International Journal of Heat and Mass Transfer 231 (2024) 125827

Contents lists available at ScienceDirect

International Journal of Heat and Mass Transfer

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

Utilization of XGBoost algorithm to predict dryout incipience quality for saturated flow boiling in mini/micro-channels

Hyeonseok Noh [a], Seunghyun Lee [a,*], Sung-Min Kim [b], Issam Mudawar [c]

[a] Two-Phase Flow and Thermal Management Laboratory, School of Mechanical Engineering, Gwangju Institute of Science and Technology, 123 Cheomdangwagi-ro, Buk-gu, Gwangju, 61005, South Korea
[b] School of Mechanical Engineering, Sungkyunkwan University, 300 Cheoncheon-dong, Suwon 16419, South Korea
[c] Purdue University Boiling and Two-Phase Flow Laboratory (PU-BTPFL), School of Mechanical Engineering, 585 Purdue Mall, West Lafayette, IN 47907, USA

ARTICLE INFO

Keywords:
mini/micro-channels
heat transfer coefficient
saturated flow boiling
dryout incipience
machine learning
extreme gradient boosting
XGBoost

ABSTRACT

Critical heat flux (CHF) is arguably the most important design parameter for applications involving cooling of high heat flux devices. Despite prior limited successes, developing accurate predictive tools for CHF remains quite challenging, but key identifiable trigger mechanisms for CHF are intermittent liquid film dryout, complete liquid film dryout, and departure from nucleate boiling (DNB). For saturated flow boiling in mini/micro-channels, dryout incipience is a necessary condition for commencement of liquid film dryout, which is marked by a substantial deterioration in the heat transfer coefficient. The primary objective of this study is to predict system parameters for dryout incipience quality using machine learning, relying on a massive database for this parameter. Predictions are achieved using eXtreme Gradient Boosting (XGBoost), one of the supervised ensemble machine learning methods. This technique is applied to the Purdue University Boiling and Two-Phase Flow Laboratory (PU-BTPFL) consolidated database for dryout incipience quality for saturated flow boiling in mini/micro-channels. This database comprises 997 datapoints amassed from 26 sources and encompasses 13 different working fluids, hydraulic diameters from 0.51 to 6.0 mm, mass velocities from 29 to 2303 kg/m^{2}s, liquid-only Reynolds numbers from 125 to 53,770, boiling numbers from 0.31×10^4 to 44.3×10^4, and reduced pressures from 0.005 to 0.78. *Optuna*, a supervised automated hyper-parameter optimization software, is used to set the best hyper-parameters in the learning process based on the consolidated database. A part of consolidated database is used to train the XGBoost algorithm, and the XGBoost machine, consisting of specific input parameters, is developed with appropriately determined hyper-parameter set after optimization. The trained XGBoost machine is shown to provide remarkable accuracy in predicting the dryout incipience quality, evidenced by a mean absolute error (MAE) of 2.45% and mean absolute deviation (MAD) of 3.57×10^{-2}.

1. Introduction

The last two decades have witnessed an unprecedented increase in efforts aimed at managing thermal loads from high heat flux electronic and power devices [1,2]. The pursuit of safe and effective thermal solutions is hindered by several factors, including extreme performance and structural sensitivity to operating temperature, and stringent packaging volume constraints. In response to these challenges, thermal engineers are relying increasingly on two-phase cooling methods, which capitalize upon the coolant's both sensible and latent heat to remove the heat while maintaining comparatively low device temperatures. Examples are a variety of methods developed at the Purdue University Boiling and Two-Phase Flow Laboratory (PU-BTPFL) and other international cooling centers, including capillary [3], pool boiling [4,5], falling film [6], channel flow boiling [7], mini/micro-channel [8], jet [9,10], and spray [11-13]. Among these methods, recent work has focused mostly on mini/micro-channel heat sinks, given their multitude of benefits, including high heat transfer coefficient, low manufacturing cost, and compactness [1,2], as well as suitability for space applications [14,15]. Aside from conventional heat sinks, recent efforts have shifted to advancing mini/micro-channel cooling performance using such methods as hybrid micro-channel/jet impingement [16,17], bidirectional counter-flow heat sinks [18], and manifold micro-channel heat sinks [19].

Despite the mentioned technological merits, two-phase mini/micro-channel heat sinks are not without shortcomings. A paramount challenge in application of these devices is difficulty accurately predicting

* Author to whom correspondence should be addressed; Tel. (+82) 062-715-2787; Website: https://tpftml.gist.ac.kr/tpftml/index.do
E-mail address: lees@gist.ac.kr (S. Lee).

https://doi.org/10.1016/j.ijheatmasstransfer.2024.125827
Received 31 January 2024; Received in revised form 15 May 2024; Accepted 6 June 2024
Available online 15 June 2024
0017-9310/© 2024 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Nomenclature

A	flow area
A_{ch}	channel cross-sectional area
A_h	heated surface area
argmin	minimum argument
Bd	bond number, $gD_h^2(\rho_f - \rho_g)/\sigma$
Bo	boiling number, q''_H/Gh_{fg}
C	classification rule; classifier
C^*	Bayes classifier
Ca	capillary number, $\mu_f G/(\rho_f \sigma)$
Cc	convective-confinement number, $Bd^{0.5}Re$
Co	confinement number, $[\sigma/(g(\rho_f - \rho_g)D_h^2)]^{0.5}$
c	arbitrary constant used in Eq. (15), $F_0(x) = \underset{c}{\text{argmin}} \sum_{i=1}^{n} L(y_i, c)$
D	diameter; terms used in Eq. (12) for loss function
D_{eq}	equivalent diameter
D_h	hydraulic diameter
E	Entropy
err	error of weak learner
Fr	Froude number, $G^2/\rho^2 gD_h$
$Fr_{g,Mori}$	Froude number used by Mori et al., $Fr_{g,Mori} = G^2/(\rho_g(\rho_f - \rho_g)gD_{eq})$
Fr^*	modified Froude number
F_m	combined weak leaner
F_M	final leaner
f	frictional factor
f_m	individual weak learner from an independent m-th tree
G	mass velocity; Gini index; $G_j = \sum_{j \in J_M} g_j$
g	gravitational acceleration; first order gradient
H	$H_j = \sum_{j \in J_M} h_j$
h	Hessian; second order gradient
h_{fg}	latent heat of vaporization
h_{tp}	two-phase heat transfer coefficient
I	identity matrix
IG	information gain
i	index for i-th sample in database
inf	infimum
J_m	total number of terminal nodes (or leaves)
j	index of terminal node (or leaf)
L	loss function
L_{gain}	gain function
L_Ω	regularized loss function
$\overline{L}$	average loss
$\widetilde{L}_\Omega$	simplified regularized loss function
l_r	learning rate
M	number of combined learner; total number of trees
m	index of trees
N	number of data
P	pressure
P_{crit}	critical pressure
P_F	wetted perimeter of channel
P_H	heated perimeter of channel
P_R	reduced pressure, P/P_{crit}
p	dimension of feature
p_i	probability of selecting data in class i
Q	heat power
q''	heat flux
q''_{CHF}	critical heat flux
q''_H	heat flux based on heated perimeter of channel
R	terminal node region
Re	Reynolds number, GD_h/μ
$\mathbb{R}^p$	set of real numbers having dimension of p
r	pseudo residual
S	dataset for entropy calculation
Su	Suratman number, $\sigma\rho D_h/\mu^2$
T	temperature
V	velocity
We	Weber number, $G^2 D_h/\rho\sigma$
w	leaf weight; sample weight
w^*	optimal weight
x_{crit}	dryout completion (CHF) quality
x_{de}	dryout end quality
x_{di}	dryout incipience quality
$\mathbf{x}_i$	i^{th} sample input
y	target feature
y_i	i^{th} sample output
$\bar{y}$	mean value of experimental data

Acronyms

ANN	artificial neural network
API	application programming interface
CART	classification and regression trees
CHF	critical heat flux
CNN	convolutional neural network
DNB	deviation from nucleate boiling
DNN	deep neural network
DWO	density wave oscillations
ID3	Iterative Dichotomiser 3
PCI	parallel channel instability
PDO	pressure drop oscillations
PFI	permutation feature importance
MAD	mean absolute deviation
$\overline{MAD}$	averaged mean absolute deviation
MAE	relative mean absolute error
MSE	mean squared error

Greek symbols

α	amount of say
β	measure of confidence
γ	regularization constant; the minimum loss reduction
ε	smallest value used in Eq. (22)
θ	percentage predicted within ±30%
λ	regularization factor
λ_{K-H}	critical Kelvin-Helmholtz instability
μ	dynamic viscosity
ξ	percentage predicted within ±50%
ρ	density
ρ_r	density ratio, ρ_g / ρ_f
σ	surface tension; standard deviation
ϕ_m	minimized objective function
Ω	equilibrium frequency of phase change; regularization term

Subscripts

ch	channel
$crit$	critical
de	dryout end
di	dryout incipience
eq	equilibrium
exp	experimental (measured)
f	saturated liquid
fo	liquid only
g	saturated vapor
go	vapor only

in	inlet	*r*	ratio
K-H	Kelvin-Helmholtz	*sat*	saturation
pred	predicted	*w*	wall
R	reduced		

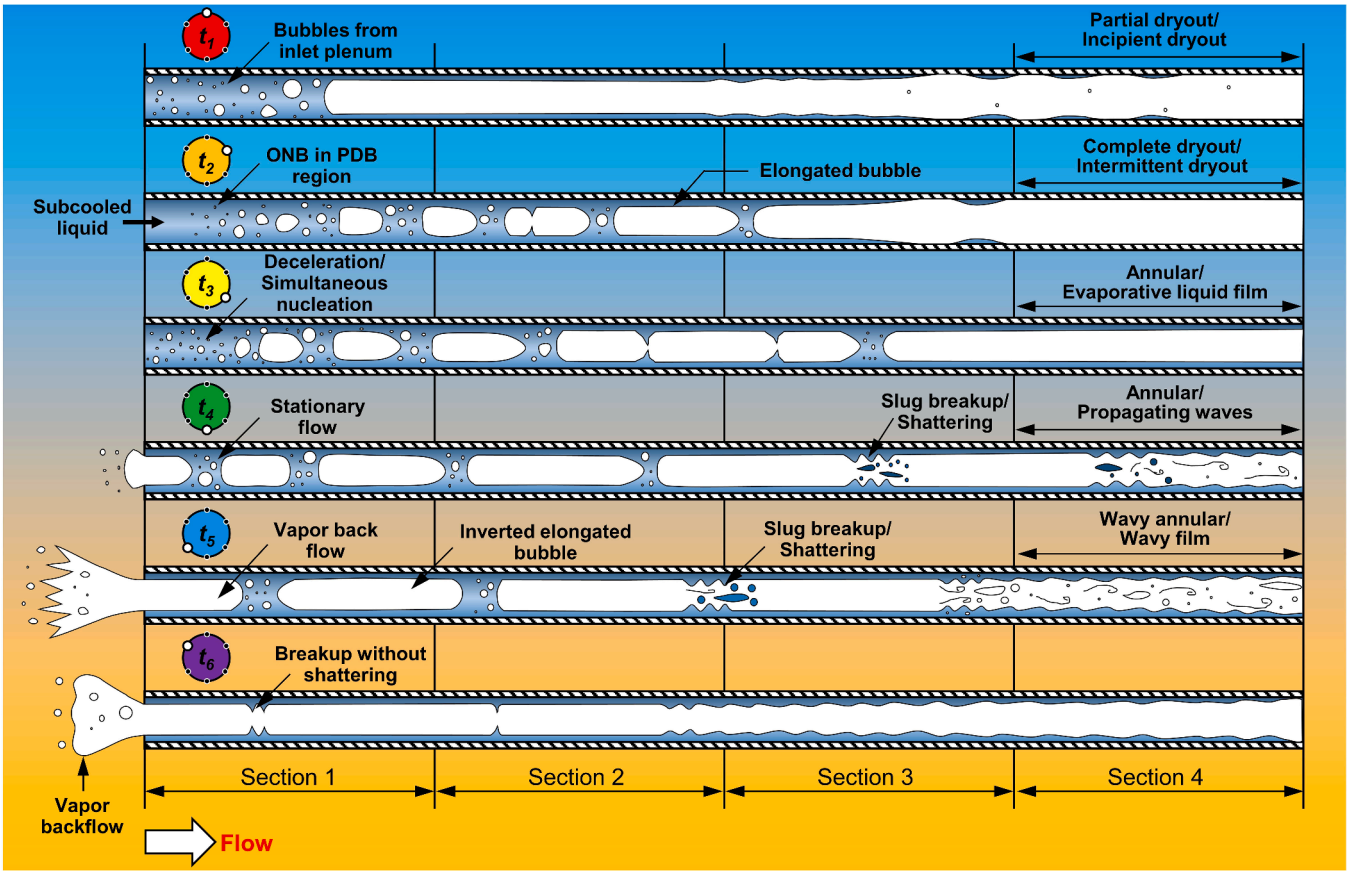


Fig. 1. Schematic renderings of transient flow patterns observed at six different times within a single periodic cycle: t_1: beginning of the cycle, t_2: forward liquid advance, t_3: rapid bubble growth, t_4: commencement of transition pattern, t_5: commencement of wavy annular pattern, and t_6: end of liquid deficient period. The indicated sections are as follows: 1: subcooled boiling dominant, 2: combined saturated nucleate and convective boiling, 3: saturated convective boiling by annular film evaporation, and 4: intermittent dryout dominant region [31].

cooling performance, which stems from such complicating two-phase phenomena as transient flow behavior [20,21], choking [22-24], flow instabilities [25-32], and perhaps most importantly, critical heat flux (CHF) [33], which can take the form of either *departure from nucleate boiling* (DNB) [34-36] or *dryout* [37-39].

Dryout causes a precipitous rise in the heating surface temperature and eventual thermal failure and is therefore considered one of the most critical design parameters for heat-flux-controlled surfaces [24]. Depending on its trigger mechanisms, dryout can be categorized into one of two types: *static* and *dynamic*. For static dryout, the interfacial shear stress induced by the velocity difference between liquid and vapor phases and ensuing Kelvin-Helmholtz instability induce a wavy pattern in the annular regime, causing partial dryout in the wave troughs [34]. Once all the wall liquid film is evaporated, it is called complete dryout, which usually happens to flows with high vapor quality in the downstream region. For dynamic dryout, the flow incurs different forms of instability, including density-wave oscillation (DWO), parallel-channel instability (PCI), and pressure drop oscillation (PDO), which can all culminate in the intermittent dryout [31]. The process of intermittent dryout resulting from flow instability is described in Fig. 1 [31], wherein and the periodic flow induces vapor back flow, which is a primary reason for the annular liquid film dryout downstream. Further details of how flow instability influences the heat transfer process in mini/micro-channel heat sinks are provided in the present authors' prior work. The intermittent dryout takes different forms depending on the flow geometry and operating conditions, and given its complicated trigger mechanisms, its impact on the system's thermal performance is highly unpredictable [27]. The *dryout incipience*, accompanied by a large reduction in the heat transfer coefficient signals the beginning of the dryout, and development of an accurate predicting tool for its occurrence is essential to the design of high heat flux devices [38].

Dryout in saturated flow boiling in mini/micro-channels is closely tied to the various complicated hydrodynamic and thermal transport phenomena occurring in the annular flow regime. The mechanisms of dryout occurrence in the *nucleate boiling dominant heat transfer* and the *convective boiling dominant heat transfer* are depicted in Fig. 2(a) and 2(b), respectively [38]. While both heat transfer modes can occur simultaneously along the channel, one or the other would typically dominate depending on operating conditions and channel length. For the nucleate boiling dominant heat transfer mode, heat transfer is dominated by bubble nucleation and departure from the heated wall. Here, higher heat transfer coefficients are observed in the upstream region, where active

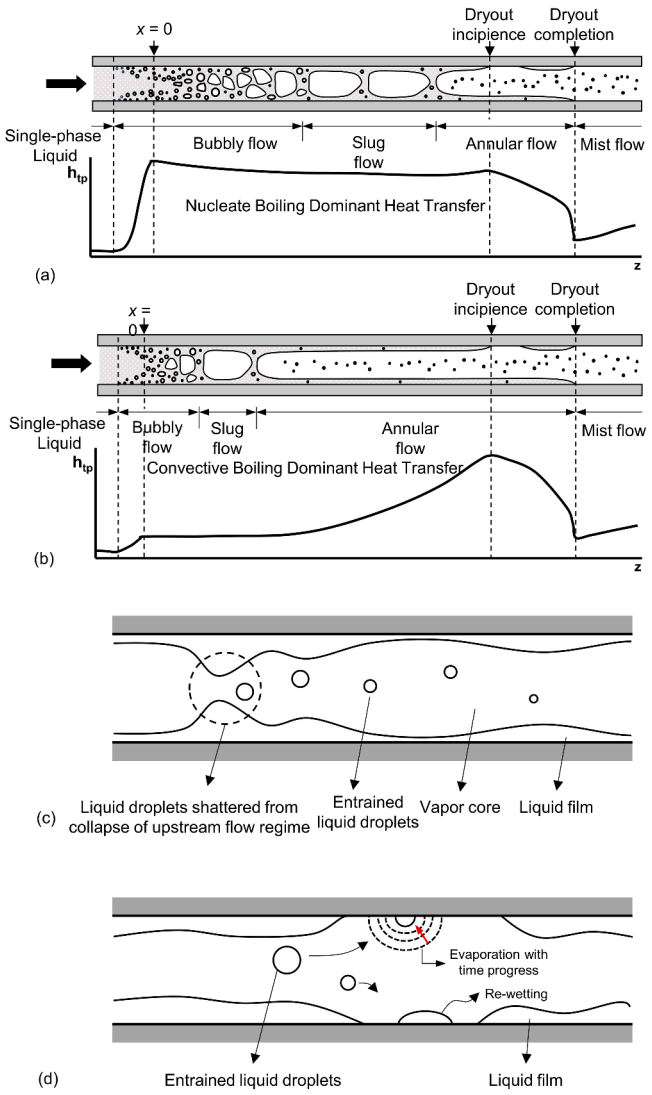


Fig. 2. Schematics of axial variation of flow regime and heat transfer coefficient along uniformly heated mini/micro-channel with wall dryout for (a) nucleate boiling dominant heat transfer [38], (b) convective boiling dominant heat transfer [38], and detailed schematic depictions of (c) shattering [40] and (d) re-wetting by collision of entrained liquid droplets in partial dryout region.

nucleation and bubbly flow prevail. This is followed by a gradual decrease in the heat transfer coefficient towards the downstream as the flow pattern transition into slug and annular wherein the nucleation is suppressed. For the convective boiling dominant heat transfer mode, which is typically encountered in relatively low heat flux scenarios, transition to annular flow is shifted upstream, causing much of the heat to be conducted across the thin liquid film before being released to the flow by interfacial evaporation.

Dryout can be initiated by uneven evaporation or partial breakup of the wavy liquid film in the annular flow regime; both of which are outcomes of the Kelvin-Helmholtz instability resulting from large velocity difference between the vapor core and the liquid film. Also associated with the velocity difference is high interfacial shear which induces liquid shattering of the wave crests and entrainment of liquid droplets, especially where flow transitions from slug to annular, as depicted in Fig. 2(c) [40]. Within the *partial dryout region* preceding the location of complete dryout, the heat transfer coefficient is compromised by a gradual increase in the heated area exposed directly to the vapor core. Nonetheless, the heat transfer is maintained to a certain extent by a combination of evaporation through the partial liquid film and wall rewetting by collision of entrained liquid droplets, as depicted in Fig. 2(d).

Prior attempts to develop predictive tools for dryout incipience quality have been based on two primary approaches: empirical correlations and analytical models [41-48]. The empirical correlations were developed mostly through reliance on dimensionless groups deemed as having direct impact on the dryout process. For relatively larger (macro) horizontal channels, Mori et al. [49], based on experimental observations, showed that the dryout process is circumferentially asymmetric. They explained that the liquid film has circumferentially varying thickness due to gravity, and begins to dry out from the top, where the film is thinnest, indicative of dryout incipience, and progresses downward to the bottom, resulting in dryout completion. For mini/micro-channels, circumferential variations in the film thickness are comparatively insignificant, leading to a dryout process that is different from that proposed by Mori et al. Instead, the dryout phenomenon was empirically investigated relative to variations in the saturation temperature and heat flux by employing the dimensionless numbers of P_R and *Bo*, respectively [44,46-48]. However, these correlations are limited to few fluids (mostly CO_2 and a few refrigerants) and narrow ranges of operating parameters. Kim and Mudawar [38] tackled these limitations by developing a universal correlation that is based on a large database spanning many fluids as well as broad ranges of relevant operating conditions.

As to the analytical models for dryout, Lavin and Young [50] suggested a mechanistic concept for transition from annular flow to the dryout region for R12 and R22 based on a relationship between Weber number and vapor Reynolds number. Kattan [51] suggested a different method for predicting the transition between annular and mist flows by employing the Steiner version of the Taitel-Dukler “A-M” curve. However, both approaches have limited application when attempting to predict dryout in mini/micro-channels.

Attempts to utilize *machine learning* techniques to analyze and predict the performance of thermal systems are quickly gaining momentum and wide acceptance [52-54]. Among these methods, *artificial neural network* (ANN) has been especially popular for modeling thermal management systems, including predicting building energy consumption [52,55], heat removal rate and temperature distribution of heat exchangers [53,54,56], and battery temperature of electric vehicles [57,58]. This method has been used successfully to predict single-phase heat transfer coefficients [59,60], however, there have also been successes in modeling two-phase flow in mini/micro-channels, including both boiling and condensation [61,62]. Other types of neural networks which have been used to predict heat transfer coefficients for flow boiling and condensation in mini/micro-channels including *deep neural networks* (DNN) and *convolutional neural networks* (CNN) [63,64]. Yet another example is use of *ensemble methods* [65,66], which aim to improve the accuracy of predictions by combining multiple models instead of using a single model, and *boosting methods*, a category of the ensemble methods [67], which combine a set of weak learners into a stronger learner to minimize training errors. Among the different boosting methods, XGBoost has recently shown excellent capability in predicting the heat transfer coefficient for flow boiling and condensation in mini/micro-channels, surpassing the performances of the neural network counterparts [62,68]. It is for this reason that XGBoost is considered the most suitable machine learning method for predicting complex two-phase flow phenomena in mini/micro-channels, including dryout incipience [69].

The primary goal of this study is to employ the boosting method to predict the dryout incipience quality for saturated flow boiling in mini/micro-channels. A total of 997 dryout incipience quality datapoints amassed from 26 sources are prepared for training and validation. Assessment of previous empirical correlations in predicting the experimental dryout incipience quality will be conducted to select the best performing correlation. The data training will be conducted using three different boosting algorithms after determining the optimal set of input

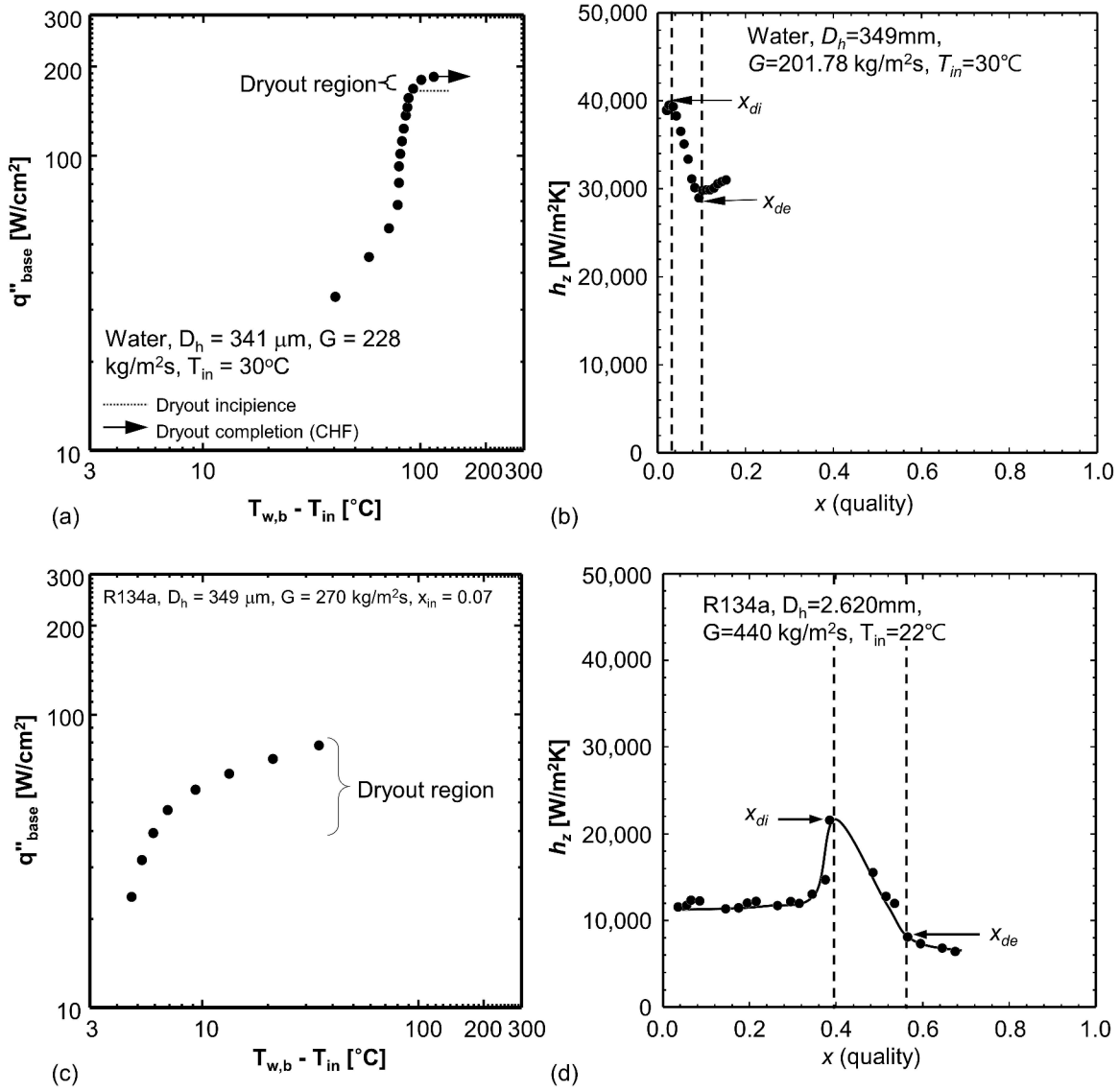


Fig. 3. Boiling curves and axial variation of heat transfer coefficient for (a) water [72] and (b) water [79], (c) R134a [98] and (d) R134a [99] flows in rectangular micro-channels.

features and corresponding optimal set of hyper-parameters through feature importance index and iterative trial-and-run process, respectively. The accuracies of trained boosting machines will be guaranteed by a validation test using randomly excluded data. A comparison between the selected boosting algorithm and empirical correlations will be performed to demonstrate the superior performance of the machine learning methodology.

2. Modeling method

2.1. Consolidated universal database of dryout incipience quality for flow boiling in mini/micro-channels

The *Consolidated Database* employed in this study consists of 997 datapoints for dryout incipience quality amassed from 26 sources. Three different types of quality are considered: x_{crit}, x_{di}, and x^*_{di}, distinguished by the manner with which they are identified. The first type, x_{crit}, corresponds to the dryout completion point where the liquid film completely dries out while minor cooling is sustained by the collision of liquid droplets entrained in the ensuing mist flow. This occurrence is accompanied by a sudden rise in the wall temperature above the saturation temperature and a large decrease in the heat transfer coefficient [70]. The dryout completion indicated by x_{crit} is essentially one of the three identifiable types of CHF: subcooled DNB, saturated DNB, and complete dryout [71]. On the other hand, x_{di} corresponds to incipient dryout associated with partial breakup of the liquid film exposing the wall directly to the vapor on a local basis and resulting in a wall temperature rise that is milder than that resulting from complete dryout. However, especially for water, there is a sharp transition from x_{di} to x_{crit} leading to rather similar values for both as the dryout completion follows closely the dryout incipience [73-78], as depicted in Fig. 3(a). Fig. 3(b) shows the variation of heat transfer coefficient corresponding to this sharp transition. The primary reason for this sharp transition is the high surface tension for water resists the shattering of liquid from the film's interface and therefore greatly diminishes the contribution of entrained droplets, causing the observed large and sudden drop in the heat transfer coefficient associated with CHF [72]. Among the 997 datapoints comprising the present Consolidated Database, 664 correspond to x_{crit} (amassed from 6 sources), all of which being for water and identified by a temperature rise from T_{sat} by 5 to 150°C [73-78].

Table 1
Consolidated Database for dryout incipience quality for saturated flow boiling in mini/micro-channel used to develop XGBoost machine. [38].

Author(s)	Channel Geometry	D_h[mm]	Fluid(s)	G[kg/m^{2}s]	Data points	Remarks
Ali and Palm (2011) [70]	C, S, V	1.22, 1.70	R134a	50 – 600	23	x_{di} identified by change of slope in boiling curve, and wall temperature
Baek and Chang (1997) [73]	C, S, V	6.0	Water	29 – 277	232	x_{crit} indentified by fast increase of T_w when T_w>250ºC
Becker (1970) [74]	C, S, V	2.4, 3.0	Water	365 – 2725	82	x_{crit} identified by fast increase of T_w
Del Col and Bortolin (2012) [81]	C, S, H	0.96	R134a, R245fa, R32	101 – 902	43	x_{di} identified by wall temperature
Ducoulombier et al. (2011) [47]	C, S, H	0.529	CO_2	200 – 1410	48	x_{di} identified by falling off of h_{tp}
Greco (2008) [82]	C, S, H	6.0	R134a, R22, R407C, R410A	199 – 1079	7	x_{di}* identified by falling off of h_{tp}
Hihara and Dang (2007) [83]	C, S, H	1.0, 2.0, 4.0, 6.0	CO_2	360 – 1440	16	x_{di} identified by falling off of h_{tp}
Karayiannis et al. (2012) [84]	C, S, V	1.1	R134a	300	3	x_{di} identified by falling off of h_{tp}
Kim et al. (2000) [75]	C, S, V	6.0	Water	99 – 277	210	x_{crit} identified by fast increase of T_w with T_w increase rate of 50 ºC/s
Lezzi et al. (1994) [76]	C, S, H	1.0	Water	776 – 2738	68	x_{crit} identified by fast increase of T_w when T_w>250 ºC
Li et al. (2012) [85]	C, S, H	2.0	R1234yf, R32	100 – 400	8	x_{di}* identified by falling off of h_{tp}
Martin-Callizo (2010) [86]	C, S, V	0.64	R134a, R22, R245fa	185 – 541	42	x_{di} identified by change of slope in boiling curve, and wall temperature
Mastrullo et al. (2012) [48]	C, S, H	6.0	CO_2, R410A	150 – 501	28	x_{di} identified by falling off of h_{tp}
Oh and Son (2011a) [87]	C, S, H	1.77, 3.36, 5.35	R134a, R22	200 – 400	6	x_{di}* identified by falling off of h_{tp}
Oh and Son (2011b) [88]	C, S, H	4.57	CO_2	600 – 900	8	x_{di}* identified by falling off of h_{tp}
Oh et al. (2011) [89]	C, S, H	1.5, 3.0	R22, R410A, R290	100 – 500	9	x_{di}* identified by falling off of h_{tp}
Ohta et al. (2009) [90]	C, S, H	0.51	FC72	107, 205	2	x_{di}* identified by falling off of h_{tp}
Roach et al. (1999) [77]	C, S, H	1.168, 1.448	Water	256 – 1037	42	x_{crit} identified by fast increase of T_w when T_w>250 ºC
Saitoh et al. (2005) [91]	C, S, H	0.51, 1.12, 3.1	R134a	150 – 300	41	x_{di}* identified by falling off of h_{tp}
Shiferaw (2008) [92]	C, S, H	1.1, 2.88, 4.26	R134a	200 – 400	13	x_{di}* identified by falling off of h_{tp}
Tibirica et al. (2012) [93]	C, S, H	1.0	R1234ze	300 – 600	4	x_{di}* identified by falling off of h_{tp}
Wang et al. (2009) [94]	C, S, H	1.3	R134a	321 – 676	9	x_{di}* identified by falling off of h_{tp}
Wu et al. (2011) [95]	C, S, H	1.42	CO_2	300 – 600	18	x_{di}* identified by falling off of h_{tp}
Yang and Fujita (2002) [96]	R, S, H	0.976	R113	100, 200	3	x_{di}* identified by falling off of h_{tp}
Yu et al. (2002) [78]	C, S, H	2.98	Water	50 – 151	30	x_{crit} identified by fast increase of T_w
Yun et al. (2005) [97]	R, M,H	1.14	CO_2	300, 400	2	x_{di}* identified by falling off of h_{tp}
Total					997	

* R: rectangular, C: circular.
** S: single-channel, M: multi-channel.
*** H: horizontal-channel, V: vertical-channel.
**** x_{crit}: critical quality data reported by original authors.
***** x_{di}: dryout incipience quality data reported by original authors.
****** x_{di}*: dryout incipience quality data identified by present authors by falling off in measured two-phase heat transfer coefficient attributed by original authors to dryout incipience.

Aside from this special case for water, x_{di} is generally associated with relatively mild and gradual reduction in the heat transfer coefficient. Here, fast the moving vapor core tends to amplify waviness of the liquid film, causing partial film breakup in the wave troughs. This process also involves shattering of liquid droplets, especially during the transition from slug to annular flow [80]. Within the Consolidated database, 333 datapoints (amassed from 20 sources) correspond to x_{di} and include refrigerants and CO_2, all of which having low surface tension [47,48,70,81-97]. The low surface tension for this fluid category is crucial to both the pronounced film waviness and easier shattering of liquid droplets. Of the x_{di} subset, 203 datapoints from 7 sources [47,48,70,81,83,84,86] were identified as dryout incipience quality by previous investigators based on three different criteria: boiling curve slope change, reduction of heat transfer coefficient, and temperature fluctuation. With inferior thermophysical properties and low CHF compared to those for water, fluids belonging to this subset exhibit a very gradual increase in slope of the boiling curve and broad dryout region, as shown in Fig. 3(c). This gradual transition is manifest in Fig. 3(d) where the heat transfer coefficient is shown decreasing over a broad range of qualities between x_{di} and x_{de}. As to relying on temperature fluctuation as indictor for dryout incipience, different ranges have been reported, 3 to 20°C [86], 5 to 20°C [70], and 1 to 3°C [81]. Of the 333 dryout incipence datapoints, 130 (330 minus 203) were identified by the present authors from 13 sources [82,85,87-97] that were missed by previous researchers. These datapoints are identified as dryout incipience quality x^*_{di} based on reduction of the heat transfer coefficient.

Table 1 provides detailed information on the Consolidated Database with its 997 datapoints spanning the following fluids and operating

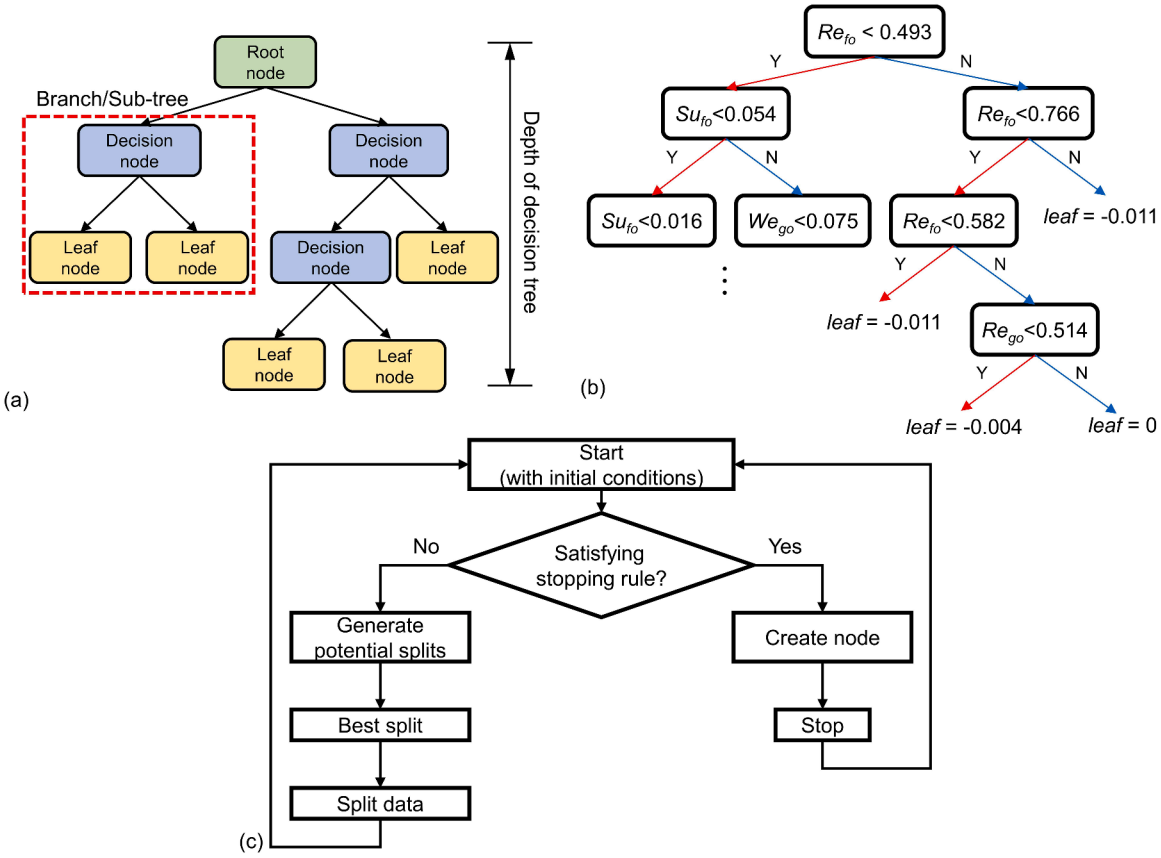


Fig. 4. (a) The simple structure of decision tree, (b) the structure of decision tree for regression, and (c) the decision tree algorithm.

conditions:

- Working fluid: CO_2, FC72, R1234yf, R1234ze, R113, R134a, R22, R245fa, R290, R32, R407c, R410a, and water
- Hydraulic diameter: $0.51 < D_h < 6.0$ mm
- Mass velocity: $29 < G < 2738$ kg/m^{2}s
- Heat flux based on heated area: $0.4666 \leq q''_H \leq 513.0$ W/cm^2
- Liquid-only Reynolds number: $125 \leq \mathrm{Re}_{fo} = GD_h/\mu_f \leq 76{,}506$
- Vapor-only Reynolds number: $4{,}720 \leq \mathrm{Re}_{go} = GD_h/\mu_g \leq 498{,}947$
- Boiling number: $0.00003 \leq Bo = q''_H/Gh_{fg} \leq 0.00443$
- Liquid-only Weber number: $0.09 \leq We_{fo} = \frac{G^2 D_h}{\rho_f \sigma} \leq 5167$
- Vapor-only Weber number: $14.99 \leq We_{go} = \frac{G^2 D_h}{\rho_g \sigma} \leq 30{,}507$
- Liquid-only Suratman number: $35{,}935 \leq Su_{fo} = \frac{\sigma \rho_f D_h}{\mu_f^2} \leq 9{,}841{,}592$
- Equilibrium frequency of phase change [100]: $0.10 \leq \Omega = \frac{q''_H P_H}{A_{ch} h_{fg}} \frac{\rho_f - \rho_g}{\rho_f \rho_g} \leq 2546.32$
- Critical Kelvin-Helmholtz instability wavelength [101]:

$$0.00291 \leq \lambda_{K-H} = \frac{2\pi}{\left[\frac{(\rho_f - \rho_g)g}{\sigma}\right]^{1/2}} \leq 0.01571 \text{ m}$$

2.2. Machine learning model: Decision tree

The decision tree, one of the supervised learning-based methodologies, is a tree-like model that predicts or classifies the target feature by utilizing its relations to input features [102]. As shown in Fig. 4(a), it features nodes, branches, and depth. The tree structure grows by branching out from decision node (parent) to leaf nodes (children) in the direction of reduced decision tree impurity, which is evaluated by two representative algorithms: ID3 (Iterative Dichotomiser 3) and CART (Classification and Regression Trees). For ID3, the degree of impurity is estimated by the entropy,

$$E(S) = \sum_{i=1}^{N} p_i \times I(\mathbf{x}_i) = -\sum_{i=1}^{N} p_i \times \log_2 p_i, \quad (1)$$

where S is the current dataset for entropy calculation, p_i the probability of selecting data in class i, $I(\mathbf{x}_i)$ the amount of information of class i, $\mathbf{x}_i$the dataset of class i, and N the number of classes in dataset of S [103]. The dataset S is split into different subsets by the attribute A, which yields the highest information gain,

$$IG = E(S) - E(S|A), \quad (2)$$

where $E(S|A)$ is total entropy, after splitting the parent data by attribute A in children nodes. This process is iteratively conducted under the condition of positive information gain, and the branching is terminated when either the information gain becomes negative with all unused attributes or the entropy of current node becomes zero. When the terminating condition is met, the decision node converts into leaf node as shown in Fig. 4(a). However, the ID3 algorithm can only be used to classify the categorical data but not conduct the regression analysis.

The CART algorithm, originally proposed by Breiman et al. [104], uses Gini index to evaluate the system impurity in the data classification process. The Gini index is calculated according to

$$G = 1 - \sum_{i=1}^{N} p_i^2, \quad (3)$$

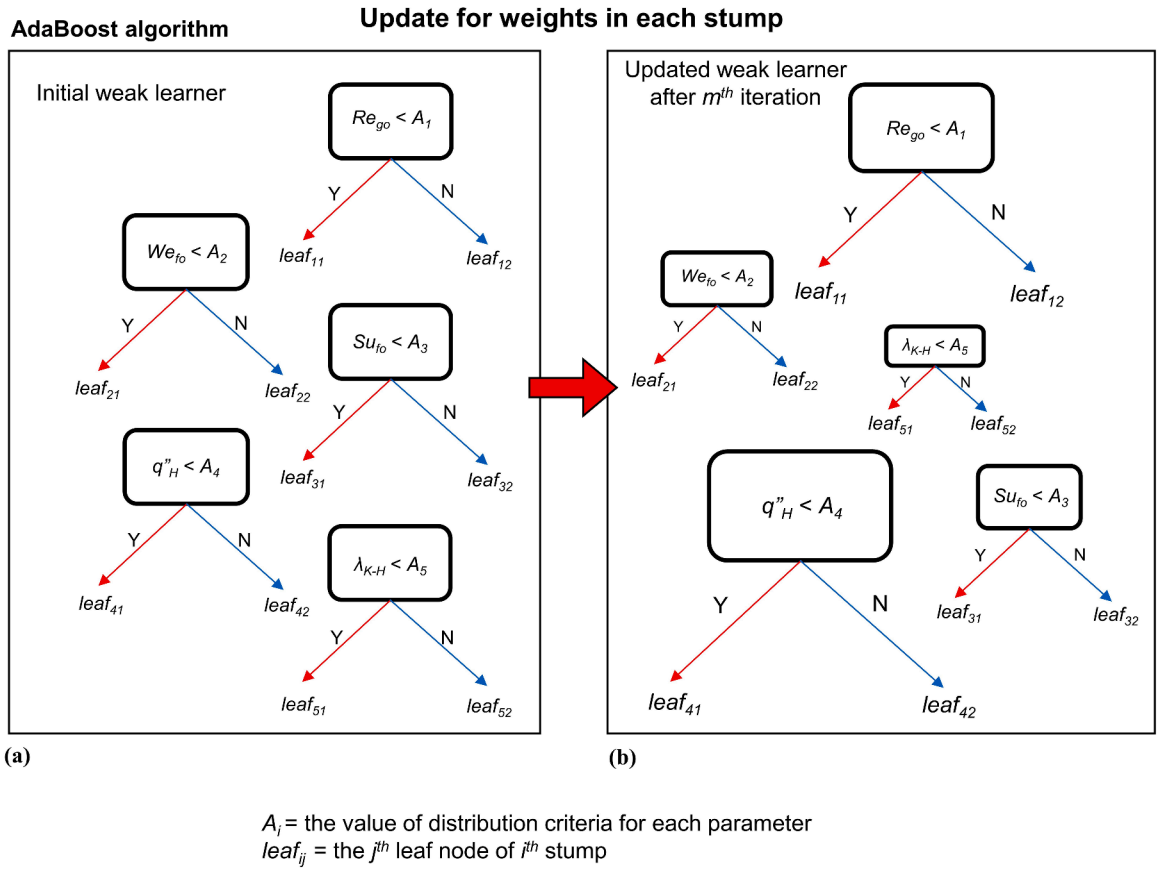


Fig. 5. The schematic of Adaboost for regression: (a) initial weak learner and (b) m^{th} iterative learner.

where p_i is the probability of selecting data in class i and N the number of classes. Among evaluated unused attributes, one yielding the lowest Gini index is selected for splitting dataset S with new children nodes with selected attribute. For the regression analysis, the CART algorithm finds the optimal point for data splitting that yields the lowest mean squared error,

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(F_M(\mathbf{x}_i) - \bar{y})^2 \tag{4}$$

where $F_M(\mathbf{x}_i)$ is the predicted value for each data point, $\bar{y}$ the mean value in split data, and N the number of split data. The working principle of CART algorithm is schematically illustrated in Fig. 4(b). In general, various decision tree algorithms, including ID3 and CART, can be customized by adding splitting criteria, stopping rules, pre-setting depth of tree, and limiting the number of nodes as shown in Fig. 4(c).

Growing of decision tree structure does not always guarantee an enhancement in the predictive performance of a trained machine. This is mostly caused by over-fitting, which occurs when a statistical model fits exactly against its training data but decreases the predictive accuracy for unseen data. This over-fitting problem can be prevented by either pre- or post-pruning processes in which unnecessary branches are removed. In the pre-pruning process, unnecessary branches are prevented from being created if specific parameters exceed pre-set limit values. In the post-pruning process, pruning is based on two criteria: complexity of the calculation and error rate. The complexity of the calculation is related to the increase in computational load attributing to the excessive number of nodes and branches, and it is usually caused by a large number or wide data range of given attributes. The error rate refers to the classification or regression performance of the decision tree. Overall, the decision tree is structured by the process of growing and pruning, which are the processes of branching off into the next nodes and removing non-critical or redundant branches to reduce the size of tree-structure. The growing process stops when the predefined stopping criterion is satisfied. The pruning process removes unnecessary or uninterpretable branches in order to avoid "over-fitting" and "under-fitting" problems [105]. Additionally, the decision tree model can be either a classification or a regression tree, depending on the type of target features, whether continuous or categorical.

2.3. Machine learning algorithms

2.3.1. Adaptive boosting (Adaboost)

Adaboost is a boosting algorithm that generates a stronger learner by the linear combination of weak learners with a sample weighting factor which is updated through the successive regression processes that adjust the weights of sample data corresponding to the error of the current prediction, as shown in Fig. 5 [106]. Boosting is one of the ensemble methods that use the combination of multiple algorithms to achieve better prediction performance than relying on a single constituent algorithm, and 'adaptive' implicates that Adaboost quickly adapts to poorly predicted abnormal data and learns predictive learners in real time. The Adaboost sets higher weighting factor for data with larger errors, and new learning data is restored and used to fit the learner. As shown in Fig. 5(a), Adaboost regressor starts with the initial weak learners and then updates the learners to compensate for the weaknesses with a different measure of confidence (α, the amount of say for classifier) observed in the previous step. The first step constructing Adaboost regressor is initializing weights, w_i, of all samples by

$$w_i = 1 \ (i = 1, 2, \cdots, N), \tag{5}$$

where N is the number of samples in the training data set, $(x_1, y_1), (x_2, y_2),$

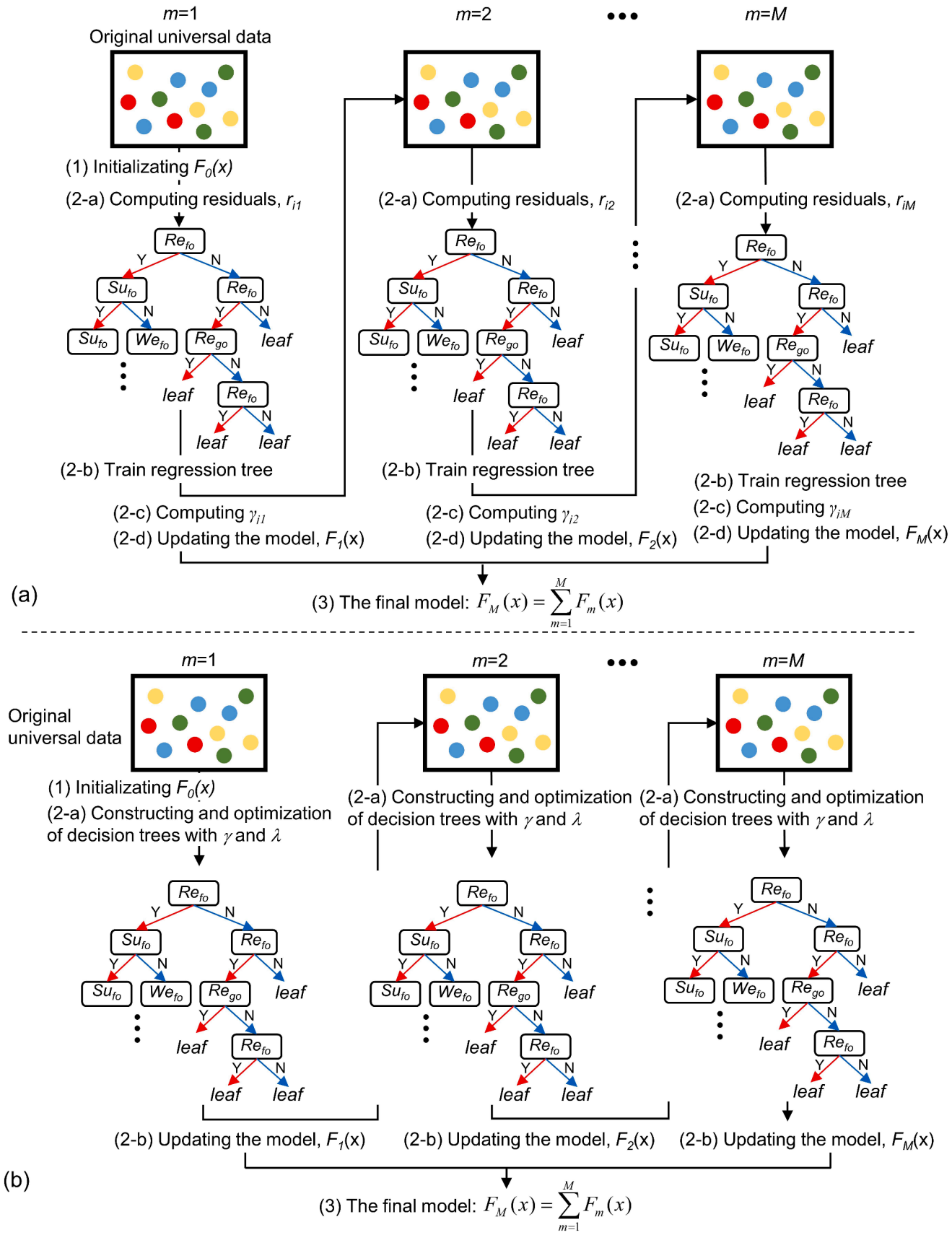
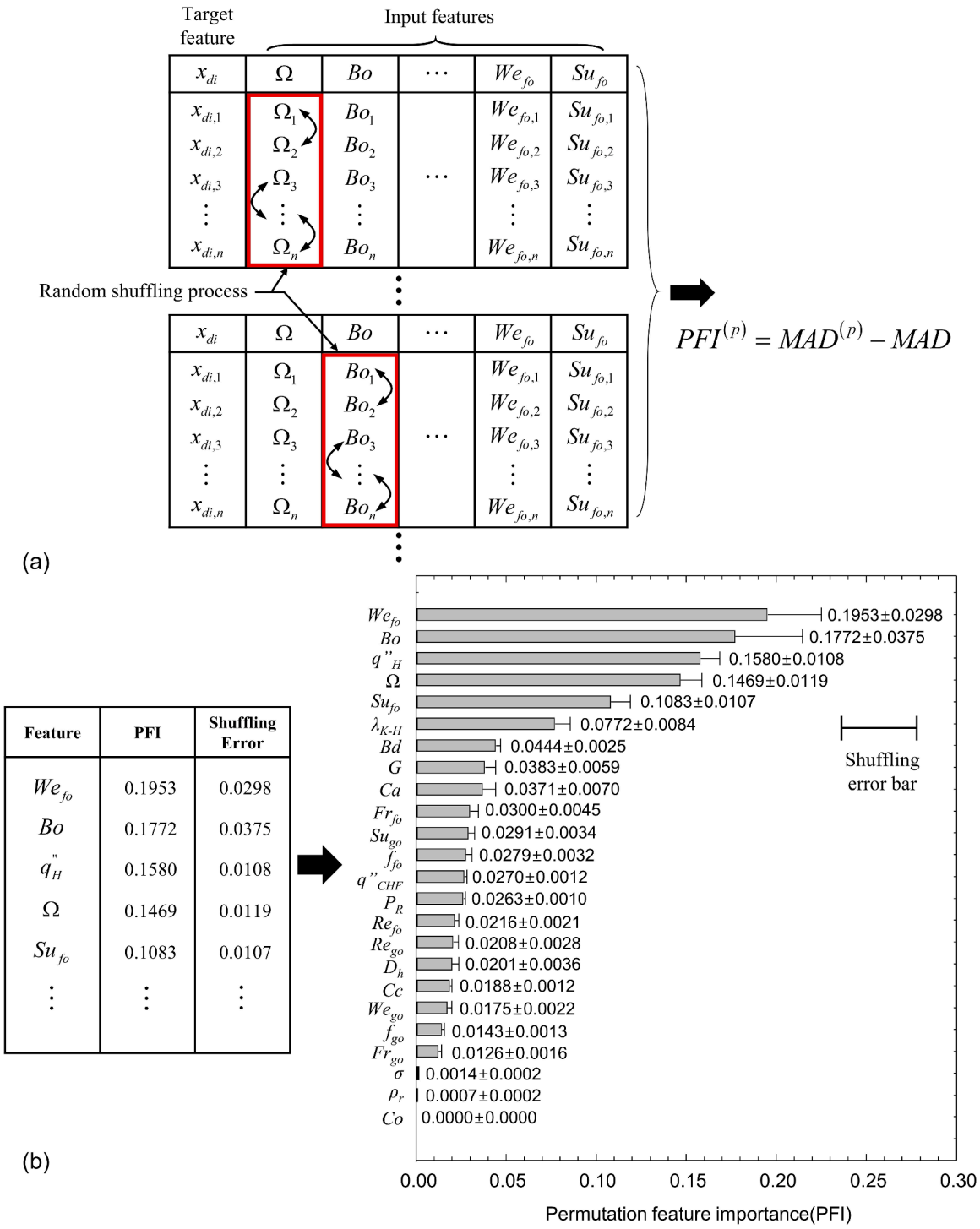


Fig. 6. The structure of machine learning algorithm for regression: (a) Gradient Boosting (GB) and (b) eXtreme Gradient Boosting (XGB).



Feature	PFI	Shuffling Error
We_{fo}	0.1953	0.0298
Bo	0.1772	0.0375
q''_H	0.1580	0.0108
Ω	0.1469	0.0119
Su_{fo}	0.1083	0.0107
⋮	⋮	⋮

Fig. 7. (a) Schematics of calculation process of permutation feature importance of input features, and (b) permutation feature importance for universal Consolidate Database sorted in descending order.

$\cdots, (x_N, y_N)$. For $m = 1$ to M, the training sequence described below is repeated until the average loss, $\overline{L} = \sum_{i=1}^{N} w_i L_i$, is less than 0.5, while updating sample weights, w_i, by

$$w_i \leftarrow w_i (\beta^{(m)})^{(1-L_i)} \ (i = 1, 2, \cdots, N), \tag{6}$$

where $\beta^{(m)}$ the measure of confidence of weak learner is expressed by

$$\beta^{(m)} = \frac{\overline{L}}{1 - \overline{L}}. \tag{7}$$

Depending on acquired $\beta^{(m)}$, the all stumps are updated to account for their importance as shown in Fig. 5(b). Here, m and M, respectively, represent the index and total number of weak learners. The loss for each sample, L_i, is calculated by

$$L_i = L[|F_m(\mathbf{x}_i) - \mathbf{y}_i|] \tag{8}$$

Table 2
Permutation feature importance (PFI) for all input features.

Feature	Symbols	Weight
Liquid only Weber number	We_{fo}	0.1953 ± 0.0298
Boiling number	Bo	0.1772 ± 0.0375
Heat flux based on heated surface	q''_H	0.1580 ± 0.0108
Equilibrium frequency of phase change	Ω	0.1469 ± 0.0119
Liquid only Suratman number	Su_{fo}	0.1083 ± 0.0107
Critical Kelvin-Helmholtz wavelength	λ_{K-H}	0.0772 ± 0.0084
Bond number	Bd	0.0444 ± 0.0025
Mass velocity	G	0.0383 ± 0.0059
Capillary number	Ca	0.0371 ± 0.0070
Liquid only Froude number	Fr_{fo}	0.0300 ± 0.0045
Vapor only Suratman number	Su_{go}	0.0291 ± 0.0034
Liquid only friction factor	f_{fo}	0.0279 ± 0.0032
Critical heat flux	q''_{CHF}	0.0270 ± 0.0012
Reduced pressure	P_R	0.0263 ± 0.0010
Liquid only Reynolds number	Re_{fo}	0.0216 ± 0.0021
Vapor only Reynolds number	Re_{go}	0.0208 ± 0.0028
Hydraulic diameter	D_h	0.0201 ± 0.0036
Convective confinement number	Cc	0.0188 ± 0.0012
Vapor only Weber number	We_{go}	0.0175 ± 0.0022
Vapor only friction factor	f_{go}	0.0143 ± 0.0013
Vapor only Froude number	Fr_{go}	0.0126 ± 0.0016
Surface tension	σ	0.0014 ± 0.0002
Density ratio	ρ_r	0.0007 ± 0.0002
Confinement number	Co	0.0000 ± 0.0000

Table 3
Default set of hyper-parameters of three boosting algorithms and their MAD for Consolidated Database of dryout incipience quality.

Type of algorithm	Hyper-parameter	Value	MAD (x 10^{-2})
XGBoost	max depth	6	4.01
	learning rate	0.3	
	n estimateors	100	
	colsample bytree	1	
	reg lambda	1	
	reg alpha	0	
	subsample	1	
	min child weight	1	
	gamma	0	
AdaBoost	n estimators	50	7.56
	learning rate	1	
Gradient Boosting	max depth	3	4.78
	min samples split	2	
	min samples leaf	1	
	max features	None	
	max leaf nodes	None	
	n estimators	100	
	learning rate	0.1	
	subsample	1	

where $F_m(\mathbf{x}_i)$ is combined weak prediction. The candidates for loss function, L, are as follows,

$$L_i = \frac{|F_m(\mathbf{x}_i) - y_i|}{D} \ (linear) \tag{9}$$

$$L_i = \frac{|F_m(\mathbf{x}_i) - y_i|^2}{D^2} \ (square) \tag{10}$$

and

$$L_i = 1 - \exp\left[\frac{-|F_m(\mathbf{x}_i) - y_i|}{D}\right] \ (exponential) \tag{11}$$

where

$$D = \sup|F_m(\mathbf{x}_i) - y_i| \ (i = 1, 2, \cdots, N). \tag{12}$$

The final prediction, $F_M(\mathbf{x})$, which is determined by combination of M number of weak learners is expressed by

Table 4
XGB model predictions for consolidated universal database with different combinations of input features and default hyper-parameter set.

Test case	Number of input features	Set of input features	MAD (x 10^{-2})
1	24	Ω, Bo, Su_{fo}, We_{fo}, q''_H, We_{go}, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Cc, Fr_{fo}, Fr_{go}, Su_{go}, σ, Co, f_{go}, D_h, ρ_r	4.45
2	23	Ω, Bo, Su_{fo}, We_{fo}, q''_H, We_{go}, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Cc, Fr_{fo}, Fr_{go}, Su_{go}, σ, f_{go}, D_h, ρ_r	4.45
3	22	Ω, Bo, Su_{fo}, We_{fo}, q''_H, We_{go}, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Cc, Fr_{fo}, Fr_{go}, Su_{go}, σ, f_{go}, D_h	4.43
4	21	Ω, Bo, Su_{fo}, We_{fo}, q''_H, We_{go}, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Cc, Fr_{fo}, Fr_{go}, Su_{go}, f_{go}, D_h	4.41
5	20	Ω, Bo, Su_{fo}, We_{fo}, q''_H, We_{go}, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Cc, Fr_{fo}, Su_{go}, f_{go}, D_h	4.45
6	19	Ω, Bo, Su_{fo}, We_{fo}, q''_H, We_{go}, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Cc, Fr_{fo}, Su_{go}, D_h	4.34
7	18	Ω, Bo, Su_{fo}, We_{fo}, q''_H, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Cc, Fr_{fo}, Su_{go}, D_h	4.32
8	17	Ω, Bo, Su_{fo}, We_{fo}, q''_H, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Fr_{fo}, Su_{go}, D_h	4.39
9	16	Ω, Bo, Su_{fo}, We_{fo}, q''_H, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, Re_{go}, f_{fo}, P_R,G, Fr_{fo}, Su_{go}	4.40
10	15	Ω, Bo, Su_{fo}, We_{fo}, q''_H, q''_{CHF}, Re_{fo}, Ca, λ_{K-H}, Bd, f_{fo}, P_R,G, Fr_{fo}, Su_{go}	4.35
11	14	Ω, Bo, Su_{fo}, We_{fo}, q''_H, q''_{CHF}, Ca, λ_{K-H}, Bd, f_{fo}, P_R,G, Fr_{fo}, Su_{go}	4.30
12	13	Ω, Bo, Su_{fo}, We_{fo}, q''_H, q''_{CHF}, Ca, λ_{K-H}, Bd, f_{fo}, G, Fr_{fo}, Su_{go}	4.39
13	12	Ω, Bo, Su_{fo}, We_{fo}, q''_H, Ca, λ_{K-H}, Bd, f_{fo}, G, Fr_{fo}, Su_{go}	4.42
14	11	Ω, Bo, Su_{fo}, We_{fo}, q''_H, Ca, λ_{K-H}, Bd, G, Fr_{fo}, Su_{go}	4.24
15	10	Ω, Bo, Su_{fo}, We_{fo}, q''_H, Ca, λ_{K-H}, Bd, G, Fr_{fo}	4.33
16	9	Ω, Bo, Su_{fo}, We_{fo}, q''_H, Ca, λ_{K-H}, Bd, G	4.33
17	8	Ω, Bo, Su_{fo}, We_{fo}, q''_H, λ_{K-H}, Bd, G	4.30
18	7	Ω, Bo, Su_{fo}, We_{fo}, q''_H, λ_{K-H}, Bd	4.20
19	**6**	Ω, Bo, Su_{fo}, We_{fo}, q''_H, λ_{K-H}	**4.16**
20	5	Ω, Bo, Su_{fo}, We_{fo}, q''_H	4.35
21	4	Ω, Bo, We_{fo}, q''_H	4.99
22	3	Bo, We_{fo}, q''_H	5.41
23	2	Bo, We_{fo}	7.48
24	1	We_{fo}	9.00

$$F_M(\mathbf{x}) = \inf\left\{F(\mathbf{x}) : \sum_{m:F_m(\mathbf{x})\le F(\mathbf{x})} \log\left(\frac{1}{\beta^{(m)}}\right) \ge \frac{1}{2}\sum_{m=1}^{M}\log\left(\frac{1}{\beta^{(m)}}\right)\right\} for\, m = 1, 2, \cdots, M \tag{13}$$

where $F_m(\mathbf{x})$ represents the prediction by m weak learner for each input x_i[107]. For the classification, the Adaboost algorithm uses same regression procedure to find final classifier, $C(\mathbf{x})$, that approximates Bayes classifier $C^*(\mathbf{x})$ by combining weak classifiers [108].

2.3.2. Gradient boosting

The Gradient boosting (GB) is supervised learning ensemble method based algorithm that generates the stronger predictive learner which has a form of a linear combination of the weak learners that are calculated through iterative regressions and successively added to the previous predictive learner. The GB algorithm, originally developed by Friedman [109], is designed to utilize the negative gradient of the loss function,

$$L = \frac{1}{2}(F_m(\mathbf{x}_i) - y_i)^2 \tag{14}$$

Table 5
List of selected eight hyper-parameters and their definitions with value of hyper-parameters determined by *Optuna*, MAE, and MAD.

Hyper-parameter	Definition	Value	MAE	MAD ($\times 10^{-2}$)
max depth	Hyper-parameter that set the maximum depth of the decision tree	8	2.45 %	3.57
learning rate	Hyper-parameter indicating the reflection level of weak learner	0.161		
n estimators	Hyper-parameter that set the number of boosting iterations	3500		
colsample bytree	Hyper-parameter indicating the ratio of features sampling required for tree generation	0.807		
subsample	Hyper-parameter indicating the data sampling ratio during training	0.948		
min child weight	Hyper-parameter indicating the minimum sum of instance weight (hessian) needed in a child	6		
gamma	Hyper-parameter indicating the minimum loss reduction value to determine further division of the leaf node	0		
Early stopping rounds	Hyper-parameter monitoring the performance of the model and stopping the training procedure once the performance on the test dataset has not improved after a fixed number of training iterations	50		

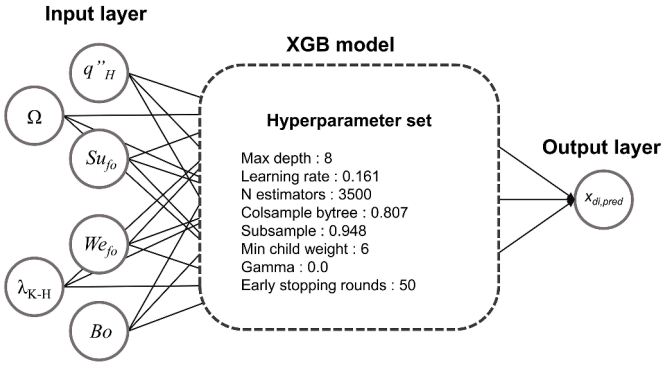


Fig. 8. Flow chart of XGB model finalized in this study with six input features and hyper-parameter set.

in calculating the residuals of the sample data, which gives the GB machine training guidance, (positive or negative) directions and magnitude, in minimizing the loss function. The squared error is usually used for loss function, L. The GB algorithm has three strong advantages including easy implementation, high accuracy by successive processes in reducing residuals, and high flexibility in selecting training models and loss functions, not sticking to decision tree and the squared errors [110].

The GB algorithm, as shown in Fig. 6(a), consists of two larger processes, initialization and iteration, and the iteration process again branches off into four sub-processes: computing residuals, training regression tree, calculating the weak prediction, and updating previous prediction. In the initialization procedure, combined weak learner, $F_m(\mathbf{x})$, is initialized by the initial learner, $F_0(\mathbf{x})$, which is calculated by finding $F_m(\mathbf{x})$ which minimizes the sum of the loss function as follows,

$$F_0(x) = \underset{c}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i, c), \tag{15}$$

where c represents arbitrary constant. Eq. (15) can be expressed as

$$\frac{\partial}{\partial F_m(\mathbf{x})} \sum_{i=1}^{N} L = \frac{\partial}{\partial F_m(\mathbf{x})} \sum_{i=1}^{N} \frac{1}{2}(F_m(\mathbf{x}_i) - y_i)^2 \tag{16}$$

and by setting it equal to zero, it can be expressed as

$$F_0(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} y_i = \bar{y} \tag{17}$$

which represents the average of the target feature data, y.

In the following iteration procedure, four sub-procedures are iterated M times, where 'm' and 'M' represent the index of each tree and total number of trees, respectively. As explained in procedure of 2-a in Fig. 6 (a), the 'm-1' learner, $F_{m-1}(\mathbf{x})$, is used to compare with original target data, y_i, and calculate the pseudo residual,

$$r_{im} = -\left[\frac{\partial L(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)}\right]_{F(\mathbf{x})=F_{m-1}(\mathbf{x})}. \tag{18}$$

Eq. (18) is simplified to

$$r_{im} = -\frac{1}{2}\frac{\partial (y_i - F_{m-1}(\mathbf{x}_i))^2}{\partial F_{m-1}} = (y_i - F_{m-1}(\mathbf{x}_i)) \tag{19}$$

in which the negative gradient yields the normal residual, $y_i - F_{m-1}(\mathbf{x}_i)$, which makes r_{im}called 'pseudo' residuals. In procedure of 2-b in Fig. 6 (a), the GB algorithm trains the trees with the input features, x, against the pseudo residuals, r_{im}, and creates the terminal nodes, R_{jm}, where each dataset, $(\mathbf{x}_i, r_{im})$, is categorized into. 'j' and 'J_m' represent the index and the total number of terminal nodes, respectively. In procedure of 2-c in Fig. 6(a), the GB searches for γ_{jm} that minimizes the loss function on each terminal node j. The sum, '$\sum_{\mathbf{x}_i \in R_{jm}}^{n} L$' implies the sum of loss functions on all dataset belong to R_{jm}. The regularization constant,γ_{jm}, for the datasets categorized into each terminal node, R_{jm}, are calculated by Eq. (20) which is expressed by

$$\gamma_{jm} = \underset{\gamma_{jm}}{\operatorname{argmin}} \sum_{\mathbf{x}_i \in R_{jm}}^{N} L\left(y_i, F_{m-1}(\mathbf{x}_i) + \gamma_{jm}\right) \text{ for } j = 1, \ldots, J_m. \tag{20}$$

Eq. (20) can be alternatively expressed as follows

$$\gamma_{jm} = \frac{1}{N_j} \sum_{\mathbf{x}_i \in R_{jm}} r_{im} \tag{21}$$

which represents the averaged residuals in terminal node, R_{jm}. The term, $N_j$means the number of datasets in 'j-th' terminal node. By taking 'ε' instead of '*argmin*' and equating to the smallest, Eq. (20) also can be expressed as

$$\frac{\partial}{\partial \gamma_{jm}} \sum_{\mathbf{x}_i \in R_{jm}} \frac{1}{2}\left(y_i - F_{m-1}(\mathbf{x}_i) - \gamma_{jm}\right)^2 = -\sum_{\mathbf{x}_i \in R_{jm}} \left(y_i - F_{m-1}(\mathbf{x}_i) - \gamma_{jm}\right) \le \varepsilon. \tag{22}$$

In procedure of 2-d in Fig. 6(a), $F_m(\mathbf{x})$ is updated by adding $F_{m-1}(\mathbf{x})$ to all previous weak learners with the learning rate of 'l_r' as follows,

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + l_r \sum_{j=1}^{J_m} \gamma_{jm} I(x \in R_{jm}), \tag{23}$$

which is intended to prevent the over-fitting problem [109]. The final learner and prediction are acquired by

$$F_M(\mathbf{x}) = \sum_{m=1}^{M} F_m(\mathbf{x}). \tag{24}$$

2.3.3. XGBoost (eXtreme gradient boosting)

The XGBoost is the modified gradient boosting algorithm that adds

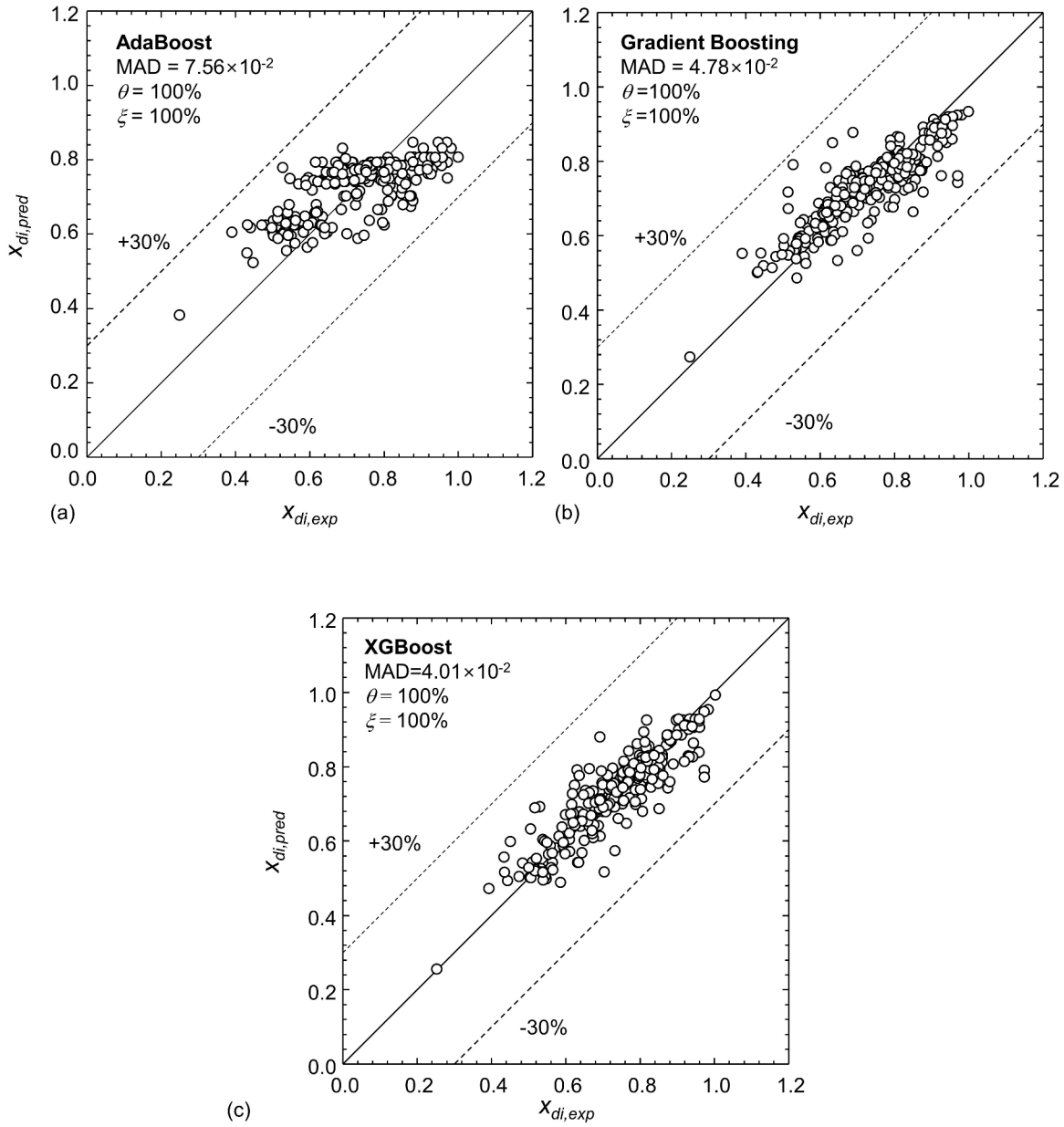


Fig. 9. Three boosting algorithm's machine predictions with six input features and default hyper-parameter settings: (a) AdaBoost, (b) Gradient Boosting, and (c) XGBoost.

the regularization term, Ω, to the differentiable loss function, L, in order to avoid over-fitting problem and to improve the computing performance, speed and accuracy [111]. For the training data set of $\{(\mathbf{x}_i, y_i)|\mathbf{x}_i \in \mathbb{R}^p, y_i \in \mathbb{R}\}$ with N samples and p features, it can be expressed as

$$L_\Omega(F_M) = \sum_{i=1}^{N} L(y_i, F_M(\mathbf{x}_i)) + \sum_{m=1}^{M} \Omega(f_m) \tag{25}$$

where

$$\Omega(f_m) = \gamma J_m + \frac{1}{2}\lambda \| w \|^2 . \tag{26}$$

In Eq. (25), L_Ω and f_m, respectively, represent the regularized loss function and individual weak learners from an independent m-th tree structure. The term, $\Omega(f_m)$, prevents the over-fitting by penalizing the complexity of tree structure. Here, γ is the minimum loss reduction and λ the regularization factor. In procedure (1) in Fig. 6(b), the weak learner is initialized by

$$F_0(x) = \underset{c}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i, c) \tag{27}$$

where c represents arbitrary constant. The regularized loss function, L_Ω, is minimized by the second-order approximation in procedure 2-a in Fig. 6(b) by

$$L_\Omega^{(m)} \simeq \sum_{i=1}^{N} \left[L(y_i, F_{(m-1)}(\mathbf{x}_i)) + g_i f_m(\mathbf{x}_i) + \frac{1}{2} h_i f_m^2(\mathbf{x}_i) \right] + \Omega(f_m) \tag{28}$$

where g_i and h_i are gradient and Hessian of loss functions that are computed, respectively, by

$$g_m(\mathbf{x}_i) = \left[\frac{\partial L(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)} \right]_{F(\mathbf{x}) = F_{m-1}(\mathbf{x})} \tag{29}$$

Table 6
Assorted empirical correlations for predicting dryout incipience quality

Author(s)	Equation
Sun (2001) [41]	$x_{crit} = 10.795(q''_H/1000)^{-0.125} G^{-0.333} (1000 D_h)^{-0.07} e^{0.01715\times10^{-5}P}$
	$for\ 4.9\ bar \le P \le 29.4\ bar$
	$x_{crit} = 19.398(q''_H/1000)^{-0.125} G^{-0.333} (1000 D_h)^{-0.07} e^{0.00255\times10^{-5}P}$
	$for\ 29.4\ bar \le P \le 98\ bar$
	$x_{crit} = 32.302(q''_H/1000)^{-0.125} G^{-0.333} (1000 D_h)^{-0.07} e^{0.00795\times10^{-5}P}$
	$for\ 98\ bar \le P \le 196\ bar,\ Fr* = \dfrac{x_{crit} G}{\sqrt{\rho_g(\rho_f-\rho_g) G \cos\theta D_h}},$
	$\theta = 0\ for\ horizontal\ flow$
Wojtan et al. (2005) [42]	$x_{crit} = x_{di} - \dfrac{8}{(2+Fr*)^2}$
	$x_{di} = 0.58\exp\left[0.52 - 0.235 We_g^{0.17} Fr_{g,Mori}^{0.37} \left(\dfrac{\rho_g}{\rho_f}\right)^{0.25} \left(\dfrac{q''_H}{q''_{crit}}\right)^{0.70}\right]$
	$We_g = \dfrac{G^2 D_{eq}}{\rho_g \sigma},\ Fr_{g,Mori} = \dfrac{G^2}{\rho_g(\rho_f-\rho_g) g D_{eq}},\ D_{eq} = \sqrt{\dfrac{4A}{\pi}}$
	$q''_{crit} = 0.131 \rho_g^{0.5} h_{fg} [g\sigma(\rho_f-\rho_g)]^{0.25}$
Cheng et al. (2006) [43]	$x_{di} = 0.58\exp\left[0.52 - 0.67 We_g^{0.17} Fr_{g,Mori}^{0.348} \left(\dfrac{\rho_g}{\rho_f}\right)^{0.25} \left(\dfrac{q''_H}{q''_{crit}}\right)^{0.70}\right]$
Del Col et al. (2007) [44]	$x_{di} = 0.4695 \left(\dfrac{4 q''_H RLL}{G D_h h_{fg}}\right)^{1.472} \left(\dfrac{G^2 D_h}{\rho_f \sigma}\right)^{0.3024} \left(\dfrac{D_h}{0.001}\right)^{0.1836} (1-P_R)^{1.239}$
	$RLL = \left[0.437 \left(\dfrac{\rho_g}{\rho_f}\right)^{0.073} \left(\dfrac{\rho_f \sigma}{G^2}\right)^{0.24} D_h^{0.72} \left(\dfrac{G h_{fg}}{q''_H}\right)\right]^{1/0.96}$
Cheng et al. (2008) [45]	$x_{di} = 0.58\exp\left[0.52 - 0.236 We_g^{0.17} Fr_{g,Mori}^{0.17} \left(\dfrac{\rho_g}{\rho_f}\right)^{0.25} \left(\dfrac{q''_H}{q''_{crit}}\right)^{0.27}\right]$
Jeong and Park (2009) [46]	$x_{di} = 6.2 \mathrm{Re}_{fo}^{-0.5} Bo^{-0.2} Bd^{-0.45}$
Ducoulombier et al. (2011) [47]	$x_{di} = 1 - 338 Bo^{0.703} P_R^{1.43}$
Mastrullo et al. (2012) [48]	$x_{di} = 1 - 20.82 q''^{0.273}_H G^{1.231} D_h^{0.252} \dfrac{\mu_f}{h_{fg}^{0.273} (\rho_f \sigma)^{1.252}} P_R^{0.721}$
Kim and Mudawar (2013) [38]	$x_{di} = 1.4 We_{fo}^{0.03} P_R^{0.08} - 15.0 \left(Bo \dfrac{P_H}{P_F}\right)^{0.15} Ca^{0.35} \left(\dfrac{\rho_g}{\rho_f}\right)^{0.06}$

and

$$h_m(\mathbf{x}_i) = \left[\frac{\partial^2 L(y_i, F(\mathbf{x}_i))}{\partial F(\mathbf{x}_i)^2}\right]_{F(\mathbf{x})=F_{m-1}(\mathbf{x})}. \tag{30}$$

Eq. (28) can be simplified by removing the constant term as follows,

$$\widetilde{L}_\Omega^{(m)} = \sum_{j=1}^{J_m} \left[\left(\sum_{j\in J_m} g_j\right) w_j + \frac{1}{2}\left(\sum_{j\in J_m} h_j + \lambda\right) w_j^2\right] + \gamma J_m. \tag{31}$$

The leaf weight at j^{th} leaf, w_j, in Eq. (31) can be updated by optimal weights that are acquired by

$$w_j^* = \frac{G_j}{\mathrm{H}_j + \lambda} \tag{32}$$

and Eq. (31) is updated as follows,

$$\widetilde{L}_\Omega^{(m)} = -\frac{1}{2}\sum_{j=1}^{J_m} \frac{G_j^2}{H_j + \lambda} + \gamma J_m \tag{33}$$

having similarity score in the first term,

$$\frac{G_j^2}{H_j + \lambda}, \tag{34}$$

which is used in gain function, L_{gain}, which is used to determine if branching is necessary. $G_j = \sum_{j\in J_M} g_j$ and $H_j = \sum_{j\in J_M} h_j$ are used in Eq. (38). The gain function is expressed as

$$L_{gain} = \frac{1}{2}\left[\frac{G_L^2}{\mathrm{H}_L + \lambda} + \frac{G_R^2}{\mathrm{H}_R + \lambda} - \frac{(G_L + G_R)^2}{\mathrm{H}_{L+R} + \lambda}\right] - \gamma, \tag{35}$$

where the subscript L and R represents left and right leaf, respectively. In procedure of 2-b in Fig. 6(b), the weak learners in each tree are updated by

$$F_m(x) = F_{m-1}(x) + l_r \cdot \phi_m(x), \tag{36}$$

where l_r is the learning rate and ϕ_m minimized objective function which is expressed as

$$\phi_m = \underset{\phi}{\mathrm{argmin}} \left[\sum_{i=1}^{N} \frac{1}{2} h_m(x_i) \left[-\frac{g_m(x_i)}{h_m(x_i)} - \phi(x_i)\right]^2 + \gamma J_m + \frac{1}{2}\lambda \| w \|^2\right]. \tag{37}$$

In procedure of 2-c in Fig. 6(b), the final combined learner and prediction are acquired by

$$F_M(x) = \sum_{m=1}^{M} F_m(x). \tag{38}$$

2.4. Model selection

2.4.1. Error estimation

The predicting accuracy of trained boosting machine is assessed by the mean absolute error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \frac{\left|x_{di,exp,i} - x_{di,pred,i}\right|}{x_{di,exp,i}} \tag{39}$$

and mean absolute deviation (MAD)

$$MAD = \frac{1}{N}\sum_{i=1}^{N} \left|x_{di,exp,i} - x_{di,pred,i}\right| \tag{40}$$

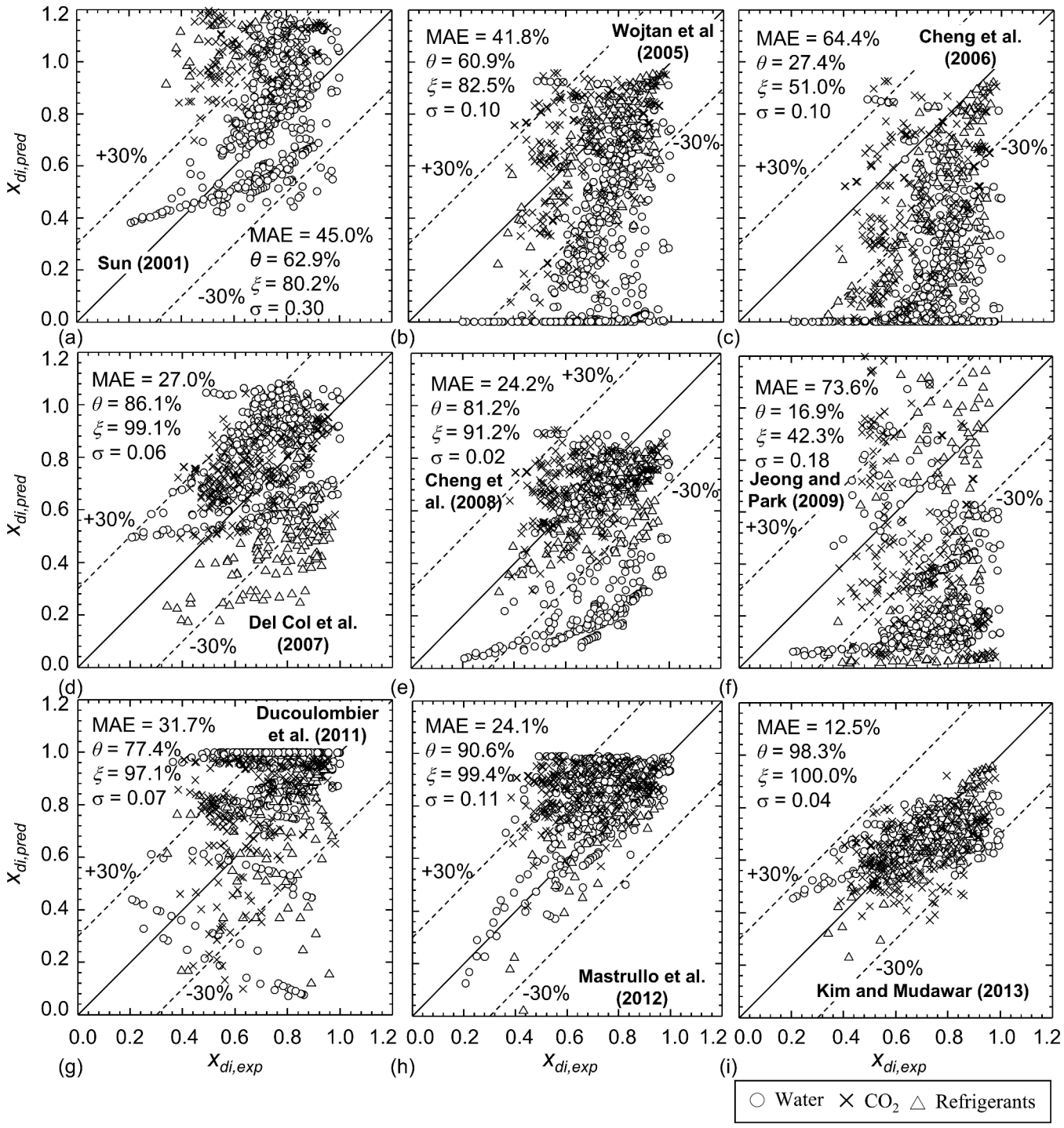


Fig. 10. Comparison of Consolidated 997-point Database with prediction of previous empirical correlations of: (a) Sun [41], (b) Wojtan et al. [42], (c) Cheng et al. [43], (d) Del Col et al. [44], (e) Cheng et al. [45], (f) Jeong and Park [46], (g) Ducoulombier et al. [47], (h) Mastrullo et al. [48], and (i) Kim and Mudawar [38].

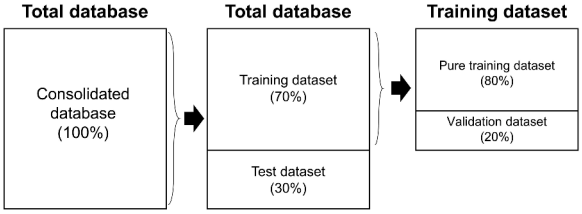


Fig. 11. Schematic diagram for dataset preparation by random selection for machine learning algorithm implementation.

where $x_{di,\exp,i}$ and $x_{di,pred,i}$ represent the i-th experimental dryout incipient data and its prediction. MAE and MAD measure how relatively and absolutely the predicted values differ from the actual value. θ and ξ are used to denote the percentage of data predicted within ±30% and ±50%, respectively. The thermophysical properties required for comparison of the universal database to predicted by empirical correlations and XGBoost are acquired by NIST's REFPROP 9.0 software [112] except for FC-72, engineering dielectric fluid, which is provided by 3M company.

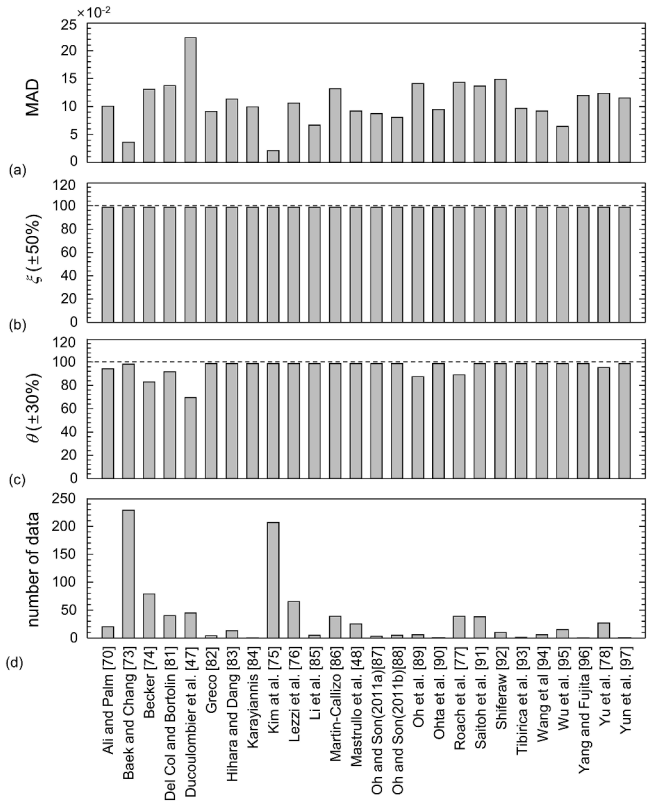


Fig. 12. Validation test results with selected hyperparameter set and six input features: (a) MAD by XGB model trained with excluded dataset and corresponding (b) ξ (c) θ, and (d) number of data.

2.4.2. *Input feature selection*

Selecting a set of features that participate in the learning process is one of the most crucial parts that influence the predictive performance of a trained model because including irrelevant features or using an excessive number of features can degrade the performance due to noise signals or heavy computational load. Selecting an appropriate set of features is key to success in the machine learning process, and 24 features were selected as candidate input features for XGBoost machine training. Among 24 features, the definitions and physical meanings of 10 key parameters are as follows:

1) Ω(Equilibrium frequency of phase change): The inverse of the characteristic frequency of phase change under the thermal equilibrium assumption [100],

$$\Omega = \frac{q''_H P_H}{A_{ch} h_{fg}} \frac{\rho_f - \rho_g}{\rho_f \rho_g}. \tag{41}$$

2) *Bo* (Boiling number): Ratio of mass of vapor generation per unit area of heat transfer surface to mass flow rate per unit cross-sectional flow area,

$$Bo = \frac{q''_H}{G h_{fg}}. \tag{42}$$

3) Su_{fo} (Liquid only Suratman number): Ratio of surface tension to the momentum transport inside a fluid,

$$Su_{fo} = \frac{\sigma \rho_f D_h}{\mu_f^2}. \tag{43}$$

4) We_{fo} (Liquid only Weber number): Ratio of disruptive hydrodynamic forces of liquid phase to the stabilizing effect to surface tension,

$$We_{fo} = \frac{G^2 D_h}{\rho_f \sigma}. \tag{44}$$

5) q''_H (Heat flux based on heated perimeter of channel): Effective heat flux averaged over heated perimeter of micro-channel,

$$q''_H = \frac{Q}{A_h} \tag{45}$$

6) We_{go} (Vapor only Weber number): Ratio of disruptive hydrodynamic forces of gas phase to the stabilizing effect to surface tension,

$$We_{go} = \frac{G^2 D_h}{\rho_g \sigma}. \tag{46}$$

7) q''_{CHF} (Critical heat flux): The heat flux at which boiling ceases to be an effective form of transferring heat from a solid surface to a liquid [113],

$$q''_{CHF} = 0.0035 G h_{fg} We_{fo}^{-0.12}. \tag{47}$$

8) Re_{fo} (Liquid only Reynolds number): Ratio of liquid only inertia to viscous force,

$$\mathrm{Re}_{fo} = \frac{G D_h}{\mu_f}. \tag{48}$$

9) *Ca* (Capillary number): Relative effect of viscous drag to surface tension force,

$$Ca = \frac{We_{fo}}{\mathrm{Re}_{fo}} = \frac{\mu_f V}{\sigma}. \tag{49}$$

10) λ_{K-H}(Critical Kelvin-Helmholtz instability wavelength): Wavelength of fluid instability that occurs along the interface between two moving fluids [101],

$$\lambda_{K-H} = \frac{2\pi}{\left[\frac{(\rho_f - \rho_g)g}{\sigma}\right]^{1/2}}. \tag{50}$$

In this study, the permutation feature importance (PFI) has been adopted as a quantitative indicator representing how closely any individual feature is related to the target feature. As shown in Fig. 7(a), the PFI measures the changes in predicting performance after shuffling a set of values in specific features. The larger difference between predictions by the original and shuffled set of data, the more important the tested feature is. The PFI measuring procedures are as follows:

1) The machine learning is conducted with an original set of data, and the prediction is conducted, yielding MAD.
2) The data of the specific feature are shuffled, and the prediction is conducted yielding $\mathrm{MAD}^{(p)}$ without implementing any additional machine learning process. $\mathrm{MAD}^{(p)}$ represents MAD with shuffling the *p*-th feature.

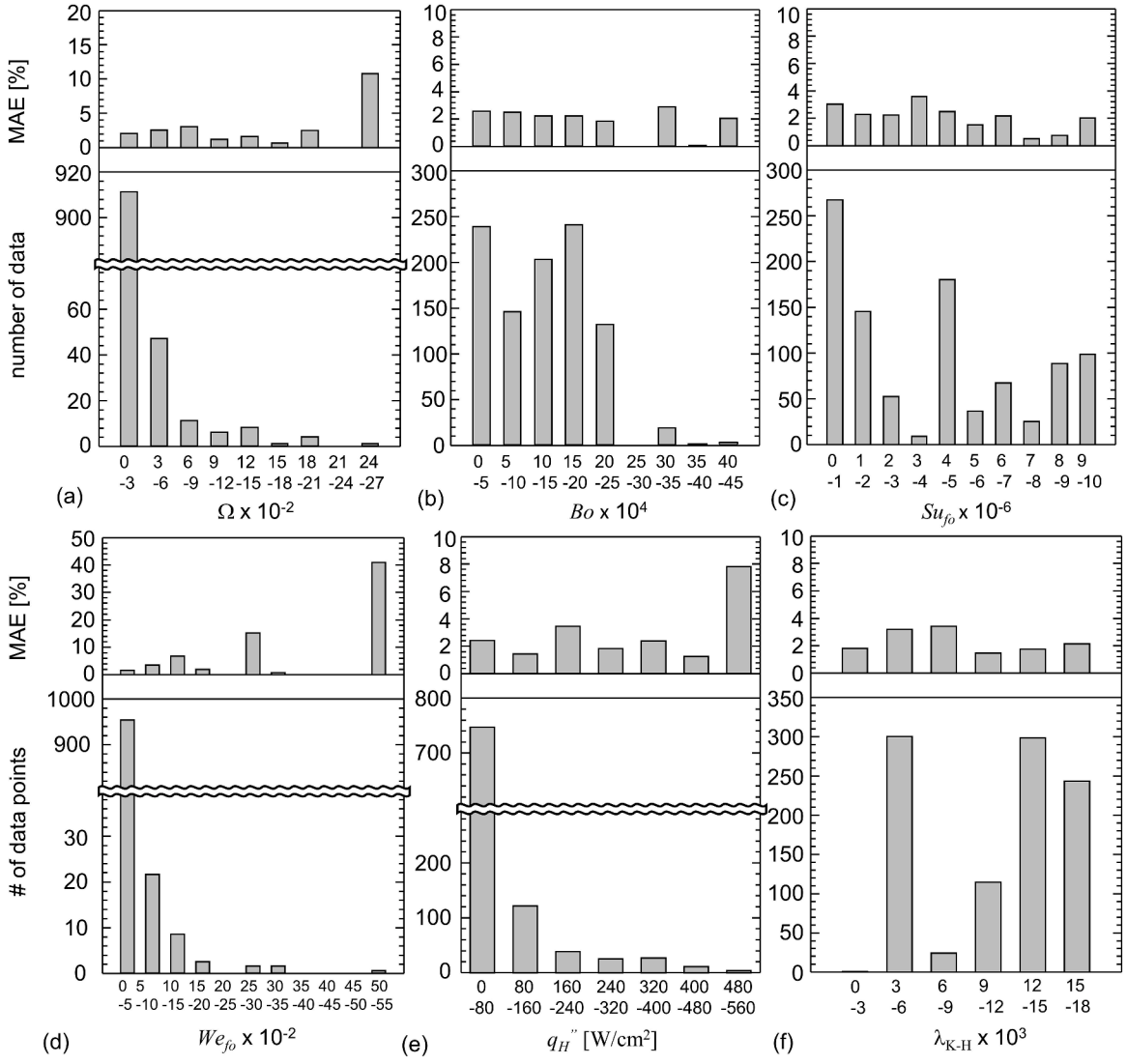


Fig. 13. Distribution of data points and MAE in predictions of newly developed XGB model for the Consolidated Database relative to: (a) equilibrium frequency of phase change, (b) boiling number, (c) liquid-only Suratman number, (d) liquid-only Weber number, (e) heat flux, and (f) Critical Kelvin-Helmholtz wavelength.

3) The PFI is calculated by $MAD^{(p)}$ - MAD.
4) Processes from (1) to (3) are repeated for the other input features.

As described in the procedures, the biggest advantage of the PFI is low computational resource utilization as it requires only one single iteration of the learning process and does not require any additional learning after shuffling. A total of 24 features are listed in the order of higher PFI as shown in Fig. 7(b) with error bars that represent the uncertainty in the shuffling process. PFI values are also provided in Table 2. The optimal set of features will be determined after investigating the effect of the number and combination of features on the learned model's predictive performance after hyper-parameter optimization.

Optuna, the automatic hyper-parameter optimization software described in the following section, is utilized for the initial settings of hyper-parameters in the optimal input feature selection process. The default hyper-parameters set by *Optuna* for three algorithms are tabulated in Table 3, and the assessment of predictive performance of models trained with different sets of input features is implemented using the default hyper-parameters set as shown in Table 4. These sets of input features are determined by removing input features one by one in the order of least important one based on PFI, and the optimal sets are determined by comparing the estimated MAD of these models. After investigation, the selected set of input features is [Ω, *Bo*, Su_{fo}, We_{fo}, q''_H, $\lambda_{K\text{-}H}$] which showed the MAD of 4.16×10^{-2} as shown in Table 4. The optimal number of input features is six determined by MAD estimation, and the predictive performance tends to decrease by over-fitting or under-fitting when an excess or insufficient number of input features is used.

2.4.3. *Hyper-parameter optimization*

The hyper-parameters, tuning parameters for the selected algorithm, should be determined before the training algorithm starts. One of characteristics of hyper-parameters that distinguishes from target and input features is their independence from the data, and the values of hyper-parameters are only used to control the data learning process of the model. The hyper-parameters are classified into model and algorithm hyper-parameters, which determine the structural settings of the model and data treatment methods, respectively. In the present XGBoost algorithm, the model hyper-parameters include max depth, learning rate, n estimators, colsample bytree, subsample, min child weight, gamma, and early stopping rounds. [114].

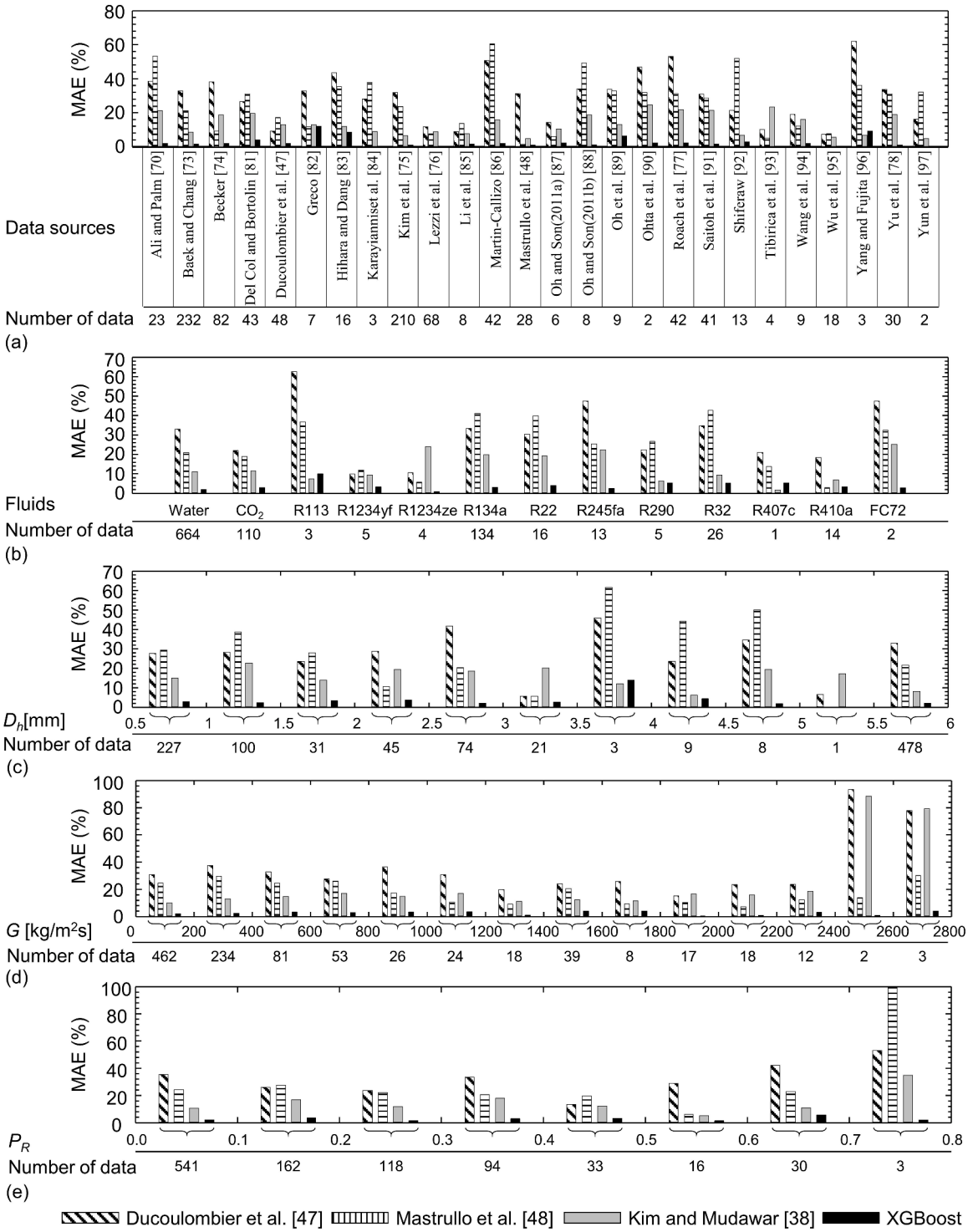


Fig. 14. Comparison of developed XGB machine with three empirical correlations in terms of predictive performance with respect to (a) data sources, (b) fluid type, (c) D_h, (d) G, and (e) P_R.

The performance of the selected model greatly varies depending on the selection of hyper-parameters, which requires the hyper-parameter optimization to find a tuple of hyper-parameters that minimizes a pre-defined loss function. In a previous study, the machine learning model with optimized hyper-parameters always performs better than model with default hyper-parameters [115]. Additionally, the hyper-parameter tuning might be a more important task than the choice of the machine learning algorithm [116]. Associated with problems in finding optimal values of hyper-parameters, the optimization methods are sorted into four categories: manual search, grid search, random search, and Bayesian optimization [117]. Bayesian optimization uses an unknown objective function and observed data to estimate the shape of the objective function, which is used to determine the next query point in the search for the extreme value, maximum or minimum. Bayesian optimization is the most widely used among the four because of its superior performance for expensive-to-evaluate data with utilization of prior knowledge [118]. A number of hyper-parameter optimization software that adopted Bayesian optimization have been developed,

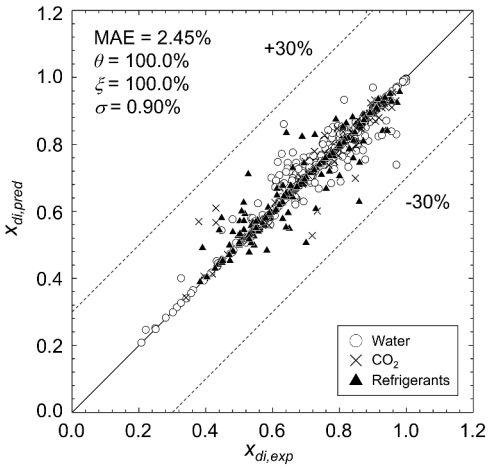


Fig. 15. Comparison of predictions of newly developed XGBoost machine against the 997-point Consolidated Database.

including Hyperopt [119], Autotune [120], and Vizier [121]. These softwares have shown technical limitations: heuristic construction of parameter-search-space, absence of a pruning strategy, and vulnerability to versatility.

In this study, an automatic hyper-parameter optimization software, *Optuna*, is employed to find the optimal values of hyper-parameters. *Optuna* features the *define-by-run* application programming interface (API) that follows a trial and study process through the respective evaluation and optimization of the objective function in the search for optimal sets of hyper-parameters. The *define-by-run* API conducts the optimization by minimizing or maximizing the objective function, which takes a set of hyper-parameters as input and returns its validation score [122]. The selected eight hyper-parameters by *Optuna* and their meanings are provided in Table 5 with MAE and MAD estimated with selected 6 features (Ω, *Bo*, Su_{fo}, We_{fo}, q''_H, $\lambda_{K\text{-}H}$). The final settings for input features and hyper-parameters and XGBoost learning procedure are schematically described in Fig. 8, and following are all results achieved by using these settings unless otherwise mentioned.

2.4.4. Assessment of predictive accuracy of candidate boosting algorithms

In the selection of the base boosting algorithm for analysis, the results of a pretest conducted for selected input features with the default hyper-parameter set provide information about the most suitable algorithm for the given database. Three well-known boosting algorithms, AdaBoost, GB and XGBoost, were tested for six input features by PFI with the universal dryout database, and the results provided with MAD are shown in Fig. 9. The XGBoost algorithm shows the best predictive accuracy with MAD of 4.01×10^{-2}, while the AdaBoost and GB show MAD values of 7.56×10^{-2} and 4.78×10^{-2}, respectively. Based on this pretest result, XGBoost is selected as the base boosting algorithm for the following detailed analysis.

3. Results and discussion

3.1. Assessment of previous empirical correlations

Previous empirical correlations were assessed for their predictive accuracies before investigating XGBoost's performance analysis, and a total of nine correlations featured for different dryout incipience cases are tabulated in Table 6. Dryout incipience occurs differently depending on flow features, including channel diameter, channel geometry, flow characteristics, and thermophysical properties of the fluid. Some correlations are designed for flows in mini/micro-channels [43-47], and one correlation is reinforced by adding CO_2 data to widen range of covered thermophysical properties [42,43,45]. The universal correlation of Kim and Mudawar which showed enhanced predicting accuracy is also included for assessment, and it was developed using the universal database of dryout incipience quality that includes various flow features and different liquids with a wide range of thermophysical properties [38].

As shown in Fig. 10, several previous correlations in Table 6 were assessed for their predictive accuracy against a universal database that was separated into three groups having very different thermophysical properties: water, refrigerant, and CO_2. The correlation of Sun [41] overpredicts the universal database except for the water data as shown in Fig. 10(a). A strong underpredicting trend especially for water data is observed for the correlations of Wojtan et al. [42] and Cheng et al. [43, 45]. Scattered predictions are observed for correlations of Del col et al. [44], Jeong and Park [46], and Ducoulombier et al. [47]. Mastrullo et al.'s correlation [48], developed using a database composed of R410A and CO_2 in mini-channels of $D = 6.00$ mm slightly over-predicts the data. The universal correlation by Kim and Mudawar [38] showed the best predictive performance among the nine correlations, with MAE of 12.5%, $\theta = 98.3$ % and $\xi = 100.0$ %, and the second smallest standard deviation, σ, in MAD for the three different groups. The standard deviation is calculated by

$$\sigma = \sqrt{\frac{\sum_{i=water,refrigerants,CO_2} (MAD_i - \overline{MAD})^2}{3}} \tag{51}$$

where

$$\overline{MAD} = \frac{\sum_{i=water,refrigerants,CO_2} MAD_i}{3}. \tag{52}$$

The superior predictive performance of Kim and Mudawar's universal correlation is attributed to the wide range of thermophysical properties, flow geometries, and flow features of the Consolidated Database.

3.2. Assessment of XGBoost algorithm predictive performance

As shown in Fig. 11, the consolidated database of dryout incipience is separated into 70% training dataset and 30% test dataset. The training dataset is used for machine learning model development, while the test dataset is used for evaluating the developed machine's predictive performance. The training dataset is further divided into 80% pure training dataset and 20% validation dataset, and the validation dataset is used to test the trained model and find optimal hyper-parameter settings. The separation of the database is randomly conducted unless the user specifies a random-state parameter, which determines the data selection pattern and fixes the test dataset during the evaluation of different input features and hyper-parameters. By setting the random-state parameter, performance evaluation with different parameters can be done using the same training and test datasets. The test for input features in Table 4 is conducted with five different random state parameters, 1 to 5, for which the averaged MAD values are presented.

The predictive performance of the XGBoost algorithm on unseen data is a crucial feature that it should be equipped with as successful thermal system design tool. The assessment of the predicting accuracy of the XGBoost model, trained with the Consolidated Database excluding one specific data source one by one was conducted by comparing it with the excluded dataset, which serves as the unseen test data. Unlike the prior training process that used 26 sources, the training dataset consists of randomly selected data from 25 sources while the excluded dataset is used for test dataset. The validation test results with excluded database are shown in Fig. 12. The model trained with the database excluding Ducoulombier et al. [47] showed the highest MAD of 22.63×10^{-2} but still acceptable accuracy. This highest MAD is attributed to the absence of a dataset that has similar flow conditions. On the other hand, the

model trained with database excluding Kim et al. [75] showed the best accuracy with the lowest MAD of 2.39×10^{-2}. The existence of Baek and Chang's [73] database which has very similar flow conditions with the database of Kim et al. allowed the model to learn the flow information of the unseen database. This result demonstrates the strong dependence of the learning process on the database, but XGBoost still shows reasonably high and uniform accuracy in predicting different unseen databases.

Fig. 13 shows the effects of data distribution within a specific feature on the predictive accuracy of the developed machine. It does so by comparing the MAE and the number of data points for individual datasets that have been segregated according to an ascending order of input features. Six individual figures, Fig. 13(a) to 13(f), for selected input features consist of upper and lower bar charts for MAE and the number of data points for segregated datasets. The developed machine shows overall good accuracy evidenced by MAE showing uniformly low values for the entire database excepting for datasets having a small number of data points in q''_H, We_{fo}, and Ω. The lower predictive accuracy for these datasets is attributed to not providing enough information to XGBoost algorithm to train machine for the given data range. Based on this analysis, the complimentary database enhancement for the present database by adding data to empty bins is necessary to increase the predicting accuracy for unseen data.

3.3. Comparison between universal correlation and XGBoost predictive machine

The developed XGBoost machine is compared with three selected empirical correlations in terms of predictive accuracy for datasets segregated into different bins with respect to data source, type of fluids, D_h, G, and P_R as shown in Fig. 14. Overall, the developed machine shows better predictive accuracy with lower MAE in most data source bins compared to the three correlations except for Yang and Fujita's [96] data as shown in Fig. 14(a). Aside from having only a small number of data, three data points from Yang and Fujita are sorted into minor bins in six features so that XGBoost algorithm cannot learn about those data during training. This also the case for seven data points from Greco et al. [82] which shows the worst predictive accuracy, with a MAE of 12.8%. Meanwhile, in spite of the small number of data, the two data points from Yun et al. (2005), sorted into dominant bins, are predicted with the best accuracy, with a MAE of 0.2%. In terms of data segregated according to type of fluid, the data for R113 and R407c, which have the smallest number of data points, are not well predicted by the XGBoost machine. They show inferior predictive accuracy compared to Kim and Mudawar's universal correlation as shown in Fig. 14(b). Three data falling into the bin of $D_h = 3.5 - 4$ are predicted by XGBoost with MAE of 14.02% which is slightly higher than Kim and Mudawar's universal correlation's MAE of 12.06% as shown in Fig. 14(c). Comparing with Fig. 13, it should be noted that XGBoost shows poor predictive accuracy (MAE of 41.65%) for one datum sorted into the bin of We_{fo}= 50-55$\times10^{-2}$ while the other two data points falling into the bins of We_{fo}=0-5 and 10-15$\times10^{-2}$ are well predicted, with MAE values of 0.27% and 0.15%, respectively. For the two parameters G and P_R, superior predictive performance of the XGBoost machine compared to the other three correlations is validated in Fig. 14(d) and 14(e), which is attributed to the uniform data distribution within the six features. Finally, the developed XGBoost machine is assessed as having excellent predictive accuracy against the entire Consolidated Database, evidenced by a MAE of 2.45%, $\theta = 100.0\%$ and $\xi = 100\%$, as shown in Fig. 15.

4. Conclusion

This study explored the use of machine learning to predict dryout incipient quality for flow boiling in mini/micro-channels. The Consolidated Database used in this study consisted of 997 data points from 26 sources. The data were randomly split into two sets: 30% for training and 70% for testing. These sets were used to train, validate, and test the machine learning methodology. The proposed methodology included a decision tree model with CART algorithm and three boosting algorithms. The XGBoost algorithm, which showed the highest predictive accuracy, was chosen as the representative algorithm. The effectiveness of the machine learning technique was validated by superior performance compared to published empirical correlations. The key findings from this study are as follows:

1. The successful use of XGBoost algorithm in predicting dryout incipience quality demonstrates the potential of applying machine learning techniques to predict complex two-phase flow and heat transfer phenomena, which expedites the proliferation of machine learning techniques in the design of two-phase thermal systems. The assessment of predictive accuracy of three boosting algorithms, Adaboost, Gradient Boosting, and XGBoost, was conducted using the Consolidated Database, and XGBoost was deemed most suitable for predicting the data, evidenced by the best predictive accuracy with MAD = 3.57×10^{-2}.
2. The predictive performance tests conducted with different input features and sets of hyper-parameters demonstrated the importance of finding optimal parameters in the XGBoost training process for best performance, let alone algorithm selection. Six input features were selected after conducting performance tests that involved one-by-one removal of features, starting from the least important one, determined by a permutation feature importance test, and eleven optimal hyper-parameters were determined by the automatic hyper-parameter optimization software *Optuna*.
3. The assessment of XGBoost machine's accuracy in predicting unseen data conducted with excluded data source revealed that the strong dependence of XGBoost's performance on whether the database includes the flow information about unseen data. However, the XGBoost model trained in this study still showed reasonably high accuracy, with MAD ranging from 2.39×10^{-2} to 22.63×10^{-2}, in the validation tests conducted using the Consolidated Database.
4. The XGBoost machine showed superior predictive performance compared to previous empirical correlations, evidenced by an overwhelmingly smaller MAE of 2.45% against the Consolidated Database compared to 12.5% for best predicting prior correlation. The data distribution within the selected features for the training algorithm is a key parameter in determining the predictive accuracy of a specific dataset.

CRediT authorship contribution statement

Hyeonseok Noh: Investigation, Software, Data curation, Writing – original draft. **Seunghyun Lee:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing. **Sung-Min Kim:** Data curation, Writing – review & editing. **Issam Mudawar:** Writing – review & editing.

Declaration of competing interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgement

The authors are grateful for financial support through the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (MSIT) of Korea government NRF-RS-2023-00280898 and acknowledge the support of the National Aeronautics and Space Administration (NASA) under grant no. 80NSSC22K0328.

References

[1] I. Mudawar, Assessment of high-heat-flux thermal management schemes, IEEE Trans. – CPMT 24 (2001) 122–141.
[2] I. Mudawar, Two-phase microchannel heat sinks: theory, applications, and limitations, J. Electron. Packag. – Trans. ASME 133 (2011), 041002-2.
[3] T.J. LaClair, I. Mudawar, Thermal transients in a capillary evaporator prior to the initiation of boiling, Int. J. Heat Mass Transf. 43 (2000) 3937–3952.
[4] S.G. Liter, M. Kaviany, Pool-boiling CHF enhancement by modulated porous-layer coating: theory and experiment, Int. J. Heat Mass Transf. 44 (2001) 4287–4311.
[5] G. Liang, I. Mudawar, Pool boiling critical heat flux (CHF) – part 2: assessment of models and correlations, Int. J. Heat Mass Transf. 117 (2018) 1368–1383.
[6] I. Mudawar, R.A. Houpt, Mass and momentum transport in smooth falling liquid films laminarized at relatively high Reynolds numbers, Int. J. Heat Mass Transf. 36 (1993) 3437–3448.
[7] C.O. Gersey, I. Mudawar, Effects of heater length and orientation on the trigger mechanism for near-saturated flow boiling critical heat flux - II. Critical heat flux model, Int. J. Heat Mass Transf. 38 (1995) 643–654.
[8] S. Mukherjee, I. Mudawar, Pumpless loop for narrow channel and micro-channel boiling from vertical surfaces, J. Electron. Packag. 125 (2003) 431–441.
[9] M.E. Johns, I. Mudawar, An ultra-high power two-phase jet-impingement avionic clamshell module, J. Electron. Packag. 118 (1996) 264–270.
[10] S.N. Joshi, E.M. Dede, Two-phase jet impingement cooling for high heat flux wide band-gap devices using multi-scale porous surfaces, Appl. Therm. Eng. 110 (2017) 10–17.
[11] S. Toda, A study of mist cooling (1st report: investigation of mist cooling), Trans. JSME 38 (1972) 581–588.
[12] W.P. Klinzing, J.C. Rozzi, I. Mudawar, Film and transition boiling correlations for quenching of hot surfaces with water sprays, J. Heat Treat. 9 (1992) 91–103.
[13] L. Lin, R. Ponnappan, Heat transfer characteristics of spray cooling in a closed loop, Int. J. Heat Mass Transf. 46 (2003) 3737–3746.
[14] S. Lee, I. Mudawar, Investigation of flow boiling in large micro-channel heat exchangers in a refrigeration loop for space applications, Int. J. Heat Mass Transf. 97 (2016) 110–129.
[15] S. Lee, V.S. Devahdhanush, I. Mudawar, Pressure drop characteristics of large length-to-diameter two-phase micro-channel heat sinks, Int. J. Heat Mass Transf. 115 (2017) 1258–1275.
[16] M.K. Sung, I. Mudawar, Single-phase and two-phase heat transfer characteristics of low temperature hybrid micro-channel/micro-jet impingement cooling module, Int. J. Heat Mass Transf. 51 (2008) 3882–3895.
[17] M.K. Sung, I. Mudawar, Single-phase and two-phase hybrid cooling scheme for high-heat-flux thermal management of defense electronics, J. Electron. Packag. 131 (2009) 021013.
[18] Y. Li, H. Wu, Experiment investigation on flow boiling heat transfer in a bidirectional counter-flow microchannel heat sink, Int. J. Heat Mass Transf. 187 (2022) 122500.
[19] R. van Erp, R. Soleimanzadeh, L. Nela, G. Kampitsis, E. Matioli, Co-designing electronics with microfluidics for more sustainable cooling, Nature 585 (2020) 211–216.
[20] S. Lee, I. Mudawar, Transient characteristics of flow boiling in large micro-channel heat exchangers, Int. J. Heat Mass Transf. 103 (2016) 186–202.
[21] S. Lee, V.S. Devahdhanush, I. Mudawar, Frequency analysis of pressure oscillations in large length-to-diameter two-phase micro-channel heat sinks, Int. J. Heat Mass Transf. 116 (2018) 273–291.
[22] M.B. Bowers, I. Mudawar, Two-phase electronics cooling using mini-channel and micro-channel heat sinks: part 2-Flow rate and pressure drop constraints, ASME J. Electron. Packag. 116 (1994) 298–305.
[23] Y. Kim, D.L. O'Neal, A comparison of critical flow models for estimating two-phase flow of HCFC22 and HFC134a through short tube orifices, Int. J. Refrig. 18 (1995) 447–455.
[24] S.M. Kim, I. Mudawar, Review of two-phase critical flow models and investigation of the relationship between choking, premature CHF, and CHF in micro-channel heat sinks, Int. J. Heat Mass Transf. 87 (2015) 497–511.
[25] H.Y. Wu, P. Cheng, Boiling instability in parallel silicon microchannels at different heat flux, Int. J. Heat Mass Transf. 47 (2004) 3631–3641.
[26] W. Qu, I. Mudawar, Measurement and prediction of pressure drop in two-phase micro-channel heat sinks, Int. J. Heat Mass Transf. 46 (2003) 2737–2753.
[27] J. Xu, J. Zhou, Y. Gan, Static and dynamic flow instability of a parallel microchannel heat sink at high heat fluxes, Energy Convers. Manag. 46 (2005) 313–334.
[28] A. Koşar, C.J. Kuo, Y. Peles, Suppression of boiling flow oscillations in parallel microchannels by inlet restrictors, J. Heat Transf. 128 (2006) 251–260.
[29] G. Wang, P. Cheng, A.E. Bergles, Effects of inlet/outlet configurations on flow boiling instability in parallel microchannels, Int. J. Heat Mass Transf. 51 (2008) 2267–2281.
[30] Y.K. Prajapati, P. Bhandari, Flow boiling instabilities in microchannels and their promising solutions – A review, Exp. Therm. Fluid Sci. 88 (2017) 576–593.
[31] S. Lee, V.S. Devahdhanush, I. Mudawar, Investigation of subcooled and saturated boiling heat transfer mechanisms, instabilities, and transient flow regime maps for large length-to-diameter ratio micro-channel heat sinks, Int. J. Heat Mass Transf. 123 (2018) 172–191.
[32] S. Lee, V.S. Devahdhanush, I. Mudawar, Experimental and analytical investigation of flow loop induced instabilities in micro-channel heat sinks, Int. J. Heat Mass Transf. 140 (2019) 303–330.
[33] J. Lee, I. Mudawar, Critical heat flux for subcooled flow boiling in micro-channel heat sinks, Int. J. Heat Mass Transf. 52 (2009) 3341–3352.
[34] R. Revellin, J.R. Thome, A theoretical model for the prediction of the critical heat flux in heated microchannels, Int. J. Heat Mass Transf. 51 (2008) 1216–1225.
[35] F. Yang, X. Dai, Y. Peles, P. Cheng, J. Khan, C. Li, Flow boiling phenomena in a single annular flow regime in microchannels (I): Characterization of flow boiling heat transfer, Int. J. Heat Mass Transf. 68 (2014) 703–715.
[36] F. Yang, X. Dai, Y. Peles, P. Cheng, J. Khan, C. Li, Flow boiling phenomena in a single annular flow regime in microchannels (II): Reduced pressure drop and enhanced critical heat flux, Int. J. Heat Mass Transf. 68 (2014) 716–724.
[37] N. Borhani, B. Agostini, J. Thome, A novel time strip flow visualization technique for investigation of intermittent dewetting and dryout in elongated bubble flow in a microchannel, Int. J. Heat Mass Transf. 53 (2010) 4809–4818.
[38] S.M. Kim, I. Mudawar, Universal approach to predicting saturated flow boiling heat transfer in mini/micro-channels – Part I. Dryout incipience quality, Int. J. Heat Mass Transf. 64 (2013) 1226–1238.
[39] L. Hsu, S. Cion, K. Lin, C. Wang, An experimental study of inclination on the boiling heat transfer characteristics of a micro-channel heat sink using HFE-7100, Int. Commun. Heat Mass Transf. 62 (2015) 13–17.
[40] S.M. Kim, I. Mudawar, Universal approach to predicting two-phase frictional pressure drop for mini/micro-channel saturated flow boiling, Int. J. Heat Mass Transf. 58 (2013) 718–734.
[41] Z. Sun, CO2 flow boiling heat transfer in horizontal tubes, Ph.D. Thesis, Purdue University, West Lafayette, IN, 2001.
[42] L. Wojtan, T. Ursenbacher, J.R. Thome, Investigation of flow boiling in horizontal tubes: Part I – a new diabatic two-phase flow pattern map, Int. J. Heat Mass Transf. 48 (2005) 2955–2969.
[43] L. Cheng, G. Ribatski, L. Wojtan, J.R. Thome, New flow boiling heat transfer model and flow pattern map for carbon dioxide evaporating inside horizontal tubes, Int. J. Heat Mass Transf. 49 (2006) 4082–4094.
[44] D. Del Col, F. Fantini, L. Rossetto, Dryout quality in a minichannel flow boiling, in: XXV UIT National Heat Transfer Conf., Italy, 2007, pp. 18–20.
[45] L. Cheng, G. Ribatski, J.M. Quiben, J.R. Thome, New prediction methods for CO_2 evaporation inside tubes: part I – a two-phase flow pattern map and a flow pattern based phenomenological model for two-phase flow frictional pressure drops, Int. J. Heat Mass Transf. 51 (2008) 111–124.
[46] S. Jeong, D. Park, Evaporative heat transfer of CO_2 in a smooth and a micro-grooved miniature channel tube, Heat Transfer Eng 30 (2009) 582–589.
[47] D. Ducoulombier, S. Colasson, J. Bonjour, P. Haberschill, Carbon dioxide flow boiling in a single microchannel – Part II: heat transfer, Exp. Therm. Fluid Sci. 35 (2011) 597–611.
[48] R. Mastrullo, A.W. Mauro, J.R. Thome, D. Toto, G.P. Vanoli, Flow pattern maps for convective boiling of CO_2 and R410A in a horizontal smooth tube: experiments and new correlations analyzing the effect of the reduced pressure, Int. J. Heat Mass Transf. 55 (2012) 1519–1528.
[49] H. Mori, S. Yoshida, K. Ohishi, Y. Kakimoto, Dryout quality and post-dryout heat transfer coefficient in horizontal evaporator tubes, in: European Thermal Science Conf., Germany, 2000, pp. 839–844.
[50] J.G. Lavin, E.H. Young, Heat transfer to evaporating refrigerants in two-phase flow, AIChE J 11 (6) (1965) 1124–1132.
[51] N. Kattan, Contribution to the heat transfer analysis of substitute refrigerants in evaporator tubes with smooth or enhanced tube surfaces, PhD thesis, Dept. of Mechanical Engineering, Swiss Federal Institute of Technology Lausanne, CH-1015 Lausanne, Switzerland, 1996.
[52] J. Yang, H. Rivard, R. Zmeureanu, On-line building energy prediction using adaptive artificial neural networks, Energy Build 37 (2005) 1250–1259.
[53] G. Diaz, M. Sen, K.T. Yang, R.L. McClain, Dynamic prediction and control of heat exchangers using artificial neural networks, Int. J. Heat Mass Transf. 44 (2001) 1671–1679.
[54] A. Pacheco-Vega, M. Sen, K.T. Yang, R.L. McClain, Neural network analysis of fin-tube refrigerating heat exchanger with limited experimental data, Int. J. Heat Mass Transf. 44 (2001) 763–770.
[55] B.B. Ekici, U.T. Aksoy, Prediction of building energy consumption by using artificial neural networks, Adv. Eng. Softw. 40 (5) (2009) 356–362.
[56] Y. Islamoglu, A new approach for the prediction of the heat transfer rate of a single layer wire-on-tube type heat exchanger – use of an artificial neural network model, Appl. Therm. Eng. 23 (2003) 243–249.
[57] F. Jaliliantabar, R. Mamat, S. Kumarasamy, Prediction of lithium-ion battery temperature in different operating conditions equipped with passive battery thermal management system by artificial neural networks, Mater. Today: Proc 48 (2022) 1796–1804.
[58] Y. Wang, X. Chen, C. Li, Y. Yu, G. Zhou, C. Wang, W. Zhao, Temperature prediction of lithium-ion battery based on artificial neural network model, Appl. Therm. Eng. 228 (2023) 120482.
[59] J. Thibault, B.P.A. Grandjean, A neural network methodology for heat transfer data analysis, Int. J. Heat Mass Transf. 34 (1991) 2063–2070.
[60] K. Jambunathan, S.L. Hartle, S. Ashforth-Frost, V.N. Fontama, Evaluating convective heat transfer coefficients using neural networks, Int. J. Heat Mass Transf. 39 (1996) 2329–2332.
[61] Y. Qiu, D. Garg, L. Zhou, C.R. Kharangate, S.M. Kim, I. Mudawar, An artificial neural network model to predict mini/micro-channels saturated flow boiling heat transfer coefficient based on universal consolidated data, Int. J. Heat Mass Transf. 149 (2020) 1192211.
[62] L. Zhou, D. Garg, Y. Qiu, S.M. Kim, I. Mudawar, C.R. Kharangate, Machine learning algorithms to predict flow condensation heat transfer coefficient in mini/micro-channel utilizing universal data, Int. J. Heat Mass Transf. 162 (2020) 120351.

[63] A. Bard, Y. Qiu, C.R. Kharangate, R. French, Consolidated modeling and prediction of heat transfer coefficients for saturated flow boiling in mini/micro-channels using machine learning methods, Appl. Therm. Eng. 210 (2022) 118305.

[64] F. Nie, H. Wang, Y. Zhao, Q. Song, S. Yan, M. Gong, A universal correlation for flow condensation heat transfer in horizontal tubes based on machine learning, Int. J. Therm. Sci. 184 (2023) 107994.

[65] C. Li, M. Liu, N. Zhao, F. Wang, Z. Zhao, S. Guo, L. Fang, X. Li, Void fraction measurement using modal decomposition and ensemble learning in vertical annular flow, Chem. Eng. Sci. 247 (2022) 116929.

[66] C. Li, Y. Song, L. Xu, N. Zhao, F. Wang, L. Fang, X. Li, Prediction of the interfacial disturbance wave velocity in vertical upward gas-liquid annular flow via ensemble learning, Energy 242 (2022) 122990.

[67] A.T. Vu, S. Gulati, P.A. Vogel, T. Grunwald, T. Bergs, Machine learning-based predictive modeling of contact heat transfer, Int. J. Heat Mass Transf. 174 (2021) 121300.

[68] Y. He, C. Hu, B. Jiang, Z. Sun, J. Ma, H. Li, D. Tang, Data-driven approach to predict the flow boiling heat transfer coefficient of liquid hydrogen aviation fuel, Fuel 324 (2022) 124778.

[69] N. Qian, X. Wang, Y. Fu, Z. Zhao, J. Xu, J. Chen, Predicting heat transfer of oscillating heat pipes for machining processes based on extreme gradient boosting algorithm, Appl. Therm. Eng. 162 (2020) 114521.

[70] R. Ali, B. Palm, Dryout characteristics during flow boiling of R134a in vertical circular minichannels, Int. J. Heat Mass Transf. 54 (2011) 3437–3446.

[71] J.G. Collier, J.R. Thome, Convective Boiling and Condensation, Oxford University Press, 1994, 3rd Ed.

[72] W. Qu, I. Mudawar, Measurement and correlation of critical heat flux in two-phase micro-channel heat sinks, Int. J. Heat Mass Transf. 47 (2004) 2045–2059.

[73] W.P. Baek, S.H. Chang, KAIST CHF data, Personal communication, 8, Korea Advanced Institute of Science and Technology, Taejon, South Korea, 1997.

[74] K.M. Becker, Burnout Measurements in Vertical Round tubes, Effect of Diameter, Aktiebolaget Atomenergi, 1970. AE-TPM-RL-1269.

[75] H.C. Kim, W.P. Baek, S.H. Chang, Critical heat flux of water in vertical round tubes at low pressure and low flow conditions, Nucl. Eng. Des. 199 (2000) 49–73.

[76] A.M. Lezzi, A. Niro, G.P. Beretta, Experimental data on CHF for forced convection water boiling in long horizontal capillary tubes, in: Proc. 10th Int. Heat Transfer Conf., vol. 7, UK, 1994, pp. 491–496.

[77] G.M. Roach Jr., S.I. Abdel-Kahlik, S.M. Ghiaasiaan, M.F. Dowling, S.M. Jeter, Low-flow critical heat flux in heated microchannels, Nucl. Sci. Eng. 131 (1999) 411–425.

[78] W. Yu, D.M. France, M.W. Wambsganss, J.R. Hull, Two-phase pressure drop, boiling heat transfer, and critical heat flux to water in a small-diameter horizontal tube, Int. J. Multiphase Flow 28 (2002) 927–941.

[79] W. Qu, I. Mudawar, Flow boiling heat transfer in two-phase micro-channel heat sinks - I. Experimental investigation and assessment of correlation methods, Int. J. Heat Mass Transf. 46 (2003) 2755–2771.

[80] S. Lee, I. Mudawar, Enhanced model for annular flow in micro-channel heat sinks, including effects of drop entrainment/deposition and core turbulence, Int. J. Heat Mass Transf. 133 (2019) 510–530.

[81] D. Del Col, S. Bortolin, Investigation of dryout during flow boiling in a single microchannel under non-uniform axial heat flux, Int. J. Therm. Sci. 57 (2012) 25–36.

[82] A. Greco, Convective boiling of pure and mixed refrigerants: an experimental study of the major parameters affecting heat transfer, Int. J. Heat Mass Transf. 51 (2008) 896–909.

[83] E. Hihara, C. Dang, Boiling heat transfer of carbon dioxide in horizontal tubes, in: Proc. 2007 ASME-JSME Thermal Eng. Summer Heat Transfer Conf., Canada, HT2007-32885, 2007, pp. 843–849.

[84] T.G. Karayiannis, M.M. Mahmoud, D.B.R. Kenning, A study of discrepancies in flow boiling results in small to microdiameter metallic tubes, Exp. Therm. Fluid Sci. 36 (2012) 126–142.

[85] M. Li, C. Dang, E. Hihara, Flow boiling heat transfer of HFO1234yf and R32 refrigerant mixtures in a smooth horizontal tube: part I. Experimental investigation, Int. J. Heat Mass Transf. 55 (2012) 3437–3446.

[86] C. Martin-Callizo, Flow boiling heat transfer in single vertical channel of small diameter, Ph.D. Thesis, Royal Institute of Technology, Sweden, 2010.

[87] H.K. Oh, C.H. Son, Evaporation flow pattern and heat transfer of R-22 and R-134a in small diameter tubes, Heat Mass Transf 47 (2011) 703–717.

[88] H.K. Oh, C.H. Son, Flow boiling heat transfer and pressure drop characteristics of CO in horizontal tube of 4.57-mm inner diameter, Appl. Therm. Eng. 31 (2011) 163–172.

[89] J.T. Oh, A.S. Pamitran, K.I. Choi, P. Hrnjak, Experimental investigation on two-phase flow boiling heat transfer of five refrigerants in horizontal small tubes of 0.5, 1.5 and 3.0 mm inner diameters, Int. J. Heat Mass Transf. 54 (2011) 2080–2088.

[90] H. Ohta, K. Inoue, M. Ando, K. Watanabe, Experimental investigation on observed scattering in heat transfer characteristics for flow boiling in a small diameter tube, Heat Transfer Eng 30 (2009) 19–27.

[91] S. Saitoh, H. Daiguji, E. Hihara, Effect of tube diameter on boiling heat transfer of R-134a in horizontal small-diameter tubes, Int. J. Heat Mass Transf. 48 (2005) 4973–4984.

[92] D. Shiferaw, Two-phase flow boiling in small- to micro-diameter tubes, Ph.D. Thesis, Brunel University, UK, 2008.

[93] C.B. Tibirica, G. Ribatski, J.R. Thome, Flow boiling characteristics for R1234ze(E) in 1.0 and 2.2 mm circular channels, ASME J. Heat Transfer 134 (2012) 020906.

[94] L. Wang, M. Chen, M. Groll, Flow boiling heat transfer characteristics of R134a in a horizontal mini tube, J. Chem. Eng. Data 54 (2009) 2638–2645.

[95] J. Wu, T. Koettig, Ch. Franke, D. Helmer, T. Eisel, F. Haug, J. Bremer, Investigation of heat transfer and pressure drop of CO_2 two-phase flow in a horizontal minichannel, Int. J. Heat Mass Transf. 54 (2011) 2154–2162.

[96] Y. Yang, Y. Fujita, Boiling Heat Transfer in Rectangular Channels of Small gaps, Memoirs of the Faculty of Engineering, 62, Kyushu University, 2002, pp. 223–239.

[97] R. Yun, Y. Kim, M.S. Kim, Convective boiling heat transfer characteristics of CO_2 in microchannels, Int. J. Heat Mass Transf. 48 (2005) 235–242.

[98] J. Lee, I. Mudawar, Two-phase flow in high-heat-flux micro-channel heat sink for refrigeration cooling applications: Part II—Heat transfer characteristics, Int. J. Heat Mass Transf. 48 (2005) 941–955.

[99] J.B. Copetti, M.H. Macagnan, F. Zinani, N.L.F. Kunsler, Flow boiling heat transfer and pressure drop of R-134a in a mini tube: an experimental investigation, Exp. Therm. Fluid Sci. 35 (2011) 636–644.

[100] P. Saha, M. Ishii, N. Zuber, An experimental investigation of the thermally induced flow oscillations in two-phase systems, J. Heat Transfer 98 (1976) 616–622.

[101] R. Revellin, J.R. Thome, A theoretical model for the prediction of the critical heat flux in heated microchannels, Int. J. Heat Mass Trans. 51 (2008) 1216–1225.

[102] A.J. Myles, R.N. Feudale, Y. Liu, N.A. Woody, S.D. Brown, An introduction to decision tree modeling, J. Chemom. 18 (6) (2004) 275–285.

[103] J.R. Quinlan, Induction of decision trees, Mach. Learn. 1 (1) (1986) 81–106.

[104] L. Breiman, Classification and Regression Trees, Routledge, 2017, 1st Ed.

[105] Y.Y. Song, Y. Lu, Decision tree methods: applications for classification and prediction, Shanghai Arch. Psychiatry 27 (2) (2015) 130–135.

[106] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, J. Comput. Syst. Sci. 55 (1) (1997) 119–139.

[107] H. Drucker, Improving regressors using boosting techniques, in: Proc. 14th Int. Mach. Learn. Conf., vol. 97, 1997, pp. 107–115.

[108] T. Hastie, S. Rosset, J. Zhu, H. Zou, Multi-class adaboost, Stat. interface 2 (3) (2009) 349–360.

[109] J.H. Friedman, Greedy boosting approximation: a gradient boosting machine, Ann. Stat. 29 (2006) 1189–1232.

[110] A. Natekin, A. Knoll, Gradient boosting machines, a tutorial, Front. Neurorob. 7 (2013) 1–21.

[111] T. Chen, C. Guestrin, Xgboost: a scalable tree boosting system, in: Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2016, pp. 785–794.

[112] E.W. Lemmon, M.L. Huber, M.O. McLinden, Reference Fluid Thermodynamic and Transport Properties – REFPROP Version 9.0, NIST, MD, 2010.

[113] A. Kosar, C. Kuo, Y. Peles, Boiling heat transfer in rectangular microchannels with reentrant cavities, Int. J. Heat Mass Transf. 48 (2005) 4867–4886.

[114] C. Wade, K. Glynn, Hands-On Gradient Boosting With XGBoost and scikit-learn: Perform accessible Machine Learning and Extreme Gradient Boosting With Python, Packt Publishing, UK, 2020.

[115] R.G. Mantovani, A.L.D. Rossi, J. Vanschoren, B. Bischl, A.C.P.L.F. de Carvalho, To tune or not to tune: Recommending when to adjust SVM hyper-parameters via meta-learning, in: ICJNN, Ireland, 2015, pp. 1–8.

[116] N. Niklas, P. Davidsson, Quantifying the impact of learning algorithm parameter tuning, in: AAAI, vol.6, 2006, pp. 395–400.

[117] B. Bischl, B. Binder, M. Lang, T. Pielok, J. Richter, S. Coors, J. Thomas, T. Ullmann, M. Becker, A.L. Boulesteix, D. Deng, M. Lindauer, Hyperparameter optimization: foundations, algorithms, best practices, and open challenges, WIREs Data Min. Knowl. Discov. 13 (2) (2023) e1484.

[118] J. Wu, X.Y. Chen, H. Zhang, L.D. Xiong, H. Lei, S.H. Deng, Hyperparamter optimization for machine learning models based on Bayesian optimization, J. Electron. Sci. Technol. 17 (1) (2019) 26–40.

[119] B. Komer, J. Bergstra, C. Eliasmith, Hyperopt-sklearn: automatic hyperparamter configuration for scikit-learn, in: ICML workshop on AutoML, vol.9, Austin, 2014, pp. 50.

[120] P. Koch, O. Golovidov, S. Gardner, B.Wujek, J. Griffin, Y. Xu, Autotune: A derivative-free optimization framework for hyperparameter tuning, in: Proc. 24th ACM SIGKDD Int. Data Min. Knowl. Discovery Conf., 2018, pp. 443–452.

[121] D. Golovin, B. Solnik, S. Moitra, G. Kochanski, Google vizier: A service for black-box optimization, in: Proc. 23rd ACM SIGKDD Int. Data Min. Knowl. Discovery Conf., 2017, pp. 1487–1495.

[122] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: A next-generation hyperparameter optimization framework, in: Proc. 25th ACM SIGKDD Int. Data Min. Knowl. Discovery Conf., Anchorage, AK, USA, July 2019, pp. 2623–2631.