



HAL
open science

Data-Driven Modeling for Multiphysics Parametrized Problems-Application to Induction Hardening Process

Khouloud Derouiche, Sevan Garois, Victor Champaney, Monzer Daoud, Khalil Traidi, Francisco Chinesta

► **To cite this version:**

Khouloud Derouiche, Sevan Garois, Victor Champaney, Monzer Daoud, Khalil Traidi, et al.. Data-Driven Modeling for Multiphysics Parametrized Problems-Application to Induction Hardening Process. *Metals*, 2021, 11 (5), pp.738. 10.3390/met11050738 . hal-03288060

HAL Id: hal-03288060

<https://hal.science/hal-03288060v1>

Submitted on 16 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data-Driven Modeling for Multiphysics Parametrized Problems-Application to Induction Hardening Process

Khoulood Derouiche ^{1,2,†}, Sevan Garois ^{1,2,†} , Victor Champany ¹, Monzer Daoud ², Khalil Traidi ³ and Francisco Chinesta ^{1,*}

¹ PIMM Lab, Arts et Metiers Institute of Technologie, 151 Boulevard de l'Hôpital, 75013 Paris, France; khoulood.derouiche@ensam.eu (K.D.); sevan.garois@ensam.eu (S.G.); victor.champany@ensam.eu (V.C.)

² French Technological Research Institute for Materials, Metallurgy and Processes (IRT- M2P), 4 rue Augustin Fresnel, 57070 Metz, France; monzer.daoud@irt-m2p.fr

³ Safran Tech, Rue des Jeunes Bois, 78117 Châteaufort, France; khalil.traidi@safrangroup.com

* Correspondence: francisco.chinesta@ensam.eu

† These authors contributed equally to this work.

Abstract: Data-driven modeling provides an efficient approach to compute approximate solutions for complex multiphysics parametrized problems such as induction hardening (IH) process. Basically, some physical quantities of interest (QoI) related to the IH process will be evaluated under real-time constraint, without any explicit knowledge of the physical behavior of the system. Hence, computationally expensive finite element models will be replaced by a parametric solution, called metamodel. Two data-driven models for temporal evolution of temperature and austenite phase transformation, during induction heating, were first developed by using the proper orthogonal decomposition based reduced-order model followed by a nonlinear regression method for temperature field and a classification combined with regression for austenite evolution. Then, data-driven and hybrid models were created to predict hardness, after quenching. It is shown that the results of artificial intelligence models are promising and provide good approximations in the low-data limit case.

Keywords: data-driven modeling; induction hardening; metamodel; proper orthogonal decomposition; artificial intelligence; hybrid model

1. Introduction

The aeronautical and automotive industries would eventually like to lighten the mechanical structures in order to meet the relative environmental concerns to carbon dioxide emissions and fuel economy while maintaining optimal mechanical properties and performances. The heat treatment provides a valuable solution to give the material its optimal microstructures and mechanical properties, corresponding to the various predefined performance criteria [1]. However, for many applications, only superficial layer material properties play an important role. In this context, surface treatments of industrial components by mechanical, thermal, or thermochemical means are particularly suitable [2,3]. These techniques are mainly used to improve the fatigue strength and the resistance to the imposed external mechanical loads by changing the properties of the critical zones.

Induction hardening (IH) is one of the most appealing surface heat treatment processes widely employed in automotive and aerospace industries [4,5]. Two main steps are at the origin of the IH: an electromagnetic induction heating and subsequent quenching. It has the advantage of providing a very short surface heat-up time, a good fatigue performance, a precise control of the treated zone, a good reproducibility, and an operating mode compatible with severe environmental requirements, unlike thermochemical treatments such as carburizing and carbonitriding [6]. In the published literature, many research works based on analytical and numerical methods were carried out on the IH process [7–10]. From the industrial point of view, the challenges behind the development of such methodologies are:

- Reducing development time;
- Reducing costs by “Do It Right The First Time (DRIFT)” (Reduction of the NRC “Non Recurring Costs” such as the costs of tools and prototypes and the RC “Recurring Cost” such as production costs);
- Understanding the involved physical mechanisms for better controlling the influencing parameters and thus being able to better define the characteristics that could be controlled;
- Anticipating the management of nonquality risks (e.g., being able to reinforce the critical zones predicted by the calculations, anticipate possible defects, etc.).

The main difficulty behind the process control and the optimization in terms of time and cost is the multidisciplinary nature of the process since it combines multiphysics (electromagnetism, metallurgical, thermal, and mechanical fields) in addition to the large number of process parameters; owing to the advanced numerical simulation tools such as FORGE[®], which is used in this study, modeling the IH process and solving the physical equations behind it is possible by using some conventional discretization methods such as the well-known and largely experienced finite element (FE) method [11]. However, passing through those methods to solve multiphysics parametrized problems is often regarded as a key issue. In practice, the need to evaluate multiple configurations (choices of parameters), before converging to an adequate result that optimizes the process performance, may make the optimization procedure very time consuming and computationally expensive since each configuration takes several hours between preparing the data setting and executing the simulation.

In order to alleviate such issues, a data-driven parametric metamodeling approach for IH process simulation constitutes an appealing alternative to standard discretization techniques given the interesting compromise in terms of execution speed, computational cost, and results accuracy. This approach enables computing the parametric-based solution from sampled data collected from FE simulation codes. The main challenge behind this work is to succeed in building a robust metamodel for some physical quantities of interest (thermal and metallurgical history as well as the hardness), by using a reduced amount of data.

This paper is decomposed into two main parts; in the first part, two models were created to predict the temporal evolution of temperature and austenite phase transformation in a spur gear of C45 steel, during induction heating and under the effect of process parameters, in almost real-time. Hence, the goal is to replace the initial physical models implemented in the FE simulation code FORGE[®] by a metamodel with sufficient accuracy and short computational time. To achieve this goal, two different approaches were considered for each QoI; however, a common point between them is the fact that both techniques rely only on data (synthetic data provided by the high-fidelity multiphysics simulation) and do not require any knowledge of the full-order formulation or intervention in the numerical FE codes. It is worth pointing out that synthetic data was collected at some sparse points in the space domain and for different values of input parameters associated with the Latin Hypercube Sampling (LHS) design of experiments [12]. For the temperature evolution, the proper orthogonal decomposition (POD), which is a widespread technique in the reduced order modeling community for the study of high dimensional problems [13], was used and then followed by a nonlinear regression model. For the austenite phase evolution, a classification combined with a regression model was applied.

The second part of this work focuses on the quenching process. Choosing a full data-driven approach, the goal is to predict the hardness at the final state of the process. Therefore, a classic machine learning pipeline to select variables and create new ones based on the evolution of the temperature of the system was applied. By using XGboost [14], a tree-boosted model was made and was able to predict, in the bulk material of the gear, the hardness. A second approach is the hybrid model where the obtained austenite phase ratio was considered to increase the knowledge for a better prediction of hardness since they seem to be well correlated. This new precision allowed for the prediction of the hardness

in the whole gear, in a global way, and for the suggestion that the hybrid approach may be the solution in any field with a lack of data and having complex systems.

The rest of the paper is organized as follows: Section 2 presents the data-driven modeling for induction heating. In this section, the process and the data generation will be first presented. Then, the applied methods, the model results, and discussions for temperature and austenite phase evolution will be illustrated. Section 3 highlights the modeling for hardness prediction with both hybrid and fully data-driven approach and Section 4 addresses the main conclusions.

2. Data-Driven Modeling for Induction Heating Process

2.1. Process and Data

Induction heating is becoming one of the preferred heating technologies at the industrial scale [15] due to its numerous advantages. Basically, it consists in applying an alternating current (AC) to a copper coil surrounding a conducting workpiece, an eddy current will be produced inside this later by the AC-based magnetic field, and consequently, the workpiece will be heated under the Joule effect. Induction heating combines electromagnetic and thermal fields in addition to the phase transformation associated with an increase in temperature. These physical phenomena can be simulated by several FE codes, in which physical models described by partial differential equations (PDE) are solved. In this paper, the workpiece is C45 steel spur gear of module 4 and 13 teeth. The dimension of the gear and the inductor is presented in Table 1.

Table 1. Dimension of the gear and the inductor.

Components	Characteristics
Gear	Tip diameter: 60 mm Root diameter: 48 mm Width: 10 mm Tooth thickness: 5 mm
Inductor	Rectangular section: 8 × 5 mm Inner diameter: 65 mm

In order to optimize and improve induction heating performances applied to the workpiece, multiple parameters can be taken into consideration such as material behavior, process parameters, and geometrical features. As a first step of this study, two important process parameters were considered while the other ones were kept constant. The selected parameters and their lower and upper limits are shown in Table 2.

Table 2. Input parameters and their lower and upper limits.

Input Parameters	Lower Limit	Upper Limit
Frequency (kHz)	10	250
Power (kW)	50	600

Changing these input parameters could highly impact the process results. In order to evaluate this effect, two important physical quantities such as temporal evolution of temperature and austenite phase transformation were taken into consideration.

Particularly, a model for each one of these quantities will be constructed. To achieve this objective, a set of precomputed high-fidelity solutions (called snapshots) are collected by solving the original full-order model for different values of input parameters using the commercial finite element software FORGE[®]. It is worth mentioning that the model has two symmetry planes and hence only half of the gear tooth was modeled to improve the computational efficiency. The implemented material properties of C45 were taken from literature and JmatPro database [16–20]. To simulate heating, FORGE[®] uses two

coupled solvers. The electromagnetic solver solves Maxwell's equations and consequently gives rise to a heating power. The latter is applied to the thermal solver to obtain the temperature field by solving the heat equation. More details can be found in [21]. Several FE simulations were carried out according to the LHS design of experiments. Despite generating a quasirandom sampling distribution, LHS guarantees a good coverage of each dimension space. According to the LHS, a total of 20 simulations have been generated as shown in Figure 1. Therefore, it is worth seeing how well the applied nonlinear regression technique performs with a reduced number of data.

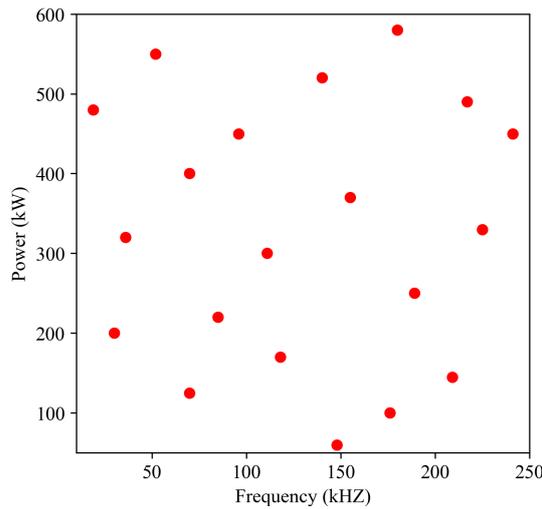


Figure 1. Latin Hypercube Sampling design of experiments.

As mentioned above, all process parameters except frequency and power were kept constant during heating. Therefore, process time, which is one of the most important parameters, was first chosen to be equal to 1 sec for all snapshots, where 1 sec is considered enough for such a rapid heating process. However, higher values of power and frequency give rise to higher temperature (more than 1000 °C) which can be reached very fast. Hence, such process time induces the divergence of the numerical computation even before reaching the final time step. To alleviate such issue, different process times were considered in order to reach an acceptable level of temperature.

For this part of the work, the temperature and austenite evolutions were extracted at 14 specific points representing the main heat-affected zones as shown in Figure 2. A model for each spatial point was constructed. It is worth mentioning that the skin effect by induction heating is dependent on the frequency and consequently different thicknesses of the heated regions will be obtained.

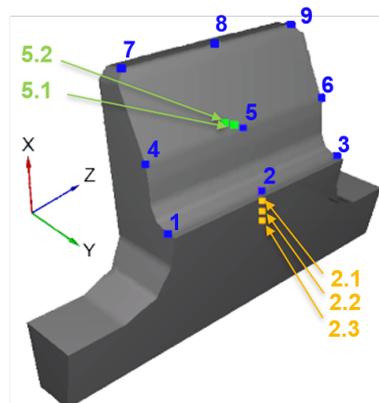


Figure 2. Measurement points (sensors).

2.2. Modeling the Time Evolution of Temperature

2.2.1. Methodology

Based on the set of computed snapshots, collected at some sparse points in the space domain and for different values of frequency and power, a dimensionality reduction approach based on POD [13] followed by a machine learning algorithm for regression was applied in order to create an approximate solution for the temperature field evolution. However, different data-lengths were obtained by using different process times. First, a normalization and a uniform discretization with respect to the final process time for each snapshot were applied in order to transfer all quantities (time, temperature, austenite, etc.) into the same discretization; however, this method introduced pollution in the regression. To overcome this issue, all the data was truncated when reaching a constant temperature and noting that the temperature evolution is monotonic, the role of time and temperature was inversed. Therefore, the time was modeled as a function of the temperature, where the maximum value of the latter was kept the same for all the experiments. Basically, a constant final temperature was chosen exclusively for each measurement point depending on its position. Despite having a constant final temperature, all the data does not have the same length; therefore, a uniform discretization was applied to data. Once the new datasets for time as a function of the temperature are established, the POD-based reduced order model is applied, followed by a nonlinear regression technique.

As a first step, a POD reduced basis was built, onto which the initial FE solution was projected and consequently a much lower dimension of the initial problem was obtained. To achieve this dimensionality reduction, consider the set of the N_s snapshots ($N_s = 20$) collected at S measurement points ($S = 14$), $\mathbf{t}_i^j = \mathbf{t}^j(F_i, P_i)_{i=1, \dots, N_s; j=1, \dots, S} \in \mathbb{R}^N$, and computed by solving the full-order FE model for different values of the input parameters, where F , P , and N are the frequency, the power, and the number of time-temperature points, respectively. The snapshot matrix $\mathbf{M} \in \mathbb{R}^{N \times N_s}$ for each measurement point is defined such that $\mathbf{M}^j = [\mathbf{t}_1^j \mathbf{t}_2^j \dots \mathbf{t}_{N_s}^j]$. To obtain the reduced basis, the singular value decomposition (SVD) is applied to \mathbf{M}^j as follows:

$$\mathbf{M}^j = \mathbf{U}^j \boldsymbol{\Sigma}^j \mathbf{V}^{jT} \quad \text{for } j = 1, \dots, S \quad (1)$$

where $\mathbf{U}^j \in \mathbb{R}^{N \times N}$ and $\mathbf{V}^j \in \mathbb{R}^{N_s \times N_s}$ are orthogonal matrices containing the left and right singular vectors of \mathbf{M}^j , respectively. $\boldsymbol{\Sigma}^j \in \mathbb{R}^{N \times N_s}$ is a rectangular diagonal matrix containing the singular values of \mathbf{M}^j sorted in a decreasing order. The reduced POD basis, $\mathbf{B}^j = [\boldsymbol{\phi}_1^j, \boldsymbol{\phi}_2^j, \dots, \boldsymbol{\phi}_R^j]$, is defined as the first R left singular vectors of \mathbf{M}^j (i.e., first R columns of \mathbf{U}^j) corresponding to the R largest singular values. Thus, singular values provide a quantitative guidance to choose the size of the POD basis. In practice, POD provides an efficient representation of the snapshot data in a low-dimensional subspace of dimension R , much lower than N , such that

$$\mathbf{t}_i^j \approx \sum_{k=1}^R \boldsymbol{\phi}_k^j \alpha_{ki}^j = \mathbf{B}^j \mathbf{a}_i^j \quad \text{for } j = 1, \dots, S \quad i = 1, \dots, N_s \quad (2)$$

where $\boldsymbol{\phi}_k^j$ and α_{ki}^j are called POD modes and POD modal coefficients, respectively. In matrix form, $\mathbf{B}^j \in \mathbb{R}^{N \times R}$ and $\mathbf{a}_i^j = (\mathbf{B}^j)^T \mathbf{t}_i^j$ where $\mathbf{a}_i^j \in \mathbb{R}^R$. Now, instead of using N_s snapshots \mathbf{t}_i^j of dimension N to fit a model to the data, the low-dimensional representation of the initial snapshots was considered. Then, the reconstruction of the solution was achieved using Equation (2).

Then, as a second step, the reduced state vectors of the snapshots data, the so-called POD modal coefficients, were considered and a nonlinear regression method called sparse Proper Generalized Decomposition (sPGD) [22], based on the separated representation approach enabling quite rich approximations for high dimensional problems in a low-data

limit [23], was used to fit the low dimensional POD modal coefficients. More details about sPGD can be found in Appendix A.

2.2.2. Results and Discussion

The POD modes and modal coefficients are computed at each measurement point as mentioned before. The left singular vectors of the snapshot matrices are truncated to the two first singular vectors which correspond to the POD modes. This choice is made in accordance with the singular values decay as shown in Figure 3. As it can be seen, more than 90% of the variance is retained with two singular values.

It is worth pointing out that only results for points (#1, #5, #7, and #2.2) in reference to Figure 2 are shown for the sake of clarity.

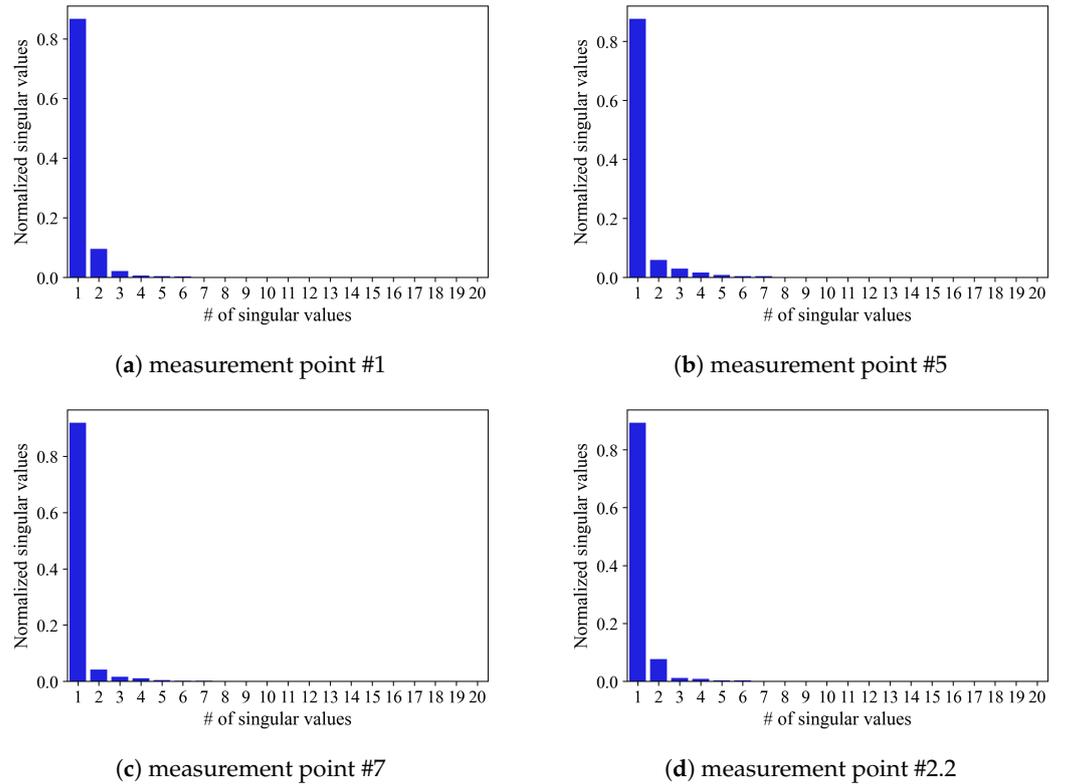
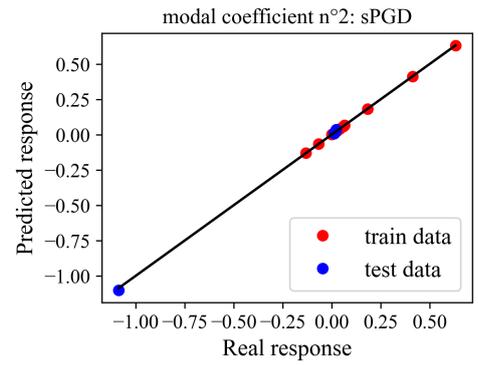
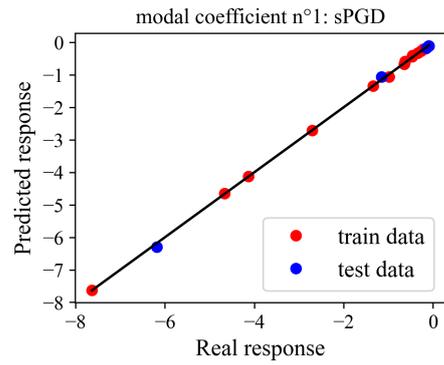
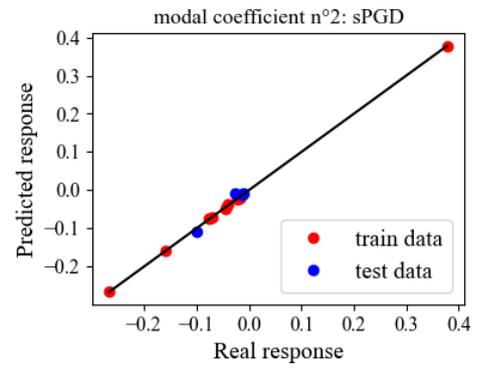
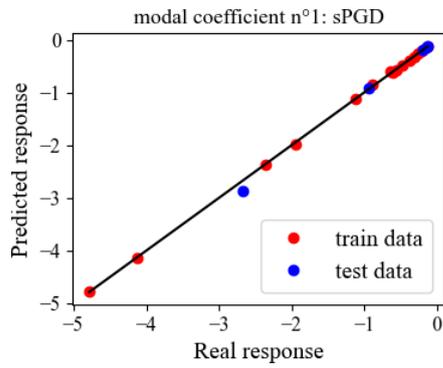


Figure 3. Normalized singular values of the thermal field for 4 snapshot matrices.

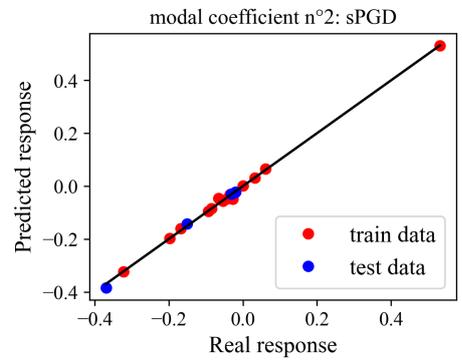
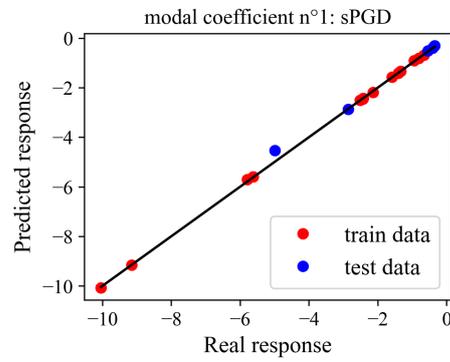
After computing the modal coefficients, a model for each one of them and for each measurement point was constructed. It is commonly known that time is inversely proportional to frequency and electric power, as a result, the inverse of the input parameters were considered when constructing the parametric regression of the modal coefficients. Besides, the standardization of the input parameters was applied to avoid problems related to units and different scaled variables, then the datasets were split into training and testing subsets (75% of data to build the models and 25% to evaluate the prediction accuracy). Figure 4 shows the real versus the predicted values of modal coefficients for points (#1, #5, #7, and #2.2) using sPGD regression model. The red points correspond to the data used to build the model while the blue ones correspond to the data used to evaluate its accuracy. When points are close to the black line, the model provides a good fit to data. Indeed, the dispersion of these points with respect to the black line gives a visual indicator of error.



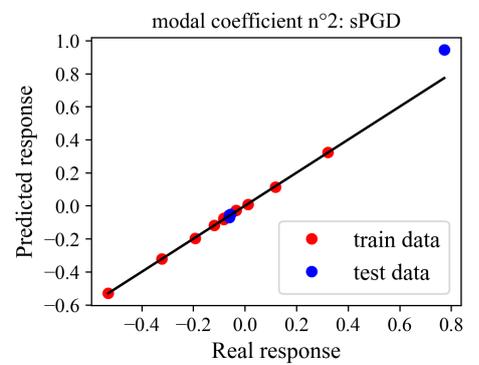
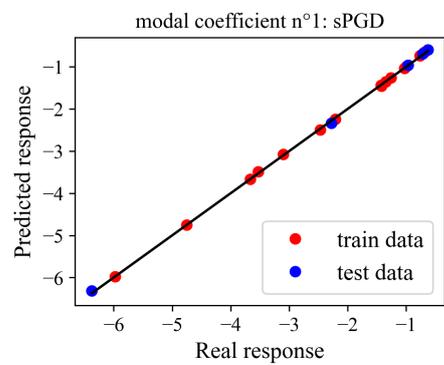
(a) measurement point #1



(b) measurement point #5



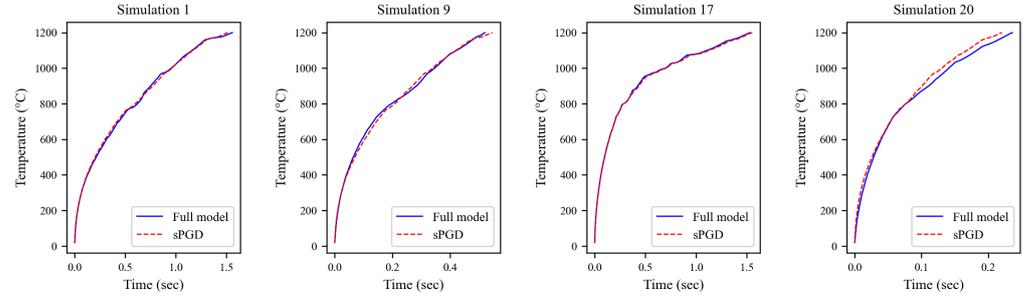
(c) measurement point #7



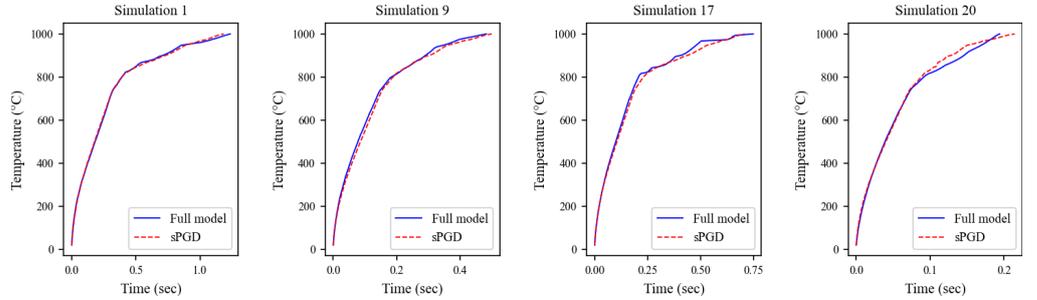
(d) measurement point #2.2

Figure 4. Real versus predicted response for 4 measurement points.

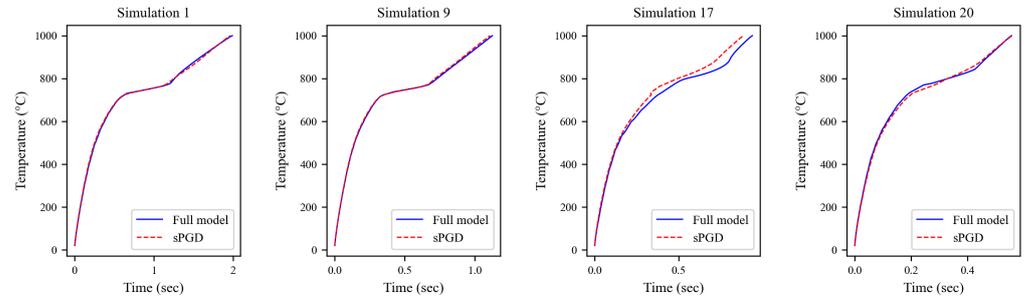
Figure 5 shows the time evolution of the temperature obtained by the full model and the sPGD regression model (i.e., using Equation (2) with the predicted values of modal coefficients as a function of the frequency and power). It can be seen that the real curves obtained by the full-order FE model (blue curves) and the predicted ones by sPGD model (dashed red curves) show the same trend and they are almost overlapped for both training and testing datasets and for the four measurement points as well.



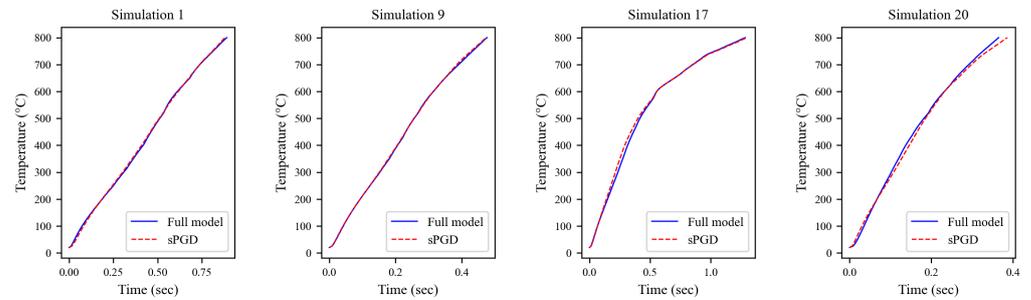
(a) measurement point #1



(b) measurement point #5



(c) measurement point #7



(d) measurement point #2.2

Figure 5. Comparison of the time evolution of temperature between the reference full model and the sPGD regression model (with the predicted values of modal coefficients) for 4 measurement points: simulations #1 and #9 are used to create the regression model and simulations #17 and #20 are used to check the regression model accuracy.

Additionally, the relative error to measure the prediction accuracy for training and testing datasets is presented in Table 3, where the computed error is defined as:

$$Error(\%) = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{t_{true}^{max}} \sqrt{\frac{1}{T^{max} - T^{min}} \int (t_{true} - t_{pred})^2 dT} \right) \times 100 \quad (3)$$

where n , t_{true}^{max} , t_{true} , t_{pred} , T^{max} , T^{min} are the number of data points, the maximum value of time, the real response values, the predicted response values, and the maximum and minimum values of temperature, respectively. The obtained results prove that the approximation is quite good.

Table 3. Relative errors of time evolution using sPGD model.

	Point #1	Point #5	Point #7	Point #2.2
Error _{train}	2.8%	2.6%	2.1%	1.6%
Error _{test}	6.5%	3.4%	4.1%	2.5%

A model to predict the time related to each temperature value in a given interval was created while keeping a constant final temperature value for all snapshots. The applied POD enables building a basis with only two modes, and the sPGD model created for the two POD modal coefficients provides an excellent prediction accuracy and a good approximate solution is obtained, even with a small amount of data.

2.3. Modeling the Temporal Evolution of Austenite

2.3.1. Methodology

During heating, austenitization consists in transforming the different metallurgical phases present at low temperature into austenite. The austenite phase provides a better comprehension of how the system evolves during heating and it will be considered in the second part of this paper. Several models for the characterization of phase transformation kinetics can be found in the literature, the Johnson–Mehl–Avrami (JMA) [24–27] model is one of the most commonly used models and it is implemented in FORGE[®] for induction heating process. However, it is worth pointing out that only data generated by FORGE[®] is needed in this work and the knowledge of the models is not required.

Besides, depending on the input parameter changes, austenite phase evolution can be observed in three different states as shown in Figure 6.

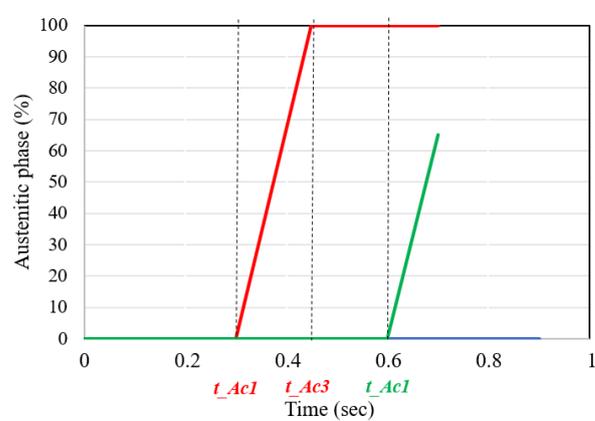


Figure 6. Three possible states of austenite phase evolution.

As can be seen, the phase transformation starts at time t_{Ac1} and the austenitization is completed at time t_{Ac3} . These instants are associated with the characteristic temperatures $Ac1$ and $Ac3$ which correspond to the start and end transformation temperatures,

respectively. In some cases, due to a short processing time or low values of process parameters, austenite phase evolution does not reach a complete austenitization (green curve) or does not even start its transformation (blue curve). Therefore, a methodology based on classification followed by regression, to avoid dealing with different shaped curves, was developed as follows:

First, the snapshots for each measurement point were classified into two classes:

- Class 1: no austenite phase transformation was produced, the austenite phase remains at zero until the end of the process (blue curve in Figure 6);
- Class 2: there is an austenite phase transformation; two cases are envisaged:
 - 100% of austenite is reached at the final time step (red curve in Figure 6);
 - less than 100% of austenite was obtained after heating (green curve in Figure 6).

Then, a constraint was imposed in order to deal only with the data in class 2.

Finally, t_{Ac1} and t_{Ac3} (if complete austenitization is reached) were collected from the data in class 2 and the nonlinear regression model sPGD was then applied to build a model of these two quantities. It is worth pointing out that the inverse of frequency and power were considered to fit a model to t_{Ac1} and t_{Ac3} .

2.3.2. Results and Discussion

A standardization of the input parameters was first applied and then the datasets were divided into training and testing subsets (75% of data was used to build the models and 25% to evaluate their prediction accuracy). Figure 7 shows the real versus the predicted values of t_{Ac1} for points (#1, #5, #7 and #2.2) using sPGD regression model and for both training and testing datasets. It can be noticed that the predicted values are in good agreement with the real ones, proving the excellent performance of the sPGD regression model.

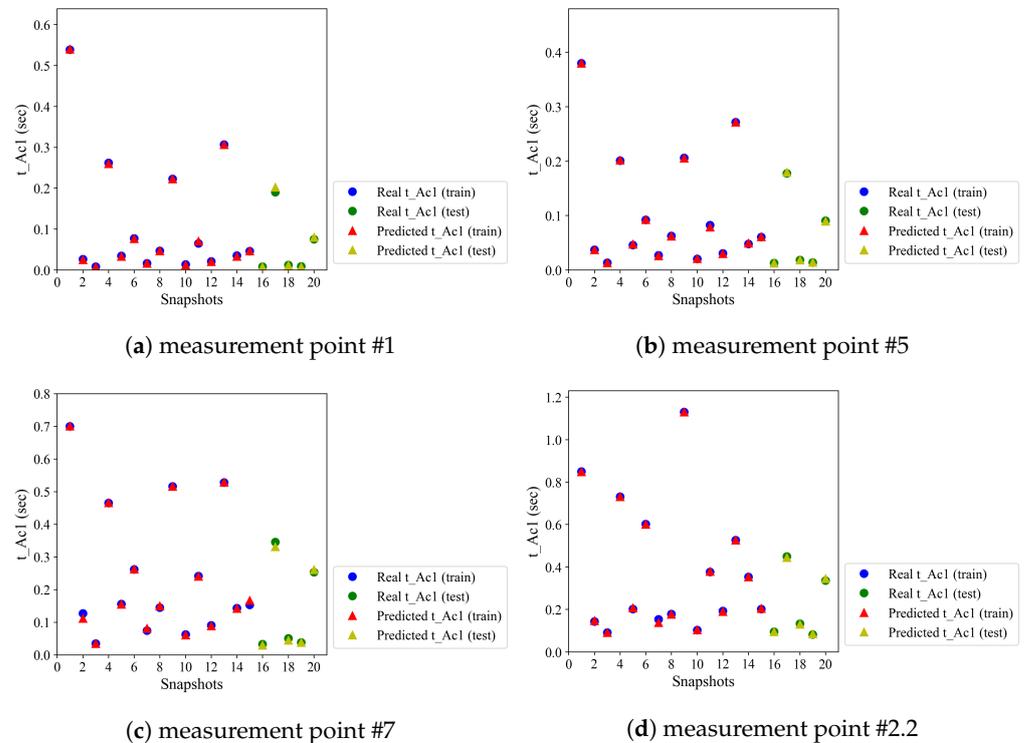


Figure 7. Real versus predicted values of t_{Ac1} for 4 measurement points.

In order to measure the prediction accuracy, the mean arctangent absolute percentage error (MAAPE) [28] was applied because of its capacity to deal with zero or close-to-zero

true values, which is the case in this study since t_{Ac1} and t_{Ac3} evolve between zero and two seconds. MAAPE is defined when predicting a quantity y as:

$$MAAPE = \frac{1}{n} \sum_{i=1}^n \arctan \left(\left| \frac{y_{true} - y_{pred}}{y_{true}} \right| \right) \times 100 \quad (4)$$

where y_{true} is the real response value, y_{pred} is the one predicted by the model, and n is the total number of data points.

The MAAPE for the training and testing datasets is shown in Table 4. As it can be noticed, the error is acceptable for all measurement points and does not exceed 6%.

Table 4. MAAPE for t_{Ac1} using sPGD model.

	Point #1	Point #5	Point #7	Point #2.2
$MAAPE_{train}$	2.3%	1.5%	2.7%	1.4%
$MAAPE_{test}$	5.2%	1.5%	5.7%	1.9%

By using the same rationale, Figure 8 shows the real versus the predicted values of t_{Ac3} for points (#1, #5, #7, and #2.2) using sPGD model and for both training and testing datasets. Again, an excellent approximation for t_{Ac3} is observed.

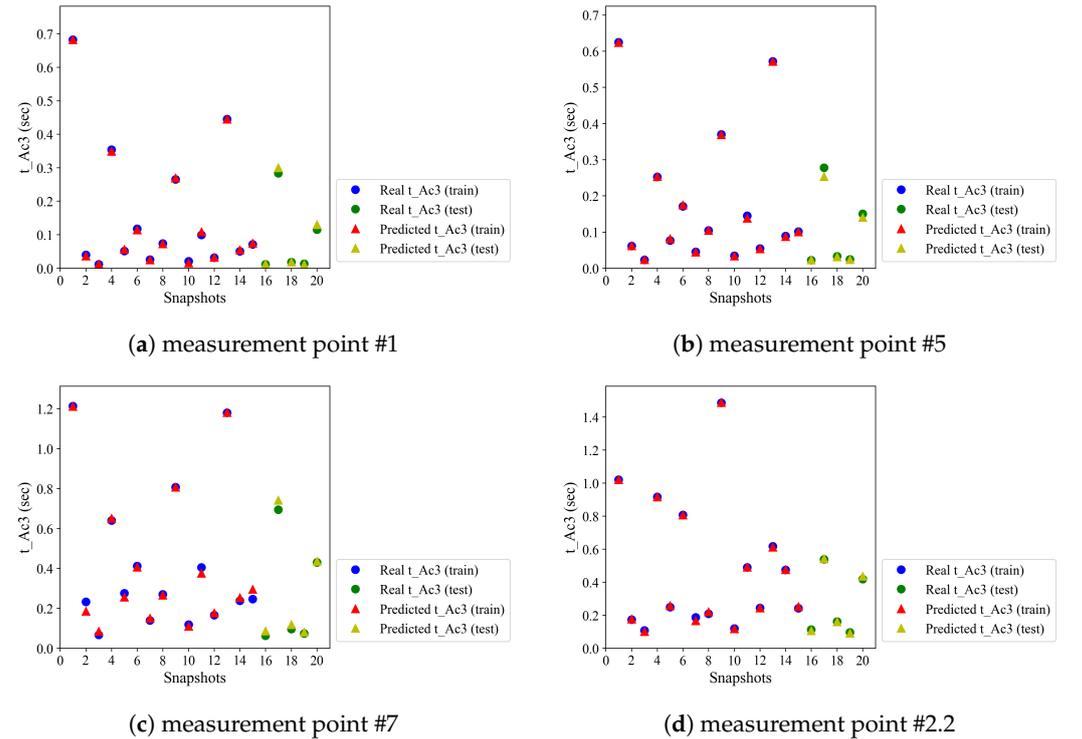


Figure 8. Real versus predicted values of t_{Ac3} for 4 measurement points.

The MAAPE for the training and testing datasets is shown in Table 5. The obtained error is acceptable, even if the error of 14.8% is obtained at the tooth tip (measurement point #7).

Table 5. MAAPE for t_{Ac3} using sPGD model.

	Point #1	Point #5	Point #7	Point #2.2
$MAAPE_{train}$	5.0%	2.0%	7.5%	2.3%
$MAAPE_{test}$	8.8%	5.6%	14.8%	3.7%

In summary, models for t_{Ac1} and t_{Ac3} were created. They provide a complete approximation of the austenite phase transformation history as follows:

- 0% of the austenite phase was obtained for all $t < t_{Ac1}$, which means that no phase transformation was initiated yet;
- 100% of the austenite phase was obtained for all $t > t_{Ac3}$, which means that a complete austenitization was reached;
- the austenite phase evolves linearly between 0% and 100% for all $t_{Ac1} < t < t_{Ac3}$.

The methodology was successfully applied for 14 sparse measurement points in the space domain, a good approximation was provided and a quite good performance of sPGD was shown as well.

3. Data-Driven and Hybrid Modeling for Hardness during Quenching

3.1. Input Analysis and Data Preprocessing

Consider the gear shown in Figure 2 at the end of the heating. Then, after a short elapsed time, a powerful quenching applies in order to cool the workpiece and induce its surface hardening. The workpiece is equipped with several measurement points (sensors), most of them located on the outer surface of the gear. Therefore, several simulations using FORGE[®] software were carried out to generate different datasets during the quenching step where the input configurations were the final state of heating. The set of the considered input variables to describe the system, the so-called feature space F , reads:

$$F = \left\{ k, t, \theta_{init}, (\theta_i)_{i \in [1;5]} \right\}, \quad (5)$$

where k is the simulation step, t is the total time of the quenching process, θ_{init} is the initial temperature at the quenching stage, $(\theta_i)_{i \in [1;5]}$ are the different temperature values collected at different times uniformly distributed on the quenching time interval t such that:

$$\forall i \in [1;5], \theta_i = \theta \left(t_i = \frac{t \times i}{5} \right), \quad (6)$$

Using these 5 temperatures provides the model with reasonable amount of information about the temperature history to describe the hardness. Using more temperature values would make the feature space too large for this low-data setting. Indeed, there is no real consensus about the optimized size of the feature space, the number \sqrt{N} , with N the number of data points, is often given but this would need an empirical analysis which is not the scope of this work. However, in this case, a more precise temperature profile brings the same information as the five temperatures. There is no need to extract more.

First preliminary predictions using naive models (simple models without optimization) showed a lack of accuracy which could be explained by a lack of information to describe the hardening process with respect to the considered variables. Hence, it was decided to define new variables from the temperature itself. Therefore, the space and time gradients of the temperature were chosen to complete the feature space F as they are expected a significant influence on the material hardening [29]. The temporal temperature gradient reads

$$C_i = \frac{\theta_{init} - \theta_i}{|t_0 - t_i|}, \quad (7)$$

where $t_0 = 0$.

In order to compute the gradient in space, it was necessary to associate a distance to ensure the frame indifference of the learned model. The proposed solution considers the distance to the surface, that is: $\forall x, d(x, x_s)$ represents the Euclidian distance from a point x

within the gear to its nearest point on the gear surface x_s . Therefore, the space gradient of the temperature field is expressed as follows, if $x \neq x_s$:

$$D_{i,x} = \frac{\theta_{i,x_s} - \theta_{i,x}}{d(x, x_s)} \quad (8)$$

The available data allows computing that gradient at the measurement point located in the bulk material, the sensor 2 in Figure 9. At the other measurement points located on the outer gear surface, the spatial temperature gradient cannot be calculated.

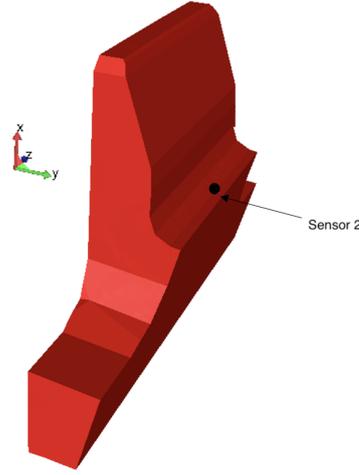


Figure 9. Location of sensor 2 in the gear.

Therefore, the initial feature space is enriched with the considered temperature gradients in (7) and (8) at the sensor 2 location, according to:

$$F = \left\{ \theta_{init}, (\theta_i)_{i \in [1;5]}, (D_{i,x_2^j})_{i \in [1;5]}, (C_i)_{i \in [1;5]}, d(x, x_2) \right\}, \quad (9)$$

where D_{i,x_2^j} is the space gradient at time i of points x_2^j in the neighborhood of x_2 , point at which sensor 2 is located on the outer surface, as illustrated in Figure 10.

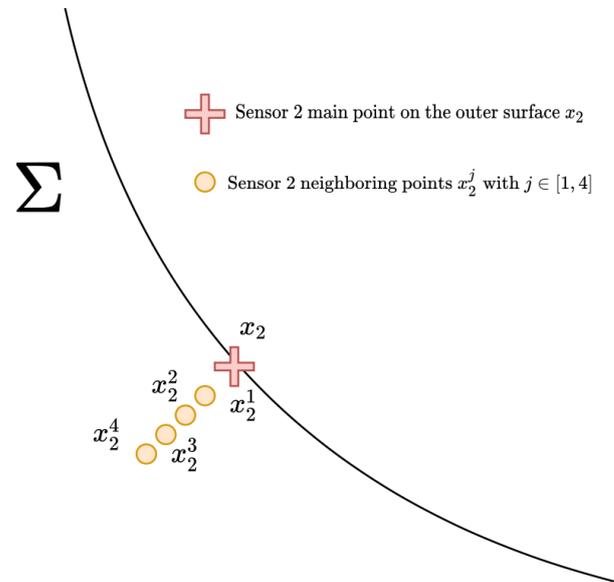


Figure 10. Schematic representation of the sensor 2 (x_2) and the associated measurement points x_2^j in the gear Σ .

The next step consists of splitting the data into the training and testing subsets. One of the most common techniques used in the field of machine learning is the K-Fold Cross Validation [30] which consists in choosing the best fold to be tested by the regression model. Nevertheless, since data distribution is not well-balanced, it brings a huge bias in the results. It can be noticed that the hardness H shows three different trends. These trends are $H < 268$, $268 < H < 721$, and $H = 721$ as shown in Figure 11.

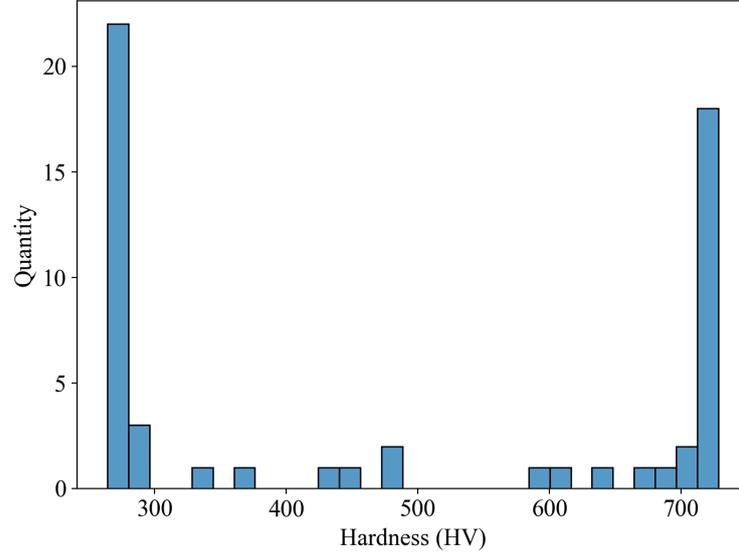


Figure 11. Hardness points histogram of the data located in the area of the sensor 2 (Figure 9).

Therefore, it was not possible to apply a K-Fold Cross Validation on these imbalanced datasets because the model performs very well on the largely represented data but lacks robustness [31]. In order to have a balanced distribution in the testing dataset, in absence of considering richer samplings, it was decided to cluster data into three classes with the K-means algorithm [32] and define a testing dataset in each cluster. Now, each trend of the distribution is represented in the testing dataset and the performance is no longer biased.

The next step is to select variables in the feature space previously defined. Considering the low number of data points, it was necessary to reduce the number of input variables [33]. The adopted technique is a statistical method called F-test. Two models are considered, M_1 and M_2 , one with a selected variable and a constant and the other one with just a constant. The F-test analyzes the least square error between M_1 and M_2 to check if it is significant or not. It leads to the F-score which is basically the significance of a variable according to the model. Then, the F-score is used to select the features. After the selection, the new space reads:

$$F_{selected} = \left\{ \theta_{init}, D_{2,x_2^j}, \theta_1, D_{1,x_2^j}, D_{0,x_2^j}, C_1, d(x, x_2), \theta_2 \right\}, \quad (10)$$

3.2. Modeling the Hardness

The type of regression model was chosen according to the needs of the studied problem. The complexity of the system excluded simple models such as linear regression. Then, the lack of data avoided the use of a classical neural network like the multilayer perceptron. Therefore, the considered model in this work was the gradient boosting algorithm [14] implemented in the XGBoost library which has recently become one of the most widely used machine learning techniques [34,35]. Parameters of this model were optimized using a random search to avoid excessive computation time and the associated issues [36]. The starting search space was defined arbitrarily with basic and foolproof hyperparameters.

3.2.1. Data-Driven Model

There are 56 data points registered at sensor 2 and its neighbor measurement points after the simulations. Therefore, the proposed dataset is small compared to the usual standards in machine learning. In fact, the more points a dataset has, the more efficient each training step is. Basically, applying machine learning algorithms to such a complex system was regarded as a challenge. This is why it was important to proceed with variables extraction, selection, and optimized hyperparameters search to maximize the performance of the model.

It is well known that models learned from data are efficient to interpolate but fail to extrapolate. Therefore, when a testing point is located in a region outside the training region of the parametric space, the error usually increases. Two points were removed from the testing dataset since there was no point close to them in the training dataset.

To avoid overfitting, a part of the training dataset was also used as an evaluation set to stop when reaching a relevant score. Moreover, another stopping criterion was set to break the loop with a chosen number of iterations [37]. This number was chosen to ensure the smallest test error.

In order to evaluate the prediction accuracy, Root Mean Squared Error (RMSE) and Root Mean Squared Percentage Error (RMSPE) were used. They are defined as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{pred} - y_{true})^2}, \quad (11)$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_{pred} - y_{true}}{y_{true}} \right)^2} \times 100, \quad (12)$$

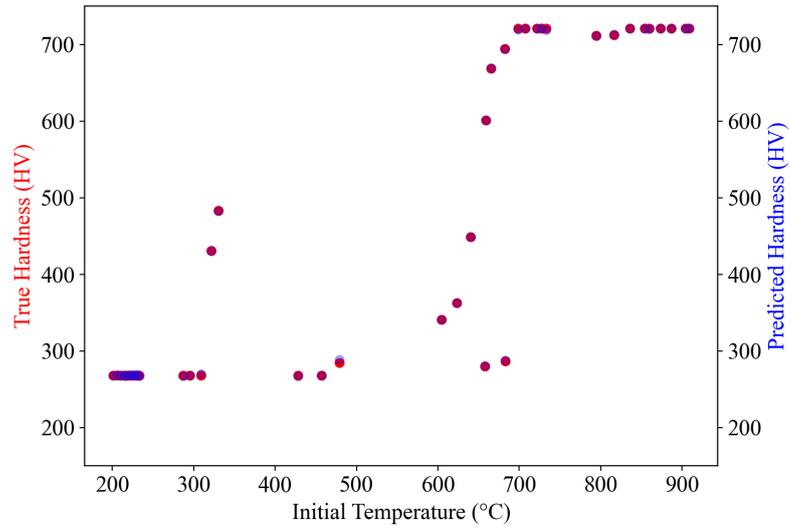
where y_{true} and y_{pred} are the real hardness and the predicted one, respectively.

The advantage of RMSPE compared to RMSE is that it shows a percentage which is more significant than an arbitrary quantity. Therefore, after fitting the regression model and testing the dedicated set on it, the results are given in Table 6.

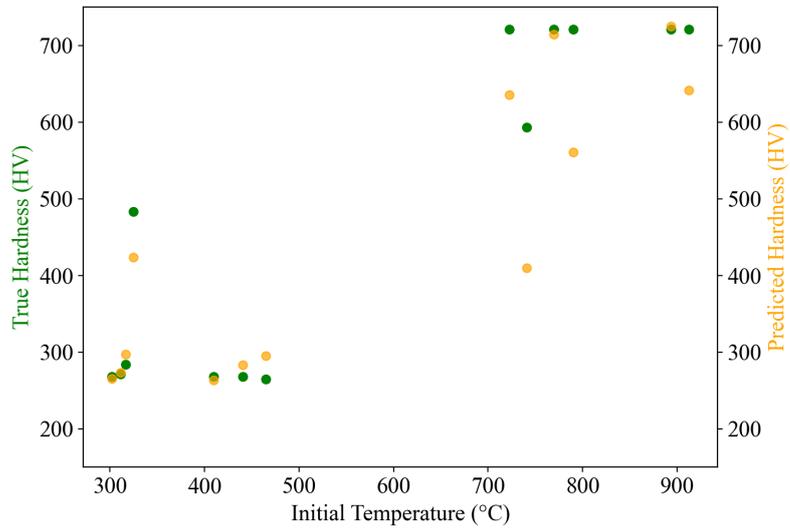
Table 6. Errors of the XGBoost and Random Forest models to compare as a simple baseline to predict hardness with sensor 2 dataset.

Model	XGBoost	Random Forest
$RMSE_{train}$	0.264	30.629
$RMSE_{test}$	77.319	151.494
$RMSPE_{train}$	0.08%	8.30%
$RMSPE_{test}$	12.58%	25.42%

It is clear that the model still has trouble generalizing on new data but those results are still relevant and give perspective to reach a better accuracy using more sampling points. The predicted data with respect to the real data was illustrated in Figure 12. It can be noticed that the trend is well described. The visualization of the numerical results shows that the training dataset was well predicted as it is almost impossible to distinguish red points (true ones), i.e., they are overlapped with the blue ones. On the other side, the testing dataset does not have the best results so far, but it would surely be possible to improve them with more data. Moreover, the fact that the maximal hardness is reached with a low temperature ($\theta_{init} = [700, 750]$ (°C) for $H = 721$ (HV)) could lead to question the considered variables. However, it is common in data science to get a large scope of different data to make the model learn on a complete profile in order to describe every behavior. Indeed, it allows to trust the results if someday an uncommon prediction is required and to avoid extrapolation which makes the prediction harder for a model learnt on small datasets.



(a)



(b)

Figure 12. Comparison between the true and predicted hardness of the sensor 2 dataset: (a) training dataset with the predicted hardness (blue) over the ground truth (red). (b) testing dataset with the predicted hardness (yellow) over the ground truth (green). Both with respect to the initial temperature of the system.

3.2.2. Hybrid Model

The lack of data, due to the high computational costs or the difficulty of collecting experimental data, makes the construction of efficient models difficult. A hybrid framework [38] could constitute an appealing compromise to ally accuracy and data frugality. Indeed, physical laws represent a valuable knowledge of the system even if in general reality is a bit more complex. Data should be used to enrich models instead of replacing them.

It is well known that the austenite ratio before quenching is highly correlated with the hardness as shown in Figure 13. Hence, the prediction of the austenite ratio at every point of the gear at the end of the heating (before quenching starts) can be used as a new input variable, represented by the variable α , to better describe the hardness.

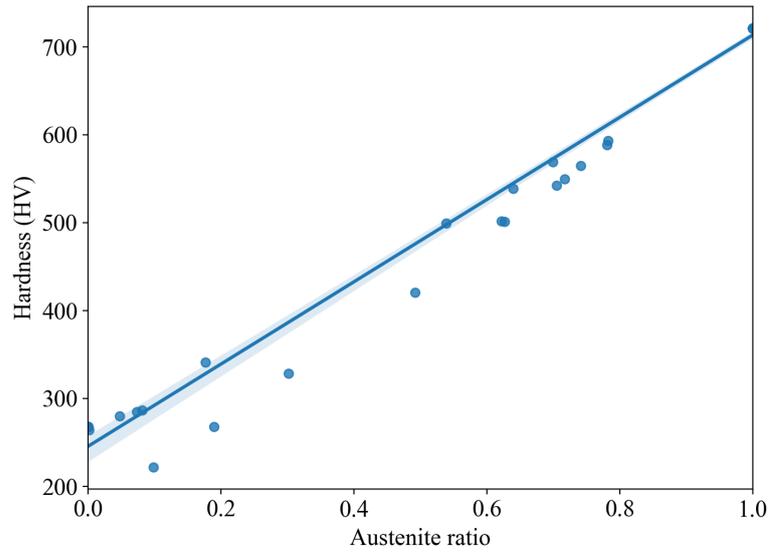


Figure 13. Regression plot of the austenite ratio with respect to the hardness.

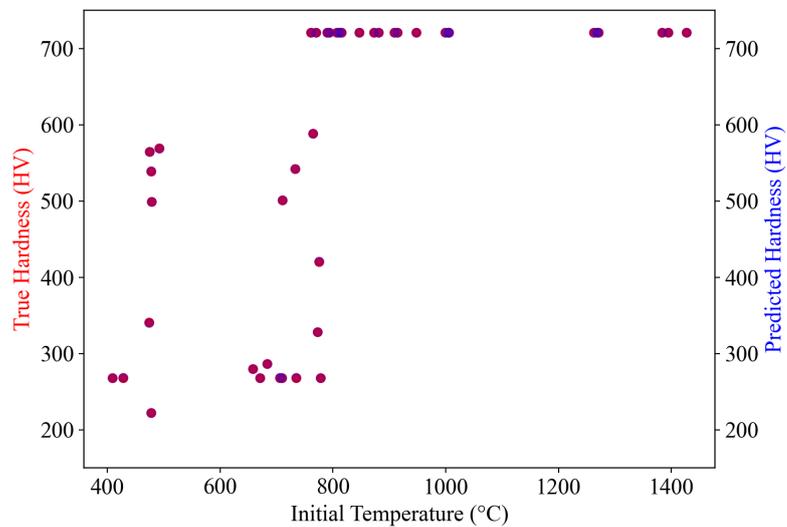
Considering α , the pipeline described in Section 3.1 is used again with a new selection of variables and model optimization. The new feature space reads:

$$F = \{\theta_{init}, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \alpha, C_1, C_2, C_4, C_5\}, \quad (13)$$

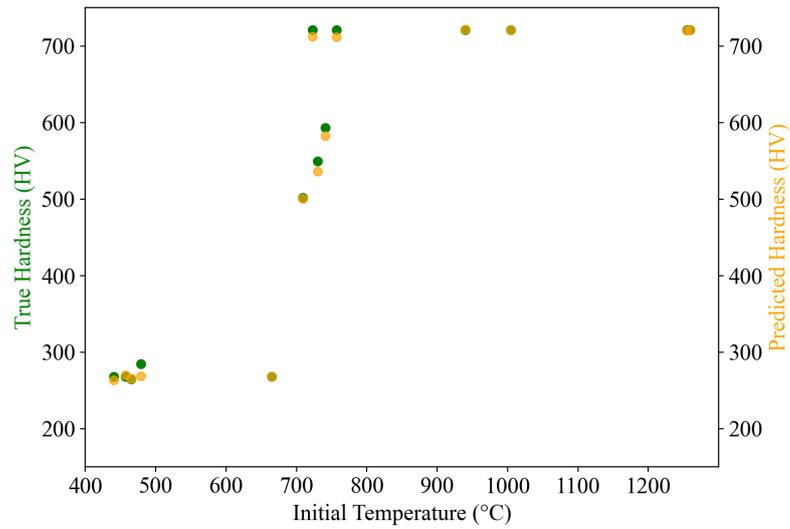
The knowledge provided by the austenite ratio offers the possibility to avoid the use of the space temperature gradients that simplifies the data collection. Results are shown in the Figure 14 and grouped in Table 7.

Table 7. Prediction error of the hardness computed with the XGBoost model, with and without the austenite ratio as a feature with the global dataset.

Model	XGBoost with Austenite	XGBoost without Austenite
$RMSE_{train}$	1.502	0.668
$RMSE_{test}$	7.249	106.229
$RMSPE_{train}$	0.44%	0.17%
$RMSPE_{test}$	1.82%	21.62%



(a)



(b)

Figure 14. Comparison between the true and predicted hardness of the global dataset: (a) training dataset with the predicted hardness (blue) over the ground truth (red). (b) testing dataset with the predicted hardness (yellow) over the ground truth (green). Both with respect to the initial temperature of the system.

3.3. Discussion

Predicting hardness at any point after the induction hardening process is an important gear validation procedure for industrial applications. It was proved in this work that a gradient-boosted tree-based model with features selection could perform satisfactorily in complex settings. However, having access to those variables (e.g., temperature gradients) seems technically a tricky issue and the lack of data combined with the high nonlinearity of the system suggest the need to introduce existing knowledge into the learning procedure. Here, the austenite ratio was considered because it could be provided by a validated reduced model, and it is a variable that exhibits a correlation with the hardness. It has been proved that the introduction of knowledge, making the model hybrid, is here mandatory to have performing results. Several questions arise in this discussion—lack of data can be solved by bringing knowledge, but if there is a lack of knowledge or if it is not available, a more enduring model is needed. The solution would be persistent with the help of the fundamental laws of physics and not a well-chosen variable of the system. This would be the next step of this work to make a more robust and sustainable model.

4. Conclusions

In this paper, the multiphysics parametrized induction hardening process was studied, and parametric metamodels were developed to predict three physical quantities: temperature and austenite phase evolution during heating and the workpiece hardness after quenching. First, an approach, based on dimensionality reduction by POD coupled with a regression technique to fit a model to the POD modal coefficients, was initially proposed to compute the time evolution as a function of temperature. The approach was successfully applied for 14 sparse measurement points in the space domain, in which a basis with two modes was built and consequently two POD modal coefficients were used for creating the prediction model. A good approximation was provided using the sPGD regression.

Second, a methodology, based on classification combined with the sPGD regression model, was applied where only data with phase transformation were considered and a model for start and complete austenitization time instants was created. A good approximation was provided using sPGD even with a small amount of data. Since these models were built for some sparse points in the space domain, the solution will be extended to

address the whole space domain using interpolation techniques. In addition, addressing geometrical parameters, such as the dimension of the workpiece and the inductor, constitutes the next step of this work. Therefore, the same approaches and techniques will be applied while considering geometrical parameters to be the new inputs for the models.

Finally, as far as the hardness is concerned, the prediction was carried out by using tree-based gradient boosted data-driven models with two approaches. On the one hand, a fully data-driven model with the extraction of new variables based on the temperature profile was used. It is only located in the bulk material because of lack of data but with encouraging performance. On the other hand, a physic-enhanced model with the austenite ratio available thanks to a reduced model was used. This one is global, i.e., in all the gear with great performance. However, more data are required to enhance prediction accuracy of data-driven model for predicting the hardness. The hybrid modeling could be improved by enforcing a thermodynamical consistency which is the fact of satisfying the physical equations.

Author Contributions: Conceptualization, F.C.; methodology, K.D., S.G. and V.C.; software, K.D. and S.G.; supervision, M.D., K.T. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was carried out as part of a project TRANSFUGE sponsored by the french institute of research and technology in materials, metallurgy and processes (IRT-M2P), France. The project benefited from funding from a consortium of industrial partners and the French PIA (Plan d'Investissement d'Avenir) granted the French National Agency for Research (ANR).

Informed Consent Statement: Not applicable.

Data Availability Statement: Provided under request.

Acknowledgments: This work was conducted with the help of the French Technological Research Institute for Materials, Metallurgy and Processes (IRT-M2P). The authors would like to acknowledge IRT-M2P and the partners of the project TRANSFUGE led by IRT-M2P.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Sparse Proper Generalized Decomposition sPGD

sPGD is a nonlinear regression model which provides a quite rich approximation for multidimensional and parametric problems with a small amount of sparse data. For ease of explanation, let us consider a parametric problem where the studied QoI lives in \mathbb{R}^2 , $u(P1, P2)$ where $P1 \in \Omega_1$ and $P2 \in \Omega_2$ are the considered parameters and u is known for certain values of $P1$ and $P2$. In order to calculate the approximate solution $\tilde{u}(P1, P2)$ of $u(P1, P2)$, the Galerkin projection is first applied such that:

$$\int_{\Omega_1 \times \Omega_2} w^*(P1, P2) (\tilde{u}(P1, P2) - u(P1, P2)) dP1 dP2 = 0 \quad (A1)$$

where w^* is an arbitrary test function. Then, following the Proper Generalized Decomposition (PGD) rationale, the approximate solution $\tilde{u}(P1, P2) \approx \tilde{u}^N(P1, P2)$ is expressed in the separated form:

$$\tilde{u}^N(P1, P2) = \sum_{i=1}^N M_i^1(P1) \cdot M_i^2(P2) \quad (A2)$$

The determination of the precise form of pairs $M_i^1(P1) \cdot M_i^2(P2)$ is done by first projecting them on a finite element basis, as shown in Equation (A3), and by employing a greedy

algorithm such that the N th order term is calculated once the approximation up to order $N - 1$ is known

$$\begin{aligned} M_i^1(P1) &= \sum_{k=1}^D N_i^k(P1) \cdot \alpha_i^k \\ M_i^2(P2) &= \sum_{k=1}^D N_i^k(P2) \cdot \beta_i^k \end{aligned} \quad (A3)$$

where N_i^k , α_i^k , and β_i^k represent the FE shape functions and the degrees of freedom of the chosen approximation, respectively. The choice of the basis in which each one of the one-dimensional function was expressed is made based on the studied problem. A choice ensuring accuracy and avoiding spurious oscillations consists of using interpolants based on Kriging techniques. It is worth mentioning that the product of the test function w^* times the objective function $u(P1, P2)$ is only evaluated at some points corresponding to the available sampled data. Since information is just known at sampling points, instead of using the test function in a finite element context, it is expressed as a set of Dirac delta distributions collocated at the sampling data (denoted by S) such that:

$$w^*(P1, P2) = \tilde{u}^*(P1, P2) \sum_{i=1}^S \delta(P1_i, P2_i) \quad (A4)$$

By combining the all equations presented above, a nonlinear system of equations is derived, due to products of terms. The alternate direction scheme is used to linearize the problem and to solve it.

References

1. Wuppermann, C.; Míček, E. Importance of heat treatment for the variety of applications of modern materials. *Prozesswärme* **2018**, *2018*, 95–101.
2. Wagner, L. Mechanical surface treatments on titanium, aluminum and magnesium alloys. *Mater. Sci. Eng. A* **1999**, *263*, 210–216. [[CrossRef](#)]
3. Nascimento, M.P.; Souza, R.C.; Pigatin, W.L.; Voorwald, H.J.C. Effects of surface treatments on the fatigue strength of AISI 4340 aeronautical steel. *Int. J. Fatigue* **2001**, *23*, 607–618. [[CrossRef](#)]
4. Rokicki, P. Induction hardening of tool steel for heavily loaded aircraft engine components. *Arch. Metall. Mater.* **2017**, *62*, 315–320. [[CrossRef](#)]
5. Candeo, A.; Ducassy, C.; Bocher, P.; Dughiero, F. Multiphysics Modeling of Induction Hardening of Ring Gears for the Aerospace Industry. *IEEE Trans. Magn.* **2011**, *47*, 918–921.
6. Rudnev, V.; Loveless, D.; Cook, R. *Handbook of Induction Heating*, 2nd ed.; CRC Press, Taylor and Francis Group: Boca Raton, FL, USA, 2017.
7. Hömberg, D. A mathematical model for induction hardening including mechanical effects. *Nonlinear-Anal-Real World APP* **2004**, *5*, 55–90. [[CrossRef](#)]
8. Spezzapria, M.; Dughiero, F.; Forzan, M.; Candeo, A. Multiphysics fem simulation of contour induction hardening process on aeronautical gears. *J. Iron Steel Res.* **2012**, *19*, 95–98.
9. Wang, K.; Chandrasekar, S.; Yang, H. Finite-element simulation of induction heat treatment. *J. Mater. Eng. Perform.* **1992**, *1*, 97–112. [[CrossRef](#)]
10. Samiuddin, M.; Muzamil, M. Experimental investigation and optimization of process parameters for through induction hardening using factorial design of experiments. *J. Eng. Res.* **2016**, *5*, 174–185.
11. Hutton, D. *Fundamentals of Finite Element Analysis*, 1st ed.; McGraw-Hill Higher Education: Boston, MA, USA, 2004.
12. Stein, M. Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics* **1987**, *29*, 143–151. [[CrossRef](#)]
13. Astrid, P. Reduction of Process Simulation Models: A Proper Orthogonal Decomposition Approach. Ph.D. Thesis, Technische Universiteit Eindhoven, Eindhoven, The Netherlands, 2004.
14. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [[CrossRef](#)]
15. Lucia, O.; Maussion, P.; Dede, E.; Burdío, J. Induction Heating Technology and Its Applications: Past Developments, Current Technology, and Future Challenges. *IEEE Trans. Ind. Electron.* **2014**, *61*, 2509–2520.
16. Magnabosco, I.; Ferro, P.; Tiziani, A.; Bonollo, F. Induction heat treatment of a ISO C45 steel bar: Experimental and numerical analysis. *Comput. Mater. Sci.* **2006**, *35*, 98–106. [[CrossRef](#)]

17. Ding, H.; Shin, Y. A Metallo-Thermomechanically Coupled Analysis of Orthogonal Cutting of AISI 1045 Steel. *J. Manuf. Sci. Eng.* **2012**, *134*, 051014. [[CrossRef](#)]
18. Geijselaers, H. Numerical Simulation of Stresses due to Solid State Transformations: The Simulation of Laser Hardening. Ph.D Thesis, University of Twente, Enschede, The Netherlands, 2003.
19. Lee, S.J.; Pavlina, E.; Vantyne, C. Kinetics modeling of austenite decomposition for an end-quenched 1045 steel. *Mater. Sci. Eng. Struct. Mater.* **2010**, *527*, 3186–3194. [[CrossRef](#)]
20. BARBA, P. Field models of induction heating for industrial applications. *Prz. Elektrotech.* **2018**, *1*, 3–7. [[CrossRef](#)]
21. Cardinaux, D. Étude et Modélisation Numérique 3D par Éléments Finis d'un Procédé de Traitement Thermique de Tôles Embouties après Chauffage par Induction: Application à un Renfort de Pied Central Automobile. Ph.D. Thesis, École Nationale Supérieure des Mines de Paris, Paris, France, 2008.
22. Ibáñez Pinillo, R.; Abisset-Chavanne, E.; Ammar, A.; González, D.; Cueto, E.; Huerta, A.; Duval, J.; Chinesta, F. A Multidimensional Data-Driven Sparse Identification Technique: The Sparse Proper Generalized Decomposition. *Complexity* **2018**, *2018*, 1–11. [[CrossRef](#)]
23. Limousin, V.; Delgerie, X.; Leroy, E.; Ibáñez, R.; Argerich, C.; Daim, F.; Duval, J.; Chinesta, F. Advanced model order reduction and artificial intelligence techniques empowering advanced structural mechanics simulations: Application to crash test analyses. *Mech. Ind.* **2019**, *20*, 804. [[CrossRef](#)]
24. Johnson, W.; Mehl, R. Reaction Kinetics in Process of Nucleation and Growth. *Trans. Am. Inst. Mining Metall. Pet. Eng.* **1939**, *135*, 416.
25. Avrami, M. Kinetics of Phase Change. I General Theory. *J. Chem. Phys.* **1939**, *7*, 1103–1112. [[CrossRef](#)]
26. Avrami, M. Kinetics of phase change II. transformation-time relations for random distribution nuclei. *J. Chem. Phys.* **1940**, *8*, 212–224.
27. Avrami, M. Kinetics of phase change III. Granulation, phase change and microstructure. *J. Chem. Phys.* **1941**, *9*, 177–184. [[CrossRef](#)]
28. Kim, S.; Kim, H. A new metric of absolute percentage error for intermittent demand forecasts. *Int. J. Forecast.* **2016**, *32*, 669–679. [[CrossRef](#)]
29. Hömberg, D.; Liu, Q.; Montalvo-Urquizo, J.; Nadolski, D.; Petzold, T.; Schmidt, A.; Schulz, A. Simulation of multi-frequency-induction-hardening including phase transitions and mechanical effects. *Finite Elem. Anal. Des.* **2016**, *121*, 86–100. [[CrossRef](#)]
30. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer Series in Statistics; Springer: New York, NY, USA, 2009.
31. Santos, M.S.; Soares, J.P.; Abreu, P.H.; Araujo, H.; Santos, J. Cross-Validation for Imbalanced Datasets: Avoiding Overoptimistic and Overfitting Approaches [Research Frontier]. *IEEE Comput. Intell. Mag.* **2018**, *13*, 59–76. [[CrossRef](#)]
32. Macqueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*; University of California: Berkeley, CA, USA, 1967; pp. 281–297.
33. Kuncheva, L.I.; Rodríguez, J.J. On feature selection protocols for very low-sample-size data. *Pattern Recognit.* **2018**, *81*, 660–673. [[CrossRef](#)]
34. Pafka, S. Benchm-ml. 2019. Available online: <https://github.com/szilard/benchm-ml> (accessed on 25 April 2021).
35. Orzechowski, P.; La Cava, W.; Moore, J.H. Where are we now? In *Proceedings of the Genetic and Evolutionary Computation Conference*, Kyoto, Japan, 15–19 July 2018
36. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
37. Prechelt, L. Early stopping-but when? In *Neural Networks: Tricks of the Trade*; Lecture Notes in Computer Science Book Series; Springer: Berlin/Heidelberg, Germany, 1998; pp. 55–69.
38. Chinesta, F.; Cueto, E.; Abisset-Chavanne, E.; Duval, J.L.; El Khaldi, F. Virtual, digital and hybrid twins: A new paradigm in data-based engineering and engineered data. *Arch. Comput. Methods Eng.* **2020**, *27*, 105–134. [[CrossRef](#)]