AN APPROACH FOR STACK-INTEGRATED FUEL CELL COOLING IN AEROSPACE; MACHINE LEARNING-BASED PERFORMANCE PREDICTION OF PULSATING HEAT PIPES

M. Illbruck*, F. Franke*, M. Kober*, S. Kazula*
* German Aerospace Center (DLR), Institute of Electrified Aero Engines, Lieberoser Straße 13a, Cottbus, Germany

Abstract

The transition towards more sustainable aviation has intensified research into novel concepts for aircraft propulsion. In this context, hydrogen-powered proton exchange membrane fuel cells (PEMFCs) represent a promising technology for aerospace applications, but require effective cooling to maintain high efficiency and minimize degradation.
Conventional liquid cooling systems require parasitic power and additional components, increasing system complexity and weight. As an alternative, pulsating heat pipes (PHPs) could be integrated into the residual space between bipolar plates, providing a passive, stack-integrated cooling solution.
For the efficient design of future PHPs, accurate performance prediction is essential. Therefore, this work expands a machine learning model based on the extreme gradient boosting (XGBoost) algorithm to predict PHP thermal resistance and employs game theory based Shapley values for model interpretation. The final model achieves a mean absolute percentage error (MAPE) of 13.6% and an R^2 value of 0.979, supporting the development of new PHP designs and their integration into fuel-cell stack cooling concepts for aerospace applications.

Keywords
Pulsating Heat Pipe; Aerospace Thermal Management Systems; Machine Learning

SYMBOLS & ABBREVIATIONS

Symbols

A_{cr}	Cross-sectional area	m^2
$A_{\mathrm{ev/cd}}$	Area ratio (EV/CD)	
α_{reg}	Regularization parameter	
Bo	Bond number	
c_{p}	Specific heat capacity	J/(kg·K)
D_{h}	Hydraulic diameter	m
ϵ	Residual	
FR	Fill ratio	%
G	Sum of first-order derivatives (gradient)	
γ	Pruning parameter	
g	Gravitational acceleration	$\mathrm{m/s}^2$
H	Sum of second-order derivatives (Hessian)	
h_{lv}	Latent heat of vaporization	J/kg
Ja	Jakob number	
k_{eff}	Effective thermal conductivity	W/(m·K)
Ku	Kutateladze number	
λ_{reg}	Regularization parameter	
$L_{\mathrm{cd\text{-}ev}}$	Distance between condenser and evaporator	m
L_{PHP}	PHP length	m
N_{turns}	Number of turns	
ν	Kinematic viscosity	m^2/s
Oh	Ohnesorge number	
$\dot{Q}$	Heat input	W
q''	Heat flux	$\mathrm{W/m}^2$
Re	Reynolds number	
ρ	Density	$\mathrm{kg/m}^3$
R_{th}	Thermal resistance	K/W
σ	Surface tension	N/m
T_{cd}	Condenser temperature	K
T_{ev}	Evaporator temperature	K
T_{∞}	Reference temperature	K

T_{sat}	Saturation temperature	K
w	Leaf weight	

Indices

cd	Condenser
ev	Evaporator
i	Sample index
j	Leaf index
l	Liquid phase
t	Tree index
v	Vapor phase

Abbreviations

BPP	Bipolar plate
CD	Condenser section
CFD	Computational Fluid Dynamics
CT	Capillary tube
CV	Cross-Validation
DLR	German Aerospace Center
EV	Evaporator section
FPPHP	Flat-Plate Pulsating Heat Pipe
LightGBM	Light Gradient Boosting Machine
Nested CV	Nested Cross-Validation
NIST	National Institute Of Standards And Technology
PEMFC	Proton Exchange Membrane Fuel Cell
PHP	Pulsating Heat Pipe
RF	Random Forest
SHAP	SHapley Additive exPlanations
VOF	Volume of Fluid Method
XGBoost	eXtreme Gradient Boosting

1. INTRODUCTION

Anthropogenic climate change is one of the most pressing challenges of the 21st century. Reducing the environmental impact of aviation therefore requires the development of alternative propulsion technologies [1]. One promising approach is the use of hydrogen-powered proton exchange membrane fuel cells (PEMFCs) for electric aircraft propulsion. To ensure the safe and reliable long-term operation of such systems, efficient thermal management of the fuel cell stacks is indispensable. State of the art liquid cooling systems can provide high cooling heat fluxes. However, they require additional auxiliary components, large heat exchangers, and significant parasitic power, for example, to operate pumps for the cooling channels. Moreover, their increased system complexity introduces additional potential single points of failure. A viable alternative that could mitigate these drawbacks while meeting thermal requirements is the integration of pulsating heat pipes (PHPs) into the fuel cell assembly [2]. PHPs are compact, two-phase heat transfer devices that have captivated substantial research interest in recent years due to their exceptional thermal conductivity, passive operation and geometrically adaptable design. These properties also position them as attractive technologies for applications with stringent space requirements like aerospace engineering [3].

PHPs, first patented by Akachi in 1990 [4], consist of capillary tubes that are evacuated and subsequently filled with a working fluid, resulting in an equilibrium of the liquid and vapor phase of the fluid. An exemplary PHP design is shown in Fig. 1.

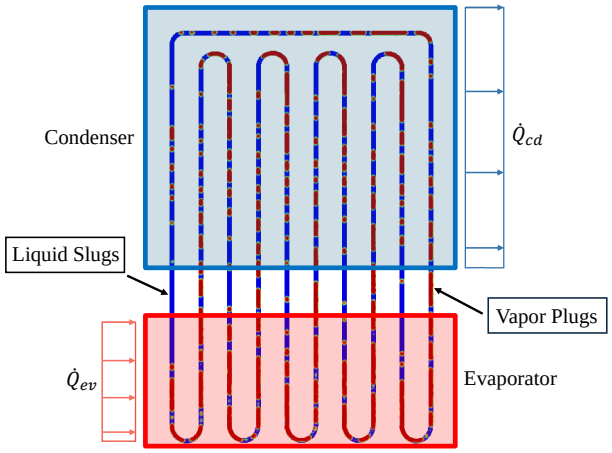


Fig. 1. Schematic representation of a PHP

Due to strong capillary forces, which act as the driving mechanism, the system operates through alternating structures of vapor and liquid phases. Fig. 2 displays an exemplary evaporator (EV) section of the PHP. The vapor and liquid phases of the working fluid are depicted in red and blue, respectively. The evaporator, receives heating power. Here, the heat causes partial evaporation of the liquid phases, leading to local pressure rises and expansions of the vapor phases. As a result, the alternating fluid structures are driven toward the condenser (CD) section. In the condenser, the vapor condenses and efficiently transfers heat, enabling fully passive operation with effective thermal conductivities far exceeding those of plain copper plates [5,6]. The central performance metric is the thermal resistance R_{th}:

$$R_{th} = \frac{T_{ev} - T_{cd}}{\dot{Q}} \tag{1}$$

where T_{ev} is the average temperature of the evaporator and T_{cd} is the condenser temperature. An extension of

this formulation is the effective thermal conductivity, which can be expressed as:

$$k_{eff} = \frac{L_{cd-ev} \cdot \dot{Q}}{A_{cr} \cdot (T_{ev} - T_{cd})}. \quad (2)$$

Here, L_{cd-ev} represents the distance between the condenser and evaporator sections, and A_{cr} denotes the cross-sectional area of the PHP [7].

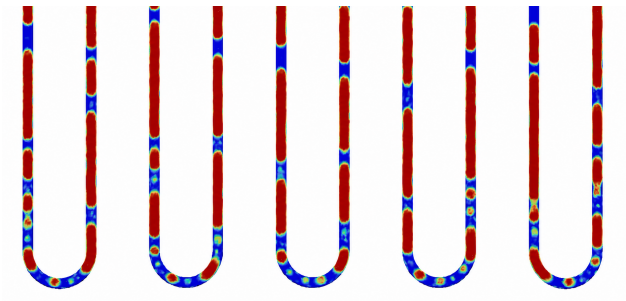


Fig. 2. Evaporator region of a simulated pulsating heat pipe (PHP) (Vapor plugs are shown in red and liquid slugs in blue)

From an external point of view the PHP is a comparatively simplistic system, but the internal coupling of the working fluid surface kinematics and thermohydrodynamics present a highly complex fluid mechanical problem. As a consequence, accurately modeling these coupled phenomena numerically requires substantial computational effort and involves considerable complexity. The liquid films that develop within the capillaries, are highly volatile and unstable. The evaporation of these liquid films is considered a major contributor to heat transfer in PHPs. Their thickness, which can be on the order of 100 µm, depends on the local temperature, pressure, velocity, and contact angle, as described by Zhang et al. [8]. To resolve these rapidly changing conditions, additional physical models and extensions beyond the conventional multiphase simulation methods are required. Moreover, only very fine computational meshes can adequately resolve these thin liquid films [9]. As this would result in prohibitively high computational costs, alternative methods for predicting the performance of pulsating heat pipes are worth considering.

With the increasing availability of machine learning algorithms, a model based on a previous study [10], utilizing the extreme gradient boosting (XGBoost) algorithm, was expanded using an extensive literature-based dataset from previous PHP experiments. The list of studies used to build the dataset is provided in the appendix 6. It includes both capillary tube (CT-) and flat-plate pulsating heat pipes (FPPHP).

2. THE XGBOOST ALGORITHM

This chapter introduces the principle of gradient boosting and provides a brief explanation of the operating principle of XGBoost and its mathematical background to facilitate the interpretation of the results of this work. Subsequently, Chapter 3 focuses on the application of this algorithm to the prediction of the thermal resistance of PHPs.

XGBoost, developed by Chen et al. (2016) [11], is an ensemble-based gradient-boosting algorithm that builds decision trees sequentially to minimize an objective function. It was selected as the most suitable model due to its extensive regularization capabilities, which help prevent overfitting, its potential for model interpretation, and its strong ability to capture complex parameter interactions [12]. These characteristics make it suitable for predicting PHP performance, while its flexible hyperparameter tuning offers substantial optimization potential without the computational overhead, typically associated with neural networks especially for smaller tabular datasets [13]. Furthermore, its performance is compared with other tree-based algorithms, namely, that of random forest (RF) and light gradient boosting machine (LightGBM) under similar experimental settings to further justify its selection [14, 15]. The results demonstrated superior performance for this specific use case, as shown in Section 5, supporting its adoption in this work.

First, a dataset comprising defined input features and a target variable is provided. The XGBoost algorithm then initializes an initial prediction of the target variable, from which the residual errors between the predicted and observed target values are determined. A decision tree is then constructed to identify patterns in these residuals. By recursively splitting the feature space based on the input features, the tree creates leaf nodes containing observations for which similar corrections to the current prediction can be made. The prediction associated with each leaf is then determined based on the residual errors of the observations assigned to that leaf.

Once the tree is completed according to a specified stopping criteria, its leaf predictions are added to the current model prediction, thereby reducing the residual errors. The updated predictions are subsequently used to calculate a new set of residuals. These residuals form the basis for the construction of the next tree, which therefore focuses on correcting the errors that remain after the previous trees. This process is repeated iteratively, with each additional tree acting as a weak learner that contributes to improving the overall prediction. Consequently, the final XGBoost model is obtained by combining the predictions of multiple sequentially constructed trees, with each tree progressively reducing the errors of the preceding model, as illustrated in Fig. 3.

The underlying mathematical formulation of this iterative optimization process is introduced in the following section. In the mentioned diagram $R_i^{(t)}$ is an individual residual at tree iteration t. The other symbols are explained in the following simplified derivation of the mathematical background of XGBoost.

For the literature-based dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{n}$ with $x_i \in \mathbb{R}^m$ representing the input feature vector of

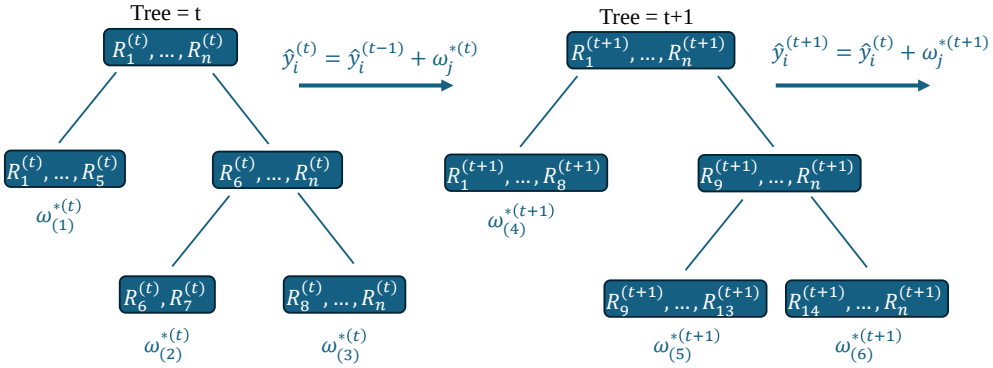


Fig. 3. XGBoost tree ensemble structure showing leaf nodes, splits, and progression from tree t to tree $t+1$ [10]

dimension m and $y_i \in \mathbb{R}$ the complementary thermal resistances and a total of n samples, the model predicts

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i), \qquad f_k \in \mathcal{F}, \tag{3}$$

where $\mathcal{F} = \{f(x) = w_{q(x)}\}$ denotes the space of regression trees ($q : \mathbb{R}^m \to T$, $w \in \mathbb{R}^T$), with q representing the tree structure and T the total number of leaf nodes. Each leaf weight w_j is found by minimizing the iteration-specific objective.

At iteration t, the objective function is minimized by finding a new regression tree f_t that reduces the prediction error of the existing ensemble while controlling model complexity:

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{4}$$

where l is the loss function measuring the discrepancy between the target y_i and the updated ensemble prediction. The regularization term

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{5}$$

penalizes overly complex trees and large leaf weights, while γ is a pruning parameter, regularizing the tree complexity and λ is the L2 regularization parameter.

The loss function l, which is the mean squared error (MSE) in this implementation, can be approximated using a second-order Taylor approximation as follows:

$$L^{(t)} \approx \sum_{i=1}^{n} \left[l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \tfrac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t) \tag{6}$$

where g_i and h_i are the first- and second-order derivatives of the loss w.r.t. the current prediction.

Setting the derivative of the simplified loss to zero

$$0 = \sum_{i \in I_j} g_i + \left(\sum_{i \in I_j} h_i + \lambda\right) f_t(x_i) \tag{7}$$

and rearranging the equation with respect to the definition of $f(x) = w_{q(x)}$, yields the optimal leaf weight

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}. \tag{8}$$

The corresponding gain for a candidate split is:

$$\text{Gain} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{9}$$

with G and H denoting the sums of g_i and h_i in the left/right child nodes. The gain determines the model's split decision by quantifying both the similarity of residuals and the potential improvement in predictive performance resulting from a candidate split. Specifically, it evaluates how well the current model explains the data through residual analysis while estimating the expected enhancement in predictive accuracy if the split is executed. The parameters γ and λ are hyperparameters that regulate tree growth and sensitivity, directly impacting the model's decisions. Hyperparameters govern the complexity of the trees and influence how the model is constructed. For example, γ controls the minimum reduction in loss required for a split to be considered worthwhile, thereby influencing tree depth and complexity. In contrast, λ penalizes leaf weights to reduce overfitting, encouraging smoother predictions and improving model stability. Further details on the model optimization methods are described in section 4.

3. MODEL FOR PREDICTING PULSATING HEAT PIPE PERFORMANCE

After gaining insight into the mathematics underlying XGBoost and how individual predictions are derived from multiple input features, this section describes how the approach is applied to predicting the thermal resistance of PHPs using selected input features.

3.1. Feature Engineering and Selection

A literature survey yielded 5,517 experimentally measured samples that span the heat-generation ranges of low- to high-temperature PEMFCs. The thermal resistance R_{th} of the PHPs is selected as the target variable for the ML model. Cases with thermal resistances outside the physically plausible interval and with fill ratios below 5% were discarded (see Tab. 1).

TAB 1. Operational envelope and data constraints for the PHP dataset

Parameter	Lower bound	Upper bound
Thermal resistance	0.001 K/W	10 K/W
Heat input	1 W	1200 W
Fill ratio	5%	100%

For all studies an equilibrium of heat fluxes between source and sink and ideal heat sinks (i.e., the cooling method of the condenser does not affect the measured resistance) are assumed. After standardizing the data, a consistent feature space shown in Tab. 2 was constructed. Features that would allow the direct calculation of the target variable were excluded to prevent data leakage. Geometric parameters, such as the hydraulic diameter, are introduced to establish a standardized basis for comparing different channel geometries and to enable generalization that allows the prediction of PHP performance of variable concepts as e.g. proposed by Franke et al. [2]. Additionally, a distinction is made between flat and regular pulsating heat pipes, as these configurations exhibit varying thermal and hydrodynamic behaviors.

The heat flux per unit area was obtained by dividing the supplied power by the projected effective evaporator area, and the heat-transfer-area ratio was defined as the non-adiabatic outer pipe area divided by the total PHP area. A dimensionless packing-density term was added to keep track of the pipe arrangement.

A fluid-property dictionary (NIST data and literature listed in the appendix A) supplied saturation-level properties for the most common working fluids used for calculating the dimensionless numbers used as input features.

For the Jakob number:

$$\mathrm{Ja} = \frac{c_p\,(T_\infty - T_{\mathrm{sat}})}{h_{lv}} \tag{10}$$

defined as the ratio of sensible to latent heat, a temperature difference of 20 K ($T_\infty - T_{\mathrm{sat}}$) was assumed, because the average evaporator temperature is not known a-priori.

At high temperatures at the evaporator the heat load can exceed the maximum heat the fluid can absorb before total evaporation and dryout occurs. Therefore, the Kutateladze number defined as:

$$\mathrm{Ku} = \frac{q''}{h_{lv}\,\rho_v\left(\sigma g(\rho_l - \rho_v)/\rho_v^2\right)^{1/4}} \tag{11}$$

was selected as it signals the onset of dryout when the applied heat flux exceeds the fluid's critical heat flux [16]. The Ohnesorge number:

$$\mathrm{Oh} = \nu_l\sqrt{\frac{\rho_l}{\sigma D_h}} \tag{12}$$

quantifies the damping of capillary-induced oscillations [17], while the Bond number:

$$\mathrm{Bo} = \frac{\rho_l g D_h^2}{\sigma} \tag{13}$$

indicates whether capillary forces dominate over gravity, a prerequisite for stable PHP operation. In the literature the critical bond number for dominating surface tension is ≈ 2 [18].

TAB 2. Input feature space for the XGB model

Feature	Type
Number of turns	Discrete
Area ratio (CD/EV)	Continuous
Fill ratio (FR)	Continuous
Inclination	Continuous
Hydraulic diameter	Continuous
Heat flux	Continuous
PHP length	Continuous
Ratio of heat-transfer area	Continuous
Jakob number	Continuous
Kutateladze number	Continuous
Bond number	Continuous
Ohnesorge number	Continuous
Pipe packing density	Continuous
Categorizing CTPHP/FPPHP	Categorical

4. HYPERPARAMETER SELECTION AND MODEL OPTIMIZATION

With the input features and target variables defined, the next step was to identify the significant hyperparameters and their respective search space, as their effects were exemplarily discussed in Sec. 2. This enabled the optimization of the model for the selected task. The hyperparameter search space used for optimizing the XGBoost model is defined by the bounds in Tab. 3.

These bounds constrain the optimization process, guiding the search toward models that balance predictive accuracy with strong generalization performance. The ranges were iteratively refined based on the model's behavior during optimization, particularly its tendency to overfit and the complexity of the resulting decision tree ensembles. The parameter optimization method selected in this work is Bayesian optimization. It is essentially based on a surrogate model and an evaluation function. Since directly evaluating the minimum

of the multidimensional objective function is computationally expensive, a surrogate function is used instead. In this implementation, a Gaussian process predicts both the mean and uncertainty for each point in the hyperparameter space. Combined with this surrogate modeling approach, which iteratively updates the expected distribution after each new evaluation, an acquisition function is used to balance exploration and exploitation, guiding the search of the ideal hyperparameters [19, 20].

To reduce the bias of a fixed training and testing dataset in final performance metrics of the model, nested cross-validation (CV) is implemented. Nested CV consists of an outer loop of k iterations in which the entire dataset is split into k subsets, shown in Fig. 4 (k=5 in this example). Here, $k-1$ subsets are selected as training data and 1 set for the testing of the model, respectively.

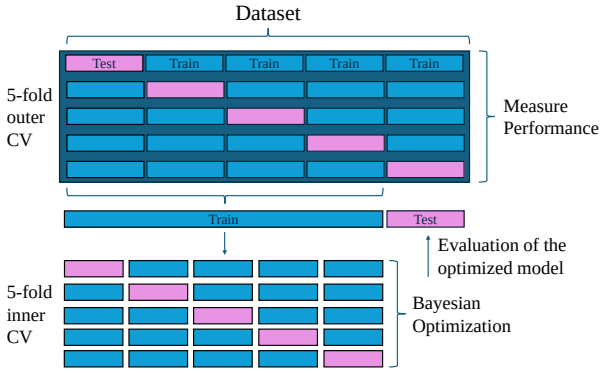


Fig. 4. Schematic of the nested cross-validation procedure ($k = 5$)

For each of the k iterations, a separate inner loop is initiated. This loop optimizes the hyperparameters based on a CV loop within the training dataset of the outer loop. When the inner loop is completed, the hyperparameter combination that leads to the best model in this iteration is validated on the test dataset from the outer loop at iteration k. The resulting metrics are compared for each of the k outer loop iterations [20, 21]. Finally, the hyperparameters that led to optimal performance among all these iterations are saved. In the implementation of this work, both the inner and outer loops use $k = 5$ folds, resulting in five different training and validation datasets. As a further method of model refinement, the thermal resistance is transformed logarithmically because the overall data are skewed, as shown in Fig. 5. The y-axis represents the density of samples within each bin. The bin size defines the width of the thermal resistance interval used to group the data into discrete ranges. In this case, each bin represents a thermal resistance range of 0.045 K/W, allowing the distribution and density of thermal resistance values within the dataset to be characterized. Multiplying the bin width by the corresponding density yields the percentage of thermal resistances within the dataset that falls into this resistance range.

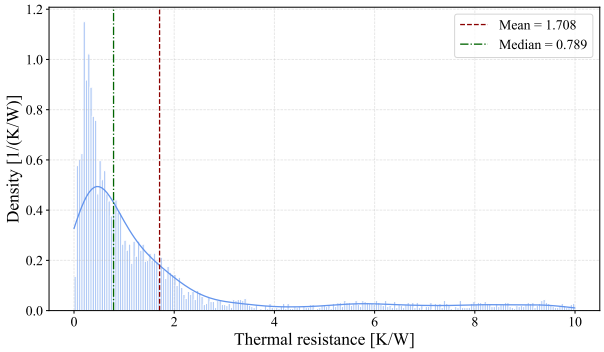


Fig. 5. Distribution of thermal resistance in the dataset (bin width ≈ 0.045 K/W)

Because the introduced logarithmic transformation changes the objective function (see Eq. 4) such that errors are evaluated in log-space and are therefore weighted approximately according to their relative magnitude rather than their absolute magnitude, a retransformation correction following Manning et al. [22] is applied. The raw residuals obtained from predicting thermal resistance using a fixed train-test split were found to deviate from a normal distribution. To account for potential heteroscedasticity in the retransformation, Duan's smearing estimate is applied in a stratified form, following the approach proposed by Manning et al., in which separate smearing factors are estimated for discrete ranges of the predicted thermal resistances. The residuals are grouped into bins based on quantiles of the predicted log-values $\hat{y}_{log}$, and a separate smearing factor is computed for each bin k:

(14)
$$\hat{E}[Y_{orig} \mid \hat{y}_{log} \in \text{Bin } k] = \exp(\hat{y}_{log}) \cdot \frac{1}{n_k} \sum_{i \in k} \exp(\hat{\epsilon}_i)$$

where n_k is the number of observations in bin k and $\hat{\epsilon}$ the residual. Under homoscedastic conditions, the smearing factor is expected to remain approximately constant across the different bins, such that stratification has little effect on the resulting predictions. In the presence of slight heteroscedasticity, however, bin-specific correction factors allow the retransformation bias to vary across the prediction range and reduce the risk of applying a correction that is disproportionately influenced by larger residuals at higher predicted values. The bin-specific smearing factors remained close to unity across most of the prediction range, indicating a limited retransformation bias for the majority of predictions.

In the case of studying PHP behavior, the interactions of the various features and their combined impact on the thermal performance are essential. To gain a comprehensive understanding of how the model evaluates the importance of different input features and how they interact in making predictions, a more detailed examination of each of the input features contributions to the model's prediction is necessary. In particular, it is of interest to identify which input features

most strongly influence the model's predictions, the magnitude of their effect on the target variable, and whether this influence varies significantly across individual cases. Such an analysis enables the identification of model uncertainties, the detection of potentially irrelevant or misleading features, and the assessment of whether the learned relationships are consistent or strongly case-dependent. A Shapley Additive exPlanations (SHAP)-based feature importance analysis, based on the concept introduced by L. S. Shapley, was conducted for this purpose. It essentially measures the contribution of each input feature to the model predictions of thermal resistance. For further information and the mathematical background it is referred to Lundberg et al. [23, 24].

TAB 3. Hyperparameter search space for the optimization of the XGBoost model

Parameter	Lower bound	Upper bound
Number of trees	100	3000
Maximum tree depth	3	12
Learning rate	0.01	0.3
Subsample of cases	0.5	0.9
Subsample of input features	0.5	0.9
λ_{reg} (L2 regularization)	1×10^{-5}	2000
α_{reg} (L1 regularization)	1×10^{-5}	100
Min. number of instances	1	10
γ (pruning parameter)	0	10
Tree method = 'hist'	–	–

5. RESULTS

The three machine learning algorithms compared for predicting thermal resistance are RF, LightGBM, and XGBoost. Although these models differ in their regularization and optimization strategies, which can affect their overall complexity, comparable hyperparameter ranges were defined prior to training to ensure a fair performance comparison. These included, for example, the number of trees, maximum tree depth, number of features considered at each split, and the minimum number of samples per leaf. The average performance metrics obtained from the nested cross-validation are summarized in Tab. 4. Across all considered metrics, XGBoost achieved the best overall performance, with an average mean absolute percentage error (MAPE) of 13.6% across the five outer folds of nested CV. This value represents the average relative deviation between the predicted and measured thermal resistance values on unseen validation data. The coefficient of determination, R^2, further confirms the strong predictive capability of the model. Since the optimal value of R^2 is 1, the achieved value of 0.979 indicates that the model substantially outperforms a baseline prediction based solely on the mean thermal resistance.

TAB 4. Model performance metrics overview of nested CV (MAE & RMSE in K/W)

Model	R^2	MAE	MAPE	RMSE
RF	0.968	0.2088	20.4%	0.4092
LightGBM	0.977	0.163	14.2%	0.345
XGBoost	0.979	0.1530	13.6%	0.3324

TAB 5. Selected hyperparameter configuration of the optimized XGBoost model following Bayesian optimization on the full dataset

Parameter	Value
Number of trees	2698
Maximum tree depth	12
Learning rate	0.3
Subsample of cases	0.9
Subsample of input features	0.9
λ_{reg} (L2 regularization)	1999.3
α_{reg} (L1 regularization)	0.0006488
Min. number of instances in leaf node	1
γ (pruning parameter)	0.0
Tree method	'hist'

The optimized XGBoost model with respect to the hyperparameters that yielded the best performance in the nested CV iterations is listed in Tab. 5.
For reference, the optimization of the XGBoost model took roughly 12 hours and 48 minutes. Once optimized, however, predictions of new cases can be conducted in a few seconds, which exhibits a significant reduction in computational time compared to conventional approaches, such as numerical simulations, for the investigation of the thermal resistance of PHP systems.
The parity plot shown in Fig. 6, compares the predicted and measured thermal resistance values obtained with the optimized XGBoost model.
The model was trained using the hyperparameter configuration of Tab. 5, with 80% of the data randomly selected for training and the remaining 20% used for testing. The achieved MAPE values were 5.886% for the training set and 13.41% for the test set. Although the difference between the training and test performance may indicate some degree of overfitting, the test performance is consistent with the average nested CV result. This suggests that the model exhibits stable generalization performance, and no substantial overfitting is suspected. Furthermore, the MAPE obtained on the test set is consistent with the average performance observed across the outer folds of the nested CV. To evaluate the predictive capability of the ML model as an alternative to conventional approaches, an entirely new PHP dataset that was not used during any phase of model development, was extracted from experimental measurements documented in prior stu-

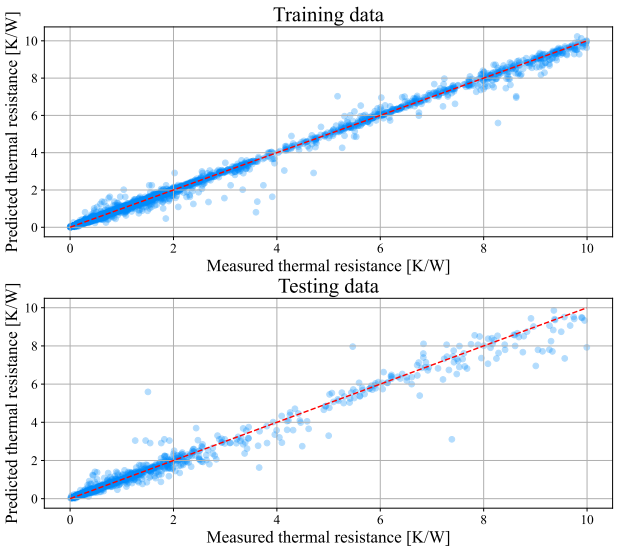


Fig. 6. Parity plot of the predicted thermal resistance of a randomly selected 80% (MAPE = 5.886%) training subset of the data and a 20% subset used for testing (MAPE = 13.41%)

dies [25]. The XGBoost model was then employed to predict the thermal resistances corresponding to these experiments. The comparative results are presented in Fig. 7. In this figure, the fill ratios of 0.4, 0.6, and 0.8 were predicted and benchmarked against the experimental results. The predictions for a fill ratio of 0.4 achieved a MAPE of 31.11%, while the predictions for fill ratios of 0.6 and 0.8 resulted in MAPE values of 19.25% and 21.97%, respectively. In addition to predicting the absolute thermal resistances, the model also correctly captures the trend that a fill ratio of 0.6 leads to lower thermal resistances than a fill ratio of 0.8 for this PHP configuration, which is consistent with the experimental measurements.

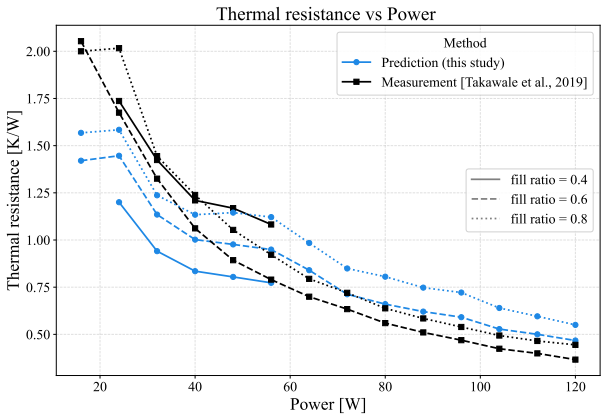


Fig. 7. Comparison of predicted and measured thermal resistances for the CTPHP experiments conducted by A. Takawale et al. [25]

Furthermore, both the machine learning model and the experiments reflect the characteristic start-up behavior, the subsequent increase in the steepness of the thermal resistance reduction with rising heat input, and the eventual flattening of the curve at higher power levels. For the fill ratio of 0.4, the experiments reached dryout at 64 W and were therefore stopped at this point. Since the current machine learning model does not yet account for or provide warnings regarding dryout phenomena, the predictions were also discontinued at this heat input. Overall, the predictions are in reasonable agreement with the experimental results. However, the accuracy is lower than the average performance obtained from nested CV. The machine learning model learns relationships between the input features and the target variable from the available training data. Consequently, its accuracy is expected to decrease in regions of the parameter space, where the distances between known data points become larger. This becomes particularly evident when entirely new geometries are predicted. In such cases, the model faces difficulties in achieving high predictive accuracy because the geometrical features exhibit the smallest variance within the feature space. Furthermore, each study is affected by its own measurement bias and experimental uncertainty. Therefore, when predictions are made for new operating regimes based on PHP measurements obtained using the same experimental setup and measurement instruments already represented in the training data, the model is expected to provide more accurate estimates. This assumption is supported by the lower MAPE achieved in nested CV, as in this approach, the entire dataset is randomly split into training and test sets, allowing the model to study and predict thermal resistances across different experimental setups. Unfortunately, this experimental bias cannot be entirely excluded when using experimental data across multiple studies and is inherent to each experimental setup.

However, diverse experimental setups not only encounter biases, they also exhibit noticeable uncertainties and measurement errors that must be accounted for while analyzing the accuracy of prediction of simulation attempts of PHP experiments. A. Takawale et al. [25] report an uncertainty of the measured thermal resistance of 11.2%.

In lower power regions where the temperature difference $T_{EV} - T_{CD}$ is small, the accuracy uncertainty can rise to approximately 24% as reported by F. Schwarz et al. [26], depending on the accuracy of the used thermocouples in the experiments. At higher temperature differences and power levels, this decreases to 9%. Other studies report uncertainties in thermal resistance measurements on the order of 3–4% [27] and 5% in [28]. Since the average MAPE, used as a measure of relative prediction error, is in some cases within or only slightly above the existing experimental uncertainty band, the results from the nested CV can be interpreted as indicating that the model achieves a predictive accuracy only slightly worse than the measurement accuracy of the existing experimental data.

It should be noted that the model is not intended solely for direct prediction of thermal resistance. As introduced previously, SHAP is used to interpret the model in greater detail. Originally Shapley or SHAP

values were used to quantify how much a specific player contributes to a total outcome relative to the other players in game theory [29]. In the context of machine learning, the players represent the input features, while the total outcome corresponds to the predicted thermal resistance in this case. The model was therefore trained on the entire dataset, after which a beeswarm plot was generated, as shown in Fig. 8, to visualize the distribution of the SHAP values across all observations. The x-axis represents the SHAP value and therefore the contribution of an individual feature to the model output for a given observation. Each dot represents the SHAP value of a feature for one observation. Positive SHAP values indicate that the feature increases the predicted thermal resistance relative to the baseline, while negative SHAP values indicate a decrease. The baseline corresponds to the expected thermal resistance based on the rest of the dataset.

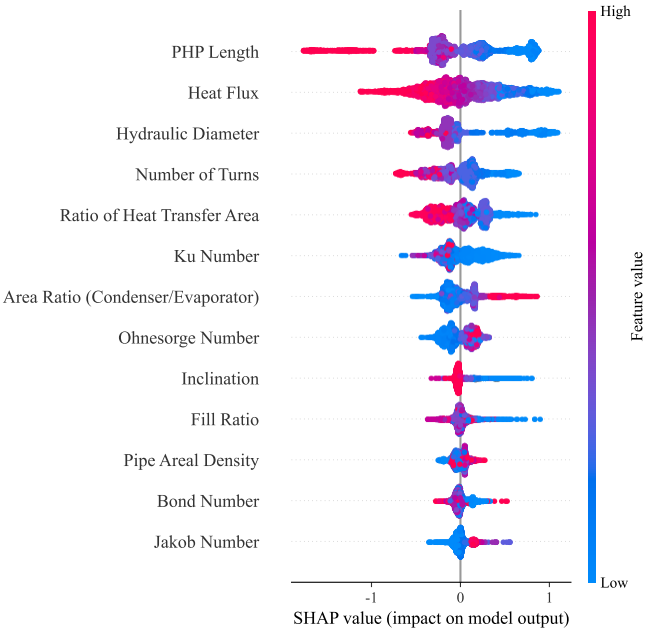


Fig. 8. Beeswarm plot of the entire dataset

Thus, if one were to add each contribution of each of the input features to the baseline value, this would result in the model prediction for that specific case. The higher the feature is in the list, the higher the overall average absolute contribution to the final prediction. This implies that, for the model predictions, the length of the PHP is the most important feature, followed by the heat flux, though it has to be said that each correlation and impact on the thermal resistance is no physical law but the correlations the model learned during training from the present dataset. The beeswarm plot also reveals trends learned by the model from the data. For example, high heat flux values, represented by red points in the feature-specific color scale, are predominantly associated with negative SHAP values, indicating that higher heat flux reduces the model-predicted thermal resistance. In the same way, medium to high hydraulic diameters also reduce it.

Another important PHP design parameter is the already discussed PHP length. The model reveals improved performance, meaning reduced thermal resistance, with increasing overall length of the straight PHP strands. To understand how the length is composited the area ratios give more insight into the individual sections present in the PHP. High ratios of (active) heat transfer area, for example, indicate that for improved PHP performance the adiabatic area should be minimized, which makes sense in reality, since if the area that actively participates in heat transfer is maximized, the heat transfer is more effective. The area ratio of the condenser to the evaporator, suggest that an increased evaporator area leads to a reduction in thermal resistance, presumably due to better evaporation effects and more homogeneously spread heat, reducing heat spikes, which would also be preferred in PHP integration into bipolar plates of fuel cells. An increased number of turns in a PHP enhances heat dissipation by distributing thermal energy across more pipes. Beyond this, the pipe area density parameter points out that the performance supposedly declines, as excessive packing density introduces geometric constraints. For instance, densely packed PHPs often feature sharp radii at meanders, which elevate wall–fluid resistance and lead to an increased pressure drop. This damping of fluid pulsation could, in turn, increase thermal resistance as experimentally shown by S. Parthasarathi [30].

According to the ML model, the Kutateladze and Ohnesorge numbers are the most relevant dimensionless non-geometrical parameters for predicting thermal resistance. A higher Kutateladze number appears to be favorable, although excessively high heat input can increase evaporation and lead to dryout, thereby reducing performance. In contrast, higher Ohnesorge numbers are associated with lower performance, which is consistent with the expected influence of increased viscous effects and flow resistance within the PHP. The Bond and Jakob number show considerably lower feature importance in the model, indicating that their influence on the predicted thermal resistance is less pronounced compared to the Kutateladze and Ohnesorge number.

Collectively, the SHAP analysis points out that the PHP length is the dominant predictor of thermal resistance for the ML model, followed by the heat flux, with both exhibiting clear trends (e.g., higher heat flux and longer PHPs reduce thermal resistance). Geometric parameters, such as the hydraulic diameter, area ratios, and number of turns, further refine predictions, while operational modes (degree of inclination) and dimensionless numbers (e.g., Ku, Bond, Oh) reveal nuanced dependencies. These insights are broadly consistent with experimental expectations and demonstrate that the model captures trends that

can be related to physically meaningful patterns, despite not being explicitly physics-informed.

6. CONCLUSION

In this study, an alternative model for the efficient and fast prediction of thermal resistances of PHPs was expanded, accelerating future design iterations and enabling their reliable use in thermal management systems for future aerospace applications. The developed ML model provides a detailed overview of the interactions between input features and their impact on thermal resistance across different PHP designs and operating regimes. Nested CV showed stable and accurate predictions for varying train and test splits, achieving an average MAPE of 13.6% and R^2 of $\approx$ 0.979. The inclusion of dimensionless thermophysical and geometrical parameters, enabled the model to capture relevant correlations more effectively. This also allowed the model to move beyond purely circular PHP cross sections.

Overall, the results show that the proposed ML approach can predict PHP performance with significantly reduced computational effort compared with conventional approaches of investigating the performance of PHPs. However, the predictive capability depends strongly on how well relevant information is represented in the training data. If the feature space is sparsely sampled in certain regions, the predictive accuracy may deteriorate, as the model does not incorporate the underlying physical laws and therefore cannot reliably extrapolate beyond the given data. Another current source of inaccuracy may be the sole use of the hydraulic diameter. While the hydraulic diameter is useful as an equivalent characteristic diameter to model non-circular cross sections, it cannot fully describe the influence of varying cross-sectional shapes on PHP thermo-hydrodynamics. Channels with identical hydraulic diameters but different shapes may exhibit different capillary pressure distributions, liquid-film thicknesses, wettability behavior, and corner-flow effects. For example, sharp corners in square cross sections can increase local capillary pressure and enhance corner wetting compared to circular channels.

Therefore, future work should extend the geometric description beyond hydraulic diameter by introducing additional shape factors that account for corner curvature, wetted perimeter distribution, and capillary confinement effects. This would likely improve the model's ability to generalize across arbitrary PHP geometries and further strengthen its use as a fast prediction and optimization tool for future PHP design.
Beyond predicting thermal resistance, the model can be leveraged as an optimization tool by framing the input features (e.g., working fluids, geometric configurations) and target thermal resistance as a multidimensional optimization problem. Under fixed constraints, the model could propose optimal parameters, such as meander radius or number of turns, to minimize thermal resistance, analogous to Bayesian optimization approaches.
A broader and denser feature space is expected to enhance the model's generalization to unseen data, reduce biases from individual experimental studies, and strengthen the identification of trends between input features and thermal performance, particularly when diverse datasets (with varying measurement methods and accuracies) are aggregated.
In conclusion, the expanded model enables the analysis of the influence of input features, including critical PHP parameters, and the correlations learned by the model with their impact on thermal resistance. Furthermore, the model can rapidly predict thermal resistance, making it a valuable tool for evaluating new operating conditions and supporting efficient analysis of PHP performance and iterative development of new PHP designs.

Contact Address:

marvin.illbruck@dlr.de

Literature

[1] Flightpath 2050: Europe's vision for aviation; maintaining global leadership and serving society's needs; report of the high-level group on aviation research. policy. Publ. Off. of the Europ. Union, Luxembourg, 2011. ISBN: 9789279197246.

[2] F. Franke, M. Kober, and S. Kazula. Preliminary design and thermal evaluation of a hexagonal pulsating heat pipe for fuel cell applications. Case Studies in Thermal Engineering, 2025. DOI: 10.1016/j.csite.2025.107272.

[3] F. Franke, A. Link, and S. Kazula. Evaluation of heat transfer technologies for high temperature polymer electrolyte membrane fuel cells as primary power source in a regional aircraft. CEAS Aeronautical Journal, 2025. DOI: 10.1007/s13272-025-00908-0.

[4] Akachi, H. Structure of a heat pipe, 1990.

[5] M. Winkler, D. Rapp, A. Mahlke, F. Zunftmeister, M. Vergez, E. Wischerhoff, J. Clade, K. Bartholomé, and O. Schäfer-Welsen. Small-sized pulsating heat pipes/oscillating heat pipes with low thermal resistance and high heat transport capability. Energies, 13(7):1736, 2020. DOI: 10.3390/en13071736.

[6] M. Madhuri and N. P. Yadav. Experimental investigation of pulsation behaviour and performance of pulsating heat pipe. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, 237(3):559–569, 2023. DOI: 10.1177/09576509221131602.

[7] M. Mameli, M. Marengo, and S. Khandekar. Local heat transfer measurement and thermo-fluid characterization of a pulsating heat pipe. International Journal of Thermal Sciences, 2014. DOI: 10.1016/j.ijthermalsci.2013.07.025.

[8] X. Zhang and V. S. Nikolayev. Physics and modeling of liquid films in pulsating heat pipes. Physical Review Fluids, 2023. DOI: 10.1103/PhysRevFluids.8.084002.

[9] V. S. Nikolayev. Physical principles and state-of-the-art of modeling of the pulsating heat pipe: A review. Applied Thermal Engineering, 2021. DOI: 10.1016/j.applthermaleng.2021.117111.

[10] F. Franke, M. Illbruck, M. Kober, L. Enghardt, and S. Kazula. Approach towards pulsating heat pipe performance prediction using gradient boosting and experimental iteration. GAMM, 2026.

[11] T. Chen and C. Guestrin. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016. DOI: 10.1145/2939672.2939785.

[12] J.-P. Lai, S. Lin, V. Lin, A. Kang, Y.-P. Wang, and P.-F. Pai. Predicting Thermal Resistance of Packaging Design by Machine Learning Models. Micromachines, 2025. DOI: 10.3390/mi16030350.

[13] L. Grinsztajn, E. Oyallon, and G. Varoquaux. Why do tree-based models still outperform deep learning on tabular data? Advances in Neural Information Processing Systems 35, 2022. DOI: 10.48550/arXiv.2207.08815.

[14] L. Breiman. Random forests. Machine Learning, (Volume 45):pages 5–32, 2001. DOI: 10.1023/A:1010933404324.

[15] G. Ke, Q. Meng, T. Wang T. Finely, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu. Lightgbm: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems 30 (NIPS 2017), 2017.

[16] M. Mameli, M. Manzoni, L. Araneo, S. Filippeschi, and M. Marengo. Pulsating Heat Pipe in Hypergravity Conditions. Heat Pipe Science and Technology, 2015. DOI: 10.1615/HeatPipeScieTech.2016013605.

[17] M. Stange, M. E. Dreyer, and H. J. Rath. Capillary driven flow in circular cylindrical tubes. Physics of Fluids, 2003. DOI: 10.1063/1.1596913.

[18] S. Khandekar, N. Dollinger, and M. Groll. Understanding operational regimes of closed loop pulsating heat pipes: an experimental study. Applied Thermal Engineering, 23(6):707–719, 2003. DOI: 10.1016/S1359-4311(02)00237-5.

[19] Optuna Contributors. sampler.py — GP Sampler Source Code (Optuna). 2026.

[20] C. E. Rasmussen and C. K. I. Williams. Gaussian processes for machine learning. Adaptive computation and machine learning. MIT Press, Cambridge, Mass., 2006. ISBN: 9780262256834. DOI: 10.7551/mitpress/3206.001.0001.

[21] M. J. Lewis, A. Spiliopoulou, K. Goldmann, C. Pitzalis, P. McKeigue, and M. R. Barnes. nestedcv: an R package for fast implementation of nested cross-validation with embedded feature selection designed for transcriptomics and high-dimensional data. International Society For Computational Biology, 2023. DOI: 10.1093/bioadv/vbad048.

[22] W. G. Manning and J. Mullahy. Estimating log models: to transform or not to transform? Journal of Health Economics, 2001. DOI: 10.1016/S0167-6296(01)00086-8.

[23] S. M. Lundberg and S.-I. Lee. A Unified Approach to Interpreting Model Predictions. arXiv:1802.03888, 2017. DOI: 10.48550/arXiv.1705.07874.

[24] S. M. Lundberg, G. G. Erion, and S.-I. Lee. Consistent Individualized Feature Attribution for Tree Ensembles. arXiv:1802.03888, 2018. DOI: 10.48550/arXiv.1802.03888.

[25] A. Takawale, S. Abraham, A. Sielaff, P. S. Mahapatra, A. Pattamatta, and P. Stephan. A comparative study of flow regimes and thermal performance between flat plate pulsating heat pipe and capillary tube pulsating heat pipe. Applied Thermal Engineering, 149:613–624, 2019. DOI: 10.1016/j.applthermaleng.2018.11.119.

[26] F. Schwarz, V. Danov, A. Lodermeyer, A. Hensler, and S: Becker. Thermodynamic analysis of the dryout limit of oscillating heat pipes. Energies, 13(23):6346, 2020. DOI: 10.3390/en13236346.

[27] K.-S. Yang, M.-Y. Jiang, C.-Y. Tseng, S.-K. Wu, and J.-C. Shyu. Experimental investigation on the thermal performance of pulsating heat pipe heat exchangers. Energies, 13(1):269, 2020. DOI: 10.3390/en13010269.

[28] S. K. Hota, K.-L. Lee, B. Leitherer, G. Elias, G. Hoeschele, and S. Rokkam. Pulsating heat pipe and embedded heat pipe heat spreaders for modular electronics cooling. Case Studies in Thermal Engineering, 49:103256, 2023. DOI: 10.1016/j.csite.2023.103256.

[29] B. Rozemberczki, L. Watson, P. Bayer, H.-T. Yang, O. Kiss, S. Nilsson, and R. Sarkar. The Shapley Value in Machine Learning. arXiv:2202.05594, 2022. DOI: 10.48550/arXiv.2202.05594.

[30] S. Parthasarathi, S. Nagarajan, S. Desai, S. U. Reddy, and G. S. V. L. Narasimham. Effect of

bend radius and insulation on adiabatic section on the performance of a single closed loop pulsating heat pipe: experimental study and heat transfer correlation. Heat and Mass Transfer, 2021. DOI: 10.1007/s00231-021-03085-y.

A. APPENDIX

TAB 6. Publications used for building the PHP dataset

First Author	Year	DOI
S. Khandekar	2004	10.18419/opus-1637
X.M. Zhang	2004	10.1080/08916150490246546
S. Liu	2007	10.1007/s11630-007-0056-8
Y.-H. Lin	2008	10.1016/j.applthermaleng.2007.10.019
Z. Lin	2009	10.1007/s11431-009-0194-1
H. Yang	2009	10.1016/j.ijthermalsci.2008.05.017
M. Shafii	2010	10.1080/01457630903547636
Y. Ji	2011	10.1016/j.expthermflusci.2011.01.007
M. Mohammadi	2012	10.1115/1.4004805
K.-H. Chien	2012	10.1016/j.ijheatmasstransfer.2012.05.068
J. Qu	2013	10.1016/j.apenergy.2013.02.030
B. Verma	2013	10.4236/jectc.2013.31004
P. R. Pachghare	2013	10.1016/j.proeng.2013.01.088
Y. Hu	2013	10.1016/j.ijheatmasstransfer.2013.11.031
H. Han	2014	10.1016/j.ijthermalsci.2014.04.003
T. Hao	2014	10.1016/j.ijheatmasstransfer.2014.01.007
M. Mameli	2014	10.1016/j.expthermflusci.2014.04.009
N. Saha	2014	10.1016/j.ijheatmasstransfer.2014.02.067
Y. Zhu	2014	10.1016/j.ijheatmasstransfer.2014.05.042
F. Laun	2014	10.5098/fhp.5.8
P. R. Pachghare	2014	10.1007/s12206-014-0751-9
M. Ebrahimi	2015	10.1016/j.applthermaleng.2015.07.040
J. Zhao	2016	10.1016/j.applthermaleng.2015.09.018
C. Y. Tseng	2016	10.1016/j.applthermaleng.2016.06.034
A. Jokar	2016	10.1007/s00231-016-1759-8
J. Zhao	2016	10.1016/j.ijheatmasstransfer.2016.03.108
S. Shi	2016	10.1016/j.applthermaleng.2016.04.014
X. Liu	2016	10.3390/app6110321
J. Wang	2016	10.1016/j.applthermaleng.2016.03.163
H. R. Goshayeshi	2016	10.1016/j.expthermflusci.2016.01.003
B. N. Mahapatra	2016	10.1016/j.applthermaleng.2016.07.137
S. Jun	2016	10.1016/j.ijheatmasstransfer.2015.12.064
J. Qu	2017	10.1016/j.expthermflusci.2017.02.022
S.W. Kang	2017	10.1016/j.applthermaleng.2017.02.051
H. R. Goshayeshi	2017	10.1016/j.physe.2017.01.014
Q. Sun	2017	10.1016/j.applthermaleng.2017.06.109

TAB 7. Publications used for building the PHP dataset (continued)

First Author	Year	DOI
W. Wang	2017	10.1007/s00231-016-1958-3
J. Lee	2017	10.1016/j.ijheatmasstransfer.2016.11.036
J. Kim	2018	10.1016/j.applthermaleng.2017.11.009
D. Bastakoti	2018	10.1016/j.ijheatmasstransfer.2017.10.075
E. Sedighi	2018	10.1016/j.ijthermalsci.2018.08.024
A. Facin	2018	Proceedings of the 19th International Heat Pipe Conference (IHPC), Pisa
J. Zhao	2018	10.1016/j.expthermflusci.2018.05.026
R. Xu	2019	10.1016/j.ijheatmasstransfer.2019.03.070
C. Czajkowski	2019	10.1016/j.applthermaleng.2019.114534
X. Lin	2019	10.1016/j.expthermflusci.2019.109871
J. Qu	2019	10.1016/j.ijheatmasstransfer.2019.03.107
S. Jun	2019	10.1016/j.ijheatmasstransfer.2019.03.163
R. Zamani	2019	10.1007/s00231-018-2418-z
A. Wei	2019	10.1016/j.ijheatmasstransfer.2019.02.021
V.M. Patel	2019	10.1016/j.enconman.2019.05.014
M. Kumar	2019	10.1080/01457632.2018.1436390
P. Srikrishna	2019	10.1007/s00231-019-02607-z
K. Mehta	2020	10.1007/s40430-020-2213-x
H. Ahmad	2020	10.1016/j.ijheatmasstransfer.2019.119245
L. Betancur	2020	10.1016/j.applthermaleng.2019.114734
Y. Xu	2020	10.1016/j.ijheatmasstransfer.2020.119727
K. Bao	2020	10.1016/j.applthermaleng.2020.115678
W. Kim	2020	10.1016/j.ijheatmasstransfer.2019.119251
P. Srikrishna	2021	10.1007/s00231-021-03085-y
L. Krambeck	2022	10.1016/j.applthermaleng.2022.118509
V. Ayel	2022	10.1016/j.ijthermalsci.2022.107599
C. Chang	2022	10.3390/mi13111949
L. Malla	2022	10.1016/j.ijheatmasstransfer.2022.123066
F. Yao	2024	10.1016/j.ijthermalsci.2024.109502
X. Chen	2024	10.1016/j.icheatmasstransfer.2024.107282

TAB 8. Sources used for the fluid properties.

Fluid	Reference	Source
Various	NIST (2026)	Link
Various	DataPhysics Instruments (2026)	Link
FC-72	3M (2026)	Link
Ethanol	NIST REFPROP (2026)	Link
Acetone	Krescher et al. (2023)	DOI
HFE-7100	Rausch et al. (2015)	DOI
HFE-7100	Manetti et al. (2021)	DOI
HFE-7100	Liu et al. (2022)	DOI