Journal of Materials Processing Tech. 303 (2022) 117534

Contents lists available at ScienceDirect

Journal of Materials Processing Tech.

journal homepage: www.elsevier.com/locate/jmatprotec

Full-field temperature recovery during water quenching processes via physics-informed machine learning

Ze Zhao [a], Michael Stuebner [b], Jim Lua [b], Nam Phan [c], Jinhui Yan [a,*]

[a] Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, United States
[b] Global Engineering & Materials, Inc, United States
[c] Naval Air Systems Command, United States

ARTICLE INFO

Associate Editor: Kunpeng Zhu

Keywords:
Physics-informed machine learning
Heat treatment
Quenching

ABSTRACT

Water quenching is an effective heat treatment process to produce high-quality metallic structures. Accurate and efficient prediction of the full-field temperature inside the part to capture and control the residual stresses and part quality remains a challenging task. This paper proposes a simple and easy-to-use model for full-field temperature recovery during water quenching processes, using physics-informed machine learning (ML). The novelty of the ML framework is that it only needs temperature measurements of sparse locations to efficiently/accurately recover the full spatio-temporal temperature field without invoking sophisticated multiphysics simulations. The ML framework consists of two tightly connected neural network (NN) models: (1) Firstly, a physics-informed neural network (PINN)-based surrogate model is constructed. The surrogate model, which approximates a high-fidelity finite element model, is responsible for quickly outputting the full-field temperature distribution from the parameterized thermal boundary conditions (BCs). (2) Then, another neural network is constructed to project the available experimental data onto the surrogate model and learn the optimal thermal BC from the parametric space, which produces the best full-field temperature prediction in the surrogate model. The proposed ML framework features high efficiency, accuracy, and universality for temperature prediction in quenching processes. These features are carefully demonstrated and the framework is validated using experimental measurements.

1. Introduction

Water quenching is an effective heat treatment process widely used in mechanical, aerospace, and automotive industries for rapid cooling of metallic workpieces to obtain specific mechanical propertie.

In water quenching, rapid and non-uniform cooling can generate undesired residual stress and distortion, which are detrimental to the structure's functionality and even cause structural failure. Thus, predicting the resulting residual stress and distortion of quenched parts has been one of the core research areas for water quenching processes. However, this task is challenging due to the fast transient thermal process that includes the strong coupling of thermal, metallurgical, and mechanical interactions. To tackle this challenge, it is imperative to obtain the accurate full spatio-temporal temperature field, which is significantly influenced by many physical parameters, including part/tank geometry, dipping rate, and flow condition in the quenching tank.

Many experiments have been conducted to acquire the temperature inside quenched parts during quenching processes. Xiao et al. (2010) measured the temperature evolution of an aluminum-based alloy during water quenching under different agitation and orientation conditions. Wang et al. (2011) studied the effects of multiple factors in gas quenching, such as gas velocity, pressure, temperature, gas mixture, part orientation, gas temperature, gas humidity, surface roughness, oxide, and material on the interfacial heat transfer. Mozumder et al. (2014) investigated the heat transfer characteristics of a rotating hollow steel cylinder during water quenching. These experiments provide valuable data to help us understand quenching physics and validate numerical models. However, quenching experiments can only measure temperature history on limited locations using thermocouples (TCs), and the measurement is not only sparse but also intrusive. It is still challenging to experimentally obtain an accurate and full temperature field.

Computational models are recently used in the quenching community to obtain the full temperature field. The most commonly used approaches rely on the heat transfer coefficients (HTCs) to obtain the

* Corresponding author.
E-mail address: yjh@illinois.edu (J. Yan).

https://doi.org/10.1016/j.jmatprotec.2022.117534
Received 9 October 2021; Received in revised form 2 February 2022; Accepted 13 February 2022
Available online 16 February 2022

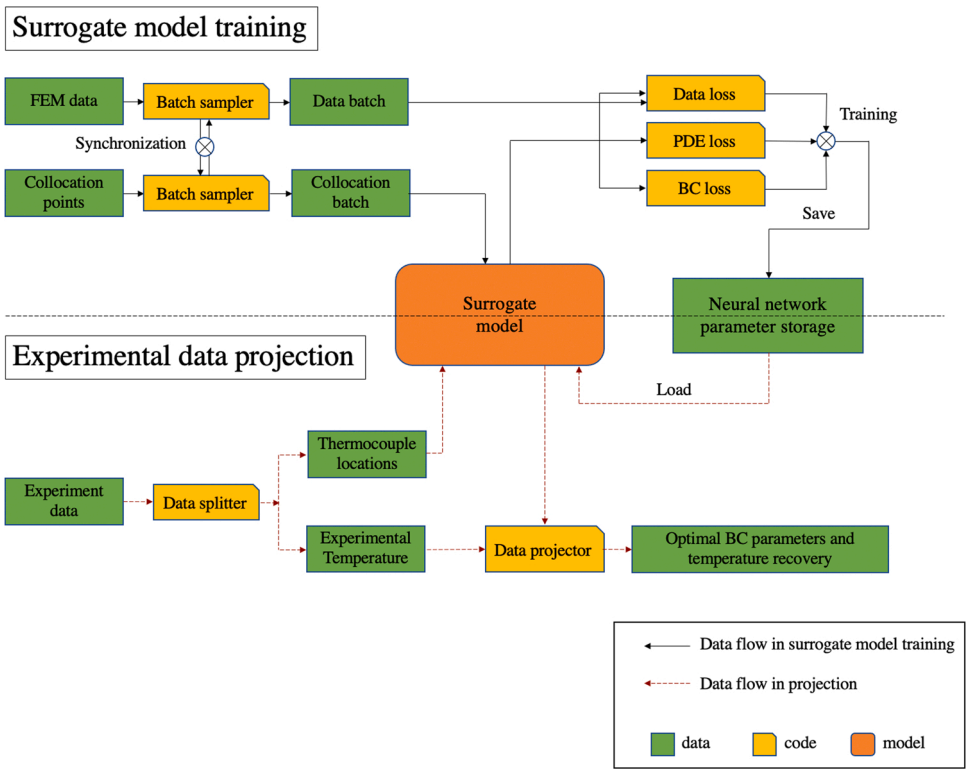

Fig. 1. Overview of the proposed ML framework.

interior temperature of the structure. This type of approach is cost-effective since the multi-physics of the quenching tank are lumped into the HTCs, and only the heat transfer inside the structure is solved. However, these models are empirical, prone to errors, and require an extensive testing matrix to calibrate the HCTs, based on measured temperature data at selected locations of the workpiece. On the other hand, computational fluid dynamics (CFD) has been attracting attention in quenching simulations in recent years. Schuettenberg et al. (2005) conducted numerical investigation of the gas quenching with jet flow field. Xiao et al. (2011) simulated air convection effects on aluminum alloy cooling. Anon (2009) studied how quenching media and structure orientation influence the quenching process. To obtain high-fidelity predictions, the CFD models explicitly resolve the multiphase flow behaviors in the quenching tank and their interaction with quenched structures, as suggested by Lua et al. (2021). The advantage of these approaches is that they can predict the temperature evolution in the structure directly from quenching parameters. However, these models require sophisticated interface-capturing methods, such as level set/-volume of fluid, to resolve the multiphase flow and phase transitions, including evaporation and condensation. The wall-bounded flows with large Reynolds number also need high mesh resolutions to capture the turbulent flows and thin boundary layers, leading to prohibitive computational cost.

Recent years have witnessed the boom of application of machine learning (ML) and data-driven models to engineering design and analysis. Liu et al. (2016) built a self-consistent clustering tool to analyze inelastic heterogeneous material properties. Balu et al. (2019) applied deep learning to bioprosthetic heart valves design. Mozaffar et al. (2019) applied deep learning to investigate path-dependent plasticity. Liu (2020) developed a data-driven material model, which enables accurate and efficient prediction of multiscale responses for heterogeneous materials with interfacial effects. Prasad et al. (2021) proposed a differentiable non-uniform rational basis spline (NURBS) module to integrate the NURBS representation of CAD models with artificial neural network (ANN). He and Tartakovsky (2021) constructed machine learning models to solve convection-diffusion systems. A number of successful applications of ML-based models to metal additive manufacturing (AM) can also be found in the literature. For example, Yan et al. (2018) proposed a data-driven framework to derive the process-structure-property relationships for metallic additive manufacturing. Paul et al. (2019) utilized an iterative machine learning approach for real-time temperature profile prediction in additive manufacturing processes. Zhu et al. (2021) developed a physics-informed neural network (PINN) to predict the temperature and melt pool fluid dynamics for selective laser beam melting processes. Xie et al. (2021) developed a mechanistic data-driven model to predict built mechanical properties of additive manufactured metals.

In light of the success of ML in other domains, this paper proposes an ML framework to accurately and efficiently recover the full spatiotemporal temperature field during water quenching processes. The ML framework consists of two tightly connected neural network (NN) models. Firstly, a physics-informed neural network Raissi et al. (2019) (PINN)-based surrogate model is constructed. The surrogate model, which approximates a high-fidelity finite element model, is responsible for quickly outputting the full-field temperature distribution from the parameterized thermal boundary conditions (BCs). Then, another NN is constructed to project the available experimental data onto the surrogate model and learn the optimal thermal BC from the parametric space, which produces the best full-field temperature prediction in the surrogate model. The ML model can be utilized in conjunction with experiments by manufacturing engineers to quickly obtain the full-field thermal history of the quenched parts to evaluate residual stress and part distortion. The novelty of the ML framework is marked by .

1. It utilizes PINN to construct a fast evaluation surrogate model, which does not rely on big training data and can efficiently output the full-field temperature for a given boundary condition. PINN provides advantages over traditional data-driven NNs that often have training difficulties in data scarcity problems. PINN encodes fundamental physical laws into the training process to relax the dependency of ML

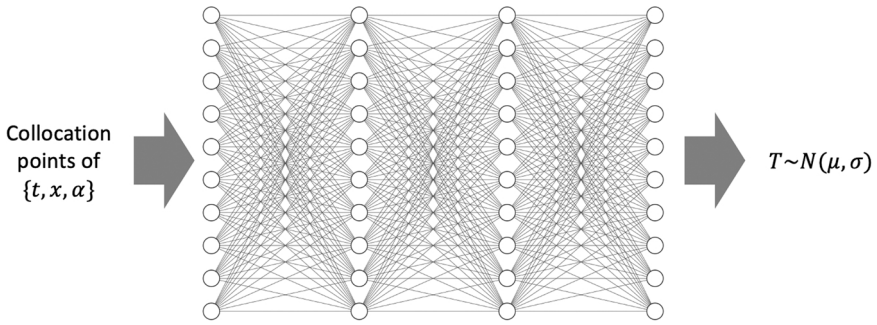

Fig. 2. PINN-based surrogate modeling utilizing a fully connected neural network.

models on data and improves the performance in sparse data regions, as many re-searchers demonstrated in the literature.
2. It can accurately recover the full-field temperature history during quenching processes using a few experimental measurements, without revoking sophisticated multiphysics simulations.

The paper is structured as follows. The formulation of the ML-based temperature recovery framework is described step by step in Section 2. The validity and accuracy of the ML framework are demonstrated in Section 3 by using the quenching experiment conducted by Worcester Polytechnic Institute (WPI). The conclusion is drawn in Section 4.

2. Machine learning based temperature recovery model

This section presents the details of the proposed ML-based full-field temperature recovery framework. The framework is formulated by integrating two physics-informed neural networks (PINN). Firstly, a surrogate model is built to map the space-time coordinates to a full temperature field for a given parameterized thermal boundary condition (BC). The surrogate model is trained by high-fidelity finite element results and informed by an energy conservation law. Then, another PINN model is utilized to explore the parametric space of the surrogate model to identify the best approximation to match measured thermal history and satisfy the energy conservation law. A flow chart of the two-step ML framework is given in Fig. 1.

2.1. Surrogate model for temperature prediction

The surrogate model approximates the behavior of a high-fidelity heat transfer model to enable fast prediction of the temperature inside the quenched part with a given boundary condition. At this stage, the model neglects the multiphase flows, phase transition (e.g., evaporation/condensation), and convective heat transfer in the quenching tank and only considers the heat transfer inside the structure, which is dominated by conduction. The finite element method (FEM) is typically used to solve the conduction-dominant heat transfer problem. However, its high computational cost makes it infeasible to be employed inside an optimization algorithm for large-scale structures. In recent years, ML-derived surrogate modeling and model reduction have been proved to be a powerful tools in engineering problems due to their well-balanced accuracy and efficiency. Nabian and Meidani (2019) applied surrogate modeling to high dimensional random differential equations. Sun et al. (2020) proposed a physical-constrained data-free deep learning model to predict flow fields. Kaneko et al. (2021) accelerated modeling of thermal cycling-induced plastic deformation using a hyper-reduction method. In this paper, PINN is utilized to construct the surrogate model, trained by high-fidelity data, informed by the energy conservation law, and agnostic to the surrounding quenching fluid, to link the full-field temperature with a given parameterized heat flux.

2.1.1. Thermal boundary condition parameterization

The parameterized heat flux q taken by the surrogate model is evaluated as

$$q(T,t) = h(T,t)(T - T_\infty) \tag{1}$$

where T is the temperature, T_∞ is the ambient temperature, h is a heat transfer coefficient (HTC), which is a function of temperature and time. Here, $h(T, t)$ is parameterized as follows. The structural surface is first decomposed into several sub-surfaces. The choice of surface decomposition depends on the geometric property of the structure. For bulk type structures, the structural surface can be decomposed into top (1), side (2), and bottom (3) sub-surfaces, which is what this paper adopts. Then, each sub-surface's HTC $H^{(i)}$ ($i = 1, \ldots, N_{face}$, where N_{face} is the number of sub-surfaces) is expressed by a linear combination of several basis functions of HTC curves, namely,

$$H^{(i)} = \sum_k^{N_k} \alpha_k^{(i)} N_k^{(i)}(T) \tag{2}$$

where $\alpha_k^{(i)}$ are the interpolation coefficients, $N_k^{(i)}$ are presumed piecewise linear basis functions of HTCs defined over several temperature intervals. This paper employs 18 temperature intervals to parameterize the HTCs, with 12 for the vapor blanket stage, 4 for the nucleate boiling stage, and 2 for the convective cooling stage. The choice of the basis functions of HTCs should ensure the positivity of HTC and need to roughly represent the characteristics of vapor blanket, nucleate boiling, and convective cooling stages during a quenching process Holman (2010). The orientation of structural surfaces can impact the heat transfer significantly. To capture this effect, instead of only using H, H is further multiplied with an orientation factor to define h as

$$h(T,t) = \sum_{i=1}^{3} O^{(i)} H^{(i)} \tag{3}$$

where the orientation factor O is defined as

$$O^{(1)}(n_z) = \frac{1}{2}(n_z + 1)n_z \tag{4}$$

$$O^{(2)}(n_z) = (1 - n_z^2) \tag{5}$$

$$O^{(3)}(n_z) = \frac{1}{2}(n_z - 1)n_z \tag{6}$$

where n_z is the z-component (gravitational direction) of the unit outward normal vector $\boldsymbol{n}$ on the structural surface.

2.1.2. High-fidelity finite element model

A high-fidelity heat transfer model based on FEM is formulated to generate necessary data to train and validate the surrogate model. The semi-discrete formulation of the finite element model is given as follows. Let Ω and Γ denote the structural domain and its boundary, $\mathbb{V}$ and $\mathbb{W}$ denote the trial and testing function spaces. With the above definitions, the FEM model finds $T \in \mathbb{V}$ such that, for $\forall w \in \mathbb{W}$,

$$\int_{\Omega} w\rho c_p \frac{\partial T}{\partial t} d\Omega + \int_{\Omega} \kappa \nabla w \cdot \nabla T d\Omega - \int_{\Gamma} wq(T,t) d\Gamma = 0 \tag{7}$$

where w is the testing function, t is the time, ρ, c_p, and κ are the density, heat capacity, and conductivity, respectively. The FEM model is solved by using linear tetrahedral elements for spatial discretization. Temporally, time integration is done by a second-order accurate and unconditionally stable scheme based on generalized-α Chung and Hulbert (1993) and Newmark-β methods Newmark (1959). A predictor-corrector iterative scheme based on the Newton-Raphson method is utilized to solve Eq. 7. A multigrid-preconditioned generalized minimal residual method (GMRES) Saad and Schultz (1986) is used to find the root of the linear system resulting from Eq. 7. Thermal BCs are sampled from the parametric space of $\boldsymbol{\alpha}$ via Latin hypercube sampling proposed by Mckay et al. (1979) and imported into the last boundary integral term of Eq. 7 to generate data to train and validate the surrogate model.

2.1.3. Physics-informed neural network for surrogate modeling

The ML-based surrogate model $\mathscr{M}$, shown in Fig. 2, utilizes artificial neural network (NN) to enable fast temperature prediction from a given set of spatiotemporal coordinate ($\boldsymbol{x}$, t) and thermal BC parameters $\boldsymbol{\alpha}$. Because only a few FEM simulations can be afforded due to the high cost, using off-the-shelf big data-based ML algorithms to train the surrogate model is unfeasible. Thus, the surrogate model is built by using a physics-informed neural network (PINN), which encodes an energy conservation law to facilitate the training process. PINN, as popular scientific machine learning (SciML) model, has been utilized in several engineering problems to overcome the challenge of lacking big data owing to the additional knowledge embedded in mechanistic/physical laws Raissi et al. (2019).

In this paper, the surrogate model treats the temperature field during water quenching as a Gaussian process, namely,

$$T \sim N(\mu, \sigma) \tag{8}$$

where μ and σ are the mean and standard deviation of the temperature field, given as

$$\mu = \mu(t, \boldsymbol{x}, \boldsymbol{\alpha}) \tag{9}$$

$$\sigma = \sigma(t, \boldsymbol{x}, \boldsymbol{\alpha}) \tag{10}$$

The PINN-based surrogate model $\mathscr{M}$ takes spatiotemporal coordinates and thermal BC parameters as inputs, and outputs mean temperature μ^M and its standard deviations σ^M, namely,

$$\mu^M, \sigma^M = \mathscr{M}(t, \boldsymbol{x}, \boldsymbol{\alpha}; \boldsymbol{W}^*, \boldsymbol{b}^*) \tag{11}$$

where $\boldsymbol{W}^*$ and $\boldsymbol{b}^*$ are the NN's weights and biases, which are learned by the following optimization problem

$$\boldsymbol{W}^*, \boldsymbol{b}^* = \underset{\boldsymbol{W},\boldsymbol{b}}{\operatorname{argmin}} L(\boldsymbol{W}, \boldsymbol{b}) \tag{12}$$

where L is the loss function, which consists of the following three components

$$L(\boldsymbol{W}, \boldsymbol{b}) = \lambda_{DAT} L_{DAT}(\boldsymbol{W}, \boldsymbol{b}) + \lambda_{PDE} L_{PDE}(\boldsymbol{W}, \boldsymbol{b}) + \lambda_{BC} L_{BC}(\boldsymbol{W}, \boldsymbol{b}) \tag{13}$$

where L_{DAT}, L_{PDE}, and L_{BC} represent the data-driven component, physics-informed component and associated boundary conditions. λ_{DAT}, λ_{BC}, and λ_{PDE} are the corresponding weights for each component. The data-driven component, which regularizes the training in the manner of maximum likelihood estimation, is defined as

$$L_{DAT}(\boldsymbol{W}, \boldsymbol{b}) = -\sum_{\boldsymbol{P}} ln(f(T_i^s; \mu(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i), \sigma(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i))) \tag{14}$$

where $\boldsymbol{P}$ is the set of labeled training data points obtained from high-fidelity FEM simulations. f is a possibility distribution function (PDF) of Gaussian profile, given as

$$f(u; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(u-\mu)^2}{2\sigma^2}) \tag{15}$$

Conventional off-the-shelf ML models only employ L_{DAT} in the training process, which relies on big data. Given the limited data available, the physics-informed components L_{PDE} and L_{BC} are added into the loss function to alleviate the dependence on big-data. The physics are encoded by penalizing the residual of the PDE of the physical principles over a set of collocation points in the interior (denoted by $\boldsymbol{S}$) and a set of collocation points on the boundary (denoted by $\partial\boldsymbol{S}$), namely,

$$L_{PDE}(\boldsymbol{W}, \boldsymbol{b}) = \sum_{S} \left| l_{PDE}(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i) \right|^2 \tag{16}$$

$$L_{BC}(\boldsymbol{W}, \boldsymbol{b}) = \sum_{\partial S} \left| l_{BC}(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i) \right|^2 \tag{17}$$

where l_{PDE} and l_{BC} are the point-wise residuals of energy conservation law and heat flux compatibility, defined as

$$l_{PDE}(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i) = \rho c_p \frac{\partial}{\partial t} \mu(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i) - \kappa \nabla^2 \mu(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i) \tag{18}$$

$$l_{BC}(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i) = \boldsymbol{n} \cdot \kappa \nabla \mu(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i) - h(\mu(t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i) - T_\infty) \tag{19}$$

where $\{t_i, \boldsymbol{x}_i, \boldsymbol{\alpha}_i\}$ is the space-time coordinate of the collocation point. Once $\boldsymbol{W}^*$ and $\boldsymbol{b}^*$ are determined, the output of $\mathscr{M}$ can be easily achieved by a feed-forward evaluation of the NN with high efficiency, since it merely costs a few matrix multiplications and vector additions.

2.2. Full-field temperature recovery model

With the surrogate model trained, another neural network model is built to learn the best approximation in the parametric space of thermal BCs to recover the full spatiotemporal temperature field. Here, the NN model is stated as the following minimization problem with respect to the thermal BC parameters $\boldsymbol{\alpha}$.

$$\boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} L^{proj}(\boldsymbol{\alpha}) \tag{20}$$

where the loss function L^{proj} is defined as

$$L^{proj}(\boldsymbol{\alpha}) = -\sum_{E} ln(f(T_j^e; \mu_j^M, \sigma_j^M)) \tag{21}$$

where $\boldsymbol{E}$ represents the set of the thermocouples locations and T_j^e is the measured temperature. μ_j^M and σ_j^M are given

$$\mu_j^M = \mu(t_j, \boldsymbol{x}_j, \boldsymbol{\alpha}; \boldsymbol{W}^*, \boldsymbol{b}^*) \tag{22}$$

$$\sigma_j^M = \sigma(t_j, \boldsymbol{x}_j, \boldsymbol{\alpha}; \boldsymbol{W}^*, \boldsymbol{b}^*) \tag{23}$$

The minimization of the loss function L^{proj} leads to the optimal thermal BC, $\boldsymbol{\alpha}^*$, which maximizes the possibility of matching measured results while obtaining the encoded energy conservation law. Once $\boldsymbol{\alpha}^*$ is identified, it will be substituted into the PINN-based surrogate model $\mathscr{M}$ to produce a full-field temperature prediction and its standard deviation, namely,

$$\mu^*(t, \boldsymbol{x}) = \mu(t, \boldsymbol{x}, \boldsymbol{\alpha}^*; \boldsymbol{W}^*, \boldsymbol{b}^*) \tag{24}$$

$$\sigma^*(t, \boldsymbol{x}) = \sigma(t, \boldsymbol{x}, \boldsymbol{\alpha}^*; \boldsymbol{W}^*, \boldsymbol{b}^*) \tag{25}$$

3. Application and results

This section deploys the proposed ML-based temperature recovery

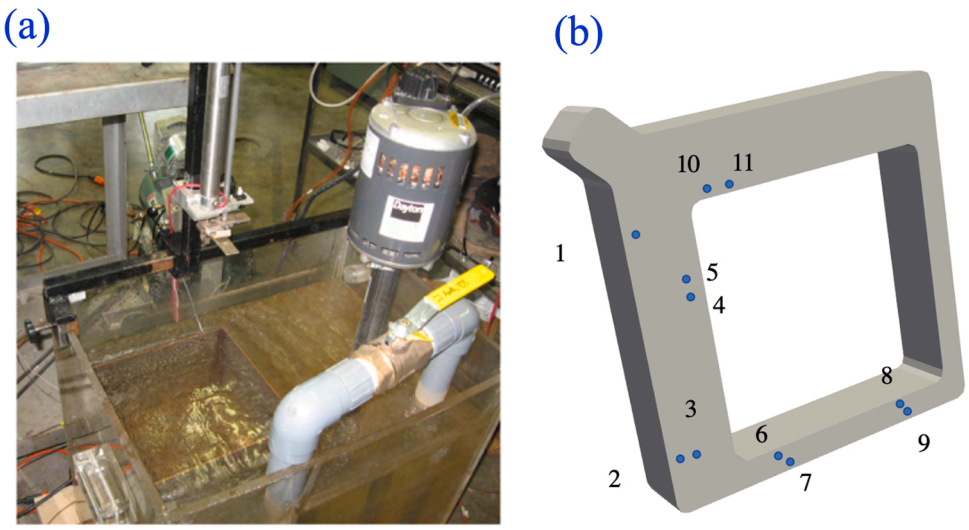

Fig. 3. Quenching experiment conducted in Xiao et al. (2010). (a): Experimental apparatus. (b): CAD model of the quenched structure, locations and numbering of the installed thermocouples (TCs).

Table 1
The properties of the quenched structure. The unit of temperature T is ℃.

Properties	Value
ρ, kg · m^{-3}	2700.0
c_p, J/(kg ·℃)	$930.334 - 0.442T + 0.003T^2 - 2.210 \times 10^{-6}\ T^3$
κ, W/(m ·℃)	$230.160 - 0.121T + 7.466 \times 10^{-4}T^2 - 2.643 \times 10^{-6}T^3 + 2.298 \times 10^{-9}T^4$

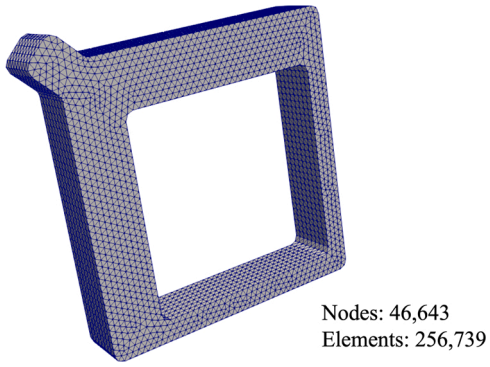

Fig. 4. The tetrahedral mesh employed in the FEM simulations.

framework to a quenching experiment conducted by Worcester Polytechnic Institute (WPI). The validity and accuracy of the proposed framework are carefully assessed.

3.1. Experimental setup and validation data

The experiment apparatus, conducted by Xiao et al. (2010) from WPI, is shown in Fig. 3(a). The experiment employs a cubic water quenching tank to quench a metallic structure based on AL-319 alloy. The structure is heated in a furnace up to 500 ˚C and then immersed into the quenching tank with the thin leg down orientation. The initial temperature of the quenching tank is 75 ˚C. The thermal properties of the quenched structure are based on Al-319 and determined by ThermoCalc. The values of these properties are listed in Table 1. The computer-aided design (CAD) model of the structure is shown in Fig. 3 (b), which also depicts the locations and numbering of installed thermocouples for temperature measurements.

3.1.1. Accuracy of surrogate modeling

The PINN-based surrogate model's accuracy of approximating high-fidelity simulations is first evaluated. FEM simulations using the model defined in Eq. 7 are performed to generate data to derive the PINN-based surrogate model. The simulations are performed using an in-house FEM-based heat transfer code, which has been validated over a set of metal AM problems, including directed energy deposition Zhao et al. (2021) and selective laser beam melting Zhu et al. (2021). Time step Δt is set to 1×10^{-2} s in the FEM simulations. A mesh refinement study has been performed to make sure the simulated results are resolution-insensitive. The final selected FEM mesh consists of 46,643 nodes and 256,739 elements. A snapshot of the mesh is shown in Fig. 4. Thermal BCs are sampled from the parametric space via Latin hypercube sampling Mckay et al. (1979) and imported into the FEM simulations to generate data to train and validate the surrogate model. In total, 500 FEM simulations are performed with different $\boldsymbol{\alpha}$, sampled from the space of [0,10] × [0,10] × [0,10]. 450 of the 500 FEM simulations serve as training data, and others are used to evaluate the surrogate model's accuracy. The FEM simulations are performed with 48 cores using Intel Xeon Platinum 8160 on Stampede2 in Texas Advanced Computing Center (TACC).

The PINN-based surrogate model consists of 4 fully connected hidden layers, each containing 80 neurons. Spatially, the number of collocation points is the same as the number of FEM nodes, and 19.8% of them are on boundaries. The model is trained by minimizing the loss function defined in Eq. 13 with respect to $\boldsymbol{W}$ and $\boldsymbol{b}$. The minimization is executed by the following procedures: (1) The spatiotemporal coordinates of collocation points and training data are substituted into Eq. 13. (2) Take the derivatives of the loss function with respect to $\boldsymbol{W}$ and $\boldsymbol{b}$. (3) Update $\boldsymbol{W}$ and $\boldsymbol{b}$ by a gradient descent method. Most of the current ML models solve the optimization problem by a stochastic gradient descent (SGD) algorithm Ruder (2016). SGD only uses a subset of collocation points, randomly sampled from the input space at each iteration, to calculate the directional gradient. Research shows that SGD works very well to skip bad local minima. One problem with SGD is the oscillation of gradient direction caused by the random selection of sampled collocation points. In this paper, the Adam method combining adaptive learning rate and momentum methods is used to improve convergence speed, as suggested by Kingma and Ba (2014). The initial learning rate is 10^{-3}, the stochastic sampling batch size is set to 20,000, and the training process costs 10,000 epochs. The proposed surrogate model is implemented with PyTorch (see Paszke et al., 2019).

To evaluate the accuracy of the trained model, we define the absolute error on a data set $\boldsymbol{B}$ as

Table 2
The error on training and test sets.

	Absolute error	Relative error
Training set $e(\boldsymbol{P})$	10.77 ˚C	2.53%
Test set $e(\boldsymbol{Q})$	17.45 ˚C	4.11%

$$e(\boldsymbol{B}) = \left(\frac{\sum_{\boldsymbol{B}} |T_i^s - \sigma(t_i, \boldsymbol{x}, \boldsymbol{\alpha})|^2}{N(\boldsymbol{B})}\right)^{\frac{1}{2}} \tag{26}$$

where $N(\boldsymbol{B})$ is the number of the elements in the set $\boldsymbol{B}$. The relative error is scaled by $T_{ref} = 425^{\circ}C$, which is the difference between initial and ambient temperature. The errors on training set $\boldsymbol{P}$ and test set $\boldsymbol{Q}$ as given in the following Table 2.

Additionally, two representative results selected among $\boldsymbol{Q}$ are presented in Fig. 5 and 6, which includes the mean full temperature field and the corresponding standard variation predicted by the trained surrogate model at several time instances. Fig. 5 shows the case with smallest error ($\alpha = [0.701,\ 1.043,\ 0.342]^T$ and the absolute error is

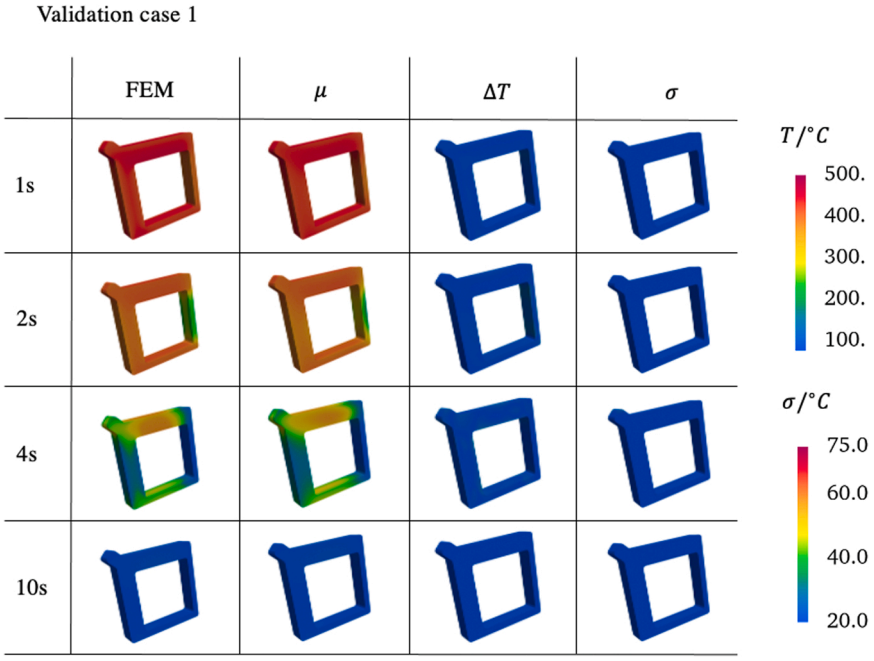

Fig. 5. Surrogate model V.S. FEM results: Case 1.

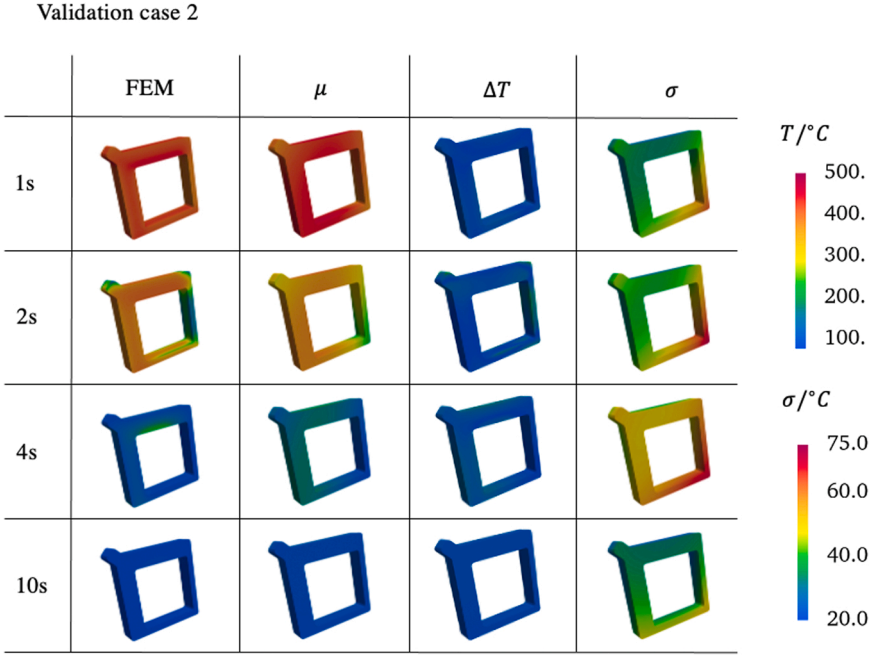

Fig. 6. Surrogate model V.S. FEM results: Case 2.

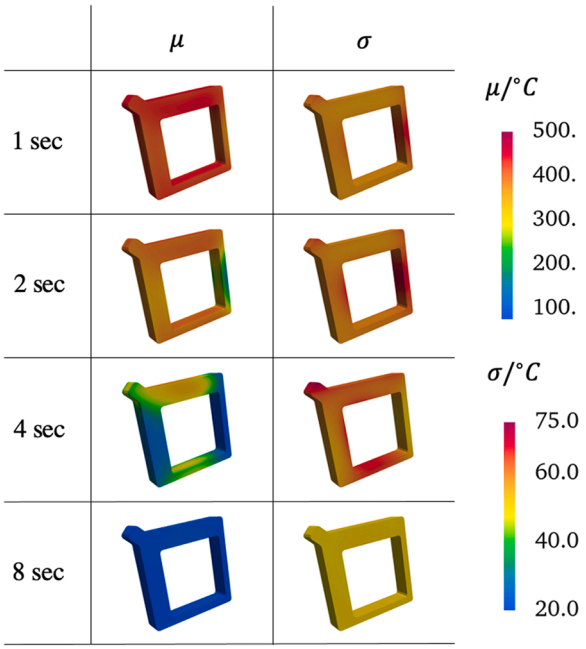

Fig. 7. Recovered full temperature field and standard deviation at several time instances.

10.99), and Fig. 6 depicts the case with largest error ($\alpha = [1.620, 0.471, 0.404]^T$ and the absolute error is 24.71). The FEM results and the temperature difference between the prediction of the FEM and the surrogate models are also shown in Fig. 5 and Fig. 6. Noticeable discrepancy is seen on thin legs. This is because a uniform distribution of collocation points is used. Therefore, there are fewer collocation points on the thin legs than the thick legs, leading to bigger overweight for the thick legs. Nonetheless, the surrogate model maintains a good overall approximation of the high-fidelity FEM simulated results achieved. Besides, one should note that the surrogate model attains a superior efficiency compared with the FEM model because it only involves a few matrix multiplications and vector additions after training.

3.1.2. *Temperature recovery and validation with experimental data*

A second layer NN model is utilized to optimize the surrogate model in the parametric space to recover the full spatiotemporal temperature field using the measured data. Among the eleven thermocouples (see Fig. 3), only seven thermocouples are used in the training process. The full temperature distribution recovered by the ML model at several time instances is shown in Fig. 7. To validate the accuracy of the recovered full temperature field, the comparison of measured temperature history at the other four thermocouples in the experiment with the ML model's recovery is shown in Fig. 8. As mentioned in the Introduction section, the quenching process has three stages: vapor blanket, nucleate boiling, and convective cooling. The ML model captures these stages with reasonable accuracy. In the vapor blanket stage, the cooling rate is relatively low because the structure is isolated from the water by the vapor blanket. In the ML model, the vapor layer stage happens when the temperature is higher than 300 ˚C. Fig. 8 shows that the temperature gradient of this stage is smaller than that of the nucleate boiling stage. In the nucleate boiling stage, the structure temperature is still higher than the boiling point yet not high enough to keep an intact vapor blanket, which means the water can directly contact the structure and take away heat. So, the cooling rate increases consequently. As seen in Fig. 8, the curves have larger slopes in the interval between 100 and 300 ˚C. When the temperature goes below 100 ˚C, it enters the convective cooling stage. Since the temperature is lower than the saturation temperature, boiling is no longer dominant, and the cooling rate decreases dramatically until the temperature reaches the ambient temperature. The predictive accuracy of TC2 is inferior compared with the predictions of other TCs. This discrepancy may be partially due to the parameterization of experimental heat flux, which does not fully represent the multiphysics in the quenching tank. The HTC parameterization used in the paper is a reduced-order representation of the real heat flux on the structure surface since only three HTC on the top, bottom, side surfaces

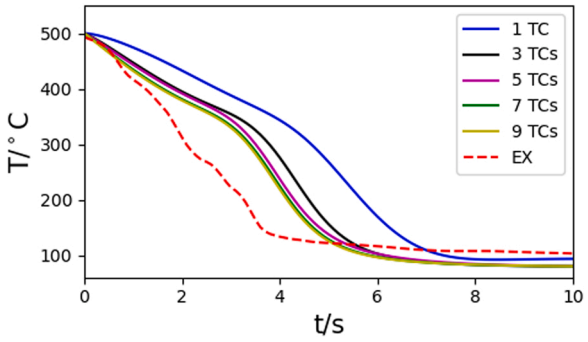

Fig. 9. Prediction of temperature of TC2 using different numbers of TCs in the training.

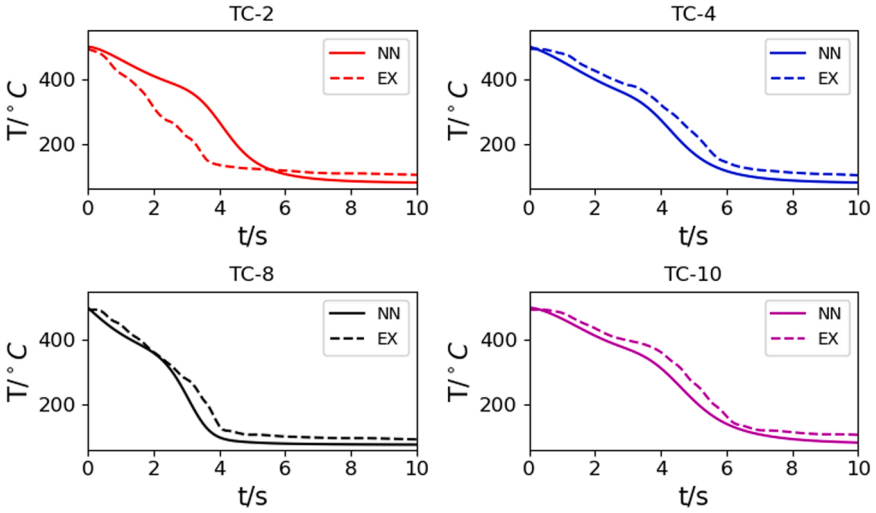

Fig. 8. Recovered temperature history compared with experimental measurements.

are measured. However, our numerical experiments still show the model's capability to converge to the measured temperature history. Fig. 9 presents the surrogate model prediction of TC2 using various numbers of TCs in training. When the number of TCs increases, the predictive accuracy is enhanced. Although the recovered predicted temperature couldn't fully agree with the experimental data since only 11 TCs are available, the trend implies that the accuracy of the surrogate model increases with more experimental data. Overall, the ML framework successfully captures the trend of cooling rate evolution. Good agreement between ML recovery and experimental measurements is obtained, despite only a few measured temperature data being used to train the ML model. The accuracy is primarily due to the additional knowledge from the physics-informed component in the ML model and the implicit effects of the surrounding quenching flow hidden in the measured temperate data, although the ML model itself is agnostic to the surrounding fluid.

4. Conclusions

This paper presented a physics-informed machine learning framework for recovering the full spatiotemporal temperature field inside quenched parts during water quenching processes. The paper was motivated to overcome the challenges of current modeling techniques pertinent to accuracy and efficiency. The key features of the proposed ML model are (1) it does not need to invoke complicated multiphysics simulation and (2) it can significantly relax ML models' dependence on big data owing to additional encoded physical principles. The proposed model is applied to a realistic water quenching process. The results showed that the model could efficiently and accurately obtain the full temperature field with only a small amount of labeled experimental data. The model can be extended to other heat treatment processing techniques and used in conjunction with experimental measurements to enable fast and efficient temperature prediction and monitoring in heat treatment processes.

CRediT authorship contribution statement

Ze Zhao: Conceptualization, Coding, Writing, and Simulations. **Michael Stuebner:** Writing and Simulations. **Jim Lua:** Writing and Editing. **Nam Phan:** Writing and Editing. **Jinhui Yan:** Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is funded by the Naval Air Warfare Center, Aircraft Division under the Contract of N68335-21-C-0057.

References

Balu, A., Nallagonda, S., Xu, F., Krishnamurthy, A., Hsu, M., Sarkar, S., 2019. A deep learning framework for design and analysis of surgical bioprosthetic heart valves. Sci. Rep. 9, 1–12.

Chung, J., Hulbert, G.M., 1993. A time integration algorithm for structural dynamics withimproved numerical dissipation: the generalized-α method. J. Appl. Mech. 60, 371–375.

He, Q., Tartakovsky, A., 2021. Physics-informed neural network method for forward andbackward advection-dispersion equations. Water Resour. Res. 57 e2020WR029479.

Holman, J.P., 2010. Heat Transfer. McGraw-Hill, Boston, MA.

AnonCFD investigation of quench media and orientation effects on structural stress induced in the intense quenching processes for aluminum cylinder heads 2009.

Kaneko, S., Wei, H., He, Q., Chen, J., Yoshimura, S., 2021. A hyper-reduction computational method for accelerated modeling of thermal cycling-induced plastic deformations. J. Mech. Phys. Solids 151, 104385.

Kingma, D., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Liu, Z., 2020. Deep material network with cohesive layers: multi-stage training and interfacial failure analysis. Comput. Methods Appl. Mech. Eng. 363, 112913.

Liu, Z., Bessa, M., Liu, W., 2016. Self-consistent clustering analysis: an efficient multi-scale scheme for inelastic heterogeneous materials. Comput. Methods Appl. Mech. Eng. 306, 319–341.

Lua, J., Yan, J., Li, P., Zhao, Z., Karuppiah, A., Stuebner, M., 2021. Novel Multi-physics-based Modeling of A Quenching Process with Thermal-metallurgical-mechanical Interactions in Aluminum Components. 77th Annual Vertical Flight Society Forum and TechnologyDisplay: The Future of Vertical Flight, FORUM 2021. Vertical Flight Society, Virtual, Online, p. 2021.

Mckay, M., Beckman, R., Conover, W., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21, 239–245.

Mozaffar, M., Bostanabad, R., Chen, W., Ehmann, K., Cao, J., Bessa, M., 2019. Deep learning predicts path-dependent plasticity. Proc. Natl. Acad. Sci. 116, 26414–26420.

Mozumder, A., Mitsutake, Y., Monde, M., 2014. Subcooled water jet quenching phenomena for a high temperature rotating cylinder. Int. J. Heat Mass Transf. 68, 466–478.

Nabian, M.A., Meidani, H., 2019. A deep learning solution approach for high-dimensional random differential equations. Probab. Eng. Mech. 57, 14–25.

Newmark, N., 1959. A method of computation for structural dynamics. J. Eng. Mech. Div. 85, 67–94.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., 2019. Pytorch: an imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. 8024–8035.

Paul, A., Mozaffar, M., Yang, Z., Liao, W., Choudhary, A., Cao, J., Agrawal, A., 2019.A real-time iterative machine learning approach for temperature profile prediction in additive manufacturing processes, in:2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA),IEEE.541–550.

Prasad, A., Balu, A., Shah, H., Sarkar, S., Krishnamurthy, A., 2021.Nurbs-diff: A differentiable nurbs layer for machine learning CAD applications.arXiv:2104.14547.

Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707.

Ruder, S., 2016.An overview of gradient descent optimization algorithms.arXiv: 1609.04747.

Saad, Y., Schultz, M., 1986. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Stat. Comput. 7, 856–869.

Schuettenberg, S., Frerichs, F., Hunkel, M., Fritsching, U., 2005. Process technology for distortion compensation by means of gas quenching in flexible jet fields. Int. J. Mater. Prod. Technol. 24, 259.

Sun, L., Gao, H., Pan, S., Wang, J., 2020. Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data. Comput. Methods Appl. Mech. Eng. 361, 112732.

Wang, Q., Chang, C., Zhang, Q., Paluch, D., 2011. Modeling of residual stresses in quenched cast aluminum components. SAE Int. J. Mater. Manuf. 4, 844–852.

Xiao, B., Wang, G., Wang, Q., Maniruzzaman, M., Sisson, R., Rong, Y., 2011. An experimental study of heat transfer during forced air convection. J. Mater. Eng. Perform. 20, 1264–1270.

Xiao, B., Wang, Q., Jadhav, P., Li, K., 2010. An experimental study of heat transfer in aluminum castings during water quenching. J. Mater. Process. Technol. 210, 2023–2028.

Xie, X., Bennett, J., Saha, S., Lu, Y., Cao, J., Liu, W., Gan, Z., 2021. Mechanistic data-driven prediction of as-built mechanical properties in metal additive manufacturing. Comput. Mater. 7, 1–12.

Yan, W., Lin, S., Kafka, O., Lian, Y., Yu, C., Liu, Z., Yan, J., Wolff, S., Wu, H., Ndip-Agbor, E., Mozaffar, M., Ehmann, K., Cao, J., Wagner, G., Liu, W., 2018. Data-driven multi-scale multi-physics models to derive process-structure-property relationships for additive manufacturing. Comput. Mech. 61.

Zhao, Z., Zhu, Q., Yan, J., 2021. A thermal multi-phase flow model for directed energy deposition processes via a moving signed distance function. Comput. Methods Appl. Mech. Eng. 373, 113518.

Zhu, Q., Liu, Z., Yan, J., 2021. Machine learning for metal additive manufacturing: predicting temperature and melt pool fluid dynamics using physics-informed neural networks. Comput. Mech. 67.