about 125 Hz due to the load on the PiBridge. This leads to about 5 ms update time on the PiBridge for each AIO module connected, which is, however, sufficient for logging of data each 0.5 s.



(a) (b)

Figure 4.5 (a) Mold for production of the 15 mm dogbone specimens with sensors installed; (b) RevPi used for the mold sensors data acquisition

Due to an error during the data acquisition during the experiment 4, data for the first run of the 8th, 9th, 10th and 13th DOE combinations was not logged and, therefore, these specimens were excluded from further analysis.

4.3.2 Dimensions data collection

To obtain the dimensions measurements (width and thickness) of the specimens coordinate measurement was performed. Zeiss DuraMax coordinate measurement machine (CMM) was used for this purpose. The CMM has accuracy of $\pm(2{,}7 + L/250)\ \mu m$, where L is the measured value [138]. A stylus with three pins with 1,5 mm diameter was assembled based on the expert's recommendation from Zeiss. Due to the low weight of the specimens, a special fixture, depicted in Figure 4.6, was designed to assist in the measurement process and provide a stable fixation of the measured parts.

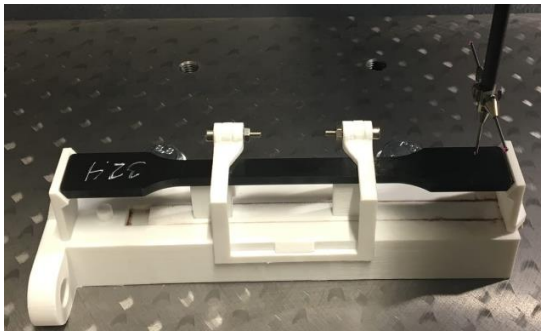


Figure 4.6. Fixture for coordinate measuring of the specimens

Width and thickness were measured in three points according to ISO 16012:2015: Plastics – Determination of linear dimensions of test specimens [142], as shown on Figure 4.7. The specimens were measured in three points within the narrow section. The final values of the dimensional features were calculated as a mean of the three corresponding measurements. The specimens were always situated with the same side facing up. It is easy to determine as the other side of the specimen includes injection molded marks for understanding of whether it is the first or the second specimen produced during the same run. In addition, the side that was connected to the runner was always on the right hand-side, when facing the measured specimen, as on Figure 4.6.

4.3.3 Mechanical properties data collection

To collect the mechanical properties of the focus parts, tensile testing was performed according to the ISO 527-1 [54]. Instron 5960 universal system with 10 kN load cell and a video extensometer were used to complete the tests. As it can be seen on Figure 4.6 the specimen has two white dots on its narrow section. The dots were added manually using an Instron stencil prior to the testing to assist the video extensometer's work. The testing speed was set to 1 mm/min until 0,25% strain and 15 mm/min after that according to the expert's from Instron recommendations. Due to an error on the extensometer, mechanical properties data for specimens 4.3.2, 28.1.1, 29.3.1 (the first number indicates the batch number, the second number the run number and the third number – the specimen number) from experiment 3 was not obtained. Therefore, these samples are not included into the dataset used for prediction of the mechanical properties of the produced parts.

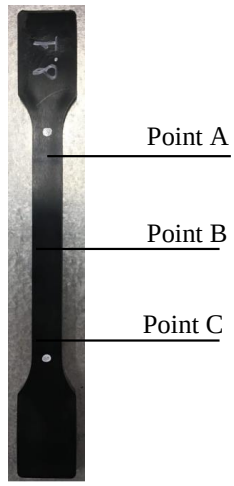


Figure 4.7. Points for measurement of width and thickness of the focus part

4.4 Data preprocessing and analysis

To conduct the necessary data analysis, prototyping and models development, Python 3 programming language and WEKA – WAIKATO Environment for Knowledge Analysis [143] were used. WEKA was often applied to test different feature selection and ML methods to see their relevance for solution of the tasks within the study. If the methods appeared to be useful, they were then reapplied using Python and the necessary libraries for execution of the algorithms of interest. Both WEKA and Python were chosen due to being open-source, easy to use and supporting execution of various complex algorithms through either a user-friendly interface (WEKA) or a "one-line" code manner (Python). This allowed to concentrate on the tasks at hand and the necessary models development rather than on programming the ML methods from scratch. Figure 4.8 shows the list of the main Python libraries used within the study.

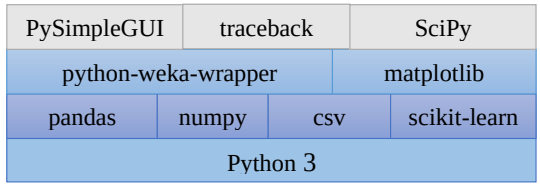


Figure 4.8. Python libraries used within the PhD study

The work with the collected data included the following main steps:

- **Data preprocessing** step is dedicated to collection of the obtained data of different types into the .csv files that are convenient for further analysis, removal

of outliers, erroneous (NaN, null) and redundant data and normalization. For these operations **pandas**, **numpy**, **csv**, **scikit-learn** Python libraries were used. As a results of this step general dataset files were created, as well as data series dataset files. Table 4.5 shows the number of samples in different datasets. The sizes of datasets used for creation of models for prediction of dimensional properties and those of physical properties differ, as the physical properties ones have outliers related to the target variable values removed. More details on these datasets will be provided in Sections 4.4.1 and 4.4.2.

- **Feature selection** step was necessary to gain a better understanding of the relationships between the process parameters and the target variables, as well as to define which parameters contain more information about the quality variables and are therefore to be included in the prediction models. Pearson correlation, Spearman's Rho correlation, Regressional ReliefF (RReliefF), CFS and RFE methods were applied with help of the **scikit-learn**, **python-weka-wrapper** Python libraries and WEKA. The above-mentioned methods were selected since some of them are better for analyzing the relations between the parameters and others for understanding which parameters are more important to be included in the prediction model.
- **Data visualization** is performed to achieve a better data understanding, illustrate different data distributions and assist in detection of the outliers. **Matplotlib, PySimpleGUI, traceback** libraries were used to create the relevant data graphs and visualizations and display them in a convenient way (through a simple user interface).
- **Data analysis** (processing) focused on application of the chosen ML methods to create the models for prediction of the dimensions and mechanical properties of the dogbone parts and optimize the process parameters. **Scikit-learn**, **SciPy** libraries and WEKA were utilized to apply MLP, Decision Trees for regression, AdaBoost Regressor, Gradient Boosting Regressor and Random Forrest. The main reasons for choosing these methods were their suitability for solving the regression tasks and ability to filter out the less important features (the last point is valid for all the methods except for the MLP).

Table 4.5. Sizes of the obtained datasets

Dataset name	**Dataset size**	
	Dimensions dataset	Physical properties dataset
Experiment 1 general parallel dataset	160	142
Experiment 1 general sequential dataset	320	296
Experiment 2 general dataset	72	65
Experiment 3 general parallel dataset	101	84
Experiment 3 general sequential dataset	202	176
Experiment 4 general parallel dataset	98	83
Experiment 4 general sequential dataset	196	177
Parallel joined dataset	359	309
Sequential joined dataset	790	714
Experiment 1 cushion data series parallel dataset	160	142
Experiment 2 cushion data series dataset	72	65
Experiment 2 mold pressure data series dataset	72	65
Experiment 2 mold temperature data series dataset	72	65

Dataset name	Dataset size	
Experiment 2 screw position data series dataset	72	65
Experiment 3 screw position data series parallel dataset	101	84
Experiment 4 screw position data series parallel dataset	98	83

4.4.1 General datasets preprocessing

To obtain high quality general datasets, the data needed to be cleaned and preprocessed before its further use. At first, the process parameter data was transformed. During the data collection process, data related to each production run was stored in a separate file in form of a data series (the data was logged each 0.5 s during the cycle). However, it was of interest to transform this data so that each parameter had one corresponding value per cycle and to store it in one file per experiment. Since different injection molding parameters logged from the IMM have different update frequencies and some of them change often during the cycle, while the others do not change at all, various parameters were treated differently. Some of them were averaged and the parameter name was then changed to "Parameter_name_average", while for those that did not change throughout the whole cycle (but vary from cycle to cycle) the first value logged was taken. For the parameters that change after the switchover point (change from the constant velocity to the constant pressure control), the value after the switchover was taken. More details on how different parameters were treated can be found in Appendix B.

Secondly, since there were three different data collection stages (parameter data collection, dimensions data collection and mechanical properties data collection) the data from the coordinate measurement and tensile testing was added to the IM process data. The list of the quality parameters that are the models' outputs can be seen in Table 4.6. The parameters columns in the .csv files were given meaningful names and it was checked that the names of the columns are the same in all the files corresponding to the different experiments.

As a third step of the data preprocessing, samples with missing data were removed. The samples that had NaN or null values were eliminated from the datasets. As mentioned before, such samples were samples of the first run of the 8^{th} (8.1.1, 8.1.2), 9^{th} (9.1.1, 9.1.2), 10^{th} (10.1.1 and 10.1.2) and 13^{th} (13.1.1 and 13.1.2) DOE combinations in the experiment 4 (due to a data acquisition error). In addition, redundant features were removed. Their redundancy was based on opinion of a human expert that looked through the list of the logged process parameters.

Table 4.6. Model outputs (quality characteristics of interest)

Model outputs/ quality characteristics
Width
Thickness
Young's modulus
Tensile strength
Tensile strain at Break

The fourth step was dedicated to the elimination of the outliers. The fifth step included feature selection, while during the last step, data normalization was performed. The data in the datasets were transformed using *z-score* method to obtain data with zero mean and unit variance. This part of the data preprocessing is depicted in Figure 4.9 (a). As a result of this preprocessing seven datasets and corresponding to them .csv files were obtained and made ready for further analysis.

Since in the experiments 1, 3-4 two specimens were produced at the same time, while one was manufactured during the experiment 2, two datasets were created for the experiments 1, 3-4 and one for the experiment 2. Two datasets for the experiments 1, 3-4 differ in the way the quality characteristics data is added. In one of the files the number of the samples is doubled, and each sample is presented as such that has the target values for only one of the produced specimens. Excerpt of such dataset is shown in Table 4.7, where in the "Code" column the first number indicates the used DOE combination, the second number shows for which time was the combination launched, and the third number the number of specimen produced during the same run. In the second file, however, each sample includes data for quality characteristics of both specimens.

Table 4.7. Example of the dataset for experiment 1

Code	Screw_speed_max	Spec_pres_switchov	Modulus Young's	Tensile strength	Tensile strain at Break	Thickness	Width
1.1.1	239.18	1058.69	944.00	26.10	22.73	3.77	9.84
1.1.2	239.18	1058.69	982.27	26.43	25.38	3.78	9.88
1.2.1	239.40	1059.35	1091.88	26.49	23.37	3.76	9.80
1.2.2	239.40	1059.35	1020.72	26.80	26.06	3.76	9.82
1.3.1	239.52	1059.62	1045.96	26.79	25.82	3.77	9.82
1.3.2	239.52	1059.62	1038.91	26.26	22.82	3.76	9.86
1.4.1	239.16	1059.09	1006.58	26.89	25.26	3.77	9.79
1.4.2	239.16	1059.09	938.96	26.57	20.67	3.76	9.81

4.4.2 Data series datasets preprocessing

In addition to the general datasets, data series datasets were also created. This was done to analyze complete sensor data profiles for material cushion in experiment 1, screw position, material cushion, mold pressure and mold temperature in experiment 2, screw position in experiments 3 and 4 (depending on the available parameters logged during the experiments). Information on which parameters were logged during each experiment is provided in Appendix B. The obtained datasets include transposed value vectors for the parameter of interest and the corresponding quality characteristics. This way the sensor measurements each 0.5 s become the model's input features or parameters, while the quality characteristics are the models' outputs or targets. Only parallel data series datasets were created. As a result, the datasets for experiments 1, 3-4 include quality data for both specimens produced during one cycle in the same data sample.

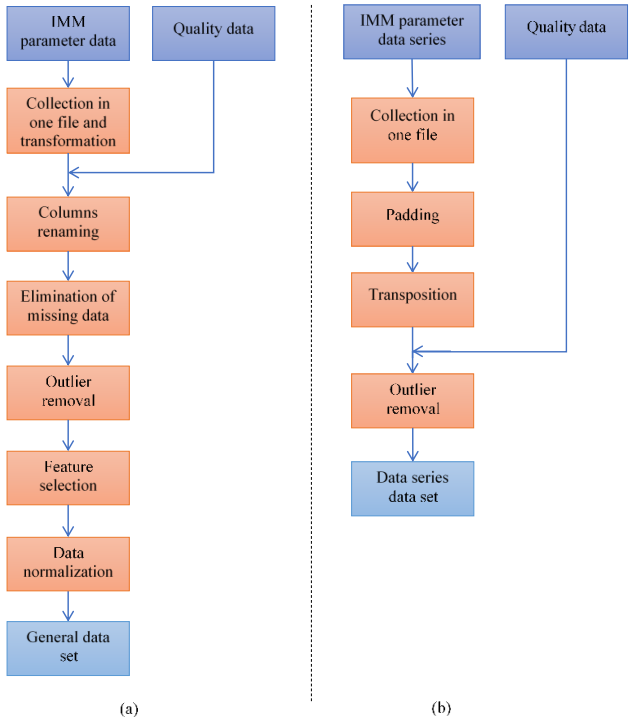


Figure 4.9. The process of the datasets creation: (a) general dataset, (b) data series dataset

To obtain these files, at first, the data series of the parameter of interest for each production cycle within the experiment were collected in one file. Due to the different cycle times for production of parts with the various DOE combinations, those data series were of different length. As a second step, to be able to work with this data, the data series that were shorter than the longest one were extended with zeros (padding technique), so that all the parameter vectors had the same length. Next, the obtained vectors were transposed, so that a value of the parameter each 0.5 s becomes a separate model input parameter. As a last step, the quality characteristics data was added to complete the data series dataset files. Figure 4.9 (b) shows a graphical representation of this process.

4.4.3 Data integration

In addition to the above-described datasets, a dataset that included data from 1, 3 and 4 experiments (parallel joined dataset) and a dataset that included data from all the experiments (sequential joined dataset) were also created. The parallel joined dataset includes data only from the machine runs where two dogbones were produced simultaneously. Since the previous datasets do not have any information on the material used in the experiment, a new column with the material type was added in this dataset. As it was mentioned, not all the specimens within the same experiment were produced in

one day. As a result, different codes were used to indicate that, even though the material type and batch were the same within one experiment, some of the material properties might have been influenced by the temperature and humidity in the storage room. For the experiment 1 M1.1, M1.2 and M1.3 were used to indicate that the samples were produced in three different days. For the experiment 2 M1.4, M1.5 and M1.6 was used following the same logics (since in both experiments 1 and 2 virgin HDPE from the same supplier was used), for the experiment 3 M2 was used, while for the experiment 4 – M3. Table 4.8 shows which material codes correspond to which experiments.

In addition, a column indicating the part type (dogbone specimen with 4 mm and 15 mm thickness) was also added. Number 1 corresponds to the 4 mm thick specimen, and 2 to the 15 mm thick specimen. All the above described datasets were then used for creation of the quality characteristics prediction models using the following ML methods: MLP, Decision Trees for regression, AdaBoost Regressor, Gradient Boosting Regressor, Random Forrest.

Table 4.8. Material codes for different experiments

Experiment number	Material code
1	M1.1, M1.2, M1.3
2	M1.4, M1.5, M1.6
3	M2
4	M3

4.5 Validation and verification of the results

The results of this work include predictive models, prototypes of the modules of the intelligent system for thermoplastics injection molding and a framework for development of this system. Therefore, validation and verification are needed. Validation is directed towards the obtained predictive models, while verification is considered the process of designing the framework for the intelligent system and its modules. The predictive models, at the same time, are parts of those modules.

The models' validation process consists of two main stages. The first stage is dedicated to application of five folds cross-validation during the models training. This helps to prevent overfitting and to create a model that is capable of generalizing on the data, rather than remembering the training set data samples. At the same time, the generalization abilities of the proposed models remain questionable, as more data from injection molding of parts with various shapes and materials is needed to create a model with greater generalization abilities.

The second step includes use of a testing dataset for testing the models' quality on previously unseen data. These datasets include samples that have not been used during the model training and are therefore new for the models. The resulting models' quality measures values (R^2, square root mean error or RMSE and correlation coefficient for the actual and predicted target variable values) are used to make decisions on the usefulness of the models and decide which of them should be used in the future work and developed further.

Chapter 5

Systems engineering for designing an intelligent control system for thermoplastics injection molding

Your best and wisest refuge from all troubles is in your science.

– Ada Lovelace

This chapter describes application of the model-based systems engineering for development of the intelligent control system for thermoplastics injection molding and can be used as a framework for creation of similar systems, it also partially answers the RQ4. In order to develop the system, it is necessary to have a good understanding of what are the functions of the system, the context in which it will be used, which modules should the system consist of and how they are connected to each other.

MBSE includes four main stages shown on Figure 5.1. The following sections cover requirements analysis stage in more details, while the rest of the stages are briefly covered. Before the requirements analysis, identification of the system's stakeholders and their needs is also presented.

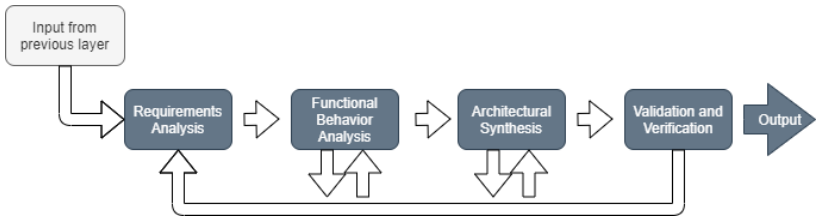


Figure 5.1. Model-based systems engineering process adopted from [141]

5.1 Identification of the system's stakeholders and their needs

According to a systems engineering framework called SPADE [144], the first stage of any system development process is identification of the system's stakeholders and their needs. Different types of stakeholders can be distinguished based on their attributes such as power, urgency and legitimacy [145]. In this study stakeholders are differentiated according to their needs and the level of their project involvement: **primary** – those that are directly involved, can influence the decision-making process and will be influenced by the project outcomes in the first place, **secondary** – those that might be influenced by the project outcomes later or by implication. Table 5.1 depicts primary and secondary stakeholders and lists their needs.

Table 5.1. Stakeholders and their needs

Stakeholders	Involvement	Needs
NTNU	Primary	• New knowledge • Increased process understanding • Systematic research • Publication and dissemination • Successful project completion

Stakeholders	Involvement	Needs
		• New funding • PhD candidate evaluation
SINTEF	Primary	• New knowledge • Increased process understanding • Systematic research • Publication and dissemination • Successful project completion • New funding
Manufacturing companies involved (Plasto AS, PipeLife Norge AS and AKVA Group ASA) and not involved in the project	Primary	• Increased process understanding • Adoption of ML methods • High quality products • Cost minimization • Reduction of the negative environmental impact • Guidelines • Increased competitiveness due to use of recycled materials and closed loop manufacturing implementation
The Norwegian Research Council	Secondary	• International recognition of Norwegian research • Stimulation of international research and cooperation between private companies and academia • Benefits obtained by the project participants • Successful completion of the project
Standards/Regulations	Secondary	• Systematic research
Environment, community, society	Secondary	• Reduction of the negative environmental impact

This PhD is part of the MegaMould project, therefore, all the project participants are stakeholders of the system under development. Their main interests differ due to NTNU and SINTEF having more research-oriented focus, while Plasto AS, PipeLife Norge AS and AKVA Group ASA being interested in business perspectives and increased efficiency of their production process.

The Norwegian Research Council is also a stakeholder due to providing the funding for this project. At the same time, other injection molding companies, as well as companies that produce IMMs not involved in the MegaMould are stakeholders that can gain additional process knowledge and guidelines on how the IM process can become more intelligent and efficient. Organizations working on the standards and regulations related to the injection molding process can also benefit from the systematic research conducted during the project. In addition, implementation of similar systems can benefit environment, community and society by reducing the negative environmental impact through increasing the production process efficiency and decreasing the amount of produced scrap.

Figure 5.2 depicts a context model for the intelligent control system for thermoplastics injection molding. Here suppliers provide materials to the manufacturing companies that use the system and the companies provide ready products to their customers. The intelligent control system is used by NTNU, SINTEF, as well as the manufacturing companies. NTNU, SINTEF and the companies provide the system with the necessary data and requirements for development and improvement. The system, in its turn, creates internal reports for NTNU, SINTEF and the manufacturing companies, as well as the relevant information and knowledge that helps to improve the research and production process. In addition, the system is also capable of providing external reports that might be useful for development of the new standards, utilized within other governmental institutions and by competitors.

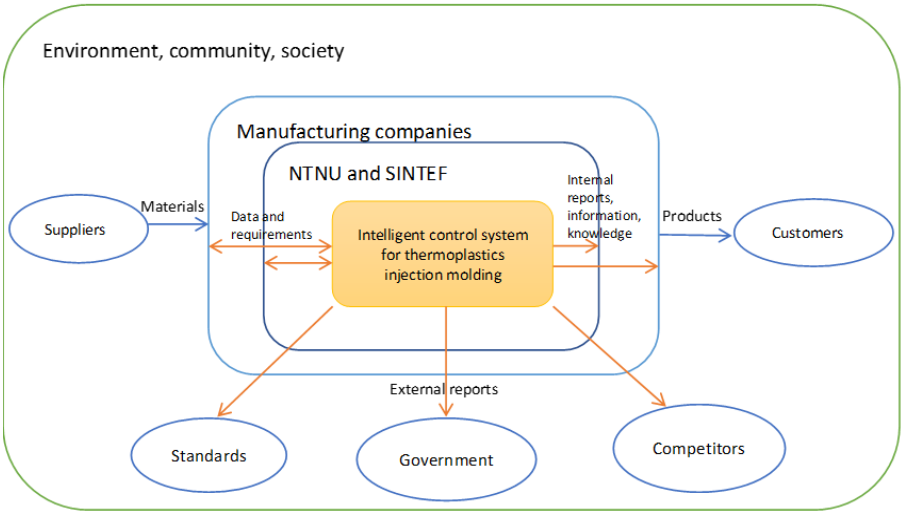


Figure 5.2. Context diagram for the intelligent control system for thermoplastics injection molding

5.2 Requirements analysis

In addition to identifying the stakeholders' needs, it is also necessary to conduct the requirements analysis, which is the first stage of the model-based systems engineering. It will allow to identify different aspects of the future system and give a better understanding of how the system should be created and which elements it needs to include. Setting clear requirements for a system, might also help to reach compromise between the system's stakeholders.

To conduct the requirements analysis the stakeholders need to be involved. Since this PhD work is part of the MegaMould project, the project description was analyzed, and the requirements of different stakeholders were identified. This way it is possible to consider the interests of the stakeholders and define what they expect to obtain in the end. To further understand how the various requirements can be incorporated into the intelligent control system for thermoplastics injection molding, the top-down MBSE principle is applied. This way the system is planned in layers, starting from a more general one and descending to the ones describing the system in more details.

5.2.1 Layer 1. General understanding of the system

On the first layer, three main components are included: the system users (customers, machine operators, scientists, etc.), the system itself and the production floor (injection molding machine(s)). Figure 5.3 depicts these connections. The users should be able to interact with the system and the system should be able to respond them through the interface and generated reports. The system should also be capable of doing the necessary analysis based on the user's request and the data obtained from the production floor. So,

communication between the system and the production floor is also essential, as the system obtains the necessary data from the IMMs, while the IMMs receive the solution produced by the intelligent system. As a result, the following functional requirements to the system can be identified:

- Accept the user input (IMM parameter values, quality data, report request, etc.);
- Analyze the user input;
- Obtain the data from the IMM(s) and additionally installed sensors (e.g. mold sensors);
- React to the user input (conduct prediction of the part quality, optimize the IMM parameter values, generate a report, etc.);
- Save the results;
- Send the results to the IMM(s) (e.g. optimized parameter values);
- Provide the result of the user input to the user (generate a report, display through the interface, etc.);
- Accept the user feedback.

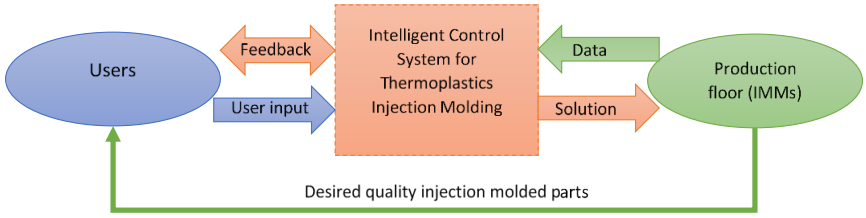


Figure 5.3. The intelligent control system for thermoplastics injection molding

5.2.2 Layer 2. Description of the main components of the system

After gaining the general understanding of the system's functionality, it is necessary to explain which components it needs to include to be able to meet the functional requirements. In this case, interface, database, application programming interface or API for communication with the IMMs and the calculation core are the most important components. The interface is needed to communicate between the user and the system: accept the user inputs, display and obtain the system's outputs. API is used to establish communication between the system and the IMMs, obtain the IM process parameter values and store them in the database. The database should be capable of storing the data that is entered by a user, uploaded as a file by a user, logged through the API from the injection molding machine, as well as save the data and models received as a result of the calculation core work. The calculation core needs to have access to the data stored in the database and to be able to process and analyze this data to create and update the quality prediction models, conduct optimization, as well as to make predictions using the models. Schematic representation of the system components and their connections is presented on Figure 5.4.

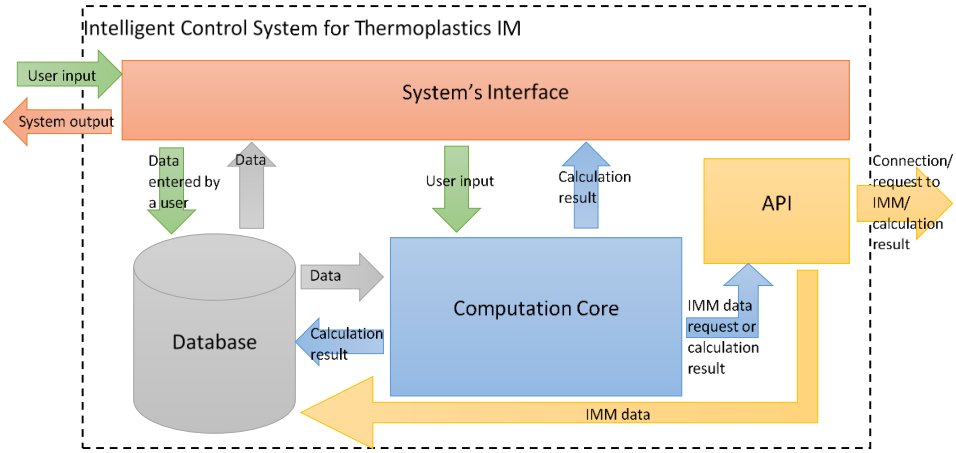


Figure 5.4. The main system components

Functional requirements to the system's interface are listed below:

- Should have an option of process or quality data file upload;
- Should have an option of manual parameter values input;
- Should have fields for entering the type of the IMM, material in use and a task to be solved (prediction of quality features, creation of a model, optimization, etc.);
- Should be able to output the task solutions in a meaningful way through the interface or as a downloadable report.

5.2.3 Layer 3. Description of the API and the database system components

In order to be able to analyze the injection molding process data, an API for connection with the IMM needs to be created. In this work, an API developed in collaboration with SINTEF Manufacturing AS was used and more details about it can be found in [17]. The API is the main component responsible for connection with the injection molding machines, obtaining IMM and additional sensors data and setting new parameter values based on the computation results, when needed. The API has the following functional requirements that are partly reflected on Figure 5.4:

- Establish connection with IMM(s);
- Obtain and process the IMM/ additional sensors data request;
- Send a request and obtain the requested data from the IMM(s)/ mold sensors;
- Send the requested data to the database;
- Obtain the calculation results (e.g. optimized parameter values) from the computation core;
- Set the parameter values on the IMM based on the calculation results.

To guarantee the correct and efficient work of the intelligent system, a proper database solution needs to be created and integrated into the system. The database should be

flexible so that the necessary changes can be easily incorporated into it, as well as fast to return the requested data in feasible time. Unfortunately, due to the project's time restrictions, the database development is out of scope for this study. However, a brief description of the data types that are to be stored in the database are provided in Figure 5.5. The functional requirements to the database are as follows:

- Receive and store data from various sources (user input data, files uploaded by a user, IMM process data);
- Store data in a structured way;
- Retrieve and provide data upon request from the other parts of the system.

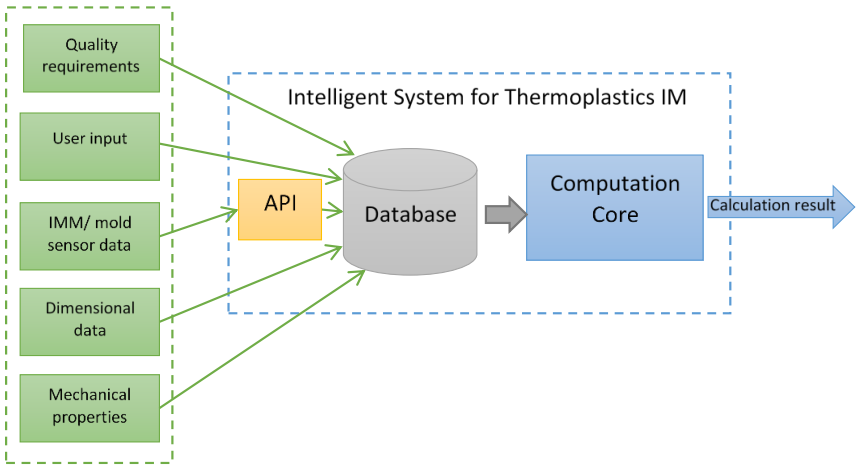


Figure 5.5. The main data types stored in the system's database

5.2.4 Layer 4. Description of the system's computation core

Development of the computation core is the main focus of this PhD work. This module is the one conducting all the necessary calculations to predict the quality of the produced injection molded parts, as well as performing the optimization of the relevant process parameters. The prediction module should be capable of predicting the dimensional properties of manufactured parts, their mechanical properties, as well as their combination using different datasets. The optimization module, on the other hand, should provide optimization of the IM process parameters to obtain the required quality of the parts. As a result, the following functional requirements can be identified for this module:

- CC1 – Predict dimensions of the produced parts based on the general datasets with the process parameter values;
- CC2 – Predict mechanical properties of the produced parts based on the general datasets with the process parameter values;
- CC3 – Conduct the model's backpropagation to identify the necessary process parameters based on the part quality requirements;

- CC4 – Predict dimensions of the produced parts based on the data series data of the process parameter values;
- CC5 – Predict mechanical properties of the produced parts based on the data series data of the process parameter values;
- CC6 – Predict dimensional deviations of the produced parts based on the general datasets with the process parameter values;
- CC7 – Conduct optimization of the process parameters to obtain the target parts quality.

Due to the time restrictions of the project, the prototypes of modules CC1, CC2, CC4 – CC6 were developed. To do this, procedures schematically visualized on Figures 5.6 and 5.7 were performed. The main stages of the procedures are described in more details in Chapter 4 and include data collection, data preprocessing and data processing. The procedure shown on Figure 5.6 was used for creation and processing of the general datasets, described in Section 4.4.1 and 4.4.3, while procedure on Figure 5.7 for dealing with data series datasets, presented in detail in Section 4.4.2. The results obtained during the modules prototyping are presented in Chapters 6-10.

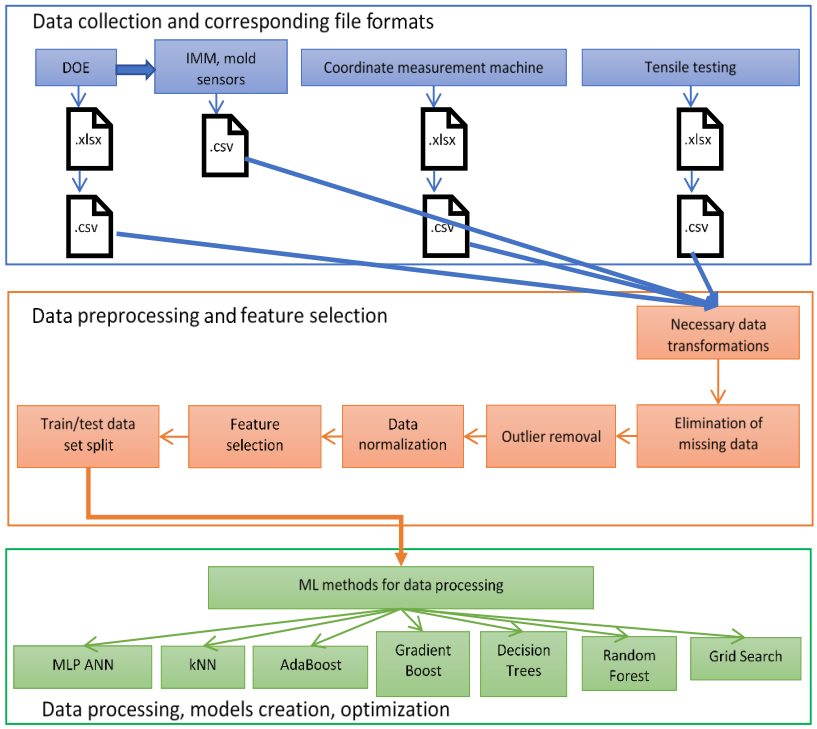


Figure 5.6. Procedure/ data pipeline used for data processing and models creation of the general datasets

5.3 Functional behavior analysis

The second stage of the model-based systems engineering process is necessary to describe the system's behavior based on the requirements defined in the previous step. System requirements form a basis for one or more behaviors, and one behavior corresponds to one or more requirements. In other words, this stage helps to understand and plan the logics of the system under development.

Within the proposed intelligent control system, the user is able to choose from two main options "work with data" (download, view, modify, upload, etc.) or "conduct the necessary calculations" (feature selection, model training, prediction, optimization, etc.). After selecting the option of interest, the user needs to confirm it and move to the next step.

On the next step the system either interacts with the data available in the database or acquired from the IMM(s) and mold sensors (selects the necessary data, uploads new data, contacts IMM to acquire requested by the user data, etc.) or makes chosen by user calculation (prediction of quality characteristics, optimization of the IM process parameters, model training, etc.). After this, the system shows result of its work to the user. If user wants to continue working in the system, he/ she confirms it and is then returned to the first page to make a new request. All the steps are repeated as many times as the user needs.

5.4 Architectural synthesis

After defining requirements to the system and understanding its logics, the third stage is carried out to plan and develop its physical structure. Here it is necessary to analyze if the proposed system is manufacturable, maintainable and supportable, as well as how to develop it so that it is such. It is important to remember that even though architecture might show elements of a system as separate entities, all of them need to be considered as a whole through adding connections between them.

Due to the time constraints of the presented PhD work, architectural synthesis has not been fully carried out. However, the Figures 5.2 – 5.7 are to be used as guidelines and a starting point for this step.

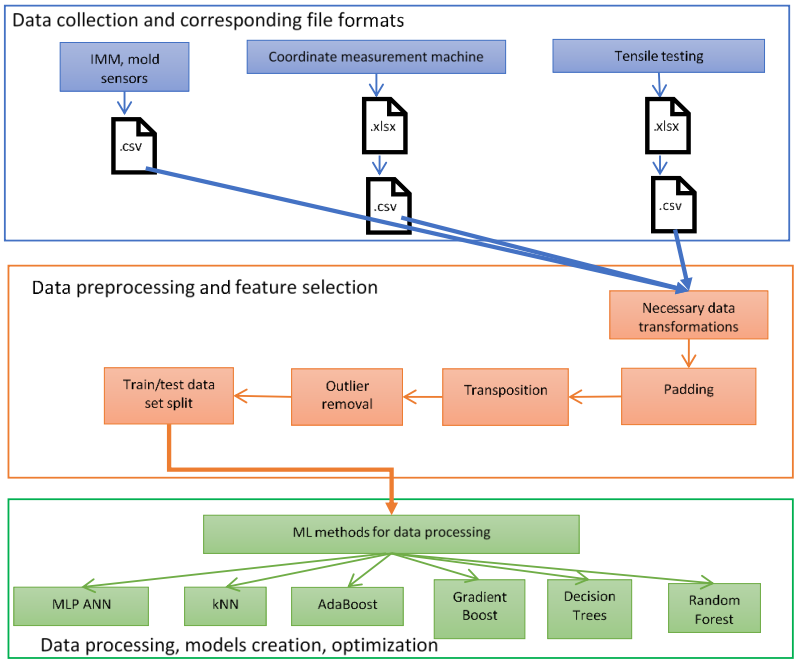


Figure 5.7. Procedure/ data pipeline used for data processing and models creation of the data series datasets

5.5 Validation and verification

Verification and validation are the last step of the MBSE approach. It is very important for successful project completion and needs to be considered on each step of design and development process. Verification is a process that ensures that the proposed system is correctly developed and complies with the necessary regulations and conditions. Validation, on the other hand, is used to define if the system satisfies the needs that have initiated the project [141].

In this work, visualization of the system's components and their connections helped to verify that all the necessary functional requirements are described for each part of the system. At the same time, validation was carried out through analysis of the MegaMould project description and the corresponding stakeholders' needs covered there.

Chapter 6

Module CC1 – Prediction of dimensions of the produced parts based on the general datasets

All models are wrong, but some are useful.

– George E. P. Box

This chapter describes development of the predictive models for dimensions of the produced parts using the general datasets described in Section 4.4.1. At first, the quality data used as the models' target value is presented. Secondly, data preprocessing and feature selection of the most important parameters to be included in the models are described. Finally, the predictive models obtained using different ML methods are presented and compared.

6.1 Data exploration

To be able to create predictive models of high quality, the data used for their training also needs to be of the corresponding quality. Therefore, different data preprocessing techniques are applied. However, before this step data exploration can be conducted in order to gain an increased understanding of the data at hand. Designs of experiments used in this work were created in a way that would cover as much of the parameters' values hyperspace as possible and result in production of parts with various quality characteristics.

At the same time, when manufacturing parts in the real industrial environment, the products need to have the best quality possible and comply with certain tolerance specifications. Therefore, the quality data obtained during the experiments is divided into different groups depending on the measured deviations of width and thickness of the dogbones. However, due to the chosen design of experiments the quality characteristics values might significantly vary. The groups are based on DIN 16742:2013 Plastic mouldings: tolerances and acceptance conditions standard [146] and are presented in Table 6.1.

Table 6.1. Measured deviation groups based on DIN 16742:2013 [146]

Measured deviation groups	3-6 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	±0.5	±1	±1.5

6.1.1 Width measurements

The nominal size of the width in the narrow section for the specimens produced in all four experiments (798 specimens in total) is 10 mm. The produced specimens were measured in three different points in the narrow section and the measured values were averaged to obtain the specimen's width. The used designs of experiments resulted in the

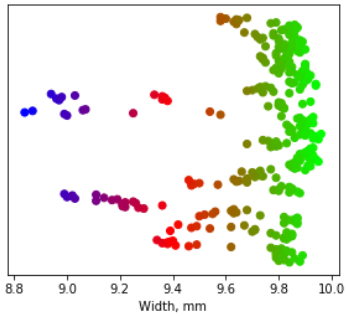
width values distribution shown on Figure 6.1 (a)-(d) for experiments 1-4 correspondingly. Table 6.2 shows the number of specimens from each of the experiments that belong to different measured deviation groups based on DIN 16742:2013 standard. In case of experiments 1, 3 and 4 more than half of the specimens belong to fine and medium groups despite variation of the parameters in the DOE. However, a significant number of the specimens do not meet requirements for the fine group. All the specimens in experiment 2, on the other hand, belong to very coarse group. One of the main reasons for this is the part's thickness of 15 mm in comparison to 4 mm thickness in experiments 1, 3-4. According to some studies the specimen's walls thickness and width significantly affect the parts shrinkage, the thicker is a part's wall – the bigger is shrinkage [147]. In addition, for semi-crystalline materials, such as HDPE, the anisotropy effects are larger and vary along the flow path [148].
Table 6.3 contains the maximum, minimum, average and standard deviation values of the width dimension measurements obtained during the experiments. The specimens produced in experiments 1, 3-4 have a significantly higher maximum value in comparison to the dogbones from the experiment 2, one of the reasons for this is the parts' thickness of 4 and 15 mm in the different experiments. Experiment 4 also has the highest minimum value, this might be due to the difference in the used material, as the virgin HDPE was used in the first two experiments, while two different types of recycled HDPE were used in experiments 3 and 4. The highest average is obtained in experiment 4, while the lowest one in experiment 2. Experiment 2, on the other hand, has the lowest variation, while the highest one is in experiment 3.

Table 6.2. Number of specimens in measured deviation groups based on the width measurements

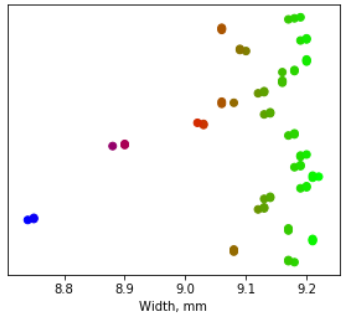
	Experiment 1	**Experiment 2**	**Experiment 3**	**Experiment 4**
Fine	49	0	63	114
Medium	129	0	57	48
Coarse	89	0	31	12
Very coarse	53	72	51	30

Table 6.3. Maximum and minimum values of width dimension, experiments 1-4

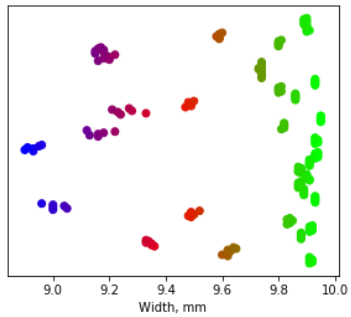
	Experiment 1	**Experiment 2**	**Experiment 3**	**Experiment 4**
Minimum width, mm	8.84	8.74	8.9	9.14
Maximum width, mm	9.96	9.22	9.95	9.98
Average, mm	9.71	9.12	9.68	9.81
Standard deviation, mm	0.25	0.11	0.3	0.22



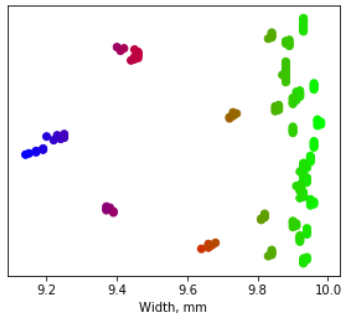
(a) Experiment 1



(b) Experiment 2



(c) Experiment 3



(d) Experiment 4

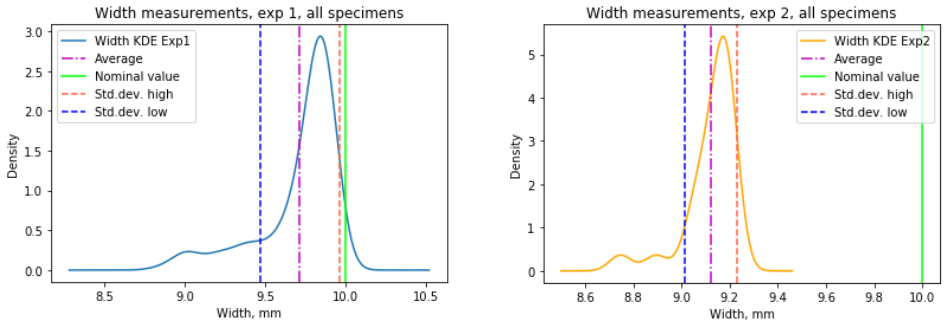
Figure 6.1. Distribution of measured width values for experiments 1-4

To further visualize the data shown in Table 6.3, Figure 6.2 depicts the kernel density estimation distributions for all four experiments and the corresponding measurements, where width's nominal value, average and standard deviation are also presented.

Figure 6.3 depicts distributions based on kernel density estimation (KDE) for each experiment in general and for different measured deviation groups. As it can be seen from Figure 6.3 (e), the width values within the very coarse group for experiment 2 have a smaller range than those from experiments 1, 3-4 within the same group. Figure 6.4 demonstrates KDE for specimens 1 and 2 separately for experiments 1, 3 and 4 only, as just one specimen was produced per run in the experiment 2.

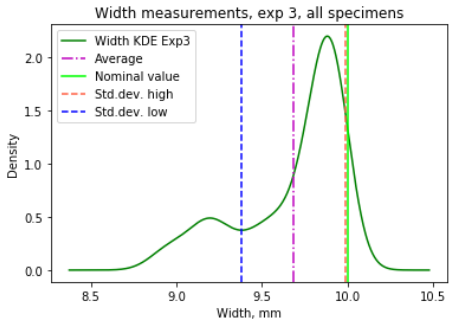
As it can be seen from Figure 6.3 (a) the width measurements follow left-skewed distributions that have one "major" and one "minor" peaks for experiments 1, 3 and 4. The experiment 2 distribution has two "minor" peaks and one "major". The presence of more than one peak signals about existence of two or more (in case of experiment 2) groups of specimens with different modes. In addition, it shows that the presented distribution might be a sum of several distributions or processes. The "minor" peaks for all four experiments are, however, significantly smaller than the "major" ones.

When we further observe Figure 6.3 (b)-(e) and Figure 6.4 the distribution for the width measurements for experiment 3 for fine, medium and coarse groups is bimodal for specimens 1 and 2. The values from experiment 4 for both specimens show bimodal behavior for medium, coarse and very coarse groups, at the same time, distributions are closer to normal for the fine measured deviation group. When it comes to experiment 1, the values for medium group for both specimens and for very coarse group for the specimen 2 are bimodal.

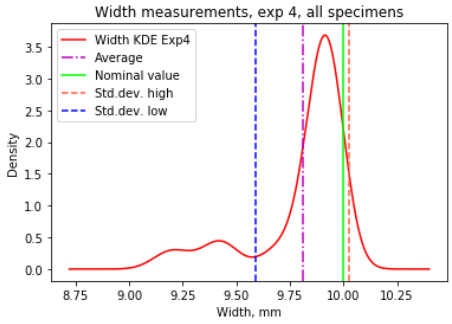


(a) Experiment 1

(b) Experiment 2



(c) Experiment 3

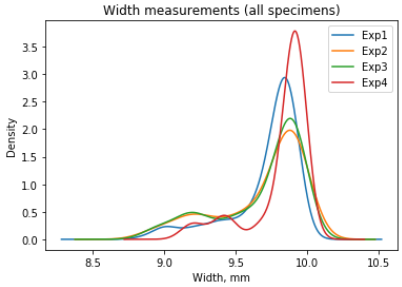


(d) Experiment 4

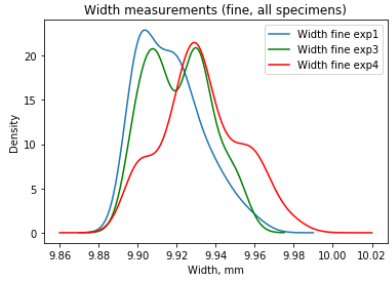
Figure 6.2. Kernel Density Estimation for width measurements, experiments 1-4 with mean and standard deviation

Some of these variations appear due to the used design of experiments that varied different production parameters. For example, it has been reported that low mold temperature reduces shrinkage and affects it in directions transversal and longitudinal to flow [147, 149]. Higher injection pressure is stated to decrease shrinkage [150], higher injection speed, at the same time increases it [151]. In addition, barrel temperature controls the melt temperature and, therefore, influences the replication ability of the process [152]. Based on the obtained experimental data, as shown on Figure 6.5 the higher holding pressure results in the higher values of the produced specimen's width. Figures

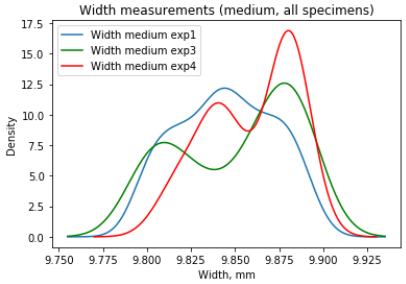
that show relationships between the rest of the varied DOE parameters and the specimen's width can be found in Appendix C, however, they do not seem to provide information as easy to interpret as in case of the holding pressure parameter.



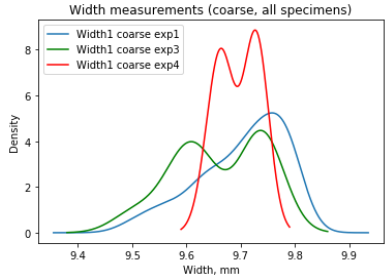
(a) KDE for all width measurements



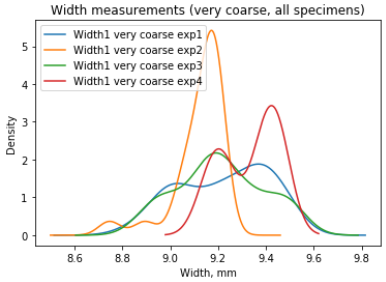
(b) KDE for fine width measurements



(c) KDE for medium width measurements



(d) KDE for coarse width measurements



(e) KDE for very coarse width measurements

Figure 6.3. Kernel Density Estimation distributions for all the width measurements, different measured deviation groups based on DIN 17642:2013

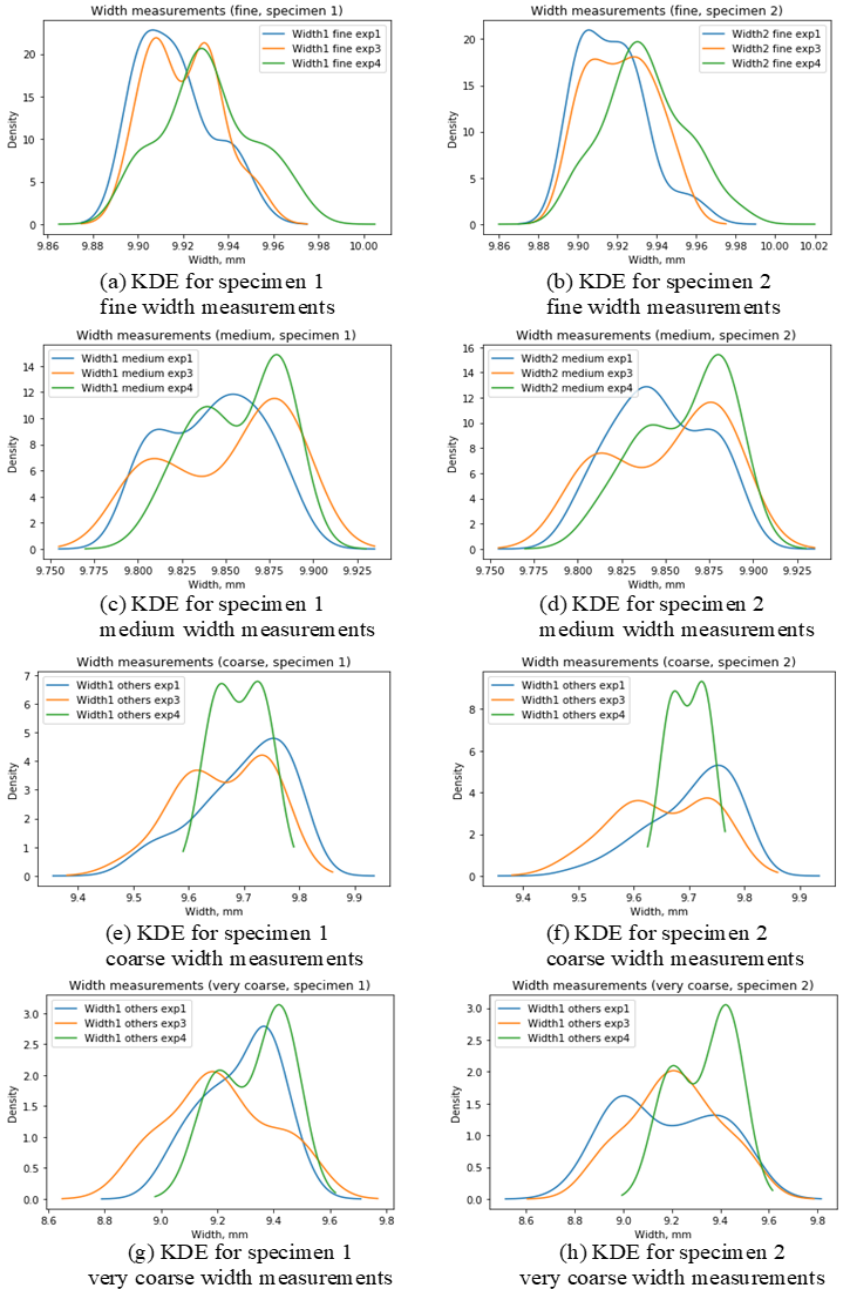
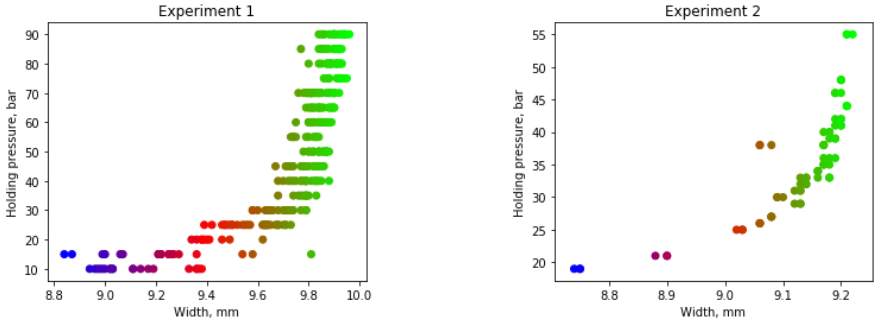


Figure 6.4. Kernel Density Estimation distributions for width measurements of the specimens 1 and 2 separately, different measured deviation groups based on DIN 17642:2013

In addition, the properties of used materials play an important role, as virgin and two types of recycled HDPE were used in different experiments. The produced part's wall thickness is also an important factor, as the thicker are the walls, the bigger shrinkage occurs [147]. Therefore, the parts with 15 mm thickness produced in experiment 2 are outside of fine and medium measured deviation groups. When it comes to the differences between the specimens 1 and 2 produced during the same machine runs, the reasons for their occurrence can be uneven temperature distribution of the mold surface and the semi-crystalline nature of HDPE that leads to uneven shrinkage along and across the plastic melt flow.



(a) Experiment 1 (b) Experiment 2

(c) Experiment 3 (d) Experiment 4

Figure 6.5. Relationship between the specimen's width and holding pressure DOE parameter

6.1.2 Thickness measurements

The nominal thickness size of produced dogbone parts is 4 mm for experiments 1, 3-4 and 15 mm for experiment 2. Similarly to the width values, the thickness of the specimens was measured in three different points spread along the narrow section of the produced part and averaged to obtain the final thickness value. Table 6.4 shows the number of specimens that belong to different measured deviation groups, while Table 6.5 includes range for the obtained thickness measurements for each of the experiments, as well as average and standard deviation values. As it is shown in Table 6.4, no specimens in all

four experiments fall into the fine and medium groups. For experiment 2 all the dogbones belong to the coarse group, while for experiment 3 to very coarse. Experiment 4 has most of the specimens in the very coarse measured deviation group with only 6 in the coarse group. According to Amaranan and Manonukul [153] such behavior might be result of significant difference between the thickness and width nominal values. Figure 6.6 depicts distribution of the obtained thickness measurements for experiments 1-4.

Table 6.4. Number of specimens in measured deviation groups based on the thickness measurements

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Fine	0	0	0	0
Medium	0	0	0	0
Coarse	84	72	0	6
Very coarse	236	0	202	198

Table 6.5. Maximum and minimum values of thickness dimension, experiments 1-4

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Minimum thickness, mm	3.3	14.61	3.25	3.31
Maximum thickness, mm	3.84	14.79	3.77	3.8
Average, mm	3.71	14.66	3.6	3.66
Standard deviation, mm	0.13	0.03	0.11	0.12

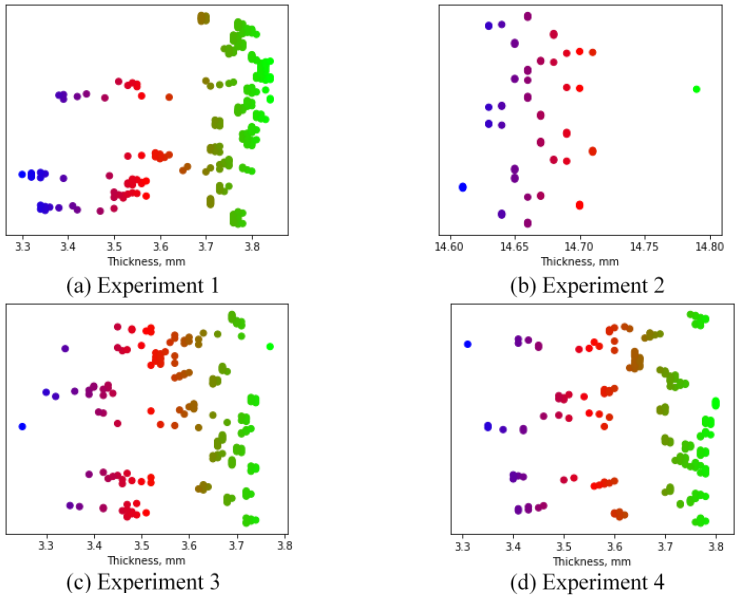
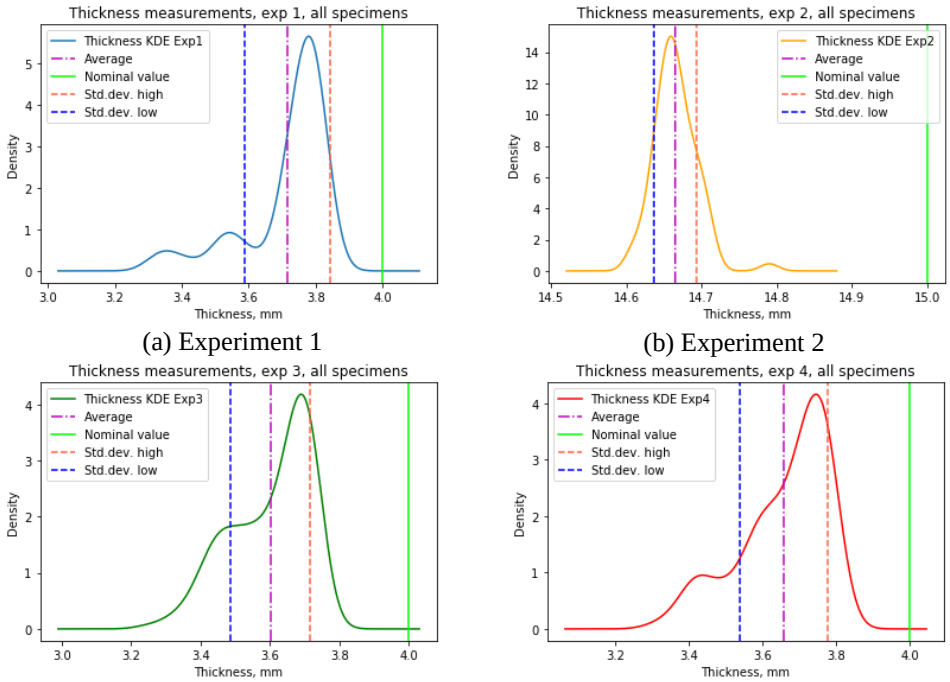


Figure 6.6. Distribution of measured thickness values for experiments 1-4

To increase understanding of the data presented in Table 6.5, Figure 6.7 depicts the kernel density estimation distributions for all four experiments and the corresponding measurements, where thickness' nominal value, average and standard deviation are also shown.



(a) Experiment 1 (b) Experiment 2

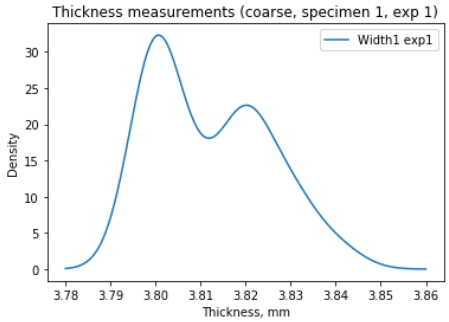
(c) Experiment 3 (d) Experiment 4

Figure 6.7. Kernel Density Estimation for thickness measurements, experiments 1-4 with mean and standard deviation

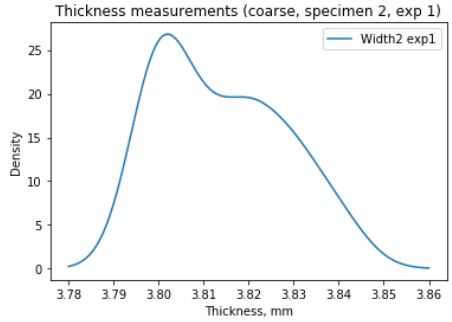
Figure 6.8 shows the corresponding KDEs for coarse and very coarse groups for specimens number 1 and 2 for experiments 1, 3-4. The experiment 2 data is not plotted on this figure, as all the width measurements from this experiment belong to the same group and their KDE can be seen on Figure 6.7 (b). Figures 6.8 (a)-(b) depict only experiment 1 data, as there are no width measurements that belong to this group in experiment 3, while experiment 4 has only 3 measurements per specimen 1 and 2 in this group, which is not enough to create a meaningful KDE.

As it can be seen from Figure 6.8, the measurements for experiments 1 and 4 have 2 "minor" and one "major" peak. For experiment 2 only one "major" peak is distinguishable, while for experiment 3 one "minor" and one "major" peaks are visible. Similarly to the width measurements this indicates possibility of having separate groups of produced specimens with different modes and presence of several separate distributions or processes. The differences between the specimens 1 and 2 for experiment 1, 3 and 4 are rather small for the very coarse measured deviation group, while for

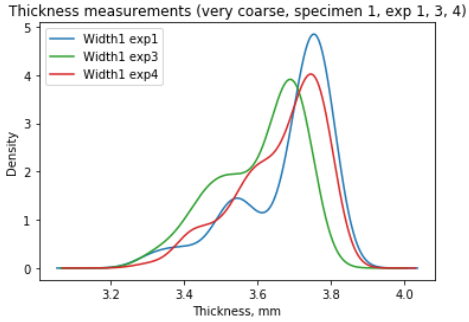
experiment 1 and the coarse group they differ more. The reasons for these observations include material differences, crystallization features of the semi-crystalline materials, as well as variations of the parameters included into the corresponding DOEs. Relationship between the thickness and the holding pressure DOE parameter is shown on Figure 6.9, behavior similar to the one on Figure 6.5 for the specimen's width is visible: the higher holding pressure is used, the higher thickness is obtained. Figures showing relationships between the rest of the DOE parameters and the thickness can be seen in Appendix C.



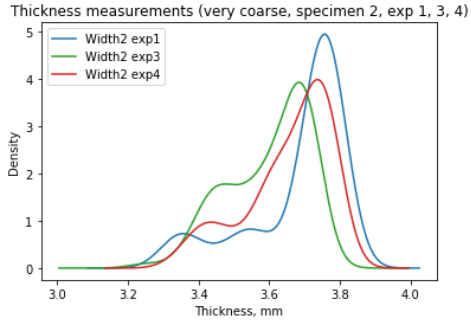
(a) KDE for specimen 1 coarse thickness measurements (experiment 1)



(b) KDE for specimen 2 coarse thickness measurements (experiment 1)

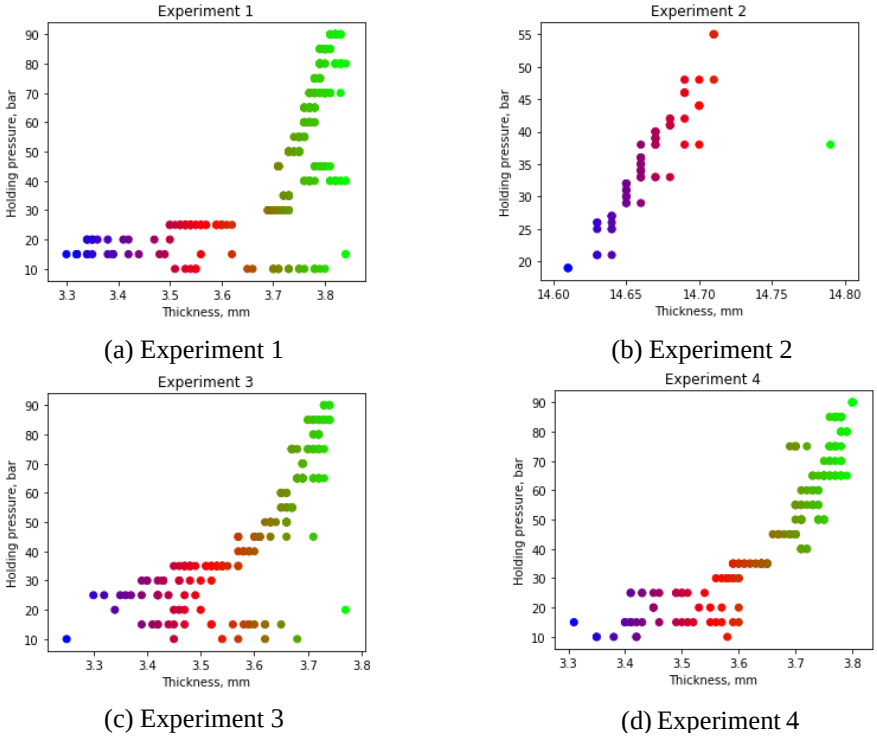


(c) KDE for specimen 1 very coarse thickness measurements (experiments 1, 3, 4)



(d) KDE for specimen 2 very coarse thickness measurements (experiments 1, 3, 4)

Figure 6.8. Kernel Density Estimation distributions for the thickness measurements, experiments 1-4



(a) Experiment 1

(b) Experiment 2

(c) Experiment 3

(d) Experiment 4

Figure 6.9. Relationship between the specimen's thickness and holding pressure DOE parameter

6.2 Data preprocessing

After conducting data exploration of the width and thickness measurements, the experimental work continued with the data preprocessing step. This step included analysis of the obtained process parameters and removal of missing data, parameters irrelevant for the further steps or the ones that did not change their values throughout the experimental runs.

The samples of the first run of DOE combinations number 8, 9, 10 and 13 from experiment 4 were removed, as there was no parameter data logged during them due to a data acquisition system error. Tables 6.6 and 6.7 include the names of "irrelevant" and constant parameters that were removed from the datasets. Explanation of the meaning of the parameters present in the tables can be found in Appendix B. The "irrelevant" parameters are such for the predictive models of interest, however, could be potentially useful when developing a model for the IMM maintenance schedule, for example, machine date, machine time and shotcounter. The constant parameters could also have been eliminated during the feature selection process, however, it was of interest to identify which of the logged parameters had constant values to know that they did not vary due to the DOE design or the general machine settings. The "x" symbol in each of the columns

in Tables 6.6 and 6.7 indicates that the specified parameter values were logged during the experiment, if the symbol is absent – this parameter is not present in the corresponding experiment dataset.

Table 6.6. Irrelevant parameters that were removed from the data for experiments 1-4

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Good parts	x	x	x	x
Machine_date	x	x	x	x
Machine_time	x	x	x	x
Machine_time_copy	x			
Bad parts	x	x	x	x
Shotcounter	x	x	x	x
Parts_count	x	x	x	x
timestamp_imm_machine_last			x	x

Table 6.7. Parameters with constant values throughout the experimental runs

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Injection_time_set_max	x	x		
Injection_time_set_min	x	x		
Ejector_pos_set_max	x			
Decomp_after_plast_vol	x		x	x
Waiting_del	x			
Shot_vol	x	x	x	x
Cushion_smallest_set_max	x	x		
Switchover_time	x	x		
Ejector_pos_set_min	x			
Plastic_time_set_max	x	x		
Plastic_time_set_min	x	x		
Cushion_ideal	x	x	x	x
Flow_n	x	x	x	x
Plastic_delay_time_set	x			
Inject_pres_limit	x			
Current_station	x		x	x
Heating_cyl1_z1_set		x		
Max_injection_speed_from_graph		x	x	x
Clamp_force_relief_time		x		
Flow_number_set		x		
Flow_number_act		x		
Hold_pressure_correction		x		
Specific_pressure_switchover_set		x		
Shot_volume_end_corrected		x		
Heating_cyl1_z3_set		x		
Fixed_plate_tool1_temp_z9		x		
Fixed_plate_tool2_temp_z10		x		
Core_movement			x	x
Demoduling_time			x	x
Start_flow_number_meas_trigger			x	x
Stop_flow_number_meas_trigger			x	x
Purging_time			x	x
Decompres_end			x	x
Flush_time			x	x

There were no outliers identified for the width and thickness data, therefore, no samples were removed due to this during this step. Later feature selection was conducted separately for width and thickness as target variables, it is described in more details in Section 6.3. As the last data preprocessing step, the data was normalized using z-score normalization technique.

6.3 Feature selection

Research related to the parameters that influence the injection molded parts quality is ongoing, as depending on the properties of material in use, part's geometry and other environmental factors, they might vary [79]. In addition, when developing machine learning models such issues as dimensionality reduction, overfitting and long training time need to be addressed. There are various approaches that can be used in order to do this, and one of them is feature selection. Selection of the most influential parameters or features can assist the injection molding practitioners to identify which process parameters to pay attention to, as well as to create better predictive models with shorter training time through elimination of irrelevant and redundant features.

This section presents how the different feature selection methods described in Chapter 3, Section 3.3 were tested with the obtained experimental data in order to select features to be included in the width and thickness prediction models.

6.3.1 Width target variable

Separate experiments datasets

To select parameters/ features that are the most relevant for predicting value of the width target variable, five different feature selection methods are compared based on the data from experiment 1. Pearson's correlation shows the level of linear correlation, while Spearman's represents the non-linear correlation. RReliefF assesses how well a feature can distinguish between the dataset instances whose values are close to each other. Correlation-based feature selection or CFS looks for features subsets that have high correlation with the predicted variable and no or low correlation with each other. Recursive features elimination or RFE recursively removes the features and builds a model including the remaining ones, the model used in this case is linear regression. The methods are described in more details in Section 3.3.

The reason for trying different feature selection methods is that there is no "rule of thumb" for choosing the most suitable feature selection method for a concrete problem. Therefore, the only way of selecting a method that fits best is to try different methods and compare the results. The selected methods use different measures to assess the features and can result in giving high scores to features that significantly vary from method to method. Tables 6.8, 6.9 and 6.10 show the scores, normalized between 0 and 1, that different feature selection methods give to the evaluated features, the "Mean" column includes the averaged methods' scores, the target variable in Table 6.8 is width1, in Table 6.9 – width2, while in Table 6.10 – width from the sequential dataset that includes both width1 and 2 in a sequence. **Width1 target variable** is the width of the first specimen produced during experiments 1, 3-4, while **width2 focus variable** is the width of the second specimen. **Width target variable** without a number includes width1 and width2 in a sequence, where the parameter values are repeated if the specimens were produced during the same run. The features are sorted from high to low based on the mean score value. To assess the methods' performance a multilayer perceptron model with the following hyperparameters was trained using scikit-learn Python library:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic',
alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = 5,
random_state = 2, momentum = 0.2);
```

To train the MLP model, the experiment 1 dataset was divided into 70% training and 30% testing subsets. The model was then trained on the training dataset using 5-folds cross-validation and tested on the testing one. Random state for dataset partitioning was set to 1. The MLP models were not tuned in this case and only the above specified model hyperparameter values were utilized.

For each of the different feature selection methods the features with scores higher than 0.2 were used to train the corresponding MLP models. The only exception is the "Cooling_time" parameter, which has a relatively high score, but has not been included in the models, as "Cooling_time" (set value) and "Cooling_time_last" (actual value) are correlated between each other. It was decided to include the actual value rather than the set one. Tables 6.8, 6.9 and 6.10 include the coefficient of determination R^2 score, RMSE for the trained models and correlation coefficient value for the real and predicted target variable value. An MLP model trained on the full feature set for the parallel dataset and width1 target variable (width of the first out of two specimens produced simultaneously) has $R^2 = 0.41$, RMSE = 0.13 and correlation coefficient of 0.81.

Table 6.8 shows that the same parameters can have high and low scores, depending on the feature selection method in use. The scoring makes sense from the injection molding process point of view, as parameters related to the cushion value, holding pressure, holding pressure time and cooling time receive high scores. Plasticizing time, closing force, clamping force at switchover, maximum speed, screw speed, last ejector position and switchover volume are some of the parameters with the lowest scores from all the FS methods for width1. It can be seen that application of all the feature selection methods except for CFS increase the MLP model quality, in case of RRelifF from R^2 equal to 0.41 to 0.84, decrease the RMSE from 0.13 to 0.007 and increase the correlation coefficient from 0.81 to 0.92.

Table 6.8. Feature selection for experiment 1 width1 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Cushion_after_hold_pres	1	0.17	0.98	1	0.67	0.76
2.	Holding_pressure	1	0.67	1	0	0.49	0.63
3.	Cushion_average	0.78	0.25	0.9	0	0.56	0.50
4.	Cushion_smallest	0.84	0.21	0.92	0	0.48	0.49
5.	Injection_work	0.24	0.71	0.38	1	0.06	0.48
6.	Screw_speed_max	0.15	0.79	0.13	0	0.85	0.38
7.	Spec_pres_switchov	0.13	0.96	0.22	0	0.58	0.38
8.	Cooling_time_last	0.27	0.54	0.36	0	0.72	0.38
9.	Cooling_time	0.31	0.5	0.39	0	0.60	0.36
10.	Injection_Speed	0.22	0.92	0.22	0	0.42	0.36
11.	Holding_pres_time	0.04	0.88	0.08	0	0.73	0.35
12.	Tool_Temperature	0.33	0.75	0.13	0	0.35	0.31

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
13.	Flow_no_plast	0	0.42	0.09	0	1.00	0.30
14.	Backpressure	0.09	0.83	0	0	0.56	0.30
15.	Last_cycle_time	0.05	1	0.18	0	0.10	0.27
16.	Heating_cyl1_z1_set	0.18	0.33	0.14	0	0.65	0.26
17.	Nozzle_tempr_z2_average	0.18	0.29	0.12	0	0.65	0.25
18.	Injection_time	0.27	0.46	0.34	0	0.15	0.24
19.	Plast_time	0.01	0.37	0.17	0	0.50	0.21
20.	Closing_force	0.09	0.58	0.2	0	0.00	0.17
21.	Clamp_force_switchov	0.04	0.63	0.1	0	0.00	0.15
22.	Speed_max	0.14	0.04	0.13	0	0.33	0.13
23.	Screw_speed	0.18	0	0.18	0	0.27	0.13
24.	Ejector_pos_last	0.07	0.13	0.15	0	0.15	0.10
25.	Switchov_vol	0	0.08	0.02	0	0.00	0.02
	MLP R^2 score	0.77	0.78	0.77	0.35	0.84	
	MLP RMSE	0.085	0.088	0.085	0.141	0.07	
	Correlation coef.	0.89	0.9	0.89	0.79	0.92	

Table 6.9 confirms observations drawn from the Table 6.8. Here, the same parallel dataset for experiment 1 has been used, however, width2 was set as a target variable (width of the second out of two specimens produced simultaneously). A model with no feature selection has $R^2 = 0.8$, RMSE = 0.102 and correlation coefficient of 0.92. Similar set of features receives the highest and the lowest scores in this case, which aligns well with having width of two specimens produced during the same IMM run as focus variables. Here use of the feature selection increases the model's R^2 from 0.8 with no feature selection to 0.9 using the RReliefF and Spearman methods, as well as decreases the RMSE value and increases the correlation between the real and predicted width2 values.

Table 6.9. Feature selection for experiment 1 width2 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Cushion_after_hold_pres	0.98	0.25	0.98	1	1.00	0.84
2.	Holding_pressure	1	0.5	1	0	0.89	0.68
3.	Cushion_smallest	0.86	0.29	0.92	0	0.54	0.52
4.	Cushion_average	0.81	0.33	0.9	0	0.48	0.50
5.	Injection_work	0.26	0.46	0.29	1	0.06	0.41
6.	Cooling_time	0.27	0.83	0.38	0	0.21	0.34
7.	Cooling_time_last	0.24	0.79	0.32	0	0.26	0.32
8.	Tool_Temperature	0.27	0.88	0.11	0	0.29	0.31
9.	Injection_Speed	0.09	0.75	0.14	0	0.43	0.28
10.	Spec_pres_switchov	0.02	0.92	0.09	0	0.35	0.28
11.	Screw_speed_max	0.07	0.96	0.08	0	0.17	0.26
12.	Flow_no_plast	0.16	0.67	0.04	0	0.35	0.24
13.	Holding_pres_time	0.07	0.71	0.01	0	0.43	0.24

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
14.	Last_cycle_time	0.05	1	0.13	0	0.00	0.24
15.	Heating_cyl1_z1_set	0.3	0.17	0	0	0.53	0.20
16.	Backpressure	0.09	0.54	0.11	0	0.25	0.20
17.	Nozzle_tempr_z2_average	0.3	0.13	0	0	0.53	0.19
18.	Injection_time	0.19	0.37	0.35	0	0.00	0.18
19.	Plast_time	0.06	0.42	0.12	0	0.23	0.17
20.	Closing_force	0.05	0.58	0.1	0	0.00	0.15
21.	Clamp_force_switchov	0.01	0.63	0.05	0	0.00	0.14
22.	Ejector_pos_last	0.14	0.21	0.02	0	0.10	0.09
23.	Screw_speed	0	0.04	0.13	0	0.10	0.05
24.	Speed_max	0.03	0	0.1	0	0.08	0.04
25.	Switchov_vol	0.03	0.08	0.04	0	0.00	0.03
	MLP R^2 score	0.85	0.82	0.9	0.57	0.9	
	MLP RMSE	0.078	0.094	0.058	0.149	0.067	
	Correlation coef.	0.94	0.93	0.95	0.83	0.96	

Similarly to the parallel dataset, application of feature selection methods for the sequential dataset with width target variable (width1 and width2 variables merged together sequentially) shown in Table 6.10 allows to increase the MLP model performance by increasing its R^2 from 0.77 with no feature selection to 0.84 with Spearman correlation measure and decreasing the RMSE from 0.106 to 0.093, while increasing the initial correlation coefficient from 0.91 to 0.93. The set of the features that get the scores higher than 0.2 is similar to the one for width1 and width2 target variables even though some of the features have higher scores in comparison to their scores for the parallel dataset and some have lower.

Table 6.10. Feature selection for experiment 1 width target variable, sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Holding_pressure	1	0.5	1	1	0.36	0.77
2.	Cushion_after_hold_pres	0.99	0.13	0.98	1	0.51	0.72
3.	Cushion_average	0.79	0.29	0.9	0	0.45	0.49
4.	Cushion_smallest	0.85	0.17	0.92	0	0.38	0.46
5.	Injection_work	0.24	0.46	0.31	1	0.01	0.40
6.	Screw_speed_max	0.01	0.92	0.08	0	1.00	0.40
7.	Cooling_time_last	0.24	0.79	0.32	0	0.65	0.40
8.	Cooling_time	0.27	0.75	0.37	0	0.57	0.39
9.	Spec_pres_switchov	0.03	0.96	0.13	0	0.54	0.33
10.	Ejector_pos_last	0.09	0.33	0.06	1	0.16	0.33
11.	Injection Speed	0.13	0.83	0.15	0	0.50	0.32
12.	Tool Temperature	0.29	0.88	0.09	0	0.32	0.32
13.	Flow_no_plast	0.08	0.42	0.04	0	0.83	0.27
14.	Injection_time	0.22	0.67	0.32	0	0.11	0.26

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
15.	Holding_pres_time	0.04	0.71	0.02	0	0.55	0.26
16.	Last_cycle_time	0.03	1	0.13	0	0.00	0.23
17.	Backpressure	0.07	0.54	0.02	0	0.47	0.22
18.	Nozzle_tempr_z2_average	0.23	0.25	0.02	0	0.48	0.20
19.	Heating_cyl1_z1_set	0.23	0.21	0.04	0	0.48	0.19
20.	Plast_time	0.01	0.37	0.12	0	0.36	0.17
21.	Closing_force	0.05	0.58	0.12	0	0.00	0.15
22.	Clamp_force_switchov	0.01	0.63	0.05	0	0.00	0.14
23.	Screw_speed	0.06	0.04	0.13	0	0.20	0.09
24.	Speed_max	0.02	0	0.09	0	0.24	0.07
25.	Switchov_vol	0	0.08	0	0	0.00	0.02
	MLP R^2 score	0.83	0.8	0.84	0.82	0.81	
	MLP RMSE	0.095	0.1	0.093	0.1	0.097	
	Correlation coef.	0.93	0.92	0.93	0.92	0.92	

Table 6.11 summarizes the MLP scores obtained with no FS and with application of the 5 chosen FS methods. The average method's score is used to choose the method that performs best on the available set of data. Some of the ML methods that are used to create the predictive models in this study do not require assistance of the FS methods, as they have "built-in" feature selection, these methods are Decision Trees, Gradient Boosting, AdaBoost and Random Forest. MLP and kNN, on the other hand, need feature selection to be performed before training the models using these methods, to filter out redundant and irrelevant features that might be acting as noise. Based on the results depicted in Table 6.11, RReliefF is the method that will be used to select the attributes for the width target variable for all the experiments. Table with the corresponding RReliefF scores for parallel and sequential datasets for experiments 1-4 with width target variable can be found in Appendix D.

Table 6.11. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (width target variable)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Experiment 1, parallel dataset, width1 target variable	MLP R^2 score	0.41	0.77	0.78	0.77	0.35	0.84
	MLP RMSE	0.13	0.085	0.088	0.085	0.141	0.07
	Correlation coef.	0.81	0.89	0.9	0.89	0.79	0.92
Experiment 1, parallel dataset, width2 target variable	MLP R2 score	0.8	0.85	0.82	0.9	0.57	0.9
	MLP RMSE	0.102	0.078	0.094	0.058	0.149	0.067
	Correlation coef.	0.92	0.94	0.93	0.95	0.83	0.96
Experiment 1, sequential dataset, width target variable	MLP R2 score	0.77	0.83	0.8	0.84	0.82	0.81
	MLP RMSE	0.106	0.095	0.1	0.093	0.1	0.097
	Correlation coef.	0.91	0.93	0.92	0.93	0.92	0.92
Average	MLP R^2 score, average	0.66	0.82	0.80	0.84	0.58	**0.85**

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
	MLP RMSE average	0.11	0.09	0.09	0.08	0.13	**0.08**
	Correlation coef. average	0.88	0.92	0.92	0.92	0.85	**0.93**

Joined dataset

In addition to analyzing data from all the experiments separately, parallel and sequential joined datasets are also created. The parallel dataset includes data from experiment 1, 3 and 4, as only these three experiments had two specimens produced per run. The sequential dataset includes data from all four experiments, where dimensional properties such as width1 and width2 (width of the first and the second specimens produced simultaneously) are stored in the width target variable. Both parallel and sequential joined datasets include only those parameters that were logged during all the experiments in the joined dataset. In addition, a material parameter was added, that includes a code for virgin HDPE, ContainerService and RePro recycled HDPE. In the sequential dataset a part type column has also been added, that indicates 1 for a 4 mm thick dogbone and 2 for the 15 mm dogbone part. More information on how the joined datasets were created can be found in Section 4.4.3. Similarly to the separate experiments datasets, five feature selection methods were used and compared. To select the feature selection method with the best performance, an MLP model was trained with the following parameters:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic', alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = (X.shape[1]+1)/2, random_state = 2, momentum = 0.2),
```

in this case `X.shape[1]` is a number of parameters considered in the model. Only parameters with the FS method scores higher than 0.2 were included into the models. To ensure that there is no overfitting, the datasets were divided into 70% training and 30% testing subsets and 5-folds cross-validation was used when training the model on the training subset. Random state for dataset partitioning was set to 0. The MLP models were not tuned in this step. Tables 6.12, 6.13 and 6.14 show the FS scores for the different parameters for width1, width2 and width target variables and the corresponding MLP models quality measures (R^2, RMSE and correlation coefficient).

The parallel joined dataset includes 20 parameters, while the sequential one – 19, which is significantly less than the full separate experiments datasets. For the parallel joined dataset and width1 as a focus variable (width of the first out of two specimens produced simultaneously) and no feature selection used, the R^2 equals to 0.86, RMSE = 0.084, correlation coefficient = 0.94. Application of the RReliefF method can increase the R^2 score to 0.87, decrease the RMSE to 0.08, while the correlation coefficient value stays the same. It is also possible to notice that similarly to the separate experiments datasets, parameters related to cushion and holding pressure receive the highest scores, while plasticizing time, screw speed and the last ejector position get the lowest ones. Material parameter gets a score higher than 0.2 from RFE, Spearman and RReliefF FS methods, meaning that it contains certain amount of information that can be useful for creation of the regression models of interest. Once again, the "Cooling_time" parameter was

excluded from the MLP model, as it is a set value of the "Cooling_time_last" (actual value) parameter.

Table 6.12. Feature selection for width1 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Cushion_after_hold_pres	1	0.16	1all	1	1.00	0.83
2.	Holding_pressure	0.95	0.79	0.99	0	0.79	0.70
3.	Cushion_smallest	0.83	0.21	0.94	0	0.39	0.47
4.	Switchov_vol	0.14	0.05	0.03	1	0.45	0.33
5.	Heating_cyl1_z1_set	0.02	0.74	0.12	0	0.74	0.32
6.	Cooling_time	0.29	0.63	0.4	0	0.29	0.32
7.	Cooling_time_last	0.25	0.68	0.35	0	0.26	0.31
8.	Speed_max	0.13	0.11	0.04	1	0.24	0.30
9.	Holding_pres_time	0.08	0.89	0.23	0	0.26	0.29
10.	Backpressure	0.15	0.84	0.23	0	0.24	0.29
11.	Spec_pres_switchov	0.11	0.95	0.12	0	0.13	0.26
12.	Material	0.14	0.32	0.38	0	0.32	0.23
13.	Injection_Speed	0.15	0.58	0.18	0	0.19	0.22
14.	Last_cycle_time	0.07	1	0	0	0.00	0.21
15.	Flow_no_plast	0.21	0.53	0.12	0	0.17	0.21
16.	Injection_time	0.26	0.47	0.22	0	0.01	0.19
17.	Injection_work	0.21	0.42	0.3	0	0.00	0.19
18.	Plast_time	0.1	0.37	0.03	0	0.13	0.13
19.	Screw_speed	0.11	0.26	0.02	0	0.22	0.12
20.	Ejector_pos_last	0	0	0.12	0	0.05	0.03
	MLP R^2 score	0.8	0.84	0.81	0.8	0.87	
	MLP RMSE	0.11	0.09	0.1	0.11	0.08	
	Correlation coef.	0.91	0.92	0.91	0.91	0.94	

For the width2 focus variable (width of the second out of two specimens produced simultaneously) and no feature selection method, R^2 is 0.92, RMSE = 0.069, correlation coefficient = 0.97. RReliefF method allows to increase the R^2 score to 0.94, lower RMSE to 0.061and increase the correlation coefficient to 0.99. It is seen that the set of parameters receiving the highest and the lowest scores for width2 is similar to that for width1. The material parameter once again receives score's higher than 0.2 from RFE, Spearman and RReliefF.

Table 6.13. Feature selection for width2 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Holding_pressure	0.97	0.53	0.99	0	0.88	0.67
2.	Cushion_after_hold_pres	1	0.16	1	1	1.00	0.83
3.	Cushion_smallest	0.84	0.21	0.95	0	0.44	0.49
4.	Cooling_time	0.28	0.79	0.4	0	0.37	0.37

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
5.	Cooling_time_last	0.24	0.84	0.34	0	0.31	0.35
6.	Injection_Speed	0.08	0.95	0.14	0	0.39	0.31
7.	Injection_work	0.2	0.63	0.27	1	0.00	0.42
8.	Holding_pres_time	0.1	0.74	0.26	0	0.31	0.28
9.	Flow_no_plast	0.26	0.68	0.14	0	0.12	0.24
10.	Last_cycle_time	0.04	1	0.01	0	0.00	0.21
11.	Spec_pres_switchov	0	0.89	0.06	0	0.20	0.23
12.	Backpressure	0.12	0.58	0.27	0	0.34	0.26
13.	Injection_time	0.22	0.42	0.21	0	0.00	0.17
14.	Material	0.17	0.32	0.36	0	0.25	0.22
15.	Heating_cyl1_z1_set	0.06	0.47	0.04	0	0.71	0.26
16.	Plast_time	0.12	0.37	0.05	0	0.13	0.13
17.	Screw_speed	0.14	0.26	0.04	0	0.21	0.13
18.	Speed_max	0.17	0.11	0.05	1	0.22	0.31
19.	Ejector_pos_last	0.05	0	0.16	1	0.02	0.25
20.	Switchov_vol	0.05	0.05	0	0	0.30	0.08
	MLP R2 score	0.87	0.89	0.9	0.8	0.94	
	MLP RMSE	0.092	0.081	0.078	0.11	0.061	
	Correlation coef.	0.94	0.95	0.96	0.92	0.99	

For the sequential joined dataset, width target variable (width1 and width2 merged sequentially in one column) and no FS applied, similarly to the parallel joined dataset the MLP quality measures scores are quite high, here $R^2 = 0.93$, RMSE = 0.083, correlation coefficient = 0.96. Only Pearson FS allowed to increase the MLP model quality scores, while the rest of the models have either almost equal characteristics or even the lower ones. For this dataset the material parameter has received one of the highest scores from each of the FS methods tested. The part type variable, at the same time, gets relatively high scores from the Pearson and Spearman methods. The rest of the scores are divided quite similarly to those for the parallel dataset.

Table 6.14. Feature selection for width target variable, joined sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Cushion_after_hold_pres	1	0.11	1	1	0.24	0.67
2.	Material	0.21	0.33	0.33	1	1.00	0.57
3.	Holding_pressure	0.83	0.89	0.93	0	0.20	0.57
4.	Cushion_smallest	0.91	0.17	0.96	0	0.15	0.44
5.	Last_cycle_time	0.46	1	0.22	0	0.00	0.34
6.	Spec_pres_switchov	0.41	0.94	0.17	0	0.12	0.33
7.	Holding_pres_time	0.49	0.67	0.4	0	0.08	0.33
8.	Screw_speed	0.18	0.28	0.1	1	0.03	0.32
9.	Cooling_time	0.53	0.83	0.04	0	0.02	0.28
10.	Speed_max	0.21	0.06	0.12	1	0.01	0.28

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
11.	Cooling_time_last	0.54	0.78	0	0	0.04	0.27
12.	Switchov_vol	0.65	0.22	0.27	0	0.06	0.24
13.	Backpressure	0.12	0.72	0.22	0	0.10	0.23
14.	Part type	0.65	0	0.5	0	0.00	0.23
15.	Plast_time	0.41	0.39	0.28	0	0.04	0.22
16.	Flow_no_plast	0.28	0.44	0.2	0	0.15	0.21
17.	Injection_Speed	0.08	0.61	0.13	0	0.10	0.18
18.	Injection_work	0	0.56	0	0	0.04	0.12
19.	Injection_time	0	0.5	0	0	0.02	0.10
	MLP R2 score	0.94	0.93	0.93	0.86	0.89	
	MLP RMSE	0.077	0.083	0.079	0.11	0.1	
	Correlation coef.	0.97	0.96	0.97	0.93	0.94	

After looking at the feature selection methods results for sequential and parallel joined datasets, Table 6.15 sums up the MLP model scores depending on the FS method applied. On average, the RReliefF method allows to build models that are as good as the models with no feature selection, while the rest of the methods result in the models whose average measures are lower than those with no feature selection. Therefore, it is decided not to use any feature selection for building the predictive models for width target variable and joined datasets. This result is most probably obtained due to all the parameters included in the joined datasets containing certain significant information about the focus variable with no need to eliminate them.

Table 6.15. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (width target variable, joined datasets)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Width1 target variable, joined parallel dataset	MLP R^2 score	0.86	0.8	0.84	0.81	0.8	0.87
	MLP RMSE	0.084	0.11	0.09	0.1	0.11	0.08
	Correlation coef.	0.94	0.91	0.92	0.91	0.91	0.94
Width2 target variable, joined parallel dataset	MLP R2 score	0.92	0.87	0.89	0.9	0.8	0.94
	MLP RMSE	0.069	0.092	0.081	0.078	0.11	0.061
	Correlation coef.	0.97	0.94	0.95	0.96	0.92	0.99
Width target variable, joined parallel dataset	MLP R2 score	0.93	0.94	0.93	0.93	0.86	0.89
	MLP RMSE	0.083	0.077	0.083	0.079	0.11	0.1
	Correlation coef.	0.96	0.97	0.96	0.97	0.93	0.94
Average	MLP R^2 score, average	**0.90**	0.87	0.89	0.88	0.82	0.90
	MLP RMSE average	**0.08**	0.09	0.08	0.09	0.11	0.08
	Correlation coef. average	**0.96**	0.94	0.94	0.95	0.92	0.96

6.3.2 Thickness target variable

Separate experiments datasets

The same procedure as described for the width target variable has been performed for the thickness focus variable. To use the 5-folds cross-validation, the random state for dataset partitioning was set to 2. Only those parameters that have a corresponding FS method score higher than 0.2 were considered in the MLP model and "Cooling_time" parameter has been disregarded either way due to being a set value of the "Cooling_time_last" parameter. The MLP model was trained with the following settings:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic',
alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = 5,
random_state = 2, momentum = 0.2);
```

Tables 6.16 – 6.18 show the parameters scores based on the different FS methods applied for thickness1, thickness2 and thickness target variables for parallel and sequential datasets. Here **thickness1 and thickness2 target variables** are thicknesses of the first and the second specimens produced during the same machine run correspondingly. **Thickness** with no number **focus variable** includes thickness1 and thickness2 merged sequentially. For the thickness1 focus variable and no feature selection method applied, the R^2 score is 0.87, RMSE = 0.042, while the correlation coefficient is 0.93. None of the FS methods in case of thickness1 target variable increase the MLP scores, however, RReliefF MLP model scores are very close to the ones with no FS used. At the same time, the parameters that get the highest scores from most of the tested feature selection algorithms are those describing holding pressure, cushion, injection work and injection speed. Those receiving the lowest scores are screw speed, clamping force at switchover, the last ejector position, maximum speed and switchover volume. This is similar to the parameters selected by the FS algorithms for the width target variable.

Table 6.16. Feature selection for experiment 1 thickness1 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Holding_pressure	1	0.63	1	1	0.93	0.91
2.	Cushion_after_hold_pres	0.99	0.42	0.91	1	1.00	0.86
3.	Tool_Temperature	0.6	0.46	0.56	1	0.98	0.72
4.	Cushion_smallest	0.87	0.13	0.81	1	0.76	0.71
5.	Holding_pres_time	0.27	0.54	0.28	1	0.66	0.55
6.	Cushion_average	0.79	0.17	0.77	0	0.75	0.50
7.	Injection_Speed	0.25	0.92	0.46	0	0.79	0.48
8.	Injection_work	0.35	0.33	0.43	1	0.27	0.48
9.	Plast_time	0.13	0.67	0.18	1	0.29	0.45
10.	Spec_pres_switchov	0.25	0.96	0.46	0	0.58	0.45
11.	Cooling_time	0.14	0.79	0.15	0	0.69	0.35
12.	Backpressure	0.02	0.71	0.03	0	0.89	0.33
13.	Cooling_time_last	0.1	0.83	0.11	0	0.55	0.32

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
14.	Screw_speed_max	0.14	0.75	0.18	0	0.42	0.30
15.	Last_cycle_time	0.15	1	0.1	0	0.00	0.25
16.	Flow_no_plast	0	0.5	0.11	0	0.58	0.24
17.	Heating_cyl1_z1_set	0.16	0.29	0.02	0	0.66	0.23
18.	Injection_time	0.32	0.37	0.31	0	0.12	0.22
19.	Nozzle_tempr_z2_average	0.16	0.25	0	0	0.66	0.21
20.	Closing_force	0.04	0.88	0.06	0	0.00	0.20
21.	Screw_speed	0.17	0.04	0.21	0	0.30	0.14
22.	Clamp_force_switchov	0.01	0.58	0.05	0	0.00	0.13
23.	Ejector_pos_last	0.24	0.08	0.11	0	0.10	0.11
24.	Speed_max	0.12	0	0.15	0	0.22	0.10
25.	Switchov_vol	0.04	0.21	0.01	0	0.00	0.05
	MLP R^2 score	0.76	0.84	0.82	0.77	0.85	
	MLP RMSE	0.053	0.047	0.046	0.056	0.044	
	Correlation coef.	0.88	0.92	0.91	0.89	0.92	

In case of the thickness2 focus characteristics (thickness of the second out of two target variables produced simultaneously) and no FS applied, R^2 equals to 0.88, RMSE is 0.041 and correlation coefficient is 0.95. As seen from Table 6.17, Spearman method allows to improve them: R^2 to 0.95, RMSE to 0.025, correlation to 0.98. Similar parameter set receives the highest and the lowest scores as in case with thickness1, which is meaningful as the parts with thickness1 and thickness2 are produced during the same production run.

Table 6.17. Feature selection for experiment 1 thickness2 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Cushion_after_hold_pres	0.98	0.63	0.91	1	0.72	0.85
2.	Holding_pressure	1	0.54	1	1	0.64	0.84
3.	Tool_Temperature	0.69	0.37	0.56	1	0.77	0.68
4.	Cushion_smallest	0.88	0.17	0.87	0	0.51	0.49
5.	Injection_work	0.33	0.46	0.34	1	0.21	0.47
6.	Cushion_average	0.8	0.21	0.82	0	0.51	0.47
7.	Spec_pres_switchov	0.3	1	0.44	0	0.35	0.42
8.	Backpressure	0.13	0.71	0.16	0	1.00	0.40
9.	Injection_Speed	0.26	0.83	0.4	0	0.50	0.40
10.	Holding_pres_time	0.26	0.75	0.29	0	0.64	0.39
11.	Screw_speed	0.26	0	0.24	1	0.36	0.37
12.	Cooling_time_last	0	0.92	0.03	0	0.44	0.28
13.	Cooling_time	0.01	0.79	0.05	0	0.51	0.27
14.	Last_cycle_time	0.17	0.96	0.13	0	0.00	0.25
15.	Injection_time	0.4	0.33	0.32	0	0.15	0.24

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
16.	Screw_speed_max	0.16	0.5	0.19	0	0.33	0.24
17.	Plast_time	0.02	0.67	0.21	0	0.21	0.22
18.	Flow_no_plast	0.08	0.42	0.15	0	0.34	0.20
19.	Closing_force	0.01	0.88	0.03	0	0.00	0.18
20.	Heating_cyl1_z1_set	0.04	0.29	0.14	0	0.42	0.18
21.	Nozzle_tempr_z2_average	0.04	0.25	0.13	0	0.42	0.17
22.	Speed_max	0.2	0.04	0.18	0	0.38	0.16
23.	Clamp_force_switchov	0	0.58	0.04	0	0.00	0.12
24.	Ejector_pos_last	0.27	0.08	0.13	0	0.02	0.10
25.	Switchov_vol	0.07	0.13	0	0	0.00	0.04
	MLP R^2 score	0.9	0.9	0.95	0.51	0.91	
	MLP RMSE	0.038	0.035	0.025	0.085	0.034	
	Correlation coef.	0.95	0.96	0.98	0.77	0.95	

For the sequential dataset and thickness target variable (thicknesses of the first and the second specimens produced simultaneously merged sequantially), the MLP model quality with no feature selection is lower than for the parallel dataset, thickness1 and thickness2: $R^2 = 0.72$, RMSE = 0.059, correlation coefficient = 0.87. The use of the Spearman method increases these scores: $R^2 = 0.78$, RMSE = 0.052, correlation coefficient = 0.89. Apart from the lower model quality, the selected parameters list with significant scores is similar to that for thickness1 and thickness2.

Table 6.18. Feature selection for experiment 1 thickness target variable, sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Holding_pressure	1	0.54	1	1	0.38	0.78
2.	Cushion_after_hold_pres	0.98	0.37	0.91	1	0.40	0.73
3.	Tool_Temperature	0.64	0.42	0.56	1	0.67	0.66
4.	Cushion_smallest	0.87	0.13	0.84	1	0.30	0.63
5.	Holding_pres_time	0.26	0.67	0.28	1	0.57	0.56
6.	Screw_speed_max	0.15	0.92	0.18	0	1.00	0.45
7.	Injection_work	0.34	0.33	0.38	1	0.11	0.43
8.	Cushion_average	0.8	0.21	0.79	0	0.35	0.43
9.	Spec_pres_switchov	0.27	0.96	0.45	0	0.38	0.41
10.	Injection_Speed	0.25	0.88	0.43	0	0.35	0.38
11.	Backpressure	0.03	0.71	0.1	0	0.67	0.30
12.	Injection_time	0.36	0.63	0.31	0	0.13	0.29
13.	Cooling_time	0.07	0.75	0.1	0	0.49	0.28
14.	Flow_no_plast	0.04	0.46	0.13	0	0.75	0.28
15.	Cooling_time_last	0.03	0.79	0.06	0	0.39	0.25
16.	Last_cycle_time	0.16	1	0.11	0	0.00	0.25

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
17.	Plast_time	0.07	0.58	0.2	0	0.33	0.24
18.	Closing_force	0.02	0.83	0.05	0	0.00	0.18
19.	Heating_cyl1_z1_set	0.1	0.29	0.08	0	0.38	0.17
20.	Nozzle_tempr_z2_average	0.1	0.25	0.06	0	0.38	0.16
21.	Ejector_pos_last	0.25	0.08	0.12	0	0.24	0.14
22.	Screw_speed	0.21	0	0.22	0	0.25	0.14
23.	Speed_max	0.15	0.04	0.16	0	0.23	0.12
24.	Clamp_force_switchov	0	0.5	0.04	0	0.00	0.11
25.	Switchov_vol	0.05	0.17	0	0	0.00	0.04
	MLP R^2 score	0.75	0.74	0.78	0.67	0.69	
	MLP RMSE	0.056	0.058	0.052	0.071	0.06	
	Correlation coef.	0.87	0.87	0.89	0.83	0.86	

Table 6.19 depicts the MLP models performance for the different datasets and FS methods, including the average values of the models' performance characteristics. For the thickness target variable, the Spearman FS method allows to increase the MLP models quality for both parallel and sequential datasets. Therefore, Spearman method will be used for FS prior to training the MLP and kNN predictive models for the thickness target variable.

Table 6.19. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (thickness target variable)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Experiment 1, parallel dataset, thickness1 target variable	MLP R^2 score	0.87	0.76	0.84	0.82	0.77	0.85
	MLP RMSE	0.042	0.053	0.047	0.046	0.056	0.044
	Correlation coef.	0.93	0.88	0.92	0.91	0.89	0.92
Experiment 1, parallel dataset, thickness2 target variable	MLP R2 score	0.88	0.9	0.9	0.95	0.51	0.91
	MLP RMSE	0.041	0.038	0.035	0.025	0.085	0.034
	Correlation coef.	0.95	0.95	0.96	0.98	0.77	0.95
Experiment 1, sequential dataset, thickness target variable	MLP R2 score	0.72	0.75	0.74	0.78	0.67	0.69
	MLP RMSE	0.059	0.056	0.058	0.052	0.071	0.06
	Correlation coef.	0.87	0.87	0.87	0.89	0.83	0.86
Average	MLP R^2 score, average	0.82	0.80	0.83	**0.85**	0.65	0.82
	MLP RMSE, average	0.05	0.05	0.05	**0.04**	0.07	0.05
	Correlation coef. average	0.92	0.90	0.92	**0.93**	0.83	0.91

Joined dataset

The same way as for the width target variable, parallel and sequential joined datasets are also used with the thickness focus variable. The same procedure is applied, random state

for dataset partitioning is set to 0 to split the training subset into 5 folds and only the parameters with the FS scores higher than 0.2 are considered during the MLP model training. The model is created using the following parameters:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic', alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = (X.shape[1]+1)/2, random_state = 2, momentum = 0.2),
```

here `X.shape[1]` is a number of parameters considered in the model. Tables 6.20 – 6.22 show the FS scores for the thickness1, thickness2 and thickness variables from the parallel and sequential joined datasets, where data from all the relevant experiments is merged. **Thickness1 and thickness2 target characteristics** are the thicknesses of the first and the second specimens produced simultaneously. **Thickness target variable**, on the other hand, includes thicknesses of both the first and the second specimens produced during the same run. For thickness1, parallel joined dataset and no feature selection the MLP scores are as follows: R^2 = 0.64, RMSE = 0.076, correlation coefficient = 0.82. In this case, RReliefF and CFS methods increase the model quality to R^2 equal to 0.68, RMSE = 0.072 and correlation coefficient equal to 0.83

Table 6.20. Feature selection for thickness1 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Holding_pressure	1	0.84	1	1	0.65	0.90
2.	Cushion_after_hold_pres	1	0.16	0.95	1	0.74	0.77
3.	Cushion_smallest	0.92	0.79	0.91	0	0.42	0.61
4.	Switchov_vol	0.38	0.05	0.38	1	1.00	0.56
5.	Material	0.25	0.21	0.38	1	0.85	0.54
6.	Ejector_pos_last	0.32	0	0.46	1	0.13	0.38
7.	Spec_pres_switchov	0.32	0.95	0.41	0	0.22	0.38
8.	Injection_work	0.36	0.11	0.33	1	0.00	0.36
9.	Injection_Speed	0.17	0.89	0.2	0	0.26	0.30
10.	Backpressure	0.23	0.74	0.22	0	0.33	0.30
11.	Heating_cyl1_z1_set	0.14	0.58	0.13	0	0.56	0.28
12.	Cooling_time	0.2	0.47	0.15	0	0.55	0.27
13.	Cooling_time_last	0.15	0.42	0.09	0	0.47	0.23
14.	Last_cycle_time	0.03	1	0	0	0.10	0.23
15.	Flow_no_plast	0.06	0.68	0.03	0	0.34	0.22
16.	Injection_time	0.24	0.37	0.2	0	0.13	0.19
17.	Holding_pres_time	0.02	0.63	0	0	0.26	0.18
18.	Plast_time	0.07	0.53	0.03	0	0.11	0.15
19.	Screw_speed	0.01	0.26	0.03	0	0.23	0.11
20.	Speed_max	0	0.32	0	0	0.20	0.10
	MLP R^2 score	0.64	0.65	0.64	0.68	0.68	
	MLP RMSE	0.076	0.075	0.076	0.072	0.072	
	Correlation coef.	0.81	0.82	0.81	0.83	0.83	

For the thickness2 variable (thickness of the second out of two specimens produced during the same run), the MLP characteristics with no feature selection are R^2 = 0.79, RMSE = 0.06, correlation coefficient = 0.9. RFE feature selection method, however, can improve it to R^2 = 0.8, RMSE = 0.057 and correlation coefficient = 0.91. At the same time, it can be seen that for both thickness1 and thickness2 the material parameter receives FS methods scores higher than 0.2, meaning that including this parameter into the datasets and later to the models can help to create prediction models of higher quality.

Table 6.21. Feature selection for thickness2 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Cushion_after_hold_pres	1	0.16	0.96	1	0.90	0.80
2.	Holding_pressure	1	0.68	1	0	0.80	0.70
3.	Switchov_vol	0.4	0.05	0.44	1	1.00	0.58
4.	Material	0.27	0.21	0.43	1	0.80	0.54
5.	Cushion_smallest	0.93	0.32	0.94	0	0.49	0.54
6.	Spec_pres_switchov	0.32	0.89	0.43	0	0.19	0.37
7.	Backpressure	0.12	0.63	0.12	0	0.92	0.36
8.	Injection_work	0.36	0.11	0.31	1	0.01	0.36
9.	Injection_Speed	0.15	0.84	0.17	0	0.23	0.28
10.	Heating_cyl1_z1_set	0.08	0.58	0.07	0	0.62	0.27
11.	Cooling_time	0.14	0.53	0.08	0	0.56	0.26
12.	Flow_no_plast	0.07	0.95	0.03	0	0.22	0.25
13.	Holding_pres_time	0.03	0.74	0.01	0	0.44	0.24
14.	Speed_max	0	0.79	0.01	0	0.40	0.24
15.	Cooling_time_last	0.09	0.47	0.03	0	0.51	0.22
16.	Last_cycle_time	0.02	1	0.01	0	0.00	0.21
17.	Ejector_pos_last	0.33	0	0.51	0	0.09	0.19
18.	Injection_time	0.25	0.26	0.2	0	0.14	0.17
19.	Screw_speed	0	0.37	0.01	0	0.40	0.16
20.	Plast_time	0.02	0.42	0	0	0.16	0.12
	MLP R2 score	0.62	0.8	0.62	0.62	0.69	
	MLP RMSE	0.082	0.057	0.082	0.082	0.074	
	Correlation coef.	0.8	0.91	0.8	0.79	0.83	

In case of the sequential joined dataset and no FS, the MLP characteristics are as follows: R^2 = 0.99, RMSE = 0.068, correlation coefficient = 0.99. Application of the Pearson FS reaches the same R^2 and correlation coefficient values, while the RMSE value is worse than with no feature selection, application of the rest of the FS methods leads to even worse performance. It can also be seen from the table that material and parameter type parameters seem to be significant for prediction of the target variable value.

Table 6.22. Feature selection for thickness target variable, joined sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Spec_pres_switchov	0.71	0.89	0.87	1	0.11	0.72
2.	Cooling_time	0.91	0.61	0.55	1	0.00	0.61
3.	Part type	1	0	0.9	1	0.00	0.58
4.	Last_cycle_time	0.76	1	0.38	0	0.02	0.43
5.	Injection_work	0.29	0.11	0.73	1	0.02	0.43
6.	Cooling_time_last	0.91	0.56	0.5	0	0.01	0.40
7.	Material	0.15	0.22	0.54	0	1.00	0.38
8.	Holding_pres_time	0.68	0.67	0.38	0	0.01	0.35
9.	Holding_pressure	0.15	0.5	1	0	0.07	0.34
10.	Plast_time	0.52	0.78	0.38	0	0.01	0.34
11.	Cushion_smallest	0.79	0.28	0.45	0	0.07	0.32
12.	Cushion_after_hold_pres	0.69	0.17	0.49	0	0.11	0.29
13.	Flow_no_plast	0.15	0.94	0.06	0	0.19	0.27
14.	Injection_time	0.3	0.33	0.53	0	0.00	0.23
15.	Switchov_vol	1	0.06	0	0	0.06	0.22
16.	Backpressure	0.03	0.83	0.12	0	0.00	0.20
17.	Injection_Speed	0	0.72	0.16	0	0.02	0.18
18.	Speed_max	0.15	0.44	0.08	0	0.00	0.13
19.	Screw_speed	0.13	0.39	0.03	0	0.00	0.11
	MLP R2 score	0.99	0.98	0.99	0.99	0.1	
	MLP RMSE	0.079	0.234	0.11	0.12	3.02	
	Correlation coef.	0.99	0.99	0.99	0.99	0.35	

Table 6.23 includes the MLP accuracy measures for the parallel and sequential joined datasets and thickness target variable. It is visible that the highest average accuracy scores are obtained with no feature selection methods, as for the FS methods even if the R^2 and correlation coefficient values are the same as with no FS, the RMSE value is higher. Therefore, no FS will be used when creating prediction models for the thickness target variable using the joined datasets. It is also worth mentioning that in general parameters that receive the highest scores from the FS methods are similar to those having high scores in case of the separate experiments' datasets.

Table 6.23. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (thickness target variable, joined datasets)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Thickness1 target variable, joined parallel dataset	MLP R^2 score	0.64	0.64	0.65	0.64	0.68	0.68
	MLP RMSE	0.076	0.076	0.075	0.076	0.072	0.072
	Correlation coef.	0.82	0.81	0.82	0.81	0.83	0.83
Thickness2 target variable, joined parallel dataset	MLP R2 score	0.79	0.62	0.8	0.62	0.62	0.69
	MLP RMSE	0.06	0.082	0.057	0.082	0.082	0.074

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
	Correlation coef.	0.9	0.8	0.91	0.8	0.79	0.83
Thickness target variable, joined parallel dataset	**MLP R2 score**	0.99	0.99	0.98	0.99	0.99	0.1
	MLP RMSE	0.068	0.079	0.234	0.11	0.12	3.02
	Correlation coef.	0.99	0.99	0.99	0.99	0.99	0.35
Average	**MLP R^2 score, average**	**0.81**	0.75	0.81	0.75	0.76	0.49
	MLP RMSE average	**0.07**	0.08	0.12	0.09	0.09	1.06
	Correlation coef. average	**0.90**	0.87	0.91	0.87	0.87	0.67

6.4 Predictive models development

After performing data exploration to gain a better understanding of our data, preprocessing and feature selection to remove features that are irrelevant for the target variables of interest, it is now possible to move to development of predictive models for dimensional target variables (width and thickness). The models will help to identify dimensions of produced parts based on the corresponding process parameters settings and avoid possible deviations.

As discussed in Chapter 2, in most of the reviewed studies there is lack of generalized approach in application of ML methods for prediction of injection molded parts quality. Some of the studies skip the data exploration step, while the others don't have preprocessing and/or feature selection, as a result, ML methods are still not widely used in the real industrial environment and the injection molding in particular, while more and more literature on this topic is being published. This work, however, presents a procedure that can be applied for industrial data analysis for the injection molding process, as well as for other manufacturing processes.

Each of the ML methods used in this chapter has its own hyperparameters that can and need to be tuned to receive a model of high quality. This is done using grid search, in order to go through various combinations of the tuned hyperparameters and select the ones with the best model performance.

The set of hyperparameters used in the grid search is the same both for width and thickness. For the MLP the next set of hyperparameters was used:

- hidden layer sizes : [10, 15, 20, 25, 30],
- activation function : 'relu' , 'logistic',
- solver : 'lbfgs', 'sgd',
- alpha (L2 penalty parameter): [0.0001, 0.05],
- learning rate init: [0.001, 0.01, 0.05, 0.1, 0.3].

For Decision Tree another set of parameters was tuned:

- criterion: 'mse', 'friedman_mse', 'mae',
- maximum tree depth: [5, 7, 10, 12, 15].

For kNN, in its turn, the following hyperparameters were tuned:

- weights: ‘uniform’, ‘distance’,
- number of neighbors: [2, 3, 4, 5, 6, 7].

For Gradient Boost Regressor a different set of hyperparameters is evaluated:

- loss: ‘ls’, ‘lad’, ‘huber’, ‘quantile’,
- learning rate: [0.001, 0.005, 0.01, 0.05, 0.1],
- number of estimators: [50, 100, 150, 200, 250, 300].

In case of AdaBoost, the same hyperparameters as for GBR are tuned, however, this algorithm uses other loss functions:

- loss: ‘linear’, ‘square’, ‘exponential’,
- learning rate: [0.001, 0.005, 0.01, 0.05, 0.1, 1],
- number of estimators': [50, 100, 150, 200, 250, 300].

The last, but not the least is Random Forest, here the number of estimators, max features and criterion hyperparameters were varied:

- number of estimators: [50, 100, 150, 200, 250, 300],
- max features: ‘auto’, ‘sqrt’, ‘log2’,
- criterion: ‘mse’, ‘mae’.

Each of the algorithms has a significant number of hypermeters that can be evaluated and tuned, however, due to the time limitations, the ones listed above are considered in this work. At the same time, it is worth mentioning that depending on the algorithm, the training time significantly varies. In case of the MLP it is up to 1-2 minutes, while for kNN and Decision Tree Regressor it is about 10 seconds. At the same time, GBR, AdaBoost and Random Forest might take up to 5 minutes training time on the sequential joined datasets considered in this dissertation. If the amount of data will increase, the training times will increase correspondingly, and kNN training time might become significantly larger due to the nature of the algorithm and its “memorizing” of the training data samples.

6.4.1 Width target variable

This section presents hyperparameters and performance measures (R^2, RMSE and correlation coefficient) for the best performing models as a result of the grid search for the width target variable. The nominal value of the width target variable is 10 mm for all the datasets. The models were created for the separate experiments’ datasets (data for experiments 1-4 saved in the separate files), as well as for the joined datasets. It is important to note that one of the reasons for doing so is that different datasets contain different parameter sets, therefore performance of the models trained on the separate datasets varies due to the parameters present in the datasets and then included in the models. In both separate and joined datasets parallel and sequential datasets are used. In the parallel datasets for experiments 1, 3-4 the data for the specimens that were produced during the same run is saved in the parallel columns. In the sequential ones, these values

follow each other in one column and duplicate the data of one IMM run twice. As a result, in the parallel datasets **width1, width2, thickness1 and thickness2 target variables** are included, these variables correspond to the width and thickness of the first and the second specimens produced during the same production runs. In the sequential datasets, on the other hand **width and thickness focus variables** include values of the corresponding measures for both specimens merged sequentially.

In case of the parallel datasets, the multi-output (2 output) models are created, where the first output corresponds to the predicted value of width1 and the second to width2. Due to this, only MLP, kNN, Decision Tree Regressor and Random Forest are used, as these are the algorithms whose implementation supports the multi-output models creation. In case of the sequential datasets, on the other hand, MLP, kNN, Decision Tree Regressor, GBR, AdaBoost and Random Forest are utilized. While training the MLP and kNN models based on the separate experiments' datasets RReliefF feature selection is used, where the threshold for including/ excluding parameters is set to 0.2. The rest of the algorithms use the full set of features in the beginning and score them during the training process. In case of the joined datasets no feature selection is used, based on the results from Section 6.3.

All the datasets are divided into 70% training and 30% testing subsets, and 5-folds cross-validation is performed on the training dataset to avoid overfitting.

Separate experiments datasets

Experiment 1

Table 6.24 presents results of the hyperparameter optimization for the used ML methods on the parallel dataset obtained from experiment 1, while Table 6.25 for the sequential dataset. The last six rows of the tables show the models' quality characteristics for the train and test datasets for the best performing set of hyperparameters. For the parallel dataset the best model quality on training set is obtained using kNN and Decision Tree algorithms, these performance characteristics are better than the ones obtained without MLP model tuning when selecting the feature selection method, as R^2 for the width1 MLP model is equal to 0.82 and for width2 – 0.9. Here, MLP's R^2 is equal to 0.92, when predicting both width1 and width2 simultaneously, while kNN's and Decision Tree's R^2 is 0.99. At the same time, when analyzing the algorithms' performance on the test set (previously unseen data), Random Forest has the best performance with $R^2 = 0.89$.

For the sequential dataset, the MLP performance on the train data with the tuned hyperparameters ($R^2 = 0.88$) is still better than the one developed for the feature selection algorithm ($R^2 = 0.82$). At the same time, for the sequential dataset Random Forest outperforms the other algorithms on both training and testing dataset. In general, all the models developed using experiment 1 data have acceptable quality and demonstrate satisfactory generalization abilities based on their performance on the train dataset.

Table 6.24. Results of predictive models hyperparameter optimization for width1 and width2, parallel dataset, experiment 1

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	30	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	3	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mae
max_depth	-	12	-	
n_estimators	-	-	-	150
max_features	-	-	-	auto
R^2 train set	0.92	0.99	0.99	0.98
RMSE train set	0.07	0.01	0.01	0.03
Correl. coef. train set	0.96	0.99	0.99	0.99
R^2 test set	0.82	0.75	0.79	0.89
RMSE test set	0.1	0.12	0.11	0.08
Correl. coef. test set	0.92	0.86	0.93	0.93

Table 6.25. Results of predictive models hyperparameter optimization for width, sequential dataset experiment 1

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	10	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.01	1	-
weights	-	-	uniform	-	-	-
number of neighbors	-	-	3	-	-	-
loss	-	-	-	lad	exponential	-
criterion	-	mae	-	-	-	mae
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	300	150	150
max_features	-	-	-	-	-	sqrt
R^2 train set	0.88	0.93	0.92	0.85	0.91	0.93
RMSE train set	0.08	0.06	0.07	0.09	0.07	0.06
Correl. coef. train set	0.94	0.96	0.96	0.94	0.96	0.97
R^2 test set	0.87	0.76	0.82	0.78	0.82	0.81
RMSE test set	0.1	0.13	0.12	0.13	0.12	0.12
Correl. coef. test set	0.94	0.87	0.91	0.91	0.91	0.9

Experiment 2

Unlike during experiments 1, 3 and 4, in the experiment 2 only one 15 mm thick dogbone specimen was produced per production run. Due to this, there is only a sequential dataset

for experiment 2. As it can be seen from Table 6.26, the best performing algorithms with the same model scores are Decision Tree, kNN and GBR. They outperform MLP, AdaBoost and Random Forest on the train set, but quite insignificantly. On the test dataset, the best performance is observed for Decision Tree and GBR algorithms. Even though the overfitting should not be present, as 5-folds cross-validation and division into the training and testing sets is used, this dataset consists of only 72 data samples and might be too small to create a model that will adequately generalize on the new data. The last six rows of the Table 6.26 show the models' performance scores for the set of the best performing hyperparameters.

Table 6.26. Results of predictive models hyperparameter optimization for width, experiment 2

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	25	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	0.005	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	7	-	-	-
loss	-	-	-	ls	exponential	-
criterion		mse		-	-	mse
max_depth		5		-	-	-
n_estimators	-	-	-	100	100	50
max_features	-	-	-	-	-	auto
R^2 train set	0.98	0.99	0.99	0.99	0.98	0.98
RMSE train set	0.01	0.01	0.01	0.01	0.01	0.01
Correl. coef. train set	0.99	0.99	0.99	0.99	0.99	0.99
R^2 test set	0.98	0.99	0.97	0.99	0.95	0.96
RMSE test set	0.02	0.01	0.02	0.01	0.02	0.02
Correl. coef. test set	0.99	0.99	0.99	0.99	0.98	0.98

Experiment 3

In case of experiment 3 and the corresponding Tables 6.27 and 6.28, for the parallel dataset all four algorithms have very high performance on the train set, however, MLP has a higher RMSE in comparison to the rest of the methods. The best test set performance is observed for the MLP and Decision Tree models. A similar situation is seen for the sequential train dataset, as only value of the RMSE is slightly worse for MLP and AdaBoost, while R^2 and the correlation coefficient are equally high for all the applied algorithms. Decision Tree and Random Forest have the best performance on the test dataset, with the highest R^2 and correlation coefficient and the lowest RMSE. In general, all the models have high scores, however, Random Forest has the best performance and is recommended to use over the Decision Tree, as it is more robust and uses a collection of decision trees instead of one, increasing its overall generalization abilities and omitting the bias. At the same time, Decision Tree model is easier to interpret and understand for a human expert.

Table 6.27. Results of predictive models hyperparameter optimization for width1 and width2, parallel dataset, experiment 3

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	20	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	10	-	
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.99	0.99	0.99	0.99
RMSE train set	0.03	0.01	0.01	0.01
Correl. coef. train set	0.99	0.99	0.99	0.99
R^2 test set	0.99	0.99	0.93	0.98
RMSE test set	0.02	0.02	0.06	0.03
Correl. coef. test set	0.99	0.99	0.97	0.99

Table 6.28. Results of predictive models hyperparameter optimization for width, sequential dataset, experiment 3

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	20	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	2	-	-	-
loss	-	-	-	huber	linear	-
criterion		mse	-	-	-	mse
max_depth		7	-	-	-	-
n_estimators	-	-	-	50	50	250
max_features	-	-	-	-	-	auto
R^2 train set	0.99	0.99	0.99	0.99	0.99	0.99
RMSE train set	0.03	0.01	0.01	0.01	0.02	0.01
Correl. coef. train set	0.99	0.99	0.99	0.99	0.99	0.99
R^2 test set	0.99	0.99	0.86	0.99	0.99	0.99
RMSE test set	0.04	0.02	0.08	0.03	0.03	0.02
Correl. coef. test set	0.99	0.99	0.94	0.99	0.99	0.99

Experiment 4

When looking at Tables 6.29 and 6.30, where the models created using the experiment 4 data are shown, it is possible to see that for the parallel train dataset all the models'

performance is very high. However, for the test data, performance of Decision Tree is significantly lower. This might be due to the data partitioning used in this case. As the data is split into train and test set randomly, it might have happened that samples describing behavior that is significantly different from all the samples in the train set were selected into the test set. At the same time, Random Forest outperforms the rest of the algorithms on the test set without any significant model performance decrease in comparison to the train set. For the sequential dataset, Random Forest, AdaBoost and GBR have the highest performance on both train and test sets and demonstrate remarkable generalization abilities.

Table 6.29. Results of predictive models hyperparameter optimization for width1 and width2, parallel dataset, experiment 4

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	30	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	5	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.99	0.99	0.99	0.99
RMSE train set	0.02	0.01	0.01	0.02
Correl. coef. train set	0.99	0.99	0.99	0.99
R^2 test set	0.94	0.45	0.86	0.98
RMSE test set	0.06	0.16	0.08	0.03
Correl. coef. test set	0.97	0.86	0.94	0.99

Table 6.30. Results of predictive models hyperparameter optimization for width, sequential dataset, experiment 4

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	10	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	3	-	-	-
loss	-	-	-	ls	linear	-
criterion	-	friedman_mse	-	-	-	mae
max_depth	-	7	-	-	-	-
n_estimators	-	-	-	100	250	100
max_features	-	-	-	-	-	auto
R^2 train set	0.99	0.99	0.99	0.99	0.99	0.99
RMSE train set	0.02	0.01	0.01	0.01	0.01	0.01

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
Correl. coef. train set	0.99	0.99	0.99	0.99	0.99	0.99
R^2 test set	0.98	0.99	0.95	0.99	0.99	0.99
RMSE test set	0.03	0.01	0.05	0.01	0.01	0.01
Correl. coef. test set	0.99	0.99	0.98	0.99	0.99	0.99

Joined datasets

Unlike separate experiments datasets, joint datasets consist of the samples from all four experiments. The parallel dataset includes only experiments 1, 3 and 4 data, while the sequential one – data from all four of them. In the parallel dataset a material column is added, it does not include the data about the material properties, such as viscosity or other characteristics, but only a denotation for virgin HDPE used on the different days of experiment 1 and two types of recycled materials used in experiments 3 and 4 correspondingly. In the sequential dataset a column for both material and product type is added. The product type column includes value of 1 for the 4 mm dogbone part and 2 for the 15 mm one.

As seen from Table 6.31, kNN has the highest model quality on the train dataset, however, Random Forest outperforms all the other models on the test dataset. For Table 6.32, the similar situation is observed for the train dataset, while AdaBoost has the best performance on the test set.

These models for both parallel and sequential joined datasets are considered more useful than those for the separate datasets, as they are trained on more data, include samples that describe different materials and even two different part types. Overall models' performance on the test data in case of the joined datasets is somewhat lower than that for the separate experiments datasets. This is due to having more scattered data that represents not only different process settings, but also various materials and even geometries. These models are more useful in the real industrial setting for support of the decision-making process in case of production of 4 and 15 mm dogbones than those trained on the separate experiments datasets. However, addition of data about more products with other geometries, materials, etc. will increase usefulness of the models and make them more universal.

Table 6.31. Results of predictive models hyperparameter optimization for width1 and width2, parallel joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	30	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	3	-
loss	-	-	-	-
criterion	-	mae	-	mse

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
max_depth	-	7	-	-
n_estimators	-	-	-	50
max_features	-	-	-	sqrt
R^2 train set	0.88	0.99	0.99	0.99
RMSE train set	0.09	0.02	0.01	0.03
Correl. coef. train set	0.94	0.99	0.99	0.99
R^2 test set	0.84	0.92	0.81	0.95
RMSE test set	0.11	0.07	0.11	0.06
Correl. coef. test set	0.92	0.96	0.9	0.98

Table 6.32. Results of predictive models hyperparameter optimization for width, sequential joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	15	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	0.1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	2	-	-	-
loss	-	-	-	lad	exponential	-
criterion	-	friedman_mse	-	-	-	mae
max_depth	-	7	-	-	-	-
n_estimators	-	-	-	200	300	300
max_features	-	-	-	-	-	sqrt
R^2 train set	0.89	0.98	0.99	0.97	0.94	0.98
RMSE train set	0.1	0.05	0.03	0.06	0.07	0.04
Correl. coef. train set	0.94	0.99	0.99	0.98	0.97	0.99
R^2 test set	0.85	0.92	0.87	0.94	0.91	0.93
RMSE test set	0.11	0.08	0.11	0.07	0.09	0.08
Correl. coef. test set	0.93	0.96	0.94	0.97	0.95	0.97

6.4.2 Thickness target variable

This section presents models with the best performance and hyperparameters tuned using grid search for the thickness target variable. The nominal values of thickness are 4 mm (data from experiments 1, 3 and 4) and 15 mm (experiment 2). The same datasets used for development of the models for the width focus variable are utilized here. As a result, models are trained on both separate experiments and joined datasets. The parallel datasets include **thickness1 and thickness2 target variables**, which are the thicknesses of the first and the second specimens produced during the same machine run. **Thickness focus variable** in the sequential datasets, at the same time, includes both thickness1 and thickness2 merged sequentially.

For development of MLP and kNN models that are trained on the separate experiments datasets, Spearman's feature selection is used. Spearman's FS was identified as the best performing FS method for the thickness target variable in Section 6.3. Once again,

threshold for acceptance/ rejection of the parameters is set to 0.2. The rest of the ML algorithms perform "inner" feature selection. No feature selection is performed prior to training models with any methods using joined datasets, based on results from Section 6.3.

The datasets are divided into 70% training and 30% testing subsets, and 5-folds cross-validation is performed on the training set. This way it is more clearly visible whether the proposed models are able to generalize on the previously unseen data. Models' performance measures are shown both for training and testing data.

Separate experiments datasets

Experiment 1

Average R^2 of the non-tuned MLP models for the training dataset based on the Spearman's FS is equal to 0.85, the tuned MLP for the parallel dataset has $R^2 = 0.86$ and for the sequential it is 0.76, as it can be seen from Tables6.33 and 6.34. The last six rows in the tables correspond to the models' quality characteristics for the set of the best performing hyperparameters. However, Random Forest model has high scores not only on the training set for the parallel dataset, but also outperforms the rest of the models on the test set. When it comes to sequential dataset, kNN has one of the best scores on the train set and the best score on the test set for experiment 1. In general, all the models have acceptable scores and can be used as a starting point for development of more robust models for prediction of thickness for 4 mm thick dogbones manufactured from virgin HDPE.

Table 6.33. Results of predictive models hyperparameter optimization for thickness1 and thickness2, parallel dataset, experiment 1

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	30	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	3	-
loss	-	-	-	-
criterion	-	mae	-	mse
max_depth	-	7	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.86	0.99	0.99	0.99
RMSE train set	0.05	0.01	0.01	0.01
Correl. coef. train set	0.91	0.99	0.99	0.99
R^2 test set	0.76	0.85	0.85	0.86
RMSE test set	0.06	0.05	0.05	0.04
Correl. coef. test set	0.87	0.92	0.92	0.93

Table 6.34. Results of predictive models hyperparameter optimization for thickness, sequential dataset, experiment 1

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	15	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	0.001	-
weights	-	-	uniform	-	-	-
number of neighbors	-	-	4	-	-	-
loss	-	-	-	lad	square	-
criterion	-	mae	-	-	-	mse
max_depth	-	7	-	-	-	-
n_estimators	-	-	-	100	100	250
max_features	-	-	-	-	-	sqrt
R^2 train set	0.76	0.96	0.93	0.92	0.92	0.96
RMSE train set	0.06	0.02	0.03	0.04	0.04	0.03
Correl. coef. train set	0.87	0.98	0.96	0.96	0.96	0.98
R^2 test set	0.6	0.7	0.82	0.76	0.8	0.78
RMSE test set	0.08	0.07	0.05	0.06	0.06	0.06
Correl. coef. test set	0.8	0.85	0.91	0.87	0.9	0.88

Experiment 2

Performance of all the models trained on the experiment 2 dataset is significantly lower on the test set than on the train set. This shows poor ability of the models to generalize on the previously unseen data. Since 5-folds cross-validation was used to evaluate the models on the train set, overfitting should not be present. One of the most probable reasons for occurrence of this phenomena is having a small number of samples (72) in this dataset. Therefore, to increase quality of these models, it is necessary to collect more relevant data on production of the 15 mm thick dogbone parts before these models can become useful for decision-making support in the real industrial setting.

Table 6.35. Results of predictive models hyperparameter optimization for thickness, experiment 2

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	30	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	0.005	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	5	-	-	-
loss	-	-	-	huber	exponential	-
criterion	-	mae	-	-	-	mae
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	100	250	50
max_features	-	-	-	-	-	auto

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
R^2 train set	0.78	0.99	0.99	0.99	0.98	0.99
RMSE train set	0.01	0.01	0.01	0.01	0.01	0.01
Correl. coef. train set	0.89	0.99	0.99	0.99	0.99	0.99
R^2 test set	0.14	0.5	0.47	0.56	0.53	0.52
RMSE test set	0.03	0.02	0.03	0.02	0.02	0.02
Correl. coef. test set	0.57	0.79	0.74	0.82	0.8	0.79

Experiment 3

The models trained on the parallel data from experiment 3 have big difference between the performance on the train and test datasets, and only MLP and Random Forest models have acceptable R^2 scores on the test set , which are 0.74 for MLP and 0.7 for Random Forest. The models trained on the sequential dataset, on the other hand, have a better performance on the test set with R^2 of up to 0.84 in the Decision Tree, GBR and Random Forest models. This can be explained by having twice as many samples as in the case of the parallel datasets and therefore having more data to train on. In addition, some of the samples included in the train set might correspond to thickness1, while samples for this machine run but for thickness2 might get included into the test one. GBR, AdaBoost and Random Forest models created for prediction of thickness of 4 mm thick dogbone parts manufactured from ContainerService recycled HDPE have a relatively high quality and can be used as a starting point for the decision-making assistance. However, addition of more relevant data will increase the models' quality.

Table 6.36. Results of predictive models hyperparameter optimization for thickness1 and thickness2, parallel dataset, experiment 3

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	20	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	4	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	7	-	-
n_estimators	-	-	-	300
max_features	-	-	-	sqrt
R^2 train set	0.87	0.99	0.99	0.96
RMSE train set	0.04	0.01	0.01	0.02
Correl. coef. train set	0.93	0.99	0.99	0.98
R^2 test set	0.74	0.37	0.54	0.7
RMSE test set	0.05	0.08	0.06	0.05
Correl. coef. test set	0.87	0.84	0.77	0.84

Table 6.37. Results of predictive models hyperparameter optimization for thickness, sequential dataset, experiment 3

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	10	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.05	0.001	-
weights	-	-	uniform	-	-	-
number of neighbors	-	-	2	-	-	-
loss	-	-	-	lad	exponential	-
criterion	-	mae	-	-	-	mae
max_depth	-	10	-	-	-	-
n_estimators	-	-	-	200	150	100
max_features	-	-	-	-	-	Log2
R^2 train set	0.81	0.93	0.91	0.89	0.89	0.93
RMSE train set	0.05	0.03	0.03	0.04	0.04	0.03
Correl. coef. train set	0.9	0.97	0.96	0.95	0.95	0.96
R^2 test set	0.74	0.84	0.79	0.84	0.8	0.84
RMSE test set	0.06	0.05	0.05	0.05	0.05	0.05
Correl. coef. test set	0.87	0.92	0.89	0.92	0.9	0.92

Experiment 4

Performance of the models trained on both parallel and sequential datasets for experiment 4 is similar to that for experiment 3, being slightly higher for the test set for the parallel dataset. The reasons for having a relatively high performance on the test dataset for the sequential dataset are similar to those for experiment 3. In general, it is possible to see that the models' performance for the separate experiments parallel and sequential datasets for the thickness dimensional property are lower than for the width target variable. At the same time, the experiment 4 models' performance is acceptable in all cases except the kNN model for the parallel dataset (due to the model's score on the test set).

Table 6.38. Results of predictive models hyperparameter optimization for thickness1 and thickness2, parallel dataset, experiment 4

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	10	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate		-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	5	-	-
n_estimators	-	-	-	150
max_features	-	-	-	sqrt
R^2 train set	0.94	0.99	0.99	0.99

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	Random Forest
RMSE train set	0.03	0.01	0.01	0.01
Correl. coef. train set	0.97	0.99	0.99	0.99
R^2 test set	0.77	0.73	0.38	0.79
RMSE test set	006	0.07	0.1	0.06
Correl. coef. test set	0.89	0.87	0.7	0.89

Table 6.39. Results of predictive models hyperparameter optimization for thickness, sequential dataset, experiment 4

Model's hyperparameter	MLP FS = Spearman	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	20	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	2	-	-	-
loss	-	-	-	lad	square	-
criterion	-	mae	-	-	-	mse
max_depth	-	7	-	-	-	-
n_estimators	-	-	-	300	200	300
max_features	-	-	-	-	-	sqrt
R^2 train set	0.95	0.99	0.99	0.99	0.98	0.99
RMSE train set	0.02	0.01	0.01	0.01	0.02	0.01
Correl. coef. train set	0.98	0.99	0.99	0.99	0.99	0.99
R^2 test set	0.85	0.81	0.84	0.89	0.89	0.87
RMSE test set	0.05	0.06	0.05	0.04	0.04	0.05
Correl. coef. test set	0.93	0.92	0.92	0.95	0.94	0.94

Joined datasets

Based on the data shown in Table 6.40 for the parallel joined dataset, Decision Tree and Random Forest algorithms have the best performance on the train dataset, while on the test set Random Forest model outperforms the rest. As for the Table 6.41 and the sequential dataset, the models show extremely high scores in terms of R^2 and the correlation coefficient both on train and test sets, while the RMSE varies. The worst RMSE is obtained from the MLP model, while the best one from GBR and Random Forest. The models' quality trained on the joined datasets is higher in comparison to the models trained on the separate experiments datasets. This confirms that the more relevant production data is gathered, the better generalization abilities will be shown by the models trained on it. At the same time, the joined datasets have relatively high performance characteristics and can be utilized as a decision-support tool for production of 4 and 15 mm thick dogbones, however, the models need to be updated with more data in order to increase their usefulness and generalization ability.

Table 6.40. Results of predictive models hyperparameter optimization for thickness1 and thickness2, parallel joined dataset

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	15	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	3	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mse
max_depth	-	10	-	-
n_estimators	-	-	-	100
max_features	-	-	-	sqrt
R^2 train set	0.64	0.98	0.99	0.98
RMSE train set	0.08	0.02	0.01	0.02
Correl. coef. train set	0.82	0.99	0.99	0.99
R^2 test set	0.61	0.8	0.67	0.87
RMSE test set	0.08	0.06	0.08	0.05
Correl. coef. test set	0.79	0.9	0.83	0.94

Table 6.41. Results of predictive models hyperparameter optimization for thickness, sequential joined dataset

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	30	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	0.05	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	5	-	-	-
loss	-	-	-	huber	exponential	-
criterion	-	mae	-	-	-	mae
max_depth	-	15	-	-	-	-
n_estimators	-	-	-	150	200	150
max_features	-	-	-	-	-	sqrt
R^2 train set	0.99	0.99	0.99	0.99	0.99	0.99
RMSE train set	0.12	0.02	0.02	0.03	0.06	0.03
Correl. coef. train set	0.99	0.99	0.99	0.99	0.99	0.99
R^2 test set	0.99	0.99	0.99	0.99	0.99	0.99
RMSE test set	0.12	0.06	0.07	0.05	0.06	0.05
Correl. coef. test set	0.99	0.99	0.99	0.99	0.99	0.99

6.4.3 Dimensional properties prediction as a vector of width and thickness

In addition to creation of the models for prediction of separate dimensional properties of the focus parts, it is also of interest to evaluate potential models for their simultaneous

prediction, as a vector. To do this, the sequential joined dataset was used, and width and thickness properties were predicted using the multi-output learning, similar to prediction of width1 and width2 or thickness1 and thickness2 (width and thickness of the first and the second specimens produced during the same machine run). However, in this case width and thickness were used as target variables simultaneously. MLP, Decision Tree Regressor, kNN and Random Forest algorithms can create the multi-output models among the six methods used in this work.

The same way as before, the dataset was divided into 70% training and 30% testing subsets, and 5-folds cross-validation was performed on the training dataset. Grid search was used to evaluate a set of models' hyperparameters to choose the one with the highest score. The evaluated hyperparameter sets are the same as presented in the beginning of Section 6.4 for the corresponding algorithms.

Table 6.42 shows the chosen hyperparameters and the corresponding model performance characteristics (R^2, RMSE and correlation coefficient). Since the used dataset includes only 19 features and based on the results from Section 6.3, no feature selection was performed prior to training the MLP and kNN models. As it can be seen from Table 6.42, models trained using all four algorithms have high performance both on training and on test sets. The best R^2 and correlation coefficient on test set are obtained by applying Decision Tree Regressor and kNN, while Random Forest has the lowest RMSE.

Table 6.42. Results of predictive models hyperparameter optimization for width, thickness, sequential joined dataset

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	35	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mae	-	mae
max_depth	-	12	-	-
n_estimators	-	-	-	200
max_features	-	-	-	auto
R^2 train set	0.98	0.99	0.99	0.99
RMSE train set	0.35	0.03	0.03	0.03
Correl. coef. train set	0.95	0.99	0.99	0.99
R^2 test set	0.99	0.99	0.99	0.95
RMSE test set	0.1	0.08	0.09	0.05
Correl. coef. test set	0.96	0.97	0.97	0.94

Figures 6.10 and 6.11 show comparison of the actual (measured) and predicted by different models width and thickness values. 15 specimens, for which prediction is made, are randomly selected from the test set. In most of the cases Decision Tree, kNN and Random Forest have very similar predictions, while MLP is the one who has the highest

deviation (it also has the lowest scores in the Table 6.42). The Decision Tree, kNN and Random Forest models can be used as a decision-support tool, while more data is obtained to increase the models' performance. It is also suggested that the Random Forest is preferred over the rest of the models due to having lower RMSE and being more robust in comparison to the others. The Decision Tree model can be used when an easy to interpret model is needed.

Figures 6.12 and 6.13 show the scoring that Decision Tree and Random Forest algorithms give to the variables that are included in the training dataset. Even though the algorithms have similar performance characteristics, they score the models parameters in a completely different way. Parameters with the highest scores from the Decision Tree algorithm are cushion after holding pressure and holding pressure. The rest of the parameters have very low values, and some are equal to zero. Random Forest, on the other hand, gives a bigger number of parameters higher scores, and the ones with the highest scores are material, cushion after holding pressure, smallest cushion value, injection time and pressure at switchover.

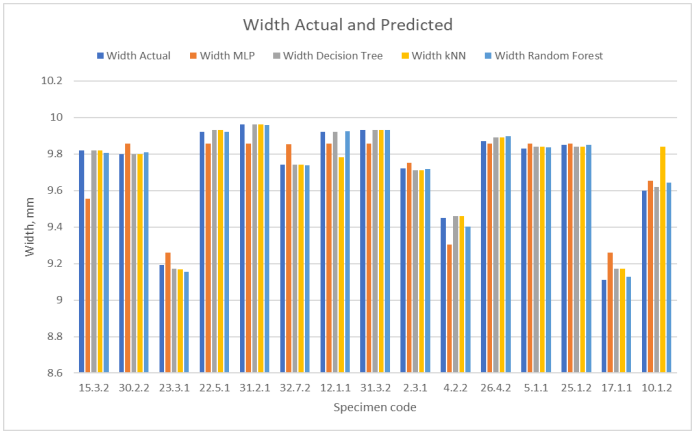


Figure 6.10. Actual and predicted values of width target variable

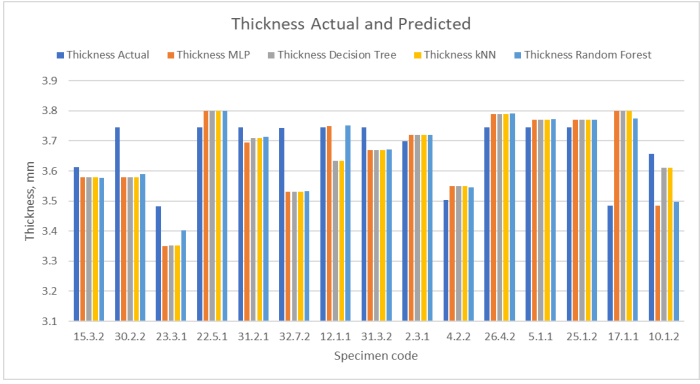


Figure 6.11. Actual and predicted values of thickness target variable

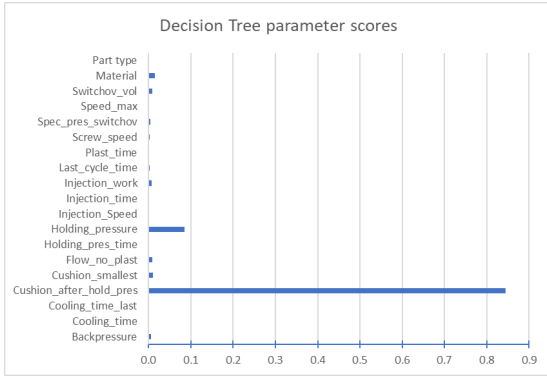


Figure 6.12. Decision Tree parameter scores

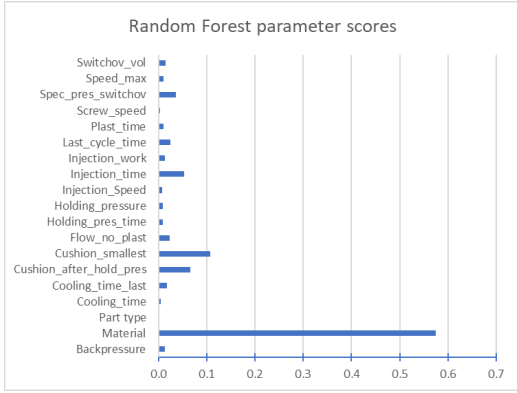


Figure 6.13. Random Forest parameter scores

Chapter 7

Module CC2 – Prediction of mechanical properties of the produced parts based on the general datasets

To be or not to be
Is a stupid question
If you have been happy at least once.

– Boris Akunin, "The Diamond Chariot"

This chapter describes development of the predictive models of mechanical properties of the produced parts using the general datasets, composition of which is explained in Section 4.4.1. The chapter includes four main sections: data exploration, data preprocessing, feature selection and description of the created models.

7.1 Data exploration

Mechanical properties data (tensile modulus, tensile strength and tensile strain at break) is part of the quality data used to create the predictive models. The ability to predict approximate mechanical properties of the produced parts is as important as prediction of the dimensions. There are many products where not only the appearance and dimensional tolerances are critical, but also the amount of load and stress that the product can potentially withstand. To gain a better understanding of the data obtained in the experiments conducted within this PhD work and identify possible outliers, data exploration of the mechanical properties data is presented.

7.1.1 Tensile modulus (Young's modulus) measurements

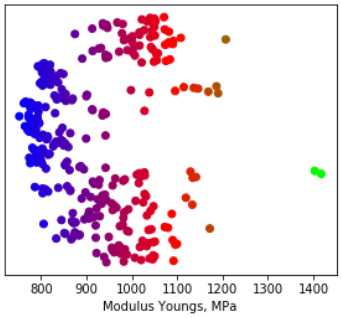
According to the data sheet for the BorSafe™ virgin HDPE, the material's tensile modulus is 800 MPa, for tensile stress tests conducted with 1 mm/min speed. For RePro recycled HDPE only the value of flexural modulus is provided – 900 MPa, while there is no such data available for ContainerService due to the differences in the materials used for production of the recycled plastic pallets. Table 7.1 shows maximum, minimum, average and standard deviation values of the tensile modulus obtained from testing the specimens at hand. The highest minimum tensile modulus is obtained for the RePro recycled HDPE, while the lowest minimum value is for BorSafe™ virgin HDPE for a 4 mm thick specimen. The highest maximum tensile modulus is registered when testing the RePro recycled material, the lowest maximum, on the other hand, is obtained from testing 15 mm thick virgin HDPE dogbone. In general, the data from Table 7.1 shows that recycled materials used in the study have higher values of tensile modulus on average in comparison to the virgin HDPE. The highest average is obtained from the experiment 4 data, while the lowest from experiment 1. The lowest standard deviation is observed for experiment 2 measurements and the highest for experiment number 1.

Table 7.1. Maximum and minimum values of tensile modulus within the experiments 1-4

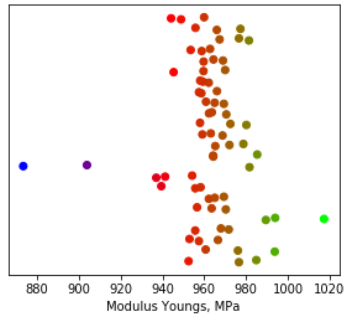
	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Minimum tensile modulus, MPa	752.22	873.43	1003.05	1061.15
Maximum tensile modulus, MPa	1417.17	1017.35	1328.27	1428.53
Average, MPa	922.16	963.3	1142.05	1260.81
Standard deviation, MPa	115.2	18.28	60.76	60.92

Figure 7.1 depicts distribution of the tensile modulus measurements. Based on the Figures 7.1 (a)-(b) for experiments 1 and 2 all the data points, where tensile modulus is higher than 1200 MPa are considered outliers and are removed from further analysis. For experiment 3 all the data points higher than 1300 MPa are considered outliers, while for experiment 4 – those above 1400 MPa. As a result, data about 3 specimens was removed from experiment 1, no specimens data was removed from experiment 2 data, 3 specimens were removed from experiment 3 and 5 from experiment 4. The outliers are examples of samples that represent Young's modulus values that significantly differ from the mean of the random variable under consideration. The reasons for their appearance might be human errors (an error of securing a specimen in the tensile testing machine's grippers) or instrumental errors (video extensometer has failed to correctly identify the white dots on the surface of a specimen). These errors might have also influenced the measurements related to the specimens that are not considered as outliers brining additional uncertainty and a source of error into further calculations and analysis.

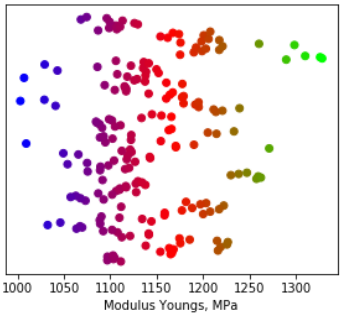
According to Figure 7.1 (a)-(d) for experiment 1 most of the values seem to lay between 800 and 1200 MPa, for experiment 2 between 920 and 1000 MPa, for experiment 3 between 1000 and 1300 MPa, while for experiment 4 in the range of 1100 and 1400 MPa. These differences are mainly based on the differences in the used materials, set processing parameters, as well as the dogbones parts thickness of 4 and 15 mm.



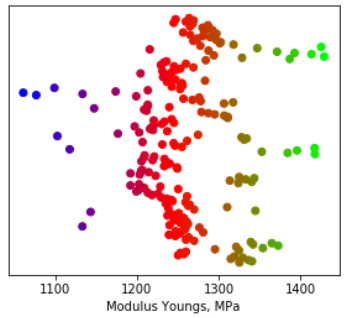
(a) Experiment 1



(b) Experiment 2



(c) Experiment 3

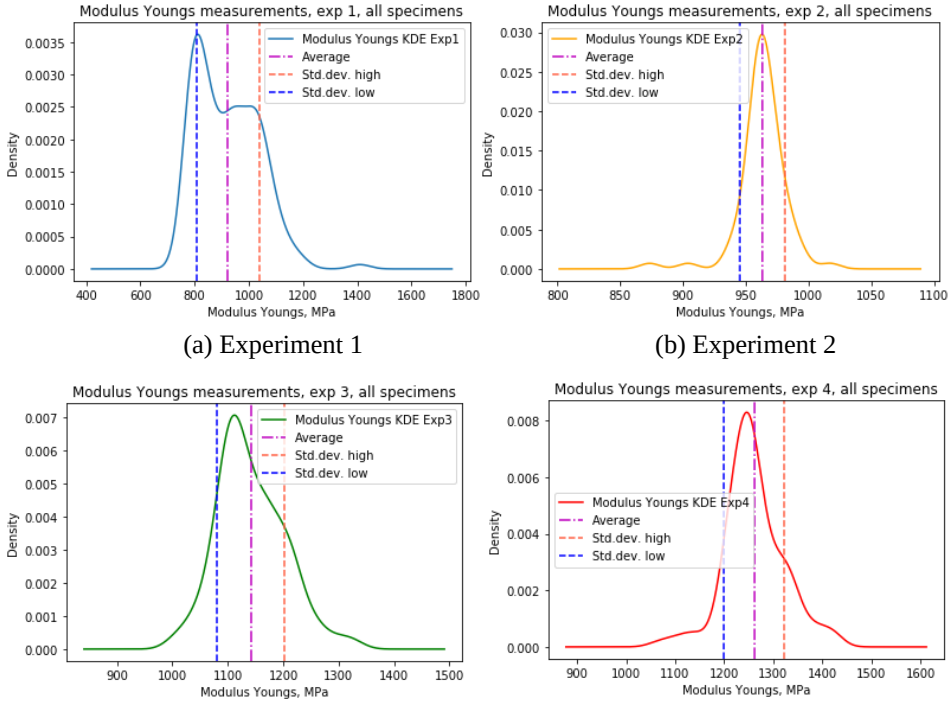


(d) Experiment 4

Figure 7.1. Distribution of tensile modulus values for experiments 1-4

Figure 7.2 helps to increase understanding of the data in Table 7.1 through depicting kernel density estimation of the tensile modulus data and the corresponding average and standard deviation values. For experiments 2-4 the KDE follow the distributions close to normal. KDE for experiment 1, on the other hand, is multimodal. Some of the reasons for this might be a larger number of machine runs in experiment 1 (and more datapoints), differences in the material properties, molding conditions and IMM parameter settings. At the same time, it is possible to see that one of the mode values for experiment 1 is close to 800 MPa, for experiment 2 to about 975 MPa, for experiment 3 it is near to 1100 MPa, while for experiment 4 – 1300 MPa.

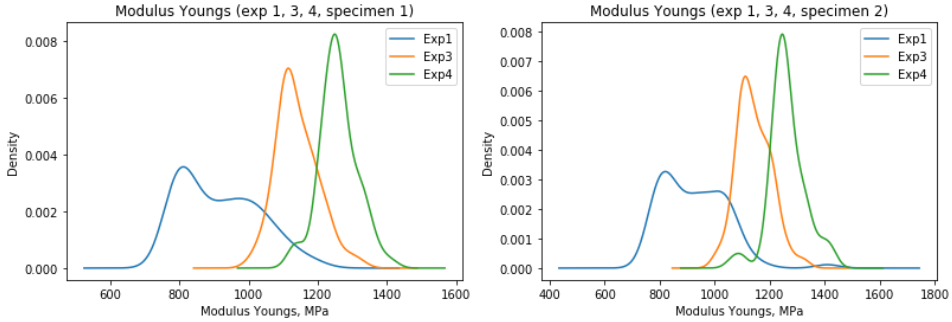
Figures 7.3 (a)-(b) depict KDE for the specimens 1 and 2 separately for experiments 1, 3 and 4. KDEs for the specimens 1 and 2 have very similar shapes, meaning that in terms of tensile modulus the specimens produced during the same run are rather similar to each other. Appendix C includes figures that show dependency of the tensile modulus or Young's modulus on the varied DOE parameters. However, it is hard to see any clear patterns when observing the graphs, as in most cases both higher and lower tensile modulus values are present for low/ high values of the varied process parameters.



(a) Experiment 1 (b) Experiment 2

(c) Experiment 3 (d) Experiment 4

Figure 7.2. Kernel Density Estimation for tensile modulus measurements, experiments 1-4 with mean and standard deviation



(a) Experiments 1, 3, 4, specimen 1 (b) Experiments 1, 3, 4, specimen 2

Figure 7.3. Kernel Density Estimation distributions for tensile modulus, experiments 1, 3, 4 and specimens 1 and 2

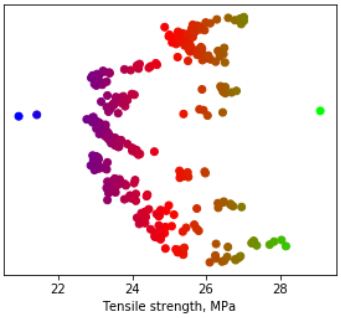
7.1.2 Tensile strength measurements

Similarly to the tensile modulus Table 7.2 includes the minimum, maximum, average and standard deviation values for the obtained tensile strength data. In this case, the largest minimum for the 4 mm specimens was observed for one of the experiment 3 dogbones and the corresponding ContainerService recycled HDPE. The largest maximum value also corresponds to experiment 3. The same way as for the tensile modulus, experiment 3 and 4 and the corresponding recycled HDPE materials have higher tensile strength values on average in comparison to experiments 1 and 2. At the same time, experiment 3 has the largest average value among the experiments, where the dogbone thickness is equal to 4 mm, the lowest value of standard deviation for the same type of dogbones is obtained for experiment 4.

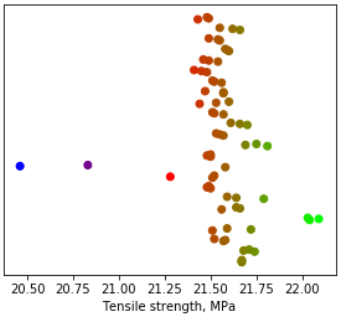
Table 7.2. Maximum and minimum values of tensile strength within the experiments 1-4

	Experiment 1	**Experiment 2**	**Experiment 3**	**Experiment 4**
Minimum tensile strength, MPa	20.93	20.46	27.3	26.7
Maximum tensile strength, MPa	29.09	22.09	37.05	32.22
Average, MPa	24.63	21.56	29.93	28.59
Standard deviation, MPa	1.36	0.21	2.06	1.22

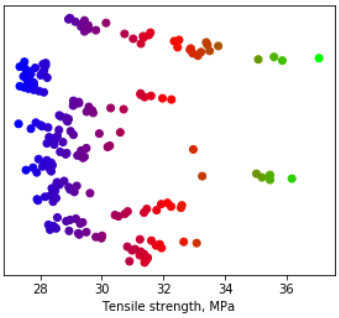
Figure 7.4 shows distribution of the tensile strength values within the conducted experiments. For experiment 1 most of the values are concentrated in the range between 23 and 28.5 MPa, for experiment 2 between 21.25 and 22 MPa, for experiment 3 between 27 and 34, while for experiment 4 between 27 and 31 MPa. As a result, tensile strength values for experiment 1 that are below 23 MPa and above 28.5 MPa are considered outliers, for experiment 2 those below 21.25 MPa and above 22 MPa, for experiment 3 all the values above 34 MPa, while for experiment 4 those higher than 31 MPa. As a result, data about 19 specimens was removed from the experiment 1 dataset, 5 from experiment 2, 9 from experiment 3 and 8 from experiment 4. Once again human and instrumental errors are the most probable sources of error, as it is in case of the Young's modulus.



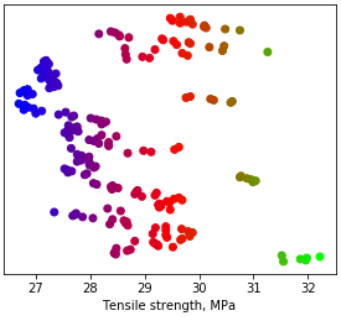
(a) Experiment 1



(b) Experiment 2



(c) Experiment 3

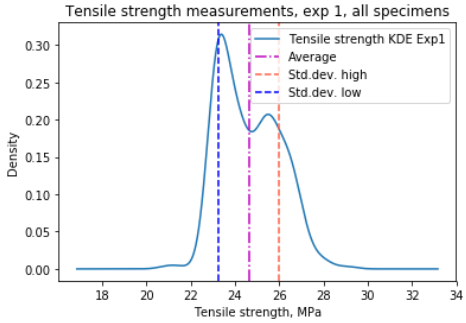


(d) Experiment 4

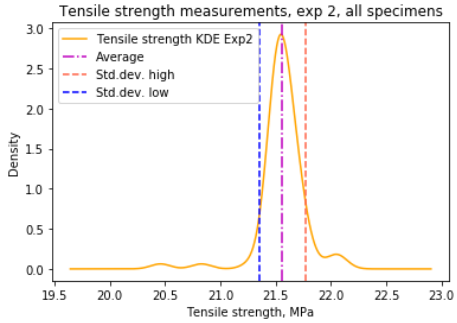
Figure 7.4. Distribution of tensile strength values for experiments 1-4

Figure 7.5 shows kernel density distributions for all the dogbones manufactured in the experiments as well as average and standard deviation values. KDEs for experiments 2 and 4 seem to be close to being normally distributed, while KDEs for experiment 1 and 3 have bimodal or multimodal distributions. The mode values for experiment 1 are around 22.5 and 25.5 MPa, for experiment 2 21.5 MPa, for experiment 3 – 28 and 32 MPa, while for experiment 4 28 MPa.

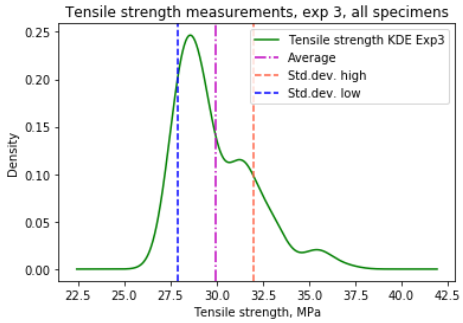
Figure 7.6 shows KDE for experiments 1, 3 and 4 and specimens 1 and 2 separately. Distributions for specimen 2 for experiments 1, 3 and 4 follow multimodal distributions, while for specimen 1 distributions for experiments 3 and 4 are close to normal. These differences might come from uneven heat distribution in the mold during the production of two dogbones at the same time. Figures depicting relationships between the DOE parameters and the tensile strength focus variable are presented in Appendix C, however, similarly to tensile modulus graphs, it is hard to find any clearly visible patterns.



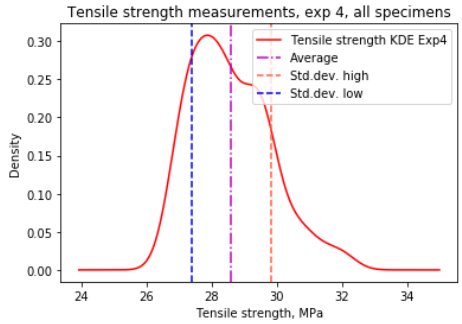
(a) Experiment 1



(b) Experiment 2

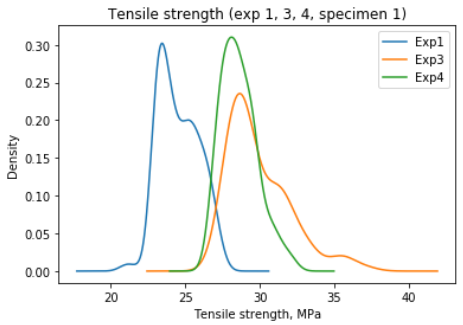


(c) Experiment 3

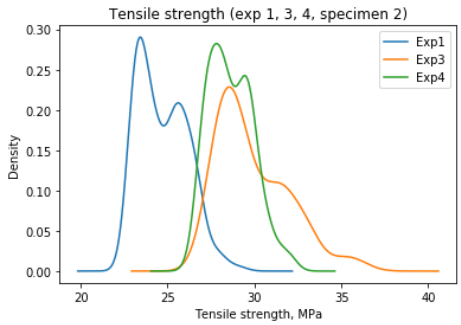


(d) Experiment 4

Figure 7.5. Kernel Density Estimation for tensile strength, experiments 1-4 with mean and standard deviation



(a) Experiments 1, 3, 4, specimen 1



(b) Experiments 1, 3, 4, specimen 2

Figure 7.6. Kernel Density Estimation distributions for tensile strength, experiments 1, 3, 4 and specimens 1 and 2

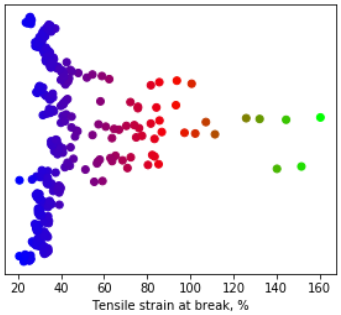
7.1.3 Tensile strain at break measurements

Table 7.3 contains range data for the tensile strain at break measurements. The highest minimum and maximum values for the tensile strain at break were obtained for experiment 2, this experiment also has the biggest range of values for this quality characteristic. The lowest values, on the other hand, were measured during testing the experiment 3 specimens, these values range is the smallest one among the four experiments. According to the data in Table 7.3 tensile strain at break values are on average higher for the experiments 1 and 2 (virgin HDPE) in comparison to those for experiments 3 and 4 (recycled HDPE). The highest average value for the tensile strain at break for experiments where the thickness value is 4 mm is gained for experiment 4, the lowest standard deviation is obtained for experiment 3.

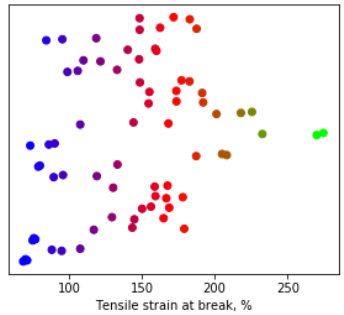
Table 7.3. Maximum and minimum values of the tensile strain at break within the experiments 1-4

	Experiment 1	**Experiment 2**	**Experiment 3**	**Experiment 4**
Minimum tensile strain at break, %	20.67	68.62	8.11	25.21
Maximum tensile strain at break, %	160.45	274.6	34.4	106.35
Average, %	41.85	143.92	22.67	47.18
Standard deviation, %	21.52	47.56	5.7	12.55

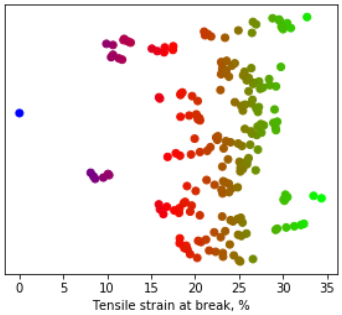
Figure 7.7 shows distribution of the tensile strain at break values. It can be seen from Figure 7.7 (a) that most of the tensile strain at break values of experiment 1 are between 20 and 120 %, while the highest concentration is observed between 20 and 40 %. For experiment 2, most of the values lay between 68 and 250 %, for experiment 3 between 5 and 30 %, while for experiment 4 between 25 and 80 %. Therefore, the values above 120 % for experiment 1, above 250% for experiment 2, below 5 and above 30 % for experiment 3 and above 80 % for experiment 4 are considered outliers and are removed from the datasets for any further analysis. For experiment 1 6 specimens were removed, for experiment 2 – 2 specimens, for experiment 3 – 17, while for experiment 4 – 6. The possible sources or error here are human, instrumental and method. In case of the method error, the break point can be calculated in different ways depending on the selected method. It might be that another break point calculation algorithm would result in less outliers.



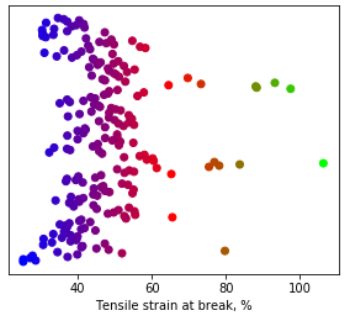
(a) Experiment 1



(b) Experiment 2



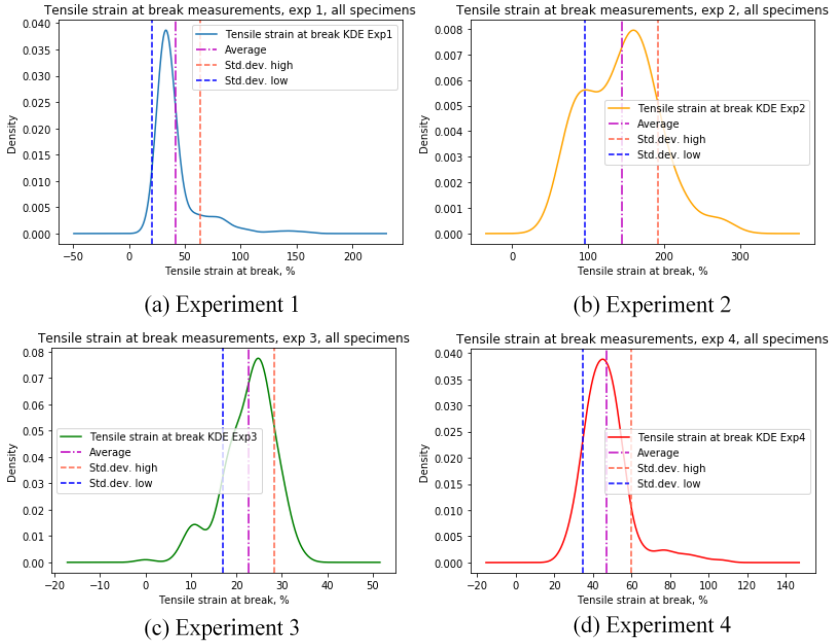
(c) Experiment 3



(d) Experiment 4

Figure 7.7. Distribution of tensile strain at break values for experiments 1-4

Figure 7.8 depicts KDEs for tensile strain at break for all the specimens manufactured during the different experiments. As it is seen from the figures, KDEs for experiments 1 and 4 follow a distribution close to normal, while experiments 2 and 3 are bimodal. Experiment 3 distribution is also the narrowest one, as it has the smallest range of the tensile strain at break values.

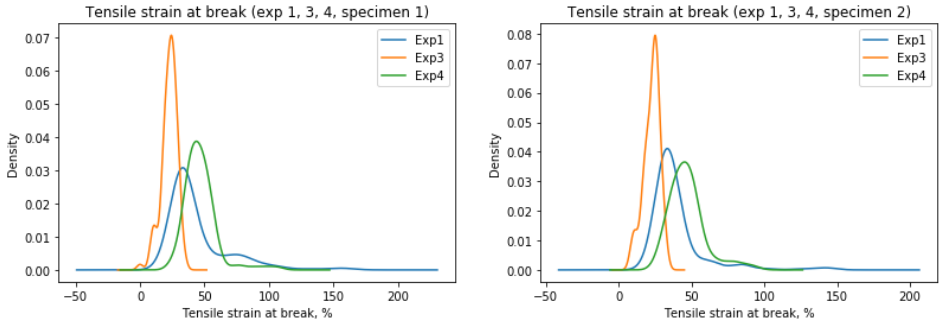


(a) Experiment 1 (b) Experiment 2

(c) Experiment 3 (d) Experiment 4

Figure 7.8. Kernel Density Estimation for tensile strain at break, experiments 1-4 with mean and standard deviation

Figure 7.9 shows KDE for specimens 1 and 2 separately. In general, distributions for specimens 1 and 2 have very similar shapes and almost do not differ. Appendix C includes figures that show relationships between the different DOE parameters and tensile strain at break. However, similarly to the figures for tensile modulus and tensile strength it is hard to see any clear trends.



(a) Experiments 1, 3, 4, specimen 1 (b) Experiments 1, 3, 4, specimen

Figure 7.9. Kernel Density Estimation distributions for tensile strain at break, experiments 1, 3, 4 and specimens 1 and 2

7.2 Data preprocessing

The same way as for the width and thickness target variables, preprocessing for the mechanical properties data has been conducted. In addition to the steps described in Section 6.2, samples number 4.3.2 (4.3) and 28.1.1 (28.1) were removed from the experiment 3 sequential and parallel datasets.

When it comes to removal of outliers for experiments 1 and 2 all the data points, where tensile modulus is higher than 1200 MPa were removed. For experiment 3 all the data points higher than 1300 MPa were eliminated, while for experiment 4 – those above 1400 MPa.

For tensile strength all the samples with values lower than 23 and higher than 28.5 MPa where removed from experiment 1 data, those lower than 21.25 and higher than 22 from the experiment 2 data, values higher than 34 MPa from experiment 3 data samples and those higher than 31 MPa from the experiment 4 data.

In addition the samples with tensile strain at break values above 120 % for experiment 1, above 250% for experiment 2, below 5 and above 30 % for experiment 3 and above 80 % for experiment 4 are considered outliers and are removed from the datasets for any further analysis. As mentioned in the previous sections. The resulting datasets sizes are presented in Table 4.5, Chapter 4. After the outliers' removal the feature selection described in detail in Section 7.3 was carried out. The last step of data preprocessing for the mechanical properties target data was normalization using z-score method.

7.3 Feature selection

Similarly to the dimensional target variables it is necessary to conduct the feature selection process for the mechanical quality characteristics. Young's modulus or tensile modulus, tensile strength and tensile strain at break characteristics are of interest, and therefore, feature selection is performed for each of them separately for separate experiments and joined datasets.

Five different feature selection methods (Pearson's correlation, Spearman's correlation, RFE, CFS and RReliefF) are tested to define which is best to be used in this case. The most fitting FS method is chosen based on application of the FS methods to the experiment 1 dataset. The procedure that is used in this case is similar to that described in Chapter 6, Section 6.3. The most suitable method is chosen based on the score of the MLP model trained using the features considered significant by the corresponding FS method.

7.3.1 Young's modulus target variable

Separate experiments datasets

The list of the parameters considered in feature selection for all the mechanical properties target variables is the same as for the dimensional ones. However, the number of samples in the datasets is a bit smaller due to elimination of the outliers mentioned in the previous section. Random state for dataset partitioning for application of 5-folds cross-validation

is set to 2. All the parameters whose score is higher than 0.2 are considered significant and are included in the MLP model. The model was trained using the following settings:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic',
alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = 12,
random_state = 2, momentum = 0.2);
```

Tables 7.4 – 7.6 show the scores obtained by using the FS methods for identifying the most influential parameters for Young's modulus1, Young's modulus2 and Young's modulus from the parallel and sequential datasets. **Young's modulus1 and Young's modulus2 target variables** are characteristics for specimens number 1 and 2 that are produced during the same production run, they are used as focus variables in the parallel datasets. For the sequential datasets, however, **Young's modulus target variable** is used, which includes Young's modulus1 and Young's modulus2 merged sequentially. When no feature selection is used with Young's modulus1 and the parallel dataset, the MLP scores are as follows: R^2 = 0.4, RMSE = 74.55, correlation coefficient = 0.66. CFS method is able to increase the model's R^2 score to 0.46, RMSE is lowered down to 71.63, while the correlation coefficient becomes 0.76. Parameters with some of the lowest scores according to any of the methods are plasticization time, injection speed, injection time and switchover volume. Nozzle temperature, holding pressure, various cushion characteristics, closing force and holding pressure time, on the other hand, are those with some of the highest scores.

Table 7.4. Feature selection for experiment 1 Young's modulus1 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Nozzle_tempr_z2_average	1	0.71	0.97	1	0.65	0.87
2.	Holding_pressure	0.68	0.92	0.72	1	1.00	0.86
3.	Heating_cyl1_z1_set	1	0.63	1	1	0.66	0.86
4.	Cushion_after_hold_pres	0.79	0.29	0.88	1	0.99	0.79
5.	Cushion_smallest	0.84	0.17	0.87	1	0.80	0.74
6.	Cushion_average	0.86	0.21	0.87	1	0.69	0.73
7.	Closing_force	0.35	0.75	0.62	1	0.48	0.64
8.	Holding_pres_time	0.23	0.67	0.18	1	0.73	0.56
9.	Clamp_force_switchov	0.56	0.37	0.66	1	0.18	0.55
10.	Screw_speed	0.5	0.08	0.53	1	0.32	0.49
11.	Ejector_pos_last	0.67	0	0.67	1	0.07	0.48
12.	Speed_max	0.52	0.04	0.55	1	0.28	0.48
13.	Screw_speed_max	0.49	1	0.5	0	0.34	0.47
14.	Backpressure	0.37	0.79	0.34	0	0.61	0.42
15.	Flow_no_plast	0.58	0.46	0.6	0	0.14	0.36
16.	Cooling_time_last	0.32	0.88	0.33	0	0.11	0.33
17.	Tool_Temperature	0.03	0.83	0.03	0	0.72	0.32
18.	Last_cycle_time	0.19	0.96	0.17	0	0.13	0.29
19.	Cooling_time	0.33	0.5	0.34	0	0.26	0.29

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
20.	Spec_pres_switchov	0.25	0.58	0.28	0	0.19	0.26
21.	Injection_work	0.39	0.25	0	0	0.58	0.24
22.	Plast_time	0.17	0.42	0.57	0	0.00	0.23
23.	Injection_Speed	0	0.54	0.02	0	0.49	0.21
24.	Injection_time	0.39	0.33	0.17	0	0.00	0.18
25.	Switchov_vol	0	0.13	0.12	0	0.20	0.09
	MLP R^2 score	0.08	0.41	0.29	0.46	0.44	
	MLP RMSE	92.38	74.6	79.55	71.63	70.97	
	Correlation coef.	0.4	0.72	0.51	0.76	0.73	

For the Young's modulus2 (characteristic of the second specimen out of two that are produced simultaneously) no feature selection results in the R^2 equal to 0.43, RMSE = 68.66 and correlation coefficient of 0.75. RReliefF and Pearson's correlation are able to improve these characteristics to R^2 0.52, RMSE of 62.72 and 62.93 correspondingly, while the correlation coefficient is either 0.76 or 0.78. Distribution of the parameter scores is quite similar to that for the Young's modulus1 with some slight differences.

Table 7.5. Feature selection for experiment 1 Young's modulus2 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Holding_pressure	0.8	0.92	0.82	1	0.79	0.87
2.	Nozzle_tempr_z2_average	1	0.33	0.97	1	0.47	0.75
3.	Heating_cyl1_z1_set	1	0.29	1	1	0.47	0.75
4.	Cushion_after_hold_pres	0.89	0.25	0.96	1	0.63	0.75
5.	Cushion_smallest	0.92	0.17	0.94	1	0.48	0.70
6.	Cushion_average	0.91	0.21	0.92	1	0.40	0.69
7.	Clamp_force_switchov	0.65	0.37	0.72	1	0.52	0.65
8.	Closing_force	0.41	0.58	0.65	1	0.55	0.64
9.	Speed_max	0.5	0.13	0.52	1	0.35	0.50
10.	Backpressure	0.45	0.88	0.42	0	0.69	0.49
11.	Screw_speed	0.5	0.08	0.51	1	0.29	0.48
12.	Ejector_pos_last	0.69	0	0.66	1	0.00	0.47
13.	Tool_Temperature	0.01	1	0.04	0	1.00	0.41
14.	Flow_no_plast	0.56	0.63	0.56	0	0.27	0.40
15.	Screw_speed_max	0.49	0.54	0.48	0	0.30	0.36
16.	Spec_pres_switchov	0.41	0.79	0.36	0	0.09	0.33
17.	Holding_pres_time	0.16	0.83	0.09	0	0.55	0.33
18.	Plast_time	0.2	0.71	0.56	0	0.00	0.29
19.	Injection_work	0.21	0.5	0.16	0	0.54	0.28
20.	Last_cycle_time	0.09	0.96	0.05	0	0.10	0.24
21.	Injection_time	0.43	0.67	0.08	0	0.00	0.24
22.	Cooling_time	0.24	0.42	0.23	0	0.10	0.20

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
23.	Cooling_time_last	0.19	0.46	0.21	0	0.12	0.20
24.	Injection_Speed	0	0.75	0	0	0.06	0.16
25.	Switchov_vol	0.02	0.04	0.09	0	0.00	0.03
	MLP R^2 score	0.52	0.15	0.34	0.49	0.52	
	MLP RMSE	62.93	96.51	74.22	66.71	62.72	
	Correlation coef.	0.78	0.25	0.62	0.75	0.76	

For the sequential dataset and no FS, the MLP model has an R^2 of 0.52, RMSE equal to 71.51, while the correlation coefficient is 0.74. However, the Pearson's correlation coefficient use can improve these scores to 0.66, 60.42 and 0.82 correspondingly.

Table 7.6. Feature selection for experiment 1 Young's modulus target variable, sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Holding_pressure	0.75	0.92	0.77	1	0.22	0.73
2.	Closing_force	0.42	0.58	0.65	1	1.00	0.73
3.	Clamp_force_switchov	0.6	0.37	0.67	1	0.93	0.71
4.	Cushion_smallest	0.88	0.54	0.9	1	0.21	0.71
5.	Nozzle_tempr_z2_average	1	0.17	0.97	1	0.21	0.67
6.	Heating_cyl1_z1_set	1	0.13	1	1	0.21	0.67
7.	Cushion_after_hold_pres	0.85	0.21	0.91	1	0.16	0.63
8.	Speed_max	0.56	0.04	0.57	1	0.31	0.50
9.	Screw_speed	0.54	0.08	0.55	1	0.24	0.48
10.	Ejector_pos_last	0.68	0	0.66	1	0.00	0.47
11.	Cushion_average	0.88	0.33	0.89	0	0.20	0.46
12.	Backpressure	0.41	0.83	0.37	0	0.54	0.43
13.	Screw_speed_max	0.53	0.88	0.52	0	0.17	0.42
14.	Flow_no_plast	0.62	0.5	0.62	0	0.10	0.37
15.	Holding_pres_time	0.24	0.79	0.17	0	0.47	0.33
16.	Last_cycle_time	0.18	1	0.16	0	0.33	0.33
17.	Cooling_time	0.35	0.63	0.34	0	0.28	0.32
18.	Spec_pres_switchov	0.38	0.71	0.36	0	0.14	0.32
19.	Cooling_time_last	0.32	0.75	0.32	0	0.19	0.32
20.	Tool_Temperature	0.07	0.96	0.03	0	0.35	0.28
21.	Plast_time	0.23	0.42	0.59	0	0.00	0.25
22.	Injection_work	0.34	0.46	0.1	0	0.27	0.23
23.	Injection_Speed	0	0.67	0	0	0.33	0.20
24.	Switchov_vol	0.1	0.29	0.15	0	0.34	0.18
25.	Injection_time	0.44	0.25	0.19	0	0.00	0.18
	MLP R^2 score	0.66	0.55	0.5	0.63	0.56	
	MLP MSE	60.42	69.1	71.91	62.66	68.64	

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
	Correlation coef.	0.82	0.75	0.77	0.8	0.77	

Based on the average MLP models scores presented in Table 7.7 RReliefF feature selection method allows to select the most relevant features and increase the overall model performance. As a result, RReliefF FS method will be used before training the MLP and kNN models for the Young's modulus target variable.

Table 7.7. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (Young's modulus target variable)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Experiment 1, parallel dataset, Young's modulus1 target variable	**MLP R^2 score**	0.4	0.08	0.41	0.29	0.46	0.44
	MLP RMSE	74.55	92.38	74.6	79.55	71.63	70.97
	Correlation coef.	0.66	0.4	0.72	0.51	0.76	0.73
Experiment 1, parallel dataset, Young's modulus2 target variable	**MLP R2 score**	0.43	0.52	0.15	0.34	0.49	0.52
	MLP RMSE	68.66	62.93	96.51	74.22	66.71	62.72
	Correlation coef.	0.75	0.78	0.25	0.62	0.75	0.76
Experiment 1, sequential dataset, Young's modulus target variable	**MLP R2 score**	0.66	0.55	0.5	0.63	0.56	0.66
	MLP RMSE	60.42	69.1	71.91	62.66	68.64	60.42
	Correlation coef.	0.82	0.75	0.77	0.8	0.77	0.82
Average	**MLP R^2 score, average**	0.50	0.38	0.35	0.42	0.50	**0.54**
	MLP RMSE, average	67.88	74.80	81.01	72.14	68.99	**64.70**
	Correlation coef. average	0.74	0.64	0.58	0.64	0.76	**0.77**

Joined dataset

It is of interest to apply feature selection methods to the joined datasets, where data from different experiments is integrated together. There is no difference in the FS procedure and the same five FS methods are tested with the datasets. To apply 5-folds cross-validation the random state for dataset partitioning is set to 0. Only those parameters that have the FS score higher than 0.2 are included in the MLP model to check performance of the feature selection methods. The model is created using the following settings:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic', alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = (X.shape[1]+1)/2, random_state = 2, momentum = 0.2),
```

as in the previous cases of the joined dataset utilization, `X.shape[1]` is a number of parameters considered in the model. For the parallel joined dataset and Young's modulus1 target variable (characteristic of the first out of two specimens produced at the same time) with no FS the MLP scores are as follows: R^2 = 0.77, RMSE = 75.3, correlation coefficient = 0.89. Use of the RFE method can improve the models' R^2 to 0.78, decrease the RMSE to 74.9, while the correlation coefficient stays 0.89. At the same

time, the material parameter has highest scores from all the FS methods except for RFE. In case of predicting mechanical properties of the produced parts, it is important to include any information about the material, especially in the datasets where there are several different materials involved.

Table 7.8. Feature selection for Young's modulus1 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	**Material**	1	0.05	1	1	1.00	0.81
2.	**Spec_pres_switchov**	0.53	0.84	0.56	1	0.08	0.60
3.	**Ejector_pos_last**	0.8	0	0.77	1	0.18	0.55
4.	**Switchov_vol**	0.37	0.42	0.28	1	0.46	0.51
5.	**Flow_no_plast**	0.18	1	0.13	1	0.03	0.47
6.	**Cushion_smallest**	0.27	0.32	0.26	1	0.04	0.38
7.	**Holding_pres_time**	0.44	0.95	0.42	0	0.00	0.36
8.	**Last_cycle_time**	0.13	0.89	0.19	0	0.00	0.24
9.	**Holding_pressure**	0.25	0.53	0.25	0	0.06	0.22
10.	**Plast_time**	0.12	0.68	0.08	0	0.01	0.18
11.	**Injection_Speed**	0.02	0.79	0.03	0	0.04	0.18
12.	**Backpressure**	0	0.74	0.02	0	0.10	0.17
13.	**Heating_cyl1_z1_set**	0.14	0.47	0.13	0	0.12	0.17
14.	**Cushion_after_hold_pres**	0.26	0.26	0.25	0	0.08	0.17
15.	**Cooling_time_last**	0.01	0.63	0.01	0	0.00	0.13
16.	**Cooling_time**	0.01	0.58	0.01	0	0.00	0.12
17.	**Injection_time**	0	0.37	0.06	0	0.03	0.09
18.	**Speed_max**	0.11	0.11	0.06	0	0.02	0.06
19.	**Screw_speed**	0.04	0.21	0	0	0.05	0.06
20.	**Injection_work**	0	0.16	0.06	0	0.05	0.05
	MLP R^2 score	0.44	0.78	0.44	-0.02	0.67	
	MLP RMSE	119.6	74.9	119.6	164.15	91.77	
	Correlation coef.	0.73	0.89	0.73	0	0.84	

For the Young's modulus2 (the second out of two specimens produced simultaneously) and no FS, the MLP characteristics are $R^2 = 0.76$, RMSE = 77.72, correlation coefficient = 0.88. None of the feature selection methods have MLP scores higher than with no FS, this might change if a lower threshold for the parameter inclusion to the model is chosen. Here once again material parameter has the highest scores from all the parameters except the RFE.

Table 7.9. Feature selection for Young's modulus2 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	**Material**	1	0.05	1	1	1.00	0.81
2.	**Spec_pres_switchov**	0.54	0.89	0.55	1	0.08	0.61
3.	**Ejector_pos_last**	0.8	0	0.78	1	0.15	0.55

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
4.	Switchov_vol	0.39	0.42	0.28	1	0.38	0.49
5.	Flow_no_plast	0.16	0.68	0.1	1	0.07	0.40
6.	Cushion_smallest	0.27	0.26	0.23	1	0.03	0.36
7.	Holding_pres_time	0.4	1	0.37	0	0.00	0.35
8.	Holding_pressure	0.25	0.63	0.22	0	0.07	0.23
9.	Last_cycle_time	0.09	0.95	0.11	0	0.00	0.23
10.	Plast_time	0.12	0.79	0.04	0	0.00	0.19
11.	Backpressure	0.01	0.74	0.01	0	0.16	0.18
12.	Injection_Speed	0.01	0.84	0	0	0.04	0.18
13.	Heating_cyl1_z1_set	0.13	0.53	0.09	0	0.09	0.17
14.	Injection_work	0.09	0.58	0.02	0	0.09	0.16
15.	Cushion_after_hold_pres	0.26	0.21	0.21	0	0.05	0.15
16.	Injection_time	0.01	0.47	0.01	0	0.08	0.11
17.	Cooling_time	0	0.37	0	0	0.00	0.07
18.	Cooling_time_last	0.02	0.32	0.02	0	0.00	0.07
19.	Speed_max	0.09	0.11	0.02	0	0.04	0.05
20.	Screw_speed	0.02	0.16	0	0	0.05	0.05

A very similar situation is seen for the Young's modulus target parameter (this parameter includes a sequence of Young's modulus1 and 2 values for specimens 1 and 2 that are produced during the same production run) from the sequential joined dataset. The MLP built with no prior FS has R^2 of 0.73, RMSE equal to 82.68 and correlation coefficient of 0.86. No FS methods contribute to the increase of the models' quality, however, with a threshold lower than 0.2 it might change.

Table 7.10. Feature selection for Young's modulus1 target variable, joined sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Material	1	0	1	1	1.00	0.80
2.	Spec_pres_switchov	0.52	0.89	0.61	1	0.12	0.63
3.	Holding_pressure	0.2	0.67	0.19	1	0.01	0.41
4.	Holding_pres_time	0.41	0.94	0.48	0	0.00	0.37
5.	Last_cycle_time	0.22	1	0.26	0	0.00	0.30
6.	Speed_max	0.12	0.17	0.11	1	0.00	0.28
7.	Flow_no_plast	0.19	0.61	0.2	0	0.09	0.22
8.	Plast_time	0.21	0.56	0.21	0	0.00	0.20
9.	Injection_time	0.05	0.78	0.12	0	0.00	0.19
10.	Injection_Speed	0.03	0.83	0.05	0	0.04	0.19
11.	Cooling_time_last	0.2	0.44	0.12	0	0.00	0.15
12.	Backpressure	0	0.72	0	0	0.02	0.15
13.	Cooling_time	0.21	0.39	0.13	0	0.00	0.15
14.	Cushion_smallest	0.01	0.5	0.09	0	0.04	0.13

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
15.	**Part type**	0.23	0.06	0.28	0	0.00	0.11
16.	**Switchov_vol**	0.2	0.11	0.13	0	0.07	0.10
17.	**Cushion_after_hold_pres**	0	0.33	0.08	0	0.05	0.09
18.	**Injection_work**	0.1	0.22	0.1	0	0.04	0.09
19.	**Screw_speed**	0.05	0.28	0.04	0	0.00	0.07

Based on the data depicted in Table 7.11 no feature selection will be used for building the MLP models based on the joined datasets. The datasets initially include a relatively low number of features and most of them seem to contain useful information about the mechanical characteristics of interest.

Table 7.11. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (Young's modulus target variable, joined datasets)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Young's modulus1 target variable, joined parallel dataset	**MLP R^2 score**	0.77	0.44	0.78	0.44	-0.02	0.67
	MLP RMSE	75.3	119.6	74.9	119.6	164.15	91.77
	Correlation coef.	0.89	0.73	0.89	0.73	0	0.84
Young's modulus2 target variable, joined parallel dataset	**MLP R2 score**	0.76	0.55	0.75	0.55	-0.02	0.54
	MLP RMSE	76.72	100.53	78.36	100.54	160.38	100.98
	Correlation coef.	0.88	0.78	0.88	0.78	0	0.78
Young's modulus target variable, joined parallel dataset	**MLP R2 score**	0.73	0	0.13	0.78	-0.02	-0.02
	MLP RMSE	82.68	155.18	147.18	74.98	161.75	161.76
	Correlation coef.	0.86	-0.14	0.23	0.88	0.17	0
Average	**MLP R^2 score, average**	**0.75**	0.33	0.55	0.59	-0.02	0.40
	MLP RMSE average	**78.23**	125.10	100.15	98.37	162.09	118.17
	Correlation coef. average	**0.88**	0.46	0.67	0.80	0.06	0.54

7.3.2 Tensile strength target variable

Separate experiments datasets

Before building the predictive models for tensile strength, feature selection is performed as in case with the other target variables. The same set of features and procedure are used. Parameters that have a feature selection scores higher than 0.2 are included in the training dataset for the MLP model. 5-folds cross-validation is applied with random state for data partitioning equal to 1. The MLP settings are identified as follows:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic',
alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = 5,
random_state = 2, momentum = 0.2);
```

Similarly to the previous cases, “Cooling_time” parameter is excluded from the model independently of its score, instead “Cooling_time_last” is included (if one of these parameter’s score is high enough) as it is actual value of the same parameter. For the parallel dataset and the tensile strength1 target variable (characteristics of the first out of two specimens produced during the same run), no feature selection application results in the following model characteristics: R^2 = 0.65, RMSE = 0.62, correlation coefficient = 0.87. Application of RReliefF allows to increase these scores to R^2 of 0.81, RMSE of 0.47 and correlation coefficient equal to 0.92.

When applying the feature selection methods such parameters as nozzle temperature, barrel temperature, cushion characteristics, screw speed and clamping force at switchover have some of the highest scores from at least four out of five FS methods. Some of these parameters had the lowest scores when evaluated in terms of the dimensional focus variables. However, in case of the mechanical properties they seem to play an important role for the high quality model development.

Table 7.12. Feature selection for experiment 1 tensile strength1 target variable, parallel dataset

#	Prameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Nozzle_tempr_z2_average	1	0.21	0.99	1	0.99	0.84
2.	Heating_cyl1_z1_set	1	0.17	1	0	1.00	0.63
3.	Cushion_average	0.68	0.29	0.7	1	0.42	0.62
4.	Cushion_after_hold_pres	0.55	0.37	0.65	1	0.48	0.61
5.	Cushion_smallest	0.66	0.25	0.7	1	0.42	0.61
6.	Clamp_force_switchov	0.4	0.67	0.48	1	0.03	0.52
7.	Screw_speed	0.43	0.08	0.46	1	0.43	0.48
8.	Speed_max	0.42	0.13	0.45	1	0.35	0.47
9.	Ejector_pos_last	0.53	0	0.54	1	0.05	0.42
10.	Holding_pressure	0.43	0.75	0.49	0	0.43	0.42
11.	Flow_no_plast	0.56	0.46	0.58	0	0.33	0.39
12.	Screw_speed_max	0.41	0.63	0.43	0	0.41	0.38
13.	Backpressure	0.26	0.79	0.26	0	0.56	0.37
14.	Closing_force	0.24	0.88	0.45	0	0.06	0.33
15.	Spec_pres_switchov	0.28	0.58	0.29	0	0.44	0.32
16.	Tool_Temperature	0.04	0.96	0.06	0	0.44	0.30
17.	Injection_work	0.34	0.5	0.18	0	0.48	0.30
18.	Holding_pres_time	0.04	0.83	0.08	0	0.46	0.28
19.	Last_cycle_time	0.08	1	0.12	0	0.15	0.27
20.	Plast_time	0.27	0.42	0.51	0	0.07	0.25
21.	Injection_Speed	0.09	0.54	0.06	0	0.47	0.23
22.	Switchov_vol	0.02	0.04	0.09	1	0.00	0.23
23.	Cooling_time_last	0	0.92	0.03	0	0.14	0.22
24.	Injection_time	0.39	0.33	0.2	0	0.12	0.21
25.	Cooling_time	0.02	0.71	0	0	0.26	0.20

#	Prameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
	MLP R^2 score	0.51	0.55	0.53	0.65	0.81	
	MLP RMSE	0.79	0.68	0.76	0.68	0.47	
	Correlation coef.	0.82	0.86	0.84	0.84	0.92	

For the same dataset and tensile strength2 (characteristic of the second out of two specimens produced at the same time), no feature selection leads to R^2 value of 0.85, RMSE equal to 0.49 and the correlation coefficient of 0.93. RReliefF application slightly improves those scores to $R^2 = 0.87$, RMSE = 0.46 and correlation coefficient = 0.95. The parameters with the highest scores are the same as for tensile strength1 target variable.

Table 7.13. Feature selection for experiment 1 tensile strength2 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	**Nozzle_tempr_z2_average**	1	0.25	0.99	1	0.99	0.85
2.	**Cushion_smallest**	0.72	0.63	0.73	1	0.49	0.71
3.	**Cushion_average**	0.73	0.29	0.72	1	0.50	0.65
4.	**Heating_cyl1_z1_set**	1	0.21	1	0	1.00	0.64
5.	**Cushion_after_hold_pres**	0.6	0.13	0.66	1	0.54	0.59
6.	**Clamp_force_switchov**	0.41	0.67	0.47	1	0.00	0.51
7.	**Screw_speed**	0.45	0.08	0.48	1	0.38	0.48
8.	**Speed_max**	0.44	0.17	0.48	1	0.29	0.48
9.	**Holding_pressure**	0.47	0.75	0.5	0	0.48	0.44
10.	**Ejector_pos_last**	0.55	0	0.54	1	0.01	0.42
11.	**Flow_no_plast**	0.57	0.46	0.59	0	0.31	0.39
12.	**Backpressure**	0.25	0.79	0.28	0	0.59	0.38
13.	**Screw_speed_max**	0.43	0.37	0.45	0	0.33	0.32
14.	**Spec_pres_switchov**	0.26	0.58	0.26	0	0.46	0.31
15.	**Holding_pres_time**	0.1	0.88	0.15	0	0.40	0.31
16.	**Closing_force**	0.22	0.83	0.42	0	0.01	0.30
17.	**Last_cycle_time**	0.14	1	0.19	0	0.09	0.28
18.	**Tool_Temperature**	0	0.96	0.06	0	0.36	0.28
19.	**Cooling_time**	0.04	0.92	0.02	0	0.36	0.27
20.	**Plast_time**	0.28	0.42	0.54	0	0.10	0.27
21.	**Injection_time**	0.38	0.54	0.2	0	0.18	0.26
22.	**Switchov_vol**	0.07	0.04	0.13	1	0.00	0.25
23.	**Injection_work**	0.3	0.33	0.22	0	0.35	0.24
24.	**Injection_Speed**	0.1	0.5	0.1	0	0.44	0.23
25.	**Cooling_time_last**	0.02	0.71	0	0	0.26	0.20
	MLP R^2 score	0.8	0.84	0.8	0.76	0.87	
	MLP RMSE	0.58	0.51	0.58	0.6	0.46	
	Correlation coef.	0.92	0.93	0.92	0.9	0.95	

When it comes to the sequential dataset, the MLP quality measures with no feature selection are: $R^2 = 0.89$, RMSE = 0.42, correlation coefficient = 0.95 for tensile strength target variable (a merged sequence of tensile strength1 and tensile strength2 characteristics for specimens 1 and 2 produced during the same machine run). Pearson's correlation application can improve the measures' values to 0.92, 0.37 and 0.96 correspondingly. Once again, the features' scores distribution in terms of the highest and the lowest ones is similar to those for the parallel dataset and the corresponding target variables.

Table 7.14. Feature selection for experiment 1 tensile strength target variable, sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Nozzle_tempr_z2_average	1	0.29	0.99	1	0.62	0.78
2.	Cushion_average	0.73	0.21	0.73	1	0.55	0.64
3.	Cushion_smallest	0.72	0.13	0.73	1	0.50	0.62
4.	Cushion_after_hold_pres	0.61	0.37	0.68	1	0.36	0.60
5.	Heating_cyl1_z1_set	1	0.25	1	0	0.63	0.58
6.	Clamp_force_switchov	0.4	0.63	0.43	1	0.07	0.51
7.	Screw_speed	0.45	0.08	0.47	1	0.50	0.50
8.	Speed_max	0.45	0.17	0.47	1	0.38	0.49
9.	Backpressure	0.25	0.79	0.25	0	1.00	0.46
10.	Flow_no_plast	0.58	0.46	0.59	0	0.55	0.44
11.	Holding_pressure	0.49	0.75	0.52	0	0.34	0.42
12.	Screw_speed_max	0.42	0.71	0.43	0	0.50	0.41
13.	Spec_pres_switchov	0.29	0.58	0.28	0	0.61	0.35
14.	Closing_force	0.25	0.96	0.41	0	0.05	0.33
15.	Tool_Temperature	0.06	0.92	0.08	0	0.57	0.33
16.	Holding_pres_time	0.08	0.83	0.1	0	0.50	0.30
17.	Injection_work	0.33	0.5	0.19	0	0.44	0.29
18.	Last_cycle_time	0.12	1	0.14	0	0.13	0.28
19.	Injection_Speed	0.1	0.54	0.07	0	0.66	0.27
20.	Cooling_time	0.01	0.88	0	0	0.45	0.27
21.	Plast_time	0.28	0.42	0.52	0	0.10	0.26
22.	Injection_time	0.41	0.33	0.22	0	0.29	0.25
23.	Switchov_vol	0.08	0.04	0.11	1	0.00	0.25
24.	Ejector_pos_last	0.54	0	0.52	0	0.09	0.23
25.	Cooling_time_last	0	0.67	0.02	0	0.35	0.21
	MLP R^2 score	0.92	0.91	0.86	0.49	0.88	
	MLP RMSE	0.37	0.38	0.46	0.83	0.43	
	Correlation coef.	0.96	0.96	0.93	0.8	0.95	

Table 7.15 sums up the MLP models' quality measures for sequential and parallel datasets and allows to conclude that application of RReliefF FS method leads to selection of the

most relevant parameters and creation of the highest quality MLP models based on the provided data. Therefore, RReliefF will be used to select parameters to be included into the predictive models of tensile strength with the separate experiments datasets.

Table 7.15. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (tensile strength target variable)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Experiment 1, parallel dataset, tensile strength1 target variable	**MLP R^2 score**	0.65	0.51	0.55	0.53	0.65	0.81
	MLP RMSE	0.62	0.79	0.68	0.76	0.68	0.47
	Correlation coef.	0.87	0.82	0.86	0.84	0.84	0.92
Experiment 1, parallel dataset, tensile strength2 target variable	**MLP R2 score**	0.85	0.8	0.84	0.8	0.76	0.87
	MLP RMSE	0.49	0.58	0.51	0.58	0.6	0.46
	Correlation coef.	0.93	0.92	0.93	0.92	0.9	0.95
Experiment 1, sequential dataset, tensile strength target variable	**MLP R2 score**	0.89	0.92	0.91	0.86	0.49	0.88
	MLP RMSE	0.42	0.37	0.38	0.46	0.83	0.43
	Correlation coef.	0.95	0.96	0.96	0.93	0.8	0.95
Average	**MLP R^2 score, average**	0.80	0.74	0.77	0.73	0.63	**0.85**
	MLP RMSE, average	0.51	0.58	0.52	0.60	0.70	**0.45**
	Correlation coef. average	0.92	0.90	0.92	0.90	0.85	**0.94**

Joined dataset

Same way as for the other target variables, the predictive models based on the joined datasets will be also trained for the tensile strength. Due to this we need to see if the feature selection methods need to be used to discard any irrelevant features. The procedure in this case is the same as previously and the feature rejection threshold is set as 0.2, as for all the previous variables and datasets. Random state for the 5-folds data partitioning is set to 0 and the following MLP settings are set:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic', alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = (X.shape[1]+1)/2, random_state = 2, momentum = 0.2)
```

here `X.shape[1]` is a number of parameters considered in the model. Tables 7.16 – 7.18 contain the scores various feature selection methods assign to the parameters from the used datasets. The parallel joined dataset and tensile strength1 focus variable (tensile strength of specimen 1 out of two that are produced at the same time) result in the MLP models with R^2 = 0.93, RMSE = 0.63, correlation coefficient = 0.97, when no FS is applied. Pearson and Spearman correlation allow to increase R^2 to 0.95, decrease RMSE to 0.56 and improve the correlation coefficient to 0.98. Parameters that have the highest scores from most of the FS methods are material, specific pressure at switchover, switchover volume, last ejector position, the smallest cushion value. The ones with the lowest scores, on the other hand, are plasticizing time, screw speed, maximum speed and injection work.

Table 7.16. Feature selection for tensile strength1 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	**Material**	0.85	0.11	0.86	1	1.00	0.76
2.	**Spec_pres_switchov**	0.66	0.68	0.67	1	0.16	0.63
3.	**Switchov_vol**	0.9	0.05	0.74	1	0.48	0.63
4.	**Ejector_pos_last**	1	0	1	1	0.15	0.63
5.	**Cushion_smallest**	0.34	0.47	0.29	1	0.22	0.46
6.	**Flow_no_plast**	0.19	0.84	0.15	1	0.14	0.46
7.	**Heating_cyl1_z1_set**	0.35	1	0.33	0	0.63	0.46
8.	**Holding_pres_time**	0.4	0.95	0.38	0	0.18	0.38
9.	**Holding_pressure**	0.17	0.74	0.14	0	0.19	0.25
10.	**Cushion_after_hold_pres**	0.3	0.32	0.24	0	0.26	0.22
11.	**Last_cycle_time**	0	0.89	0.04	0	0.07	0.20
12.	**Backpressure**	0.03	0.79	0	0	0.10	0.18
13.	**Injection_Speed**	0.07	0.63	0.03	0	0.08	0.16
14.	**Cooling_time**	0.08	0.58	0.09	0	0.04	0.16
15.	**Cooling_time_last**	0.09	0.53	0.11	0	0.02	0.15
16.	**Injection_time**	0	0.42	0.07	0	0.09	0.12
17.	**Plast_time**	0.14	0.37	0.04	0	0.00	0.11
18.	**Screw_speed**	0	0.26	0.04	0	0.17	0.09
19.	**Speed_max**	0.07	0.21	0.02	0	0.15	0.09
20.	**Injection_work**	0.05	0.16	0.15	0	0.05	0.08
	MLP R^2 score	0.95	0.82	0.95	0.82	0.94	
	MLP RMSE	0.56	1.05	0.56	1.05	0.63	
	Correlation coef.	0.98	0.91	0.98	0.91	0.97	

For tensile strength2 (characteristics of specimen 2 out of two that are produced simultaneously), on the other hand, no FS method leads to having an MLP model with R^2 = 0.93, RMSE = 0.66, correlation coefficient = 0.97. Selection of only those parameters that have Pearson's correlation value higher than 0.2, improves these values to 0.94, 0.63 and 0.97 correspondingly. The feature scores have a similar distribution as for the tensile strength1 target variable. For both focus variables the material parameter is one of those with highest score.

Table 7.17. Feature selection for tensile strength2 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	**Material**	0.84	0.11	0.87	1	1.00	0.76
2.	**Spec_pres_switchov**	0.66	0.84	0.67	1	0.16	0.67
3.	**Switchov_vol**	0.9	0.05	0.75	1	0.56	0.65
4.	**Ejector_pos_last**	1	0	1	1	0.15	0.63
5.	**Flow_no_plast**	0.2	1	0.16	1	0.12	0.50
6.	**Heating_cyl1_z1_set**	0.37	0.89	0.34	0	0.67	0.45

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
7.	Cushion_smallest	0.37	0.26	0.3	1	0.28	0.44
8.	Holding_pres_time	0.38	0.58	0.36	0	0.13	0.29
9.	Cushion_after_hold_pres	0.32	0.37	0.24	0	0.31	0.25
10.	Last_cycle_time	0	0.95	0.02	0	0.07	0.21
11.	Holding_pressure	0.19	0.47	0.15	0	0.23	0.21
12.	Injection_Speed	0.07	0.79	0.04	0	0.07	0.19
13.	Injection_work	0.05	0.74	0.15	0	0.02	0.19
14.	Cooling_time	0.07	0.68	0.11	0	0.05	0.18
15.	Cooling_time_last	0.09	0.63	0.13	0	0.02	0.17
16.	Backpressure	0	0.53	0	0	0.12	0.13
17.	Plast_time	0.15	0.42	0.05	0	0.00	0.12
18.	Injection_time	0	0.32	0.06	0	0.10	0.10
19.	Screw_speed	0	0.21	0.03	0	0.14	0.08
20.	Speed_max	0.07	0.16	0.02	0	0.13	0.08
	MLP R2 score	0.94	0.77	0.93	0.83	0.92	
	MLP RMSE	0.63	1.2	0.65	1.02	0.72	
	Correlation coef.	0.97	0.89	0.97	0.92	0.96	

When the sequential dataset is used with no feature selection, the model's characteristics improve in comparison to those for the parallel dataset and are: $R^2 = 0.97$, RMSE = 0.51, correlation coefficient = 0.98. No feature selection methods are able to improve these scores. The parameters' scores are similar to those given to them when the parallel dataset is used.

Table 7.18. Feature selection for tensile strength target variable, joined sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Spec_pres_switchov	1	0.83	1	1	0.17	0.80
2.	Material	0.88	0.11	0.97	1	1.00	0.79
3.	Holding_pres_time	0.76	0.94	0.68	1	0.00	0.68
4.	Part type	0.81	0	0.79	1	0.00	0.52
5.	Flow_no_plast	0.29	0.89	0.29	1	0.08	0.51
6.	Plast_time	0.51	0.39	0.42	1	0.00	0.46
7.	Last_cycle_time	0.6	1	0.37	0	0.01	0.40
8.	Cooling_time	0.71	0.67	0.3	0	0.00	0.34
9.	Cooling_time_last	0.71	0.61	0.28	0	0.00	0.32
10.	Speed_max	0.17	0.17	0.14	1	0.04	0.30
11.	Cushion_smallest	0.46	0.5	0.12	0	0.12	0.24
12.	Switchov_vol	0.76	0.06	0.25	0	0.07	0.23
13.	Injection_time	0.24	0.44	0.36	0	0.04	0.22
14.	Injection_Speed	0.08	0.78	0.05	0	0.12	0.21
15.	Cushion_after_hold_pres	0.37	0.33	0.16	0	0.13	0.20
16.	Backpressure	0	0.72	0	0	0.08	0.16

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
17.	Injection_work	0.26	0.28	0.19	0	0.04	0.15
18.	Holding_pressure	0.04	0.56	0.03	0	0.07	0.14
19.	Screw_speed	0.09	0.22	0.05	0	0.01	0.07
	MLP R2 score	0.94	0.88	0.9	0.8	0.69	
	MLP RMSE	0.69	0.98	0.91	1.31	1.61	
	Correlation coef.	0.97	0.94	0.95	0.9	0.83	

Based on information presented in Table 7.19, even though Pearson correlation allows to slightly increase the MLP performance measures for the tensile strength1 and 2 variables (characteristics of specimens 1 and 2 produced during the same machine run), the best average score is obtained when no FS is performed. Therefore, no feature selection will be used when training the prediction models with parallel and sequential joined datasets.

Table 7.19. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (tensile strength target variable, joined datasets)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Tensile strength1 target variable, joined parallel dataset	MLP R^2 score	0.93	0.95	0.82	0.95	0.82	0.94
	MLP RMSE	0.63	0.56	1.05	0.56	1.05	0.63
	Correlation coef.	0.97	0.98	0.91	0.98	0.91	0.97
Tensile strength2 target variable, joined parallel dataset	MLP R2 score	0.93	0.94	0.77	0.93	0.83	0.92
	MLP RMSE	0.66	0.63	1.2	0.65	1.02	0.72
	Correlation coef.	0.97	0.97	0.89	0.97	0.92	0.96
Tensile strength target variable, joined parallel dataset	MLP R2 score	0.97	0.94	0.88	0.9	0.8	0.69
	MLP RMSE	0.51	0.69	0.98	0.91	1.31	1.61
	Correlation coef.	0.98	0.97	0.94	0.95	0.9	0.83
Average	MLP R^2 score, average	**0.94**	0.94	0.82	0.93	0.82	0.85
	MLP RMSE average	**0.60**	0.63	1.08	0.71	1.13	0.99
	Correlation coef. average	**0.97**	0.97	0.91	0.97	0.91	0.92

7.3.3 Tensile strain at break target variable

Separate experiments datasets

The same way as for the other target variables, five feature selection methods are used to identify which method suits best for selection of parameters to be included in the predictive models for tensile strain at break. The methods' performance is evaluated through building of an untuned MLP model. The training dataset is divided into 5 folds to apply cross-validation procedure when evaluating the obtained model performance. Only parameters that have feature selection scores higher than 0.2 are included into the model. To do this random state for dataset partitioning is set to 5. The MLP model is created using the following settings:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic',
alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = 5,
random_state = 2, momentum = 0.2);
```

Tensile strain at break1 target variable (characteristics of specimen number 1 out of two that are produced at the same time) from the parallel dataset allows to build an MLP model with $R^2 = 0.55$, RMSE = 10.18, correlation coefficient = 0.81 if no feature selection is applied. All feature selection methods, except CFS, are able to improve these scores. The most significant improvement is achieved using the Pearson's correlation, as in this case R^2 becomes as high as 0.69, RMSE as low as 8.09 and the correlation coefficient for the predicted and actual values is 0.89. Barrel temperature, the smallest cushion value, holding pressure, injection time, average nozzle temperature are some of the crucial parameters in the predictive model. Last cycle time, injection speed, last cooling time, injection work and switchover volume, on the other hand, are those that are excluded from the model to increase its performance.

Table 7.20. Feature selection for experiment 1 tensile strain at break1 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Heating_cyl1_z1_set	0.8	0.42	0.94	1	0.52	0.74
2.	Cushion_smallest	1	0.21	0.94	1	0.34	0.70
3.	Holding_pressure	0.96	0.79	0.94	0	0.76	0.69
4.	Injection_time	0.78	0.08	0.39	1	0.94	0.64
5.	Nozzle_tempr_z2_average	0.8	0.63	0.93	0	0.52	0.58
6.	Cushion_average	1	0.5	0.9	0	0.34	0.55
7.	Cushion_after_hold_pres	0.91	0.37	1	0	0.35	0.53
8.	Speed_max	0.39	0.17	0.36	1	0.48	0.48
9.	Flow_no_plast	0.43	0.83	0.47	0	0.46	0.44
10.	Screw_speed_max	0.38	0.92	0.32	0	0.49	0.42
11.	Tool_Temperature	0.09	0.75	0.19	0	1.00	0.41
12.	Spec_pres_switchov	0.51	0.58	0.31	0	0.61	0.40
13.	Closing_force	0.29	1	0.62	0	0.00	0.38
14.	Backpressure	0.23	0.88	0.42	0	0.31	0.37
15.	Ejector_pos_last	0.32	0	0.45	1	0.00	0.35
16.	Holding_pres_time	0.2	0.46	0.01	0	0.77	0.29
17.	Screw_speed	0.39	0.13	0.37	0	0.41	0.26
18.	Cooling_time	0.19	0.71	0.05	0	0.33	0.26
19.	Plast_time	0.26	0.25	0.47	0	0.28	0.25
20.	Clamp_force_switchov	0.3	0.33	0.56	0	0.00	0.24
21.	Last_cycle_time	0.08	0.96	0	0	0.12	0.23
22.	Injection_Speed	0.18	0.54	0.01	0	0.42	0.23
23.	Cooling_time_last	0.17	0.67	0.03	0	0.27	0.23
24.	Injection_work	0.19	0.29	0.03	0	0.00	0.10
25.	Switchov_vol	0	0.04	0.07	0	0.00	0.02

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
	MLP R^2 score	0.69	0.63	0.58	0.12	0.56	
	MLP RMSE	8.09	9.19	9.33	17.99	9.08	
	Correlation coef.	0.89	0.84	0.84	0.45	0.87	

In case of the tensile strain at break2 target variable (characteristics of the second out of two specimens produced during the same machine run), no feature selection MLP scores are unsatisfactory: R^2 = 0.19, RMSE = 7.6, correlation coefficient = 0.67. However, RReliefF FS improves the situation and the MLP model performance: R^2 = 0.76, RMSE = 8.1, correlation coefficient = 0.87.

Table 7.21. Feature selection for experiment 1 tensile strain at break2 target variable, parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	**Cushion_after_hold_pres**	0.92	0.54	1	0	0.21	0.53
2.	**Holding_pressure**	0.93	0.46	0.92	0	0.55	0.57
3.	**Heating_cyl1_z1_set**	0.83	0.25	1	0	0.36	0.49
4.	**Cushion_average**	0.99	0.17	0.91	0	0.14	0.44
5.	**Cushion_smallest**	1	0.13	0.93	1	0.17	0.65
6.	**Nozzle_tempr_z2_average**	0.83	0.21	0.98	0	0.36	0.48
7.	**Closing_force**	0.31	0.96	0.62	0	0.00	0.38
8.	**Clamp_force_switchov**	0.42	0.67	0.59	0	0.00	0.34
9.	**Screw_speed_max**	0.34	0.92	0.36	0	0.52	0.43
10.	**Flow_no_plast**	0.41	0.58	0.5	0	0.39	0.38
11.	**Backpressure**	0.14	0.88	0.38	0	0.53	0.39
12.	**Spec_pres_switchov**	0.25	0.79	0.32	0	0.63	0.40
13.	**Plast_time**	0.38	0.42	0.49	0	0.51	0.36
14.	**Last_cycle_time**	0.13	1	0.12	0	0.16	0.28
15.	**Injection_time**	0.43	0.37	0.36	0	0.71	0.37
16.	**Cooling_time_last**	0.11	0.83	0.1	0	0.32	0.27
17.	**Holding_pres_time**	0.32	0.5	0.13	1	0.69	0.53
18.	**Tool_Temperature**	0.06	0.75	0.13	0	1.00	0.39
19.	**Ejector_pos_last**	0.39	0	0.5	1	0.00	0.38
20.	**Cooling_time**	0.13	0.63	0.15	0	0.40	0.26
21.	**Screw_speed**	0.34	0.08	0.39	0	0.44	0.25
22.	**Speed_max**	0.34	0.04	0.39	1	0.51	0.46
23.	**Injection_Speed**	0.02	0.71	0	0	0.64	0.27
24.	**Injection_work**	0.31	0.33	0.06	0	0.00	0.14
25.	**Switchov_vol**	0	0.29	0.09	0	0.00	0.08
	MLP R^2 score	-0.21	-0.7	0.49	-0.86	0.76	
	MLP RMSE	9.07	9.3	6.22	10.62	8.1	
	Correlation coef.	0.73	0.66	0.82	0.52	0.87	

For the sequential dataset use of all the parameters in the dataset also does not allow to build a good model, as its characteristics are: R^2 = 0.25, RMSE = 11.5, correlation coefficient = 0.74. RReliefF, however, once again allows to improve it, this way R^2 becomes 0.52, RMSE = 9.87 and the correlation coefficient is increased to 0.79. This might not seem like a high-quality model, however, it is better than the one created with no prior feature selection.

Table 7.22. Feature selection for experiment 1 tensile strain at break target variable, sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Cushion_smallest	1	0.42	0.94	1	0.29	0.73
2.	Cushion_average	0.99	0.37	0.92	1	0.30	0.72
3.	Nozzle_tempr_z2_average	0.82	0.21	0.95	1	0.25	0.65
4.	Injection_time	0.61	0.5	0.41	1	0.65	0.63
5.	Cushion_after_hold_pres	0.92	0.54	1	0	0.34	0.56
6.	Holding_pressure	0.95	0.33	0.93	0	0.48	0.54
7.	Spec_pres_switchov	0.39	0.79	0.32	0	0.87	0.47
8.	Heating_cyl1_z1_set	0.82	0.17	0.97	0	0.25	0.44
9.	Speed_max	0.39	0.04	0.39	1	0.34	0.43
10.	Tool_Temperature	0	0.92	0.12	0	1.00	0.41
11.	Screw_speed_max	0.38	1	0.35	0	0.27	0.40
12.	Flow_no_plast	0.45	0.58	0.51	0	0.40	0.39
13.	Ejector_pos_last	0.36	0	0.48	1	0.06	0.38
14.	Plast_time	0.31	0.29	0.49	0	0.80	0.38
15.	Closing_force	0.32	0.96	0.6	0	0.00	0.38
16.	Backpressure	0.21	0.71	0.38	0	0.50	0.36
17.	Cooling_time	0.18	0.75	0.13	0	0.47	0.31
18.	Holding_pres_time	0.27	0.46	0.07	0	0.63	0.29
19.	Last_cycle_time	0.18	0.88	0.05	0	0.26	0.27
20.	Injection_Speed	0.08	0.67	0	0	0.60	0.27
21.	Cooling_time_last	0.15	0.63	0.09	0	0.47	0.27
22.	Injection_work	0.27	0.83	0.08	0	0.12	0.26
23.	Screw_speed	0.39	0.13	0.39	0	0.36	0.25
24.	Clamp_force_switchov	0.35	0.25	0.53	0	0.00	0.23
25.	Switchov_vol	0.06	0.08	0.12	0	0.00	0.05
	MLP R^2 score	0.15	0.41	0.42	0.28	0.52	
	MLP RMSE	10.74	11.23	10.15	11.71	9.87	
	Correlation coef.	0.81	0.75	0.79	0.75	0.79	

As a result of the previous analysis summed up in Table 7.23, it is easy to see that application of RReliefF can improve the tensile strain at break prediction models. Therefore, it is decided that this FS method will be used to select the most influential parameter for creation of the corresponding predictive models.

Table 7.23. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (tensile strain at break target variable)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Experiment 1, parallel dataset, tensile strain at break1 target variable	**MLP R^2 score**	0.55	0.69	0.63	0.58	0.12	0.56
	MLP RMSE	10.18	8.09	9.19	9.33	17.99	9.08
	Correlation coef.	0.81	0.89	0.84	0.84	0.45	0.87
Experiment 1, parallel dataset, tensile strain at break2 target variable	**MLP R2 score**	0.19	-0.21	-0.7	0.49	-0.86	0.76
	MLP RMSE	7.6	9.07	9.3	6.22	10.62	8.1
	Correlation coef.	0.67	0.73	0.66	0.82	0.52	0.87
Experiment 1, sequential dataset, tensile strain at break target variable	**MLP R2 score**	0.25	0.15	0.41	0.42	0.28	0.52
	MLP RMSE	11.5	10.74	11.23	10.15	11.71	9.87
	Correlation coef.	0.74	0.81	0.75	0.79	0.75	0.79
Average	**MLP R^2 score, average**	0.33	0.21	0.11	0.50	-0.15	**0.61**
	MLP RMSE, average	9.76	9.30	9.91	8.57	13.44	**9.02**
	Correlation coef. average	0.74	0.81	0.75	0.82	0.57	**0.84**

Joined dataset

Similarly to the other focus variables, joined datasets are used to create the predictive models for tensile strain at break. The procedure and threshold for the parameters' elimination are the same. The random state for data partitioning is set to 0 for utilization of 5-folds cross-validation, while the MLP setting is as specified below:

```
model = MLPRegressor(solver = 'lbfgs', activation = 'logistic', alpha = 1e-5, learning_rate_init = 0.3, hidden_layer_sizes = (X.shape[1]+1)/2, random_state = 2, momentum = 0.2)
```

Tables 7.24 – 7.26 contain scores that the various feature selection methods give to the considered parameters for tensile strain at break1, tensile strain at break2 and parallel joined dataset (target variables that represent characteristics of the first and second specimens produced during the same production run), as well as tensile strain at break (includes tensile strain at break1 and 2 merged sequentially) and sequential joined dataset. When no FS is applied for the tensile strain at break1, the trained MLP has the following performance characteristics: R^2 = 0.65, RMSE = 8.77, correlation coefficient = 0.84. In case of tensile strain at break2: R^2 = 0.66, RMSE = 7.85, correlation coefficient = 0.84.

Table 7.24. Feature selection for tensile strain at break1 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	**Switchov_vol**	1	0	1	1	1	0.8
2.	**Heating_cyl1_z1_set**	0.61	0.84	0.5	1	0.01	0.59
3.	**Cushion_smallest**	0.79	0.32	0.64	1	0.05	0.56
4.	**Cushion_after_hold_pres**	0.71	0.47	0.58	0	0.02	0.36

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
5.	Holding_pressure	0.65	0.63	0.43	0	0	0.34
6.	Spec_pres_switchov	0.5	0.74	0.27	0	0.08	0.32
7.	Material	0.22	0.16	0.37	0	0.82	0.32
8.	Ejector_pos_last	0.65	0.05	0.69	0	0.11	0.3
9.	Flow_no_plast	0.29	0.95	0.2	0	0.054	0.3
10.	Speed_max	0.23	0.11	0.14	1	0	0.3
11.	Holding_pres_time	0	1	0.09	0	0	0.22
12.	Last_cycle_time	0.06	0.89	0.04	0	0.05	0.21
13.	Backpressure	0.12	0.79	0.12	0	0	0.21
14.	Plast_time	0.23	0.42	0.16	0	0.03	0.17
15.	Injection_Speed	0.11	0.68	0.05	0	0	0.17
16.	Injection_time	0.37	0.21	0.12	0	0.13	0.16
17.	Cooling_time	0.09	0.58	0.01	0	0	0.14
18.	Cooling_time_last	0.03	0.53	0.04	0	0	0.12
19.	Screw_speed	0.2	0.26	0.12	0	0	0.12
20.	Injection_work	0.12	0.37	0	0	0	0.1
	MLP R^2 score	0.69	0.27	0.62	0.51	0.45	
	MLP RMSE	8.1	12.72	8.96	10.68	11.36	
	Correlation coef.	0.87	0.61	0.82	0.76	0.69	

Table 7.25. Feature selection for tensile strain at break2 target variable, joined parallel dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Switchov_vol	1	0.05	1	1	1.00	0.81
2.	Material	0.47	0.21	0.47	1	0.72	0.57
3.	Cushion_smallest	0.54	0.32	0.53	1	0.01	0.48
4.	Heating_cyl1_z1_set	0.47	0.47	0.45	1	0.00	0.48
5.	Cushion_after_hold_pres	0.51	0.58	0.49	0	0.00	0.32
6.	Last_cycle_time	0.09	1	0.06	0	0.07	0.24
7.	Ejector_pos_last	0.5	0	0.63	0	0.09	0.24
8.	Spec_pres_switchov	0.24	0.74	0.18	0	0.04	0.24
9.	Holding_pressure	0.4	0.42	0.35	0	0.00	0.23
10.	Flow_no_plast	0.13	0.79	0.13	0	0.03	0.22
11.	Backpressure	0.06	0.89	0.11	0	0.00	0.21
12.	Cooling_time	0.03	0.95	0	0	0.00	0.20
13.	Holding_pres_time	0.06	0.84	0.07	0	0.00	0.19
14.	Injection_time	0.15	0.63	0.11	0	0.05	0.19
15.	Injection_Speed	0	0.68	0.04	0	0.00	0.14
16.	Plast_time	0.16	0.37	0.1	0	0.04	0.13
17.	Cooling_time_last	0.02	0.53	0.03	0	0.00	0.12
18.	Injection_work	0.16	0.26	0	0	0.00	0.08

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
19.	Speed_max	0.11	0.11	0.1	0	0.00	0.06
20.	Screw_speed	0.07	0.16	0.06	0	0.00	0.06
	MLP R2 score	0.68	0.48	0.64	0.65	0.53	
	MLP RMSE	7.51	9.46	8.02	8.08	9.3	
	Correlation coef.	0.85	0.75	0.83	0.81	0.73	

In case of the sequential dataset and no feature selection, the trained MLP model has R^2 = 0.84, RMSE = 13.02, correlation coefficient = 0.92. In all cases except for the tensile strain at break1 with the parallel dataset, application of FS to the data from the joined dataset only decreases the resulting MLP model performance.

Table 7.26. Feature selection for tensile strain at break target variable, joined sequential dataset

#	Parameter name	Pearson	RFE	Spearman	CFS	RReliefF	Mean
1.	Part type	1	0	1	1	0.00	0.60
2.	Spec_pres_switchov	0.81	0.89	0.74	0	0.12	0.51
3.	Injection_time	0.44	0.11	0.54	1	0.20	0.46
4.	Holding_pres_time	0.64	1	0.49	0	0.14	0.45
5.	Cooling_time	0.92	0.56	0.51	0	0.16	0.43
6.	Cooling_time_last	0.92	0.5	0.46	0	0.17	0.41
7.	Last_cycle_time	0.71	0.83	0.35	0	0.15	0.41
8.	Plast_time	0.48	0.61	0.6	0	0.07	0.35
9.	Material	0.02	0.28	0.27	0	1.00	0.31
10.	Flow_no_plast	0.15	0.94	0.32	0	0.15	0.31
11.	Switchov_vol	0.98	0.06	0.35	0	0.07	0.29
12.	Cushion_smallest	0.7	0.39	0	0	0.13	0.24
13.	Cushion_after_hold_pres	0.6	0.44	0	0	0.13	0.23
14.	Backpressure	0.04	0.72	0.1	0	0.22	0.22
15.	Injection_Speed	0.04	0.78	0.01	0	0.20	0.21
16.	Holding_pressure	0	0.67	0.22	0	0.11	0.20
17.	Injection_work	0.19	0.33	0.34	0	0.06	0.18
18.	Speed_max	0.13	0.17	0.27	0	0.09	0.13
19.	Screw_speed	0.11	0.22	0.21	0	0.11	0.13
	MLP R2 score	0.74	0.81	0.8	0.68	0.48	
	MLP RMSE	16.65	14.09	14.5	18.64	23.62	
	Correlation coef.	0.87	0.91	0.91	0.89	0.72	

As it is seen from the Table 7.27 it is not worth applying any of the feature selection techniques to the joined datasets as the highest average model score is obtained with no feature selection application. Therefore, no feature selection will be used prior to the models' development for tensile strain at break variable and joined datasets. It is also

interesting to note that the material parameter for the parallel dataset has higher FS algorithms scores in comparison to the score for the sequential dataset. In the sequential dataset, however, the part type parameter is acknowledged by the high FS from Pearson's and Spearman's correlation and CFS methods.

Table 7.27. MLP average accuracy measures (R^2, RMSE, correlation coefficient) for feature selection with different methods (tensile strain at break target variable, joined datasets)

	Accuracy measure	No FS	Pearson	RFE	Spearman	CFS	RReliefF
Tensile strain at break1 target variable, joined parallel dataset	**MLP R^2 score**	0.65	0.69	0.27	0.62	0.51	0.45
	MLP RMSE	8.77	8.1	12.72	8.96	10.68	11.36
	Correlation coef.	0.84	0.87	0.61	0.82	0.76	0.69
Tensile strain at break2 target variable, joined parallel dataset	**MLP R2 score**	0.66	0.68	0.48	0.64	0.65	0.53
	MLP RMSE	7.85	7.51	9.46	8.02	8.08	9.3
	Correlation coef.	0.84	0.85	0.75	0.83	0.81	0.73
Tensile strain at break target variable, joined parallel dataset	**MLP R2 score**	0.84	0.74	0.81	0.8	0.68	0.48
	MLP RMSE	13.02	16.65	14.09	14.5	18.64	23.62
	Correlation coef.	0.92	0.87	0.91	0.91	0.89	0.72
Average	**MLP R^2 score, average**	**0.72**	0.70	0.52	0.69	0.61	0.49
	MLP RMSE average	**9.88**	10.75	12.09	10.49	12.47	14.76
	Correlation coef. average	**0.87**	0.86	0.76	0.85	0.82	0.71

7.4 Predictive models development

After performing data exploration for the mechanical properties target variables in order to gain a better understanding of data at hand, data preprocessing, where constant features and outliers were removed, as well as feature selection, it is now possible to train predictive models for the target variables of interest. The procedure followed in this step is similar to that used for creation of prediction models for the dimensional target variables. MLP, Decision Tree Regressor, kNN, GBR, AdaBoost and Random Forest methods are used to train the models. The models are trained for the separate experiments datasets, as well as for the joined datasets.

The preprocessing step led to elimination of outliers from the datasets at hand. Therefore, they contain smaller number of samples than those used for development of the dimensional properties models. Results of the feature selection are used for creation of MLP and kNN models (for the separate experiments datasets), as they do not have internal mechanisms for filtering out the irrelevant features. When training models using the joined datasets, no prior FS is used based on the results from Section 7.3.

The datasets are divided into 70% training and 30% testing subsets, where 5-folds cross-validation is performed on the training set. Models with the highest performance characteristics are searched for using grid search on the corresponding sets of the

hyperparameters. The same hyperparameter sets are used for training the mechanical properties models as for prediction of the dimensional characteristics.

As a reminder, the list of the hyperparameters and their tested values is presented once more. For the MLP the following set of hyperparameters was considered:

- hidden layer sizes : [10, 15, 20, 25, 30],
- activation function : ‘relu’ , ‘logistic’,
- solver : ‘lbfgs’, ‘sgd’,
- alpha (L2 penalty parameter): [0.0001, 0.05],
- learning rate init: [0.001, 0.01, 0.05, 0.1, 0.3].

For Decision Tree the criterion and maximum tree depth parameters were evaluated:

- criterion: ‘mse’, ‘friedman_mse’, ‘mae’,
- maximum tree depth: [5, 7, 10, 12, 15].

For kNN, in its turn, the following hyperparameters were tuned:

- weights: ‘uniform’, ‘distance’,
- number of neighbors: [2, 3, 4, 5, 6, 7].

For Gradient Boost Regressor a different set of hyperparameters is considered:

- loss: ‘ls’, ‘lad’, ‘huber’, ‘quantile’,
- learning rate: [0.001, 0.005, 0.01, 0.05, 0.1],
- number of estimators: [50, 100, 150, 200, 250, 300].

In case of AdaBoost, the same hyperparameters as for GBR are tuned, however, this algorithm uses different loss function types:

- loss: ‘linear’, ‘square’, ‘exponential’,
- learning rate: [0.001, 0.005, 0.01, 0.05, 0.1, 1],
- number of estimators': [50, 100, 150, 200, 250, 300].

The last, but not the least is Random Forest, here the number of estimators, max features and criterion hyperparameters were varied:

- number of estimators: [50, 100, 150, 200, 250, 300],
- max features: ‘auto’, ‘sqrt’, ‘log2’,
- criterion: ‘mse’, ‘mae’.

The same way as while tuning hyperparameters for the dimensional properties, the overall models’ training time significantly varies for the different algorithms. With the data amount at hand, it takes the least time to perform grid search for kNN and Decision Tree Regressor models (up to 30 seconds), followed by MLP (1-3 minutes) and GBR, AdaBoost and Random Forest (up to 10 minutes).

7.4.1 Young's modulus target variable

Similarly to Chapter 6 this section includes results of the models' hyperparameters tuning for the Young's modulus (tensile modulus) target variable. There is no nominal value that can be identified for the tensile modulus, however, the data at hand includes the minimum value of Young's modulus of about 750 MPa, while the largest after the outlier removal is close to 1200 MPa. The datasets for the separate experiments, as well as joined datasets are utilized. Their sizes are slightly smaller than those used for development of models for dimensional target variables, as certain outliers were removed, as described in Section 7.2. The threshold for including or excluding the parameters scored by RReliefF FS is set to 0.2, unless stated otherwise. The FS is used only for MLP and kNN models (separate experiments datasets), while the rest of the algorithms receive a full set of parameters as an input.

Separate experiments datasets

Experiment 1

Tables 7.28 and 7.29 show the best perfroming combinations of hyperparameters for the parallel and sequential datasets correspondingly. As it can be seen from the tables, both for parallel and sequential datasets, the MLP scores are higher than those obtained without the parameter tuning in Section 7.3. The best performance, however, is shown by kNN and Random Forest models, and not by MLP. In case of the sequential dataset, Random Forest model performance stands out among the rest of the models.

Table 7.28. Results of predictive models hyperparameter optimization for Young's modulus1 and Young's modulus2, parallel dataset, experiment 1

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	10	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	7	-
loss	-	-	-	-
criterion	-	mae	-	mae
max_depth	-	5	-	-
n_estimators	-	-	-	200
max_features	-	-	-	sqrt
R^2 train set	0.62	0.91	0.99	0.95
RMSE train set	60.86	30.14	0.01	21.51
Correl. coef. train set	0.8	0.95	0.99	0.98
R^2 test set	0.59	0.57	0.71	0.7
RMSE test set	78.67	80.39	66.06	66.83
Correl. coef. test set	0.8	0.78	0.85	0.85

Table 7.29. Results of predictive models hyperparameter optimization for Young's modulus, sequential dataset, experiment 1

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	10	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.01	0.05	-
weights	-	-	uniform	-	-	-
number of neighbors	-	-	3	-	-	-
loss	-	-	-	huber	exponential	-
criterion	-	mse	-	-	-	mae
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	300	50	150
max_features	-	-	-	-	-	sqrt
R^2 train set	0.66	0.86	0.86	0.87	0.81	0.91
RMSE train set	60.82	39.74	39.72	38.31	45.89	31.81
Correl. coef. train set	0.81	0.93	0.93	0.94	0.9	0.95
R^2 test set	0.67	0.6	0.74	0.77	0.71	0.78
RMSE test set	64.39	70.31	57.08	53.8	60.46	52.56
Correl. coef. test set	0.82	0.8	0.86	0.88	0.84	0.88

Experiment 2

Unlike results for the experiment 1 model training, all the models for experiment 2 have unacceptably low performance, especially on the test set. This means that their generalization abilities are not high enough and more data needs to be obtained before they can be used for the assistance in the decision-making process. This dataset consists of only 65 samples, which does not seem to be enough to create a meaningful model for prediction of the Young's modulus.

Table 7.30. Results of predictive models hyperparameter optimization for Young's modulus, experiment 2

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	30	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.005	0.1	-
weights	-	-	uniform	-	-	-
number of neighbors	-	-	7	-	-	-
loss	-	-	-	ls	linear	-
criterion	-	Friedman_mse	-	-	-	mse
max_depth	-	10	-	-	-	-
n_estimators	-	-	-	150	100	100
max_features	-	-	-	-	-	sqrt
R^2 train set	0.51	0.4	0.28	0.57	0.9	0.88
RMSE train set	7.64	8.34	9.25	7.13	3.43	3.75
Correl. coef. train set	0.71	0.67	0.55	0.9	0.96	0.98

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
R^2 test set	0.25	0.07	0.09	0.2	0.37	0.16
RMSE test set	12.73	13.75	11	11.91	12.34	11.22
Correl. coef. test set	0.21	0.05	0.35	0.16	0.05	0.21

Experiment 3

For experiment 3, the feature selection threshold was decreased from 0.2 to 0.05, as the 0.2 threshold was too high in this case. However, even after doing so, the obtained models for parallel and sequential datasets had unacceptably low performance, in some cases both on the training and test sets. There might be several reasons for this. The first reason is necessity to collect more relevant data, and the second one is general complications related to prediction of mechanical properties such as Young's modulus for plastic materials, due to the materials' nature, dependence of the mechanical properties on the carbon chains orientation, etc.

Table 7.31. Results of predictive models hyperparameter optimization for Young's modulus1 and Young's modulus2, parallel dataset, experiment 3

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	30	-	-	-
solver	lbfgs	-	-	-
alpha	0.0001	-	-	-
learning rate	0.001	-	-	-
weights	-	-	uniform	-
number of neighbors	-	-	5	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mse
max_depth	-	5	-	-
n_estimators	-	-	-	50
max_features	-	-	-	sqrt
R^2 train set	0.12	0.87	0.5	0.89
RMSE train set	54.13	19.28	38.27	17.66
Correl. coef. train set	0.22	0.93	0.71	0.96
R^2 test set	0.08	0.08	0.05	0.2
RMSE test set	51.11	59.79	59.04	43.93
Correl. coef. test set	0.04	0.26	0.09	0.5

Table 7.32. Results of predictive models hyperparameter optimization for Young's modulus, sequential dataset, experiment 3

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	10	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.01	0.1	-
weights	-	-	uniform	-	-	-

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
number of neighbors	-	-	2	-	-	-
loss	-	-	-	huber	linear	-
criterion	-	mae	-	-	-	mae
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	200	50	300
max_features	-	-	-	-	-	sqrt
R^2 train set	0.48	0.75	0.8	0.74	0.71	0.84
RMSE train set	38.08	26.43	23.47	26.94	28.58	21.21
Correl. coef. train set	0.69	0.87	0.9	0.88	0.84	0.92
R^2 test set	0.33	0.01	0.34	0.37	0.3	0.36
RMSE test set	43.69	53.26	43.65	42.42	44.89	42.8
Correl. coef. test set	0.6	0.45	0.59	0.61	0.6	0.61

Experiment 4

The same result as for experiments 2 and 3 is observed for experiment 4. The models' performance is low, especially on the test set. If the model has high performance on the train set and a low one on the test set it means that the model is not capable of making meaningful predictions on the previously unseen data and, therefore, cannot be used in its current state. To increase the models' quality, it is necessary to collect more relevant data and if possible add material characteristics (such as viscosity) of the material batches used for the focus parts production.

Table 7.33. Results of predictive models hyperparameter optimization for Young's modulus1 and Young's modulus2, parallel dataset, experiment 4.

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	relu	-	-	-
hidden layer neurons	20	-	-	-
solver	lbfgs	-	-	-
alpha	0.0001	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	7	-
loss	-	-	-	-
criterion	-	12	-	mse
max_depth	-	mae	-	-
n_estimators	-	-	-	250
max_features	-	-	-	log2
R^2 train set	0.43	0.78	0.99	0.91
RMSE train set	39.05	0.05	0.01	15.62
Correl. coef. train set	0.66	0.82	0.99	0.98
R^2 test set	0.06	0.04	0.19	0.38
RMSE test set	73.7	52.32	46.19	40.58
Correl. coef. test set	0.12	0.46	0.42	0.65

Table 7.34. Results of predictive models hyperparameter optimization for Young's modulus, sequential dataset, experiment 4

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	15	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.01	0.005	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	7	-	-	-
loss	-	-	-	lad	square	-
criterion	-	mae	-	-	-	mse
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	250	50	50
max_features	-	-	-	-	-	log2
R^2 train set	0.45	0.76	0.85	0.61	0.67	0.79
RMSE train set	38.98	25.51	20.25	32.87	30.34	23.96
Correl. coef. train set	0.67	0.88	0.92	0.83	0.83	0.9
R^2 test set	0.36	0.33	0.28	0.4	0.41	0.45
RMSE test set	45.05	46.09	48.03	43.76	43.59	42.09
Correl. coef. test set	0.61	0.62	0.56	0.66	0.66	0.68

Joined datasets

When the data from three experiments (joined parallel dataset) and data from four experiments (joined sequential dataset) is added to create the joined datasets and predictive models are created thereafter, the obtained models' performance is significantly higher. The models' scores increase as more data is used to train them. In addition, simplified material and part type parameters are added, which also seem to be helpful. RMSE value for the Young's modulus is much higher than for the dimensional target variables. This is due to tensile modulus having a significantly larger range of values (from 750 to 1400 MPa). As a result, Random Forest has the best performing model for the parallel dataset, while GBR has the highest performance scores for the sequential dataset, with Random Forest having the second best score for both train and test sets. These models, unlike the ones for the separate datasets for experiments 2-4, can be taken into consideration when selecting parameters for production of dogbone specimens.

Table 7.35. Results of predictive models hyperparameter optimization for Young's modulus1 and Young's modulus2, parallel joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	relu	-	-	-
hidden layer neurons	10	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	6	-

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
loss	-	-	-	-
criterion	-	mae	-	mse
max_depth	-	5	-	-
n_estimators	-	-	-	150
max_features	-	-	-	auto
R^2 train set	0.76	0.94	0.99	0.99
RMSE train set	81.57	41.52	0.01	19.16
Correl. coef. train set	0.88	0.97	0.99	0.99
R^2 test set	0.67	0.86	0.61	0.9
RMSE test set	87.62	56.24	94.6	49
Correl. coef. test set	0.83	0.93	0.79	0.95

Table 7.36. Results of predictive models hyperparameter optimization for Young's modulus, sequential joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	25	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.0001	-	-	-	-	-
learning rate	0.001	-	-	0.05	0.1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	6	-	-	-
loss	-	-	-	lad	linear	-
criterion	-	mae	-	-	-	mse
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	300	100	200
max_features	-	-	-	-	-	auto
R^2 train set	0.65	0.93	0.98	0.95	0.91	0.97
RMSE train set	92.66	40.6	22.55	36.03	46.07	25.35
Correl. coef. train set	0.81	0.97	0.99	0.97	0.96	0.99
R^2 test set	0.62	0.89	0.83	0.91	0.89	0.91
RMSE test set	101.43	54.56	69.01	48.75	54.52	50.03
Correl. coef. test set	0.8	0.95	0.91	0.96	0.95	0.96

7.4.2 Tensile strength target variable

Tensile strength is another target variable for which the predictive models are created and tuned using grid search. The procedure does not differ from that described for the dimensional target variable and tensile modulus. Tensile strength values obtained from the experiments vary from 21.5 to 34 MPa. Due to having a larger range of values in comparison to those for the dimensional properties, higher values of the RMSE are to be expected.

Separate experiments dataset

Experiment 1

The most suitable FS method was selected by training MLP models with and without the different FS methods. For the train parallel dataset R^2 for tensile strength1 (characteristics of the first specimen out of two that are produced during the same machine run) was equal to 0.81 and 0.87 for tensile strength2 (characteristics of the second specimen out of two that are produced during the same machine run) with RReliefF feature selection method. For the train sequential dataset, the obtained R^2 was 0.88. If these values are compared to the ones presented in Tables 7.37 and 7.38 after applying the grid search, we can see that they are outperformed. For the parallel dataset, the best performance is shown by Random Forest with the R^2 value of 0.98 and 0.91 for train and test sets correspondingly. For the sequential dataset, all the models except MLP have comparatively high scores for both train and test set, with kNN slightly outperforming the rest.

Table 7.37. Results of predictive models hyperparameter optimization for tensile strength1 and tensile strength2, parallel dataset, experiment 1

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	10	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	7	-
loss	-	-	-	-
criterion	-	mse	-	mse
max_depth	-	7	-	-
n_estimators	-	-	-	50
max_features	-	-	-	sqrt
R^2 train set	0.8	0.99	0.99	0.98
RMSE train set	0.57	0.11	0.01	0.16
Correl. coef. train set	0.9	0.98	0.99	0.98
R^2 test set	0.81	0.88	0.89	0.91
RMSE test set	0.57	0.46	0.43	0.39
Correl. coef. test set	0.9	0.94	0.94	0.96

Table 7.38. Results of predictive models hyperparameter optimization for tensile strength, sequential dataset, experiment 1

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	15	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	0.1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	5	-	-	-
loss	-	-	-	ls	exponential	-
criterion	-	friedman_mse	-	-	-	mae
max_depth	-	15	-	-	-	-
n_estimators	-	-	-	50	100	250
max_features	-	-	-	-	-	auto

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
R^2 train set	0.89	0.97	0.97	0.96	0.92	0.96
RMSE train set	0.43	0.23	0.23	0.26	0.36	0.25
Correl. coef. train set	0.94	0.98	0.98	0.98	0.96	0.98
R^2 test set	0.89	0.92	0.93	0.93	0.91	0.92
RMSE test set	0.42	0.37	0.35	0.35	0.39	0.36
Correl. coef. test set	0.95	0.96	0.96	0.96	0.96	0.96

Experiment 2

For the experiment 2 data, Decision Tree, kNN, GBR, AdaBoost and Random Forest models have rather high scores on the train set. On the test set, however, none of the models show acceptable performance. The main reason for this is the same as described for the Young's modulus target variable – a small number of samples in the dataset. Therefore, before obtaining a meaningful predictive model for tensile strength of 15 mm thick dogbones, more relevant data needs to be obtained and incorporated in the model creation process.

Table 7.39. Results of predictive models hyperparameter optimization for tensile strength, experiment 2

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	25	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.05	1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	6	-	-	-
loss	-	-	-	lad	exponential	-
criterion	-	mse	-	-	-	mae
max_depth	-	7	-	-	-	-
n_estimators	-	-	-	100	250	300
max_features	-	-	-	-	-	log2
R^2 train set	0.67	0.99	0.99	0.9	0.96	0.93
RMSE train set	0.05	0.01	0.01	0.03	0.02	0.02
Correl. coef. train set	0.82	0.99	0.99	0.97	0.98	0.98
R^2 test set	0.2	0.31	0.21	0.44	0.41	0.36
RMSE test set	0.09	0.08	0.09	0.07	0.08	0.08
Correl. coef. test set	0.576	0.71	0.66	0.81	0.77	0.74

Experiment 3

Unlike in case with tensile modulus models trained on the experiment 3 data, the models for tensile strength have a slightly better quality. When it comes to the parallel dataset, the only model that has good performance characteristics on both train and test sets is Random Forest with R^2 of 0.97 and 0.81 correspondingly. Decision Tree and kNN, at the same time, have extremely good scores on the training set, but low scores on the test one,

this means that overfitting is most probably present. For the sequential dataset the performance of AdaBoost and Random Forest is almost equally good. The mentioned models can be used as a starting point for development of even more robust models for the decision-making support.

Table 7.40. Results of predictive models hyperparameter optimization for tensile strength1 and tensile strength2, parallel dataset, experiment 3

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	20	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	6	-
loss	-	-	-	-
criterion	-	Friedman_mse	-	mae
max_depth	-	10	-	-
n_estimators	-	-	-	50
max_features	-	-	-	sqrt
R^2 train set	0.78	0.99	0.99	0.97
RMSE train set	0.82	0.01	0.01	0.3
Correl. coef. train set	0.89	0.99	0.99	0.99
R^2 test set	0.6	0.45	0.65	0.81
RMSE test set	0.97	1.14	0.91	0.66
Correl. coef. test set	0.79	0.74	0.84	0.91

Table 7.41. Results of predictive models hyperparameter optimization for tensile strength, sequential dataset, experiment 3

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	15	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.05	1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	3	-	-	-
loss	-	-	-	lad	square	-
criterion	-	friedman_mse	-	-	-	mse
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	100	250	100
max_features	-	-	-	-	-	sqrt
R^2 train set	0.54	0.97	0.98	0.95	0.96	0.97
RMSE train set	1.14	0.27	0.21	0.39	0.33	0.27
Correl. coef. train set	0.73	0.99	0.99	0.98	0.98	0.99
R^2 test set	0.38	0.83	0.76	0.81	0.84	0.84
RMSE test set	1.38	0.71	0.87	0.76	0.71	0.7
Correl. coef. test set	0.62	0.92	0.89	0.92	0.92	0.93

Experiment 4

Unlike models based on the data for experiments 2 and 3, the models for the experiment 4 data have relatively higher performance characteristics both for the parallel and sequential datasets. For both cases MLP and kNN have a noticeably lower R^2 on the test sets. Random Forest and Decision Tree algorithms have the highest performing models for the parallel dataset, while for the sequential one Decision Tree, GBR and Random Forest have the highest scores. As a result, these models can be used for decision support and reference when producing the focus parts from RePro recycled HDPE.

Table 7.42. Results of predictive models hyperparameter optimization for tensile strength1 and tensile strength2, parallel dataset, experiment 4

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	20	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	5	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.91	0.98	0.99	0.97
RMSE train set	0.29	0.15	0.01	0.17
Correl. coef. train set	0.96	0.99	0.99	0.99
R^2 test set	0.88	0.89	0.71	0.83
RMSE test set	0.39	0.37	0.62	0.47
Correl. coef. test set	0.94	0.95	0.87	0.94

Table 7.43. Results of predictive models hyperparameter optimization for tensile strength, sequential dataset, experiment 4

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	logistic	-	-	-	-	-
hidden layer neurons	15	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.1	1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	2	-	-	-
loss	-	-	-	lad	linear	-
criterion	-	mse	-	-	-	mae
max_depth	-	12	-	-	-	-
n_estimators	-	-	-	300	150	250
max_features	-	-	-	-	-	log2
R^2 train set	0.89	0.99	0.99	0.98	0.95	0.98
RMSE train set	0.35	0.12	0.12	0.15	0.23	0.16

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
Correl. coef. train set	0.95	0.99	0.99	0.99	0.98	0.99
R^2 test set	0.88	0.94	0.81	0.94	0.93	0.94
RMSE test set	0.36	0.26	0.52	0.26	0.28	0.25
Correl. coef. test set	0.94	0.97	0.91	0.97	0.97	0.98

Joined dataset

Tables 7.44 and 7.45 show results of the hyperparameter optimization using greed search for the parallel and sequential joined datasets and tensile strength target variable. The observed situation is similar to that with the Young's modulus models' performance on the joined datasets. Decision Tree and Random Forest models are able to predict tensile strength with $R^2 = 0.99$ on the train set and $R^2 = 0.92$ and 0.94 for the corresponding methods on the test set of the parallel dataset. For the sequential dataset, on the other hand, GBR and Random Forest outperform other methods on the test set. This testifies that both models are able to generalize on the previously unseen data and therefore can have a value for the real industrial environment. Unlike the models for the separate experiments datasets, they are more useful, as they are trained on larger datasets and the samples they are trained on include data about dogbones with two thicknesses and three different materials.

Table 7.44. Results of predictive models hyperparameter optimization for tensile strength1 and tensile strength2, parallel joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	relu	-	-	-
hidden layer neurons	15	-	-	-
solver	lbfgs	-	-	-
alpha	0.0001	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mse
max_depth	-	12	-	-
n_estimators	-	-	-	250
max_features	-	-	-	auto
R^2 train set	0.77	0.99	0.99	0.99
RMSE train set	1.2	0.01	0.01	0.19
Correl. coef. train set	0.88	0.99	0.99	0.99
R^2 test set	0.69	0.92	0.64	0.94
RMSE test set	1.52	0.79	1.64	0.68
Correl. coef. test set	0.85	0.96	0.82	0.97

Table 7.45. Results of predictive models hyperparameter optimization for tensile strength, sequential joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	20	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.0001	-	-	-	-	-
learning rate	0.001	-	-	0.05	0.1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	7	-	-	-
loss	-	-	-	huber	exponential	-
criterion	-	mse	-	-	-	mae
max_depth	-	12	-	-	-	-
n_estimators	-	-	-	300	300	100
max_features	-	-	-	-	-	auto
R^2 train set	0.84	0.99	0.99	0.99	0.95	0.99
RMSE train set	1.2	0.2	0.21	0.28	0.68	0.24
Correl. coef. train set	0.92	0.99	0.99	0.990	0.98	0.99
R^2 test set	0.81	0.96	0.95	0.97	0.94	0.97
RMSE test set	1.25	0.59	0.64	0.46	0.71	0.48
Correl. coef. test set	0.9	0.98	0.97	0.99	0.97	0.99

7.4.3 Tensile strain at break target variable

In addition to tensile modulus and tensile strength, tensile strain at break is also a focus variable of interest. Therefore, this section presents results of the hyperparameter optimization with grid search, but for tensile strain at break target variable. The values of this variable in the data at hand vary from 5 to 250 %. Therefore, the values of the RMSE will be higher than in case of the dimensional target variables, as their range is significantly smaller. The procedure followed in this case is the same as for the other focus variables. RReliefF is used as a FS method to retain the most relevant variables before training the MLP and kNN prediction models on the separate experiments datasets. RRelieF is selected based on the results presented in the Section 7.3. The threshold for including/ excluding a process parameter into the model is set to 0.2. At the same time, all the datasets are divided into 70% train and 30% test datasets. 5-folds cross-validation is performed on the training set to avoid overfitting and confirm the models' ability to perform meaningful predictions on new data.

Separate experiments dataset

Experiment 1

For experiment 1 data, the models trained on the parallel dataset have a better performance in comparison to those trained on the sequential dataset. At the same time, MLP and kNN models have a better performance in comparison to the untuned MLP models that were used to select the feature selection method for training the tensile strain at break models. The untuned models have R^2 = 0.56 and 0.76 for the tensile strain at break1 and 2 target variables on train set, while the ones obtained after the grid search

have an R^2 up to 0.99 on the train set. When it comes to the sequential dataset, the untuned MLP model has R^2 = 0.52, the tuned MLP model has a value of 0.64 on the train set, while Random Forest R^2 reaches 0.92.

Table 7.46. Results of predictive models hyperparameter optimization for tensile strain at break1 and tensile strain at break2, parallel dataset, experiment 1

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	relu	-	-	-
hidden layer neurons	30	-	-	-
solver	lbfgs	-	-	-
alpha	0.0001	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	5	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mae
max_depth	-	5	-	-
n_estimators	-	-	-	100
max_features	-	-	-	sqrt
R^2 train set	0.64	0.97	0.99	0.96
RMSE train set	8.57	2.5	0.01	2.77
Correl. coef. train set	0.79	0.98	0.99	0.98
R^2 test set	0.57	0.66	0.62	0.71
RMSE test set	11.02	9.78	10.38	9.05
Correl. coef. test set	0.72	0.83	0.77	0.83

Table 7.47. Results of predictive models hyperparameter optimization for tensile strain at break, sequential dataset, experiment 1

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	10	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.0001	-	-	-	-	-
learning rate	0.001	-	-	0.05	0.1	-
weights	-	-	uniform	-	-	-
number of neighbors	-	-	3	-	-	-
loss	-	-	-	ls	square	-
criterion	-	friedman_mse	-	-	-	mse
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	50	100	50
max_features	-	-	-	-	-	sqrt
R^2 train set	0.64	0.88	0.94	0.88	0.83	0.92
RMSE train set	9.64	5.69	4.02	5.51	6.67	4.67
Correl. coef. train set	0.8	0.94	0.97	0.95	0.91	0.96
R^2 test set	0.48	0.37	0.25	0.52	0.49	0.51
RMSE test set	13.36	10.67	11.63	9.36	9.59	9.4
Correl. coef. test set	0.72	0.7	0.61	0.73	0.72	0.73

Experiment 2

kNN has the best performance on both train and test sets for the experiment 2 dataset, with $R^2 = 0.72$ on the test set. The rest of the models have lower scores, even though their performance on the train set is rather high. However, the test set scoring is extremely important, as it reflects the model's ability to predict target variable values on previously unseen data. Once again, it is possible to conclude that more data is needed in order to obtain models of a better quality.

Table 7.48. Results of predictive models hyperparameter optimization for tensile strain at break, experiment 2

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	30	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.05	1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	3	-	-	-
loss	-	-	-	lad	linear	-
criterion	-	mse	-	-	-	mae
max_depth	-	10	-	-	-	-
n_estimators	-	-	-	250	50	200
max_features	-	-	-	-	-	log2
R^2 train set	0.85	0.99	0.99	0.97	0.97	0.96
RMSE train set	16.06	0.01	0.01	7.36	7.2	8.62
Correl. coef. train set	0.92	0.99	0.99	0.99	0.99	0.98
R^2 test set	0.21	0.66	0.72	0.62	0.69	0.61
RMSE test set	43.8	28.68	25.87	30.13	27.23	30.51
Correl. coef. test set	0.61	0.85	0.85	0.85	0.88	0.85

Experiment 3

Tables 7.49 and 7.50 contain characteristics of the models that have the highest performance characteristics based on the grid search. The models trained on the sequential dataset have a better performance in comparison to the multi-output ones obtained using the parallel dataset. At the same time, Random Forest algorithm outperforms the other methods in both cases with the corresponding R^2 equal to 0.96 and 0.92 for the train set, as well as 0.69 and 0.87 for the test set. The model for the sequential dataset can be used as a s starting point for development of a more robust model for prediction of tensile strain at break for 4 mm thick dogbones produced with ContainerService recycled HDPE material.

Table 7.49. Results of predictive models hyperparameter optimization for tensile strain at break1 and tensile strain at break2, parallel dataset, experiment 3

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	relu	-	-	-

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
hidden layer neurons	30	-	-	-
solver	lbfgs	-	-	-
alpha	0.0001	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	4	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	7	-	-
n_estimators	-	-	-	200
max_features	-	-	-	sqrt
R^2 train set	0.78	0.99	0.99	0.96
RMSE train set	2.05	0.52	0.01	0.86
Correl. coef. train set	0.88	0.99	0.99	0.99
R^2 test set	0.43	0.61	0.62	0.69
RMSE test set	2.79	2.33	2.29	2.05
Correl. coef. test set	0.76	0.8	0.79	0.84

Table 7.50. Results of predictive models hyperparameter optimization for tensile strain at break, sequential dataset, experiment 3

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	10	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.0001	-	-	-	-	-
learning rate	0.001	-	-	0.01	0.1	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	2	-	-	-
loss	-	-	-	huber	square	-
criterion	-	mae	-	-	-	mae
max_depth	-	7	-	-	-	-
n_estimators	-	-	-	300	50	50
max_features	-	-	-	-	-	sqrt
R^2 train set	0.78	0.91	0.94	0.89	0.86	0.92
RMSE train set	2.08	1.34	1.09	1.48	1.66	1.29
Correl. coef. train set	0.88	0.95	0.97	0.95	0.93	0.96
R^2 test set	0.8	0.81	0.7	0.84	0.85	0.87
RMSE test set	1.92	1.85	2.34	1.72	1.67	1.51
Correl. coef. test set	0.9	0.92	0.85	0.92	0.92	0.94

Experiment 4

In case of the data obtained during the experiment 4, none of the models have high performance neither for parallel, nor for sequential dataset. In both cases, even if a model has a relatively high sore on the train set (Random Forest for the parallel dataset), the model's score drops significantly when it is tested on the test set. As it can be seen from the models' quality for all three mechanical characteristics, it is important to obtain more

data and if possible, add the relevant material information about the material characteristics of the material batch used for production of the focus parts.

Table 7.51. Results of predictive models hyperparameter optimization for tensile strain at break1 and tensile strain at break2, parallel dataset, experiment 4

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	relu	-	-	-
hidden layer neurons	15	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	uniform	-
number of neighbors	-	-	5	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mse
max_depth	-	5	-	-
n_estimators	-	-	-	250
max_features	-	-	-	sqrt
R^2 train set	0.51	0.72	0.53	0.9
RMSE train set	5.46	4.09	5.36	2.5
Correl. coef. train set	0.72	0.85	0.74	0.97
R^2 test set	0.51	0.25	0.43	0.53
RMSE test set	6.38	7.87	6.86	6.26
Correl. coef. test set	0.73	0.53	0.68	0.77

Table 7.52. Results of predictive models hyperparameter optimization for tensile strain at break, sequential dataset, experiment 4

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	20	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.05	-	-	-	-	-
learning rate	0.001	-	-	0.01	0.001	-
weights	-	-	uniform	-	-	-
number of neighbors	-	-	5	-	-	-
loss	-	-	-	ls	linear	-
criterion	-	mse	-	-	-	mae
max_depth	-	5	-	-	-	-
n_estimators	-	-	-	150	200	100
max_features	-	-	-	-	-	sqrt
R^2 train set	0.45	0.69	0.64	0.65	0.65	0.76
RMSE train set	7.12	5.32	5.73	5.68	5.63	4.67
Correl. coef. train set	0.67	0.83	0.81	0.84	0.81	0.88
R^2 test set	0.52	0.49	0.51	0.5	0.53	0.49
RMSE test set	5.63	5.8	5.72	5.79	5.58	5.79
Correl. coef. test set	0.72	0.71	0.73	0.71	0.73	0.71

Joined dataset

When it comes to the joined datasets, as Table 7.53 and 7.54 shows, there are no models trained on the parallel dataset that have acceptable characteristics. The Random Forest model has a relatively good performance, but it can and needs to be improved by adding more relevant data and the real material characteristics to the dataset. For the sequential dataset, the GBR and Random Forest models have acceptable scores of R^2 equal to 0.94 and 0.93 correspondingly, with the RMSE of 8.35 and 8.69. These models are more useful than those trained on the separate experiments datasets, as they have "learned" using more samples of data and as a result have better generalization abilities.

Table 7.53. Results of predictive models hyperparameter optimization for tensile strain at break1 and tensile strain at break2, parallel joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	relu	-	-	-
hidden layer neurons	10	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	4	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mse
max_depth	-	7	-	-
n_estimators	-	-	-	200
max_features	-	-	-	auto
R^2 train set	0.39	0.95	0.99	0.97
RMSE train set	11.19	3.18	0.01	2.32
Correl. coef. train set	0.59	0.97	0.99	0.99
R^2 test set	0.16	0.51	0.02	0.72
RMSE test set	15.6	10.11	14.32	7.72
Correl. coef. test set	0.26	0.77	0.43	0.85

Table 7.54. Results of predictive models hyperparameter optimization for tensile strain at break, sequential joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
activation	relu	-	-	-	-	-
hidden layer neurons	30	-	-	-	-	-
solver	lbfgs	-	-	-	-	-
alpha	0.0001	-	-	-	-	-
learning rate	0.001	-	-	0.05	0.05	-
weights	-	-	distance	-	-	-
number of neighbors	-	-	7	-	-	-
loss	-	-	-	huber	exponential	-
criterion	-	friedman_mse	-	-	-	mae
max_depth	-	12	-	-	-	-
n_estimators	-	-	-	200	50	300
max_features	-	-	-	-	-	auto
R^2 train set	0.81	0.99	0.99	0.97	0.91	0.98
RMSE train set	16.07	3.98	3.87	6.29	11.02	4.97

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	GBR	AdaBoost	Random Forest
Correl. coef. train set	0.9	0.99	0.99	0.99	0.95	0.99
R^2 test set	0.78	0.9	0.88	0.94	0.9	0.93
RMSE test set	15.94	10.67	11.64	8.35	10.88	8.69
Correl. coef. test set	0.88	0.96	0.94	0.97	0.95	0.97

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

It is also of interest to train models for prediction of the Young's modulus, tensile strength and tensile strain at break as a vector. This can be done using MLP, Decision Tree Regressor, kNN and Random Forest, as these algorithms can be utilized for the multi-output learning.

In order to do this the sequential joined dataset is used and the models are tuned using the grid search, similarly to how it was done in the previous sections. The hyperparameter set used for the grid search is the same as presented in the beginning of Section 7.4. This dataset is smaller than the one used for the same procedure for the dimensional properties, as some outliers were removed here as described in Section 7.2. No feature selection is applied prior to the model training, as the dataset includes only 19 features and based on the results presented in Section 7.3. Once again, the 70% – 30% training to testing ratio is used in combination with the 5-folds cross- validation on the training set. Table 7.55 shows the models that are obtained from the grid search. It is possible to see that Random Forest has the best performance in terms of the model's scores on the test set, while Decision Tree is the second best.

Table 7.55. Results of predictive models hyperparameter optimization for Young's modulus, tensile strength, tensile strain at break, sequential joined dataset

Model's hyperparameter	MLP FS = RReliefF	Decision Tree Regressor	kNN	Random Forest
activation	relu	-	-	-
hidden layer neurons	30	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	6	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	7	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.6	0.96	0.98	0.97
RMSE train set	58.91	18.93	13.21	15.31
Correl. coef. train set	0.6	0.97	0.99	0.99
R^2 test set	0.57	0.87	0.83	0.92
RMSE test set	63.71	35.53	40.43	28.32
Correl. coef. test set	0.56	0.94	0.94	0.97

Figures 7.10 – 7.12 show comparison of the actual and predicted values of the mechanical target variables of interest. The results are shown for 15 randomly selected test set samples. It is possible to see that the model predicts Young's modulus and tensile strength with higher precision in comparison to tensile strain at break. The reason for this might be that during the tensile testing, it is often hard to calculate the break point precisely and different methods might have different values of this characteristics. It is also visible that in most of the cases Random Forest values are the closest to the actual value of the focus variable, even for the tensile strain at break.

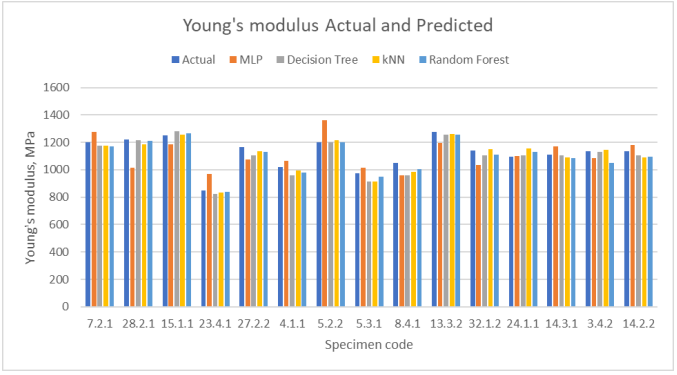


Figure 7.10. Actual and predicted values of Young's modulus target variable

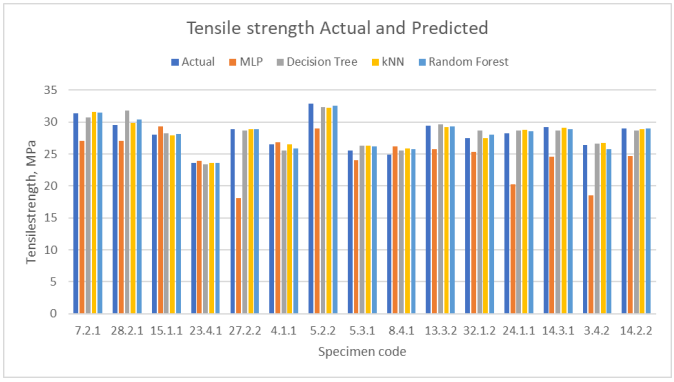


Figure 7.11. Actual and predicted values of tensile strength target variable

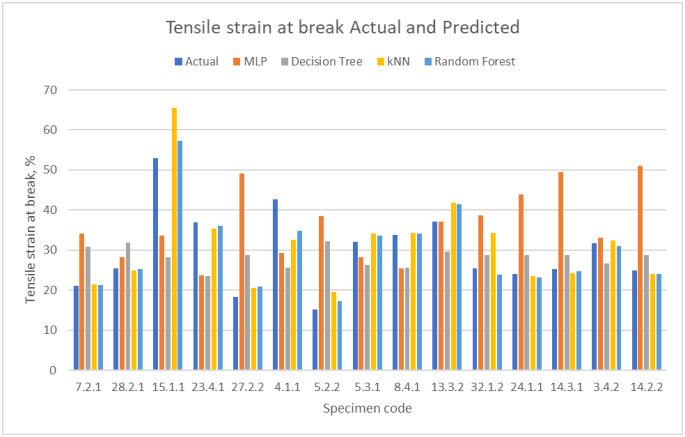


Figure 7.12. Actual and predicted values of tensile strain at break target variable

Since Decision Tree Regressor and Random Forest have very similar models' quality and a rather high performance, it was of interest to see if they use the same process parameters to make the predictions. Figures 7.13 and 7.14 are presented for this purpose. Both of the models give the material variable the highest score, which is very meaningful since the models are used for prediction of mechanical properties. Both models give some of the highest scores to the smallest cushion value, cushion after holding pressure, injection time and pressure value at switchover. Random Forest model, however, seems to be utilizing a higher number of parameters in comparison to the Decision Tree. At the same time, it is worth mentioning, that the Random Forest model is more robust and can be more useful in the long term perspective, as it uses many decision tree learners to do the prediction, while Decision Tree Regressor uses a single tree, and therefore might have a higher bias. At the same time, if a model with high interpretability is needed, Decision Tree Regressor model is the best choice.

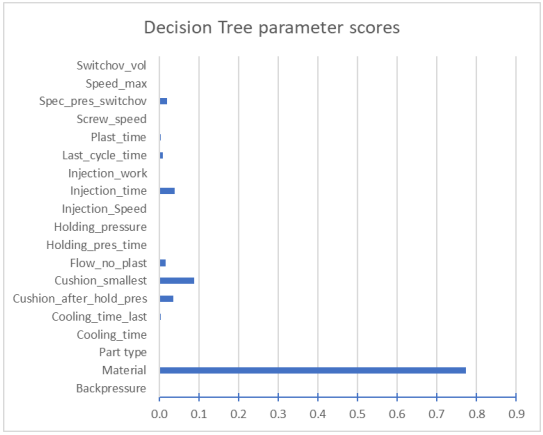


Figure 7.13. Decision Tree parameter scores

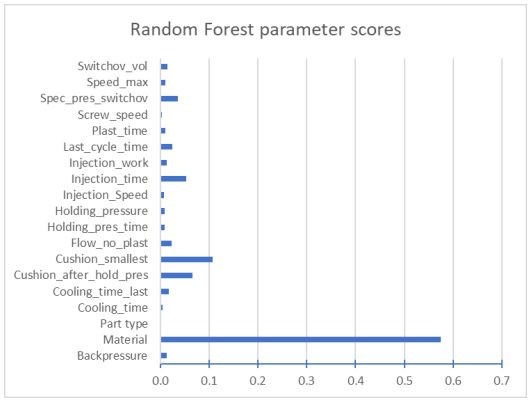


Figure 7.14. Random Forest parameter scores

Chapter 8

CC4 – Prediction of dimensional properties based on the data series datasets

Dear future generations:
Please accept our apologies. We were rolling drunk on petroleum.

– Kurt Vonnegut

This chapter describes development of the predictive models for the width and thickness dimensional properties of the produced parts based on the data series datasets. Composition of these datasets is described in more details in Section 4.4.2. The chapter shows how width and thickness can be predicted using the data series data of such parameters as cushion, mold pressure, mold temperature and screw position.

8.1 Preliminaries

As it has been mentioned previously, some of the injection molding process variables stay constant during production cycle, while the others are described with a curve rather than with a single number. Examples of such variables are cushion, mold pressure and temperature. Unlike datasets used in Chapters 6 and 7, where a set of various process variables is used as a model input to describe the injection molding process, in this chapter the datasets include data series that correspond to a curve that describes variation of a single process parameter. Here all the model's input features are the chosen process parameter values logged each 0.5 seconds during a production run. The number of input parameters in this case corresponds to the cycle time in seconds multiplied by 2 (one parameter per each 0.5 seconds). Cycle time is calculated as the time between the mold closing until its opening. The output/ focus variables are width1, width2, thickness1 and thickness2 (quality characteristics of specimens 1 and 2 that are produced during the same machine run). The data series datasets are only parallel datasets, no sequential or joined datasets for several experiments are used.

These prediction models are different from the ones trained in the previous chapters, as values of only one parameter need to be logged during the production cycle. As the cycle length might vary, in order to have the same number of features in all the data samples, cycles with the shorter cycle time are completed with zeros (padding) to become the same length as the samples with the longest cycle time. The nominal value of the width target variable is 10 mm for all experiments, while for thickness it is 4 mm for experiments 1, 3, 4 and 15 mm for experiment 2. No outliers are removed in the datasets used for training predictive models for the dimensional target variables.

Unlike the procedure described in Chapters 6 and 7, here no feature selection is performed, and all parameters are used as an input for training the models. This is done as removal of information regarding the value of the variable that describes the injection molding process at any point of the process might negatively influence the dataset's usefulness. Depending on the variables logged during the different experiments, various

parameters are used as a base for the dataset. In case of experiment 1, the cushion values during the production cycle are logged each 0.5 seconds, transposed, padded and used as the input parameters of the dataset. As it has been shown in the previous chapters various cushion characteristics, such as the minimum cushion value, cushion at switchover, etc. are often considered valuable by the feature selection methods. During experiment 2, such parameters as cushion, mold pressure, mold temperature and screw position are logged. Therefore, four different datasets based on the data series of the corresponding parameter are created to develop the prediction models. Mold pressure and temperature are considered as "fingerprints" of the process. It has been proven that the mold pressure trajectory can reflect variation of melt temperature, screw rotational speed, injection pressure and other useful process parameters [59]. For experiments 3 and 4 the screw position is used to develop the predictive models, as it is also one of the parameters that are representative of the production cycle's flow.

8.2 Predictive models development

Since the multi-output models are developed in this chapter, only MLP, Decision Tree Regressor, kNN and Random Forest methods are utilized. Apart from not using the FS methods prior to training the MLP and kNN models, the procedure is similar to that presented in Chapters 6 and 7. The dataset is divided into 70% train and 30% test set, and 5-folds cross-validation is performed on the train set. Grid search is once again used to select the most meaningful model through tuning the relevant hyperparameters. For the MLP the following hyperparameters and their values are considered:

- hidden layer sizes : [20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70],
- activation function : 'relu' , 'logistic',
- solver : 'lbfgs', 'sgd',
- alpha (L2 penalty parameter): [0.0001, 0.05],
- learning rate init: [0.001, 0.01, 0.05, 0.1, 0.3].

For Decision Tree the criterion and maximum tree depth parameters were evaluated:

- criterion: 'mse', 'friedman_mse', 'mae',
- maximum tree depth: [5, 7, 10, 12, 15, 20, 25].

For kNN the next hyperparameters were used:

- weights: 'uniform', 'distance',
- number of neighbors: [2, 3, 4, 5, 6, 7].

In case of the Random Forest, the number of estimators, max features and criterion hyperparameters were varied:

- number of estimators: [50, 100, 150, 200, 250, 300],
- max features: 'auto', 'sqrt', 'log2',
- criterion: 'mse', 'mae'.

It is important to note that time needed for the grid search for selection of the best performing MLP and Random Forest models has increased. With use of the data series

datasets it takes up to 20 minutes to find the best set of hyperparameters. Time for the hyperparameters selection for the Decision Tree Regressor and kNN did not vary.

8.2.1 Experiment 1

Table 8.1 shows results of training the chosen ML methods using the cushion data series dataset and the hyperparameters tuning with grid search. The model can predict values of width and thickness for the first and the second specimen molded during the same production run for experiments 1, 3-4 or for only one specimen produced per run in case of experiment 2. The models do not have extremely high performance, however, it isn't low either. kNN has the highest model's scores both on the train and the test sets with R^2 = 0.79, RMSE = 0.09 and the correlation coefficient of 0.92 on the test set. Random Forest also has a rather high performance and is only slightly outperformed by the kNN. These two models demonstrate that they can to certain extent generalize on the previously unseen data. Their performance is comparative with that of the models trained on experiment 1 parallel dataset (Sections 6.4.1 and 6.4.2). However, only cushion process variable data is used in the presented model, instead of over 10 parameters utilized in the previously trained models. It is assumed that addition of more samples to the training dataset will lead to increase of the model's quality and development of a useful decision-making support tool.

Table 8.1. Results of predictive models hyperparameter optimization for width1, width2, thickness1 and thickness2, cushion data series dataset for experiment 1

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	45	-	-	-
solver	lbfgs	-	-	-
alpha	0.0001	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mse
max_depth	-	7	-	-
n_estimators	-	-	-	100
max_features	-	-	-	auto
R^2 train set	0.87	0.99	0.99	0.98
RMSE train set	0.07	0.02	0.01	0.03
Correl. coef. train set	0.92	0.99	0.99	0.99
R^2 test set	0.66	0.69	0.79	0.77
RMSE test set	0.11	0.11	0.09	0.09
Correl. coef. test set	0.83	0.87	0.92	0.89

8.2.2 Experiment 2

Four different datasets are used to train the predictive models in this section based on the data obtained from experiment 2. The model scores with hyperparameters selected using grid search are shown in Tables 8.2 – 8.5. They use values of the selected variable logged

each 0.5 seconds during the production cycle as inputs, the variables are: cushion, mold pressure, mold temperature or screw position. The models based on cushion and mold pressure have rather high performance for train and test sets, except for MLP trained on the mold pressure data. When looking at the models trained on the mold temperature data, only Decision Tree Regressor and Random Forest have relatively acceptable scores, while MLP and kNN are not able to generalize on the test set data. In case of the screw position, kNN and Random Forest show good results. The models trained on the experiment 2 dataset to predict width and thickness in Chapter 6 have rather high scores. However, it has been mentioned that obtaining more data would increase the models' performance even more. In case of the data series datasets, which have significantly more features (one per each 0.5 seconds of the production cycle), it is shown that it is also possible to obtain models with acceptable quality that can predict both width and thickness simultaneously. If we compare the models' performance depending on which process parameter was used as a base for the dataset, it is possible to see that the cushion data series dataset has the highest average score. The models trained using the mold temperature data series dataset have the lowest score on the test sets. kNN and Random Forest models created based on the mold pressure and screw position data series datasets have relatively good scores on both train and test sets, while the rest do not have acceptable quality.

Table 8.2. Results of predictive models hyperparameter optimization for width and thickness, cushion data series dataset for experiment 2

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	50	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mae	-	mse
max_depth	-	10	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.97	0.99	0.99	0.97
RMSE train set	0.02	0.01	0.01	0.01
Correl. coef. train set	0.89	0.99	0.99	0.99
R^2 test set	0.84	0.83	0.92	0.88
RMSE test set	0.03	0.03	0.02	0.03
Correl. coef. test set	0.71	0.86	0.9	0.87

Table 8.3. Results of predictive models hyperparameter optimization for width and thickness, mold pressure data series dataset for experiment 2

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	60	-	-	-
solver	lbfgs	-	-	-

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	4	-
loss	-	-	-	-
criterion	-	mse	-	mse
max_depth	-	7	-	-
n_estimators	-	-	-	50
max_features	-	-	-	log2
R^2 train set	0.87	0.99	0.99	0.99
RMSE train set	0.03	0.01	0.01	0.01
Correl. coef. train set	0.83	0.99	0.99	0.99
R^2 test set	0.62	0.93	0.89	0.91
RMSE test set	0.05	0.02	0.03	0.02
Correl. coef. test set	0.67	0.89	0.86	0.85

Table 8.4. Results of predictive models hyperparameter optimization for width and thickness, mold temperature data series dataset for experiment 2

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	70	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mae	-	mae
max_depth	-	10	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.84	0.99	0.99	0.95
RMSE train set	0.03	0.01	0.01	0.02
Correl. coef. train set	0.9	0.99	0.99	0.98
R^2 test set	0.32	0.66	0.17	0.72
RMSE test set	0.11	0.05	0.08	0.04
Correl. coef. test set	0.35	0.71	0.67	0.74

Table 8.5. Results of predictive models hyperparameter optimization for width and thickness, screw position data series dataset for experiment 2

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	20	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mse
max_depth	-	5	-	-
n_estimators	-	-	-	150
max_features	-	-	-	log2
R^2 train set	0.42	0.99	0.99	0.98
RMSE train set	0.06	0.01	0.01	0.01
Correl. coef. train set	0.79	0.99	0.99	0.99
R^2 test set	0.2	0.52	0.92	0.84
RMSE test set	0.14	0.06	0.02	0.03
Correl. coef. test set	0.16	0.5	0.9	0.85

8.2.3 Experiment 3

Table 8.6 shows that kNN and Random Forest models developed based on the screw position trajectories data for the experiment 3 have high values of R^2 and correlation coefficient and a low value of the RMSE. They are able to predict the produced part's width and thickness simultaneously based only on the screw position data obtained during the production cycle without any other process parameter information. The models presented in Section 6.4.1 (width target variable), have even higher model characteristics scores. However, those obtained in Section 6.4.2 (width focus variable) underperform the kNN and Random Forest models presented in this chapter. Such models can be used for the decision-making support and their quality can be increased as more data is added to the training set.

Table 8.6. Results of predictive models hyperparameter optimization for width1, width2, thickness1 and thickness2, screw position data series dataset for experiment 3

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	50	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	10	-	-
n_estimators	-	-	-	300
max_features	-	-	-	log2
R^2 train set	0.98	0.99	0.99	0.98
RMSE train set	0.03	0.01	0.01	0.04
Correl. coef. train set	0.99	0.99	0.99	0.99
R^2 test set	0.61	0.67	0.91	0.89
RMSE test set	0.11	0.1	0.05	0.06
Correl. coef. test set	0.89	0.9	0.98	0.96

8.2.4 Experiment 4

Based on the models' scores seen in Table 8.7, the random Forest is the only model that has a high score on the training set and a somewhat acceptable performance on the test set. At the same time, even this score might not be good enough to use this model for an accurate decision support while producing 4 mm thick dogbones from the RePro recycled HDPE material. The models are trained using screw position parameter logged during the parts' production cycle. The models presented in Chapter 6 for both width and thickness target variables outperform the models obtained in this section.

Table 8.7. Results of predictive models hyperparameter optimization for width1, width2, thickness1 and thickness2, screw position data series dataset for experiment 4

Model's hyperparameter	**MLP FS = No FS**	**Decision Tree Regressor**	**kNN**	**Random Forest**
activation	logistic	-	-	-
hidden layer neurons	45	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	7	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	7	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.96	0.99	0.99	0.97
RMSE train set	0.03	0.01	0.01	0.03
Correl. coef. train set	0.99	0.99	0.99	0.99
R^2 test set	0.32	0.43	0.63	0.75
RMSE test set	0.15	0.14	0.11	0.09
Correl. coef. test set	0.8	0.83	0.83	0.9

Chapter 9

CC5 – Prediction of mechanical properties based on the data series datasets

When you're seventeen you know everything. When you're twenty-seven if you still know everything you're still seventeen.

– Ray Bradbury, "Dandelion Wine"

Chapter 9 presents how predictive models for tensile modulus, tensile strength and tensile strain at break target variables were developed using the data series datasets. The datasets development and structure are described in Section 4.4.2. Depending on the experiment number, data series for cushion, mold pressure, mold temperature or screw position are utilized.

9.1 Preliminaries

Similarly to the case with the dimensional target variables, it is of interest to compare quality of the predictive models based on the data series datasets with those developed using the general datasets. While creating the datasets, the same outliers as mentioned in Chapter 7 were removed. The input variables are points from cushion, mold pressure, mold temperature or screw speed trajectories, depending on the experiment number. Instead of width and thickness, the mechanical properties are used as the target variables. Multi-output models are trained, where values of all the parameters for one (experiment 2) or two (experiments 1, 3-4) specimens produced during one run are predicted simultaneously. MLP, Decision Tree Regressor, kNN and Random Forest methods are used to develop the models, as the rest of the methods used in this work cannot be used to train the multi-output models.

Young's modulus has a range of values between 750 to 1200 MPa, tensile stress varies between 21.25 and 34 MPa, while tensile strain at break between 5 and 250 % based on the data at hand. The presented range corresponds to the target variables values across all four experiments. Due to a large spread of the values, the RMSE is expected to be significantly higher for the models obtained in this chapter. No prior feature selection is applied before training MLP or kNN models.

9.2 Predictive models development

The procedure used to train the models is similar to that described in Chapter 8. The dataset is divided into 70-30 % ratio for the train and test set, 5-folds cross-validation is used on the train set. Grid search is utilized to find the best performing model on the set of the chosen hyperparameters. The hyperparameters search space is the same as the one used in Chapter 8 and is therefore not repeated here.

9.2.1 Experiment 1

Table 9.1 contains the MLP, Decision Tree Regressor, kNN and Random Forest model scores and hyperparameter values selected using the grid search. The obtained models scores are similar to those from Sections 7.4.1 (Young's modulus target variable) and 7.4.3 (tensile strain at break target variable), while the models for separate prediction of tensile strength have higher model characteristics. The models obtained in this section are trained only on the cushion value cycle data, while for the models from Chapter 7 over 10 model parameters are used within a prediction model. Random Forest outperforms the other methods but does not have high enough scores on the test set to be used on practice. If more data is obtained, the models' quality will increase, and their usability will improve correspondingly. Young's modulus1 and 2, tensile strength1 and 2 and tensile strain at break 1 and 2 are characteristics of specimen 1 and 2 that are produced during the same production cycle.

Table 9.1. Results of predictive models hyperparameter optimization for Young's modulus1, Young's modulus2, tensile strength1 and tensile strength2, tensile strain at break1, tensile strain at break2 cushion data series dataset for experiment 1

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	55	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	5	-
loss	-	-	-	-
criterion	-	mse	-	mse
max_depth	-	7	-	-
n_estimators	-	-	-	300
max_features	-	-	-	auto
R^2 train set	0.68	0.91	0.99	0.94
RMSE train set	32.74	17.3	2.08	13.77
Correl. coef. train set	0.8	0.91	0.99	0.98
R^2 test set	0.58	0.63	0.65	0.7
RMSE test set	46.17	43.25	42.03	39.05
Correl. coef. test set	0.88	0.83	0.88	0.89

9.2.2 Experiment 2

Tables 9.2 – 9.5 describe the models obtained through use of the data series datasets for cushion, mold pressure, mold temperature and screw position. Unfortunately, the only model with a relatively acceptable performance is a Decision Tree Regressor trained on the mold temperature data series dataset (Table 9.4). The rest of the models regardless of the dataset they were trained on have very low performance characteristics and need to be further worked on. It is a huge difference between the models' quality for the same experiment datasets, but for the dimensional properties target variables. Prediction of mechanical properties is more challenging and requires more data. All the datasets used

in this chapter are slightly smaller than those used in Chapter 8, due to the outlier's removal. However, for the experiment 2 dataset the difference is only in 7 samples. At the same time, the dataset obtained from the experiment 2 is in general relatively small and this is the main reason for obtaining models of low quality. The models developed within this section have performance characteristics that are similarly low to those presented in Sections 7.4.1 for tensile modulus, in 7.4.2 for tensile strength and in 7.4.3 for tensile strain at break using the separate experiment 2 dataset. Since in experiment 2 only one specimen was produced per production cycle, mechanical properties of one specimen are predicted by the models.

Table 9.2. Results of predictive models hyperparameter optimization for Young's modulus, tensile strength, tensile strain at break cushion data series dataset for experiment 2

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	55	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	6	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	7	-	-
n_estimators	-	-	-	200
max_features	-	-	-	auto
R^2 train set	0.04	0.99	0.99	0.94
RMSE train set	23.96	2.42	0.01	6.14
Correl. coef. train set	0.07	0.94	0.99	0.98
R^2 test set	0.03	0.41	0.17	0.7
RMSE test set	29.44	22.33	26.51	15.9
Correl. coef. test set	0.25	0.01	0.15	0.31

Table 9.3. Results of predictive models hyperparameter optimization for Young's modulus, tensile strength, tensile strain at break mold pressure data series dataset for experiment 2

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	50	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	6	-
loss	-	-	-	-
criterion	-	mse	-	mse
max_depth	-	5	-	-
n_estimators	-	-	-	100
max_features	-	-	-	auto

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
R^2 train set	0.12	0.97	0.99	0.89
RMSE train set	22.85	3.95	0.01	8.13
Correl. coef. train set	0.01	0.83	0.99	0.99
R^2 test set	0.05	0.04	0.53	0.44
RMSE test set	28.29	28.39	19.91	21.63
Correl. coef. test set	0.11	0.14	0.51	0.45

Table 9.4. Results of predictive models hyperparameter optimization for Young's modulus, tensile strength, tensile strain at break mold temperature data series dataset for experiment 2

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	70	-	-	-
solver	lbfgs	-	-	-
alpha	0.0001	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	5	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mse
max_depth	-	5	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.07	0.86	0.99	0.86
RMSE train set	23.49	9.08	0.01	9.2
Correl. coef. train set	0.32	0.74	0.99	0.98
R^2 test set	0.03	0.79	0.17	0.56
RMSE test set	29.48	8.96	26.37	19.18
Correl. coef. test set	0.24	0.72	0.15	0.43

Table 9.5. Results of predictive models hyperparameter optimization for Young's modulus, tensile strength, tensile strain at break screw position data series dataset for experiment 2

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	55	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	6	-
loss	-	-	-	-
criterion	-	mae	-	mae
max_depth	-	5	-	-
n_estimators	-	-	-	50
max_features	-	-	-	auto
R^2 train set	0.22	0.92	0.99	0.9

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
RMSE train set	21.53	6.99	0.01	7.57
Correl. coef. train set	0.29	0.78	0.99	0.97
R^2 test set	0.21	0.22	0.29	0.53
RMSE test set	31.69	25.62	24.38	19.9
Correl. coef. test set	0.18	0.03	0.24	0.25

9.2.3 Experiment 3

Table 9.6 indicates that all the models obtained as a result of application of selected ML methods to the screw position data series dataset for experiment 3 are of extremely low quality, as they are not able to perform adequately on the test set. MLP model has a rather low performance even on the train set. If these models are compared to the ones for the parallel dataset of experiment 3 used in Sections 7.4.1 – 7.4.3 with different target variables, it can be seen that those models for tensile strength and tensile strain at break have a better performance on the test set, while those for Young's modulus have very low scores. The models presented in this chapter, however, predict all those target variables simultaneously. Due to this and to the relatively small dataset size, the models cannot be considered for further use in their current state and need to be further worked on.

Table 9.6. Results of predictive models hyperparameter optimization for Young's modulus1, Young's modulus2, tensile strength1 and tensile strength2, tensile strain at break1, tensile strain at break2 screw position data series dataset for experiment 3

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	55	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	4	-
loss	-	-	-	-
criterion	-	friedman_mse	-	mae
max_depth	-	7	-	-
n_estimators	-	-	-	300
max_features	-	-	-	sqrt
R^2 train set	0.37	0.86	0.99	0.92
RMSE train set	24.82	11.68	0.01	9.13
Correl. coef. train set	0.69	0.93	0.99	0.98
R^2 test set	0.01	0.02	0.29	0.25
RMSE test set	28.54	28.67	23.94	24.62
Correl. coef. test set	0.29	0.63	0.81	0.69

9.2.4 Experiment 4

When it comes to the data from experiment 4, the models that were trained using the separate parallel dataset have low performance for Young's modulus and tensile strain at

break, while the models for tensile strength have relatively good results. In case of the screw position data series dataset from experiment 4, Table 9.7 shows that the models have low scores for the test set and are not useable in their current form. The reasons and ways to help this are the same as for the other experiments described in this chapter (collection of more relevant data, addition of information about the material properties of the used material batch, etc.).

Table 9.7. Results of predictive models hyperparameter optimization for Young's modulus1, Young's modulus2, tensile strength1 and tensile strength2, tensile strain at break1, tensile strain at break2 screw position data series dataset for experiment 4

Model's hyperparameter	**MLP FS = No FS**	**Decision Tree Regressor**	**kNN**	**Random Forest**
activation	logistic	-	-	-
hidden layer neurons	60	-	-	-
solver	lbfgs	-	-	-
alpha	0.0001	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	6	-
loss	-	-	-	-
criterion	-	mae	-	mae
max_depth	-	10	-	-
n_estimators	-	-	-	150
max_features	-	-	-	log2
R^2 train set	0.38	0.99	0.99	0.92
RMSE train set	23.74	2.5	0.01	8.37
Correl. coef. train set	0.62	0.98	0.99	0.98
R^2 test set	0.07	0.03	0.28	0.14
RMSE test set	29.04	37.3	25.56	32.21
Correl. coef. test set	0.36	0.55	0.43	0.52

Chapter 10

CC6 – Prediction of dimensional deviations using the general datasets

This too shall pass.

– Persian adage

This chapter describes development of models for prediction of dimensional deviations of focus parts. The models are trained using parallel and sequential joined datasets. More information about these datasets can be found in Section 4.4.3.

10.1 Preliminaries

Chapter 6 describes development of predictive models for the dimensional properties of focus parts trained on separate and joined datasets. It is of interest to compare quality of those models with the models that predict dimensional deviations. The models are trained only on joined datasets, as it has been seen based on results described in Chapters 6 and 7 that the separate experiments datasets are often too small to create a meaningful model.

There were no outliers removed from the joined datasets for prediction of dimensions and dimensional deviations. No feature selection is applied prior to the model development as parallel joined dataset contains only 20 features, while the sequential one only 19. In addition, results from Section 6.3 have shown that the quality of the models increase if all the features from the joined datasets are included. The dataset was divided into 70% train and 30% test sets, where 5-folds cross-validation was performed on the train set. MLP, Decision Tree Regressor, kNN and Random Forest are used for the models' development, as they are capable of creating multi-output models. Similarly to Chapters 6 – 9, grid search is utilized to tune various hyperparameters. The set of the tuned hyperparameters for selected ML methods is the same as the one used in Chapters 8 and 9 and can be found in Section 8.2. Similar models are not proposed for prediction of the mechanical properties, as no nominal value was identified for them. In case of width the nominal value is 10 mm for all experiments, while the nominal thickness is 4 mm for experiments 1, 3-4 and 15 mm for experiment 2.

10.2 Predictive models for dimensional deviations, parallel joined dataset

Table 10.1 contains information about the models' hyperparameters selected in the process of grid search application. Since the parallel joined dataset is utilized here, the predicted variables are **thickness1, thickness2, width1 and width2** – these are dimensional characteristics of specimen 1 and 2 produced during the same machine run. Random Forest model outperforms the rest of the models with $R^2 = 0.99$ on the train and $R^2 = 0.94$ on the test set. Values of the correlation coefficient for this model are also high. RMSE at the same time is considerably high, since the dimensional deviations and not the dimensions value itself is predicted. The second-best model is the one created with Decision Tree Regressor method. It has $R^2 = 0.99$ on the train set and $R^2 = 0.88$ on the

test set. Values of the correlation coefficient are also acceptable, while the RMSE is slightly higher than the one for the Random Forest. These results are not compared with those presented in Section 6.4.3 as sequential joined dataset is utilized there.

Table 10.1. Results of predictive models hyperparameter optimization for deviations of thickness1, thickness2, width1 and width2 target variables, parallel joined dataset

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	25	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	3	-
loss	-	-	-	-
criterion	-	mae	-	mse
max_depth	-	12	-	-
n_estimators	-	-	-	200
max_features	-	-	-	auto
R^2 train set	0.77	0.99	0.99	0.99
RMSE train set	0.1	0.01	0.01	0.02
Correl. coef. train set	0.74	0.99	0.99	0.99
R^2 test set	0.68	0.88	0.78	0.94
RMSE test set	0.12	0.07	0.1	0.05
Correl. coef. test set	0.64	0.88	0.83	0.94

Figures 10.1 – 10.4 depict comparison of the actual and predicted values of the dimensional deviations for each of the variables predicted by various models. Here, thickness1, thickness2, width1 and width2 are dimensional characteristics of specimens 1 and 2 produced simultaneously. The values are shown for 15 randomly selected data samples from the test set. The figures show that the Random Forest is the one that usually predicts values that are closest to the actual one. In case of sample 11.2, however, all the models have a significant difference between the actual and the predicted values. MLP is the model that underperforms the most in case of all the target variables predictions.

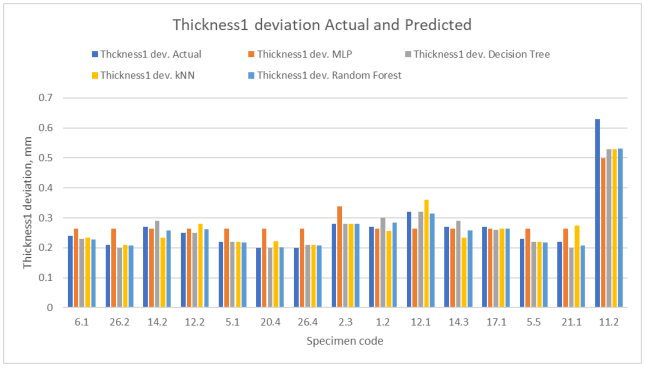


Figure 10.1. Actual and predicted values of thickness1 deviation target variable

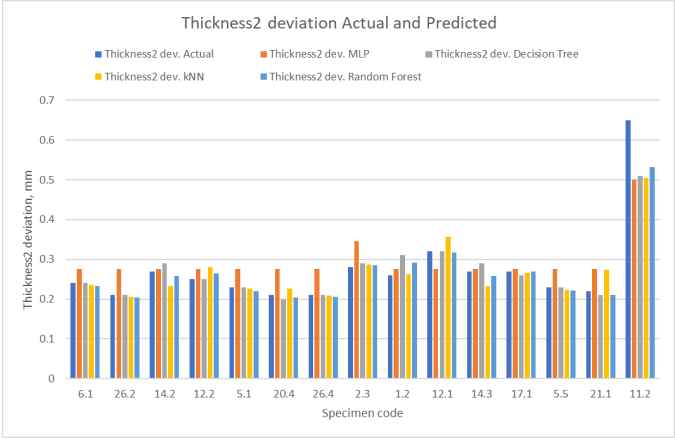


Figure 10.2. Actual and predicted values of thickness2 deviation target variable

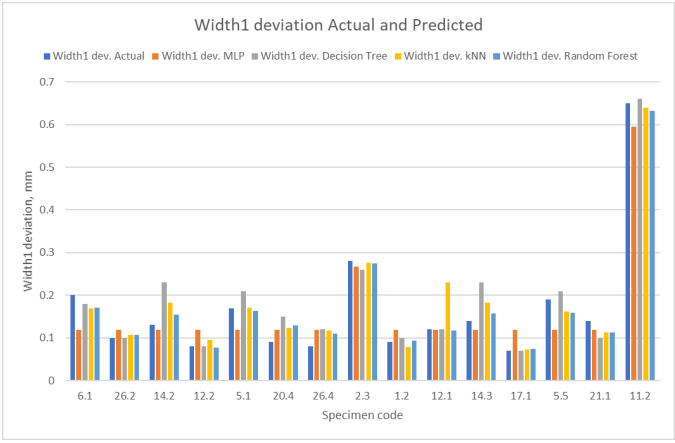


Figure 10.3. Actual and predicted values of width1 deviation target variable

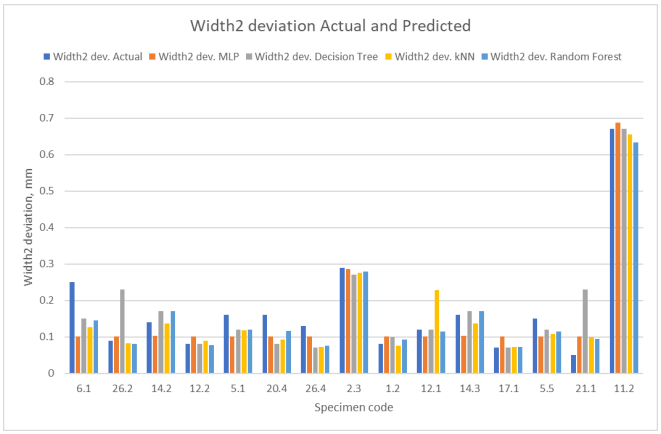


Figure 10.4. Actual and predicted values of width2 deviation target variable

Since Random Forest and Decision Tree Regressor have the best and the second-best model qualities, it was of interest to see which features do these models utilize and how do they score them. Figures 10.5 and 10.6 show parameter scores of Decision Tree and Random Forest models correspondingly. Both models give the highest scores to holding pressure, cushion after holding pressure and material. The Decision Tree gives more parameters significant scores, such as the smallest cushion value, injection work, last ejector position and backpressure. The Random Forest model, on the other hand, has many parameters with very low scores and some that are equal to zero.

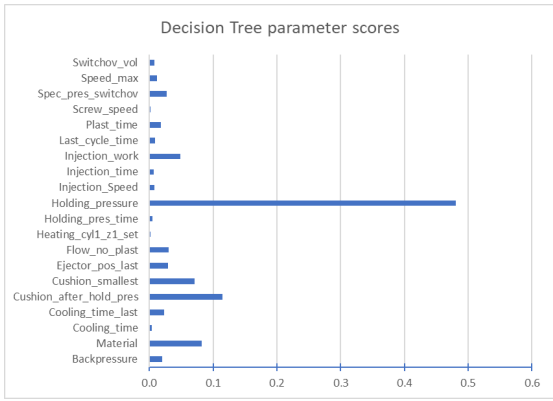


Figure 10.5. Decision Tree parameter scores, parallel joined dataset

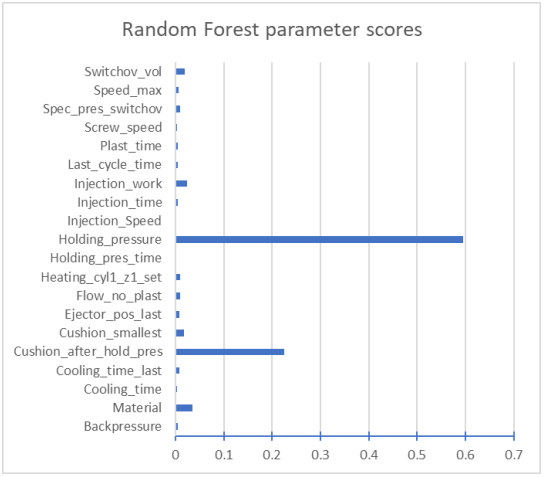


Figure 10.6. Random Forest parameter scores, parallel joined dataset

10.3 Predictive models for dimensional deviations, sequential joined dataset

Table 10.2 shows results of the hyperparameter tuning for the models with **width and thickness target variables** (include width1, width2 and thickness1, thickness2 correspondingly merged sequentially) for the sequential joint dataset. Since this dataset is larger than the parallel one and more data is available for training, the predictive models' quality is slightly higher than for the models trained on the parallel joined dataset. Once again Random Forest and Decision Tree Regressor algorithms outperform the rest in terms of their performance on the test set, at the same time kNN has the highest scores on the train set. The models able to predict the dimensional properties values described in Section 6.4.3 have better scores than the ones presented in this section. However, Random Forest and Decision Tree Regressor still have a worthy performance and can be used as a starting point for further development of the models for prediction of the dimensional properties' deviations.

Table 10.2. Results of predictive models hyperparameter optimization for thickness and width deviation target variables, sequential joined dataset

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
activation	logistic	-	-	-
hidden layer neurons	70	-	-	-
solver	lbfgs	-	-	-
alpha	0.05	-	-	-
learning rate	0.001	-	-	-
weights	-	-	distance	-
number of neighbors	-	-	2	-
loss	-	-	-	-
criterion	-	mse	-	mae
max_depth	-	5	-	-
n_estimators	-	-	-	100

Model's hyperparameter	MLP FS = No FS	Decision Tree Regressor	kNN	Random Forest
max_features	-	-	-	sqrt
R^2 train set	0.86	0.95	0.99	0.98
RMSE train set	0.09	0.05	0.03	0.03
Correl. coef. train set	0.86	0.96	0.99	0.99
R^2 test set	0.84	0.92	0.85	0.92
RMSE test set	0.09	0.06	0.09	0.06
Correl. coef. test set	0.87	0.94	0.9	0.94

Figures 10.7 and 10.8 reflect on the difference between the actual and predicted values of the thickness and width deviations target variables correspondingly. The values are shown for 15 samples that are randomly selected from the test dataset. In most of the cases all four models can predict the deviation values close to the actual one. However, for some of the samples, such as 17.1.1 for the thickness deviation, all models' predictions significantly differ from the real value.

Figures 10.9 and 10.10, at the same time, depict the parameter scores that Decision Tree Regressor and Random Forest methods give to the input parameters. These two models have the best performance on the test set and are therefore compared. Unlike the models trained on the parallel dataset, Decision Tree has more parameters with low and zero scores, while Random Forest utilizes most of the input parameters. Similarly to the parallel dataset models, cushion after holding pressure, holding pressure and material features are those with the highest scores in both models.

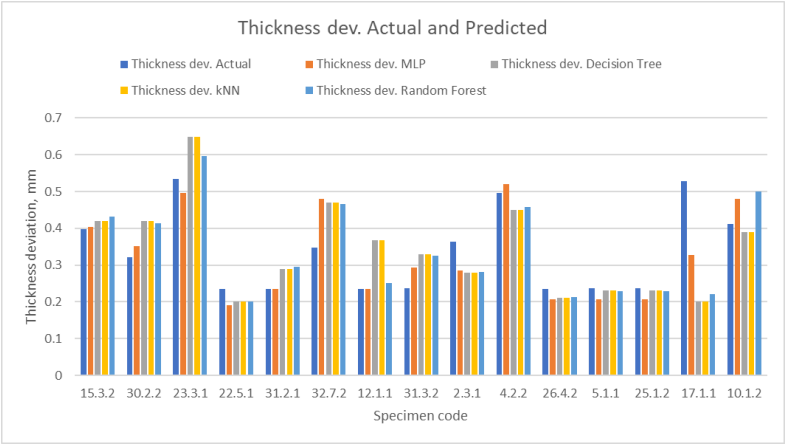


Figure 10.7. Actual and predicted values of thickness deviation target variable

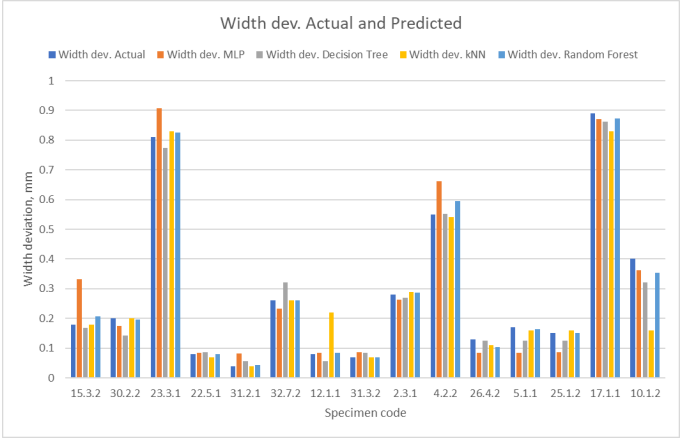


Figure 10.8. Actual and predicted values of width deviation target variable

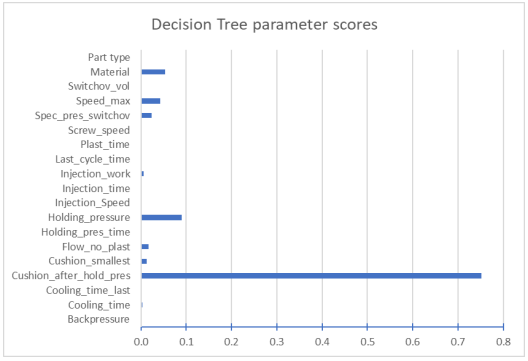


Figure 10.9. Decision Tree parameter scores, sequential joined dataset

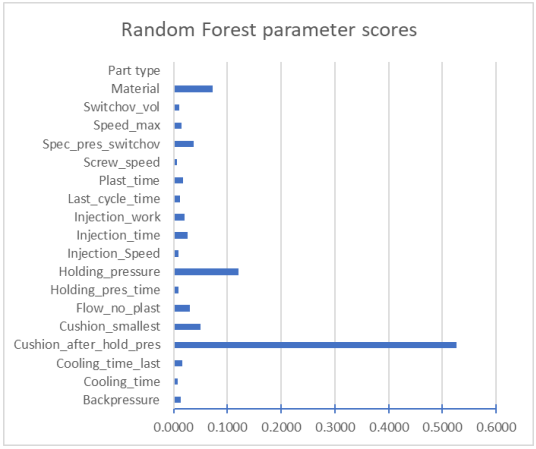


Figure 10.10. Decision Tree parameter scores, sequential joined dataset

Chapter 11

Discussion

No one travels
Along this way but I,
This autumn evening.

– Matsuo Bashō

This chapter discusses results obtained in Chapters 6 – 10, provides answers to the research questions formulated in Section 1.3 of Chapter 1, as well as elaborates on the use of the research philosophy paradigms applied in this study.

In Section 4.1 of Chapter 4, it has been stated that combination of positivism and pragmatism was used in this PhD work. Positivism is often adopted by natural scientists as it "*entails working on observable social reality to produce law-like generalizations*" [135]. The focus of this paradigm emphasizes using data obtained from the objective reality to draw facts uninfluenced by human interpretation from it. In this work, the injection molding data was acquired to create prediction models for various quality characteristics of focus parts. It is attempted to create those models so that they can describe the injection molding process and relationships between the process and target variables clearly and concisely. The research questions are formulated so that the answers to them can be used by other researchers in a similar way as described in this study.

Pragmatism concentrates on identification of a research problem and finding a practical solution to it by applying methods that bring the best results. This paradigm, however, suggests that there might be more than one method that suits the problem at hand, depending on what kind of data is available and what type of knowledge needs to be created. At the same time, pragmatists agree that different interpretations of the reality exist and that one single point of view can never provide the entire picture [135]. In this work this point of view is adopted. The research problem of development of an intelligent control system for thermoplastics injection molding is addressed and the research questions that would help to find a corresponding practical solution are formulated. It is proposed to use ML methods to develop the necessary prediction models, as these methods are reported to have a better performance in comparison to other more conventional modelling techniques.

Even though the combination of positivism and pragmatism fits well the quantitative study at hand, interpretivism also plays an important role here. Despite the fact that in most cases this paradigm is considered as the one used in the qualitative or social studies, it can to certain extent be applied when conducting the technical studies as well. Positivistic paradigm claims that a scientist needs to be neutral and detached of what is researched. However, how can one be completely sure of fully meeting this requirement is not completely understandable. The way reality is viewed by a scientist adopting the positivistic or pragmatic paradigm is influenced by the person's interpretation of reality and problem at hand. Problem formulation, selection of relevant method, data analysis and conclusions based on it are integral parts of any study, including the quantitative ones.

However, all these stages are influenced by the scientist's view on the surrounding world, his or her background and previous experiences. Therefore, interpretivism is also partially involved in the described PhD study, as the person conducting it was formulating the research questions, selecting the methods to be applied and analyzing the obtained results based on her past experiences. Any mathematical model will always be a more or less simplified representation of an abstract "truth" and it is up to the end user to decide to what extent it is useful. The author means that machine learning can be a useful approach for prediction of quality of IM parts, given enough samples for training of the models are provided.

11.1 RQ1

How to select parameters that influence the injection molded part quality using the ML methods?

In most of the research studies reviewed in Chapters 2 and 3 important process parameters are selected based on the theoretical knowledge, experience related to the injection molding process or using ANOVA [87]. A limited number of selected process parameters is chosen and then their values are logged, then these parameters are further included in the prediction model. Such approach might be relevant only in cases when the expert involved in the parameter selection is completely sure that these are the most important process features or in cases when no access to the rest of the parameters can be obtained. In some other cases, ANOVA is used. However, it also has its downsides, for example, it assumes that the parameters under analysis are independent and normally distributed, which is not always the case.

The amounts of logged process data are increasing, and the number of obtained process parameters raises correspondingly. In some cases, theoretical background, previous experience or ANOVA application might not be enough to identify the most relevant parameters among dozens of the obtained ones, as identification of the most important process features is an ongoing field of research [79]. Therefore, it is suggested that feature selection methods are put in use for these purposes. FS methods can be used not only for identification of the most relevant process parameters, but in order to increase understanding of relationships between the target characteristic and the process features, as well as for dimensionality reduction.

Various studies suggest different process parameters to be important for prediction of dimensional characteristics of focus parts. In [63] injection time, switchover time and mold temperature are used for development of a mathematical prediction models for warpage. Holding pressure and cooling time are among the parameters considered in [58] for shrinkage prediction. Mold temperature, holding time and pressure, as well as injection time are considered for minimization analysis of shrinkage and warpage in [88].

In this study 41 various machine and process parameters were logged during experiment 1, 65 during experiment 2 and 52 during experiments 3 and 4. Some of these parameters can be more relevant for predictive maintenance of IMMs, while the others can be used for development of models for prediction of dimensional and mechanical properties of injection molded parts. Five feature selection methods (Pearson's correlation, RFE,

Spearman's Correlation, CFS and RReliefF) were used to eliminate the parameters that do not contain significant amount of information about the corresponding target variables. Their performance was compared through training of an MLP model on the full parameters set and on the sets created based on results of the FS methods.

When it comes to prediction of width and thickness dimensional variables, based on Section 6.3, some of the important parameters selected by the feature selection methods are various characteristics of cushion variable (cushion value after holding pressure, average value, the smallest value), holding pressure, pressure at switchover, cooling time, material, mold temperature, holding pressure time, injection speed, backpressure and flow number of plastics (specific parameter utilized by the ENGEL IMM used during the experimental work). While some of these parameters were extensively utilized by other scholars, some of them are rarely or almost never considered. RReliefF and Spearman's FS methods were identified as the ones with the highest performance for width and thickness focus variables based on the collected data.

When mechanical properties rather than dimensional ones are in focus, the reviewed studies highlight such process parameters as holding pressure, injection pressure, melt temperature, injection time, holding time and cooling time to be influencing the final part's mechanical characteristics [86, 90]. Based on the results presented in Section 7.3 and utilization of FS methods parameters that are considered relevant for prediction of mechanical target variables are: mold pressure, holding pressure, average nozzle temperature, cushion characteristics, last ejector position, screw speed, material, clamping force at switchover, barrel temperature and closing force. RReliefF FS method has shown the best performance on the experimental data at hand, when compared to untuned MLP models developed based on the features selected by four other methods.

Some of the parameters that receive the highest scores by the FS methods in cases of both dimensional and mechanical target variables are rarely mentioned or used by other scholars for development of the corresponding prediction models. This is despite the fact that literature on the injection molding process itself often stresses importance of such parameters as cushion or barrel temperature. Based on the theoretical foundations of the injection molding process, it is possible to say that parameters selected by the FS methods are relevant to the corresponding target variables. Therefore, when in need to identify those that are the most influential for IM or other manufacturing processes the FS methods might be useful.

The FS methods advantage over application of, for example, ANOVA is that not all of these methods assume that the input parameters are normally distributed, independent or have only linear dependency with the target variable. In addition, these methods can be used for identification of process problem places, as some of them are based on analysis of variance, as well as for dimensionality reduction or improved understanding of relationships between the target variable and process parameters.

11.2 RQ2

How can machine learning methods be used for prediction of dimensions of the injection molded parts?

In this work results of development of prediction models for the dimensional target variables (width, thickness) and the corresponding deviations are described in Chapters 6, 8 and 10. The models are developed using various ML techniques. There are examples of such models creation using more traditional methods as FEM, RSM, numerical and stochastic simulations, analysis of P-V-T relationship and Taguchi methodology [63, 74, 79, 80]. However, as mentioned in Chapter 3, Section 2.2.3, most of them have certain disadvantages that make their application for the predictive models' creation too cumbersome and ineffective. For example, FEM requires significant computation efforts, numerical modeling might include simplifications that hinder possibility of proper reflection of the non-linear relationships due to the material behavior, etc. ML methods, on the other hand were reported to require smaller computational time, be capable of processing large amounts of data and model non-linear relationships. Therefore, in this dissertation application of six ML methods was investigated (MLP, Decision Tree Regressor, kNN, GBR, AdaBoost and Random Forest).

Shrinkage in the injection molding process is inherent, as it is caused by the difference in the polymer density at the processing temperature and the ambient temperature. As a result, final part's shrinkage might be up to 20% of the initial part size, depending on the process parameter settings. Low injection pressure, short holding or cooling times might be some of the main causes of extensive shrinkage and variation of the dimensional target characteristics. The task of considering various parameters that influence width and thickness is not easy. However, ML methods are capable of developing the models that describe a dynamic system with a significant number of important parameters. If the number is too large, FS methods should be applied as discussed in the previous section.

To make the Machine Learning methods work effectively and develop high quality models, large amounts of process data need to be acquired. This is necessary due to the fact that the ML methods are statistical methods and a small dataset will not allow to create a model that can adequately generalize on the previously unseen data. The data can be obtained on the daily basis, when the IMMs are running to produce focus parts at the shop floor. Later the data can be used to create or update the existing models to increase their initial quality characteristics. The latter is possible as the ML methods allow to adjust the models to the fast-changing environment, when more relevant data is acquired. However, standardized routines for the data acquisition and processing need to be established.

After obtaining the IM process data, this work proposes to go through four main steps: preprocess the data (remove missing data, eliminate outliers, normalize the data), perform feature selection (removal of constant, irrelevant features), apply the chosen ML methods and validate the models. This approach will allow to train the models on high quality data that doesn't have missing pieces of information or features that act like noise and lower quality of the overall model. In addition, it will allow to verify the final models' quality.

Based on the results presented in this dissertation, the separate experiments datasets might be too small to create meaningful models. However, when the data is integrated together and material and part type parameters are added, the obtained models often have a rather high performance and can be used as a starting point for further development of the necessary decision-support tools. Performance of the models trained on joined parallel

and sequential datasets for prediction of dimensional characteristics is comparable. The joined sequential dataset models usually slightly outperform the joined parallel models. In addition, the predictive models for the dimensional characteristics have higher performance than those created for prediction of the dimensional deviations. Random Forest was the method with the highest performance on the test set in most of the cases, showing that the obtained models have good generalization abilities and can make predictions on new data samples.

To confirm the models' usefulness validation routines are very important. In this work prior to the ML method application, used dataset was divided into 70% train and 30% test sets. This way the model was trained on 70% of data from the dataset, while 30% of it were used to show the model's performance on new data. At the same time, 5-folds cross-validation was applied on the train data. This way a two-step process for avoiding overfitting was used. In some cases, it has been observed no overfitting on the train dataset, but when testing the model on the test set, its generalization abilities were not reaching the acceptable level.

Three model characteristics were used in this work: R^2, square root mean error and correlation coefficient. R^2 shows how close the real data is to the fitted regression line. RMSE reflects on difference between predicted and observed (real) values. Correlation coefficient (Pearson's in this case) measures the strength and direction of a linear relationship between the predicted and measured values. Each of these measures reflects on the model's quality from a different perspective and if two of the scores are acceptable, while the third one is not, the obtained model's quality should be questioned. It is important to utilize several such measures rather than one, as a single measure might not always show the whole picture. For example, in some of the cases the correlation coefficient value is quite high, however, R^2 is too low, while the RMSE is too high. This means that the model systematically predicts values that are higher or lower than the observed ones. In such cases if only the correlation coefficient performance score is used, the obtained model's quality might be considered acceptable. However, in the reality it is not.

11.3 RQ3

How can machine learning methods be used for prediction of physical properties of the injection molded parts?

Quality is a complex concept and its definition might vary depending on the application field [43]. In case of injection molding, some of the most important quality issues are dimensional tolerances and mechanical properties [37]. Due to this, prediction models for mechanical properties (tensile modulus, tensile strength, tensile strain at break) are also developed in this work using the ML methods.

The data used for the models' development was collected during four different experiments and in the two first experiments virgin HDPE was used, in experiment 3 recycled HDPE from ContainerService supplier was utilized, while in experiment 4 – recycled HDPE from the RePro supplier. The created models' quality show that it is possible to create the relevant prediction models not only using the virgin material data,

but also utilizing the recycled materials data. Moreover, including this data into the training datasets for the prediction models will potentially facilitate the recycled materials use for production of various parts.

For each of the experiments a DOE was created, where certain process parameters were varied in order to gather data that reflects to certain extent on the variations within the injection molding process. This way a more diversified information about the process was obtained. The received datasets were used for development of predictive models for both dimensional and mechanical properties. The only difference between the datasets used for the width and thickness models training and the mechanical properties models development, is that in the second case, outliers related to the mechanical properties values had to be removed from the datasets. Appearance of the outliers might be caused by the errors in the data acquisition system, measuring devices or human errors. This, however, shows that the same data can be reused for creation of different models and for various purposes.

Due to the focus on the mechanical properties, material parameter was added to the joined datasets. It was not a value of a certain material characteristic, but just a code for different batches for the same material used on different days or the different material. When looking at the scores various FS and ML methods were giving to the dataset parameters, it is possible to see that the material parameter was often scored rather high. When further developing the models and gathering the relevant data, it would be useful to add a parameter for a real material property, such as for example viscosity rather than just a material code.

To collect the mechanical properties data a destructive tensile stress test needs to be conducted. When collecting the real production data, this might often be impossible to do, as only several control parts out of each batch are tested. In addition, it is understandable that no one wants to produce actual parts to apply destructive tests on all of them. At the same time, if a large number of products needs to be evaluated, this method becomes expensive and time consuming. This provides challenges regarding collection of the relevant mechanical properties data, and further investigation of their solutions is needed.

Most of the models created based on the sequential joined dataset have rather high quality and can be used as a starting point for further development of the necessary decision support tool. However, the current models might not be general enough and in case of the separate experiments datasets do not have high enough quality characteristics to be utilized in their current form. The fact that the data was collected using the same IMM, only three material types and two parts whose main difference is in the thickness value, pose too many limitations. One of the solutions for this might be transfer learning, as it allows to use the same data to solve new tasks and addition of data from different sources to find a suitable solution to the task at hand.

11.4 RQ4

How to create an intelligent control system for thermoplastics injection molding?

Chapter 5 presents application of the model-based systems engineering approach for development of the intelligent control system for thermoplastics injection molding. The chapter includes a framework that can be used as a basis/ support for planning and creation of such systems. Some of the important functions that the system should be able to perform are: acceptance of various user inputs, the input analysis, acquisition of machine and process data from the IMM built-in and additional sensors, prediction of the requested characteristics of the parts' quality, optimization of the IMM parameter values, reporting, etc. The corresponding CC (computational core) modules are described that should cover most of the system's functions. At the same time, the examples of the computational units' development and operation are provided in the results Chapters from 6 to 10. These chapters formulate a basis and a rough example of how a decision-support system for enhancement of the injection molding process can be created.

For such system to be successful, high quality and timely data acquisition is a must. If no relevant data can be acquired – no further analysis is possible. As it has been mentioned in the previous chapters, there is possibility for destructive and non-destructive data collection. Non-destructive data collection allows to log the data through use of sensors that are already available on an IMM without installation of additional ones, as well as in some cases through use of mathematical models and simulation. In addition, nondestructive sensors such as ultrasonic probes, dielectrometers and displacement transducers can be applied. On the one hand, IMMs and their built-in sensors are the most important source of the relevant data [7]. On the other hand, there is a significant number of scholars that underline importance of acquisition of mold data, such as mold cavity pressure and temperature, since cavity pressure trajectory during the IM process is claimed to be "fingerprints" of the process [59].

In this work, models based on a number of parameters, whose data was acquired from the machine built-in sensors, were developed, as well as models based on the data from the destructive mold pressure and temperature sensors. In addition, cushion and screw position trajectories logged from the machine built-in sensors were used to propose the corresponding prediction models. When comparing quality of the models developed using the multiple parameters data (joined datasets) and single parameter trajectories data, it is possible to see that the models' quality for prediction of the dimensional target variables is comparable, even though the models trained on the general joined datasets outperform those trained on the data series datasets. However, when it comes to development of the models for prediction of mechanical properties, the general joined datasets models significantly outperform the data series ones. This might be due to the fact that joined datasets models have more data or due to not having enough information about the process, when a dataset includes data of only one parameter trajectory. One of the disadvantages of using the data series datasets is that length of the production cycle might significantly vary depending on the produced part size, geometry and material. Therefore, the length of the obtained data samples and features differs for the data series dataset. Padding or interpolation techniques can be applied to make sure that all the data samples are of the same length, however, this requires additional data processing and increases the dataset size in terms of the features space.

Development of an intelligent control system for thermoplastics injection molding might not only help to directly optimize the IM process, but also contribute to elimination of unnecessary plastics waste and energy consumption. If trial and error method is used less often and prediction models are utilized to select the proper process parameters settings instead, the amount of produced scrap and energy consumed will be reduced. This can to certain extent contribute to the Goal 9 "Industry, innovation and infrastructure" and Goal 12 "Responsible consumption and production" of the Sustainable Development Goals (SDGs) presented by the United Nations [154]. The Goal number 9 focuses on development of infrastructure that provides the basic physical systems, industrialization for driving the economic growth and innovation to advance the technological capabilities and to prompt development of new skills. The Goal 12 is focused on sustainable development that requires minimization of natural resources and toxic materials used. The intelligent control system can address both of them in a sense that it contributes to the technology advancement, minimization of plastic material use and the corresponding energy consumption.

The process of designing the framework for the intelligent system development described in Chapter 5 is considered as part of the verification in this work. The predictive models that are developed within the study are parts of the corresponding modules of the proposed system and are, in their turn, validated using the 70%-30% training-test set division, as well as 5-folds cross-validation on the train set. These steps help to show if the models face challenges related to the overfitting and if they are able to generalize on the previously unseen data. Connection of the developed parts of the system and its application in the real industrial setting will help to confirm the system's and the framework's validity in future. This, however, is not addresses in this study due to the time limitations.

11.5 Validation and limitation of results

The results presented in this work have certain limitations in terms of conducted quality characteristics measurements. The main types of measurement errors are systematic and random. While random errors are extremely hard to avoid and are related to things like equipment precision, systematic errors appear due to how well/ correct this equipment is used or an experiment is controlled. Some of the main sources of systematic errors are humans, instruments/ equipment and selected methods. A human might make an error due to a distraction or tiredness, equipment might not be properly calibrated, while selected method might be not the most suitable for the task at hand.

In case of the dimensional characteristics' measurements, both width and thickness were measured in three points along the narrow section of specimens and these values were then averaged to obtain the desired characteristic. This might be a source of error, because if the measurements were conducted in more points, the obtained value might have been more precise. At the same time, decision to take the measurements in three points was due to the limited amount of time provided to perform the study and a large number of samples at hand (in total 798 specimens were measured). Another source of error here was putting each of the specimens in the fixture designed to hold them during the

measurement. Since this operation was performed by a human, a slight deviation in placement of the specimens in the fixture might have been present.

When it comes to definition of mechanical properties and the corresponding tensile testing, several possible sources of errors were present. The first source is a human performing the tests. It was necessary to add two white dots on each of the specimens to assist the video extensometer's work. This was done by hand using an Instron stencil. Next, the placement of the specimens in the tensile testing machine was also performed by a human and might have led to slight deviations in placement relative to the machine's grippers. The second source of error is equipment and the video extensometer that does not always precisely recognize the white dots on the specimen's surface. The third source is the method. For example, a method that is used for calculation of the break point of the tested specimen, as different methods might calculate it differently. This source might be a reason for the obtained performance of the tensile strain at break predictive models.

All of these sources of errors add to the overall uncertainty of the performed measurements and potential errors in the datasets. It was attempted to avoid these errors through removing the data samples that appeared to be outliers, as well as through controlling the experimental process along the way. However, it can not be guaranteed that all of them were successfully eliminated. Possible presence of the measurement errors in the data is one of the limitations of this work, which directly influences the quality of the obtained models.

Chapter 12

Conclusions and future work

The moment you doubt whether you can fly, you cease for ever to be able to do it.

– J. M. Barrie, "Peter Pan"

This chapter concludes presented in this dissertation work and describes ideas that need to be addressed in the future.

12.1 Conclusions

In this PhD study a new approach to control of the injection molding process is presented. It was investigated how various IM process parameters influence dimensional and mechanical properties of focus parts. The main challenge that has been addressed was how to develop an intelligent system that would allow to avoid use of trial and error method for selection of process settings when fabricating injection molded parts.

The thesis is based on analysis of data acquired during four experiments, where dogbone specimens with 4 mm thickness are produced in experiments 1, 3 and 4, and 15 mm thick specimens during experiment 2. Virgin HDPE material is used in the first two experiments, while recycled material from two suppliers is utilized in experiments 3 and 4 correspondingly. DOE that includes variation of 8 parameters is created for experiment 1, 6 parameters for experiment 2 and the same DOE with variation of 7 parameters is used in experiments 3 and 4. During experiment 1, 41 machine and process parameters are logged per production cycle, during experiment 2 – 65 parameters, and 52 parameters during experiments 3-4. A procedure for the IMM data acquisition and preprocessing is proposed, as before developing the predictive models, the data needs to go through several steps to obtain a high-quality dataset. If the dataset quality is not acceptable, no matter which methods are applied, the predictive models will not have high characteristics scores. Therefore, data cleaning, integration, normalization and feature selection become important steps within the ML approach.

Model-based systems engineering approach was applied to develop a framework for designing an intelligent control system for thermoplastics injection molding. The system includes 7 computational core (CC) modules, where each is responsible for a certain prediction or optimization task. The connections between the modules are also described and used as guidelines for the data preprocessing and analysis. Prototypes for 5 modules are presented in this work, as well as the corresponding prediction models for dimensional and mechanical characteristics. The models are trained using the data obtained from experiments. In total 798 specimens were produced during the corresponding 435 IMM runs. Producing a large number of specimens with different process settings allowed to demonstrate that it is possible to create prediction models for quality characteristics of focus parts.

Prior to development of the modules and the corresponding prediction models, removal of missing, redundant data, as well as outliers is performed as part of the data pre-

processing. Next the FS is applied to select the most influential process parameters that need to be included into the model. Five feature selection methods were used on the available data: Pearson's correlation, RFE, Spearman's Correlation, CFS and RReliefF. According to the parameters scoring by the selected FS methods, parameters that are to be included in the models are cushion characteristics (cushion value after holding pressure, average value, the smallest value), holding pressure, pressure at switchover, cooling time, material, mold temperature, holding pressure time, injection speed, backpressure and flow number of plastics (specific parameter utilized by the ENGEL IMM used during the experimental work). Most of these parameters are relevant for creation of models for dimensional and mechanical characteristics prediction. RReliefF has shown to be the best performing method in cases of FS for width, Young's modulus, tensile strength, tensile strain at break, while Spearman's correlation for thickness. Application of the FS methods does not only help to create models of better quality, but also contributes to the knowledge of which parameters need to be logged at all.

A different number of parameters was used to develop models based on the individual experiments datasets, depending on the experiment number. When it comes to joined datasets, the parallel dataset included 20 parameters and the sequential one 19 parameters. As a result, 12 models were proposed:

4. A predictive model with $R^2 = 0.95$, RMSE = 0.05 and correlation coefficient = 0.94 (test set of sequential joined dataset) for dimensional properties (width and thickness) was developed using Random Forest ML method. The models created using MLP, Decision Tree Regressor and kNN methods have higher R^2 and correlation coefficient, but also worse RMSE. The model belongs to the CC1 module of the intelligent control system for IM.
5. A Random Forest predictive model with $R^2 = 0.92$, RMSE = 28.32 and correlation coefficient = 0.97 for the test set of sequential joined dataset for mechanical properties (tensile modulus, tensile strength, tensile strain at break) is proposed. The model has a significantly higher RMSE in comparison to the dimensional properties model due to having larger values of the focus variables. The model is part of the CC2 module of the proposed intelligent system.
6. Predictive models for dimensional properties based on individual experiments data series datasets have lower quality characteristics than those developed using general datasets. For example, a model trained on the experiment 1 cushion data series dataset has $R^2 = 0.79$, RMSE = 0.09, correlation coefficient = 0.92 created using kNN ML method. Models developed using data series datasets from the rest of experiments have similar quality scores. In most of the cases models trained using the general datasets show better results than those trained on the data series datasets. These models are part of the CC4 module.
7. Predictive models for the mechanical properties using the separate experiments data series datasets are also proposed. However, their scores are too low, and these models require more work. They were developed as part of the CC5 module of the intelligent system.
8. In addition, models for prediction of dimensional deviations are created as part of the CC6 module. The best performing model is developed using the Random Forest machine learning method and sequential joined dataset. It has $R^2 = 0.92$,

RMSE = 0.06 and correlation coefficient of 0.94. A Decision Tree Regressor model has the same scores and can also be used.

In cases when several ML algorithms have models with similar quality scores, a model created by the most robust method is suggested for utilization. For example, when choosing between a Random Forest model and a Decision Tree Regressor model, a model by Random Forest is better to be used, as this method is usually more robust in terms of overfitting. At the same time, if a highly interpretable model is needed, Decision Tree Regressor should be chosen instead.

All the proposed models have constraints related to possible measurement errors, being trained on data collected from the same IMM, using only virgin and two types of recycled HDPE material, as well as being limited to the same geometry with two thickness values.

12.2 Future work

Based on the results and limitations of this work, the following aspects should be considered:

- More experimental work needs to be conducted, and more relevant data collected. Different machines, focus parts geometries and materials should be taken into consideration. Collection of data from a real production environment, where parts with complex geometry are manufactured would be beneficial. The datasets and the models are to be updated correspondingly. Possibility of using simulation and experimental data together needs to be investigated to see if the datasets size can be increased this way.
- Other ways of obtaining the quality data need to be reviewed. For example, if a final part has a more complicated geometry with various values of width and thickness along the part's body, procedures for taking this into account need to be established. In addition, new ways of obtaining the mechanical properties data that do not involve destructive testing are to be investigated.
- Transfer learning approach should be taken into account so that the obtained knowledge is stored and applied for solving similar tasks on different IMMs, for different parts and materials.
- The material parameter needs to be included into the datasets not by using a material code, but rather a real material characteristic, such as, for example, viscosity.
- A database module that would store data logged during various experiments needs to be developed. This is important for storing data in an organized and concise manner to facilitate its reuse and sharing.
- All the remaining intelligent system modules need to be developed and smooth connection between them should be established.
- Fuzzy methods should be considered for future development and improvement of the intelligent control system for thermoplastics injection molding. This would allow to predict a range of values for the desired quality characteristic of focus part rather than a single value. A range for correction can be provided this way [4].

The author of this work hopes that it will bring certain value and additional knowledge to the injection molding and manufacturing community. And that more intelligent machines capable of providing decision-support and higher degree of autonomy are soon a reality.

List of publications

1. Ogorodnyk, O., Lyngstad, O. V., Larsen, M., Martinsen, K. (2021). Prediction of Width and Thickness of Injection Molded Parts Using Machine Learning Methods. In EcoDesign and Sustainability (pp. 455-469). Springer, Singapore.

2. Ogorodnyk, O., Larsen, M., Martinsen, K, Lyngstad, O. V. (2021). Development of Application Programming Interface Prototype for Injection Molding Machines. Procedia CIRP, 97, 453-458.

3. Ogorodnyk, O., Larsen, M., Lyngstad, O. V., Martinsen, K. (2020). Towards a general application programming interface (API) for injection molding machines. PeerJ Computer Science, 6, e302.

4. Ogorodnyk, O., Lyngstad, O. V., Larsen, M., Martinsen, K. (2019). Application of feature selection methods for defining critical parameters in thermoplastics injection molding. Procedia CIRP, 81, 110-114.

5. Ogorodnyk, O., Lyngstad, O. V., Larsen, M., Wang, K., Martinsen, K. (2018). Application of machine learning methods for prediction of parts quality in thermoplastics injection molding. In International Workshop of Advanced Manufacturing and Automation (pp. 237-244). Springer, Singapore.

6. Ogorodnyk, O., Martinsen, K. (2018). Monitoring and control for thermoplastics injection molding – A review. Procedia CIRP, 67, 380-385.

Bibliography

1. Osswald, T.A., Hernandez-Ortiz, J.P. Polymer Processing Modeling and Simulation, Hanser Gardner Publication. Inc, Cincinnati. 2006:140-50.

2. Todd, R.H., Allen, D.K., Alting, L. Manufacturing processes reference guide: Industrial Press Inc.; 1994.

3. Mendibil, X., Llanos, I., Urreta, H., Quintana, I. In process quality control on micro-injection moulding: the role of sensor location. The International Journal of Advanced Manufacturing Technology. 2016:1-10.

4. Gao, H., Zhang, Y., Fu, Y., Mao, T., Zhou, H., Li, D. Process parameters optimization using a novel classification model for plastic injection molding. The International Journal of Advanced Manufacturing Technology. 2018;94(1-4):357-70.

5. Lotti, C., Ueki, M., Bretas, R. Prediction of the shrinkage of injection molded iPP plaques using artificial neural networks. Journal of Injection Molding Technology. 2002;6(3):157.

6. Guilong, W., Guoqun, Z., Huiping, L., Yanjin, G. Analysis of thermal cycling efficiency and optimal design of heating/cooling systems for rapid heat cycle injection molding process. Materials & Design. 2010;31(7):3426-41.

7. Zhao, P., Zhou, H., He, Y., Cai, K., Fu, J. A nondestructive online method for monitoring the injection molding process by collecting and analyzing machine running data. The International Journal of Advanced Manufacturing Technology. 2014;72(5-8):765-77.

8. Dumitrescu, O.R., Baker, D.C., Foster, G.M., Evans, K.E. Near infrared spectroscopy for in-line monitoring during injection moulding. Polymer testing. 2005;24(3):367-75.

9. Vrabič, R., Kozjek, D., Butala, P. Knowledge elicitation for fault diagnostics in plastic injection moulding: A case for machine-to-machine communication. CIRP Annals. 2017;66(1):433-6.

10. Negri, E., Fumagalli, L., Macchi, M. A Review of the Roles of Digital Twin in CPS-based Production Systems. Procedia Manufacturing. 2017;11:939-48.

11. Kull, H. Mass customization: Opportunities, methods, and challenges for manufacturers: Apress; 2015.

12. Saldivar, A.A.F., Goh, C., Li, Y., Yu, H., Chen, Y., editors. Attribute identification and predictive customisation using fuzzy clustering and genetic search for Industry 4.0 environments. Software, Knowledge, Information Management & Applications (SKIMA), 2016 10th International Conference on; 2016: IEEE.

13. Lee, J., Kao, H.A., Yang, S.H. Service innovation and smart analytics for Industry 4.0 and big data environment. Proc Cirp. 2014;16:3-8.

14. Mourtzis, D. Simulation in the design and operation of manufacturing systems: state of the art and new trends. International Journal of Production Research. 2020;58(7):1927-49.

15. Tao, F., Qi, Q., Liu, A., Kusiak, A. Data-driven smart manufacturing. Journal of Manufacturing Systems. 2018;48:157-69.

16. Mourtzis, D., Vlachou, E., Milas, N. Industrial big data as a result of IoT adoption in manufacturing. Procedia cirp. 2016;55:290-5.

17. Ogorodnyk, O., Larsen, M., Lyngstad, O.V., Martinsen, K. Towards a general application programming interface (API) for injection molding machines. PeerJ Computer Science. 2020;6:e302.

18. Tellaeche, A., Arana, R. Rapid data acquisition system for complex algorithm testing in plastic molding industry. International Journal of Mechanical, Aerospace, Industrial, Mechatronic Manufacturing Engineering. 2013;7(7):1391-5.

19. Tellaeche, A., Arana, R.J.I.J.o.M., Aerospace, Industrial, Mechatronic, Engineering, M. Rapid data acquisition system for complex algorithm testing in plastic molding industry. 2013;7(7):1391-5.

20. Cheng, H.-P., Cheng, C.-S., editors. A support vector machine for recognizing control chart patterns in multivariate processes. Proceedings the 5th Asian Quality Congress; 2007.

21. Yin, S., Ding, S.X., Xie, X., Luo, H.J.I.T.o.I.E. A review on basic data-driven approaches for industrial process monitoring. 2014;61(11):6418-28.

22. Haq, N.F., Onik, A.R., Hridoy, M.A.K., Rafni, M., Shah, F.M., Farid, D.M. Application of machine learning approaches in intrusion detection system: a survey. IJARAI-International Journal of Advanced Research in Artificial Intelligence. 2015;4(3):9-18.

23. Al'Aref, S.J., Anchouche, K., Singh, G., Slomka, P.J., Kolli, K.K., Kumar, A., et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. European heart journal. 2019;40(24):1975-86.

24. Peters, J. Machine learning for motor skills in robotics. KI-Künstliche Intelligenz. 2008;2008(4):41-3.

25. Carbonneau, R., Laframboise, K., Vahidov, R. Application of machine learning techniques for supply chain demand forecasting. European Journal of Operational Research. 2008;184(3):1140-54.

26. Wuest, T., Weimer, D., Irgens, C., Thoben, K.-D. Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research. 2016;4(1):23-45.

27. Schreiber, A. Regelung des Spritzgießprozesses auf Basis von Prozessgrößen und im Werkzeug ermittelter Materialdaten: RWTH Aachen; 2011.

28. Hopmann, C., Ressmann, A., Reiter, M., Stemmler, S., Abel, D. A self-optimising injection moulding process with model-based control system parameterisation. International Journal of Computer Integrated Manufacturing. 2016;29(11):1190-9.

29. Kim, W.-W., Gang, M.G., Min, B.-K., Kim, W.-B. Experimental and numerical investigations of cavity filling process in injection moulding for microcantilever structures. The International Journal of Advanced Manufacturing Technology. 2014;75(1-4):293-304.

30. Jurkovic, Z., Cukor, G., Brezocnik, M., Brajkovic, T. A comparison of machine learning methods for cutting parameters prediction in high speed turning process. Journal of Intelligent Manufacturing. 2018;29(8):1683-93.

31. Dang, X.-P. General frameworks for optimization of plastic injection molding process parameters. Simulation Modelling Practice and Theory. 2014;41:15-27.

32. Ma, M.-D., Wong, D.S.-H., Jang, S.-S., Tseng, S.-T. Fault detection based on statistical multivariate analysis and microarray visualization. IEEE Transactions on Industrial Informatics. 2010;6(1):18-24.

33. Yao, Y., Chen, T., Gao, F. Multivariate statistical monitoring of two-dimensional dynamic batch processes utilizing non-Gaussian information. Journal of Process Control. 2010;20(10):1188-97.

34. Scikit-Learn. Choosing the right estimator [Available from: http://scikit-learn.org/stable/tutorial/machine_learning_map/index.html.

35. Yang, Y., Gao, F. Injection molding product weight: online prediction and control based on a nonlinear principal component regression model. Polymer Engineering & Science. 2006;46(4):540-8.

36. ISO. ISO 527-2:2012 Tensile Properties of Plastics 2012 [Available from: https://www.iso.org/standard/56046.html.

37. Zheng, R., Tanner, R.I., Fan, X.-J. Injection molding: integration of theory and modeling methods: Springer Science & Business Media; 2011.

38. Malloy, R.A. Plastic part design for injection molding: Hanser Publishers New York; 1994.

39. Rubin, I.I., Rubin, A.A. Injection molding: theory and practice: Wiley New York; 1972.

40. Vaupel, E. Arthur Eichengrün—Tribute to a Forgotten Chemist, Entrepreneur, and German Jew. Angewandte Chemie International Edition. 2005;44(22):3344-55.

41. Merrill, A.M. Plastics Technology: Rubber/Automotive Division of Hartman Communications, Incorporated; 1955.

42. Bryce, D.M. Plastic injection molding: manufacturing process fundamentals: Society of Manufacturing Engineers; 1996.

43. Reeves, C.A., Bednar, D.A. Defining quality: alternatives and implications. Academy of management Review. 1994;19(3):419-45.

44. Bryce, D.M. Plastic injection molding: Mold design and construction fundamentals: Society of Manufacturing Engineers; 1998.

45. TorayPlastics. Common Injection molding defects and their countermeasures 2020 [cited 2020 10.11.2020]. Available from: https://www.toray.jp/plastics/en/amilan/technical/tec_014.html.

46. Polyplastics. Measures for preventing molding defects 2020 [cited 2020 10.11.2020]. Available from: https://www.polyplastics.com/en/support/mold/durafide/pps07.html.

47. GPTmold. Injection molding defects and countermeasures 2020 [cited 2020 10.11.2020]. Available from: http://www.gptmold.com/molding-defects-and-countermeasures.html.

48. Ecomolding. Injection molding defects' causes and countermeasures 2018 [cited 2020 10.11.2020]. Available from: https://www.ecomolding.com/injection-molding-defects/.

49. Wikibooks. Plastics Molding & Manufacturing/Defects 2020 [cited 2020 10.11.2020]. Available from: https://en.wikibooks.org/w/index.php?title=Plastics_Molding_%26_Manufacturing/Defects&oldid=3752684.

50. McCrum, N.G., Buckley, C., Bucknall, C.B., Bucknall, C. Principles of polymer engineering: Oxford University Press, USA; 1997.

51. Gedde, U. Polymer physics: Springer Science & Business Media; 1995.

52. Encyclopaedia_Britannica. Polyethylene: Encyclopædia Britannica, inc.; 2019 [updated 15.11.2019; cited 2020 21.04.2020]. Available from: https://www.britannica.com/science/polyethylene.

53. Colak, O.U., Dusunceli, N. Modeling Viscoelastic and Viscoplastic Behavior of High Density Polyethylene (HDPE). Journal of Engineering Materials and Technology. 2006;128(4):572-8.

54. ISO. ISO 527-1:2019. Plastics - Determination of tensile properties - Part 1: General Principles. 2019.

55. Kurt, M., Kamber, O.S., Kaynak, Y., Atakok, G., Girit, O. Experimental investigation of plastic injection molding: Assessment of the effects of cavity pressure and mold temperature on the quality of the final products. Materials & Design. 2009;30(8):3217-24.

56. Kurt, M., Kaynak, Y., Kamber, O.S., Mutlu, B., Bakir, B., Koklu, U. Influence of molding conditions on the shrinkage and roundness of injection molded parts. The International Journal of Advanced Manufacturing Technology. 2010;46(5-8):571-8.

57. Gao, R.X., Tang, X., Gordon, G., Kazmer, D.O. Online product quality monitoring through in-process measurement. CIRP Annals-Manufacturing Technology. 2014;63(1):493-6.

58. Panchal, R.R., Kazmer, D.O. In-situ shrinkage sensor for injection molding. Journal of Manufacturing Science and Engineering. 2010;132(6):064503.

59. Zhang, Y., Mao, T., Huang, Z., Gao, H., Li, D. A statistical quality monitoring method for plastic injection molding using machine built-in sensors. The International Journal of Advanced Manufacturing Technology. 2016;85(9-12):2483-94.

60. Hopmann, C., Abel, D., Heinisch, J., Stemmler, S. Self-optimizing injection molding based on iterative learning cavity pressure control. Production Engineering. 2017:1-10.

61. Karbasi, H., Reiser, H., editors. Smart Mold: Real-Time in-Cavity Data Acquisition. First Annual Technical Showcase & Third Annual Workshop, Canada; 2006: Citeseer.

62. Tsai, K.-M., Luo, H.-J. An inverse model for injection molding of optical lens using artificial neural network coupled with genetic algorithm. Journal of Intelligent Manufacturing. 2017;28(2):473-87.

63. Guo, W., Hua, L., Mao, H., Meng, Z. Prediction of warpage in plastic injection molding based on design of experiments. Journal of Mechanical Science and Technology. 2012;26(4):1133-9.

64. Rosato, D.V., Rosato, M.G. Injection molding handbook: Springer Science & Business Media; 2012.

65. Kitayama, S., Yokoyama, M., Takano, M., Aiba, S. Multi-objective optimization of variable packing pressure profile and process parameters in plastic injection molding

for minimizing warpage and cycle time. The International Journal of Advanced Manufacturing Technology. 2017;92(9-12):3991-9.

66. Hermann, M., Pentek, T., Otto, B., editors. Design principles for industrie 4.0 scenarios. 2016 49th Hawaii international conference on system sciences (HICSS); 2016: IEEE.

67. Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., et al. Cyber-physical systems in manufacturing. CIRP Annals. 2016;65(2):621-41.

68. Lee, H., Ryu, K., Cho, Y. A framework of a smart injection molding system based on real-time data. Procedia Manufacturing. 2017;11:1004-11.

69. MacDougall, W. Industrie 4.0: Smart manufacturing for the future: Germany Trade & Invest; 2014.

70. Mamat, A., Trochu, T., Sanschagrin, B. Analysis of shrinkage by dual kriging for filled and unfilled polypropylene molded parts. Polymer Engineering & Science. 1995;35(19):1511-20.

71. Bushko, W.C., Stokes, V.K. Solidification of thermoviscoelastic melts. Part I: Formulation of model problem. Polymer Engineering & Science. 1995;35(4):351-64.

72. Ogorodnyk, O., Martinsen, K. Monitoring and Control for Thermoplastics Injection Molding A Review. Procedia CIRP. 2018;67:380-5.

73. Zhao, C., Gao, F., Sun, Y. Between-phase calibration modeling and transition analysis for phase-based quality interpretation and prediction. AIChE Journal. 2013;59(1):108-19.

74. Berti, G., Monti, M. A virtual prototyping environment for a robust design of an injection moulding process. Computers & Chemical Engineering. 2013;54:159-69.

75. Fu, J., Ma, Y. Computer-aided engineering analysis for early-ejected plastic part dimension prediction and quality assurance. The International Journal of Advanced Manufacturing Technology. 2018;98(9-12):2389-99.

76. Ansys. Ansys 2020 [cited 2020 10.11.2020]. Available from: https://www.ansys.com/.

77. Zhao, P., Wang, S., Ying, J., Fu, J. Non-destructive measurement of cavity pressure during injection molding process based on ultrasonic technology and Gaussian process. Polymer Testing. 2013;32(8):1436-44.

78. Johnston, S., McCready, C., Hazen, D., VanDerwalker, D., Kazmer, D. On-line multivariate optimization of injection molding. Polymer Engineering & Science. 2015;55(12):2743-50.

79. Zhang, S., Dubay, R., Charest, M. A principal component analysis model-based predictive controller for controlling part warpage in plastic injection molding. Expert Systems with Applications. 2015;42(6):2919-27.

80. Wang, J., Xie, P., Ding, Y., Yang, W. On-line testing equipment of P–V–T properties of polymers based on an injection molding machine. Polymer Testing. 2009;28(3):228-34.

81. Yang, Y., Yang, B., Zhu, S., Chen, X. Online quality optimization of the injection molding process via digital image processing and model-free optimization. Journal of Materials Processing Technology. 2015;226:85-98.

82. Park, H.S., Nguyen, T.T. Optimization of injection molding process for car fender in consideration of energy efficiency and product quality. Journal of Computational Design and Engineering. 2014;1(4):256-65.

83. Zhang, J.Z. Development of an in-process Pokayoke system utilizing accelerometer and logistic regression modeling for monitoring injection molding flash. The International Journal of Advanced Manufacturing Technology. 2014;71(9-12):1793-800.

84. Xu, Y., Zhang, Q., Zhang, W., Zhang, P. Optimization of injection molding process parameters to improve the mechanical performance of polymer product against impact. The International Journal of Advanced Manufacturing Technology. 2015;76(9-12):2199-208.

85. Zhang, T., Chen, K., Liu, G., Zheng, X., editors. Injection Molding Process Optimization of Polypropylene using Orthogonal Experiment Method Based on Tensile Strength. IOP Conference Series: Materials Science and Engineering; 2019: IOP Publishing.

86. Ozcelik, B. Optimization of injection parameters for mechanical properties of specimens with weld line of polypropylene using Taguchi method. International Communications in Heat and Mass Transfer. 2011;38(8):1067-72.

87. Kuo, C.-F.J., Su, T.-L., Li, Y.-C. Construction and analysis in combining the Taguchi method and the back propagation neural network in the PEEK injection molding process. Polymer-Plastics Technology and Engineering. 2007;46(9):841-8.

88. Heidari, B.S., Oliaei, E., Shayesteh, H., Davachi, S.M., Hejazi, I., Seyfi, J., et al. Simulation of mechanical behavior and optimization of simulated injection molding process for PLA based antibacterial composite and nanocomposite bone screws using central composite design. Journal of the mechanical behavior of biomedical materials. 2017;65:160-76.

89. Fernandez, A., Muniesa, M., Javierre, C. In-line rheological testing of thermoplastics and a monitored device for an injection moulding machine: Application to raw and recycled polypropylene. Polymer testing. 2014;33:107-15.

90. Mehat, N.M., Kamaruddin, S. Investigating the effects of injection molding parameters on the mechanical properties of recycled plastic parts using the Taguchi method. Materials and Manufacturing Processes. 2011;26(2):202-9.

91. Shi, H., Xie, S., Wang, X. A warpage optimization method for injection molding using artificial neural network with parametric sampling evaluation strategy. The International Journal of Advanced Manufacturing Technology. 2013;65(1-4):343-53.

92. Chen, W.-C., Tai, P.-H., Wang, M.-W., Deng, W.-J., Chen, C.-T. A neural network-based approach for dynamic quality prediction in a plastic injection molding process. Expert systems with Applications. 2008;35(3):843-9.

93. Park, H.S., Phuong, D.X., Kumar, S. AI based injection molding process for consistent product quality. Procedia Manufacturing. 2019;28:102-6.

94. Manjunath, P.G., Krishna, P., editors. Prediction and optimization of dimensional shrinkage variations in injection molded parts using forward and reverse mapping of artificial neural networks. Advanced Materials Research; 2012: Trans Tech Publ.

95. Liu, J., Chen, X., Lin, Z., Diao, S. Multiobjective optimization of injection molding process parameters for the precision manufacturing of plastic optical lens. Mathematical Problems in Engineering. 2017;2017.

96. Tsai, K.-M., Luo, H.-J. Comparison of injection molding process windows for plastic lens established by artificial neural network and response surface methodology. The International Journal of Advanced Manufacturing Technology. 2015;77(9-12):1599-611.

97. Yin, F., Mao, H., Hua, L., Guo, W., Shu, M. Back propagation neural network modeling for warpage prediction and optimization of plastic products during injection molding. Materials & design. 2011;32(4):1844-50.

98. Nagorny, P., Pillet, M., Pairel, E., Le Goff, R., Loureaux, J., Wali, M., et al., editors. Quality prediction in injection molding. 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA); 2017: IEEE.

99. Zhu, J., Chen, J.C. Fuzzy neural network-based in-process mixed material-caused flash prediction (FNN-IPMFP) in injection molding operations. The International Journal of Advanced Manufacturing Technology. 2006;29(3-4):308-16.

100. Guo, F., Zhou, X., Liu, J., Zhang, Y., Li, D., Zhou, H. A reinforcement learning decision model for online process parameters optimization from offline data in injection molding. Applied Soft Computing. 2019;85:105828.

101. Kozjek, D., Kralj, D., Butala, P., Lavrač, N. Data mining for fault diagnostics: A case for plastic injection molding. Procedia CIRP. 2019;81:809-14.

102. Woll, S.L., Cooper, D.J., Souder, B.V. Online pattern-based part quality monitoring of the injection molding process. Polymer Engineering & Science. 1996;36(11):1477-88.

103. NIST/SEMATECH. NIST/SEMATECH e-Handbook of Statistical Methods 2013 [cited 2020 10.11.2020]. Available from: http://www.itl.nist.gov/div898/handbook/.

104. Chen, W.-C., Nguyen, M.-H., Chiu, W.-H., Chen, T.-N., Tai, P.-H. Optimization of the plastic injection molding process using the Taguchi method, RSM, and hybrid GA-PSO. The International Journal of Advanced Manufacturing Technology. 2016;83(9-12):1873-86.

105. Zhao, J., Cheng, G., Ruan, S., Li, Z. Multi-objective optimization design of injection molding process parameters based on the improved efficient global optimization algorithm and non-dominated sorting-based genetic algorithm. The International Journal of Advanced Manufacturing Technology. 2015;78(9-12):1813-26.

106. Cheng, J., Liu, Z., Tan, J. Multiobjective optimization of injection molding parameters based on soft computing and variable complexity method. The International Journal of Advanced Manufacturing Technology. 2013;66(5-8):907-16.

107. Kitayama, S., Natsume, S. Multi-objective optimization of volume shrinkage and clamping force for plastic injection molding via sequential approximate optimization. Simulation Modelling Practice and Theory. 2014;48:35-44.

108. Tercan, H., Guajardo, A., Heinisch, J., Thiele, T., Hopmann, C., Meisen, T. Transfer-learning: Bridging the gap between real and simulation data for machine learning in injection molding. Procedia CIRP. 2018;72:185-90.

109. Chaves, M., Sánchez-González, L., Díez, E., Pérez, H., Vizán, A. Experimental assessment of quality in injection parts using a fuzzy system with adaptive membership functions. Neurocomputing. 2019.

110. Liau, Y., Lee, H., Ryu, K., editors. Digital Twin concept for smart injection molding. IOP Conference Series: Materials Science and Engineering; 2018: IOP Publishing.

111. Mitchell, T.M. Machine learning. McGraw-hill New York; 1997.

112. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. From data mining to knowledge discovery in databases. AI magazine. 1996;17(3):37-.

113. Theodoridis, S., Koutroumbas, K. Pattern Recognition. Fourth edition ed: Academic Press; 2009.

114. Kononenko, I., Kukar, M. Machine learning and data mining: introduction to principles and algorithms: Horwood Publishing; 2007.

115. Huber, P. Robust Statistics: John Wiley & Sons; 1981.

116. Rubin, D.B. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.

117. Schafer, J.L., Graham, J.W. Missing data: our view of the state of the art. Psychological methods. 2002;7(2):147.

118. Hall, M.A. Correlation-based feature selection for machine learning. Hamilton, New Zealand: University of Waikato; 1999.

119. Jović, A., Brkić, K., Bogunović, N., editors. A review of feature selection methods with applications. Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2015 38th International Convention on; 2015: IEEE.

120. Sánchez-Maroño, N., Alonso-Betanzos, A., Tombilla-Sanromán, M., editors. Filter methods for feature selection–a comparative study. International Conference on Intelligent Data Engineering and Automated Learning; 2007: Springer.

121. SciPy. Pearson correlation coefficient 2020 [cited 2020 10.11.2020]. Available from: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.pearsonr.html.

122. SciPy. Spearman correlation coefficient 2020 [cited 2020 10.11.2020]. Available from: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.spearmanr.html.

123. Kira, K., Rendell, L.A. A practical approach to feature selection. Machine Learning Proceedings 1992: Elsevier; 1992. p. 249-56.

124. Robnik-Šikonja, M., Kononenko, I. Theoretical and empirical analysis of ReliefF and RReliefF. Machine learning. 2003;53(1-2):23-69.

125. Scikit-Learn. Recursive feature elimination 2020 [cited 2020 10.11.2020]. Available from: https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html.

126. Scikit-Learn. Decision Tree Regressor 2020 [cited 2020 10.11.2020]. Available from: https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html#sklearn.tree.DecisionTreeRegressor.

127. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. Scikit-learn: Machine learning in Python. the Journal of machine Learning research. 2011;12:2825-30.

128. Kotsiantis, S.B. Decision trees: a recent overview. Artificial Intelligence Review. 2013;39(4):261-83.

129. Freund, Y., Schapire, R.E., editors. A desicion-theoretic generalization of on-line learning and an application to boosting. European conference on computational learning theory; 1995: Springer.

130. Breiman, L. Random forests. Machine learning. 2001;45(1):5-32.

131. Liaw, A., Wiener, M. Classification and regression by randomForest. R news. 2002;2(3):18-22.

132. Creswell, J.W. Research design: Qualitative and mixed methods approaches: London and Thousand Oaks: Sage Publications; 2009.

133. Blaikie, N. Designing social research 2nd Edition: Cambridge: Polity Press; 2010.

134. Goldkuhl, G. Pragmatism vs interpretivism in qualitative information systems research. European journal of information systems. 2012;21(2):135-46.

135. Saunders, M., Lewis, P., Thornhill, A. Understanding research philosophies and approaches. Research methods for business students. 2009;4(106-135).

136. Esteco. modeFrontier 2020 [cited 2020 23.11.2020]. Available from: https://www.esteco.com/modefrontier.

137. KunbusGmbH. RevPi 2020 [cited 2020 23.11.2020]. Available from: https://revolution.kunbus.com/revpi-core/.

138. Zeiss. Zeiss DuraMax 2020 [cited 2020 23.11.2020]. Available from: https://www.zeiss.no/metrology/products/systems/coordinate-measuring-machines/production-cmms/duramax.html.

139. McKay, M.D., Beckman, R.J., Conover, W.J. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics. 1979;21(2):239-45.

140. INCOSE. What is Systems Engineering? [Available from: https://www.incose.org/AboutSE/WhatIsSE.

141. Long, D., Scott, Z. A primer for model-based systems engineering: Lulu. com; 2011.

142. ISO. ISO 16012:2015 Plastics — Determination of linear dimensions of test specimens. 2015. p. 9.

143. WEKA - Waikato Environment for Knowledge Analysis 2020 [cited 2020 10.11.2020]. Available from: https://www.cs.waikato.ac.nz/ml/weka/.

144. Haskins, C. Systems engineering analyzed, synthesized, and applied to sustainable industrial park development. 2008.

145. Mitchell, R.K., Agle, B.R., Wood, D.J. Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. Academy of management review. 1997;22(4):853-86.

146. Standardization, G.I.f. DIN 16742:2013 Plastics mouldings: tolerances and acceptance conditions. 2013.

147. Annicchiarico, D., Alcock, J.R. Review of factors that affect shrinkage of molded part in injection molding. Materials and Manufacturing Processes. 2014;29(6):662-82.

148. Jansen, K., Van Dijk, D., Husselman, M. Effect of processing conditions on shrinkage in injection molding. Polymer Engineering & Science. 1998;38(5):838-46.

149. Chen, C.-C., Su, P.-L., Lin, Y.-C. Analysis and modeling of effective parameters for dimension shrinkage variation of injection molded part with thin shell feature using response surface methodology. The International Journal of Advanced Manufacturing Technology. 2009;45(11-12):1087.

150. Fernandes, C., Pontes, A., Viana, J., Gaspar-Cunha, A. Using multiobjective evolutionary algorithms in the optimization of operating conditions of polymer injection molding. Polymer Engineering & Science. 2010;50(8):1667-78.

151. Pomerleau, J., Sanschagrin, B. Injection molding shrinkage of PP: Experimental progress. Polymer Engineering & Science. 2006;46(9):1275-83.

152. Sha, B., Dimov, S., Griffiths, C., Packianather, M. Investigation of micro-injection moulding: Factors affecting the replication quality. Journal of Materials Processing Technology. 2007;183(2-3):284-96.

153. Amaranan, S., Manonukul, A. Aspect ratio of green MIM parts on shrinkage during sintering. The Minerals, Metals and Materials Society – 3rd International Conference on Processing Materials for Properties; Bangkok, Thailand2008.

154. UN. Sustainable Development Goals [cited 2020 24.11.2020]. Available from: https://www.un.org/sustainabledevelopment/sustainable-development-goals/

Appendix A

Design of experiments for experiments 1-4

Table A.1. Design of experiment for experiment 1

Holding pressure	Holding pressure time	Backpressure	Cooling time	Injection speed	Screw speed	Barrel temp./ Heating_cyl1_z1_set	Tool temp.
65	60	20	30	25	0.45	190	50
30	30	37	10	75	0.733	190	70
40	20	4	5	10	0.55	190	70
55	50	58	4.9	60	0.6	195	35
70	65	11	25	65	0.35	195	40
60	45	7	4.6	30	0.7	200	45
45	25	95	4.6	15	0.733	200	60
60	80	40	5	60	0.733	205	60
30	55	78	30	40	0.5	200	65
40	30	78	20	35	0.2	210	70
20	75	68	20	50	0.733	210	25
45	50	31	20	20	0.7	205	25
25	25	30	5	60	0.25	215	28
50	80	18	30	30	0.65	215	28
25	45	26	15	80	0.3	225	30
15	65	53	15	35	0.8	230	30
10	35	63	4.76	75	0.6	235	30
50	80	82	5	45	0.55	240	35
25	55	75	10	65	0.65	245	35
85	35	49	10	50	0.4	245	28
75	40	46	15	75	0.4	240	25
85	70	93	25	70	0.55	215	50
35	35	53	10	80	0.35	220	50
90	20	62	25	20	0.4	225	45
65	40	0	20	55	0.25	235	45
80	60	44	30	45	0.45	245	45
15	50	13	20	40	0.35	230	40
70	45	70	4.78	55	0.65	235	55
10	70	86	10	25	0.2	250	55
90	75	100	25	10	0.2	250	60
80	25	24	25	15	0.5	225	65
70	70	90	15	30	0.3	220	65

Table A.2. Design of experiment for experiment 2

Holding pressure	Holding pressure time	Backpressure	Cooling time	Injection speed	Screw speed
35	99	58	119	67	0.21
27	121	45	97	48	0.26
44	63	64	77	62	0.26
38	88	72	149	37	0.3
19	132	39	131	40	0.31
31	104	81	115	42	0.33
32	143	48	104	34	0.34
42	82	53	99	43	0.35
55	96	22	135	47	0.36
39	151	36	109	61	0.37
46	107	67	123	33	0.38
21	92	43	91	54	0.4
40	114	63	87	19	0.41
25	129	91	124	50	0.42
33	116	55	127	25	0.43
38	49	31	73	31	0.43
29	91	58	69	46	0.45
34	111	49	111	38	0.46
33	70	70	82	55	0.48
48	124	60	154	52	0.49
30	58	55	176	49	0.51
41	139	51	104	57	0.54
26	75	41	144	45	0.55
36	192	68	47	40	0.63

Table A.3. Design of experiment for experiments 3 and 4

Holding pressure	Holding pressure time	Backpressure	Cooling time	Injection speed	Screw speed	Barrel temp./ Heating_cyl1_z1_set
85	48	10	30	75	0.68	190
85	43	70	30	20	0.68	190
60	43	100	9	40	0.59	195
20	60	40	26	20	0.55	195
15	35	100	13	35	0.2	195
45	60	20	5	15	0.55	200
35	27	60	13	60	0.42	200
75	52	90	13	65	0.33	205
70	18	60	22	65	0.24	205
35	56	80	9	45	0.42	205
25	31	40	22	50	0.72	210

Holding pressure	Holding pressure time	Backpressure	Cooling time	Injection speed	Screw speed	Barrel temp./ Heating_cyl1_z1_set
65	31	70	30	55	0.29	210
75	23	50	18	80	0.33	215
50	18	90	30	55	0.46	215
30	43	100	5	30	0.37	215
15	14	80	22	30	0.5	220
85	39	20	22	80	0.42	220
65	60	10	18	45	0.63	225
55	43	0	9	30	0.2	225
50	14	50	13	40	0.63	230
55	10	20	18	10	0.46	230
65	23	30	5	15	0.55	230
10	39	0	9	55	0.24	235
80	18	30	26	25	0.29	235
15	31	60	9	70	0.24	235
45	56	60	26	70	0.37	240
90	27	10	22	15	0.72	240
25	52	80	13	75	0.46	245
30	35	40	5	60	0.33	245
40	23	50	26	50	0.72	250
75	10	0	30	10	0.5	250
35	48	80	18	40	0.59	250

Appendix B

List of logged process parameters and their description

Table B.1. List of logged process parameters and their description

#	Column name	Actual/Set	Units	Description	Python processing	Experiments where parameter is logged
1.	Last_cycle_time	Actual	seconds	Cycle time of the current production cycle	The 1st value of the series	1-4
2.	Ejector_pos	Actual	cm	Current ejector position	The last value of the series	1, 3, 4
3.	Plast_time	Actual	seconds	Plasticizing time value	The second value of the series (after its updated from the previous cycle)	1-4
4.	Flow_no_plast	Actual	int	Plasticizing number	The second value of the series (after its updated from the previous cycle)	1-4
5.	Flow_n	Set	int	Flow number	The 1st value of the series	1-4
6.	Switchover_time	Set	seconds	Switchover time	The 1st value of the series	1, 2
7.	Screw_speed	Actual	u/min	Screw speed value	Maximum value	1, 2
8.	Injection_time	Actual	seconds	Injection time	The second value of the series (after its updated from the previous cycle)	1-4
9.	Injection_time_set_max	Set	seconds	Maximum set injection time	The 1st value of the series	1, 2
10.	Inject_time_set_min	Set	seconds	Minimum set injection time	The 1st value of the series	1, 2
11.	Plastic_time_set_max	Set	seconds	Maximum set plasticizing time	The 1st value of the series	1, 2
12.	Plastic_time_set_min	Set	seconds	Minimum set plasticizing time	The 1st value of the series	1, 2
13.	Clamp_force_switchov	Actual	kN	Mold clamping force value at switchover	The second value of the series (after its updated from the previous cycle)	1, 2
14.	Cushion_mater	Actual	cm^3	Material cushion value	Average value	1, 2
15.	Cushion_smallest	Actual	cm^3	Smallest cushion value	The second value of the series (after its updated from the previous cycle)	1-4

#	Column name	Actual/Set	Units	Description	Python processing	Experiments where parameter is logged
16.	Cushion_after_hold_pres	Actual	cm^3	Cushion after holding pressure value	The second value of the series (after its updated from the previous cycle)	1-4
17.	Cushion_ideal	Set	cm^3	Ideal set cushion value	The 1st value of the series	1-4
18.	Cushion_smallest_set_max	Set	cm^3	Maximum set cushion value	The 1st value of the series	1, 2
19.	Decomp_after_plast_vol	Actual	cm^3	Decompression volume after plasticizing	The 1st value of the series	1, 3, 4
20.	Switchov_vol	Actual	cm^3	Switchover volume	The second value of the series (after its updated from the previous cycle)	1-4
21.	Spec_pres_switchov	Actual	cm^3	Specific pressure at switchover value	The second value of the series (after its updated from the previous cycle)	1-4
22.	Injection_work	Actual	kWh	Injection work	The second value of the series (after its updated from the previous cycle)	1-4
23.	Speed_peak	Actual	m/s	Speed peak value	Maximum value	1-4
24.	Shot_vol	Set	cm^3	Shot volume	The 1st value of the series	1-4
25.	Holding_pres_time	Set	seconds	Holding pressure time	The 1st value of the series	1, 2
26.	Closing_force	Actual	int	Closing force	The second value of the series (after its updated from the previous cycle)	1, 2
27.	Current_station	Set	int	Current station number	The 1st value of the series	1, 3, 4
28.	Machine_date	Actual	text	Machine date	The 1st value of the series	1-4
29.	Machine_time	Actual	text	Machine time	The last value of the series	1-4
30.	Bad_parts	Actual	int	Bad parts counter	Maximum value	1-4
31.	Shotcounter	Actual	int	Shot counter	Maximum value	1-4
32.	Good_parts	Actual	int	Good parts counter	Maximum value	1-4
33.	Heating_cyl1_z1	Actual	C	Barrel temperature actual value	Average value	1, 2
34.	Heating_cyl1_z1_set	Set	C	Barrel temperature set value	The 1st value of the series	1, 2
35.	Cooling_time	Actual	seconds	Cooling time actual value	The last value of the series	1-4
36.	Parts_count	Actual	int	Parts counter	Maximum value	1-4

#	Column name	Actual/Set	Units	Description	Python processing	Experiments where parameter is logged
37.	Ejector_pos_set_min	Set	cm	Minimum set ejector position	The 1st value of the series	1
38.	Ejector_pos_set_max	Set	cm	Maximum set ejector position	The 1st value of the series	1
39.	Plastic_delay_time_set	Set	cm	Plasticizing delay time	The 1st value of the series	1
40.	Inject_pres_limit	Set	bar	Injection pressure limit	The 1st value of the series	1
41.	Waiting_del	Set	seconds	Waiting delay	The 1st value of the series	1
42.	Cooling_cycletime	Set	seconds	Set cooling time	The 1st value of the series	2-4
43.	Screw_pos	Actual	mm	Screw position value	The last value of the series	2-4
44.	Cycletime_hold_pressure	Set	seconds	Set holding pressure time	The 1st value of the series	2-4
45.	Cushion_end_hold_pres	Actual	cm^3	Cushion value in the end of the holding pressure	The 1st value of the series	2-4
46.	Max_injection_speed_from_graph	Actual	mm/s	Maximum injection speed from the graph	The 1st value of the series	2-4
47.	Plast_volume	Actual	cm^3	Plasticizing volume	The second value of the series (after its updated from the previous cycle)	2-4
48.	Backpressure_peak	Actual	bar	Backpressure peak value	The second value of the series (after its updated from the previous cycle)	2-4
49.	Hold_pres_peak	Actual	bar	Holding pressure peak value	The second value of the series (after its updated from the previous cycle)	2-4
50.	Pre_pres_peak	Actual	bar	Pre-pressure peak value	The second value of the series (after its updated from the previous cycle)	2-4
51.	Injection_rate	Actual	cm/s	Injection rate of the shot	The second value of the series (after its updated from the previous cycle)	2-4
52.	Shot_vol_act	Actual	cm^3	Actual shot value	The second value of the series (after its updated from the previous cycle)	2-4
53.	Torque_mean_cycle	Actual	Nm	Torque mean value	The second value of the series (after its updated from the previous cycle)	2-4
54.	Torque_peak	Actual	Nm	Torque peak value	The second value of the series (after its updated from the previous cycle)	2-4
55.	Clamp_force_relief_time	Set	seconds	Clamping force relief time	The 1st value of the series	2

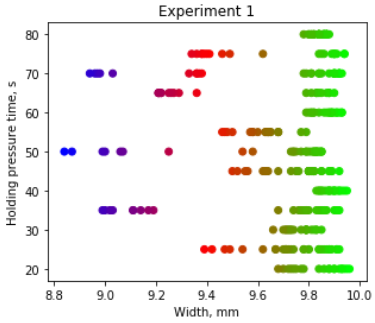
#	Column name	Actual/Set	Units	Description	Python processing	Experiments where parameter is logged
56.	Flow_number_set	Set	bar	Flow number set value	The 1st value of the series	2
57.	Flow_number_act	Actual	bar	Flow number actual value	The 1st value of the series	2
58.	Hold_pressure_correction	Set	bar	Correction of holding pressure	The 1st value of the series	2
59.	Specific_pressure_switchover_set	Set	bar	Set specific pressure at switchover value	The 1st value of the series	2
60.	Injection_speed_peak	Actual	cm/s	Injection speed peak value	The second value of the series (after its updated from the previous cycle)	2
61.	Shot_volume_end_corrected	Set	int	Corrected shot volume value in the cycle end	The 1st value of the series	2
62.	Clamping_force_peak	Actual	kN	Peak clamping force value	The second value of the series (after its updated from the previous cycle)	2
63.	Mold_protection_time	Actual	seconds	Mold protection time	The second value of the series (after its updated from the previous cycle)	2
64.	Open_stroke_peak	Actual	int	Mold opening stroke peak value	The 1st value of the series	2
65.	Heating_cyl1_z2	Actual	C	Actual value barrel temperature zone 2	The 1st value of the series	2
66.	Heating_cyl1_z3	Actual	C	Actual value barrel temperature zone 3	The 1st value of the series	2
67.	Heating_cyl1_z3_set	Set	C	Set value barrel temperature zone 3	The 1st value of the series	2
68.	Nozzle_temp_z5	Actual	C	Nozzle temperature zone 5	The 1st value of the series	2
69.	Nozzle_temp_z6	Actual	C	Nozzle temperature zone 6	The 1st value of the series	2
70.	Fixed_plate_tool1_temp_z9	Set	C	Fixed plate temperature zone 9	The 1st value of the series	2
71.	Fixed_plate_tool2_temp_z10	Set	C	Fixed plate temperature zone 10	The 1st value of the series	2
72.	Mold_pressure	Actual	bar	Mold pressure	Average value	2
73.	Mold_tempature	Actual	C	Mold temperature	Average value	2
74.	Core_movement	Actual	seconds	Cycle time core movement	The 1st value of the series	3-4
75.	Demoduling_time	Actual	seconds	Demolding time	The 1st value of the series	3-4
76.	Closinig_time	Actual	seconds	Closing time	The 1st value of the series	3-4
77.	Cycletime_close	Actual	seconds	Cycle time at the mold closing	The 1st value of the series	3-4

#	Column name	Actual/Set	Units	Description	Python processing	Experiments where parameter is logged
78.	Clamp_force_buildup_time	Actual	seconds	Clamping force buildup time	The 1st value of the series	3-4
79.	Cycletime_open	Actual	seconds	Cycle time at the mold opening	The 1st value of the series	3-4
80.	Cycletime_removal	Actual	seconds	Cycle time	The 1st value of the series	3-4
81.	Cycletime_until_demoulding_end	Actual	seconds	Cycle time value until demolding end	The 1st value of the series	3-4
82.	Ejector_cycletime	Actual	seconds	Ejector cycle time	The 1st value of the series	3-4
83.	Start_flow_number_meas_trigger	Set	cm^3	Start measuring flow number value	The 1st value of the series	3-4
84.	Stop_flow_number_meas_trigger	Set	cm^3	Stop measuring flow number value	The 1st value of the series	3-4
85.	Purging_time	Set	seconds	Purging time	The 1st value of the series	3-4
86.	Decompres_end	Set	cm^3	Decompression end	The 1st value of the series	3-4
87.	Flush_time	Set	seconds	Flushing time	The 1st value of the series	3-4
88.	Cycletime_nozzle	Set	seconds	Nozzle cycle time	The 1st value of the series	3-4
89.	timestamp_imm_machine	Actual	text	IMM timestamp	The last value of the series	3-4

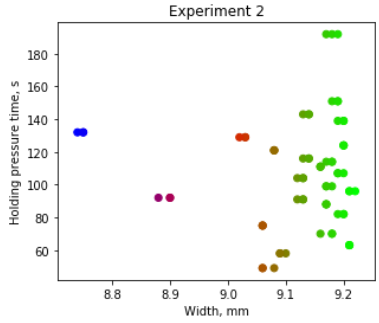
Appendix C

Figures depicting relationships between the DOE parameters and the target quality parameters

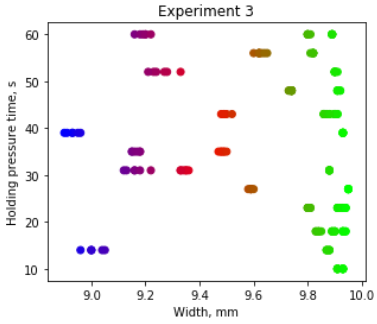
C.1 Width figures



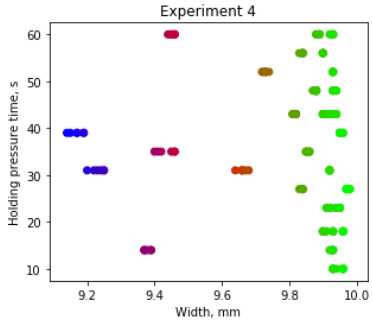
(a) Experiment 1



(b) Experiment 2

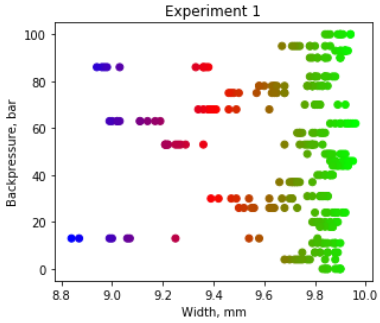


(c) Experiment 3

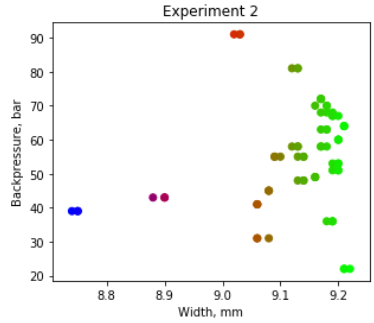


(d) Experiment 4

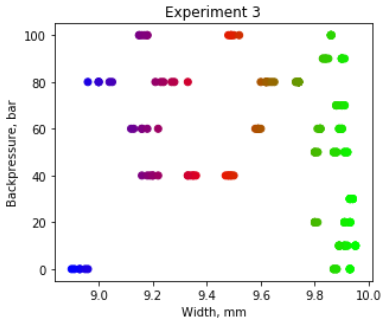
Figure C.1.1. Relationship between the specimen's width and holding pressure time DOE parameter



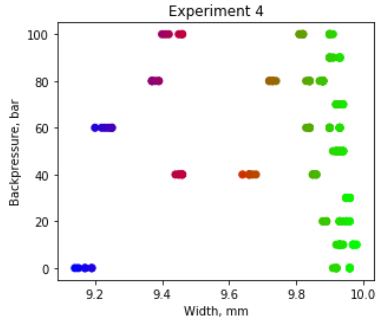
(a) Experiment 1



(b) Experiment 2

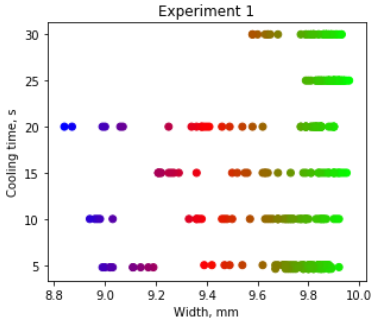


(c) Experiment 3

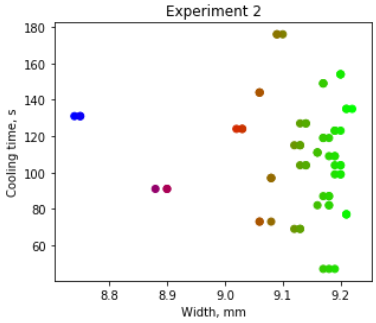


(d) Experiment 4

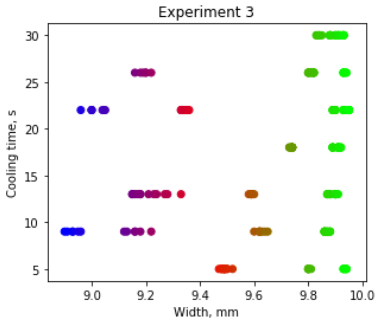
Figure C.1.2. Relationship between the specimen's width and backpressure DOE parameter



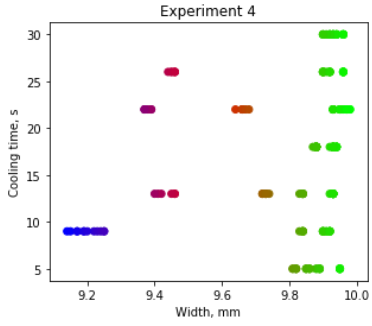
(a) Experiment 1



(b) Experiment 2

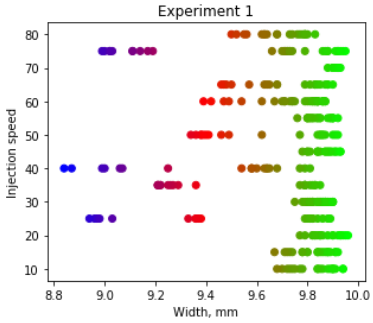


(c) Experiment 3

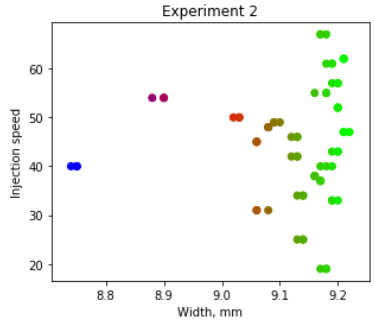


(d) Experiment 4

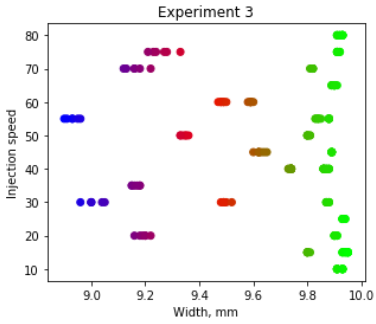
Figure C.1.3. Relationship between the specimen's width and cooling time DOE parameter



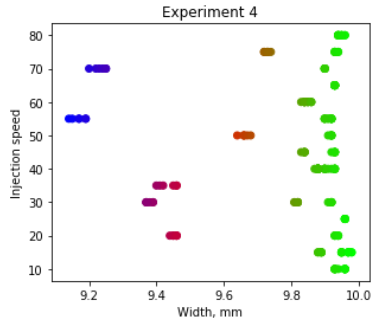
(a) Experiment 1



(b) Experiment 2

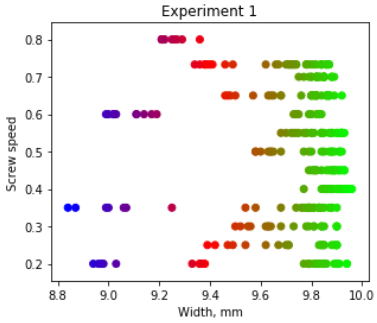


(c) Experiment 3

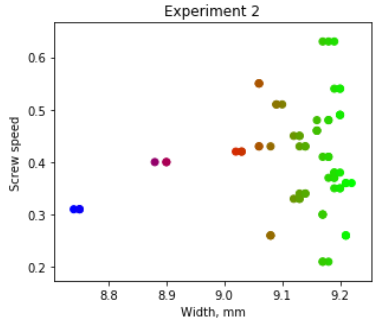


(d) Experiment 4

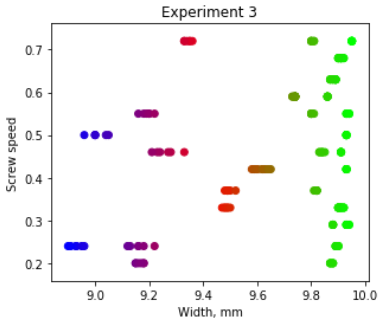
Figure C.1.4. Relationship between the specimen's width and injection speed DOE parameter



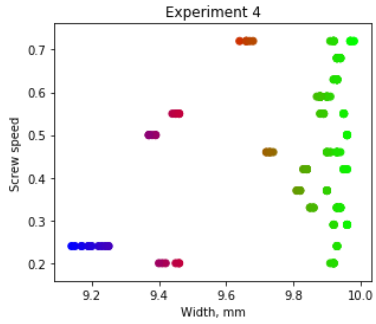
(a) Experiment 1



(b) Experiment 2

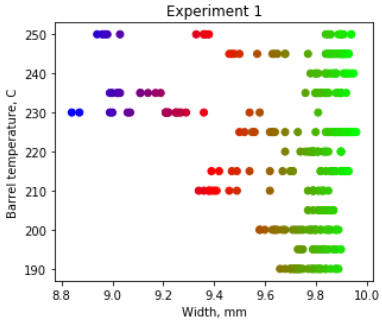


(c) Experiment 3

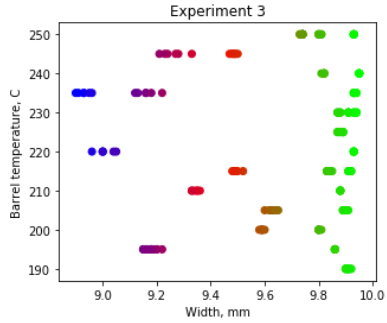


(d) Experiment 4

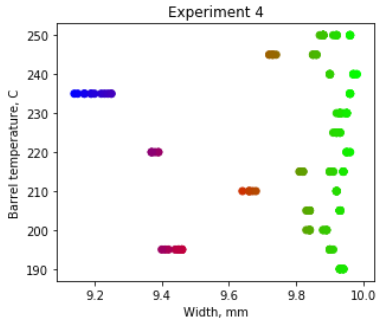
Figure C.1.5. Relationship between the specimen's width and screw speed DOE parameter



(a) Experiment 1



(b) Experiment 3



(c) Experiment 4

Figure C.1.6. Relationship between the specimen's width and barrel temperature DOE parameter

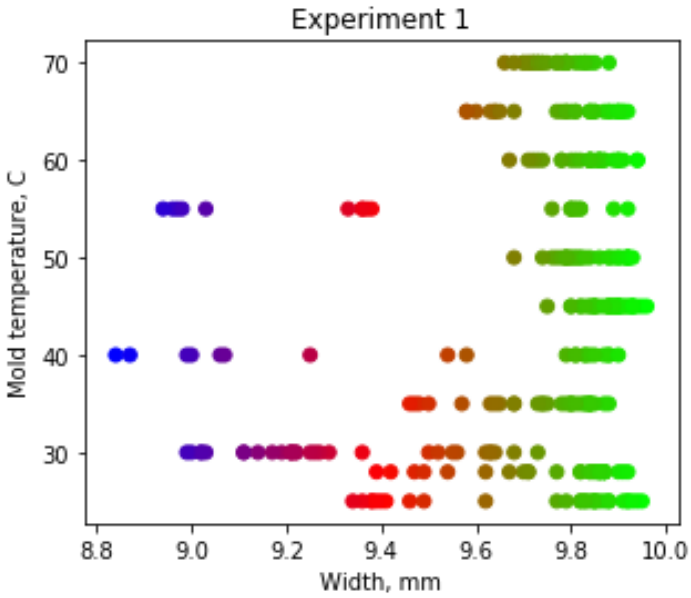
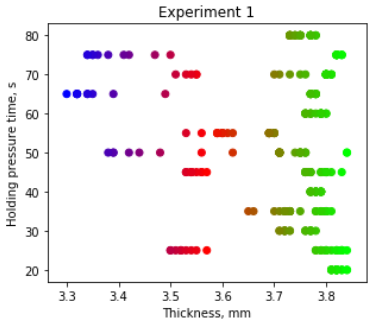
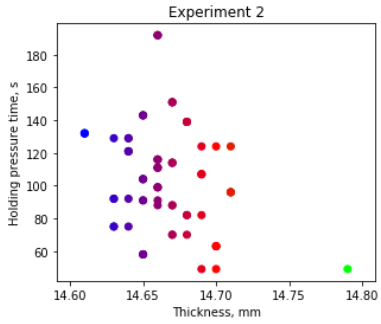


Figure C.1.7. Relationship between the specimen's width and mold temperature DOE parameter

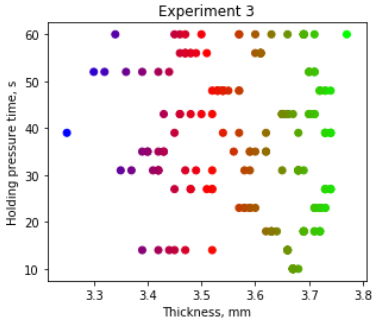
C.2 Thickness figures



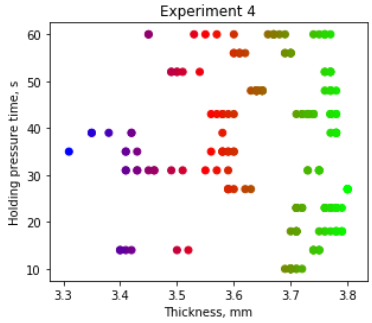
(a) Experiment 1



(b) Experiment 2

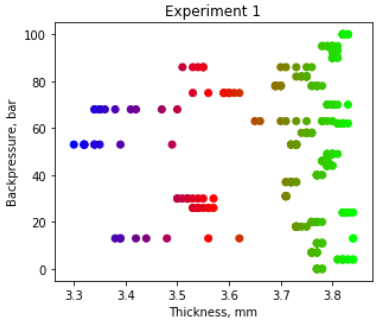


(c) Experiment 3

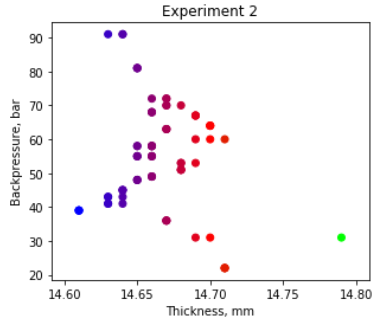


(d) Experiment 4

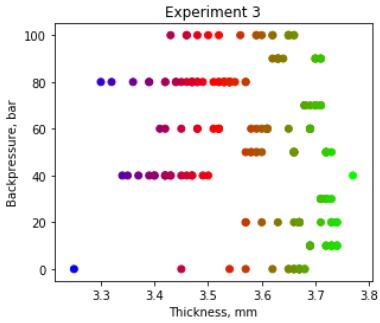
Figure C.2.1. Relationship between the specimen's thickness and holding pressure time DOE parameter



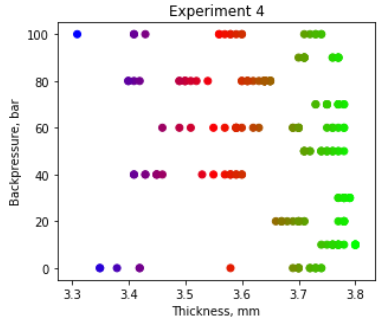
(a) Experiment 1



(b) Experiment 2

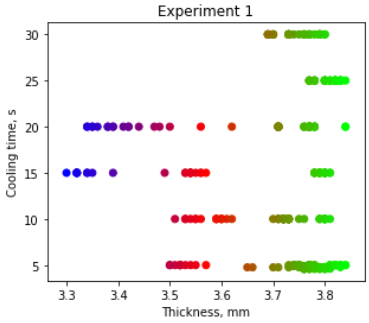


(c) Experiment 3

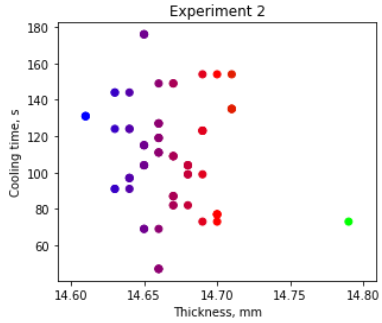


(d) Experiment 4

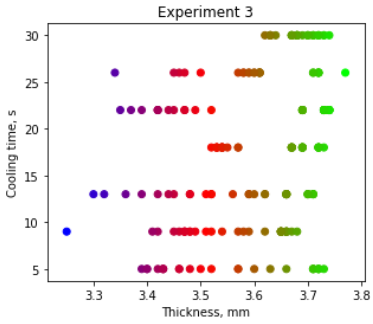
Figure C.2.2. Relationship between the specimen's thickness and backpressure DOE parameter



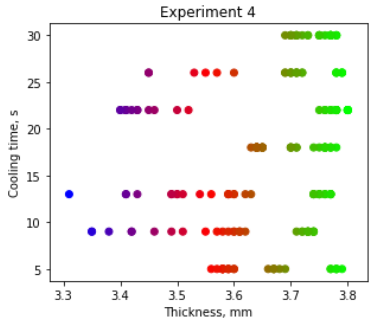
(a) Experiment 1



(b) Experiment 2

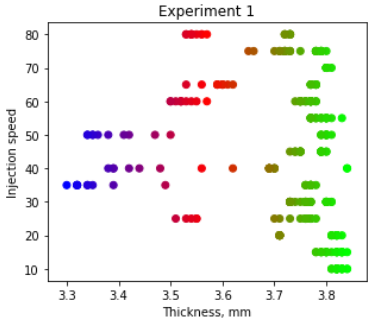


(c) Experiment 3

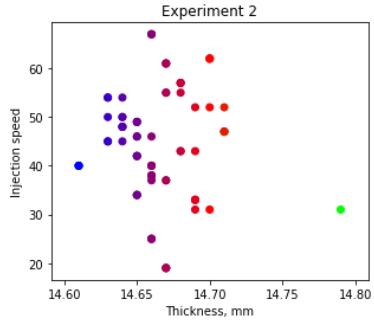


(d) Experiment 4

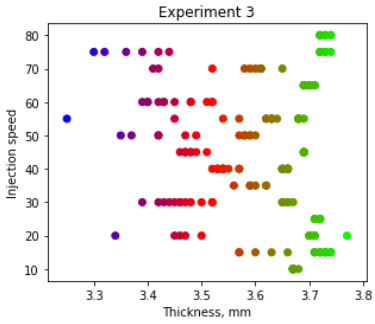
Figure C.2.3. Relationship between the specimen's thickness and cooling time DOE parameter



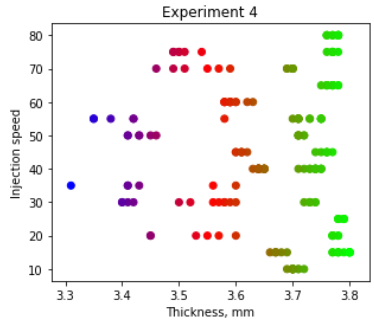
(a) Experiment 1



(b) Experiment 2

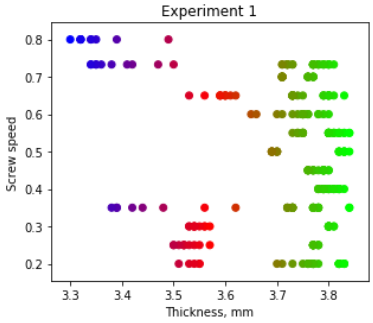


(c) Experiment 3

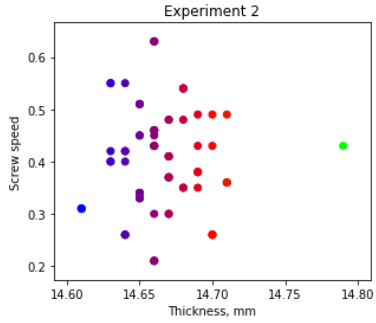


(d) Experiment 4

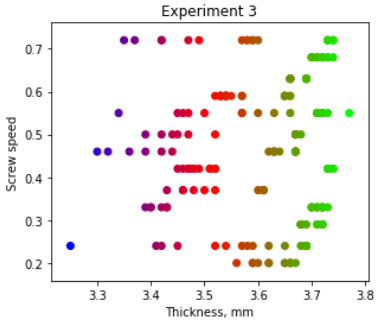
Figure C.2.4. Relationship between the specimen's thickness and injection speed DOE parameter



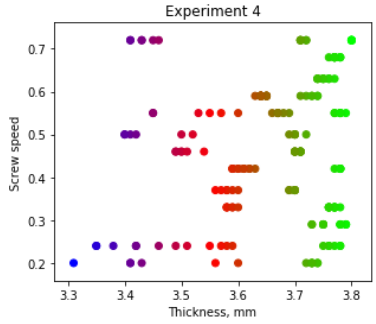
(a) Experiment 1



(b) Experiment 2

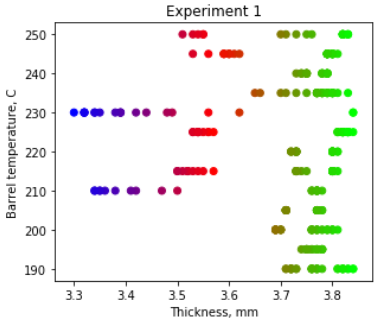


(c) Experiment 3

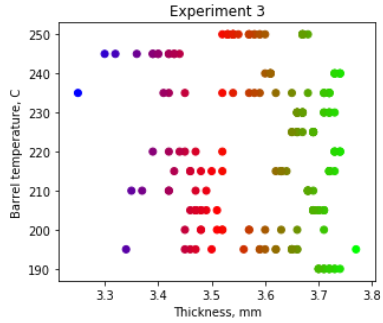


(d) Experiment 4

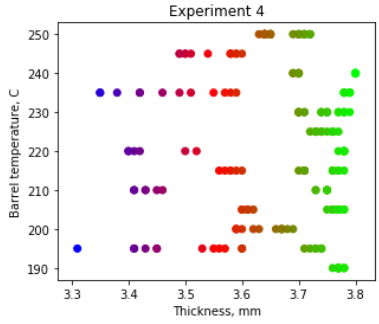
Figure C.2.5. Relationship between the specimen's thickness and screw speed DOE parameter



(a) Experiment 1



(b) Experiment 3



(c) Experiment 4

Figure C.2.6. Relationship between the specimen's thickness and barrel temperature DOE parameter

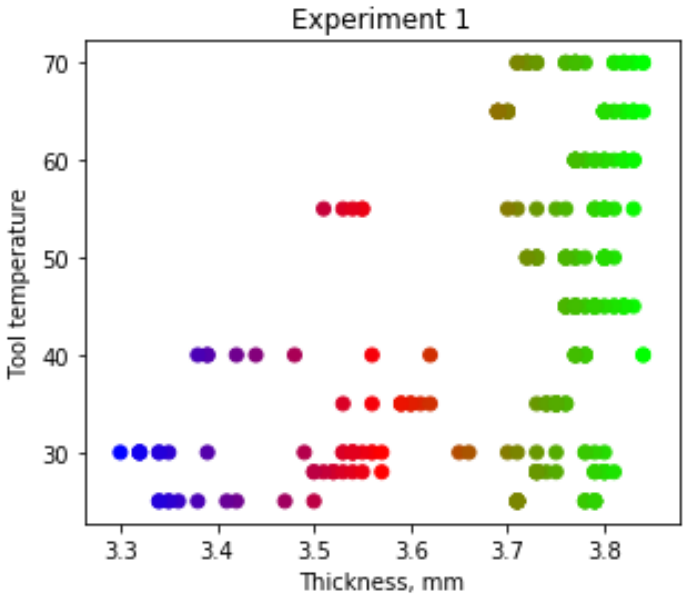
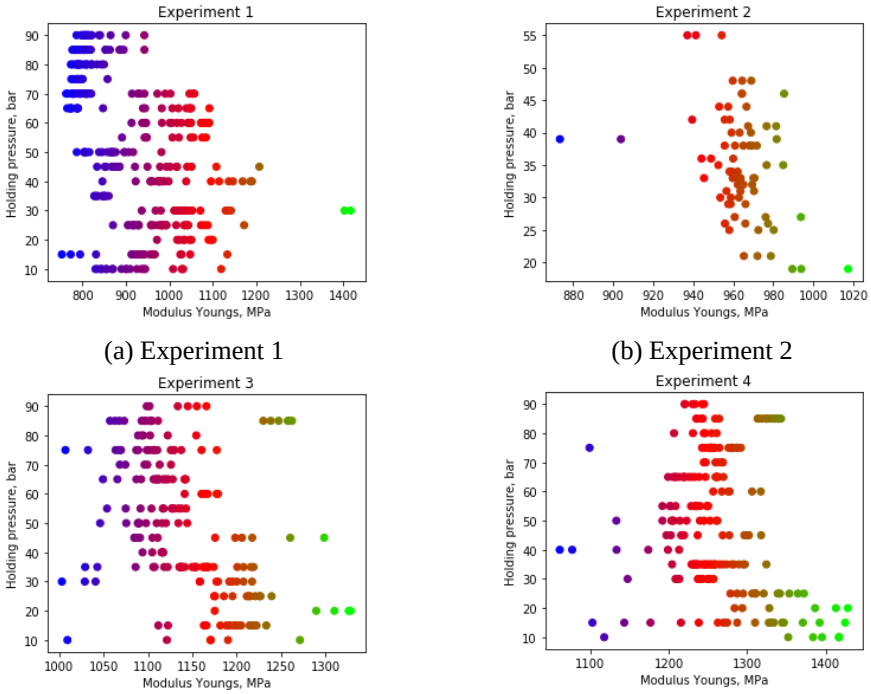


Figure C.2.7. Relationship between the specimen's thickness and mold temperature DOE parameter

C.3 Tensile modulus (Young's modulus) figures



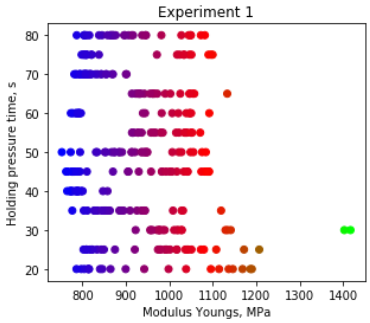
(a) Experiment 1

(b) Experiment 2

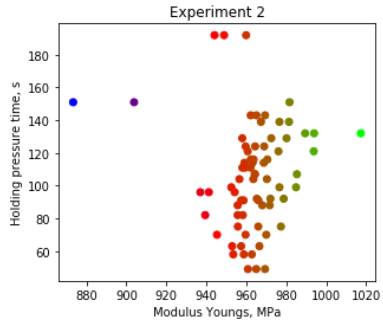
(c) Experiment 3

(d) Experiment 4

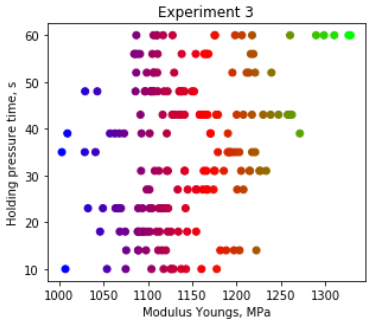
Figure C.3.1. Relationship between tensile modulus (Young's modulus) and holding pressure DOE parameter



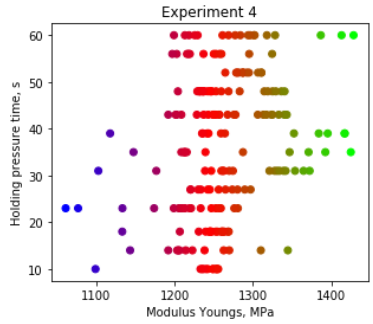
(a) Experiment 1



(b) Experiment 2

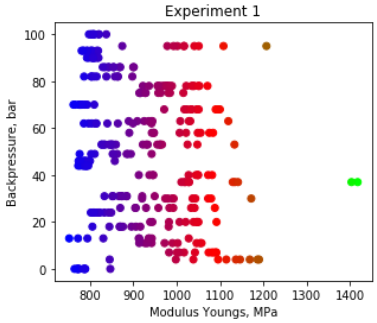


(c) Experiment 3

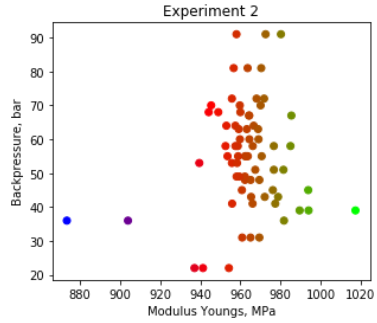


(d) Experiment 4

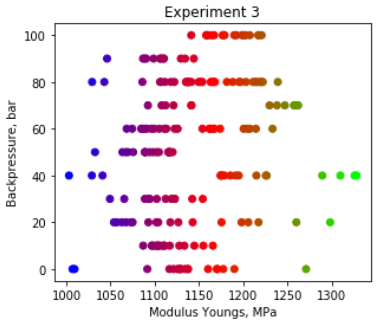
Figure C.3.2. Relationship between tensile modulus (Young's modulus) and holding pressure time DOE parameter



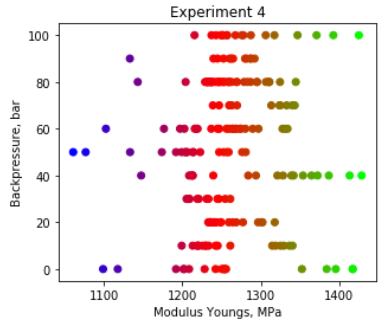
(a) Experiment 1



(b) Experiment 2

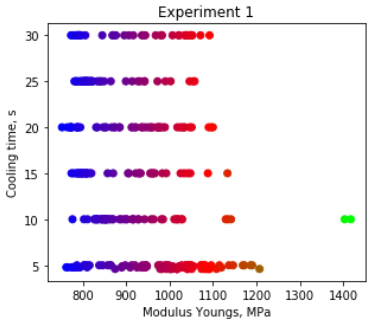


(c) Experiment 3

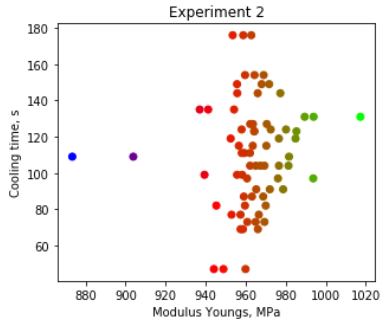


(d) Experiment 4

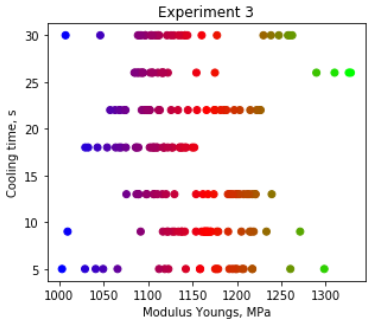
Figure C.3.3. Relationship between tensile modulus (Young's modulus) and backpressure DOE parameter



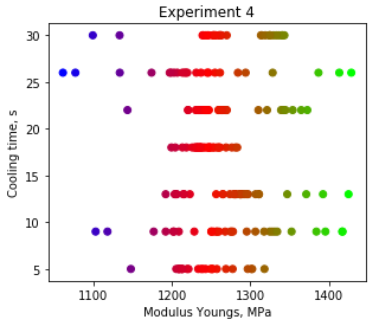
(a) Experiment 1



(b) Experiment 2

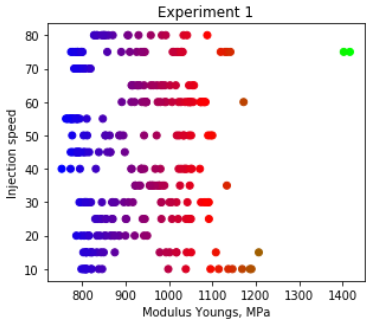


(c) Experiment 3

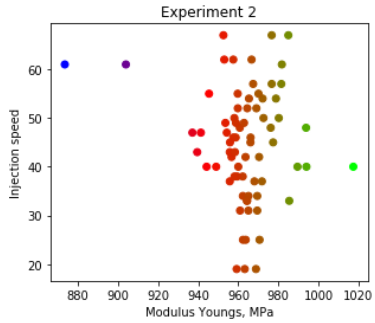


(d) Experiment 4

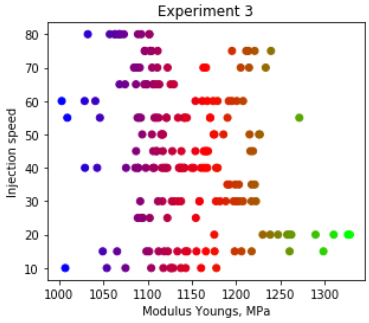
Figure C.3.4. Relationship between tensile modulus (Young's modulus) and cooling time DOE parameter



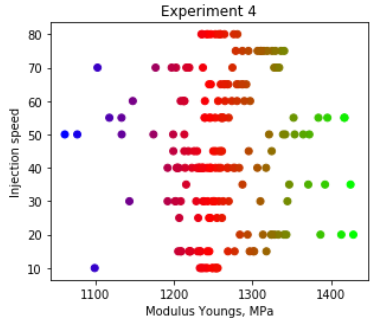
(a) Experiment 1



(b) Experiment 2

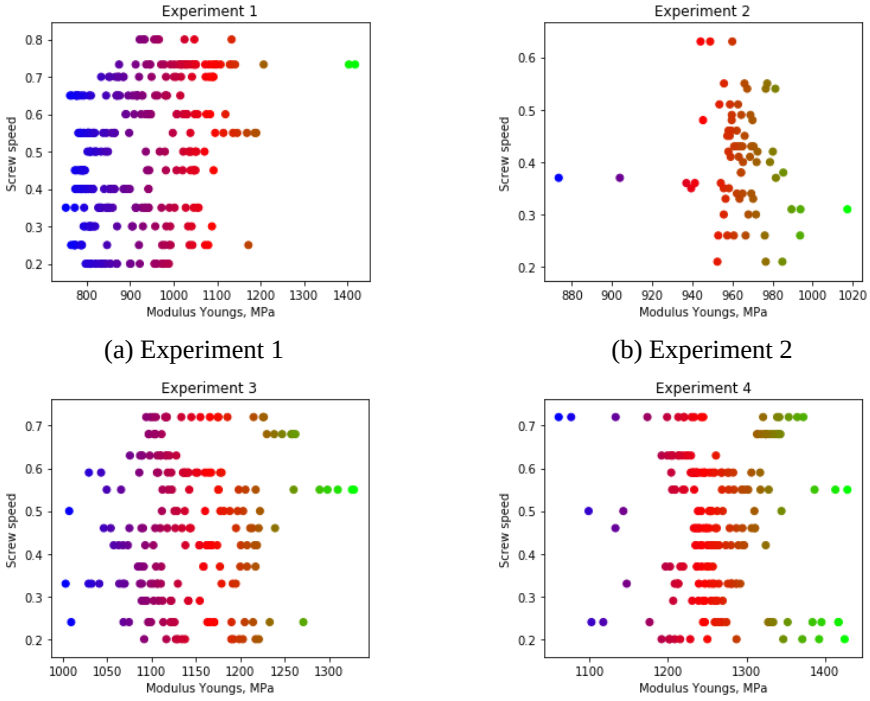


(c) Experiment 3



(d) Experiment 4

Figure C.3.5. Relationship between tensile modulus (Young's modulus) and injection speed DOE parameter



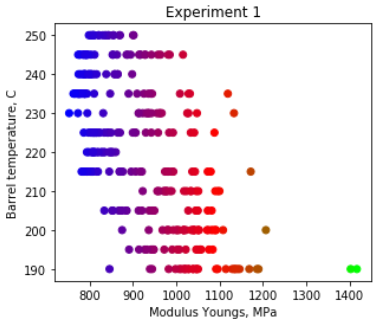
(a) Experiment 1

(b) Experiment 2

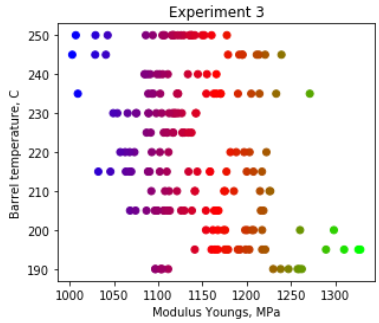
(c) Experiment 3

(d) Experiment 4

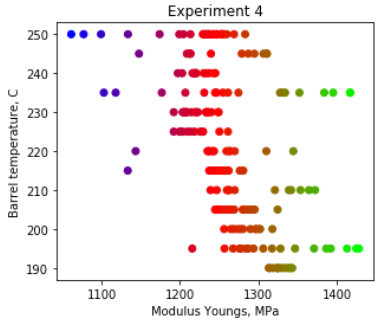
Figure C.3.6. Relationship between tensile modulus (Young's modulus) and screw speed DOE parameter



(a) Experiment 1



(b) Experiment 3



(c) Experiment 4

Figure C.3.7. Relationship between tensile modulus (Young's modulus) and barrel temperature DOE parameter

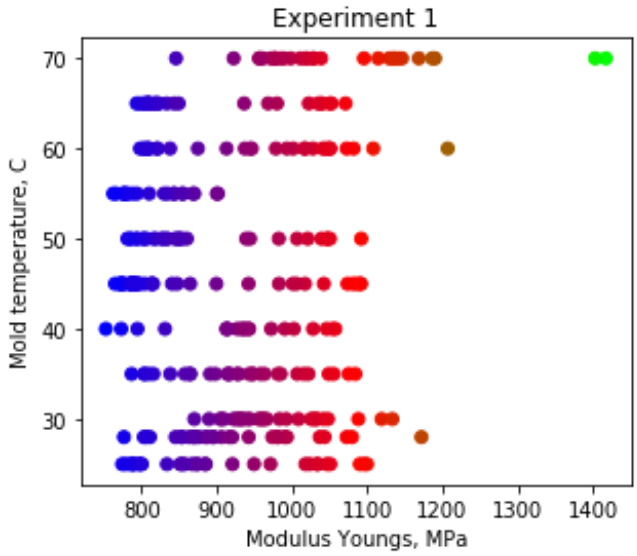
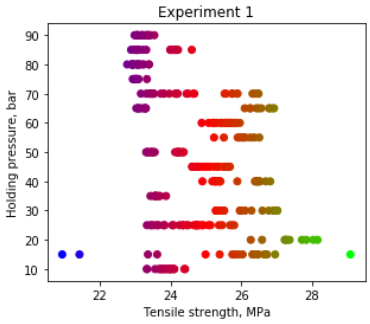
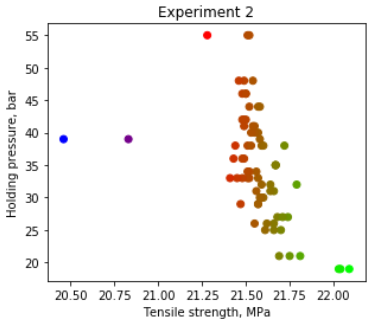


Figure C.3.8. Relationship between tensile modulus (Young's modulus) and mold temperature DOE parameter

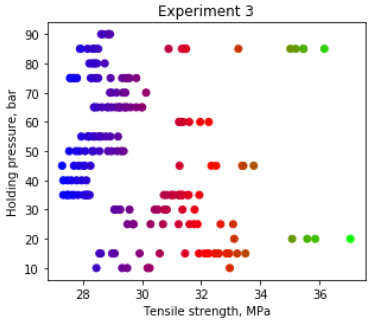
C.4 Tensile strength figures



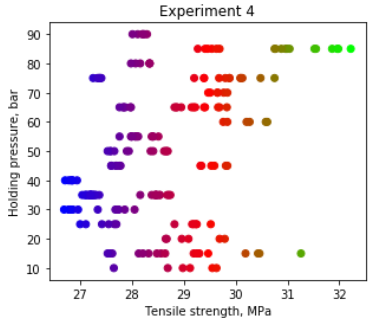
(a) Experiment 1



(b) Experiment 2

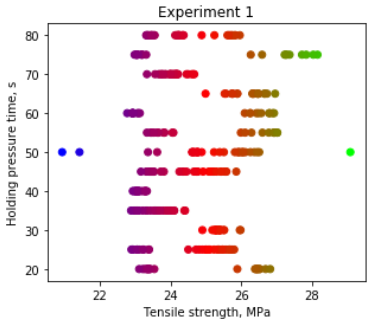


(c) Experiment 3

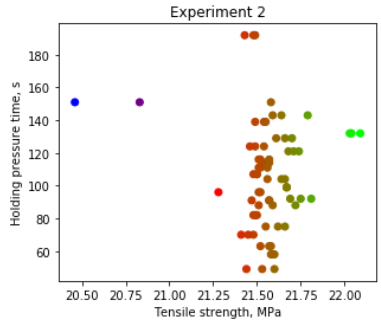


(d) Experiment 4

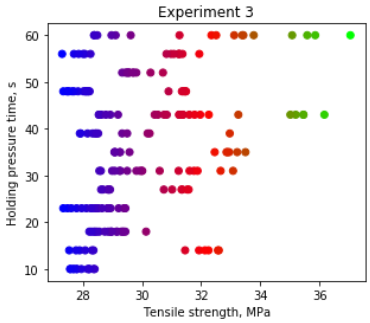
Figure C.4.1. Relationship between tensile strength and holding pressure DOE parameter



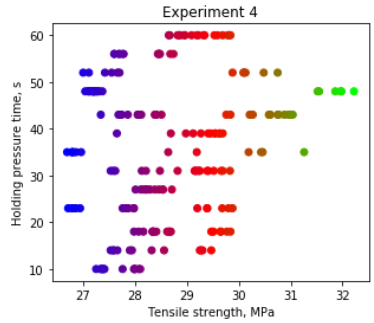
(a) Experiment 1



(b) Experiment 2

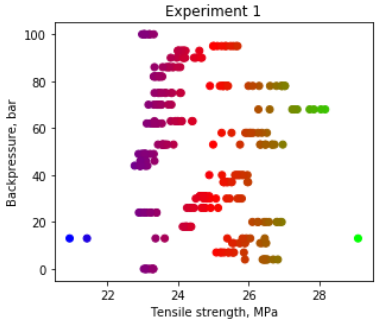


(c) Experiment 3

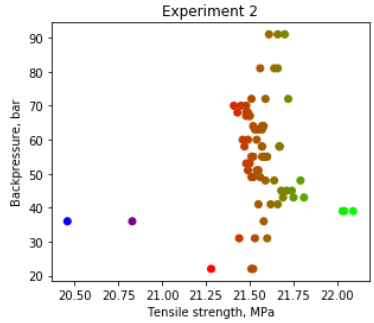


(d) Experiment 4

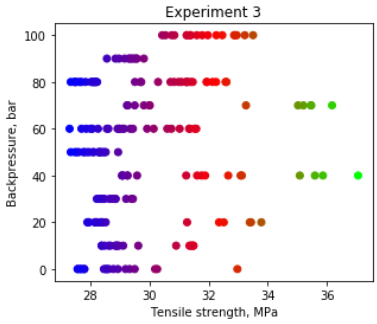
Figure C.4.2. Relationship between tensile strength and holding pressure time DOE parameter



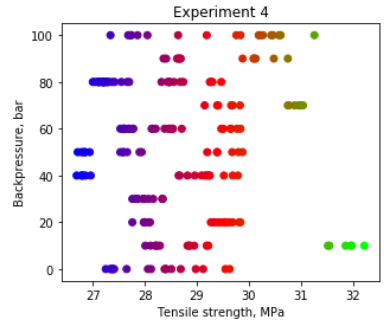
(a) Experiment 1



(b) Experiment 2

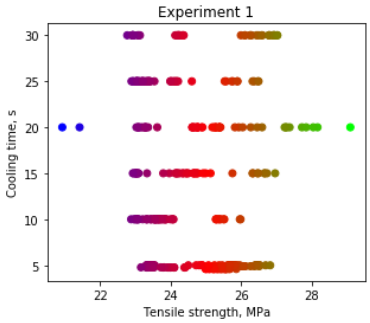


(c) Experiment 3

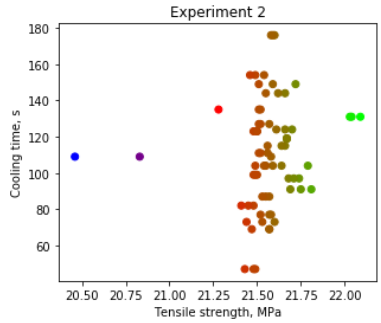


(d) Experiment 4

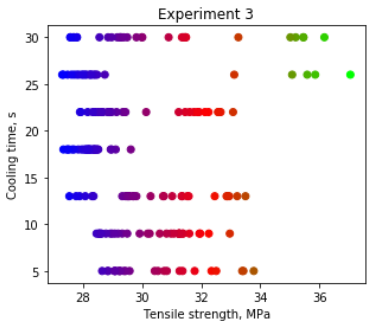
Figure C.4.3. Relationship between tensile strength and backpressure DOE parameter



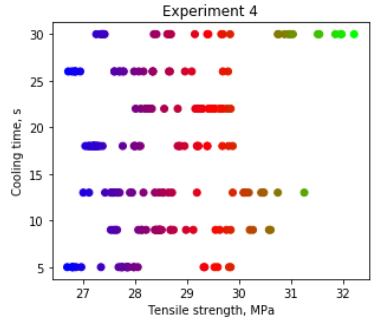
(a) Experiment 1



(b) Experiment 2

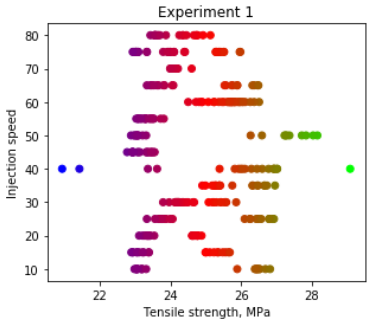


(c) Experiment 3

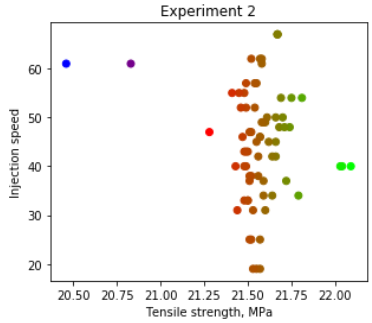


(d) Experiment 4

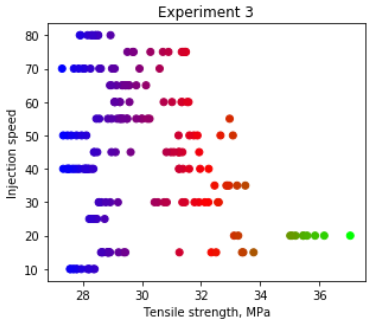
Figure C.4.4. Relationship between tensile strength and cooling time DOE parameter



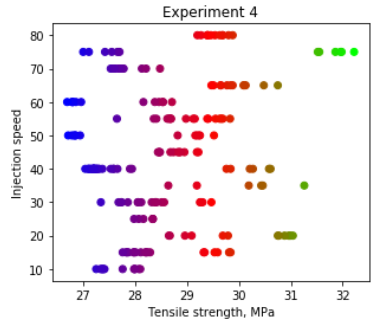
(a) Experiment 1



(b) Experiment 2

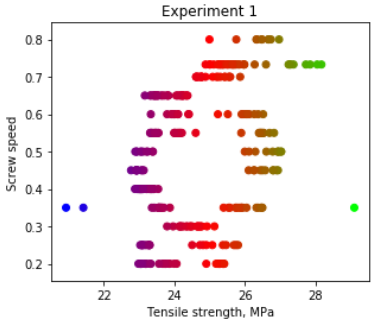


(c) Experiment 3

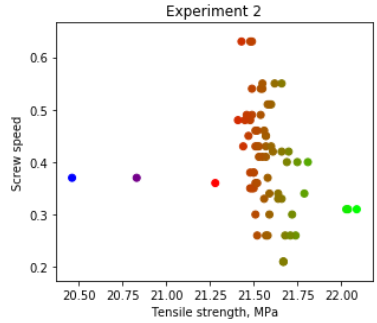


(d) Experiment 4

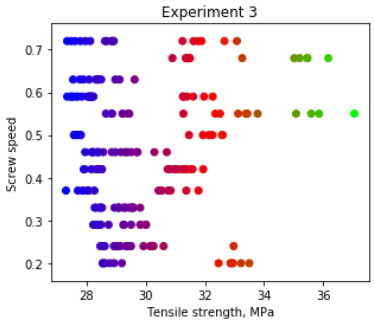
Figure C.4.5. Relationship between tensile strength and injection speed DOE parameter



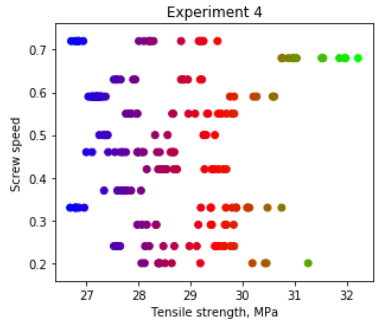
(a) Experiment 1



(b) Experiment 2

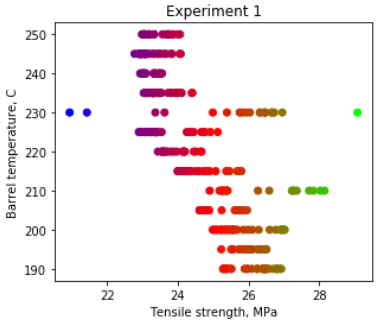


(c) Experiment 3

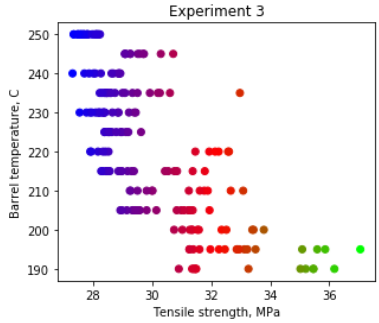


(d) Experiment 4

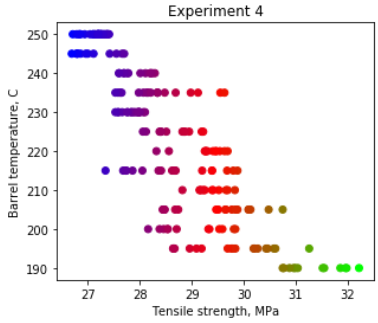
Figure C.4.6. Relationship between tensile strength and screw speed DOE parameter



(a) Experiment 1



(b) Experiment 3



(c) Experiment 4

Figure C.4.7. Relationship between tensile strength and barrel temperature DOE parameter

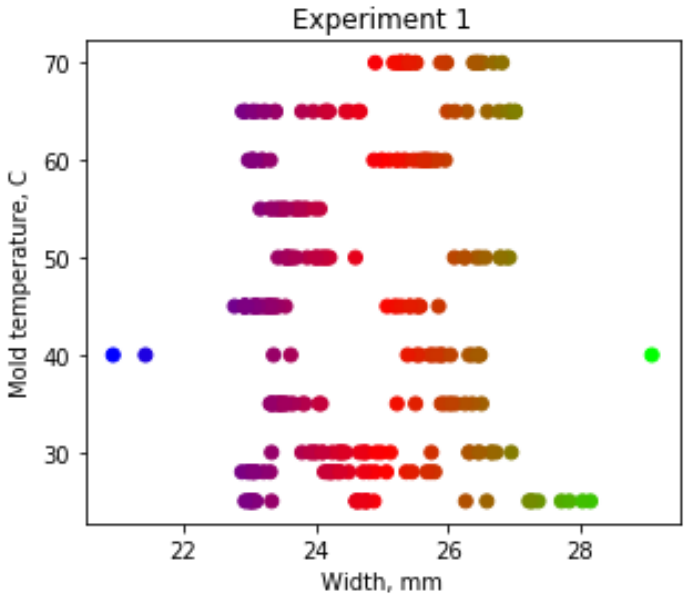
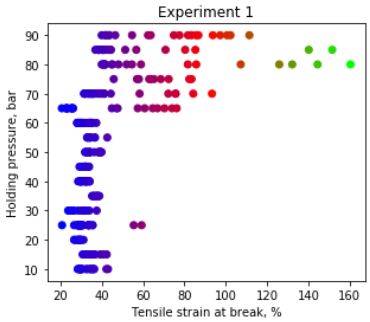
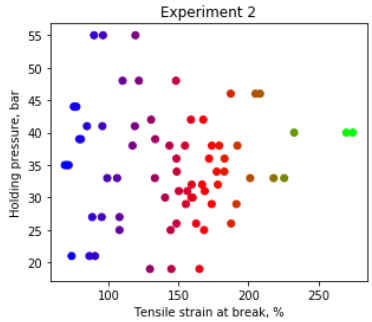


Figure C.4.8. Relationship between tensile strength and mold temperature DOE parameter

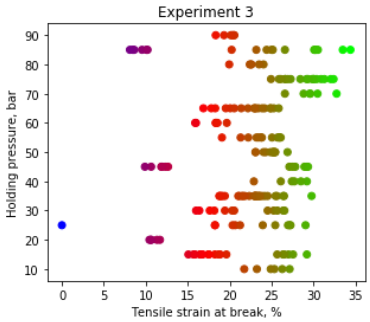
C.5 Tensile strain at break figures



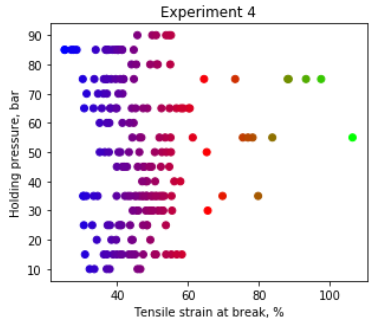
(a) Experiment 1



(b) Experiment 2

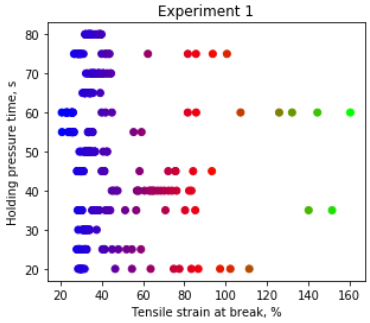


(c) Experiment 3

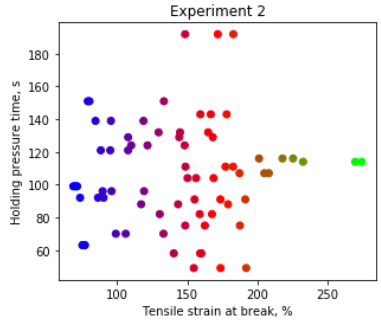


(d) Experiment 4

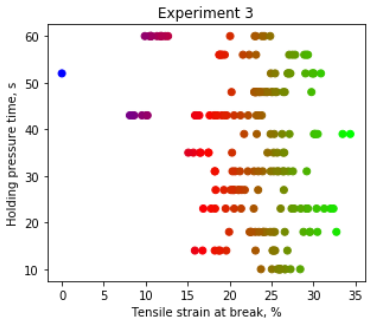
Figure C.5.1. Relationship between tensile strain at break and holding pressure DOE parameter



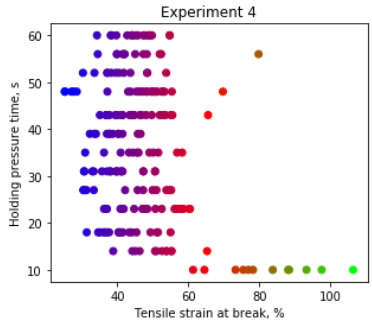
(a) Experiment 1



(b) Experiment 2

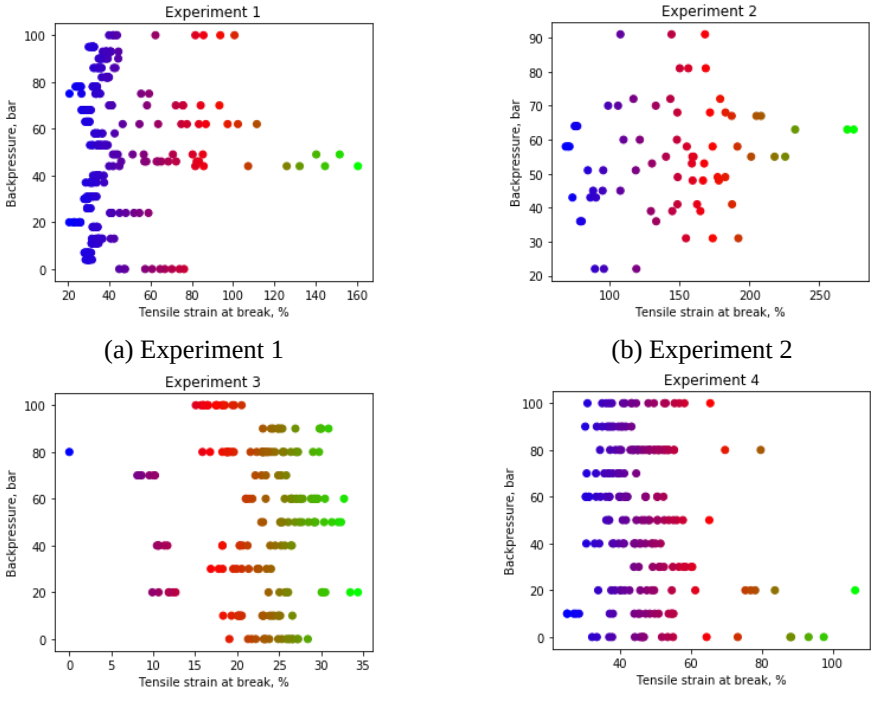


(c) Experiment 3



(d) Experiment 4

Figure C.5.2. Relationship between tensile strain at break and holding pressure time DOE parameter



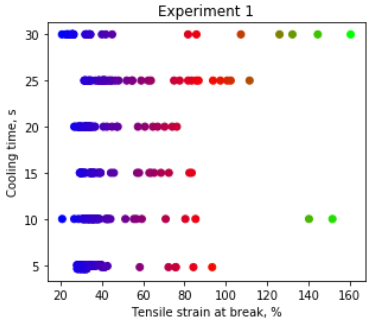
(a) Experiment 1

(b) Experiment 2

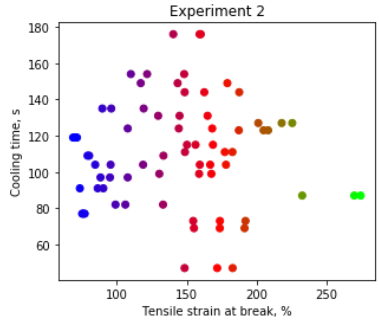
(c) Experiment 3

(d) Experiment 4

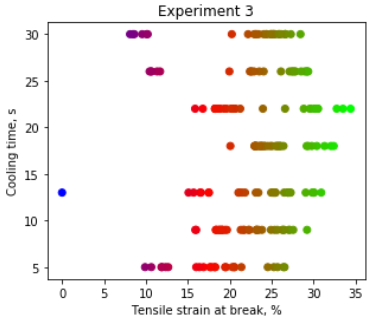
Figure C.5.3. Relationship between tensile strain at break and backpressure DOE parameter



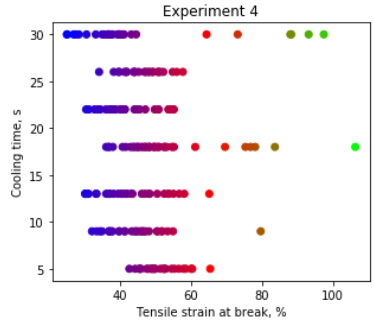
(a) Experiment 1



(b) Experiment 2

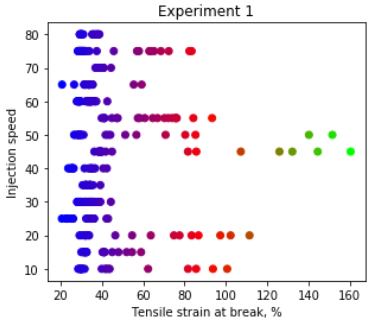


(c) Experiment 3

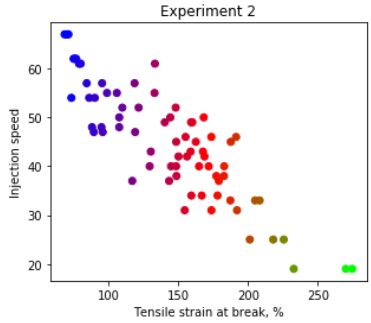


(d) Experiment 4

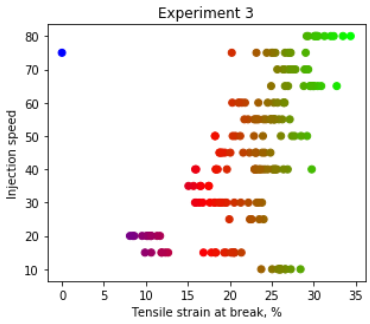
Figure C.5.4. Relationship between tensile strain at break and cooling time DOE parameter



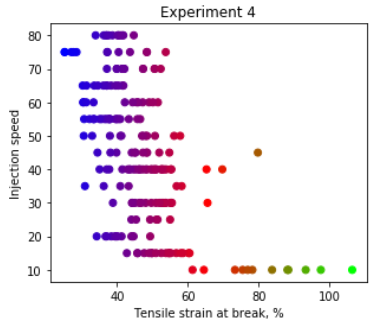
(a) Experiment 1



(b) Experiment 2

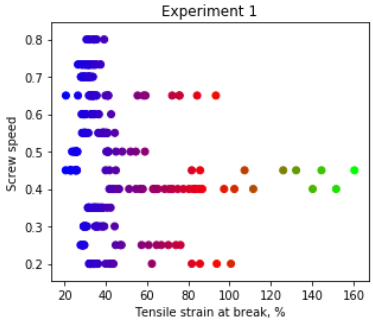


(c) Experiment 3

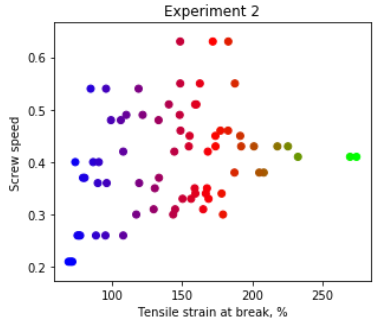


(d) Experiment 4

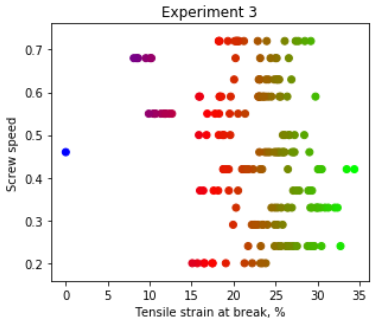
Figure C.5.5. Relationship between tensile strain at break and injection speed DOE parameter



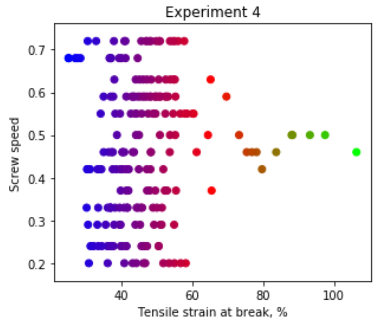
(a) Experiment 1



(b) Experiment 2

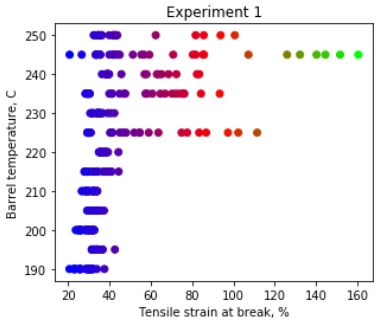


(c) Experiment 3

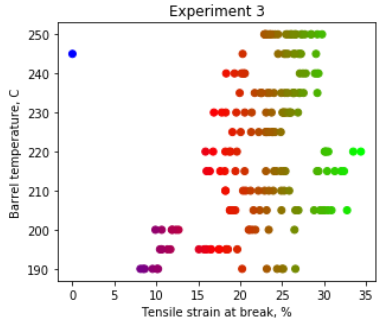


(d) Experiment 4

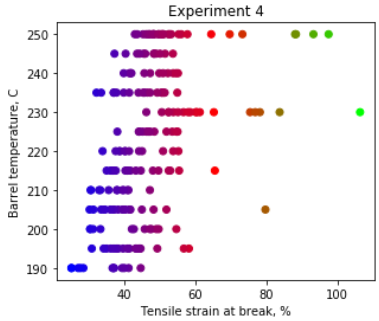
Figure C.5.6. Relationship between tensile strain at break and screw speed DOE parameter



(a) Experiment 1



(b) Experiment 3



(c) Experiment 4

Figure C.5.7. Relationship between tensile strain at break and barrel temperature DOE parameter

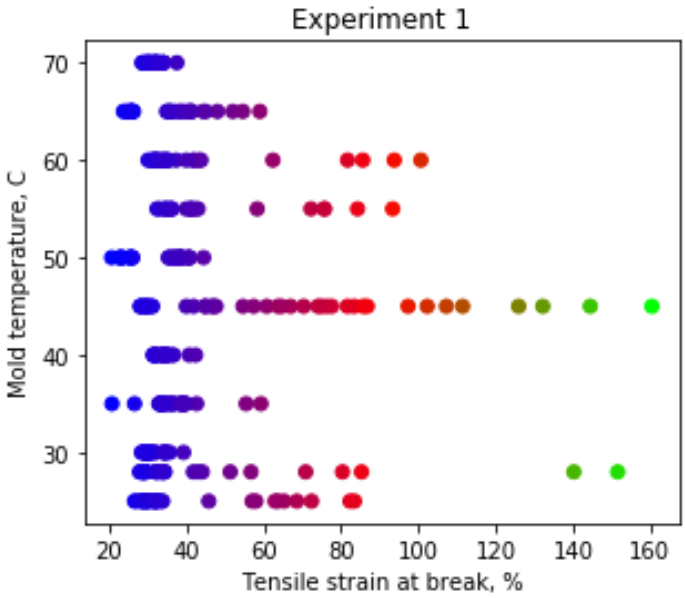


Figure C.5.8. Relationship between tensile strain at break and mold temperature DOE parameter

Appendix D

Feature selection scores for the best performing feature selection methods

Table D.1. RReliefF feature selection method scores for experiments 1-4 parallel and sequential datasets, width target variable

Parameter name	Exp. 1, width, sequential dataset	Exp. 1, width1, parallel dataset	Exp. 1, width2, parallel dataset	Exp. 2, width	Exp. 3, width, sequential dataset	Exp. 3, width1, parallel dataset	Exp. 3, width2, parallel dataset	Exp. 4, width, sequential dataset	Exp. 4, width1, parallel dataset	Exp. 4, width2, parallel dataset
Cushion_smallest	0.38	0.48	0.54	1	0.57	0.45	0.45	0.47	0.44	0.44
Cushion_after_hold_pres	0.51	0.67	1	0.99	1	1	1	0.97	1	1
Shot_vol_act	x	x	x	0.99	1	1	1	0.97	1	1
Cushion_average	0.45	0.56	0.48	0.75	x	x	x	x	x	x
Holding_pressure	0.36	0.49	0.89	0.57	0.9	0.78	0.79	0.83	0.73	0.73
Hold_pres_peak	x	x	x	0.55	0.83	0.72	0.75	1	0.94	0.94
mold_pressure_average	x	x	x	0.55	x	x	x	x	x	x
Backpressure_peak	x			0.52	0.23	0.04	0.06	0.32	0.12	0.16
mold_temperature_average	x	x	x	0.48	x	x	x	x	x	x
Backpressure	0.47	0.56	0.25	0.43	0.36	0.08	0.1	0.43	0.14	0.15
Injection_work	0.01	0.06	0.06	0.34	0.28	0.09	0.09	0.2	0.13	0.13
Plast_volume	x	x	x	0.31	0	0	0	0	0.06	0.06
Cushion_end_hold_pres	x	x	x	0.3	0.7	0.7	0.7	0.66	0.62	0.62
Cooling_time_last	0.65	0.72	0.26	0.25	0.75	0.34	0.31	0.31	0.02	0.02
Cooling_time	0.57	0.6	0.21	0.25	0.79	0.39	0.36	0.33	0.05	0.06
Pre_pres_peak	x			0.21	0.28	0.13	0.12	0.12	0	0
Spec_pres_switchov	0.54	0.58	0.35	0.2	0.28	0.13	0.12	0.12	0	0
Switchov_vol	0	0	0	0.15	0	0	0	0.04	0	0
Holding_pres_time	0.55	0.73	0.43	0.14	0.42	0.07	0.05	0.16	0.04	0.03

Parameter name	Exp. 1, width, sequential dataset	Exp. 1, width1, parallel dataset	Exp. 1, width2, parallel dataset	Exp. 2, width	Exp. 3, width, sequential dataset	Exp. 3, width1, parallel dataset	Exp. 3, width2, parallel dataset	Exp. 4, width, sequential dataset	Exp. 4, width1, parallel dataset	Exp. 4, width2, parallel dataset
Plast_time	0.36	0.5	0.23	0.14	0.33	0.24	0.26	0.56	0.44	0.44
Speed_max	0.24	0.33	0.08	0.12	0.27	0.13	0.14	0.36	0.22	0.23
Screw_speed	0.2	0.27	0.1	0.12	0.26	0.16	0.15	0.33	0.24	0.24
Torque_mean_cycle	x	x	x	0.11	0.11	0	0	0.29	0.27	0.27
Screw_pos_last	x	x	x	0.11	0	0	0	0	0.13	0.13
Torque_peak	x	x	x	0.1	0.12	0	0	0.26	0.23	0.23
Nozzle_temp_z5_average	x	x	x	0.1	x	x	x	x	x	x
Mold_protection_time	x	x	x	0.1	x	x	x	x	x	x
Last_cycle_time	0	0.1	0	0.09	0	0	0	0	0	0
Flow_no_plast	0.83	1	0.35	0.09	0.07	0	0	0.11	0.2	0.21
Injection_rate	x	x	x	0.09	0.2	0.09	0.09	0.26	0.22	0.23
Screw_speed_max	1	0.85	0.17	0.08	x	x	x	x	x	x
Nozzle_temp_z3_average	x	x	x	0.08	x	x	x	x	x	x
Nozzle_tempr_z2_average	0.48	0.65	0.53	0.07	x	x	x	x	x	x
Injection_Speed	0.5	0.42	0.43	0.07	0.42	0.36	0.32	0.33	0.11	0.13
Injection_speed_peak	x	x	x	0.07	x	x	x	x	x	x
Nozzle_temp_z6_average	x	x	x	0.04	x	x	x	x	x	x
Open_stroke_peak	x	x	x	0.01	x	x	x	x	x	x
Cooling_cycletime	x	x	x	0	0.25	0.03	0.01	0.11	0	0
Cycletime_hold_pressure	x	x	x	0	0.18	0	0	0	0	0
Closing_force	0	0	0	0	x	x	x	x	x	x
Clamp_force_switchov	0	0	0	0	x	x	x	x	x	x
Nozzle_temp_z4_average	x	x	x	0	x	x	x	x	x	x
Clamping_force_peak	x	x	x	0	x	x	x	x	x	x

Parameter name	Exp. 1, width, sequential dataset	Exp. 1, width1, parallel dataset	Exp. 1, width2, parallel dataset	Exp. 2, width	Exp. 3, width, sequential dataset	Exp. 3, width1, parallel dataset	Exp. 3, width2, parallel dataset	Exp. 4, width, sequential dataset	Exp. 4, width1, parallel dataset	Exp. 4, width2, parallel dataset
Injection_time	0.11	0.15	0	0	0.22	0.05	0.04	0.1	0	0
Heating_cyl1_z1_set	0.48	0.65	0.53	x	0.71	0.46	0.44	0.46	0.22	0.23
Tool_Temperature	0.32	0.35	0.29	x	x	x	x	x	x	x
Ejector_pos_last	0.16	0.15	0.1	x	0.12	0.02	0.02	0.23	0.08	0.07
Cycletime_open	x	x	x	x	0.45	0.21	0.19	0.26	0.1	0.1
Ejector_cycletime	x	x	x	x	0.22	0.04	0.05	0.19	0.12	0.12
Cycletime_close	x	x	x	x	0	0	0	0	0	0
Closinig_time	x	x	x	x	0	0	0	0	0.03	0.01
Cycletime_removal	x	x	x	x	0	0	0	0	0	0
Cycletime_until_demoulding_end	x	x	x	x	0	0	0	0	0	0
Clamp_force_buildup_time	x	x	x	x	0	0	0	0.15	0.1	0.1
Cycletime_nozzle	x	x	x	x	0	0	0	0	0.13	0.13

Table D.2. Spearman feature selection method scores for experiments 1-4 parallel and sequential datasets, thickness target variable

Parameter name	Exp. 1, thickn., seq. dataset	Exp. 1, thickn.1, paral. dataset	Exp. 1, thickn.2, paral. dataset	Exp. 2, thickn.	Exp. 3, thickn., seq. dataset	Exp. 3, thickn.1, paral. dataset	Exp. 3, thickn.2, paral. dataset	Exp. 4, thickn., seq. dataset	Exp. 4, thickn.1, paral. dataset	Exp. 4, thickn.2, paral. dataset
mold_pressure_average	x	x	x	1	x	x	x	x	x	x
Hold_pres_peak	x	x	x	0.93	0.95	0.95	0.95	0.95	0.95	0.95
Holding_pressure	1	1	1	0.97	1	1	1	1	1	1
Cushion_end_hold_pres	x	x	x	0.55	0.72	0.7	0.75	0.78	0.78	0.78
Cushion_smallest	0.84	0.81	0.87	0.81	0.94	0.91	0.96	0.94	0.94	0.94
Cushion_average	0.79	0.77	0.82	0.71	x	x	x	x	x	x
Shot_vol_act	x	x	x	0.81	0.95	0.94	0.97	0.96	0.96	0.96
Cushion_after_hold_pres	0.91	0.91	0.91	0.81	0.95	0.94	0.97	0.96	0.96	0.96
Holding_pres_time	0.28	0.28	0.29	0.2	0.2	0.2	0.22	0.17	0.16	0.17
Cycletime_hold_pressure	x	x	x	0.2	0	0	0.05	0.02	0.02	0.02
Injection_work	0.38	0.43	0.34	0.47	0.36	0.37	0.38	0.33	0.31	0.35
mold_temperature_average	x	x	x	0.29	x	x	x	x	x	x
Cooling_time	0.1	0.15	0.05	0.13	0.36	0.4	0.36	0.33	0.34	0.32
Last_cycle_time	0.11	0.1	0.13	0.07	0	0.01	0.04	0.07	0.06	0.06
Cooling_cycletime	x	x	x	0.08	0.26	0.29	0.27	0.24	0.23	0.24
Nozzle_temp_z4_average	x	x	x	0.13	x	x	x	x	x	x
Flow_no_plast	0.13	0.11	0.15	0.03	0.14	0.16	0.15	0.24	0.24	0.24
Screw_speed_max	0.18	0.18	0.19	0.03	x	x	x	x	x	x
Cooling_time_last	0.06	0.11	0.03	0.13	0.31	0.35	0.31	0.23	0.23	0.22
Clamp_force_switchov	0.04	0.05	0.04	0.01	x	x	x	x	x	x
Nozzle_temp_z5_average	x	x	x	0.11	x	x	x	x	x	x
Closing_force	0.05	0.06	0.03	0.06	x	x	x	x	x	x
Clamping_force_peak	x	x	x	0.01	x	x	x	x	x	x

Parameter name	Exp. 1, thickn., seq. dataset	Exp. 1, thickn.1, paral. dataset	Exp. 1, thickn.2, paral. dataset	Exp. 2, thickn.	Exp. 3, thickn., seq. dataset	Exp. 3, thickn.1, paral. dataset	Exp. 3, thickn.2, paral. dataset	Exp. 4, thickn., seq. dataset	Exp. 4, thickn.1, paral. dataset	Exp. 4, thickn.2, paral. dataset
Backpressure	0.1	0.03	0.16	0.03	0.42	0.45	0.41	0.33	0.36	0.3
Backpressure_peak	x	x	x	0.07	0.45	0.48	0.44	0.27	0.3	0.25
Injection_time	0.31	0.31	0.32	0.02	0.08	0.14	0.05	0.07	0.09	0.05
Mold_protection_time	x	x	x	0.11	x	x	x	x	x	x
Torque_peak	x	x	x	0.01	0.11	0.17	0.1	0.21	0.2	0.21
Plast_time	0.2	0.18	0.21	0.01	0.02	0.08	0.01	0.21	0.21	0.21
Open_stroke_peak	x	x	x	0.04	x	x	x	x	x	x
Torque_mean_cycle	x	x	x	0.01	0.06	0.11	0.05	0.22	0.21	0.22
Plast_volume	x	x	x	0.09	0.23	0.26	0.23	0.07	0.08	0.06
Injection_speed_peak	x	x	x	0.01	x	x	x	x	x	x
Injection_Speed	0.43	0.46	0.4	0.01	0.04	0.11	0.02	0.08	0.09	0.06
Screw_pos_last	x	x	x	0.07	0	0.01	0.03	0.01	0.03	0.01
Injection_rate	x	x	x	0.01	0.06	0.12	0.05	0.24	0.24	0.24
Nozzle_tempr_z2_average	0.06	0	0.13	0.1	x	x	x	x	x	x
Nozzle_temp_z6_average	x	x	x	0.02	x	x	x	x	x	x
Nozzle_temp_z3_average	x	x	x	0.13	x	x	x	x	x	x
Switchov_vol	0	0.01	0	0	0.08	0.15	0.05	0.12	0.12	0.1
Screw_speed	0.22	0.21	0.24	0.02	0.04	0.09	0.04	0.21	0.21	0.21
Pre_pres_peak	x	x	x	0.08	0.01	0.03	0.04	0	0.02	0.02
Spec_pres_switchov	0.45	0.46	0.44	0.08	0.01	0.03	0.04	0	0.02	0.02
Speed_max	0.16	0.15	0.18	0.01	0.04	0.08	0.05	0.2	0.2	0.2
Ejector_pos_last	0.12	0.11	0.13	x	0.01	0.06	0	0.08	0.1	0.06
Heating_cyl1_z1_set	0.08	0.02	0.14	x	0.15	0.21	0.14	0.09	0.08	0.09
Tool_Temperature	0.56	0.56	0.56	x	x	x	x	x	x	x

Parameter name	Exp. 1, thickn., seq. dataset	Exp. 1, thickn.1, paral. dataset	Exp. 1, thickn.2, paral. dataset	Exp. 2, thickn.	Exp. 3, thickn., seq. dataset	Exp. 3, thickn.1, paral. dataset	Exp. 3, thickn.2, paral. dataset	Exp. 4, thickn., seq. dataset	Exp. 4, thickn.1, paral. dataset	Exp. 4, thickn.2, paral. dataset
Cycletime_nozzle	x	x	x	x	0.02	0.06	0.04	0.14	0.15	0.14
Cycletime_until_demoulding_end	x	x	x	x	0	0.01	0.03	0	0	0
Cycletime_removal	x	x	x	x	0.05	0.08	0.07	0.01	0.01	0.01
Cycletime_close	x	x	x	x	0.14	0.16	0.16	0.12	0.11	0.12
Clamp_force_buildup_time	x	x	x	x	0.13	0.15	0.14	0.07	0.07	0.07
Closinig_time	x	x	x	x	0.2	0.22	0.22	0.11	0.11	0.11
Cycletime_open	x	x	x	x	0.13	0.18	0.11	0.04	0.03	0.04
Ejector_cycletime	x	x	x	x	0.08	0.12	0.08	0.15	0.15	0.14

Table D.3. RReliefF feature selection method scores for experiments 1-4 parallel and sequential datasets, Young's modulus target variable

Parameter name	Exp. 1, Young's mod., seq. dataset	Exp. 1, Young's mod.1, paral. dataset	Exp. 1, Young's mod.2, paral. dataset	Exp. 2, Young's mod.	Exp. 3, Young's mod., seq. dataset	Exp. 3, Young's mod.1, paral. dataset	Exp. 3, Young's mod.2, paral. dataset	Exp. 4, Young's mod., seq. dataset	Exp. 4, Young's mod.1, paral. dataset	Exp. 4, Young's mod.2, paral. dataset
Plast_volume	x	x	x	1.00	0.00	0	0	0.20	0	0.14
Hold_pres_peak	x	x	x	0.75	0.10	0	0.36	0.83	0.67	0.74
Holding_pressure	0.25	1.00	0.79	0.66	0.16	0.06	0.17	0.79	0.57	0.63
Backpressure	0.56	0.61	0.69	0.59	0.35	0.00	0.69	0.78	0.50	0.75
mold_pressure_average	x	x	x	0.55	x	x	x	x	x	x
Injection_speed_peak	x	x	x	0.54	x	x	x	x	x	x
Injection_Speed	0.35	0.49	0.06	0.54	0.62	0.77	0.24	0.10	0.16	0.00
Backpressure_peak	x	x	x	0.53	0.03	0	0.24	0.59	0.32	0.44
Torque_peak	x	x	x	0.41	0.03	0	0.04	0.36	0.31	0.29
Plast_time	0.02	0.00	0.00	0.37	0.10	0	0.00	0.52	0.42	0.43
Torque_mean_cycle	x	x	x	0.34	0.00	0	0.00	0.42	0.31	0.36
Holding_pres_time	0.49	0.73	0.55	0.33	0.00	0	0.21	0.37	0.27	0.18
Screw_pos_last	x	x	x	0.27	0.00	0	0.00	0.31	0.30	0.39
Flow_no_plast	0.13	0.14	0.27	0.23	0.41	0	0.79	0.44	0.15	0.53
Shot_vol_act	x	x	x	0.18	0.32	0.46	0.52	0.98	1.00	0.98
Cushion_after_hold_pres	0.19	0.99	0.63	0.18	0.32	0.46	0.52	0.98	1.00	0.98
Cushion_smallest	0.24	0.80	0.48	0.17	0.00	0.00	0.13	0.41	0.19	0.33
Screw_speed	0.26	0.32	0.29	0.16	0.06	0.09	0.00	0.28	0.22	0.15
Speed_max	0.33	0.28	0.35	0.15	0.31	0.74	0.00	0.29	0.11	0.16
Cushion_average	0.23	0.69	0.40	0.14	x	x	x	x	x	x
Injection_rate	x	x	x	0.11	0.10	0.01	0.03	0.32	0.20	0.29
Screw_speed_max	0.20	0.34	0.30	0.09	x	x	x	x	x	x

Parameter name	Exp. 1, Young's mod., seq. dataset	Exp. 1, Young's mod.1, paral. dataset	Exp. 1, Young's mod.2, paral. dataset	Exp. 2, Young's mod.	Exp. 3, Young's mod., seq. dataset	Exp. 3, Young's mod.1, paral. dataset	Exp. 3, Young's mod.2, paral. dataset	Exp. 4, Young's mod., seq. dataset	Exp. 4, Young's mod.1, paral. dataset	Exp. 4, Young's mod.2, paral. dataset
Cooling_cycletime	x	x	x	0.05	0.48	0.00	0.69	0.84	0.40	1.00
Injection_work	0.29	0.58	0.54	0.04	0.07	0.00	0.00	0.00	0.00	0.00
Nozzle_temp_z6_average	x	x	x	0.00	x	x	x	x	x	x
Closing_force	1	0.48	0.55	0.00	x	x	x	x	x	x
Injection_time	0.00	0.00	0.00	0.00	0.31	0.08	0.63	0.00	0.00	0.00
Cycletime_hold_pressure	x	x	x	0.00	0.08	0	0.16	0.00	0.00	0.10
Cushion_end_hold_pres	x	x	x	0.00	1	1.00	0.50	1.00	0.76	0.81
Last_cycle_time	0.35	0.13	0.10	0.00	0.05	0.00	0.37	0.08	0.04	0.09
Nozzle_temp_z5_average	x	x	x	0.00	x	x	x	x	x	x
Clamp_force_switchov	0.94	0.18	0.52	0.00	x	x	x	x	x	x
Mold_protection_time	x	x	x	0.00	x	x	x	x	x	x
Open_stroke_peak	x	x	x	0.00	x	x	x	x	x	x
Cooling_time	0.30	0.26	0.10	0.00	0.51	0.15	1.00	0.28	0.01	0.04
Cooling_time_last	0.21	0.11	0.12	0.00	0.51	0.13	0.98	0.55	0.33	0.40
Nozzle_temp_z4_average	x	x	x	0.00	x	x	x	x	x	x
Clamping_force_peak	x	x	x	0.00	x	x	x	x	x	x
mold_temperature_average	x	x	x	0.00	x	x	x	x	x	x
Nozzle_temp_z3_average	x	x	x	0.00	x	x	x	x	x	x
Switchov_vol	0.36	0.20	0.00	0.00	0.00	0.00	0.00	0.29	0.17	0.40
Nozzle_tempr_z2_average	0.23	0.65	0.47	0.00	x	x	x	x	x	x
Pre_pres_peak	x	x	x	0.00	0.26	0.26	0.10	0.05	0.12	0.00
Spec_pres_switchov	0.17	0.19	0.09	0.00	0.26	0.26	0.10	0.05	0.12	0.00
Tool_Temperature	0.37	0.72	1.00	x	x	x	x	x	x	x
Heating_cyl1_z1_set	0.23	0.66	0.47	x	0.00	0.12	0.00	0.84	0.78	0.89
Ejector_pos_last	0.04	0.07	0.00	x	0.14	0.07	0.00	0.36	0.29	0.16

Parameter name	Exp. 1, Young's mod., seq. dataset	Exp. 1, Young's mod.1, paral. dataset	Exp. 1, Young's mod.2, paral. dataset	Exp. 2, Young's mod.	Exp. 3, Young's mod., seq. dataset	Exp. 3, Young's mod.1, paral. dataset	Exp. 3, Young's mod.2, paral. dataset	Exp. 4, Young's mod., seq. dataset	Exp. 4, Young's mod.1, paral. dataset	Exp. 4, Young's mod.2, paral. dataset
Clamp_force_buildup_time	x	x	x	x	0.56	0.08	0.33	0.55	0.41	0.75
Cycletime_nozzle	x	x	x	x	0.09	0.00	0.27	0.47	0.28	0.60
Cycletime_open	x	x	x	x	0.00	0.00	0.31	0.67	0.64	0.76
Cycletime_close	x	x	x	x	0.00	0.00	0.45	0.00	0.08	0.00
Ejector_cycletime	x	x	x	x	0.00	0.00	0.12	0.52	0.55	0.62
Cycletime_removal	x	x	x	x	0.00	0.00	0.00	0.00	0.12	0.00
Cycletime_until_demoulding_end	x	x	x	x	0.00	0.00	0.00	0.00	0.12	0.00
Closinig_time	x	x	x	x	0.00	0.00	0.29	0.00	0.15	0.11

Table D.4. RReliefF feature selection method scores for experiments 1-4 parallel and sequential datasets, tensile strength target variable

Parameter name	Exp. 1, tens. strength, seq. dataset	Exp. 1, tens. strength1, paral. dataset	Exp. 1, tens. strength2, paral. dataset	Exp. 2, tens. strength	Exp. 3, tens. strength, seq. dataset	Exp. 3, tens. strength1, paral. dataset	Exp. 3, tens. strength2, paral. dataset	Exp. 4, tens. strength, seq. dataset	Exp. 4, tens. strength1, paral. dataset	Exp. 4, tens. strength2, paral. dataset
mold_pressure_average	x	x	x	1.00	x	x	x	x	x	x
mold_temperature_average	x	x	x	0.81	x	x	x	x	x	x
Clamp_force_switchov	0.07	0.03	0.00	0.61	x	x	x	x	x	x
Open_stroke_peak	x	x	x	0.60	x	x	x	x	x	x
Clamping_force_peak	x	x	x	0.58	x	x	x	x	x	x
Cushion_end_hold_pres	x	x	x	0.56	0.49	0.44	0.52	0.43	0.15	0.24
Closing_force	0.05	0.06	0.01	0.56	x	x	x	x	x	x
Cushion_smallest	0.50	0.42	0.49	0.50	0.44	0.39	0.45	0.41	0.37	0.38
Cushion_after_hold_pres	0.36	0.48	0.54	0.49	0.61	0.55	0.65	0.67	0.43	0.48
Shot_vol_act	x	x	x	0.49	0.61	0.55	0.65	0.67	0.43	0.48
Holding_pressure	0.34	0.43	0.48	0.43	0.38	0.27	0.34	0.64	0.52	0.57
Holding_pres_time	0.50	0.46	0.40	0.39	0.48	0.38	0.30	0.75	0.62	0.65
Nozzle_temp_z3_average	x	x	x	0.38	x	x	x	x	x	x
Hold_pres_peak	x	x	x	0.38	0.31	0.20	0.26	0.64	0.41	0.48
Backpressure	1.00	0.56	0.59	0.36	0.39	0.13	0.17	0.69	0.49	0.49
Mold_protection_time	x	x	x	0.31	x	x	x	x	x	x
Last_cycle_time	0.13	0.15	0.09	0.26	0.06	0.04	0.00	0.08	0.01	0.03
Nozzle_temp_z5_average	x	x	x	0.25	x	x	x	x	x	x
Cooling_time	0.45	0.26	0.36	0.23	0.27	0.21	0.05	0.38	0.12	0.08
Cooling_time_last	0.35	0.14	0.26	0.23	0.27	0.21	0.06	0.43	0.19	0.14
Plast_time	0.10	0.07	0.10	0.20	0.15	0.05	0.01	0.42	0.09	0.11
Backpressure_peak	x	x	x	0.20	0.31	0.18	0.16	0.56	0.44	0.42
Cushion_average	0.55	0.42	0.50	0.19	x	x	x	x	x	x

Parameter name	Exp. 1, tens. strength, seq. dataset	Exp. 1, tens. strength1, paral. dataset	Exp. 1, tens. strength2, paral. dataset	Exp. 2, tens. strength	Exp. 3, tens. strength, seq. dataset	Exp. 3, tens. strength1, paral. dataset	Exp. 3, tens. strength2, paral. dataset	Exp. 4, tens. strength, seq. dataset	Exp. 4, tens. strength1, paral. dataset	Exp. 4, tens. strength2, paral. dataset
Torque_peak	x	x	x	0.16	0.13	0.05	0.00	0.34	0.17	0.13
Torque_mean_cycle	x	x	x	0.15	0.06	0.00	0.00	0.36	0.18	0.14
Injection_rate	x	x	x	0.14	0.13	0.03	0.00	0.32	0.09	0.08
Flow_no_plast	0.55	0.33	0.31	0.14	0.04	0.05	0.00	0.30	0.21	0.17
Screw_speed	0.50	0.43	0.38	0.13	0.15	0.05	0.00	0.36	0.10	0.08
Plast_volume	x	x	x	0.13	0.00	0.00	0.00	0.15	0.21	0.19
Screw_speed_max	0.50	0.41	0.33	0.11	x	x	x	x	x	x
Speed_max	0.38	0.35	0.29	0.08	0.01	0.00	0.00	0.33	0.03	0.03
Screw_pos_last	x	x	x	0.01	0.00	0.00	0.00	0.05	0.02	0.00
Nozzle_tempr_z2_average	0.62	0.99	0.99	0.01	x	x	x	x	x	x
Injection_work	0.44	0.48	0.35	0.00	0.24	0.21	0.20	0.38	0.34	0.36
Switchov_vol	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.05	0.45	0.48
Cooling_cycletime	x	x	x	0.00	0.11	0.03	0.00	0.45	0.28	0.26
Cycletime_hold_pressure	x	x	x	0.00	0.15	0.00	0.00	0.18	0.02	0.00
Nozzle_temp_z6_average	x	x	x	0.00	x	x	x	x	x	x
Spec_pres_switchov	0.61	0.44	0.46	0.00	0.39	0.36	0.35	0.33	0.28	0.28
Pre_pres_peak	x	x	x	0.00	0.39	0.36	0.35	0.33	0.28	0.28
Injection_speed_peak	x	x	x	0.00	x	x	x	x	x	x
Injection_Speed	0.66	0.47	0.44	0.00	0.45	0.31	0.32	0.45	0.22	0.21
Nozzle_temp_z4_average	x	x	x	0.00	x	x	x	x	x	x
Injection_time	0.29	0.12	0.18	0.00	0.20	0.08	0.07	0.17	0.12	0.11
Heating_cyl1_z1_set	0.63	1.00	1.00	x	1.00	1.00	1.00	1.00	1.00	1.00
Tool_Temperature	0.57	0.44	0.36	x	x	x	x	x	x	x
Ejector_pos_last	0.09	0.05	0.01	x	0.05	0.04	0.01	0.23	0.15	0.12
Cycletime_open	x	x	x	x	0.80	0.86	0.83	0.72	0.79	0.74

Parameter name	Exp. 1, tens. strength, seq. dataset	Exp. 1, tens. strength1, paral. dataset	Exp. 1, tens. strength2, paral. dataset	Exp. 2, tens. strength	Exp. 3, tens. strength, seq. dataset	Exp. 3, tens. strength1, paral. dataset	Exp. 3, tens. strength2, paral. dataset	Exp. 4, tens. strength, seq. dataset	Exp. 4, tens. strength1, paral. dataset	Exp. 4, tens. strength2, paral. dataset
Ejector_cycletime	x	x	x	x	0.47	0.35	0.28	0.52	0.54	0.51
Clamp_force_buildup_time	x	x	x	x	0.32	0.20	0.18	0.20	0.32	0.34
Closinig_time	x	x	x	x	0.26	0.18	0.13	0.17	0.11	0.29
Cycletime_close	x	x	x	x	0.17	0.14	0.10	0.09	0.07	0.15
Cycletime_until_demoulding_end	x	x	x	x	0.02	0.04	0.00	0.02	0.04	0.07
Cycletime_removal	x	x	x	x	0.02	0.02	0.00	0.00	0.00	0.00
Cycletime_nozzle	x	x	x	x	0.00	0.01	0.00	0.15	0.04	0.03

Table D.5. RReliefF feature selection method scores for experiments 1-4 parallel and sequential datasets, tensile strain at break target variable

Parameter name	Exp. 1, tens. strain at break, seq. dataset	Exp. 1, tens. strain at break1, paral. dataset	Exp. 1, tens. strain at break2, paral. dataset	Exp. 2, tens. strain at break	Exp. 3, tens. strain at break, seq. dataset	Exp. 3, tens. strain at break1, paral. dataset	Exp. 3, tens. strain at break2, paral. dataset	Exp. 4, tens. strain at break, seq. dataset	Exp. 4, tens. strain at break1, paral. dataset	Exp. 4, tens. strain at break2. paral. dataset
Injection_Speed	0.60	0.42	0.64	1.00	1.00	0.99	0.92	0.56	0.30	0.45
Injection_speed_peak	x	x	x	0.99	x	x	x	x	x	
Backpressure_peak	x	x	x	0.81	0.72	0.88	0.64	0.08	0.03	0.00
Injection_time	0.65	0.94	0.71	0.69	0.76	0.88	0.79	0.60	0.33	0.48
Backpressure	0.50	0.31	0.53	0.67	0.77	0.95	0.63	0.17	0.08	0.00
Spec_pres_switchov	0.87	0.61	0.63	0.44	0.54	0.53	0.56	0.54	0.58	0.51
Pre_pres_peak	x	x	x	0.44	0.54	0.53	0.56	0.54	0.58	0.51
Nozzle_temp_z4_average	x	x	x	0.41	x	x	x	x	x	x
Injection_work	0.12	0.00	0.00	0.36	0.35	0.28	0.28	0.40	0.25	0.25
Plast_volume	x	x	x	0.36	0.00	0.00	0.00	0.00	0.00	0.10
Cooling_time	0.47	0.33	0.40	0.31	0.34	0.32	0.27	0.05	0.08	0.24
Cooling_time_last	0.47	0.27	0.32	0.30	0.35	0.34	0.28	0.11	0.27	0.16
Hold_pres_peak	x	x	x	0.27	0.24	0.23	0.15	0.27	0.13	0.00
Cushion_average	0.30	0.34	0.14	0.26	x	x	x	x	x	x
Shot_vol_act	x	x	x	0.23	0.41	0.43	0.40	0.36	0.16	0.00
Cushion_after_hold_pres	0.34	0.35	0.21	0.23	0.41	0.43	0.40	0.36	0.16	0.00
Screw_pos_last	x	x	x	0.23	0.00	0.00	0.00	0.00	0.06	0.00
Cushion_smallest	0.29	0.34	0.17	0.23	0.45	0.61	0.51	0.30	0.15	0.02
Open_stroke_peak	x	x	x	0.19	x	x	x	x	x	x
Switchov_vol	0.00	0.00	0.00	0.18	0.01	0.28	0.09	0.00	0.21	0.52
Holding_pressure	0.48	0.76	0.55	0.17	0.29	0.30	0.24	0.19	0.09	0.00
mold_pressure_average	x	x	x	0.15	x	x	x	x	x	x

Parameter name	Exp. 1, tens. strain at break, seq. dataset	Exp. 1, tens. strain at break1, paral. dataset	Exp. 1, tens. strain at break2, paral. dataset	Exp. 2, tens. strain at break	Exp. 3, tens. strain at break, seq. dataset	Exp. 3, tens. strain at break1, paral. dataset	Exp. 3, tens. strain at break2, paral. dataset	Exp. 4, tens. strain at break, seq. dataset	Exp. 4, tens. strain at break1, paral. dataset	Exp. 4, tens. strain at break2. paral. dataset
Cycletime_hold_pressure	x	x	x	0.15	0.22	0.00	0.08	0.78	0.20	1.00
Torque_mean_cycle	x	x	x	0.14	0.16	0.14	0.11	0.13	0.10	0.00
Torque_peak	x	x	x	0.14	0.22	0.20	0.18	0.14	0.10	0.03
Screw_speed	0.36	0.41	0.44	0.13	0.35	0.21	0.24	0.14	0.00	0.00
Nozzle_temp_z5_average	x	x	x	0.13	x	x	x	x	x	x
Plast_time	0.80	0.28	0.51	0.12	0.24	0.16	0.03	0.06	0.00	0.00
Speed_max	0.34	0.48	0.51	0.09	0.17	0.14	0.00	0.25	0.00	0.06
Cooling_cycletime	x	x	x	0.08	0.05	0.00	0.04	0.18	0.03	0.00
Injection_rate	x	x	x	0.07	0.33	0.24	0.16	0.12	0.00	0.00
mold_temperature_average	x	x	x	0.07	x	x	x	x	x	x
Nozzle_temp_z3_average	x	x	x	0.06	x	x	x	x	x	x
Nozzle_temp_z6_average	x	x	x	0.06	x	x	x	x	x	x
Screw_speed_max	0.27	0.49	0.52	0.04	x	x	x	x	x	x
Flow_no_plast	0.40	0.46	0.39	0.02	0.06	0.00	0.00	0.09	0.05	0.00
Clamping_force_peak	x	x	x	0.00	x	x	x	x	x	x
Last_cycle_time	0.26	0.12	0.16	0.00	0.07	0.05	0.00	0.05	0.03	0.00
Clamp_force_switchov	0.00	0.00	0.00	0.00	x	x	x	x	x	x
Holding_pres_time	0.63	0.77	0.69	0.00	0.63	0.46	0.45	1.00	0.53	0.80
Cushion_end_hold_pres	x	x	x	0.00	0.18	0.17	0.18	0.00	0.05	0.00
Nozzle_tempr_z2_average	0.25	0.52	0.36	0.00	x	x	x	x	x	x
Mold_protection_time	x	x	x	0.00	x	x	x	x	x	x
Closing_force	0.00	0.00	0.00	0.00	x	x	x	x	x	x
Tool_Temperature	1.00	1.00	1.00	x	x	x	x	x	x	x
Heating_cyl1_z1_set	0.25	0.52	0.36	x	0.92	1.00	1.00	0.76	0.99	0.46

Parameter name	Exp. 1, tens. strain at break, seq. dataset	Exp. 1, tens. strain at break1, paral. dataset	Exp. 1, tens. strain at break2, paral. dataset	Exp. 2, tens. strain at break	Exp. 3, tens. strain at break, seq. dataset	Exp. 3, tens. strain at break1, paral. dataset	Exp. 3, tens. strain at break2, paral. dataset	Exp. 4, tens. strain at break, seq. dataset	Exp. 4, tens. strain at break1, paral. dataset	Exp. 4, tens. strain at break2. paral. dataset
Ejector_pos_last	0.06	0.00	0.00	x	0.00	0.00	0.00	0.15	0.10	0.04
Cycletime_open	x	x	x	x	0.74	0.81	0.83	0.54	1.00	0.44
Clamp_force_buildup_time	x	x	x	x	0.34	0.12	0.02	0.03	0.17	0.35
Ejector_cycletime	x	x	x	x	0.30	0.11	0.22	0.34	0.39	0.34
Cycletime_nozzle	x	x	x	x	0.14	0.02	0.08	0.00	0.00	0.00
Closinig_time	x	x	x	x	0.04	0.00	0.05	0.00	0.28	0.02
Cycletime_close	x	x	x	x	0.04	0.00	0.00	0.04	0.27	0.17
Cycletime_until_demoulding_end	x	x	x	x	0.00	0.00	0.00	0.17	0.10	0.11
Cycletime_removal	x	x	x	x	0.00	0.00	0.00	0.04	0.00	0.32

ISBN 978-82-326-5810-7 (printed ver.)
ISBN 978-82-326-5906-7 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology