

RESEARCH ARTICLE | SEPTEMBER 30 2024

Predicting sessile droplet evaporation kinetics via cascaded deep networks and tree-based machine learning approach

Arnov Paul  ; Purbarun Dhar  



Physics of Fluids 36, 092031 (2024)

<https://doi.org/10.1063/5.0230332>



View
Online



Export
Citation

Articles You May Be Interested In

Evaporation kinetics of sessile droplets morphed by substrate curvature

Physics of Fluids (December 2021)

Transients of Marangoni and Stefan advection dynamics during generic sessile droplet evaporation

Physics of Fluids (October 2023)

Substrate concavity influenced evaporation mechanisms of sessile droplets

Physics of Fluids (August 2021)

Predicting sessile droplet evaporation kinetics via cascaded deep networks and tree-based machine learning approach

Cite as: Phys. Fluids **36**, 092031 (2024); doi: [10.1063/5.0230332](https://doi.org/10.1063/5.0230332)

Submitted: 23 July 2024 · Accepted: 8 September 2024 ·

Published Online: 30 September 2024



View Online



Export Citation



CrossMark

Arnov Paul^{a)}  and Purbarun Dhar^{b)} 

AFFILIATIONS

Hydrodynamics and Thermal Multiphysics Lab (HTML), Department of Mechanical Engineering, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India

^{a)}Electronic mail: arnov0381@gmail.com

^{b)}Author to whom correspondence should be addressed: purbarun@mech.iitkgp.ac.in

ABSTRACT

The intricate nature of sessile droplet evaporation phenomenon makes detailed experimental studies time consuming and requires sophisticated apparatus; while complete numerical simulations are computationally expensive and also time-intensive. In this article, for the first time, we explore the applicability of machine learning (ML) approaches to predict the evaporation kinetics of generalized sessile droplets under various conditions. An in-house dataset, obtained using an experimentally well validated numerical model, is used to develop the ML models: deep artificial neural network (ANN) and decision tree algorithms such as Random Forest (RF) and extreme gradient boosting (XGB). The structures of these models are modified by cascading the output features according to the physics involved. This distinctive approach results in better prediction of the evaporation kinetics than basic ML models. The models are trained by a large set of input parameters for the target variables: viz. droplet evaporation rate, velocity scale, and temperature drop in the solid and liquid domain. Finally, the performance of these ML models is assessed by comparing their predictions with that of physics-based, experimentally validated, numerical models. Results show that the inclusion of additional features obtained using feature engineering significantly improve the prediction performance of ML algorithms, and consistently accurate predictions of droplet evaporation kinetics are obtained. Among various algorithms considered here, the ANN outperforms in term of various error matrices for most of the cases, followed by XGB, and RF models. Also, the highest mean average error (MAE) yield by the ML models for evaporation rate, velocity scale and temperature drop in liquid remains within $\sim 12.5\%$. In the case of temperature drop for the solid, the MAE is considerably higher due to large variability of the same target variable. Overall, the work clearly shows that ML algorithms can be used to obtain physically consistent predictions for sessile droplet evaporation parameters.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0230332>

24 October 2024 18:46:36

I. INTRODUCTION

The evaporation problem of sessile droplets is relevant to several natural, industrial, and practical applications. For instance, it plays a crucial role in fog, dew, and rain formation mechanisms.^{1,2} It is pertinent to cooling systems,³ printing technology,⁴ biomedical applications,⁵ and microfluidic devices.⁶ For this reason, their evaporation induced transport phenomena have been widely studied widely (see Refs. 7 and 8 for a review). These studies report that the sessile droplet evaporation behavior greatly depend on droplet-substrate geometry,^{9–11} wetting state,^{12,13} ambient conditions,¹⁴ and thermophysical properties of the liquid, and solid substrate.^{15,16} For instance, the mass loss rate remains nearly constant during evaporation of droplet over

hydrophilic surfaces;¹² whereas on hydrophobic and superhydrophobic (SH) surfaces, mostly nonlinear evaporation rate are observed, especially at latter stages of evaporation.¹³ The dynamic evolution of the contact line during evaporation also significantly depends on the surface energy of underlying substrate. Bormashenko *et al.*¹⁷ reported that droplets evaporate mostly via constant contact radius (CCR) mode on surfaces with high surface energy, whereas on low surface energy surfaces the constant contact angle (CCA) or mixed mode is dominant.

In addition, the evaporation rate also depends on the thermal conductively of the substrate: the mass loss rate is higher on thermally conductive substrates due to elevated droplet base temperature.¹⁵ This

altered mass loss rate substantially affects the evaporative cooling and the corresponding internal advection. For instance, the evaporative cooling is substantially enhanced on thermally conductive substrates.¹⁶ The effects are further amplified on SH surfaces due to centrally concentrated mass flux. Here, the large evaporative flux at the apex region cools down the droplet core to a greater extent than hydrophilic surfaces.¹⁸ This thermal imbalance results in buoyancy driven advection within the evaporating droplets, and the temperature gradient at the liquid–vapor interface triggers Marangoni flow. This bulk motion of fluid affects heat transport phenomenon and thereby the internal temperature distribution in droplets.^{16,19} This altered thermal profile ultimately influences the diffusive mass transport phenomenon due to strong dependence of saturated vapor concentration on temperature conditions. The mass transfer rate may further be affected by the buoyancy driven flow due to evaporative cooling effects in the gaseous domain.²⁰ In this context, Carle *et al.*²¹ showed that the ambient convective transport in the vapor phase has significant contribution to the overall mass transfer rate during sessile drop evaporation.

In recent times, machine learning (ML) algorithms have made significant inroads into fields of fluid dynamics and thermal sciences research,^{22–25} due its caliber in performing complex nonlinear regression tasks. Several recent studies^{26,27} have shown that ML models can effectively resolve heat conduction problems and accurately predict the transport coefficients. In fact, even in the presence of nonlinear convective effects, the artificial neural network (ANN) model can yield good prediction accuracy with a mean average error (MAE) of within 12.5%.^{27,28} Jadrich *et al.*²⁹ applied principal component analysis (PCA) and have shown that PCA autonomously discovers parameters to quantify the solid–liquid and liquid–gas phase transition phenomena. To predict frictional pressure drop in two-phase flow, Khosravi *et al.*³⁰ applied a number of ML models such as neural network, support vector machines, and group method of data handling with the problem geometry and initial conditions as input features. Later, Qiu *et al.*³¹ reported that incorporating appropriate dimensionless numbers as the input features significantly improves the prediction accuracy of ML models. Nie *et al.*³² trained ANN and extreme gradient boosting (XGB) models to generalize this pressure drop prediction in horizontal tubes. They have shown that these ML models can predict the unknown data from previous studies with a mean absolute deviation of $\sim 8.59\%$ and regression fit (R^2) of ~ 0.988 or higher.

The volume of literature on sessile droplet evaporation problem substantiates the strong influence of several input parameters on the transport phenomena in a complicated, nonlinear fashion. This intricate nature of a problem with multiple parameters necessitates sophisticated and time-intensive experimental studies, which needs advanced instrumentation to precisely regulate environmental conditions and capture real-time data. Further, due to the presence of highly coupled transient nonlinear transport processes, the numerical modeling of the same becomes computationally expensive. A trade-off between the numerical accuracy, experimental and computational resources can be realized using ML models to predict droplet evaporation behavior under varied conditions, from a set of not-so-extensive data. Here, we employ 600 such in-house numerical data points to train several ML algorithms: deep ANN and decision tree structures based on random forest (RF) and XGB. The simulations are performed under varied input parameters in a fully coupled manner in COMSOL Multiphysics platform. The ML models are trained by a large set of input parameters

for the target variables as: droplet evaporation rate, internal velocity scale and temperature drop in solid and liquid domain. The performance of these models is assessed using several error matrices to analyze how the ML algorithms fare with the conventional physics-based models. It is worthwhile to mention that due to the limited number of data points considered, the entire training and testing session of the ML models can be performed in minutes, even on standard desktop computers. The primary advantage of this approach is the extremely low computation cost compared to conventional physics-based numerical simulations.

II. METHODS

A. Data base constructed from numerical study

1. Description of physical problem

We consider the evaporation phenomena of a sessile droplet on solid substrates in an initially quiescent environment. The problem involves coupled transport of mass, momentum, and energy across the evolving liquid–vapor interface of drying droplets along with other interfacial features such as contact line dynamics, Marangoni flow, natural convection. We formulate the physical problem in r - z coordinate as shown in Fig. 1(a). The initial profile of the liquid–vapor interface of sessile droplets under varied conditions is obtained using level-set (LS) method similar to our recent study.³³ In this approach, the level-set (LS) equations coupled with the Navier–Stokes (NS) equations are solved with appropriate boundary condition given as:³⁴

$$\frac{\partial \psi}{\partial t} + \nabla(u_{LS}\psi) = \xi \nabla \cdot \left(\epsilon \nabla \psi - \left[\psi(1-\psi) \frac{\nabla \psi}{|\nabla \psi|} \right] \right), \quad (1)$$

$$\rho \left(\frac{\partial u_{LS}}{\partial t} + (u_{LS} \cdot \nabla) u_{LS} \right) = \nabla \cdot \left[-pI + \mu \left(\nabla u_{LS} + (\nabla u_{LS})^T \right) \right] + \rho g, \quad (2)$$

$$\sigma_{nt} = \sigma_n - (\sigma_n \cdot \vec{n}_{wall}) \vec{n}_{wall}$$

$$= \gamma \delta(\vec{n}_{int} \cdot \vec{n}_{wall} - \cos(\theta_0)) (\vec{n}_{int} \cdot \vec{t}_{wall}) \vec{t}_{wall} - \frac{\mu}{\beta} u_{slip}, \quad (3)$$

$$u_{slip} = u_{LS} - (u_{LS} \cdot \vec{n}_{wall}) \vec{n}_{wall}, \quad (4)$$

$$\vec{n}_{wall} \left(\epsilon \nabla \psi - \psi(1-\psi) \frac{\nabla \psi}{|\nabla \psi|} \right) = 0. \quad (5)$$

Here, the level set parameter ψ denotes the mass fraction of liquid phase. Also, u_{LS} , u_{slip} , β , ϵ , ξ , g represents the velocity in level set method, slip velocity, slip length, parameter controlling interfacial thickness, re-initialization velocity, and gravitational acceleration, respectively. The tangential and normal component of viscous stress tensor are expressed as $\sigma_{nt} = \sigma_n - (\sigma_n \cdot \vec{n}_{wall}) \vec{n}_{wall}$ and σ_n , respectively. Also, the unit vectors along normal and tangential direction of the solid substrate are given by \vec{n}_{wall} and \vec{t}_{wall} , respectively. The position of the liquid–vapor interface is always located at zero value of the level set parameter ($\psi = 0$). Also, it is worth mentioning that we employ smeared out Heaviside function given by $\rho = \rho_l + (\rho_v - \rho_l)H(\Phi)$ (where $H(\zeta)$ is the Heaviside unit step function) to render a smooth change in thermophysical properties at the liquid–vapor interface. A more detailed discussion regarding the selection of various parameters in the LS method can be obtained from our recent study.³³

Next, this numerically obtained initial profile of sessile droplet serves as the input for geometrical specifications in the evaporation

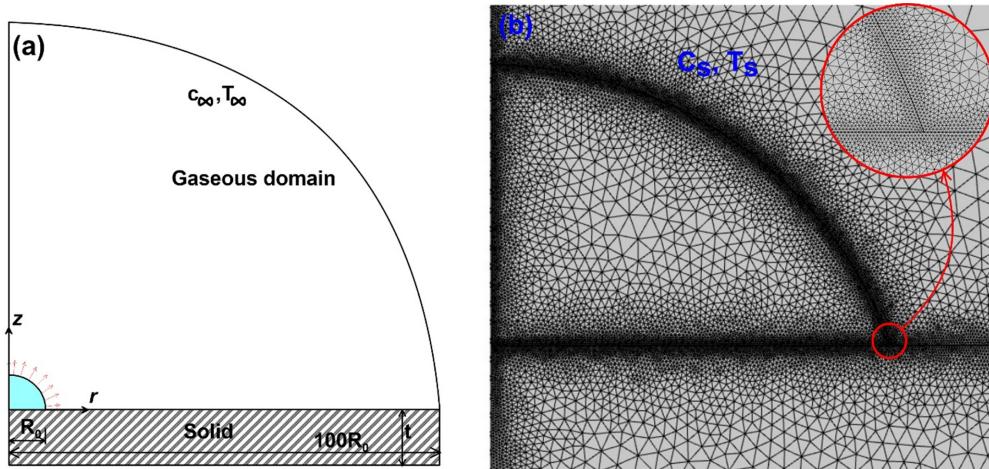


FIG. 1. (a) Schematic of the sessile droplet geometric model considered; (b) distribution of the triangular meshes near the droplet interfaces (The inset figure shows mesh details near the contact line).

problem. The evaporation phenomenon is mathematically modeled in r - z coordinate [see Fig. 1(a)] based on the Arbitrary Lagrangian–Eulerian (ALE) framework.^{35,36} The transport of mass, momentum and energy during the evaporation phenomenon can be modeled using the advection-diffusion equation, NS equations and energy equations given as:

$$\frac{\partial c}{\partial t} + (u_c \cdot \nabla)c = \nabla \cdot [D\nabla c], \quad (6)$$

$$\nabla \cdot U = 0, \quad (7)$$

$$\rho \left(\frac{\partial U}{\partial t} + (u_c \cdot \nabla)U \right) = \nabla \cdot [-pI + \mu(\nabla U + (\nabla U)^T)] + G, \quad (8)$$

$$\rho C_p \left(\frac{\partial T}{\partial t} + (u_c \cdot \nabla)T \right) = \nabla \cdot [k\nabla T] \quad (\text{For liquid and gaseous domain}), \quad (9)$$

$$\rho C_p \left(\frac{\partial T}{\partial t} \right) = \nabla \cdot [k\nabla T] \quad (\text{For solid domain}). \quad (10)$$

In these equations D , C_p , k , μ , ρ , u_c , c , p , T represents diffusivity of water vapor in air, specific heat, thermal conductivity, viscosity, density, convection velocity in ALE framework, vapor concentration, pressure, and temperature, respectively. In the NS equations, the gravity force is considered using the term: $G = (\rho + \Delta\rho)g$; where $\Delta\rho$ is the change in density due to thermal effects. In the case of liquid and gaseous domain this change in density is modeled as:

$$\Delta\rho_l = \alpha\rho_{l\infty}(T - T_{\infty}), \quad (11)$$

$$\Delta\rho_g = \left(\frac{P_g M_a}{RT} + cM_l \right) - \left(\frac{P_{g\infty} M_a}{RT_{\infty}} + c_{\infty} M_l \right). \quad (12)$$

Here, $P_g = P_0 - cRT$ is the partial pressure of air-vapor mixture; where P_0 represents far field atmospheric pressure. Also α , R , M_a , and M_l denotes thermal expansion coefficient (K^{-1}) of water, universal gas constant, molecular weight of air and water vapor, respectively. The subscripts "g," "l," and " ∞ " denotes the liquid phase, gaseous phase,

and a reference state at far field ambient, respectively. Also, it is worth mentioning that, this natural convection model is similar to Boussinesq's approximation considering the fact that the change in density does not tamper the incompressible flow assumptions.³⁷

These governing equations for the evaporating droplets during both CCR and CCA mode are solved by applying appropriate boundary conditions. It can be mentioned that the CCR mode is primarily considered for $\theta \leq \pi/2$ whereas for $\theta > \pi/2$, we assume CCA model to maintain physical consistency with previous works.^{11,17} Under such assumptions, we consider the presence of saturated vapor concentration at the liquid–vapor interface:

$$c_{sat}(r) = \frac{P_{sat}}{RT_i} = \exp \left[A - \frac{B}{T_i(r) - C} \right] \times \frac{10^6}{RT_i(r)}. \quad (13)$$

Here, $T_i(r)$ is the interfacial temperature in K. Also, A , B , and C are component specific constants from Antoine equation and for water droplets the values are given as $A = 9.487$, $B = 3893$ K, and $C = 42.68$ K.³⁸ At the far field, the vapor concentration depends on the relative humidity (ϕ) conditions given as: $c_{\infty} = \phi c_s$. This difference in vapor concentration results in interfacial mass transfer across droplet surface:

$$J = \vec{n} \cdot (-D\nabla c + U \cdot c)M_l, \quad (14)$$

$$U_g - U_l = J \left(\frac{1}{\rho_l} - \frac{1}{\rho_g} \right). \quad (15)$$

Here, J is the interfacial evaporative flux and U_g and U_l denotes the velocity component in liquid and gaseous domain, respectively. This evaporated liquid molecules take away latent heat of vaporization for which we apply energy balance across the droplet surface:

$$n \left(k_g (\nabla T)_g - k_l (\nabla T)_l \right) = -JL_h, \quad (16)$$

where L_h is the latent heat of vaporization. This thermal imbalance triggers thermal Marangoni flow or buoyancy driven advection in both

liquid and gaseous domain. In such conditions, the normal and tangential stresses are balanced at the liquid–vapor interface:

$$(\vec{n}_{\text{int}} \cdot S_g - \vec{n}_{\text{int}} \cdot S_l) \cdot \vec{n}_{\text{int}} = \frac{\gamma}{r_o} \cdot \vec{n}_{\text{int}}, \quad (17)$$

$$(\vec{n}_{\text{int}} \cdot S_g - \vec{n}_{\text{int}} \cdot S_l) \cdot \vec{t}_{\text{int}} = -\gamma_T \nabla_g T, \quad (18)$$

where $\nabla_g = (I - \vec{n} \cdot \vec{n}^T) \nabla$, γ , S , t_{int} , n_{int} , r_o , γ_T are the surface gradient operator, surface tension, stress tensor, unit vector along tangential and normal direction at interface, overall curvature radius and temperature derivative of surface tension, respectively. Also, \vec{n}_{int} and \vec{t}_{int} represents unit vectors along normal and tangential direction to the interface. In addition, we apply no penetration vapor condition at the solid–vapor interface in the gaseous domain and no slip velocity condition at solid surface in both liquid and gaseous domain. The temperate at the substrate base (T_s) and far field (T_∞) are considered as constant: $T_s = T_\infty = 293.15\text{K}$. The initial values of velocity and relative humidity are considered as 0 m/s, and 50%, respectively. The numerical values of different thermophysical properties of water vapor and air is obtained from standard handbook.³⁹

2. Numerical simulations and validation

In present study, the numerical simulations are performed in two stages. In the first step, the sessile droplet profiles are obtained using

the level set method. In these simulations, we specify the droplet volume and the angle at three-phase contact line to iteratively obtain the initial stable profile of droplet interface under influence of gravity. In the next step, this stable profile of sessile droplet is used to build the droplet geometry for evaporation problem. The corresponding governing equations are solved in a fully coupled manner based on Arbitrary Lagrangian–Eulerian (ALE) framework. This approach allows us to precisely trace the position of liquid–vapor interface during evaporation under varied contact line dynamics (i.e., both CCA and CCR mode). However, appropriate boundary conditions are required for these different modes of evaporation. For instance, in the case of CCR mode, we apply zero mesh mobility at the contact line as the same is pinned to the solid walls. For the CCA mode, the contact line speed is specified along with the constant contact angle at solid–liquid interface. For this purpose, the tangential velocity components at the nearest grid points from wall are conveniently used.⁴⁰ A more detailed discussion regarding the ALE implementation and boundary conditions near contact line can be found in our recent studies.^{19,33}

For both the simulations, we employ transient, fully coupled solver provided in the commercial package COMSOL Multiphysics. The computational domain is considered as a hemisphere with radius $100r_0$; where r_0 is the initial radius of sessile droplets. This geometry was meshed using unstructured triangular meshes with the finest mesh

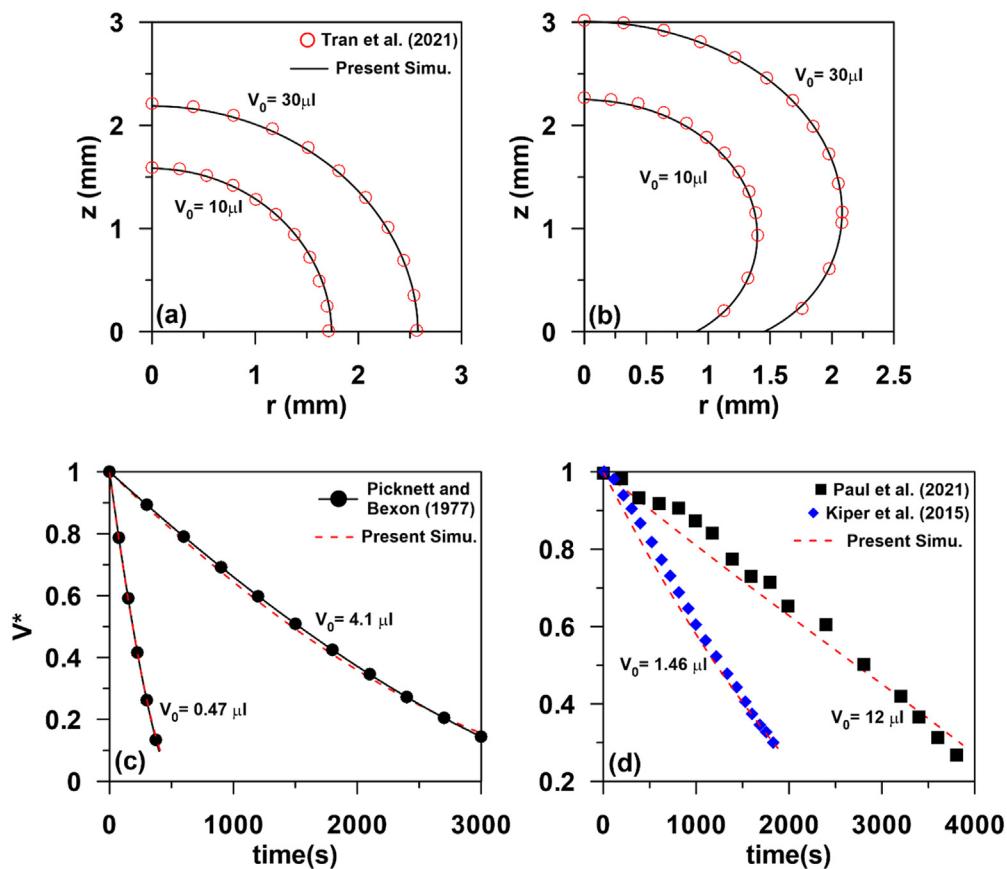


FIG. 2. Validation of initial stable profiles of sessile droplets obtained from present study with that of Tran *et al.*⁴¹ for contact angles (a) $\pi/2$ and (b) $5\pi/6$; Comparisons of the transient sessile droplet volume obtained from present study with that of previous works for (c) CCA and (d) CCR mode of evaporation.

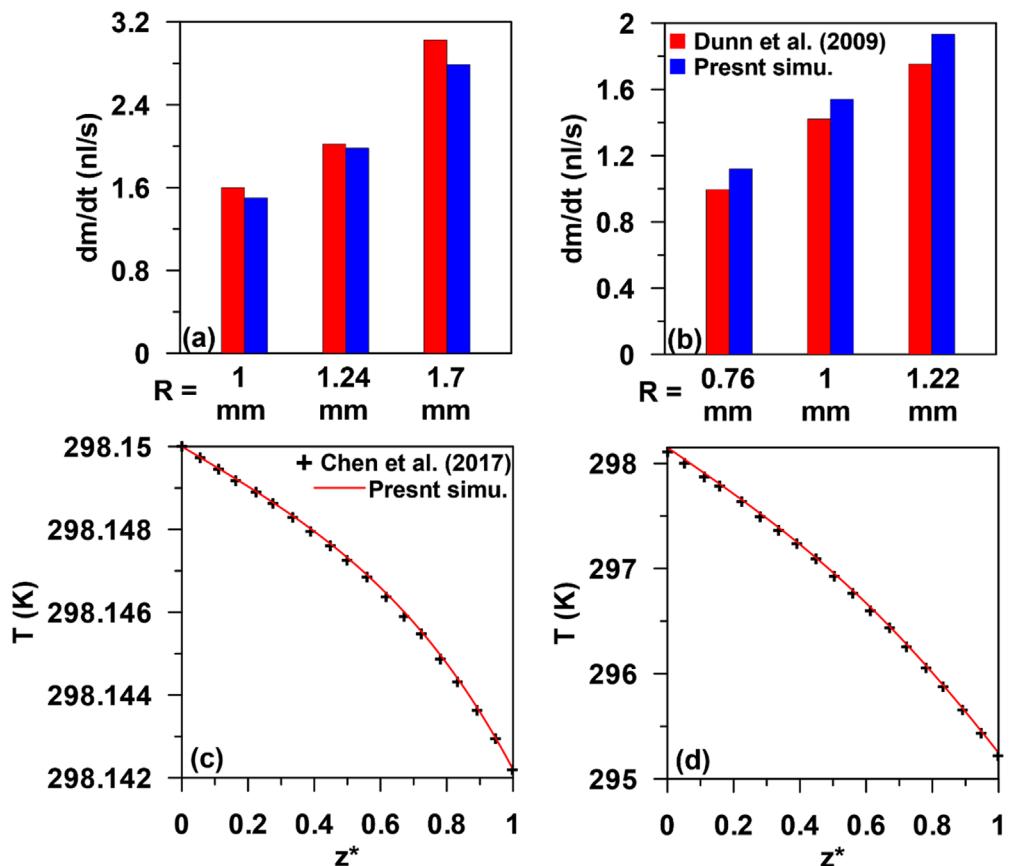


FIG. 3. Comparisons of the evaporation rate and interfacial temperature distribution of drying droplets obtained from present simulations with that of previous works over (a) and (c) aluminum and (b) and (d) ptfte substrate.

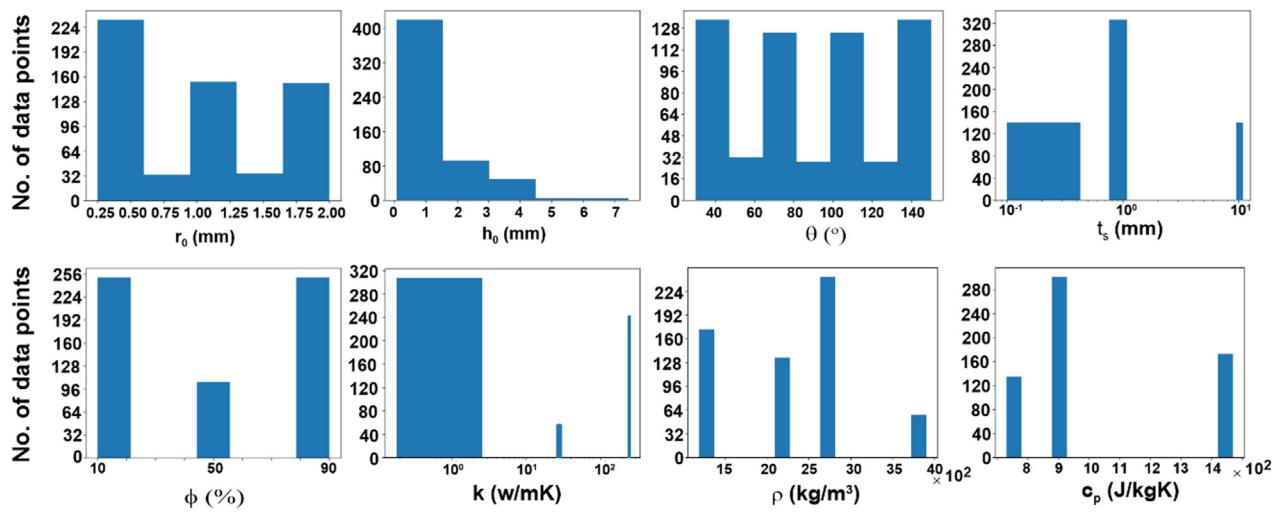


FIG. 4. Distribution of the number of data points with respect to geometric features (r_0 , h_0 , θ , t_s), ambient conditions (ϕ), and thermophysical properties of solid substrate (k , ρ , c_p).

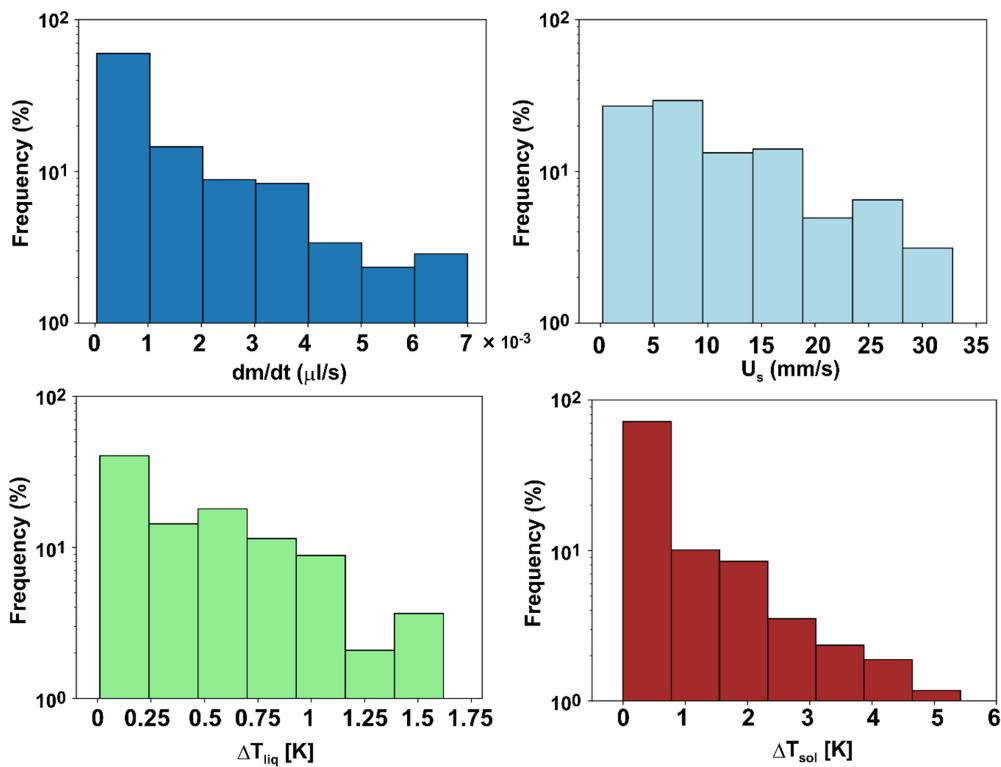


FIG. 5. Distribution of the target variables in dataset: evaporation rate (dm/dt), velocity scale (U_s), temperature drop in liquid (ΔT_{liq}), and solid domain (ΔT_{sol}).

(about 200 times finer than the droplet radius) being located near the liquid–vapor interface and at the three-phase contact line [see Fig. 1(b)]. The time stepping was considered as automatic based on backward difference formula (BDF) with the maximum time step was limited to 0.1 s. A more details about parameter discretization, stability, numerical accuracy, and mesh independence results can be obtained in our recent study.¹⁹

To validate present LS model, we compare the initial stable profiles of sessile droplet obtained from present study with that of Tran *et al.*⁴¹ under varied conditions as shown in Figs. 2(a) and 2(b). For the ALE based evaporation model, the instantaneous droplet volume during evaporation obtained from present study is compared with that of Picknett and Bexon,⁴² Paul *et al.*,⁴³ and Kiper *et al.*⁴⁴ Figures 2(c) and 2(d) shows these comparisons for CCA and CCR mode of evaporation,

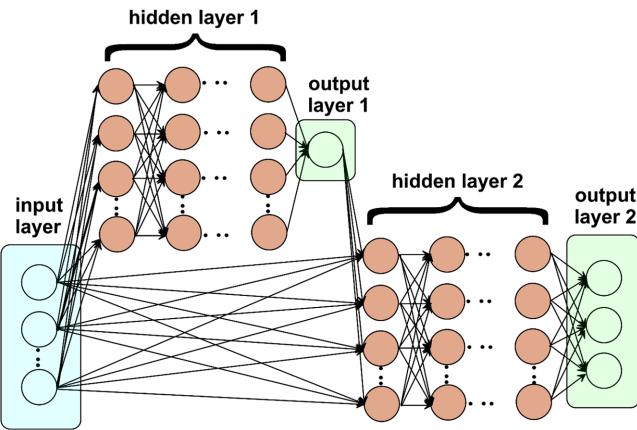


FIG. 6. Schematic representation of cascaded ANN architecture employed in present study.

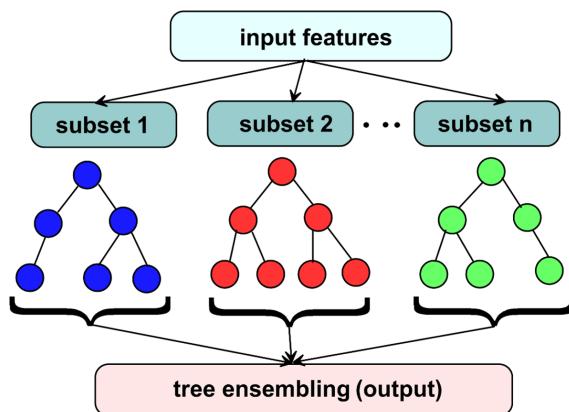


FIG. 7. Schematic representation of simplified decision tree structure.

respectively. It is observed that the droplet profile and evaporation rate obtained from present model agrees quite well with the previous works under varied conditions. For the CCR mode, we observe [see Fig. 2(d)] marginal deviation between the present simulation results and previous experimental findings. This discrepancy may be attributed to the inaccuracies involved in experiments as reported by Paul *et al.*⁴³ We also notice that the V^* vs t plot remains fairly linear up to $V^* \approx 0.5$ indicating nearly constant evaporation rate at this initial regime. We compare this evaporation rate obtained from present study with that of Dunn *et al.*¹⁵ over aluminum and Teflon substrates [see Figs. 3(a) and 3(b)]. The temperature drops along the liquid–vapor interface for droplets evaporating over these substrates are also compared with the numerical results of Chen *et al.*¹⁶ at stable conditions as shown in Figs. 3(c) and 3(d). We notice that the present results are in good agreement with the previous works under varied thermophysical properties of underlying substrate. Further, a detailed validation of

the internal velocity field and temperature distribution has already been reported in our recent study⁴⁵ and will not be repeated here for brevity.

3. Dataset distribution

We vary the parametric features in the droplet evaporation problem to obtain a consolidated database for developing ML models. For instance, in the case of geometric features, we vary the initial droplet contact radius (r_0), contact angle (θ), substrate thickness (t_0) in the range 0.25–2 mm, 30°–150°, 0.1–10 mm, respectively. Similarly, the ambient relative humidity (ϕ) was varied over a wide range 10–90%. Also, we consider aluminum, alumina, glass and acrylic as the underlying solid substrate to consider the influence of the thermophysical properties. The thermophysical properties of these substrates considered in present study is obtained from.⁴⁶ Under these varied input parameters, a total 600 numerical simulations are performed to obtain

TABLE I. Summary of model parameters considered in present study.

ML model	Parameter description	Value/type
Artificial neural network (ANN)	Activation Function	ReLU
	Optimizer	Adam
	Batch Size	42
	L2 Regularization parameter	0.0001
	Tolerance	0.0001
	Learning rate	0.002
	Exponential decay rate for estimates of first moment vector	0.9
	Exponential decay rate for estimates of first moment vector	0.999
	No. of Hidden layer for output1 (no. of nodes in each layer)	4 [200, 100, 50, 10] ^a
	No. of Hidden layer for output2 (no. of nodes in each layer)	10 [800, 600, 400, 200, 150, 100, 75, 50, 25, 10] ^a
Extreme gradient boosting (XGB)	Maximum training cycle	10 000 epochs
	Early stopping	2000 epochs
	No. of estimators	600
	Maximum depth of trees	8
	Minimum child weight	0.8
	Learning rate	0.1
	Minimum loss reduction for splitting	0.1
	Sub-sample ratio	1
	L1 regularization term on weights	0.5
	L2 regularization term on weights	0.5
	Refresh leaf	1
	Growth policy	Depth wise
	Scale Pos weight	1
	No. of estimators	600
Random forest (RF)	Maximum depth of trees	8
	bootstrap	True (Bagging)
	Random state	35
	Minimum samples split	2
	Minimum samples leaf	1
	Maximum features	Auto

^aConsidered based on manual search technique.

the database for ML algorithms. Figure 4 shows the distribution of the data points for these various input parameters considered in the numerical simulations. The input features are varied over a range to construct considerably robust ML models with four different target variables (output): evaporation rate (dm/dt), peak internal velocity (U_s), and highest temperature drop in the liquid (ΔT_{liq}) and solid (ΔT_{sol}) domain at stable conditions. The frequency distributions of these target variables are illustrated in Fig. 5. It is noticed that due to wide range of input parameters, the target variables also vary over large ranges. In addition to the geometric, ambient, and solid thermophysical features, we obtain additional features using feature engineering to enhance the performance of ML models. The details regarding feature engineering and the influence of input features on individual target variables are depicted in Sec. III.

B. Machine learning approaches

We apply the consolidated database obtained from the validated numerical study to develop supervised ML models. In this approach, an algorithm is trained using labeled input–output pairs sampled from a dataset of unknown distribution. The primary objective of this procedure is to minimize the loss or error function between the predicted and target variables. Next, the algorithm is assessed on cross-validation data to tune the hyperparameters, preventing overfitting and ensuring the algorithm generalizes well to the unseen data. Finally, the model is deployed to make predictions based on fresh and unseen test data. Our

input features and target variables for droplet evaporation problem are nonlinearly interdependent on one another. To model such complicated phenomenon, we employ three different ML algorithms: ANN, XGB, and RF.

In general, ANN model consists of interconnected layers of neurons at their core; each performing simple calculations based on activation function to process input data and pass it through to subsequent layers. Based on suitable choice of activation function and hidden layer properties, any physical phenomenon can effectively be modeled using deep ANN algorithms. Here, the term “deep” refers to the presence of multiple hidden layers in ANN architecture with several neurons in each layer. The other ML models considered here, that is, XGB and RF, both belong to decision tree classes. In such models, the regression tasks are performed at various nodes in tree structures on datasets sampled randomly from the input data. For RF models, outputs from each tree structure are simply ensembled to obtain the result. Whereas, in the case of XGB, the models are sequentially improved by focusing on the residuals obtained in previous step. This approach helps in capturing the nonlinear complexities in dataset in a better manner. A more detailed discussion regarding these ML models can be found in the literature.^{47–49}

1. Model architecture and parametric optimization

In the evaporation problem, we observed that one of the output variables (i.e., the evaporation rate) significantly influence the others

TABLE II. Model predictions for varied combination of input features.

Input features	ML model	Target variables	MSE	R ²	EV
$r, h, \theta, \phi, t, k, \rho, c_p$	ANN	dm/dt	1.82×10^{-8}	0.9881	0.9943
		U_s	1.3396	0.9781	0.9921
		ΔT_{liq}	0.00714	0.9493	0.9749
		ΔT_{sol}	0.0654	0.9469	0.9784
		dm/dt	3.84×10^{-8}	0.982	0.9826
	XGB	U_s	0.544	0.988	0.987
		ΔT_{liq}	0.0037	0.963	0.9637
		ΔT_{sol}	0.0196	0.983	0.9834
		dm/dt	4.97×10^{-8}	0.977	0.977
		U_s	0.531	0.97	0.971
$r, h, \theta, \phi, t, k, \rho, c_p, \alpha, h/t$	ANN	dm/dt	1.42×10^{-8}	0.9929	0.993
		U_s	0.06455	0.9894	0.99
		ΔT_{liq}	0.003563	0.978	0.977
		ΔT_{sol}	0.002584	0.979	0.9784
		dm/dt	2.3×10^{-8}	0.989	0.9891
	XGB	U_s	0.054	0.991	0.9916
		ΔT_{liq}	0.0051	0.965	0.9652
		ΔT_{sol}	0.0118	0.99	0.992
		dm/dt	4×10^{-8}	0.982	0.9818
		U_s	0.172	0.972	0.974

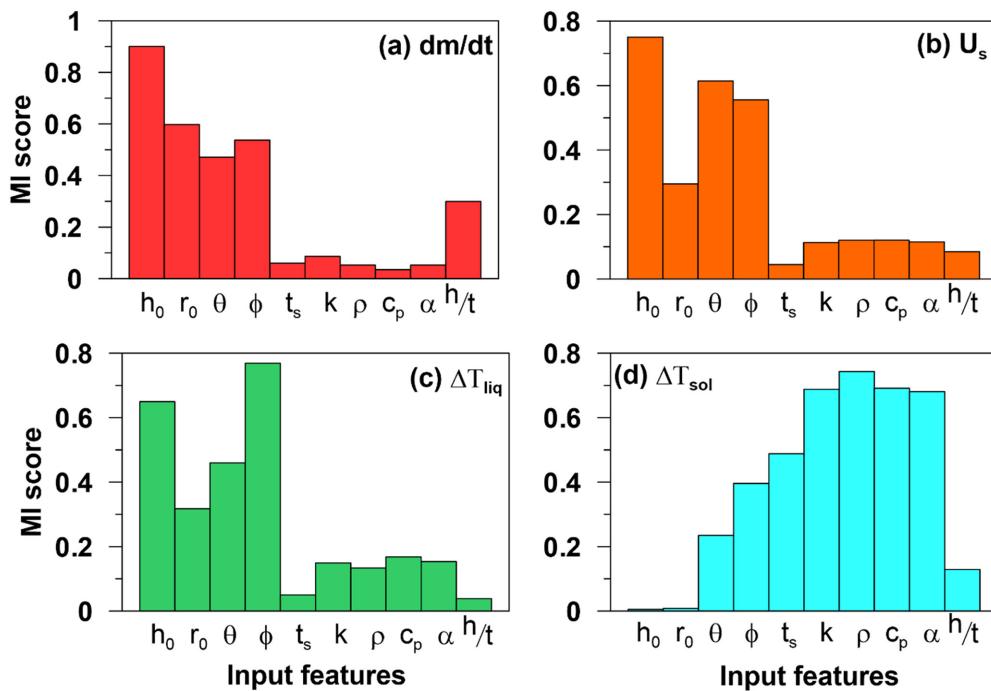


FIG. 8. Mutual info. score of various input features for the target variables: (a) evaporation rate (dm/dt), (b) Velocity scale (U_s), temperate drop in (c) liquid (ΔT_{liq}), and (d) solid (ΔT_{sol}).

(such as evaporative cooling effects, velocity scale). To consider this effect, we propose a cascaded structure of neural network (see Fig. 6). It consists of two deep ANN models with the output of first one being input along with other features to the second. More specifically, first the input features are fed to all the neurons in the hidden layer 1 to obtain the output 1. In the next step, we feed this output 1 along with all the input features to the hidden layer 2 to attain the predictions from output layer 2. In present study, we obtain the evaporation rate from output layer 1 and employ the same along with other features to obtain the velocity scale and cooling effects from output layer 2. This approach ensures the ANN architecture to remain physically consistent with the problem description. In the case of decision tree analysis, we may employ a multivariate tree ensembled structure. However, such structure fits one regressor per target and ultimately wraps them.

Hence, to ensure structural simplicity, we employ separate regressor per target output for both RF and XGB algorithms. Figure 7 depicts the simplified schematic structure for decision tree analysis. Also, in these decision tree analyses, the cascading is done manually, that is, the output from the tree structure of evaporation rate serves as input along with other features to the decision trees of velocity scale and evaporative cooling effects.

Next, we discuss about the parametric features employed in present study to optimize the training and validation process. For this purpose, the consolidated database obtained from numerical study is randomly split into train, cross-validation, and test data at a ratio of 70%, 15%, and 15%, respectively. The ML algorithms are trained on the training dataset adjusting the weight and bias parameters while the cross-validation data are used to tune the model's hyperparameters.

TABLE III. ANN model predictions for output 1 (dm/dt) under varied combinations of hidden layer.

Test no.	No. of hidden layer (no. of neurons in each layer)	MAE (%)	MSE	R^2
1	2 [10, 5]	15.65	4.82×10^{-8}	0.977
2	3 [20, 10, 5]	13.15	4.19×10^{-8}	0.9806
3	4 [50, 20, 10, 5]	12.65	3.1×10^{-8}	0.986
4	4 [100, 50, 25, 10]	9.43	1.5×10^{-8}	0.9912
5	4 [200, 100, 50, 10] ^a	6.29	1.42×10^{-8}	0.9934
6	5 [200, 100, 50, 20, 5]	7.89	2.60×10^{-8}	0.988
7	6 [200, 150, 100, 50, 20, 5]	7.49	2.56×10^{-8}	0.987
8	8 [400, 200, 150, 100, 50, 25, 10, 5]	7.67	2.4×10^{-8}	0.987

^aOptimized configuration.

These hyperparameters greatly depend on nature of data (such as distribution and feature dimensionality), algorithm specific factors (algorithm type and optimizer) and complexities involved. We use ReLu activation function with Adam optimizer⁵⁰ to update the weights and biases in ANN model since the same combines the advantages of multiple extensions of the classic stochastic gradient descent procedure. In present study, the hyperparameters are selected based on manual search technique similar to Ref. 51 and the final optimum values are summarized in Table I. The effects of some of the important optimization parameters are discussed in a detailed manner in Sec. III. Further, to compare the model performances during hyperparameter tuning and testing procedure, we consider several error matrices as discussed in Sec. II B 2.

2. Algorithm implementation and performance evaluation

Here, the proposed ANN and decision tree models are designed using Python, NumPy, pandas, TensorFlow, and scikit-learn⁵² libraries. The numerical data are preprocessed to scale the input features for better convergence of the models. For this purpose, the MinMaxScaler () module from scikit-learn library is employed. During training session, the Mean Squared Error (MSE) is considered as the loss function for all ML algorithms adopted here.⁵³ Now, because of limited considered cases (600 data points) compared to traditional ML models, the present computational cost is significantly low. For instance, the entire training, validation, and testing procedure can be conducted within minutes, even with moderately powerful desktop CPUs. To evaluate performance and compare the accuracies of different ML models, we consider several error metrics such as R^2 -square (R^2), mean absolute error (MAE), MSE, explained variance regression score (EV) and Kendall correlation coefficient (τ). Here, MAE and MSE represent the mean value of absolute and squared deviations between the target (Y) and predicted data (\hat{Y}). Thus, a lower value of these two parameters indicate lower deviation of the overall ML predictions from simulation

data for various target variables (such as evaporation rate, velocity scale, and temperature drop). R^2 represents the goodness of fit in terms of proportion of the variance in the dependent variable that can be predicted from the independent variable. This determination coefficient is similar to the explained variance score: $EV = 1 - [\text{var}(Y - \hat{Y})/\text{var}(Y)]$ when the ensemble deviation becomes zero. Thus, as R^2 and EV increase, the ML predictions accurately fit the simulation results for the given range depicting the complex, nonlinear physical phenomena of sessile droplet evaporation in a better manner. Also, the Kendall correlation coefficient (τ)⁵⁴ measures the similarity of orders between the target and predicted values. Thus, a higher τ value (close to +1) suggests that the model is performing well in terms of maintaining the relative ordering of target and predicted variables. Also, we use an additional parameter f_{30} to monitor the model performance which indicates the percentage of data predicted within $\pm 30\%$ range of the target value.

III. RESULTS AND DISCUSSION

A. Feature engineering and parametric influence

We noticed that the explicit input features primarily consist of geometrical parameters of droplet-substrate system, ambient conditions, and thermophysical properties of solid–liquid. These features are primarily manifested based on physical description of the problem. In feature engineering, these available input variables are modified and organized to create new input features that improves the model performance. We do this by forming dimensionless numbers from the raw input features, such as thermal diffusivity of solid ($\alpha = \frac{k}{\rho c_p}$) and the ratio of liquid to solid column (h/t) during evaporation. The nondimensionalization of the input features is primarily done based on the physics involved in the problem. For instance, both α and h/t significantly influence the individual and conjugate heat transport in the solid and liquid domain; and thereby the evaporative mass transfer and associated thermofluidic phenomena in drying droplets.^{15,16} To assess the influence of these additional features on the performance of the various ML algorithms we look into the error matrices. Table II

TABLE IV. ANN model predictions for output 2 under varied combinations of hidden layer.

Test no.	No. of hidden layer (no. of nodes in each layer)	MAE (%)			MSE			R^2		
		U_s	ΔT_{liq}	ΔT_{sol}	U_s	ΔT_{liq}	ΔT_{sol}	U_s	ΔT_{liq}	ΔT_{sol}
1	4 [50, 20, 10, 5]	9.67	9.37	13.37	1.33	0.007	0.032	0.968	0.951	0.973
2	6 [100, 75, 50, 25, 10, 5]	8.79	9.20	13.28	1.25	0.007	0.029	0.979	0.957	0.972
3	8 [200, 150, 100, 75, 50, 25, 10, 5]	8.71	11.27	17.99	0.98	0.007	0.035	0.978	0.954	0.962
4	8 [400, 200, 150, 100, 50, 25, 10, 5]	8.4	11.8	14.86	1.28	0.007	0.026	0.979	0.966	0.973
5	10 [400, 250, 200, 150, 100, 75, 50, 25, 10, 5]	6.87	8.48	16.15	0.88	0.005	0.028	0.984	0.970	0.975
6	10 [600, 400, 200, 150, 100, 75, 50, 25, 10, 5]	7.25	8.89	16.76	1.15	0.005	0.031	0.983	0.969	0.972
7	10 [800, 600, 400, 200, 150, 100, 75, 50, 25, 10] ^a	6.45	7.19	14.64	1.01	0.004	0.027	0.985	0.976	0.975
8	12 [800, 600, 500, 400, 300, 200, 150, 100, 75, 50, 25, 10]	9.00	9.35	15.89	1.37	0.005	0.042	0.977	0.965	0.965
9	14 [1000, 800, 700, 600, 500, 400, 300, 200, 150, 100, 75, 50, 25, 10]	8.49	9.19	20.8	1.77	0.005	0.053	0.971	0.962	0.957
10	16 [1200, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 150, 100, 75, 50, 25, 10]	10.9	11.8	22.6	1.86	0.006	0.061	0.969	0.952	0.954

^aFinal selections for ANN model configuration.

depicts the performances of the ML models applied on the test data with and without considering the additional input features. It is noticed that the inclusion of these features significantly improves the forecasting capabilities for all ML models, especially so for thermofluidic predictions. For instance, it is noticed that R^2 and MSE of the prediction of ΔT_{sol} for ANN model significantly improves from 0.9469 and 0.065 to 0.979 and 0.0025, respectively. It is worth mentioning that the error matrices are calculated under varied architecture of the ML models (as discussed later) and the results shown here are obtained from the optimum one.

The results shown in Table II also imply that the input features (both raw and derived dimensionless number) may differently affect individual target variables under varied conditions. To probe the preferential influence of these input features in a more detailed manner, we look into the mutual information (MI) score of all the input features for each of the output variables as shown in Fig. 8. MI score assess the dependency between variables by qualifying the amount of information obtained about one random variable (target) through another variable (feature).⁵⁵ A higher relative

MI score indicates strong influence of input feature on target variable; whereas a smaller value suggests weak correlation between them. We observe that, the MI scores of geometrical parameters and ambient conditions are significantly higher for evaporation rate (dm/dt) indicating a substantial influence; whereas the thermophysical properties of solid substrate has a comparatively minor role. For the velocity scale and temperature drop in liquid domain, the MI scores of the solid substrate properties increase to a substantial extent; thereby suggest influential role in addition to geometrical features and ambient conditions. In the case of solid domain temperature drop, the MI scores are lower for the geometrical features and the same is primarily enhanced for thermophysical properties of solid substrate [see Fig. 8(d)]. These findings suggest a strong correlation between the thermophysical properties and temperature drop in solid substrate. Thus, we observe that, the relative MI scores of the input features vary over a wide range for different target variables. To consider these effects, the architecture of the ML models shall be carefully designed such that it retains robustness while being computationally efficient.

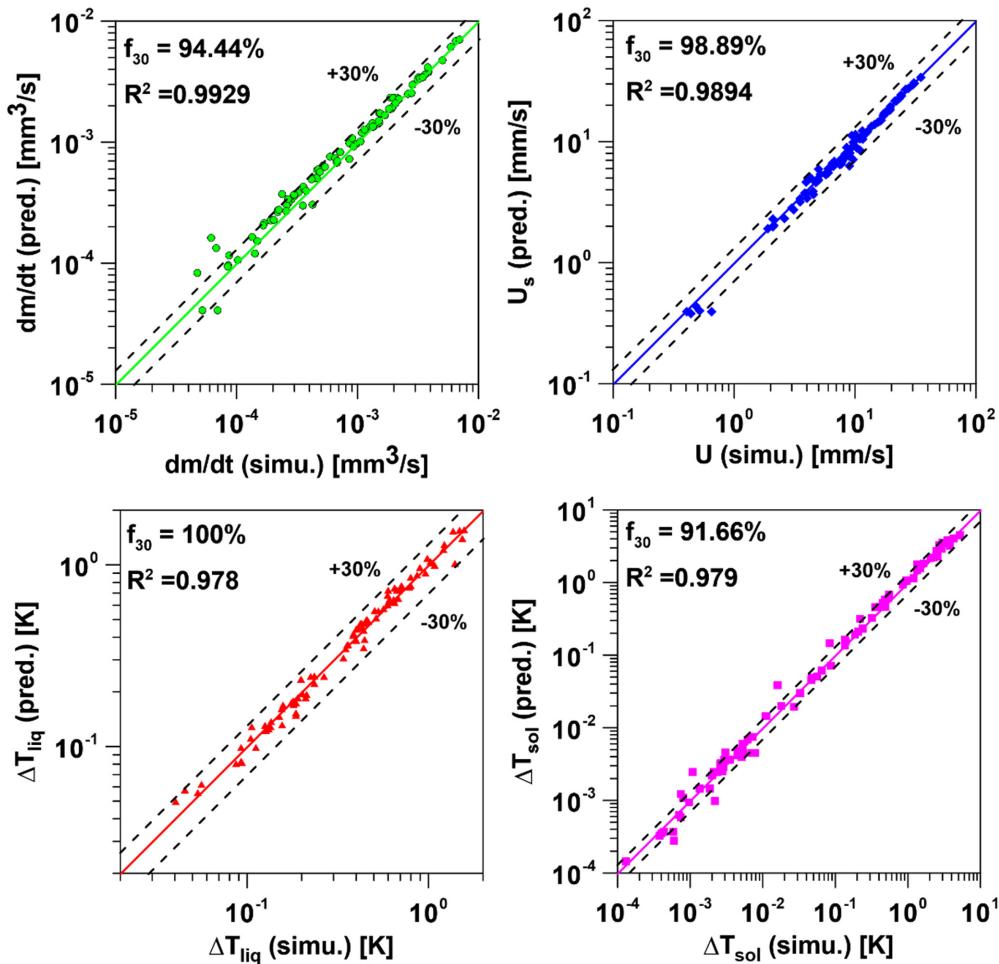


FIG. 9. Comparison of simulation data (x-axes) and predicted results (y-axes) of ANN model in the case of test dataset for target variables: (a) evaporation rate (dm/dt), (b) velocity scale (U_s), temperature drop in (c) liquid, and (d) solid domain.

B. Optimum model architecture

To obtain a trade-off between the model robustness and the computation cost, we vary the architecture of various ML models and select the optimum one. For instance, in the case of ANN model, the computation cost significantly increases with the number of layers and the number of neurons or nodes in each layer.⁴⁷ However, enhanced layer and increased nodes helps in resolving the nonlinearities involved in the problem thereby improving the predicting performance of model. This trade-off relation also exists in the case of decision tree models (such as number of predictors and depth of trees) for efficient use of computational resources. Here, we vary these model architectures over a wide range. Tables III and IV show the performance of ANN model on cross-validation data for different combination of model architectures. It is noticed that developing models with simplified structures levy higher toll on R^2 , MAE, and MSE. Now, as the number of nodes in layers increase, we obtain satisfactorily higher R^2 values (>0.96) despite considerably large MAE and MSE. When both the number of layers and nodes increase to an optimal level, the R^2 value improves

significantly while a considerably lower MAE and MSE is obtained. Beyond a critical point, bringing extra nodes or addition of redundant layers result in degradation of the R^2 value and augments the residual error due to overfitting. This behavior of the error matrices exists for all the target variable considered in present study. A similar optimum condition also exists in the case of decision tree models. In such cases, the number of predictors and the maximum depth of the decision tree are optimized (not shown here for brevity of work) to improve the performance of RF and XGB models. These optimum architectures of various ML models are primarily obtained using manual search technique⁵² and the results for hidden layer optimization are pointed in Tables III and IV. Finally, these appropriate model structures are applied to predict the target variables in the test dataset as described in Sec. III C.

C. Comparison of model performances

In this subsection, we test the predictions from various ML models considered in present study with the physics-based full-scale

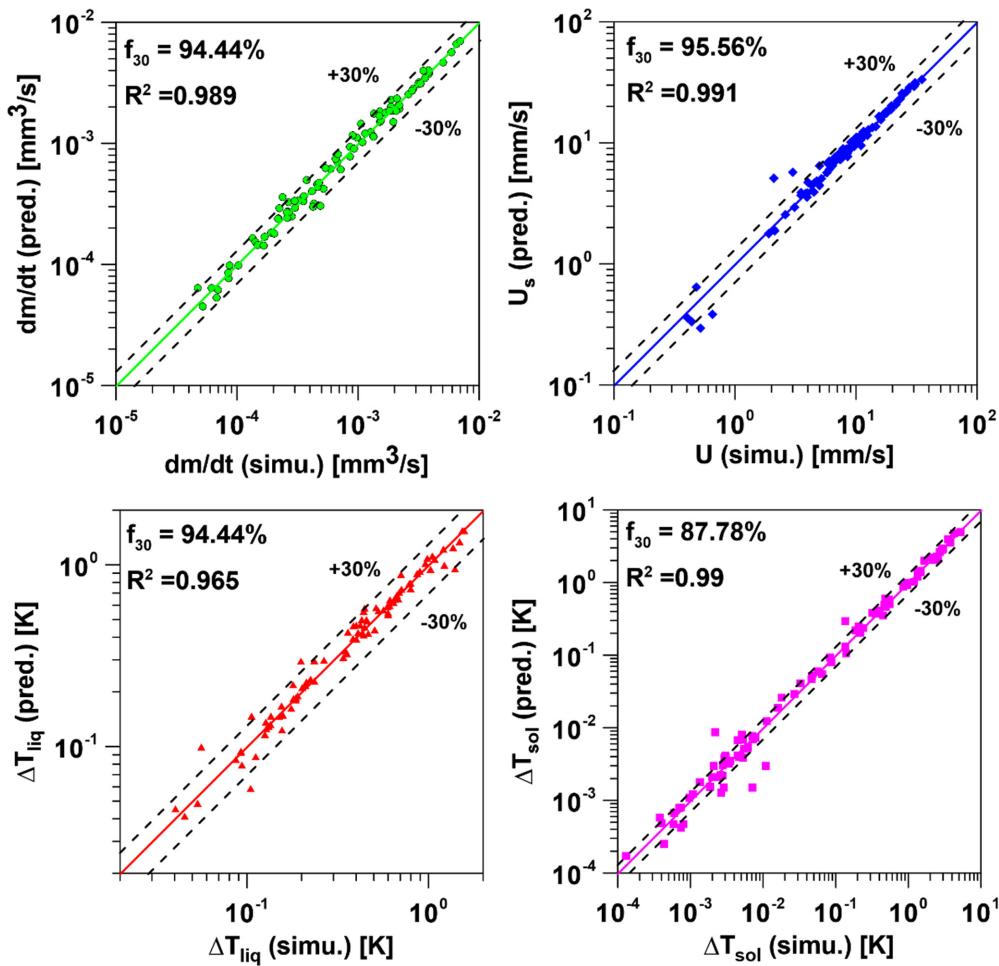


FIG. 10. Comparison of simulation data (x-axes) and predicted results (y-axes) of XGB model in the case of test dataset for target variables: (a) evaporation rate (dm/dt), (b) velocity scale (U_s), temperature drop in (c) liquid, and (d) solid domain.

numerical simulations. Figure 9 shows this comparison for various target variables in the case of ANN model; where on the y axis we show prediction results and the x axis denotes corresponding cases of simulation data. Hence, the ideal case when the ML predictions are exactly same as numerical data are represented by a straight line of gradient +1 as shown in Fig. 9. It is noticed that in the case of ANN model, ML predictions closely follow this ideal straight line for the entire range of various target variables. The goodness of fit parameter, R^2 yields a lowest value of 0.978 for the predictions of temperature drop in liquid domain and the same attains its highest value ($R^2 = 0.9929$) in the case of mass transfer rate. These satisfactorily higher R^2 values suggest the robustness of the ML model in predicting the variability of all the target variables. The MSE values (as seen in the Tables III and IV) in these cases are sufficiently low suggesting the absence of large-scale deviations between the predicted and simulation data. We also notice that the percentage of data predicted within $\pm 30\%$ error is satisfactorily high with the lowest value being noted as 91.66% for temperature drop in solid. These results indicate the superiority of deep ANN

model in predicting the intricacy of sessile droplet evaporation phenomenon with high accuracy.

In the case of decision tree models, these performance comparison plots are illustrated in Figs. 10 and 11 for XGB and RF algorithms, respectively. It is noticed that, the XGB model yields R^2 values similar to ANN model in the case of evaporation rate, velocity scale and temperature drop in solid. However, in the case of temperature drop in liquid domain the predictions are worse. Also, the percentage of data lying outside the $\pm 30\%$ error bar is significantly increased for all the target variables. These results indicate the overall degraded performance of XGB as compared to the ANN model. The model performance matrices (both R^2 and f_{30}) are degraded further in the case of RF algorithm especially for the temperature drop in solid and liquid domain. For instance, the parameter f_{30} for RF model is reduced to 87.78% in the case of temperature drop in solid domain. This worse performance can be attributed to the large variability of the solid domain temperature drop as compared to other target variables (as discussed in Sec. IIID). Moreover, the absence of boosting framework

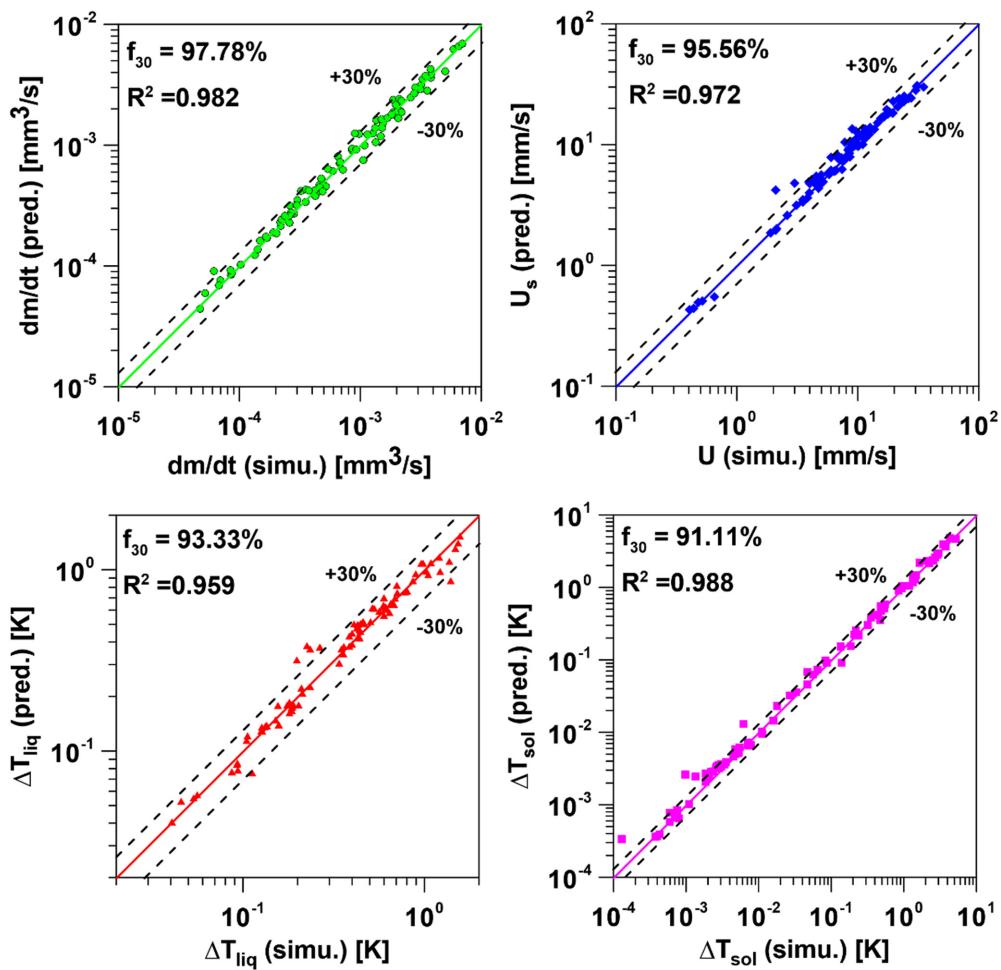


FIG. 11. Comparison of simulation data (x-axes) and predicted results (y-axes) of RF model in the case of test dataset for target variables: (a) evaporation rate (dm/dt), (b) velocity scale (U_s), temperature drop in (c) liquid, and (d) solid domain.

degrades the performance matrices further. In boosting algorithms, each new model aims to improve prediction by focusing on the errors of the previous models.⁴⁸ This iterative process creates a comparatively robust final model that can capture inherent complexities in the evaporation phenomena especially the associated cooling effects. Moreover, XGB considers Newton's method that applies higher order alternative to optimization methods.⁴⁸ RFs, on the other hand, simply build multiple decision trees independently and then aggregate their results for prediction. Hence, the overall XGB predictions are comparatively better than that of RF model. However, both these decision tree algorithms are unable to match the performance of deep ANN model because of its inherent capability of capturing complex relationships between the target variables and input features manifested in the hidden layers. Also, the ability of autonomous feature learning and scalability of large range data in ANN models helps in resolving the complex interrelation in a better manner especially for the temperature drop both in solid and liquid domain.

D. Error analysis

In Sec. III C, we observed that, the ANN model outperforms the decision tree models in terms of R^2 and f_{30} values. Now, although these parameters manifest the proportion of explained variance, that is, how well model predictions fit the actual data; the same does not provide any information regarding the actual size of error. To obtain a clear and direct measure of average prediction error, we look into the MAE for various ML models. It is worthwhile to mention that we consider the MAE as the primary evolution parameter for error analysis over MSE, since the former is less sensitive to outliers. Also, to assess the ordinal relationship between the actual and predicted values we look into the Kendall correlation coefficient (τ). Figure 12 show these MAE

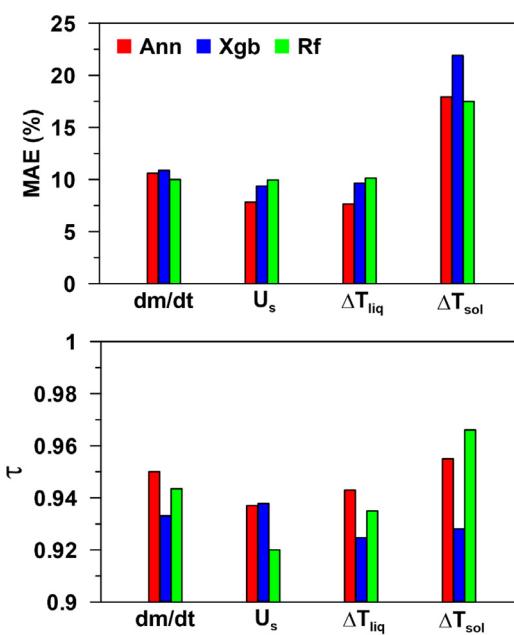


FIG. 12. Comparison of MAE and Kendall correlation coefficient (τ) for various ML models.

and τ values for various target variables and different ML models considered in present study. It is observed that the MAE for the evaporation rate, velocity scale and temperature drop in liquid domain is less than 12.5% of for the all ML models. For the temperature drop in solid, the MAE values are considerably higher with a peak value of 22% in the case of the XGB model.

The poor performance of all the ML models for the temperature drop in solid domain can be attributed to the large variability range of the same as compared to other target variables [see Figs. 9(d), 10(d), and 11(d)]. We observe that, in solid domain, the temperature drop varies by five orders of magnitude under different input cases considered in present study. These large range may exhibit large variability of errors, that is, the errors may be relatively small in some cases and quite large for others. At the larger values, the absolute errors may be significantly higher as compared to small scale values. Also, in the case of lower scale of target variables, the data points may be scattered over a wider range for different combination of input features; causing it harder for the ML models to identify the actual relationship between input and output variables. Therefore, in a long-range scenario, the absolute errors tend to be larger, driving up the MAE. Further, it can be seen that the τ value for all considered the cases is greater than 0.92. These higher τ values evince strong ordinal relationship between the target and predicted variables. Now, among various ML models, we observe that these τ values are considerably higher in the case of ANN model as compared to others for most of cases. A similar behavior can also be observed in the case of MAEs, that is, the ANN model yields lower MAE for most of the considered cases. Also, in term of other matrices (such as R^2 and f_{30}), the performance of ANN models is comparatively better. Hence, we infer that for this problem, the ANN algorithm surpasses other ML models considered.

IV. CONCLUSIONS

In this study, we employed deep ANN and other decision tree models (RF and XGB) to predict the evaporation behavior of sessile droplets. To maintain physical consistence and make accurate predictions, we adopted a cascaded structure for these ML models, that is, one of the outputs is considered as an input for other target variables. A consolidated database of 600 points for droplet evaporation problem are obtained from a well-validated numerical model that includes wide range of geometrical features (such as contact radius, height, contact angle), ambient conditions and thermophysical properties of solid substrate. Based on these data, the ML algorithms are trained and tested to predict droplet evaporation rate, peak velocity and temperature drop in solid and liquid domain. The key insights from present work can be summarized as:

- The inclusion of additional input features (such as thermal diffusivity) based on feature engineering significantly improve the performance of ML models epically for the temperature drop predictions. This primarily occurs due to higher mutual information score obtained in case of thermal features for these thermophysical properties.
- Among various ML models, the ANN algorithm outperforms others in term of R^2 and f_{30} for all the target variables due to its ability to learn complex relationships between the input and output variables with deep networks. In case of XGB and RF, the former provides comparatively better prediction for evaporation rate and internal velocity scale due to boosting framework and

sequential building of models. However, in case of thermal state prediction, the accuracy of both these models curtail due to inherent problem complexity.

- Among different target variables, the percentage of MAE for all ML models are substantially higher in case of temperature drop in the solid domain. This worse behavior is attributed to the large variability of the same target variable as compared to others. Also, due to the sparse data especially in lower range the models might not be able to generalize the underlying behavior well.

We observed that, despite deep network structures and boosting algorithms in the decision tree models, the error magnitude between predicted and target variables is noticeably high for some cases. These inaccuracies can be alleviated by considering more advanced ML models or by increasing the training data size. For instance, in the case of problems (such as evaporation process) with high dimensional and complex nonlinear PDEs, Physics-Informed Neural Network (PINN) or Fourier Neural Operators (FNO) may be employed. These approaches help in learning the solution to a complex physical problem while respecting its underlying physics with high-dimensional input-output spaces. Moreover, the volume of training set data may be increased to for better prediction of these advanced ML models. These approaches may be considered as a potential outlook for generalization of the present study.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Arnov Paul: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal).

Purbarun Dhar: Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – original draft (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1. R. Tardif and R. M. Rasmussen, "Evaporation of nonequilibrium raindrops as a fog formation mechanism," *J. Atmos. Sci.* **67**(2), 345–364 (2010).
2. I. O. Ucar and H. Y. Erbil, "Use of diffusion controlled drop evaporation equations for dropwise condensation during dew formation and effect of neighboring droplets," *Colloids Surf. A* **411**, 60–68 (2012).
3. J. Kim, "Spray cooling heat transfer: The state of the art," *Int. J. Heat Fluid Flow* **28**(4), 753–767 (2007).
4. H. Minemawari, T. Yamada, H. Matsui, J. Y. Tsutsumi, S. Haas, R. Chiba *et al.*, "Inkjet printing of single-crystal films," *Nature* **475**(7356), 364–367 (2011).
5. A. R. Martin and W. H. Finlay, "Nebulizers for drug delivery to the lungs," *Expert Opin. Drug Deliv.* **12**(6), 889–900 (2015).
6. T. Thorsen, R. W. Roberts, F. H. Arnold, and S. R. Quake, "Dynamic pattern formation in a vesicle-generating microfluidic device," *Phys. Rev. Lett.* **86**(18), 4163 (2001).
7. S. K. Wilson and H. M. D'Ambrosio, "Evaporation of sessile droplets," *Annu. Rev. Fluid Mech.* **55**(1), 481–509 (2023).
8. H. Gelderblom, C. Diddens, and A. Marin, "Evaporation-driven liquid flow in sessile droplets," *Soft Matter* **18**(45), 8535–8553 (2022).
9. P. Dhar, R. K. Dwivedi, and A. R. Harikrishnan, "Surface declination governed asymmetric sessile droplet evaporation," *Phys. Fluids* **32**(11), 112010 (2020).
10. A. Paul and P. Dhar, "Evaporation kinetics of sessile droplets morphed by substrate curvature," *Phys. Fluids* **33**(12), 122010 (2021).
11. A. Paul, R. K. Dash, and P. Dhar, "Phenomenology and kinetics of sessile droplet evaporation on convex contours," *Int. J. Therm. Sci.* **187**, 108194 (2023).
12. H. Hu and R. G. Larson, "Evaporation of a sessile droplet on a substrate," *J. Phys. Chem. B* **106**(6), 1334–1344 (2002).
13. H. Gelderblom, A. G. Marin, H. Nair, A. Van Houselt, L. Lefferts, J. H. Snoeijer, and D. Lohse, "How water droplets evaporate on a superhydrophobic substrate," *Phys. Rev. E* **83**(2), 026306 (2011).
14. D. Brutin, "Influence of relative humidity and nano-particle concentration on pattern formation and evaporation rate of pinned drying drops of nanofluids," *Colloids Surf. A* **429**, 112–120 (2013).
15. G. J. Dunn, S. K. Wilson, B. R. Duffy, S. David, and K. Sefiane, "The strong influence of substrate conductivity on droplet evaporation," *J. Fluid Mech.* **623**, 329–351 (2009).
16. Y. H. Chen, W. N. Hu, J. Wang, F. J. Hong, and P. Cheng, "Transient effects and mass convection in sessile droplet evaporation: The role of liquid and substrate thermophysical properties," *Int. J. Heat Mass Transfer* **108**, 2072–2087 (2017).
17. E. Bormashenko, A. Musin, and M. Zinigrad, "Evaporation of droplets on strongly and weakly pinning surfaces and dynamics of the triple line," *Colloids Surf. A* **385**(1–3), 235–240 (2011).
18. T. A. Nguyen, S. R. Biggs, and A. V. Nguyen, "Analytical model for diffusive evaporation of sessile droplets coupled with interfacial cooling effect," *Langmuir* **34**(23), 6955–6962 (2018).
19. A. Paul and P. Dhar, "Transients of Marangoni and Stefan advection dynamics during generic sessile droplet evaporation," *Phys. Fluids* **35**(10), 102022 (2023).
20. F. Carle, B. Sobac, and D. Brutin, "Experimental evidence of the atmospheric convective transport contribution to sessile droplet evaporation," *Appl. Phys. Lett.* **102**(6), 061603 (2013).
21. F. Carle, S. Semenov, M. Medale, and D. Brutin, "Contribution of convective transport to evaporation of sessile droplets: Empirical model," *Int. J. Therm. Sci.* **101**, 35–47 (2016).
22. S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine learning for fluid mechanics," *Annu. Rev. Fluid Mech.* **52**(1), 477–508 (2020).
23. K. Fukami, K. Fukagata, and K. Taira, "Super-resolution reconstruction of turbulent flows with machine learning," *J. Fluid Mech.* **870**, 106–120 (2019).
24. P. P. Popov, D. A. Buchta, M. J. Anderson, L. Massa, J. Capecelatro, D. J. Bodony, and J. B. Freund, "Machine learning-assisted early ignition prediction in a complex flow," *Combust. Flame* **206**, 451–466 (2019).
25. M. T. Hughes, G. Kini, and S. Garimella, "Status, challenges, and potential for machine learning in understanding and applying heat transfer phenomena," *J. Heat Transfer* **143**(12), 120802 (2021).
26. J. Thibault and B. P. Grandjean, "A neural network methodology for heat transfer data analysis," *Int. J. Heat Mass Transfer* **34**(8), 2063–2070 (1991).
27. K. Jambunathan, S. L. Hartle, S. Ashforth-Frost, and V. N. Fontama, "Evaluating convective heat transfer coefficients using neural networks," *Int. J. Heat Mass Transfer* **39**(11), 2329–2332 (1996).
28. E. Cho, H. Lee, M. Kang, D. Jung, G. Lee, S. Lee *et al.*, "A neural network model for free-falling condensation heat transfer in the presence of non-condensable gases," *Int. J. Therm. Sci.* **171**, 107202 (2022).
29. R. B. Jadrich, B. A. Lindquist, and T. M. Truskett, "Unsupervised machine learning for detection of phase transitions in off-lattice systems. I. Foundations," *J. Chem. Phys.* **149**(19), 194109 (2018).
30. A. Khosravi, J. J. G. Pabon, R. N. N. Koury, and L. Machado, "Using machine learning algorithms to predict the pressure drop during evaporation of R407C," *Appl. Therm. Eng.* **133**, 361–370 (2018).
31. Y. Qiu, D. Garg, S. M. Kim, I. Mudawar, and C. R. Kharangate, "Machine learning algorithms to predict flow boiling pressure drop in mini/micro-channels based on universal consolidated data," *Int. J. Heat Mass Transfer* **178**, 121607 (2021).
32. F. Nie, S. Yan, H. Wang, C. Zhao, Y. Zhao, and M. Gong, "A universal correlation for predicting two-phase frictional pressure drop in horizontal tubes based on machine learning," *Int. J. Multiphase Flow* **160**, 104377 (2023).

- ³³A. Paul, A. Roy, and P. Dhar, "External Stefan and internal Marangoni thermo-fluid dynamics for evaporating capillary bridges," *Langmuir* **40**(10), 5255–5269 (2024).
- ³⁴E. Olsson, G. Kreiss, and S. Zahedi, "A conservative level set method for two phase flow II," *J. Comput. Phys.* **225**(1), 785–807 (2007).
- ³⁵J. Donea, A. Huerta, J. P. Ponthot, and A. Rodríguez Ferran, "Arbitrary Lagrangian-Eulerian methods," in *Encyclopedia of Computational Mechanics* (Wiley, 2004).
- ³⁶M. Souli and J. P. Zolesio, "Arbitrary Lagrangian-Eulerian and free surface methods in fluid mechanics," *Comput. Methods Appl. Mech. Eng.* **191**(3–5), 451–466 (2001).
- ³⁷B. Sobac and D. Brutin, "Thermal effects of the substrate on water droplet evaporation," *Phys. Rev. E* **86**(2), 021602 (2012).
- ³⁸F. Thomas, Jr., *Steam and Gas Tables with Computer Equations* (Elsevier, 2012).
- ³⁹R. C. Weast and M. J. Astle, *CRC Handbook of Chemistry and Physics*, 62nd ed. (CRC Press, Boca Raton, FL, 1981–1982).
- ⁴⁰Y. Sui, H. Ding, and P. D. Spelt, "Numerical simulations of flows with moving contact lines," *Annu. Rev. Fluid Mech.* **46**, 97–119 (2014).
- ⁴¹D. T. Tran, N. K. Nguyen, P. Singha, N. T. Nguyen, and C. H. Ooi, "Modelling sessile droplet profile using asymmetrical ellipses," *Processes* **9**(11), 2081 (2021).
- ⁴²R. G. Picknett and R. Bexon, "The evaporation of sessile or pendant drops in still air," *J. Colloid Interface Sci.* **61**(2), 336–350 (1977).
- ⁴³A. Paul, G. Khurana, and P. Dhar, "Substrate concavity influenced evaporation mechanisms of sessile droplets," *Phys. Fluids* **33**(8), 082003 (2021).
- ⁴⁴I. Kiper, R. Fulcrand, C. Pirat, G. Simon, B. Stutz, and S. M. Ramos, "Sessile drop evaporation on (super) hydrophobic surfaces: Effect of low pressure on the contact line dynamics," *Colloids Surf. A* **482**, 617–623 (2015).
- ⁴⁵A. Paul, S. Mondal, and P. Dhar, "Droplet evaporation on curved surfaces: Transients of internal and external transport phenomena," *Proc. R. Soc. A* **480**(2287), 20230613 (2024).
- ⁴⁶C. L. Yaws, *Yaws' Handbook of Thermodynamic and Physical Properties of Chemical Compounds* (Knovel, 2003).
- ⁴⁷A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer* **29**(3), 31–44 (1996).
- ⁴⁸D. Nielsen, "Tree boosting with XGBoost—why does XGBoost win ‘every’ machine learning competition?" M.S. thesis (NTNU, 2016).
- ⁴⁹T. K. Ho, "Random decision forests," in *Proceedings of 3rd International Conference on Document Analysis and Recognition* (IEEE, 1995), Vol. 1, pp. 278–282.
- ⁵⁰D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
- ⁵¹Y. Qiu, D. Garg, L. Zhou, C. R. Kharangate, S. M. Kim, and 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 Transfer* **149**, 119211 (2020).
- ⁵²F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
- ⁵³H. Li, Z. Xu, G. Taylor, C. Studer, and T. Goldstein, "Visualizing the loss landscape of neural nets," in *Proceedings of the Advances in Neural Information Processing Systems* (2018), pp. 6389–6399.
- ⁵⁴M. S. Schaeffer and E. E. Levitt, "Concerning Kendall's tau, a nonparametric correlation coefficient," *Psychol. Bull.* **53**(4), 338 (1956).
- ⁵⁵D. J. MacKay, *Information Theory, Inference and Learning Algorithms* (Cambridge University Press, 2003).