algorithms

MDPI

Article

Optimal Multi-Physics Synthesis of a Dual-Frequency Power Inductor Using Deep Neural Networks and Gaussian Process Regression

Paolo Di Barba [1], Arash Ghafoorinejad [1], Maria Evelina Mognaschi [1,*], Fabrizio Dughiero [2], Michele Forzan [2] and Elisabetta Sieni [3]

[1] Department of Electrical Computer and Biomedical Engineering, University of Pavia, 27100 Pavia, Italy; paolo.dibarba@unipv.it (P.D.B.); arash.ghafoorinejad01@universitadipavia.it (A.G.)
[2] Department of Industrial Engineering, University of Padova, 35131 Padova, Italy; fabrizio.dughiero@unipd.it (F.D.); michele.forzan@unipd.it (M.F.)
[3] Department of Theoretical and Applied Sciences, University of Insubria, 21100 Varese, Italy; elisabetta.sieni@uninsubria.it
* Correspondence: eve.mognaschi@unipv.it; Tel.: +39-0382-985785

Abstract: In this paper, a multi-physics case study belonging to the class of induction heating problem is considered. Finite Element Analysis is used to evaluate the temperature along a line on a graphite disk heated by two power inductors. In order to build a surrogate field model of the device, i.e., to compute the temperature profile on the disk, given the amplitudes and frequencies of the supply currents, three methods have been used (Support Vector Regression (SVR), fully connected Neural Network (NN) and Gaussian Process Regression (GPR)). In turn, to solve the inverse problem, i.e., to identify frequencies and currents of the two coils, given a prescribed temperature profile, two approaches have been implemented. The former is an optimization approach based on a multi-objective formulation, solved by means of the NSGA-II algorithm; the latter is a two-step procedure, based on fully connected Deep Neural Networks (DNNs), solving an optimal design problem first and, subsequently, an optimal control problem.

Keywords: induction heating; multi-physics domain; finite element analysis; deep neural networks; multi-objective optimization

Academic Editors: Van-Hai Bui, Xuan Zhou, Wencong Su and Akhtar Hussain

Received: 29 November 2024
Revised: 29 December 2024
Accepted: 31 December 2024
Published: 2 January 2025

Citation: Di Barba, P.; Ghafoorinejad, A.; Mognaschi, M.E.; Dughiero, F.; Forzan, M.; Sieni, E. Optimal Multi-Physics Synthesis of a Dual-Frequency Power Inductor Using Deep Neural Networks and Gaussian Process Regression. *Algorithms* **2025**, *18*, 10. https://doi.org/10.3390/a18010010

1. Introduction

Induction heating is a technique commonly applied to heat materials in a contactless and efficient way because it provides precise control over the temperature of the material [1]. It is an advanced technology utilized in several industrial applications for efficient and targeted heating. There are many fields of application, ranging from automotive to aerospace, from metalworking to manufacturing.

This contactless method is based on electromagnetic induction: a time-varying current flowing in an inductor induces eddy currents in a conductive piece of material, which is heated thanks to the Joule effect [2,3].

This highly localized heating allows precise temperature control and efficient energy transfer, leading to reduced processing times and enhanced energy efficiency compared to conventional heating methods [4].

This technique is very efficient, with an efficiency higher than 90% [5], as energy is directly transferred to the material, minimizing losses associated with heat conduction to other objects, not intended to be heated. Additionally, induction heating provides

rapid heating and precise temperature control, allowing uniform heating in the whole workpiece, which is essential in processes such as hardening, annealing, and brazing in metalworking. Furthermore, induction heating systems are clean and safe because they do not involve open flames or toxic byproducts, enhancing workplace safety and meeting stringent industry standards for emissions and pollutants.

However, effective induction heating requires precise calibration of parameters, including frequency and power. A wrong setup of the inductors or badly designed inductors deteriorates the efficiency of the process. Hence, one of the main challenges in induction heating is the design of inductors and their parameter tuning.

To this end, automated optimal designs and control of inductors have been proposed in the literature in the last few decades. Moreover, numerical methods such as Finite Element Analysis have replaced analytical formulations, increasing the accuracy of the solution [6,7]. Specifically, the field problem requires the solution of coupled electromagnetic and thermal problems to simulate the temperature in the workpiece generated by the power supplied by an inductor. In complicated geometry, this task requires a big effort in terms of computation time and hardware resources since an electromagnetic problem is coupled to a thermal problem [8,9]. These features made this problem an interesting and well-known benchmark, solved in different areas of computational electromagnetism in a comparative way [10].

With the advent of innovative algorithms based, e.g., on Neural Networks (NNs), it is possible to build reliable surrogate models for predicting the temperature distribution in a workpiece during induction heating processes [11,12]. Actually, surrogate models help very much in the design of inductors, because a typical optimization routine, using, e.g., genetic algorithms, requires the solution of a large number of parametric geometries to find a set of improved solutions [13–15]. For example, when a multi-objective optimization is considered, the Pareto approach coupled with FEA is very time-expensive; therefore, in this area, where the field problem involves both electromagnetic and temperature fields, giving rise to a coupled problem, it is convenient to use a surrogate model to reduce the number of required FEA analyses [13,16–18].

In terms of applications, in this paper, an induction heating system for epitaxial growth of silicon is considered, as described in two patents (Forzan, M.; Crippa, D.; Preti, S. Deposition Reactor with Inductors and Electromagnetic Shields 2021, EP3870734A1; Ogliari, V.; Forzan, M.; Preti, S. Inductively Heatable Susceptor and Epitaxial Deposition Reactor 2017, WO2017137872A1) and a study [19]. The proposed geometry is a pancake inductor that uniformly heats a graphite disk. The aim of the device design is to obtain a prescribed temperature profile, which should be as uniform as possible along the disk surface and should reach a given temperature. Examples of geometries similar to the one proposed in the paper are reported in [15,20], where the device for induction heating is optimized. The peculiarity of the proposed device is that the power inductor is composed of two coils, which contribute to the creation of the magnetic field and hence to the eddy currents in the workpiece; the two coils are fed with two currents, each characterized by a different magnitude and frequency. The study of the magneto-thermal field problem given by two inductors fed by different frequencies is not straightforward and not easily solvable with commercial software. Hence, in this case, the use of a surrogate model is suitable.

In this paper, the thermal behavior of the power inductor is investigated by means of several algorithms. Specifically, NNs, Gaussian Process Regression (GPR), and Support Vector Regression (SVR) algorithms are considered. Based on this study, a subsequent optimal synthesis of the electrical parameters of the coils follows; the final aim is to obtain a prescribed quasi-homogenous temperature profile on the workpiece surface. Specifically, the surrogate model is then used to evaluate the objective functions, in the frame of a multi-objective optimization, performed by means of a well-known algorithm NSGA-II [21–24].

Another approach, based on a deep learning algorithm, is then set up and the obtained results are promising.

The rest of this paper is organized as follows: in Section 2, the forward problem in terms of field problem and related surrogate model is described. In Section 3, the inverse problem and the proposed approaches to solve it are described. In Section 4, the results are shown and in Section 5, they are discussed. Finally, a conclusion follows.

2. Forward Model

2.1. The Device

A benchmark induction-heating device with cylindrical symmetry (Figure 1) is considered as the case study [19]. The device features a flat, circular graphite disk that has an electrical resistivity of 960 $\times$ 10^{-8} Ωm and thermal conductivity of 150 $Wm^{-1}K^{-1}$ at 1200 Celsius degrees. This graphite disk has a diameter of 500 mm. Wound around it, there are two copper coils consisting of 7 and 6 turns, respectively, encased in a quartz container. These coils are fed separately: coil 1 carries an electric current varying between 100 Arms and 500 Arms and frequency between 20 kHz and 100 kHz, coil 2 carries an electric current varying between 400 Arms and 2 kArms and frequency between 1 kHz and 8 kHz.

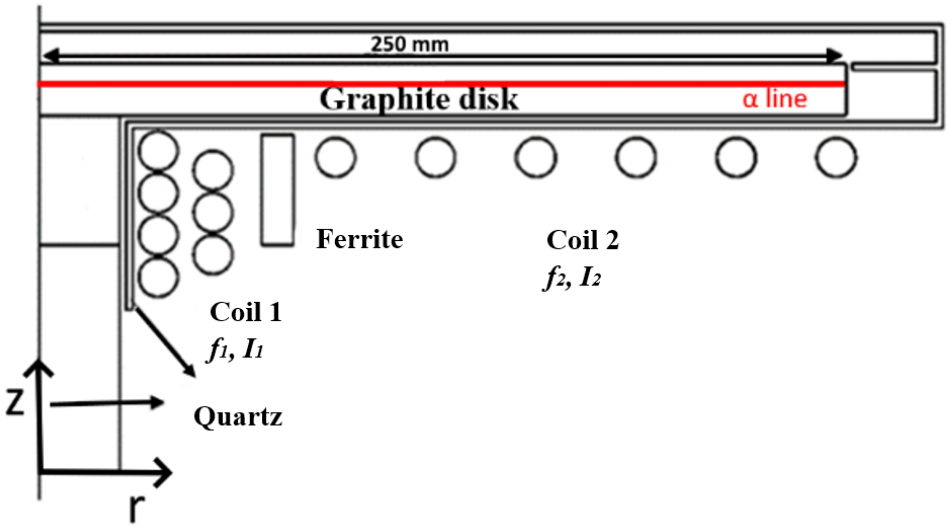

Figure 1. Geometry of the induction heating device.

2.2. Finite Element Model

In order to simulate the temperature distribution in the graphite disk for specific current and frequency values, a Finite Element (FE) model, with cylindrical symmetry, is set up. A weakly coupled magneto-thermal problem is solved: the magnetic problem is solved in time-harmonic conditions with the A–V formulation. In particular, the magnetic problem, solved in terms of the phasor of the magnetic vector potential, $\dot{\boldsymbol{A}}$, involves the equation [25,26]:

$$\nabla^2 \dot{\boldsymbol{A}} - j\omega\mu\sigma\dot{\boldsymbol{A}} = -\mu\dot{\boldsymbol{J}} \quad (1)$$

where $\dot{\boldsymbol{J}}$ is the phasor of the current density, μ and σ are the material permeability and conductivity, respectively, and ω the angular frequency relevant to the frequency f of the current. The model is solved for each applied frequency, in this case, 2 frequencies were considered. Finally, the power in the disk is evaluated using the superposition approach.

In turn, the thermal problem is solved in the steady-state condition, considering the power density in the disk computed after the magnetic field analysis, to calculate the heat source. Specifically, the thermal problem is solved only in the graphite disk domain and

takes into account both convective and radiative contributions to heat exchange with the surrounding environment. The power density $\sigma\omega^2 A^2$ calculated by solving the magnetic problem is the thermal source, i.e., the Joule term in the Fourier equation. The latter is used to evaluate the distribution of the temperature T in the material [27]:

$$-\nabla(k\nabla \mathrm{T}) = \omega^2\sigma\|\dot{\boldsymbol{A}}\|^2 \tag{2}$$

where k is the thermal conductivity of the material. Along the disk boundary, the thermal model is subject to the following boundary conditions:

$$\frac{\partial T}{\partial n} = 0 \quad at \quad r = 0 \tag{3}$$

where n is the normal direction and

$$-k\frac{\partial T}{\partial n} = h(T - T_0) + \varepsilon k_B\left(T^4 - T_0^4\right) \tag{4}$$

where k_B is the Stefan-Boltzmann constant, T_0 is the external temperature, and h and ε are the convective exchange coefficient and emissivity coefficient, respectively. As far as the convective exchange coefficient h and emissivity coefficient ε are concerned, their identification is not trivial due to several factors affecting them, for instance, the surface temperature. For the purpose of modeling, values of h = 10 $\mathrm{Wm^{-2}K^{-1}}$ and ε = 0.8 have been considered, identified by means of some measurement experiences on a prototype. The external temperature, T_0, is set to 850 °C. The coupled problem is solved with the commercial software COMSOL 6.0 [28]. The values of the electrical, magnetic, and thermal characteristics of the model materials are shown in Table 1.

Table 1. Material properties.

Material	Electrical Properties	Magnetic Properties	Thermal Properties
Disk Graphite	$\sigma = 1.04 \times 10^5$ S/m	$\mu_r = 1$	K = 150 W/(m·K)
Coil Copper	$\sigma = 5.998 \times 10^7$ S/m	$\mu_r = 1$	-
Ring Ferrite	$\sigma = 10^{-9}$ S/m	$\mu_r = 1500$	-

The mesh used for solving the FE problem is composed of both mapped domains and domains with tetrahedral mesh; it exhibits around 19,000 elements with a quadratic variation in potential. Figure 2 shows details of the mesh. For the graphite and ferrite parts, a mapped mesh has been used, which means that a structured mesh composed of rectangular elements discretizes relevant sub-domains. Typical magnetic field and temperature maps are shown in Figure 3.

The solution of a Finite Element model with two coils fed by currents at different frequencies is not straightforward [29], hence, the superposition approach has been chosen. Two magnetic simulations have been implemented, each with just one coil on, fed by a single frequency. This way, the mutual inductance between the two coils is neglected.

The power densities calculated from each simulation are superposed and they are the heat source for the thermal simulation: only one thermal simulation is subsequently solved. The superposition is possible if the following relationship between the two frequencies is fulfilled:

$$m_1 f_1 = m_2 f_2 \tag{5}$$

where f is the frequency of each coil and m_1 and m_2 are positive integer numbers.

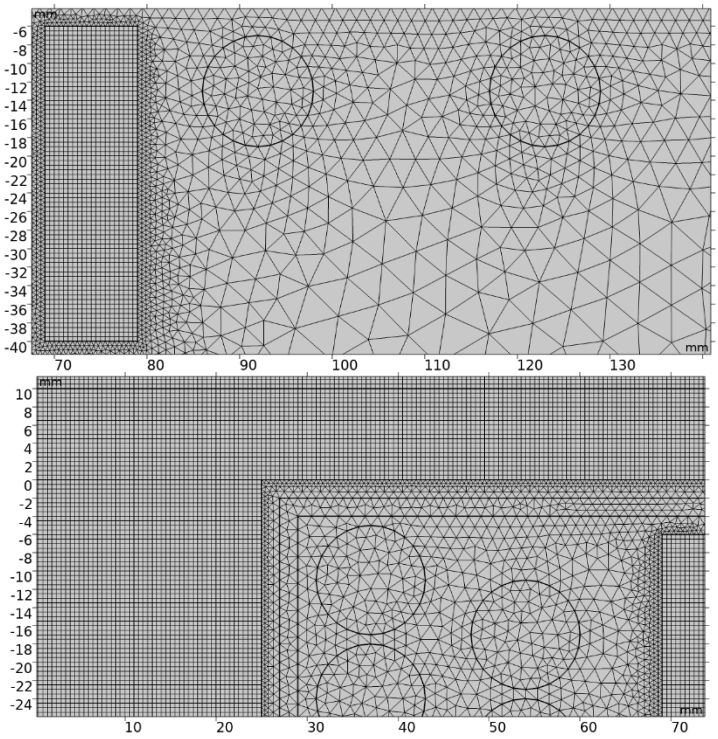

Figure 2. Mesh detail of ferrite and six turns (**bottom**), and mesh detail of the susceptor and the disk (**top**).

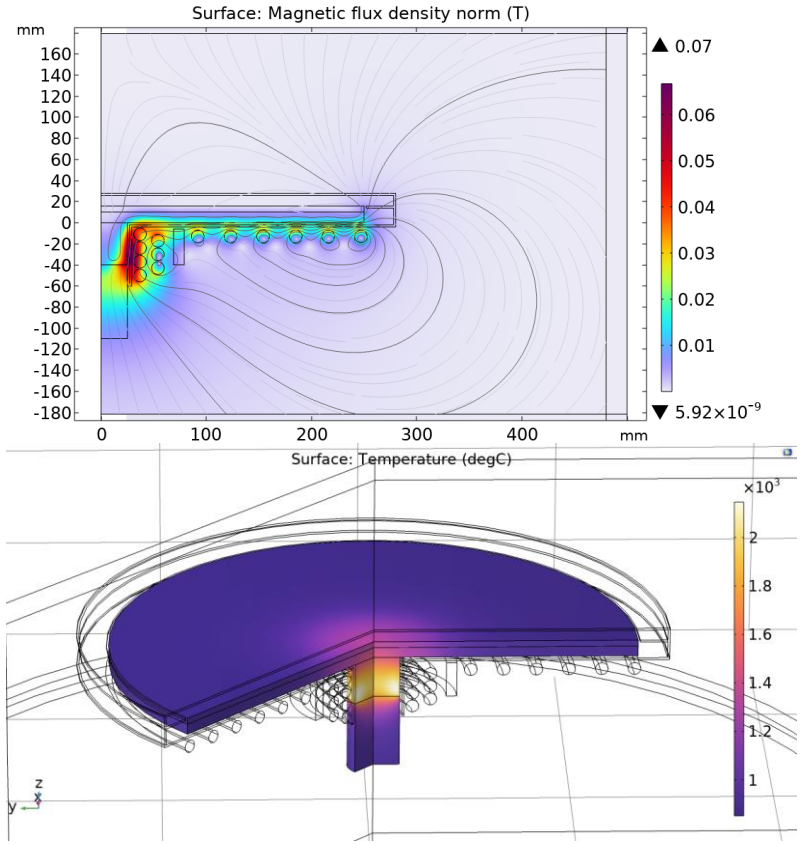

Figure 3. Colored map of the norm of the magnetic flux density in Tesla (T) (**top**), and Temperature (°C) field map (**bottom**).

2.3. Surrogate Model

In order to solve the coupled-field problem in a fast way, a surrogate model is implemented. Specifically, for predicting the temperature on the disk profile and knowing the amplitudes and frequencies of the current inductors, three methods are investigated: Support Vector Regression (SVR), Gaussian Process Regression (GPR), and a fully connected Neural Network (NN).

The SVR predicts the outputs as:

$$f(x^*) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x^*) + b \tag{6}$$

where α_i, α_i^* are the variables determined during training, $K(x_i, x^*)$ is the kernel function, b is the bias term, which will be optimized during training, and x_i is a support vector (subset of the training vector points that are on or outside the epsilon margin). In the paper, the cubic kernel is utilized:

$$K(x_i, x^*) = \left(x_i^T x^*\right)^3 \tag{7}$$

where x^* is the input and x_i is a training point. The epsilon-insensitive loss function is used:

$$L(y, \hat{y}) = \begin{cases} 0 & if\ |y - \hat{y}| \leq \epsilon \\ |y - \hat{y}| - \epsilon & \text{otherwise} \end{cases} \tag{8}$$

where y is the true value, $\hat{y} = f(x)$ is the predicted value, and ϵ is the tolerance margin and the input features were standardized to zero mean and unit variance before training. The training procedure solves a dual optimization problem to find α_i, α_i^* [30]. The SVR algorithm used in this study is available in Matlab 2024b [31].

In this study, GPR was implemented with a constant basis function (β_0) and an isotropic exponential kernel function, which defines the covariance:

$$K(x, x\prime) = \sigma_f^2 e^{\left(-\frac{\|x - x\prime\|}{l}\right)} \tag{9}$$

where σ_f^2 is the signal variance and l is the length scale and the same length scale is used for all input dimensions. The exponential kernel is commonly used in GPR because it is computationally efficient. GPR's objective function is an optimization process which focuses on optimizing the hyperparameters (σ_f^2, l, noise variance σ_n^2) and the basis function (β_0) by maximizing the log marginal likelihood [32]. Using training data $\{X, y\}$, for a new point x^*, the predicted average and variance are:

$$\mu(x^*) = \beta_0 + k^T \left[K + \sigma_n^2 I\right]^{-1} (y - \beta_0) \tag{10}$$

$$\sigma^2(x^*) = K(x^*, x^*) - k^T \left[K + \sigma_n^2 I\right]^{-1} k \tag{11}$$

where the k is the covariance vector between x^* and the training points, K is the covariance matrix of the training data, and $K(x^*, x^*)$ is the covariance of x^* with itself.

The GPR algorithm used in this study is available in Matlab [33].

The neural network that is used in this study for the forward model is a fully connected deep neural network (DNN) with three hidden layers with 100 neurons in each one. In all layers, the ReLU (Rectified Linear Unit) activation function is used:

$$ReLU(x) = \max(0, x), \tag{12}$$

The loss function is the well-known mean squared error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{13}$$

where y_i is the actual value and $\hat{y}_i$ is the predicted value. LBFGS (Limited-memory Broyden–Fletcher–Goldfarb–Shanno) is chosen for the solver, which utilizes an estimation of the Hessian matrix. This solver is selected because it has some of the advantages of Newton-like methods while it has a much lower computational cost [34]. The maximum number of iterations for training is set to 1000.

A database, namely DB1, based on the FE model, is built for training purposes. It is composed of N_s = 5832 samples. Each sample contains two vectors, one with the current amplitudes and frequencies and another with the relevant N_p = 250 temperature values along the α line (see Figure 1). All-in-one N_s couples of vectors compose the database.

Each FE magneto-thermal analysis takes about 147 s to be solved on a computer equipped with an Intel(R) Core (TM) i7-10700KF CPU @ 3.80GHz, 128 GB of RAM, and an NVIDIA Quadro RTX-4000 graphics card. Building the whole DB1 database takes about 266 h.

The database is divided into two subsets, following the rule of thumb 70-15-15 i.e., 70% (N_t = 4082 couples of vectors) are utilized for training, 15% (N_v = 875 couples of vectors) are utilized for validation, and 15% (N_e = 875 couples of vectors) are utilized for testing.

In order to evaluate the quality of the prediction, the Mean Absolute Percentage Error MAPE (%) was calculated considering the $N = N_v + N_e$ vectors of temperature values of both validation and test set:

$$\text{MAPE} = 100\frac{1}{N}\sum_{j=1}^{N}\frac{1}{N_r}\sum_{i=1}^{N_p}\frac{\left|\hat{Y}_{j,i} - Y_{j,i}\right|}{\left|Y_{j,i}\right|} \tag{14}$$

where Y is the vector of N_r true values calculated with the FE model, and $\hat{Y}$ is the vector of N_r values predicted by the relevant surrogate model. The Root Mean Square Error (RMSE) and R-squared are also evaluated:

$$\text{RMSE} = \sqrt{\sum_{j=1}^{N}\sum_{i=1}^{N_p}\frac{\left(\hat{Y}_{j,i} - Y_{j,i}\right)^2}{NN_p}} \tag{15}$$

$$R^2 = 1 - \frac{\sum_{j=1}^{N}\sum_{i=1}^{N_p}\left(\hat{Y}_{j,i} - Y_{j,i}\right)^2}{\sum_{j=1}^{N}\sum_{i=1}^{N_p}\left(\hat{Y}_{j,i} - Y_{j,mean}\right)^2} \tag{16}$$

They are computed on the N vectors of temperature values used for the model test or train, with values normalized in the range $[-1, 1]$.

3. Inverse Problem

The inverse problem reads as follows: knowing the geometries and the material properties of the susceptor and given a prescribed temperature profile, identify the coil parameters in terms of currents (I_1 and I_2) and frequencies (f_1 and f_2). To solve this problem, two different approaches have been selected; the first is based on a multi-objective formulation of the problem, while the second is a procedure with two problems solved in cascade (an optimal design and an optimal control problem, sequentially). For process purposes, a homogenous temperature profile is needed. Specifically, for the sake of an example, considering the silicon crystal growth process, the temperature profile should be as much as possible homogeneous and equal to the target temperature T_t = 1100 °C.

3.1. First Approach: Multi-Objective Optimization Problem

The inverse problem is formulated in terms of a multi-objective optimization problem, based on the GPR surrogate model for the evaluation of the objective functions. Specifically, two objective functions have been considered and minimized according to the Pareto optimization [20]:

$$v_1 = N\{1\ if\ T_i : |T_i - T_t| > \alpha\}, \mathrm{i} = 1, \ldots, 200 \tag{17}$$

with

$$\alpha = \frac{\delta\ T_t}{100}, \tag{18}$$

and $\delta = 1\%$, while

$$v_2 = \max_i T_i - \min_i T_i \tag{19}$$

The objective function v_1 is the number of points along the reference line (Figure 1), the temperature of which falls out of the range $\pm 1\%$ with respect to the target temperature T_t of 1100 °C, to be minimized. The objective function v_2 is the discrepancy between maximum and minimum values of the actual temperature profile, to be minimized. So, with the first objective function, the algorithm tries to find a temperature profile that is as close to 1100 °C as possible, and with the second, the algorithm tries to find a homogeneous response.

This multi-objective optimization is performed by utilizing the well-known Non-Dominated Sorting Genetic Algorithm NSGA-II [21] available from the Mathworks file exchange website [35]. It is a genetic optimizer that exploits the concept of solution ranking by means of the non-dominated sorting algorithm: the population is subdivided into several Pareto fronts and a fitness value is assigned to each front depending on the vicinity to the Utopia point (the closer, the better). Within each Pareto front, a fitness value is assigned to each solution, depending on the crowding distance i.e., the distance of the current solution from the closest solutions (the wider, the better). Iteration after iteration, by applying crossover, mutation, and tournament selection operators, just one Pareto front results, and the solutions are usually equally distributed along the front.

3.2. Second Approach: Deep Neural Network Approach

Results from multi-objective optimization prove the solvability of the inverse problem, which is now reformulated by means of an innovative approach, based on Deep Neural Networks (DNNs). To this end, two auxiliary inverse problems are defined and subsequently solved: an optimal design problem, first and an optimal control problem, second. In Figure 4, the process diagram of this procedure is depicted.

3.2.1. Optimal Design Problem

The optimal design problem reads as follows: given the prescribed temperature profile, the current and frequency (I_{2OD}, f_{2OD}) of the second coil are identified.

The solution to this problem is found using a DNN, implemented in Matlab, and trained utilizing a database, here called DB2, that contains a subset of samples of the database DB1 used for training the forward models. Specifically, DB2 contains the solutions obtained by filtering DB1 in a way that retains the solutions that guarantee an approximately uniform heating of the central part of the disk. The rationale behind this choice is that heating the central region is not straightforward due to the presence of the axial hole of the disk.

Solving the optimal design problem, we obtain the parameters of coil 2, mainly responsible for heating the outer part of the disk, namely f_{2OD} and I_{2OD}.

Moreover, it is possible to investigate the solutions in DB2 to search, e.g., for the most likely value of frequency f_1; in this case, the average value is considered. This value is called f_{1OD} and it is taken for granted.

Finally, at the end of the optimal design problem, the values f_{1OD} and f_{2OD} for both the frequencies f_1 and f_2 are identified. Specifically, frequency f_{1OD} comes from searching for the most frequent value identified in DB2, while f_{2OD} comes from the results of the optimal design problem.

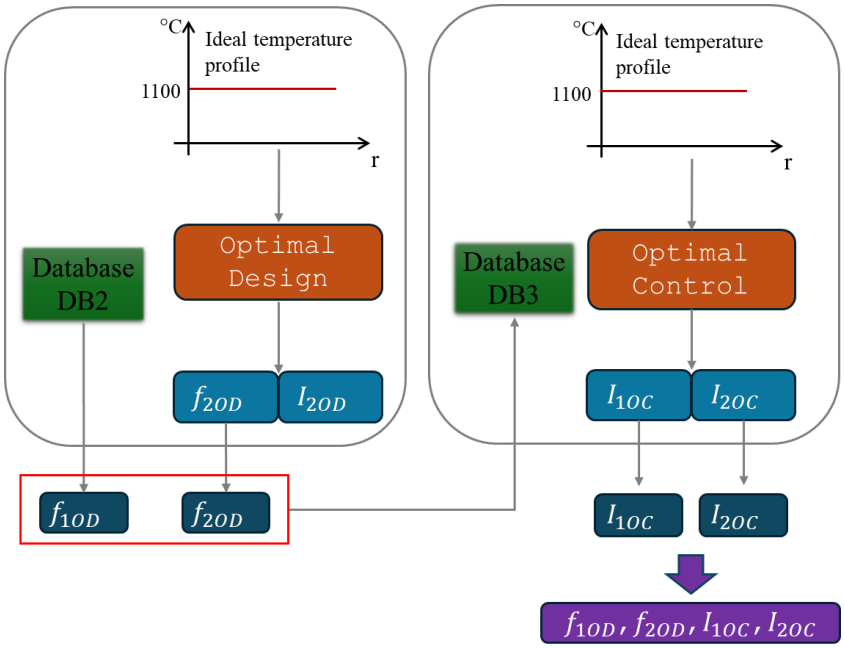

Figure 4. Cascade procedure: optimal design followed by optimal control problems are solved.

3.2.2. Optimal Control Problem

In terms of a global optimum, the current control problem is eventually considered and solved. The problem can be described as follows: for a given temperature profile and frequencies f_{1OD} and f_{2OD}, find the current amplitudes I_1 and I_2 such that the discrepancy between actual and prescribed temperature profile is minimal.

Knowing the optimal values f_{1OD} and f_{2OD} from the optimal design problem, a new database DB3 is prepared. Specifically, new samples are generated utilizing the forward surrogate model GPR, by varying the currents of the two coils, with constant frequencies f_{1OD} and f_{2OD}. Each sample of DB3 contains two vectors, one with the current amplitudes I_1 and I_2 and another with the relevant $N_p = 250$ temperature values along the α line (see Figure 1).

The solution to this problem is found using a DNN, implemented in Matlab and trained utilizing the database DB3. The solution to this problem is given by the couple I_{1OC} and I_{2OC} of the optimal current intensity of the two coils.

At the end of this Deep Neural Network approach, the four optimal parameters f_{1OD}, f_{2OD}, I_{1OC}, and I_{2OC} are found with the relevant optimal temperature profile (see Figure 4).

4. Results

4.1. Surrogate Models

The accuracies of the investigated surrogate models are shown in Table 2, in terms of the three metrics R-squared, RMSE, and MAPE. Specifically, the three metrics are evaluated

for the test set. In Table 2, the training time is shown. All three algorithms have been trained on a computer equipped with an Intel(R) Core (TM) i7-10700KF CPU @ 3.80GHz, 128 GB of RAM, and an NVIDIA Quadro RTX-4000 graphics card.

Table 2. Results of surrogate models.

Parameter	SVR	NN	GPR
R-square	0.9822	0.9991	0.9999
RMSE	41.116	6.149	0.292
MAPE (%)	1.945	0.290	0.009
Time (h)	26	7	54

Figure 5 shows an example of the temperature profile prediction using surrogate models compared to the actual profile obtained using Finite Element simulation.

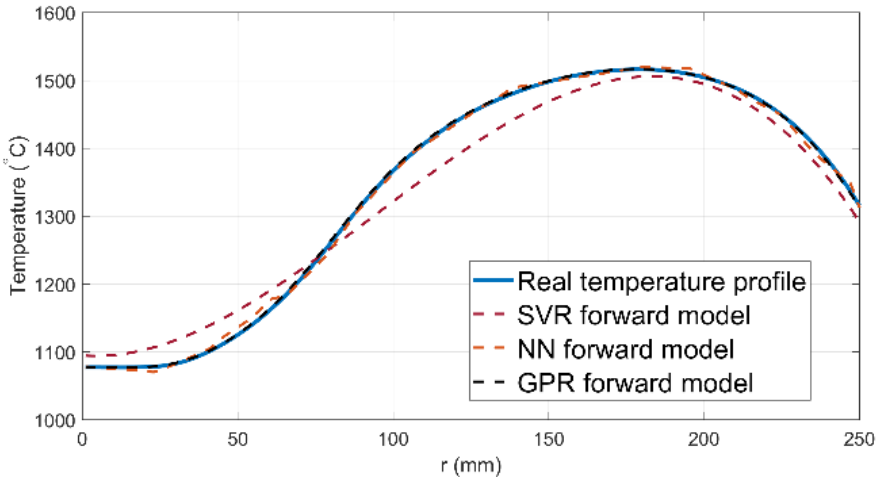

Figure 5. True and predicted temperature profiles for input vector [f_1 = 70.589 kHz, I_1 = 139 Arms, f_2 = 2.949 kHz, I_2 = 1.275 kArms].

Based on the results, the Gaussian process regression (GPR) model was selected as the surrogate field model to use in the subsequent approaches. The computational time for one call to GPR was 7 s.

4.2. First Approach: Multi-Objective Optimization Problem

The NSGA-II algorithm was applied, with 50 generations, 50 individuals, and a random starting population. The following parameters were set up in the algorithm: size of mating pool equal to 25, tournament size equal to 2, probability of crossover equal to 90%, probability of mutation equal to 10%, and distribution indices for both crossover and mutation equal to 0.002.

Figure 6 shows the first and last generations of the optimization. The two endpoints P_1 (minimum of v_1) and P_2 (minimum of v_2) and the knee point P_3, which is a tradeoff between P_1 and P_2, are highlighted in Figure 6, and the relevant temperature profiles are depicted in Figure 7.

In Figure 7, the gray zone indicates the range of ±1% of 1100 °C (the target temperature).

The design variables of points P_1, P_2, and P_3 are given in Table 3.

The whole optimization process, based on the GPR surrogate model, takes about 327 h.

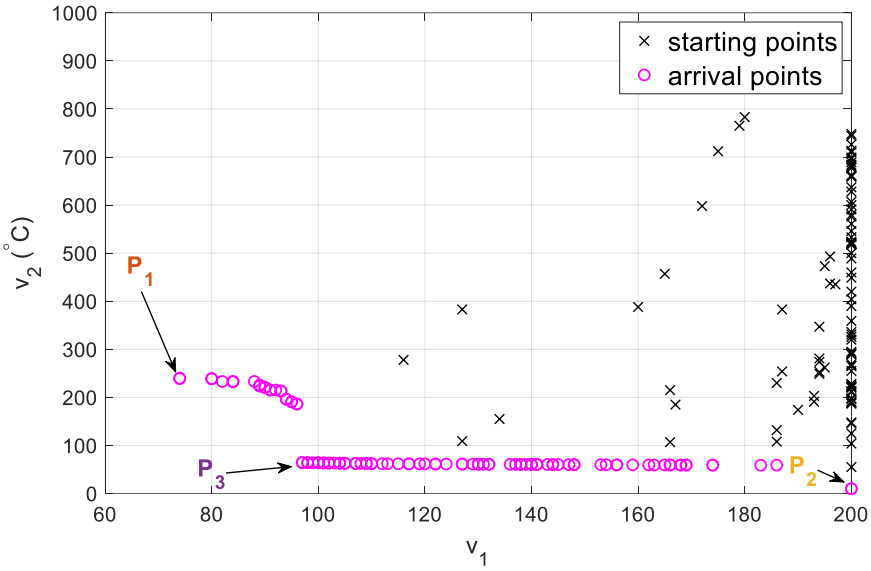

Figure 6. Starting and arrival population of the multi-objective optimization.

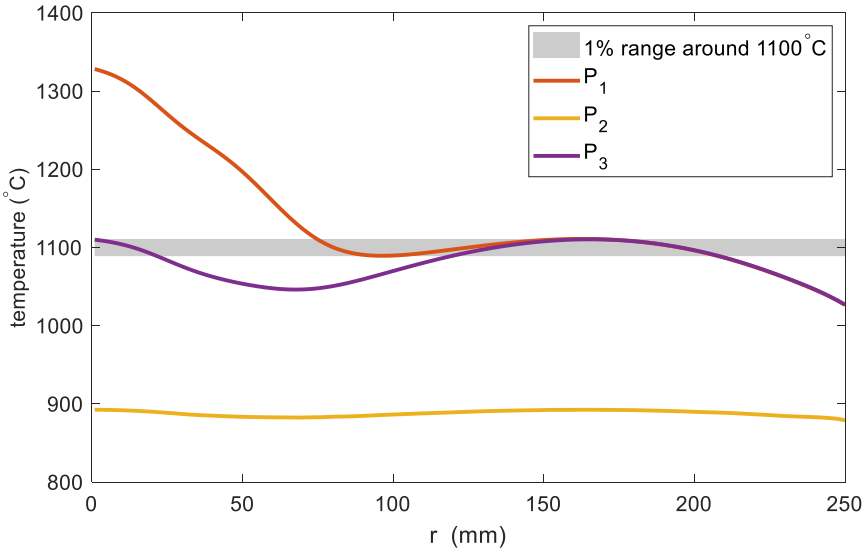

Figure 7. Temperature profiles of 3 optimal individuals P_1, P_2, and P_3 (see Figure 6).

Table 3. Optimal points of the Pareto front.

Points	v_1	v_2 (°C)	f_1 (kHz)	I_1 (A)	f_2 (kHz)	I_2 (A)
P_1	74	238.95	100	305	1	1137
P_2	200	10.02	20	100	1.04	400
P_3	97	64.16	20	288	1.01	1134

4.3. Second Approach: Deep Neural Network Approach

The optimal design problem was solved using a fully connected DNN with 6 layers with 80, 40, 30, 20, 10, and 8 neurons per layer from the first to the last layer, respectively. The following statement is assumed: the input of the NN is the prescribed temperature profile and the outputs are the current and frequency of the second coil. The NN is trained utilizing the database DB2. Specifically, DB2 contains those solutions with temperature profiles around 1100 °C with ±5% range, considering the first 25 mm of the graphite disk. In fact, the main effect of the first coil is on the first 25 mm, and it guarantees that for all

DB2 samples (of which there are 998), the first coil acts in the selected range. In DB2, the average f_{1od} of the frequency f_1 of the samples is considered and it is equal to 38.5 kHz.

The DNN is trained using 70% of samples for training, 15% for validation, and 15% for testing. The RMSE and MAPE errors, calculated with values normalized in the range $[-1, 1]$, for the validation and test sets, are 418 and 9.72%, respectively. In Figure 8, the predicted versus actual values of the validation and test sets are plotted.

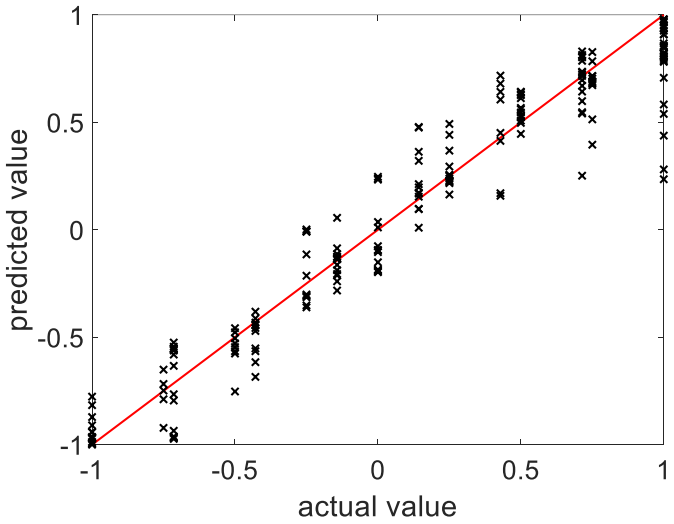

Figure 8. Optimal design DNN actual values vs. predicted values (test and validation sets).

By solving the optimal design problem with the prescribed temperature profile being homogenous and equal to 1100 °C, the optimal values of f_2 and I_2 are identified: f_{2od} = 1.38 kHz and I_{2od} = 1.02 kA.

Once the optimal design problem has been solved, the optimal control problem is handled.

Knowing the optimal values of f_1 (being the average of the frequencies in DB2) and f_2 (obtained by solving the optimal design problem), the database DB3 is prepared. Specifically, new samples are generated utilizing the forward surrogate model GPR, considering f_{1od} = 38.5 kHz, f_{2od} = 1.38 kHz, and the current varying between 100 and 400 A for the first coil and between 400 and 2000 A for the second one. This way, the new database DB3 with 2665 samples is created.

The optimal control problem is solved using a fully connected NN with 6 layers with 80, 40, 30, 20, 10, and 8 neurons per layer from the first to the last layer, respectively. The following statement is assumed: the input of the NN is the prescribed temperature profile and the outputs are the current intensity of the two coils.

The DNN is trained using 70% of samples for training, 15% for validation, and 15% for testing. The RMSE and MAPE errors, calculated with values normalized in the range $[-1, 1]$, for the validation and test sets, are 1.95 and 0.21%, respectively. In Figure 9, the predicted versus actual values of the validation and test sets are plotted.

By solving the optimal control problem with the prescribed temperature profile being homogenous and equal to 1100 °C, the optimal values of I_1 and I_2 are identified: I_{1oc} = 255 A and I_{2oc} = 996 A. Based on the result of the cascade approach, the optimal temperature profile is shown in Figure 10.

The whole cascade approach takes about 330 h.

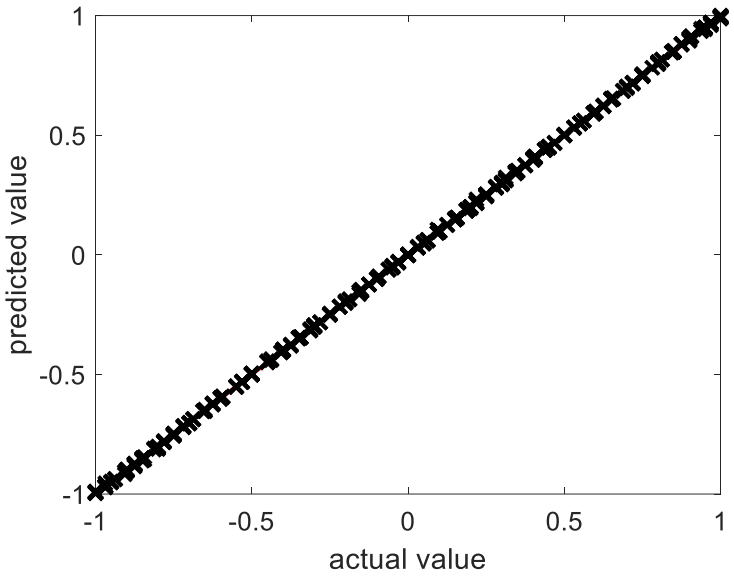

Figure 9. Optimal control DNN actual values vs. predicted values (test and validation sets).

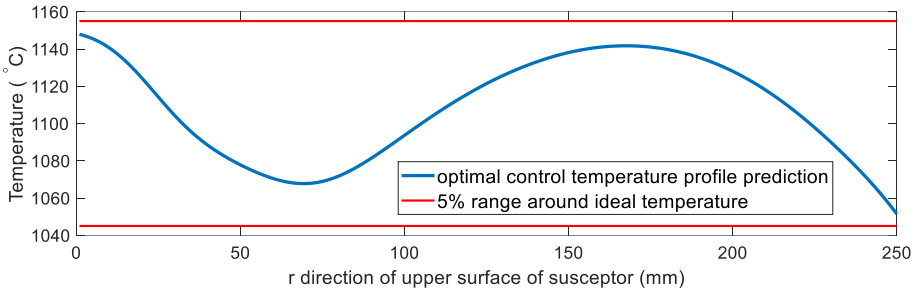

Figure 10. Optimal temperature profile obtained using the DNN approach [f_{1od} = 38.5 kHz, I_{1od} = 255 A, f_{2oc} = 1.38 kHz, I_{2oc} = 996 A].

5. Discussion

The three surrogate models trained for solving the coupled field problem are able to provide a rather accurate temperature profile along a line of the graphite disk, specifically, the α line, with known source terms i.e., magnitude and frequency of the current of each of the two inductors.

Considering the evaluation parameters in Table 2, they are in agreement for each surrogate model in terms of accuracy: SVR shows the worst performance, with its R-squared being the lowest among the three methods and its RMSE and MAPE being the highest. Overall, the performance of SVR is the worst for our case study. The NN model shows values of R-squared, RMSE, and MAPE in the middle of the three methods. The best method is GPR because its R-squared is the closest to 1 among the three methods and its RMSE and MAPE errors are the lowest.

This performance can be also assessed from the example shown in Figure 5, where the temperature profile for one case, approximated by means of the three methods, is shown and compared to the one calculated by means of the Finite Element method. The worst approximation is given by the SVR model, while NN and GPR better approximate the actual profile. However, the temperature profile given by the NN is less smooth than the one approximated by GPR, at a few points.

Considering the training time, the GPR model is the most expensive in terms of computational burden, while the NN model is the lightest.

However, based on the accuracy, the GPR model is the best one, therefore, it was used as the surrogate model in the solution of the inverse problem, applied with both of the two approaches.

The multi-objective approach found a Pareto front, depicted in Figure 6, which can be divided into two parts. The right part of the front, characterized by low v_2 values, is representative of homogeneous temperature profiles, with a small difference between the maximum and lowest values along the profile. However, these profiles have many points outside the gray zone ($\pm 1\%$ of 1100 °C, see Figure 7), which means high values for the objective function v_1. An example of this behavior is shown by point P_2 in Figure 7: a rather flat temperature profile, far from the target temperature.

Conversely, the left part of the front is characterized by low values of v_1, which means a small number of points of the temperature profile outside the gray zone, but high v_2 values, i.e., a high discrepancy between the minimum and maximum values of the temperature profile. The temperature profile of point P_1, belonging to this part of the front, is shown in Figure 7: it is not a flat profile but many of its points are close to the target temperature.

Finally, point P_3 is a trade-off between P_1 and P_2, being in the middle, also called the knee-point, of the front. P_3 shows a rather homogenous profile, as depicted in Figure 7, and it is close to the target temperature.

Looking at Table 3, where the design variables of the three solutions P_1, P_2, and P_3 are shown, it can be noted that the two solutions with a rather homogenous profile, P_2 and P_3, show a low value of frequency f_1 (inner coil). While they have a significantly different current I_1 (higher for P_3, which has higher temperatures close to the disk center, specifically close to the target temperature). Considering the outer coil, points P_1 and P_3, which have temperature profiles close to the target temperature, have a higher current I_2 with respect to point P_2, which has a temperature profile far from the target temperature. The frequency f_2 has roughly the same value for each of the three points. Overall, it can be stated that the best solution obtained by this approach is solution P_3.

In terms of computational costs, the entire process based on GPR takes about 327 h. If the same optimization process were based on FEAs, it would take about 150 h. Therefore, using the surrogate model might seem inefficient; however, the following points should be considered. With a surrogate model, the optimization problem can be run multiple times, considering different populations or iterations, with each objective function call taking about 7 s. This flexibility is not possible with FEAs, as each FEA run takes about 147 s. Additionally, the surrogate model can be applied to other design problems using the same device. Specifically, the surrogate model is a field model that provides the temperature profile given the source. If a different target temperature or temperature profile is required, another optimization problem can be solved, and the same GPR model remains valid. In this case, only the optimization cost needs to be considered, which takes approximately 7 h to solve and obtain the optimal solutions.

The deep learning approach is divided into two problems solved in cascade. The first problem is the optimal design problem, which returns the optimal parameters of coil 2, given a prescribed temperature profile. The trained DNN used for solving this problem shows some discrepancies between true and predicted values, as shown in Figure 8, probably because of the difficulty in identifying parameters of coil 2; the reason could be that the problem is less sensitive to those parameters. However, after solving the optimal design problem, the values that need to be included with the deep learning approach are the two frequencies f_{1od} and f_{2od}, the latter being the output of the DNN, while f_{1od} was found after a manual search among DB2 solutions. These two optimal frequencies are close to the frequencies found by the multi-objective approach for points P_2 and P_3, which show a rather flat temperature profile.

The second problem is the optimal control problem, which is solved using a DNN trained using database DB3, based on f_{1od} and f_{2od}. The optimal control problem returns the optimal currents I_{1oc} and I_{2oc} of the two coils. The values so found (I_{1oc} = 255 A and I_{2oc} = 996 A) are close to those found by the multi-objective approach for point P_3 (I_1 = 288 A and I_2 = 1134 A). The relevant temperature profile is all inside the range of ±5% of 1100 °C (target temperature).

In terms of computational time, it is worth noting that the multi-objective approach and the deep learning approach take approximately the same time (327 vs. 330 h) to obtain the optimal solutions.

6. Conclusions

The problem of optimizing the uniformity of the temperature profile in a dual-frequency susceptor was successfully solved by means of surrogate models and DNNs. Specifically, the field problem was approximated by means of Support Vector Regression SVR, Deep Neural Network DNN, and Gaussian Process Regression GPR models. SVR was the least accurate and the training was time-consuming, while DNN was fast to train but less accurate than GPR. In turn, GPR was the most accurate; so, this model was chosen to be used as the surrogate forward model in a multi-objective optimization problem for solving the inverse problem (to find the coil parameters that give uniform temperature profile) with good results.

Finally, the inverse problem was solved by means of a procedure based on DNNs and the GPR surrogate model.

Author Contributions: Conceptualization, P.D.B. and F.D.; methodology, P.D.B., M.F. and M.E.M.; software, A.G. and E.S.; validation, M.F. and M.E.M.; formal analysis, P.D.B. and M.F.; investigation, A.G., M.E.M. and E.S.; data curation, A.G. and E.S.; writing—original draft preparation, A.G. and E.S.; writing—review and editing, M.F. and M.E.M.; visualization, A.G.; supervision, P.D.B. and F.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Dataset available on request from the authors.

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

References

1. Baake, E.; Nacke, B. Efficient Heating by Electromagnetic Sources in Metallurgical Processes: Recent Applications and Development Trends. *Przegląd Elektrotech.* **2010**, *86*, 11–14.
2. Rudnev, V.; Loveless, D.; Cook, R. Handbook of Induction Heating. In *Manufacturing Engineering and Materials Processing*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2017; ISBN 978-1-4665-5395-8.
3. Rudnev, V.; Totten, G.E. Induction Heating of Selective Regions. In *Induction Heating and Heat Treatment*; ASM International: Materials Park, OH, USA, 2014; pp. 346–358, ISBN 978-1-62708-167-2.
4. Fisk, M. Induction Heating. In *Encyclopedia of Thermal Stresses*; Hetnarski, R.B., Ed.; Springer: Dordrecht, The Netherlands, 2014; pp. 2419–2426, ISBN 978-94-007-2738-0.
5. Rapoport, E.; Pleshivtseva, Y. *Optimal Control of Induction Heating Processes*; Mechanical Engineering; CRC/Taylor & Francis: Boca Raton, FL, USA, 2007; ISBN 978-0-8493-3754-3.
6. Brazhnik, D.S.; Bolotin, K.E. Different Approaches to Taking Joule Heat into Induction Heating of Graphite Crucible. In Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), St. Petersburg and Moscow, Russia, 27–30 January 2020; pp. 616–618.
7. Mannanov, E.; Galunin, S. Numerical Simulation of the Induction Heating Process of a Disk Profile. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *643*, 012065. [CrossRef]
8. Fisk, M.; Ristinmaa, M.; Hultkrantz, A.; Lindgren, L.-E. Coupled Electromagnetic-Thermal Solution Strategy for Induction Heating of Ferromagnetic Materials. *Appl. Math. Model.* **2022**, *111*, 818–835. [CrossRef]

9. Jankowski, T.A.; Pawley, N.H.; Gonzales, L.M.; Ross, C.A.; Jurney, J.D. Approximate Analytical Solution for Induction Heating of Solid Cylinders. *Appl. Math. Model.* **2016**, *40*, 2770–2782. [CrossRef]
10. Di Barba, P.; Dughiero, F.; Lupi, S.; Savini, A. Optimal Shape Design of Devices and Systems for Induction-Heating: Methodologies and Applications. *COMPEL Int. J. Comput. Math. Electr. Electron. Eng.* **2003**, *22*, 111–122. [CrossRef]
11. Gong, R.; Tang, Z. Training Sample Selection Strategy Applied to CNN in Magneto-Thermal Coupled Analysis. *IEEE Trans. Magn.* **2021**, *57*, 1–4. [CrossRef]
12. Di Barba, P.; Mognaschi, M.E.; Cavazzini, A.M.; Ciofani, M.; Dughiero, F.; Forzan, M.; Lazzarin, M.; Marconi, A.; Lowther, D.A.; Sykulski, J.K. A Numerical Twin Model for the Coupled Field Analysis of TEAM Workshop Problem 36. *IEEE Trans. Magn.* **2023**, *59*, 1–4. [CrossRef]
13. Naar, R.; Bay, F. Numerical Optimisation for Induction Heat Treatment Processes. *Appl. Math. Model.* **2013**, *37*, 2074–2085. [CrossRef]
14. Karban, P.; Kotlan, V.; Dolezel, I. Numerical Model of Induction Shrink Fits in Monolithic Formulation. *IEEE Trans. Magn.* **2012**, *48*, 315–318. [CrossRef]
15. Zgraja, J. The Optimisation of Induction Heating System Based on Multiquadric Function Approximation. *COMPEL Int. J. Comput. Math. Electr. Electron. Eng.* **2005**, *24*, 305–313. [CrossRef]
16. Amestoy, P.; Buttari, A.; Joslin, G.; L'Excellent, J.-Y.; Sid-Lakhdar, M.; Weisbecker, C.; Forzan, M.; Pozza, C.; Perrin, R.; Pellissier, V. Shared-Memory Parallelism and Low-Rank Approximation Techniques Applied to Direct Solvers in FEM Simulation. *IEEE Trans. Magn.* **2014**, *50*, 517–520. [CrossRef]
17. Chen, H.-C.; Huang, K.-H. Finite Element Analysis of Coupled Electromagnetic and Thermal Fields within a Practical Induction Heating Cooker. *Int. J. Appl. Electromagn. Mech.* **2008**, *28*, 413–427. [CrossRef]
18. Vassent, E.; Meunier, G.; Sabonnadier, J.C. Simulation of Induction Machine Operation Using Complex Magnetodynamic Finite Elements. *IEEE Trans. Magn.* **1989**, *25*, 3064–3066. [CrossRef]
19. Forzan, M.; Maccalli, G.; Valente, G.; Crippa, D. Design of an Innovative Heating Process System for the Epitaxial Growth of Silicon Carbide Layers Wafer. In Proceedings of the MMP-Modelling for Material Processing, Riga, Latvia, 8–9 June 2006.
20. Di Barba, P.; Dughiero, F.; Forzan, M.; Mognaschi, M.E.; Sieni, E. New Solutions to a Multi-Objective Benchmark Problem of Induction Heating: An Application of Computational Biogeography and Evolutionary Algorithms. *Arch. Electr. Eng.* **2018**, *67*, 139–149. [CrossRef] [PubMed]
21. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]
22. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; Wiley-Interscience Series in Systems and Optimization; Wiley: Chichester, UK; Weinheim, Germany, 2004; ISBN 978-0-471-87339-6.
23. Zitzler, E.; Thiele, L. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. *IEEE Trans. Evol. Comput.* **1999**, *3*, 257–271. [CrossRef]
24. Srinivas, N.; Deb, K. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evol. Comput.* **1994**, *2*, 221–248. [CrossRef]
25. Meunier, G. *The Finite Element Method for Electromagnetic Modeling*; ISTE: London, UK; Wiley: Hoboken, NJ, USA, 2008; ISBN 978-1-84821-030-1.
26. Binns, K.J.; Lawrenson, P.J.; Trowbridge, C.W. *The Analytical and Numerical Solution of Electric and Magnetic Fields*; Wiley: Chichester, UK, 1992; ISBN 0-471-92460-1.
27. Carslaw, H. *Conduction of Heat in Solids*, 2nd ed.; Clarendon Press: New York, NY, USA; Oxford University Press: Oxford, UK, 1986; ISBN 978-0-19-853368-9.
28. COMSOL. Multiphysics Software for Optimizing Designs. Available online: https://www.comsol.com/ (accessed on 11 November 2024).
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. Smola, A.J.; Schölkopf, B. A Tutorial on Support Vector Regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]
31. Train Kernel Approximation Model Using Regression Learner App. Available online: https://it.mathworks.com/help/stats/train-kernel-approximation-using-regression-learner-app.html (accessed on 24 December 2024).
32. Rasmussen, C.E.; Williams, C.K.I. Adaptive Computation and Machine Learning. In *Gaussian Processes for Machine Learning*; 3. print.; MIT Press: Cambridge, MA, USA, 2008; ISBN 978-0-262-18253-9.
33. Regression Learner. Available online: https://it.mathworks.com/help/stats/regressionlearner-app.html (accessed on 24 December 2024).

34. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; The MIT Press: Cambridge, MA, USA; London, UK, 2016; ISBN 978-0-262-03561-3.
35. NSGA-II: A Multi-Objective Optimization Algorithm. Available online: https://it.mathworks.com/matlabcentral/fileexchange/10429-nsga-ii-a-multi-objective-optimization-algorithm (accessed on 24 December 2024).