POLITECNICO DI TORINO
Repository ISTITUZIONALE

A Physics-Informed Neural Network for the Thermal–Hydraulic Modeling of Normal Operation in Superconducting Coils for Fusion Applications

Original
A Physics-Informed Neural Network for the Thermal–Hydraulic Modeling of Normal Operation in Superconducting Coils for Fusion Applications / Zappatore, A., Sprò, M.. - In: APPLIED SCIENCES. - ISSN 2076-3417. - 16:20(2026). [10.3390/app162010441]

Availability:
This version is available at: 11583/3016868 since: 2026-10-31T11:58:19Z

Publisher:
MDPI

Published
DOI:10.3390/app162010441

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

27 November 2026

applied sciences

MDPI

Article

A Physics-Informed Neural Network for the Thermal–Hydraulic Modeling of Normal Operation in Superconducting Coils for Fusion Applications

Andrea Zappatore * and Mauro Sprò

NEMO Group, Dipartimento Energia, Politecnico di Torino, 10129 Turin, Italy
* Correspondence: andrea.zappatore@polito.it

Abstract

Transient conjugate heat transfer in coupled solid–fluid systems is governed by diffusion, advection, and heat transfer processes that can make repeated simulations expensive and time-consuming in the design phase. This study develops a physics-informed neural network for one-dimensional coupled solid–fluid heat transfer. The model takes space, time, and operating parameters as inputs and predicts the temperature fields in both the solid and fluid domains. The governing equations, heat transfer, boundary conditions, and initial conditions are incorporated into the training through differentiable residual losses, reducing the reliance on high-fidelity training data. The proposed approach is formulated in nondimensional form and evaluated using metrics that quantify conservation, boundary conditions satisfaction, and prediction accuracy. Particular attention is paid to the model behavior within the training range and when operating conditions extend beyond it. The resulting framework provides a basis for the thermal–hydraulic assessment of the performance of superconducting coils for fusion applications in different operating scenarios under evaluation during the design optimization phase using the physics-informed neural network approach. The analyzed case study is the Central Solenoid of the DTT facility currently under construction at ENEA Frascati in Italy.

Keywords: conjugate heat transfer; physics-informed neural networks; reduced-order modeling; superconducting magnet; surrogate modeling; transient heat transfer

1. Introduction

The Divertor Tokamak Test (DTT) facility is a superconducting tokamak being built in Frascati and is intended to address one of the central technological issues on the path toward fusion power plants: the power exhaust of large plasma heat fluxes under reactor-relevant conditions [1]. In order to fulfill this role, DTT has been designed to operate with a high degree of scenario flexibility, including different magnetic configurations and operating conditions relevant to the study of plasma exhaust and divertor solutions. Each candidate plasma scenario corresponds to a specific time evolution of the currents in the superconducting magnet system and, therefore, to a specific distribution of AC losses in the coils.

Among the magnet system components, the Central Solenoid (CS) and Poloidal Field (PF) coils are particularly sensitive to scenario changes because they are operated in the pulsed mode. Their time-varying currents induce AC losses in the cable-in-conduit conductors (CICCs), leading to a transient temperature increase and corresponding reduction in

Check for updates

Academic Editors: Simone Minucci and Alessandro Lampasi

Received: 29 September 2026
Revised: 16 October 2026
Accepted: 20 October 2026
Published: 21 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

the conductor temperature margin [2], i.e., the difference between current sharing temperature and conductor temperature. In DTT, the CS consists of six independently powered Nb_3Sn modules (see Figure 1), whereas the PF system includes both Nb_3Sn and NbTi CICC coils (an example of a relevant Nb_3Sn CICC layout is detailed in [3]), all cooled by forced-flow supercritical helium at 4.5 K, 5 bar. The assessment of the minimum temperature margin in these conductors is therefore a necessary step in the validation of any relevant plasma scenario.

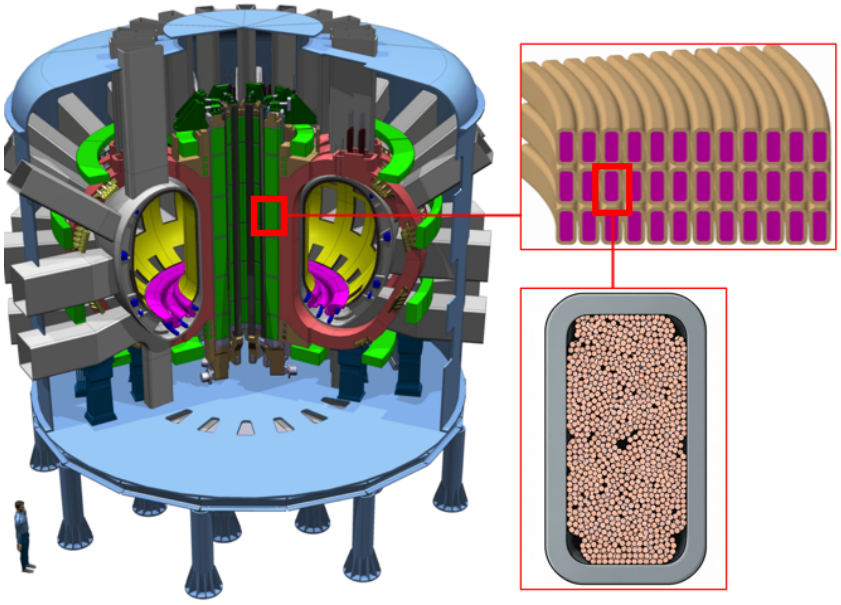

Figure 1. CAD view of the DTT magnet system and a portion of the pancake-wound CS module. Sketch of the CICC cross-section foreseen for the DTT CS, showing the superconducting strands and stainless-steel jacket.

High-fidelity thermal–hydraulic tools, such as the 4C code [4], provide the temperature margin distribution required for the verification of superconducting magnet performance. In previous DTT studies, the 4C code was used to simulate the transient thermal–hydraulic response of the CS and PF coils during reference plasma scenarios, accounting for AC losses, conductor layout, helium cooling, and interturn/interpancake thermal coupling [2]. These tools solve a set of partial differential equations to account for helium flow dynamics, providing the evolution of the distribution of velocity, pressure, and temperature in the fluid, as well as for heat conduction in the solids, providing the evolution of the temperature distribution in the conductor.

However, the same level of detail that makes high-fidelity thermal–hydraulic simulations reliable also makes them poorly suited when many slightly different scenarios need to be assessed. At the present stage of DTT development, the conductor and coil designs have been substantially defined, whereas plasma scenarios and operating waveforms (typically computed with plasma equilibrium codes [5]) may still undergo further refinement. Consequently, while a preliminary magnet sizing study would require parameterizing geometric features, the iterative engineering loop addressed here focuses on evaluating many candidate current waveforms on a fixed geometry. Extending the input vector to include dimensionless geometric groups remains a natural future extension for preliminary reactor sizing. A typical engineering loop may therefore require a new set of coil current waveforms to be transferred from the plasma scenario or magnet design teams to the thermal–hydraulic team, followed by input preparation, execution of detailed simulations over the relevant CS modules and PF coils, extraction of temperature margin maps, and feedback to the scenario designers. Repeating this procedure for several small changes in

the operating scenario can make the simulation turnaround time a significant part of the overall design workflow.

A reduced-order predictive model could therefore provide a useful intermediate step between scenario definition and detailed thermal–hydraulic analysis. Once the conductor geometry, cooling configuration, and material properties are fixed, such a model could rapidly estimate the thermal response to a new current waveform. Its purpose is not to replace high-fidelity simulations for final design verification—it should act as a fast-screening model, capable of identifying whether a proposed current scenario is likely to be thermally acceptable, where the minimum margin is expected to occur, and which time intervals are most critical. A reliable surrogate model would allow for plasma scenario designers to receive immediate feedback on the thermal–hydraulic consequences of waveform changes, reserving full 4C simulations for selected candidate scenarios and final confirmation.

Machine-learning surrogates are natural candidates for this task. Within the applied superconductivity and fusion communities, artificial intelligence (AI), machine learning (ML), and deep learning (DL) techniques have gained increasing traction for design optimization, quench detection, and surrogate modeling [6]. Traditional, purely data-driven artificial neural networks (ANNs) have already demonstrated their capabilities in these fields, successfully predicting pulsed heat loads in the ITER Central Solenoid [7], evaluating real-time AC losses via neural ODEs [8], estimating magnetic levitation forces in superconducting bulks [9], and performing tokamak anomaly detection [10]. However, purely data-driven models are difficult to justify in large superconducting magnet applications because they heavily rely on large, representative training datasets and inherently lack physical consistency. To overcome these limitations, physics-informed neural networks (PINNs) [11] provide a powerful alternative by embedding governing equations, boundary conditions, and physical constraints directly into the training process (see Figure 2). PINNs have emerged as a robust paradigm and have been successfully applied across a broad set of physics domains [12,13], including advanced thermal–hydraulic system simulations, conjugate heat transfer, and multiphysics applications [14–17]. In the superconductivity field, PINNs have recently been proposed for the transient thermal analysis of superconducting tapes, demonstrating their ability to reproduce heat diffusion dynamics while reducing the need for repeated finite-element simulations [18]. Recent advances in physics-informed and coupled thermal modeling highlight the growing necessity of hybrid machine-learning strategies to capture complex thermal effects [19]. Drawing from these methodological foundations, it is clear why the accurate resolution of coupled systems requires embedding physical constraints, rather than relying on sparse data alone.

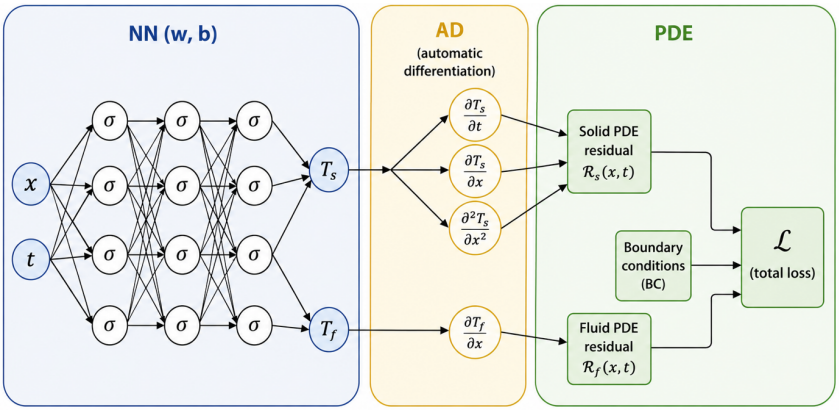

Figure 2. Sketch of a PINN architecture showing the neural network (NN) composed of neurons and layers, featuring weights (w) and bias (b), the differentiation step, and the computation of losses on the basis of which the network is trained.

The thermal–hydraulic response of forced-flow CICC magnets introduces additional effects that are central to superconducting coil operation. The conductor temperature is determined by the coupled evolution of the solid strands and jacket temperature, along with the helium coolant temperature due to strong heat transfer between solid and fluid. For pulsed DTT coils, this coupling is crucial because the temperature margin depends on the local conductor temperature relative to the current sharing temperature, whereas the conductor temperature itself depends on both AC loss deposition and helium cooling. A PINN-based surrogate for this application must therefore include the solid–fluid thermal coupling, rather than solving the solid heat equation alone.

In this work, a physics-informed neural network is developed to predict the transient 1D coupled solid–fluid thermal response of a CICC subjected to AC losses. The model couples a solid energy equation with a 1D helium energy balance, with temperature-dependent material properties and a spatially non-uniform heat source profile. The neural network is trained using a set of current waveforms derived from realistic DTT CS operating scenarios, including variations in waveform duration and amplitude, and is validated using previously unseen current profiles.

The study assesses the ability of the proposed surrogate to reproduce the coupled thermal response and generalize across different current histories. Particular attention is paid to its capability to capture thermally relevant features, including the early temperature peak and the late high-$|I|$ interval. The computational cost of the trained surrogate is also compared with that of the finite-difference reference solver to quantify the potential inference time advantage for repeated scenario evaluations. It is important to emphasize that this work represents a methodological proof-of-concept for a reduced-order PINN. At this stage, the formulation employs a simplified 1D thermal–hydraulic coupled model (quasi-steady helium energy balance with constant mass flow rate). This constant-flow assumption is physically adequate for slow normal-operation scenarios, where severe flow pulsations or compressibility-driven pressure spikes (e.g., during plasma disruptions or Loss-of-Flow Accidents) do not dominate. Extending the network architecture to embed the full unsteady compressible Euler/Navier–Stokes equations for such fast off-normal events remains a subject for future development. Consequently, this framework is not intended to replace high-fidelity tools such as 4C, which account for fluid momentum dynamics, compressibility, and complex AC loss scaling. Rather, it is proposed as an intermediate fast-screening layer for normal-operation transients to filter candidate scenarios before final high-fidelity verification.

The remainder of this paper is organized as follows. Section 2 introduces the physical problem and the thermal–hydraulic model, including the conductor geometry, material properties, heat source formulation, current scenarios, and finite-difference reference model. Section 3 presents the proposed physics-informed neural network formulation and the corresponding training framework. Section 4 assesses the surrogate performance on unseen current profiles and discusses its accuracy, computational cost, and potential application to rapid scenario screening. Finally, Section 5 summarizes the main conclusions of this study.

2. Case Study and Physical Model

2.1. Thermal–Hydraulic Quantity of Interest

One of the key engineering quantities for superconducting magnet operation is the conductor temperature margin, in particular, its minimum value. At a given location and time, the local temperature margin can be written as:

$$\Delta T_{\mathrm{marg}}(x,t) = T_{\mathrm{CS}}(x,t) - T_{\mathrm{cond}}(x,t), \tag{1}$$

where T_{CS} is the current sharing temperature and T_{cond} is the conductor temperature. The minimum temperature margin over the conductor and throughout the pulse is then:

$$\Delta T_{marg}^{min} = \min_{x,t}[T_{CS}(x,t) - T_{cond}(x,t)]. \quad (2)$$

T_{CS} can be computed from the superconducting critical surface scaling law once the local magnetic field, operating current density, and strain state are known. For Nb_3Sn conductors, for example, the critical current is commonly parameterized as a function of magnetic field, temperature, and axial strain [20]. Therefore, for a fixed conductor and coil design, the thermal–hydraulic calculation mainly needs to provide the transient conductor temperature field $T_{cond}(x,t)$. The objective of the reduced-order model developed in this work is thus to predict the transient conductor temperature distribution rapidly and accurately for a family of imposed coil current scenarios.

In this work, the term conductor temperature refers to the equivalent solid temperature representing the CICC composite along the hydraulic path. This is an approximation with respect to the 1D model implemented in high-fidelity tools where the strand and jacket are typically considered as different solid regions. However, during transients, such as the coil normal operation, which are slow with respect to the timescale of heat transfer in the conductor cross-section, the temperature difference between the strand and jacket is negligible; thus they are considered together in a single region here. This approximation would be less justified in case of faster transients, such as quench propagation; however, this is beyond the scope of this work.

2.2. Reference Conductor Geometry and Modeling Assumptions

The model considers a 1D conductor segment of length L, representative of a CICC hydraulic path. The axial coordinate is denoted by $x \in [0, L]$. The conductor is described by two coupled thermal fields: the equivalent solid temperature $T_s(x,t)$, representative of the strands and jacket, and the helium temperature $T_f(x,t)$, representative of the forced-flow coolant. The solid and fluid domains exchange heat through an effective wetted perimeter P_w and an effective heat transfer coefficient h. The solid and fluid cross-sectional areas are denoted by A_s and A_f, respectively.

The following simplifying assumptions are adopted:

1. The thermal problem is reduced to one dimension along the conductor hydraulic path (as it is typically done also in high fidelity thermal–hydraulic tools);
2. The conductor geometry and cooling layout are fixed;
3. The helium mass flow rate is prescribed and constant: this is justified by the relatively slow transient under analysis here, where compressibility effects are not expected to be relevant;
4. The helium pressure dependence on the thermophysical properties is embedded in fitted property laws at the reference pressure, because a small variation in pressure does not cause a large variation in helium properties, whereas the opposite is true for the temperature, thus accounting for the dependence of the helium properties on its temperature (see Section 2.4);
5. The heat deposition is represented by a spatial shape function multiplied by a time-dependent source amplitude derived from the coil current waveform: this accounts for the actual distribution and evolution of AC losses;
6. The magnetic field and strain distributions are not solved by the PINN, because their role in the final margin calculation enters through T_{CS}, which can be evaluated separately from the superconducting scaling law. Furthermore, those distributions are typically computed by electromagnetic and mechanical formulae or numerical tools and are typically known once the thermal–hydraulic verification is required.

The geometric parameters used in the present implementation are summarized in Table 1. They correspond to the representative conductor model used throughout the numerical experiments.

Table 1. Geometric and hydraulic parameters of the representative conductor model.

Quantity	Symbol	Value
Conductor length	L	39 m
Helium flow area	A_f	88 mm^2
Solid cross-sectional area	A_s	605 mm^2
Wetted perimeter	P_w	1.1585 m
Helium inlet temperature	T_{in}	4.5 K
Helium mass flow rate	$\dot{m}$	1.0 g/s

2.3. Thermal–Hydraulic Model

The solid temperature is governed by the following 1D transient heat equation, which includes axial conduction, volumetric heat generation, and heat exchange with the helium coolant:

$$\rho_s c_{p,s}(T_s)\frac{\partial T_s}{\partial t} = \frac{\partial}{\partial x}\left[k_s(T_s)\frac{\partial T_s}{\partial x}\right] + q'''(x,t) - \frac{hP_w}{A_s}\Big(T_s - T_f\Big), \tag{3}$$

where ρ_s is the equivalent solid density, $c_s(T_s)$ is the equivalent solid specific heat, $k_s(T_s)$ is the equivalent solid thermal conductivity, and $q'''(x,t)$ is the volumetric heat source.

The helium is modeled by a 1D quasi-steady energy balance at each time instant:

$$\dot{m}c_{p,f}(T_f)\frac{dT_f}{dx} = hP_w\Big(T_s - T_f\Big), \tag{4}$$

where $\dot{m}$ is the helium mass flow rate and $c_{p,f}(T_f)$ is the temperature-dependent specific heat of helium. As demonstrated in previous validations of superconducting current leads [21], when solid thermal inertia dominates and the pressure drop is small relative to the absolute pressure, neglecting local fluid accumulation in favor of advective enthalpy transport introduces negligible error while greatly simplifying the mathematical problem. The same equation may also be written in terms of helium enthalpy:

$$\dot{m}\frac{dH_f}{dx} = hP_w\Big(T_s - T_f\Big), \qquad H_f(T_f) = \int_{T_{ref}}^{T_f} c_{p,f}(T)\,dT. \tag{5}$$

The enthalpy form is particularly useful because the helium heat capacity varies significantly in the cryogenic temperature range considered here.

The initial condition is a uniform solid temperature equal to the helium inlet temperature:

$$T_s(x,0) = T_{in}. \tag{6}$$

At the helium inlet, the fluid temperature is imposed as

$$T_f(0) = T_{in}. \tag{7}$$

For the solid, adiabatic boundary conditions are applied at both ends:

$$\left.\frac{\partial T_s}{\partial x}\right|_{x=0} = 0, \qquad \left.\frac{\partial T_s}{\partial x}\right|_{x=L} = 0. \tag{8}$$

These boundary conditions correspond to those already used in high-fidelity models.

2.4. Material Property Laws

The solid thermal conductivity and specific heat are fitted in the operating temperature range as

$$k_s(T) = a_k T + b_k, \tag{9}$$

and

$$c_{p,s}(T) = a_c T^2 + b_c T + c_c. \tag{10}$$

The corresponding coefficients are reported in Table 2.

Table 2. Equivalent material property coefficients used in the reduced-order conductor model.

Coefficient	Value
a_k	8.9250
b_k	1.2941
a_c	0.0178
b_c	0.1483
c_c	0.2683

Note that these equivalent solid material properties are weighted on the mass and cross-section, respectively, for the specific heat and thermal conductivity of the superconducting strands and stainless-steel jacket.

For helium, a fitted temperature-dependent heat capacity is used at the reference pressure in the range 4–10 K, as it is expected not to overcome this threshold during normal operation:

$$c_{p,\mathrm{f}}(T) = \sum_{i=1}^{4} A_i \exp\left[-\left(\frac{(T - T_{\mathrm{mean}})/T_{\mathrm{std}} - B_i}{C_i}\right)^2\right]. \tag{11}$$

The corresponding coefficients are reported in Table 3.

Table 3. Fit coefficients for helium specific heat. The temperature input is normalized as $\widetilde{T} = (T - T_{\mathrm{mean}})/T_{\mathrm{std}}$, with $T_{\mathrm{mean}} = 6.75$ K and $T_{\mathrm{std}} = 1.303$ K.

Coefficient	A_i	B_i	C_i
1	7.5568×10^3	0.9605	3.2076
2	2.8201×10^3	−0.1990	0.4473
3	660.4374	−0.3621	0.2377
4	3.8741×10^3	−0.0505	0.8480

An analytical fit is preferred over tabulated data because it provides smooth, differentiable material property functions, avoiding interpolation-induced discontinuities during the evaluation of the PINN residuals. This was needed in order to exploit GPU implementation and avoid bottlenecks caused by lookup tables or external package calling during training.

2.5. Heat Source Model

In pulsed superconducting coils, the dominant transient heat deposition is associated with AC losses induced by current and magnetic field variations [22]. In a detailed 4C analysis, the local loss density depends on the magnetic field history, conductor architecture, coupling time constant, superconducting critical surface, and operating current. In the reduced-order model, the heat source is approximated in separable form:

$$q'''(x,t) = q_{\mathrm{ref}} S(x) g(t), \tag{12}$$

where q_{ref} is a reference volumetric power density, $S(x)$ is a prescribed spatial deposition shape, and $g(t)$ is the normalized temporal driver.

The spatial function $S(x)$ is chosen to mimic the concentration of AC losses in the most loaded conductor region. A decreasing piecewise-constant profile is adopted, with the largest source intensity in the first turn:

$$S(x) = S_j, \qquad x_{j-1} \leq x < x_j, \qquad j = 1, \ldots, N_s. \tag{13}$$

This piecewise distribution is rather accurate in the case of solenoids, because AC losses are proportional to magnetic field variation. The distribution of the latter, in a pancake-wound solenoid, is piecewise (along the hydraulic path), starting from the largest value at the innermost turn. This is also why the end of the first turn, located at x = 3.5 m, is adopted as the reference hotspot location throughout the validation. This position is where the peak solid temperature is expected, together with the lowest current sharing temperature (T_{CS}), which translates at the point where the minimum temperature margin is also expected. The temporal part of the driver is obtained from the current waveform as:

$$g(t) = \frac{|dI/dt|}{\max_t |dI/dt|}. \tag{14}$$

This is justified as the main variation of the magnetic field, which in turn generates AC losses, being primarily driven by the self field, which is governed by the coil current. The magnitude of the AC losses is therefore approximated as proportional to the time derivative of the coil current under investigation.

2.6. Current Scenarios and Synthetic Waveforms

The current scenarios used in this work are based on six realistic current waveforms representative of DTT CS operation, i.e., actual scenarios that have been produced and investigated recently [23]. These profiles define the source family from which the training and validation waveforms are generated. Each current waveform is tabulated as a time-current sequence and is converted into a heat source history through (14).

Using only the original current profiles would mainly assess interpolation and would provide limited information on the network ability to generalize. Therefore, the waveform set is augmented using constrained synthetic profiles. A generic synthetic current profile is written as:

$$I_{\mathrm{syn}}(t) = s \sum_{i=1}^{N_I} w_i I_i(t), \qquad \sum_{i=1}^{N_I} w_i = 1, \qquad w_i \geq 0, \tag{15}$$

where $I_i(t)$ is the original current scenario, w_i is the convex-mixture weight, and s is a scaling factor. In the present implementation, it has been assumed that $s \in [0.90, 1.10]$. The synthetic profiles are clipped to remain within the current envelope spanned by the original scenarios. Furthermore, the convex mixture formulation inherently provides statistical variance reduction against high-frequency sensor noise, as the variance of the combined signal scales with the sum of the squared fractional weights ($\sum w_i^2 < 1$). This mathematical smoothing, coupled with the physical thermal inertia of the solid—which acts as a natural low-pass filter embedded directly into the network's PDE loss—ensures that random current fluctuations do not translate into non-physical thermal spikes in the predicted temperature.

The waveform generation procedure is constrained to avoid nonphysical profiles. Independent pointwise perturbations of the current are discarded because they introduce artificial oscillations and, consequently, artificial spikes in $|dI/dt|$. Instead, the final augmentation strategy mixes short- and long-duration scenarios. This preserves the observed qualitative structure of the current histories: an initial current decay, a late negative current minimum or plateau, and a final recovery through two or three ramps (see Figure 3). The synthetic profiles allow for exploring short, intermediate, and long scenarios in the training set.

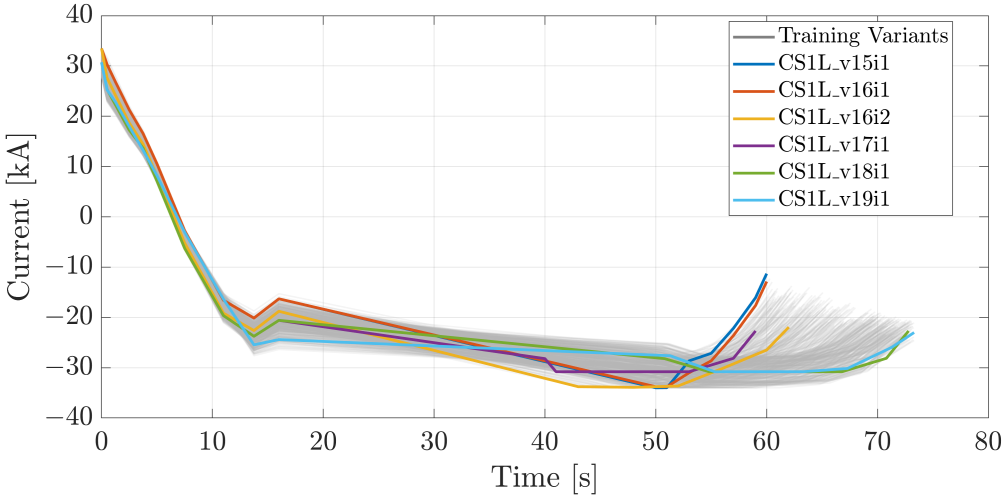

Figure 3. Current waveform family used for training. The grey curves are the training variants, and the solid colored curves are the original current scenarios.

The validation set consists of ten unseen synthetic profiles generated with the same physical constraints, but excluded from the training, see Figure 4. These profiles are evaluated separately to assess the ability of the surrogate to predict conductor temperatures for unseen current histories.

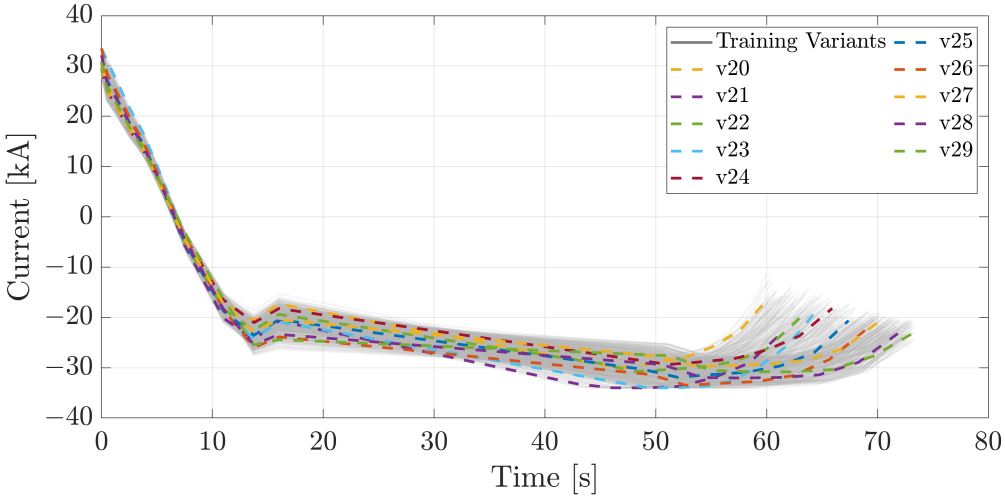

Figure 4. On top of the training variants in grey, the ten held-out unseen profiles for validation are displayed as dashed, colored curves.

2.7. Reference Finite-Difference Model

A finite-difference solver for (3) and (4) is used as a reference model. The finite-difference model solves the same thermal–hydraulic problem used by the PINN, with the same geometry, material properties, heat-source definition, and boundary conditions. It serves two purposes.

First, it provides validation data for assessing the trained PINN on unseen current profiles. Second, sparse finite difference solutions are used as anchor points during training. These anchors are introduced because the coupled transient problem contains localized thermal events that are difficult to capture accurately with residual minimization alone. The anchor data are used only during training. During inference, the PINN is evaluated directly, and no finite difference solve is performed.

It is important to distinguish this reference model from the full 4C model. The 4C code remains the high-fidelity tool for final engineering validation of DTT superconducting magnet performance. The finite-difference model used here is instead an intermediate reference for developing and testing the reduced-order PINN methodology.

3. Physics-Informed Neural Network Formulation

3.1. PINN Architecture

Physics-informed neural networks (PINNs) approximate the solution of a differential problem by means of a neural network whose training loss includes residuals of the governing equations and boundary and/or initial conditions [11]. In a conventional supervised surrogate, the network is trained to reproduce a database of precomputed solutions: this approach is extremely costly, as it requires first the production of such a database; in this case, this would require multiple runs of the 4C code. This procedure had already been explored, e.g., in [24], but it required several weeks just to produce the reference database. In a PINN, instead, the neural approximation is constrained to satisfy the physical model at a set of collocation points sampled in the spatio-temporal domain.

For the present problem, the network receives, as input, the axial coordinate, time, and descriptors of the imposed current waveform, and returns the corresponding solid and helium temperatures. The governing equations introduced in Section 2 are enforced by the automatic differentiation [25] of the network output with respect to space and time. The network architecture consists of a shared trunk with four hidden layers of 112 neurons each, followed by two separate branches for the solid and fluid domains, each comprising three hidden layers of 72 neurons. A hyperbolic-tangent activation function is applied across all hidden layers. The tanh activation has been selected because it is smooth and differentiable, which is convenient for evaluating the first- and second-order derivatives required by the PINN residuals through automatic differentiation. Training is performed with the Adam optimizer, starting from a learning rate of 8×10^{-4} and progressively reducing it during the final part of the optimization down to 2×10^{-5}. To ensure full reproducibility of the PINN training procedure, the complete set of hyperparameters, network architecture details, and sampling strategies are summarized in Table 4. The loss function weights (λ_i) were selected to balance the residual magnitudes during the initial learning phase and were adjusted via a step-decay scheduling. Specifically, the data anchor weight λ_{FD} was set to 120.0 for the first 2200 epochs (warm-up) to quickly guide the network toward the physical temperature range, reduced to 40.0 up to epoch 11,500, and finally relaxed to 22.0 to ensure that the PDE residuals dominated the final optimization phase. A hyperparameter sensitivity study (see Table 5 for a detailed comparison) confirmed that this specific multi-task learning split prevents gradient conflicts between the mathematically distinct differential operators without requiring an over-parameterized shared trunk; consequently, this lightweight topology minimizes inference latency and perfectly bounds error metrics below the 0.1 K engineering target, fulfilling the primary requirements of a fast-screening surrogate.

Table 4. Hyperparameters and network architecture.

Parameter	Value/Description
Architecture (Shared Trunk)	4 layers, 112 neurons each
Architecture (Solid/Fluid Branches)	3 layers each, 72 neurons each
Activation function	Tanh
Optimizer	Adam
Learning Rate Schedule	Initial: 8.0×10^{-4} Epoch 3000: 3.0×10^{-4} Epoch 8000: 1.0×10^{-4} Epoch 11,000: 5.0×10^{-5} Epoch 13,000: 2.0×10^{-5}
Total Epochs	14,000
Collocation Points	26,880 (Space–time int.) per epoch
Loss Weights (λ)	PDE (Solid/Fluid): 5.0 Boundary Conditions: 20.0 Step Continuity: 12.0 Fluid Energy: 2.0 Data Anchor: Scheduled ($120.0 \rightarrow 40.0 \rightarrow 22.0$)

Table 5. Hyperparameter sensitivity analysis.

Configuration	Shared Trunk	Branches	RMSE at x =3.5 m [K]
0 (Baseline)	4×112	3×72	0.079
1 (Deep Trunk)	6×112	3×72	0.092
2 (Deep Branches)	4×112	5×72	0.066
3 (Wide Trunk)	4×160	3×72	0.076
4 (Wide Branches)	4×112	3×112	0.078
5 (Trunk Compression)	4×56	3×72	0.075
6 (Asymmetric)	4×112	4×96 (S)/2×48 (F)	0.088
7 (Minimalist)	2×64	2×48	0.080
8 (Over-Parameterized)	5×128	4×96	0.071

$(L \times N)$ denotes L hidden layers with N neurons each. (S) and (F) denote solid and fluid branches, respectively, when asymmetric.

3.2. Nondimensional Variables

Nondimensionalization is critical in PINNs to ensure numerical stability and well-conditioned gradient flow during training. Scaling all input features and output fields to orders of magnitude close to unity, i.e., $\mathcal{O}(1)$, prevents exploding gradients and avoids optimization bias caused by significant scale discrepancies among physical quantities.

The spatial coordinate and time are nondimensionalized as:

$$\xi = \frac{x}{L}, \qquad \tau = \frac{t}{t_{\text{ref}}}, \tag{16}$$

where L is the conductor length and t_{ref} is the reference time corresponding to the maximum duration of the current scenarios considered in the training and validation set. The solid and fluid temperatures are represented through nondimensional variables,

$$\theta_s = \frac{T_s - T_{\text{in}}}{\Delta T}, \qquad \theta_f = \frac{T_f - T_{\text{in}}}{\Delta T}, \tag{17}$$

where T_{in} is the helium inlet temperature and ΔT is a reference temperature scale chosen to cover the expected conductor temperature rise, typically equal to 1 K.

The neural network approximates the mapping:

$$\mathcal{N}_{\boldsymbol{\theta}} : (\xi, \tau, \boldsymbol{\eta}) \mapsto \Big(\theta_s, \theta_f\Big), \tag{18}$$

where ing$\boldsymbol{\eta}$ denotes the set of waveform-conditioning features described in Section 3.3. The dimensional temperatures are recovered as:

$$T_s = T_{\text{in}} + \Delta T\, \theta_s, \qquad T_f = T_{\text{in}} + \Delta T\, \theta_f. \tag{19}$$

The initial thermal equilibrium condition is imposed by construction through a time-dependent factor that multiplies the raw neural output. In particular, the solid temperature is forced to satisfy:

$$T_s(x, 0) = T_{\text{in}}, \tag{20}$$

whereas the remaining boundary and residual constraints are enforced through the loss terms described below. This hard initial condition treatment reduces the burden on the optimizer and prevents the network from spending training capacity on satisfying the trivial initial state.

3.3. Waveform Conditioning and Current Family Representation

The heat source is driven by the current derivative; therefore, the network must be conditioned on the imposed current history. The waveform set contains the original real scenarios (denoted as $v17$–$v19$), synthetic training variants (see Figure 3), and the held-out unseen validation profiles (denoted as $v20$–$v29$) (see Figure 4). During training, only the original and synthetic training profiles are used (the resulting training set contains 450 synthetic profiles + 6 real scenarios); the validation profiles are included in the waveform table only to allow for post-training interrogation.

For a given waveform coordinate, the temporal source feature is:

$$g(\tau) = \frac{|dI/dt|}{g_{\text{ref}}}, \tag{21}$$

where g_{ref} is the maximum value used for normalization over the waveform set.

Event indicators are introduced as Gaussian functions of time centered around physically important instants:

$$E_k(\tau) = \exp\left[-\left(\frac{\tau - \tau_k}{\sigma_k}\right)^2\right], \tag{22}$$

where τ_k is the event time and σ_k controls the event width. Two event classes are used. The first targets the early solid temperature peak at the end of the first spatial power step (end of the first turn). The second targets the late-window maximum-$|I|$ event, searched between 40 and 73 s. This second event is distinct from the global maximum-$|I|$ event because several current profiles have a large positive initial current at t = 0 s, whereas the thermally relevant late event is associated with the negative current plateau or minimum.

The final input vector may therefore be written schematically as:

$$\mathbf{z} = \Big[\xi,\ \tau,\ g,\ E_{\text{early}},\ E_{\text{late}},\ S(\xi)\Big], \tag{23}$$

where $S(\xi)$ is the spatial heat source shape function. It is crucial to clarify that the event indicators $E_{early}(\tau)$ and $E_{late}(\tau)$ do not require prior knowledge of the thermal solution, making them fully compatible with forward-predictive or real-time screening applications. These indicators are extracted exclusively from the known input current waveform $I(t)$. Specifically, the early peak event is defined by a fixed time window covering the first initial

current decay (e.g., around 11 s for the tested waveforms). The late event is determined solely by finding the time of maximum absolute current $|I(t)|$ within the late-phase operational window (40–73 s). As the input current scenario is entirely known *a priori*, these temporal markers can be automatically generated before the PINN inference begins.

3.4. PINN Losses Definition

All loss components are written in terms of the nondimensionalized governing equations. The solid energy balance is normalized by the reference volumetric power density q_{ref}, yielding the dimensionless solid residual $\tilde{R}_{\mathrm{s}}$:

$$\tilde{R}_{\mathrm{s}} = \left[\frac{\rho_{\mathrm{s}} c_{\mathrm{p,s}}(\theta_{\mathrm{s}})\Delta T}{q_{\mathrm{ref}} t_{\mathrm{ref}}}\right]\frac{\partial \theta_{\mathrm{s}}}{\partial \tau} - \left[\frac{1}{q_{\mathrm{ref}} L^2}\right]\frac{\partial}{\partial \xi}\left[k_{\mathrm{s}}(\theta_{\mathrm{s}})\Delta T\frac{\partial \theta_{\mathrm{s}}}{\partial \xi}\right] - S(\xi)g(\tau) + \left[\frac{hP_{\mathrm{w}}\Delta T}{A_{\mathrm{s}} q_{\mathrm{ref}}}\right](\theta_{\mathrm{s}} - \theta_{\mathrm{f}}). \tag{24}$$

The bracketed terms represent the characteristic dimensionless groups governing the solid thermal dynamics. The latter physically scale the problem and prevent any single residual term from numerically dominating the optimization. The primary groups governing the solid–fluid dynamics evaluated under reference conditions (T_{ref} = 5.25 K) are:

- Fourier Number (Fo): ratio of the rate of heat conduction to the rate of thermal energy storage, $Fo = \frac{k_s t_{ref}}{\rho_s c_{p,s} L^2} \sim \mathcal{O}(10^{-4})$, indicating that axial conduction is slow relative to the transient duration. At T_{ref} = 5.25 K, $Fo = 1.68 \times 10^{-4}$;
- Dimensionless Heat Source (Q): $Q = \frac{q_{ref} t_{ref}}{\rho_s c_{p,s} \Delta T} \sim \mathcal{O}(10^1)$, scaling the volumetric heat deposition against the thermal capacity. At T_{ref} = 5.25 K, Q = 80.6;
- Solid–Fluid Coupling Number (C_{sf}): $C_{sf} = \frac{hP_w t_{ref}}{A_s \rho_s c_{p,s}} \sim \mathcal{O}(10^1)$, representing the strength of the convective heat transfer compared to solid thermal inertia. At T_{ref} = 5.25 K, C_{sf} = 51.0;
- Number of Transfer Units (NTU): $NTU = \frac{hP_w L}{\dot{m} c_{p,f}} \sim \mathcal{O}(10^1)$, dictating the spatial fluid temperature evolution. At T_{ref} = 5.25 K, NTU = 40.0.

By embedding these groups into the residuals, all terms in the loss function are scaled to an order of magnitude of $\mathcal{O}(1)$, ensuring well-conditioned gradients. All spatial and temporal derivatives in (24) are evaluated by automatic differentiation. The corresponding solid residual loss is:

$$\mathcal{L}_s = \frac{1}{N_s}\sum_{j=1}^{N_s}\left|\tilde{R}_s(\xi_j, \tau_j)\right|^2. \tag{25}$$

Each collocation point j in space and time is sampled together with one current scenario, from which the corresponding values of $g(\tau)$, $E_k(\tau)$, and $S(\xi)$ are computed and passed to the network.

For the fluid, the residual is enforced in an energy-consistent form. Normalizing the 1D fluid energy balance by the characteristic heat transfer rate per unit length, $hP_{\mathrm{w}}\Delta T$, the dimensionless form is obtained:

$$\left[\frac{\dot{m} c_{p,f}(\theta_f)}{hP_{\mathrm{w}} L}\right]\frac{d\theta_f}{d\xi} = \left(\theta_s - \theta_f\right). \tag{26}$$

This is equivalent to the enthalpy form:

$$\left[\frac{\dot{m}}{hP_{\mathrm{w}} L\Delta T}\right]\frac{dH_f(\theta_f)}{d\xi} = \left(\theta_s - \theta_f\right). \tag{27}$$

The enthalpy form is preferred here because of the strong temperature-dependence of $c_{p,f}$ in the cryogenic range. The dimensionless fluid residual $\tilde{R}_f$ is therefore written as:

$$\tilde{R}_f = \left[\frac{\dot{m}}{hP_{\mathrm{w}}L\Delta T}\right]\frac{dH_f(\theta_f)}{d\xi} - \left(\theta_s - \theta_f\right), \tag{28}$$

with the corresponding fluid loss:

$$\mathcal{L}_f = \frac{1}{N_f}\sum_{j=1}^{N_f}\left|\tilde{R}_f(\xi_j,\tau_j)\right|^2. \tag{29}$$

Concerning the boundary conditions, their associated losses are easily nondimensionalized as:

$$\mathcal{L}_{\mathrm{BC},s} = \frac{1}{N_{\mathrm{BC},s}}\sum_{j=1}^{N_{\mathrm{BC},s}}\left[\left(\frac{\partial\theta_s}{\partial\xi}\bigg|_{\xi=0,\tau_j}\right)^2 + \left(\frac{\partial\theta_s}{\partial\xi}\bigg|_{\xi=1,\tau_j}\right)^2\right]. \tag{30}$$

The helium inlet condition takes advantage of the chosen thermal shift (i.e., $\theta_f = 0$ at the inlet), leading to a simplified formulation:

$$\mathcal{L}_{\mathrm{BC},f} = \frac{1}{N_{\mathrm{BC},f}}\sum_{j=1}^{N_{\mathrm{BC},f}}\left[\theta_f(0,\tau_j)\right]^2. \tag{31}$$

Note that, although the helium equation is quasi-steady and contains no time derivative, it is solved as a 1D energy balance for each time instant of the transient solid problem. The inlet helium temperature is therefore imposed at $\xi = 0$ for every sampled time τ_j.

The total boundary condition loss is:

$$\mathcal{L}_{\mathrm{BC}} = \mathcal{L}_{\mathrm{BC},s} + \mathcal{L}_{\mathrm{BC},f}. \tag{32}$$

An additional integral energy consistency loss is imposed on the fluid. By integrating the physical fluid energy balance over the conductor length and dividing by the total characteristic transferred power $hP_{\mathrm{w}}L\Delta T$, the dimensionless energy residual is formed:

$$\begin{aligned}\tilde{R}_E(\tau) = &\left[\frac{\dot{m}}{hP_{\mathrm{w}}L\Delta T}\right][H_{\mathrm{f}}(\theta_{\mathrm{f}}(1)) - H_{\mathrm{f}}(0)]\\ &- \int_0^1[\theta_{\mathrm{s}}(\xi) - \theta_{\mathrm{f}}(\xi)]\mathrm{d}\xi,\end{aligned} \tag{33}$$

with the corresponding loss:

$$\mathcal{L}_E = \frac{1}{N_E}\sum_{j=1}^{N_E}\left|\tilde{R}_E(\tau_j)\right|^2. \tag{34}$$

This term helps to constrain the outlet helium enthalpy rise and improves the consistency of the solid–fluid coupling.

In the present coupled conjugate heat transfer problem, the combination of a piecewise-discontinuous spatial heat deposition profile $S(\xi)$ and an impulsive temporal driver $g(\tau)$ introduces stiff, localized spatio-temporal temperature gradients. As demonstrated in the recent literature [26–28], neural networks with smooth activation functions, such as tanh, exhibit a "spectral bias": they possess a strong tendency to learn globally smooth features first, struggling significantly to capture high-frequency variations or sharp local gradients. Consequently, in a purely residual-driven PINN training, the optimizer initially attempts to smear the solution. In regions of localized thermal stiffness, minimizing the PDE residual against a sharp heat source creates a complex optimization landscape full of local

minima. The optimizer can easily become trapped in these sub-optimal valleys, resulting in a systematic underprediction of the sharp temperature peaks, especially for accurately capturing localized transient events [29] (e.g., the early temperature peak at the end of the first spatial power step and late high-$|I|$ interval). Rather than relying permanently on a massive dataset of reference solutions, sparse finite-difference (FD) reference solutions are introduced as anchor data strictly for a "warm-up" phase. The anchor loss is written as:

$$\mathcal{L}_{\mathrm{FD}} = \frac{1}{N_D} \sum_{j=1}^{N_D} w_j \left[\left(\theta_{s,\theta}^{(j)} - \theta_{s,\mathrm{FD}}^{(j)} \right)^2 + \left(\theta_{f,\theta}^{(j)} - \theta_{f,\mathrm{FD}}^{(j)} \right)^2 \right], \tag{35}$$

where $\theta_{s,\theta}$ and $\theta_{f,\theta}$ are the PINN predictions, $\theta_{s,\mathrm{FD}}$ and $\theta_{f,\mathrm{FD}}$ are the scaled finite difference reference temperatures, and w_j are event-dependent weights.

The anchor points are selected only from the training current profiles. The ten held-out validation profiles are excluded from the residual sampling, boundary condition sampling, energy consistency loss, and FD anchor loss. Therefore, the finite difference solver is used during training and post-training validation, but it is not used when the trained PINN is interrogated on an unseen profile.

The anchor weights are increased near:

1. The end of the first spatial power step, i.e., first turn, $x = 3.5$ m;
2. The early temperature peak window;
3. The late window maximum-$|I|$ event.

Preliminary training showed that a low global error could coexist with biased predictions at these events, motivating this event-focused weighting.

The total loss minimized during training is:

$$\mathcal{L} = \lambda_s \mathcal{L}_s + \lambda_f \mathcal{L}_f + \lambda_{\mathrm{BC}} \mathcal{L}_{\mathrm{BC}} + \lambda_E \mathcal{L}_E + \lambda_D \mathcal{L}_{\mathrm{FD}}. \tag{36}$$

The weights λ_i are chosen to balance the different terms and to emphasize the contribution of the FD anchor during the initial part of the training. The anchor weight is then progressively reduced so that the final network remains constrained by the governing equations while preserving the physically relevant transient features learned from the sparse reference points.

The network is trained with a stochastic gradient-based optimizer. At each training step, batches of collocation points are sampled in the (ξ, τ, g) domain. The sampling is not uniform. Instead, it combines:

1. Source weighted samples, concentrated where $|dg/d\tau|$ is large;
2. Early event samples near the first step temperature peak;
3. Late event samples near the maximum $|I|$ between 40 s and 73 s;
4. Uniformly distributed samples for global coverage.

4. Results

Figure 5 reports the evolution of the main normalized loss components during training. The late-stage reduction in the fluid and energy losses indicates improved consistency of the solid–fluid coupling. The total loss decreased from 2.08×10^{1} at the first recorded epoch to 9.55×10^{-1} after 14,000 epochs. The final training run required approximately 4 min on an NVIDIA H100 GPU for the configuration considered here. The GPU log indicated that, during active computation, the utilization was typically about 80%, with a peak value of 84%; the memory footprint remained low, around 1.1 GB. These values indicate that the present implementation is already computationally light compared with the available hardware, leaving margin for larger networks or collocation batches in future extensions.

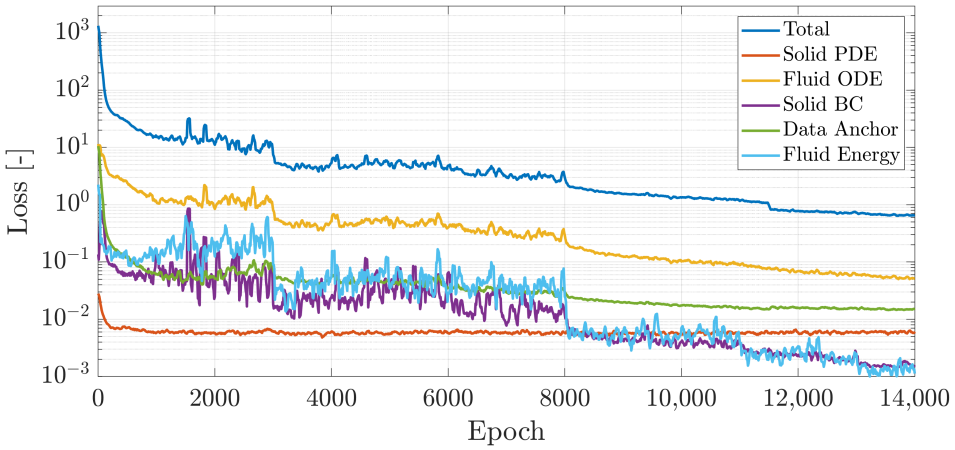

Figure 5. Training history of the loss components.

4.1. Global Accuracy on Unseen Current Profiles

The trained PINN was evaluated on the ten held-out current profiles $v20$–$v29$. The global metric is the relative L^2 error over the evaluated spatio-temporal grid:

$$\varepsilon_{2,s} = \frac{\|T_{s,\theta} - T_{s,\mathrm{FD}}\|_2}{\|T_{s,\mathrm{FD}}\|_2}, \qquad \varepsilon_{2,f} = \frac{\left\|T_{f,\theta} - T_{f,\mathrm{FD}}\right\|_2}{\left\|T_{f,\mathrm{FD}}\right\|_2}. \tag{37}$$

Table 6 summarizes the main error metrics. The relative L^2 error on the solid temperature ranged from 2.21% to 4.95%, with an average value of 3.26%. The corresponding fluid temperature error ranged from 1.62% to 4.52%, with an average value of 2.81%. The largest global errors were obtained for $v21$, $v23$, and $v26$ (mean: 0.079 K, 0.057 K, 0.027 K; maximum: 0.311 K, 0.331 K, and 0.117 K, respectively), which were also among the cases showing the largest late-window local error. The duration of the scenario (t_{end}) and the instant at which maximum $|I|$ was reached are shown in Table 6, confirming that the held-out profiles were scanning short-, intermediate-, and long-duration scenarios, thus covering the span of possible scenarios of interest in this phase of the DTT project.

Table 6. Prediction error and event metrics for the ten held-out unseen profiles. The late event is the maximum value of $|I|$ in the interval 40–73 s.

Profile	t_{end} [s]	t_{late} [s]	MAE [K]	RMSE [K]	Max Error [K]
v20	60.0	47.8	0.038	0.049	0.422
v21	61.5	46.1	0.061	0.083	0.432
v22	63.0	48.1	0.033	0.049	0.405
v23	64.4	49.3	0.058	0.087	0.506
v24	65.9	51.1	0.029	0.041	0.388
v25	67.4	52.0	0.033	0.048	0.416
v26	68.9	52.7	0.053	0.069	0.408
v27	70.3	53.7	0.024	0.037	0.380
v28	71.8	57.7	0.036	0.049	0.386
v29	73.3	58.6	0.029	0.039	0.335
Mean	–	–	0.039	0.055	0.408

In superconducting magnet operation, the ultimate safety quantity is the minimum temperature margin, $\Delta T_{marg} = T_{CS} - T_{cond}$. As T_{CS} is dictated by local magnetic field maps extracted from separate electromagnetic codes, the PINN is specifically tasked with predicting T_{cond}. For the DTT Central Solenoid, nominal temperature margins are strictly

greater than 1.0–1.5 K. Although the global maximum absolute error across the entire space–time domain can locally peak around 0.5 K (primarily near the spatial boundaries, as shown in the residual maps displayed in Figure 6), this does not compromise the component assessment. We define the engineering acceptance criterion based on the critical hotspot regions (e.g., the end of the first spatial power step) where the minimum margin actually occurs. At these critical event locations, the surrogate maintains a maximum absolute temperature error ($\epsilon_{max} = |T_{pred} - T_{FD}|$) of less than 0.15 K. An error within this threshold consumes less than 10–15 % of the available safety margin. Consequently, the PINN accurately preserves the scenario classification (safe vs. unsafe) during the preliminary fast-screening phase. For this reason, the validation also focuses on two specific event classes: the early temperature peak at the end of the first spatial power step and the late high-$|I|$ interval.

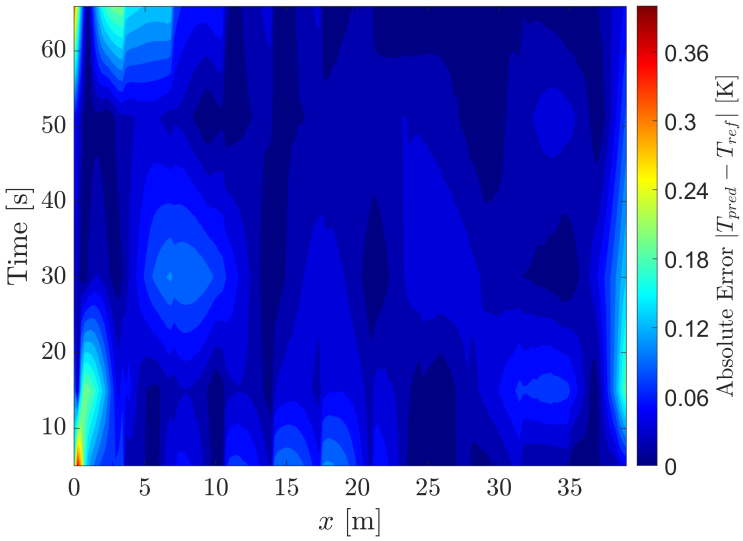

Figure 6. Spatial–temporal map of the absolute temperature error ($|T_{pred} - T_{FD}|$) for the unseen scenario v24. The maximum localized error occurs predominantly near the spatial boundaries, whereas the critical inner regions maintain high accuracy.

4.2. Early and Late High Current Temperature Peaks

Figure 7 compares the finite-difference and PINN time traces of the solid temperature at $x = 3.5$ m, which corresponds to the end of the first turn, where the minimum temperature margin is usually located in a pancake-wound coil. To provide a clearer view of the transient response, the representative unseen scenario v24 has been isolated from the global overview of Figure 7 and is detailed separately in Figure 8. The event-focused training strategy significantly improved the reproduction of the early peak. As summarized in Figure 9, the predicted peak temperature evaluated at the finite-difference peak time differed by less than 0.12 K in all validation cases. The mean signed error at the finite-difference peak time was –0.026 K, indicating a negligible underprediction. The predicted peak time coincided with the reference peak time or slightly earlier; the average peak time error was –0.29 s, with the largest early shift equal to –0.83 s. These values were small relative to the overall pulse duration and indicated that the event-focused sampling and anchor weighting corrected the systematic underprediction observed in earlier training attempts.

The second validation event was the late maximum-$|I|$ time, defined as the time at which $|I|$ is maximum in the interval 40–73 s. This definition avoids selecting the initial condition, where the current may be large, but the conductor has not yet experienced the relevant thermal load. The selected event times range from 46.1 s for $v21$ to 58.6 s for $v29$, reflecting the increasing duration of the unseen profiles.

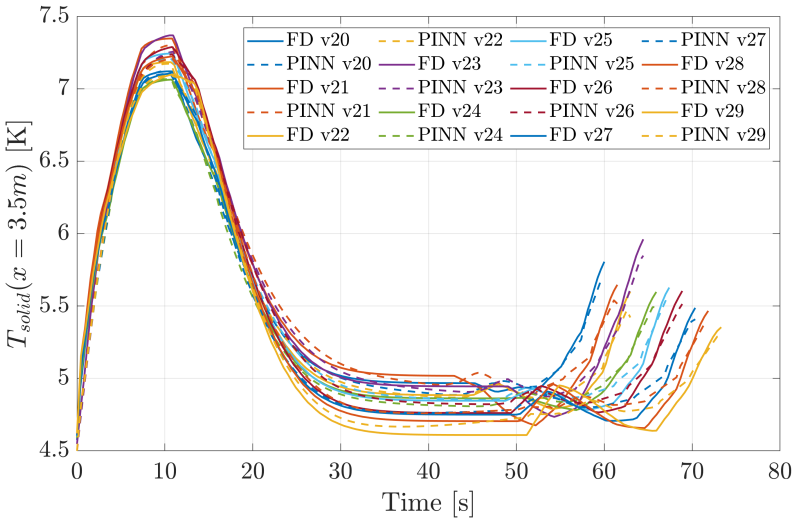

Figure 7. Solid temperature at the end of the first spatial power step, x = 3.5 m, for the ten held-out unseen profiles. Solid and dashed lines report the finite-difference reference solution and the PINN prediction, respectively.

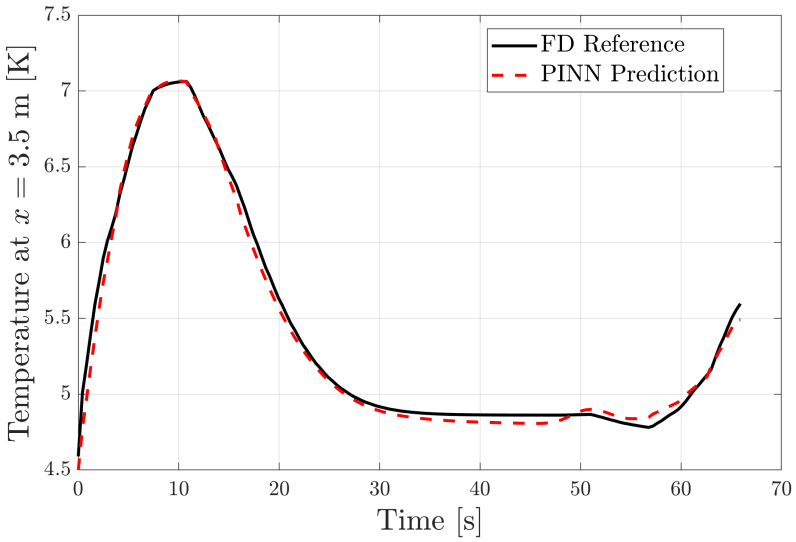

Figure 8. Solid temperature at the end of the first spatial power step, x = 3.5 m, for the v24 unseen scenario.

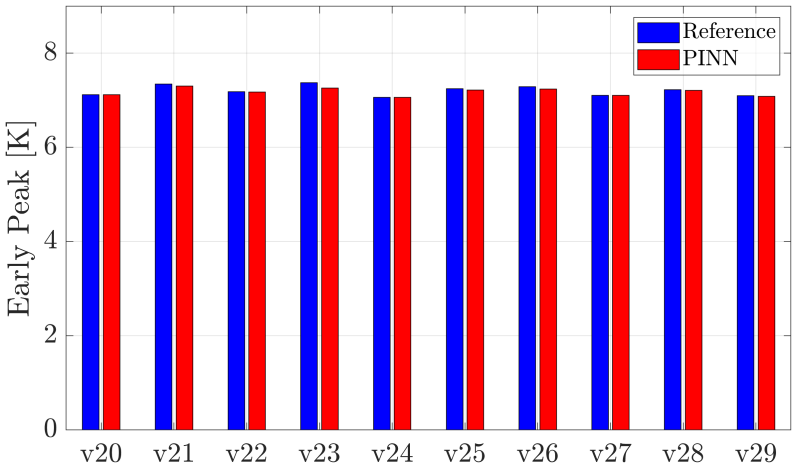

Figure 9. Early peak temperature at x = 3.5 m. The PINN value is evaluated at the time where the finite difference reference reaches its peak.

At the end of the first spatial power step, the late-event temperature error remains limited (see Figure 10). The signed error in $T_s(x = 3.5\,\mathrm{m})$ at the late maximum-$|I|$ event ranges from −0.035 K to 0.106 K, with an average value of 0.027 K. As the typical temperature margin sought for these coils was in the range 1–1.5 K, a maximum error of 0.1 K on the conductor temperature was perfectly acceptable.

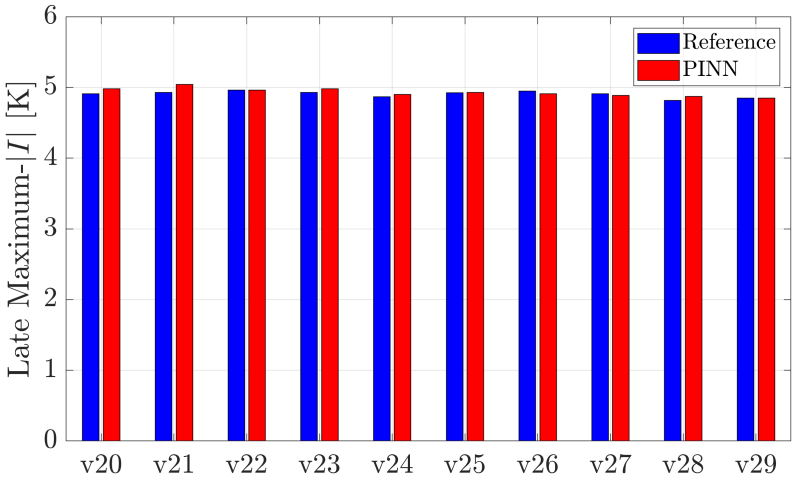

Figure 10. Solid temperature at x = 3.5 m at the late maximum-$|I|$ event.

To verify the physical fidelity of the reduced 1D model, a direct benchmark against the high-fidelity 4C code was performed for scenario v24 (see Figure 11). The 4C code, which solves the fully compressible unsteady flow equations, predicted slightly lower peak temperatures. This discrepancy was physically attributable to the transient expansion of the supercritical helium, which temporarily increased the local mass flow rate and enhanced the convective cooling at the hotspot. Consequently, the PINN slightly overestimate the conductor temperature (MAE = 0.1140 K), providing a conservative prediction of the minimum temperature margin highly desirable for fast-screening. Furthermore, parametric off-design sweeps demonstrated high numerical stability (RMSE < 0.095 K): higher mass flow rates (+25%, +50%) accurately produced enhanced convective cooling, leading to lower peaks and faster thermal recovery, conversely to the delayed cooling and higher peaks observed at reduced flow rates. Meanwhile, shifting the inlet temperature (±0.2 K) resulted in a physically consistent shift in the overall temperature profile, correctly capturing the non-linear thermodynamic behavior of the materials and the fluid in the cryogenic regime as a consequence of the non-linearity of the material properties.

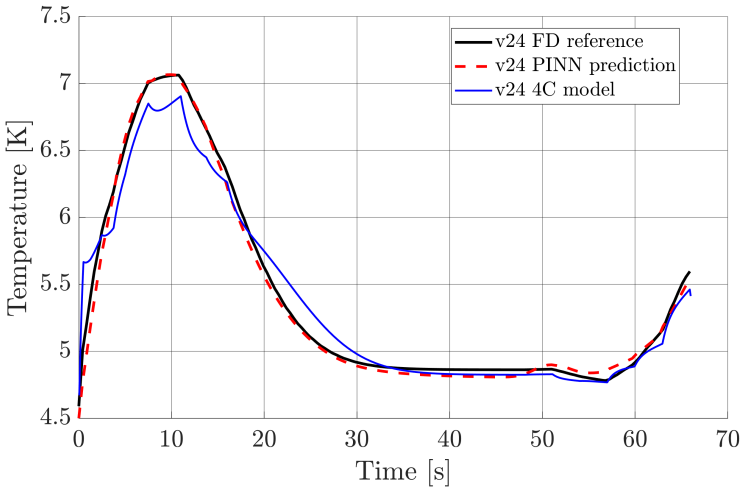

Figure 11. Comparison between FD solver, PINN prediction and 4C model, for the v24 unseen scenario.

4.3. Representative Spatial Transients

To complement the aggregate metrics, Figure 12 reports representative spatial temperature profiles for one unseen case, for the sake of compactness and clarity. The comparison includes several time snapshots: an early time during the initial heat pulse, the vicinity of the first thermal peak, an intermediate cooling/reheating phase, the late maximum-$|I|$ time, and the final available time for the profile.

The spatial profiles show that the PINN captures the main propagation of the thermal wave along the conductor. At early times, the hottest region is close to the first source step. At later times, the maximum temperature moves downstream as heat is transferred to and convected by the helium. The agreement remains close over most of the spatial domain. The largest local discrepancies occur near the downstream high-temperature region (see the last meters of the orange and green curves). This suggests that the current event-focused formulation is adequate for predicting the response of the first turn, where minimum margin is expected to occur, but that further refinement of the downstream portion may be useful if the outlet temperature is important. In this layout, i.e., pancake-wound coil, the T_{CS} at the outlet is substantially larger than that at the inner turns. Therefore, even though the conductor had the largest temperature in some instants, the minimum temperature margin was expected elsewhere. Certainly, the absolute minimum (in time and space), that is, the driving design criterion, would never be in the considered layout. For this reason, event-focused training was concentrated in the first turns and in key instants of the scenario, without needing to refine the outlet region.

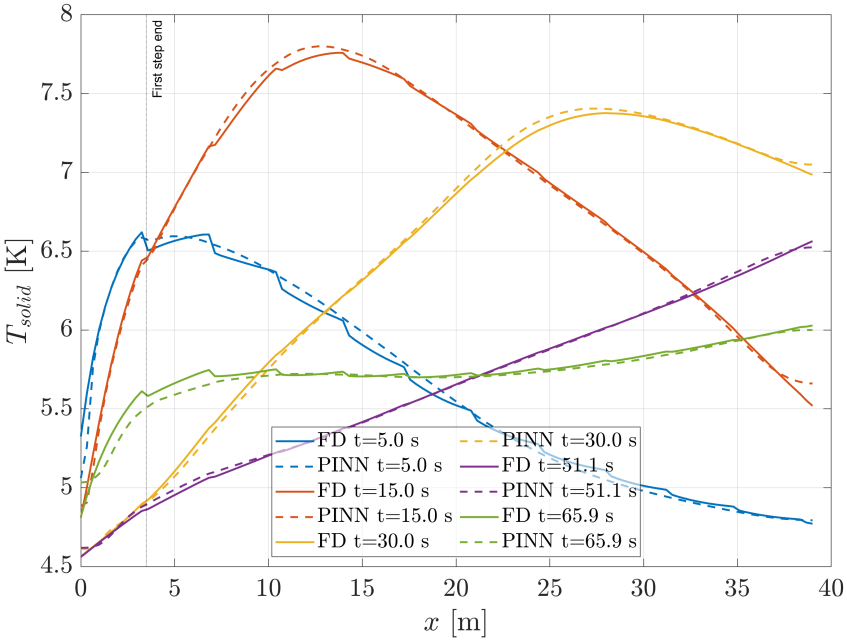

Figure 12. Conductor (solid) temperature distribution for unseen profile $v24$. Solid lines report the finite-difference reference solution and dashed lines the PINN prediction at selected times.

4.4. Inference Time

The computational cost of the trained PINN is compared with the finite-difference reference solver in Figure 13. For the snapshot evaluation used in the validation, the finite-difference solver required approximately 10.2 s per unseen profile, whereas the PINN required approximately 1.8×10^{-2} s. The corresponding average speed-up was approximately 5.7×10^{2}.

These timings should be interpreted as online inference costs. The offline training cost is not included in the comparison and is paid once for a given conductor design, material

property set, and waveform family. This distinction is important for the intended use case: the PINN is not intended to accelerate a single final verification run, but rather, to provide rapid feedback when many candidate current scenarios must be screened during the design loop. The very short interrogation time of the network could make them interesting for implementation directly inside the plasma scenario calculation tools that compute plasma equilibrium finding the required coil currents. The PINN would provide, during the calculation itself, feedback on the feasibility of a given scenario for a given coil, significantly shortening the time taken for the design loop. Another potential use could be in the control algorithm of the actual machine before or during plasma scenario: the PINN would almost immediately provide feedback on the temperature levels in the conductors that a given plasma scenario could produce or is producing, in the case of an online interrogation of the PINN. This would require further development, for example, enabling the PINN to receive real-time feedback from sensors in the magnet system on the mass flow rate or temperatures, but the feasibility of such an approach has been demonstrated here and the room for improvement is considerable, given that the presented PINN is relatively simple and can be trained in a few minutes on suitable hardware.

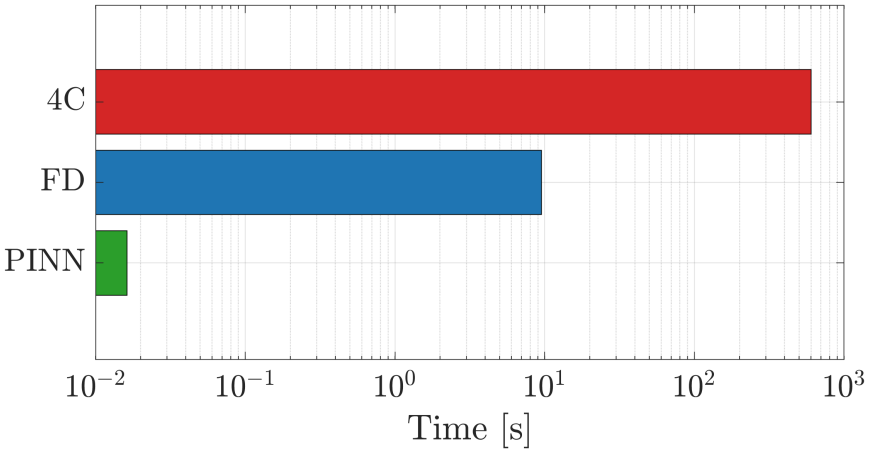

Figure 13. Comparison between finite-difference computational time, PINN inference time, and 4C model single conductor runtime for one held-out unseen profile. The PINN time corresponds to direct neural network evaluation after training.

5. Conclusions and Perspective

In this study, a physics-informed neural network to solve a thermal–hydraulic problem relevant to fast-screening current scenarios in superconducting coils for fusion applications has been developed and applied to the Central Solenoid of DTT. From the point of view of implementation, the main challenges successfully tackled were (1) the coupling of a partial differential equation in space and time with an ordinary differential equation in space and (2) the strong variation of the material properties, especially the specific heat of cryogenic helium. The PINN was trained on several different current scenarios, based on a set of real ones, to be able to accurately predict the conductor temperature evolution in the case of previously unseen current scenarios. The tested PINN had errors lower than 5% and, in terms of temperature, below 0.1 K, which is perfectly sound for a first assessment in this field, where the margins are larger than 1 K.

This approach demonstrates the potential of PINNs for rapid offline screening of candidate plasma scenarios, offering inference times orders of magnitude shorter than those of numerical solvers. The model maintains physical conservation and robustness across off-design mass flow rate and inlet temperature variations, providing a conservative, safe-side temperature estimation compared with high-fidelity 4C benchmarks. In perspective, to transition this proof of concept toward preliminary reactor sizing or real-time diagnostics,

future work will focus on extending the input vector to include dimensionless geometric parameters and coupling dynamic mass flow rates with unsteady compressible fluid conservation laws.

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

Funding: This research received no external funding.

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study. Requests to access the datasets should be directed to the corresponding author.

Acknowledgments: The authors acknowledge CINECA and the EUROfusion consortium for awarding this project access to the high-performance computing resources on the Pitagora cluster hosted at CINECA.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Romanelli, F.; Abate, D.; Acampora, E.; Agguiaro, D.; Agnello, R.; Agostinetti, P.; Agostini, M.; Aimetta, A.; Albanese, R.; Alberti, G.; et al. Divertor Tokamak Test facility project: Status of design and implementation. *Nucl. Fusion* **2024**, *64*, 112015. [CrossRef]
2. Bonifetto, R.; Di Zenobio, A.; Muzzi, L.; Turtù, S.; Zanino, R.; Zappatore, A. Thermal-hydraulic analysis of the DTT CS and PF pulsed coil performance during AC operation. *Fusion Eng. Des.* **2021**, *173*, 112836. [CrossRef]
3. della Corte, A.; Corato, V.; Di Zenobio, A.; Fiamozzi Zignani, C.; Muzzi, L.; Polli, G.M.; Reccia, L.; Turtù, S.; Bruzzone, P.; Salpietro, E.; et al. Successful performances of the EU-AltTF sample, a large size Nb3Sn cable-in-conduit conductor with rectangular geometry. *Supercond. Sci. Technol.* **2010**, *23*, 045028. [CrossRef]
4. Savoldi Richard, L.; Casella, F.; Fiori, B.; Zanino, R. The 4C code for the cryogenic circuit conductor and coil modeling in ITER. *Cryogenics* **2010**, *50*, 167–176. [CrossRef]
5. Albanese, R.; Ambrosino, R.; Mattei, M. CREATE-NL+: A robust control-oriented free boundary dynamic plasma equilibrium solver. *Fusion Eng. Des.* **2015**, *96–97*, 664–667. [CrossRef]
6. Yazdani-Asrami, M.; Song, W.; Morandi, A.; De Carne, G.; Murta-Pina, J.; Pronto, A.; Oliveira, R.; Grilli, F.; Pardo, E.; Parizh, M.; et al. Roadmap on artificial intelligence and big data techniques for superconductivity. *Supercond. Sci. Technol.* **2023**, *36*, 043501. [CrossRef]
7. Savoldi Richard, L.; Bonifetto, R.; Carli, S.; Froio, A.; Foussat, A.; Zanino, R. Artificial Neural Network (ANN) modeling of the pulsed heat load during ITER CS magnet operation. *Cryogenics* **2014**, *63*, 231–240. [CrossRef]
8. Basei, R.; Pase, F.; Lucchini, F.; Torchio, R.; Toso, F. A Structured Neural ODE Approach for Real-Time Evaluation of AC Losses in 3-D Superconducting Tapes. *IEEE Trans. Appl. Supercond.* **2026**, *36*, 5901014. [CrossRef]
9. Alipour Bonab, S.; Xing, Y.; Russo, G.; Fabbri, M.; Morandi, A.; Yazdani-Asrami, M. Estimation of magnetic levitation and lateral forces in MgB_2 superconducting bulks with various dimensional sizes using cascade-forward neural network. *Supercond. Sci. Technol.* **2024**, *37*, 075008. [CrossRef]
10. Paret, J.; Iglesias, D.; Sabio Ruiz, D.; Meloni, D.; Antonione, A.; Bertazzoni, R.; Carannante, G.; Ferrari, M.; Ortiz de Zúñiga, M.; Cavinato, M.; et al. Machine Learning-Based Anomaly Detection for ITER's Tokamak Systems Monitor: A Gyrotron Case Study. *IEEE Trans. Plasma Sci.* **2026**, *54*, 2394–2400. [CrossRef]
11. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
12. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-Informed Machine Learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [CrossRef]
13. Cuomo, S.; Di Cola, V.S.; Giampaolo, F.; Rozza, G.; Raissi, M.; Piccialli, F. Scientific Machine Learning Through Physics–Informed Neural Networks: Where We Are and What's Next. *J. Sci. Comput.* **2022**, *92*, 88. [CrossRef]
14. Zhang, X.; Tu, C.; Yan, Y. Physics-informed neural network simulation of conjugate heat transfer in manifold microchannel heat sinks for high-power IGBT cooling. *Int. Commun. Heat Mass Transf.* **2024**, *159*, 108036. [CrossRef]

15. Wang, T.; Wang, Z.; Huang, Z.; Xi, G. Multi-domain physics-informed neural network for solving heat conduction and conjugate natural convection with discontinuity of temperature gradient on interface. *Sci. China Technol. Sci.* **2022**, *65*, 2442–2461. [CrossRef]
16. Cai, S.; Mao, Z.; Wang, Z.; Yin, M.; Karniadakis, G.E. Physics-Informed Neural Networks (PINNs) for Heat Transfer Problems. *J. Heat Transf.* **2021**, *143*, 060801. [CrossRef]
17. Zhao, Z.; Wang, Y.; Zhang, W.; Ba, Z.; Sun, L. Physics-informed neural networks in heat transfer-dominated multiphysics systems: A comprehensive review. *Eng. Appl. Artif. Intell.* **2025**, *157*, 111098. [CrossRef]
18. Alipour Bonab, S.; Song, W.; Yazdani-Asrami, M. Physics-informed neural network model for transient thermal analysis of superconductors. *Supercond. Sci. Technol.* **2025**, *38*, 08LT01. [CrossRef]
19. Du, Z.; Lu, R. Physics-Informed Neural Networks for Advanced Thermal Management in Electronics and Battery Systems: A Review of Recent Developments and Future Prospects. *Batteries* **2025**, *11*, 204. [CrossRef]
20. Bottura, L.; Bordini, B. $J_C(B, T, \varepsilon)$ Parameterization for the ITER Nb_3Sn Production. *IEEE Trans. Appl. Supercond.* **2009**, *19*, 1521–1524. [CrossRef]
21. Zappatore, A.; Heller, R.; Savoldi, L.; Zanino, R. Assessment of the performance of a 20 kA REBCO current lead. *Cryogenics* **2018**, *95*, 95–101. [CrossRef]
22. Wilson, M.N. *Superconducting Magnets*; Monographs on Cryogenics; Clarendon Press: Oxford, UK, 1987.
23. Castaldo, A.; Acampora, E.; Ambrosino, R.; Albanese, R.; Aucone, L.; Bonanomi, N.; Casiraghi, I.; Marzo, V.D.; Mantica, P. Electromagnetic feasibility studies of plasma scenarios of DTT tokamak. *Fusion Eng. Des.* **2026**, *226*, 115689. [CrossRef]
24. Carli, S.; Bonifetto, R.; Savoldi, L.; Zanino, R. Incorporating Artificial Neural Networks in the dynamic thermal–hydraulic model of a controlled cryogenic circuit. *Cryogenics* **2015**, *70*, 9–20. [CrossRef]
25. Baydin, A.G.; Pearlmutter, B.A.; Radul, A.A.; Siskind, J.M. Automatic Differentiation in Machine Learning: A Survey. *J. Mach. Learn. Res.* **2018**, *18*, 1–43.
26. Rahaman, N.; Baratin, A.; Arpit, D.; Draxler, F.; Lin, M.; Hamprecht, F.; Bengio, Y.; Courville, A. On the spectral bias of neural networks. In *Proceedings of the 36th International Conference on Machine Learning*; PMLR: Cambridge, MA, USA, 2019; Volume 97, pp. 5301–5310. Available online: https://proceedings.mlr.press/v97/rahaman19a.html (accessed on 4 October 2026).
27. Wang, S.; Teng, Y.; Perdikaris, P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM J. Sci. Comput.* **2021**, *43*, A3055–A3081. [CrossRef]
28. Wang, S.; Yu, X.; Perdikaris, P. When and why PINNs fail to train: A neural tangent kernel perspective. *J. Comput. Phys.* **2022**, *449*, 110768. [CrossRef]
29. Harmening, J.H.; Pioch, F.; Fuhrig, L.; Peitzmann, F.J.; Schramm, D.; el Moctar, O. Data-assisted training of a physics-informed neural network to predict the separated Reynolds-averaged turbulent flow field around an airfoil under variable angles of attack. *Neural Comput. Appl.* **2024**, *36*, 15353–15371. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.