APLICACIÓN DE UNA METODOLOGÍA DE MACHINE LEARNING PARA OPTIMIZAR EL PROCESO DE FABRICACIÓN ADITIVA

Trabajo final
Ingeniería Mecánica

Autor:
Carola Zingarelli

16 de diciembre, 2022

APLICACIÓN DE UNA METODOLOGÍA DE MACHINE LEARNING PARA OPTIMIZAR EL PROCESO DE FABRICACIÓN ADITIVA

Trabajo final
Ingeniería Mecánica

Autor:
Carola Zingarelli

16 de diciembre, 2022

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Resumen

Uno de los principales desafíos de la fabricación aditiva es la importante deformación de la pieza producida, lo cual induce diferencias significativas entre las dimensiones de diseño y la forma final de la pieza. La simulación por elementos finitos del proceso ayuda a reducir considerablemente el número de pruebas y errores necesario para obtener una pieza con dimensiones dentro de los límites de tolerancia. Sin embargo, este tipo de simulaciones complejas requieren tiempos de simulación importantes. Cambios en los parámetros de proceso o del material, que afecten el resultado de la simulación, inducen un aumento del tiempo necesario durante la fase de diseño del proceso. El uso de inteligencia artificial en el área de la fabricación aditiva permite construir modelos que rápidamente vinculen los parámetros de entrada con la forma final de la pieza fabricada, acelerando la definición de la correcta configuración del proceso. El objetivo de este trabajo es definir una metodología de machine learning para predecir la deformacion, vinculando dos softwares de la compañía Hexagon: Digimat-AM y ODYSSEE CAE, a través de un programa en lenguaje Python. Se desarrolló un modelo capaz de predecir la forma final del componente producido y el tiempo de enfriamiento de la cámara de impresión, para diferentes tipos de procesos de fabricación aditiva. Para lograrlo, se entrenaron modelos de orden reducido en ODYSSEE CAE, utilizando una base de datos conformada por simulaciones realizadas en Digimat-AM. Los tipos de proceso analizados son el de sinterización selectiva por láser y modelado por deposición fundida. ODYSSEE CAE cuenta con múltiples algoritmos para estudiar el problema planteado, abarcando los métodos de interpolación directa, modelos de órden reducido y clustering. Ofrece, además, la posibilidad de entrenar múltiples algoritmos y comparar la calidad de las predicciones sobre un conjunto de data de validación. También se entrenan modelos usando redes neuronales y redes neuronales bayesianas, por ser el tipo de algoritmo más utilizado en la bibliografía consultada sobre la aplicación de inteligencia artificial en el campo de la fabricación aditiva. Los resultados muestran que la aplicación de esta metodología permite obtener predicciones de buena precisión, requiriéndose solamente 5 sets de data de entrenamiento para lograr optimizar la deformación de la pieza variando parámetros de proceso y posición. Por otro lado, la data de entrenamiento debe estar conformada por al menos 60 simulaciones cuando se introduce la variación de la orientación de la pieza, ya que la relación subyacente entre su rotación y forma final es más compleja. El enfriamiento de la pieza impresa también es predecible con esta metodología. Los algoritmos de redes neuronales no presentaron mejoras significativas en cuanto a la calidad de los resultados. En términos de tiempo computacional, el uso de modelos de machine learning es ventajoso por sobre las simulaciones por elementos finitos en casos que requieran pequeños conjuntos de entrenamiento y, por lo tanto, un tiempo reducido para la generación de la base de datos.

Abstract

One of the main challenges in additive manufacturing process is the warpage on the produced part, inducing significant differences between the as-designed and as-printed part shapes. Finite elements (FE) simulation greatly helps to reduce the number of trials and errors to get the printed part within acceptable tolerances. However, the simulation time of such advanced process simulation is important so that any change in process parameters, material parameters,

or any other inputs affecting the simulation results induce additional elapsed time along the process design phase. The use of artificial intelligence to build up surrogate models, quickly linking input parameters to predict the as-printed part shape, opens new opportunities to speed up the identification of the right process settings. The goal of this work is to define a machine learning methodology to predict warpage, linking two softwares from Hexagon company: Digimat-AM and ODYSSEE CAE, through python scripts. A model capable of predicting the final warpage of printed parts and cooldown time for different manufacturing processes, by training ROM models in ODYSSEE CAE with data provided by simulations performed in Digimat-AM has been developed. The results show that the application of this methodology allows to get good accuracy predictions, requiring only 5 training sets to perform warpage optimizations for varying process and position parameters. Around 60 simulations are needed when part orientation is varied as the underlying relationship is more complex. Cooldown behaviour of the printed part can also be predicted using this methodology. In terms of computational time, using ROM is advantageous for cases requiring small training sets and therefore reduced time for data generation.

1 INTRODUCTION

1.1 ABOUT THE COMPANY

e-Xstream Engineering, part of Hexagon's Manufacturing Intelligence division, is a material & engineering services company 100% focused on materials. It offers the industry an integrated solution portfolio to leverage the full potential of Integrated Computational Material Engineering (ICME). The solutions are built on an integrated stack of state-of-the-art software, hardware and engineering expertise to model materials, manufacturing process and final part performance.

The company's ICME Solution is centered around its leading nonlinear multi-scale modeling technology, Digimat and the leading data management system, MaterialCenter. In addition Hexagon is also developing software to leverage AI/ML methodologies when applied to CAE.

e-Xstream was founded in 2003 in Belgium by Dr. Roger Assaker and Prof. Issam Doghri from University of Louvain. Solvay group in Belgium and Goodyear in Luxembourg were the first to manifest their interest in e-Xstream product - Digimat, advanced material modelling software. The number of customers and turnover kept growing fast throughout the years and nowadays e-Xstream engineering has become the key player in the materials modelling segment with 9 out of 10 largest OEMs (original equipment manufacturer) and most material suppliers among its clients.

With software enhancements coming out on average every six months and continuous development, e-Xstream has become the leading specialist in composite materials and plastics for all types of industries.

With the acquisition of the company by MSC Software, the international developments were intensified. In three years span, the company tripled its size and increased its turnover by three and a half times, all this while preserving its historical structure. In 2017 MSC Software was acquired by Swedish group Hexagon, which brought additional value and allowed to connect the virtual and the physical worlds.

e-Xstream engineering is part of Hexagon, a leading global provider of information technologies that drive productivity and quality across geospatial and industrial enterprise applications.

1.2 Industrial Context

Additive Manufacturing (AM) enables the manufacturing of complex structural components and systems, allowing them at the same time to be lighter, stronger, easier to produce, and more cost-effective parts. Advancements in machinery, materials, and software made this technology more and more prevalent in the industry, making it possible to extend its principal use from creating rapid prototypes to functional end-use parts [6]. Multiple reports, such as the "3d Printing Trend Report 2022" from Hubs company [7], revealed that the additive manufacturing market grew by 19.8% in 2020, and is expected to continue to grow by, at least, 17% percent annually over the next three years. Multiple industries, including healthcare, automotive, aerospace and defense are expected to drive the additive manufacturing market growth [8], as it represents an efficient and cost effective solution for these industries, specially in today's context of innovation competition and industrialized rapid manufacturing taking over as a process to mass produce products [9].

On the other hand, extending these benefits to large-scale and load-bearing industrial applications requires a deeper understanding of the mechanical behaviour of these materials. Some important challenges in the AM field include the warpage and shrinkage that occur during the 3D printing of polymer parts and the inconsistency of the quality, leading to important differences between the as-designed and as-printed part shapes. One way to address this challenge is conducting high-fidelity simulations [10]. In this context, Hexagon offers **Digimat-AM**, a solution of Digimat Software that uses FEA to simulate and optimize the printing process and predict the as-printed part warpage, residual stresses, process-induced microstructure and its thermal response. It allows to greatly reduce the number of trials and errors necessary to keep the component within acceptable tolerances.

However, some shortcomings are also found in the use of FE methods for simulation of complex physical systems. The major disadvantage is that it requires solving a system of equations at each time step, being time-consuming. Any change in the process or material parameters, or any other input that affects the simulation results, is translated in the need of performing a new simulation, increasing the computation time intended for the design phase. Machine Learning (ML) technologies have proven to provide effective ways to process optimization, modelling and quality control of complex systems. The use of Artificial Intelligence (AI) in the AM field to build up surrogate models, quickly linking input parameters to predicted as-printed part shape, opens new opportunities to speed up the definition of the right process settings. The application of ML for AM has been reviewed in many studies [10, 11, 12]. Qi et al. [13] overviews the progress of applying the Neural Networks (NN) algorithm to several aspects of the AM whole chain, including model design, *in situ* monitoring and quality evaluation. Current challenges in applying NN to AM and potential solutions are also reviewed. Noriega et al. [9] propose a NN method for predicting dimensions and therefore improve accuracy of the distance between parallel faces

on FDM manufactured prismatic parts. The results show that the application of this methodology allows to reduce the manufacturing error up to approximately 50% for external dimensions and 30% for internal dimensions of prismatic parts. Koeppe et al. [14] combine experiments, FE simulations and NN algorithms to predict 3D-printed lattice-cell structures. The experimental results validate the FE model, which is used to build the database. Finally, a NN is trained to predict stresses, proving to be significantly faster than a full FE simulation. Hajializadeh and Ince [15] propose an approach integrating NN with FE analysis to predict the residual stresses distribution of metal parts built by Direct Metal Deposition (DMD) process. Dataset is generated from thermo-mechanical simulations of DMD deposition of all parts, extracting the nodal temperature history and resultant stress distribution of the medium, which are the inputs and outputs of the NN model, respectively. Results of predictions showed that the novel approach is capable of accurate and efficient prediction of residual stress distributions and the computational time is significantly improved.

Data-driven approaches based on ML technologies have been increasingly adopted in a wide range of fields (such as computer science, aviation, healthcare, and the manufacturing industry) to build highly complex relationships in digital system [12]. The software **ODYSSEE CAE**, also part of Hexagon, is a platform that allows to apply Machine Learning, Artificial Intelligence, Reduced Order Modelling (ROM) and Design Optimization to workflows seeking to accelerate product design and development via real-time parametric simulation and optimization.

The general motivation of this work is to develop and implement an AI-augmented methodology using Python scripts that links ODYSSEE CAE and Digimat-AM to optimize the additive manufacturing process assuring the quality of printed parts, providing an alternative tool that reduces computation time over FE simulations. From the reported literature works, NN models are mainly chosen when applying ML to AM, while it was not possible to find studies establishing the possibility of applying other algorithms. This project aims to focus on the application of ROM models, comparing it's performance with NN and Bayesian neural networks (BNN), to optimize the manufacturing process by determining the optimal value of the parameters, which include laser power, scan speed and spacing, part's position and orientation. The proposed approach combines both FE analysis and ML models to predict warpage of different parts and the thermal cooldown, with reasonable accuracy and reduction in computation time. Previously, a proof of concept had been developed by Salmi [16] in the application of ODYSSEE for asserting warpage of a 3D printed part using Machine Learning (ML) and Digimat-AM. A ROM model to predict warpage on a specific geometry when varying three process parameters in an SLS process: scan speed, scan spacing and laser power was developed, using at least 15 AM simulations as data. The current work starts from the work done, expanding the proof of concept to multiple geometries, AM processes and parameters and also exploring the cooldown behaviour prediction.

2 ADDITIVE MANUFACTURING

The ASTM International [17] defines Additive Manufacturing as "a process of joining materials to make objects from 3D model data, usually layer upon layer, as opposed to subtractive manufacturing methodologies". Despite AM was initially used for rapid prototyping, nowadays its use has extended to end-use parts, becoming a rapid manufacturing process. However, the main disadvantage of AM is the dimensional inaccuracy in the printed parts. The as-printed shape still does not meet the as-designed model, as the model accuracy is a function of many different factors. One of the main factors of part inaccuracy in SLS is shrinkage during sintering [1]. There are multiple additive manufacturing systems available today. Digimat AM allows to work with SLS, FDM and FFF.

Selective laser sintering (SLS) is a three-dimensional printing process in which a powder is sintered or fused by the application of a carbon dioxide laser beam [2]. The main advantage of this process is that it fabricates the parts without using a support structure and it is applicable to a wide range of powder materials. In general, for any manufacturing process, the part aimed to manufacture is specified in an STL file. An STL is, according to ASTM [17], the file format for 3D model data used by machines to build physical parts. The STL format uses triangular facets to approximate the shape of an object. The format lists the vertices, ordered by the right-hand rule, and unit normals of the triangles, excluding CAD model attributes.

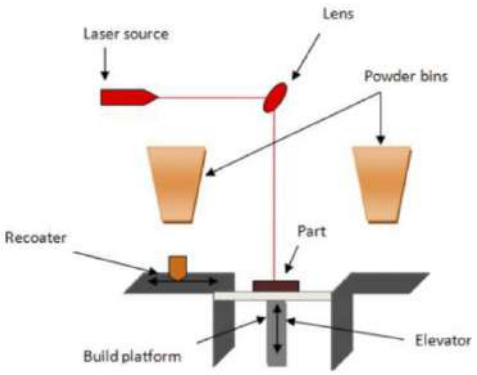


Figure 2.1: Schematic view of SLS process [1].

Fused deposition modeling (FDM) and fused filament fabrication (FFF) are additive manufacturing processes of layer by layer deposition of fused material beads. A thin filament of plastic feeds a machine where a print head melts and extrude it [2]. When manufacturing parts through these processes not only the STL file needs to be specified but also a toolpath file is required. It indicates the path through space that the tip follows on its way to produce the desired geometry of the workpiece. The pattern defined on the toolpath greatly influences various properties of the final printed part, such as its ultimate strength and elastic modulus [18, 19]. Optimization of toolpath design strategies are not covered in this work.

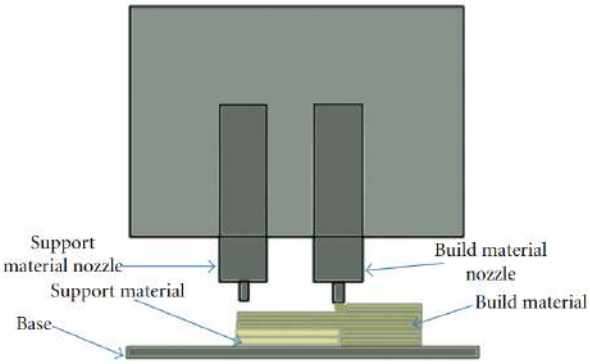


Figure 2.2: Schematic view of FDM [2].

Process parameters have a strong influence in the final shape of printed parts, therefore the selection and control of these parameters is key to achieve high dimensional accuracy by minimizing shrinkage. The relationship between SLS parameters and dimensional accuracy of produced parts has been studied by many researchers. Wang et al. [20] found that percentage shrinkage increases with increase in the scanning speed and hatch spacing, but decreases with increasing layer thickness, laser power, part bed temperature and delay time. They concluded that it's suitable to apply NN approach to study the SLS process. Negi and Sharma [1] used response surface methodology (RSM) and neural networks (NN) for predicting warpage and performed a sensitivity analysis to determine the most significant parameter for a composite. They concluded that shrinkage is more sensitive to the scan speed variations, but also part bed temperature and scan spacing are dominant parameters. Also, they highlighted the interactions between parameters affecting the final shrinkage. Raghunath and Pandey [21] used the Taguchi method to study the relation between shrinkage along X, Y and Z directions in cuboids. They observed that there is reduction in shrinkage as the energy density is increased. In addition, they reported that shrinkage in X-direction is mostly affected by laser power and scan length. Along Y-direction, by laser power and beam speed and along Z-direction by beam speed, hatch spacing and part bed temperature. Wang et al. [22] applied a genetic algorithm to determine the optimal SLS process parameters. NN models the relation between process parameters and shrinkage and is validated using experimental data.

On the other hand, there are other manufacturing processes available and not only process parameters affect the final warpage. Senthilkumaran et al. [23] investigated the shrinkage behaviour of polyamide parts with changes in part orientation, exposure strategies and other compensations. It was found that part orientation influenced the shrinkage pattern, and it was reported that the shrinkage is highly non-uniform along the Y (width) direction than along the X (length) direction and high shrinkage is observed at a lower scan length.

Not only the study of warpage is important to assess the quality of the printed part, temperature evolution is a key factor affecting the mechanical properties of

polymeric parts. Thermal history experienced by the parts during the SLS process could strongly affect to overcome incomplete particle melting and coalescence. These microstructural defects could further decrease mechanical performance of the SLS parts. Kiani et al. [24] studied the cooling-down process and part orientation, as influencing factors on the thermal history of SLS parts.

3 MACHINE LEARNING

Machine Learning (ML) is a subfield of AI that allows a machine or system to learn from data automatically and make decisions or predictions without being explicitly programmed [11]. It's structured in three main families: Supervised Learning, Unsupervised Learning and Reinforcement Learning [16].

Supervised learning is simply a formalization of the idea of learning from examples. The training set consists of n ordered pairs $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, where each x_i is some measurement or set of measurements of a single example data point, and y_i is the label for that data point. The validation data is a set of labelled data not seen by the model during the training stage, and the goal is to identify the labels of the test data with the highest possible accuracy.

Unsupervised learning is a data-driven ML technique which can uncover hidden patterns or group similar data together in a given dataset. It infers from unlabelled data [11].

Reinforcement learning allows the model to take actions in an environment in order to maximize the notion of cumulative reward. It differs from Supervised learning in that it does not require a training dataset, instead it is bound to learn from its experience.

As there are well-defined inputs and outputs for the AM processes, Supervised Learning is appropriate for them [25].

3.1 REDUCED ORDER MODELLING

Reduced Order Model is a data-driven model. Three main steps can be recognized: decomposition, reduction and reconstruction. ROM are based on modal decomposition, which allows to catch specific behaviors, and provide interpolating functions based on physical content of the signal, which are called modes (Figure 3.1). The main objective of all ROM techniques is to compress or reduce the volume of data to be handled and in consequence, the number of floating-point operations required for the response prediction [3].

On the following paragraphs the ROM steps are shortly described.

3.1.1 DECOMPOSITION

The aim is to establish a function (the ROM operator) which relates output data Y to input data X. Odyssee CAE proposes three different decomposition methods: Proper Orthogonal Decomposition (POD), Clustering (CLS) and Fast Fourier Transform (FFT).

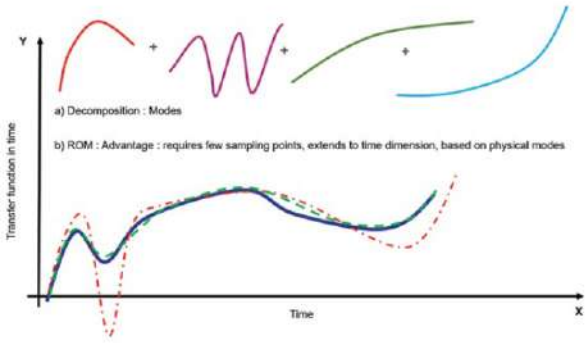


Figure 3.1: ROM modal decomposition [3]

In the first step, the response of the system is divided into multiple simpler responses to catch specific behaviors. The aim is to find a low-order representation of the system, meaning less computational requirements.

Knowing a set of data X inside a Design of Experiments (DoE), $X \in \mathbb{C}^{nxm}$, that represents a phenomena Y. It is possible to define the decomposition:

$$Y = G.S.H \tag{3.1}$$

Where G is a "spatial" matrix, S is a transfer diagonal or semi-diagonal matrix and H is a "modal" matrix.

A popular method used at this phase is Singular Value Decomposition (SVD), in which case the S matrix contains the singular values of Y. The POD is the SVD algorithm applied to partial differential equations (PDE) and it is one of the most important dimensional reduction techniques available to study complex, spatio-temporal systems [26]. Clustering takes the entire dataset and finds “groups of similar entities” within the dataset. There are several clustering methods, such as $Y = \sum \alpha\ y_p$. It can be interpreted as a weighted linear sum of a set of distributions, where y_p are weights or cluster functions and α are coefficients, both obtained via an appropriate learning or clustering algorithm.

In all cases, the model is projected from the Cartesian reference frame on to low-rank subspace with very useful properties such as orthogonality (in case of SVD decomposition) or locally compactness (in case of clustering), where it becomes simpler.

3.1.2 Reduction

The volume of data is reduced, keeping only a reduced number of modes/clusters that are enough to predict the expected curve within a reasonable accuracy. In POD, matrix S contains a set of singular values of descending order, that corresponds to the effect of the corresponding singular value on the complete reconstruction of the original base. The matrix S can be truncated, keeping only dominant singular values based on some tolerance criteria, and so subsets of G and H are used. If the original

data X and Y are of dimensions $X_{m,n}$ and $Y_{m,p}$ (assuming p»m), then the decomposed matrices may have the dimensions $G_{m,r}$, $S_{r,r}$ and $H_{r,p}$, where r represents the number of retained modes or clusters.

3.1.3 RECONSTRUCTION

Finally, during the reconstruction, the system is built back up again from the number of "modes" kept, after solving the simpler system.

To predict the results Y_N for a new parameter set X_N, G or H can be replaced by their modified updates G' or H', considering the change in the parameters, and obtaining a new set of response corresponding to slightly modified new positions X_N compared to the original X. To make predictions of the effect of moving X to X_N it's only necessary to consider the effect of G moving to G' or H moving to H'. The updated G' or H' may be simply obtained by interpolation techniques. RBF, Kriging or Inverse Distance (InvD) are used in this last step.

3.2 NEURAL NETWORKS

Artificial Neural Networks (ANN) is the most used ML algorithm due to its ability to capture nonlinear functions using large datasets. ANN can represent highly complex relationships with non-linearity between inputs and outputs. It contains three types of layers: an input layer, one or several hidden layers and an output layer [4]. Each layer contains neurons, which are called units, and they are interconnected from one layer to another. These connections have weights that determine the influence of one unit on another unit. The input layer receives data that the neural network needs to analyze or learn about. Then this data passes through one or multiple hidden layers that transform the input into data that is valuable for the output layer. Finally, the output layer provides an output in the form of a response of the Artificial Neural Networks to input data provided.

Throughout this work, the mention of NN may be referring to ANN.

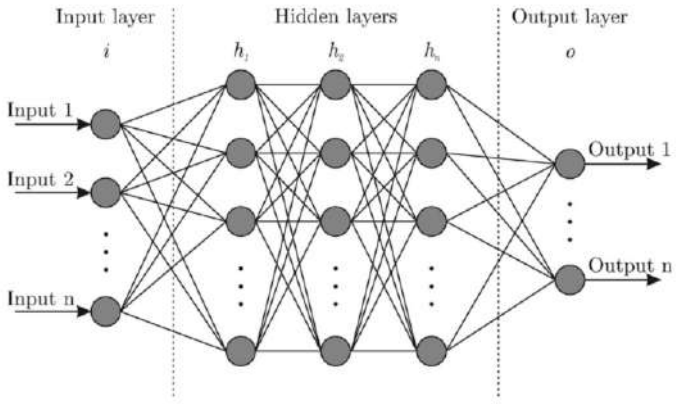


Figure 3.2: Structure of ANN [4]

3.3 Bayesian Neural Network

Bayesian neural network (BNN) combines neural network with Bayesian inference. It quantifies the uncertainty introduced by the model in terms of output and weights so as to explain the trustworthiness of the prediction. It has two advantages over standard NN. First, the built-in variability in BNNs makes them resistant to model overfitting. Secondly, inputs where the model is uncertain of its prediction can be identified. The main disadvantage of BNN compared to NN is that it is significantly more complex, thereby requiring large datasets and also higher computational power for training such networks.

3.4 ODYSSEE CAE

Odyssee CAE is a platform that allows to apply Machine Learning, Artificial Intelligence, Reduced Order Modelling (ROM) and Design Optimization to workflows through real-time predictive modeling and optimization for both CAE simulation data and physical test data. The software proposes to exploit available data on the input-output relationship to predict new responses on unseen inputs using ROM models. These methods require a Design of Experiments (DoE). A specific design includes the variables in a model and their range of application, in order to control the environment where the experiments are run and thus the data-generating process. The different variables settings are tested and their results are saved to be used in the predictive model, conforming the database that the algorithm uses to learn. The proper selection of these points needs to assure a good cover of the design space, in order to reduce interpolation errors. Various methods providing different distributions exists, such as uniform, normal or Latin Hypercube distributions.

In ODYSSEE CAE, a ROM model can be created based on csv files (X-file.csv for input data and Y-file.csv for output data), corresponding to variables and results from simulation models for different variable settings in the design space (DoE) [5].

The given problem can be simply explained as follows: given $F(X) = Y$, the objective is to find $F(X_{new})$. In Figure 3.3, the response y associated with the combination of variables X, represented by the green dots, is known; whereas the response $y = F(X_{new})$ associated with the new set of variables X_{new}, represented by the red dots, will be predicted.

On the following sections, the solver algorithms offered by the program are detailed, and an introduction to the basis of ROM is made.

3.4.1 Predictive models available in ODYSSEE CAE

The different solvers existing in Lunar are presented in table 3.1, extracted from [5]. The solver choice depends on the size of data analysed, the required accuracy, and the CPU performance.

Spatial interpolation can be classified into two main categories; point interpolation and area interpolation. The methods reviewed correspond to the first group. Point interpolation is divided into two subcategories; exact methods and approximate methods, respectively, whether or not they preserve the original point values. In that

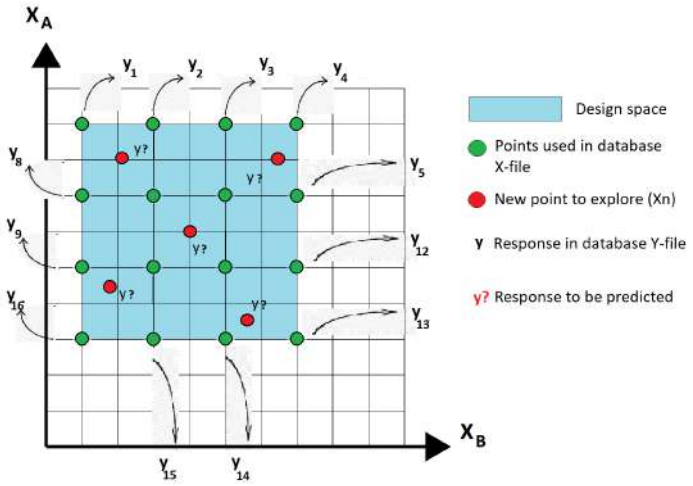


Figure 3.3: DoE and database distribution [5]

Methods	Solver	Type of use
Interpolation methods (direct interpolation)	Kriging	Time response Scalar response
	RBF ARBF	Time response Scalar response
	InvD	Time response Scalar response
Reduced Order Modelling method	POD RBF POD Kriging POD ARBF POD InvD	Animation Time response Scalar response
	Clustering RBF Clustering Kriging Clustering ARBF	Images Animation Quick for big data (>1e+06)
	FFT RBF FFT Kriging FFT ARBF FFT InvD	Periodic time response
Clustering Method (classification method)	SVM	Time response Scalar response

Table 3.1: Available solver algorithms in Odyssee CAE [5]

sense, Kriging is a exact method, whereas Inverse Square Distance is an approximate one [27].

As defined in [27], Kriging states the statistical surface as a regionalized variable, with a certain degree of continuity. Its main characteristic is the generalized covariance *k(d)*, which for a two-dimensional case is defined as:

$$k(d) = \begin{cases} 1 + \frac{1}{c_1}(\frac{dc_0}{d_{max}})^2 \ln(\frac{dc_0}{d_{max}}) & if\ d \leq d_{max} \\ 0 & if\ d > d_{max}, \end{cases} \tag{3.2}$$

where d_{max} is the maximum correlation distance considered, c_0 is the value for which *k(d)* function has its minimum, which equals -c_1. The statistical surface is obtained with the formula:

$$f(x,y) = a_1 + a_2 x + a_3 y + \sum_{j=1...N} b_j k(d_j), \tag{3.3}$$

where a_i, b_j are calculated resolving the following linear system:

$$\begin{bmatrix} k(d_{1,1}) & k(d_{1,2}) & ... & k(d_{1,N}) & 1 & x_1 & y_1 \\ k(d_{2,1}) & k(d_{2,2}) & ... & k(d_{2,N}) & 1 & x_2 & y_2 \\ ... & ... & ... & ... & ... & ... & ... \\ k(d_{N,1}) & k(d_{N,2}) & ... & k(d_{N,N}) & 1 & x_N & y_N \\ 1 & 1 & ... & 1 & 0 & 0 & 0 \\ x_1 & x_2 & ... & x_N & 0 & 0 & 0 \\ y_1 & y_2 & ... & y_N & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ ... \\ b_N \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ ... \\ v_N \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.4}$$

The first N equations express the condition that Kriging is an exact interpolator, while the further three find the coefficients of the plane that fits data points.

The Inverse Square distance method (InvD) gives more weight to nearby points than to distant points. Stating *f(x, y)* the analytical expression of the surface, it is possible to define:

$$f(x,y) = \frac{\sum_{j=1...N} w(d_j) v_j}{\sum_{j=1...N} w(d_j)}, \tag{3.5}$$

where N is the number of data points, v_j is point j value, d_j is the Euclidean distance with point j, and *w(d)* is the weighting function, defined:

$$w(d) = \begin{cases} \frac{1}{d_{min}^2} & if\ d \leq d_{min} \\ \frac{1}{d^2} & if\ d_{min} < d d_{max} \\ 0 & if\ d d_{max}, \end{cases} \tag{3.6}$$

where d_{min} is minimum distance and d_{max} is maximum distance. Index d_{min} prevents infinite weight values for $d = 0$, while index d_{max} avoids using too distant points. If no points fall into the circle of radius d_{max}, average data value is taken.

Rippa [28] defines a Radial Basis Function (RBF) as a distance-based function with dimension-independent and mesh-free features, and it is suitable for the surface simulation of scattered data. Suppose that $F = F(x)$, $x \in \mathbb{R}^d$ is a real-valued function in d dimensions. Therefore, an RBF approximation to F is a function S of the form:

$$S(x) = \sum_{i=1}^{N} a_i \Phi(\| x - x_i \|), \quad x \in \mathbb{R}^d, \tag{3.7}$$

where $\Phi(t)$, $t \geqslant 0$, is a fixed real-valued function and $\|.\|$ denotes the Euclidean norm. The points x_i, $i = 1, ..., N$, are called the center of the RBF interpolation. N is the number of interpolation points and a_i is the undetermined coefficient. The multi-quadratic (MQ) function is one of the most commonly used types of RBFs [29]. On the adaptive radial basis function interpolation algorithm (ARBF) the shape factors of MQ functions are determined adaptively by the local point densities of the points to be interpolated.

3.5 Quality indicators

All ML models need a performance metric to judge performance. In order to evaluate regression models, the R-squared and root mean squared error metrics are used in this work.

3.5.1 R-squared score

R-squared (R^2) is a metric for assessing the performance of ML models. It measures how much of the dependent variable variation is explained by the independent variables in the model. It is then a measure of fit [30, 31] to judge how predicted values fit the measured ones. It can be appreciated graphically on how well points fit a line in a plot of predicted versus actual values, as Figure 5.2 showed later on this work. The formula for calculating R^2 score is:

$$R^2 = 1 - \frac{\sum (y_{pred} - y_{actual})^2}{\sum (y_{actual} - \overline{y_{actual}})^2}, \tag{3.8}$$

where y_{pred} are the predicted values, y_{actual} are simulation values, and overline indicates average value. The value of R^2 ranges from 0 to 1, from none of the variance being explained to all variance being explained, respectively. Generally, an R-squared value between 0.75 to 1 can be interpreted as a significant amount of variance in the dependent variable is explained by the independent variables in the regression model.

3.5.2 Root Mean Squared Error (RMSE)

RMSE is a measure of model performance given on the same scale as the prediction target. In a general sense, *RMSE* can be interpreted as the average error that the model's predictions have in comparison with the actual, with extra weight added to larger prediction errors. It is the square root of the mean squared error between the predicted and actual values. It helps understand the model performance over the whole dataset [32]. The formula for calculating *RMSE* is:

$$RMSE = \sqrt{\frac{\sum (y_{actual} - y_{pred})^2}{Number\ of\ observations}} \tag{3.9}$$

RMSE is a metric which ranges from 0 to infinity, where the closer the score is to 0 the better performing the model is. So the *RMSE* value should be as low as possible. The metric that *RMSE* produces is on the same scale as the unit being predicted, so a good *RMSE* value can only be evaluated within the dataset context.

4 PROCEDURE

4.1 DEFINITION OF THE PREDICTION MODELS

Three independent tasks has been defined in this work, according to the type of manufacturing process (SLS, FDM and FFF) studied, the input parameters and the simulation output aimed to predict:

1. Predict warpage for varying process parameters, part position and orientation for SLS process
2. Predict warpage for varying part position for FDM process
3. Predict temperature evolution for selected nodes for SLS process to estimate cooldown time.

This section describes the general workflow defined to apply Machine Learning models to Additive Manufacturing process, seen in Figure 4.1, the definition of the input variable parameters and their range of values, the output aimed to predict at each case and the design of experiments (DoE).

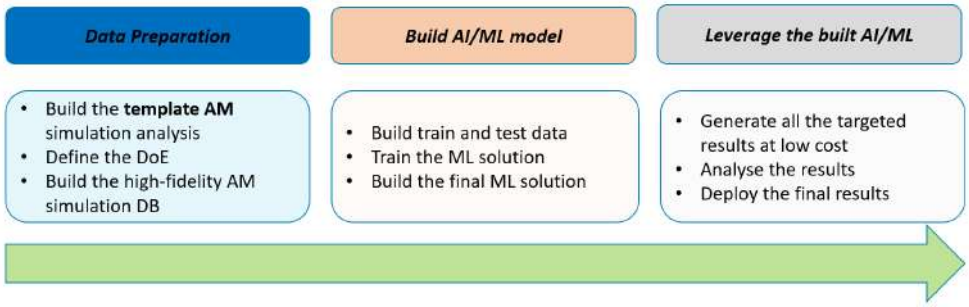


Figure 4.1: Typical workflow for AI/ML applied to CAE and specified to AM

Three main steps can be recognised. The data preparation process consists of defining the DoE, the output data (Y data) and the generation of the database itself, which throughout this work will be conformed by simulation results, not experimental data, since the objective of this study is to offer an alternative to the work done by Digimat-AM, reducing the time needed in the design phase. The AM template is the file gathering all the important information defining the manufacturing process that Digimat-AM uses to run the simulations conforming the database. Once the database is generated, a ROM model is trained using Odyssee CAE, and the best solver

algorithm is determined. Predictions on the validation data help to assess the model's accuracy by comparing the values to the ones obtained through simulations. The quality indicators used to determine the accuracy of predictions are R^2 and *RMSE*.

The workflow is developed by using a Python script that calls both Digimat-AM and Odyssee-CAE, in order to perform the following tasks:

1. Define and generate input values inside a given range, following a determined distribution
2. Run the specified number of simulations in Digimat-AM with each dataset, varying for each one the value of selected process parameters
3. Arrange and organize results into csv files to create the X and Y input simulations database
4. Run ODYSSEE CAE to train the ML model, determine best algorithm to make predictions on data and fit the validation data.

The ROM model trained is exclusive to the geometry used, so the workflow shall be applied to different geometries to assess robustness. However, results displayed in following sections corresponds to predictions made upon component in Figure 4.2.

Figure 4.2: Aerospace rocker model obtained by additive manufacturing

4.2 VARIABLE PARAMETERS AND RANGE OF VALUES

The variable parameters include process, position and orientation parameters, each group including 3 inputs variables. The first one contains the laser power, scan speed and spacing, while the second and third are the position in X, Y and Z direction inside the printing chamber and the rotation of the part along the 3 axis, represented by angles α, β and γ for rotation along Z, Y and X axes respectively. Figure 4.3 shows the relevant parameters of SLS process.

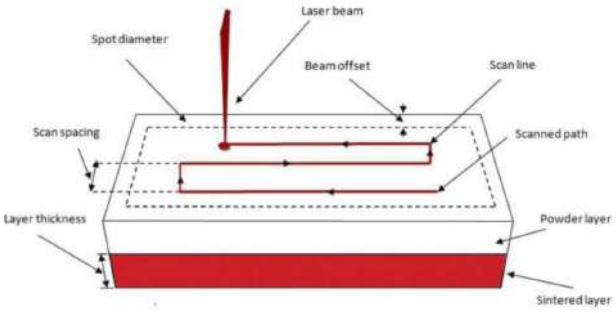


Figure 4.3: Schematic diagram of SLS process, showing the relevant parameters [1].

For the first task, to predict warpage for SLS process, it is desired to study the influence of the variation of all 9 parameters in the deflection of the nodes. To do so, it is divided into two subtasks.

A Study only the contribution of process parameters, while position and orientation values remain fixed.

B Study the influence on warpage when varying all 9 parameters

For the second task, as the studied process is FDM and the toolpath file should not be modified, the variable parameters are only 2, position along the X and Y axis. Finally, to build the ML model to predict cooldown time for SLS, only the 3 process parameters variation has influence on the temperature evolution of the printed part.

The definition of the application range varies for each group of parameters. In the case of position parameters, it depends on the size of the printing chamber and of the part itself. For the study cases analyzed in this work it is defined fully manually, visualizing the part on Digimat-AM user interface. Automatizing this step could be an update on future versions of the python script. For the manufacturing parameters, the definition of the range is carried out by introducing the energy density concept and its acceptable range of values. An increase in energy density leads to a better fusion of particles. For small energy density values, the sintering of the powder does not occur in the SLS process, but too high values of energy density could degrade the material [1, 21].

$$Energy\,density = \frac{Power}{speed * spacing * thickness * volume\,factor} \tag{4.1}$$

To avoid bad quality printing, energy density values are kept between 0,15 and 0,25 J/mm3 and adjusted working ranges are defined in Table 4.1, in the basis of previous studies and suggestion of domain experts.

Inputs		**Outputs**
Parameters	**Range**	
Laser power	[10000-50000] mW	Warpage at each STL node
Scan speed	[1000-12500] mm/s	Thermal evolution at 5 nodes
Scan spacing	[0.1-0.4] mm	
Part position in x, y, z	Manually determined [mm]	
Part rotation in x, y, z	[0° - 360°]	

Table 4.1: Input parameters and output conforming the X and Y database

4.3 MINIMUM NUMBER OF SAMPLING POINTS

Lunar User Guide [5] recommends to use a polynomial formula to estimate the minimum number of sampling points required to model an input-output relationship for a n dimensional input and 1 dimensional output.

$$Min\ number\ of\ sampling\ points = C_d^{n+d}, \tag{4.2}$$

where n is the number of variables and d is the order of the formula (order d = 2).

For the analysed cases, the minimum number of sampling depending on the number of inputs is displayed in Table 4.2. These are suggested values, and so a bigger database is usually built in order to assess the evolution of the accuracy of the model with the number of training samples.

Number of inputs n	**Minimum number of sampling points**
2	6
3	10
9	55

Table 4.2: Minimum number of sampling points for different number of variables.

4.4 DESIGN OF EXPERIMENTS, TRAINING AND VALIDATION DATA

The training data corresponds to input variables and results from simulation models for different variable settings distributed in the design space (DoE). The information in the train data forms the basis for training the model.

In order to evaluate the performance of the predictive model, for each trained algorithm 5 datasets are generated and not used to train the model. These sets are referred as the validation data. The total number of simulations performed must

include not only the training data, but also the sets considered as validation data. Dividing the database into training and validation subsets is done randomly, in order to guarantee that both subsets are random samples from the same distribution.

The training and validation sets are formed by a series of simulations whose results are deflection at each node of the STL file for the first two models, and the thermal history at 5 randomly selected nodes for the last case. The distribution of these simulations in the space should be homogeneous, so that the greatest number of possible interactions can be taken into account. Therefore, the limits of the DoE coincide with the limits of the ranges for the input variables of the model. Prediction on extrapolated values of inputs are not guaranteed by the ML models.

The type of distribution of the simulations depends on the number of inputs considered. As mentioned in the previous section, for the first case, the variable parameters has been separated so that, on the first place, a study considering only three inputs is made. Here, a uniform distribution of the input values is chosen, while when increasing the number of inputs up to nine variables, it has been observed that using a Latin hypercube sampling reduces the number of datasets needed at the training stage to get good predictions [33].

In uniform distribution, samples are distributed randomly in the design space, while in Latin hypercube sampling, samples are distributed evenly over a sample space. For two independent variables, x and y, the sample space of each variable is divided into n evenly spaced regions and a random sample from each sample space is picked to obtain random values across two dimensions. This is easily extended to more dimensions. Each variable is simply split into evenly spaced regions and random samples are then chosen from each region to obtain a controlled random sample.

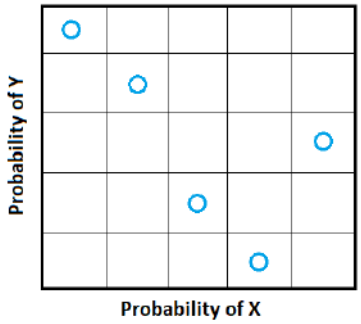


Figure 4.4: Two-dimensional Latin hypercube sampling

5 RESULTS

5.1 WARPAGE PREDICTION FOR SLS

5.1.1 VARIATION OF PROCESS PARAMETERS

Figure 5.1 shows the evolution of the average values for the R^2 score and the *RMSE* indicators with the number of simulation sets used as training data, for both the training and validation sets. The fitting of the training data is best when using 10 and 15 training sets. When observing the validation sets, R^2 score for 10 training samples is above 98.7%, and for 15 is above 99.5%. Even though the curve for both indicators becomes fairly stable from 15 simulations and above, there is a small difference between the score values for 10 and 15 simulations.

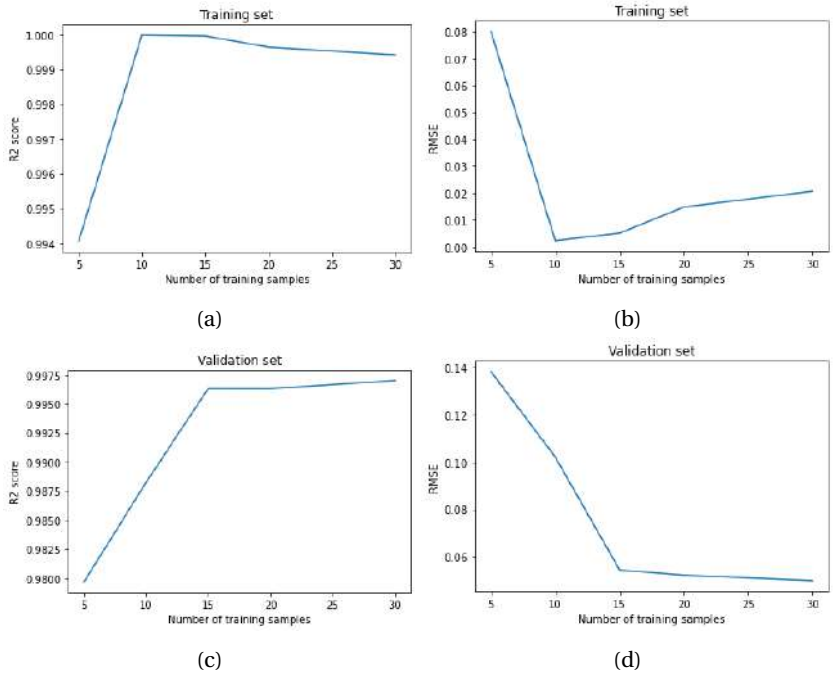


Figure 5.1: Evolution of quality indicators with the number of training sets. a) R^2 and b) *RMSE* on the training set and c) R^2 and d) *RMSE* on the validation set.

Figure 5.2 shows the predicted values versus the simulation values for the deflection output, for the ROM model trained with 15 datasets, and the values of the *RMSE* and R^2 for the validation set can be seen in Table 5.1.

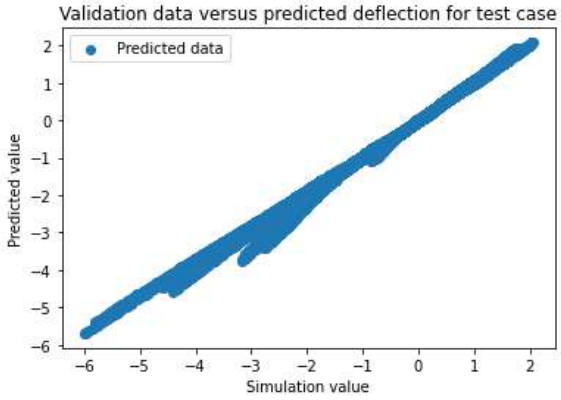


Figure 5.2: Fitting for the validation set between predicted values and simulation results, reported for the deflection output

Output	Number of training samples	R2	RMSE
Deflection in x, y and z axis	15	99.63%	0.055

Table 5.1: Quality indicators for the prediction of the deflection over the validation set when training the model with 15 sets.

Figure 5.3 shows the warped geometry for one validation set, comparing a) simulation and b) predicted outputs. The difference between both can be seen in c).

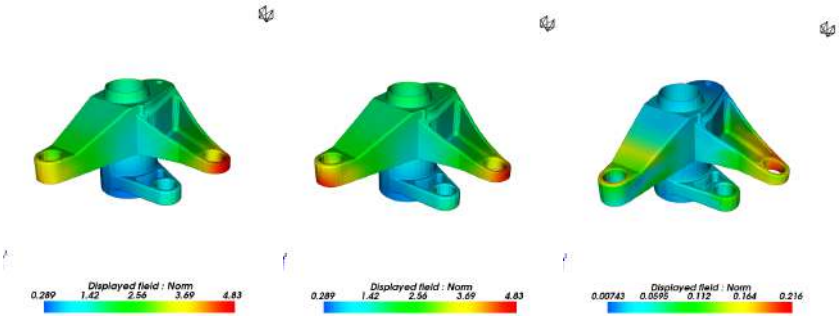


Figure 5.3: Warped geometry a) from simulation and b) predicted. c) displays the difference between predicted and simulated part.

Models using Neural Networks and Bayesian Neural Networks are also trained. Figure 5.4 shows the comparison between these results and the ones obtained with ODYSSEE, where it can be seen that BNN displays a R2 above 99% already from 5 training sets, while NN stabilizes only from 20 training sets.

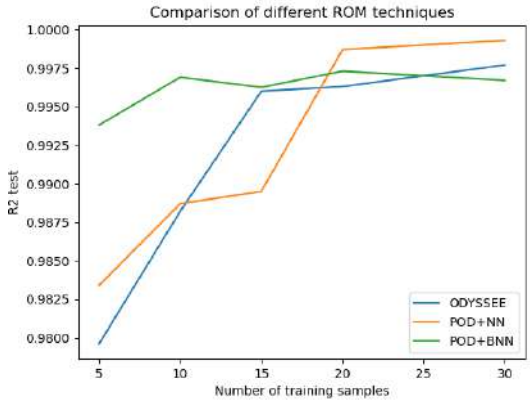


Figure 5.4: Effect of increasing the number of training samples for different algorithms.

5.1.2 Variation of process, position and orientation parameters

From the first use case, the training of the model when varying process parameters does not represent big difficulties. In the Appendix several models are trained for different geometries in order to asses robustness of the model. Results for all the studied models confirm that there is a straightforward relationship between warpage and the analyzed process parameters, being easy to the ML model to make predictions over it. The problem changes when the orientation of the part inside the printing chamber is introduced as a variable parameter. As a first warning, the number of inputs to the ROM model increases from 3 to 9, which is already expected to significantly increase the minimum number of training sets required, from 10 to at least 55 samples, following Equation 4.2. Therefore, the number of simulations performed in the stage of database generation increases, and with it, the time required by this step.

From Figure 5.5 and Table 5.2 (labeled as case A), it can be seen that when using 120 sets as training data, the R^2 score obtained is 11.6%. Therefore, the predicted values for the validation set do not match the simulation values, and the model is not reliable.

In order to understand the weight of the different input parameters two other models are built:

B. varying only the rotation of the part in the three directions while the process and position parameters remain fixed (variable parameters and inputs are alpha, beta and gamma)

C. varying all 9 parameters but only considering the rotation angles as inputs for the ROM model.

The R^2 scores obtained are displayed in Table 5.2.

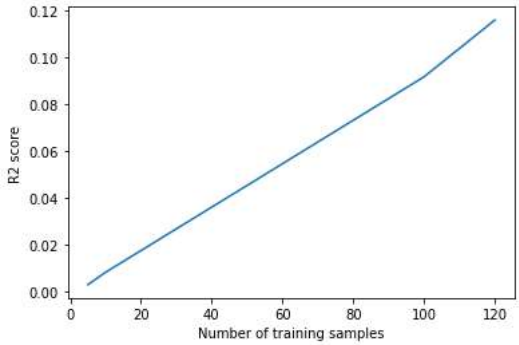


Figure 5.5: Evolution of R2 with the number of training sets for the ROM model built with 9 input parameters

Case	N = 5	N = 10	N = 60	N = 100	N = 120
A. All inputs	2.7	0.8	-	9.17	11.6
B. All angles	26.98	37.28	76.95	81.24	89.84
C. All parameters variation Only angles for ROM	4.4	6.9	93.51	98.82	99.27

Table 5.2: R2 score for validation set for different models considering part's orientation.

Figure 5.6 a) and b) shows the evolution of the average values for the *RMSE* and *R2* indicators for the training set, when all parameters are varied but only the rotation angles are used as input data. The evolution of indicators shows that *RMSE* achieves a minimum when using 30 samples. Simultaneously, R^2 stabilizes at 30 samples. Both indicators reach a fairly stable value for a number of sets above 30. Nevertheless, when observing the evolution of the indicators for the validation set (Figure 5.6 c) and d)), the R^2 value for 30 training samples is below 80%. This level of accuracy is generally unsatisfactory for the customer needs, so at least 60 training sets should be used in order to assure R^2 values above 90%. After 80 training samples, the curves for the validation set start to stabilize into fairly stable values.

After the study of different cases, the final approach adopted when optimizing the manufacture process will depend on the customer needs. If the customer needs to optimize the value of all 9 parameters to minimize the final warpage of the part, then the easiest way to do so in terms of time and computational power required found in this study is by steps. Fixing orientation parameters, process and position parameters can be optimized with a model using using a small number of training data. Once this is done, the model to optimize rotation is built.

Optimization of all 9 parameters simultaneously could also be done by building a higher database. This study did not went further on assessing an accurate model to optimize process, position and orientation parameters together because the budget and computational power required exceed the scope of this project.

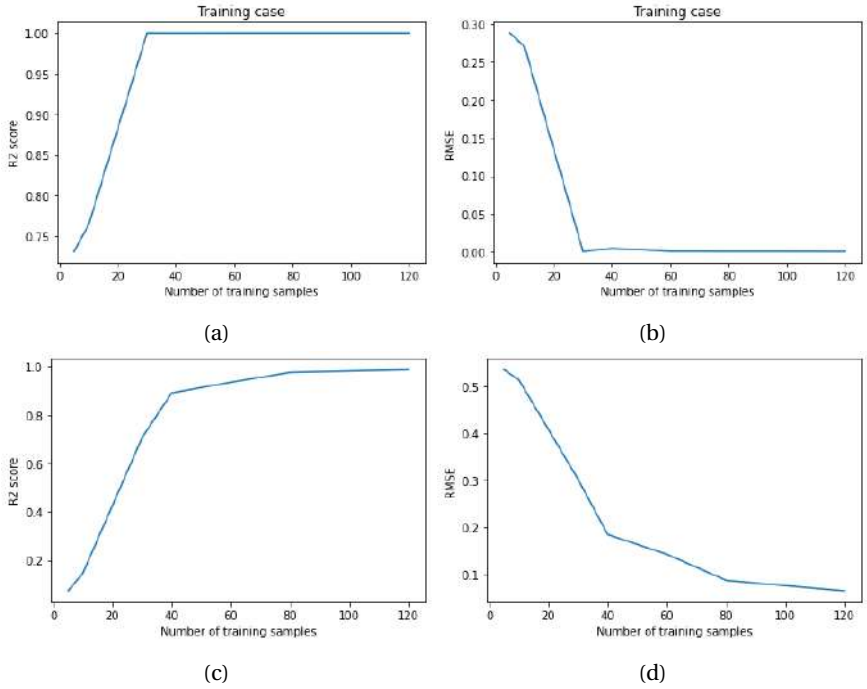


Figure 5.6: Evolution of quality indicators with the number of training sets. a) R^2 and b) *RMSE* on the training set and c) R^2 and d) *RMSE* on the validation set. The database is built varying process, position and orientation parameters but only the rotation angles are considered as inputs for the ROM model.

5.1.3 WARPAGE VISUALIZATION

The influence of process parameters on deflection appears to be fairly easy to predict, getting good results even for a small number of training data. Figure 5.7 shows the deflected geometry for two different simulations, whose corresponding process parameters values are in Table 5.3. The changes in deflection when varying process parameters are analyzed.

Sim	**Laser power [mW]**	**Scan spacing [mm]**	**Scan speed [mm/s]**	**Energy Density [J/mm3]**
1	53009	0.149	12189.3	0.242
4	46535.1	0.157	11390.3	0.217

Table 5.3: Process parameters values for 2 different simulations on the validation set.

The zone of maximum warpage does not seem to change when varying process parameters and deflection along the Z direction is the most susceptible to energy density's variation. Along the Z direction, the maximum variations occur close to the

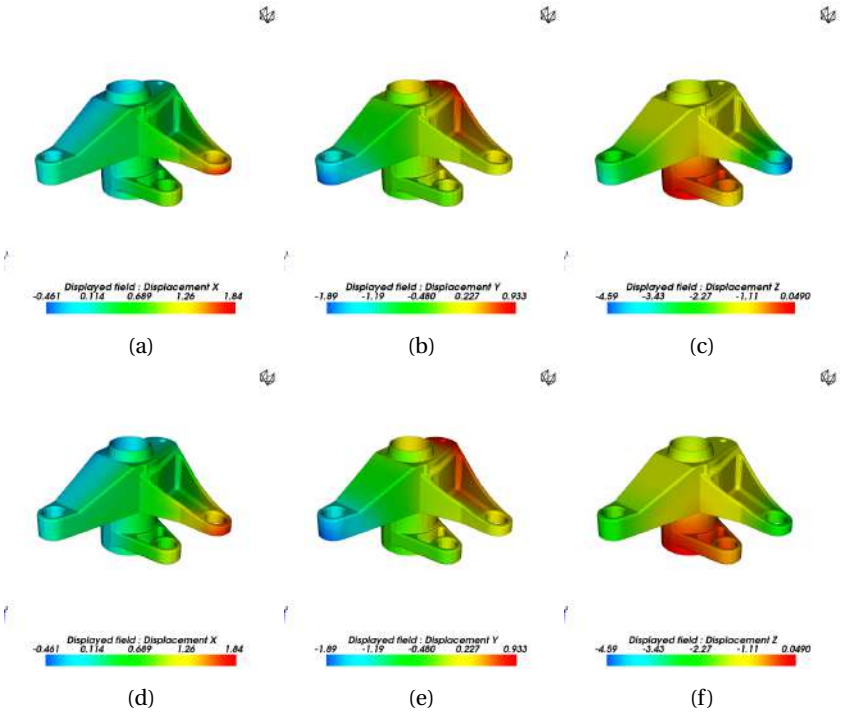


Figure 5.7: Warpage for validation case 1 along the a) X, b) Y and c) Z direction and for validation case 4 along the d) X, e) Y and f) Z direction.

edges of the part, representing between 25% and 40% of the deflection value.

On the other hand, when observing the deflection distribution for different orientations of the part, Figure 5.8, the region of high warpage is very sensitive to part orientation. 5.8 a), b) and c) correspond to warpage along the X, Y an Z axes respectively for part obtained by simulation 7 of the database, and e), f) and g) represent warpage for part obtained by simulation 11 (see Table 5.4). Note that the geometry is kept in the same orientation for both simulations for better visualization. Figure 5.9 shows how the part is printed for these two analyzed cases. The results are consistent with the difficulty introduced on the predictions by the variation of rotation angles.

Sim	**Energy Density**	**Alpha**	**Beta**	**Gamma**
7	0.226	193.23	40.85	99.36
11	0.166	240.9	233.02	310.12

Table 5.4: Energy density and orientation parameters for two different simulations in the database.

Deflection's distribution along the Y axis seems to be the less susceptible to the orientation of the part, while the warpage distribution along the X and Z axes is highly

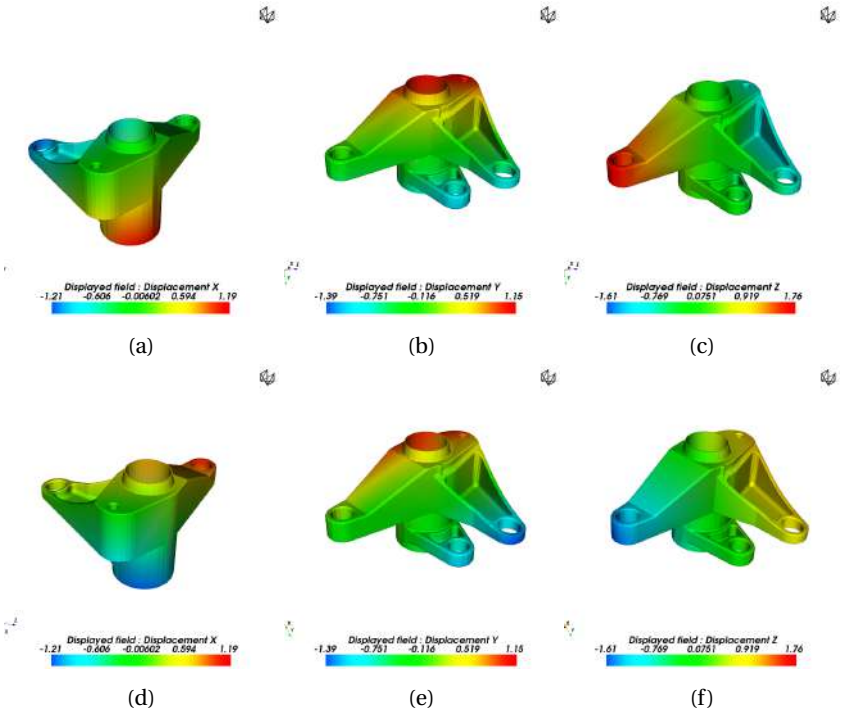


Figure 5.8: Warpage for simulation 7 along the a) X, b) Y and c) Z direction and for simulation 11 along the d) X, e) Y and f) Z direction.

different for both cases, being the bottom and edges of the part the most susceptible areas.

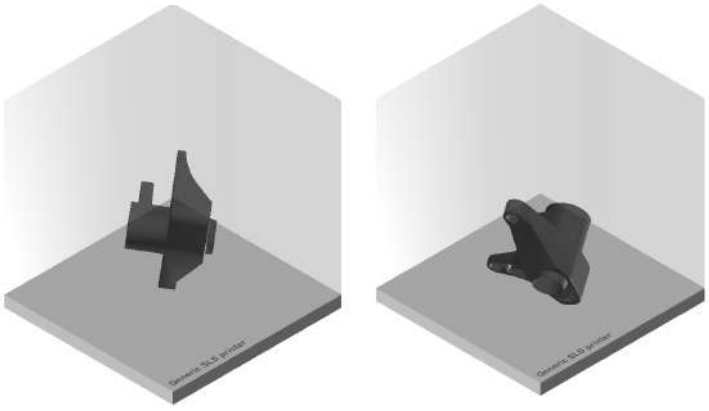
5.2 Warpage prediction for FDM

The minimum number of training samples suggested from Equation 4.2 when working with 2 input parameters is 6. From Figure 5.10, it can be concluded that only 5 training sets are enough to predict warpage in FDM process when varying position along the X and Y axes, as the R^2 score is already 99.99% .

The values for *RMSE* and R^2 for the validation sets for the model trained with 5 simulations are shown in Table 5.5.

Output	**Number of training samples**	**R2**	**RMSE**
Deflection in STL nodes	5	99.99%	0.0018

Table 5.5: Quality indicators for the prediction of warpage in FDM process over the validation set, when using 5 simulations as training data.



(a) (b)

Figure 5.9: Position inside the printing chamber for 2 different orientation angles configurations, a) displays simulation 7 and b) simulation 11.

In Figure 5.11 the total warpage for one validation case is presented, comparing a) simulation output and b) the predicted output. The difference between simulation and prediction is represented in c).

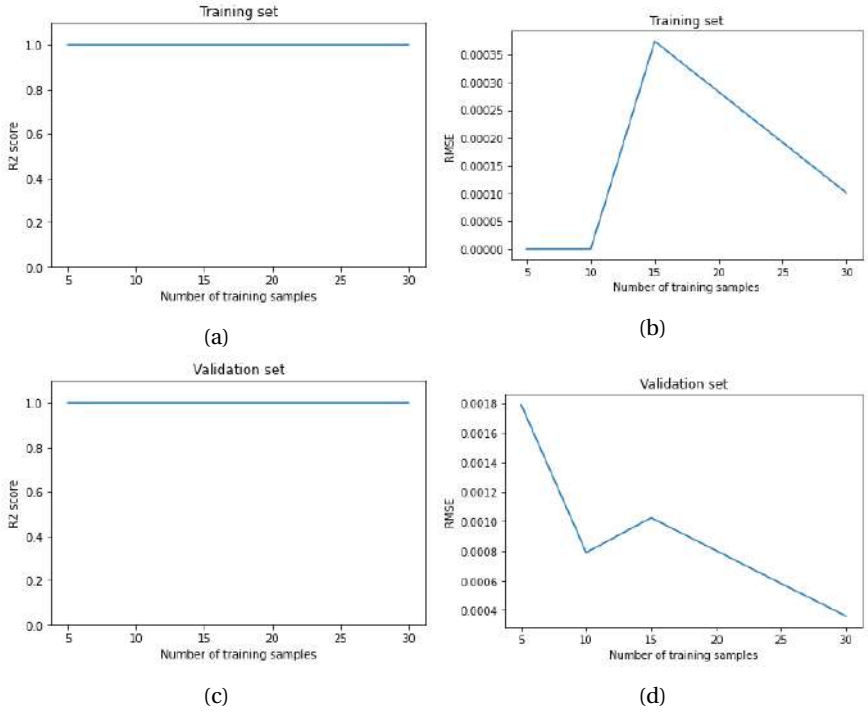


Figure 5.10: Evolution of quality indicators with the number of training sets. a) R^2 and b) *RMSE* on the training set and c) R^2 and d) *RMSE* on the validation set. The database is built varying position parameters for FDM process.

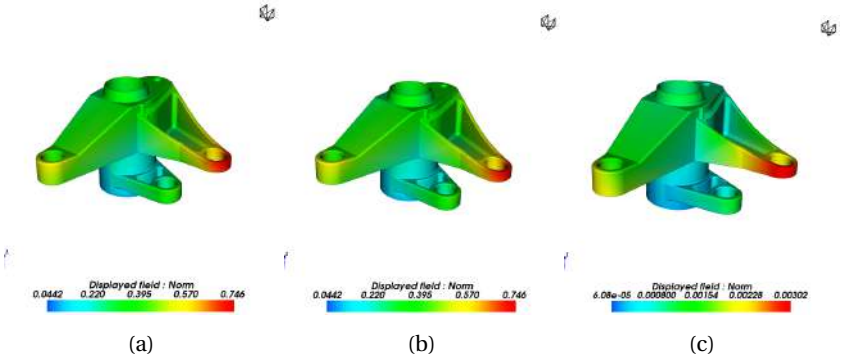


Figure 5.11: Warped geometry from FDM process a) from simulation and b) predicted. c) displays the difference between predicted and simulated parts.

Neural Networks and Bayesian Neural Networks algorithms are also tried, obtaining good predictions already with 5 training sets, but with a lower *R2* score than

ODYSSEE models (see Figure 5.12 and Table 5.6).

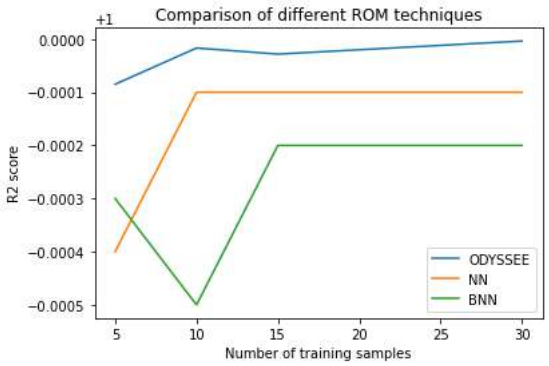


Figure 5.12: Effect of increasing the number of training samples for different algorithms

	R2	Relative error
NN	99.96%	0.017
BNN	99.97%	0.018

Table 5.6: Quality indicators for warpage prediction in FDM process with Neural Networks (NN) and Bayesian Neural Networks (BNN) models, using 5 simulations as trining data.

5.2.1 WARPAGE VISUALIZATION

Figure 5.13 shows the deflection in the X, Y and Z direction along the part for two different simulations varying part position, confirming that the region of maximal warpage is not sensitive to the position of the part inside the printer chamber. This could explain the easiness to fit the model. The warpage's variation with position represents between 5%-12% of the deflection in all 3 directions.

Sim	X [mm]	Y [mm]
1	-136.02	-80.8
3	78.5	113.5

Table 5.7: Position parameters values for 2 different simulations inside the DoE.

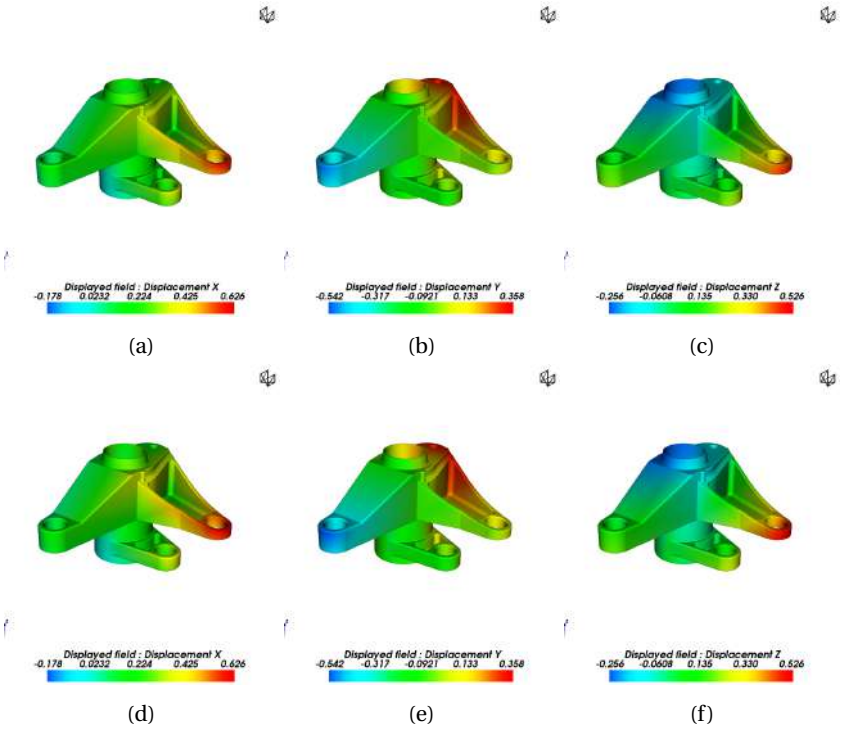


Figure 5.13: Warpage visualization for simulation with different part's position obtained by FDM process

5.3 TEMPERATURE PREDICTION FOR SLS PROCESS

Figure 5.14 shows the evolution of the quality indicators for both training and validation sets. R-squared scores above 99% are already reached for both the training and validation sets with 5 training samples.

The quality indicators R^2 and *RMSE* calculated for the validation set in Table 5.8 show that a ROM model trained with 5 sets already gives good predictions on thermal history for different process parameters in the printed part. However, when analyzing the curve for the predicted temperature versus the validation temperature values, Figure 5.15, a difficulty to predict high temperatures is found.

Output	**Number of training samples**	**R2**	**RMSE**
Deflection in x, y and z axis	5	99.91%	1.89

Table 5.8: Quality indicators on the validation set for the temperature prediction when training the model with 5 sets.

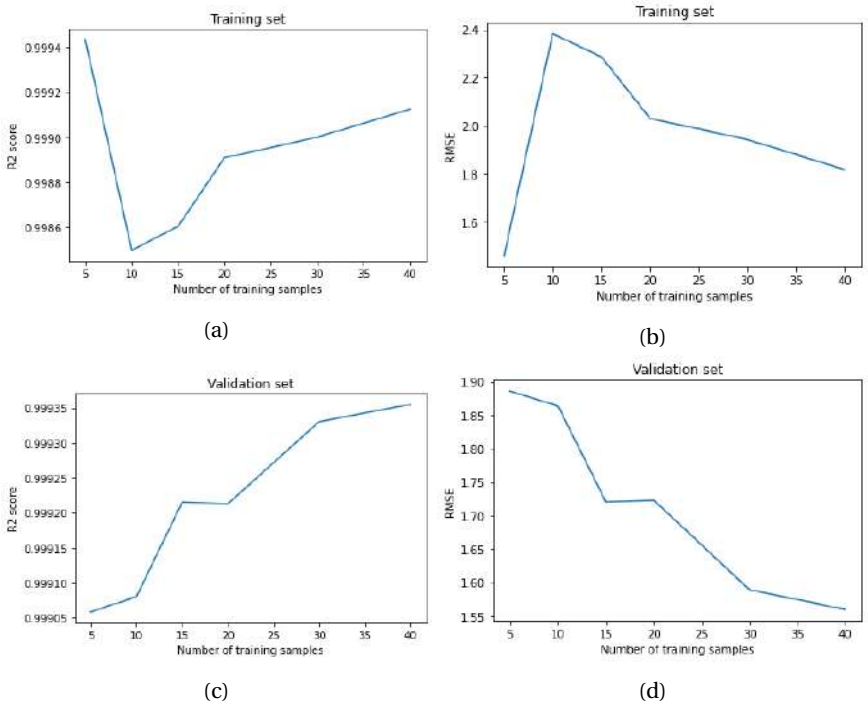


Figure 5.14: Evolution of quality indicators with the number of training sets. a) R^2 and b) *RMSE* on the training set and c) R^2 and d) *RMSE* on the validation set. The ROM model predicts thermal evolution on selected nodes of the part when varying process parameters in SLS process.

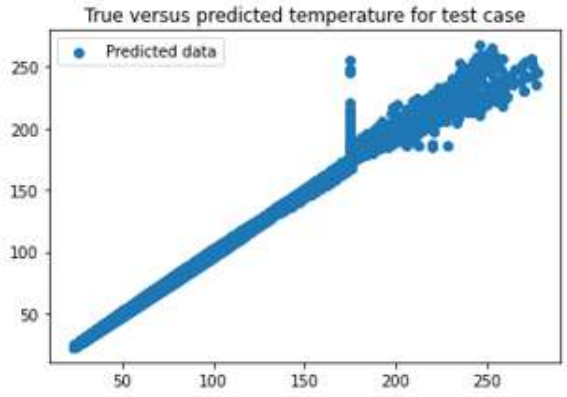


Figure 5.15: Predicted values versus simulation results for temperature evolution of a given node.

When observing the thermal evolution during the SLS process of a given node, Figure 5.16, it is possible to conclude that the model's struggle to predict high temperatures is in fact due to the high temperature variations in a short time interval. The high temperatures the model struggles to predict correspond to the peak in the curve, being this sharp discontinuity really hard to reconstruct by the model. It represents the moment when the laser reaches the node. The time it occurs depends on the scan speed and spacing, while the maximum temperature achieved varies with the energy density (see Table 5.9). The final time to reach the room temperature of 23° does not present significant variations for the 3 cases presented. It varies from 14.8 hours for simulation 2 to 15 hours for simulation 20.

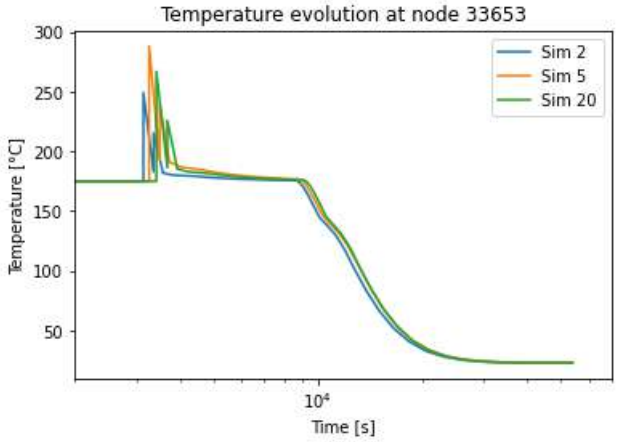
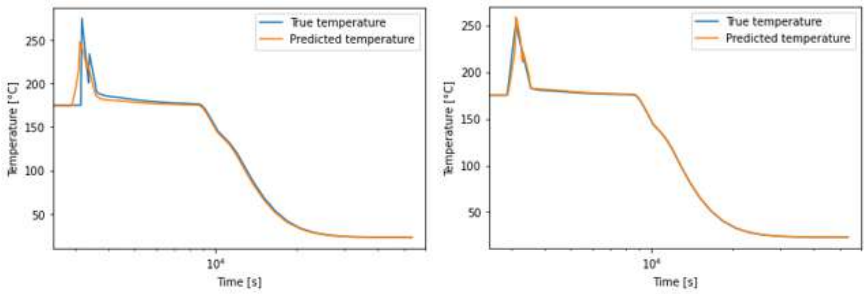


Figure 5.16: Temperature evolution of a node during the SLS process for different process parameters

Sim	Laser power [mW]	Scan spacing [mm]	Scan speed [mm/s]	Energy Density [J/mm3]
2	27242.83	0.35	4181.2	0.156
5	24049.17	0.39	2073.5	0.249
20	12006.8	0.15	3358.4	0.198

Table 5.9: Process parameter configuration for three different simulations on the database.

When superposing the predicted and simulation's thermal behaviour for different energy density values in Figure 5.17, it can be confirmed that the model is capable to predict the cooldown region accurately. On the contrary, the peak's prediction is bad due to the difficulty that it represents to decompose and reconstruct a sharp discontinuity for the ROM model [9].



(a) Energy density = 0.23 J/mm3 (b) Energy density = 0.16 J/mm3

Figure 5.17: Predicted thermal evolution of a node during the SLS process for different energy density values.

5.4 REQUIRED TIME TO COMPLETE THE WORKFLOW

In previous sections, it has been established that it's possible to fit a ROM model to make predictions on the additive manufacturing process. The solver algorithms offered by ODYSSEE CAE are able to perform with the same, or even better, accuracy than NN and BNN, which are the most common ML algorithms applied to AM in the available bibliography. However, to consider the application of ML to AM as a superior alternative to FEA, the computation time required to create the database and train a ROM model needs to be analyzed. Table 5.10 gives information about the time required to perform the different stages of the defined workflow for the different tasks defined.

Time to complete workflow

	Simulation [min/sim]	**Build and train model [min]**		
Task		**ODYSSEE**	**NN**	**BNN**
1.A	1.5	1	1	1
1.B	100	4	-	-
2	6	9	1.5	2
3	12	1	1	1

Table 5.10: Computation time required to complete established workflow.

The time required to complete Digimat-AM simulations depends mainly on the number of voxels of the model, while the training and determination of the best solver algorithm in ODYSSEE CAE depends on the size of the Y data. The value of fixed parameters that affect the simulation time are defined in Table 5.11 and the time consumed by each simulation expressed in Table 5.10 is referred to those settings.

The first step of the defined workflow is the most time-consuming, and the total time dedicated to the generation of the database will finally depend on the customers need. The higher the accuracy expected, the higher the number of training data needed.

Database - Digimat AM				
Task	**1.A**	**1.B**	**2**	**3**
Simulation type	Warpage - High fidelity			Thermal
Chamber size [mm]	400x400x400	150x150x150	914.4x609.6x914.4	100x100x100
Mesh size [mm]	3	2.3	1.27	2.4
Voxels	2854	287496	24623	74088
CPUs job generation	1	4	1	4
CPUs warpage/thermal	4	4	4	4
ROM model				
Nodel on STL	8061	8061	122355	8061
Y vector	24183	24183	367065	10000

Table 5.11: Value of fixed parameters used for different simulation cases.

Studying the influence of part's orientation in SLS process is the most challenging case regarding both accuracy and computation time. The time consumed to generate the database can be optimized by increasing the number of CPUs and using a cluster to run multiple simulations in parallel. Also when analyzing rotation angles to optimize the quality of the printed part, users may not consider a full rotation range along the 3 axes, but only some principal orientations, such as 45° and 90° rotations along the 3 directions and some other few angles. Reducing the DoE could significantly reduce the number of sampling points needed on the database.

Even if temperature during cooldown can be predicted accurately using AI, it can be computationally expensive with respect to the added benefit, so the current Digimat-AM approach to estimate cooldown time using Digimat-AM is not improved by the use of ML.

5.5 SIMULATION'S ACCURACY

Table 5.11 introduces the question on the level of refinement required for the simulation (the number of voxels) in order to obtain representative results. The refinement of the mesh size depends on the property aimed to study. From Digimat-AM experts experience, a coarse mesh size is already enough to catch the thermal evolution inside the printer chamber accurately. When studying warpage, the mesh size should be fine enough to catch the features of the geometry. Finally, to study the post-processing residual stresses a really fine mesh is recommended.

If the simulation parameters are defined correctly and the mesh size is chosen wisely, then Digimat-AM is expected to simulate the manufacturing process with an error that represents around 20% of the real distortion of the printed part. Calculating the accuracy of the simulation is a tricky task, since it strongly depends on the material model, the size and shape of the geometry, the repeatability of the printer itself (meaning how much the properties of the printing part vary when printing the same part multiple times with the exact same printer's configuration), the process parameters configuration and other factors.

6 CONCLUSIONS

This work presents a methodology for predicting warpage and cooldown behaviour by leveraging FE simulations and machine learning in the AM field, by developing a tool that links Hexagon softwares Digimat-AM and ODYSSEE CAE.

A model capable of predicting the final warpage of printed parts and cooldown time for different manufacture processes, by training ROM models in ODYSSEE CAE with data provided by simulations performed in Digimat-AM has been developed.

The best solver algorithm is evaluated by analysis of two quality indicators: R-squared score and *RMSE*. A good predictive model has been obtained for all the studied cases, requiring only 5 training sets to perform warpage optimizations varying process and position parameters and to predict cooldown time varying position parameters. Good accuracy to minimize warpage varying orientation parameters can also be achieved but increasing considerably the number of training sets, from 5 to at least 60 simulation used at the training stage.

The generation of the database is the most time consuming step representing at least the 70% of the total time required to complete the established workflow, as expected when defining ML models.

For tasks 1.A and 2, since the size of the database is small, there is an improvement on the computation time with respect to the FEA approach. On the other hand, predicting cooldown time using ROM models does not prove to reduce the computation time required by Digimat-AM, and so the use of ML for this particular use case is dismissed.

Applying this methodology to different geometries and manufacturing processes demands choosing different variable parameters and the definition of a new DoE. The robustness of the approach is evaluated on the Appendix section, confirming that the results obtained are general and can be extended to multiple geometries without further inconvenience. The computation time required can vary for each geometry, assuming the same hardware. For the generation of the database, the simulation's time depends on the number of voxels. For building the ROM model, it depends on the number of training sets and mainly on the size of a single Y vector. As ODYSSEE CAE tries more than 600 algorithms including POD, clustering and FFT models, a way to reduce the time consumed on the step of building the ROM can be reducing the number of tested algorithms, based on the studied problem.

Following these results, Hexagon can continue to develop a tool aiming at increasing manufacturing accuracy by reducing the time required in the design phase, improving the manufacture's experience. Next step is already defined: studying the mechanical performance of the printed part accounting material properties and microstructure uncertainty, combining more Digimat tools with ODYSSEE CAE. Also, further investigation on other thermal applications, such as control of the cooldown profile varying the chamber temperature to assess the final degree of crystallization of the part, can be defined. The possibility of training a ROM model independent of the geometry for SLS thermal analysis would be of value for studying the thermal cooldown.

7 APPENDIX

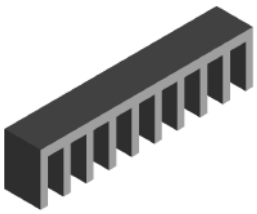
In this appendix different prediction models built for different geometries are presented. Results support what is presented throughout the report, proving robustness of the proposed workflow for different parts and DoE.

7.1 TASK 1: PREDICT WARPAGE FOR VARYING PROCESS PARAMETERS, PART POSITION AND ORIENTATION FOR SLS PROCESS

Four new use cases are evaluated for the variation of process parameters. On the other hand, only one new use case was established for the variation of orientation parameters, due to schedule limits introduced by the computation time consumed on the database generation step.

For all the use cases, all the ROM models built have an R^2 score above 99.99% already from 5 training samples. The graphs are omitted since there is no significant variation from Figure 7.7 a) and c).

NN and BNN were also tried for each use case, obtaining scores already above 99.99% when using 5 training sets. The computation time required to complete the workflow is shown in Table 7.5.



Database - Digimat-AM	
Simulation type	Warpage - high fidelity
Mesh size	3
Voxels	1350
CPUs job generation	1
CPUs warpage/thermal	4
ROM model	
Nodes on STL	123
Points on a single Y vector	369

Table 7.1: Beam cantilever use case. Variation of process parameters.

Database - Digimat-AM	
Simulation type	Warpage - high fidelity
Mesh size	6
Voxels	5918
CPUs job generation	1
CPUs warpage/thermal	4
ROM model	
Nodes on STL	4703
Points on a single Y vector	14109

Table 7.2: Lever bracket use case. Variation of process parameters.

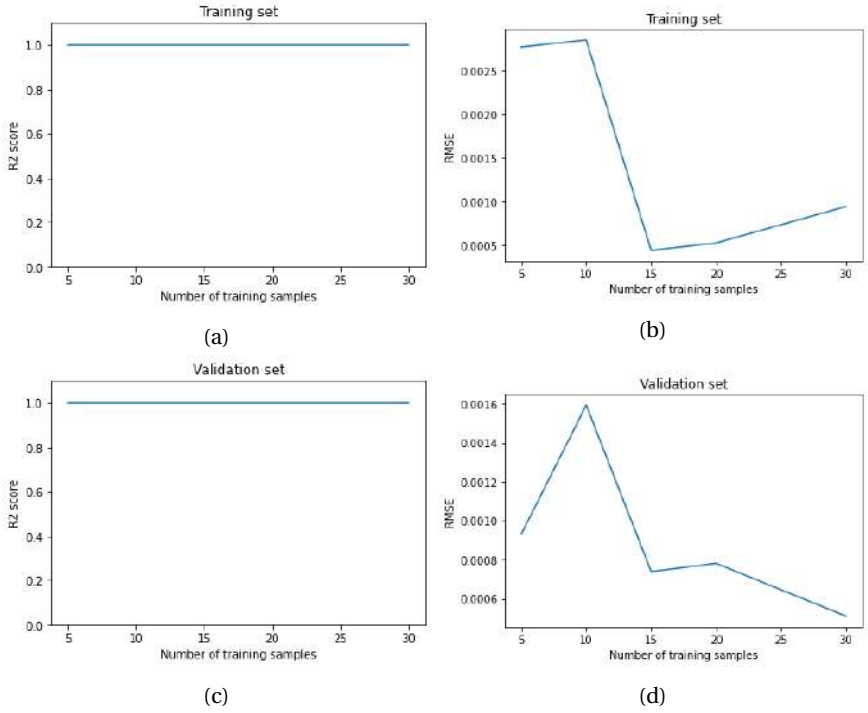


Figure 7.1: Evolution of quality indicators with the number of training sets. a) R^2 and b) *RMSE* on the training set and c) R^2 and d) *RMSE* on the validation set for the beam cantilever use case.

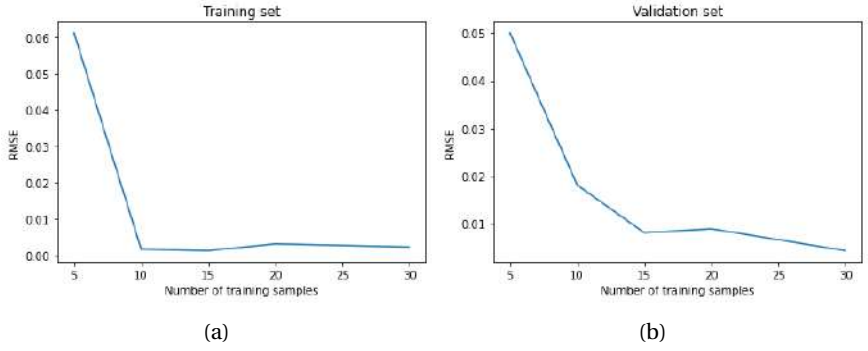
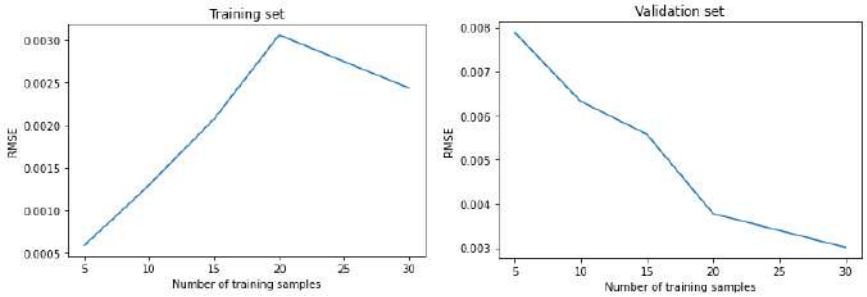


Figure 7.2: Evolution of *RMSE* value a)for the training and b) validation sets with the number of training sets for the lever bracket use case.

As mentioned, ODYSSEE CAE contains a script that tries and compare over 700 ROM models and choses the best algorithm based on the R2 score for the validation set. The greater the number of points conforming the Y data, the greater the time the

Database - Digimat-AM	
Simulation type	Warpage - high fidelity
Mesh size	6
Voxels	7639
CPUs job generation	1
CPUs warpage/thermal	4
ROM model	
Nodes on STL	13147
Points on a single Y vector	39441

Table 7.3: Valve's body case. Variation of process parameters.



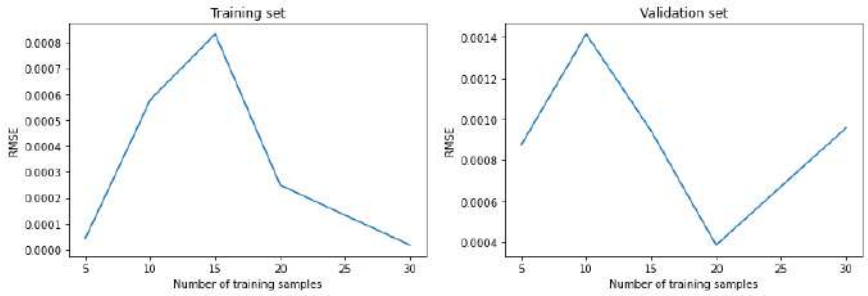
(a) (b)

Figure 7.3: Evolution of *RMSE* value a)for the training and b) validation sets with the number of training sets for the valve's body use case.

Database - Digimat-AM		
Task	**1.A**	**1.B**
Simulation type	Warpage - high fidelity	
Mesh size	1.7	3.5
Voxels	13422	222912
CPUs job generation	1	4
CPUs warpage/thermal	4	4
ROM model		
Nodes on STL	168702	
Points on a single Y vector	506106	

Table 7.4: Ring nut spanner case. Variation of process parameters (Task 1.A) and variation of all parameters (Task 1.B).

program needs to try all the different parameter's combinations. As a consequence, the required computation time to build and train the model can be significantly bigger than when using algorithms such as NN and BNN. One proposed solution is to reduce the number of configurations the program tries in order to select the best



(a) (b)

Figure 7.4: Evolution of *RMSE* value a)for the training and b) validation sets with the number of training sets for the ring nut spanner use case.

method, reducing it to POD or CLS methods only, for example. NN algorithms are expected to be introduced in future versions of ODYSSEE CAE.

Time to complete workflow

	Simulation [min/sim]	**Build and train model [min]**		
Case		**ODYSSEE**	**NN**	**BNN**
Beam cantilever	<1	<3	<1	<1
Lever bracket	1.5	<5	<1	<1
Valve	2.2	5	<1	2
1.A Ring nut spanner	1.4	15	<1	<4
1.B Ring nut spanner	190	15	<1	<4

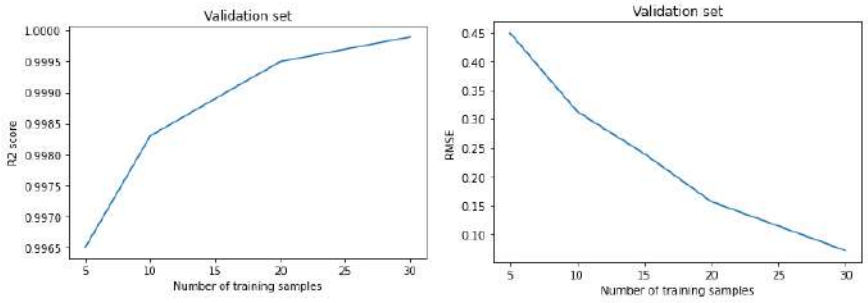
Table 7.5: Computation time to complete workflow for the first defined task for different parts.

7.2 TASK 2: PREDICT WARPAGE FOR VARYING PART POSITION FOR FDM PROCESS

Two new use cases are evaluated for warpage optimization in FDM process varying position parameters. Both a thermoelastic and a thermoviscoelastic material are used.

Database - Digimat-AM	
Simulation type	Warpage - high fidelity
Mesh size	3.5
Voxels	38473
CPUs job generation	1
CPUs warpage/thermal	4
ROM model	
Nodes on STL	41062
Points on a single Y vector	123186

Table 7.6: Curved duct's case. Variation of position parameters for FDM.



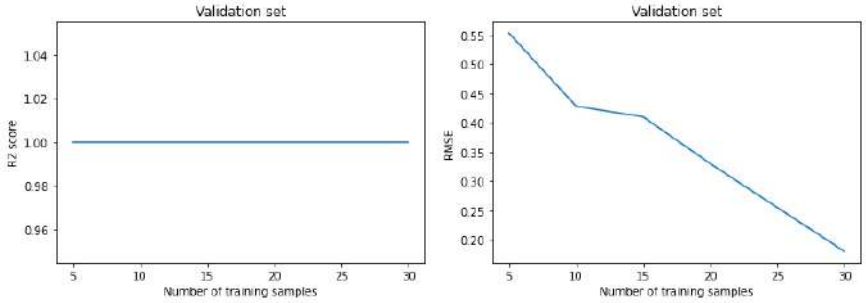
(a) (b)

Figure 7.5: Evolution of quality indicators a) R^2 and b) *RMSE* with the number of training sets for the validation data for the curved duct's case by FDM process.

Database - Digimat-AM	
Material	**TVE**
Simulation type	Warpage - high fidelity
Mesh size	5
Voxels	21105
CPUs job generation	1
CPUs warpage/thermal	4
ROM model	
Nodes on STL	59791
Points on a single Y vector	179373

Table 7.7: Duct's case. Variation of position parameters for FDM.

Parts with thermoelastic and thermoviscoelastic materials can be predicted equally good. Neural networks and Bayesian networks typically give lower R2 score than ODYSSEE models and take similar or less time to define the models.



(a) (b)

Figure 7.6: Evolution of quality indicators a) R^2 and b) *RMSE* with the number of training sets for the validation data for the duct's case by FDM process.

Time to complete workflow - Task 2		
	Simulation [min/sim]	**Build and train model [min] ODYSSEE**
Curved duct	15	6
Duct	13	6

Table 7.8: Computation time to complete workflow for the second defined task for different parts.

7.3 TASK 3: PREDICT TEMPERATURE EVOLUTION FOR SELECTED NODES FOR SLS PROCESS TO ESTIMATE COOLDOWN TIME

A new use case is evaluated for thermal prediction in SLS process varying process parameters.

Database - Digimat-AM	
Simulation type	Thermal
Mesh size	5.5
Voxels	92455
CPUs job generation	4
CPUs warpage/thermal	4
ROM model	
Nodes on STL	4703
Points on a single Y vector	10000

Table 7.9: Lever bracket use case for thermal analysis. Variation of process parameters.

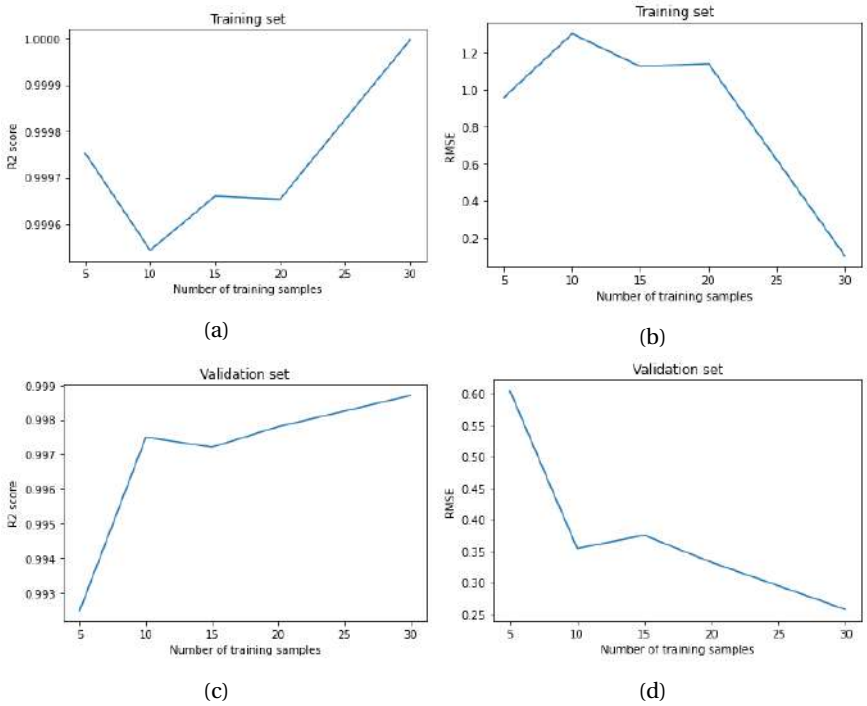


Figure 7.7: Evolution of quality indicators with the number of training sets. a) R^2 and b) *RMSE* on the training set and c) R^2 and d) *RMSE* on the validation set for the lever bracket use case on thermal analysis.

A 99% R^2 fit can be obtained using only 5 points for training. Adding more data improves the R^2 score for the validation set, whereas the training set improves from above 25 simulations as training data. Predictions are good for cooldown region [0-180]°C but they are not good in peak temperature region [180-250]°C due to the spike in temperature.

Time to complete workflow - Task 3

	Simulation [min/sim]	**Build and train model [min] ODYSSEE**
Lever bracket	16	1

Table 7.10: Computation time to complete workflow for the third defined task.

REFERENCES

[1] Sushant Negi and Rajesh Kumar Sharma. Study on shrinkage behaviour of laser sintered pa 3200gf specimens using rsm and ann. *Rapid Prototyping Journal*, 22:645–659, 6 2016.

[2] Kaufui V. Wong and Aldo Hernandez. A review of additive manufacturing. *ISRN Mechanical Engineering*, 2012:1–10, 8 2012.

[3] Kambiz Kayvantash and Mike Lee. Odyssee (cae a-eye): An engineering solution for exploiting data-based techniques (ml/ai/rom) for improved and efficient engineering.

[4] Muhammad Arif Mahmood, Anita Ioana Visan, Carmen Ristoscu, and Ion N. Mihailescu. Artificial neural network algorithms for 3d printing, 1 2021.

[5] Hexagon. ODYSSEE CAE Lunar User Guide, 2022.

[6] Daniel R. Eyers and Andrew T. Potter. Industrial additive manufacturing: A manufacturing systems perspective. *Computers in Industry*, 92-93:208–218, 11 2017.

[7] Isaac Simon, Merritt Gurley, and Samantha Van Roosenbeek. Market changes and technological shifts in the 3d printing market.

[8] Knowledge sourcing intelligence. Additive manufacturing market size: forecasts from 2022 to 2027, 6 2022.

[9] A. Noriega, D. Blanco, B. J. Alvarez, and A. Garcia. Dimensional accuracy improvement of fdm square cross-section parts using artificial neural networks and an optimization algorithm. *International Journal of Advanced Manufacturing Technology*, 69:2301–2313, 12 2013.

[10] Lingbin Meng, Brandon McWilliams, William Jarosinski, Hye Yeong Park, Yeon Gil Jung, Jehyun Lee, and Jing Zhang. Machine learning in additive manufacturing: A review, 6 2020.

[11] C. Wang, X. P. Tan, S. B. Tor, and C. S. Lim. Machine learning in additive manufacturing: State-of-the-art and perspectives, 12 2020.

[12] Jian Qin, Fu Hu, Ying Liu, Paul Witherell, Charlie C.L. Wang, David W. Rosen, Timothy W. Simpson, Yan Lu, and Qian Tang. Research and application of machine learning for additive manufacturing, 4 2022.

[13] Xinbo Qi, Guofeng Chen, Yong Li, Xuan Cheng, and Changpeng Li. Applying neural-network-based machine learning to additive manufacturing: Current applications, challenges, and future perspectives, 8 2019.

[14] Arnd Koeppe, Carlos Alberto Hernandez Padilla, Maximilian Voshage, Johannes Henrich Schleifenbaum, and Bernd Markert. Efficient numerical modeling of 3d-printed lattice-cell structures using neural networks. *Manufacturing Letters*, 15:147–150, 1 2018.

[15] F. Hajializadeh and A. Ince. Integration of artificial neural network with finite element analysis for residual stress prediction of direct metal deposition process. *Materials Today Communications*, 27, 6 2021.

[16] Moncef Salmi. Data science and artificial intelligence applied in the field of materials & icme.

[17] Standard terminology for additive manufacturing technologies: F2792 12a.

[18] Mojtaba Mozaffar, Ablodghani Ebrahimi, and Jian Cao. Toolpath design for additive manufacturing using deep reinforcement learning. 9 2020.

[19] John C. Steuben, Athanasios P. Iliopoulos, and John G. Michopoulos. Implicit slicing for functionally tailored additive manufacturing. *CAD Computer Aided Design*, 77:107–119, 8 2016.

[20] Rong Ji Wang, Lingling Wang, Lihua Zhao, and Zijian Liu. Influence of process parameters on part shrinkage in sls. *International Journal of Advanced Manufacturing Technology*, 33:498–504, 6 2007.

[21] N. Raghunath and Pulak M. Pandey. Improving accuracy through shrinkage modelling by using taguchi method in selective laser sintering. *International Journal of Machine Tools and Manufacture*, 47:985–995, 5 2007.

[22] Wang Rong-Ji, Li Xin-Hua, Wu Qing-Ding, and Wang Lingling. Optimizing process parameters for selective laser sintering based on neural network and genetic algorithm. *International Journal of Advanced Manufacturing Technology*, 42:1035–1042, 6 2009.

[23] K. Senthilkumaran, Pulak M. Pandey, and P. V.M. Rao. Influence of building strategies on the accuracy of parts in selective laser sintering. *Materials and Design*, 30:2946–2954, 9 2009.

[24] Ali Kiani, Saeed Khazaee, Mohsen Badrossamay, Ehsan Foroozmehr, and Mehdi Karevan. An investigation into thermal history and its correlation with mechanical properties of pa12 parts produced by selective laser sintering process. *Journal of Materials Engineering and Performance*, 29:832–840, 2 2020.

[25] Erik G Learned-Miller. Introduction to supervised learning, 2014.

[26] Steven L. Brunton and J. Nathan Kutz. *Data Driven Science Engineering: Machine Learning, Dynamical Systems, and Control*.

[27] C. Caruso. Interpolation methods comparison, 1998.

[28] Shmuel Rippa. An algorithm for selecting a good value for the parameter c in radial basis function interpolation, 1999.

[29] Kaifeng Gao, Gang Mei, Salvatore Cuomo, Francesco Piccialli, and Nengxiong Xu. Arbf: adaptive radial basis function interpolation algorithm for irregularly scattered point sets. *Soft Computing*, 24:17693–17704, 12 2020.

[30] Stephen Allwright. What is a good r-squared value?, 2022.

[31] Aman Kharwal. R2 score in machine learning, 6 2021.

[32] Stephen Allwright. What is a good rmse value? simply explained, 7 2022.

[33] Charles Elkan. Evaluating classifiers 1. test sets and validation sets, 2012.