energies

MDPI

Article

Layer-Aware Physics-Informed Neural Networks with Condition Embedding for Electro-Thermal Coupled Temperature-Field Modeling of XLPE HVDC Cables

Jia-Xun He [1], Ya Zhang [1], Jun-Jie Ding [1], Kang-Jie Ruan [1], Shuo-Han Jing [1], Hai-Yan Yang [1], Ling-Zhi Zhu [1], Hong-Shuo Zhang [2] and Wei Lu [2,*]

[1] Chongqing Taishan Cable Co., Ltd., Chongqing 401120, China
[2] School of Control Science and Engineering, Dalian University of Technology, Dalian 116023, China
* Correspondence: luwei@dlut.edu.cn

Abstract

The conductor temperature of cross-linked polyethylene (XLPE) high-voltage direct-current (HVDC) cables governs ampacity assessment and insulation life management, yet it cannot be measured in service, and finite-element simulation is too expensive for real-time use. This paper presents a physics-informed neural network (PINN) that embeds the transient heat-conduction equation, a temperature-dependent Joule source, and the boundary and initial conditions into the training loss of a neural surrogate. Three ingredients adapt the framework to power cables: a layer-aware material mapping over the eight heterogeneous cable layers; an electro-thermal coupling through the temperature dependence of the conductor conductivity, handled during training by a convergent Picard-type evaluation of the Joule source; and a condition-embedding input treating the load current and ambient temperature as continuous parameters so that a single network covers the admissible current–ambient envelope of the studied cable configuration. Validated against finite-element references under fifteen operating conditions, the model attains a root-mean-square error of 0.0024 K (mean over five training seeds) on a held-out condition relative to a finite-element reference whose mesh-discretization error a refinement study bounds at about 0.04 K while reducing the governing-equation residual by approximately 28-fold relative to an identically sized data-driven network at statistically indistinguishable pointwise accuracy. The physics prior also renders degradation under training-data reduction more graceful and improves extrapolation to unseen ambient temperatures, whereas current extrapolation remains the most challenging transfer. The differentiable surrogate identifies the load current and the unmeasurable conductor hotspot from ten surface sensors within seconds, at below 9 ms per 10^5 queries. A loss-weight sensitivity study and a three-dimensional cable-end-effect case on a second material configuration are also reported. All reference data are numerical; experimental cable-loop validation remains for future work.

Keywords: physics-informed neural network; layer-aware modeling; condition embedding; HVDC cable; XLPE insulation; electro-thermal coupling; temperature field; surrogate model; inverse problem; digital twin

Check for updates

Academic Editor: Pavlos S. Georgilakis

Received: 13 June 2026
Revised: 8 August 2026
Accepted: 10 August 2026
Published: 12 August 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.

1. Introduction

High-voltage direct-current (HVDC) cables insulated with cross-linked polyethylene (XLPE) are widely deployed in long-distance bulk power transmission and offshore wind

integration. Their transmissible power is bounded by a thermal constraint: the conductor temperature must remain below the 90 °C limit of XLPE to avoid accelerated insulation aging [1,2]. Because the conductor is buried beneath multiple insulation and sheath layers, its temperature cannot be measured directly in service; it must be inferred from models driven by the load current, the ambient condition, and—at best—temperature readings on the cable surface.

Two families of models dominate current practice. Analytical thermal-circuit methods standardized in IEC 60287 [1] are fast but restricted to steady state and simplified geometries. Finite-element method (FEM) simulation, typically performed in COMSOL Multiphysics with version 6.2.

Multiphysics, resolves the full transient, multi-layer, electro-thermally coupled problem with high fidelity [3,4], but a single transient run over a realistic load cycle requires minutes to hours. This cost hinders the direct use of FEM in real-time tasks of growing importance in modern grids, such as dynamic cable rating, online hotspot tracking, and digital-twin state estimation [5]. Data-driven surrogates trained on FEM results offer millisecond inference and have been applied to cable ampacity and temperature prediction [6]. Beyond feed-forward regressors, Gaussian-process regression and physics-augmented recurrent networks have likewise been employed for cable thermal prediction [7,8], and, more broadly, simulation-trained neural surrogates now provide fast field prediction across transport and continuum-mechanics problems, from hemodynamic wall-shear-stress estimation [9] to label-free temperature-field prediction with physics-informed convolutional networks [10]. This work belongs to that class of simulation-trained field surrogates; its distinguishing feature is not the replacement of FEM by a network per se but the combination of a parametric condition embedding, layer-aware physics residuals, and a differentiable formulation usable for inverse problems within one surrogate. However, purely data-driven networks carry no representation of the underlying physics. They interpolate well inside the training distribution but tend to degrade under sparse data or unseen operating conditions, and their predictions may deviate from energy conservation—behavior that is problematic for safety-related thermal assessment.

Physics-informed neural networks (PINNs) [11] combine the two paradigms by penalizing the residual of the governing partial differential equation (PDE), as evaluated by automatic differentiation, alongside a conventional data-fitting loss. PINNs have been applied to heat conduction [12,13], power system dynamics [14,15], thermal modeling of power equipment [16–18], and inverse thermal problems [19]. Treating problem parameters as additional network inputs, as done here for the load current and ambient temperature, is, itself, an established idea: it underlies operator-learning frameworks such as DeepONet [20] and the Fourier neural operator [21] and has been used for parameterized PDEs within the PINN literature. The present condition embedding is therefore deliberately lightweight; the claimed contribution is its combination with the layer-aware residual and the temperature-dependent Joule source in a single differentiable surrogate whose condition inputs also serve gradient-based inverse identification. Nevertheless, three obstacles complicate their direct application to HVDC cables. First, the cable cross-section comprises eight material layers whose thermal conductivities span three orders of magnitude (0.30–400 $\mathrm{W\,m^{-1}\,K^{-1}}$), producing steep interface gradients that can destabilize vanilla PINNs. Second, the heat source is not prescribed: Joule heating depends on the conductor conductivity, which, itself, depends on temperature, creating a two-way electro-thermal coupling. Third, engineering use favors a single surrogate valid over the continuous range of admissible load currents and ambient temperatures rather than a network retrained per condition.

This paper addresses these three aspects and makes the following contributions:

1. A layer-aware PINN formulation in which the eight-layer cable geometry is encoded as a radial look-up map supplying the local density, specific heat, and thermal conductivity to the PDE residual at arbitrary collocation points, avoiding both domain-decomposition overhead and trainable material parameters;
2. An electro-thermal coupling with a convergent Picard treatment: The Joule source embeds the temperature coefficient of the copper conductivity and is evaluated during training at the lagged temperature iterate in a Picard (operator-splitting) manner; the training fixed point satisfies the fully coupled equation, and the contraction constant of the underlying iteration is far below unity for the studied cable;
3. A condition-embedding input layer that augments the spatio-temporal coordinates with the load current and ambient temperature as continuous problem parameters, enabling one network to interpolate and extrapolate across the operating envelope and to address inverse problems by gradient descent over the condition variables;
4. A systematic four-axis evaluation—accuracy, physics consistency, data efficiency, and generalization—against a purely data-driven and a purely physics-driven ablation of an identical architecture, together with inverse-problem, robustness, and runtime studies; a one-decade sensitivity study of the four loss weights; and a three-dimensional cable-end-effect case on a second material configuration over an extended operating envelope, which clarifies where the physics prior is beneficial.

The remainder of the paper is organized as follows. Section 2 formulates the parametric coupled problem. Section 3 details the proposed PINN architecture, the composite loss, and the training algorithm. Section 4 presents the numerical studies—the reference dataset, ablation baselines, evaluation protocol, and results—and Section 5 draws conclusions and discusses limitations.

2. Problem Formulation

2.1. Cable Geometry and Materials

A typical XLPE HVDC cable is composed, from the center outward, of the following structural elements (Figure 1): the current-carrying conductor; an inner semi-conductive layer (conductor screen) that smooths the electric field at the conductor surface; the XLPE insulation withstanding the operating voltage; an outer semi-conductive layer (insulation screen) grading the field at the insulation boundary; a metallic screen providing the fault-current return path and a radial moisture barrier; an inner sheath and an armor layer for mechanical protection, present depending on the installation environment; and a polymeric outer sheath. Each element occupies an annular region with distinct electrical and thermal properties, so the XLPE HVDC cable constitutes a radially layered heterogeneous medium—the feature that dominates its temperature field and must be preserved by any credible thermal model.

For numerical modeling, this layered construction is idealized as a set of concentric, homogeneous annular layers, each carrying constant material properties. The studied single-core XLPE HVDC land cable maps onto eight such layers, indexed as $\ell \in \{0, \ldots, 7\}$ from the conductor outward; Table 1 summarizes the correspondence, together with the radial ranges. Most structural elements translate one to one: the inner and outer semi-conductive layers become the conductor screen ($\ell = 1$) and the insulation screen ($\ell = 3$), respectively. Two deviations from the generic schematic deserve mention. First, the metallic-screen function is realized jointly by two model layers—a semi-conductive water-blocking buffer ($\ell = 4$) and an extruded aluminum sheath ($\ell = 5$)—with the sheath acting simultaneously as the screen, fault-current return path, and radial moisture barrier. Second, being a land design, the cable carries no separate inner sheath or steel-wire armor—layers typical of submarine installations; mechanical protection is provided by the polymeric

outer sheath ($\ell = 6$) alone, and a thin outer semi-conductive coating ($\ell = 7$) is added on top of it for after-installation sheath-integrity testing. Under this idealization, the material fields are piecewise constant and fully determined by the layer-index map, i.e.,

$$\ell(x,y) = \min\{\, i \mid \sqrt{x^2+y^2} \le r_i \,\}, \qquad \rho(\boldsymbol{x}) = \rho_{\ell(x,y)}, \;\; c(\boldsymbol{x}) = c_{\ell(x,y)}, \;\; k(\boldsymbol{x}) = k_{\ell(x,y)}, \quad (1)$$

where x and y are Cartesian coordinates in the cable cross-section with origin on the cable axis; r_i denotes the outer radius of layer i; $\boldsymbol{x} = (x, y, z)$ is the spatial position, with z representing the axial coordinate along the cable; and ρ, c, and k are the density, specific heat, and thermal conductivity, respectively, with ρ_ℓ, c_ℓ, and k_ℓ denoting their constant values in layer ℓ. Equation (1) can be evaluated analytically at any point, so the local material properties are available in closed form throughout the cross-section.

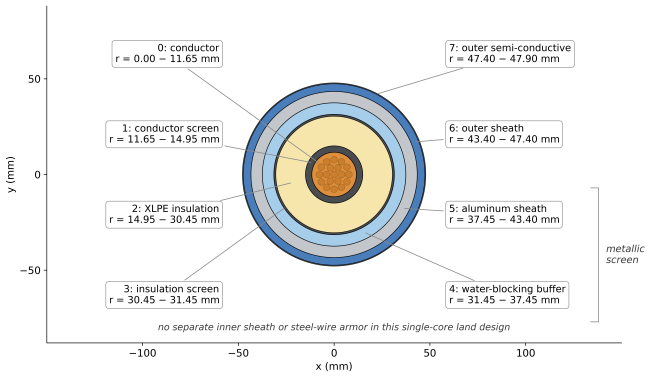


Figure 1. Cross-section of the studied XLPE HVDC cable: the eight model layers with their radial ranges, labeled as in Table 1.

Table 1. Correspondence between the generic cable structure of Figure 1 and the eight-layer model.

Generic Structural Element	Model Layer (ℓ)	Radial Range (mm)	Equivalence in the Studied Cable
Conductor	0	0–11.65	Stranded copper conductor, $\sigma_{20} = 5.8 \times 10^7$ S m^{-1}
Inner semi-conductive layer	1	11.65–14.95	Conductor screen (identical)
XLPE insulation	2	14.95–30.45	Identical
Outer semi-conductive layer	3	30.45–31.45	Insulation screen (identical)
Metallic screen	4, 5	31.45–43.40	Semi-conductive water-blocking buffer (31.45–37.45 mm) plus extruded aluminum sheath (37.45–43.40 mm) acting as screen and radial moisture barrier
Inner sheath/armor	—	—	Not present in the single-core land design; mechanical protection provided by the outer sheath
Outer sheath	6	43.40–47.40	Polymeric outer sheath (identical)
—	7	47.40–47.90	Thin outer semi-conductive coating for sheath-integrity testing; not part of the generic schematic

With the model structure fixed, the remaining ingredients are the dimensions and material constants of the specific XLPE HVDC cable investigated here. The cable has a length of $L = 9.0$ m and an outer radius of $R_{\max} = 47.90$ mm; the conductor radius of $r_0 = 11.65$ mm gives the conductor a cross-sectional area of $A_c = \pi r_0^2 \approx 4.265 \times 10^{-4}$ m^2. Table 2 lists, for each layer, the density (ρ_ℓ), specific heat (c_ℓ), and thermal conductivity (k_ℓ); the thermal conductivities span three orders of magnitude, from 0.30 W m^{-1} K^{-1} in the XLPE insulation to 400 W m^{-1} K^{-1} in the copper conductor.

Table 2. Geometry and material properties of the eight cable layers.

ℓ	Layer	Outer Radius (mm)	ρ_ℓ (kg m^{-3})	c_ℓ (J kg^{-1}K^{-1})	k_ℓ (W m^{-1} K^{-1})
0	Conductor (Cu)	11.65	8960	385	400.0
1	Conductor screen	14.95	1000	2000	0.32
2	XLPE insulation	30.45	930	2200	0.30
3	Insulation screen	31.45	1000	1800	0.40
4	Water-blocking buffer	37.45	1050	1900	0.45
5	Aluminum sheath	43.40	2700	900	205.0
6	Outer sheath	47.40	1000	2000	0.32
7	Outer semi-conductive coating	47.90	1000	1900	0.32

2.2. Governing Equation and Electro-Thermal Coupling

Within the layered medium defined by the index map (1), the temperature field ($T(\boldsymbol{x}, t)$) of the XLPE HVDC cable ($\boldsymbol{x} \in \Omega$), obeys the transient heat-conduction equation, i.e.,

$$\rho(\boldsymbol{x})\, c(\boldsymbol{x})\, \frac{\partial T}{\partial t} = \nabla \cdot \left[k(\boldsymbol{x}) \nabla T\right] + Q(\boldsymbol{x}, T, I), \tag{2}$$

where T is the temperature, t is the time, $\Omega \subset \mathbb{R}^3$ is the cylindrical domain occupied by the XLPE HVDC cable, ∇ is the spatial gradient operator, Q is the volumetric heat-source density, and I is the load current; the material fields take the layer values $(\rho_\ell, c_\ell, k_\ell)$ of Table 2. Since these values are constant within each annulus, Equation (2) reduces inside layer ℓ to the constant-coefficient form of $\rho_\ell c_\ell\, \partial T/\partial t = k_\ell \nabla^2 T + Q$, with ∇^2 representing the Laplace operator, and the eight layers are coupled through the seven internal interfaces, where the temperature and the normal heat flux are continuous: $[T] = 0$ and $[k\, \partial T/\partial r] = 0$, with $[\cdot]$ denoting the jump of the enclosed quantity across an interface and $r = \sqrt{x^2 + y^2}$ representing the radial coordinate. Because k jumps by up to three orders of magnitude across an interface, the flux condition forces a corresponding jump in the radial temperature gradient. The field therefore inherits the layered character of the construction: the metallic layers ($\ell = 0$ and $\ell = 5$) are nearly isothermal, while the polymeric layers—above all the XLPE insulation—sustain almost the entire radial temperature drop. The layer structure equally governs the transient response. The thermal diffusivities ($\alpha_\ell = k_\ell/(\rho_\ell c_\ell)$—not to be confused with the temperature coefficient (α_{Cu}) of the copper conductivity introduced in Equation (3))) range from 1.2×10^{-4} m^2 s^{-1} in the copper conductor to 1.5×10^{-7} m^2 s^{-1} in the XLPE insulation, so the conductor equilibrates internally within seconds, whereas heat storage and release in the polymeric annuli set the time scale on which the XLPE HVDC cable, as a whole, approaches a steady state.

The heat source is tied to the same layer structure: Joule dissipation occurs exclusively in the conductor layer ($\ell = 0$). For DC operation, the current density ($\boldsymbol{J}$) is uniform over the conductor cross-section, i.e., $|\boldsymbol{J}| = I/A_{\mathrm{c}}$, and the volumetric source selected by the layer map is

$$Q(\boldsymbol{x}, T, I) = \begin{cases} \dfrac{I^2}{A_{\mathrm{c}}^2\, \sigma(T)}, & \ell(x, y) = 0, \\ 0, & \text{otherwise}, \end{cases} \qquad \sigma(T) = \frac{\sigma_{20}}{1 + \alpha_{\mathrm{Cu}}\,(T - T_{\mathrm{ref}})}, \tag{3}$$

where $\sigma(T)$ is the electrical conductivity of the copper conductor, $\sigma_{20} = 5.8 \times 10^7$ S m^{-1} is its value at the reference temperature of $T_{\mathrm{ref}} = 293.15$ K, and $\alpha_{\mathrm{Cu}} = 3.93 \times 10^{-3}$ K^{-1} is the temperature coefficient of copper. All heat generated in the innermost layer must traverse the seven outer annuli—dominated by the high-thermal-resistance polymeric layers—before reaching the environment, so the layered construction fixes the thermal path from the source to the environment. The temperature dependence of σ closes an

electro-thermal feedback loop confined to the conductor: a temperature rise lowers the conductivity, raises the losses, and further increases the temperature, rendering Equation (2) nonlinear in layer $\ell = 0$. At the level of the problem formulation, this two-way coupling is contained in the governing equation itself: because the DC current density is uniform over the conductor cross-section, no separate electromagnetic field problem arises, and Equations (2) and (3) constitute a single nonlinear initial-boundary-value problem in the temperature field alone.

2.3. Boundary and Initial Conditions

The lateral surface of the XLPE HVDC cable exchanges heat with the environment by convection (Robin condition), i.e.,

$$-k_7 \left.\frac{\partial T}{\partial n}\right|_{r=R_{\max}} = h\,(T - T_{\mathrm{amb}}), \qquad \frac{\partial T}{\partial n} = \frac{x}{r}\frac{\partial T}{\partial x} + \frac{y}{r}\frac{\partial T}{\partial y}, \tag{4}$$

where $\partial T/\partial n$ is the derivative of the temperature along the outward unit normal ($\boldsymbol{n}$) of the lateral surface; k_7 is the thermal conductivity of the outermost layer ($\ell = 7$); h is the convective heat-transfer coefficient, equal to 5.6 $\mathrm{W\,m^{-2}\,K^{-1}}$ and uniform over the lateral surface in the baseline configuration; and T_{amb} is the ambient temperature.

Two end conditions are considered at the faces expressed as $z = 0$ and $z = L$. In the baseline configuration, the end faces are adiabatic, i.e., $\partial T/\partial z = 0$; because the source (3), the lateral condition (4), and the initial state are all independent of z, the solution is axially uniform, and the three-dimensional problem is formally equivalent to a two-dimensional cross-sectional problem. The formulation equally admits an axially and layerwise heterogeneous cooling, in which the convective coefficient of the lateral surface varies along the cable, while the end faces are adiabatic over the conductor, as well as its screen and convective, with layer-dependent coefficients over the outer annuli—a model of end sections and terminations exposed to a better-ventilated environment. The axial gradients that this produces make the problem genuinely three-dimensional; the coefficient values are operating specifications and are therefore not fixed at the level of the formulation. The XLPE HVDC cable starts from thermal equilibrium, i.e.,

$$T(\boldsymbol{x}, 0) = T_{\mathrm{amb}}, \qquad \forall\, \boldsymbol{x} \in \Omega. \tag{5}$$

2.4. Parametric Problem Statement

The load current (I) and ambient temperature (T_{amb}) are not fixed constants but operating parameters that vary continuously during the actual operation of the XLPE HVDC cable. They are therefore treated as coordinates of a parametric family of initial-boundary-value problems: for every pair (I, T_{amb}) in the admissible operating envelope,

$$\mathcal{E} = \left[0,\, I_{\max}\right] \times \left[T_{\mathrm{amb}}^{\min},\, T_{\mathrm{amb}}^{\max}\right], \tag{6}$$

where $I_{\max}$ is the maximum admissible load current and $T_{\mathrm{amb}}^{\min}$ and $T_{\mathrm{amb}}^{\max}$ bound the admissible ambient temperature. Equations (2)–(5) define a unique temperature field for the XLPE HVDC cable. Thus, the modeling task addressed in this work is the approximation, by a single neural surrogate, of the parametric solution operator, i.e.,

$$(\boldsymbol{x}, t, I, T_{\mathrm{amb}}) \longmapsto T(\boldsymbol{x}, t; I, T_{\mathrm{amb}}), \qquad (\boldsymbol{x}, t) \in \Omega \times [0, t_{\max}],\ \ (I, T_{\mathrm{amb}}) \in \mathcal{E}, \tag{7}$$

where $t_{\max}$ denotes the time horizon of interest rather than the solution of one instance of the problem. For the XLPE HVDC cable considered here, the envelope bounds are $I_{\max} = 200$ A, $T_{\mathrm{amb}}^{\min} = 15\ ^{\circ}\mathrm{C}$, $T_{\mathrm{amb}}^{\max} = 25\ ^{\circ}\mathrm{C}$, and $t_{\max} = 3.6 \times 10^4$ s.

3. Proposed PINN for Temperature-Field Modeling of XLPE HVDC Cables

3.1. Overall Framework

The proposed method approximates the parametric solution operator (7) of the XLPE HVDC cable temperature field as a single neural surrogate trained under two complementary supervision signals. The framework consists of three components arranged in a forward–backward loop. (i) A condition-embedding neural network maps the normalized spatio-temporal coordinates and operating parameters $(\boldsymbol{x}, t, I, T_{\mathrm{amb}})$ to a temperature prediction $(\hat{T}_\theta)$ through a physics-based output parametrization so that one set of weights (θ) represents the temperature field of the XLPE HVDC cable over the entire operating envelope $(\mathcal{E})$. (ii) A layer-aware physics module evaluates how well $(\hat{T}_\theta)$ satisfies the governing equations: it differentiates the network exactly by automatic differentiation, retrieves the local material properties of the eight cable layers through the layer-index map (1), assembles the Joule source (3), and returns the residuals of the heat-conduction equation (2), the convective boundary condition (4), and the initial condition (5). (iii) A composite loss combines these physics residuals, evaluated on randomly resampled collocation points, with a data-fitting term on finite-element reference samples; its gradient drives the update of θ. The data term anchors the prediction to the reference solution under the sampled conditions, while the physics term constrains the network on the whole continuous domain—including conditions and regions for which no reference data exist. Figure 2 depicts the three components and the forward–backward loop that connects them. In what follows, each component of the proposed PINN is detailed in turn.

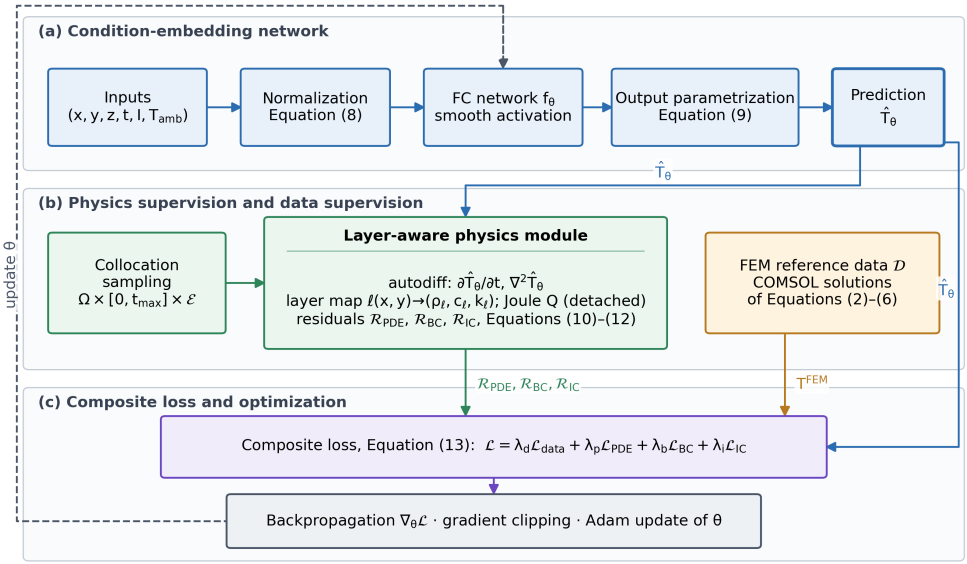


Figure 2. Overall architecture of the proposed PINN framework for XLPE HVDC cable temperature field modeling: (**a**) condition-embedding network with physics-based output parametrization; (**b**) layer-aware physics module evaluating the PDE, boundary, and initial residuals on resampled collocation points, alongside the FEM reference data; (**c**) composite loss and Adam update of the network parameters (θ), with the dashed feedback path closing the forward–backward training loop.

3.2. Network Architecture and Physics-Based Output Parametrization

The first component is a fully connected network (f_θ) that maps the normalized six-dimensional input, i.e.,

$$\boldsymbol{u} = \left(\frac{x}{R_{\max}}, \frac{y}{R_{\max}}, \frac{z}{L}, \frac{t}{t_{\max}}, \frac{I}{I_{\max}}, \frac{T_{\mathrm{amb}} - 273.15\,\mathrm{K}}{\Theta} \right), \tag{8}$$

to a scalar, where $\Theta = T_{\text{amb}}^{\max} - 273.15\,\text{K}$ is the normalization scale of the ambient temperature and the remaining scales are the geometric and operational bounds ($R_{\max}$, L, $t_{\max}$, and $I_{\max}$). The depth and width of f_θ are design hyperparameters; its activation functions, however, must be smooth, because the physics module differentiates the network twice with respect to its inputs. Including I and T_{amb} as continuous inputs (condition embedding) makes the network an approximation of the parametric solution operator (7) rather than of a single solution instance: one set of weights represents the XLPE HVDC cable's temperature field for every condition in $\mathcal{E}$, and the prediction is differentiable with respect to the operating conditions, which additionally enables gradient-based identification of unknown conditions from temperature measurements.

Rather than emitting the absolute temperature, the network predicts a normalized temperature rise that is reassembled as

$$\hat{T}_\theta(\boldsymbol{x}, t, I, T_{\text{amb}}) = T_{\text{amb}} + f_\theta(\boldsymbol{u})\,\Delta T_{\max}\left(\hat{I}^2 + \varepsilon\right), \qquad \hat{I} = \tfrac{I}{I_{\max}}, \tag{9}$$

where $\Delta T_{\max}$ is a preset upper bound of the attainable temperature rise of the XLPE HVDC cable and ε is a small relaxation constant. This parametrization (i) anchors the prediction at T_{amb} for zero current up to the small relaxation (ε), consistent with the unloaded steady state and the initial condition; (ii) injects the quadratic current dependence of Joule heating as an inductive bias; and (iii) shifts the learning target from the absolute temperature, whose numerical range on the kelvin scale is far from the origin and weakly sensitive to the load-induced rise, to a bounded normalized rise, which accelerates convergence. The quadratic factor deserves a qualification: the proportionality of the temperature rise to I^2 is exact only at DC steady state, whereas the transient response also involves the heat capacities ($\rho_\ell c_\ell$) and the diffusion time scales. The factor is therefore an inductive prior rather than a hard constraint—the network f_θ, itself, receives $\hat{I}$ as an input and can represent departures from the quadratic law during the transient—so the parametrization biases but does not restrict the attainable solution family.

3.3. Layer-Aware PDE Residual via Automatic Differentiation

The second component translates the layered physics of the XLPE HVDC cable into differentiable training signals. All derivatives in Equations (2)–(4) are evaluated by automatic differentiation of $\hat{T}_\theta$, with the chain-rule rescaling of $\partial_x = R_{\max}^{-1}\partial_{x_n}$ and $\partial_x^2 = R_{\max}^{-2}\partial_{x_n}^2$ implied by the normalization (8), where x_n denotes the normalized counterpart of the x coordinate. At an arbitrary collocation point, the residual of the governing equation reads

$$\mathcal{R}_{\text{PDE}}[\hat{T}_\theta] = \rho_\ell\, c_\ell\, \frac{\partial \hat{T}_\theta}{\partial t} - k_\ell\left(\frac{\partial^2 \hat{T}_\theta}{\partial x^2} + \frac{\partial^2 \hat{T}_\theta}{\partial y^2} + \frac{\partial^2 \hat{T}_\theta}{\partial z^2}\right) - Q\left(\boldsymbol{x}, \hat{T}_\theta, I\right), \tag{10}$$

where $(\rho_\ell, c_\ell, k_\ell)$ are obtained from the look-up map (1), so the residual automatically carries the correct material properties of whichever of the eight cable layers contains the collocation point. The material parameters are frozen physical constants, not trainable quantities; the network learns only the temperature field, which avoids data–physics conflicts and ensures that the residual is evaluated with respect to the stated governing equation of the XLPE HVDC cable.

Inside the Joule source, the temperature argument is detached from the computation graph: at optimization step k, the source is evaluated as $Q\left(\boldsymbol{x}, \hat{T}_{\theta_k}, I\right)$, with $\hat{T}_{\theta_k}$ treated as a constant, so the electro-thermal nonlinearity is handled by a Picard-type (operator-splitting) iteration interleaved with the optimizer updates rather than by simultaneous differentiation through the source. Three properties delimit the consequences of this choice. First, the detachment alters gradients only, never values: at any parameter vector (θ), the residual (10)

coincides with the residual of the fully coupled equation because the source is evaluated at the network's own prediction; all residual metrics reported in this work therefore refer to the fully coupled Equation (2). Second, the gradient contribution suppressed by the detachment, i.e.,

$$-\frac{2\lambda_p}{N_p}\sum_{\mathcal{P}}\frac{\mathcal{R}_{\mathrm{PDE}}}{(\rho_\ell c_\ell)^2}\,\frac{\partial Q}{\partial \hat{T}_\theta}\,\nabla_\theta \hat{T}_\theta,$$

with λ_p representing the weight of the physics loss and $\mathcal{P}$ denoting the set of N_p interior collocation points, is proportional to the residual itself and vanishes together with it; thus, any parameter vector at which the fully coupled residual vanishes is also a stationary point of the detached physics loss, so the lagged scheme and the fully coupled formulation share their zero-residual fixed points. Third, the iteration is strongly contractive: the source (3) is affine in temperature, i.e., $Q = \big(I^2/(A_c^2\sigma_{20})\big)\big(1+\alpha_{\mathrm{Cu}}(T-T_{\mathrm{ref}})\big)$, and the loop gain of the electro-thermal feedback evaluates to $\alpha_{\mathrm{Cu}}\,\Delta T_{\mathrm{rise}}$, where ΔT_{rise} denotes the steady-state conductor temperature rise, which bounds the transient rise from above. Over the operating envelope of the baseline configuration (Section 2.4; $\Delta T_{\mathrm{rise}} \lesssim 2$ K), the gain is below 10^{-2}, and it remains below 0.3, even at the XLPE thermal limit, so the lagged evaluation converges geometrically to the self-consistent solution of the fully coupled problem. Boundary and initial residuals follow directly from Equations (4) and (5):

$$\mathcal{R}_{\mathrm{BC}} = -k_7\Big(\tfrac{x}{R_{\max}}\partial_x \hat{T}_\theta + \tfrac{y}{R_{\max}}\partial_y \hat{T}_\theta\Big) - h\big(\hat{T}_\theta - T_{\mathrm{amb}}\big), \tag{11}$$

$$\mathcal{R}_{\mathrm{IC}} = \hat{T}_\theta(\boldsymbol{x}, 0, I, T_{\mathrm{amb}}) - T_{\mathrm{amb}}. \tag{12}$$

3.4. Composite Loss and Training Algorithm

The third component assembles the supervision signals into a single objective. Let $\mathcal{D} = \{(\boldsymbol{x}_i, t_i, I_i, T_{\mathrm{amb},i}, T_i^{\mathrm{FEM}})\}_{i=1}^{N_d}$ be the N_d finite-element reference samples of the XLPE HVDC cable's temperature field, where T_i^{FEM} is the reference temperature at the i-th sample, and let $\mathcal{P}$, $\mathcal{B}$, $\mathcal{T}_0$ denote sets of N_p, N_b, and N_i collocation points drawn uniformly at random—at every optimization step—from the cable interior (Ω), its lateral surface, and the $t = 0$ slice, respectively, with condition coordinates (I, T_{amb}) sampled from the continuous envelope ($\mathcal{E}$). The training objective is

$$\mathcal{L}(\theta) = \lambda_d \underbrace{\frac{1}{N_d}\sum_{\mathcal{D}}\big(\hat{T}_\theta - T^{\mathrm{FEM}}\big)^2}_{\mathcal{L}_{\mathrm{data}}} + \lambda_p \underbrace{\frac{1}{N_p}\sum_{\mathcal{P}}\Big(\frac{\mathcal{R}_{\mathrm{PDE}}}{\rho_\ell c_\ell}\Big)^2}_{\mathcal{L}_{\mathrm{PDE}}} + \lambda_b \underbrace{\frac{1}{N_b\,C_b^2}\sum_{\mathcal{B}}\mathcal{R}_{\mathrm{BC}}^2}_{\mathcal{L}_{\mathrm{BC}}} + \lambda_i \underbrace{\frac{1}{N_i}\sum_{\mathcal{T}_0}\mathcal{R}_{\mathrm{IC}}^2}_{\mathcal{L}_{\mathrm{IC}}}, \tag{13}$$

where the four terms quantify the fidelity to the reference data and the violation of the three physical constraints of the XLPE HVDC cable temperature field: $\mathcal{L}_{\mathrm{data}}$ is the mean-squared deviation of the prediction ($\hat{T}_\theta$) from the FEM reference temperatures over $\mathcal{D}$, anchoring the surrogate to the reference solution; $\mathcal{L}_{\mathrm{PDE}}$ is the mean-squared, $\rho_\ell c_\ell$-normalized residual (10) of the heat-conduction Equation (2) over the interior collocation set ($\mathcal{P}$), penalizing departures of the prediction from the governing equation inside the cable; $\mathcal{L}_{\mathrm{BC}}$ is the mean-squared residual (11) of the convective boundary condition (4) over the surface set ($\mathcal{B}$), enforcing the heat exchange with the environment; and $\mathcal{L}_{\mathrm{IC}}$ is the mean-squared residual (12) of the initial condition (5) over the set ($\mathcal{T}_0$), enforcing the thermal-equilibrium start at $t = 0$. The scalar weights (λ_d, λ_p, λ_b, and λ_i) balance the four terms and are treated as tunable hyperparameters, with subscripts d, p, b, and i denoting the data, PDE, boundary, and initial contributions throughout, with the same convention for the batch sizes (N_d, N_p, N_b, and N_i); division by $\rho_\ell c_\ell$ rescales the PDE residual to a temperature rate across layers whose volumetric heat capacities differ between materials; and C_b is a scaling constant that brings the heat-flux units of the boundary term to the same order as the remaining terms. The four terms carry different physical units—K^2 for $\mathcal{L}_{\mathrm{data}}$ and $\mathcal{L}_{\mathrm{IC}}$, $(\mathrm{K\,s^{-1}})^2$ for $\mathcal{L}_{\mathrm{PDE}}$, and

dimensionless for $\mathcal{L}_{\mathrm{BC}}$ after division by C_b^2—so the weights are understood to carry the reciprocal units required to render $\mathcal{L}$ homogeneous; their numerical values reported below refer to this convention. The constant of $C_b = 10^3\ \mathrm{W\,m^{-2}}$ carries the units of a heat flux; its magnitude is set by the characteristic boundary flux ($h\,\Delta T_{\max} \approx 5.6 \times 10^2\ \mathrm{W\,m^{-2}}$), rounded to the nearest decade. The weights ($(\lambda_d, \lambda_p, \lambda_b, \lambda_i) = (1, 0.05, 1, 1)$) were tuned manually, guided by the validation accuracy and the magnitude of the sampled residual rather than by an automated search; systematic alternatives such as neural-tangent-kernel-based balancing and self-adaptive weighting [22,23] were not employed here and remain compatible drop-in refinements.

The two supervision signals play complementary roles in this objective. The data term acts only under the discrete conditions and at locations where reference samples exist, whereas the physics terms act on the entire continuous domain ($\Omega \times [0, t_{\max}] \times \mathcal{E}$) of the XLPE HVDC cable. Moreover, because the collocation sets are resampled at every step, the physics constraints are enforced on a continually refreshed set of points covering, in particular, conditions and regions for which no reference data are available—thereby supplying physical supervision precisely where labeled data are absent.

The objective (13) is minimized by a stochastic gradient-based optimizer (Adam in this work): each step propagates a mini-batch of reference samples and freshly sampled collocation points through the network, evaluates the four loss terms with the layer-aware physics module, and updates θ by backpropagation, with standard stabilization measures such as learning-rate decay and gradient-norm clipping. Algorithm 1 summarizes the complete procedure, which realizes the forward–backward loop of the overall framework.

Algorithm 1 PINN training for XLPE HVDC cable temperature-field modeling

Require: FEM dataset $\mathcal{D}$; operating envelope $\mathcal{E}$; layer map $\ell(x, y)$ and material table $(\rho_\ell, c_\ell, k_\ell)$; loss weights $(\lambda_d, \lambda_p, \lambda_b, \lambda_i)$; number of epochs N_E; batch sizes $|\mathcal{B}_d|$, N_p, N_b, N_i

Ensure: Trained network parameters θ^*

Initialize θ; initialize the optimizer with learning rate η_0
for epoch $= 1$ **to** N_E **do**
  for each mini-batch $\mathcal{B}_d \subset \mathcal{D}$ **do**
    Evaluate $\hat{T}_\theta$ on $\mathcal{B}_d$ via Equations (8) and (9); form $\mathcal{L}_{\mathrm{data}}$
    Sample N_p collocation points in $\Omega \times [0, t_{\max}] \times \mathcal{E}$
    Compute $\partial \hat{T}_\theta / \partial t$ and $\nabla^2 \hat{T}_\theta$ by automatic differentiation
    Look up $(\rho_\ell, c_\ell, k_\ell)$ via Equation (1); evaluate Q via Equation (3) at the lagged iterate $\hat{T}_\theta$ (detached; Picard step)
    Form $\mathcal{R}_{\mathrm{PDE}}$ via Equation (10); form $\mathcal{L}_{\mathrm{PDE}}$
    Sample N_b boundary and N_i initial points; form $\mathcal{R}_{\mathrm{BC}}$, $\mathcal{R}_{\mathrm{IC}}$ via Equations (11) and (12); form $\mathcal{L}_{\mathrm{BC}}$, $\mathcal{L}_{\mathrm{IC}}$
    $\mathcal{L} \leftarrow \lambda_d \mathcal{L}_{\mathrm{data}} + \lambda_p \mathcal{L}_{\mathrm{PDE}} + \lambda_b \mathcal{L}_{\mathrm{BC}} + \lambda_i \mathcal{L}_{\mathrm{IC}}$
    Backpropagate $\nabla_\theta \mathcal{L}$; clip the gradient norm; update θ with the optimizer
  end for
  Adjust the learning rate according to the chosen decay schedule
end for
return $\theta^* = \theta$

4. Numerical Studies

This section reports the numerical studies of the proposed PINN framework for XLPE HVDC cable temperature-field modeling: the reference dataset, the ablation baselines, and the evaluation protocol are described first, followed by the results.

4.1. Reference Dataset and Validation Conditions

To generate reference data and assess the surrogate, fifteen discrete operating conditions were selected from the envelope ($\mathcal{E}$) by combining five load currents ($I \in \{0, 50, 100, 150, 200\}$ A) with three ambient temperatures ($T_{\mathrm{amb}} \in \{15, 20, 25\}$ °C). These conditions serve exclusively for model training supervision and performance verification; the surrogate itself is defined over the continuous envelope ($\mathcal{E}$).

Reference solutions were computed in COMSOL Multiphysics 6.2 (build 6.2.0.290) for each of the fifteen conditions by solving Equations (2)–(5) on the three-dimensional, eight-layer geometry, with the Joule source supplied by a stationary electric-current interface, equivalent to Equation (3) for the axially uniform DC conductor. The domain is discretized by prismatic elements obtained by sweeping a free triangular source-face mesh that conforms to the layer interfaces; the per-layer size settings grade from extremely coarse in the conductor, the XLPE insulation, and the water-blocking buffer, through extra coarse in the aluminum and outer sheaths, to finer in the two thinnest annuli (insulation screen and outer coating) and extra fine in the conductor screen, with the azimuthal resolution determined automatically by the source-face triangulation. The sweep uses a uniform axial distribution, yielding 30 node stations. The temperature field is discretized with linear Lagrange elements and the electric-current interface with quadratic elements; the exported node set comprises 66 cross-sectional nodes at 30 axial stations, corresponding to 1980 mesh nodes in total. The time-dependent solver uses tolerance-based termination (at most, 10 nonlinear iterations and tolerance factor of 1). Each simulation covers $0 \leq t \leq 600$ min sampled at 1 min intervals (601 instants including $t = 0$), and temperatures were exported at all mesh nodes, yielding 1.78×10^7 space–time samples over the fifteen conditions. For training, the time axis was subsampled by a factor of 10 (61 instants per condition), while all mesh nodes were retained at full spatial resolution; excluding the condition (20 °C, 100 A) held out for testing, yielding 1.69×10^6 training samples over the remaining fourteen conditions. All reported test errors are computed at the original 1 min temporal resolution.

To quantify the discretization uncertainty of this reference, a mesh-refinement study was performed under the held-out condition (20 °C, 100 A): the per-layer element sizes of the cross sectional source-face mesh were reduced in two stages—first, by one predefined size level per layer, then to one quarter of the baseline values via custom size settings—raising the cross-sectional resolution from 66 to 282 and 1442 nodes per section, with the axial sweep, element orders, and solver settings retained; the refined solutions were evaluated at the 1980 baseline export nodes at all 601 instants. The finest reference differs from the baseline by, at most, 0.040 K (RMS 0.012 K); the difference grows monotonically toward steady state and concentrates in the conductor and its screen, where it amounts to a nearly uniform offset of the conductor plateau. Between the two refinement stages, these figures change by only a few millikelvin (0.038 versus 0.040 K in the maximum and 0.011 versus 0.012 K in the RMS), and the refinement drives the mesh volume of the discretized cylinder from 5.84×10^7 through 6.41×10^7 to 6.49×10^7 mm^3, against the exact value of $\pi R_{\mathrm{max}}^2 L = 6.49 \times 10^7$ mm^3; the 10% volume deficit of the baseline mesh corresponds to the coarse azimuthal resolution of its source-face triangulation noted in Section 4.4. The discretization error of the baseline reference is therefore about 0.04 K. This uncertainty exceeds the surrogate-to-reference errors reported below: those errors quantify the fidelity of the surrogate to the discrete reference on which it is trained and evaluated, while its accuracy with respect to the mesh-converged temperature field is bounded by the reference uncertainty of about 0.04 K—well within engineering requirements for cable thermal assessment.

4.2. Ablation Baselines

To isolate the contribution of each supervision signal to XLPE HVDC cable temperature-field modeling, all subsequent experiments compare three models that share the identical architecture, data, and optimizer of Algorithm 1 and differ only in the loss weights of Equation (13): MLP sets $\lambda_p = \lambda_b = \lambda_i = 0$ (data-supervised only), PINN-phys sets $\lambda_d = 0$ (purely physics-driven, no FEM supervision), and the full model is denoted as PINN-full. Restricting the comparison to loss-weight switches removes architecture capacity as a confounder, so any performance difference is attributable to the supervision signals themselves.

One qualification applies to the terminology used below: because the three models share one architecture, the MLP baseline also retains the physics-based output parametrization (9), including its $\hat{I}^2$ factor and its anchoring at T_{amb}. It is therefore free of physics losses rather than free of physical structure, and the label of "purely data-driven" used for brevity in what follows should be read in that restricted sense.

4.3. Evaluation Protocol

Four complementary axes are evaluated for XLPE HVDC cable temperature-field modeling: (i) accuracy on a held-out condition $(20\,^{\circ}\mathrm{C}, 100\,\mathrm{A})$ excluded from training, using MAE, RMSE, maximum error, and the conductor-hotspot error; (ii) physics consistency, the root-mean-square of the $\rho_\ell c_\ell$-normalized residual (10) over 10^4 random collocation points, together with a convergence diagnostic of the electro-thermal Picard iteration recorded during training; (iii) data efficiency, test RMSE when the training set is subsampled to $\{100, 50, 25, 10, 5\}\%$; and (iv) generalization, with current interpolation (train $I \in \{0, 50, 150, 200\}$ A, test 100 A), current extrapolation (train $I \leq 100$ A, test $\{150, 200\}$ A), and ambient-temperature extrapolation (train $\{15, 20\}$ °C, test 25 °C).

Inference latency is measured on batches of 10^5 queries with device synchronization, and inverse-problem accuracy is assessed by recovering the load current and the conductor hotspot from sparse virtual sensors. The implementation instantiates the general framework as follows: five hidden layers of 128 neurons with tanh activation; output-parametrization constants of $\Delta T_{\max} = 100$ K and $\varepsilon = 10^{-4}$; loss weights of $(\lambda_d, \lambda_p, \lambda_b, \lambda_i) = (1, 0.05, 1, 1)$ with $C_b = 10^3\ \mathrm{W\,m^{-2}}$; and Adam with an initial learning rate of 10^{-3}, a decay factor of 0.7 every 500 epochs, gradient-norm clipping at 1, data mini-batches of 4096 samples, and $N_p = 1024$, $N_b = N_i = 256$ collocation points per step. Full-budget trainings run for 1000 epochs; the data-efficiency and generalization experiments use an identical, reduced budget of 500 epochs per retraining. The accuracy and physics-consistency metrics, the inverse-problem results and the robustness study are reported as mean $\pm$ one standard deviation over five random training seeds; the remaining experiments use a single fixed seed. Unless noted otherwise, all neural-network training and inference were performed on a single Apple M3 Max system on chip with 128 GB of unified memory, through the Metal Performance Shaders backend of PyTorch 2.1.3; all reported training times and inference latencies refer to this hardware. The finite-element reference solutions were computed in COMSOL Multiphysics 6.2 on an Intel Core i7-10700 CPU (2.90 GHz, 8 physical cores, 16 GB RAM). The data-efficiency and condition-transfer studies use a further reduced training set—a spatially and temporally subsampled set of about 4.6×10^5 samples, roughly a quarter of the full-resolution set used for Table 3—together with the reduced budget stated above; both models share that configuration and are compared within it, so the reported comparisons and trends are unaffected by this difference. All reported surrogate errors are defined with respect to the reference solution at the stated discretization.

4.4. Predictive Accuracy

Table 3 summarizes the errors on the held-out condition after 1000 training epochs. PINN-full attains an MAE of 0.0018 ± 0.0004 K, an RMSE of 0.0024 ± 0.0004 K, and a maximum pointwise error of 0.0138 ± 0.0011 K; the corresponding relative L_2 error, referred to the absolute temperature field, is 8.2×10^{-6}, which places the prediction well within typical engineering requirements for XLPE HVDC cable temperature-field modeling. The purely data-driven MLP attains a statistically indistinguishable accuracy (RMSE of 0.0023 ± 0.0001 K and maximum error of 0.0135 ± 0.0009 K): the difference in mean RMSE values is below 2×10^{-4} K before rounding, i.e., smaller than the seed-induced standard deviation of either model. PINN-phys, trained without any FEM sample, still recovers the field to an RMSE of 0.221 K, with a maximum error of 0.425 K, indicating that the embedded physics alone constrain the solution to within a few tenths of a kelvin.

Table 3. Accuracy under the held-out condition (20 °C, 100 A) and physics-consistency metrics. The PDE residual is the RMS of Equation (10) normalized by $\rho_\ell c_\ell$ and therefore carries the units of a temperature rate. Errors and residuals for PINN-full and MLP are mean ± one standard deviation over five random seeds; PINN-phys is trained once, and the spread quoted for its inference latency is the dispersion of repeated timing runs of that single model rather than a seed spread. Training times and inference latencies refer to the Apple M3 Max hardware. The training time of PINN-phys is not reported: that run shared the machine with a concurrent training workload, so its wall-clock time is not comparable.

Model	MAE (K)	RMSE (K)	Max Error (K)	PDE Residual (K s^{-1})	Training Time (h)	Inference (ms/10^5 pts)
PINN-full	0.0018 ± 0.0004	0.0024 ± 0.0004	0.0138 ± 0.0011	$(7.9 \pm 1.9) \times 10^{-3}$	10.8 ± 0.2	8.44 ± 0.23
MLP	0.0017 ± 0.0002	0.0023 ± 0.0001	0.0135 ± 0.0009	$(2.24 \pm 0.30) \times 10^{-1}$	9.2 ± 0.9	8.62 ± 0.25
PINN-phys	0.189	0.221	0.425	1.09×10^{-4}	—	8.56 ± 0.32

Beyond the scalar metrics, Figure 3 visualizes the predicted temperature field of the XLPE HVDC cable over the cross-section under a representative condition ($I = 150$ A, $T_{\text{amb}} = 20$ °C). The sequence reproduces the qualitative behavior established by the governing equations: the field starts uniform at the ambient temperature, heat accumulates in and around the conductor as Joule dissipation sets in, the isotherms organize into concentric rings following the layer interfaces, and the field approaches steady state on the slow time scale set by heat storage in the polymeric annuli—the snapshots at 300 and 600 min are nearly indistinguishable, with a final conductor temperature rise of about 1 K, consistent with the moderate load level. A single forward pass of the surrogate yields each of these full-field snapshots, which is the basis of the field-level monitoring envisaged for digital-twin use.

A quantitative cut through these fields is given in Figure 4, which compares predicted and reference radial profiles under the same condition. The prediction reproduces the layered character of the field: the nearly isothermal conductor plateau enforced by the high copper conductivity, the steep gradient across the conductor screen and the XLPE insulation, and the flattening toward the outer layers. The transient behavior is equally captured—the profiles rise progressively from the ambient start and approach steady state between 300 and 600 min. Small deviations below 0.1 K are visible during the early transient ($t \leq 60$ min), where the reference profiles exhibit a slight non-monotonic bump inside the conductor that the surrogate smooths out; these deviations decay as the field approaches steady state.

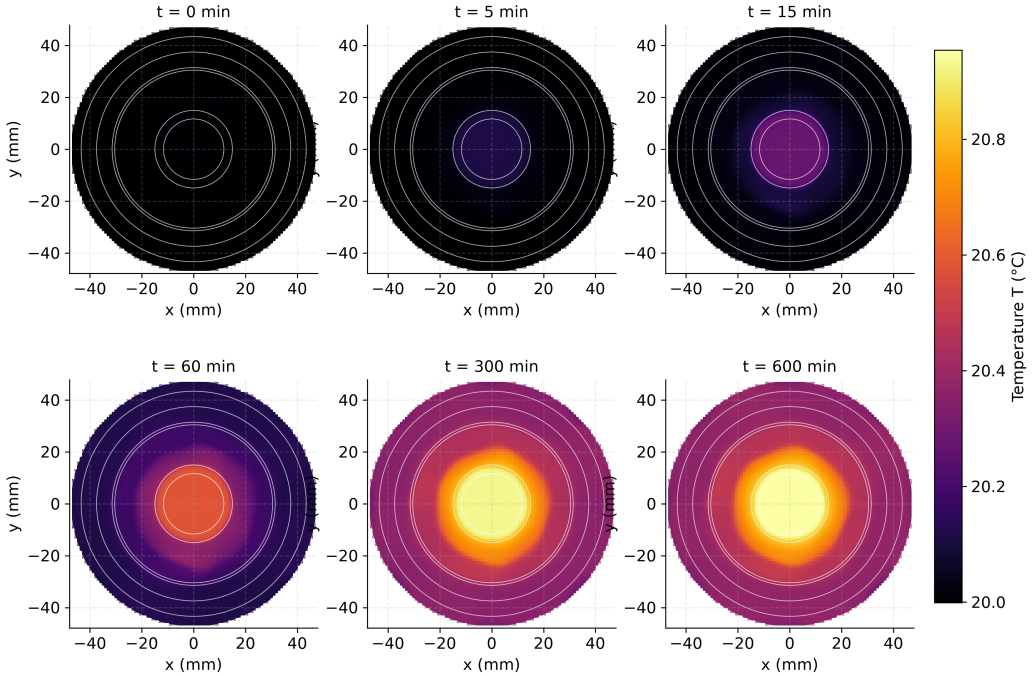


Figure 3. Predicted temperature field of the XLPE HVDC cable over the cross-section at $z = 4.5$ m for $I = 150$ A and $T_{\text{amb}} = 20$ °C, at six instants ($t = 0$, 5, 15, 60, 300, and 600 min). White circles mark the layer interfaces.

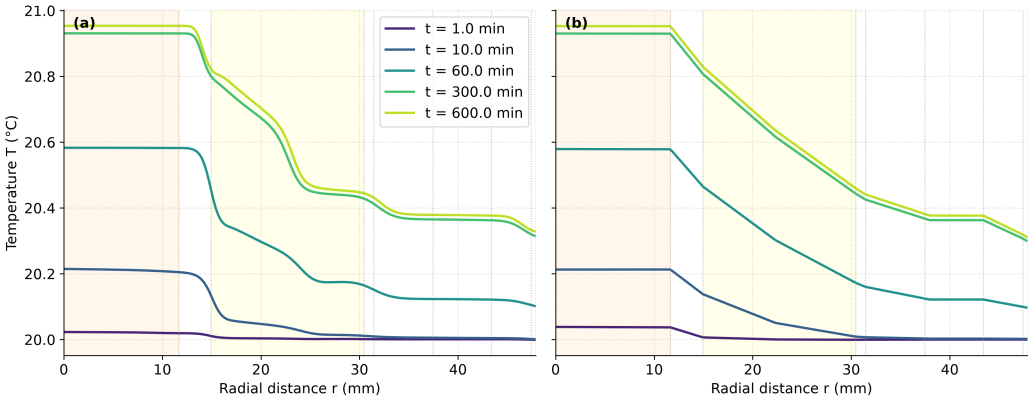


Figure 4. Radial temperature profiles at $z = 4.5$ m for $I = 150$ A and $T_{\text{amb}} = 20$ °C: (**a**) PINN-full prediction and (**b**) FEM reference at the same successive instants, with matching colors across the two panels (legend in panel (**a**)). Vertical dotted lines mark layer interfaces.

Finally, Figure 5 localizes the prediction error in space by mapping the absolute deviation ($|\hat{T}_\theta - T^{\text{FEM}}|$) over the cross-section at $t = 60$ min, the most demanding instant of the early transient identified above; since this map refers to the higher-load representative condition rather than the held-out condition, its local deviations exceed the tabulated metrics of Table 3. Even at this instant, the deviation stays below 0.1 K over most of the cross-section and is smallest in the conductor core; the largest values, up to about 0.18 K, are confined to isolated spots near the outer surface. Notably, the seven material interfaces do not act as error concentrators: apart from a faint ring of moderate deviation along the conductor screen, no systematic accumulation is visible where the thermal conductivity jumps by up to three orders of magnitude, indicating that the layer-aware residual handles the interface discontinuities without degrading the solution there. Under the uniform volumetric sampling of the collocation sets, the thin layers receive proportionally few collocation points—about 28 of the $N_p = 1024$ points per step for the insulation screen

($\ell = 3$, 2.7% of the cross-sectional area) and about 21 for the outer coating ($\ell = 7$, 2.1%)—but because the sets are resampled at every optimization step, the cumulative coverage of each thin layer over training approaches 10^7 points, and the per-layer residual statistics reported in Table 4 show no residual concentration in the thin annuli: for PINN-full, the insulation screen and the outer semi-conductive coating carry residuals no larger than those of the thicker polymeric annuli, and the residual instead concentrates mildly in the conductor and the aluminum sheath—the source-bearing layer and the layer of largest conductivity contrast. The purely data-driven MLP concentrates its residual at exactly those two layers (6.6×10^{-1} and 2.1×10^{-1}, factors of about 48 and 19 above PINN-full), whereas PINN-phys, which optimizes the residual alone, attains near-vanishing values in every layer except the source-bearing conductor, where its residual of 4.8×10^{-4} still lies about a factor of 30 below that of PINN-full. The per-layer values also reconcile arithmetically with the overall RMS of Table 3: under the uniform volumetric sampling, the squared overall residual is the volume-fraction-weighted mean of the squared per-layer values, i.e., $(\mathrm{RMS})^2 = \sum_\ell f_\ell \, \mathrm{RMS}_\ell^2$, with f_ℓ representing the volume fraction of layer ℓ; for PINN-phys, the conductor ($f_0 \approx 5.9\%$) dominates this sum, and $\sqrt{f_0} \times 4.8 \times 10^{-4} \approx 1.2 \times 10^{-4}$ recovers the tabulated 1.09×10^{-4} to within the sampling variability of the two independent 10^4-point draws. The lobed azimuthal pattern of the deviation admits three candidate contributions: smooth undulations of the Cartesian network representation of a nearly axisymmetric field, the finite collocation density near the layer interfaces, and the coarse azimuthal resolution of the free-triangular reference mesh, whose outer-ring vertices fall on multiples of 45°; the pattern decays together with the overall error level as the field approaches steady state, and none of the three contributions concentrates the error at the material interfaces. Thus, the error map complements the scalar metrics of Table 3 by showing that the reported accuracy is achieved broadly over the layered geometry rather than by averaging out localized failures.

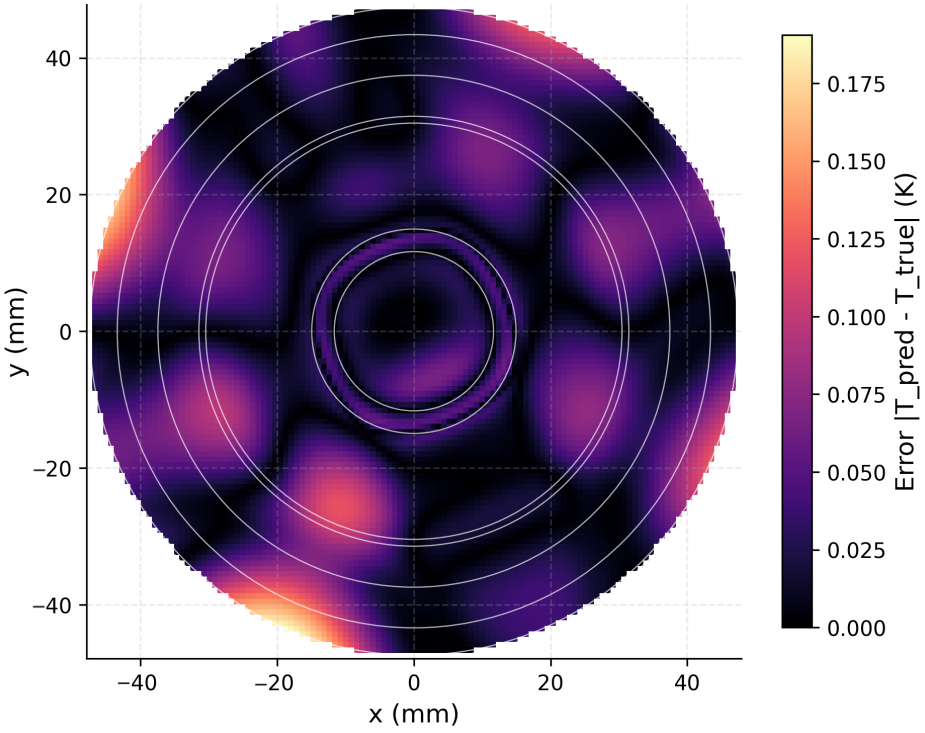


Figure 5. Absolute prediction error ($|\hat{T}_\theta - T^{\mathrm{FEM}}|$) over the cross-section of the XLPE HVDC cable at $z = 4.5$ m for $I = 150$ A and $T_{\mathrm{amb}} = 20\,^\circ$C at $t = 60$ min. White circles mark the layer interfaces.

Table 4. Per-layer RMS of the normalized PDE residual ($\mathrm{K\,s^{-1}}$) for the three models, evaluated on 10^4 interior collocation points sampled over the space–time–condition domain and grouped by layer, with one representative training seed per model; the sampling and normalization follow the residual metric of Table 3. Layer indices follow Table 2.

Layer	PINN-Full	MLP	PINN-Phys
0–Conductor	1.4×10^{-2}	6.6×10^{-1}	4.8×10^{-4}
1–Conductor screen	6.0×10^{-3}	9.2×10^{-4}	1.1×10^{-10}
2–XLPE insulation	2.4×10^{-3}	4.5×10^{-4}	1.0×10^{-10}
3–Insulation screen	1.9×10^{-3}	6.6×10^{-4}	1.1×10^{-10}
4–Water-blocking buffer	9.8×10^{-4}	5.9×10^{-4}	1.1×10^{-10}
5–Aluminum sheath	1.1×10^{-2}	2.1×10^{-1}	1.1×10^{-8}
6–Outer sheath	6.6×10^{-4}	3.8×10^{-4}	1.0×10^{-10}
7–Outer semi-conductive coating	1.2×10^{-3}	2.2×10^{-4}	1.0×10^{-10}

4.5. Physics Consistency and the Residual–Accuracy Trade-Off

The PDE-residual column of Table 3 and the residual distributions of Figure 6 jointly characterize how faithfully the three ablation baselines satisfy the governing equation of the XLPE HVDC cable. Figure 6 shows the density of $\log_{10}|\mathcal{R}_{\mathrm{PDE}}|$ over 10^4 random collocation points. PINN-phys is separated from the two data-supervised models across the whole distribution, whereas PINN-full and the MLP overlap over much of their bulk and differ mainly in the upper tail; the ordering of PINN-phys, PINN-full, and MLP from small to large residuals therefore refers to the tabulated RMS values, which, being dominated by the largest residuals, sit on the upper flank of each distribution. The separations of these RMS values—nearly two orders of magnitude between PINN-phys and PINN-full (1.09×10^{-4} versus 7.9×10^{-3}) and a further factor of about 28 up to the MLP—are consistent with that reading. Because the residual of Equation (10) is normalized by $\rho_\ell c_\ell$, it carries the units of a temperature rate ($\mathrm{K\,s^{-1}}$): the tabulated values measure the spurious local heating or cooling rate implied by the predicted field beyond the governing equation, which is precisely the quantity that physics consistency controls. The lower tails common to all three models extend toward numerical precision and arise at collocation points where the signed residual changes sign; they carry no information about model quality.

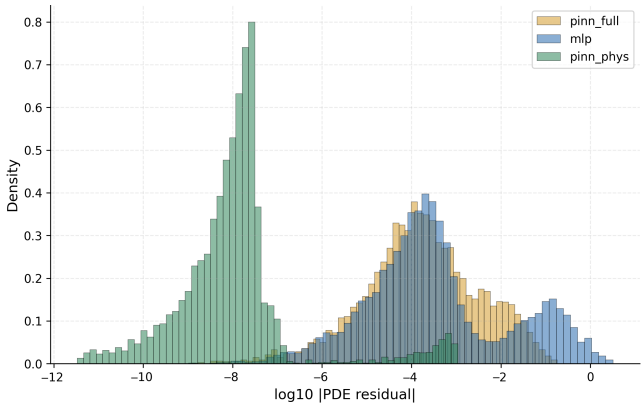


Figure 6. Distribution of $\log_{10}|\mathcal{R}_{\mathrm{PDE}}|$ over 10^4 random collocation points for the three models.

The joint reading of accuracy and residual reveals a structured trade-off. The MLP, despite its comparable pointwise accuracy, violates the governing equation by a factor of about 28 relative to PINN-full ($(2.24 \pm 0.30) \times 10^{-1}$ vs. $(7.9 \pm 1.9) \times 10^{-3}$): it fits the reference data closely without satisfying the underlying dynamics, so its predictions between and beyond the sampled conditions are unconstrained by the physics of the XLPE HVDC

cable. Conversely, PINN-phys attains the smallest residual (1.09×10^{-4}) yet the largest RMSE of 0.221 K. The combination is not contradictory; it reflects the non-uniqueness of the boundary-value problem under soft constraints. The physics losses admit a family of near-solutions of Equation (2), and without data anchoring, the optimizer may converge to a member whose absolute level is offset from the reference field while satisfying the differential operator almost exactly. PINN-full resolves the ambiguity: the data term selects the appropriate member of the family, and the physics term regularizes it toward equation consistency—a 28-fold residual reduction over the MLP at an accuracy cost that is within the seed-to-seed variability, i.e., statistically indistinguishable from zero. This residual–accuracy structure supports the interpretation that the physics loss is genuinely minimized during training rather than acting only as a generic regularizer.

Finally, the convergence of the Picard-type treatment of the electro-thermal coupling is verified empirically. Figure 7 tracks the relative epoch-to-epoch change of the lagged Joule source, i.e., $\|Q(\hat{T}_{\theta_{k+1}}) - Q(\hat{T}_{\theta_k})\|_2 / \|Q(\hat{T}_{\theta_k})\|_2$, evaluated on a fixed set of 4096 conductor collocation points under the corner condition (15 °C, 200 A), where the electro-thermal feedback is strongest; the per-epoch values reflect the combined effect of the Picard lag and the stochastic optimizer updates. The change decreases from a level of order 10^{-3} at the beginning of training to a median of 1.7×10^{-5} over the final 100 epochs; isolated excursions caused by individual stochastic updates remain below 5.1×10^{-3} and decay within a few epochs, and no divergence occurs at any stage. Thus, the lagged source stabilizes, training converges to the self-consistent fixed point, and the physics-consistency metrics of Table 3 accordingly quantify the residual of the converged surrogate with respect to the fully coupled governing equation. An ablation in which the detachment is removed and the source is differentiated through yields no benefit: in a single-seed controlled pair under an identical initialization and training budget, the fully differentiated variant attains a comparable held-out RMSE (0.0024 K versus 0.0020 K for the lagged scheme; these are single-run values of that controlled pair and are not the five-seed means of Table 3) and a roughly twofold larger PDE residual (1.27×10^{-2} versus 6.13×10^{-3}), while its own source change converges to the same level (1.6×10^{-5}), empirically confirming that the lagged treatment is not a limiting approximation for this problem.

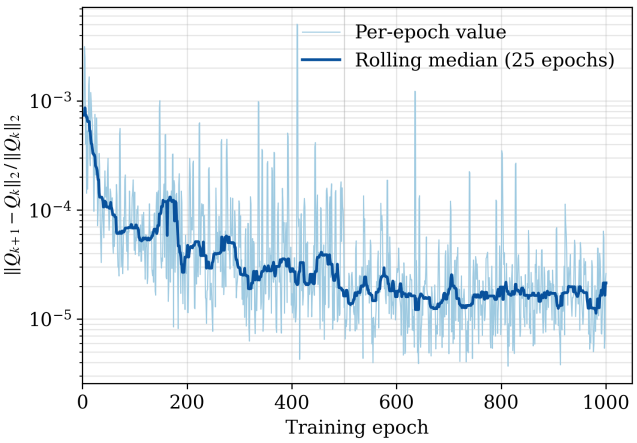


Figure 7. Relative epoch-to-epoch change of the lagged Joule source over training, evaluated on a fixed set of 4096 conductor collocation points at (15 °C, 200 A); the dark line is a rolling median over 25 epochs.

4.6. Data Efficiency

Table 5 and Figure 8 report the held-out RMSE when the training set is subsampled to fractions from 100% down to 5%, with both models retrained from scratch for each fraction under an identical, reduced training budget of 500 epochs; the 100% entries therefore differ

from the fully converged results of Table 3, which additionally employ the full-resolution training set and the full budget. Over the whole tested range, both models degrade monotonically and gracefully: even at 5% of the training data, the errors remain below 0.07 K, well within engineering requirements for XLPE HVDC cable temperature-field modeling. In absolute terms, the purely data-driven MLP remains the more accurate model at every tested fraction. This is attributable to the nature of the problem: the temperature field of the XLPE HVDC cable is smooth and of low amplitude, so even the 5% subset still contains roughly 2×10^4 labeled samples—enough for a pure regressor to interpolate such a field accurately.

Table 5. Test RMSE (K) versus training-data fraction. Obtained with the reduced training set of about 4.6×10^5 samples at the 500-epoch budget.

Model	100%	50%	25%	10%	5%
PINN-full	0.0063	0.0113	0.0354	0.0613	0.0684
MLP	0.0033	0.0058	0.0127	0.0493	0.0533

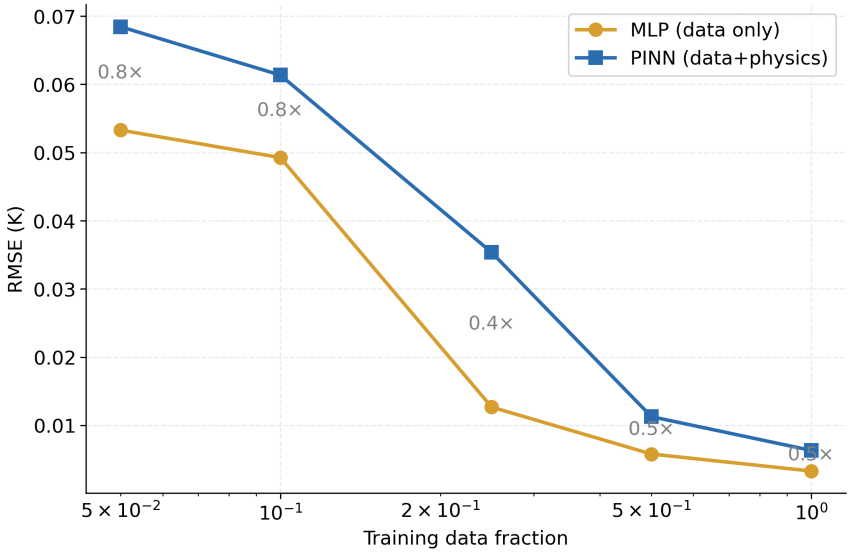


Figure 8. Data-efficiency response: We have modified the formats in the figure. curves: held-out RMSE versus training-data fraction (logarithmic abscissa). Gray annotations give the MLP-to-PINN error ratio at each fraction.

The picture changes, however, when the degradation is examined in relative terms. Between 100% and 5%, the MLP error grows by a factor of 16.2, whereas PINN-full grows by a factor of 10.8, and the MLP-to-PINN error ratio narrows accordingly from about 0.5 at full data to about 0.8 at 5%. Thus, the relative value of the physics prior increases as labeled data become scarcer, consistent with its role of supplying supervision where reference samples are missing; within the data regime tested here, its principal contribution is equation consistency rather than raw pointwise accuracy.

4.7. Generalization to Unseen Operating Conditions

Because the surrogate is defined over the continuous envelope ($\mathcal{E}$), its behavior under operating conditions absent from the training data is of direct practical interest for XLPE HVDC cable temperature-field modeling. Three condition-transfer scenarios are examined, with both models retrained from scratch per scenario under the same reduced budget of 500 epochs: current interpolation, current extrapolation, and ambient-temperature

extrapolation. Table 6 reports RMSE and MAE; Figure 9 shows the prediction–reference scatter for the two current-transfer scenarios.

Table 6. Generalization to unseen conditions: RMSE (MAE) in K. Obtained with the reduced training set of about 4.6×10^5 samples at the 500-epoch budget; each scenario uses its own retraining, so the entries refer to three separately trained surrogates per model.

Scenario	Train/Test Conditions	PINN-Full	MLP
Current interpolation	$I \in \{0, 50, 150, 200\}$ A/$I = 100$ A	0.006 (0.005)	0.007 (0.006)
Current extrapolation	$I \leq 100$ A/$I \in \{150, 200\}$ A	0.338 (0.178)	0.161 (0.113)
Ambient extrapolation	$T_{\text{amb}} \in \{15, 20\}$ °C/25 °C	0.102 (0.050)	0.150 (0.057)

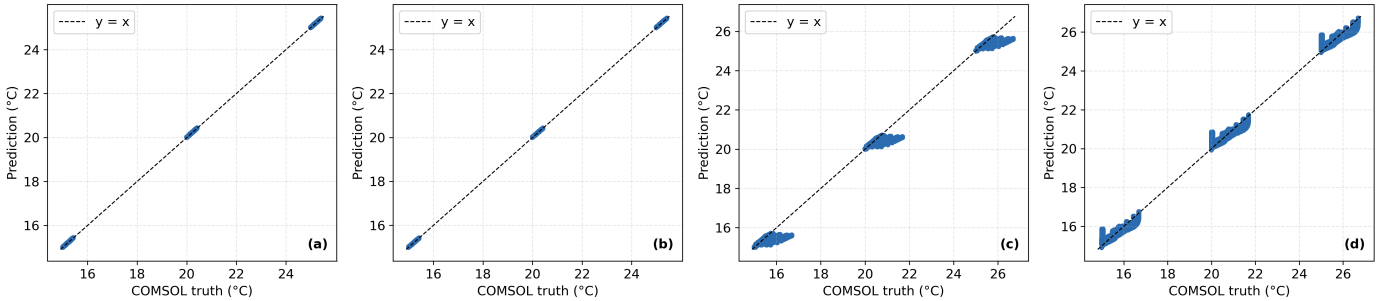


Figure 9. Predicted versus reference temperatures, with the dashed line marking $y = x$ and each panel reporting its RMSE: (**a**) PINN-full and (**b**) MLP under current interpolation (test $I = 100$ A); (**c**) PINN-full and (**d**) MLP under current extrapolation (test $I \in \{150, 200\}$ A).

Under current interpolation, both models predict the unseen 100 A condition accurately, but PINN-full holds a consistent advantage on both metrics: its RMSE of 0.006 K is about 14% below the 0.007 K of the MLP, with a comparable MAE margin (0.005 K versus 0.006 K, about 17%). The advantage is attributable to physics supervision: the collocation points are sampled over the continuous envelope and therefore also at intermediate currents, supplying equation-consistent supervision precisely in the gap between the training conditions, where the MLP must rely on the smoothness of the network alone. Thus, within the sampled range, the physics prior refines the interpolation of the condition embedding beyond what the shared output parametrization already provides. The two extrapolation scenarios, however, separate the models in opposite directions, and the contrast is instructive. At the unseen ambient temperature, PINN-full is the more accurate model (RMSE of 0.102 K versus 0.150 K, about 32% below); the much smaller MAE margin (0.050 K versus 0.057 K, about 12%) indicates that the MLP error distribution carries heavier tails. The advantage is explained by the structure of the physics supervision: the ambient temperature enters the convective boundary residual (11) explicitly, so the boundary term supplies correct, condition-specific supervision at ambient temperatures for which no reference data exist. This conclusion is, however, tied to the narrow 10 K ambient window of the studied envelope, over which the temperature field is quasi-linear in T_{amb}; seasonal soil or air variations of 30–50 K and a convective coefficient that depends on wind speed would render the ambient dependence nonlinear, and the demonstrated advantage may not transfer to problems with nonlinear ambient dependence. Current extrapolation is the hardest scenario for both models, and here, the ranking reverses: the MLP degrades to 0.161 K, whereas PINN-full reaches 0.338 K—about twice as large, with the scatter of Figure 9c revealing a systematic saturation of the PINN-full prediction below the reference at the highest loads. The mechanism is consistent with the non-uniqueness of the boundary-value problem under soft constraints: the load current acts through the nonlinear Joule source confined to the conductor, and at unseen currents, the soft PDE constraint—unanchored

by data—admits near-solutions of reduced amplitude toward which the optimizer drifts. The MLP, by contrast, benefits from the quadratic current dependence built into the shared output parametrization (9), which closely matches the nearly quadratic load response of the DC problem.

Taken together, the three scenarios delineate where the physics prior aids in generalization in XLPE HVDC cable temperature-field modeling: it improves in-range interpolation and transfer to conditions that enter the embedded constraints explicitly, as for the ambient temperature, but does not, by itself, resolve the solution-family ambiguity introduced by the nonlinear heat source at unseen currents. In every scenario, the condition-embedded surrogate remains usable over the envelope, with a worst-case RMSE of 0.34 K across the three transfer settings.

4.8. Computational Efficiency

The computational figures of Table 3 quantify the cost of the proposed framework on both sides of the workflow, measured on the Apple M3 Max hardware. On the training side, the physics supervision is inexpensive: PINN-full trains in 10.8 h, against 9.2 h for the purely data-driven MLP, so evaluating the layer-aware residuals on resampled collocation points adds about 18% to the training cost. This one-off cost is amortized over all subsequent evaluations of the surrogate.

On the inference side, the physics supervision is free by construction: once trained, all three models reduce to the same forward pass through the network, and their measured latencies are statistically indistinguishable (8.4–8.6 ms per 10^5 space–time–condition queries, i.e., roughly 85 ns per query in batched evaluation). Thus, equation consistency is obtained at zero inference cost. For comparison, a transient FEM run of a single operating condition over the 600 min horizon requires about 70 s in COMSOL Multiphysics 6.2 on the Intel Core i7-10700 CPU with the reference mesh. Reconstructing the full space–time field of one condition from the surrogate—1.19×10^6 mesh-node–instant queries in twelve batches—takes about 0.10 s, corresponding to a speed-up of nearly three orders of magnitude; for the isolated pointwise queries required by monitoring and rating tasks, the advantage grows further, since the finite-element model cannot be evaluated at a single point without solving the entire initial-boundary-value problem. The comparison is indicative rather than strict: it spans different hardware and execution modes (batched inference on the M3 Max against a CPU solve under the solver's default threading), and the finite-element run yields the complete field by construction. The trained surrogate is also compact, comprising roughly 6.7×10^4 parameters (about 0.3 MB in single precision) in place of gigabyte-scale FEM meshes and solution files. The measured latency readily satisfies the sub-second budgets of dynamic-rating and digital-twin loops for XLPE HVDC cable temperature-field modeling.

4.9. Inverse Problems from Sparse Measurements

Because $\hat{T}_\theta$ is differentiable with respect to its condition inputs, an unknown condition ($\xi \in \{I,\ T_{\mathrm{amb}}\}$) can be identified from N_s sparse measurements ($\{T_j^{\mathrm{meas}}\}$) by freezing the trained parameters (θ^*) and descending

$$\xi^* = \arg\min_{\xi} \ \frac{1}{N_s} \sum_{j=1}^{N_s} \left[\hat{T}_{\theta^*}(\boldsymbol{x}_j, t_j; \xi) - T_j^{\mathrm{meas}}\right]^2, \qquad \xi \in \mathcal{E}, \tag{14}$$

where ξ^* is the identified condition, N_s is the number of sensors, and T_j^{meas} is the temperature measured at the j-th sensor location $(\boldsymbol{x}_j, t_j)$. This capability is exercised in three scenarios under the reference condition $(20\,^\circ\mathrm{C}, 150\,\mathrm{A})$, each using the same array of $N_s = 10$ virtual surface sensors extracted from the FEM solution, whose exact locations and reading instants are listed in Table 7; Table 8 summarizes the results.

Table 7. Locations and reading instants of the ten virtual surface sensors, drawn once from the FEM export grid and shared by all inverse scenarios and seeds; φ denotes the azimuth in the cable cross-section. Sensors lie on the outermost two node rings ($r = 47.9$ mm, the outer surface, and $r = 47.4$ mm, the sheath–coating interface 0.5 mm beneath it; the temperature difference across the thin coating is below 5×10^{-3} K).

j	r (mm)	φ (deg)	z (m)	t (min)
1	47.4	0	0.62	281
2	47.9	0	7.14	321
3	47.9	45	0.62	521
4	47.4	135	5.90	211
5	47.4	−45	3.72	431
6	47.9	180	4.03	511
7	47.9	−90	6.52	471
8	47.4	135	0.62	261
9	47.4	−135	1.86	471
10	47.4	180	8.07	191

Table 8. Inverse-problem results with $N_s = 10$ virtual surface sensors at (20 °C, 150 A), reported as mean ± one standard deviation over the five independently trained surrogates of Table 3. The sensor array (Table 7) is identical across seeds and scenarios, so the spread isolates training-seed variability. Relative errors of temperature quantities are referred to the Celsius scale.

Scenario	Ground Truth	Identified	Abs. Error	Rel. Error
Conductor hotspot (soft sensing)	20.957 °C	20.955 ± 0.001 °C	$(2.5 \pm 1.5) \times 10^{-3}$ K	$0.012 \pm 0.007\%$
Load-current identification (init. 100 A)	150.0 A	150.21 ± 0.88 A	0.74 ± 0.39 A	$0.49 \pm 0.26\%$
Ambient-temperature identification (init. 22 °C)	20.0 °C	20.001 ± 0.004 °C	$(3.0 \pm 1.6) \times 10^{-3}$ K	$0.015 \pm 0.008\%$

The first scenario addresses the quantity of greatest engineering interest for an XLPE HVDC cable in service: the conductor hotspot, which is inaccessible to direct measurement. Here, the surrogate acts as a soft sensor—its agreement with the surface measurements validates the operating state, and a direct forward query at the conductor then infers the hotspot to within 0.0025 ± 0.0015 K of the FEM reference (0.0041 K in the worst seed). The remaining two scenarios identify unknown operating conditions through Equation (14). Starting from a deliberately distant initial guess of 100 A, the gradient descent recovers the true load current of 150 A to 150.21 ± 0.88 A ($0.49 \pm 0.26\%$) in about 2 s; the equivalent ambient temperature, initialized at 22 °C, is recovered to 20.001 ± 0.004 °C (0.015%). A comparable FEM-based identification would typically require repeated forward solves and correspondingly longer computation. Combined with the millisecond-scale inference, these results make the surrogate a suitable candidate for state estimation within a digital twin of the XLPE HVDC cable.

The placement of Table 7 also clarifies why the inverse problems are well posed despite the sparsity of the array: each identification concerns a single scalar, the surface temperature is monotone in both the load current and the ambient temperature, and the sensors span the axial extent and several azimuths, so the array is informationally redundant for the noise-free case; the value of multiple sensors lies in noise averaging.

Table 9 examines the identification under realistic measurement imperfections. Two regimes emerge. The ambient temperature remains robust: a uniform shift of all readings is essentially a mean estimate, and the error at $\sigma = 0.5$ K (0.121 K) stays at the level of the $\sigma/\sqrt{N_s}$ benchmark of 0.16 K. The load current, by contrast, is noise-limited in the studied operating envelope: the surface-temperature rise at 150 A is only about 0.3 K, so thermocouple-level noise exceeds the signal itself, and the errors follow the $\sigma/\sqrt{N_s}$ scaling of a noise-limited scalar estimate to within about 10% across $N_s \in \{3, 5, 10\}$. Two

consequences follow. First, the single-snapshot setting studied here is a worst case: because the errors scale as $\sigma/\sqrt{N}$, routine temporal averaging—some 10^4 readings from the same ten sensors, i.e., tens of minutes at typical monitoring sampling rates—restores ampere-level current accuracy. Second, the reduced envelope of the present study is, itself, the adverse regime; at rated loads, the surface signal is one to two orders of magnitude larger, and the same noise corresponds to a relative current error of a few tenths of a percent. The bias case exposes a distinct mechanism: a uniform $+0.5$ K calibration offset is observationally indistinguishable from an ambient-temperature shift and is absorbed by the ambient estimate almost exactly (0.501 ± 0.004 K), whereas the current estimate maps the same offset quadratically into a 94 A excursion. Sensor calibration therefore matters more than sensor precision for current identification, and the chained hotspot error is, at most, about 0.7 K under all considered noise cases—small against the 90 °C XLPE limit of the insulation—while reaching 1.6 K only under the uncompensated bias.

Table 9. Robustness of the inverse identification to sensor noise, reduced sensor count, and calibration bias at (20 °C, 150 A). Gaussian noise of standard deviation (σ) is added to the sensor readings; reduced counts keep the first k sensors of Table 7, emulating partial sensor failure; the bias case offsets all readings by $+0.5$ K. Entries are mean $\pm$ one standard deviation over five training seeds, each averaged over ten noise realizations. The hotspot error refers to the practically relevant chain of identifying the current from the noisy sensors, then querying the conductor temperature.

Case	Current Error (A)	Ambient Error (K)	Hotspot Error (K)
$\sigma = 0$, $N_s = 10$ (reference)	0.74 ± 0.39	0.003 ± 0.002	0.008 ± 0.005
$\sigma = 0.5$ K, $N_s = 10$	32.2 ± 3.2	0.121 ± 0.018	0.38 ± 0.05
$\sigma = 1.0$ K, $N_s = 10$	63.6 ± 12.7	0.241 ± 0.036	0.70 ± 0.13
$\sigma = 0.5$ K, $N_s = 5$	41.5 ± 10.9	0.149 ± 0.023	0.44 ± 0.09
$\sigma = 0.5$ K, $N_s = 3$	57.6 ± 16.7	0.203 ± 0.042	0.56 ± 0.10
Bias $+0.5$ K, $N_s = 10$	94.0 ± 1.5	0.501 ± 0.004	1.57 ± 0.03

4.10. Sensitivity to the Loss Weights

The robustness of the composite objective to its weighting is quantified by retraining PINN-full with each weight (λ_d, λ_p, λ_b, and λ_i) increased and decreased by one order of magnitude around the default $(1, 0.05, 1, 1)$, holding every remaining hyperparameter fixed. Because the balance between the data term and the physics terms is most consequential precisely when labeled data are scarce, the study deliberately adopts a reduced, sensitivity-only configuration—250 randomly selected spatial nodes per condition (2.1×10^5 samples), a 100-epoch budget, and a single fixed seed—which makes it a conservative stress test of the weighting. The nine trainings were performed on an NVIDIA H200 GPU rather than on the Apple M3 Max used elsewhere; since the study compares the nine settings only against one another and the reduced configuration already makes the absolute values incomparable to Table 3, the change of device does not affect the conclusions drawn here. The held-out condition and the evaluation metrics are unchanged. Table 10 reports the held-out accuracy metrics and the sampled PDE residual for the nine resulting settings.

Three observations emerge. First, the default weighting attains the lowest value of all three accuracy metrics, and none of the eight one-decade perturbations improves on it; thus, the sweep provides an independent check on the manual selection of the base weights, which it does not displace. The held-out RMSE of the perturbed settings exceeds the default by factors of between about 1.1 and 1.8. Because each cell is a single training run, the ordering among the mildest perturbations should not be over-interpreted; the informative outcome is the bounded response. Second, no perturbed setting diverges or destabilizes training, and the worst-case RMSE across the full decade-wide range of every weight remains near 0.10 K, so the method does not hinge on a fragile weight choice. Third, the $\lambda_i \times$

10 setting attains the smallest sampled residual (3.9×10^{-4}) while simultaneously yielding the largest held-out error, echoing the residual–accuracy dissociation exposed by the data-free ablation: a small sampled residual alone does not certify an accurate temperature field. Systematic weighting schemes such as neural-tangent-kernel balancing and self-adaptive weighting [22,23] remain compatible drop-in refinements should substantially broader envelopes require re-balancing.

Table 10. Sensitivity of the held-out metrics to one-decade changes of each loss weight around the default ($(\lambda_d, \lambda_p, \lambda_b, \lambda_i) = (1, 0.05, 1, 1)$). All runs use a sensitivity-only reduced configuration (250 randomly selected spatial nodes per condition, 2.1×10^5 samples, a 100-epoch budget, and a single fixed seed); absolute values are therefore not comparable to those reported in Table 3. The PDE residual is the same $\rho_\ell c_\ell$-normalized quantity as in Table 3, in $\mathrm{K\,s^{-1}}$.

Setting	MAE (K)	RMSE (K)	Max (K)	PDE Residual ($\mathrm{K\,s^{-1}}$)
Default	0.0425	0.0571	0.170	2.6×10^{-2}
$\lambda_d \times 10$	0.0603	0.0737	0.184	3.7×10^{-2}
$\lambda_d/10$	0.0650	0.0870	0.208	6.5×10^{-3}
$\lambda_p \times 10$	0.0856	0.1017	0.240	2.9×10^{-2}
$\lambda_p/10$	0.0785	0.0974	0.230	2.8×10^{-2}
$\lambda_b \times 10$	0.0638	0.0824	0.229	1.8×10^{-2}
$\lambda_b/10$	0.0487	0.0612	0.189	3.2×10^{-2}
$\lambda_i \times 10$	0.0807	0.1026	0.236	3.9×10^{-4}
$\lambda_i/10$	0.0592	0.0728	0.191	3.2×10^{-2}

4.11. The Cable-End Effect over an Extended Operating Envelope

This subsection exercises the genuinely three-dimensional capability of the framework on an axially and layerwise heterogeneous cooling configuration. On the lateral surface, $h = 12\ \mathrm{W\,m^{-2}\,K^{-1}}$ acts over the two end meters against $7.5\ \mathrm{W\,m^{-2}\,K^{-1}}$ over the mid-span; the end faces are adiabatic at the conductor and its screen and convective at the outer annuli ($9.8\ \mathrm{W\,m^{-2}\,K^{-1}}$ for layers 2–4 and $15\ \mathrm{W\,m^{-2}\,K^{-1}}$ for layers 5–7). The cooling contrast induces axial transition layers of analytically estimated width ($\delta_{\mathrm{ax}} = \sqrt{k_0 A_{\mathrm{c}} R'} \approx 0.38$ m, where $R' \approx 1.08\ \mathrm{K\,m\,W^{-1}}$ denotes the mid-span radial thermal resistance per unit of length). The study simultaneously exercises a second cable configuration: the layer radii of Table 2 are retained, but the material set differs throughout (Table 11), led by an alloy conductor with $k_0 = 320\ \mathrm{W\,m^{-1}\,K^{-1}}$, so the thermal-conductivity contrasts across the layers change as well. This second set is a controlled variation rather than a cataloged material specification: the conductor retains the density, specific heat, and electrical conductivity of copper while its thermal conductivity is lowered so that the effect of the altered conductivity contrasts is isolated from that of the heat capacities. Its Joule source uses the constant conductivity of $\sigma_{20} = 5.8 \times 10^7\ \mathrm{S\,m^{-1}}$ without temperature dependence, so the electro-thermal feedback and its Picard treatment pertain to the baseline configuration only. Had the temperature dependence of the conductivity been retained in this configuration, the loop gain ($\alpha_{\mathrm{Cu}}\,\Delta T_{\mathrm{rise}}$) of Section 3.3 would have reached about 0.06 under the held-out condition (conductor rise of 16.2 K) and about 0.26 in the 1200 A corner (rise of about 65 K)—beyond the 10^{-2} figure quoted for the baseline envelope, though still within the sub-0.3 contraction bound stated there; no such iteration arises here, since the conductivity is constant. The reference dataset spans a substantially extended envelope of $I \in \{0, 200, \ldots, 1200\}$ A and $T_{\mathrm{amb}} \in \{-20, -10, \ldots, 40\}$ °C (49 conditions) so that the highest-load corner brackets the 90 °C XLPE limit (steady mid-span conductor rise of about 65 K at 1200 A, following the exactly quadratic scaling implied by the temperature-independent conductivity of this configuration). The reference data use a dedicated swept mesh with a three-zone axial distribution of 3–9–3 linear elements, refined toward the end sections, giving 1056 mesh

nodes per condition (66 cross-sectional nodes at 16 axial stations), with the cross-sectional discretization of the baseline reference data retained; the refined end zones resolve the transition layers of width $\delta_{ax} \approx 0.38$ m, which lie entirely within the two end meters.

Table 11. Material properties of the second cable configuration used in the end-effect study; the layer radii coincide with those of Table 2.

Layer (ℓ)	k_ℓ (W m^{-1} K^{-1})	ρ_ℓ (kg m^{-3})	c_ℓ (J kg^{-1} K^{-1})
0 Conductor (Cu alloy)	320.0	8960	385
1 Conductor screen	0.35	1020	1800
2 XLPE insulation	0.30	930	2200
3 Insulation screen	0.35	1050	1800
4 Water-blocking buffer	0.45	950	1500
5 Aluminum sheath	205.0	2700	900
6 Outer sheath	0.22	950	1600
7 Outer semi-conductive coating	0.26	950	1400

A single surrogate is trained on the 48 remaining conditions (3.09×10^6 samples) for 300 epochs (6.2 h on the Apple M3 Max) and evaluated on the held-out condition (10 °C, 600 A); the architecture and losses are identical to the baseline, with only the boundary residuals and the normalization scales exchanged. Table 12 summarizes the errors. The surrogate reconstructs the held-out field to an RMSE of 0.077 K and a maximum pointwise error of 0.39 K, against an axial temperature contrast of 4.9 K and a conductor rise of 16.2 K under this condition—a relative L_2 error of 2.7×10^{-4}, again referred to as the absolute temperature field. The error is essentially uniform between the end regions (0.073 K RMSE) and the mid-span (0.080 K), so the three-dimensional transition layers do not act as error concentrators. Figure 10 shows the reconstructed r–z field near steady state, and Figure 11 compares the axial centerline profiles with the FEM reference at three instants: the early-transient profile at 61 min carries only a weak axial contrast, which subsequently develops into the fully formed plateau–transition structure, and the surrogate tracks the reference through all stages.

Table 12. Region-split errors under the held-out condition (10 °C, 600 A) of the end-effect study. End regions denote $z < 1$ m and $z > 8$ m; the mid-span denotes the remainder.

Region	MAE (K)	RMSE (K)	Max Error (K)
Overall	0.061	0.077	0.39
End regions	0.055	0.073	0.39
Mid-span	0.065	0.080	0.36

Two independent physics checks corroborate the reconstruction. First, the mid-span conductor rise of 16.19 K in the reference data (at $t = 591$ min) agrees to within 3% with the thermal-circuit prediction of 15.73 K obtained from the line heat rate at 600 A and the radial resistances of the second configuration. Second, the end-region conductor plateau of 11.76 K falls below the lateral-only local equilibrium of 13.31 K; the deficit is the direct signature of the additional heat extracted through the convective end faces, and the transition between the two plateaus occurs over a length consistent with the analytical width of $\delta_{ax} \approx 0.38$ m. Together with the extended envelope, the study demonstrates that one condition-embedded surrogate accommodates a genuinely three-dimensional temperature field, a second material configuration, and a spatially varying convective coefficient without structural changes to the framework.

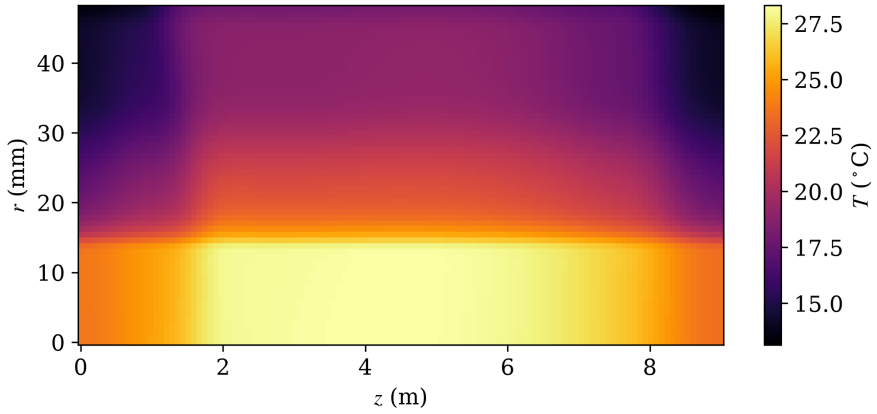


Figure 10. Predicted temperature of the end-effect configuration over the r–z plane at $t = 591$ min (near steady state) under the held-out condition (10 °C, 600 A); the enhanced cooling of the end sections and the axial transition layers adjacent to $z = 1$ and $z = 8$ m is reproduced.

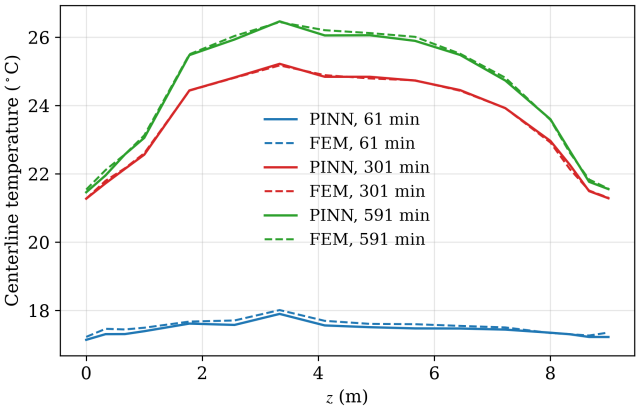


Figure 11. Axial centerline temperature ($T(0, 0, z)$) of the end-effect configuration at 61, 301, and 591 min: PINN (solid) versus FEM reference (dashed) under the held-out condition (10 °C, 600 A).

5. Conclusions

This paper presented a physics-informed neural network for XLPE HVDC cable temperature-field modeling, combining a layer-aware material mapping, a temperature-dependent Joule source with a convergent Picard treatment, a physics-based output parametrization, and a condition embedding spanning a continuous current–ambient envelope. Against finite-element references under fifteen operating conditions, it achieved an RMSE of 0.0024 K (mean over five training seeds) relative to the discrete finite-element reference (mesh-discretization error of about 0.04 K) under a held-out condition while reducing the governing-equation residual by approximately 28-fold relative to an identically sized data-driven network, at about 18% extra training cost and with no inference overhead. The data-free ablation exposed a residual–accuracy trade-off rooted in the non-uniqueness of the problem under soft constraints.

The physics prior made degradation under data reduction more graceful and improved ambient-temperature extrapolation within the studied quasi-linear window; current extrapolation, acting through the nonlinear Joule source, is not resolved by the soft constraint alone. At below 9 ms per 10^5 queries, the surrogate identified the load current to 0.49% and the unmeasurable conductor hotspot to within 0.0025 K from ten noise-free virtual sensors. Under realistic sensor noise, the ambient estimate and hotspot soft sensing remain robust, whereas single-snapshot current identification in this low-load envelope becomes noise-limited, and sensor calibration matters more than sensor precision. The reported spreads are seed-to-seed variability over five independently trained surrogates rather than

calibrated predictive uncertainty, and the evidence covers two fixed cable designs with prescribed convection and purely numerical reference data. Future work will parameterize geometry and convection as additional operator inputs, strengthen current extrapolation, develop calibrated uncertainty quantification, and validate the framework on experimental cable-loop data.

Author Contributions: Conceptualization, J.-X.H. and Y.Z.; methodology, J.-X.H.; software, resources, and validation, K.-J.R.; validation, J.-J.D. and S.-H.J.; formal analysis, H.-Y.Y.; investigation, L.-Z.Z.; methodology and validation, H.-S.Z.; writing—original draft preparation, J.-X.H.; writing—review and editing, Y.Z. and W.L.; supervision, W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of China Electrical Equipment Group Co., Ltd. (Development of Digital Twin Platform for Power Transmission and Transformation Equipment, Twin Construction, and Service Performance Evolution Services).

Data Availability Statement: The cable design parameters and the finite-element reference data originate from an industrial partner and are subject to confidentiality obligations. Reference temperature data regenerated from the parameters published in this article, covering the key operating conditions, including the held-out conditions, are openly available at 31 December 2025 https://github.com/luweidut/xlpe-hvdc-cable-pinn, with the approval of Chongqing Taishan Cable Co., Ltd. Further data supporting the reported results are available from the corresponding author upon reasonable request, subject to the same approval.

Conflicts of Interest: Authors Jia-Xun He, Ya Zhang, Jun-Jie Ding, Kang-Jie Ruan, Shuo-Han Jing, Hai-Yan Yang, and Ling-Zhi Zhu were employed by Chongqing Taishan Cable Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare that this study received funding from Science and Technology Project of China Electrical Equipment Group Co., Ltd. (Development of Digital Twin Platform for Power Transmission and Transformation Equipment, Twin Construction, and Service Performance Evolution Services). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

References

1. *IEC 60287-1-1:2023*; Electric Cables–Calculation of the Current Rating–Part 1-1: Current Rating Equations (100% Load Factor) and Calculation of Losses–General. International Electrotechnical Commission: Geneva, Switzerland, 2023.
2. Mazzanti, G.; Marzinotto, M. *Extruded Cables for High-Voltage Direct-Current Transmission: Advances in Research and Development*; Wiley: Hoboken, NJ, USA, 2013.
3. De Leon, F.; Anders, G.J. Effects of backfilling on cable ampacity analyzed with the finite element method. *IEEE Trans. Power Deliv.* **2008**, *23*, 537–543. [CrossRef]
4. Zhang, Y.; Chen, X.; Zhang, H.; Liu, J.; Zhang, C.; Jiao, J. Analysis on the temperature field and the ampacity of XLPE submarine HV cable based on electro-thermal-flow multiphysics coupling simulation. *Polymers* **2020**, *12*, 952. [CrossRef] [PubMed]
5. Bragatto, T.; Cresta, M.; Gatta, F.; Geri, A.; Maccioni, M.; Paulucci, M. A 3-D nonlinear thermal circuit model of underground MV power cables and their joints. *Electr. Power Syst. Res.* **2019**, *173*, 112–121. [CrossRef]
6. Zhou, X.; Li, Y.; Zhao, S.; Wang, X.; Chen, L.; Cheng, M.; Huang, L. Cable Temperature Prediction Algorithm Based on the MSST-Net. *Electricity* **2026**, *7*, 6. [CrossRef]
7. Han, W.; Hao, C.; Kong, D.; Yang, G. Cable Temperature Prediction Based on RF-GPR for Digital Twin Applications. *Appl. Sci.* **2023**, *13*, 7700. [CrossRef]
8. Zhao, L.; Lin, G.; Wang, Y.; Xie, B.; Wan, C.; Zhang, H.; Tian, S.; Bai, Z.; Zhang, M.; Jin, X. Research on temperature prediction of shearer cable based on bidirectional long short-term memory. *Int. J. Therm. Sci.* **2025**, *210*, 109597. [CrossRef]
9. Sarkhosh, M.H.; Edrisnia, H.; Raveshi, M.R.; Sharbatdar, M. Prediction of time averaged wall shear stress distribution in coronary arteries' bifurcation varying in morphological features via deep learning. *Front. Physiol.* **2025**, *16*, 1518732. [CrossRef] [PubMed]
10. Zhao, X.; Gong, Z.; Zhang, Y.; Yao, W.; Chen, X. Physics-informed convolutional neural networks for temperature field prediction of heat source layout without labeled data. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105516. [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. Cai, S.; Wang, Z.; Wang, S.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks for heat transfer problems. *J. Heat Transf.* **2021**, *143*, 060801. [CrossRef]
13. Xu, J.; Wei, H.; Bao, H. Physics-informed neural networks for studying heat transfer in porous media. *Int. J. Heat Mass Transf.* **2023**, *217*, 124671. [CrossRef]
14. Misyris, G.S.; Venzke, A.; Chatzivasileiadis, S. Physics-informed neural networks for power systems. In *Proceedings of the 2020 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, Canada, 2–6 August 2020*; IEEE: Piscataway, NJ, USA, 2020; pp. 1–5.
15. Stiasny, J.; Zhang, B.; Chatzivasileiadis, S. Pinnsim: A simulator for power system dynamics based on physics-informed neural networks. *Electr. Power Syst. Res.* **2024**, *235*, 110796. [CrossRef]
16. Bragone, F.; Morozovska, K.; Hilber, P.; Laneryd, T.; Luvisotto, M. Physics-informed neural networks for modelling power transformer's dynamic thermal behaviour. *Electr. Power Syst. Res.* **2022**, *211*, 108447. [CrossRef]
17. Odeback, O.W.; Bragone, F.; Laneryd, T.; Luvisotto, M.; Morozovska, K. Physics-informed neural networks for prediction of transformer's temperature distribution. In *Proceedings of the 2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA), Nassau, The Bahamas, 12–14 December 2022*; IEEE: Piscataway, NJ, USA, 2022; pp. 1579–1586.
18. Laneryd, T.; Bragone, F.; Morozovska, K.; Luvisotto, M. Physics informed neural networks for power transformer dynamic thermal modelling. *IFAC-PapersOnLine* **2022**, *55*, 49–54. [CrossRef]
19. Jagtap, A.D.; Mao, Z.; Adams, N.; Karniadakis, G.E. Physics-informed neural networks for inverse problems in supersonic flows. *J. Comput. Phys.* **2022**, *466*, 111402. [CrossRef]
20. Lu, L.; Jin, P.; Pang, G.; Zhang, Z.; Karniadakis, G.E. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nat. Mach. Intell.* **2021**, *3*, 218–229. [CrossRef]
21. Li, Z.; Kovachki, N.; Azizzadenesheli, K.; Liu, B.; Bhattacharya, K.; Stuart, A.; Anandkumar, A. Fourier Neural Operator for Parametric Partial Differential Equations. In Proceedings of the International Conference on Learning Representations (ICLR), Virtual, 3–7 May 2021.
22. 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]
23. McClenny, L.D.; Braga-Neto, U.M. Self-adaptive physics-informed neural networks. *J. Comput. Phys.* **2023**, *474*, 111722. [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.