Energy and AI 20 (2025) 100482

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.sciencedirect.com/journal/energy-and-ai

Physics-informed machine learning for enhanced prediction of condensation heat transfer

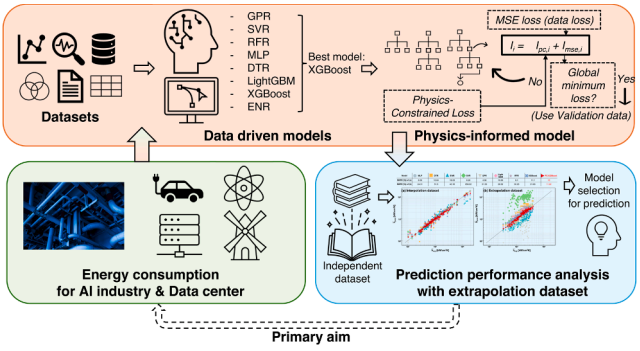
Haeun Lee [a,b], Cheonkyu Lee [c], Hyoungsoon Lee [b,*]

[a] Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, USA
[b] School of Mechanical Engineering, Chung-Ang University, Seoul 06974, South Korea
[c] Industrial Energy R&D Department, Research Institute of Sustainable Development Technology, Korea Institute of Industrial Technology (KITECH), South Korea

HIGHLIGHTS

- New model introduces a physics-informed ML approach for predicting condensation HTCs.
- Improved performance of PC-XGBoost shows better extrapolation than data-driven ML methods.
- Comprehensive evaluation assesses eight data-driven ML methods with a physics-constrained approach.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Physics-constrained
Deep learning
Heat transfer
Condensation
Nusselt model
XGBoost

ABSTRACT

Developing a universal model for predicting condensation heat transfer coefficients remains challenging, particularly for steam–non-condensable gas mixtures, owing to the intricate nonlinear interactions between multiphase flow, heat, and mass transfer phenomena. Data-driven machine learning (ML) shows promise in efficiently and accurately predicting condensation heat transfer coefficients. Research has employed various ML methods—multilayer perceptron neural networks, convolutional-neural-network–based DenseNet, back-propagation neural networks, etc.—to investigate steam condensation with non-condensable gases. However, these exhibit limited extrapolation ability and heavily rely on data quantity owing to their black-box nature. This study proposes a physics-informed ML model that combines physical constraints derived from the modified Nusselt model with conventional data-driven ML techniques. The model's predictive performance is evaluated using a comprehensive database (879 datapoints from 13 studies). A physics-constrained and eight data-driven ML methods are assessed. The results reveal that the physics-constrained approach combined with XGBoost significantly outperforms conventional ML methods on extrapolation datasets (199 datapoints from 3 studies), achieving a mean absolute percentage error of 11.22 %, which is approximately half that of the best-performing fully data-driven model at 21.63 %. The model demonstrates consistent and reliable performance across diverse datasets, making it an effective tool for predicting heat transfer coefficients in steam–non-condensable gas

* Corresponding author. School of Mechanical Engineering, Chung-Ang University, Seoul, 06974, South Korea.
E-mail address: leeh@cau.ac.kr (H. Lee).

https://doi.org/10.1016/j.egyai.2025.100482

Available online 8 February 2025

mixtures. By deepening the understanding of the underlying physical processes, the proposed model supports the development of precise and efficient engineering solutions for condensation heat transfer.

Nomenclature

C_p	specific heat [J/kg·K]
D_h	hydraulic diameter [mm]
F	degradation factor [-]
h	heat transfer coefficient [W/(m^2·K)]
h_{fg}	enthalpy [J]
I	error [-]
k	thermal conductivity [W/m·K]
L	length [mm]
M	molecular weight [g/mole]
$\dot{m}$	mass flow rate [kg/s]
T_m	mean temperature [K]
N	number of data
Nu	Nusselt number [-]
P_{tot}	system pressure [kPa]
q''	heat flux [W/m^2]
Re	Reynolds number [-]
R''	total thermal resistance [K•cm^2/W]
T_{sat}	saturation temperature [K]
T_w	wall temperature [K]
u	fluid velocity [m/s]
W	mass fraction [-]
X	mole fraction [-]
ΔP	pressure drop [Pa]
ΔT_{sub}	subcooling temperature [K]

Greek symbols

β	expansion coefficient [K^{-1}]
δ	film thickness [m]
Γ	mass flow rate per unit film width [kg/m·s]
λ	adjustment factor [-]
μ	dynamic viscosity [Pa·s]
ν	kinematic viscosity [m^2/s]
ρ	density [kg/m^3]
σ	surface tension [N/m]
τ^*	dimensionless shear stress [-]
ψ	node [-]

Superscripts

−	average component

Subscripts

deg	degradation model
*	dimensionless
exp	experiment
f	fluid
film	film
g	vapor
h	heater
mse	mean squared error
m	mixture
nc	non-condensable gas
Nu	Nusselt model
pc	physics-constrained
pred	prediction
tot	total
v	steam vapor

1. Introduction

Condensation heat transfer is a crucial process for achieving efficient heat transfer; in this process, substantial latent heat is leveraged for a wide range of energy applications, including energy conversion, [1] seawater desalination, [2] water harvesting systems, [3,4] thermal management systems, [5] environmental control, [6] and nuclear power generation. [7] However, predicting heat transfer becomes increasingly complex with the presence of non-condensable gases (NCGs) owing to the intricate nonlinear interactions between multiphase flow, heat, and mass transfer phenomena. Despite these challenges, accurately predicting steam condensation in the presence of NCGs remains essential for practical engineering applications, as NCGs significantly reduce the condensation rate, posing difficulties in scenarios where their presence is unavoidable.

Numerous experimental studies have been conducted to access condensation heat transfer in the presence of NCGs, as numerical models frequently demonstrate limited predictive accuracy. However, these experimental studies are often cost-prohibitive, largely owing to the complex interplay of geometric and operational parameters, such as subcooling (ΔT_{sub}), system pressure (P_t), NCG fraction (W_{nc}), tube length (L), and diameter (D). [8]

The theoretical foundation for filmwise condensation (FWC) in the presence of NCG has been achieved by many researchers. Although Nusselt's pioneering work laid the initial foundation for understanding condensation phenomena, his model was limited in scope, as it was applicable only to laminar flow scenarios involving pure steam condensation. [9] Recognizing these limitations, subsequent researchers expanded on Nusselt's framework to incorporate the effects of NCGs, which significantly alter the condensation process. They developed advanced models using approaches such as heat and mass transfer analogy (HMTA) and degradation factor methods. These approaches have been crucial in capturing the complex physical mechanisms introduced by NCGs, broadening the applicability and improving the predictive accuracy of condensation models. HMTA-based models, in particular, provide more comprehensive insights and yield more reliable predictions than traditional empirical correlations. Dehbi's work [10] is notable in this regard; he used HMTA and curvature correction with the Popiel model to derive a general model for the condensation heat transfer coefficient (HTC) of vapor–air mixtures over vertical walls. The suction effect was considered using Bird et al.'s model, [11] and subsequent improvements incorporated a new curvature enhancement model based on parametric numerical studies. Other researchers, such as Bae et al., [12] have modified these models using experimental data, and Wang et al. [13] investigated the effects of gas compressibility, thermal resistance of the condensate film, and convection heat transfer on modeling steam condensation using HMTA. Nevertheless, the inherent nature of heat and mass analogy, which predicts condensation heat transfer by linking momentum, mass, and heat transfer, necessitates the use of experimentally derived correlations for mass transfer. This reliance on empirical data highlights a limitation in that theoretical models cannot yet fully replace traditional empirical condensation HTC correlations. Regarding the degradation factor model, Kuhn et al. [14] developed a degradation factor correlation based on the Reynolds number of film thickness and NCG mass fraction, including precise measurements of local heat flux and extensive FWC data for NCG

mixtures. Although the degradation factor method demonstrated satisfactory engineering accuracy, its applicability was restricted to specific experimental conditions. Accordingly, despite extensive efforts, developing a universal model for predicting condensation HTC in NCG mixtures remains a significant challenge owing to the complex, nonlinear nature of multiphase flow, heat, and mass transfer processes. In addition, conventional empirical correlations are often limited to the datasets from which they were derived and involve intricate mathematical expressions that are difficult to generalize across a wide range of geometrical and operational conditions.

Data-driven machine learning (ML) has recently emerged as a promising tool for addressing these challenges; it can predict condensation HTC with high efficiency and accuracy. The ML approach has been actively used in the broader field of heat transfer, facilitating advancements in various applications. [15–21] In particular, to investigate steam condensation with NCG, several studies have utilized various ML methods, including multilayer perceptron (MLP) neural networks, [22] convolutional-neural-network–based DenseNet, [23] and back-propagation neural networks. [24] Although these data-driven ML models offer powerful predictive capabilities, they have limitations such as limited extrapolation beyond the training data and strong reliance on large datasets owing to their black-box nature. These models generally perform well within the scope of the training data but often face challenges when confronted with data outside that range. To overcome these challenges, physics-informed, physics-constrained, or physics-guided ML approaches have received increasing research attention. These methods integrate the predictive power of ML with physical models by embedding physical principles into loss functions, allowing for more reliable predictions even for data beyond the training range. This integration can involve directly solving governing conservation equations or incorporating theoretical models and empirical correlations based on experimental data. One notable example is the physics-informed neural network (PINN) introduced by Raissi et al., [25] which solves partial differential equations (PDEs) by incorporating conservation and boundary conditions into the neural network's loss function. By combining ML with mathematical theory, Raissi's approach enhances conventional data-driven models, thus enabling accurate predictions even in regimes with limited data.

In this study, we employ a physics-informed ML approach that integrates the mechanistic relationships of condensation and thus reduces data dependence to enhance the prediction accuracy of condensation HTC for steam–NCG mixtures over a vertical tube. This approach has demonstrated higher accuracy and lower risk of overfitting than other classical ML methods, such as MLP, random forest regression (RFR), support vector regression (SVR), decision tree regression (DTR), Gaussian processes regression (GPR), elastic net regression (ENR), and light gradient boosting machine (LightGBM), and extreme gradient boosting (XGBoost). [26] To incorporate physical constraints into these conventional ML models, we developed a novel loss function that combines mean squared error (MSE) loss with a physics-constrained loss based on our newly developed degradation Nusselt model. Our method, referred to as physics-constrained XGBoost (PC-XGBoost), is further optimized with a Bayesian algorithm to improve its predictive performance and extrapolation capabilities, addressing the limitations of purely data-driven models. By integrating physical constraints from the degradation Nusselt model with standard ML techniques, this approach offers more robust and reliable predictions of HTC, surpassing the accuracy and consistency afforded by traditional data-driven models.

2. Methodology

2.1. Database and preprocessing

A database of 856 experimentally measured condensation HTC data points for steam–air or steam–nitrogen mixtures over a vertical plate has been compiled by Cho et al. [22] After including recently published data, [27,28] the database now contains 879 experimental interpolation data points collected from 13 sources. Each dataset includes 17 input parameters representing environmental and geometric conditions (D_h, L, M_{nc}, P_{tot}, W_{nc}, ΔT_{sub}, T_{sat}, k_f, k_g, $C_{p,f}$, $C_{p,g}$, h_{fg}, μ_f, μ_g, σ, ρ_f, ρ_g) and 1 output parameter (HTC). The input parameters were selected based on their significant correlation with HTC, the ability of the ML method to handle variables with different units through normalization, and the importance of capturing physical properties such as steam characteristics and the mass fraction of non-condensable gases, as air and nitrogen dominate most practical scenarios. In addition, uncertainty in the data was indirectly addressed [29] by ensuring data collection under consistent conditions with free-falling NCGs, incorporating inherent uncertainty handling through GPR, and conducting a robust error analysis using multiple metrics such as the mean absolute percentage error (MAPE) and determination coefficient (R^2). All input parameters were preprocessed using min–max normalization, and the output parameter was log-transformed to reflect the distribution characteristics of the HTC data.

The database of condensation HTC was divided into training (60 %), validation (20 %), and test (20 %) datasets for all ML models. In addition, for validating the predictive accuracy for extrapolation data, 199 data points were selected from three additional sources. [30–32] These data points were obtained from independent experimental conditions that were not involved in the model development. The sources used for both the interpolation and extrapolation datasets are detailed in Appendix Tables A.1 and A.2.

Table 1

2.2. Proposed new degradation Nusselt model

Condensed, free-falling films exhibit different heat transfer characteristics depending on their flow regimes, which can be laminar, wavy-laminar, or turbulent, as determined by the film's Reynolds number (see Fig. 1). In 1916, Nusselt [9] introduced a theoretical model for predicting laminar film condensation, taking into account the thermophysical properties of the working fluid and the condensation length. The derived equation from the force balance can be summarized as follows.

$$\frac{du}{dy} = \frac{g(\rho_f - \rho_g)}{\mu_f}(\delta - y) \tag{1}$$

The model starts with the differential equation for velocity within the condensation film, derived from the balance of forces, including gravity and viscous forces. Here, u represents the velocity within the condensation film, whereas δ and y denote the film thickness and the distance from the wall, respectively. By applying continuity and energy balance equations, where the total heat transfer is the sum of the latent heat of vaporization lost by condensation and the subcooling of the same mass of liquid to the mean fluid temperature, we obtain the following equation with an assumption of linear temperature profile within the liquid film:

$$q''_w \Delta z = h_{fg}\frac{d\Gamma}{dz}\Delta z + c_{p,f}(\rho_f)(T_{sat} - T_m)\frac{d\Gamma}{dz}\Delta z, \tag{2}$$

where $c_{p,f}(\rho_f)(T_{sat} - T_m) = \frac{\int_0^\delta \rho_f u c_{p,f}(T_{sat}-T)dy}{\int_0^\delta \rho_f u dy}$.

By rearranging the equation and applying Rohsenow's modified parabolic temperature profile [33] with a linear assumption, the film thickness and local HTC can be expressed as follows.

$$\delta_{Nu} = \sqrt[4]{\frac{4k_f\mu_f\Delta Tz}{g\rho_f(\rho_f - \rho_g)h'_{fg}}} \tag{3}$$

Table 1
Thermo-physical and operating ranges of data in the consolidated database.

Dataset	T_{sat} [°C]	ρ_f [kg/m^3]	ρ_g [kg/m^3]	$\mu_f \times 10^4$ [kg/m s]	$\mu_g \times 10^4$ [kg/m s]	k_f [W/m K]	k_g [W/m K]	$C_{p,f}$ [J/kg K]	$C_{p,g}$ [J/kg K]	h_{fg} [kJ/kg]	σ [N/m]	Re_{film}
Interpolation [22,27,28]	62–210	887–988	0.15–9.6	0.23–5.5	0.11–0.16	0.64–0.68	0.022–0.042	4182–4405	1969–3153	1899–2352	0.036–0.065	8–7160
Extrapolation [30–32]	78–196	924–995	0.27–7.5	1.9–7.5	0.12–0.16	0.61–0.67	0.023–0.039	4182–4294	1977–2762	1954–2314	0.039–0.063	20–2148

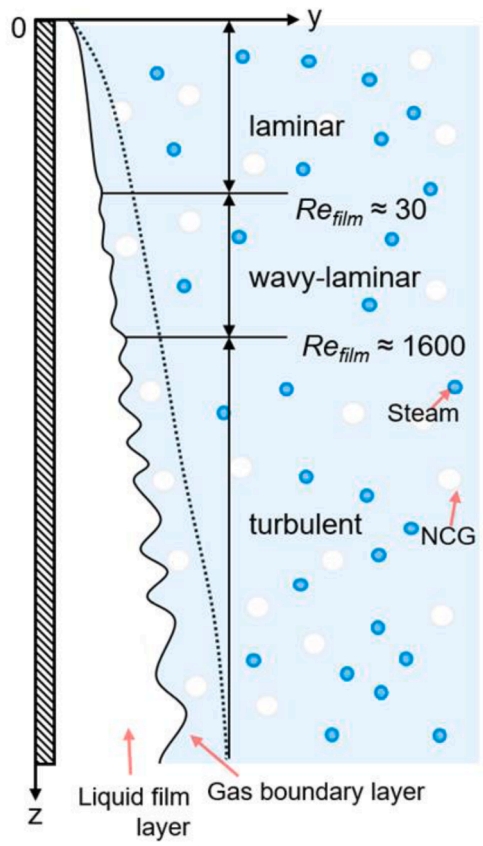


Fig. 1. | Conceptual schematic of free-falling condensation heat transfer on vertical surface.

$$h_{Nu} = \sqrt[4]{\frac{g\rho_f(\rho_f - \rho_g)k_f^3 h'_{fg}}{4\mu_f \Delta T z}} \tag{4}$$

Here, $h'_{fg} = h_{fg} + 0.68c_{p,f}\Delta T$.

Finally, deriving the average HTC from the local HTC yields the following expression.

$$\overline{h}_{Nu} = \frac{1}{L}\int_0^L h_{Nu}dz = \frac{4}{3}h_L = 0.943\left(\frac{g\rho_f(\rho_f - \rho_g)k_f^3 h'_{fg}}{\mu_f \Delta T L}\right)^{1/4} \tag{5}$$

Although the Nusselt model is widely recognized for its accuracy in predicting condensation HTC in the laminar flow regime, it tends to underestimate HTC in the wavy-laminar and turbulent regimes, as it omits interfacial wave and turbulence effects. [34] To address this shortcoming, modified models for wavy-laminar and turbulent flow regimes that incorporate wave dynamics and turbulence effects, either through turbulent models or empirical approaches, have been proposed. [10–14] However, these turbulence-based models are often complex, lack universal applicability, and tend to inaccurately predict heat transfer at low Prandtl numbers owing to their inability to account for increased film thickness. [35]

Here, we chose the degradation factor method to account for NCG. Although this method relies on empirical correlations, it directly constrains predictions to analytical solutions, offering deeper insights into the physical relationships in condensation heat transfer. The concept of a degradation factor for NCG condensation has been developed and refined by several researchers. Vierow [36] introduced the degradation factor as the ratio of experimental to pure steam condensation HTC based on the Nusselt model, correlating it with the mixed-gas Reynolds number and NCG mass fraction. Kuhn [14] developed a correlation for the degradation factor in air–helium FWC experiments using the ratio of film thickness, film Reynolds number, and NCG mass fraction. Lee and

Kim [37] extended this work by exploring local heat transfer in a U-tube during reflux condensation with NCG, incorporating film Reynolds number, NCG mass fraction, and shear force. Their degradation factors and related parameters are summarized in Table 2.

Nevertheless, challenges in adopting these degradation Nusselt models as a physics-based approach remain. All three aforementioned models have limitations for universal applicability. First, they are derived under restricted experimental conditions; second, the parameters used are not always appropriate. In our consolidated database, which largely comprises free-falling natural convection scenarios, defining the mixed gas Reynolds number is challenging. Moreover, few studies have measured liquid film thickness, further limiting the usability of these models. To address this, we redefined the parameters to derive a degradation factor that is more universally applicable and suitable for our database. Fig. 2 shows the ratio of the HTC predicted by the Nusselt model to the experimentally obtained HTC as a function of Refilm, with the color of the symbols representing the mass fraction of NCGs. The Nusselt model consistently overestimates HTC as the NCG fraction (Wnc) increases. In addition, the model tends to overestimate HTC as Refilm decreases because, in the laminar regime, turbulent effects diminish. Conversely, as Refilm increases into the turbulent regime, mixing within the film and the increased thermal conductivity owing to turbulent eddies cause the actual HTC to increase, leading the Nusselt model to underestimate HTC. These findings suggest that both Wnc and Refilm are essential degradation factors for predicting free-falling condensation in the presence of NCG, and they must be considered separately for the laminar, wavy regions, and turbulent regimes. The equation for the determined factor is as follows.

$$\mathrm{Re}_{film} \equiv \frac{4\Gamma}{\mu_f} = \frac{4}{\mu_f}\frac{h_{Nu} L \Delta T_{\mathrm{sub}}}{h_{fg'}} \tag{6}$$

$$F = \frac{h_{\mathrm{exp}}}{h_{Nu}} = f_1 f_2 = \left(1 + a\mathrm{Re}_{film}^{b}\right)\left(1 - cW_{nc}^{d}\right) \tag{7}$$

The values of a, b, c, and d were derived using the GRG nonlinear method [38] with a convergence criterion of 1e-12 based on an interpolation dataset of 879 data points. This process led to the selection of the new degradation Nusselt model as the physical model for this study.

2.3. ML approach

2.3.1. Fully data-driven methods

To evaluate the performance of the physics-constrained method, we compared it with eight well-known fully data-driven ML methods that are highly regarded for their regression modeling capabilities:

(1) XGBoost [26] is an ensemble learning technique that combines predictions from multiple weak learners to create a more robust model. It efficiently implements gradient boosting, using second-order gradients for optimization and regularization to prevent overfitting. In this study, XGBoost was implemented using the XGBoost library.

(2) LightGBM [39] is a gradient boosting framework that utilizes tree-based learning algorithms. It is designed for high performance, high training speed, low memory usage, and high accuracy. In our experiments, LightGBM was implemented using the LightGBM library.

(3) SVR [40] is a type of support vector machine (SVM) that supports both linear and non-linear regression. It aims to fit the error within a certain threshold, making it robust to outliers. The scikit-learn library was used to implement SVR in this study.

(4) Also known as the deep neural network, MLP [41] is a class of feedforward artificial neural networks comprising multiple layers of nodes in a directed graph, with each layer fully connected to the next. We implemented MLP using the TensorFlow library.

(5) RFR [42] is an ensemble learning method that constructs multiple decision trees during training and outputs the mean prediction of the individual trees. It is particularly useful for handling high-dimensional data. The scikit-learn library was used for implementing RFR.

(6) ENR [43] combines the properties of both Lasso and Ridge regressions, making it useful in the presence of multiple correlated features. This method was implemented using the scikit-learn library.

(7) DTR [44] uses a decision tree to model the relationship between input features and target values. It splits the data into subsets based on the value of input features, resulting in a tree-like model of decisions. We implemented DTR using the scikit-learn library.

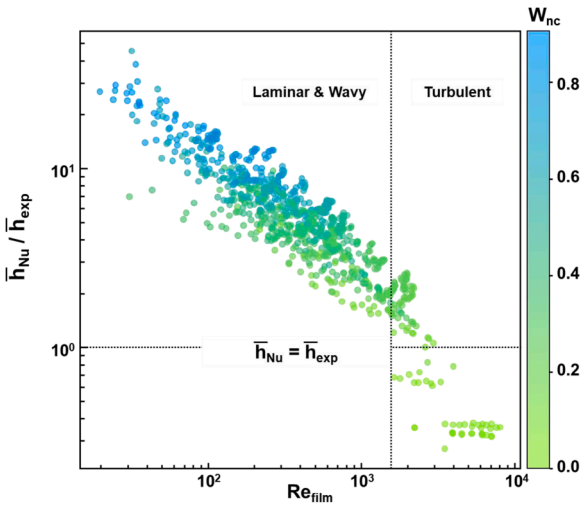


Fig. 2. | Determination of degradation factor parameters. The ratio of the HTC by Nusselt number to the experimentally obtained HTC according to film Reynolds number and NCG mass fraction.

Table 2
The existing and new degradation factor for the Nusselt model.

Author(s)	Correlation	Range		a	b	c	d
Vierow et al. (1990) [36]	$F = (1 + a\mathrm{Re}_m^b)(1 - cW_{nc}^d)$	Air	$W_{nc} < 0.063$	2.88E-05	1.18	10	1
	Re_m: Mixture Reynolds number$(\rho_m u D/\mu_m)$		$0.063 < W_{nc} < 0.6$			0.94	0.13
	W_{nc}: NCG mass fraction		$0.6 < W_{nc}$			1	0.22
Kuhn et al. (1997) [14]	$F = \frac{\delta_{\mathrm{exp}}}{\delta_{Nu}}(1 + a\mathrm{Re}_{film})(1 - bW_{nc}^c)$	Air	$W_{nc} < 0.1$	7.32E+04	2.6	0.708	
			$0.1 < W_{nc}$		1	0.202	
	Re_{film}: Film Reynolds number $\left(\Gamma/\mu_f\right)$	Helium	$0.003 < W_{nc} < 0.01$		−35.8	1.074	
	δ_{exp}: Experimental thickness		$0.01 < W_{nc} < 0.1$		−2.09	0.457	
	δ_{Nu}: Nusselt thickness		$0.1 < W_{nc}$		−1	0.137	
Lee and Kim (2008) [37]	$F = \tau_g^{*a}(1 + b\mathrm{Re}_{film})(1 - W_{nc}^c)$	Air	$0.038 < W_{nc} < 0.814$	0.3124	0.964	0.402	
	τ_g^*: Dimensionless of shear stress $(\tau_g/(\rho g L))$		$0.06 < \tau_g^* < 46.65$				

(8) GPR [45] is a non-parametric, kernel-based probabilistic model. It predicts the output by assuming a Gaussian process prior over the latent function and using the observed data to update this prior. The scikit-learn library was used to implement GPR in our experiments.

By employing these diverse ML methods, we thoroughly evaluated the performance and robustness of the proposed physics-constrained approach against various fully data-driven techniques. Each method was implemented using well-established libraries in Python to ensure reliability and consistency in the results.

2.3.2. Physics-constrained XGBoost method

In this study, to implement a physics-constrained method, we employed XGBoost, as it demonstrated the most stable performance on extrapolation datasets relative to fully data-driven ML models. A comprehensive analysis of the results is provided in Section 3.2, where the performance of the physics-constrained and data-driven ML models is compared. The integration model for a decision tree can be described as follows.

$$\widehat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in F \tag{8}$$

In this context, y_i represents the predicted value, K is the total number of decision trees, f_k denotes the output from the k-th regression tree, F signifies the space of all regression trees, and x_i is the i th input feature.

The regularized objective function of the XGBoost model is expressed as follows.

$$\mathscr{L}(\phi) = \sum_i l(\widehat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{9}$$

$$\Omega(f) = \gamma T_l + \frac{1}{2}\tau \| \omega \|^2 \tag{10}$$

Here, l is a differentiable convex loss function that measures the difference between the predicted value y_i and the target y_i and is calculated by l. Ω is the regularization term, which penalizes the complexity of the model. α and λ are penalty factors; T_l is the number of leaf nodes in the tree; and w is the weight of the leaf nodes. The XGBoost model is incrementally trained to minimize the objective function. By adding an incremental function $f_t(x)$ at the t-th iteration, the following objective function is obtained.

$$\mathscr{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \widehat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{11}$$

A second-order approximation is applied to Eqs. (11), and constant terms are removed to optimize the objective:

$$\mathscr{L}^{(t)} \approx \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} k_i f_t^2(x_i) \right] + \Omega(f_t), \tag{12}$$

where g_i and k_i are the first- and second-order gradient statistics on the loss function, respectively. Eqs. (12) can be rewritten by expanding Ω:

$$\mathscr{L}^{(t)} = \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} k_i + \lambda \right) \omega_j^2 \right] + \gamma T_l, \tag{13}$$

where I_j is the instance set of leaf j.

Substituting the optimal weight of $\omega_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}$ into Eqs. (13), we obtain:

$$\mathscr{L}^{(t)} = \frac{1}{2} \sum_{j=1}^{T} \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} k_i + \lambda} + \gamma T_l. \tag{14}$$

Here, we employed a novel loss function that combines the MSE loss function and a physics-constrained loss function. The new degradation HTC model proposed in Section 2.1 was used for the physics-constrained loss function.

The MSE between the values obtained from Eqs. (14) and XGBoost is used as the physics-constrained loss.

$$l_{MSE,i} = \frac{1}{2}(h_i - \widehat{h}_i)^2 \tag{15}$$

$$\overline{h}_{\mathrm{deg}} = F\overline{h}_{Nu} = \left(1 + a\mathrm{Re}_{film}^{b}\right)\left(1 - cW_{nc}^{d}\right) 0.943 \left[\frac{g\rho_f(\rho_f - \rho_g)k_f^3 h_{fg'}}{\mu_f L \Delta T_{\mathrm{sub}}} \right]^{1/4} \tag{16}$$

$$l_{PC,i} = \frac{1}{2}\left[\left(1 + a\mathrm{Re}_{film}^{b}\right)\left(1 - cW_{nc}^{d}\right) 0.943 \left[\frac{g\rho_f(\rho_f - \rho_g)k_f^3 h_{fg'}}{\mu_f L \Delta T_{\mathrm{sub}}} \right]^{1/4} - \widehat{h}_i \right]^2 \tag{17}$$

Here, h'_{fg} is the modified latent heat given by $h'_{fg} = h_{fg} + 0.68 C_{pf} \Delta T_{\mathrm{sub}}$ for a nonlinear temperature profile. [33]

A new hyperparameter, the adjustment factor λ, is introduced to regulate the relative influence of $l_{\mathrm{pc,i}}$ and $l_{mse,\mathrm{i}}$. Consequently, the new loss function can be expressed as follows.

$$l_i = l_{mse,i} + \lambda l_{pc,i} \tag{18}$$

In summary, the XGBoost model was modified to include physical constraints by integrating the degradation Nusselt model into the MSE loss function. An adjustable parameter λ was used to balance the experimental data with physical principles, improving generalization and ensuring predictions align with physical laws, particularly on extrapolation datasets.

The schematic of the Physics constrained (PC)-XGBoost model for predicting the condensation HTC of steam–NCG mixtures is depicted in Fig. 3. The training dataset is employed to train the PC-XGBoost model, and the validation dataset is utilized to evaluate the model at each iteration. The test dataset is used to assess the accuracy and reliability of the ML model using metrics such as MAPE, root mean square error (RMSE), R^2, and the percentage of data points predicted within ±30 % (P_{30}).

$$\mathrm{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\overline{h}_{\mathrm{pre}} - \overline{h}_{\mathrm{exp}}}{\overline{h}_{\mathrm{exp}}} \right| \times 100 \tag{19}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\overline{h}_{\mathrm{pre}} - \overline{h}_{\mathrm{exp}} \right)^2} \tag{20}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \left(\overline{h}_{\mathrm{pre}} - \overline{h}_{\mathrm{exp}} \right)^2}{\sum_{i=1}^{N} \left(\overline{h}_{\mathrm{exp}} - \overline{\overline{h}}_{\mathrm{exp}} \right)^2} \tag{21}$$

N is the number of data point, $\overline{h}_{\mathrm{exp}}$ and $\overline{h}_{pred}$ are the experimental and predicted values of condensation HTC, respectively, and $\overline{\overline{h}}_{\mathrm{exp}}$ is the average of $\overline{h}_{\mathrm{exp}}$ overall data points. The present ML models are developed under the Python 3.10.12 environment using the TensorFlow 2.15 and Hyperopt packages. [46]

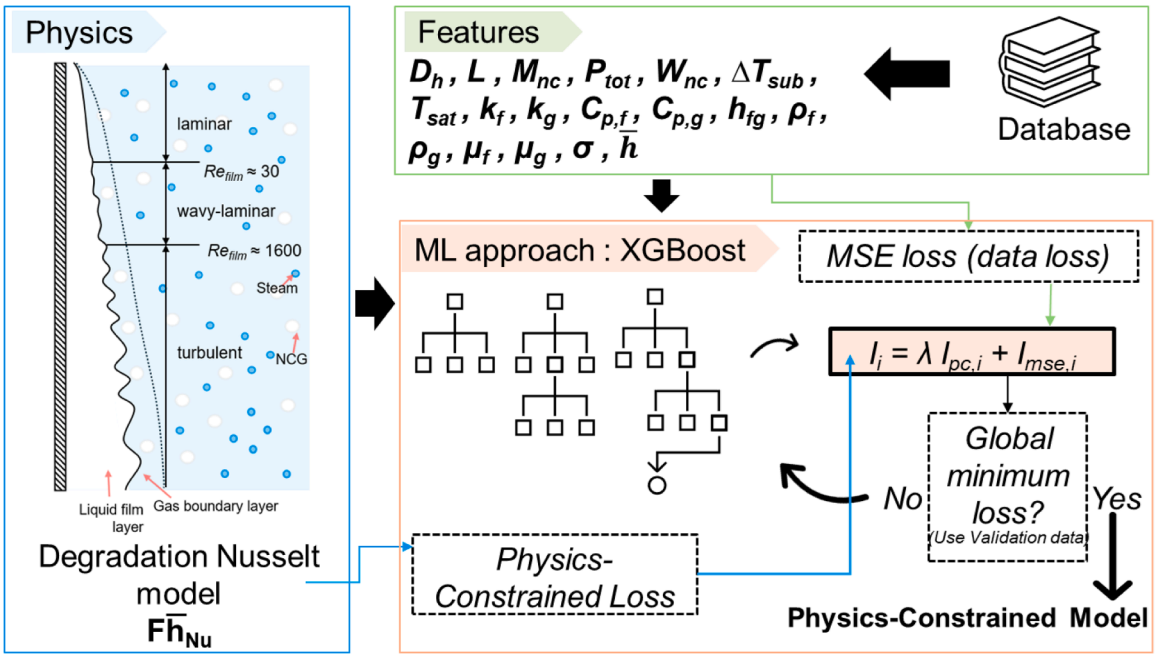


Fig. 3. | The schematic of the PC-XGBoost model. The final loss is computed by adjusting the data loss and physics loss with a factor guiding model retraining.

2.4. Hyperparameter optimization

All ML methods were optimized using the Hyperopt library, which utilizes a Bayesian optimization technique. Each model was trained on 100 different configurations. Given the varying hyperparameters for each ML method, we focused on optimizing those that significantly influence model performance.

In particular, for the PC-XGBoost model, hyperparameters such as the physics penalty factor, learning rate, evaluation metric, colsample by tree, subsample, max depth, L1 regularization, L2 regularization, minimum child weight, and minimum loss reduction were optimized using the tree-structured Parzen estimator (TPE). The optimized hyperparameters for this model are presented in Table 3. TPE is a Bayesian optimization method that enhances search efficiency using the performance of previously evaluated parameters to predict the most promising next steps. The objective function, designed to minimize the RMSE of the validation dataset using K-fold cross-validation ($k = 5$), [47] evaluates the performance of the hyperparameters. The Hyperopt library facilitated this optimization process. To ensure fair comparison, the same Bayesian optimization approach with consistent constraints was applied to all ML models. The use of a unified framework ensured unbiased evaluation under identical conditions, and the TPE method effectively balanced exploration and exploitation, providing a reliable basis for comparing performance across the interpolation and extrapolation datasets. [48]

3. Results and discussion

3.1. Degradation Nusselt model

Based on the procedures outlined in the methodology, the proposed degradation Nusselt model is presented below.

$$Re_{film} < 1600 \quad h_{deg} = Fh_{Nu} = \left(1 + 1.19\mathrm{Re}_{film}^{0.39}\right)\left(1 - 0.994\, W_{nc}^{0.011}\right) \tag{22a}$$

$$Re_{film} \geq 1600 \quad h_{deg} = Fh_{Nu} = \left(1 + 0.001\mathrm{Re}_{film}^{0.837}\right)\left(1 - 1.27\, W_{nc}^{0.01}\right) \tag{22b}$$

Fig. 4 compares the experimental HTC values with the HTC predictions from (a) the original Nusselt model and (b) the proposed degradation Nusselt model. In Fig. 4(a), the comparison reveals that the Nusselt model tends to overestimate HTC values in the lower range likely owing to the correlation between low HTC values at lower Re_{film} and a higher mass fraction of NCGs. Conversely, as HTC approaches 10,000, the Nusselt model tends to underestimate the values. This underestimation can be attributed to the increasing liquid film thickness, which leads to a transition from laminar to wavy-laminar or turbulent flow in the liquid film, thereby increasing the HTC.

Fig. 4(b) displays the results predicted by the proposed degradation Nusselt model. As illustrated in this figure, the semi-theoretical model shows strong alignment with various experimental conditions, significantly reducing the MAPE from 522.3 % to 15.7 %. The degradation model underpredicted the HTC, particularly for the data by Kim et al. (2009), [49] because the turbulent Re_{film} had a nitrogen gas mass fraction below 0.1. This discrepancy likely stems from the limited data available for turbulent Re_{film} compared with laminar and wavy Re_{film}. In the turbulent regime, most test cases involved either pure vapor or air with a nitrogen gas mass fraction above 0.2, which led to inaccuracies in determining the NCG mass fraction coefficient. Nonetheless, these cases account for <2 % of the overall dataset. When these instances are excluded, the degradation model consistently outperformed the original Nusselt model.

3.2. Comparison of physics-constrained and data-driven ML models

The effectiveness of each ML model was assessed using interpolation test data, with evaluation metrics including MAPE, RMSE, R^2, and P30, as depicted in Fig. 5. First, considering MAPE, which allows for an intuitive comparison of model performance, the PC-XGBoost model achieved a MAPE of 10 %. These results are slightly better than those of GPR, SVR, RFR, and MLP. RMSE is highly influenced by large errors; hence, GPR shows higher RMSE values than SVR and RF. The PC-XGBoost model also includes data with significant errors, resulting in a higher RMSE than DTR, despite having a lower MAPE.

The R^2 metric assesses the extent to which the predicted values capture the variability of the actual data. Models such as XGBoost, GPR, SVR, RFR, MLP, DTR, and LightGBM demonstrate strong predictive

Table 3
Selected hyperparameters for the PC-XGBoost model.

Model	L1 regularization	Col sample by tree	Minimum loss reduction	L2 regularization	Learning rate	Maximum depth	Minimum child weight	Physics penalty factor	Subsample
PC-XGBoost	0.0116	0.7076	0.0025	15.3319	0.003	14.0	7.0770	1.0851	0.6327

performance, with R^2 values exceeding 0.97, indicating their effectiveness in explaining data variability. In contrast, the ENR model, which relies on linear regression, exhibits a MAPE of 18 % and an R^2 of 0.67, indicating suboptimal performance. This disparity is likely attributable to the model's inherent limitations in addressing nonlinear patterns and high-dimensional datasets. Given the complexity of our dataset, which includes 17 input parameters, the ENR model was not well-suited to accurately learn the intricate relationships between these variables.

Based on the interpolation test data alone, the PC-XGBoost model did not demonstrate distinct advantages over other fully data-driven models, ranking fifth in MAPE performance among the nine models assessed. However, evaluating model performance solely on this test dataset has inherent limitations. Although the test data were not used for model training, the training and validation datasets include data from the same experimental setup and similar test conditions, which may obscure the signs of overfitting due to the dependence on the dataset split. To address this issue and provide a more accurate assessment of predictive performance, we extracted an extrapolation dataset from three independent studies unrelated to the interpolation dataset. This approach enables a more robust and unbiased comparison of model performance.

3.3. Extrapolation capability for physics-constrained ML models

Fig. 6 uses kernel density estimation (KDE) to compare the data distribution of key parameters between the interpolation and extrapolation datasets. KDE is a non-parametric method that estimates the probability density function of a dataset by applying a kernel function to each data point and averaging these to estimate the overall distribution. In Fig. 6(a) and 6(b) shows that values of ΔT_{sub} and T_{sat}, respectively, are more frequently distributed at higher temperatures in the extrapolation dataset. In addition, both the mean and variance are larger in the extrapolation than interpolation dataset. As observed in Fig. 6(c), the latent heat of vaporization (h_{fg}) has a lower mean and higher variance in the extrapolation dataset than in the interpolation dataset. In Fig. 6(d), the NC mass fraction covers the same range from 0 to 1.0 in both datasets. However, the interpolation and extrapolation datasets respectively show higher density of values between 0.3 and 0.5 and between 0.15 and 0.25. These differences in data composition highlight the distinct conditions and geometries under which the datasets were collected. This variation underscores the importance of considering both interpolation and extrapolation datasets for robust model evaluation and indicates that overcoming the limitations of test data may be key to improving model generalization.

Fig. 7 compares the performance of each model on (a) the interpolation and (b) the extrapolation dataset. On the interpolation dataset, the GPR model showed the best performance. However, when applied to the extrapolation dataset, it underestimated the values, resulting in a significant increase in MAPE from 4.96 % to 51.93 %. Similarly, the second-best-performing model on the interpolation dataset, SVR, overestimated the values on the extrapolation dataset, causing its MAPE to jump from 6.26 % to 656.62 %.

The GPR model, which relies on data distribution for predictions, tends to converge toward the mean of the data when faced with out-of-range values, as seen in Fig. 7. This results in predicted values lower than the actual ones. In contrast, the SVR model, which classifies data based on hyperplanes, tends to maintain the slope of the hyperplane outside the data range, leading to overestimation on the extrapolation dataset.

XGBoost, which performed poorly on the interpolation dataset (only better than ENR), showed the best performance among fully data-driven models on the extrapolation dataset. This improvement is attributed to XGBoost's boosting algorithm, which iteratively corrects errors, and its parallel processing and pruning techniques that prevent overfitting.

The primary focus of this study, PC-XGBoost, exhibited moderate performance on the interpolation dataset. However, on the extrapolation dataset, it uniquely showed only a minor increase in error, with

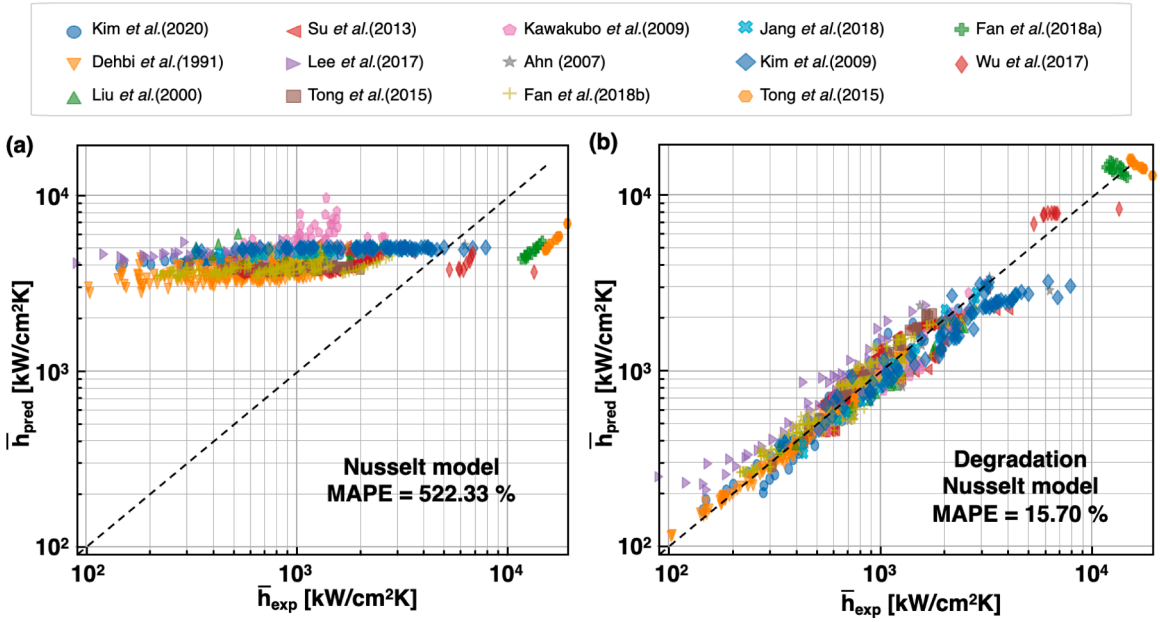


Fig. 4. | Comparison of the predictive accuracy of the original and proposed degradation Nusselt models. The predicted HTC ($\bar{h}_{pred}$) versus the experimental HTC ($\bar{h}_{exp}$) is shown for (a) the original Nusselt model and (b) the degradation Nusselt model.

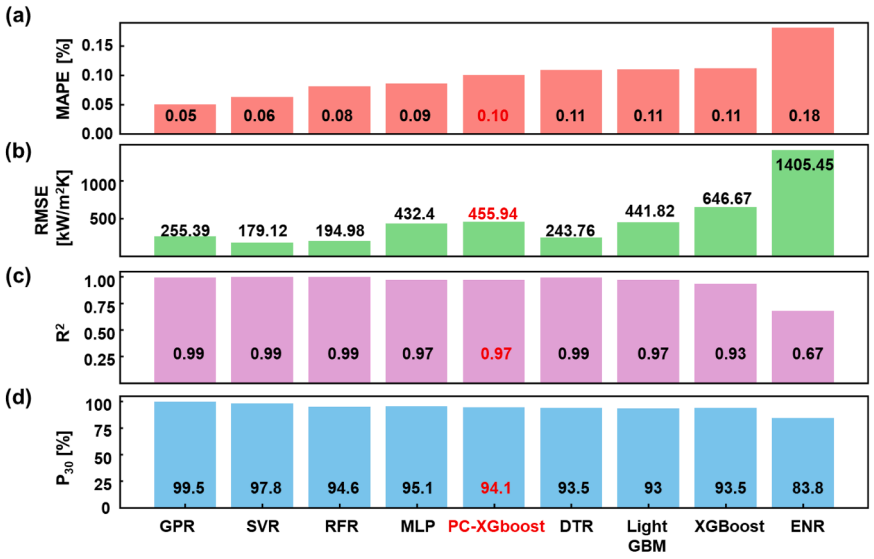


Fig. 5. | Interpolation dataset—Comparison of each ML model for different error metrics. (a) MAPE, (b) RMSE, (c) R^2, and (d) P_{30}.

MAPE rising by just 1.2 % to 11.22 %. This indicates its strong ability to effectively predict beyond the data range. The degradation factor in the physics model, which incorporates mass fraction and film Reynolds number into the Nusselt model derived from energy and mass balance, likely provides the necessary physical correlations. For the extrapolation dataset, the physics model had a MAPE of 18.07 %, outperforming all data-driven ML models. This suggests that the physics model's integration into PC-XGBoost significantly enhanced its prediction accuracy and reliability. This advantage is attributable to the physics-constrained framework of PC-XGBoost, which reduces dependence on training data alone and leverages physical principles to provide robust predictions, minimizing the overfitting commonly observed in fully data-driven models.

The performance of each model on the extrapolation dataset can be examined through various performance metrics in Fig. 8. Although the MAPE error, which provides an intuitive evaluation of model performance, decreased for all models on the extrapolation dataset compared with the interpolation dataset, the PC-XGBoost model showed only a 1.2 % increase, maintaining nearly the same performance. In particular, for the PC-XGBoost model, the RMSE significantly improved from 455.94 on the interpolation dataset to 190.05 on the extrapolation dataset. This improvement is likely attributable to the presence of outlier errors in the interpolation test dataset. In other words, the extrapolation dataset lacks extremely high or low HTC values. A similar trend is observed in XGBoost, where the RMSE improved from 646.67 to 325.25, indicating better performance owing to the reduced impact of outliers in the extrapolation dataset.

In contrast, the R^2 metric, which evaluates the suitability of the models, showed negative values for the DTR, GPR, and SVR models, indicating that these models are inadequate because they converge to the mean for out-of-range data owing to their tendency to maintain the learned decision boundaries. Models with positive R^2 values showed

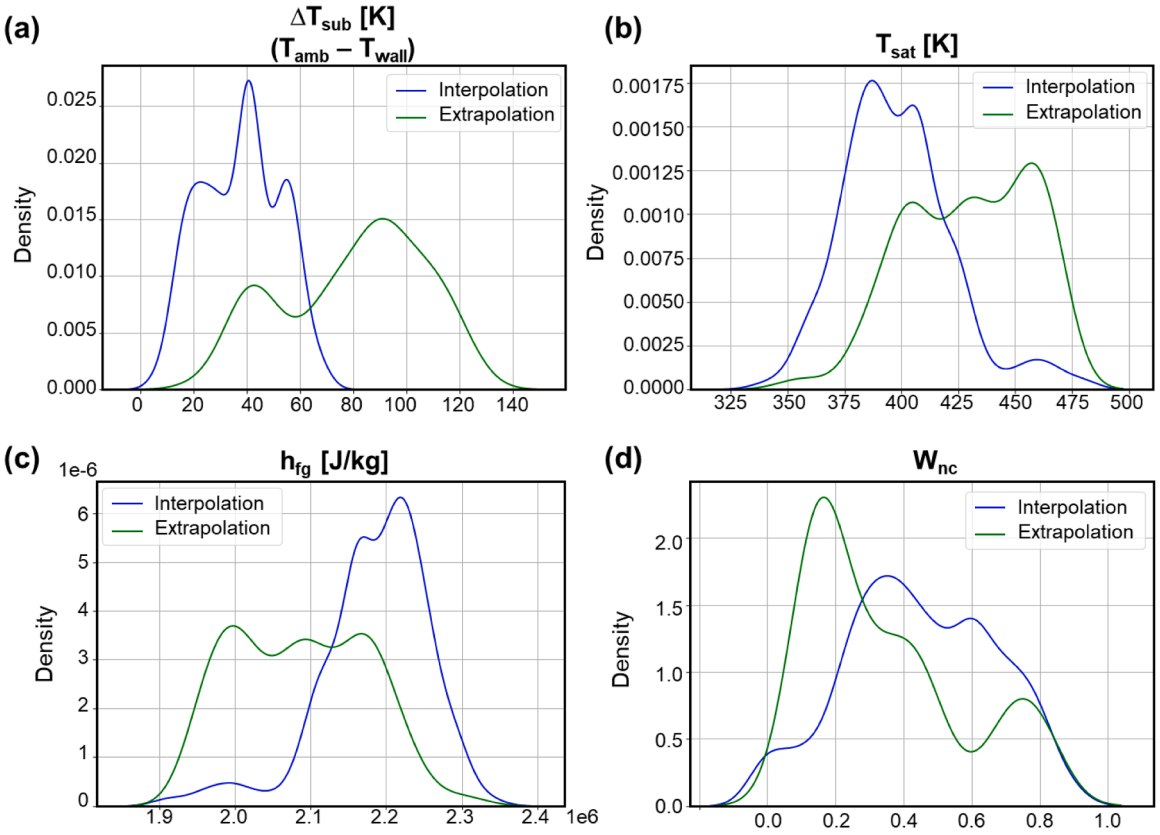


Fig. 6. | Kernel density estimation of the number of data points for the interpolation and extrapolation dataset. (a) ΔT_{sub}, (b) T_{sat}, (c) h_{fg}, and (d) W_{nc}.

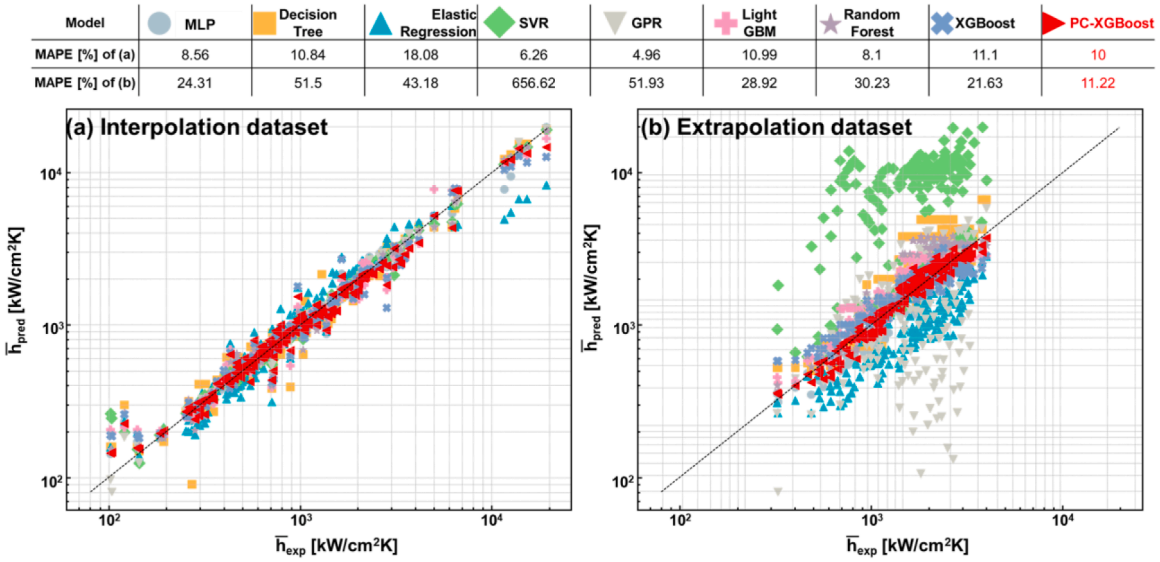


Fig. 7. | The predicted HTC ($\bar{h}_{pred}$) according to the experimental HTC ($\bar{h}_{exp}$) for each ML model. (a) Interpolation and (b) extrapolation dataset.

high MAPE values, confirming their appropriateness for predicting data beyond the training range.

In conclusion, across all performance metrics, including MAPE, RMSE, R^2, and P30, the PC-XGBoost model exhibited the best performance on the extrapolation dataset. Among the fully data-driven ML methods, XGBoost showed the best performance, highlighting its robustness in handling extrapolation tasks.

Fig. 9 examines the changes in HTC with respect to the W_{nc} mass fraction using part of the interpolation dataset from Kim et al. (2020) [27] and part of the extrapolation dataset from Kang et al. (2021). [30] On the interpolation dataset, the degradation Nusselt model generally underestimates the values for Kim et al.'s data; as a result, the PC-XGBoost model also underestimates owing to its reliance on this model. In addition, in the mass fraction range of 0.1–0.3, the predictions across different models show minimal differences. This is because other values from the same experimental set are included in the training and validation datasets, leading to similar predictions among the models.

In contrast, on the extrapolation dataset, the trends for the GPR, SVR,

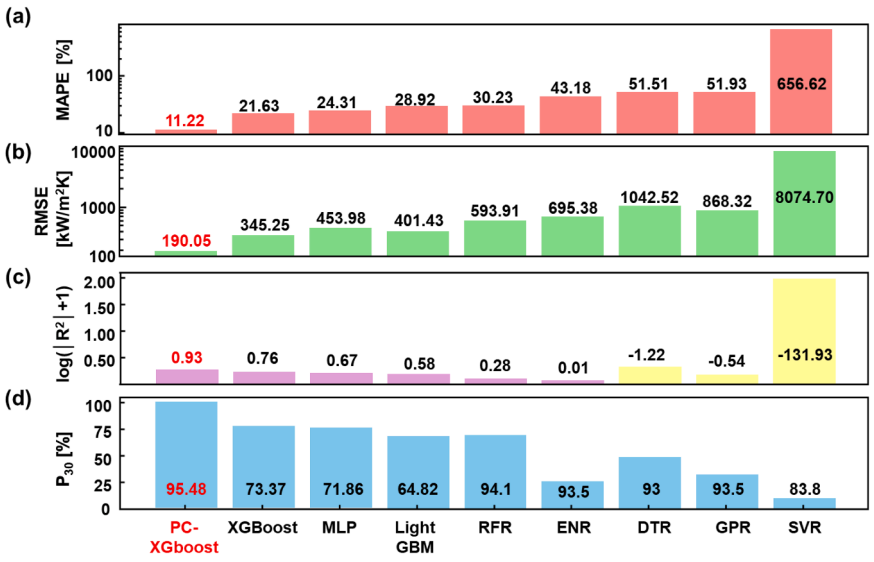


Fig. 8. | Extrapolation dataset–Comparison of each ML model for different error metrics. (a) MAPE, (b) RMSE, (c) R^2, and (d) P30.

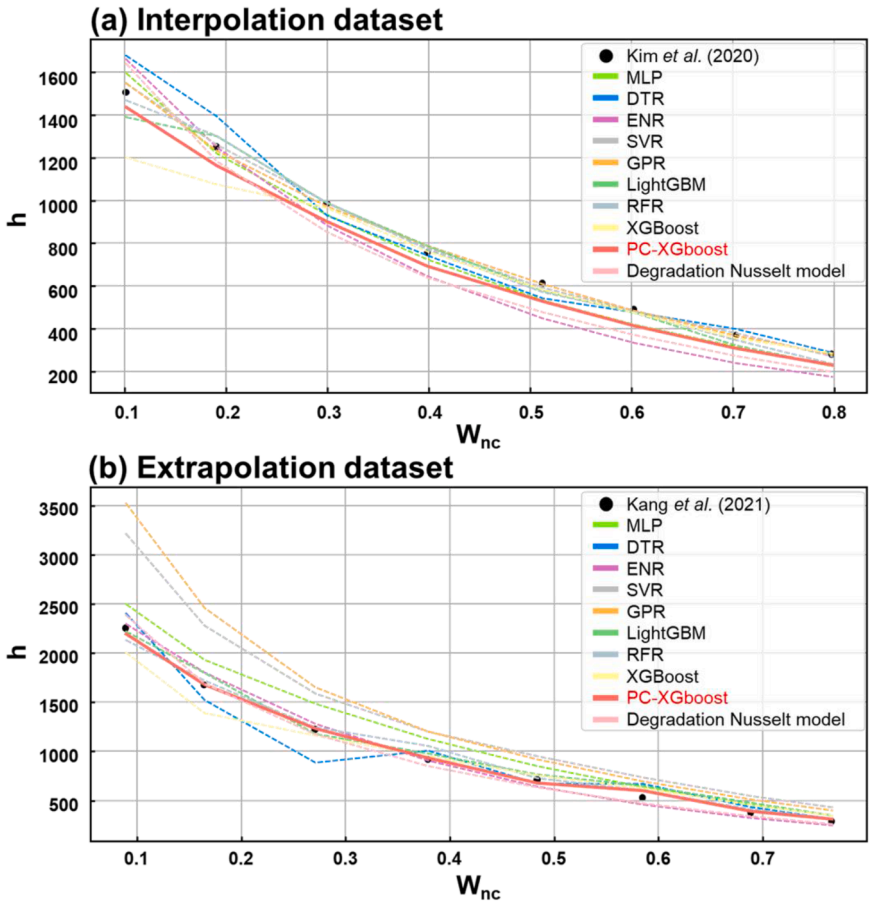


Fig. 9. | Variation of condensation HTC with NCG fraction. (a) Interpolation dataset and (b) extrapolation dataset.

and DTR models significantly deviate in the 0.1–0.3 mass fraction range. This divergence can be attributed to the absence of similar subcooling and saturation temperature ranges in the interpolation dataset, causing these models to heavily rely on the temperature ranges present within the interpolation data to predict HTC. The degradation model, however, performs relatively well on the extrapolation dataset. Consequently, the PC-XGBoost model also operates effectively across various mass fraction ranges without biased errors in specific ranges, demonstrating robust performance regardless of the mass fraction range.

4. Conclusions

In this study, we developed a PC-XGBoost model combined with a newly proposed degradation Nusselt model for enhanced prediction of

condensation HTC in steam–NCG mixtures. Our approach integrates physical constraints from the Nusselt model and incorporates key parameters such as mass fraction and film Reynolds number into the degradation factor. The main findings are as follows:

1. A new degradation Nusselt model was developed, addressing the limitations of existing models that rely on empirical data and simplified assumptions. By incorporating mass fraction and film Reynolds number, the model better captures the physical phenomena of condensation in the presence of NCGs.
2. Although fully data-driven models such as GPR and SVR performed well on interpolation datasets, their accuracy sharply declined in extrapolation scenarios, demonstrating their dependency on the training data range. In contrast, XGBoost, although weaker on the interpolation dataset, showed the best performance on the extrapolation dataset owing to its error correction and overfitting prevention abilities.
3. The PC-XGBoost model, which integrates physical constraints, significantly improved HTC predictions in steam–NCG mixtures, particularly in extrapolation scenarios. Although its interpolation performance was moderate, PC-XGBoost maintained accuracy beyond the training range with only a minor increase in MAPE (from 10 % to 11.22 %). This demonstrates the effectiveness of incorporating physical constraints into the ML framework to obtain a more reliable and accurate prediction model for HTC.
4. These results highlight the value of incorporating physical models into ML frameworks, particularly for resolving complex engineering problems with multiple interacting parameters. The PC-XGBoost model offers a reliable tool for addressing multiphase heat transfer challenges in energy applications, such as heat exchangers, by enhancing prediction accuracy and providing deeper insights into the underlying physical processes of the heat transfer.
5. Future work should expand the dataset to cover a broader range of experimental conditions, particularly in turbulent regimes with varying NCG mass fractions, to improve model generalization. In addition, the incorporation of uncertainty analysis using techniques such as Bayesian neural networks could enhance the model's reliability by providing confidence intervals alongside predictions.

CRediT authorship contribution statement

Haeun Lee: Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization. **Cheonkyu Lee:** Writing – review & editing, Supervision, Conceptualization. **Hyoungsoon Lee:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

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.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024–00353227 and RS-2024–00411577).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.egyai.2025.100482.

Data availability

Data will be made available on request.

References

[1] Beér JM. High efficiency electric power generation: the environmental role. Prog Ener Combust Sci 2007;33:107–34.
[2] Khawaji AD, Kutubkhanah IK, Wie JM. Advances in seawater desalination technologies. Desalination 2008;221:47–69.
[3] Love JC, Estroff LA, Kriebel JK, Nuzzo RG, Whitesides GM. Self-assembled monolayers of thiolates on metals as a form of nanotechnology. Chem Rev 2005;105:1103–70.
[4] Andrews HG, Eccles EA, Schofield WCE, Badyal JPS. Three-dimensional hierarchical structures for fog harvesting. Langmuir 2011;27:3798–802.
[5] Miljkovic N, Wang EN. Modeling and optimization of hybrid solar thermoelectric systems with thermosyphons. Sol Ener 2011;85:2843–55.
[6] Pérez-Lombard L, Ortiz J, Pout C. A review on buildings energy consumption information. Ener Build 2008;40:394–8.
[7] Bae BU, Kim S, Park YS, Kang KH. Experimental investigation on condensation heat transfer for bundle tube heat exchanger of the PCCS. Ann Nucl Ener 2020;139:107285.
[8] Huang J, Zhang J, Wang L. Review of vapor condensation heat and mass transfer in the presence of non-condensable gas. Appl Therm Eng 2015;89:469–84. https://doi.org/10.1016/j.applthermaleng.2015.07.005.
[9] Nusselt W. Die oberflachenkondensation des wasserdamphes. VDI-Z 1916;60:541.
[10] Dehbi A. Out-of-sample validation of a general correlation for steam condensation in the presence of air over a vertical slender tube in turbulent free convection. Nucl Eng Des 2022;395:111865. https://doi.org/10.1016/j.nucengdes.2021.111865.
[11] Bird RB. Transport phenomena. Appl Mech Rev 2002;55. https://doi.org/10.1115/1.1404624.
[12] Bae BU, Cho JU, Kim YS, Jang C, Kim JH. Experimental investigation on condensation heat transfer for bundle tube heat exchanger of the PCCS. Ann Nucl Ener 2020;139:107285. https://doi.org/10.1016/j.anucene.2019.107285.
[13] Wang T, Tong L, Cao X. A universal correlation development for film-wise condensation in the presence of non-condensable gases on vertical walls under turbulent free convection. Ann Nucl Energy 2023;187:109789. https://doi.org/10.1016/j.anucene.2023.109789.
[14] Kuhn SZ, Schrock VE, Peterson PF. An investigation of condensation from steam–gas mixtures flowing downward inside a vertical tube. Nucl Eng Des 1997;177:53–69. https://doi.org/10.1016/S0029-5493(97)00115-1.
[15] Lee H, Kang M, Wook K, Kharangate CR, Lee S, Iyengar M, et al. An artificial neural network model for predicting frictional pressure drop in micro-pin fin heat sink. Appl Therm Eng 2021;194:117012. https://doi.org/10.1016/j.applthermaleng.2021.117012.
[16] Kim K, Lee H, Kang M, Lee G, Jung K, Kharangate CR, et al. A machine learning approach for predicting heat transfer characteristics in micro-pin fin heat sinks. Int J Heat Mass Transf 2022;194:123087. https://doi.org/10.1016/j.ijheatmasstransfer.2022.123087.
[17] Lee H, Lee G, Kim K, Kong D, Lee H. Multimodal machine learning for predicting heat transfer characteristics in micro-pin fin heat sinks. Case Stud Therm Eng 2024;57:104331. https://doi.org/10.1016/j.csite.2024.104331.
[18] Lee H, Kim K, Kong D, Ahn MH, Lee D, Jun H, et al. Dynamic artificial neural network model for ultralow temperature prediction in hydrogen storage tank. J Ener Stor 2023;69:107866. https://doi.org/10.1016/j.est.2023.107866.
[19] Zobeiry N, Humfeld KD. A physics-informed machine learning approach for solving heat transfer equation in advanced manufacturing and engineering applications. Eng Appl Artif Intell 2021;101:104232. https://doi.org/10.1016/j.engappai.2021.104232.
[20] Tizakast Y, Bouzidi M, Khanafer K, Vafai K. Machine learning based algorithms for modeling natural convection fluid flow and heat and mass transfer in rectangular cavities filled with non-Newtonian fluids. Eng Appl Artif Intell 2023;119:105750. https://doi.org/10.1016/j.engappai.2023.105750.
[21] Uzel AR, Edwards RJ, Button BL. An expert system for convective heat transfer measurements using a transient analysis. Eng Appl Artif Intell 1989;2:40–8. https://doi.org/10.1016/0952-1976(89)90006-1.
[22] Cho E, Lee H, Kang M, Jung D, Lee G, Lee S, et al. A neural network model for free-falling condensation heat transfer in the presence of non-condensable gases. Int J Therm Sci 2022;171:107202. https://doi.org/10.1016/j.ijthermalsci.2021.107202.
[23] Lee DH, Kim YS, Choi J, Lee S, Lee J, Ha SJ, et al. Application of the machine learning technique for the development of a condensation heat transfer model for a passive containment cooling system. Nucl Eng Technol 2022;54:2297–310. https://doi.org/10.1016/j.net.2022.03.005.
[24] Cao H, Zhang J, Liu W, Li Y, Huang T. Prediction of heat transfer coefficients for steam condensation in the presence of air based on ANN method. Int J Adv Nucl React Des Technol 2023;5:77–85. https://doi.org/10.1016/j.ijandr.2023.02.005.
[25] Raissi M, Perdikaris P, Karniadakis GE. 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. https://doi.org/10.1016/j.jcp.2018.10.045.
[26] Chen T, Guestrin C. Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. p. 785–94. https://doi.org/10.1145/2939672.2939785.

[27] Kim UK, Yoo JW, Jang YJ, Lee YG. Measurement of heat transfer coefficients for steam condensation on a vertical 21.5-mm-O.D. tube in the presence of air. J Nucl Sci Technol 2020;57:905–16. https://doi.org/10.1080/00223131.2020.1736200.
[28] Wu L, Jia HJ, Ma XZ, et al. Research on the effect of Reynolds correlation in natural convection film condensation. Nucl Sci Tech 2017;28:85. https://doi.org/10.1007/s41365-017-0240-9.
[29] Hosseini SH, Karami H, Panahi A, Mostoufi N, Ahmadi G. New smart models for minimum fluidization velocity forecasting in the tapered fluidized beds based on particle size distribution. Ind Eng Chem Res 2021;60(42):15289–300. https://doi.org/10.1021/acs.iecr.1c02682.
[30] Kang J, Kim H, Bak J, Lim SG, Yun B. Condensation of steam mixed with non-condensable gas on vertical heat exchanger tubes in circumstances with free convection. Int J Heat Mass Transf 2021;169:120925. https://doi.org/10.1016/j.ijheatmasstransfer.2021.120925.
[31] Cao B, Li Y, Lu Y, Zhou S, Bian H, Ding M. Experimental study of air–steam condensation on the influence of tube diameter and inclination angle. Nucl Eng Des 2021;381:111357. https://doi.org/10.1016/j.nucengdes.2021.111357.
[32] Bian H, Lu Y, Li C, Liang T, Ding M. Comprehensive parameter analyses on steam-air condensation at pressures up to 1.6 MPa. Nucl Eng Des 2021;385:111536. https://doi.org/10.1016/j.nucengdes.2021.111536.
[33] Rohsenow WM, Webber JH, Ling AT. Effect of vapor velocity on laminar and turbulent-film condensation. Trans ASME 1956;78:1637–42.
[34] Zhang W, Wang S, Lianbo M. Analytical modeling for vapor condensation in the presence of noncondensable gas and experimental validation. J Heat Transf 2021;143:011601. https://doi.org/10.1115/1.4047481.
[35] Lin YT. Turbulent film condensation in the presence of non-condensable gases over a horizontal tube. Int J Therm Sci 2009;48:1777–85. https://doi.org/10.1016/j.ijthermalsci.2009.02.003.
[36] Vierow KM. Behavior of steam-air systems condensing in cocurrent vertical downflow. Berkeley: University of California; 1990 [MS Thesis].
[37] Lee KY, Kim MH. Experimental and empirical study of steam condensation heat transfer with a noncondensable gas in a small-diameter vertical tube. Nucl Eng Des 2008;238:207–16. https://doi.org/10.1016/j.nucengdes.2007.06.021.
[38] Smith S, Lasdon L. Solving large sparse nonlinear programs using GRG. ORSA J Comput 1992;4:2–15.
[39] Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. Lightgbm: a highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst 2017;30:3146–54.
[40] Awad M, Khanna R. Efficient learning machines: theories, concepts, and applications for engineers and system designers. Springer 2015:67–80. https://doi.org/10.1007/978-1-4302-5990-9_4.
[41] Popescu MC, Balas VE, Perescu-Popescu L, Mastorakis N. Multilayer perceptron and neural networks. WSEAS Trans Circ Syst 2009;8:579–88.
[42] Segal M.R. Machine learning benchmarks and random forest regression. Tech Rep 2004.
[43] Hans C. Elastic net regression modeling with the orthant normal prior. J Am Stat Assoc 2011;106:1383–93. https://doi.org/10.1198/jasa.2011.tm09993.
[44] Xu M, Watanachaturaporn P, Varshney PK, Arora MK. Decision tree regression for soft classification of remote sensing data. Remote Sens Environ 2005;97:322–36. https://doi.org/10.1016/j.rse.2005.05.008.
[45] Taki M, Rohani A, Soheili-Fard F, Abdeshahi A. Assessment of energy consumption and modeling of output energy for wheat production by neural network (MLP and RBF) and gaussian process regression (GPR) models. J Clean Prod 2018;172:3028–41. https://doi.org/10.1016/j.jclepro.2017.11.148.
[46] Bergstra J, Komer B, Eliasmith C, Yamins D, Cox DD. Hyperopt: a python library for model selection and hyperparameter optimization. Comput Sci Discov 2015;8:014008. https://doi.org/10.1088/1749-4699/8/1/014008.
[47] Anguita D, Ghelardoni L, Ghio A, Oneto L, Ridella S. The 'K' in K-fold cross validation. In: Proceedings of the 20th International Conference on Neural Information Processing; 2012.
[48] Ozaki Y, Tanigaki Y, Watanabe S, Onishi M. Multiobjective tree-structured parzen estimator for computationally expensive optimization problems. Proc Genet Evol Comput Conf 2020:533–41. https://doi.org/10.1145/3377930.3390234.
[49] Kim JW, Lee YG, Ahn HK, Park GC. Condensation heat transfer characteristic in the presence of noncondensable gas on natural convection at high pressure. Nucl Eng Des 2009;239:688–98. https://doi.org/10.1016/j.nucengdes.2008.12.011.