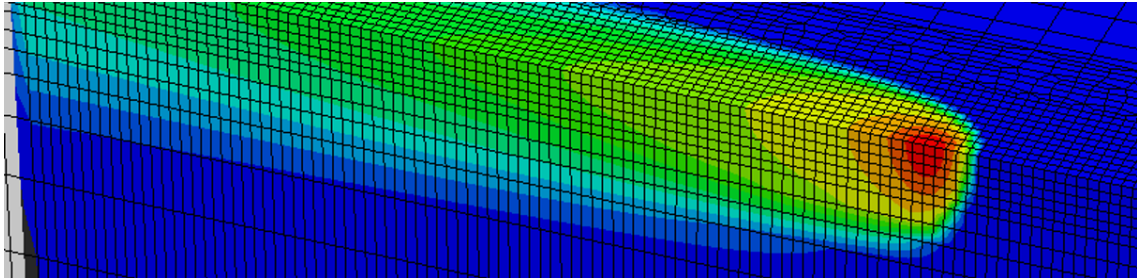Master Thesis

Surrogate modelling for 3D multiscale thermal simulations of powder-bed additive manufacturing

Natalia Smatsi

Advisors: Prof. Dr. E. Mazza, Dr. E. Hosseini, P. Gh. Ghanbari

Experimental Continuum Mechanics Lab, EMPA
Department of Mechanical Engineering, ETH Zürich

31.12.2020

Abstract

In this master thesis, the potential of using a surrogate model to reduce the computational cost of 3D thermal simulations for selective laser melting is explored. This work is part of a general strategy developed in *Empa* of using a multiscale approach for solving finite element (FE) simulations more efficiently. This is done by using a coarse mesh solution in the areas far away from the laser and a fine mesh solution for the areas close to the melt pool where finer spatial and temporal detail is needed. The surrogate model aims to substitute the fine mesh solution under the assumption that close to the melt pool, the temperature field is influenced mainly by the laser, the initial temperature profile, and the local geometry/topology. For this reason, a FE model representing a local geometry under the influence of a single laser pass is set up and validated. The FE thermal model was developed for Hastelloy-X and was tuned to capture the phase transition from powder to liquid and then solid by validating the melt pool size with experimentally measured values. Regarding the surrogate model, principal component analysis was used to decrease the dimension of the temperature fields and polynomial chaos expansion to model the dependence between the input (initial temperature, geometry) and output data (temperature profile under the influence of a single laser pass). The surrogate model temperature profile was compared to the corresponding finite element simulations for a range of initial temperatures, relevant to preheating temperatures used in selective laser melting, and a range of different geometries. The latter involved distributing pockets or blocks of powder in the section beneath the laser pass and using as input the position and size of these blocks to predict the temperature field.

Acknowledgments

Firstly, I would like to thank Prof. Edoardo Mazza for giving me the opportunity to write this thesis at the Experimental Continuum Mechanics Lab and allowing me to conclude my Master at ETH Zurich with this work.

I am very grateful to Dr. Ehsan Hosseini and Pooriya Ghanbari for their supervision and help all through these months. They not only shared their expertise with me regarding the SLM process, Abaqus thermal simulations and surrogate modelling, but also supported me through the challenges that were often presented. In a generally challenging year, I would be forever grateful for our personal interactions. Special thanks to Filippo Kusch for sharing his Latex template and our useful discussions regarding his additive manufacturing experimental insights.
Finally, to Athina, Lefteris, Sophia, Iason, and my parents, I could not find a better way to express my feelings rather than with a simple but heartfelt $\varepsilon\upsilon\chi\alpha\rho\iota\sigma\tau\acute{\omega}$.

<div align="right">Natalia Smatsi, Zürich, December 2020.</div>

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

_____

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):   **First name(s):**

With my signature I confirm that
  − I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
  − I have documented all methods, data and processes truthfully.
  − I have not manipulated any data.
  − I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date   **Signature(s)**

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.

Contents

List of Figures

List of Tables

1 Introduction

Selective laser melting (SLM) is a powder bed fusion (PBF) metal additive manufacturing (MAM) technique aimed at small volume production of highly customized metal components built in a layer-by-layer fashion [1]. Specifically, a layer of fine metal powder is melted and then selectively solidified by a moving laser source under an inert gas (1) atmosphere. The laser (2) is focused and directed to the correct location by a scanner system (3) following the information of a sliced CAD model. A heated bed (4) supports the solidified layer and moves as consecutive powder layers are deposited by a recoater blade or roller (5), melted and solidified until the final geometry is achieved. Any excess powder after the recoating is gathered in a storage container (6). A schematic of the process is shown in Figure 1.
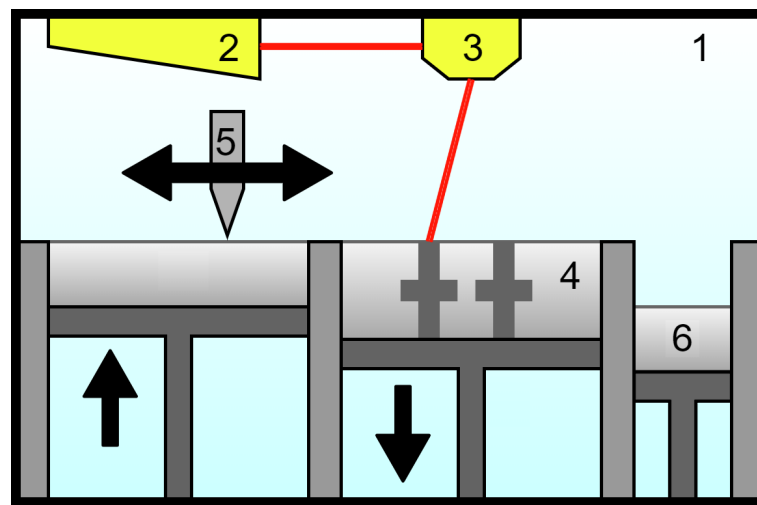
Figure 1: SLM machine schematic.

Due to the additive nature of the process, SLM is characterized by freedom of design, short lead time between concept phase and production, near-net-shape fabrication, and limited material waste, while it is suitable for processing of various metals, among which the most common are iron, titanium, aluminum, and nickel-based alloys [2]. This makes SLM very attractive across a range of industries, including aerospace, automotive, and biomedical, used for rapid prototyping, manufacturing of complex shapes and tooling. A recent example is the use of SLM parts in projects involved in the MARS 2020 mission at the Jet Propulsion Laboratory [3], such as the concept of the Mars helicopter rotor bracket shown in Figure 2.

The main issues associated with SLM are the often poor-as-built mechanical properties, mainly concerning fatigue strength, which require the complex control and optimization of multiple process parameters [4]. These challenges are closely linked to the high heating and cooling rates in the range of 10^3 to 10^8 K s^{-1} [5] which can cause the development of high residual stresses, deformations, and possible fractures. Furthermore, the temperature transitions can greatly affect the microstructure and, therefore, the material-specific mechanical and thermal properties of the final parts. To achieve highly precise, pore-free parts with minimum residual stresses, trial-and-error experimental approaches are often needed. These usually consist of iterating and repeating multiple prints until the required dimensional accuracy, material characteristics, and mechanical properties are achieved [6]. On this front, numerical simulations could offer a distinct advantage
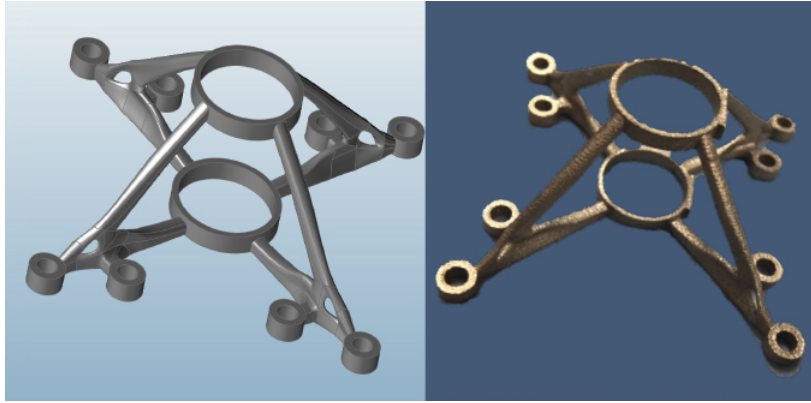
Figure 2: Mars helicopter rotor bracket. Image source: NASA-JPL [3].

for replacing in part these costly manufacturing iterations by providing an insight onto the effect of process parameters, as well as the physical phenomena that take place during SLM printing.

Therefore, since the temperature profiles are also used as input for mechanical and microstructural simulations, the starting point for reliable SLM simulations is an accurate and computationally efficient thermal model. The majority of SLM thermal simulations employ Finite Element Modelling (FEM), which discretizes the model domain with mesh elements with representative physical properties [2]. The laser is then modelled as a volumetric heat source with a Gaussian intensity distribution with the shape chosen according to the selected heat source model, i.e., spherical, cylindrical, ellipsoid, etc, [7]. Concerning the computational cost of these simulations, the challenge is directly connected to the laser modelling, which requires both a refined space and time discretization due to its highly concentrated and fast-moving energy input. The simulation is further complicated by the temperature dependence of the thermal material properties and the influence of a variety physical phenomena spanning from the microscale to the macroscale (i.e. powder melting and sintering, meltpool evolution and cooling rate of the solidified metal due to convection, conduction and radiation) [8].

To tackle these issues, various computational methods for simplifying the process have been proposed. Specifically for SLM, the powder layer and its interaction with the laser is often facilitated by modelling the former as a continuum medium with adjusted material properties to make up for neglecting the detailed modelling of fluid-related phenomena in the vicinity of the melt pool [8]. Furthermore, many works [9, 10, 11] employ a re-meshing strategy, where a fine mesh solution following the laser pass is combined with a coarse mesh solution far away from the melt pool where the temperature transitions are more gradual. Ghanbari et al.[10] have developed an adaptive local-global approach, where the global model solved the thermal problem using a coarse mesh. The temperature profile was then "corrected" in the vicinity of the melt pool by multiple local models with a fine mesh which capture the physics and fast-kinetics of the laser. This method was applied successfully for 2D simulations of laser passes, reducing the computational cost of simulations by an order of magnitude with minimal loss in accuracy [10].

To further decrease the computational cost of these local models, a surrogate modelling approach was proposed in the master thesis of Keller Fabian [12]. The goal in [12] was to construct a cheap computational model that emulates the behavior of the expensive FEM simulations. The FEM simulations were treated as a black-box. Namely, the surrogate model was trained in a data-driven way, using a set of input and output data from FEM without knowledge about the

internal model parameters. The surrogate model was trained to replace local models identical in geometry that differ in initial and boundary conditions. The initial and boundary temperature conditions were used as inputs, and the required output was the temperature distribution of the whole local model. The surrogate model used a combination of principal component analysis (PCA) and polynomial chaos expansion (PCE) to decrease the dimensionality of the temperature distributions and map the dependency between the input and output distribution, respectively. The framework was developed for 2D geometries predicting the resultant temperature field successfully. However, it was noted that the temperature profiles, apart from the initial and boundary temperatures, were also influenced by the bulk geometry of the local model, e.g., the presence of holes, overhangs, and in general volumes of unsolidified powder in the bulk of the part. It was, therefore, suggested that this geometrical information should be incorporated as input to the surrogate model in future works to further increase its accuracy and applicability.

This thesis explores the potential of the surrogate modelling approach for thermal simulations in 3D geometries. On these grounds, a 3D finite element model representing the local simulations in the vicinity of the melt pool was set up and validated. The thermal model was developed for the printing of Hastelloy-X parts employing experimentally relevant process and material parameters. The transition from 2D to 3D led to an inherent increase in computational cost requiring a careful mesh optimization study. Regarding the surrogate modelling, the goal was to emulate the temperature field caused by a single laser pass under the influence of different initial temperatures and bulk geometries by using a combination of PCA and PCE similar to [12].

The remainder of the thesis is structured as follows; in Section 2 an overview of SLM modelling techniques as well as a description of the PCA and PCE methods used in the surrogate model in this work are presented. Then, in Section 3, a description of the set-up of the finite element thermal model and the surrogate model is included. Finally, in Sections 4 the obtained results are presented and discussed in detail, reaching to the conclusions and possible research recommendations for future works in Section 5.

2 Theory

2.1 Modelling of SLM

The SLM process modeling can be divided into three ranges as presented by Markl et al. [8]: microscale, mesoscale, and macroscale.
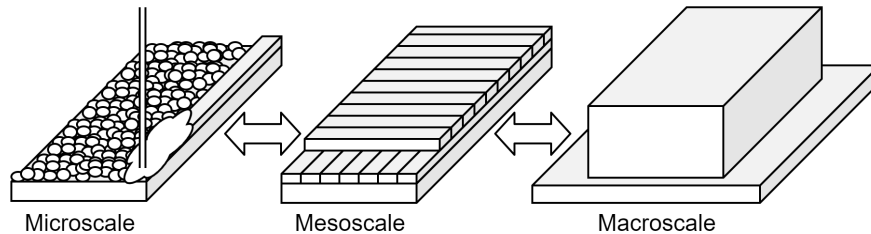
Figure 3: The three main scales involved in the thermo-mechanical modeling and simulation of the SLM process.

The microscale modelling captures phenomena happening in the scale of the powder particles and the melt pool of the metal (nm-µm). This usually involves the explicit modelling of the powder particles, their interactions with the laser, the evolution of the melt pool, and the phases of the metal alloys under-cooling and reheating of the solidified parts [13]. The melt pool size, properties, and evolution are critical factors for predicting the microstructure and porosity, along with surface roughness and overall melting process stability. These, in turn, affect the mechanical properties and, therefore, the resultant stresses and macroscopic deformations of the final parts. Depending on the focus of the microstructure simulations, various computational methods have been employed in the literature: cellular automata [14] and phase-field modeling for phase and microstructure evolution [15] and, Monte Carlo for grain size/morphology simulations [16].

The mesoscale modelling employs continuous models to study the thermal and mechanical behavior of regions with size ranging from a single laser pass to a hatch or up to a full layer (10µm to 1mm) [2]. The powder layer is modelled as a continuous medium with adjusted properties tuned to be comparable with the high detailed representations. An example of this adjustment is the so-called effective thermal conductivity of the liquefied powder, which is artificially increased to consider the various convective heat transfer modes existing inside the melt pool [17]. Mesoscale simulations can be used to generate a surrogate model [12, 18] or an equivalent load model [19] to be used later in macroscale simulations.

Finally, the macroscale modelling refers to the simulations of the entire process, including the prediction of part-level stresses and deformations, as well as the thermal history of complete components. These simulations scale from mm up to m and are obtained with continuous models that stack up some layers, partially waiving the resolution on the effective thermal gradients and the microscale phenomena. In this scale, complete functional components are modeled, including the build-plate and the material supports [20]. Therefore, these simulations can be used to optimize the print settings by estimating the quality of the final parts without the need for trial-and-error iterations. Surrogate models or equivalent loads obtained from mesoscale simulations are two approximating approaches leveraged in macroscale part-level simulations and are presented in detail below.

2.1.1 Equivalent loads

The idea behind this approach is to substitute the loads (thermal and/or mechanical) applied on the component with a simplified set of equivalent loads that approximate the original case and reduce the number of steps and iterations needed. This set of equivalent loads can be created analytically or by running a set of high detail simulations. Specifically for SLM, this strategy can be applied to simplify the high detailed laser passes with an equivalent layer-based body heat flux. In the work of Li et al. [19, 21] this was implemented using a so-called temperature-thread multiscale approach described in the three modelling scales of SLM presented in the previous section, e.g. microscale, mesoscale, and macroscale. In the microscale laser scan model, the temperature field of the melt pool was calculated via a moving heat source over a single track. In the mesoscale layer hatch model, the temperature history obtained from the microscale model is extended to a deposited powder layer. In the macroscale part model, a complete part is built layer upon layer by applying the thermal load calculated from the mesoscale model. Figure 4 shows a simple schematic of a moving Gaussian heat source with high fidelity hatching and its equivalent body heat flux applied to a complete layer. The model was in good agreement with experimental results regarding a simple plate. However, problems could arise in more complex geometries requiring extended microscale modeling of features such as overhangs and cavities. Also, this model cannot capture the effect of interlayer scanning strategy, which can significantly affect the anisotropy of the mechanical properties of the printed part and, therefore, the development of residual stresses [22].
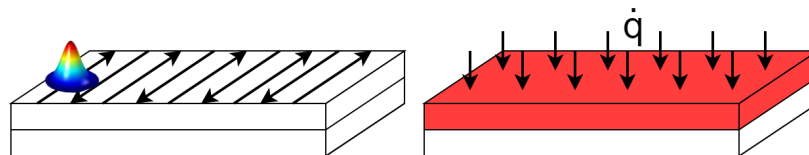
Figure 4: A high fidelity moving heat source and the equivalent load substituting the high-detail model. Adapted from [19].

2.1.2 Surrogate modelling

Surrogate modelling is a meta-modelling technique that aims to capture the dependency of a model output on its inputs [23]. They are used to replace computationally expensive models or experiments by treating them as black-boxes without knowledge of their internals. Surrogate models work in a data-driven way as they are based only on pairs of inputs and outputs produced by the model they aim to replace. Standard surrogate models include Kriging [24], Polynomial Chaos Expansion (PCE) [12], Support Vector Machines (SVM) [24] and Neural Networks (NNs) [25]. However, constructing an accurate surrogate model can be very computationally demanding in itself due to the need to generate the input-output pairs using the original expensive model. Another challenge is the handling and processing of the often high-dimensional input and output data.

   Both these challenges are very relevant to SLM thermal modelling. Specifically for the mesoscale FE approach for parts with dimensions in the order of a few centimeters, a high fidelity simulation of a laser pass requires an element size of 10μm (in the vicinity of the melt pool) and time increments of 10^{-6}s [10]. Therefore, the object of interest being a field quantity,

namely the temperature field over many timesteps and nodes, is inherently high dimensional. Additionally, the computational gain is directly related to the surrogate model depending on a sparse set of runs from the finite element model. In the work of Keller [12] this was successfully tackled with a combination of Principal Component Analysis (PCA) to reduce the dimensionality of the input and output temperature distributions and a Polynomial Chaos Expansion (PCE), as the surrogate model to map the dependency of the reduced input to the reduced output. A similar scheme was used in this work, and the two methods are presented in detail in the following sections.
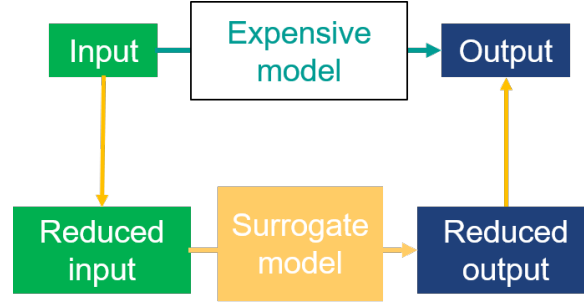
Figure 5: Surrogate models can replace computationally expensive models through training on a set of inputs and outputs produced by the latter.

2.2 Principal component analysis

The principal component analysis (PCA) method is a standard tool in modern data analysis that allows a non-parametric approach to the extraction of relevant information from confusing or noisy data sets [26]. It has also been used for computational reduction schemes, which were applied to the simulation of metal additive manufacturing [18]. In the following section, the method is explained, following the tutorial from Jonathon Shlens [26].

To be able to process the problem, the data is gathered in a single real matrix $\mathbf{X} \in \mathbb{R}^{m,n}$, where n is the number of observations in the system, and m the number of properties or measurements assigned to each observation $\vec{X}_i$.

$$\vec{X}_i = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad \in \mathbb{R}^m \tag{2.1}$$

The data matrix $\mathbf{X}$ containing all observations can then be created by joining the column vectors of the n observations in the system.

$$\mathbf{X} = [\vec{X}_1, ..., \vec{X}_n] \quad \in \mathbb{R}^{m,n} \tag{2.2}$$

By multiplying the original matrix $\mathbf{X}$ by a suitable matrix $\mathbf{P} \in \mathbb{R}^{m,m}$ a linear transformation onto the new space represented by the matrix $\hat{\mathbf{X}}$ is performed:

$$\hat{\mathbf{X}} = \mathbf{P}\mathbf{X} \tag{2.3}$$

The rows of the matrix $\mathbf{P}$, $\{\vec{P}_1, ..., \vec{P}_m\}$ contain the orthonormal basis of the transformed space onto which the original data is projected. The columns of the matrix $\hat{\mathbf{X}}$ contain the so called *component scores* which are the transformed measurements of each observation $0, ..., n$. Figure 6 shows how the original data, which is not aligned with the main axes and is spread, can be projected on a reduced space.
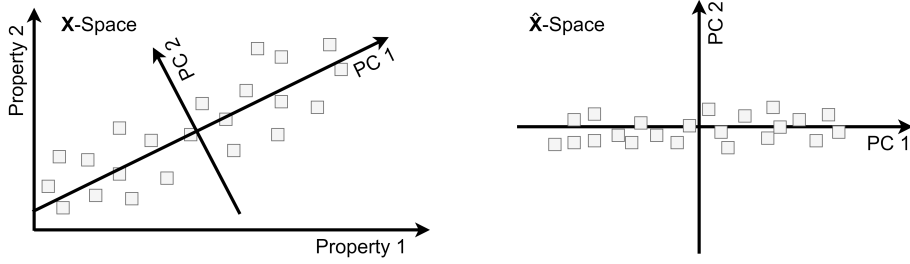
Figure 6: Original and PCA transformed data.

To find the suitable transformations, a search for the orthonormal directions, along which the variance is maximized, must be done. By calculating the associated variances, it is possible to rank the components according to their importance. To choose the appropriate basis, the covariance matrix $\mathbf{C_X}$ needs to be considered, where the diagonal terms are the values of the variance for the corresponding feature:

$$\mathbf{C_X} = \frac{1}{n}\mathbf{XX}^T \in \mathbb{R}^{m,m} \tag{2.4}$$

Shlens[26] presents two approaches for solving this problem: using the eigenvector decomposition or using the singular value decomposition (SVD). SVD is the default technique employed in the *pca* algorithm of the *MATLAB* library (which is implemented in this work) and is, therefore, presented below.

Considering a matrix $\mathbf{Y}$ the relation with its eigenvectors $\hat{v}_i$ and eigenvalues λ_i with $i \in 0, ..., r$ is:

$$(\mathbf{Y}^T\mathbf{X})\vec{v}_i = \lambda_i\vec{v}_i \tag{2.5}$$

and define the singular values σ_i and the vectors $\vec{u}_i$

$$\sigma_i = \sqrt{\lambda_i} \tag{2.6}$$

$$\vec{u}_i = \frac{1}{\sigma_i}\mathbf{Y}\vec{v}_i \tag{2.7}$$

$$\vec{u}_i \cdot \vec{u}_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \tag{2.8}$$

$$\|\mathbf{Y}\vec{v}_i\| = \sigma_i \tag{2.9}$$

Which can be rewritten as

$$\mathbf{Y}\vec{v}_i = \sigma_i\vec{u}_i \tag{2.10}$$

By expressing the problem in matrix form (with $\mathbf{V}$ containing the eigenvectors and $\mathbf{U}$ the vectors $\vec{u}$, and the matrix $\mathbf{\Sigma}$ the singular values in it diagonal) and rewriting the equations the following is obtained:

$$\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{2.11}$$

$$\mathbf{V}^T\mathbf{Y}^T = \mathbf{U}^T\mathbf{\Sigma} \tag{2.12}$$

$$\mathbf{V}^T\mathbf{Y}^T = \mathbf{Z} \tag{2.13}$$

By choosing appropriately $\mathbf{Y}$ and calculating its SVD, the found $\mathbf{V}$ matrix contains the eigenvectors of $\mathbf{C_X}$, thus the columns of $\mathbf{V}$ are the principal components of $\mathbf{X}$.

$$\mathbf{Y} = \frac{1}{\sqrt{n}}\mathbf{X}^T \tag{2.14}$$

$$\mathbf{Y}^T\mathbf{Y} = \mathbf{C_X} \tag{2.15}$$

2.3 Polynomial chaos expansion

Polynomial Chaos Expansion (PCE) is a spectral decomposition similar to the Fourier series expansion [23]. It is often used to construct a cheap computational model that approximates an original, computationally expensive one. It treats the latter as a black box and uses a set of input-output data of the original model to emulate its behavior. PCE can be seen as a subset of polynomial approximation methods as it is based on polynomial functions forming a basis representation on which the model is projected.

The following sections are based on the documentation of the UQLab MATLAB library developed by Stefano Marelli and Bruno Sudret [23]. Consider we want to approximate a finite variance model with result $Y = \mathcal{M}(\mathbf{X})$ the polynomial chaos expansion can be written as

$$Y = \mathcal{M}(\mathbf{X}) = \sum_{\alpha \in \mathbb{N}^M} y_\alpha \Psi_\alpha(\mathbf{X}) \tag{2.16}$$

The functions Ψ_α form an infinite dimensional polynomial basis and $y_\alpha \in \mathbb{R}$ are their coefficients for any multi-index $\alpha \in \mathbb{N}^M$. The argument $\mathbf{X} \in \mathbb{R}^M$ of the computational model follows a multivariate probability distribution characterized by the density $f_\mathbf{X}$. Then the basis functions Ψ_α have to be orthonormal with respect to $f_\mathbf{X}$ and the model must have a finite variance

$$\int_{\mathbb{R}^M} \mathcal{M}(x) f_\mathbf{X}(x) dx < \infty \tag{2.17}$$

For some classical probability distributions, the family of univariate polynomials to which they are orthonormal have specific names i.e., the Legendre polynomials for the uniform distribution and the Hermite polynomials for the Gaussian distribution. For the case of multivariate distributions, the orthonormal multivariate polynomials can then be constructed as the tensor product of their univariate counterparts.

2.3.1 Truncation schemes

For practical applications, the infinite sum of polynomials must be truncated to get a tractable expression. The finite set of multi-indices considered in the truncated expansion is written as

$\mathcal{A} \subset \mathbb{N}^M$. Then the expansion becomes

$$Y \approx \mathcal{M}^{PC}(\mathbf{X}) = \sum_{\alpha \in \mathcal{A}} y_\alpha \Psi_\alpha(\mathbf{X}) \tag{2.18}$$

For the truncation, we can consider polynomials up to a certain total degree p. The total degree of a multivariate polynomial is the sum of the degrees of the underlying univariate polynomials.

$$\mathcal{A}^{M,p} = \{\alpha \in \mathbb{N}^M : |\alpha| \le p\} \tag{2.19}$$

The number of elements of this set is expressed by the cardinality of the corresponding basis.

$$\text{card } \mathcal{A}^{M,p} = \binom{M+p}{p} \tag{2.20}$$

Two additional non-mutually exclusive truncation schemes available in the UQLab library and also used in this work are the maximum interaction and the hyperbolic truncation.

The maximum interaction truncation scheme aims at reducing the interaction between different components of the random input by limiting the non-zero entries of the elements of the set $\mathcal{A}^{p,M}$ defined in (2.19). For a maximum interaction coefficient of r this truncation set is written as

$$\mathcal{A}^{M,p,r} = \{\alpha \in \mathcal{A}^{M,p} : ||\alpha||_0 \le r\} \tag{2.21}$$

with the zero-norm giving the number of non-zero elements $||\alpha||_0 = \sum_{i=1}^{M} \mathbf{1}_{a_i > 0}$. With this basis, it is possible to reduce the number of elements while retaining polynomials of relatively high degrees. This is especially effective for high dimensional input vectors.

Regarding the hyperbolic truncation scheme, the modification of equation 2.19 is proposed by using the q-norm.

$$||\alpha||_q = \left(\sum_{i=1}^{M} \alpha_i^q\right)^{1/q} \tag{2.22}$$

$$\mathcal{A}^{M,p,q} = \{\alpha \in \mathcal{A}^{M,p} : ||\alpha||_q \le p\} \tag{2.23}$$

For $q=1$, the standard truncation scheme is recovered. For $q<1$, the hyperbolic truncation includes all the high-degree terms in each single variable but discourages equivalently high order interaction terms.

2.3.2 Computation of the coefficients

The next step after calculating the truncated polynomial basis is the calculation of the coefficients $\{y_\alpha\}_{\alpha \in \mathcal{A}}$. In the UQLab library, only non-intrusive methods are available, e.g., the projection method and the least-squares method. The non-intrusiveness stems from the fact that these methods are based on the post-processing of carefully selected input/output pairs provided by the original expensive model. This set of model evaluation values is also referred to as the experimental design.

In the projection method, we take the expectation of the polynomial chaos expansion (Eq. (2.16)) multiplied by the polynomial Ψ_β (equivalent to integrating over the domain of $\mathbf{X}$). Using

the orthonormality of the polynomial basis, we get

$$\mathbb{E}[\Psi_\beta(\mathbf{X})\mathcal{M}(\mathbf{X})] = \int_{\mathcal{D}_X} \Psi_\beta(x)\mathcal{M}(x)f_{\mathbf{X}}(x)dx, \tag{2.24}$$

$$= \int_{\mathcal{D}_X} \Psi_\beta(X) \left(\sum_{\alpha \in \mathcal{A}} y_\alpha \Psi_\alpha(x)\right) f_{\mathbf{X}}(x)dx, \tag{2.25}$$

$$= \sum_{\alpha \in \mathcal{A}} y_\alpha \int_{\mathcal{D}_X} \Psi_\alpha(x)\Psi_\beta(x)f_{\mathbf{X}}(x)dx, \tag{2.26}$$

$$= y_\beta, \tag{2.27}$$

For calculating the expectation, numerical integration schemes can be used, such as the Gaussian quadrature. For the case of multivariate integrals, the tensor-product of univariate integration is used. Therefore the number of integration nodes increases rapidly with the number of input variables. To tackle that, sparse quadrature rules can be applied, which use a sparse grid to reduce the number of quadrature points.

In the least-squares regression method, we rewrite the computational model as a sum of the truncated expansion and a residual r:

$$\mathcal{M}(\mathbf{X}) = \mathcal{M}^{PC}(\mathbf{X}) + r. \tag{2.28}$$

Then the goal is to find coefficients y_α that minimize the mean value of the squared residuals. This can be written as

$$\hat{\mathbf{y}} = \underset{\mathbf{y}}{argmin} \, \mathbb{E}[r^2] = \underset{\mathbf{y}}{argmin} \, \mathbb{E}\left[\left(\mathcal{M}(\mathbf{X}) - \sum_{\alpha \in \mathcal{A}} y_\alpha \Psi_\alpha(\mathbf{X})\right)^2\right]. \tag{2.29}$$

The minimization problem can then be solved using the Ordinary Least Squares (OLS) methods using all the points of the experimental design. The main advantage of the least-square minimization method, as opposed to the Gaussian quadrature, lies in the fact that an arbitrary number of points can be used to calculate the coefficients, as long as they are a representative sample of the input vector X reducing, therefore, the computational cost.

2.3.3 Adaptive PCE

Using methods for assessing the error of a PCE, it is possible to construct adaptive methods for building the final PCE. Specifically, in the UQLab, a leave-one-out cross-validation scheme is used to calculate the generalization error of the fitting. The generalization error is then used in both adaptive algorithms implemented in the UQLab library; the basis-adaptive PCE and the sparse PCE. In the basis adaptive PCE, the starting point is a small candidate polynomial basis, and new elements are added (e.g., by increasing the maximum polynomial degree in the truncation scheme). The best PCE in terms of generalization error is then selected. For the sparse PCE, the starting polynomial basis includes a subset of the most relevant polynomials, while the coefficients of all the others are set to 0. The adaptive PCE uses both the truncation methods presented above and employs a penalization scheme, similar to Lasso or Ridge regression, where a regularization term is added in the least-squares minimization that favors solutions with small coefficients.

3 Model development

3.1 Finite element modelling

The thermal simulation was carried out using the Additive Manufacturing Modeler (AM-Modeler) plug-in [27] and the Abaqus/Standard 2020 finite element solver [28]. To reduce the computational cost of the thermal simulations, a mesh optimization study was performed. The optimized mesh was designed using Altair HyperMesh [29] and then imported into Abaqus. The process parameters were chosen to match experimentally printed Hastelloy X samples. These were the laser power P=200W, the scanning speed V=900mm/s and the powder layer thickness l_{powder}=30µm.

3.1.1 Model geometry

(a)

(b)

(c)

(d)
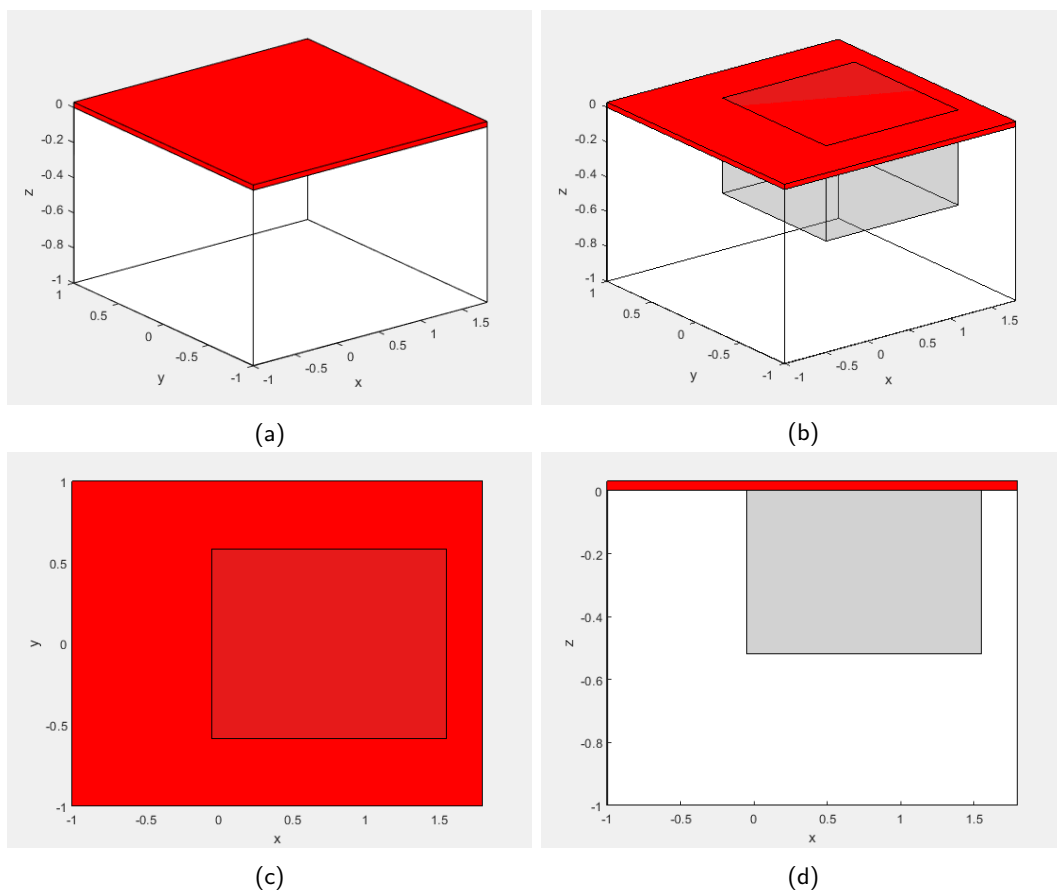
Figure 7: Schematic of the model geometry in mm.

The model was set up as a rectangular box with dimensions 2.8mm x 2mm x 1.03mm. A schematic of the box in mm is shown in Figure 7a and b where the top 0.03mm along the z-axis illustrated with red represent the deposited powder layer before the laser pass and the white cuboid underneath represent the already solidified layers. For the purpose of this work, this is going to be referred to as substrate being beneath the powder layer that will be selectively solidified by the laser pass. However, the temperature field was not recorded for the whole volume of the cuboid. A smaller rectangular box was defined with dimensions 1.6mm x 0.5mm x 0.53mm, illustrated with gray. A top and side view of the schematic is shown in Figure 7c and d. This smaller rectangular box will be referred to in the rest of this work as the inner local model. The larger cuboid was used to precondition the model. The laser pass is applied in the

middle of the powder layer along the positive direction of the x-axis with the center of the laser beam following the trajectory with coordinates $x=[-1,2.8]$, $y=0$, $z=0.03$ and a constant scanning speed of $V=900$mm/s.
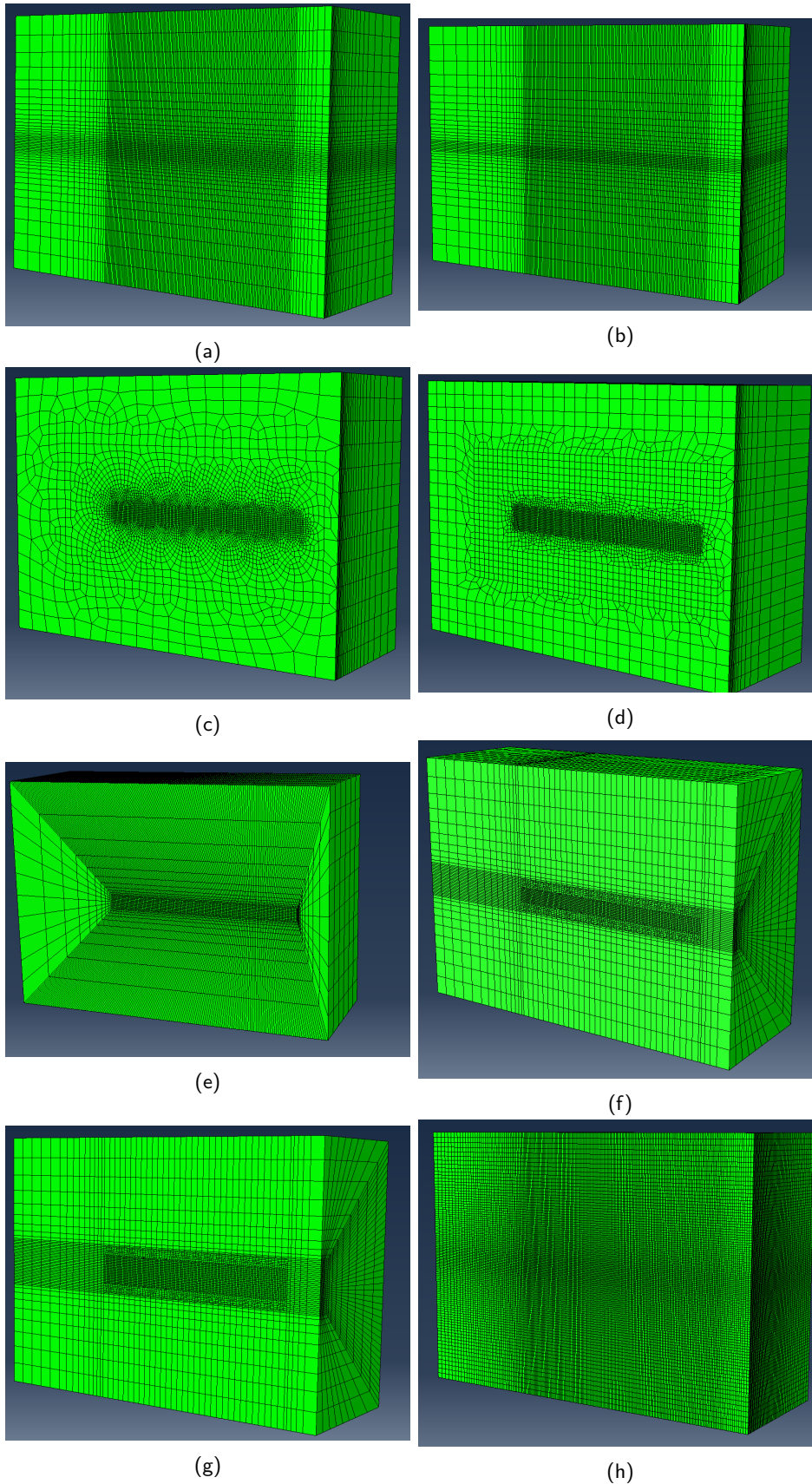
Figure 8: Mesh strategies.

3.1.2 Mesh optimization

The next step was to construct a mesh for the FE model. The element size proposed in the literature for the area close to the laser is 10µm in order to capture the melt pool dynamics accurately [10]. This is however, not necessary for the area far away from the laser where larger elements can be used. There was, therefore, the need to define the optimal mesh refinement scheme so that the temperature field is not affected while the number of elements and nodes is computationally reasonable. The various meshing strategies were compared to a model with a uniform mesh size of 10µm, which will be referred to as reference mesh. This fine mesh with 10µm element size resulted in a local model with 1.32 million elements and 1.36 million nodes.

All meshes constructed in this work consisted of hexagonal elements (DC3D8) to make the model compatible with structural simulations that will be performed at a later stage [2]. The various strategies employing structured and unstructured elements (and their combinations) are illustrated in Figure 30. Figures 30a and b show structured meshes which become gradually less dense for elements away from the region of interest (laser track). Similarly, Figures 30c and d show unstructured meshes. Figures 30e to g show a mesh strategy we named "extruded" mesh in this work, as this mesh was constructed by starting from the finer mesh in the vicinity of the laser track and "extruding" it to the sides of the rectangular. The extruded mesh in Figure 30e consisted of only structured elements, whereas Figure 30f and g had a combination of structured and unstructured elements. Finally, Figure 30h illustrates the reference mesh with 10µm in the whole geometry.
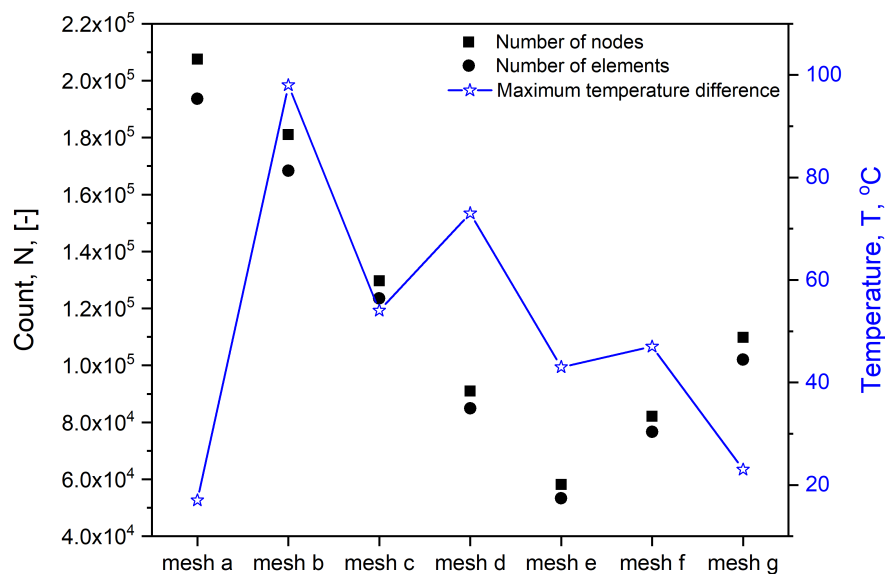
Figure 9: Comparison of the various mesh strategies with respect to the number of elements and nodes and the maximum temperature difference compared to the model with the reference mesh.

A preliminary run with a single laser track and simplified material description was implemented to evaluate the optimal mesh strategy. To speed up the process no phase transition mechanism was implemented and the whole domain was assumed to be solid. The goal was to compare the temperature fields of the reduced meshes (Figure 30a to g) to the reference mesh (Figure

30h). In Figure 9 the mesh evaluation is presented: the number of elements and nodes of each optimized mesh solution is shown on the left y-axis, and the maximum temperature difference between the optimized meshes and the reference mesh is shown on the right y-axis. In the x-axis the labels of the various meshes are shown consistently with Figure 30.

The mesh solution selected in this work is mesh g with 102000 elements and 109800 nodes. It consisted of 10µm elements close to the laser track up to 150µm elements on the outer boundaries of the cuboid. As reported earlier, the mesh was constructed using Altair Hypermesh due to the limited flexibility of the Abaqus mesh toolbox. The meshed model is presented in Figure 10, where the inner local model is illustrated with a lighter shade of blue. The inner local model consisted of 71280 elements and 77363 nodes.
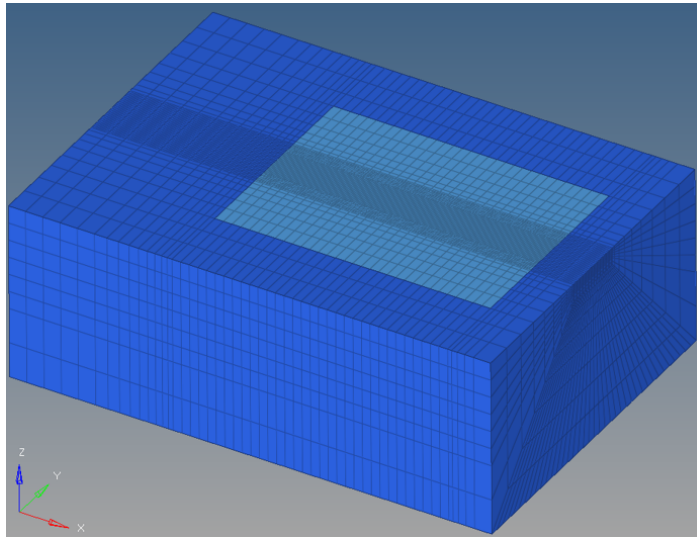
Figure 10: Optimized mesh.

3.1.3 Process parameters, heat source, and material modelling

The process parameters, which were common for the simulation and the experimental setup of SLM printed Hastelloy X parts, are shown in Table 1.

Table 1: Process parameters

Property	Symbol	Value	Unit
Laser power	P	200	W
Laser beam diameter	d_{laser}	55	µm
Layer thickness	l_{powder}	30	µm
Scan speed	V	900	$\mathrm{mm\,s^{-1}}$

Generally, the governing equation of a 3D heat transfer model[7] is

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k_x \frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k_y \frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k_z \frac{\partial T}{\partial z}\right) + Q\left(x, y, z, t\right) \qquad (3.1)$$

where ρ, c, k_x, k_y, k_z, T and Q are the material density [kg/m^3], the specific heat [J/kg°C], the thermal conductivity [W/m°C] in x, y, z directions, the temperature [°C] and the internal heat generation [W/m^3] respectively.
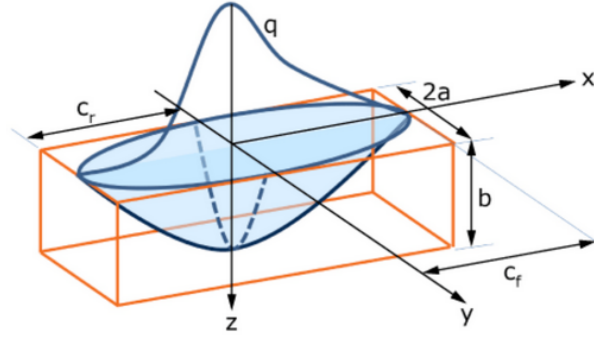
Figure 11: Semi ellipsoidal heat source.

Regarding the boundary conditions, convective heat transfer q_{conv} and radiative heat transfer q_{rad} were applied for the top surface of the cuboid

$$q_{conv} = -h_c \left(T - T_{amb}\right) \tag{3.2}$$

$$q_{rad} = -\epsilon\sigma \left(T^4 - T^4_{amb}\right) \tag{3.3}$$

where h_c was the heat transfer coefficient set to 25W/m^2°C [30], T_{amb} was the ambient temperature set to 25°C, ϵ was the powder emissivity set to 0.7 [31], and σ is the Stefan-Boltzmann constant. The rest of the rectangular sides were considered insulated. Both the heat transfer coefficient and the powder emissivity values were taken from works on Inconel 718 due to the absence of respective results for Hastelloy X. Hastelloy X [32] and Inconel 718 [33] present considerable similarities in composition and thermal properties, making this assumption acceptable under the current premises.

Finally, the laser was modelled as a 3D moving Gaussian heat source using the Abaqus AM plugin Goldak definition [34] with a semi-ellipsoidal shape, as shown in Figure 11 adapted from [28]. The volumetric heat flux of the laser source q_s moved along the positive x-axis and was defined as

$$q_{f/r} = \frac{6\sqrt{3}f_{f/r}\eta P}{abc\pi\sqrt{\pi}} \exp\left[-3\frac{x^2}{c^2_{f/r}}\right] \exp\left[-3\frac{y^2}{a^2}\right] \exp\left[-3\frac{z^2}{b^2}\right] \tag{3.4}$$

$$q_s = q_f \quad when \quad x \geq 0 \tag{3.5}$$

$$q_s = q_r \quad when \quad x < 0 \tag{3.6}$$

where P was the laser power set to 200W, η was the laser absorptivity set to 0.7 [35], a, b, c_f/r were the respective radii of the sides, the rear and the front of the ellipsoid, and f_f/r were the portion of the heat deposited respectively in the front and rear ellipsoid. In this work we set $c_f = c_r = a = d_{laser}/2$ and $f_f = f_r = 1$ resulting in a semi-spherical heat source shape. Finally, b, which represented the laser penetration depth, was set to 100µm [35].

Finally, a critical part of the SLM process is the material phase transition from powder to liquid under the heat input from the laser and then solid during cooling. In this work, temperature-dependent material properties (density ρ, thermal conductivity k, and specific heat c) character-

izing these three states were used, and the phase transition was defined by a cut-off temperature, namely the solidus temperature of Hastelloy X at 1260°C [32]. In the simulation, the phase transition was implemented using the Fortran-based user-defined UFIELD [28], and the material states were described by a Field Variable (FV) which took the value 0 in the powder state and 1 in the bulk state (solid or liquid) and is defined in the integration points of the element. To avoid sharp changes in material properties when the FV changed from 0 to 1, a new mechanism was proposed in this work based on the phenomenon of sintering that takes place as the temperature of powder increases towards the solidus temperature [13]. Specifically, the FV of powder takes values between 0 and 1 depending on a linear interpolation between the sintering temperature T_{sint} and the solidus temperature T_{solid}. As the temperature increases approaching the solidus temperature, the FV approaches the value of 1. The algorithm for the phase transition is presented below in Algorithm 1.

Algorithm 1: Phase transition

Variable: FV : Field variable, all the FV are initialized to 0;
T_{solid}: Solidus temperature, material specific value;
T_{sint}: Sintering temperature, currently set by the user;
if $timestep = 1$ *and* $increment = 1$ **then**
    if $z_{coord} < 0$ **then**
        We are in the substrate which is already solid;
        Set $FV = 1$;
    end if
end if
if $T > T_{solid}$ **then**
    We are in the liquid state;
    Set $FV = 1$;
    else if $T > T_{sint}$ **then**
        Define a sintering ratio between 0 and 1 based on the temperature scale;
        Set $sintratio = (T - T_{\text{sint}})/(T_{\text{solid}} - T_{\text{sint}})$;
        if $FV < sintratio$ **then**
            The FV can be smaller than the sintering ratio when it is transitioning from powder to bulk;
            Set $FV = sintratio$;
        end if
    end if
end if

The sintering temperature T_{sint} used in Algorithm 1 was set to 25 °C. The material properties were defined as a function of temperature and Field Variable. To simplify the material description and reduce the computational cost of the simulation, the density ρ and specific heat c were kept constant for all temperatures and material states at 8220 kg/m^3 [32] and 605 J/kg°C [32] respectively. Additionally, latent heat L was assumed to be released over a range of temperatures from a lower (solidus) temperature to an upper (liquidus) temperature. This caused an increase of the specific heat over this range of temperatures. The latent heat of Hastelloy X was set to 276kJ/kg [36] and applied between 1260 and 1660°C; the upper limit was set higher than the liquidus temperature of Hastelloy X (T_{liquid}=1350°C [32]) in order to avoid a sharp heat input, which could cause numerical instabilities. The thermal conductivity for bulk and unsintered powder is shown in Figure 12. We defined the thermal conductivity values for specific temperature points (marked with circles for powder and triangles for bulk), and Abaqus interpolated between

them. Mainly, the thermal conductivity values for the unsintered powder and bulk followed the values used by Shahabad et al.[35] However, the bulk thermal conductivity at 1660 °C was artificially increased to consider the effect of heat convection due to fluid motion inside the melt pool. The enhancement factor was 16-fold as proposed by the work of Ladani et al. [17] on Inconel 718. Finally, for the sintered powder, Abaqus interpolated between the properties of bulk and unsintered powder according to the value of the FV, which was calculated from the sintering ratio.
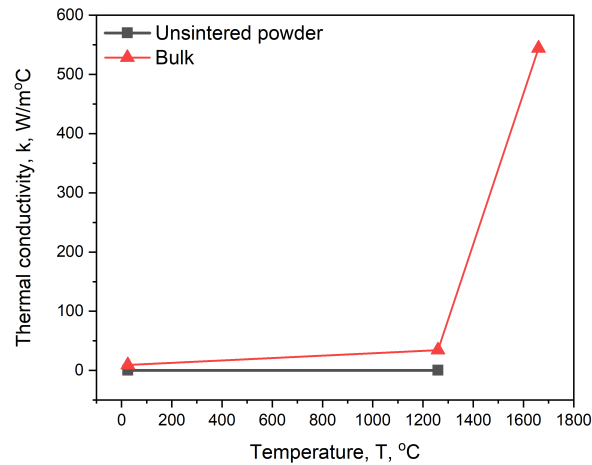
Figure 12: Thermal conductivity of powder and bulk as a function of temperature.

3.1.4 Thermal simulation

As described above, the center of the laser followed a trajectory with coordinates x=[-1,2.8], y=0, z=0.03 with a constant scanning speed. The simulation consisted of three steps: two heating and one cooling step. The region of interest was the internal local model, as described in Section 3.1.1. The first heating step was discarded as it included the time before the laser beam entered the internal local model, and it was used to precondition the laser. The second heating step recorded the temperature field when the laser was inside the internal local model. Finally, the cooling step recorded the temperature field during the cooling phase after the laser had left the internal local model. The duration of the second heating step and the cooling step was 0.0017s and 0.0011s, respectively, and the temperature field was recorded every 0.0001s resulting in 17 frames during heating and 11 frames during cooling. These were later used as the target output for the surrogate model presented in the following section. Finally, an initial temperature of 25°C was chosen for the model.

3.2 Surrogate model setup for different initial temperatures

The goal of the surrogate model was to replace the expensive FE simulations; in this work, the target output was to predict the temperature field of the internal local model during heating and cooling. According to the simulation setup described in the previous section, the internal local model consisted of 77363 nodes, on which the temperature was tracked over 17 frames during heating and 11 frames during cooling. This resulted in an array of 77363x28=2166164 data points for every FE run. The main source of non-linearities in the temperature field was

introduced by the radiative and convective heat fluxes on the top surface of the model and the material-dependent thermal properties (mainly the thermal conductivity), which were altered during the phase transition caused by the laser heat input. As a first step, the input data included only the initial temperature, which was chosen to be the same for all the nodes of the model.

As previously mentioned, the surrogate model was set up similarly to the work of Keller [12]. A PCA scheme was used to reduce the dimensionality of the input/output data, and PCE was used to find and approximate a map between the reduced input and output. An inverse PCA was then used to recover the output data from the reduced output predicted by the PCE. Here, a PCA was only implemented for the output temperature distribution. The input data was described with a single value (the initial temperature). Therefore, only a standardization with mean 0 and variance 1 was applied to the array with the initial temperatures of the FE runs, before using it as input to the PCE. The coefficients of the PCE were determined and optimized internally by the UQLab library functions using training temperature data from the FE model. The user-set parameters determined the range for the truncation of the PCE; the adaptive polynomial degree p ranging from 1 to 20, the adaptive q-norm truncation with q ranging from 0.1 to 0.9 with a step of 0.1, and the maximum interaction order of 2. Finally, the output temperature distribution was first standardized with a 0 mean and variance 1, and then the principal components were calculated. No truncation of the principal components was used in this preliminary work to minimize the error from the reconstruction of the reduced output during the inverse PCA transform.

3.2.1 Surrogate model evaluation

At this point, it is important to note that the PCA basis was computed from the same training data that were used to train the PCE. This highlights the importance of defining a well-representative dataset for training. As it was shown in the work of Keller[12], it was critical to define how much training data is needed. Due to the high computational cost of the FE simulations, which built up the training set, the goal was to use as few runs as possible for training without compromising the fidelity of the surrogate model. Keller[12] emphasized the need for defining an equidistantly distributed training set that represents the whole range of the experimental design. The dataset in this work consisted of 10 FE runs with different initial temperatures $T_{initial}$ [°C] with

- $T_{initial} \in \{25, 100, 150, 200, 250, 300, 350, 400, 450, 500\}$

To evaluate the performance of the model, the dataset was split into a training set and a test set. The training set was used to obtain the principal components and the coefficient of the PCE. When the model was built, it was evaluated on the test set, as an independent set of observations from the training. The performance of the model on the test set was used to obtain the prediction error. To generalize the performance of the model, a hold-out validation scheme was implemented, where the dataset was split randomly into a training and test set according to a predefined split ratio. To reduce the variance induced by the randomness in the splits, the errors were averaged over ten such splits. Finally, the implementation of the prediction error followed a commonly used error measure (also implemented in the work of Keller), namely, the normalized root mean square error (RMS), which is defined as

$$RMS = \sqrt{\frac{\sum_{i=1}^{n_{val}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{val}} (y_i - \bar{y}_i)^2}}. \tag{3.7}$$

where n_{val} is the size of the evaluation dataset, y_i are the true observations, $\hat{y}_i$ are the model observations and $\bar{y}_i = 1/n_{val} \sum_{i=1}^{n_{val}} y_i$ is the mean value of the true observations. The RMS error computes the L^2 deviation of the model observations from the true observations, normalized by the standard deviation of the true observations.

Finally, to test the effect of training on an equidistantly distributed training set instead of using a random split of the data, we trained the model on a set that always included runs with the minimum and maximum initial temperature values, e.g., 25°C and 500°C. We then increased the runs in the training set, choosing consistently equidistantly distributed values. We evaluated the model performance on the test set that was recovered after the split, using the RMS. The detailed data split is presented in Appendix A.

3.3   Surrogate model setup for different powder distributions

The second goal in this work was to build up a surrogate model that emulated the temperature field caused by a single laser pass for different bulk geometries. For this reason, we kept the initial temperature constant at 250°C, and distributed a predefined volume of powder inside the substrate. To simplify the problem and provide realistic cases used in the SLM process, we placed the "pockets" of powder on the edges of the large rectangular box, and linearly varied the amount of powder along each axis. A schematic of this method is shown in Figure 13, where we marked with purple the volume of the substrate replaced by powder. The black arrows indicate how the powder pocket volume was varied along the x, y, and z-direction shown in 13a, b and c respectively. The description of the powder pocket position, which was later used as input for the surrogate model, was given by six values. The coordinates of the center of the rectangular containing the powder pocket (x_c, y_c, z_c) and its sides x_{size}, y_{size}, z_{size}.

| (a) | (b) | (c) |
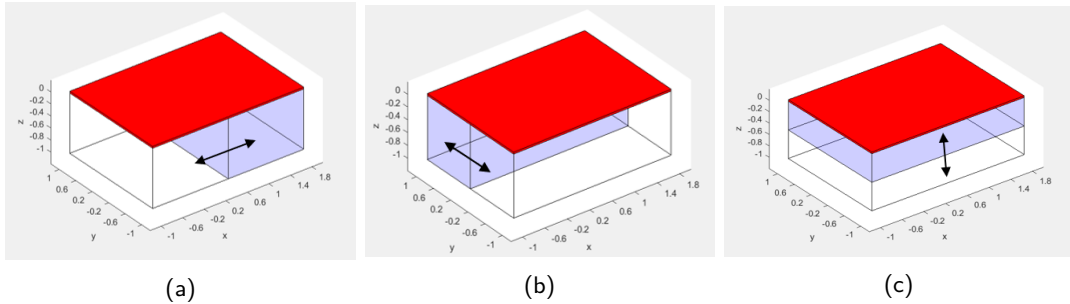
Figure 13: Model setups with different powder distributions in the substrate.

The powder pocket in the substrate was initialized with FV=0 in the FE model. Fourteen runs were set up for each case shown in Figure 13; the schematics of the model geometry of every run are shown in Appendix A. Four surrogate models were trained and tested; one for every case (variation in x, y, and z) and one containing runs from all three cases. By combining the cases, we wanted to test how the surrogate model would perform on a database with more variability. The same scheme of PCA and PCE that was used in the surrogate model for initial temperature was implemented. Additionally, a PCA transform was applied on the input vector containing the 6 values that described the position of the powder pocket. The surrogate model performance was evaluated using the RMS error, and the datasets were split into training and test sets both randomly and using equidistantly distributed values; the splits are shown in Appendix A. To further expand the database of different geometries, a new set of runs was carried out where the

(a)   (b)   (c)
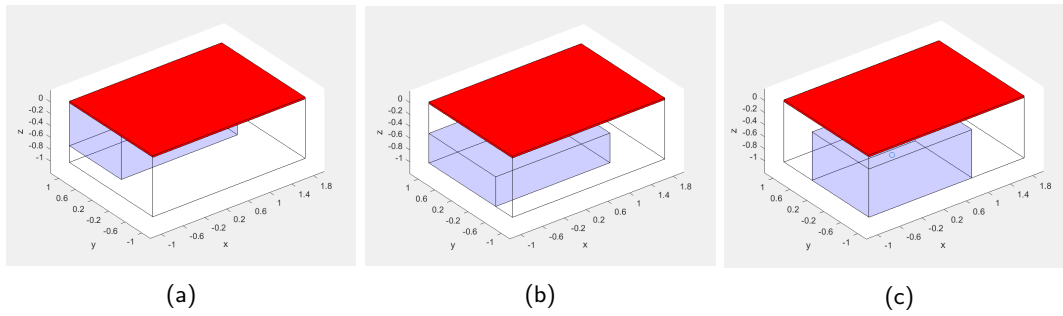
Figure 14: Examples of powder pockets placed on the corners of the substrate.

powder pockets were placed on the corners of the larger rectangular box. The dataset consisted of 32 runs (4 different powder pockets in every corner), where the size of the powder pockets was randomly selected. Examples of the powder pocket configurations are shown in Figure 14 and the full database is included in Appendix A. The restriction that was implemented was that the sides of the boxes containing the powder did not exceed the preset maximum and minimum value. This values were the same that were also employed in the previous powder pocket database shown in Figure 13, and are listed in Table 2. Finally, the surrogate model and its evaluation followed the scheme of the previous powder pocket database. An additional leave-one-out scheme was used for each corner to calculate an average error; the training set contained three runs for every corner (random split) and the fourth run was used in the test set.

Table 2: Maximum and minimum values of the sides of the box containing powder in the substrate.

Side	Minimum [mm]	Maximum [mm]
x_{size}	0.56	2.24
y_{size}	0.4	1.6
z_{size}	0.2	0.8

4 Results

4.1 Validation of the FE model

Before setting up the surrogate model, the first step was to validate the FE simulations to other works in the literature. The two results derived from the thermal simulation, which could be directly comparable, were the peak temperature and the melt pool dimensions. Here, the challenges were not only the limited published experimental and computational works on Hastelloy X but also the large variation of process and simulation parameters, which did not match the respective parameters used in this work. For this reason, we attempted to identify general trends between SLM process parameters and melt pool size. At the same time, literature on Inconel 718 was used for comparison in the absence of respective results for Hastelloy X. The latter is an acceptable assumption for this work since Hastelloy X and Inconel 718 present considerable similarities in composition and thermal and mechanical properties, shown in [32] and haynesinc respectively.

4.1.1 Peak temperature

In experiments, the peak temperature depends mainly on the laser parameters. This is mainly regulated by the physical phenomena inside the melt pool, which are controlled by mass and heat transfer (e.g., heat convection and radiation, material vaporization, Marangoni convection, sintering, phase change, capillary effects, etc.) [8]. The majority of these phenomena are not directly considered in thermal FE mesoscale simulations such as the one used in this work as described in section 3. Instead, some thermal properties and simulation parameters were used as fitting values to approximate the effect of these phenomena, such as the enhanced liquid thermal conductivity [17], the laser absorption, and the penetration depth of the heat source [35]. Specifically, the thermal conductivity of the liquid phase was enhanced by 16-fold following the work of Ladani et al. [17] on Inconel 718, where the simulated melt pool dimensions compared favorably with the experimental values after implementing the enhancement factor. They reported a peak temperature between 1800 and 2100°C in their simulations. A similar enhancement factor (15-fold) again for Inconel 718 was reported by Luo et al. [30] with peak temperature between 2000 and 2300°C. By following a similar strategy in the enhancement factor, the reported peak temperature in this work at 2043°C was well within the expected values as shown in Figure 15.
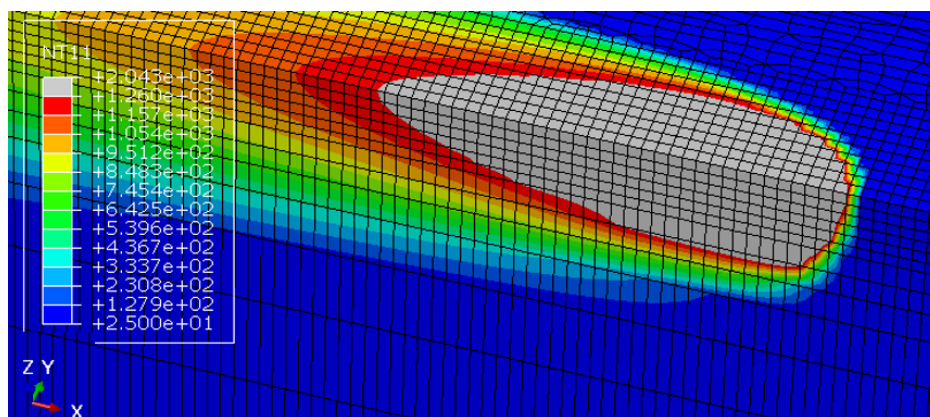

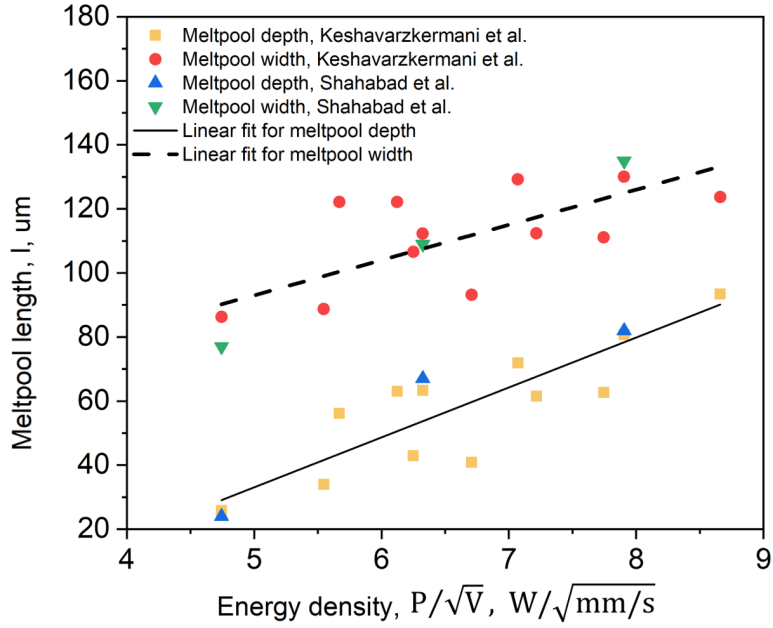Figure 15: Temperature distribution near the melt pool from the FE model.

Figure 16: Experimental melt pool length as a function of the energy density. The experimental data were taken from the work of Keshavarzkermani et al. [37] and Shahabad et al. [35]

4.1.2 Meltpool dimensions

The melt pool size, usually characterized by the melt pool depth and width, is a commonly reported experimental measurement in SLM works as it is directly connected to the quality of the parts [37]. Partial or incomplete melting, as well as reheating and/or remelting of the solidified substrate influence the microstructure and the mechanical properties of the final part [38]. In general, to optimize the quality of the printed parts, an investigation for the optimal process parameters takes place, mainly of the laser power, scanning speed, and layer thickness [6, 37]. In the literature, there is a clear correlation between laser power and the scanning speed, and the melt pool size. Namely, the melt pool increases with increasing laser power and decreases with increasing scanning speed; this behavior has been reported extensively for many materials used in SLM [6, 7, 39, 37].

To combine the effect of these two parameters in order to come up with guidelines for optimal printing conditions, many works propose an energy density expression, which is defined as a ratio of the laser power over the scanning speed, with both values raised to an exponent. The exponents can vary between different works in the literature [37, **?**]. Here, we defined the energy density as a ratio of the laser power P over the square root of the scanning speed V with units $W/\sqrt{mm/s}$.

$$E_{\text{density}} = \frac{P}{\sqrt{V}} \tag{4.1}$$

This energy density expression has been proposed by the Multi-Scale Additive Manufacturing Lab in the University of Waterloo [7]. It has been used as a guideline for parameter optimization in both simulation and experimental works for multiple materials processed with SLM. Specifically, Zhang et al. [7] and Shahabad et al. [35] propose a linear relation can be derived between the

melt pool depth and width, and the energy density expression (Equation 4.1) for Stainless Steel 17-4PH and Hastelloy X, respectively. Further experimental measurements of melt pool depth and width for Hastelloy X found in the work of Keshavarzkermani et al. [37] showed that there is no clear correlation of the layer thickness and the melt pool size. We combined the experimental data of both works on Hastelloy X (Shahabad et al. [35], and Keshavarzkermani et al. [37]) in Figure 33. We used these data to calculate the coefficients of a linear relation of the meltpool depth and width for Hastelloy X as a function of the energy density. Detailed parameters of the linear fittings can be found in Appendix B.

$$d_{\mathsf{meltpool}} = 15.57 * \frac{P}{\sqrt{V}} - 44.75 \qquad (4.2)$$

$$w_{\mathsf{meltpool}} = 10.99 * \frac{P}{\sqrt{V}} + 38.04 \qquad (4.3)$$

We used these linear relations to estimate an expected meltpool depth and width, and their standard error, for the energy density used in this work with laser parameters $P = 200W$ and $V = 900mm/s$. The meltpool depth and width extracted from the FE model in this work were well within the predicted range suggested by the experimental data as shown in Table 3.

Table 3: Comparison of meltpool dimensions

| Source	d_{meltpool} [µm]	w_{meltpool} [µm]
This work - FE simulation	50	120
Experiments	59.1 ± 11.1	111.3 ± 13.2

4.2   Surrogate modelling for different initial temperatures

The goal for the surrogate model built was to be able to predict the temperature field of the inner local model under the influence of a single laser pass. The temperature field was tracked over 28 frames; 17 frames were recorded during the heating step where the laser was inside the inner local model, and 11 frames were recorded during cooling when the laser had left the inner local model. The first step was to predict the temperature field for different initial temperatures. As described in Methods, the dataset contained ten runs with initial temperatures ranging from 25 up to 500°C. These values are consistent with the range of preheating temperatures commonly used in SLM [40]. Using a 70/30 split ratio (70% of the observations were used for training and 30% for testing), the normalized RMS was calculated on the test set and averaged over ten random splits. The average and standard deviation of the RMS was 3.8 ± 5.2 %. This performance was in part expected since the difference in initial temperature did not introduce large deviations in the temperature profile. This was also validated when looking at the percentage of variance explained by the principal components used to reduce the dimensionality of the output (temperature field) of the surrogate model; the first principle component recovered by the PCA could explain 99.9% of the variance. In terms of computational cost, the duration of each FE run using 24 threads on an Intel Xeon Gold 5118 processor in the Euler cluster was calculated on average to 8 hours (wall clock time). On the other hand, once the database was built, the training and evaluation of the surrogate model were approximately 1-2 minutes on a local PC.

   In Figure 17, a comparison of the temperature profile of a FE run and a surrogate model
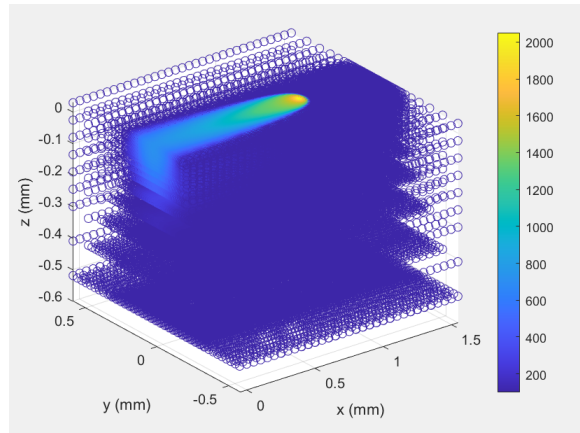
(a) FE model
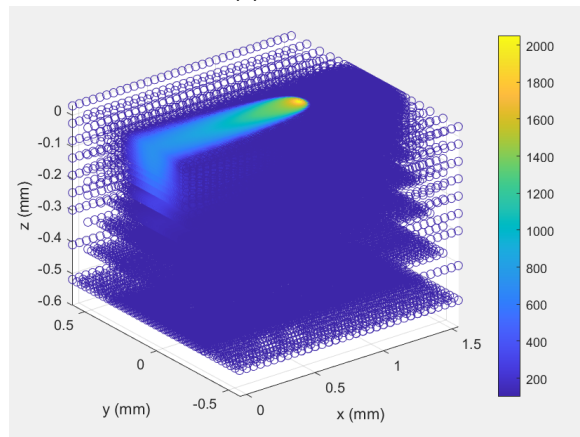
(b) Surrogate model
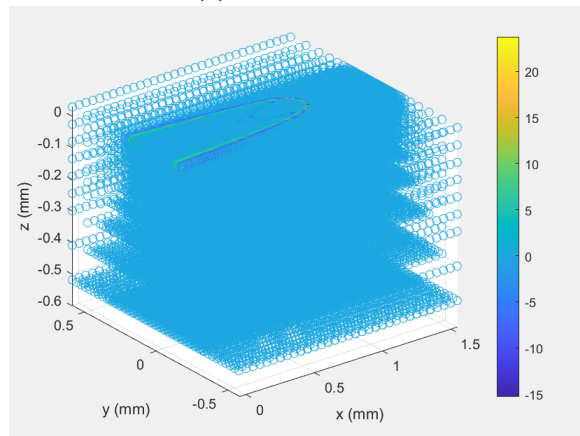
(c) Absolute temperature difference

Figure 17: Comparison of the temperature profile of the FE model and the surrogate model during heating.

for initial temperature 150°C is shown. The RMS for this run was calculated at 2.8%. The temperature field in (a),(b) correspond to frame 10 during the heating step produced by the FE model and the surrogate model, respectively. In (c), the absolute difference between the previously mentioned models is plotted. Similarly, Figures 18 (a),(b),(c) correspond to frame 20 during cooling for the FE model, surrogate model, and their absolute temperature difference, respectively. The largest temperature difference is shown to be on the edges of the melt pool for both cases. The reason for the deviation could be the larger spatial thermal gradients, which are common in the region. Another explanation is that the edges of the melt pool also
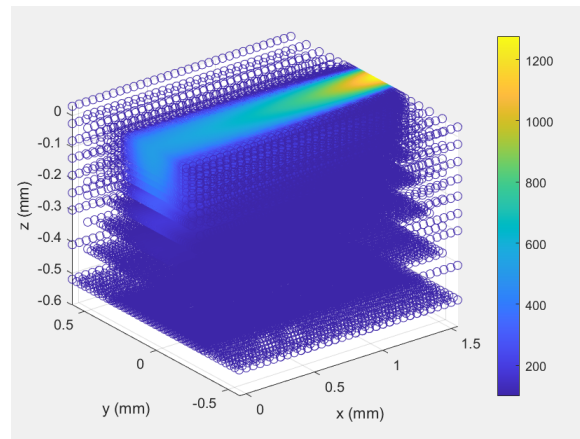
(a) FE model
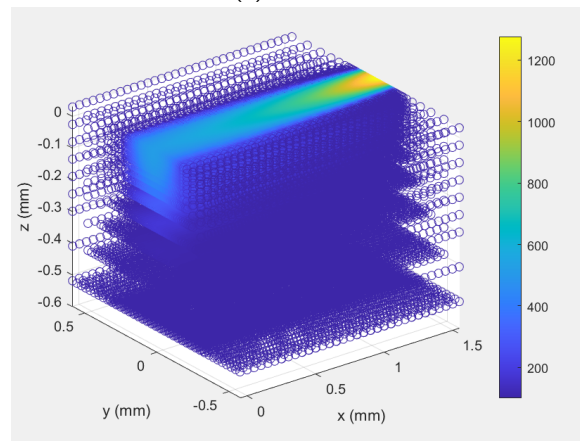
(b) Surrogate model
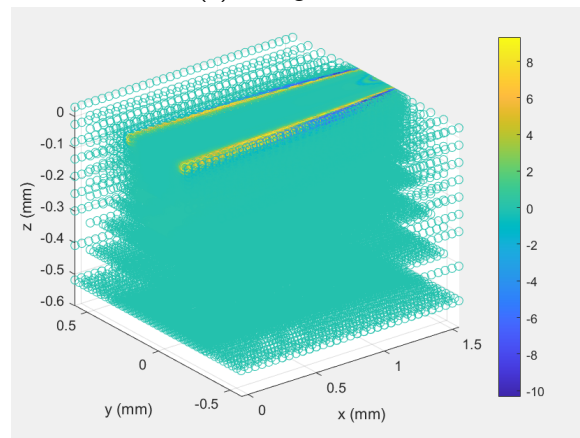
(c) Absolute temperature difference

Figure 18: Comparison of the temperature profile of the FE model and the surrogate model during cooling.

correspond to the area were the mesh transitions from 10µm elements to increasingly larger ones using unstructured elements. The nodes of the unstructured elements could cause some inconsistencies in the way the temperature profile is saved in the data arrays, which in turn could be challenging for the PCE to map the dependency of these nodes to the temperature profiles.

To characterize better the surrogate model, we wanted to evaluate the influence of the experimental design on its performance. For this reason, we used training sets of increasing size (from 3 to 8 observations). The RMS error of the surrogate model was computed using a hold-out validation scheme, where the dataset was split randomly into a training and test set according

to a predefined split ratio. To reduce the variance induced by the randomness in the splits, the errors were averaged over 10 such splits. The results are shown in Figure 19. The solid line shows the average RMS error on the test set over ten different random splits, whereas the shaded region depicts the minimum and maximum values of the error. The error decreases with increasing size of the training set, which was expected as the more data the surrogate model learns, the easier it is to find the map between the input and the output of the underlying physical model. By including more than five runs in the training set, the error falls under 5%. However, the maximum error value does not decrease significantly and is shown to exceed 15% in some cases. These were often the case when the training set did not involve the extreme values of the dataset, e.g., the runs with initial temperature 25 and 500°C.
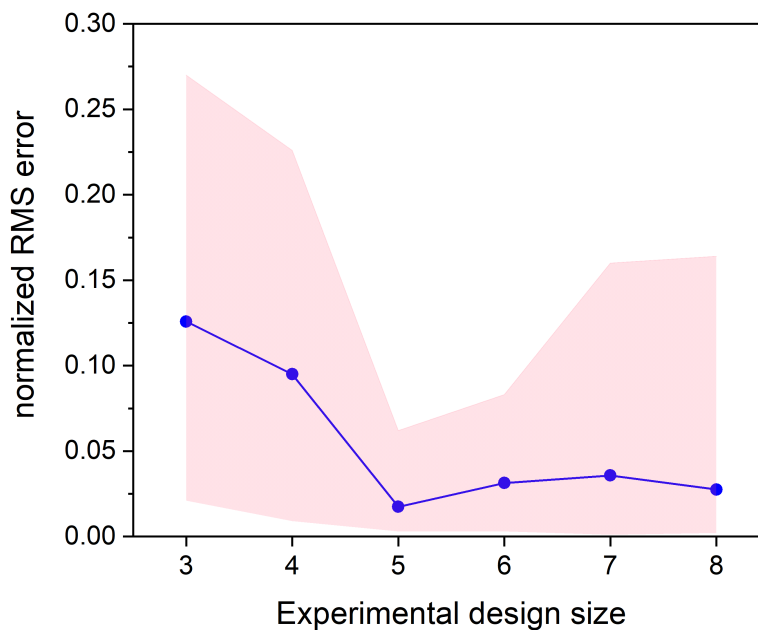
Figure 19: Normalized RMS error plotted against the size of the experimental design for the training of
the surrogate model. The experimental design was selected using random splits of the available
dataset.

By investigating the cases where the error was high, a pattern was observed. The surrogate model performed worse when the runs used for training did not correspond to initial temperature values well distributed within the range of the dataset. Namely, if the experimental design (training set) did not represent the dataset correctly, the performance of the surrogate model would not be consistent. For this reason, we implemented new training and evaluation sets based on the size of the experimental design, where the training set was not selected from random splits. Instead, we handpicked the runs in the training set so that they are equidistantly distributed in terms of initial temperature and always contained the runs with the minimum and maximum initial temperature. For example, the experimental design containing three observations was trained on the runs with initial temperature 25, 250, and 500°C. The experimental design containing four observations was trained on the runs with initial temperature 25, 150, 350, and 500°C. The user-selected experimental designs are included in Appendix A. The respective test sets included the runs that remained in the dataset after the experimental design was selected. The results of

this process are shown in Figure 20. The solid line represents the average error on the respective test set, and the shaded area the maximum and minimum error on each test set. There is a clear indication that by wisely selecting the experimental design, the prediction error is minimized under 1% for experimental design size over 5. In general, this is consistent with the surrogate model theory and the observations in the work of Keller. However, here, an experimental size of 3 seems to perform better than an experimental size of 4. This inconsistency could be due to the limited and user-selected number of runs in this work and it indicates that some runs could be more suitable than other for the training of the surrogate model. Therefore, a more systematic work is needed to extract statistically relevant guidelines for designing a future database for SLM processes.

Figure 20: Normalized RMS error plotted against the size of the user-selected experimental design for the training of the surrogate model.

4.3 Surrogate modelling for different powder distributions

The second challenge addressed in this work was how the temperature field was affected by the existence of powder pockets inside the bulk of the substrate. The goal was to test whether the surrogate model could map the position and volume of the powder pocket to the output temperature distribution. The first step was to use powder pockets normal to the three Cartesian axes and vary their volume along the respective axis, as shown in Appendix A. The surrogate model was firstly evaluated on a randomly split experimental design (training set), as shown in Figure 21. The RMS error on the test set is plotted against the increasing size of the training set, where the solid line represents the average error over ten random splits, and the shaded area depicts the minimum and maximum values of the error. It is shown that the surrogate model performs similarly for powder pockets varied along the x and y-axis (RMS is approximately 8%), whereas for the case of the variation along the z-axis, the RMS drops to 2%.

Similarly to the surrogate model built for the initial temperature variation, we tested the effect

(a) Powder pocket volume varied along the x-axis

(b) Powder pocket volume varied along the y-axis
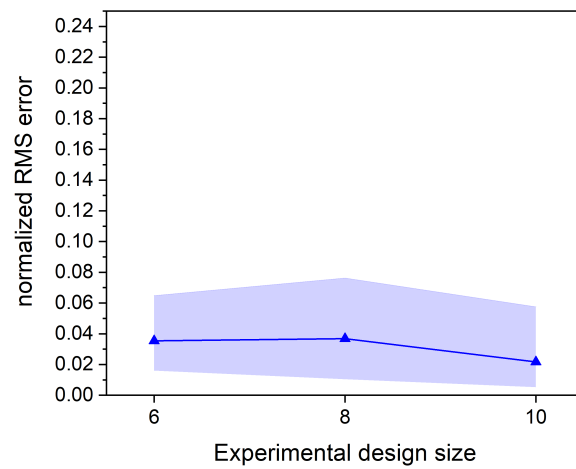
(c) Powder pocket volume varied along the z-axis

Figure 21: Normalized RMS error plotted against the size of the randomly split experimental design for the training of the surrogate model.
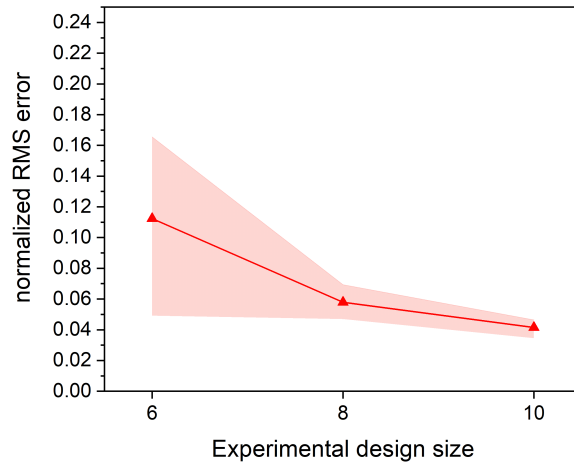
(a) Powder pocket volume varied along the x-axis
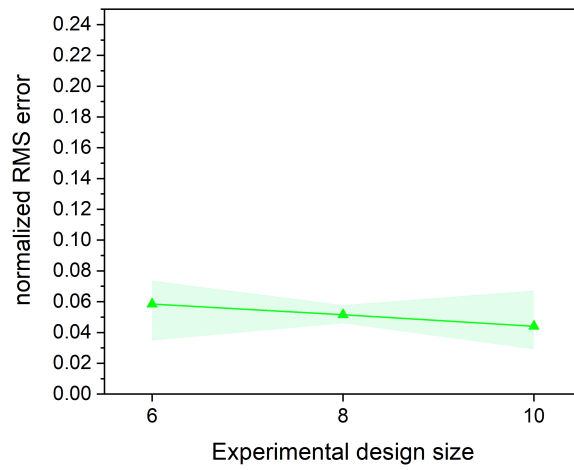
(b) Powder pocket volume varied along the y-axis
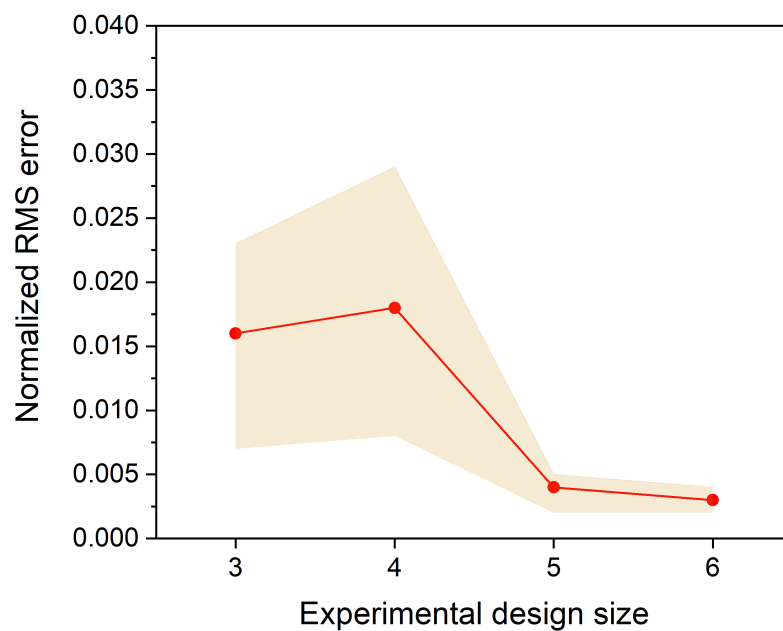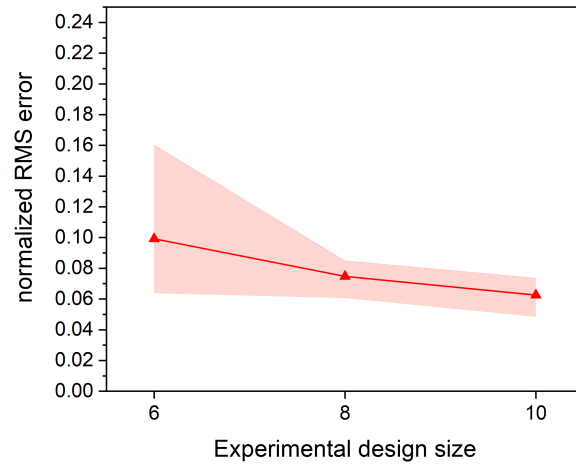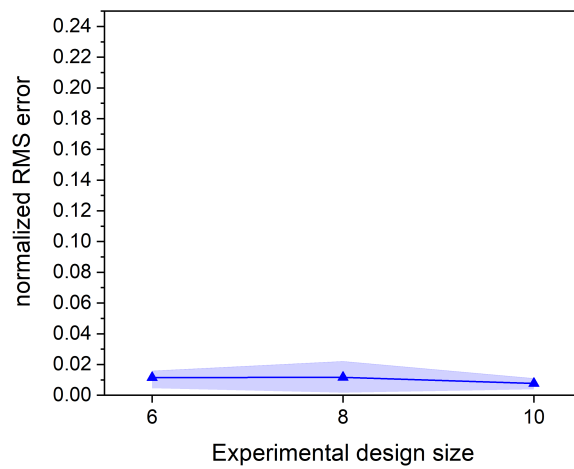

(c) Powder pocket volume varied along the z-axis

Figure 22: Normalized RMS error plotted against the size of the experimental design for the user-selected training of the surrogate model.
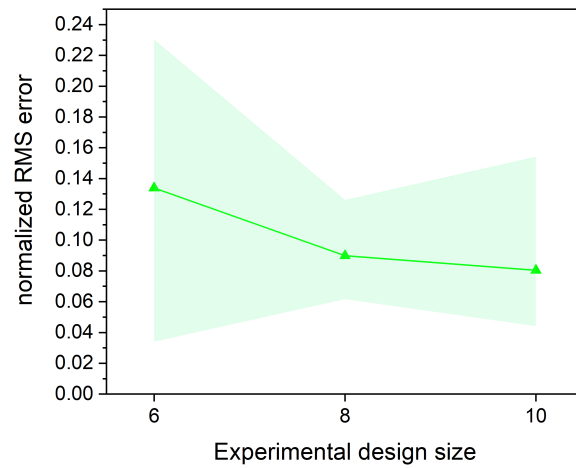
of using a user-defined equidistantly distributed training set, the results for which are shown in Figure 22. As was expected, the results of the surrogate model improved significantly for all three cases. The size of the experimental design produced by a 70/30 split (10 runs in the training set) was shown to be optimal, decreasing the error to 4% for the powder pockets varied along the x and y-axis and to 1% for the respective z-axis. An example of each case corresponding to the training set with 10 observations is shown in Figures 23, 24 and 25 during heating (frame 10). Specifically, in (a), the power distribution schematic is illustrated for each example. In (b), the absolute temperature difference between the FE model with the powder pocket and the respective surrogate model is plotted. Finally, the absolute temperature difference between the FE model with the powder pocket and the FE model without the powder pocket (all solid substrate) is shown in (c). The last image was added to showcase how much the temperature field changed by introducing the powder pocket in the substrate.

Until now, three separate surrogate models were trained for each case of varying the powder pocket along the three Cartesian axes. By combining all of them in a common dataset, the goal was to see how the surrogate model would perform when trained for a future database containing characteristic cases of different powder distributions. This was tested for the case of equidistantly distributed training sets containing ten runs for each case. In Table 4 the RMS error on individual test runs is shown. Runs 1-4, 5-8, and 8-12 have powder pockets varied along the x, y, z-axis, respectively. The RMS error from the surrogate model trained on each separate dataset and a common one is compared with good agreement showing the potential of building a common database of local model runs.

Table 4: RMS error of surrogate models trained on three separate and one common training database

Test run index	RMS with separate training databases	RMS with common training database
1	0.042	0.035
2	0.035	0.036
3	0.046	0.050
4	0.043	0.047
5	0.067	0.023
6	0.029	0.030
7	0.042	0.073
8	0.038	0.026
9	0.005	0.014
10	0.006	0.013
11	0.002	0.017
12	0.004	0.008

The next goal was to see how the surrogate model would perform in more complicated cases where the powder pockets were placed on a corner of the local model. Instead of building a separate surrogate model for every corner, we combined all runs in a shared database. The performance of the surrogate model is shown in Figure 26. The RMS of the surrogate model trained on a database with powder pockets in corners is shown in red. To further assist the surrogate model, we added in the experimental design the cases where the powder pockets were varied linearly along the Cartesian axes. The results of the extended database are shown in blue. Finally, an RMS error was calculated between the FE model with the powder pocket and the FE model without considering the powder (all solid substrate) and is shown in black. This was done

to define a measure of how much the existence of the powder pocket caused a change in the temperature field. Since four runs were available for every corner, an averaged error and standard deviation were calculated by a leave-one-out scheme, where the training set included three runs and the test set one run for each corner. As shown in Figure 26 expanding the databases lowers in most cases the RMS of the surrogate model. However, a weakness of the surrogate model is also shown here. It is shown that when powder pockets were placed in the lower corners (corner indices 1,3,5 and 7), the presence of the powder did not affect the temperature distribution, showing a deviation of the temperature field close to 0, something the surrogate model was not able to detect.

After evaluating the surrogate model on different powder distributions in the substrate, some possible explanations for the shortcomings of the current approach are addressed. The surrogate model was shown to perform better when the experimental design (training set) included extreme cases based on the input, and, subsequently, when the test set included intermediate cases that could be "interpolated." This was shown for both the surrogate model for the initial temperature variation and for the powder pockets. However, as shown more clearly in Figure 26, the presence of powder did not always affect the temperature field. If the powder pocket had a certain distance from the laser pass, the powder pocket volume had no significant influence. The surrogate model should be able to recognize that there is a cut-off distance for which the geometry of the powder pocket (given here as the input to our surrogate model) will result in the same temperature field. Under the current input, this would require the surrogate model to be trained in many examples where the powder pocket is far away from the substrate. This was in part true for the powder pockets where the volume was varied along the z-axis, which could be the reason for the RMS error of the surrogate model being close to 1% compared to the cases where the powder pocket volume was varied along the x and y-axis. Parallel to increasing the size of the experimental design, it would be beneficial incorporate also temporal information as input to the surrogate model. This could be done in the form of the absolute distance of the center of the powder pocket from the laser source for every time frame. Finally, even with the above mentioned shortcomings, the benefits of the surrogate modelling approach for lowering the computational cost are clear; a FE simulation run with a powder pocket in the substrate required on average 11 hours (wall clock time) to be completed, whereas the respective surrogate model training and evaluation required 3 to 4 minutes on a local PC. The submission in Euler cluster involved using 24 threads on an Intel Xeon Gold 5118 processor.
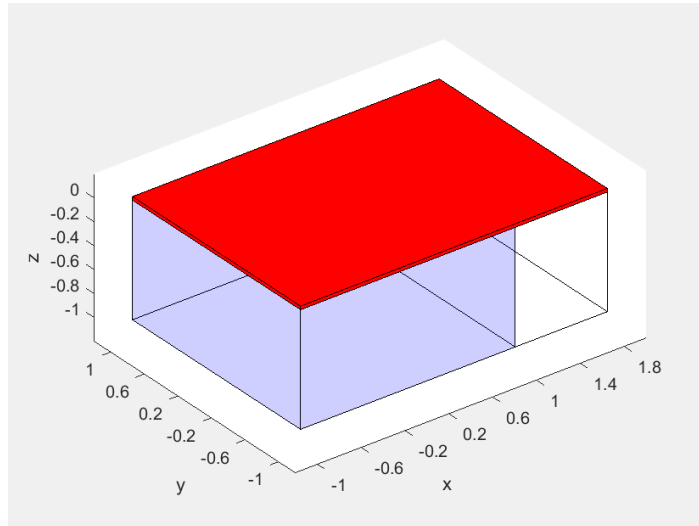
(a) Schematic of powder distribution in the substrate.
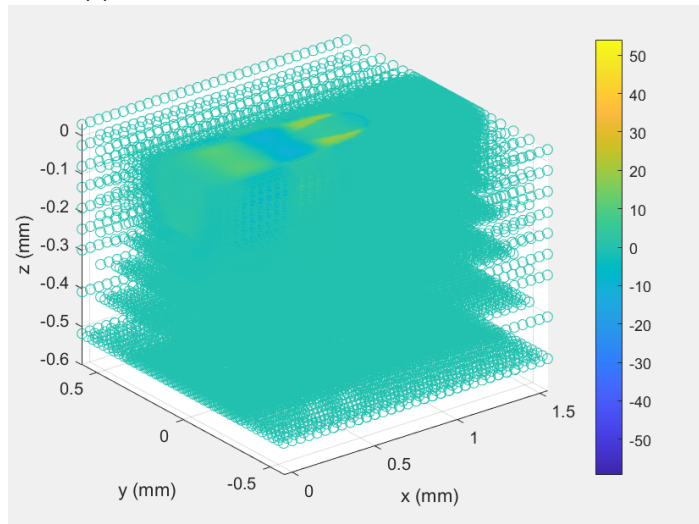
(b) Absolute temperature difference of FE model with powder and
    surrogate model.
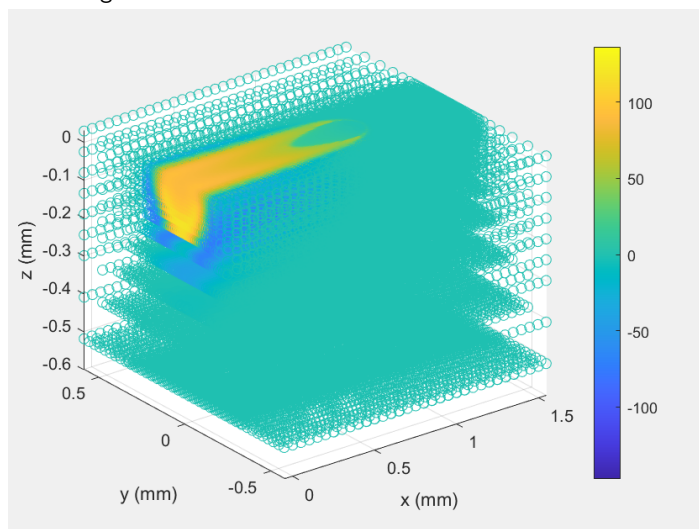
(c) Absolute temperature difference of FE model with powder and
    FE model without powder.

Figure 23: Comparison of the temperature profile of the FE model and the surrogate model with the
           presence of a powder pocket volume in the substrate. The RMS error between the FE model
           and the surrogate model was calculated at 3.5%.
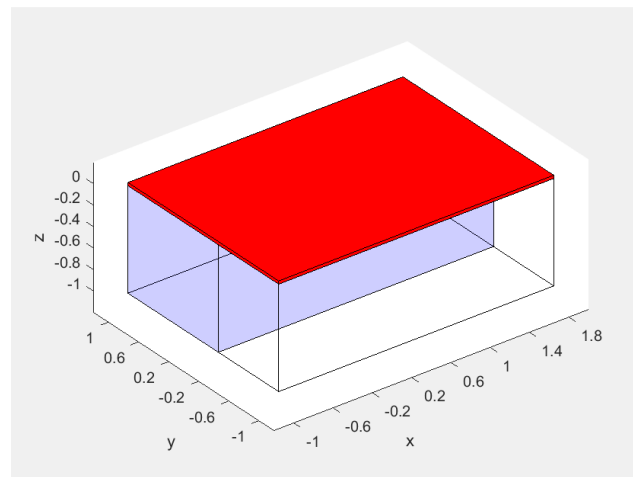
(a) Schematic of powder distribution in the substrate.
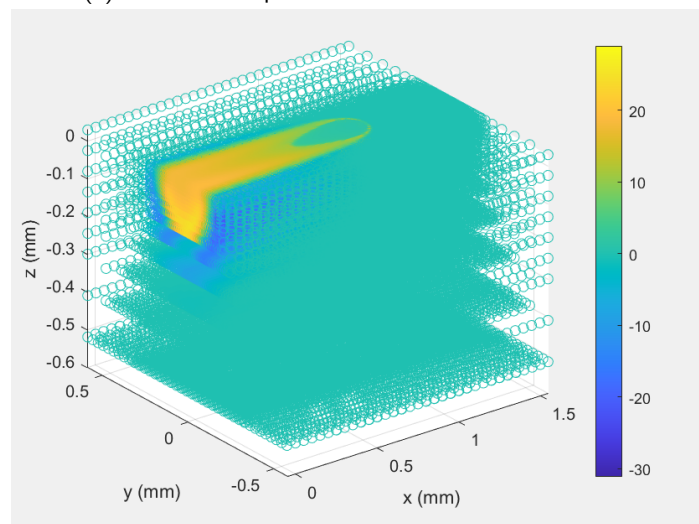
(b) Absolute temperature difference of FE model with powder and
    surrogate model.
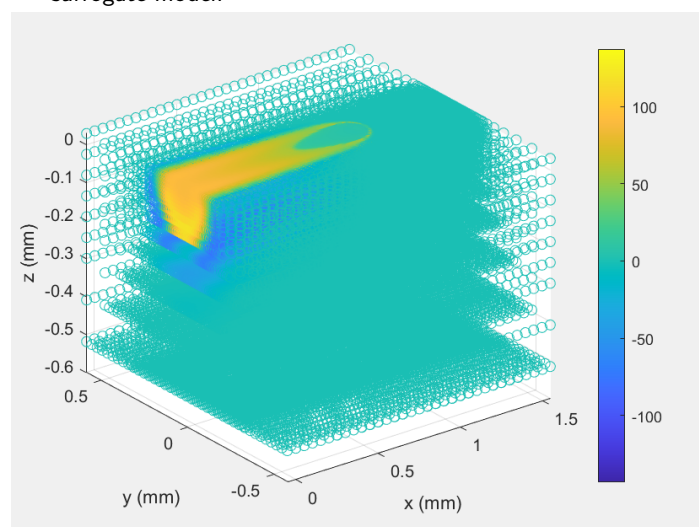
(c) Absolute temperature difference of FE model with powder and
    FE model without powder.

Figure 24: Comparison of the temperature profile of the FE model and the surrogate model with the presence of a powder pocket volume in the substrate. The RMS error between the FE model and the surrogate model was calculated at 2.9%
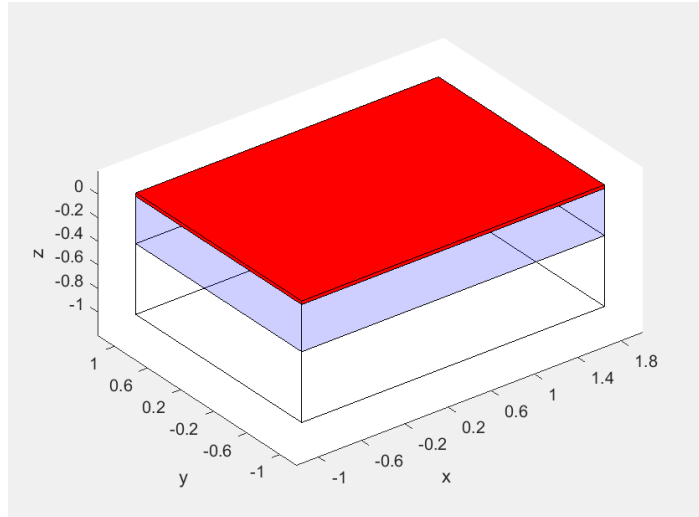
(a) Schematic of powder distribution in the substrate.
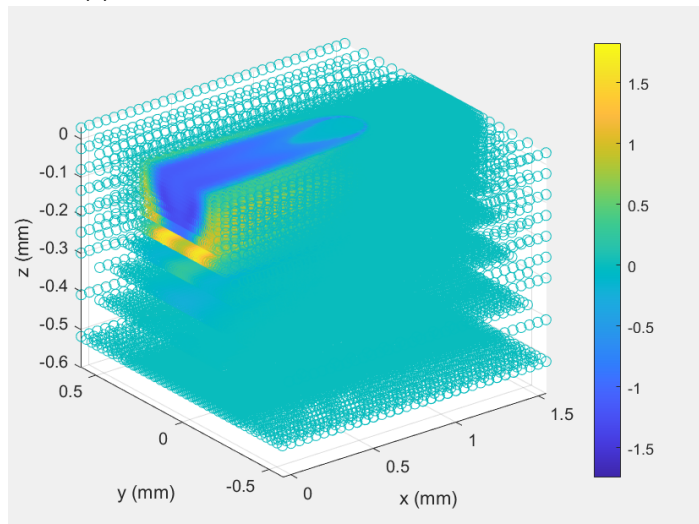
(b) Absolute temperature difference of FE model with powder and
surrogate model.
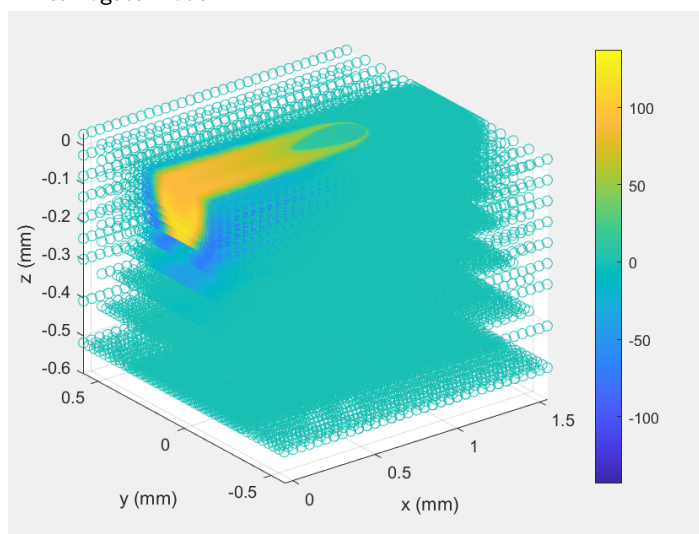
(c) Absolute temperature difference of FE model with powder and
FE model without powder.

Figure 25: Comparison of the temperature profile of the FE model and the surrogate model with the
presence of a powder pocket volume in the substrate. The RMS error between the FE model
and the surrogate model was calculated at 0.2%.
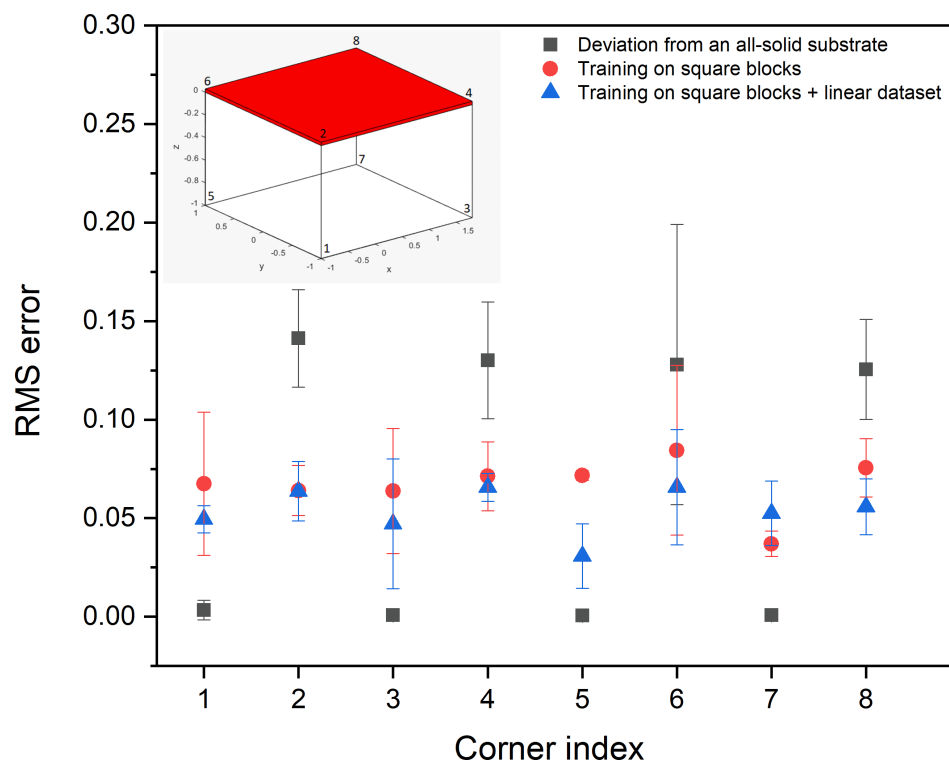
Figure 26: Normalized RMS error for powder pockets placed on the corners of the local model.

5 Conclusions and Outlook

The goal of this thesis was to explore the potential of using a surrogate model to replace costly 3D thermal simulations for powder-bed additive manufacturing and, more specifically, selective laser melting. This work is part of a larger multiscale framework developed in the High-Temperature Integrity Group in *Empa* for solving finite element (FE) simulations more efficiently. Specifically, an adaptive glocal-local approach is proposed where a coarse mesh is used to solve the thermal model in the areas far away from the laser pass, and a fine mesh was used in the vicinity of the melt pool to capture the physics and kinetics of the highly concentrated and fast-moving heat source (laser). These two solutions are combined by solving a global model of the part with a coarse mesh and correcting the temperature profile in the vicinity of the melt pool by multiple local models that follow the laser pass. To further decrease the computational cost of these local models, a surrogate model can be trained to emulate the behavior of the expensive FEM simulations based on the initial conditions of the model (initial temperature profile) and the local geometry. The latter is typically characterized by the existence of unsolidified volumes of powder (powder pockets) in the substrate. In a previous master thesis[12], the surrogate modelling was successfully applied for different 2D geometries using a combination of principal component analysis (PCA) and polynomial chaos expansion (PCE) to decrease the dimensionality of the temperature distributions and map the dependency between the input and output distribution, respectively. In this work, we explored how a similar approach could be applied for the 3D case.

The first step was to define a 3D local model and setup the respective FE thermal simulations. The thermal simulation consisted of a single laser pass, and the temperature profile was tracked during heating and cooling. The mesh of the local model was optimized using different meshing strategies in order to minimize the number of elements and nodes without compromising the temperature profile of the FE simulations. The thermal model was developed for the printing of Hastelloy X and was tuned to capture the phase transition from powder to liquid and then solid based on experimentally relevant process parameters and material properties. The FE model was validated by comparing the simulated melt pool size with respective experimentally measured values, which was in excellent agreement. Additionally, the peak temperature reported in our model was consistent with other simulation works using a similar methodology. Two types of FE simulations were carried out. The first series of runs involved a solid substrate, and a single laser pass selectively melting a powder layer for different initial temperatures. The initial temperature was the same for all the nodes of each run and was used as a single value input to the surrogate model. The target output was the temperature field of the local model during heating and cooling. A PCA was used to calculate the principal components of the output temperature distribution. Subsequently, a PCE was used to map the dependency of the initial temperature to the reduced output distribution. The surrogate model was trained on a range of FE runs with different initial temperatures and evaluated on an independent test set using as measure the normalized root mean squared (RMS) error. In general, the surrogate model was successful in predicting the temperature distribution for different initial temperatures. However, the RMS error was reduced below 1% when the training set included runs corresponding to initial temperature values well distributed within the range of the whole dataset. Namely, the training set should contain values that represent well the range of observations that the surrogate model is expected to emulate. This can be achieved either by making sure that the training set is trained in a large dataset of

randomly distributed runs or, in our case, where the FE runs are computationally expensive and limited, representative runs can be selected by the user.

The second series of FE runs introduced pockets of unsolidified powder inside the substrate. The initial temperature was kept constant over all the runs, and cuboids of powder (powder pockets) were distributed in a consistent way. Specifically, pockets of powder were placed both on the corners and on the sides of the cuboid surrounding the local model, and their volume was varied along the three Cartesian axes. The position and size of the powder pockets were used as input for the surrogate model. The target output was the temperature profile during heating and cooling under the influence of a single laser pass. A combination of PCA and PCE was used, and the performance of the surrogate model was evaluated with the RMS error on an independent test set. The surrogate model was trained and evaluated both on a randomly distributed training set and an equidistantly distributed one. The latter proved superior, showing that the surrogate model performed better when the experimental design (training set) included cases well distributed within the input database. Finally, the current surrogate model's weakness is that a powder pocket far away from the laser does not affect the temperature profile independent of the powder's volume. The current surrogate model could not always determine successfully the cut-off distance over which any geometry of powder pocket would result in the same temperature field. Last but not least, regarding the computational reduction, we should note that each FE simulation was run using 24 threads on an Intel Xeon Gold 5118 processor in the Euler cluster. The duration (wall clock time) was in average 8 and 11 hours for the runs with varying initial temperature and with powder pockets, respectively. Once the database was built the training and evaluation of the respective surrogate model required 2 to 4 minutes on a local PC.

Regarding future research on the topic, we suggest a few key points about improving the current model, mainly regarding the powder pocket database. Firstly we suggest that the size of the powder pockets should be built as multiple of the powder layer thickness, which is more relevant and realistic for the SLM process. Secondly, the current size of the database should include more cases so that the surrogate model learns to recognize the cut-off distance over which the powder volume does not affect the temperature distribution. Another recommendation is to provide additional, i.e. temporal, information as input to the surrogate model, such as the absolute distance of the center or the edges of the powder pocket to the center of the laser for each time frame in the simulation. This could help map better the dependence of the output temperature field on the position of the powder pocket. Finally, more sophisticated models (such as neural networks) could be investigated that can capture better the non-linearity introduced by the different powder configurations.

Another topic that needs to be improved is the assessment of the surrogate model. Firstly, the sources of final error measure should be isolated, namely which error is caused by the PCA or the PCE, which would help optimize the surrogate model setup. Furthermore, the performance assessment of the surrogate model should be extended. In addition to the RMS error, more physically relevant error measures could be introduced, such as the size evolution or volume evolution of the melt pool and the cooling rate in selected positions inside the local model. Regarding the RMS error, it should be calculated in a reduced area around the melt pool instead of all the nodes of the local model to provide larger but also more representative variations between the different test cases.

Once the surrogate model is optimized, more elaborate temperature profiles on the boundaries

of the local model could be introduced and not only a constant initial temperature. This could be a way of defining a heat flux input across the boundaries of the local model, which is a more realistic approach, especially since the final goal is to implement the local models in a multiscale framework.

Finally, more effort should be directed towards the validation of the FE model, i.e. by comparing the sensitivity of the simulated meltpool dimensions with experimental data for varying process parameters. On these grounds, a surrogate model could be built to predict the sensitivity of the melt pool size to experimental process parameters. The findings should be compared to the analytical dependence of melt pool width and depth to the energy density, as well as with experimental measurements both from the literature [35, 37] but also from printed Hastelloy X samples produced in *Empa*.

Bibliography

[1] E. Hosseini and V. Popovich, "A review of mechanical properties of additively manufactured inconel 718," *Additive Manufacturing*, vol. 30, p. 100877, 2019.

[2] L. Bertini, F. Bucchi, F. Frendo, M. Moda, and B. D. Monelli, "Residual stress prediction in selective laser melting," *The International Journal of Advanced Manufacturing Technology*, vol. 105, pp. 609–636, Aug. 2019.

[3] "Ongoing projects - JPL additive manufacturing center." `https://scienceandtechnology.jpl.nasa.gov/amc`. Accessed: 12.10.2020.

[4] Y. M. Arısoy, L. E. Criales, T. Özel, B. Lane, S. Moylan, and A. Donmez, "Influence of scan strategy and process parameters on microstructure and its optimization in additively manufactured nickel alloy 625 via laser powder bed fusion," *The International Journal of Advanced Manufacturing Technology*, vol. 90, no. 5-8, pp. 1393–1417, 2017.

[5] D. Gu, Y.-C. Hagedorn, W. Meiners, G. Meng, R. J. S. Batista, K. Wissenbach, and R. Poprawe, "Densification behavior, microstructure evolution, and wear performance of selective laser melting processed commercially pure titanium," *Acta Materialia*, vol. 60, no. 9, pp. 3849 – 3860, 2012.

[6] Z. Dong, Y. Liu, W. Wen, J. Ge, and J. Liang, "Effect of hatch spacing on melt pool and as-built quality during selective laser melting of stainless steel: Modeling and experimental approaches," *Materials*, vol. 12, 2018.

[7] Z. Zhang, Y. Huang, A. R. Kasinathan], S. I. Shahabad], U. Ali, Y. Mahmoodkhani, and E. Toyserkani, "3-dimensional heat transfer modeling for laser powder-bed fusion additive manufacturing with volumetric heat sources based on varied thermal conductivity and absorptivity," *Optics & Laser Technology*, vol. 109, pp. 297 – 312, 2019.

[8] M. Markl and C. Körner, "Multiscale modeling of powder bed–based additive manufacturing," *Annual Review of Materials Research*, vol. 46, no. 1, pp. 93–123, 2016.

[9] "Abaqus documentation - about adaptive remeshing." `https://abaqus-docs.mit.edu/2017/English/SIMACAEANLRefMap/simaanl-c-adpover.htm`. Accessed: 16.07.2020.

[10] P. Gh Ghanbari, E. Mazza, and E. Hosseini, "Adaptive local-global multiscale approach for thermal simulation of the selective laser melting process," *Additive Manufacturing*, vol. 36, p. 101518, 2020.

[11] A. Olleak and Z. Xi, "Efficient L-PBF process simulation using finite element modeling with adaptive remeshing for distortions and residual stresses prediction," *Manufacturing Letters*, 2020.

[12] F. Keller, "Surrogate modeling for multiscale thermal simulation of powder-bed additive manufacturing," Master Thesis, ETH Zürich and Empa, Zürich, Switzerland, 2020.

[13] J. Zhang, Y. Zhang, W. H. Lee, L. Wu, H.-H. Choi, and Y.-G. Jung, "A multi-scale multi-physics modeling framework of laser powder bed fusion additive manufacturing process," *Metal Powder Report*, vol. 73, no. 3, pp. 151 – 157, 2018.

[14] Y. Lian, Z. Gan, C. Yu, D. Kats, W. K. Liu, and G. J. Wagner, "A cellular automaton finite volume method for microstructure evolution during additive manufacturing," *Materials & Design*, vol. 169, p. 107672, 2019.

[15] *Mesoscale Multi-Physics Simulation of Solidification in Selective Laser Melting Process Using a Phase Field and Thermal Lattice Boltzmann Model*, vol. Volume 1: 37th Computers and Information in Engineering Conference of *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 08 2017. V001T02A027.

[16] T. M. Rodgers, J. D. Madison, and V. Tikare, "Simulation of metal additive manufacturing microstructures using kinetic monte carlo," *Computational Materials Science*, vol. 135, pp. 78 – 89, 2017.

[17] L. Ladani, J. Romano, W. Brindley, and S. Burlatsky, "Effective liquid conductivity for improved simulation of thermal transport in laser beam melting powder bed technology," *Additive Manufacturing*, vol. 14, pp. 13 – 23, 2017.

[18] M. Vohra, P. Nath, S. Mahadevan, and Y.-T. Tina Lee, "Fast surrogate modeling using dimensionality reduction in model inputs and field output: Application to additive manufacturing," *Reliability Engineering & System Safety*, vol. 201, p. 106986, 2020.

[19] C. Li, J. Liu, X. Fang, and Y. Guo, "Efficient predictive model of part distortion and residual stress in selective laser melting," *Additive Manufacturing*, vol. 17, pp. 157 – 168, 2017.

[20] C. Li, C. Fu, Y. Guo, and F. Fang, "A multiscale modeling approach for fast prediction of part distortion in selective laser melting," *Journal of Materials Processing Technology*, vol. 229, pp. 703 – 712, 2016.

[21] C. Li, J. Liu, and Y. Guo, "Prediction of residual stress and part distortion in selective laser melting," *Procedia CIRP*, vol. 45, pp. 171 – 174, 2016. 3rd CIRP Conference on Surface Integrity.

[22] H. Ali, H. Ghadbeigi, and K. Mumtaz, "Effect of scanning strategies on residual stress and mechanical properties of selective laser melted ti6al4v," *Materials Science and Engineering: A*, vol. 712, pp. 175 – 187, 2018.

[23] S. Marelli and B. Sudret, "UQLab: a framework for uncertainty quantification in MATLAB," in *Proc. 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014), Liverpool, United Kingdom*, 2014.

[24] "The uqlab user manuals." `https://www.uqlab.com/user-manuals`. Accessed: 08.12.2020.

[25] F. W. Baumann, A. Sekulla, M. Hassler, B. Himpel, and M. Pfeil, "Trends of machine learning in additive manufacturing," *International Journal of Rapid Manufacturing*, vol. 7, no. 4, p. 310, 2018.

[26] J. Shlens, "A tutorial on principal component analysis," *arXiv preprint arXiv:1404.1100*, 2014.

[27] "Use abaqus to simulate additive manufacturing – printing a hip implant.." https://info.simuleon.com/blog/using-abaqus-to-simulate-additive-manufacturing-printing-an-optimized-hip-implant. Accessed: 15.12.2020.

[28] *ABAQUS/Standard User's Manual, Version 2019*. United States: Dassault Systèmes Simulia Corp, 2019.

[29] "Altair-hyperworks." https://www.altair.com/hyperworks/. Accessed: 15.12.2020.

[30] Z. Luo and Y. Zhao, "Efficient thermal finite element modeling of selective laser melting of inconel 718," *Computational Mechanics*, vol. 65, pp. 763–787, Nov. 2019.

[31] Y. Zhou, Z. Zhang, Y. Wang, G. Liu, S. Zhou, Y. Li, J. Shen, and M. Yan, "Selective laser melting of typical metallic materials: An effective process prediction model developed by energy absorption and consumption analysis," *Additive Manufacturing*, vol. 25, pp. 204 – 217, 2019.

[32] "hastelloy® x alloy - haynes international ag." http://www.haynes.ch/doc/HASTELLOY_X.pdf. Accessed: 20.12.2020.

[33] "Haynes® 718 alloy - haynes international." https://www.haynesintl.com/alloys/alloy-portfolio_/High-temperature-Alloys/haynes718-alloy/principal-features.aspx. Accessed: 20.12.2020.

[34] J. Goldak, A. Chakravarti, and M. Bibby, "A new finite element model for welding heat sources," *Metallurgical Transactions B*, vol. 15, pp. 299–305, June 1984.

[35] S. I. Shahabad, Z. Zhang, A. Keshavarzkermani, U. Ali, Y. Mahmoodkhani, R. Esmaeilizadeh, A. Bonakdar, and E. Toyserkani, "Heat source model calibration for thermal analysis of laser powder-bed fusion," *The International Journal of Advanced Manufacturing Technology*, vol. 106, pp. 3367–3379, Jan. 2020.

[36] J. J. Valenica and P. N. Quested, *Thermophysical properties*, pp. 468–481. ASM Handbook Committee, 2008.

[37] A. Keshavarzkermani, E. Marzbanrad, R. Esmaeilizadeh, Y. Mahmoodkhani, U. Ali, P. D. Enrique, N. Y. Zhou, A. Bonakdar, and E. Toyserkani, "An investigation into the effect of process parameters on melt pool geometry, cell spacing, and grain refinement during laser powder bed fusion," *Optics & Laser Technology*, vol. 116, pp. 83–91, Aug. 2019.

[38] A. Gusarov, M. Pavlov, and I. Smurov, "Residual stresses at laser surface remelting and additive manufacturing," *Physics Procedia*, vol. 12, pp. 248 – 254, 2011. Lasers in Manufacturing 2011 - Proceedings of the Sixth International WLT Conference on Lasers in Manufacturing.

[39] J. Dilip, S. Zhang, C. Teng, K. Zeng, C. Robinson, D. Pal, and B. Stucker, "Influence of processing parameters on the evolution of melt pool, porosity, and microstructures in ti-6al-4v alloy parts fabricated by selective laser melting," *Progress in Additive Manufacturing*, vol. 2, no. 3, pp. 157–167, 2017.

[40] M. Gouge and P. Michaleris, "Chapter 1 - an introduction to additive manufacturing processes and their modeling challenges," in *Thermo-Mechanical Modeling of Additive Manufacturing* (M. Gouge and P. Michaleris, eds.), pp. 3 − 18, Butterworth-Heinemann, 2018.

A Appendix

A.1 Input files for runs with powder pockets

The description of the powder pocket position, which was used as input for the surrogate model, was given by six values. The coordinates of the center of the rectangular containing the powder pocket (x_c, y_c, z_c) and its sides x_{size}, y_{size}, z_{size}.

Table 5: Powder volume varied along the x-axis

Run ID	x_c [mm]	y_c [mm]	z_c [mm]	x_{size} [mm]	y_{size} [mm]	z_{size}[mm]
1	-0.72	0.0	-0.5	0.56	2	1
2	-0.58	0.0	-0.5	0.84	2	1
3	-0.44	0.0	-0.5	1.12	2	1
4	-0.3	0.0	-0.5	1.4	2	1
5	-0.16	0.0	-0.5	1.68	2	1
6	-0.02	0.0	-0.5	1.96	2	1
7	0.12	0.0	-0.5	2.24	2	1
8	1.52	0.0	-0.5	0.56	2	1
9	1.38	0.0	-0.5	0.84	2	1
10	1.24	0.0	-0.5	1.12	2	1
11	1.1	0.0	-0.5	1.4	2	1
12	0.96	0.0	-0.5	1.68	2	1
13	0.82	0.0	-0.5	1.96	2	1
14	0.68	0.0	-0.5	2.24	2	1

Table 6: Powder volume varied along the y-axis

Run ID	x_c [mm]	y_c [mm]	z_c [mm]	x_{size} [mm]	y_{size} [mm]	z_{size}[mm]
1	-0.4	-0.8	-0.5	2.8	0.4	1
2	-0.4	-0.7	-0.5	2.8	0.6	1
3	-0.4	-0.6	-0.5	2.8	0.8	1
4	-0.4	-0.5	-0.5	2.8	1.0	1
5	-0.4	-0.4	-0.5	2.8	1.2	1
6	-0.4	-0.3	-0.5	2.8	1.4	1
7	-0.4	-0.2	-0.5	2.8	1.6	1
8	-0.4	0.8	-0.5	2.8	0.4	1
9	-0.4	0.7	-0.5	2.8	0.6	1
10	-0.4	0.6	-0.5	2.8	0.8	1
11	-0.4	0.5	-0.5	2.8	1.0	1
12	-0.4	0.4	-0.5	2.8	1.2	1
13	-0.4	0.3	-0.5	2.8	1.4	1
14	-0.4	0.2	-0.5	2.8	1.6	1

Table 7: Powder pockets volume varied along the z-axis

Run ID	x_c [mm]	y_c [mm]	z_c [mm]	x_{size} [mm]	y_{size} [mm]	z_{size}[mm]
1	0.4	0.0	-0.9	2.8	2	0.2
2	0.4	0.0	-0.85	2.8	2	0.3
3	0.4	0.0	-0.8	2.8	2	0.4
4	0.4	0.0	-0.75	2.8	2	0.5
5	0.4	0.0	-0.7	2.8	2	0.6
6	0.4	0.0	-0.65	2.8	2	0.7
7	0.4	0.0	-0.6	2.8	2	0.8
8	0.4	0.0	-0.1	2.8	2	0.2
9	0.4	0.0	-0.15	2.8	2	0.3
10	0.4	0.0	-0.2	2.8	2	0.4
11	0.4	0.0	-0.25	2.8	2	0.5
12	0.4	0.0	-0.3	2.8	2	0.6
13	0.4	0.0	-0.35	2.8	2	0.7
14	0.4	0.0	-0.4	2.8	2	0.8

(a) Run ID 1   (b) Run ID 2   (c) Run ID 3

(d) Run ID 4   (e) Run ID 5   (f) Run ID 6

(g) Run ID 7   (h) Run ID 8   (i) Run ID 9

(j) Run ID 10   (k) Run ID 11   (l) Run ID 12

(m) Run ID 13   (n) Run ID 14
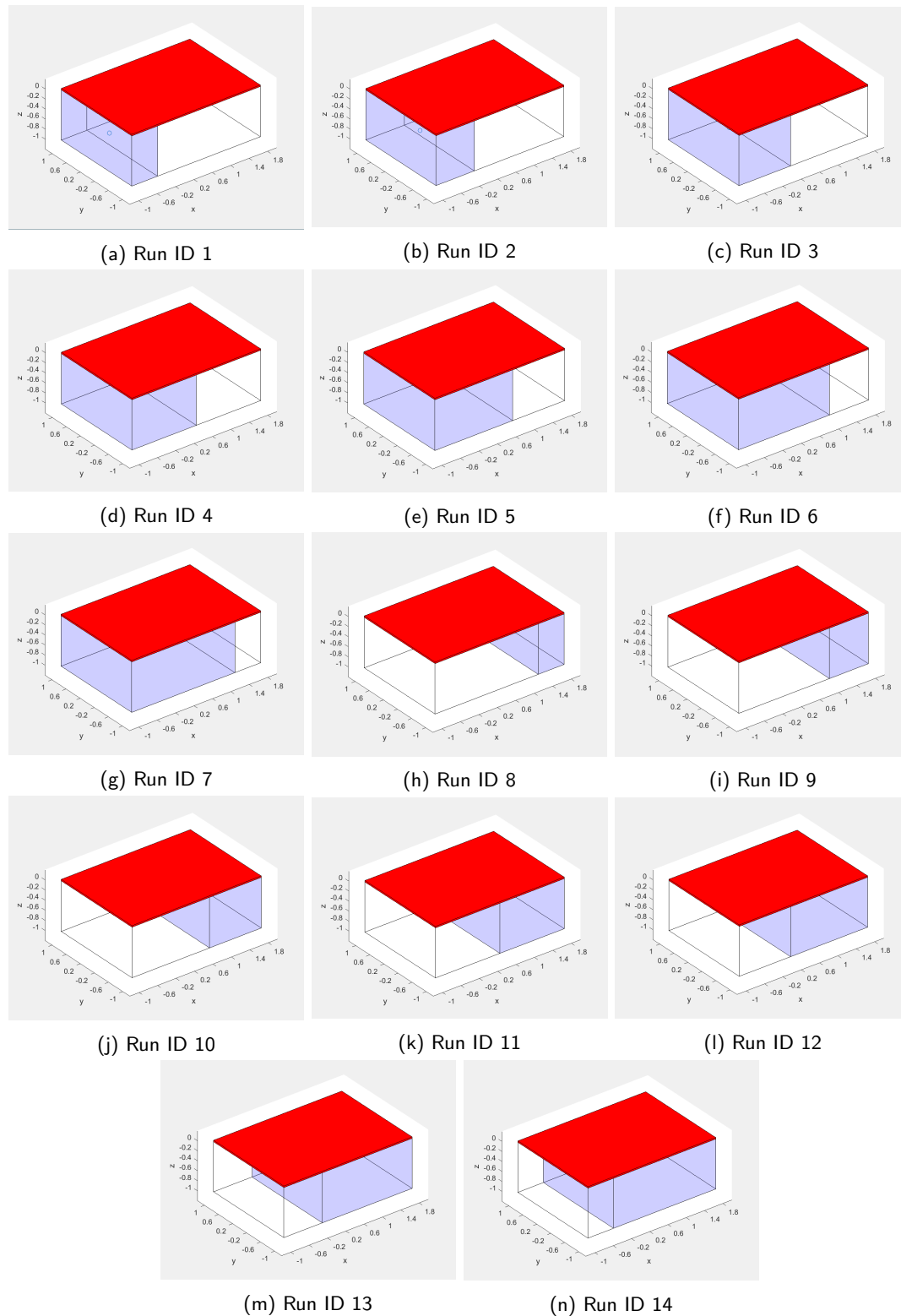
Figure 27: Powder pocket volume varied along the x-axis.

A.2 User-selected experimental design

Here, instead of dividing the database into training and test set using random splits, we hand-picked the runs in the training set so that they are equidistantly distributed in terms of the input values in the respective surrogate models. The runs that were used in the user-defined training sets (experimental design) are listed in Table 9. The test sets included the rest of the runs from

(a) Run ID 1   (b) Run ID 2   (c) Run ID 3

(d) Run ID 4   (e) Run ID 5   (f) Run ID 6

(g) Run ID 7   (h) Run ID 8   (i) Run ID 9

(j) Run ID 10   (k) Run ID 11   (l) Run ID 12

(m) Run ID 13   (n) Run ID 14
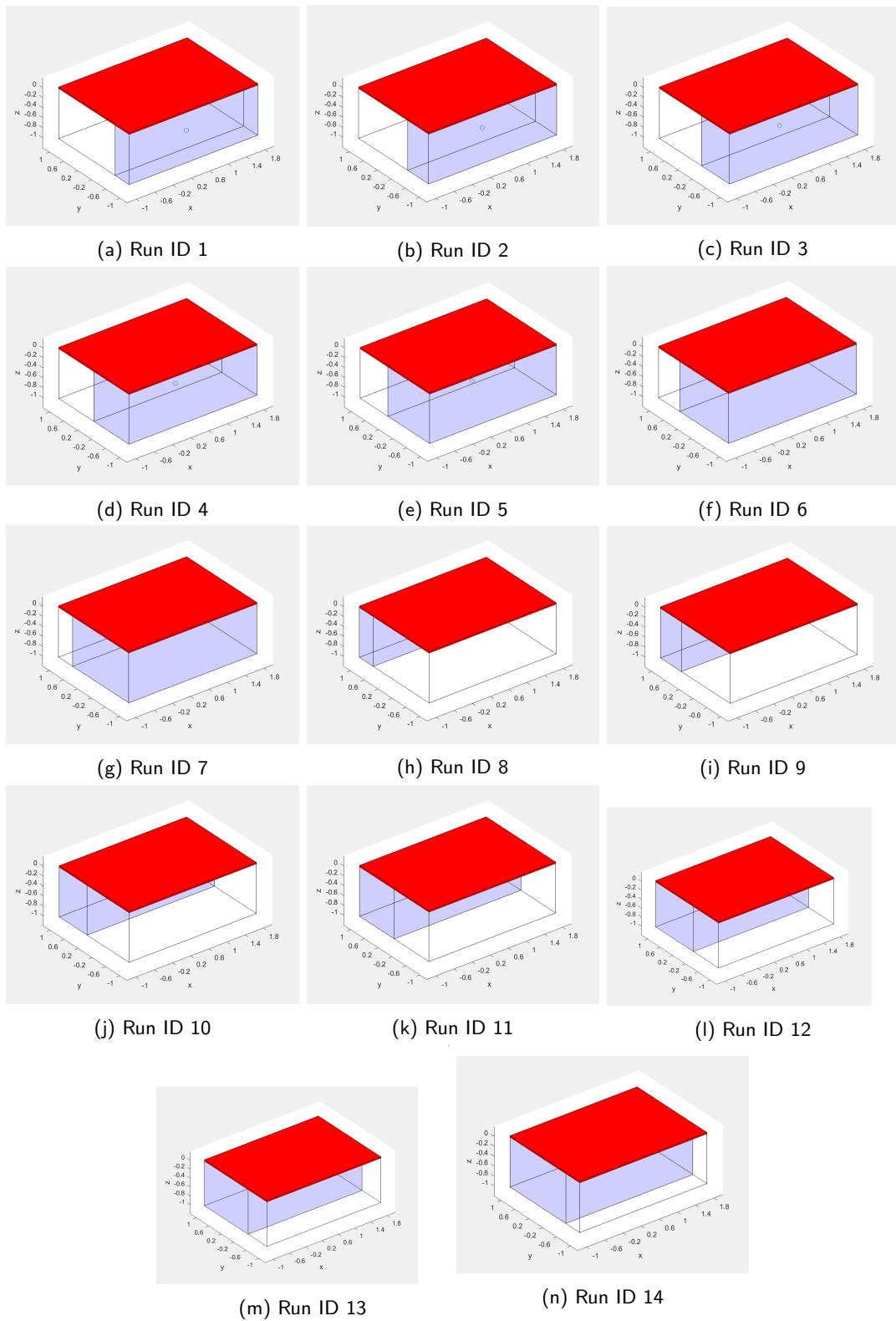
Figure 28: Powder pocket volume varied along the y-axis.

the original dataset after constructing the respective experimental design.

(a) Run ID 1   (b) Run ID 2   (c) Run ID 3

(d) Run ID 4   (e) Run ID 5   (f) Run ID 6

(g) Run ID 7   (h) Run ID 8   (i) Run ID 9

(j) Run ID 10   (k) Run ID 11   (l) Run ID 12

(m) Run ID 13   (n) Run ID 14

Figure 29: Powder pocket volume varied along the z-axis.

Table 8: Powder pockets on the corners of the local model.

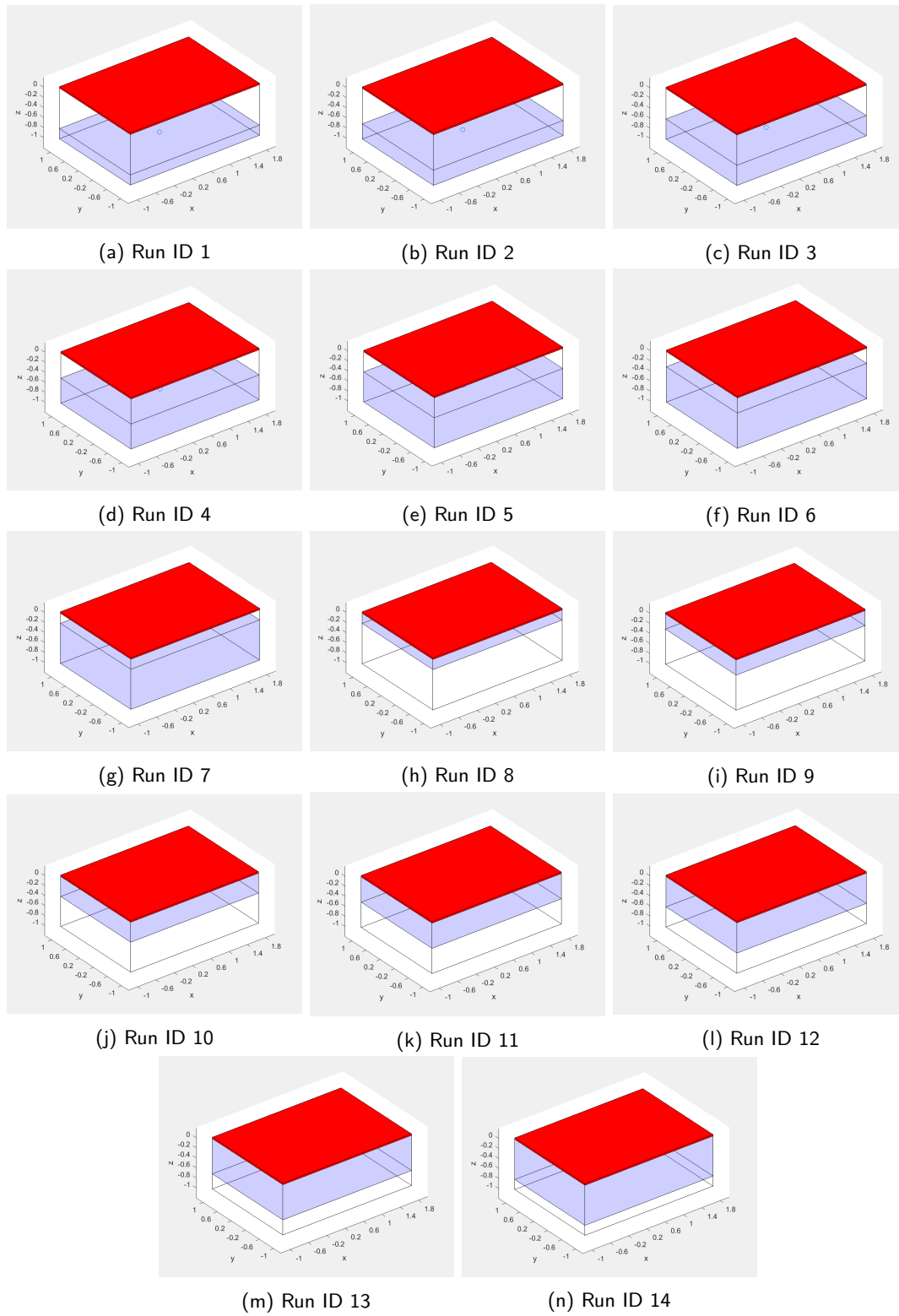| Run ID	x_c [mm]	y_c [mm]	z_c [mm]	x_{size} [mm]	y_{size} [mm]	z_{size} [mm]
Run ID 1	-0.18	-0.23	-0.65	1.64	1.54	0.7
Run ID 2	-0.48	-0.43	-0.64	1.04	1.14	0.72
Run ID 3	-0.28	-0.39	-0.76	1.44	1.22	0.48
Run ID 4	-0.07	-0.33	-0.6	1.86	1.34	0.8
Run ID 5	-0.36	-0.3	-0.25	1.28	1.4	0.5
Run ID 6	-0.04	-0.24	-0.15	1.92	1.52	0.3
Run ID 7	-0.24	-0.25	-0.33	1.52	1.5	0.66
Run ID 8	-0.01	-0.33	-0.2	1.98	1.34	0.4
Run ID 9	0.8	-0.53	-0.61	2	0.94	0.78
Run ID 10	0.87	-0.4	-0.73	1.86	1.2	0.54
Run ID 11	1.07	-0.44	-0.61	1.46	1.12	0.78
Run ID 12	1.07	-0.29	-0.77	1.46	1.42	0.46
Run ID 13	0.98	-0.37	-0.28	1.64	1.26	0.56
Run ID 14	1.02	-0.3	-0.23	1.56	1.4	0.46
Run ID 15	1.18	-0.46	-0.37	1.24	1.08	0.74
Run ID 16	0.93	-0.51	-0.27	1.74	0.98	0.54
Run ID 17	0.01	0.46	-0.64	2.02	1.08	0.72
Run ID 18	0.07	0.45	-0.78	2.14	1.1	0.44
Run ID 19	0.05	0.2	-0.75	2.1	1.6	0.5
Run ID 20	-0.09	0.44	-0.75	1.82	1.12	0.5
Run ID 21	-0.01	0.32	-0.24	1.98	1.36	0.48
Run ID 22	0.04	0.37	-0.39	2.08	1.26	0.78
Run ID 23	0.07	0.38	-0.36	2.14	1.24	0.72
Run ID 24	-0.56	0.37	-0.39	0.88	1.26	0.78
Run ID 25	0.73	0.22	-0.66	2.14	1.56	0.68
Run ID 26	1.06	0.29	-0.76	1.48	1.42	0.48
Run ID 27	0.9	0.23	-0.69	1.8	1.54	0.62
Run ID 28	0.94	0.21	-0.63	1.72	1.58	0.74
Run ID 29	1.39	0.32	-0.36	0.82	1.36	0.72
Run ID 30	0.82	0.48	-0.32	1.96	1.04	0.64
Run ID 31	0.72	0.48	-0.26	2.16	1.04	0.52
Run ID 32	0.9	0.44	-0.32	1.8	1.12	0.64

Table 9: User-defined experimental design

Surrogate model	Size of experimental design	ID of runs in the experimental design
Initial temperature variation	3	T={25,250,500}
	4	T={25,150,350,500}
	5	T={25,150,250,350,500}
	6	T={25,100,200,300,400,500}
Powder volume varied along x-axis	6	Run ID={1,4,7,8,11,14}
	8	Run ID={1,3,5,7,8,10,12,14}
	10	Run ID={1,3,4,5,7,8,9,11,13,14}
Powder volume varied along y-axis	6	Run ID={1,4,7,8,11,14}
	8	Run ID={1,3,5,7,8,10,12,14}
	10	Run ID={1,3,4,5,7,8,9,11,13,14}
Powder volume varied along z-axis	6	Run ID={1,4,7,8,11,14}
	8	Run ID={1,3,5,7,8,10,12,14}
	10	Run ID={1,3,4,5,7,8,9,11,13,14}

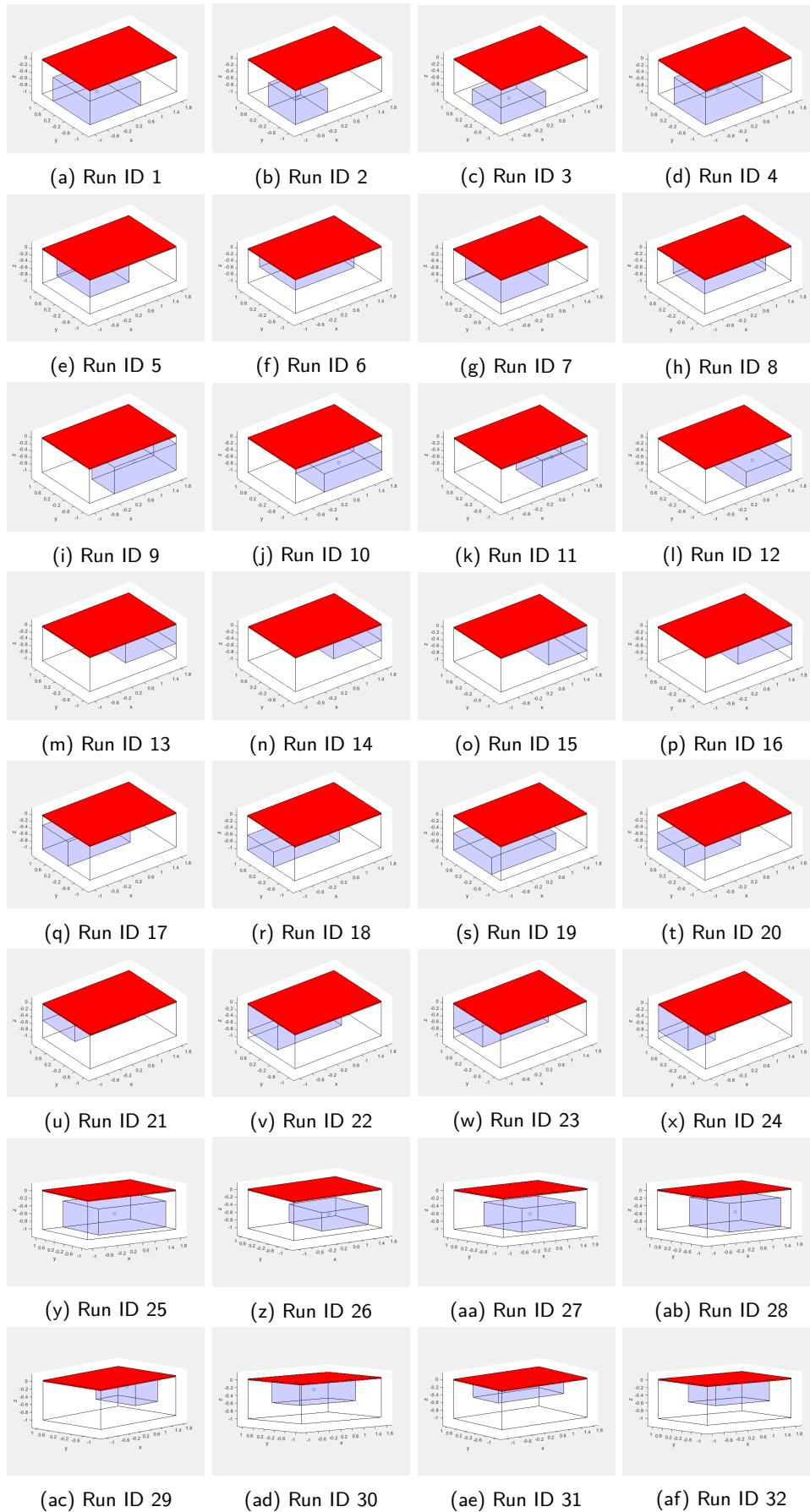|			
(a) Run ID 1	(b) Run ID 2	(c) Run ID 3	(d) Run ID 4
(e) Run ID 5	(f) Run ID 6	(g) Run ID 7	(h) Run ID 8
(i) Run ID 9	(j) Run ID 10	(k) Run ID 11	(l) Run ID 12
(m) Run ID 13	(n) Run ID 14	(o) Run ID 15	(p) Run ID 16
(q) Run ID 17	(r) Run ID 18	(s) Run ID 19	(t) Run ID 20
(u) Run ID 21	(v) Run ID 22	(w) Run ID 23	(x) Run ID 24
(y) Run ID 25	(z) Run ID 26	(aa) Run ID 27	(ab) Run ID 28
(ac) Run ID 29	(ad) Run ID 30	(ae) Run ID 31	(af) Run ID 32

Figure 30: Powder pocket volume placed on the corners of the local model.

B Appendix
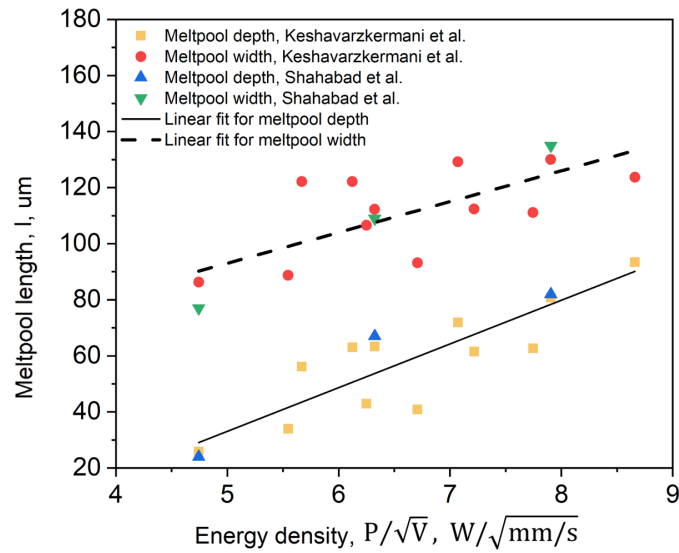
B.1 Linear fit on experimental melt pool dimensions

Figure 31: Experimental melt pool dimensions as a function of the energy density.

B.1.1 Melt pool depth

Equation	y = a + b*x
Data	Melt pool depth
Weight	No Weighting
Intercept	38.04307 ± 19.02085
Slope	10.99554 ± 2.84272
Residual Sum of Squares	1990.8753
Pearson's r	0.73148
R-Square (COD)	0.53507
Adj. R-Square	0.49931

Figure 32: Linear fitting parameters for melt pool depth

B.1.2 Melt pool width

Equation	y = a + b*x
Data	Melt pool width
Weight	No Weighting
Intercept	-44.75044 ± 15.99975
Slope	15.57168 ± 2.39121
Residual Sum of Squares	1408.67454
Pearson's r	0.87486
R-Square (COD)	0.76537
Adj. R-Square	0.74732

Figure 33: Linear fitting parameters for melt pool width.