Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Machine-learning prediction of flow boiling heat transfer in Al 2 O 3 /water and TiO 2 /water nanofluids: base-fluid anchoring and grouped validation

**Adel Bouali**

boualiadelgm2020@gmail.com

University Yahia Fares of Medea

**Hamza Ramdani**

Center of Research in Mechanics

**Khaoula Nedjem**

University Amar Telidji

**Taqiy Eddine Boukelia**

Mohamed Seddik Benyahia University

**Aissa Laouissi**

University of Bordj Bou Arreridj

**Mourad Nouioua**

King Fahd University of Petroleum and Minerals (KFUPM)

**Research Article**

**Keywords:** Nanofluid, Flow boiling, Machine learning, Heat transfer coefficient, Cross-validation, Base-fluid anchoring

**Posted Date:** November 24th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11664409/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Machine-learning prediction of flow boiling heat transfer in $Al_2O_3$/water and $TiO_2$/water nanofluids: base-fluid anchoring and grouped validation

Adel Bouali[1*], Hamza Ramdani[2], Khaoula Nedjem[3], Taqiy Eddine Boukelia[4,5], Aissa Laouissi[6], Mourad Nouioua[7],

[1*]Laboratory of Biomaterials and Transport Phenomena (LBMPT), University of Medea, 26000 Medea, Algeria.

[2]Mechanics Research Centre, P.O. Box 73B, 25000, Constantine, Algeria.

[3]Laboratory of Mechanics, University Amar Telidji, 03000 Laghouat, Algeria.

[4]Mechanical Engineering Laboratory, Faculty of Science and Technology, Mohamed Seddik Benyahia University, Jijel, Algeria.

[5]Mechanical and Advanced Materials Laboratory, Polytechnic School of Constantine, Constantine, Algeria.

[6]Department of Mechanical Engineering, Faculty of Sciences and Technology, University of Bordj Bou Arreridj, 34000 Bordj Bou Arreridj, Algeria

[7]Interdisciplinary Research Center for Intelligent Manufacturing and Robotics, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, 31261, Saudi Arabia

*Corresponding author(s). E-mail(s): boualiadelgm2020@gmail.com;

**Abstract**

Machine-learning models of nanofluid flow boiling are usually trained and tested on neighbouring records of one test section. This study examines how validation design and target formulation change their apparent accuracy. Eight regression models predicted 186 heat transfer coefficients of $Al_2O_3$/water and $TiO_2$/water nanofluids in a horizontal annulus at 1 bar. They predicted the coefficient directly or as an enhancement ratio anchored to the measured base-fluid coefficient. Accuracy was assessed with the coefficient of determination ($R^2$), the root mean square error, the mean bias error (MBE), and the mean absolute percentage error under random and grouped cross-validation. Under random cross-validation, gradient boosting reached $R^2$ values of 0.995 and 0.992. When a mass-flux level was withheld, they fell to 0.545 and 0.105, and MBE values of +4.06 and +4.02 $kW\,m^{-2}\,K^{-1}$ revealed a systematic over-prediction. Smooth models kept $R^2$ values of 0.89 to 0.96. Under the same scheme, anchoring raised $R^2$ to at least 0.990 and lowered the best percentage errors to 1.20 and 0.56%. The

measured base fluid supplied about 95% of the target variance. With estimated base-fluid coefficients, this gain disappeared. Prediction intervals were not calibrated at extreme levels. The results apply to withheld levels of one test section.

**Keywords:** Nanofluid, Flow boiling, Machine learning, Heat transfer coefficient, Cross-validation, Base-fluid anchoring

# 1 Introduction

Nanofluids, suspensions of particles smaller than 100 nm in a base liquid, were proposed by Choi [12] as coolants with a higher thermal conductivity than their base fluids. For copper particles, that study estimated large reductions of the pumping power of heat exchangers. In the International Nanofluid Property Benchmark Exercise, more than 30 organisations measured identical samples [8]. Most results lay within ±10% of the sample mean, and effective-medium theory of the Maxwell type reproduced the measured conductivity, so no anomalous enhancement was found. Changes of the boiling heat transfer coefficient (HTC) that exceed these property effects must therefore have another origin.

Experimental studies of nanofluid flow boiling report changes of both signs. Kim et al. [29] measured subcooled flow boiling at atmospheric pressure with alumina, zinc oxide, and diamond nanofluids of up to 0.1 vol%. The nanofluid HTC stayed within ±20% of that of water and increased with mass flux and heat flux. Particle deposition altered the micro-cavities and the wettability of the heated surface, but no consistent correlation was found between the nucleation site density and the HTC. Wang et al. [58] found enhancement and deterioration with 0.1 vol% SiC/water (−24.7 to +30.6%) and graphite/water (−29.4 to +13.3%) nanofluids. They attributed the sign to a balance between property and surface effects that assist boiling and a reduction of active boiling sites that resists it. Mukherjee et al. [40] reported a higher HTC for $Al_2O_3$/water nanofluids of 0.01 to 1 wt% in subcooled flow boiling in a horizontal tube. Their performance evaluation criterion reached 94% above water, although the pressure drop increased with concentration. Recent experiments and simulations point to the heated surface. Soleimani and Ziabasharhagh [53] measured subcooled flow boiling of a 0.1 vol% $Al_2O_3$/water nanofluid on aluminium, brass, and copper heaters at three velocities and three roughness levels. Brass gave the highest and copper the lowest heat transfer. Wang et al. [57] coated heaters with a porous layer of silica nanoparticles or with zinc oxide nanowires, and both coatings raised the CHF of subcooled flow boiling at 1 bar by about 15%. Hou et al. [21] reported CHF and HTC gains of up to 32 and 26% for $Al_2O_3$ nanofluids in microchannel flow boiling, with lower performance at higher concentration. In a simulation that treated the deposited layer as a porous medium, Mohebali and Baniamerian [38] found that deposition reduced the HTC by 31% and the CHF by 16% on average. Eulerian simulations of subcooled boiling in a horizontal tube gave a maximum HTC gain of 10.56% for $Al_2O_3$/water [9].

The records modelled in the present study were measured by Dadhich et al. [14] with $Al_2O_3$/water and $TiO_2$/water nanofluids in a horizontal annulus at 1 bar. Both nanofluids exceeded the HTC of water, $Al_2O_3$ more than $TiO_2$, and the ratio to the water coefficient increased with concentration, heat flux, and mass flux. Correlations and an artificial neural network were fitted to these records. In the same facility at 1 to 2.5 bar, Dadhich et al. [15] reported a maximum gain of 76.7% for $TiO_2$/water at 0.20 vol%, 2.5 bar, and a Reynolds

number of 24,000. Reviews describe the contradictions between such studies. Barber et al. [3] found conflicting effects of nanoparticles on the nucleate boiling HTC but an almost unanimous enhancement of the critical heat flux (CHF), which several studies linked to particle deposition. Fang et al. [17] reviewed the HTC, CHF, pressure drop, stability, and flow patterns of nanofluid flow boiling and identified contradictions between published results. Kamel et al. [23] reviewed the effects of operating conditions and particle morphology on the HTC and CHF, and Dadhich et al. [16] reviewed more than 100 articles on flow boiling of water, nanofluids, and nano-refrigerants. For microchannels, Adio et al. [1] reported CHF gains of up to 49% together with unresolved problems of stability, clogging, fouling, and pumping power.

Classical flow boiling correlations combine a convective and a nucleate boiling contribution. Chen [11] multiplied a liquid convective coefficient by an enhancement factor and a nucleate boiling coefficient by a suppression factor. Gungor and Winterton [20] fitted a general correlation to more than 4300 points of seven fluids, including subcooled boiling and annuli, with mean deviations of 21.4% for saturated and 25.0% for subcooled boiling. Kandlikar [24] introduced a fluid-dependent parameter and correlated 5246 points of ten fluids, with a mean deviation of 15.9% for water. Liu and Winterton [33] correlated saturated and subcooled boiling in tubes and annuli with an explicit nucleate pool boiling term. Shah [49] predicted 4852 points of 30 fluids in channels of 0.38 to 27.1 mm with a mean absolute deviation of 18.6%. None of these correlations contains a term for the surface modification produced by nanoparticles. Their nanofluid versions rely on effective properties, such as the conductivity model of Maxwell [37] and the viscosity model of Brinkman [7].

Data-driven models avoid the assumed structure of correlations. Qiu et al. [43] trained an artificial neural network on 16,953 saturated flow boiling points of 16 fluids in mini- and micro-channels from 50 sources. It reached a mean absolute error of 14.3% on test points but predicted poorly for a fluid whose data were excluded from training. Bard et al. [4] applied imputation and feature selection to the same database and reached a mean absolute percentage error of 11.3%. For the CHF, Zhao et al. [63] trained a model on the mismatch between measurements and a model based on domain knowledge, and this hybrid model generalised better than standalone learners. Hughes et al. [22] reviewed machine learning for heat transfer and discussed when data-driven models are appropriate. Recent studies tested the generalisation of such models more directly. Shourehdeli and Gholipour [51] trained four learners on 24,614 saturated flow boiling points from 120 sources. Their mean absolute relative deviation was below 4% on random test points but about 30% on excluded data subsets. Kim et al. [27] derived an explicit correlation by symbolic regression from 12,406 points, which reached a MAPE of 13.6% on 2138 points from eight unseen sources. XGBoost, random forest, and neural network models were less consistent in these extrapolative cases. Phan et al. [42] added four universal correlations to the loss function of a neural network, which then extrapolated better to unseen fluids, especially with limited training data. Upadhyay et al. [55] identified data collection, feature selection, and performance metrics as the keys to reliable models of boiling and argued that physics-inspired techniques improve generalisation. Suh et al. [54] reviewed artificial intelligence for boiling and condensation and outlined physics-centred learning as a direction.

For nanofluids, Bahiraei et al. [2] reviewed artificial-intelligence methods for predicting and optimising nanofluid properties and performance. Ma et al. [35] reviewed machine learning for nanofluid heat transfer in renewable energy systems and stressed the non-linearity and the contradictory results of nanofluid data. Maleki et al. [36] found that the accuracy of models of hybrid-nanofluid viscosity and conductivity depended on the algorithm, the training data, and the choice of inputs. For flow boiling, Bouali et al. [5] trained a multilayer perceptron with eight hidden neurons on 354 literature records of five water-based nanofluids and reported a coefficient of determination of 0.9992. For $Al_2O_3$/water in an annulus, Shekhar et al. [50] tuned gradient boosting with a jellyfish search optimiser and reported 0.98453 on the test records. Serrao et al. [48] compared SVR, multilayer perceptron, and random forest models of the CHF of nanofluids in pool boiling, and the multilayer perceptron gave the highest mean $R^2$. Li et al. [32] measured the pool boiling HTC of $Al_2O_3$/water under hypergravity and reached $R^2 = 0.998$ with XGBoost. Riyadi et al. [45] reviewed machine learning for nanofluid heat transfer in porous media and heat exchangers, including models trained on data from computational fluid dynamics. Such records lie on grids of heat-flux, mass-flux, and concentration levels, where random cross-validation measures interpolation between neighbouring grid points.

Roberts et al. [46] showed that random cross-validation of structured data underestimates the error for new data and recommended blocked splits when new regions of the predictor space are predicted. Kapoor and Narayanan [25] found data leakage in 294 papers of 17 fields. In one case study, correcting it removed the apparent advantage of complex models over logistic regression. Varma and Simon [56] showed that tuning on the evaluation folds biases the error estimate, and that nested cross-validation removes this bias. The indicators of accuracy also matter. Willmott [61] argued that correlation-based measures are inadequate for model evaluation and recommended the root mean square error (RMSE). Legates and McCabe [30] showed that such measures are insensitive to additive and proportional bias, and Chai and Draxler [10] recommended reporting several complementary error measures. Loyola-Fuentes et al. [34] listed the data-splitting philosophy as one of six steps of a regression framework for heat transfer data, and they reported overfitting and limited extrapolation of the fitted models. For the CHF, Saraswat [47] found that the error of gradient boosting rose from 4.69% under random splits to 11.91% when complete experimental campaigns were withheld. Hybrids that learned the residual of a correlation reached 3.94 and 4.96% under the same validation. Furlong et al. [19] combined CHF correlations with neural-network ensembles, Bayesian neural networks, and deep Gaussian processes, and these hybrids outperformed purely data-driven models, especially with scarce data. Mondal and Sharma [39] reported coverages of 94.7 and 93.9% for the nominal 95% intervals of two probabilistic CHF models on a random test split.

Hybrid models that learn a correction to a physical baseline are an established approach in physics-guided machine learning [26, 60]. Nanofluid experiments offer a measured baseline, because the enhancement is judged against base-fluid runs at the same conditions. A model of the ratio between the nanofluid and base-fluid coefficients applies the residual idea to this measured baseline. For nanofluid flow boiling, its advantage has not been quantified when complete operating levels are withheld, and the part of the gain supplied by the measurement itself is not known. Prediction intervals, available from Gaussian process regression [44] and conformal prediction [31], are also rarely reported for such models.

The present study uses the $Al_2O_3$/water and $TiO_2$/water records to examine how validation design, target formulation, and model structure change the apparent and the actual accuracy. The anchored formulation examined here predicts the enhancement ratio for a measured base-fluid HTC at the same heat flux and mass flux. It is not an independent predictor of the nanofluid HTC at a new condition: the measured base fluid supplies the dependence on heat flux, mass flux, geometry, and regime. Within this scope, the contributions are as follows:

1. a test of base-fluid anchoring under schemes that withhold complete boiling curves or operating levels, with separate summaries for interior and extreme levels;
2. an evaluation of eight models with the coefficient of determination, the RMSE, the mean bias error, the mean absolute percentage error, and error-band fractions, including the sign of the bias at each withheld level;
3. a variance decomposition and an estimated-anchor test that attribute the gain of anchoring to the measured base-fluid coefficient at the predicted condition;
4. learning curves of both formulations, stated with the base-fluid measurements that the anchored formulation requires;
5. evidence that neither Gaussian process nor grouped conformal intervals were calibrated when an extreme operating level was withheld;
6. monotonicity-constrained models and a comparison with classical correlations evaluated with effective nanofluid properties;
7. a one-parameter calibration between the two nanofluids on all records and on their common coverage region.

# 2 Data

## 2.1 Source, variables, and coverage

The records originate from flow boiling experiments with water-based $Al_2O_3$ and $TiO_2$ nanofluids in a horizontal annular test section at 1 bar [14]. The inputs are the volumetric concentration $\varphi$, the mass flux $G$, and the heat flux $q$, and the output is the HTC, $h$. The design comprises six heat-flux levels (6.25, 26.04, 54.69, 93.75, 130.21, and 143.23 $\mathrm{kW\,m^{-2}}$) and five mass-flux levels (338.13 to 1014.3 $\mathrm{kg\,m^{-2}\,s^{-1}}$). Its five concentrations are 0 (water), 0.05, 0.10, 0.15, and 0.20 vol%. The records give neither the annulus dimensions, the heated length, the inlet temperature, the local quality, nor the measurement uncertainty.

Table 1 and Fig. 1 summarise the records. $Al_2O_3$/water covers the complete design for $q \leq 93.75$ $\mathrm{kW\,m^{-2}}$, three of the four concentrations at 130.21 $\mathrm{kW\,m^{-2}}$, and no record at 143.23 $\mathrm{kW\,m^{-2}}$. $TiO_2$/water covers all six heat-flux levels at the two lowest mass fluxes but lacks 29 conditions of the high-mass-flux, high-heat-flux corner, which hold the highest coefficients of the design. The two datasets share 80 conditions $(q, G, \varphi)$, their common coverage region.

**Table 1** Records used in the analysis. Ranges refer to the nanofluid records; $E$ is the ratio to the base-fluid coefficient at the same $q$ and $G$.

| Quantity | $Al_2O_3$/water | $TiO_2$/water |
|---|---|---|
| Nanofluid records | 95 | 91 |
| Base-fluid records | 25 | 29 |
| Heat flux $q$ (kW m$^{-2}$) | 6.25–130.21 | 6.25–143.23 |
| Mass flux $G$ (kg m$^{-2}$ s$^{-1}$) | 338.13–1014.3 | 338.13–1014.3 |
| Concentration $\varphi$ (vol%) | 0.05–0.20 | 0.05–0.20 |
| HTC $h$ (kW m$^{-2}$ K$^{-1}$) | 7.29–56.54 | 5.59–23.76 |
| Enhancement ratio $E$ | 1.606–2.418 | 1.231–1.632 |

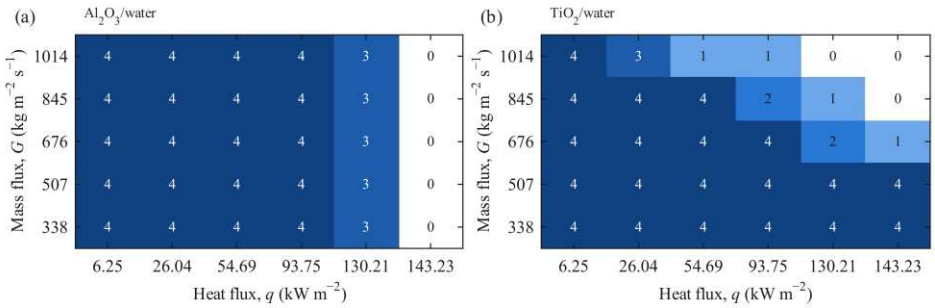



**Figure 1** Number of nanofluid concentrations available at each heat-flux and mass-flux level: (a) $Al_2O_3$/water; (b) $TiO_2$/water.

The base-fluid records supplied with the two nanofluids agree at 24 of their 25 common conditions and differ by 0.17 kW m$^{-2}$ K$^{-1}$ (1.05%) at the remaining one. Each nanofluid was referred to its own base-fluid records, and the enhancement ratio was defined as

$$E = \frac{h(q, G, \varphi)}{h_{\mathrm{bf}}(q, G)}, \tag{1}$$

where $h_{\mathrm{bf}}$ is the base-fluid coefficient measured at the same heat flux and mass flux. It is available for all 186 nanofluid records.

## 2.2 Structure of the records

The base-fluid coefficient depends mainly on mass flux (Fig. 2a). Between the lowest and highest heat flux, $h_{\mathrm{bf}}$ increases by 12 to 22% at the two lowest mass fluxes and by 50 to 84% at the three highest. The effects of ln $q$ and ln $G$ are therefore not additive. The enhancement ratio lies between 1.606 and 2.418 for $Al_2O_3$/water and between 1.231 and 1.632 for $TiO_2$/water, and most of it appears at the lowest concentration (Fig. 2b).

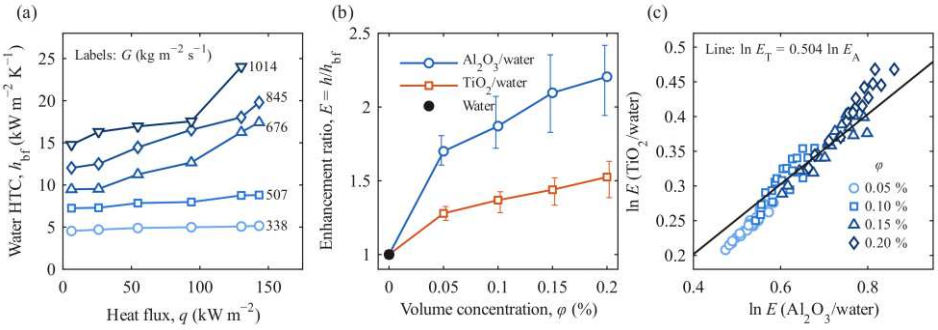



**Figure 2** Structure of the records: (a) base-fluid HTC versus heat flux, labelled by mass flux; (b) mean enhancement ratio versus concentration, with range bars; (c) log enhancement ratios at common conditions.

On all 39 boiling curves with more than one record, $E$ at the highest heat flux exceeds $E$ at the lowest (Fig. 3). The ratio also increases with mass flux at every heat flux and concentration. At fixed mass flux and concentration, $E$ changes by at most 16.7% ($Al_2O_3$/water) and 15.7% ($TiO_2$/water) along heat flux. At fixed heat flux, it changes by at most 13.1 and 9.0% along mass flux. Over the same ranges, $h_{bf}$ changes by up to 84% along heat flux and 373% along mass flux. For $Al_2O_3$/water at 0.15 and 0.20 vol%, the two highest mass fluxes lie 0.08 to 0.16 above the other curves, for a reason that the records do not identify.

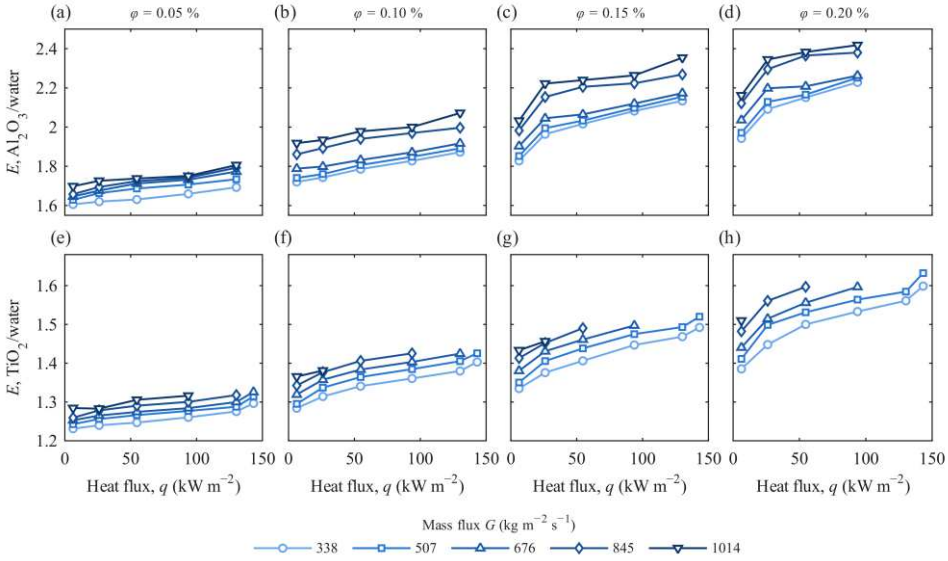



**Figure 3** Enhancement ratio versus heat flux for each mass flux: (a–d) $Al_2O_3$/water at 0.05, 0.10, 0.15, and 0.20 vol%; (e–h) $TiO_2$/water at the same concentrations.

At the 80 common conditions, ln $E$ of $TiO_2$/water is proportional to ln $E$ of $Al_2O_3$/water (Fig. 2c). The least-squares factor through the origin is 0.504, and the correlation coefficient

is 0.980. Fitted per concentration, the factor is 0.460, 0.508, 0.502, and 0.534 at 0.05, 0.10, 0.15, and 0.20 vol%, so the proportionality is an approximation.

# 3 Methods

## 3.1 Formulations and scope

In the direct formulation, a model $f$ predicts $\ln h$ from $\mathbf{x} = (\ln q, \ln G, \varphi)$. In the anchored formulation, a model $g$ predicts $\ln E$, and the coefficient is recovered as

$$\hat{h}(\mathrm{x}) = h_{\mathrm{bf}}(q, G)\exp\left[\hat{g}(\mathrm{x})\right]. \tag{2}$$

Here, $h_{\mathrm{bf}}$ was the measured base-fluid coefficient at the heat flux and mass flux of the predicted record. The anchored formulation therefore predicts the enhancement ratio given this measurement. It is not an independent predictor of the nanofluid HTC at a new condition, because the anchor supplies the dependence on heat flux, mass flux, geometry, and boiling regime. Both formulations were fitted to the 186 nanofluid records only.

## 3.2 Models

The physically structured log-linear regression (PLR) expresses the target $y$ ($\ln h$ or $\ln E$) as

$$y = \beta_0 + \beta_1 \ln G + \beta_2 \ln q + \beta_3 \ln G \ln q + \beta_4 (\ln q)^2 + \beta_5 \varphi + \beta_6 \varphi^2 + \beta_7 \varphi \ln G + \beta_8 \varphi \ln q, \tag{3}$$

with nine coefficients estimated by least squares. Its power-law and product terms follow classical boiling correlations [13, 33] and the interactions of Fig. 2. The monotone variant PLR-M was fitted subject to $\partial y/\partial \ln q \geq 0$ and $\partial y/\partial \ln G \geq 0$. Both derivatives are affine in $(\ln q, \ln G, \varphi)$, so 12 linear constraints at the corners of the design box enforce them in the whole box. In the anchored formulation, the constraints make $E$ non-decreasing in $q$ and $G$, which is sufficient for $\hat{h}$ to be non-decreasing wherever $h_{\mathrm{bf}}$ is. The measured $h$ and $h_{\mathrm{bf}}$ are non-decreasing along $q$ and $G$ at every consecutive step, and $E$ at every step but one (−0.12%).

Support vector regression (SVR) used a radial basis function kernel [52] with $C \in \{1, 10, 100, 1000\}$, $\gamma \in \{0.05, 0.2, 0.5, 1.5\}$, and $\varepsilon \in \{0.005, 0.02, 0.05\}$. The neural network (MLP) had one hidden layer of eight neurons [5], an $L_2$ penalty in $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$, and five averaged initialisations. Both were tuned in every training fold by an inner grouped cross-validation with at most five folds. Gaussian process regression (GPR) used a Matérn kernel ($\nu = 5/2$) with one length scale per input and white noise, fitted by maximum marginal likelihood [44]. The random forest (RF) had 500 trees [6]. Gradient boosting (GBM) had 400 trees of depth 3, a learning rate of 0.05, and a subsampling fraction of 0.8 [18]. The monotone histogram gradient boosting HGB-M had the same number of trees, depth, and learning rate, at least three records per leaf, and non-decreasing constraints in $\ln q$ and $\ln G$. The unconstrained booster served as its control, and all models used scikit-learn [41].

## 3.3 Validation schemes

Resubstitution and random 10-fold cross-validation, repeated ten times, were compared with four grouped schemes. Leave-one-curve-out (LOCC) withholds one of the 20 boiling curves ($G, \varphi$) of a nanofluid, which always lies inside the training ranges of $G$ and $\varphi$. LO$q$O, LO$G$O, and LO$\varphi$O withhold all records at one level of $q$, $G$, or $\varphi$. Folds that withhold the lowest or highest level of their variable are called extreme-level folds and require extrapolation; the other folds are interior-level folds. The extreme levels are 6.25 and 130.21 or 143.23 kW m$^{-2}$, 338.13 and 1014.3 kg m$^{-2}$ s$^{-1}$, and 0.05 and 0.20 vol%. Hyper-parameters were selected with the training records of each fold only.

Two tests addressed the role of the anchor. Under LO$q$O and LO$G$O, the base-fluid records of the withheld level were also withheld. The coefficient $h_{\mathrm{bf}}$ at that level was then estimated in two ways. The first estimate interpolated or extrapolated ln $h_{\mathrm{bf}}$ linearly in ln $q$ or ln $G$ from the two nearest measured levels. The second fitted the same model family to the remaining base-fluid records. A concentration-mean baseline without heat-flux or mass-flux information assigned to each withheld record the mean ln $E$ of the training records at its concentration, interpolated linearly in $\varphi$ under LO$\varphi$O.

## 3.4 Statistical performance indicators

For each nanofluid, formulation, model, and scheme, the pooled out-of-fold predictions $\hat{h}_i$ of the $n$ records were compared with the measured coefficients $h_i$ through

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{h}_i - h_i)^2}{\sum_{i=1}^{n}(h_i - \bar{h})^2}, \qquad \mathrm{RMSE} = \left[\frac{1}{n}\sum_{i=1}^{n}(\hat{h}_i - h_i)^2\right]^{1/2}, \tag{4}$$

$$\mathrm{MBE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{h}_i - h_i), \qquad \mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{|\hat{h}_i - h_i|}{h_i}, \tag{5}$$

where $\bar{h}$ is the mean measured coefficient. $R^2$ is the coefficient of determination of the predictions, not the squared correlation coefficient, and it becomes negative for predictions worse than $\bar{h}$. The mean bias error (MBE) is positive for over-prediction. RMSE and MBE are given in kW m$^{-2}$ K$^{-1}$ and, where stated, relative to $\bar{h}$. The shares of records with an absolute percentage error of at most 5, 10, and 20% describe the distribution of the errors. For the grouped schemes, the signed mean relative error (MRE) of the records of each withheld level, $100 n_k^{-1}\sum(\hat{h}_i - h_i)/h_i$, gives the direction of the bias at that level. Under random cross-validation, each indicator was computed for each of the ten repetitions and averaged. The indicators were reported together, because correlation-based measures are insensitive to additive and proportional bias [30], RMSE weights the largest coefficients, and MAPE weights all records equally [10, 62].

## 3.5 Variance decomposition and scatter of the records

A complete factorial sub-grid holds one record per combination of levels. On such a sub-grid, the sum of squares of $\ln h$ or $\ln E$ divides exactly into three main effects, three two-factor interactions, and a three-factor term. The levels were treated as categories. Complete sub-grids exist for $Al_2O_3$/water at $q \leq 93.75$ kW m$^{-2}$ (80 records), for $TiO_2$/water at $q \leq 93.75$ kW m$^{-2}$ and $G \leq 676.27$ kg m$^{-2}$ s$^{-1}$ (48 records), and for both at $q \leq 54.69$ kW m$^{-2}$ and $G \leq 845.33$ kg m$^{-2}$ s$^{-1}$ (48 records each). The three-factor term contains the random scatter of the records and any genuine three-factor interaction. Its mean square per degree of freedom is therefore an upper estimate of the random scatter of a record, but it excludes systematic errors shared by a level.

## 3.6 Learning curves and prediction intervals

For eleven numbers of training boiling curves between $m = 3$ and 19, 100 random subsets were drawn (all 20 for $m = 19$). The six unconstrained models were fitted in both formulations to the same subsets. The anchored models used the measured $h_{bf}$ at every training and withheld condition, that is, at all 25 ($Al_2O_3$/water) or 27 ($TiO_2$/water) heat-flux and mass-flux conditions.

Nominal 90% prediction intervals were evaluated under LOCC, LO$q$O, and LO$G$O. GPR gave the predictive mean plus or minus 1.645 predictive standard deviations of the logarithmic target. For PLR and SVR, grouped conformal intervals [31] used the leave-one-group-out residuals inside the training records of each outer fold as conformity scores. Their quantile of order $\lceil 0.9(n + 1) \rceil / n$ gave the half-width. As a calibration check, the factor $c_{90}$ by which the nominal half-width must be multiplied to cover 90% of the withheld records was computed. Coverage was also evaluated separately for interior-level and extreme-level folds. Because $c_{90}$ uses the withheld records, it measures the miscalibration and is not a calibration method.

## 3.7 Calibration between the two nanofluids

Motivated by Fig. 2c, a one-parameter calibration was tested,

$$\ln E_T(\mathrm{x}) = \beta \ln \hat{E}_S(\mathrm{x}). \tag{6}$$

Here, $\hat{E}_S$ is the prediction of an anchored model fitted to all records of the source nanofluid. The factor $\beta$ was estimated by least squares from $k = 1$ to 32 records of the target nanofluid, with 200 random draws for each $k$. The calibration was compared with GPR and PLR models fitted to the same $k$ target records. The calibration and the comparisons between the nanofluids were repeated on their 80 common conditions.

## 3.8 Correlations with effective properties

Without the geometry, inlet temperature, and local quality, flow boiling correlations cannot be evaluated for $h_{bf}$. The enhancement ratio that a correlation predicts, $E_{corr}$, was therefore computed as the ratio of the correlation with effective nanofluid properties to the same correlation with water properties. Both were evaluated at equal $q$, $G$, hydraulic diameter,

and quality. Conductivity followed the Maxwell model [37], viscosity the Brinkman model [7], and density and heat capacity volume-weighted mixing. The vapour properties, surface tension, and latent heat were those of saturated water at 1 atm. The nominal particle properties were 3970 kg m$^{-3}$, 765 J kg$^{-1}$ K$^{-1}$, and 40 W m$^{-1}$ K$^{-1}$ for $Al_2O_3$ and 4250 kg m$^{-3}$, 686.2 J kg$^{-1}$ K$^{-1}$, and 8.95 W m$^{-1}$ K$^{-1}$ for $TiO_2$. Their sensitivity was examined over 3000 to 5000 kg m$^{-3}$, 500 to 1000 J kg$^{-1}$ K$^{-1}$, and 5 W m$^{-1}$ K$^{-1}$ to infinity. The correlations of Chen [11], Gungor and Winterton [20], Liu and Winterton [33], and Kandlikar [24] were evaluated at every design heat flux and mass flux. The hydraulic diameter ranged from 2 to 30 mm and the quality from 0.001 to 0.5.

# 4 Results

## 4.1 Apparent and actual accuracy

Table 2 and Fig. 4 give the MAPE of all models, formulations, and schemes. In the direct formulation, resubstitution and random cross-validation give 0.26 to 5.44%. Gradient boosting has the lowest resubstitution error for $Al_2O_3$/water (0.30%) but 37.44% under LO*G*O. Under LO*G*O, the three tree ensembles reach 37.4 to 43.3% for both nanofluids. The SVR error for $Al_2O_3$/water increases from 2.74% under resubstitution to 4.01% under random cross-validation and 8.34% under LO*G*O. Every direct model with a resubstitution MAPE below 1% has errors of 4.99 to 43.31% when a heat-flux or mass-flux level is withheld. The PLR has the highest resubstitution error of the unconstrained direct models (4.25 and 3.44%) but the lowest LO*G*O error (5.29 and 5.15%). No direct model is the most accurate under all grouped schemes.

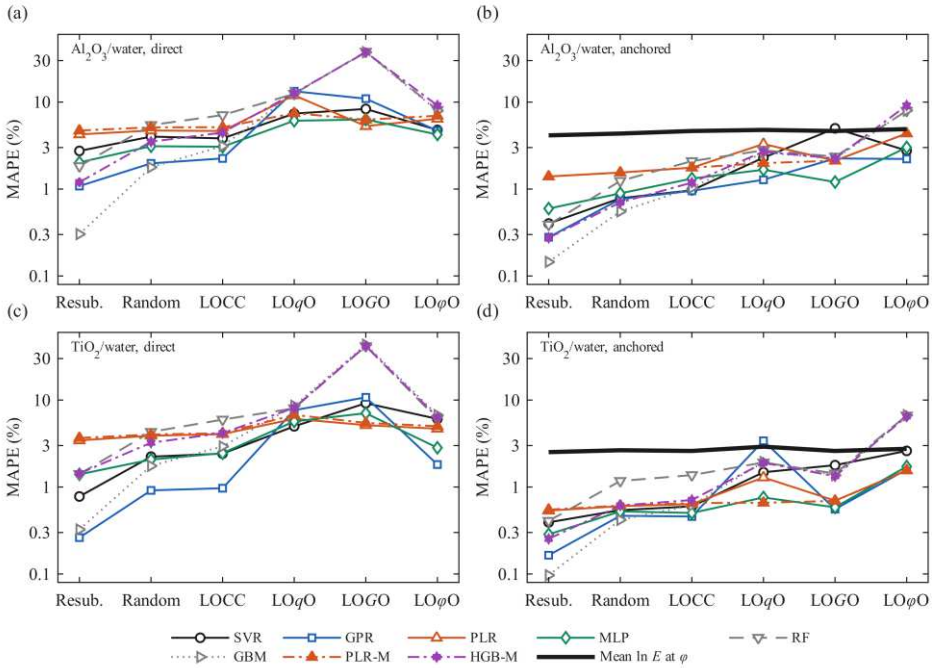



**Figure 4** MAPE of the out-of-fold predictions versus validation scheme: (a) $Al_2O_3$/water, direct; (b) $Al_2O_3$/water, anchored; (c) $TiO_2$/water, direct; (d) $TiO_2$/water, anchored. Thick line, concentration-mean baseline.

**Table 2** MAPE (%) of the out-of-fold predictions. The lowest model value in each column and formulation is underlined; the $\varphi$-mean baseline uses no heat-flux or mass-flux information.

| | $Al_2O_3$/water | | | | | | $TiO_2$/water | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | Res. | Rand. | LOCC | LO$q$O | LO$G$O | LO$\varphi$O | Res. | Rand. | LOCC | LO$q$O | LO$G$O | LO$\varphi$O |
| Direct formulation | | | | | | | | | | | | |
| PLR | 4.25 | 4.71 | 4.65 | 11.98 | 5.29 | 6.42 | 3.44 | 3.89 | 4.06 | 6.09 | 5.15 | 4.68 |
| SVR | 2.74 | 4.01 | 3.82 | 7.38 | 8.34 | 4.78 | 0.78 | 2.23 | 2.42 | 4.99 | 9.20 | 6.08 |
| GPR | 1.08 | 1.97 | 2.25 | 13.31 | 10.95 | 4.68 | 0.26 | 0.92 | 0.97 | 7.65 | 10.75 | 1.82 |
| MLP | 2.02 | 3.11 | 3.07 | 6.10 | 6.31 | 4.24 | 1.40 | 2.07 | 2.45 | 5.68 | 7.06 | 2.83 |
| RF | 1.83 | 5.44 | 7.08 | 12.45 | 37.98 | 8.12 | 1.41 | 4.30 | 5.96 | 8.04 | 41.76 | 6.05 |
| GBM | 0.30 | 1.78 | 3.12 | 12.66 | 37.44 | 7.94 | 0.32 | 1.75 | 2.93 | 8.59 | 43.31 | 6.72 |
| PLR-M | 4.71 | 5.13 | 5.14 | 7.46 | 6.26 | 6.97 | 3.66 | 4.02 | 4.15 | 6.76 | 5.46 | 5.01 |
| HGB-M | 1.20 | 3.50 | 4.46 | 12.65 | 37.43 | 9.14 | 1.43 | 3.23 | 4.17 | 8.16 | 42.02 | 6.35 |
| Anchored formulation | | | | | | | | | | | | |
| PLR | 1.39 | 1.55 | 1.76 | 3.30 | 2.12 | 4.36 | 0.54 | 0.60 | 0.64 | 1.29 | 0.70 | 1.55 |
| SVR | 0.40 | 0.78 | 0.97 | 2.30 | 4.99 | 2.77 | 0.39 | 0.54 | 0.60 | 1.48 | 1.78 | 2.60 |
| GPR | 0.28 | 0.77 | 0.95 | 1.28 | 2.25 | 2.22 | 0.16 | 0.47 | 0.46 | 3.40 | 0.56 | 1.59 |
| MLP | 0.60 | 0.89 | 1.32 | 1.67 | 1.20 | 3.01 | 0.29 | 0.52 | 0.50 | 0.76 | 0.58 | 1.72 |
| RF | 0.39 | 1.24 | 2.10 | 2.79 | 2.36 | 7.94 | 0.41 | 1.17 | 1.37 | 1.92 | 1.43 | 6.64 |
| GBM | 0.15 | 0.55 | 1.03 | 2.64 | 2.24 | 8.01 | 0.10 | 0.42 | 0.62 | 1.94 | 1.36 | 6.69 |
| PLR-M | 1.39 | 1.55 | 1.76 | 1.99 | 2.12 | 4.36 | 0.55 | 0.61 | 0.65 | 0.66 | 0.69 | 1.56 |
| HGB-M | 0.28 | 0.70 | 1.19 | 2.70 | 2.25 | 9.15 | 0.25 | 0.61 | 0.71 | 1.90 | 1.33 | 6.49 |
| $\varphi$-mean | 4.14 | 4.35 | 4.65 | 4.79 | 4.65 | 4.89 | 2.52 | 2.65 | 2.60 | 2.91 | 2.60 | 2.75 |

## 4.2 Coefficient of determination, RMSE, and mean bias error

Table 3 gives $R^2$, RMSE, and MBE under random cross-validation and LO*G*O, and Fig. 5 gives them under all six schemes. Under random cross-validation, the direct models reached $R^2$ values of 0.952 to 0.996 and RMSE values of 0.49 to 2.40 kW m$^{-2}$ K$^{-1}$, and their MBE lay between −0.20 and +0.05 kW m$^{-2}$ K$^{-1}$. By $R^2$, GPR and gradient boosting ranked first, with 0.991 to 0.996.

**Table 3** $R^2$, RMSE, and MBE (kW m$^{-2}$ K$^{-1}$) of the out-of-fold predictions under random cross-validation (mean of ten repetitions) and LO*G*O.

| | $Al_2O_3$/water | | | | | | $TiO_2$/water | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Random | | | LO*G*O | | | Random | | | LO*G*O | | |
| Model | $R^2$ | RMSE | MBE | $R^2$ | RMSE | MBE | $R^2$ | RMSE | MBE | $R^2$ | RMSE | MBE |
| Direct formulation | | | | | | | | | | | | |
| PLR | 0.9628 | 2.12 | −0.05 | 0.9578 | 2.26 | 0.03 | 0.9592 | 1.14 | −0.03 | 0.9411 | 1.37 | −0.08 |
| SVR | 0.9518 | 2.40 | −0.05 | 0.9571 | 2.28 | 0.19 | 0.9830 | 0.73 | −0.02 | 0.9256 | 1.54 | −0.20 |
| GPR | 0.9958 | 0.71 | 0.04 | 0.8911 | 3.63 | −0.03 | 0.9912 | 0.53 | 0.05 | 0.9177 | 1.62 | 0.01 |
| MLP | 0.9781 | 1.62 | −0.05 | 0.9618 | 2.15 | 0.39 | 0.9874 | 0.62 | −0.03 | 0.9428 | 1.35 | −0.02 |
| RF | 0.9611 | 2.17 | −0.20 | 0.5461 | 7.41 | 4.00 | 0.9702 | 0.97 | −0.06 | 0.2165 | 4.99 | 3.68 |
| GBM | 0.9953 | 0.75 | −0.01 | 0.5450 | 7.41 | 4.06 | 0.9923 | 0.49 | −0.02 | 0.1050 | 5.33 | 4.02 |
| PLR-M | 0.9583 | 2.24 | −0.08 | 0.9546 | 2.34 | 0.01 | 0.9537 | 1.21 | −0.03 | 0.9371 | 1.41 | −0.01 |
| HGB-M | 0.9782 | 1.62 | −0.08 | 0.5494 | 7.38 | 4.08 | 0.9747 | 0.89 | −0.05 | 0.2042 | 5.03 | 3.79 |
| Anchored formulation | | | | | | | | | | | | |
| PLR | 0.9978 | 0.52 | −0.03 | 0.9967 | 0.63 | −0.12 | 0.9997 | 0.10 | 0.00 | 0.9996 | 0.12 | −0.01 |
| SVR | 0.9994 | 0.27 | −0.01 | 0.9899 | 1.11 | 0.46 | 0.9996 | 0.11 | 0.00 | 0.9985 | 0.22 | 0.08 |
| GPR | 0.9993 | 0.29 | −0.01 | 0.9970 | 0.60 | 0.17 | 0.9997 | 0.09 | 0.00 | 0.9997 | 0.10 | −0.01 |
| MLP | 0.9992 | 0.31 | −0.01 | 0.9985 | 0.43 | 0.08 | 0.9997 | 0.10 | 0.00 | 0.9997 | 0.10 | 0.01 |
| RF | 0.9983 | 0.46 | −0.03 | 0.9955 | 0.74 | 0.19 | 0.9986 | 0.21 | −0.03 | 0.9984 | 0.22 | 0.02 |
| GBM | 0.9996 | 0.21 | −0.01 | 0.9960 | 0.70 | 0.23 | 0.9998 | 0.09 | −0.01 | 0.9988 | 0.20 | 0.11 |
| PLR-M | 0.9978 | 0.52 | −0.03 | 0.9967 | 0.63 | −0.12 | 0.9996 | 0.11 | 0.00 | 0.9996 | 0.12 | −0.01 |
| HGB-M | 0.9995 | 0.24 | −0.03 | 0.9960 | 0.70 | 0.24 | 0.9995 | 0.12 | −0.02 | 0.9988 | 0.19 | 0.10 |

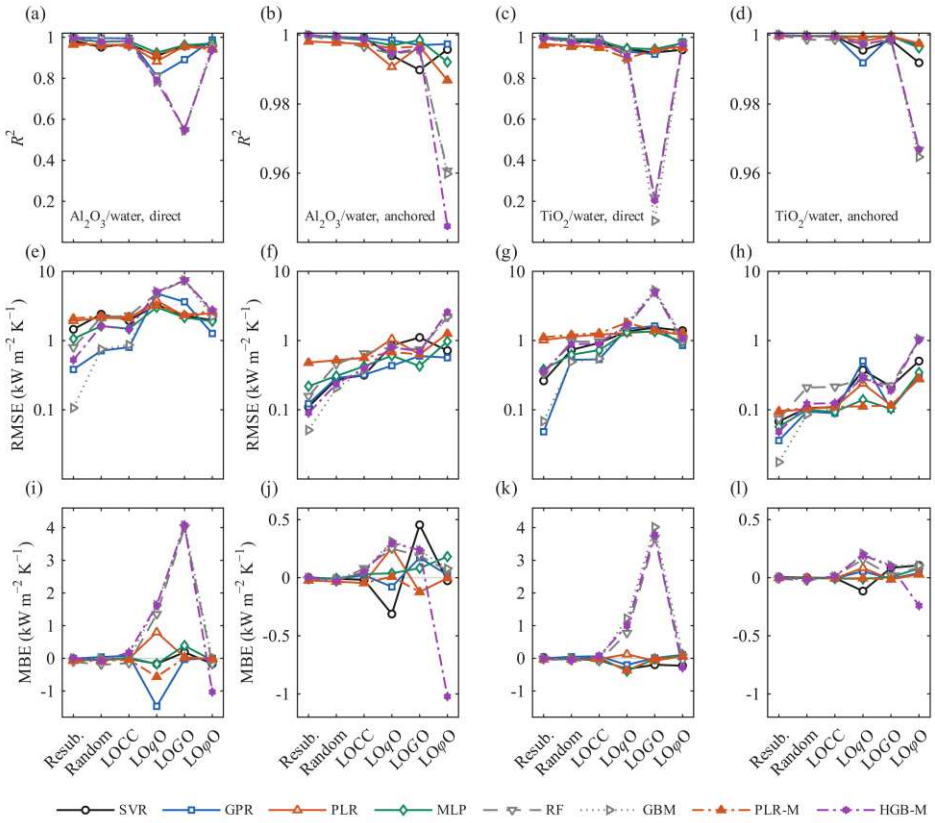



**Figure 5** $R^2$ (a–d), RMSE (e–h), and MBE (i–l) of the out-of-fold predictions versus validation scheme, for each nanofluid and formulation.

Under LO*G*O, the three tree ensembles fell to $R^2$ values of 0.545 to 0.549 for $Al_2O_3$/water and 0.105 to 0.216 for $TiO_2$/water, with RMSE values of 4.99 to 7.41 kW m$^{-2}$ K$^{-1}$. Their MBE of +3.68 to +4.08 kW m$^{-2}$ K$^{-1}$, 17.9 to 29.8% of the mean coefficient, shows a systematic over-prediction rather than random scatter. The smooth direct models kept $R^2$ values of 0.891 to 0.962 and MBE values between −0.20 and +0.39 kW m$^{-2}$ K$^{-1}$, although their MAPE rose to 5.15 to 10.95%. A high $R^2$ therefore did not indicate a small relative error at a withheld level: GPR reached $R^2 = 0.891$ for $Al_2O_3$/water with a MAPE of 10.95%. $R^2$ and MAPE also ranked the models differently, because $R^2$ is dominated by the largest coefficients. Under LO*G*O for $Al_2O_3$/water, MLP had the highest direct $R^2$ (0.962) but a MAPE of 6.31%, against 0.958 and 5.29% for PLR.

Under LO*q*O, the sign of the bias depended on the model. The direct GPR under-predicted $Al_2O_3$/water by 1.47 kW m$^{-2}$ K$^{-1}$ (−6.6% of the mean coefficient), and the direct PLR over-predicted it by 0.80 kW m$^{-2}$ K$^{-1}$ (+3.6%). The tree ensembles over-predicted both nanofluids by 0.77 to 1.62 kW m$^{-2}$ K$^{-1}$ (5.7 to 9.0%).

Anchoring raised $R^2$ under LO*G*O to 0.990 to 0.9997 and lowered the RMSE to 0.43 to 1.11 kW m$^{-2}$ K$^{-1}$ for $Al_2O_3$/water and 0.10 to 0.22 kW m$^{-2}$ K$^{-1}$ for $TiO_2$/water. Its largest

MBE, +0.46 kW m$^{-2}$ K$^{-1}$ for SVR with $Al_2O_3$/water, came from the lowest mass flux (Section 4.3). Under LO$\varphi$O, the anchored tree ensembles had the lowest $R^2$ values of the anchored formulation, 0.945 to 0.967, and HGB-M under-predicted $Al_2O_3$/water by 1.02 kW m$^{-2}$ K$^{-1}$ (−4.6%).

Figure 6 shows the predictions behind these indicators. Under random cross-validation (first of the ten repetitions), the gradient-boosting predictions lie close to the diagonal for every mass flux. Under LO*G*O, the predictions of each withheld mass-flux level form a separate band, above the diagonal for the four lower levels and below it for the highest. The direct SVR predictions scatter about the diagonal, with the largest departures at the lowest and highest mass flux. The anchored SVR predictions lie within ±10% except at the lowest mass flux of $Al_2O_3$/water.

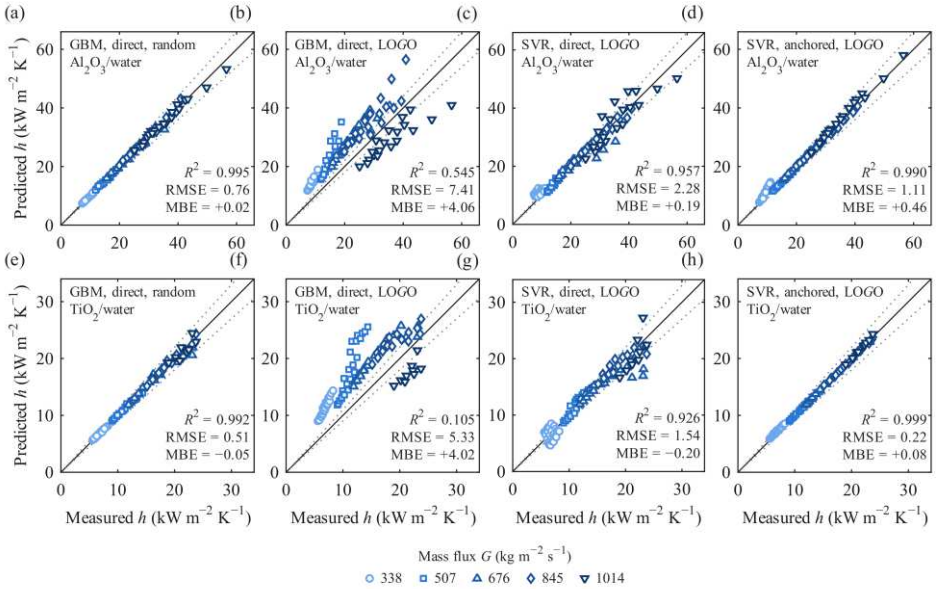



**Figure 6** Predicted against measured HTC for gradient boosting (direct) under random cross-validation and LO*G*O, and SVR (direct, anchored) under LO*G*O: (a–d) $Al_2O_3$/water; (e–h) $TiO_2$/water. Dotted lines, ±10%.

Figure 7 gives the distribution of the absolute percentage errors under LO*G*O. In the direct formulation, PLR and PLR-M kept 84.2 to 92.3% of the records within ±10%, and MLP kept 77.9 and 79.1%. SVR and GPR kept 63.7 to 73.7%, and the tree ensembles 5.5 to 11.0%. In the anchored formulation, every model placed all records within ±10%, except SVR for $Al_2O_3$/water (87.4%), and MLP placed all records of both nanofluids within ±5%.

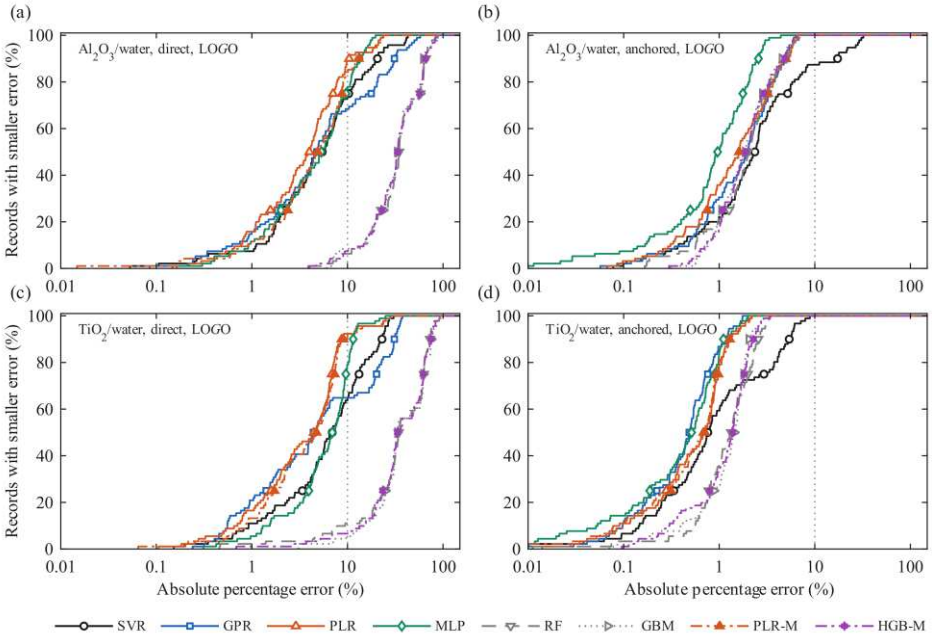



**Figure 7** Cumulative distribution of the absolute percentage error under LO*G*O: (a) $Al_2O_3$/water, direct; (b) $Al_2O_3$/water, anchored; (c) $TiO_2$/water, direct; (d) $TiO_2$/water, anchored. Markers, quartiles and 90th percentile.

## 4.3 Interior and extreme levels

For the five smooth models (PLR, SVR, GPR, MLP, and PLR-M), the extreme-level error exceeds the interior-level error by factors of 1.1 to 6.1 (Table 4). This holds in all 60 combinations of model, nanofluid, formulation, and scheme. Withholding the lowest or highest mass flux raises the direct SVR error from 4.12 to 14.68% for $Al_2O_3$/water and from 6.04 to 14.77% for $TiO_2$/water. The direct PLR under LO*q*O for $Al_2O_3$/water has 5.23% at interior and 23.56% at extreme levels, where its quadratic term in $\ln q$ is extrapolated. The monotonicity constraint reduces the extreme-level error to 11.69%. Anchoring keeps this pattern, as the anchored SVR error for $Al_2O_3$/water is 16.9% at the lowest mass flux and 1.3 to 2.5% at the interior levels (Fig. 8).

**Table 4** MAPE (%) of the records of interior-level folds (Int.) and extreme-level folds (Ext.), which withhold the lowest or highest level of the variable.

| | $Al_2O_3$/water | | | | | | $TiO_2$/water | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | LO*q*O | | LO*G*O | | LO$\varphi$O | | LO*q*O | | LO*G*O | | LO$\varphi$O | |
| Model | Int. | Ext. | Int. | Ext. | Int. | Ext. | Int. | Ext. | Int. | Ext. | Int. | Ext. |
| Direct formulation | | | | | | | | | | | | |
| PLR | 5.23 | 23.56 | 4.20 | 6.93 | 4.98 | 8.02 | 4.35 | 9.79 | 4.72 | 5.91 | 3.13 | 6.14 |
| SVR | 6.11 | 9.56 | 4.12 | 14.68 | 4.18 | 5.45 | 3.45 | 8.29 | 6.04 | 14.77 | 3.59 | 8.40 |
| GPR | 10.13 | 18.76 | 6.74 | 17.26 | 2.58 | 7.01 | 5.89 | 11.40 | 6.49 | 18.23 | 0.85 | 2.72 |
| MLP | 5.26 | 7.55 | 4.55 | 8.95 | 3.83 | 4.69 | 3.79 | 9.74 | 5.80 | 9.28 | 1.82 | 3.77 |
| RF | 12.06 | 13.13 | 35.70 | 41.39 | 6.26 | 10.19 | 8.78 | 6.43 | 35.31 | 53.10 | 4.64 | 7.37 |
| GBM | 13.29 | 11.60 | 35.14 | 40.89 | 7.56 | 8.36 | 9.97 | 5.65 | 37.76 | 53.08 | 6.44 | 6.98 |
| PLR-M | 5.00 | 11.69 | 5.01 | 8.15 | 5.34 | 8.78 | 4.67 | 11.23 | 4.88 | 6.49 | 3.21 | 6.68 |
| HGB-M | 13.28 | 11.57 | 35.16 | 40.84 | 9.82 | 8.38 | 9.57 | 5.14 | 35.73 | 53.07 | 5.75 | 6.92 |
| Anchored formulation | | | | | | | | | | | | |
| PLR | 1.74 | 5.96 | 1.68 | 2.77 | 2.53 | 6.40 | 0.59 | 2.78 | 0.63 | 0.81 | 0.81 | 2.25 |
| SVR | 1.01 | 4.50 | 1.68 | 9.96 | 2.12 | 3.49 | 0.81 | 2.91 | 0.63 | 3.82 | 0.92 | 4.18 |
| GPR | 1.00 | 1.75 | 1.46 | 3.43 | 1.81 | 2.68 | 3.26 | 3.71 | 0.49 | 0.67 | 0.80 | 2.33 |
| MLP | 1.17 | 2.53 | 1.13 | 1.32 | 2.03 | 4.10 | 0.58 | 1.13 | 0.50 | 0.73 | 1.06 | 2.34 |
| RF | 2.02 | 4.10 | 2.50 | 2.14 | 7.63 | 8.29 | 1.49 | 2.83 | 1.19 | 1.86 | 6.50 | 6.77 |
| GBM | 1.97 | 3.77 | 2.58 | 1.73 | 7.69 | 8.36 | 1.66 | 2.53 | 1.26 | 1.55 | 6.56 | 6.82 |
| PLR-M | 1.67 | 2.53 | 1.68 | 2.77 | 2.53 | 6.40 | 0.52 | 0.95 | 0.62 | 0.82 | 0.80 | 2.27 |
| HGB-M | 2.07 | 3.76 | 2.60 | 1.72 | 9.85 | 8.36 | 1.62 | 2.49 | 1.23 | 1.51 | 6.13 | 6.82 |

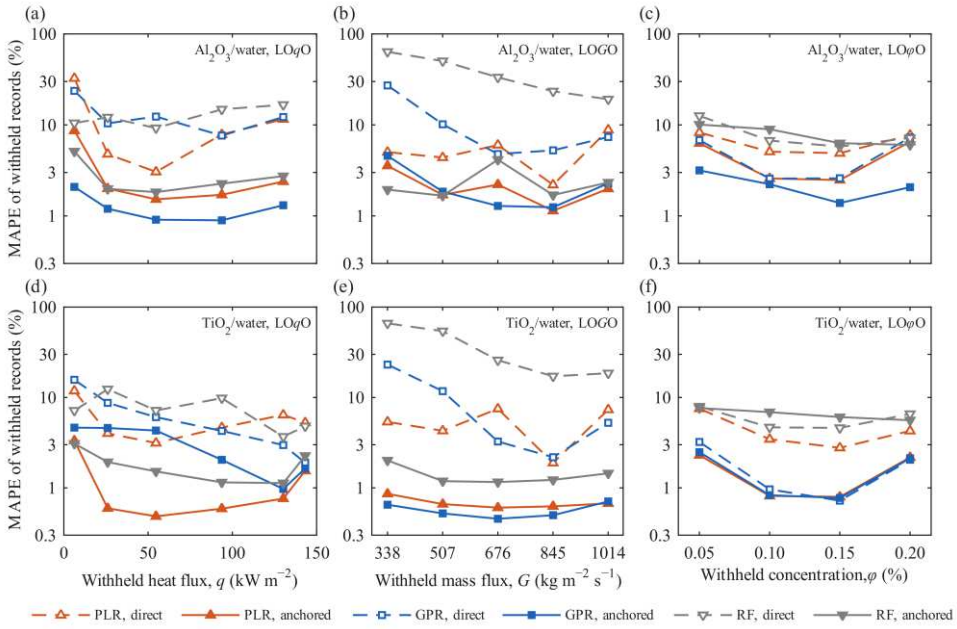



**Figure 8** MAPE of the records of each withheld level under (a, d) LO*q*O, (b, e) LO*G*O, and (c, f) LO$\varphi$O: (a–c) $Al_2O_3$/water; (d–f) $TiO_2$/water. Open markers, direct; filled, anchored.

The tree ensembles fail at interior as well as extreme mass-flux levels, with direct LO*G*O errors of 35.1 to 37.8% and 40.8 to 53.1%. Their piecewise-constant predictions assign a withheld level the prediction of an adjacent training level, and adjacent mass-flux levels differ in $h$ by a median of 25 to 63%. Their anchored LOCC errors of 0.62 to 2.10%

show that this failure concerns widely spaced levels, not interpolation between close conditions.

Figure 9 gives the signed mean relative error of each withheld level and shows which neighbour the tree ensembles chose. Under LO*G*O, the direct RF and GBM over-predicted the lowest mass flux by 63.3 to 66.1% and the three interior levels by 15.8 to 57.2%. They under-predicted the highest mass flux by 18.5 to 19.2%. The over-prediction at interior levels follows from the design. The mass-flux levels are equally spaced in $G$, so each interior level lies closer to its upper than to its lower neighbour in $\ln G$, the input of the models. A split placed midway between the two neighbours in $\ln G$ therefore assigns the withheld level to the upper neighbour, whose coefficient is higher. The pooled MBE of +3.68 to +4.08 kW m$^{-2}$ K$^{-1}$ is the sum of these level biases. Under LO$\varphi$O, the withheld interior concentrations lie exactly midway between their neighbours. In the single-precision arithmetic of the tree implementation, 0.10 vol% fell below and 0.15 vol% above such a split. RF and GBM therefore assigned them the lower and the upper neighbour, respectively. Their errors were −9.0 to −4.7% at 0.10 vol% and +1.8 to +6.3% at 0.15 vol%.

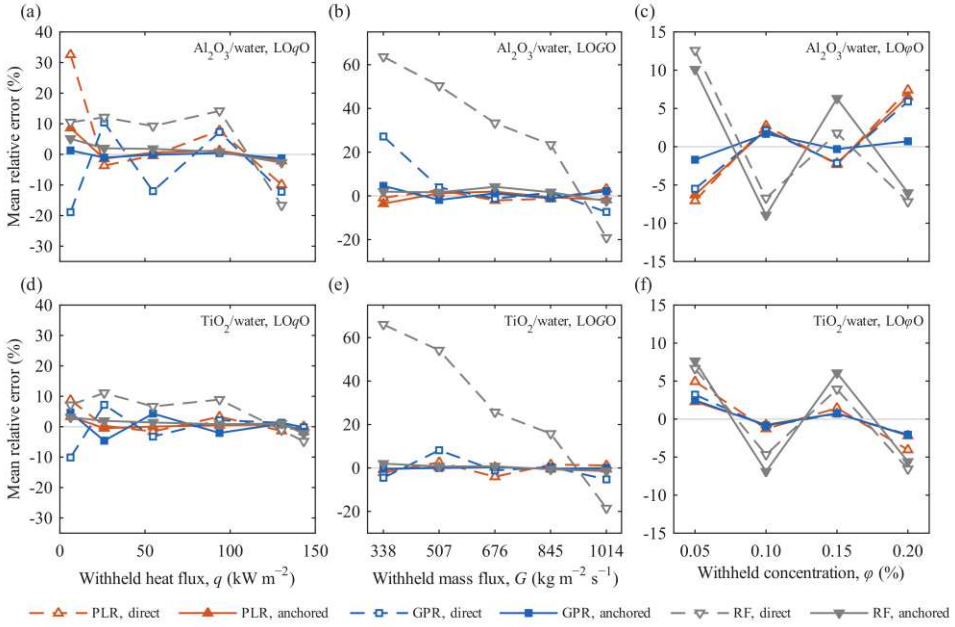



**Figure 9** Signed mean relative error of the records of each withheld level under (a, d) LO*q*O, (b, e) LO*G*O, and (c, f) LO$\varphi$O: (a–c) $Al_2O_3$/water; (d–f) $TiO_2$/water.

The smooth models extrapolated with biases of both signs. Under LO*q*O, all direct models under-predicted the highest heat flux of $Al_2O_3$/water by 9.9 to 16.7%. At the lowest heat flux, the direct PLR over-predicted by 32.6%, where its quadratic term in $\ln q$ is extrapolated, whereas the direct GPR under-predicted by 18.9%. In the anchored formulation, the level biases of the smooth models stayed between −4.9 and +8.6% under LO*q*O and LO*G*O. The exception was SVR at the lowest mass flux of $Al_2O_3$/water (+16.9%).

## 4.4 Anchoring as a change of the prediction task

Anchoring reduces the grouped errors of every model (Fig. 4b and d). Under LO*G*O, the lowest error falls from 5.29 to 1.20% for $Al_2O_3$/water and from 5.15 to 0.56% for $TiO_2$/water. The tree-ensemble errors fall from 37 to 43% to 1.33 to 2.36%. Three results show that these gains describe a change of the prediction task rather than a better learner.

First, the measured base fluid carries most of the variance of the direct target (Fig. 10). On the complete sub-grids, anchoring removes 94.7 and 96.1% of the variance of the logarithmic target. Mass flux carries 90.4 and 89.3% of the variance of ln *h*, and concentration carries 80.8 and 81.5% of that of ln *E*. The cells missing from the sub-grids lie at high heat and mass flux, so the heat-flux shares of ln *h* are probably underestimated.

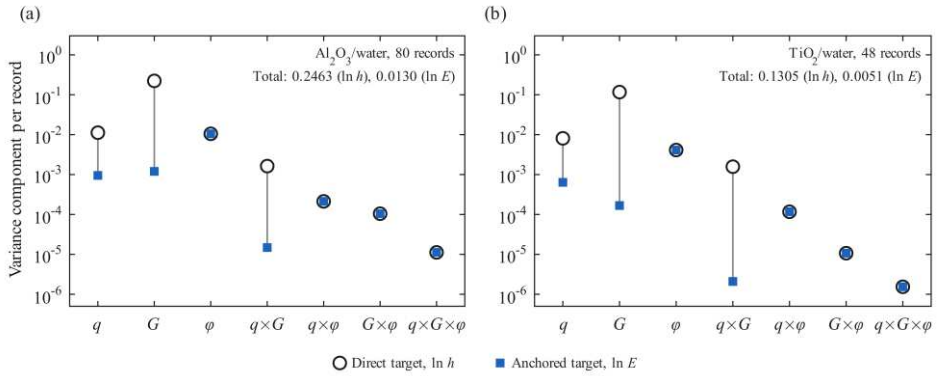



**Figure 10** Variance components of ln *h* and ln *E* on complete sub-grids: (a) $Al_2O_3$/water, $q \le 93.75$ kW m$^{-2}$, 80 records; (b) $TiO_2$/water, also $G \le 676.27$ kg m$^{-2}$ s$^{-1}$, 48 records.

Second, the concentration-mean baseline gives 4.14 to 4.89% for $Al_2O_3$/water and 2.52 to 2.91% for $TiO_2$/water under the six schemes (last row of Table 2). Under LO*q*O and LO*G*O, these values lie below the lowest direct errors, and the anchored models reduce them further by factors of 3.7 to 4.6.

Third, the gain requires the measured anchor at the withheld level (Table 5). With the base-fluid records of the withheld level also withheld, the interpolated $h_{bf}$ alone had errors of 5.63 to 8.67%. The lowest anchored errors were then 5.19 to 8.52% with the interpolated anchor and 4.14 to 6.16% with the same model family. These values are of the same order as the lowest direct errors of 4.99 to 6.10%.

**Table 5** Lowest MAPE (%) over the eight models under LO*q*O and LO*G*O for different sources of the base-fluid coefficient at the withheld level; model in parentheses.

| | $Al_2O_3$/water | | $TiO_2$/water | |
|---|---|---|---|---|
| Source of $h_{bf}$ | LO*q*O | LO*G*O | LO*q*O | LO*G*O |
| Direct (no $h_{bf}$) | 6.10 (MLP) | 5.29 (PLR) | 4.99 (SVR) | 5.15 (PLR) |
| Anchored, measured | 1.28 (GPR) | 1.20 (MLP) | 0.66 (PLR-M) | 0.56 (GPR) |
| Anchored, measured, $\varphi$-mean baseline | 4.79 | 4.65 | 2.91 | 2.60 |
| Anchored, interpolated | 7.18 (GPR) | 7.29 (MLP) | 5.19 (GPR) | 8.52 (MLP) |
| Anchored, same model family | 6.16 (MLP) | 5.21 (PLR) | 4.14 (MLP) | 5.62 (PLR) |
| Interpolated $h_{bf}$ with measured *E* | 7.23 | 7.24 | 5.63 | 8.67 |

Concentration is the remaining difficulty of the anchored formulation. LO$\varphi$O gives the largest anchored error in 13 of the 16 combinations of model and nanofluid (1.55 to 9.15%). This is consistent with the share of concentration in the variance of ln $E$. The extreme concentrations give the largest of these errors (Table 4).

## 4.5 Number of training curves

Figure 11 gives the MAPE on withheld boiling curves against the number of training curves $m$; at $m$ = 19 it reproduces the LOCC values. With three training curves, about 14 records, the direct errors are 13.3 to 50.8% and the anchored errors 1.93 to 8.04%. At this size, 55 to 56% of the withheld records lie outside the training range of mass flux or concentration. The anchored errors stay at or below 5% from $m$ = 3 to 6, depending on the model. The direct errors do so only from $m$ = 8 to 16, and those of RF not at all.

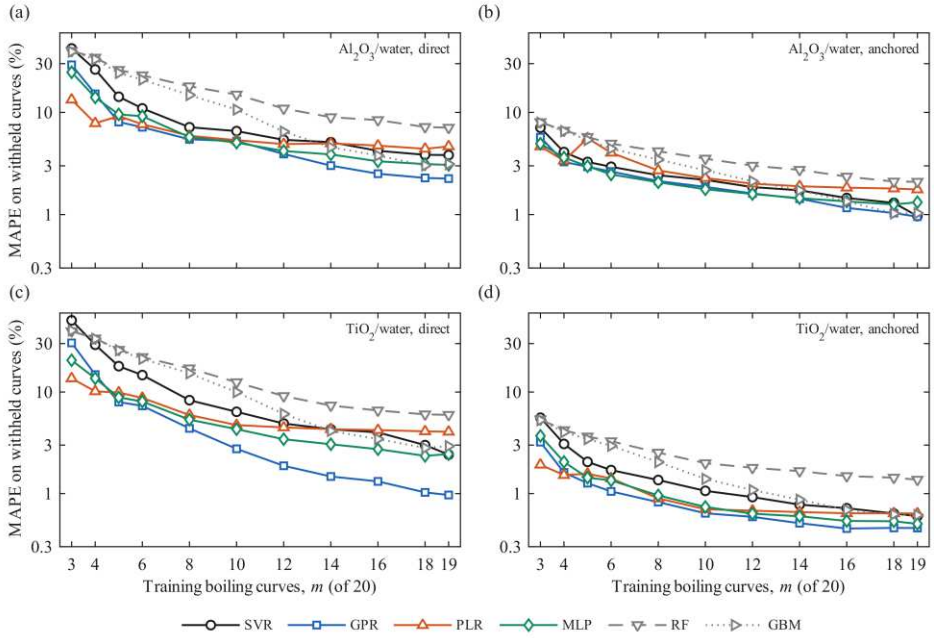



**Figure 11** MAPE on withheld boiling curves versus the number of training curves, pooled over random subsets: (a) $Al_2O_3$/water, direct; (b) $Al_2O_3$/water, anchored; (c) $TiO_2$/water, direct; (d) $TiO_2$/water, anchored.

With four training curves, the anchored models are more accurate than the direct models trained on twelve curves in 11 of the 12 combinations of model and nanofluid. This comparison counts nanofluid records only. The anchored models with four curves used on average 19.1 and 18.1 nanofluid records, together with measured base-fluid coefficients at all 25 and 27 conditions. The direct models with twelve curves used 57.1 and 54.7 nanofluid records and no base-fluid record. The learning curves therefore compare two measurement programmes, and they do not show that four nanofluid curves suffice without base-fluid runs at every condition.

## 4.6 Prediction intervals

Neither interval method is calibrated when a complete level is withheld (Fig. 12). The GPR intervals cover 54.7 to 87.4% of the withheld records under LO*q*O and LO*G*O. Their standard deviation would have to be multiplied by 1.14 to 4.70 to cover 90%. Under LOCC, they cover 84.2 to 90.1%. The grouped conformal intervals cover 83.2 to 95.6% (PLR) and 73.6 to 95.6% (SVR) of the records of a scheme.

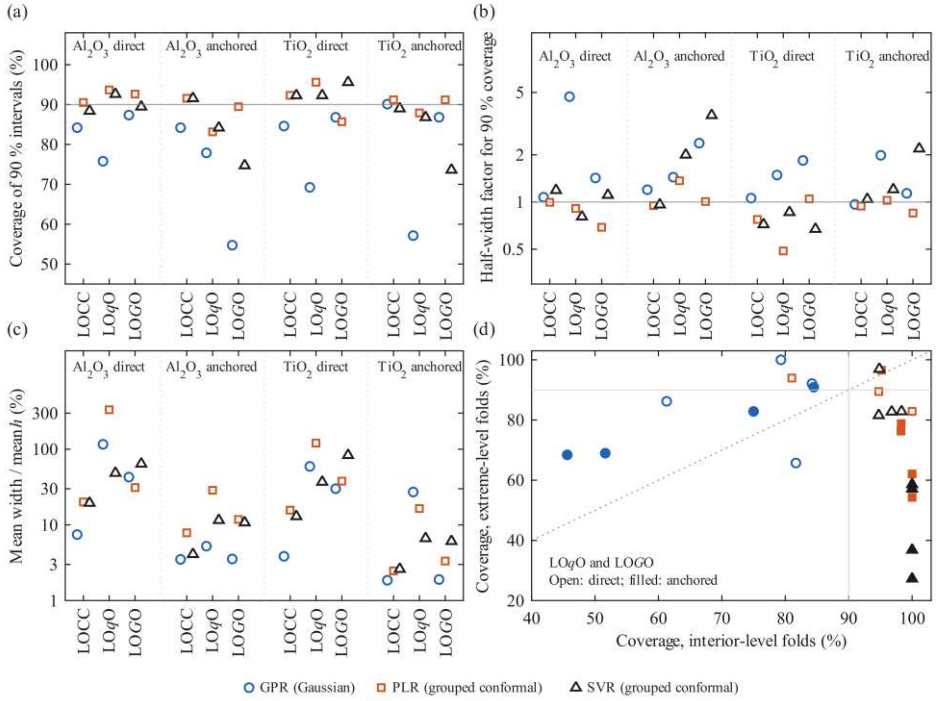



**Figure 12** Nominal 90% prediction intervals under grouped validation: (a) coverage; (b) half-width factor giving 90% coverage; (c) mean width relative to the mean HTC; (d) extreme-level against interior-level coverage.

These pooled values hide a dependence on the position of the withheld level (Fig. 12d). In the anchored formulation, the conformal intervals cover 98.2 to 100% of the records of interior-level folds but only 27.3 to 78.8% of those of extreme-level folds. Their conformity scores mix interior and extreme training groups. The half-width is therefore too wide for interpolation and too narrow for extrapolation, where the withheld group is not exchangeable with the training groups. The anchored GPR intervals cover 45.6 to 84.5% at interior and 68.4 to 90.9% at extreme levels. None of the methods therefore gives a calibrated uncertainty for a new extreme level. Anchoring narrows the PLR intervals under LO*q*O for $Al_2O_3$/water from 332 to 28.6% of the mean coefficient (Fig. 12c), but no width includes the unknown measurement uncertainty.

## 4.7 The two nanofluids and the one-parameter calibration

On the 80 common conditions (Table 6), the range limits and means of $E$ change by less than 0.06. Anchoring removes 93.8 and 97.1% of the variance on the common complete sub-grid. The lowest anchored errors remain lower for $TiO_2$/water under every grouped scheme, at 0.36 to 1.59% against 0.72 to 2.63%. The lower errors of $TiO_2$/water are therefore not a consequence of its smaller coverage, and they are consistent with the smaller variance of its enhancement ratio.

**Table 6** The two nanofluids on all records and on their 80 common conditions. MAPE rows (%) give the lowest value over the eight models; calibration rows use the PLR source.

| | $Al_2O_3$/water | | $TiO_2$/water | |
|---|---|---|---|---|
| Quantity | All | Common | All | Common |
| Nanofluid records | 95 | 80 | 91 | 80 |
| Range of $E$ | 1.606–2.418 | 1.606–2.365 | 1.231–1.632 | 1.231–1.597 |
| Mean $E$ | 1.956 | 1.911 | 1.393 | 1.382 |
| Variance of ln $h$ removed (%)[a] | 94.7 | 93.8 | 96.1 | 97.1 |
| Direct, LO$q$O | 6.10 | 6.23 | 4.99 | 5.88 |
| Direct, LO$G$O | 5.29 | 5.02 | 5.15 | 4.67 |
| Anchored, LOCC | 0.95 | 0.72 | 0.46 | 0.36 |
| Anchored, LO$q$O | 1.28 | 1.47 | 0.66 | 0.43 |
| Anchored, LO$G$O | 1.20 | 1.38 | 0.56 | 0.42 |
| Anchored, LO$\varphi$O | 2.22 | 2.63 | 1.55 | 1.59 |
| Calibration, $k$ = 4 | 3.78 | 3.69 | 1.88 | 1.76 |
| Calibration, $k$ = 32 | 3.44 | 3.35 | 1.72 | 1.61 |
| Calibration, measured $E$, $k$ = 80 | – | 3.46 | – | 1.75 |

[a]Complete sub-grids of Fig. 10 (All) and at $q \le 54.69$ kW m$^{-2}$ and $G \le 845.33$ kg m$^{-2}$ s$^{-1}$ (Common).

Applying the anchored $Al_2O_3$/water model to $TiO_2$/water without calibration gives 38.1% (Fig. 13a). With the calibration of Eq. (6), one $TiO_2$/water record reduces the error to 2.35% (PLR source), four to 1.88%, and 32 to 1.72%. Target-only models are less accurate up to eight records (2.24% for GPR with eight) and more accurate with 16 or 32 records (0.65 to 0.91%). The calibration error does not fall below about 1.7%, because a single factor cannot represent the increase of $\beta$ with concentration.

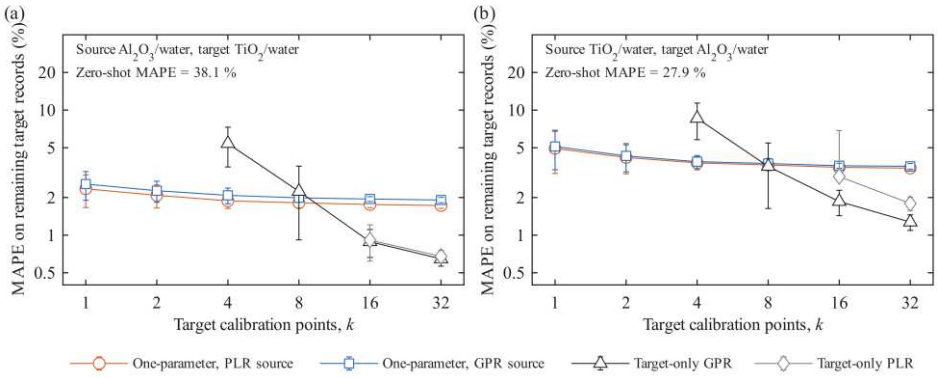



**Figure 13** MAPE on the remaining target records versus the number of calibration records: (a) $Al_2O_3$/water to $TiO_2$/water; (b) $TiO_2$/water to $Al_2O_3$/water. Bars show one standard deviation over 200 draws.

In the reverse direction, the calibration gives 4.94% with one record and 3.44% with 32 (Fig. 13b). On the common conditions, the two directions give 1.61 and 3.35% with 32

records, and even the measured ratios, with all 80 conditions used for calibration, give 1.75 and 3.46%. The asymmetry is therefore not caused by the coverage of the source records. It follows from the relation itself, whose residual in ln $E$ is multiplied by the reverse factor, about 2, when $Al_2O_3$/water is the target. The calibration rests on two oxides tested in one annulus and has not been tested for other particle materials or facilities.

## 4.8 Physical consistency and monotonicity constraints

The models were fitted to all records and evaluated on a grid of 40 heat-flux, 40 mass-flux, and 31 concentration values. Table 7 gives the share of adjacent grid steps along which the predicted $h$ (direct) or $E$ (anchored) decreases. In the direct formulation, the four smooth unconstrained models decrease along heat flux on 4.1 to 22.4% of the steps. In the anchored formulation, SVR, GPR, and MLP decrease along heat flux or mass flux on up to 18.6%. PLR-M and HGB-M have no decreasing step along $q$ or $G$.

**Table 7** Share (%) of adjacent grid steps along which the predicted $h$ (direct) or $E$ (anchored) decreases; models fitted to all records.

| | | Direct, $h$ | | | Anchored, $E$ | | |
|---|---|---|---|---|---|---|---|
| Nanofluid | Model | along $q$ | along $G$ | along $\varphi$ | along $q$ | along $G$ | along $\varphi$ |
| $Al_2O_3$/water | PLR | 12.96 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | SVR | 8.95 | 0.00 | 0.78 | 6.23 | 18.57 | 0.00 |
| | GPR | 21.56 | 0.00 | 0.00 | 0.00 | 5.69 | 0.00 |
| | MLP | 5.25 | 0.00 | 0.00 | 0.07 | 3.60 | 0.00 |
| | RF | 0.00 | 0.00 | 0.02 | 0.62 | 0.00 | 0.00 |
| | GBM | 0.37 | 0.42 | 0.24 | 0.31 | 0.37 | 0.08 |
| | PLR-M | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | HGB-M | 0.00 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 |
| $TiO_2$/water | PLR | 11.71 | 0.00 | 0.00 | 0.75 | 0.00 | 0.00 |
| | SVR | 20.07 | 0.61 | 1.25 | 0.33 | 0.03 | 0.00 |
| | GPR | 22.36 | 1.67 | 0.00 | 14.54 | 0.00 | 0.00 |
| | MLP | 4.11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | RF | 0.34 | 0.09 | 0.56 | 0.00 | 0.84 | 0.00 |
| | GBM | 0.26 | 0.28 | 0.31 | 0.50 | 0.34 | 0.06 |
| | PLR-M | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | HGB-M | 0.00 | 0.00 | 0.06 | 0.00 | 0.00 | 0.00 |

Figure 14 shows the direct models along heat flux at 0.10 vol%. Between the two lowest heat-flux levels, the widest gap in ln $q$, the SVR curves for $TiO_2$/water fall 5.8 and 11.7% below the lower measured value. The GPR curves fall by up to 5.8%. The PLR curves have a minimum near 15 kW m$^{-2}$ at the lowest mass flux, imposed by the quadratic term. The PLR-M curves are non-decreasing. Like the PLR curves, they depart from the records by up to 16.3 and 21.0% near the highest heat fluxes of the higher mass fluxes (Fig. 14b and d).

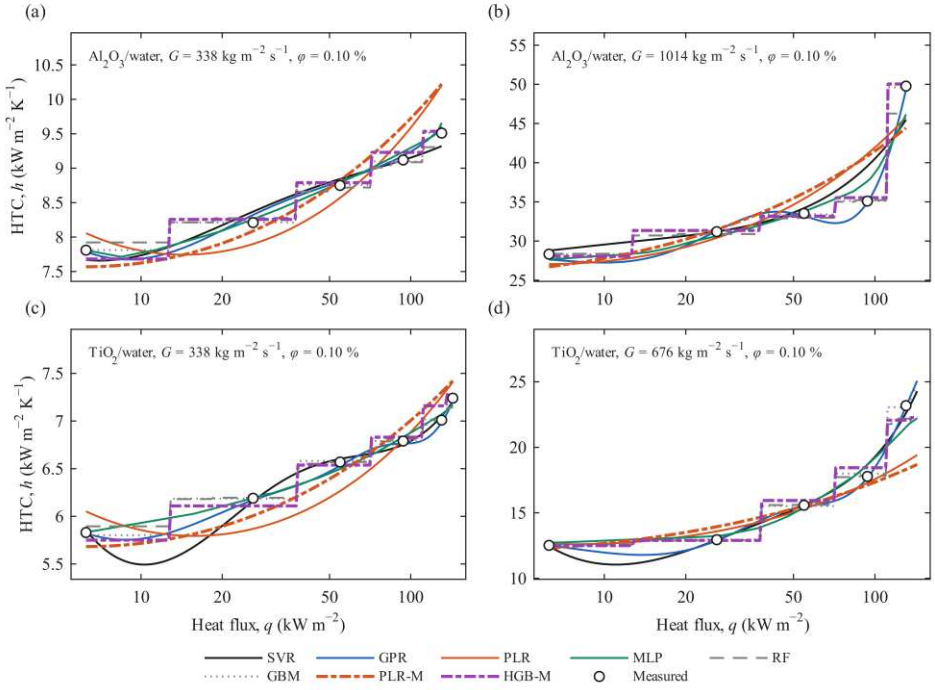



**Figure 14** Direct-formulation predictions along heat flux of the eight models fitted to all records, at 0.10 vol%: (a, b) $Al_2O_3$/water; (c, d) $TiO_2$/water. Circles, measured coefficients.

The effect of the constraints on grouped accuracy is mixed. PLR-M reduces the direct LO*q*O error for $Al_2O_3$/water from 11.98 to 7.46% and the anchored LO*q*O errors from 3.30 to 1.99% and from 1.29 to 0.66%. It raises the other direct grouped errors by 0.09 to 0.97 percentage points. HGB-M and the unconstrained booster differ by at most 0.8 percentage points under every grouped scheme. The constraints therefore remove decreases between measured levels without a systematic loss of grouped accuracy, but they do not repair the failure of piecewise-constant models between widely spaced levels.

## 4.9 Classical correlations with effective properties

At 0.05 to 0.20 vol%, the effective properties differ from those of water by at most 0.74%. The single-phase liquid term then gives $E_{corr}$ of 0.9979 to 0.9997, and the nucleate term of the Chen correlation 1.0006 to 1.0031 at equal wall superheat. The pool boiling term of Cooper [13] contains no transport property and gives exactly 1. The particle-property range changes these values by at most 0.002. The four correlations give $E_{corr}$ of 0.998 to 1.002 over all examined diameters, qualities, heat fluxes, and mass fluxes, against measured ratios of 1.231 to 2.418. Anchored to the measured base fluid, they reproduce the nanofluid coefficients with MAPE of 48.1 to 48.3% ($Al_2O_3$/water) and 27.7 to 28.0% ($TiO_2$/water), the values obtained by assuming no enhancement.

The base-fluid records also depart from nucleate boiling. The heat-flux exponent of $h_{bf}$, fitted at each mass flux, is 0.04 to 0.19, against 0.67 in the correlation of Cooper, and the mass-flux exponent is 1.06 to 1.56. Over the examined diameters and qualities, the four correlations give water coefficients of 2.2 to 220 $kW\,m^{-2}\,K^{-1}$. A comparison with the models under the same schemes therefore requires the geometry and inlet conditions, which are not available.

# 5 Discussion

## 5.1 Scope of the anchored formulation

For these records, the accuracy on the fitting data or under random cross-validation does not measure the accuracy for a new operating level. Reports of small-data boiling models would be more informative if they stated which levels the test and training records share, and whether the withheld levels are interior or extreme.

The anchored formulation is more accurate because the measured base fluid supplies about 95% of the variance of the logarithmic target, the part that is hardest to extrapolate. Its gain is a property of the task rather than of the learners. A concentration mean outperformed every direct model under LO*q*O and LO*G*O, and the gain disappeared when the base-fluid coefficient of the withheld level was estimated. The anchored models therefore predict the nanofluid HTC at a new level only when a base-fluid run at that level is available. They do not carry the dependence on geometry and boiling regime to another test section.

Relative to residual learning in hybrid models [26, 47, 63], the present study does not introduce a new principle. Unlike the model-based baselines of those studies, the baseline here is a measurement. It adds a quantitative test of an experimental anchor under withheld operating levels, an explanation of its gain, and the conditions under which the gain does not hold. These conditions are concentration extrapolation and extreme levels, where the intervals were not calibrated. The learning curves translate the gain into measurement effort, and four nanofluid curves sufficed only because the base fluid was measured at every condition.

## 5.2 What the statistical indicators add

Each indicator answered a different question for these records. $R^2$ stayed above 0.89 for the smooth direct models under every grouped scheme except LO*q*O, although their MAPE rose from random cross-validation to LO*G*O by factors of 1.1 to 11.7. A reported $R^2$ near 0.99, as in earlier models of nanofluid flow boiling [5, 50], therefore does not establish a small relative error at a new operating level. The MBE separated systematic from random error. It was within $\pm 0.20\ kW\,m^{-2}\,K^{-1}$ for every direct model under random cross-validation but reached +3.68 to +4.08 $kW\,m^{-2}\,K^{-1}$ for the tree ensembles under LO*G*O. Its decomposition by withheld level then identified the cause, the assignment of a withheld level to its upper neighbour. RMSE in absolute units weighted the high-mass-flux records, which $TiO_2$/water lacks, so the RMSE values of the two nanofluids are not directly comparable. $R^2$, RMSE, MBE, a relative error, and the share of records within an error band, all from the same grouped out-of-fold predictions, describe the accuracy more completely

than any one of them. This conclusion rests on two nanofluids from one facility and does not establish which indicator is most informative for other datasets.

## 5.3 Model differences and the scatter of the records

The three-factor term of the variance decomposition gives a random scatter per degree of freedom of 0.50% of $h$ for $Al_2O_3$/water (36 degrees of freedom) and 0.20% for $TiO_2$/water (18). The rounding of $h$ contributes at most 0.09%. This scatter bounds the random error of a record from above, but it excludes systematic errors, so the total measurement uncertainty remains unknown and may be larger. Several anchored errors differ by less than this scatter, for example 0.56, 0.58, and 0.70% for GPR, MLP, and PLR under LO*G*O for $TiO_2$/water. Differences below about 1% therefore do not rank the models, and anchored errors below 1% do not establish an accuracy better than that of the measurements. At the scale of the direct errors, model structure matters more than flexibility when a complete level is withheld. This conclusion is specific to three inputs and to the smooth, multiplicative structure of the present records.

## 5.4 Physical interpretation of the enhancement

Correlations with effective properties predict ratios within 0.2% of unity, so the measured enhancement is not a bulk-property effect. The records contain no surface characterisation, wettability measurement, test sequence, or flow-regime information, and the mechanism cannot be identified. Particle deposition, which changes the roughness, wettability, and nucleation of the heated surface [28, 59], is consistent with the size of the effect and with its concentration at the lowest loading. An effect confined to nucleation is less consistent with the records. In superposition correlations, the nucleate share decreases with mass flux at fixed heat flux and quality, whereas the measured $E$ increases with $G$ at every $q$ and $\varphi$. The heat-flux exponent of $h_{bf}$, 0.04 to 0.19, also indicates that forced convection contributes a large part of the base-fluid coefficient.

If the enhancement reflects the evolution of the heated surface during the campaign, the anchored models learn a facility-specific surface state rather than a transferable property of the nanofluid. The calibration between the oxides would then apply only to a surface with the same deposition history. This possibility limits the transfer to other facilities more than the choice of model does. It can be tested only with surface measurements or with repeated runs after cleaning.

## 5.5 Limitations

Five limitations apply, listed from the most restrictive.

1. All records come from one annular test section at 1 bar, and no independent dataset with another geometry, pressure, or laboratory was available. The grouped folds are internal to one campaign, so the results establish accuracy for withheld operating levels of this campaign only. This is the most restrictive limitation.
2. The anchored formulation requires a measured base-fluid coefficient at every predicted condition; with estimated base-fluid values, its lowest errors were 4.14 to 8.52%.

3. The $TiO_2$/water records lack the 29 conditions with the highest coefficients, and the $Al_2O_3$/water records lack the highest heat flux. The accuracies therefore apply to the covered region, although the comparisons between the nanofluids were unchanged on their 80 common conditions.
4. The records give no measurement uncertainty, surface characterisation, or regime information. Errors below about 1% are at the scale of the record scatter, and the origin of the enhancement cannot be identified.
5. Neither interval method was calibrated at extreme levels, where the anchored coverage was 27.3 to 90.9% against a nominal 90%. The one-parameter calibration rests on two oxides from one facility.

# 6 Conclusions

Flow boiling HTCs of $Al_2O_3$/water and $TiO_2$/water nanofluids in a horizontal annulus were predicted with eight machine-learning and regression models, two target formulations, and six validation schemes. The predictions were assessed with $R^2$, RMSE, MBE, and MAPE. The main findings are as follows.

1. In the direct formulation, withholding a mass-flux level raised the MAPE of SVR from 2.74% (resubstitution) to 8.34%, and that of the tree ensembles to 37.4 to 43.3%.
2. Under random cross-validation, the direct models reached $R^2$ values of 0.952 to 0.996 with MBE values within ±0.20 kW m$^{-2}$ K$^{-1}$. Under LO*G*O, the tree ensembles fell to $R^2$ values of 0.105 to 0.549 with MBE values of +3.68 to +4.08 kW m$^{-2}$ K$^{-1}$, whereas the smooth models kept 0.891 to 0.962.
3. The tree-ensemble bias was systematic: they over-predicted the lowest and the three interior mass-flux levels by 15.8 to 66.1%, because each interior level lies closer to its upper neighbour in ln *G*.
4. For the five smooth models, extreme-level folds had 1.1 to 6.1 times the error of interior-level folds in all 60 comparisons.
5. Anchoring predicts the enhancement ratio for a measured base-fluid coefficient at the same heat flux and mass flux. Under LO*G*O, it raised $R^2$ to 0.990 to 0.9997 and reduced the lowest MAPE to 1.20% ($Al_2O_3$/water) and 0.56% ($TiO_2$/water). The anchor supplied 94.7 and 96.1% of the variance of the logarithmic target, and a concentration mean alone reached 4.65 and 2.60%.
6. With the base-fluid coefficient of the withheld level estimated instead of measured, the lowest anchored errors were 4.14 to 8.52%, against 4.99 to 6.10% for the direct models.
7. Anchored models trained on 4 of 20 boiling curves, with base-fluid values measured at all 25 or 27 conditions, outperformed direct models trained on 12 curves in 11 of 12 cases.
8. GPR intervals covered 54.7 to 87.4% of the records of withheld heat-flux and mass-flux levels against a nominal 90%. Anchored conformal intervals covered 98.2 to 100% at interior but 27.3 to 78.8% at extreme levels.

9. Monotonicity constraints removed all decreasing steps along heat flux and mass flux and reduced the direct PLR error for a withheld heat-flux level of $Al_2O_3$/water from 11.98 to 7.46%.
10. Correlations with effective properties predicted enhancement ratios of 0.998 to 1.002, against 1.231 to 2.418 measured.
11. With four records, a one-parameter calibration between the nanofluids reached errors of 1.88% for $TiO_2$/water and 3.78% for $Al_2O_3$/water, and its factor rose from 0.460 to 0.534 with concentration.
12. Differences below about 1% between anchored models were at the scale of the record scatter of 0.20 to 0.50%.

The results do not establish the accuracy of any model for other test sections, pressures, or nanoparticle materials, or for the nanofluid HTC without a base-fluid measurement at the same condition. The next steps are the evaluation of the anchored formulation on the complete design of the source experiments and on independent datasets with reported geometry, inlet conditions, and uncertainty. Additional particle materials with surface characterisation are needed to test whether the calibration factor relates to measurable particle or surface properties.

# Statements and Declarations


### Funding

No external funding was received for this research.


### Competing interests

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

### Ethics approval and consent

Not applicable. The study uses published experimental records and involves no human participants or animals.

### Consent to participate

Consent to Participate declaration: not applicable

### Consent to publish

Consent to Publish declaration: not applicable

### Data availability

The records analysed in this study were reported by Dadhich et al. [14]. The processed records, the out-of-fold predictions, and the result tables are available from the corresponding author on reasonable request.

**Code availability**

The analysis code and the scripts used to generate the figures and tables are available from the corresponding author on reasonable request.

**Author contributions**

Conceptualization and methodology: Adel Bouali, Hamza Ramdani, and Taqiy Eddine Boukelia. Software, data curation, and formal analysis: Adel Bouali. Validation and investigation: Adel Bouali and Khaoula Nedjem. Writing—original draft: Adel Bouali and Hamza Ramdani. Writing—review and editing: Hamza Ramdani, Khaoula Nedjem, Taqiy Eddine Boukelia, Aissa Laouissi, and Mourad Nouioua. Supervision: Hamza Ramdani and Taqiy Eddine Boukelia. All authors approved the final manuscript and its submission.

# References


[1] Adio SA, Atofarati EO, Muritala AO, et al (2025) Nanofluids flow boiling and convective heat transfer in microchannels: a systematic review and bibliometrics analysis. Journal of Thermal Analysis and Calorimetry 150(12):8879–8911. https://doi.org/10.1007/s10973-025-14265-x

[2] Bahiraei M, Heshmatian S, Moayedi H (2019) Artificial intelligence in the field of nanofluids: A review on applications and potential future directions. Powder Technology 353:276–301. https://doi.org/10.1016/j.powtec.2019.05.034

[3] Barber J, Brutin D, Tadrist L (2011) A review on boiling heat transfer enhancement with nanofluids. Nanoscale Research Letters 6(1):280. https://doi.org/10.1186/1556-276X-6-280

[4] Bard A, Qiu Y, Kharangate CR, et al (2022) Consolidated modeling and prediction of heat transfer coefficients for saturated flow boiling in mini/micro-channels using machine learning methods. Applied Thermal Engineering 210:118305. https://doi.org/10.1016/j.applthermaleng.2022.118305

[5] Bouali A, Mohammedi B, Hanini S (2022) A smart model for the prediction of heat transfer coefficient during flow boiling of nanofluids in horizontal tube. Nano Hybrids and Composites 36:89–102. https://doi.org/10.4028/p-9ge01g

[6] Breiman L (2001) Random forests. Machine Learning 45(1):5–32. https://doi.org/10.1023/A:1010933404324

[7] Brinkman HC (1952) The viscosity of concentrated suspensions and solutions. The Journal of Chemical Physics 20(4):571. https://doi.org/10.1063/1.1700493

[8] Buongiorno J, Venerus DC, Prabhat N, et al (2009) A benchmark study on the thermal conductivity of nanofluids. Journal of Applied Physics 106(9):094312. https://doi.org/10.1063/1.3245330

[9] Chai M, Hong Z, Li N, et al (2025) Numerical modeling and analysis of nanofluids subcooled boiling in a horizontal circular tube. Numerical Heat Transfer, Part B: Fundamentals 86(11):3656–3677. https://doi.org/10.1080/10407790.2024.2364775

[10] Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding RMSE in the literature. Geoscientific Model Development 7(3):1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

[11] Chen JC (1966) Correlation for boiling heat transfer to saturated fluids in convective flow. Industrial & Engineering Chemistry Process Design and Development 5(3):322–329. https://doi.org/10.1021/i260019a023

[12] Choi SUS (1995) Enhancing thermal conductivity of fluids with nanoparticles. In: Siginer DA, Wang HP (eds) Developments and Applications of Non-Newtonian Flows. American Society of Mechanical Engineers, New York, FED-Vol. 231/MD-Vol. 66, pp 99–105, https://doi.org/10.1115/IMECE1995-0926

[13] Cooper MG (1984) Heat flow rates in saturated nucleate pool boiling – a wide-ranging examination using reduced properties. In: Advances in Heat Transfer, vol 16. Elsevier, p 157–239, https://doi.org/10.1016/S0065-2717(08)70205-3

[14] Dadhich M, Prajapati OS, Rohatgi N (2020) Flow boiling heat transfer analysis of $Al_2O_3$ and $TiO_2$ nanofluids in horizontal tube using artificial neural network (ANN). Journal of Thermal Analysis and Calorimetry 139(5):3197–3217. https://doi.org/10.1007/s10973-019-08674-y

[15] Dadhich M, Prajapati OS, Sharma V (2021a) Investigation of boiling heat transfer of titania nanofluid flowing through horizontal tube and optimization of results utilizing the desirability function approach. Powder Technology 378:104–123. https://doi.org/10.1016/j.powtec.2020.09.077

[16] Dadhich M, Prajapati OS, Sharma V (2021b) A systematic review on the heat transfer investigation of the flow boiling process. Chemical Engineering and Processing – Process Intensification 165:108425. https://doi.org/10.1016/j.cep.2021.108425

[17] Fang X, Wang R, Chen W, et al (2015) A review of flow boiling heat transfer of nanofluids. Applied Thermal Engineering 91:1003–1017. https://doi.org/10.1016/j.applthermaleng.2015.08.100

[18] Friedman JH (2001) Greedy function approximation: A gradient boosting machine. The Annals of Statistics 29(5):1189–1232. https://doi.org/10.1214/aos/1013203451

[19] Furlong A, Zhao X, Salko RK, et al (2025) Physics-based hybrid machine learning for critical heat flux prediction with uncertainty quantification. Applied Thermal Engineering 279:127447. https://doi.org/10.1016/j.applthermaleng.2025.127447

[20] Gungor KE, Winterton RHS (1986) A general correlation for flow boiling in tubes and annuli. International Journal of Heat and Mass Transfer 29(3):351–358. https://doi.org/10.1016/0017-9310(86)90205-X

[21] Hou J, Wu J, Zhao X, et al (2024) Microfluidic flow synthesis of $Al_2O_3$ nanofluids for efficient phase-change boiling heat transfer enhancement of electronic devices. Case Studies in Thermal Engineering 58:104458. https://doi.org/10.1016/j.csite.2024.104458

[22] Hughes MT, Kini G, Garimella S (2021) Status, challenges, and potential for machine learning in understanding and applying heat transfer phenomena. Journal of Heat Transfer 143(12):120802. https://doi.org/10.1115/1.4052510

[23] Kamel MS, Lezsovits F, Hussein AK (2019) Experimental studies of flow boiling heat transfer by using nanofluids: A critical recent review. Journal of Thermal Analysis and Calorimetry 138(6):4019–4043. https://doi.org/10.1007/s10973-019-08333-2

[24] Kandlikar SG (1990) A general correlation for saturated two-phase flow boiling heat transfer inside horizontal and vertical tubes. Journal of Heat Transfer 112(1):219–228. https://doi.org/10.1115/1.2910348

[25] Kapoor S, Narayanan A (2023) Leakage and the reproducibility crisis in machine-learning-based science. Patterns 4(9):100804. https://doi.org/10.1016/j.patter.2023.100804

[26] Karpatne A, Atluri G, Faghmous JH, et al (2017) Theory-guided data science: A new paradigm for scientific discovery from data. IEEE Transactions on Knowledge and Data Engineering 29(10):2318–2331. https://doi.org/10.1109/TKDE.2017.2720168

[27] Kim J, Lee S, Kim SM, et al (2026) A symbolic regression-based correlation for the heat transfer coefficient of saturated flow boiling in mini/micro-channels. Applied Thermal Engineering 307:133210. https://doi.org/10.1016/j.applthermaleng.2026.133210

[28] Kim SJ, Bang IC, Buongiorno J, et al (2007) Surface wettability change during pool boiling of nanofluids and its effect on critical heat flux. International Journal of Heat and Mass Transfer 50(19–20):4105–4116. https://doi.org/10.1016/j.ijheatmasstransfer.2007.02.002

[29] Kim SJ, McKrell T, Buongiorno J, et al (2010) Subcooled flow boiling heat transfer of dilute alumina, zinc oxide, and diamond nanofluids at atmospheric pressure. Nuclear Engineering and Design 240(5):1186–1194. https://doi.org/10.1016/j.nucengdes.2010.01.020

[30] Legates DR, McCabe GJ (1999) Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. Water Resources Research 35(1):233–241. https://doi.org/10.1029/1998WR900018

[31] Lei J, G'Sell M, Rinaldo A, et al (2018) Distribution-free predictive inference for regression. Journal of the American Statistical Association 113(523):1094–1111. https://doi.org/10.1080/01621459.2017.1307116

[32] Li G, Tian A, Chen Y, et al (2026) Experimental investigation and data-driven modeling of nanofluid pool boiling under rotational hypergravity. Heat and Mass Transfer 62(5):51. https://doi.org/10.1007/s00231-026-03680-x

[33] Liu Z, Winterton RHS (1991) A general correlation for saturated and subcooled flow boiling in tubes and annuli, based on a nucleate pool boiling equation. International Journal of Heat and Mass Transfer 34(11):2759–2766. https://doi.org/10.1016/0017-9310(91)90234-6

[34] Loyola-Fuentes J, Nazemzadeh N, Diaz-Bejarano E, et al (2024) A framework for data regression of heat transfer data using machine learning. Applied Thermal Engineering 248:123043. https://doi.org/10.1016/j.applthermaleng.2024.123043

[35] Ma T, Guo Z, Lin M, et al (2021) Recent trends on nanofluid heat transfer machine learning research applied to renewable energy. Renewable and Sustainable Energy Reviews 138:110494. https://doi.org/10.1016/j.rser.2020.110494

[36] Maleki A, Haghighi A, Mahariq I (2021) Machine learning-based approaches for modeling thermophysical properties of hybrid nanofluids: A comprehensive review. Journal of Molecular Liquids 322:114843. https://doi.org/10.1016/j.molliq.2020.114843

[37] Maxwell JC (1873) A Treatise on Electricity and Magnetism. Clarendon Press, Oxford, 2 vols.

[38] Mohebali MM, Baniamerian Z (2024) Effect of nanofluid sedimentation on heat transfer and critical heat flux in boiling flows. Journal of Thermal Analysis and Calorimetry 149(15):8225–8244. https://doi.org/10.1007/s10973-024-13303-4

[39] Mondal A, Sharma SL (2026) Probabilistic machine learning framework for predicting critical heat flux in uniformly heated vertical tubes with uncertainty quantification. Nuclear Technology 212(11):2912–2929. https://doi.org/10.1080/00295450.2026.2629142

[40] Mukherjee S, Jana S, Mishra PC, et al (2021) Experimental investigation on thermo-physical properties and subcooled flow boiling performance of $Al_2O_3$/water nanofluids in a horizontal tube. International Journal of Thermal Sciences 159:106581. https://doi.org/10.1016/j.ijthermalsci.2020.106581

[41] Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12:2825–2830. https://jmlr.org/papers/v12/pedregosa11a.html

[42] Phan TH, Wang H, Kharangate CR (2026) Correlation-informed neural networks for predicting flow boiling heat transfer coefficient in mini/micro-channels. International Journal of Heat and Mass Transfer 270:129254. https://doi.org/10.1016/j.ijheatmasstransfer.2026.129254

[43] Qiu Y, Garg D, Zhou L, et al (2020) An artificial neural network model to predict mini/micro-channels saturated flow boiling heat transfer coefficient based on universal consolidated data. International Journal of Heat and Mass Transfer 149:119211. https://doi.org/10.1016/j.ijheatmasstransfer.2019.119211

[44] Rasmussen CE, Williams CKI (2006) Gaussian Processes for Machine Learning. MIT Press, Cambridge, MA, https://doi.org/10.7551/mitpress/3206.001.0001

[45] Riyadi TWB, Herawan SG, Tirta A, et al (2024) Nanofluid heat transfer and machine learning: Insightful review of machine learning for nanofluid heat transfer enhancement in porous media and heat exchangers as sustainable and renewable energy solutions. Results in Engineering 24:103002. https://doi.org/10.1016/j.rineng.2024.103002

[46] Roberts DR, Bahn V, Ciuti S, et al (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40(8):913–929. https://doi.org/10.1111/ecog.02881

[47] Saraswat SP (2026) Campaign-disjoint validation of machine-learned critical heat flux models: leakage quantification and physics-anchored residual transfer. Energy and AI 26:100889. https://doi.org/10.1016/j.egyai.2026.100889

[48] Serrao BP, Kim KM, Duarte JP (2023) Analysis of the effects of different nanofluids on critical heat flux using artificial intelligence. Energies 16(12):4762. https://doi.org/10.3390/en16124762

[49] Shah MM (2017) Unified correlation for heat transfer during boiling in plain mini/micro and conventional channels. International Journal of Refrigeration 74:606–626. https://doi.org/10.1016/j.ijrefrig.2016.11.023

[50] Shekhar, Sambyo K, Prajapati OS, et al (2026) Development of a XGBJSO hybrid machine learning model for predicting heat transfer coefficient of water and $Al_2O_3$ nanofluid in flow boiling. Nanotechnology Reviews 15(1):20250287. https://doi.org/10.1515/ntrev-2025-0287

[51] Shourehdeli SA, Gholipour H (2026) Harnessing multi-source data to develop universal machine learning models for saturated flow boiling heat transfer in mini/macro channels. International Communications in Heat and Mass Transfer 170:110010. https://doi.org/10.1016/j.icheatmasstransfer.2025.110010

[52] Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. Statistics and Computing 14(3):199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

[53] Soleimani B, Ziabasharhagh M (2026) Experimental investigation of surface material impact on nanofluid subcooled flow boiling. Thermal Engineering 73(5):351–361. https://doi.org/10.1134/s0040601524600949

[54] Suh Y, Chandramowlishwaran A, Won Y (2024) Recent progress of artificial intelligence for liquid-vapor phase change heat transfer. npj Computational Materials 10(1):65. https://doi.org/10.1038/s41524-024-01223-8

[55] Upadhyay A, Hazra SK, Assam A, et al (2024) Review of the current status and the potential of machine learning tools in boiling heat transfer. Numerical Heat Transfer, Part B: Fundamentals 85(11):1489–1532. https://doi.org/10.1080/10407790.2023.2266770

[56] Varma S, Simon R (2006) Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics 7:91. https://doi.org/10.1186/1471-2105-7-91

[57] Wang C, Su G, Akinsulire O, et al (2024) Investigation of critical heat flux enhancement on nanoengineered surfaces in pressurized subcooled flow boiling using infrared thermometry. Heat Transfer Engineering 45(4–5):417–432. https://doi.org/10.1080/01457632.2023.2191441

[58] Wang Y, Deng K, Wu J, et al (2020) A mechanism of heat transfer enhancement or deterioration of nanofluid flow boiling. International Journal of Heat and Mass Transfer 158:119985. https://doi.org/10.1016/j.ijheatmasstransfer.2020.119985

[59] Wen D (2012) Influence of nanoparticles on boiling heat transfer. Applied Thermal Engineering 41:2–9. https://doi.org/10.1016/j.applthermaleng.2011.08.035

[60] Willard J, Jia X, Xu S, et al (2023) Integrating scientific knowledge with machine learning for engineering and environmental systems. ACM Computing Surveys 55(4):1–37. https://doi.org/10.1145/3514228

[61] Willmott CJ (1981) On the validation of models. Physical Geography 2(2):184–194. https://doi.org/10.1080/02723646.1981.10642213

[62] Willmott CJ, Matsuura K (2005) Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research 30(1):79–82. https://doi.org/10.3354/cr030079

[63] Zhao X, Shirvan K, Salko RK, et al (2020) On the prediction of critical heat flux using a physics-informed machine learning-aided framework. Applied Thermal Engineering 164:114540. https://doi.org/10.1016/j.applthermaleng.2019.114540