Proceedings of the 17th International Heat Transfer Conference, IHTC-17
14 - 18 August 2023, Cape Town, South Africa

IHTC-17 | ID: 673

A MACHINE LEARNING APPROACH TO THE PREDICTION OF HEAT TRANSFER COEFFICIENTS IN MICROCHANNELS

**Tullio Traverso [1,2], Francesco Coletti [3,4], Luca Magri [2,1,5], Tassos G. Karayiannis [6],
and Omar K. Matar [7,1]***

[1]The Alan Turing Institute, British Library, 96 Euston Road, London, NW1 2DB, UK

[2]Department of Aeronautics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

[3]Department of Chemical Engineering, Brunel University London, Kingston Lane, Uxbridge, Middlesex UB8 3PH, UK

[4] Hexxcell Ltd., Foundry Building, 77 Fulham Palace Rd., London, W6 8AF, UK

[5] Department of Aerospace Engineering, Technion, Haifa, Israel (visiting)

[6] Department of Mechanical and Aerospace Engineering, Brunel University London, Kingston Lane, Uxbridge, Middlesex UB8 3PH, UK

[7] Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

ABSTRACT

The accurate prediction of the two-phase heat transfer coefficient (HTC) as a function of working fluids, channel geometries and process conditions is key to the optimal design and operation of compact heat exchangers. Advances in artificial intelligence research have recently boosted the application of machine learning (ML) algorithms to obtain data-driven surrogate models for the HTC. For most supervised learning algorithms, the task is that of a nonlinear regression problem. Despite the fact that these models have been proven capable of outperforming traditional empirical correlations, they have key limitations such as overfitting the data, the lack of uncertainty estimation, and interpretability of the results. To address these limitations, in this paper, we use a multi-output Gaussian process regression (GPR) to estimate the HTC in microchannels as a function of the mass flow rate, heat flux, system pressure and channel diameter and length. The model is trained using the Brunel Two-Phase Flow database of high-fidelity experimental data. The advantages of GPR are data efficiency, the small number of hyperparameters to be trained (typically of the same order of the number of input dimensions), and the automatic trade-off between data fit and model complexity guaranteed by the maximization of the marginal likelihood (Bayesian approach). Our paper proposes research directions to improve the performance of the GPR-based model in extrapolation.

KEYWORDS: Gaussian process regressions, data-driven modelling, flow boiling, micro-tubes.

1. INTRODUCTION

The dissipation of high heat fluxes from small surface areas is one of the main challenges in the thermal design of electronic devices. Dissipation can be achieved effectively by using two-phase flow boiling in small-to-micro diameter channels, a solution that has the advantage of dissipating high heat fluxes at nearly uniform substrate surface temperatures [1]. The accurate prediction of the heat transfer coefficient (HTC) is key to the optimal design and operation of such compact heat exchangers. However, due to the nonlinear behaviour in phase-change systems such as flow boiling, the accurate prediction of the HTC is challenging

*Corresponding Author: o.matar@imperial.ac.uk

[2]. To this end, physics-based models and empirical correlations have been developed as design tools for a large number of fluid flows and heat transfer configurations. However, the developed micro-scale correlations may not be extrapolated with confidence outside their applicability ranges, while few physics-based models have been developed. This is the case especially at the micro- and nano-scales, where the underlying physical phenomena are complex and remain relatively poorly understood [3].

A more flexible approach is to use data-driven numerical methods to perform regressions on the available data, and infer the HTC under different working conditions or for new heat exchanger designs. Recent advances in the field of artificial intelligence are making this option suitable for solving heat transfer problems [4, 5]. Specifically, machine learning algorithms have the potential for extracting trends from intricate datasets, generating robust optimization frameworks, and providing accurate prediction, often outperforming traditional methods [6, 7]. Despite their remarkable performance, commonly employed supervised ML models, such as support vector regression, random forest, and artificial neural networks (ANN), present significant limitations. In fact, these models do not provide uncertainty quantification, may overfit data due to the large number of parameters, and might require tedious fine tuning of their hyperparameters [4]. Their predictions are not readily interpretable, providing limited insight into understanding the data.

In this work, we train Gaussian process regression models to infer the two-phase HTC in micro-tubes as a function of five design parameters, and two working fluids, using part of the *Brunel Two-Phase Flow database*. We demonstrate how this family of ML models present key features, which make them a promising method to perform this task. First, GPR are non-parametric models that are designed not to overfit the data, dramatically reducing the degree of arbitrariness introduced at the model design stage (here, reduced to choosing the appropriate kernel function). Moreover, the kernel parameters are often easily interpretable, providing a powerful tool for understanding the trends in the data, as discussed in section 2. Second, GPR models are data-efficient becasue they require relatively small datasets to achieve a good predictive performance. Finally, GPR models are based on a Bayesian framework, which naturally provides uncertainty estimation. Conversely, if the data uncertainty is known, the uncertainty information can be consistently incorporated into the model prediction. We end the paper by providing directions to enable the extrapolation to unseen scenarios.

2. GAUSSIAN PROCESS REGRESSION

Gaussian process models, or simply Gaussian processes (GPs), are popular because they can be used in a regression framework to approximate nonlinear functions with probabilistic estimates of their uncertainties [8]. This section provides an overview of Gaussian process regression (GPR).

A Gaussian process is a collection of random variables $\{f(\mathbf{x})|\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^n\}$ for which, given any finite set of N inputs $\mathbf{X} = \{\mathbf{x}_i \in \mathcal{X} | i = 1,\dots,N\}$, the collection $f(\mathbf{X}) \equiv \{f(\mathbf{x}_i)|i = 1,\dots,N\}$ has a joint multivariate Gaussian distribution. This joint distribution is denoted as

$$f(\mathbf{X}) \sim GP(m(\mathbf{X}), k(\mathbf{X},\mathbf{X})) \tag{1}$$

where $m(\mathbf{X})$ is the N-dimensional vector with elements $m(\mathbf{x}_i) = \mathbb{E}\left[f(\mathbf{x}_i)\right]$, i.e., the mean of f at $\mathbf{x}_i$, and $k(\mathbf{X},\mathbf{X})$ is the N-by-N covariance matrix defined as

$$k_{ij} = k(\mathbf{x}_i,\mathbf{x}_j) = \mathbb{E}\left[\left(f(\mathbf{x}_i) - m(\mathbf{x}_i)\right)\left(f(\mathbf{x}_j) - m(\mathbf{x}_j)\right)\right] \tag{2}$$

The GP is defined by the knowledge of its mean and covariance.

Before regression is performed (i.e., before data is considered), the functions m and k reflect prior knowledge about the unknown function f. In this work, the form of the kernel $k(\mathbf{x},\mathbf{x}')$ is chosen a priori and then tuned by adjusting a finite set of continuously varying hyperparameters $\boldsymbol{\theta}$. This form of training is known as *type II maximum likelihood approximation* or ML-II. The prior mean function is assumed to be zero unless otherwise stated, that is $m(\mathbf{x}) = 0$ for all $\mathbf{x}$.

For a GPR model, we consider a Gaussian process f and a data set D. The data set comprises of a vector of N noisy observations $\mathbf{y} = [y_1, \ldots, y_N]$ at the input locations $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$. These observations are obtained from the true value $f(\mathbf{x}_i)$ plus additive independent and identically distributed Gaussian noise ϵ with variance σ_n^2, that is

$$y_i = f(\mathbf{x}_i) + \epsilon \tag{3}$$

$$\epsilon \sim \mathcal{N}(0, \sigma_n^2) \tag{4}$$

The model is *trained* by maximizing the marginal likelihood of the models hyperparameters ($\hat{\boldsymbol{\theta}}$) given the data, $p(\mathbf{y}|\mathbf{X}, \hat{\boldsymbol{\theta}})$. The models hyperparameters define the kernel, and the noise variance, i.e., $\hat{\boldsymbol{\theta}} = \{\boldsymbol{\theta}, \sigma_n^2\}$. The resulting optimization problem is

$$\underset{\hat{\boldsymbol{\theta}}}{\arg\max} \, p(\mathbf{y}|\mathbf{X}, \hat{\boldsymbol{\theta}}) \tag{5}$$

which is solved by using a gradient based approach in the GP framework GPy [9]. Finally, by conditioning the joint distribution on the training data and the testing inputs $\mathbf{x}_*$, we derive the predictive distribution

$$\mathbf{f}_*|\mathbf{x}, \mathbf{x}_*, \mathbf{y} \sim \mathcal{N}(\bar{\mathbf{f}}_* | cov(\mathbf{f}_*)) \tag{6}$$

where $\bar{\mathbf{f}}_*$ and $cov(\mathbf{f}_*)$ are given by [10]

$$\bar{\mathbf{f}}_* = k(\mathbf{x}_*, \mathbf{x}) \left[k(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I} \right]^{-1} \mathbf{y} \tag{7}$$

$$cov(\mathbf{f}_*) = k(\mathbf{x}_*, \mathbf{x}_*) - k(\mathbf{x}_*, \mathbf{x}) \left[k(\mathbf{x}, \mathbf{x}) + \sigma_n^2 \mathbf{I} \right]^{-1} k(\mathbf{x}, \mathbf{x}_*) \tag{8}$$

and $\mathbf{I}$ is the N-by-N identity matrix.

In this work, we chose the squared exponential kernel, which is defined as

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left(-\frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^{\mathrm{T}} M (\mathbf{x}_i - \mathbf{x}_j) \right) \tag{9}$$

with hyperparameters $\boldsymbol{\theta} = \{M, \sigma_f^2\}$, where $M = \mathrm{diag}(\boldsymbol{\lambda})^{-2}$, and $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_n]$. The hyperparameters of the covariance function in equation (9) have a clear interpretation, which is of great importance when trying to understand the data. Specifically, the λ_i hyperparameters play the role of characteristic length-scales indicating the degree to which a change in the i-th design parameter is necessary in order to realise a significant change in the output. Such a kernel thus implements the automatic relevance determination (ARD) [11] since the inverse of the length-scale determines the relevance of an input.

3. DATA SET

Our data set consists of 15000 high-fidelity data points that are part of the *Brunel Two-Phase Flow database*. They relate the pipe length (L), its diameter (d), the mass flow rate (G), the inlet pressure (P_{in}), and the supplied heat flux (q) with the local two-phase HTC (α_L). The pipes used are made of stainless-steel with circular cross sections. Two working fluids, i.e., R134a and R245fa, are considered. Data with R134a are obtained in [3] and [12], details regarding the test rig can be found in [13], while experiments with refrigerant R245fe are presented in [14] and [15].

The experimental facilities consist of the closed-loop refrigerant main circuit, data acquisition and control, and cooling and heating systems. Direct electric heating was applied to the test section: fifteen K-type thermocouples were soldered to the outside of the tube at equal distances to provide the wall temperatures, and pressure transducers were used to measure inlet and outlet temperatures and pressures. The local heat transfer coefficient at each thermocouple point was estimated by

$$\alpha_L = \frac{q}{T_w - T_s} \tag{10}$$

where T_w is the local inner wall temperature, T_s is the local saturation temperature, and q is the inner wall heat flux to the fluid. The temperature T_s was deduced from the fluid pressure which, in turn, was determined based on the assumption of a linear pressure drop through the test section at the start of saturated boiling; T_w was calculated by solving the one-dimensional heat conduction equation with volumetric heat generation in cylindrical coordinates. This makes T_w a function of the outside surface temperature recorded by the thermocouples, the heat flux, and the tube wall thermal resistance. Further details on the calculation procedure are available in [16].

Before training, the data points are pre-processed by averaging the value of the local two-phase HTC along the pipe length. This yields approximately 1000 points associating the input design parameters (L, d, G, P_{in}, q) with the average HTC, denoted by α. Finally, we note that all the training performed in this study takes only a few seconds on a standard laptop, therefore, the computational cost is negligible. Unlike more traditional ML tasks, where scalability with the size of the data set is a primary concern [17], in engineering applications the main challenge often consist in making the most out of a relatively small number of reliable data points.

4. RESULTS

This study has two main thrusts. The first is to characterize the prediction capabilities of the model using data for both refrigerants, and study its performance as a function of the ratio between training and test points. The second is to characterize the model's performance when extrapolating. This means that the test points are chosen at the edges of the input space, as opposed to being randomly sampled. The performance of the models are evaluated by the mean absolute error

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \frac{|\alpha_{p,i} - \alpha_{e,i}|}{\alpha_{e,i}} \tag{11}$$

where $\alpha_{e,i}$ is the average HTC obtained from the i-th data point. The corresponding mean prediction by the GPR model is denoted by $\alpha_{p,i}$. The specifications *tot* and *test* after the MAE in figures 1 and 2 refers to whether equation (11) was evaluated on the full data set or on test points only, respectively.

4.1 Data efficiency of GPRs models To study the performance of the model as a function of the ratio between the size of the training and test sets, the data points related to the two refrigerants are merged. No labels related to the working fluid are provided, i.e., the GPR model is unaware of which data point is obtained with which fluid, and the test points are selected randomly. Figure 1 shows the parity plots relating the value of α measured experimentally with the value predicted by the mean of the GPR model. The plots shown in figure 1(a)-(c) refer to three different models trained by setting the percentage of data points used for testing to 70%, 50%, and 30%, respectively.

As shown in figure 1(a), the use of a test set of 70% randomly sampled data points, corresponding to just 30% of the data used for training, already provides a reasonable prediction, corresponding to $MAE = 23.1\%$ on the test data and 16.6% on the full data set. The performance significantly increases with the size of the training set: for the common choice of 30% of the data used for testing, the MAE on the test set is reduced to 11.5%, and to 3.4% on the full data set, with nearly all the predictions within the dashed ±30% error bounds (see figure 1(c). The results shown in figure 1, in particular those with the 70% test set, demonstrate that our GPR models are data efficient.

4.2 Automatic trade-off between data-fit and model complexity The parity plot in figure 1(a) shows an excellent data fit of the training set, as all the blue squares lie on the parity line. The good fitting on the test set (red circles) of the model trained with only 30% of the data gives a strong indication that the model does not overfit the data.

We remark that, unlike parametric models such as deep artificial neural networks, GPR models provide an intrinsic trade-off between data fit and model complexity [10]. In other words, the optimization problem in equation (5) is such that the optimal solution favours the simpler "explanation" of the data that offers the best fit. This avoids the problem of having to adjust the many hyperparameters of the ML algorithm to avoid overfitting (e.g., number and width of layers).
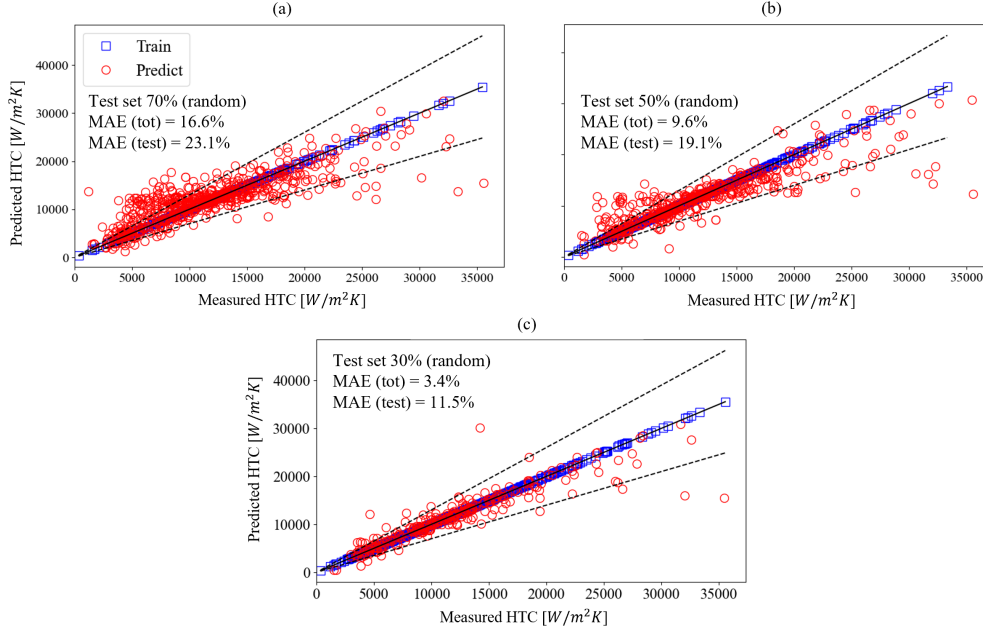
Fig. 1 Randomly selected test points chosen from the datasets associated with refrigerants R134a and R245fa: parity plots showing a comparison of the experimental results and the model predictions for the cases with 70% (a), 50% (b), and 30% (c) of test points.

4.3 Extrapolation performance We further test the extrapolation abilities of the model by deliberately selecting test points at the edges of the experimental design space, that is, we ask the model to extrapolate on all the test points. To avoid extrapolating also on different refrigerants, we restrict the training set to the data points relative to R245fa. Then, we exclude from the training set all the data points available for the smaller pipe diameter, $d = 1.1$ mm, resulting in 58% points used for testing. Figure 2(a) shows the parity plot associated with this case, while figure 2(b)-(f) shows the percentage error of the model prediction relative to the measured value, $err = 100|\alpha_{p,i} - \alpha_{e,i}|/\alpha_{e,i}$, as a function of the design parameters. As shown in figure 2(c)-(f), by eliminating smaller diameters we extrapolate also along other design parameters, i.e., the mass flow rate (G), pipe length (L), inlet pressure (P_{in}), and the heat flux (q).

As figure 2 (a) shows, the MAE is nearly 66.5% on the test set, and 38.7% on the full data set. This result is unsurprising considering the large amount of test data compared to the total, and that the model is extrapolating along four out of five input directions. Despite this, the model still favours the simplest answer, namely it predicts nearly the same value of α for all the test points, providing a prediction within the dashed ±30% error lines for approximately half of the test points. As discussed in the next section, we expect to improve the extrapolation performance by incorporating part of our physical knowledge into the model prior mean.
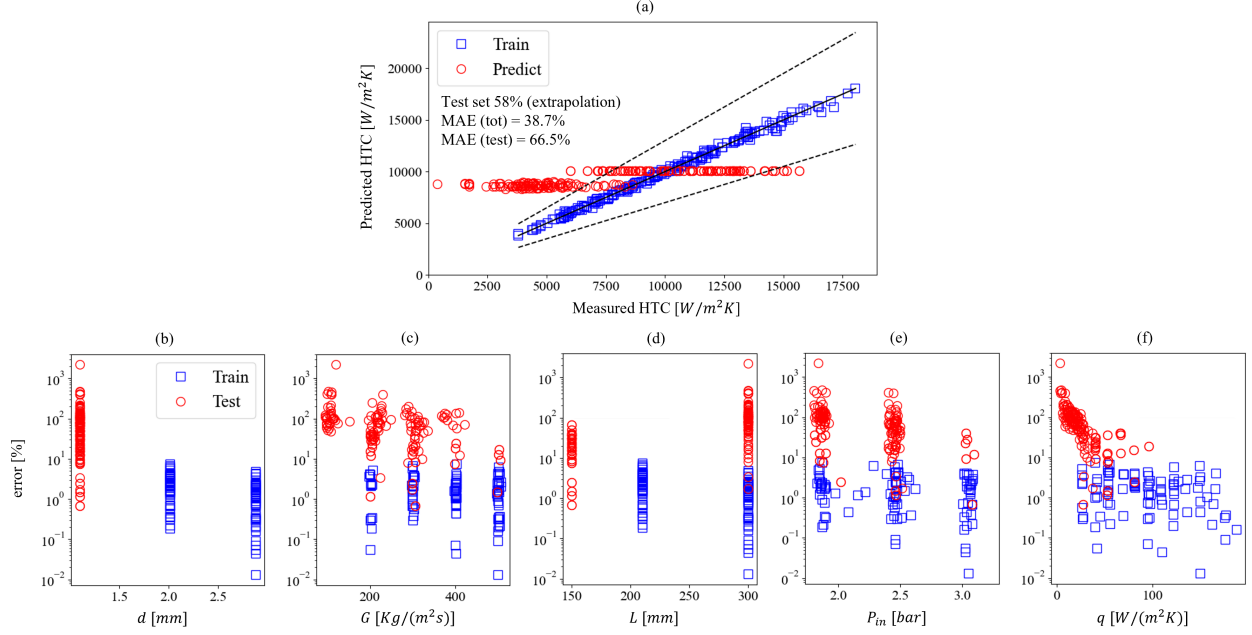
Fig. 2 Using the GPR model for extrapolation using test points associated with the refrigerant R245fa: (a) parity plot showing a comparison between the experimental results and the model predictions based on 58% of the dataset used for testing; (b)-(f) percentage error (defined in the main text) of the model prediction relative to the measured value as a function of the five design parameters considered, d, G, L, P_{in}, and q, respectively.

5. CONCLUSIONS AND FUTURE WORK

In this work, we demonstrate the potential of Gaussian Process Regression (GPR) models in the prediction of the two-phase heat transfer coefficients in micro-pipes. We start by investigating how the performance of GPR models changes with the size of the training set. Our models are found to perform well even when small fractions of the data set are used for training. Being data-efficient makes this approach particularly suitable for engineering applications due to the large cost of obtaining high-fidelity data. Regardless of the data used for training, the model does not overfit. This automatic trade-off between the model's capacity (i.e., data-fit) and complexity is a key feature to facilitate the deployment of GPR models. Practically, this allows the end user to employ the same algorithm to train different models on different data set, thus simplifying the job of carefully hand-tuning the model's hyperparameters when either the task or the data change, a process for which clear and general guidelines are case-dependent.

Finally, we show how the model's performance significantly deteriorates when it is tested on data at the edges of the design space, i.e., when we try to extrapolate, an issue not unique to GPR models. Still, we argue that the framework of GPR has the potential to handle the task of extrapolating more successfully. While we leave this for future work, we observe how a first improvement could be obtained by using existing correlations for the HTC as prior mean of the GPR. In the current implementation the prior mean of the GPR is assumed to be zero, meaning that we are not exploiting our prior knowledge about the system. Other current developments include the use of multi-output (or vector-valued) GPR models to extrapolate to different fluids or other categorical (i.e., not continuous) design variables.

ACKNOWLEDGMENT

NOMENCLATURE

$\dot{\alpha}_L$	local HTC	[W/(m^{2}K)]	n	dimensions of GPR's input space	
α	HTC averaged along pipe	[W/(m^{2}K)]			
α_p	predicted HTC	[W/(m^{2}K)]	N	number of observations	
α_e	experimental HTC	[W/(m^{2}K)]	$\mathcal{N}$	Gaussian distribution	
d	pipe diameter	[mm]	P_{in}	inlet pressure	[bar]
ϵ	output noise	[W/(m^{2}K)]	q	heat flux	[W/(m^{2}K)]
err	relative prediction error	[-]	$\mathbb{R}$	real numbers	
f	surrogate function	[W/(m^{2}K)]	σ_f	kernel's variance	
$\mathbf{f}_*$	GPR posterior prediction	[W/(m^{2}K)]	σ_n	noise's variance	[W/(m^{2}K)]
$\bar{\mathbf{f}}_*$	GPR mean prediction	[W/(m^{2}K)]	θ	kernel's hyperparameters	
G	mass flux	[kg/(m^{2}s)]	$\hat{\theta}$	GPR's hyperparameters	
GP	Gaussian process		T_s	saturation temperature	[K]
$\mathbf{I}$	N-by-N identity matrix		T_w	wall temperature	[K]
k	covariance function		$\mathcal{X}$	GPR's input space	
λ	characteristic length-scales		$\mathbf{x}_i$	i-th input location	
			$\mathbf{x}_*$	test input location	
L	pipe length	[mm]	$\mathbf{X}$	All input locations	
m	mean of GPR	[W/(m^{2}K)]	$\mathbf{y}_i$	i-th output	[W/(m^{2}K)]
MAE	mean absolute error				

REFERENCES

[1] Karayiannis, T. G. and Mahmoud, M. M., "Flow boiling in microchannels: Fundamentals and applications", *Applied Thermal Engineering*, 115, pp. 1372–1397 (2017).

[2] Ribatski, G., Zhang, W., Consolini, L., Xu, J. and Thome, J. R., "On the Prediction of Heat Transfer in Micro-Scale Flow Boiling", *Heat Transfer Engineering*, 28(10), pp. 842–851, doi:10.1080/01457630701378267, Available at `https://doi.org/10.1080/01457630701378267` (2007).

[3] Mahmoud, M. M. and Karayiannis, T. G., "Heat transfer correlation for flow boiling in small to micro tubes", *International Journal of Heat and Mass Transfer*, 66, pp. 553–574 (2013).

[4] Hughes, M. T., Kini, G. and Garimella, S., "Status, Challenges, and Potential for Machine Learning in Understanding and Applying Heat Transfer Phenomena", *Journal of Heat Transfer*, 143(12), doi:10.1115/1.4052510, Available at `https://doi.org/10.1115/1.4052510`, 120802 (2021).

[5] Loyola-Fuentes, J., Pietrasanta, L., Marengo, M. and Coletti, F., "Machine Learning Algorithms for Flow Pattern Classification in Pulsating Heat Pipes", *Energies*, 15(6), p. 1970 (2022).

[6] Hughes, M. T., Fronk, B. M. and Garimella, S., "Universal condensation heat transfer and pressure drop model and the role of machine learning techniques to improve predictive capabilities", *International Journal of Heat and Mass Transfer*, 179, p. 121712 (2021).

[7] Zhou, L., Garg, D., Qiu, Y., Kim, S. M., Mudawar, I. and Kharangate, C. R., "Machine learning algorithms to predict flow condensation heat transfer coefficient in mini/micro-channel utilizing universal data", *International Journal of Heat and Mass Transfer*, 162, p. 120351 (2020).

[8] Swiler, L. P., Gulian, M., Frankel, A. L., Safta, C. and Jakeman, J. D., "A survey of constrained Gaussian process regression: Approaches and implementation challenges", *Journal of Machine Learning for Modeling and Computing*, 1(2) (2020).

[9] GPy, "GPy: A Gaussian process framework in python", `http://github.com/SheffieldML/GPy` (since 2012).

[10] Williams, C. K. and Rasmussen, C. E., *Gaussian processes for machine learning*, vol. 2, MIT press Cambridge, MA (2006).

[11] Neal, R. M., *Bayesian Learning for Neural Networks*, Ph.D. thesis, Dept. of Computer Science, University of Toronto (1994).

[12] Shiferaw, D., Karayiannis, T. G. and Kenning, D. B. R., "Flow boiling in a 1.1 mm tube with R134a: Experimental results and comparison with model", *International Journal of Thermal Sciences*, 48(2), pp. 331–341 (2009).

[13] Huo, Y., Tian, S. and Karayiannis, T. G., "R134a flow boiling heat transfer in small diameter tubes", RT Edwards Inc. (2007).

[14] Pike-Wilson, E. A. and Karayiannis, T. G., "Flow boiling of R245fa in 1.1 mm diameter stainless steel, brass and copper tubes", *Experimental Thermal and Fluid Science*, 59, pp. 166–183 (2014).

[15] Gaheeshi, A. M. R. A., *Flow Boiling in Vertical Small to Micro Scale Tubes*, Ph.D. thesis, College of Engineering, Design and Physical Sciences Brunel University London (2018).

[16] Mahmoud, M. M., Karayiannis, T. G. and Kenning, D. B. R., "Surface effects in flow boiling of R134a in microtubes", *International Journal of Heat and Mass Transfer*, 54(15-16), pp. 3334–3346 (2011).

[17] Goodfellow, I., Bengio, Y. and Courville, A., *Deep learning*, MIT press (2016).