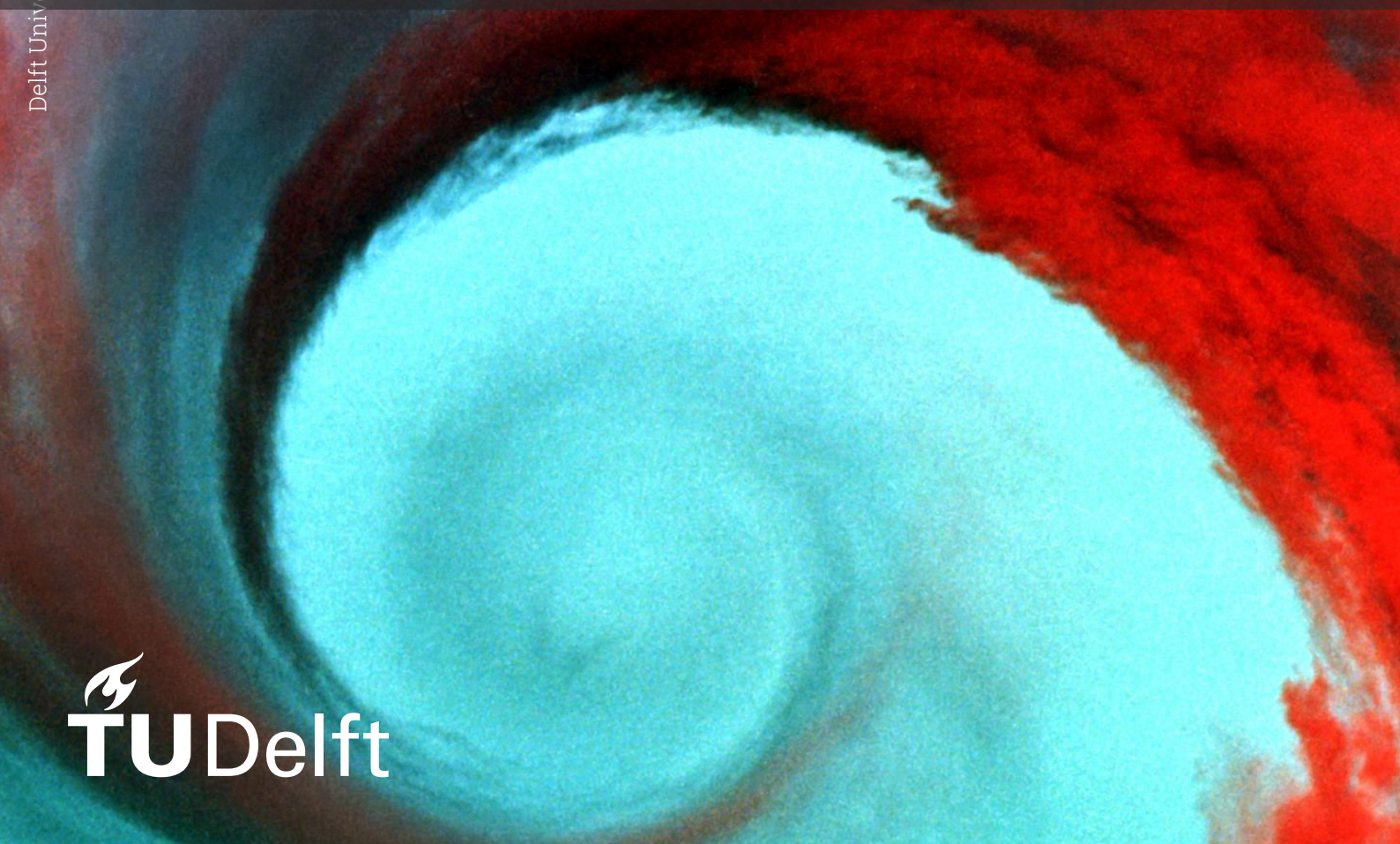
# Prediction of heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated vertical tube

Master thesis

Floris Krijgsman

Delft University of Technology



TU Delft

# Prediction of heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated vertical tube

Master thesis

by

Floris Krijgsman

| Student Name | Student Number |
| --- | --- |
| Floris Krijgsman | 4602706 |


Supervisors: Prof.dr. R. Pecnik, Dr.ir. J.W.R. Peeters, Dr. H. Caesar
Project Duration: May, 2023 - May, 2024
Faculty: Faculty of Mechanical, Maritime and Materials Engineering, Delft

Cover: Vortex art by NASA Langley Research Center [1]

TU Delft

# Summary

Supercritical $CO_2$ ($SCO_2$) is a promising alternative to traditional working fluids in heat pumps and power cycles due to its high density, thermal efficiency, and stability. These properties allow for the design of more compact and efficient equipment. However, accurately modeling supercritical heat transfer, especially near its pseudocritical point, is challenging because of extreme variations in its thermophysical properties. As a result of these inaccuracies in modeling, large fault margins must be taken into account, leading to overdesign of equipment.

Existing methods for estimating heat transfer in such systems include empirical measurements, computational fluid dynamics (CFD), and Nusselt correlations. Although measurements provide accurate data, it is not a scalable solution. CFD methods offer this scalability. However, CFD cannot be applied in complex scenarios such as modeling of $SCO_2$ due to the trade-off between computational cost and accuracy. Nusselt correlations have low computational demand and are a scalable solution but come with low accuracy, making Nusselt correlations unreliable for equipment design.

To overcome these limitations, this research applies a new model to predict heat transfer to $SCO_2$ flowing upward in a heated vertical tube. Current prediction methods primarily rely on artificial neural networks (ANN). This research applies a convolutional neural network (CNN) for its ability to capture spatial context from the heat transfer trajectory. This property enables CNN to capture global and local patterns important for accurate prediction in $SCO_2$ heat transfer. In addition, the multiple kernels per layer enables the model to extract various features relevant to heat transfer. The CNN model has shown superior performance, achieving an $R^2$ score of 0.970 and an MSE of 1477, outperforming existing prediction methods and Nusselt correlations.

A feature importance study was performed to identify the nondimensional parameters that are important for heat transfer prediction. Furthermore, the experiments show that normalization is vital to enhance model performance, countering issues such as bias and exploding or vanishing gradients.

The findings in this research suggest great potential for using machine learning models to design more effective and compact heat transfer equipment. Future work will focus on a feature importance study for normalized, dimensional, and nondimensional parameters for improved predictions. In addition, more synthetic data should be generated in the heat transfer detoriation and heat transfer enhancement areas for better prediction of those parts. Finally, a study into the application of machine learning models for designing heat transfer equipment should be done to show the benefits of such models.

# Contents

# Nomenclature

## Abbreviations

| Abbreviation | Definition |
|---|---|
| AAD | Average absolute deviation |
| ANN | Artificial neural network |
| CFD | Computational fluid dynamics |
| CNN | Convolutional neural network |
| $CO_2$ | Carbon dioxide |
| DNS | Direct numerical simulations |
| HTD | Heat transfer detoriation |
| HTE | Heat transfer enhancement |
| LES | Large eddy simulations |
| MAE | Mean absolute error |
| MAPE | Mean average precision error |
| MSE | Mean squared error |
| RANS | Reynolds-averaged Navier-Stokes |
| ReLU | Rectified linear unit |
| RF | Random forest |
| $SCO_2$ | Supercritical carbon dioxide |
| SVR | Support vector regression |
| val | Validation |

## Symbols

| Symbol | Definition | Unit |
|---|---|---|
| $Bu$ | Buoyancy parameter | [-] |
| $c_p$ | Specific heat capacity | [J/kg · K] |
| $\overline{c}_p$ | $\overline{c}_p = \frac{h_w - h_b}{T_w - T_b}$ | |
| $d$ | Diameter of the tube | [m] |
| $Fr$ | Froude number | [-] |
| $G$ | Mass flux | [kg/m$^2$ · s] |
| $Gr$ | Grashof number | [-] |
| $h$ | Specific enthalpy | J/kg |
| $k$ | Thermal conductivity | [W /m · K] |
| $K_v$ | Flow acceleration parameter | [-] |
| $\dot{m}$ | Mass flow | kg/s |
| $Nu$ | Nusselt number | [-] |
| $P$ | Pressure | [Pa] |
| $Pr$ | Prandtl number | [-] |
| $\overline{Pr}$ | $\overline{Pr} = \frac{\mu \cdot \overline{c_p}}{k}$ | [-] |
| $Q$ | Heat transfer rate | [W] |
| $q$ | Heat flux | [W/m$^2$] |
| $Q^+$ | Local heat flux parameter | [-] |
| $q^+$ | Nondimensional heat flux | [-] |
| $Re$ | Reynolds number | [-] |
| $R^2$ | Coefficient of determination | [-] |

| Symbol | Definition | Unit |
|---|---|---|
| $T$ | Temperature | [K] |
| $u$ | Velocity | [m/s] |
| $x/d$ | Dimensionless length | [-] |
| $\alpha$ | Heat transfer coefficient | [W/m$^2$K] |
| $\mu$ | Dynamic viscosity | [kg/m s] |
| $\rho$ | Density | [kg/m$^3$] |
| $\overline{\rho}$ | $\overline{\rho} = \frac{1}{T_w - T_b} \int_{T_B}^{T_w} \rho \, dT$ | |

# Indices

| Symbol | Definition |
|---|---|
| $b$ | bulk |
| $c$ | critical |
| $exp$ | experimental |
| $h$ | hydraulic |
| $in$ | inlet |
| $out$ | outlet |
| $pc$ | pseudocritical |
| $pred$ | predicted |
| $r$ | reduced |
| $w$ | wall |

# 1 Introduction

Supercritical $CO_2$ ($SCO_2$) is emerging as a promising fluid to accelerate the energy transition. $SCO_2$ offers significant improvements in heat pumps and power cycle performance due to its unique properties:

- High density: This allows for a more compact turbomachinery and heat exchanger design, which reduces the cost and footprint.
- High thermal efficiency: Compared to steam, $SCO_2$ has a much higher power density, resulting in an increase in thermal efficiency.
- High thermal stability: $SCO_2$ can maintain stable temperature and pressure under a wide range of conditions, making it suitable for use in processes that require consistent heat transfer.
- Clean: $SCO_2$ is a clean, nontoxic, nonhazardous, and stable fluid in contrast to current refrigerants being used. Furthermore, refrigerants used in heat pumps and power cycles are phased out due to environmental regulations [13], positioning $SCO_2$ as the perfect solution.

## 1.1. Motivation

This section outlines the challenges of designing equipment that uses $SCO_2$ as a working fluid, which motivates the focus of this research. $SCO_2$ refers to $CO_2$ in a supercritical state. A fluid enters the supercritical state once it passes its critical point. The critical point is a combination of a specific value for pressure and temperature. For $CO_2$, the critical point is at $P_c$ = 7.38 MPa, $T_c$ = 304.2 K. Within this supercritical state, the fluid can be in a fluid-like state or a gas-like state depending on the temperature and pressure. When $CO_2$ is at supercritical pressure in the fluid-like state and is heated, it transitions to the gas-like state. During this transition, $CO_2$ passes the pseudocritical point. The pseudocritical point is a point above the critical pressure at pseudocritical temperature. Around the pseudocritical point extreme variations in thermophysical properties occur [42]. Figure 1.2 shows the variations in the properties around the pseudocritical point. This nonlinear behavior can result in heat transfer detoriation (HTD) or heat transfer enhancement (HTE). HTD describes a situation where the heat transfer coefficient at the wall is lower compared to normal heat transfer conditions. This results in higher wall temperatures. Conversely, HTE occurs when the heat transfer coefficient at the wall is higher than normal. This leads to lower wall temperatures [32]. These heat transfer behaviors make it very challenging to model the fluid and lead to overdesign of equipment because a large fault margin must be taken into account. To address the challenge of accurately modeling heat transfer near the pseudocritical point, this research will focus on the prediction of heat transfer to $SCO_2$. In this research, the test case of $SCO_2$ flowing upward within a uniformly heated vertical tube will be considered. The goal is to develop accurate models for predicting heat transfer within the supercritical region, specifically focusing on the heat transfer behavior around the pseudocritical point.

## 1.2. Current methods

In this section, the methods that are currently used to approximate heat transfer will be discussed: Measurements, computational fluid dynamics (CFD), and Nusselt correlations.

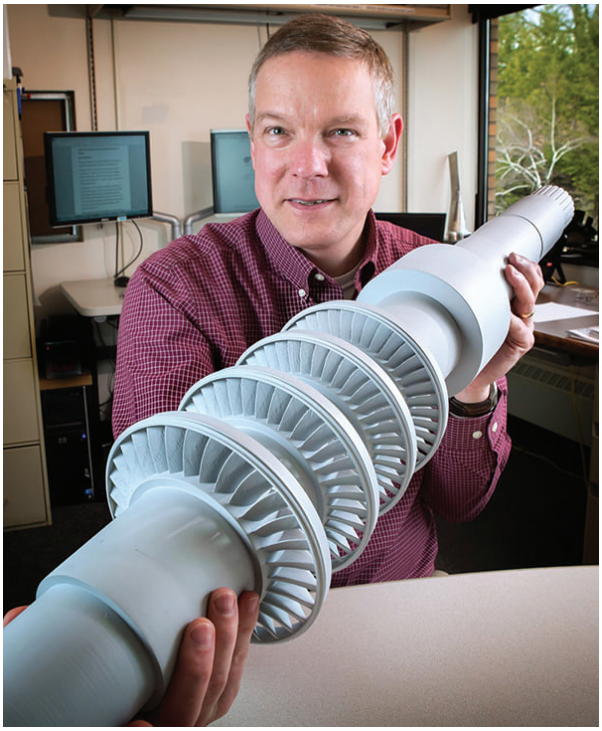


**Figure 1.1:** A 10 MW $SCO_2$ turbine demonstrating the compact design achievable with supercritical $CO_2$. This turbine is 20 times smaller than a steam turbine and can power 10.000 homes [38].

## Measurements

Measurements play an important role in data collection. Generally, temperature, pressure, and mass flow are measured to find the characteristics of a fluid at all locations in the process. Obtaining precise measurements is essential to discover the behavior of $SCO_2$ under various conditions. Table 1.1 lists the typical measurement equipment used in a lab setup to monitor these parameters. Most studies

**Table 1.1:** Typical range and accuracy of measurement equipment in a test assembly [21].

| Instrument | Range | Accuracy |
|---|---|---|
| Thermocouple (K-type) | $0 \sim 1260^\circ$ C | $\pm$ 0.75 % |
| Pressure transmitter | $0 \sim 16$ MPa | $\pm$ 0.25 % |
| Differential pressure transmitter | $0 \sim 58.8$ kPa | $\pm$ 0.055 % |
| Mass flow meter (Coriolis type) | $0 \sim 680$ kg/h | $\pm$ 0.15 % |

choose a tube to measure the $SCO_2$. This experimental setting is relatively simple compared to measurements in a channel or heat exchanger. Figure 1.3 illustrates a typical test assembly used to gather the measurements. While measurements can be used to design equipment, it is not a feasible solution to build a test assembly for every piece of equipment. This limitation shows the need for models and simulations to complement empirical data, allowing for a more efficient design process in $SCO_2$ applications.

## Computational Fluid Dynamics

Fluid flow problems are described by the Navier-Stokes equations. These equations, based on the principles of conservation of mass, momentum, and energy, are too complicated to solve analytically. Computational fluid dynamics (CFD) offers a method for numerically solving fluid flow problems. Three dominant methods are used in CFD modeling: Reynolds-averaged Navier-Stokes (RANS), large eddy simulations (LES), and direct numerical simulations (DNS) each varying in resolution as depicted in Figure 1.4.

- RANS provides the lowest resolution, relying on simplifications and assumptions to achieve high

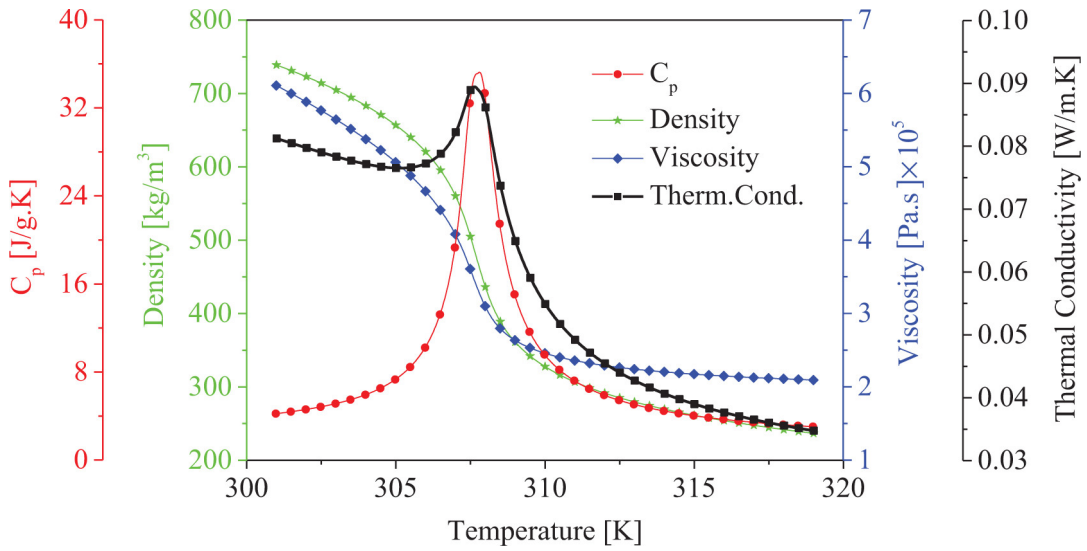



**Figure 1.2:** Thermophysical characteristics around the pseudocritical temperature of $CO_2$ [12].

computational efficiency. RANS can be used to model turbulence in $SCO_2$. However, the current turbulence models developed for RANS can only model isothermal, incompressible fluids accurately [7] [34]. These turbulence models show low accuracy when applied to $SCO_2$ flows, as they are not built to model turbulence within $SCO_2$. Although methods have been developed to improve turbulence models [28], it is still necessary to increase the performance of these models.

- LES offers a compromise between accuracy and computational efficiency by resolving large-scale vortices and modeling smaller scales.
- DNS provides the most detailed, accurate representation of physical processes and a detailed grid by solving all the vortices from the small- to the large-scale. However, this higher resolution comes at the expense of high computational costs, limiting its applicability to flows with high Reynolds numbers.

Existing methods such as LES and DNS are restricted to scenarios with very low Reynolds numbers due to their computational demands [25]. $SCO_2$ flows have high Reynolds numbers limiting the applicability of DNS and LES. While RANS simulations can be applied, their accuracy is insufficient for design of equipment like turbomachinery and heat exchangers. These limitations show the need for alternative methods capable of accurately predicting heat transfer at low computational cost.

## Nusselt correlations

Nusselt correlations offer a straightforward approach to estimating heat transfer by calculation of the Nusselt number. The Nusselt number $Nu$ can be used to determine the heat transfer coefficient $\alpha$ as shown in Equation (1.1).

$$\alpha = \frac{Nuk}{d} = \frac{q_w}{T_w - T_b} \tag{1.1}$$

where $k$ is the thermal conductivity, $d$ is the diameter, $q_w$ is the wall heat flux and, $T_w$ and $T_b$ are the wall- and bulk temperature, respectively. According to research by Pethukov and Kirillov [31], a function $f$ containing the following thermophysical properties is essential to approximate the Nusselt number.

$$Nu = f\left(Re, Pr, \frac{\mu_w}{\mu_b}, \frac{k_w}{k_b}, \frac{c_{p,w}}{c_{p,b}}, \frac{\rho_w}{\rho_b}\right) \tag{1.2}$$

where $Re$ is the Reynolds number, $Pr$ is the Prandtl number, $\mu$ is the dynamic viscosity, $k$ is the thermal conductivity, $c_p$ is the specific heat capacity and $\rho$ is the density. $w$ and $b$ represent the wall and the bulk, respectively. These properties form the basis for most Nusselt correlations used in supercritical heat transfer. The correlations are developed by finding a form of Nusselt correlation that can indicate supercritical heat transfer. Subsequently, the scalars of the correlation are found using line fitting techniques [16]. Unfortunately, as shown in Figures 2.1 and 4.4, the Nusselt correlations are lacking in

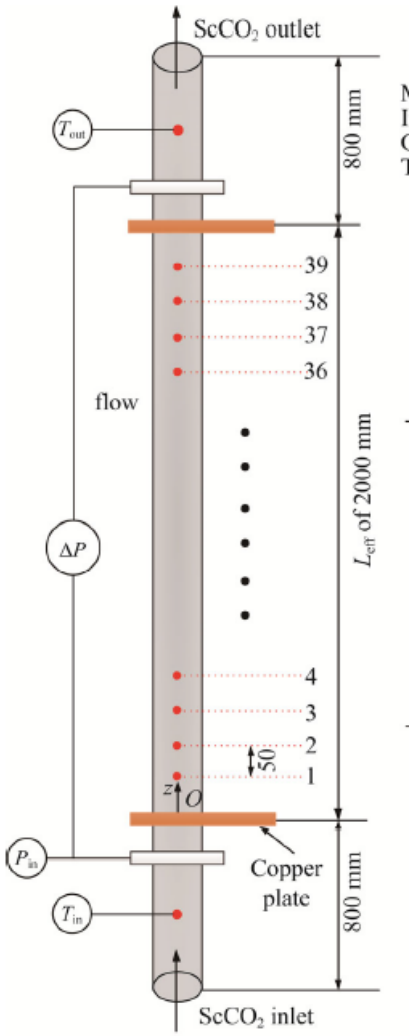



**Figure 1.3:** Vertically heated tube containing $CO_2$ flowing upward [48]. The test assembly shows a thermocouple and a pressure transducer at the inlet and the outlet. In addition, the heated section contains 39 thermocouples to measure the wall temperature. The test assembly uses a Coriolis mass flow meter to measure the mass flow rate.

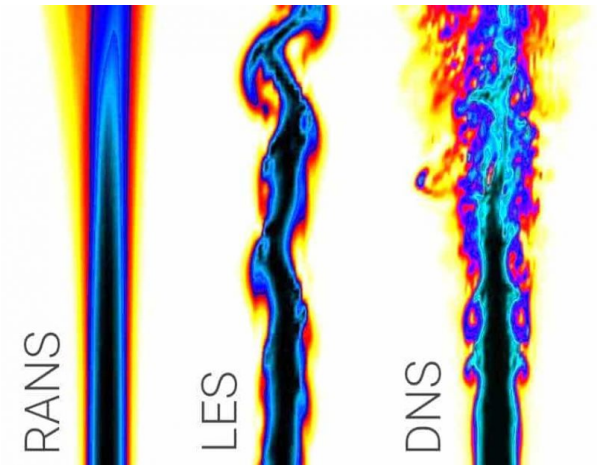



**Figure 1.4:** Comparison of three methods for CFD modeling of turbulent flow [14].

accuracy. Furthermore, Peeters' research [30] shows that a heat exchanger design based on Nusselt correlations requires a significant size increase (up to 42%) to account for modeling inaccuracies and achieve the desired performance. Therefore, it can be concluded that Nusselt correlations can only be used when a large error margin is taken into account, and it also underscores the need for more accurate prediction methods.

## 1.3. Research questions

Considering the limitations of existing methods such as measurements, CFD and Nusselt correlations, a new method is necessary to predict heat transfer. Therefore, machine learning methods will be investigated to find a solution. The method should be sufficiently accurate for reliable equipment design while maintaining computational efficiency. To achieve this goal, the following research questions are formulated:

- What type of model is best suited for predicting heat transfer?
- How can we incorporate spatial context into the prediction of heat transfer?
- How does the proposed model compare to current methods?
- What are the important features for predicting supercritical heat transfer?

## 1.4. Research objectives

This section describes the research objectives that will lead to the answers of the research questions.

The first research objective is to develop a model that incorporates the spatial context of a heat transfer trajectory. The hypothesis is that such a model will lead to better results compared to current methods, such as Nusselt correlations and prediction methods. These methods only consider one point on the trajectory, neglecting valuable information from other locations.

The second research objective focuses on identifying the features that influences HTD and HTE. Therefore, a parameter study will be done to find nondimensional parameters that can explain the heat transfer in a supercritical fluid. Nondimensionalization is crucial for ensuring training convergence, as it brings parameters to a comparable scale. Unlike statistical dimensionality reduction techniques, this approach, based on thermophysical properties, retains the physical meaning of the parameters while reducing their dimensionality.

## 1.5. Contributions

This section describes the contributions of this research, which address the challenge of accurately predicting supercritical heat transfer in $SCO_2$ flows. In this research, a convolutional neural network (CNN) is developed to predict heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated vertical tube. A CNN model is chosen because of its ability to capture the spatial context of the heat transfer trajectory. by leveraging spatial information, the model achieves significantly better performance compared to traditional methods. The developed model achieves a coefficient of determination ($R^2$) of 0.970 and a mean squared error (MSE) of 1477, outperforming existing artificial neural networks (ANN) and Nusselt correlations. The contributions are as follows.

- CNN model: A CNN with its ability to learn spatial context leads to better predictions of heat transfer compared to existing methods. Additionally, the proposed model is applicable to a wider range of parameters than traditional Nusselt correlations, making it a more scalable solution.
- Nondimensional parameter study: A feature importance framework is developed to identify the most relevant nondimensional parameters for CNN training. Through the feature importance framework, a set of nondimensional parameters is selected, $Q^+_{pc}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$, $\frac{K_v}{K_{v,pc}}$, $\frac{Pr_b}{Pr_{pc}}$, $\frac{Re_b}{Re_{pc}}$, which shows high performance in predicting Nusselt numbers.
- Normalization of features: Normalization of features significantly improves the performance of models by reducing the bias and preventing issues such as exploding or vanishing gradients. Normalization contributes to the stability of model predictions and introduces a smoothing effect that results in more consistent outputs. Furthermore, models trained with normalized dimensional features consistently outperform those trained with normalized nondimensional features. This

suggests that nondimensionalization might not be essential for achieving high performance, and using dimensional features with normalization can be a viable approach.

# 2

# Related works

This section describes the various methods employed for prediction of heat transfer. First, applications of condensation heat transfer are investigated. Second, it examines a method for the discovery of parameters. Third, it discusses methods for the application of supercritical transfer.

## 2.1. Condensation heat transfer

The prediction of heat transfer has been a focus of research in various domains. One of these domains is the prediction of condensation heat transfer [47] [17] [27]. A comparative review of recent studies, shown in Table 2.1, reveals a variety of models applied. For example, Zhou et al. [47] report superior results with an artificial neural network (ANN), achieving a mean average error (MAE) of 6.8% and an r-squared score of 0.98, while Hughes et al. [17] apply support vector regression (SVR) and achieve a mean average precision error (MAPE) of 5.0%. Nie et al. [27] demonstrate that a convolutional neural network (CNN) has the highest MARD of 5.82%, but XGBoost, based on gradient boosting, has the highest R-squared score of 0.98.

However, the applicability of these studies is limited. The models are trained on measurements taken from various fluids and geometries. When the data set is narrowed down to a single fluid and geometry, only a few hundred training samples are left. Consequently, comparisons between models become unreliable due to the limited data available for each specific fluid and geometry. For example, an ANN will beat an SVR in performance when provided with enough data, as ANNs contain more trainable parameters. In contrast, an SVR, having fewer trainable parameters, can achieve high accuracy with a substantially smaller data set.

In addition, each study employs unique metrics, which complicates the comparison of these methods. This highlights the need for standardized metrics that are suitable for heat transfer prediction.

## 2.2. Parameter discovery

The study by Scalabrin and Piazza [35] is one of the first to employ neural networks for supercritical heat transfer. The research applies a multilayer feedforward neural network to find the optimal form of the heat transfer correlation. The optimal form is evaluated using the average absolute deviation (AAD).

The study proposes the following feature vector: Reduced pressure $P_r$, reduced temperature $T_r$, mass flow $\dot{m}$ and wall temperature against bulk temperature $\frac{T_w}{T_b}$. This results in an AAD of 2.67% for the proposed method versus 4.30% for a conventional equation defined in the research. The application

**Table 2.1:** Machine learning models used for prediction of condensation heat transfer. RF stands for Random Forest.

| Author | Models | Mediums | Training samples |
|---|---|---|---|
| Zhou et al. [47] | ANN, AdaBoost, RF, XGBoost | 17 | 4882 |
| Hughes et al. [17] | ANN, SVR, RF | 9 | 4000 |
| Nie et al. [27] | KNN, ANN, CNN, RF, XGBoost | 28 | 6604 |

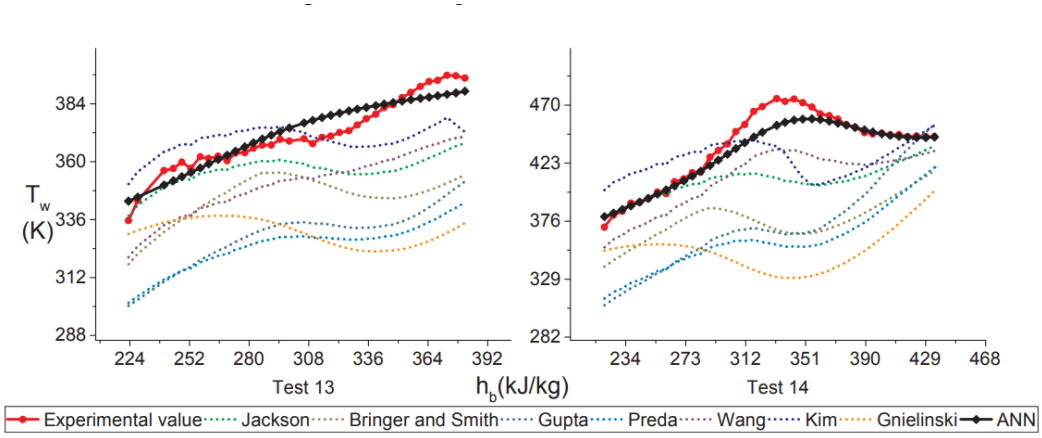



**Figure 2.1:** Prediction of two heat transfer trajectories are shown in the graphs. The ANN of Ye et al. [44] is compared with various correlations and the experimental value.

of the proposed feature vector in this research is not possible, as it contains a wall temperature. In this research, heat transfer needs to be predicted using bulk properties, mass flux, wall heat flux, and pressure, as an iterative method to approximate wall temperature cannot be used.

## 2.3. Supercritical heat transfer

Building on the research by Scalabrin and Piazza [35], several researchers have developed models for the prediction of supercritical heat transfer. The research is listed in Table 2.2. Most of these studies, mentioned in Table 2.2, use pressure, mass flux, heat flux, bulk temperature, and tube diameter to predict wall temperature. All studies employ an ANN to make predictions.

Two studies investigate the prediction of heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated vertical tube [44] [51]. The study by Ye et al. [44] presents graphs in which wall temperature is plotted against bulk enthalpy, demonstrating that their ANN model works better than the classical correlations. However, these graphs also reveal an increase in error as trajectories become more nonlinear, as depicted in Figure 2.1. Additionally, their study excludes trajectories with significant buoyancy and flow acceleration effects, limiting the scope of their method.

The study by Zhu et al. [50] shows that it can accurately track most HTD and HTE cases, except for cases with very steep peaks and troughs. In this research, the model developed by Zhu et al. [50], is applied to predict wall temperature based on the data set. Figures C.6 and C.7 in Appendix C.4 show the prediction of their model on our data set. Although the ANN can follow the linear trends, it fails to predict the HTD and HTE areas.

In conclusion, several issues have been identified with existing approaches to supercritical heat transfer.

- The studies mentioned in Table 2.2 are based solely on information from a single measurement location to predict the wall temperature. This approach does not take advantage of the sequential nature of the data, which could provide contextual information for accurate predictions.
- The studies lack a systematic approach to feature selection. Identifying the most relevant parameters as input features for the prediction model can significantly improve its performance and efficiency.
- Existing research overlooks the potential of applying nondimensional features to improve prediction. These features allow direct input into prediction models, thereby preserving information that might otherwise be lost through normalization processes.

The evaluation of the aforementioned research demonstrates that there is a need to improve the prediction of supercritical heat transfer. In particular, HTD cases should be investigated. Moreover, conducting an in-depth study on features to identify the most effective parameters is also crucial. Finally, research into nondimensional parameters can contribute to the predictive capability of these models.

**Table 2.2:** This table shows the research papers where an ANN is developed for prediction of heat transfer for supercritical fluids. The parameter ranges, the type and amount of data is described.

| Author | $P_r(-)$ | $T_b$(K) | $G$(kg/m$^2$s) | $q_w$(kW/m$^2$) | $d$(mm) | Type of data | Data points |
|---|---|---|---|---|---|---|---|
| Ye et al. [44] | $1.02-1.25$ | $267-388$ | $100-3079$ | $0.479-616.3$ | $2-22$ | Measurements of $SCO_2$ in a heated upward tube | 7313 |
| Zhu et al. [48] | $1.02-2.82$ | 500 | $488-1500$ | $84-350$ | $2-16$ | Measurements of $SCO_2$ in a heated upward tube | 2674 |
| Chu et al. [9] | $1.08-1.19$ | $288-310$ | $33-233$ | $5-30$ | $2-10$ | DNS of $SCO_2$ in a heated horizontal tube | 2100 |
| Chen et al. [6] | $1.06-1.78$ | $290-340$ | $250-850$ | $12-66$ | 10.9 | Measurements of $SCO_2$ in a heated horizontal tube | 1115 |
| Dhanuskodi et al. [11] | $0.99-1.82$ | $348-881$ | $2,083-5170$ | $48.5-290$ | 8.2 | Measurements of SH2O in boiler | 10367 |
| Chang et al. [5] | $1.02-1.13$ | $340-473$ | $700-3500$ | $300-1600$ | $8-20$ | Measurements of SH2O in a heated horizontal tube | 5280 |

# 3
# Method

This chapter describes the method to predict heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated vertical tube. The method involves selecting and preparing the data set followed by a parameter study on nondimensional features. Subsequently. data preprocessing and a feature importance method are presented. A convolutional neural network (CNN) is proposed to predict the Nusselt number. A comparison is made with an artificial neural network (ANN) and Nusselt correlations. Finally, a study on the normalization of parameters is done.

## 3.1. Data selection

This method describes the process of selecting the data. The data is selected based on the following requirements: It should contain measurements or DNS data for $CO_2$ in the supercritical state. Furthermore, the heat transfer should occur for supercritical $CO_2$ flowing upward within a uniformly heated vertical tube.

The data set for this research is obtained from the study by Ye et al. [44]. Their data set contains data from 15 research articles with the specific papers and their parameter ranges presented in Table A.1 in Appendix A. The data set by Ye et al. [44] relies on data captured from figures in the studies mentioned in Appendix A. The extracted data contains pressure, diameter, bulk temperature, mass flux, wall heat flux, wall temperature, heat transfer coefficient, and Nusselt number. In some cases, the location of the measurement, measured from the start of the heated pipe, and the inlet temperature are extracted. The researchers used between 20-40 thermocouples to measure heat transfer in 3-9 m test sections. In all studies, pressure, mass flow, and wall-, bulk-, and inlet temperature are measured. The data set contains 9957 measurements. Table 3.1 shows the parameter ranges for the data set. In addition, Figure A.1 in Appendix A shows the data distribution of the data set.

**Table 3.1:** Parameter ranges of the data set used in this study which is developed by Ye et al. [44].

| Author | $P_r(-)$ | $T_{in}$(K) | $G$(kg/m$^2$s) | $q_w$(kW/m$^2$) | $d$(mm) | Type of data | Data points |
|---|---|---|---|---|---|---|---|
| Ye et al. [44] | $1-1.25$ | $267-388$ | $230-2027$ | $19.6-436.4$ | $4.4-9$ | Measurements | $9957$ |

## 3.2. Data preparation

This section describes the methods used to prepare the data for future use. Initially, heat transfer trajectories are plotted for validation against the corresponding research articles listed in Appendix A, ensuring alignment with the source data. The data set includes thermophysical properties, such as thermal conductivity and dynamic viscosity. These thermophysical properties are verified using CoolProp, a property database, to confirm their accuracy. The heat transfer coefficient and the Nusselt number are recalculated using Equations (3.1) and (3.2), respectively.

$$\alpha = \frac{q_w}{T_w - T_b} \tag{3.1}$$

$$Nu = \frac{\alpha d}{k} \tag{3.2}$$

The data set occasionally lacks measurement locations. In that case, the location of the measurement within the tube is derived from the bulk enthalpy, given in Equation (3.3).

$$x_n = \frac{G(h_{b,n} - h_{b,0})}{4q_w} + x_0 \tag{3.3}$$

where $h_b$ is the bulk enthalpy, $x$ is the location in the pipe, $G$ is the mass flux, $0$ represents the first measurement point and $n$ a specific measurement point. Additionally, thermophysical properties, along with Reynolds and Prandtl numbers, are appended to the data set using CoolProp and relevant formulas. Table 3.2 shows all the parameters available after the data set is enhanced with the aforementioned methods.

**Table 3.2:** Parameters of the data set

| Parameters |
|---|
| $h_b$, $T_b$, $\rho_b$, $\mu_b$, $k_b$, $c_{p_b}$, $Pr_b$, $Re_b$, |
| $h_w$, $T_w$, $\rho_w$, $\mu_w$, $k_w$, $c_{p_w}$ |
| $h_{pc}$, $T_{pc}$, $c_{p_{pc}}$, $k_{pc}$, $\mu_{pc}$, $\rho_{pc}$, $Pr_{pc}$, $Re_{pc}$, |
| $x/d$, $d$, $q_w$, $G$, $T_{out}$, $T_{in}$, $P$, $Nu$, $\alpha$. |

## 3.3. Scaling of thermophysical properties

In this section, the thermophysical properties are scaled to improve learning efficiency of the model and prevent issues such as exploding or vanishing gradients [15]. This process involves scaling the properties to make them nondimensional, a strategy chosen over the use of normalization to avoid the loss of vital thermophysical information that could lead to inaccurate model predictions. Most properties are scaled using the pseudocritical values of the heat transfer trajectory. The physical location within the system is scaled using the pipe diameter, allowing for a standardized reference across different system sizes.

More advanced nondimensional parameters are used for the wall heat flux and mass flux. Both parameters have a large influence on the heat transfer behavior and are therefore important to scale correctly. Table 3.3 shows the formulas for these nondimensional parameters, and the motivation for the parameters is given below:

- The nondimensional heat flux, $q^+$, serves as a scaling option for the wall heat flux. The formula consists of the wall heat flux, the specific heat capacity, the bulk temperature, and the mass flux. Variations in the specific heat capacity and temperature along the axial direction provide information on whether HTD or HTE occurs and can help predict the Nusselt number. Therefore, these parameters are selected to scale the wall heat flux.
- The nondimensional heat flux, $\frac{q_w}{G \cdot h_b}$, of the Nusselt correlation by Yang et al. [43] is chosen because it scales the wall heat flux with a length aspect of the tube, the bulk enthalpy. Equation (3.3) shows how the bulk enthalpy is correlated with the length of a tube. When the wall heat flux is scaled with the bulk enthalpy, the heat flux becomes a function of the location in the pipe.
- The Froude number, $Fr$, is also incorporated to scale the mass flux, characterizing the ratio of inertia to gravity. This number can provide information on flow instability or relaminarization effects resulting from buoyancy-driven flows around the pseudocritical point. The buoyancy-driven flows can result from density differences near the wall that are influenced by gravity and can lead to HTD or HTE [29].
- Liu et al. [24] state that a flow acceleration parameter, $K_v$, can be used to indicate the transition from turbulent to laminar flow. This parameter contains the wall heat flux and the mass flux, both of which need to be scaled.
- The local heat flux parameter, $Q^+$, is developed by McEligot et al. [26] to make the wall heat flux variable with its location by scaling it with local bulk properties. This formula should account for the transition from turbulent to laminar flow according to the research.

Combining these parameters gives insight into the behavior of HTD and HTE, which might help the algorithm predict both situations. All of these parameters are added to the data set to expand the options of parameters from which to learn.

**Table 3.3:** Nondimensional parameters and their formulas. These parameters are added to the data set.

| Author | Name | Formula |
|---|---|---|
| Liu et al. [24] | Nondimensional heat flux ($q^+$) | $q^+ = \frac{q_w}{c_{p,b} \cdot T_b \cdot G}$ |
| Yang et al. [43] | Nondimensional heat flux with length aspect | $\frac{q_w}{G \cdot h_b}$ |
| Nie et al. [27] | Froude number ($\mathrm{Fr}_{pc}$) | $\mathrm{Fr}_{pc} = \frac{G^2}{g \cdot \rho_{pc}^2 \cdot d}$ |
| Liu et al. [24] | Flow acceleration parameter ($K_v$) | $K_v = 4\frac{q^+}{\mathrm{Re}_b}$ |
| McEligot et al. [26] | Local heat flux parameter ($Q^+$) | $Q^+ = \frac{q_w d}{2 k_b T_b}$ |

## 3.4. Data preprocessing

The data set contains varying lengths of heat transfer trajectories. However, CNNs require input data of uniform dimensions, as they are based on fixed-size layers and filters to apply convolutions across the input data set. Therefore, an algorithm must be designed that brings these trajectories to the same size. Three methods are tested to find the best performing method:

- Segmentation and padding: This method involves splitting a heat transfer trajectory when it exceeds a predetermined length and appending zeros to the end of the trajectory if it falls short of the required length.
- Overlapped segments with interpolation and padding: This method addresses the challenge of preserving sequential relationships, opposing the potential loss of sequential information when the data is split. By introducing overlaps between segments, this method ensures that sequential data from the heat transfer trajectories is not lost. If the trajectories are too short, interpolation is applied before creating the overlapped trajectories of a data set.
- Segmentation and interpolation: This method increases the length of the trajectories by interpolation, rather than padding. If trajectories are exceeding the maximum length, they are split. This approach ensures that the trajectories maintain their physical properties throughout their entire length, potentially offering a more accurate representation of the data for model training.

These methods ensure that all heat transfer trajectories, regardless of their original size, are transformed into a uniform format suitable for CNN training. The length of the trajectories is set at 38 as 77% of the trajectories have that length. The other trajectories are transformed using one of the aforementioned methods.

## 3.5. Feature importance

This section employs various methods to assess feature importance, thereby distinguishing influential features from noise. These methods span from statistical techniques to model-based approaches, offering insight into feature importance. In the following, we describe the methods utilized in the feature importance analysis.

### Correlation analysis

Correlation analysis involves the calculation of pairwise correlations between each feature and the output variable utilizing the Pearson correlation coefficient. This coefficient measures the linear correlation between these two vectors. As both vectors are normalized, the coefficient can have a value between -1 and 1. The following shows what the coefficient means:

- **1** implies a perfect positive relationship: as one variable increases, the other variable increases by a proportionate amount.
- **-1** implies a perfect negative relationship: As one variable increases, the other variable decreases by a proportionate amount.
- **0** implies no linear relationship: There is no linear trend between the variables.

The formula to calculate the Pearson correlation coefficient $r$ between two variables, $X$ and $Y$, is given by:

$$r = \frac{\sum (X - \overline{X})(Y - \overline{Y})}{\sqrt{\sum (X - \overline{X})^2 \sum (Y - \overline{Y})^2}} \tag{3.4}$$

where $X$ and $Y$ are the individual sample points, and $\overline{X}$ and $\overline{Y}$ are the means of the variables$X$ and $Y$, respectively. The closer a feature is to 1, the higher the correlation with the target value. When a feature has a score of 0 there is no correlation to be found. The feature can also be negatively correlated. Again, the closer the feature is to -1, the higher the correlation.

The Pearson correlation is not an ideal method as it only measures the linear correlation, whereas our data set also contains a lot of nonlinear parts. However, it is used as a first insight to assess whether there is any correlation between input features and the Nusselt number.

### Tree-based models evaluation

Tree-based models, such as Random Forest, Gradient Boosting Machines (GBM), and XGBoost, offer insight into feature importance. The feature importance can be assessed from how features are used to split data across trees and their contribution to improving model performance. The steps we take are as follows. First, a random forest regression (RF) model is trained [4]. Next, the feature importance is calculated. The process involves evaluating the decrease in node impurity across all trees in the model, adjusted by the probability of reaching that node, which is determined by the proportion of samples passing through the node. Node impurity quantifies how much variation there is in a node. Essentially, features that lead to significant decreases in impurity are considered more important. This method offers a direct metric to assess the value each feature brings to the model's predictions, with higher values indicating features that more effectively split the data, thus improving the model's accuracy. A disadvantage of using this method is that we are using feature importance based on a trained model. When the model is overfitting on the data set and we try to employ feature importance, it may lead to inaccurate results.

### Permutation feature importance

Permutation feature importance is a model-agnostic approach that evaluates the impact of each feature on the performance of the model. The method was first proposed in the work of Breiman et al. [4] for RF. In this work, it is used for both the RF and the CNN. For each feature, the values are randomly shuffled across all samples, which affects the feature-target relationship. The model performance is then recalculated on this perturbed dataset. By randomly shuffling the values of a feature and measuring the change in model performance, this method determines the importance of a feature. The idea is that significant features, when permuted, will lead to a noticeable increase in prediction error, indicating their importance for prediction.

### Lasso Regularization

Least Absolute Shrinkage and Selection Operator Regularization, also known as Lasso Regularization, is a type of linear regression that uses shrinkage [39]. Shrinkage is where data values are shrunk towards a central point, like the mean. The lasso technique encourages simple and sparse models (i.e., models with fewer parameters). The method introduces a penalty term to the loss function equal to the absolute value of the magnitude of the coefficients. This approach can shrink some coefficients to zero, effectively performing feature selection. The objective function of Lasso regression is given by:

$$\min_{\beta} \left\{ \frac{1}{2n} \sum_{i=1}^{n} (y_i - X_i\beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{3.5}$$

where $n$ is the number of samples, $X_i$ are the feature vectors, $y_i$ are the target values, $\beta$ are the coefficients, $p$ is the number of features, and $\lambda$ is the regularization parameter controlling the strength of the penalty.

The intrinsic property of Lasso regularization to perform feature selection can be used for our application. Features that are not important will have coefficients that are reduced to zero, whereas important features will have a larger magnitude coefficient. The coefficients per feature are extracted to show which features are important. Furthermore, the method will shrink the coefficients of features that are highly correlated to zero. The advantage is that features that contribute exactly the same information get a low score. However, the disadvantage is that, as these features will have low coefficients due to similarity, it cannot be assessed whether they contribute to predicting the target. They will get a low coefficient independent of the importance of the feature in relation to the target. Another disadvantage

of Lasso regularization is that it might not capture complex relationships as it prefers to reduce the feature space to a simple representation.

### Input features and output

The feature importance study resulted in an input vector that can predict the Nusselt number. The input vector consists of $Q^+_{pc}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$, $\frac{K_v}{K_{v,pc}}$, $\frac{Pr_b}{Pr_{pc}}$, $\frac{Re_b}{Re_{pc}}$. The Nusselt number is chosen as the output because it is already a nondimensional number.

### Procedure

The following procedure is used to find the best features for CNN training.

1. Correlation Analysis: Initially, features are analyzed for their correlation with the target variable to identify linear relationships.
2. Random Forest Analysis: The feature vector containing all the nondimensional numbers is used to train an RF model.
   (a) Feature importance: The feature importance provided by the RF model is analyzed. This method ranks features based on their contribution to improving model accuracy, highlighting those with the most predictive power.
   (b) Permutation Feature Importance: Following the initial feature importance ranking, permutation feature importance is applied to the trained RF model. This technique evaluates the impact of shuffling each feature on model accuracy, further refining the understanding of feature relevance.
   (c) Lasso regularization: This method helps to identify features that can be eliminated (those with coefficients reduced to zero), thus focusing on the most informative ones.
3. CNN-based analysis: A CNN is trained using the feature vector to further explore feature importance for the CNN. After training, permutation feature importance is applied. This method assesses how the permutation of each feature affects the performance of the model.
4. Cumulative tree-based importance: A cumulative score is compiled from the correlation analysis and the tree-based methods, the cumulative tree-based importance.
5. Feature selection and performance: The seven feature vectors with the highest performance are selected for each method. Subsequently, a baseline model is trained for each method based on this selection. The metrics are reported for each method.
6. Feature refinement: A subsequent round of permutation feature importance identifies the subset of features that contribute significantly to model performance. Top-performing features are selected on the basis of a criterion: Only those with scores of at least 10% of the highest performing feature's score are chosen. The selected features are then used to re-train the baseline model to refine the metrics and identify the best-performing features.

## 3.6. Convolutional neural network

The goal of the model is to predict the Nusselt number based on the properties available through measurements. Traditionally, a CNN is used in computer vision tasks where the kernel is shifted over an image to extract features. This research aims to leverage the ability of CNNs to capture spatial information. The data analyzed in this research, heat transfer trajectories, are of a sequential nature. Predicting the Nusselt number based solely on isolated data points would not fully capture the underlying dynamics of the process. Therefore, a CNN is proposed to capture this information.

### Architecture

The core architecture of the CNN comprises seven convolutional layers. The kernels of the convolutional layers slide over the input data, applying filters to capture patterns like heat transfer detoriation, heat transfer enhancement, or other relevant characteristics in segments of the data. Fully-connected layers are used to convert these extracted features into the desired output, the Nusselt number. Figure 3.1 shows the architecture of the algorithm. The architecture is found through a hyperparameter study. The dimensions of the input layer can be explained as follows. The length of the trajectory is 38. The

width of the trajectory is one, as we only have measurements in the axial direction of the tube. The last dimension, 7, stands for the amount of input features. All the convolutional layers have a length of 38 and a width of 1. The last dimension of the convolutional layers represents the number of kernels that are used. The last dimension in the fully-connected layer stands for the amount of neurons. Furthermore, the kernels are all one-dimensional, with the first dimension giving the length of the kernel.

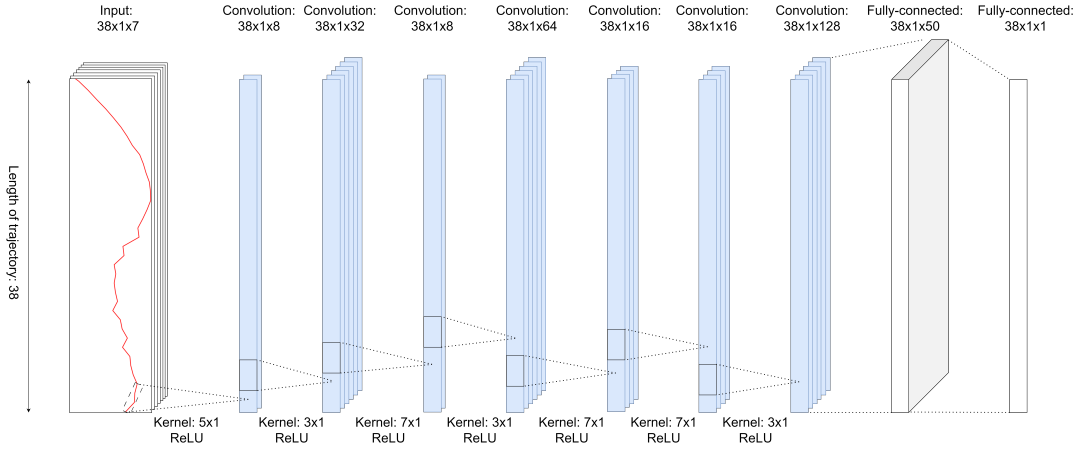



**Figure 3.1:** Architecture of the proposed convolutional neural network

Pooling layers are usually applied in a CNN. They can be used to down sample the data so that the data is a summarized version of the original input. It results in a model being invariant to local translation, reduces overfitting, and speeds up computation [15]. However, we aim to learn from the spatial information and, therefore, would like to keep the full size of the data. Furthermore, our data set is relatively small, thus speed up of the model is not necessary. On a more practical note, we want to output the same trajectory size as the input size. Therefore, the reduction in size is undesirable.

### Activation function

Rectified Linear Activation Unit (ReLU) is used as activation function. The function returns the input value if it is greater than zero and returns a zero if it is less than zero. Most models today use the ReLU activation function due to its easy calculation. Furthermore, it is more robust to noise and overfitting compared to the Sigmoid or Tanh function. Another advantage is that a ReLU can introduce sparsity into the model by setting neurons to zero. In practice, this means that kernels specialized in heat transfer enhancement are activated only when such behavior occurs. A disadvantage of the ReLU function is that when the weights in the model lead to negative input into a ReLU function, the ReLU function will output zero, also known as the ”dying ReLU” problem [10]. Once the ReLU only outputs zero, the weights will not be updated. However, this output only occurs when all inputs become zero. If only one input is nonzero. The weight will be updated, and the ReLU activation function will stay alive. Using the Adam optimizer, which adapts the learning rate, the dying ReLU problem is also limited, as it becomes only severe once the learning rate is high. Small learning rates can prevent the gradients from driving ReLU activations below zero, thus reducing the likelihood of ”dying ReLUs”.

### Bias initializer

Bias of the output layer is initialized using the mean of the training output as proposed by Karpathy [19]. This approach is intended to start off the predictions of the model in the statistical context of the data, accelerating the convergence.

### Loss function

The Mean Squared Error (MSE) has been selected as a loss function. This function is chosen because it punishes large errors more severely. In the case of supercritical heat transfer, the model will have difficulty predicting when HTE or HTD occurs because this behavior comes with large variations in the parameters. A loss function that penalizes these errors in the large variations more severely will likely have a positive effect on the prediction capability. The MSE has a limitation to evaluate the Nusselt

number predictions. First, MSE offers limited interpretability, providing only a single value that does not reveal how the model performs across the data set. Secondly, MSE is sensitive to the scale of the Nusselt number. If the average value is high, errors for those high values will dominate the MSE even if they are proportionally similar to errors for lower values. This behavior can lead to prioritization of accurate prediction of high Nusselt numbers at the expense of lower ones.

### Optimizer

The model uses an Adam optimizer that is a combination of AdaGrad and RMSProp developed by Kingma et al. [23]. The optimizer adjusts the learning rate for each parameter by calculating the moving averages of the gradients and the squared gradient to adjust the learning rate. The Adam optimizer leads to faster and more stable convergence with minimal hyperparameter tuning. A disadvantage of an Adam optimizer is that it is susceptible to noise. Through the hyperparameter study, an initial learning rate of 0.001 is found for the Adam optimizer.

## 3.7. Nusselt correlations

Nine Nusselt correlations are selected, specifically designed for the approximation of heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated vertical tube. These correlations are used to compare performance. The Nusselt correlations used in this research are shown in Table B.1 in Appendix B. Most correlations have some form of Equation (1.2). Their parameter ranges are shown in Table 3.4.

All Nusselt correlations require an iterative method to find the wall temperature. Therefore, the iterative method, developed by Peeters [29], is used to converge on the wall temperature. The iterative method works as follows: Starting with an initial range of wall temperature between a lower bound, bulk temperature, and an arbitrarily chosen upper bound temperature, the method iteratively narrows down this range to find the temperature that satisfies a certain threshold.

**Table 3.4:** Parameter Ranges for Nusselt Correlations

| Author | Pressure [MPa] | Mass Flux [kg/m$^2$s] | Wall Heat Flux [kW/m$^2$] | Diameter [mm] |
|---|---|---|---|---|
| Gupta [16] | 7.4 - 8.8 | 706 - 3169 | 9.3 - 616.6 | 8 |
| Kim [22] | 7.75, 8.12, 8.85 | 400 - 1200 | 0 - 150 | 4.4 |
| Jackson [18] | 7.85, 9.81 | 1100 - 7500 | 900 - 2600 | 4.4 |
| Preda [33] | 7.58 - 9.58 | 419 - 1200 | 20 - 130 | 0.948 - 9 |
| Pioro [32] | 8.38 - 8.8 | 700 - 3200 | 18.4 - 161.2 | 8 |
| Wang [41] | 7.4 - 10.3 | 102 - 3400 | 2.6 - 1053 | 0.95 - 29 |
| Yang [43] | 7.6, 8.4, 8.8 | 900 - 3000 | 0 - 600 | 8 |
| Zhang [46] | 7.5 - 10.5 | 50 - 200 | 5 - 60 | 16 |
| Zhu [49] | 7.52 - 21 | 315 - 2000 | 18.4 - 893 | 2 - 26 |

# 4

# Experiments

In this chapter, the experiments are described. The CNN is evaluated for prediction of heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated vertical tube. The experiments are structured to incrementally build upon the findings of the preceding ones, encompassing:

- Development of a baseline model to serve as a reference point
- Comparison of three data pre-processing methods
- Feature importance study
- Hyperparameter search
- Evaluation of the impact of various model configurations through ablation studies
- Benchmarking the proposed CNN against ANN and Nusselt correlations
- Influence of data normalization on model performance

Each experiment follows a structured format. Initially, the aim of the experiment is described, followed by a detailed description of the methodologies employed. Finally, the results are reported along with a discussion of the experimental findings.

Two metrics are used for the evaluation of the developed models, the MSE and the $R^2$ score, also known as the coefficient of determination. MSE is chosen because it gives us insight on the performance of the model on the outliers in the data, the HTE and HTD areas. The $R^2$ score is defined as the amount of variance explained by the prediction model. The $R^2$ score measures the difference between the predicted values and the data. A higher $R^2$ score means that the variance is better explained. What should be taken into account when interpreting the $R^2$ value is that it does not account for bias. Therefore, the metric is combined with MSE to assess model performance.

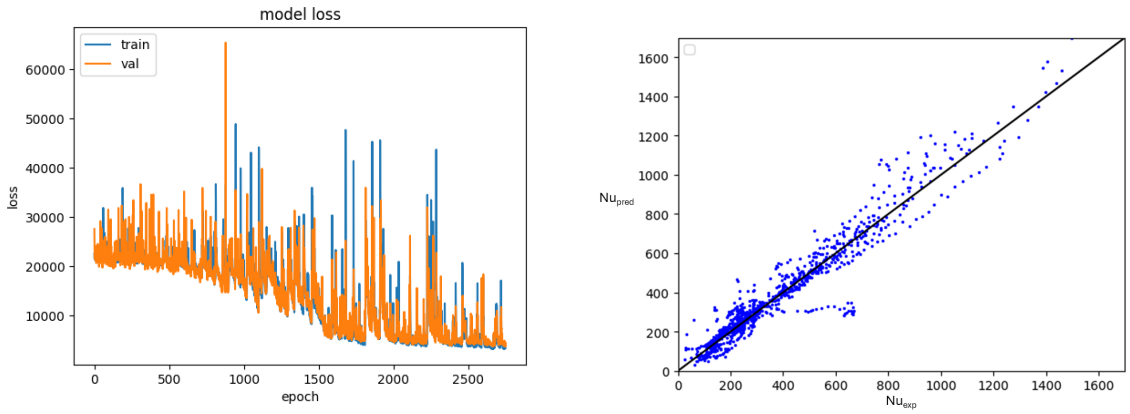
The data set is divided into three parts: training, validation and testing, following an 80-10-10 percentage split. Training and validation data sets are used for model training, while the test data set is reserved for qualitative and quantitative evaluations of the model performance. A random state 42 is chosen to ensure the same split at all times.

## 4.1. Experiment 1: Baseline model

The primary objective of this experiment is to create a simple and effective model that serves as a reference point to evaluate the performance of more complex models in future experiments. Establishing this baseline is crucial to providing a clear benchmark against which the effectiveness of other methods can be measured, ensuring that any improvements can be quantitatively assessed. The baseline method is found through trial and error by varying the number of layers until acceptable results were found.

### Methodology

In this experiment, a CNN is developed to be used as a baseline model. The baseline consists of three 1D convolutional layers, each with 32 filters, and two 1D convolutional layers with 64 filters each. Each layer contains a kernel size of 3 and ReLU as an activation function. The last two layers consist




**(a)** Loss curve

**(b)** Experimental Nusselt number vs. predicted Nusselt number

**Figure 4.1:** Baseline model performance

of two fully-connected layers with 50 and 1 neurons, respectively. MSE is chosen as a loss function. As metrics, $R^2$ and MSE are selected. The model is set at 3000 epochs and a batch size of 32. In addition, it uses an early stopping function to prevent overfitting. A bias initializer is incorporated in the last fully-connected layer to ensure fast convergence.

The thermophysical properties of the Nusselt correlations, given in Equation (1.2), are used as inspiration for the feature vector. The wall heat flux and the mass flux are also added to the feature vector. The properties are made nondimensional using thermophysical properties at pseudocritical temperature. The resulting feature vector contains the following features: $\frac{T_b}{T_{pc}}$, $\frac{\mu_b}{\mu_{pc}}$, $\frac{k_b}{k_{pc}}$, $\frac{\rho_b}{\rho_{pc}}$, $q^+_{pc}$, $Fr_{pc}$. The goal is to predict the Nusselt number. The data is prepared using the segmentation and interpolation method described in Section 3.4.

## Results

Table 4.1 shows the performance of the baseline model in terms of MSE and $R^2$ scores.

**Table 4.1:** Performance of the baseline model on the train, validation and test data set.

| Type | MSE (train-val-test) | $R^2$ (train-val-test) |
|---|---|---|
| Baseline model | 3310 3643 5479 | 0.934 0.916 0.896 |

Furthermore, Figure 4.1 presents two visualizations. First, it displays the training loss curves of the baseline model, providing insight into the learning process of the model over time. Second, it provides a scatter plot of the predicted Nusselt numbers versus the experimental Nusselt numbers, highlighting the accuracy of the model in the range of the test data set. Figure C.1 in Appendix C.1 shows the heat transfer trajectories, providing qualitative information on the overall performance of the model.

## Discussion

Analyzing the data quantitatively, Table 4.1 reveals that the model generalizes effectively between the train and the validation data. This generalization is indicated by the marginal increase between the training and validation MSE. However, the model shows poor generalizability to the test data. This behavior is indicated by the significantly higher test MSE. Furthermore, the decrease in $R^2$ scores from training through validation to the test data set suggests that model generalization to unseen data is inadequate. Therefore, a hyperparameter study will be carried out in experiment 4.4 to increase the generalizability of the model.

The loss curves in Figure 4.1a decrease during training. A decrease in training loss indicates that the model can learn and adapt to the training data. The loss curve also shows that the train and validation losses decrease at the same rate and stay close to each other, indicating that the model is not overfitting. However, the presence of large fluctuations within the loss curves points to a potential instability in the learning process. The fluctuations indicate that the batch size or the learning rate should be modified. At the end of the training, the loss curves level off, indicating that the model has

reached its learning capacity. In the hyperparameter study in Section 4.4, the batch size and learning rate of the model will be modified. Also, the learning capacity of the model will be increased.

Figure 4.1b shows a comparison between the predicted and experimental Nusselt numbers. Generally, the predictions are close to the experimental Nusselt numbers, with minor discrepancies. However, a significant deviation is observed around the location: $Nu_{ex}$ = 600 where $Nu_{pred}$ = 350. Moreover, the greatest discrepancies are predominantly seen in the lower range of Nusselt numbers. Therefore, these areas should be the focus of improvement during training.

Figure C.1 presents the predictions of the Nusselt number for the test data set. The model is able to capture the linear trend in the heat transfer trajectories. However, the model performs poorly in the troughs and peaks associated with HTD and HTE, respectively. Therefore, these discrepancies need to be further examined in subsequent experiments.

## 4.2. Experiment 2: Data preprocessing methods

The CNN needs trajectories of the same length. The heat transfer trajectories in the data set are characterized by trajectories of varying lengths. Therefore, in this experiment, three data preprocessing methods are evaluated that standardize the trajectories.

### Methodology

In this section, three methods are described to standardize the data set to a specified trajectory length. The three methods are evaluated by training the CNN baseline model. The method with the lowest MSE and highest $R^2$ score will be employed in subsequent experiments. The following methods are applied:

- Segmentation and padding
- Overlapped segments with interpolation and padding
- Segmentation and interpolation

### Results

Table 4.2 shows the performance of the three methods.

**Table 4.2:** Performance of the baseline model trained with three different preprocessing methods. The MSE and $R^2$ score are used as metrics.

| Category | MSE (train-val-test) | $R^2$ (train-val-test) |
|---|---|---|
| Segmentation and padding | 1101 1733 **2461** | 0.978 0.955 **0.956** |
| Overlapped segments with interpolation and padding | 554 1311 3607 | 0.983 0.958 0.813 |
| Segmentation and interpolation | 3093 3671 6296 | 0.940 0.917 0.888 |

### Discussion

The MSE loss and the $R^2$ score of the overlapped segments with interpolation and padding shows low performance. Although the train MSE is very low, the model is overfitting on the test data, as indicated by the significantly higher test MSE. Overfitting is due to overrepresentation in certain parts of the heat transfer trajectory as a result of the overlapping sections. Segmenting and padding the data gives the highest performance as indicated by the MSE and $R^2$ score. This performance demonstrates that zero padding gives better results than using interpolation to fill the missing values in the heat transfer trajectories. Again, interpolation leads to overrepresentation in certain areas, resulting in a negative effect on the generalizability of the model as shown by the MSE and $R^2$ results. Therefore, zero-padding will be used in the hyperparameter search to optimize the model.

## 4.3. Experiment 3: Feature importance

This experiment focuses on evaluating the feature importance of the nondimensional parameters within the data set. The aim is to identify the most significant features that increase the accuracy of the model in predicting heat transfer.

## Methodology

In this section, the method is outlined to find the best-performing feature vector. The nondimensional numbers, described in Section 3.3, are evaluated in this feature importance study. The procedure applied in this experiment is described in detail in Section 3.5. An overview of the procedure is provided below:

1. Correlation analysis
2. Random forest analysis
   (a) Feature importance
   (b) Permutation feature importance
   (c) Lasso regularization
3. CNN-based analysis
4. Cumulative tree-based importance
5. Feature selection and performance
6. Feature refinement

## Results

Figure C.5 in Appendix C.3 shows the results of correlation analysis, random forest analysis, CNN-based analysis and cumulative tree-based importance as described in Steps 1-4. The figure shows a bar plot with compounded normalized scores per feature. Table 4.3 presents the feature vectors and their performance as described in Step 5. The permutation feature importance, applied in Step 5, is

**Table 4.3:** Performance and description of the feature vectors selected after employing feature importance methods

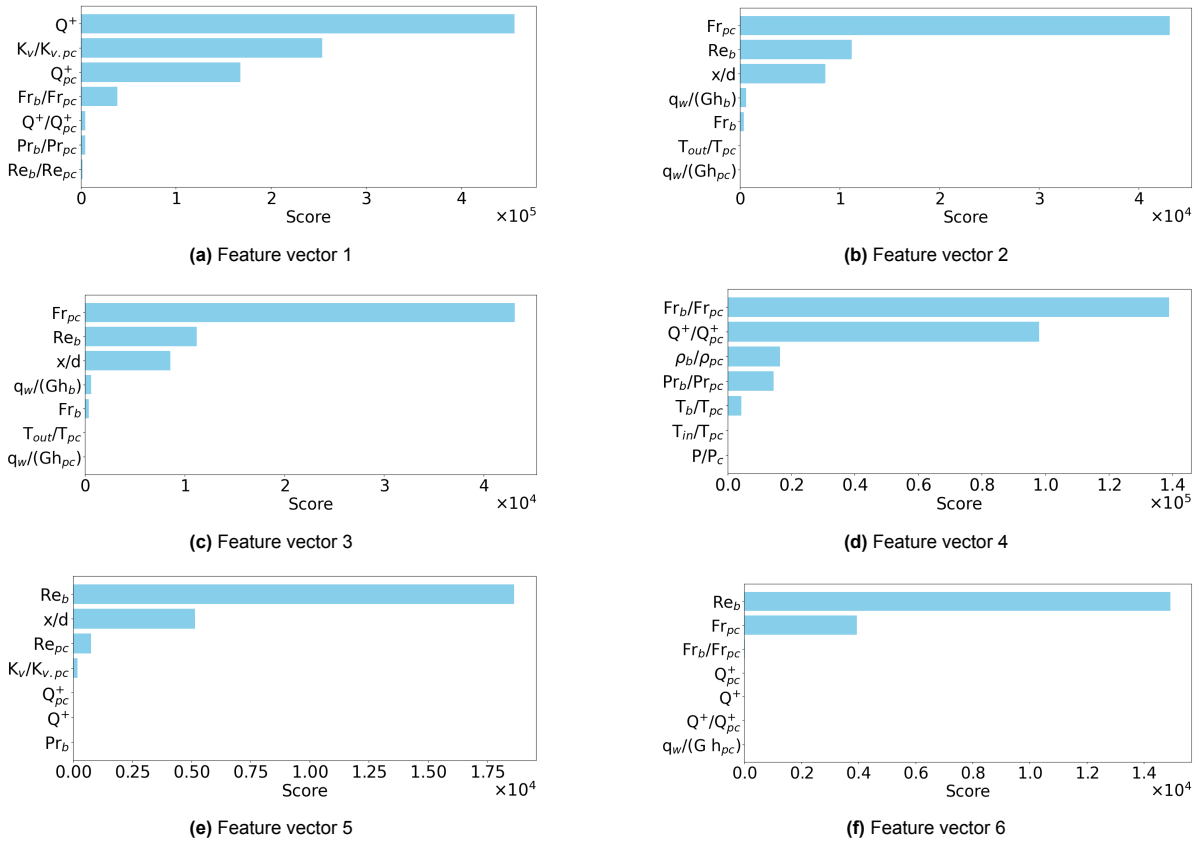
| Vector | Features | MSE | $R^2$ | Feature importance method |
|---|---|---|---|---|
| 1 | $Q^+_{pc}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$, $\frac{K_v}{K_{v,pc}}$, $\frac{Pr_b}{Pr_{pc}}$, $\frac{Re_b}{Re_{pc}}$ | **5483** | **0.892** | Correlation analysis |
| 2 | $Fr_{pc}$, $Re_b$, $\frac{q_w}{Gh_{pc}}$, $\frac{T_{out}}{T_{pc}}$, $Fr_b$, $\frac{T_b}{T_{pc}}$, $\frac{q_w}{Gh_b}$ | 26 529 | 0.660 | Feature importance |
| 3 | $Fr_{pc}$, $Re_b$, $\frac{q_w}{Gh_{pc}}$, $\frac{T_{out}}{T_{pc}}$, $Fr_b$, $\frac{x}{d}$, $\frac{q_w}{Gh_b}$ | 27 816 | 0.645 | Permutation feature importance |
| 4 | $\frac{P}{P_c}$, $\frac{\rho_b}{\rho_{pc}}$, $\frac{T_b}{T_{pc}}$, $\frac{T_b}{T_{pc}}$, $\frac{Pr_b}{Pr_{pc}}$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$ | 58 390 | 0.131 | Lasso regularization |
| 5 | $Re_{pc}$, $Re_b$, $\frac{x}{d}$, $Q^+$, $Q^+_{pc}$, $Pr_b$, $\frac{K_v}{K_{v,pc}}$ | 46 283 | 0.261 | CNN permutation feature importance |
| 6 | $Fr_{pc}$, $Re_b$, $Q^+_{pc}$, $\frac{q_w}{Gh_{pc}}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$ | 49 081 | 0.220 | Cumulative tree-based importance |

shown in Figure 4.2. The bar plots show a score per feature per category. The permutation feature importance was applied after training the baseline model with the selected feature vectors from Step 1-4.

Table 4.4 shows the performance of the CNNs after applying the 10% threshold as described in Step 6.

## Discussion

This experiment provides insight into the most critical features. Figure C.5 shows the cumulative score of all features. The pseudocritical Froude number has the highest cumulative score. Additionally, the Reynolds numbers at bulk- and pseudocritical temperatures achieve a high score for the permutation




**(a)** Feature vector 1

**(b)** Feature vector 2

**(c)** Feature vector 3

**(d)** Feature vector 4

**(e)** Feature vector 5

**(f)** Feature vector 6

**Figure 4.2:** Evaluation of features as described in Step 5

feature importance. The local heat flux parameter, $Q^+$, for bulk and pseudocritical properties is of importance in correlation analysis. Lasso regularization scores almost all features the same, with reduced pressure, $\frac{P}{P_c}$, being an exception, scoring significantly higher than other features. As the reduced pressure, $\frac{P}{P_c}$, alone cannot be used to predict the Nusselt number, this shows that Lasso regularization gives us limited information and therefore is not the best feature importance method.

Table 4.3 shows the selected features and their performance after Steps 1-4. Feature vector 1 performs best by a large margin. Figure 4.2 shows the permutation feature importance scores after prediction with the trained CNN as described in Step 5. The Froude number, $Fr$, the Reynolds number, $Re$, the local heat flux parameter, $Q^+$, and the nondimensional location, $x/d$, are the parameters that score high for the permutation feature importance. Feature vector 9 in Table 4.4 emphasizes the essence of the Reynolds number and the Froude number in combination with the nondimensional location. Comparing Feature vector 9 with Feature vector 8 shows that only the addition of the nondimensional location significantly increases the prediction performance.

Following the feature importance study, Feature vector 1 will be chosen because it has the lowest MSE. The importance of the features are described below.

- The importance of the Froude number, $Fr$, can be attributed to the fact that it scales the mass flux to characterize the ratio of inertia to gravity, providing information on the flow instability or relaminarization as a result of buoyancy-driven flows under pseudocritical conditions. Therefore, this number can help to find areas where HTD or HTE might develop.
- The Reynolds number, $Re$, gives insight into the amount of turbulence in a fluid. Turbulence is directly related to the amount of heat that can be transferred from one medium to another and therefore offers insight into HTE, HTD or normal heat transfer.
- The local nondimensional heat flux, $Q^+$, was developed to scale the wall heat flux, $q_w$, with the diameter, $d$, and make it variable with its location through the local bulk properties. The parameter gives us insight into the heat flux at the wall compared to heat transfer at a bulk location. The

**Table 4.4:** Performance and description of the feature vectors selected as described in Step 6

| Vector | Features | MSE | $R^2$ | feature importance method |
|---|---|---|---|---|
| 7 | $Q^+_{pc}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{K_v}{K_{v,pc}}$ | 6445 | 0.887 | Correlation analysis |
| 8 | $Fr_{pc}$, $Re_b$ | 44 285 | 0.281 | Feature importance |
| 9 | $Fr_{pc}$, $Re_b$, $\frac{x}{d}$ | **5517** | **0.908** | Permutation feature importance |
| 10 | $\frac{\rho_b}{\rho_{pc}}$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$ | 58 108 | 0.130 | Lasso Regularization |
| 11 | $Re_b$, $\frac{x}{d}$ | 43 387 | 0.298 | CNN permutation feature importance |
| 12 | $Fr_{pc}$, $Re_b$ | 48 391 | 0.275 | Cumulative tree-based importance |

scaling with the diameter allows for a better comparison of the wall heat flux between the heat transfer trajectories. Local bulk properties will help to indicate the amount of heat flow to $CO_2$.

- The Prandtl number, $Pr$, provides insight into the speed with which the thermal energy diffuses compared to the momentum of the fluid.
- The flow acceleration parameter, $K_v$, compares the nondimensional heat flux with the Reynolds number. The nondimensional heat flux, $q^+$, provides insight into the influence of the heat flux on the local bulk parameters. The Reynolds number is a measure of the potential for turbulent flow. Together, these parameters can indicate the potential for heat transfer rate.

In addition, Five out of the seven nondimensional parameters are normalized using their corresponding values at the pseudocritical temperature of the fluid. The scaling brings the parameters to a value around one. This makes it much easier for the CNN to predict, as there is fewer chances of exploding or vanishing gradients. $Q^+$ and $Q^+_{pc}$ are included because these parameters have a magnitude close to one. This scale also makes it easier for the network to make predictions.

In conclusion, correlation analysis is a sufficient method to use as the feature importance method and Feature vector 1 will be used for further training and optimization of the CNN.

## 4.4. Experiment 4: Hyperparameter search

This experiment aims to find the optimal hyperparameters for the CNN. Hyperparameter search involves searching through a range of hyperparameter values to find the most optimal configuration for the CNN.

### Methodology

A random search is used to find the optimal combination of hyperparameters. Random search is selected because it is computationally more efficient than a grid search in which all combinations of hyperparameters are tested. Random search chooses the set of hyperparameters at random from a selection of hyperparameters to find the optimal configuration. The model is based on the baseline model from Section 4.1. The following components are modified to find more optimal hyperparameters:

1. One-dimensional convolutional layer: 3, 5, 7
2. Number of filters: 8, 16, 32, 64, 128
3. Kernel size: 3, 5, 7
4. Learning rate: $10^{-1}$, $10^{-2}$, $10^{-3}$, $10^{-4}$

The other components from the model are kept as is. The model is trained for 3000 epochs. To prevent overfitting and minimize training time, early stopping is incorporated with a patience of 200. In addition, the batch size is optimized by running the optimal configuration for batch size of 8, 16, 32, 64, 128.

### Results

After running the hyperparameter search, the following hyperparameter selection is found:

1. One-dimensional convolutional layers: 7
2. Number of filters: 8-32-8-64-16-16-128

3. Kernel size: 5-3-7-3-7-7-3
4. Learning rate: $10^{-2}$

Table 4.5 shows the performance of the CNN with optimal hyperparameters.

**Table 4.5:** MSE and $R^2$ values of model resulting from the hyperparameter search

| Type | MSE (train-val-test) | $R^2$ (train-val-test) |
|---|---|---|
| Optimal model | $1999$ $3039$ $\mathbf{3523}$ | $0.963$ $0.933$ $\mathbf{0.925}$ |

### Discussion

The hyperparameter search shows that the deepest model with 7 convolutional layers is capable of achieving the best performance. This performance can be attributed to the ability of a deeper model to uncover more complex features, resulting in better predictive performance.

In the architecture of the optimized model, an alternating amount of kernel size can be observed. Smaller kernels can capture more localized features, while larger kernels can capture more general features. By alternating the kernel size, the receptive field is varied and can result in capturing both global and local patterns, which can be important for the predictive capability. Third, the smallest batch size gives the best performance in terms of MSE and $R^2$ score.

## 4.5. Experiment 5: Ablation studies

Ablation studies are performed to understand the contribution of each component or hyperparameter to the overall performance of a model. By systematically removing or changing parts of the model, the impact on performance can be observed.

### Methodology

The following parameters will be altered to validate the performance of our method

1. Loss function: MAE
2. Layers: +2,-2 convolutional layers
3. Normalization of basic feature vector as seen in the studies in Section 2.3
4. Normalization of the nondimensional feature vector

### Results

Table 4.6 shows the performance of the methods in this ablation study.

**Table 4.6:** MSE and $R^2$ performance of the ablation studies

| Method | MSE | $R^2$ |
|---|---|---|
| Nondimensional feature vector | $3523$ | $0.925$ |
| MAE | $4408$ | $0.919$ |
| +2 convolutional layers | $6289$ | $0.892$ |
| -2 convolutional layers | $5875$ | $0.893$ |
| Normalized basic feature vector | $\mathbf{1477}$ | $\mathbf{0.970}$ |
| Normalized nondimensional feature vector | $2808$ | $0.942$ |

### Discussion

The ablation study shows the performance of the methods against the proposed model trained with the nondimensional feature vector. Applying MAE as a loss function leads to lower performance, as it fails to heavily penalize outliers. Specifically, MAE does not adequately address the sparse regions associated with HTE and HTD. Consequently, when the model follows the linear trend and only deviates significantly during HTE and HTD, such errors are not significantly penalized, hindering the ability to learn from these critical behaviors. Furthermore, increasing or decreasing the number of convolutional layers decreases the performance, proving that the model has the right number of layers.

The performance of the normalized feature vectors is a significant finding. Both perform much better than the nondimensional feature vector extracted from the experiment 4.3. As the values in the nondimensional feature vector have very different sizes, it can lead to exploding or vanishing gradients [15]. Normalization of the nondimensional feature vector can mitigate this problem.

## 4.6. Experiment 6: Benchmarking CNN against ANN and Nusselt correlations

The model in this research is compared to the ANN of Zhu et al. [48]. This model is selected because it has relatively high performance for HTD and HTE and is developed to predict heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated tube. The model in this research is also compared with nine Nusselt correlations.

### Methodology

The model proposed by Zhu et al. [48] has the following characteristics:

- Two hidden layers with each 40 nodes
- Activation function: Hyperbolic tangent
- Optimizer: ADAM with learning rate = 0.01
- Batch size: 32
- Training iterations: 100000

The following features are used as input: Pressure, mass flux, heat flux, bulk enthalpy, diameter. The model attempts to predict the wall temperature. Both the input and output vectors are normalized between zero and one. The model is trained for 100000 iterations. Early stopping is also implemented to reduce overfitting. As the output of the model is the wall temperature, the Nusselt number must be calculated to compare it with the CNN. CNN is trained on two feature vectors. The nondimensional feature vector, presented in Section 3.5, contains $Q^+_{pc}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$, $\frac{K_v}{K_{v,pc}}$, $\frac{Pr_b}{Pr_{pc}}$, $\frac{Re_b}{Re_{pc}}$. The normalized basic feature vector, presented in the ablation studies in Section 4.5, contains $P$, $T_b$, $G$, $q_w$, $d$. Furthermore, nine Nusselt correlations are tested on the data set to validate that the model performs better than the correlations. Most of these correlations also have specific parameter ranges, which limits the applicability to a subset of the data set.

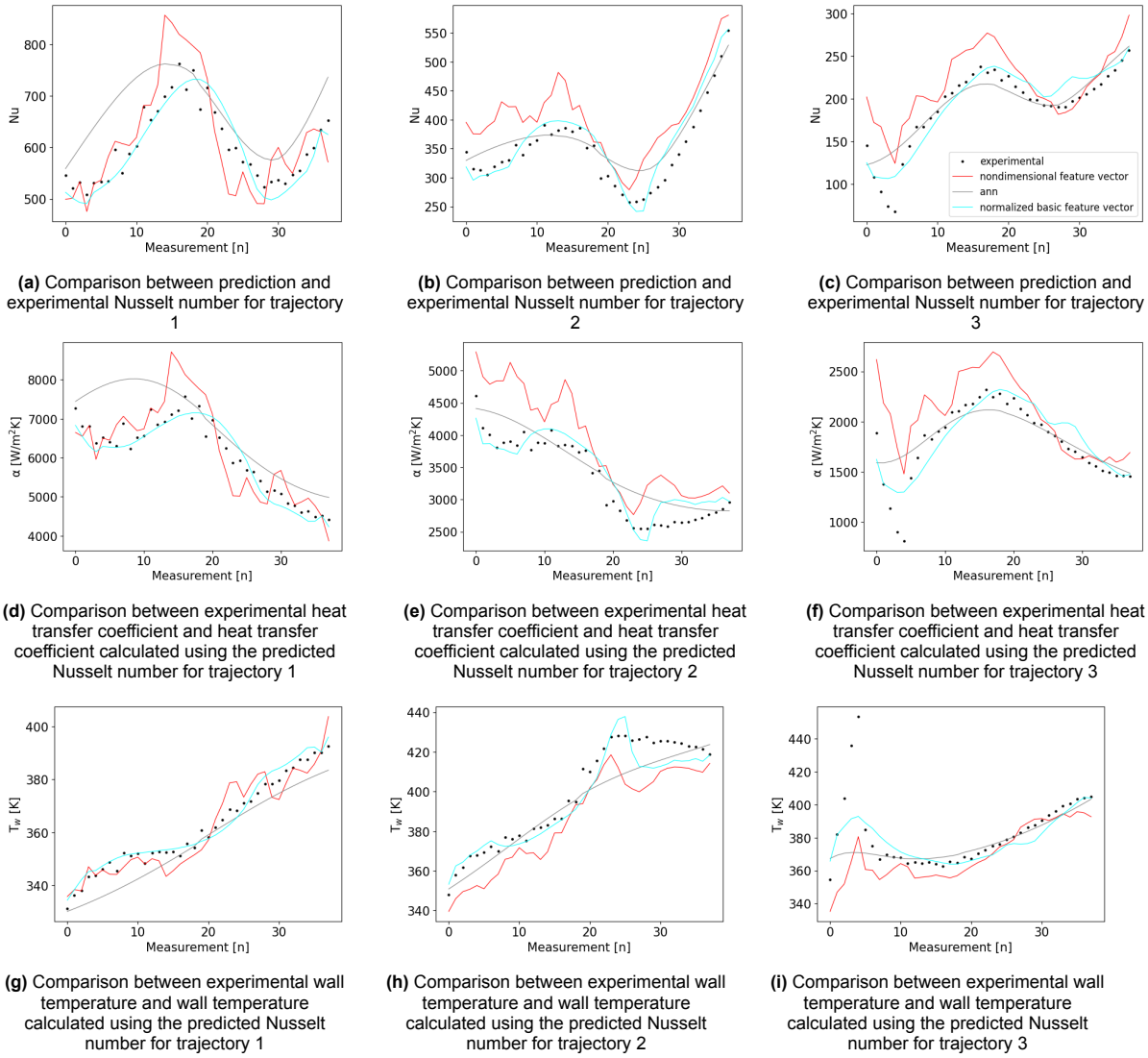
### Results

Table 4.7 shows the performance of the ANN developed by Zhu et al. [50] compared to the methods developed in this research. The graphs in Figure 4.3 show the trajectories taken from the test data set

**Table 4.7:** Metrics for Nusselt number, wall temperature and heat transfer coefficient of the ANN compared to the model developed in this research

| Type | Property | MSE | $R^2$ |
|---|---|---|---|
| ANN [50] | $Nu$ | 23 176 | 0.691 |
| | $T_w$ | 53 | 0.95 |
| | $\alpha$ | 2 840 545 | 0.691 |
| CNN trained with a nondimensional feature vector | $Nu$ | 3523 | 0.925 |
| | $T_w$ | 69 | 0.987 |
| | $\alpha$ | 690 650 | 0.924 |
| CNN trained with a normalized basic feature vector | $Nu$ | **1506** | **0.970** |
| | $T_w$ | **39** | **0.993** |
| | $\alpha$ | **323 922** | **0.966** |

to evaluate the HTD for the ANN and the proposed model and method. Furthermore, the performance of the Nusselt correlations is shown in Table 4.8. The other correlations are not included as they have parameter ranges outside the test data set. In addition, Figure 4.4 shows an overview of the performance of the Nusselt correlation. Finally, Figure C.8 in Appendix C.4 shows the iterative method




**(a)** Comparison between prediction and experimental Nusselt number for trajectory 1

**(b)** Comparison between prediction and experimental Nusselt number for trajectory 2

**(c)** Comparison between prediction and experimental Nusselt number for trajectory 3

**(d)** Comparison between experimental heat transfer coefficient and heat transfer coefficient calculated using the predicted Nusselt number for trajectory 1

**(e)** Comparison between experimental heat transfer coefficient and heat transfer coefficient calculated using the predicted Nusselt number for trajectory 2

**(f)** Comparison between experimental heat transfer coefficient and heat transfer coefficient calculated using the predicted Nusselt number for trajectory 3

**(g)** Comparison between experimental wall temperature and wall temperature calculated using the predicted Nusselt number for trajectory 1

**(h)** Comparison between experimental wall temperature and wall temperature calculated using the predicted Nusselt number for trajectory 2

**(i)** Comparison between experimental wall temperature and wall temperature calculated using the predicted Nusselt number for trajectory 3

**Figure 4.3:** Evaluation of the performance of the ANN for heat transfer detoriation trajectories

from Peeters' [29] in comparison to directly inputting the wall temperature into the Nusselt correlation. Direct input of the wall temperature into the Nusselt correlation can be seen as the ground truth for the iterative method. Figure C.9 in Appendix C.4 shows the correlation calculation by Wang et al. [41] compared to the calculation by Zahlan et al. [45].

## Discussion

First, performance is compared in Table 4.7 for the Nusselt number, wall temperature and heat transfer coefficient.

- Performance on predicting the Nuselt number shows that the ANN from Zhu et al. [50] has a low performance for the Nusselt number compared to both CNNs.
- Interestingly, for wall temperature predictions, the CNN trained with the nondimensional feature vector is initially outperformed by the ANN. However, when a normalized basic feature vector is used for training, the proposed model surpasses the ANN, showcasing the importance of normalizing the features.
- CNN model trained with the normalized basic feature vector performs best on the heat transfer coefficient. The model trained with the nondimensional feature vector has an error twice as large as the model trained with the normalized basic feature vector. However, the error magnitude of

**Table 4.8:** MSE and $R^2$ performance of Nusselt correlation

| Method | MSE | $R^2$ |
|---|---|---|
| Nondimensional feature vector | **3523** | **0.925** |
| Kim et al. [22] | 4886 | 0.521 |
| Preda et al. [33] | 180 816 | 0.540 |
| Pioro et al. [32] | 48 443 | 0.520 |
| Wang et al. [40] | 47 705 | 0.505 |
| Zhu et al. [48] | 10 625 | 0.643 |

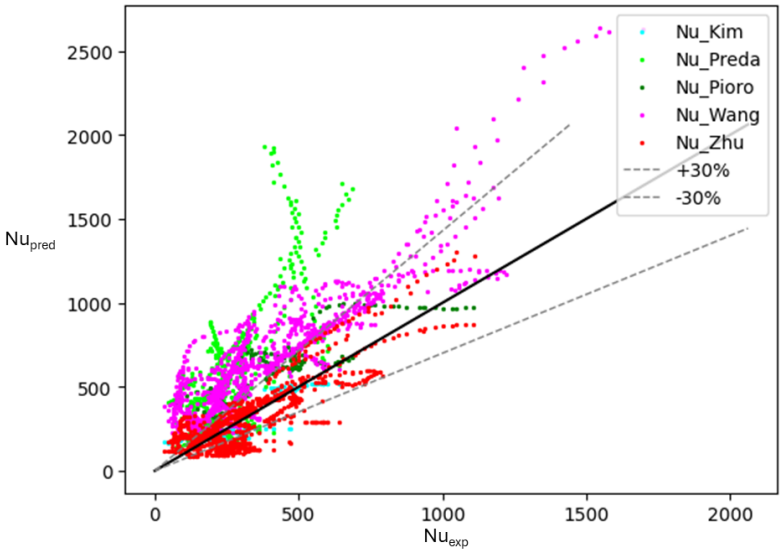



**Figure 4.4:** Experimental Nusselt numbers against predicted Nusselt number for all correlations that fall within test parameter ranges

the ANN is nearly tenfold larger than the error of the normalized basic feature vector. Therefore, it can be concluded that the ANN model does not translate well to the heat transfer coefficient.

The qualitative result can be observed in Figure 4.3. From the qualitative analysis, the following can be concluded:

- The ANN captures linear trends effectively, but has difficulty accurately capturing the troughs and peaks associated with HTD and HTE. This behavior is particularly visible in the prediction of wall temperatures, where, despite the high accuracy in linear fits, it fails to mimic the peaks and troughs present in the experimental data. This observation indicates a lack of complexity within the ANN architecture. Enhancing the ANN by adding more layers could introduce the required complexity, or alternatively, employing a CNN could offer a better solution to capture the complexity within the measurements.
- Although the ANN can mimic the global trend of the data, it is unable to accurately recognize local trends. On the contrary, the CNN model demonstrates its ability to recognize global and local trends in the data. Figure 4.3f shows that the CNN recognizes the HTD in the trajectory, although the predicted magnitude in the trough shows some deviation. However, the ANN does not recognize the HTD in that area. In Figure 4.3d the ANN predicts a trajectory opposite to the heat transfer trajectory in the first HTD area. This prediction shows that the ANN is limited in predicting these local trends in the data. The reason CNN predicts local trends better is due to the use of kernels. The kernels enable CNN to comprehend the spatial context, allowing for a more accurate representation of the data. The multiple kernels per layer allow for learning various features of the data. This advantage is exhibited by its consistent alignment with data trends, despite occasional deviations in magnitude.

Figure 4.4 displays 5 correlations. The correlations from Kim et al. [22] and Zhu et al. [48] have the highest performance with many of the data points within the 30% range. Furthermore, Table 4.8 shows the MSE and $R^2$ scores of the correlations. The MSE of those correlations is relatively low compared to the other Nusselt correlations. In general, the correlations perform worse than the proposed method. Furthermore, it should be noted that the correlations can only approximate the Nusselt number in a specified parameter range, while the proposed method covers all the correlations.

Figure C.8 in Appendix C.4 shows the comparison between the iterative method and the direct calculation. The iterative method can follow the linear trend of the correlation. However, it is not exact as intended by the correlations in the nonlinear parts. Furthermore, Figure C.9 in Appendix C.4 shows the correlation calculation by Wang et al. [41] compared to the calculation by Zahlan et al. [45]. The wall temperature is accurately calculated using the iterative method. Unfortunately, the calculation of the heat transfer coefficient is less accurate. The results show that the iterative algorithm should be refined to improve the convergence of wall temperature.

In conclusion, the CNN can predict heat transfer more accurately than the ANN. Especially if normalized features are used. In addition, the CNN also shows higher performance than the Nusselt correlations. However, it should be noted that the iterative method is not as accurate as it should be.

## 4.7. Experiment 7: Impact of feature normalization on model performance

A final experiment is performed to compare the performance of three different feature vectors. The CNN will be trained on three different feature vectors. Normalizing the features leads to more stable results, as is shown in the ablation study in Section 4.5. Therefore, three feature vectors are investigated to learn more about the influence of normalization.

### Methodology

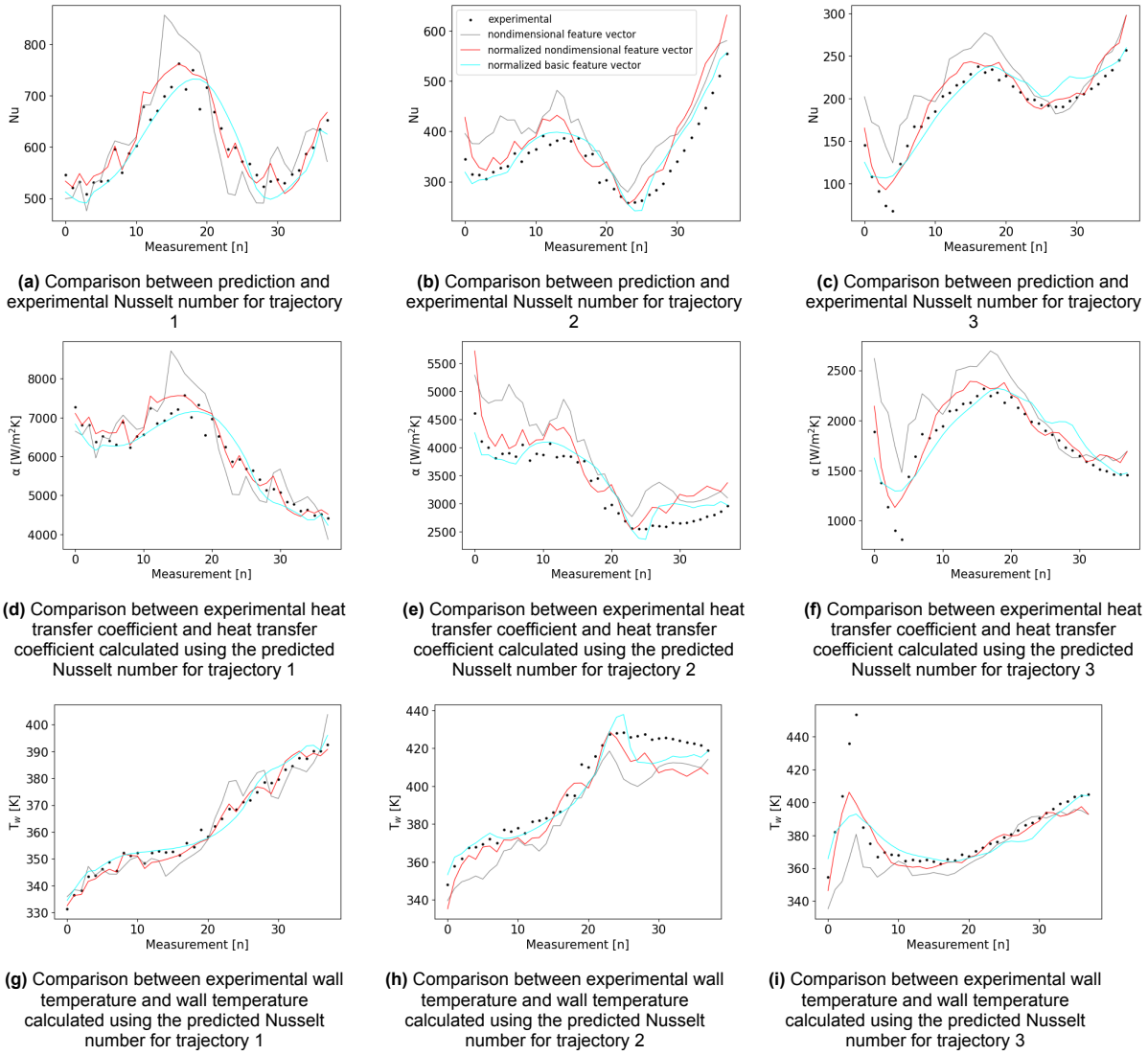
The proposed model, presented in Section 3.6, will be trained on three feature vectors. The following feature vectors are compared:

- Nondimensional feature vector, presented in Section 3.5
  - $Q^+_{pc}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$, $\frac{K_v}{K_{v,pc}}$, $\frac{Pr_b}{Pr_{pc}}$, $\frac{Re_b}{Re_{pc}}$
- Normalized nondimensional feature vector
  - $Q^+_{pc}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$, $\frac{K_v}{K_{v,pc}}$, $\frac{Pr_b}{Pr_{pc}}$, $\frac{Re_b}{Re_{pc}}$
- Normalized basic feature vector as proposed in Section 4.5
  - $T_b$, $G$, $q_w$, $d$, $P$

Normalization is achieved by subtracting the mean and dividing by the standard deviation, transforming all inputs to have a zero mean and a standard deviation of 1. The wall temperature and heat transfer coefficient will also be calculated to see how the predicted Nusselt numbers translate into these properties. Finally, three trajectories that contain heat transfer detoriation will be compared.

### Results

Table 4.9 shows the performance of the proposed model using three different feature vectors to predict the Nusselt number. Additionally, the wall temperature and heat transfer coefficient are calculated to show the performance of those properties.




**(a)** Comparison between prediction and experimental Nusselt number for trajectory 1

**(b)** Comparison between prediction and experimental Nusselt number for trajectory 2

**(c)** Comparison between prediction and experimental Nusselt number for trajectory 3

**(d)** Comparison between experimental heat transfer coefficient and heat transfer coefficient calculated using the predicted Nusselt number for trajectory 1

**(e)** Comparison between experimental heat transfer coefficient and heat transfer coefficient calculated using the predicted Nusselt number for trajectory 2

**(f)** Comparison between experimental heat transfer coefficient and heat transfer coefficient calculated using the predicted Nusselt number for trajectory 3

**(g)** Comparison between experimental wall temperature and wall temperature calculated using the predicted Nusselt number for trajectory 1

**(h)** Comparison between experimental wall temperature and wall temperature calculated using the predicted Nusselt number for trajectory 2

**(i)** Comparison between experimental wall temperature and wall temperature calculated using the predicted Nusselt number for trajectory 3

**Figure 4.5:** Evaluation prediction of the properties for heat transfer detoriation

**Table 4.9:** Metrics for Nusselt number, wall temperature and heat transfer coefficient which are predicted and calculated using three different feature vectors

| Type | Property | MSE | $R^2$ |
|---|---|---|---|
| CNN trained with nondimensional feature vector | $Nu$ | 3523 | 0.925 |
| | $T_w$ | 69 | 0.987 |
| | $\alpha$ | 690 650 | 0.924 |
| CNN trained with normalized nondimensional feature vector | $Nu$ | 2808 | 0.942 |
| | $T_w$ | 47 | 0.991 |
| | $\alpha$ | 577 420 | 0.934 |
| CNN trained with normalized basic feature vector | $Nu$ | **1506** | **0.970** |
| | $T_w$ | **39** | **0.993** |
| | $\alpha$ | **323 922** | **0.966** |

Figure 4.5 shows the trajectories selected from the test data set to evaluate the HTD.

## Discussion

Figure 4.5a, 4.5b and 4.5c illustrate the prediction of the Nusselt number for three distinct trajectories. The normalized basic feature vector leads to the highest accuracy across the three trajectories. The normalized basic feature vector closely matches the experimental Nusselt numbers, exhibiting minimal deviation except in Figure 4.5c, where it encounters difficulty matching a trough. In addition, the trajectory of the predictions is quite smooth. These results suggest that the normalized basic feature vectors leads to models with lower bias and variance compared to other feature vectors. The normalized nondimensional feature vector has slightly lower performance in all three trajectories compared to the normalized basic feature vector. The model trained with the normalized nondimensional feature vector seems to be more reactive towards the variations in the heat transfer trajectory. Between measurement 0 and 10 in Figure 4.5a, the prediction follows the variation in the measurement data quite accurately, while the normalized basic feature vector follows the trend. This behavior might not be desirable as it indicates that the model is overfitting. Figure 4.5c shows that the normalized nondimensional feature vector is better able to capture the trough around measurement 4. The nondimensional feature vector has lowest performance in every trajectory. The nondimensional feature vector has a large overshoot in the cases of a peak. In addition, the predictions show bias in all trajectories. Table 4.9 shows the quantitative performance of the methods. The metrics show that performance improves when normalization is implemented. Comparing the nondimensional feature vector with the normalized nondimensional feature vector shows that performance increases significantly. These results show that normalization has a positive effect on prediction.

Figures 4.5d, 4.5e, and 4.5f illustrate the performance of heat transfer coefficient calculations based on the predicted Nusselt number for the three trajectories. All feature vectors follow the trend. The normalized basic feature vector demonstrates smooth trajectories. The normalized nondimensional feature vector shows overfitting in all the trajectories as it tries to closely fit every measurement. In addition, the normalized nondimensional feature vector seems to perform best for prediction of the HTD in Figure 4.5f. The nondimensional feature vector again has a large overshoot in the HTE peaks and has large bias in all trajectories. In general, the significant dip observed in Figure 4.5f and the prediction of measurement 25 to 35 in Figure 4.5e is challenging for every feature vector.

Figures 4.5g, 4.5h and 4.5i show the calculated wall temperature using the predicted Nusselt number for the three trajectories. The normalized basic feature vector consistently delivers strong performance across all trajectories, although it encounters a significant error in Figure 4.5h. This error is characterized by a spike and subsequent dip in wall temperature. The normalized nondimensional feature vector closely matches the measurement data in Figures 4.5g and 4.5h. In addition, the normalized nondimensional feature vector predicts the HTD peak in Figure 4.5i best. Again, the nondimensional feature vector presents very noisy results, as is also seen for the Nusselt number and heat transfer coefficient.

In conclusion, the results show that data normalization is essential to achieve a lower bias in the predictions. The nondimensional feature vector, which lacks normalization, shows a large bias, whereas the normalized feature vectors are able to follow the trend in the data much better.

# 5

# Conclusion

$SCO_2$ shows promising applications in power generation and heat pumps due to the high thermal efficiency and compact size of the equipment. However, $SCO_2$ cannot be accurately modeled around the pseudocritical point. These modeling challenges lead to overdesign of equipment using $SCO_2$ as a working fluid. Therefore, this research proposes to predict heat transfer using machine learning. To approach these challenges the research is guided by the following research questions:

- What type of model is best suited for predicting heat transfer?
- How can we incorporate spatial context into the prediction of heat transfer?
- How does the proposed model compare to current methods?
- What are the important features for predicting supercritical heat transfer?

In this research, a model is developed for the prediction of heat transfer to supercritical $CO_2$ flowing upward within a uniformly heated vertical tube. The model uses a CNN architecture to predict heat transfer. The architecture is chosen because it incorporates the spatial context of the heat transfer trajectory to predict the Nusselt number at one location. The developed model achieves an R-squared score of 0.970 and an MSE of 1506, an MAE of 38 for the Nusselt number, and an MAPE of 10.7%. The comparison in Section 4.6 shows that CNN outperforms the ANN of Zhu et al. [50]. This improved performance can be attributed to the fact that CNN can learn local and global trends, while ANN can learn only global trends in the heat transfer trajectory. The ability of CNN to learn these local trends underscores the added value of kernels in a CNN. Furthermore, the presence of multiple kernels within each CNN layer facilitates the extraction of various features from the data. Finally, the model outperforms the Nusselt correlations. Moreover, Nusselt correlations can only be applied in a small parameter range. The proposed model is applicable in a wide range of parameters $P_r$ = 1-1.25, $T_{in}$ = 267-388 K, $q_w$ = 19.6-436.4 kW/m$^2$, $d$ = 4.4-9 mm. This broad applicability makes the proposed model a highly scalable solution.

A nondimensional parameter study is conducted and subsequently a feature importance method is used, leading to the following feature vector: $Q^+_{pc}$, $Q^+$, $\frac{Q^+}{Q^+_{pc}}$, $\frac{Fr_b}{Fr_{pc}}$, $\frac{K_v}{K_{v,pc}}$, $\frac{Pr_b}{Pr_{pc}}$, $\frac{Re_b}{Re_{pc}}$. This feature vector demonstrated superior performance compared to the nondimensional feature vector derived from correlations as mentioned in Section 4.1. Furthermore, the feature importance study shows the critical role of the Reynolds number, $Re_b$, and Froude number, $Fr_{pc}$, in combination with the nondimensional location, $x/d$. Prediction with only these parameters results in relatively high performance.

The ablation studies discussed in Section 4.5 demonstrate that normalization of features significantly improves the performance of the model. Normalization helps improve performance by preventing issues such as bias and exploding or vanishing gradients. The model trained with nondimensional feature vector that is not normalized exhibits erratic and jittery predictions and higher bias. In addition, the model trained with normalized dimensional features outperforms the model trained with normalized nondimensional features, indicating that nondimensionalization is not necessary to achieve high performance.

In future work, the following research is proposed to improve the field of prediction of heat transfer for supercritical fluids.

- This research was based on the use of nondimensional features to predict heat transfer. Nondimensional features were a focus point as it is a way to keep thermophysical information contained, which otherwise is lost when normalization is applied. However, Section 4.7 shows that normalization gives better results. Therefore, normalized dimensional and nondimensional features should be taken into account as well to select the best feature vector. In addition, dropout layers should be added which increase stability and can handle the exploding and vanishing gradients. This addition makes it possible to use a combination of dimensional, nondimensional and normalized features.
- To improve the prediction of very deep troughs and peaks (HTD/HTE), two approaches can be considered. First, a loss function could be specifically designed to penalize these underrepresented regions more heavily. However, the major problem with these areas is that they are represented by only a few data points per trajectory. On top of that, there are also many linear trajectories in the data set that provide no information on those particular cases. This underrepresentation of the HTD and HTE areas in the data set should be addressed. Synthetic data should be sampled for the underrepresented areas. The method of Shi et al. [36] can be applied to generate additional data using Nusselt correlations, but other methods for sampling data should also be investigated.
- The application of machine learning models in the design of heat transfer equipment has not been explored in the existing literature. A study on the applicability of our model can show the benefits for desiging equipment and accelerate the field of applications of $SCO_2$.

# Acknowledgements

I would like to thank Prof.dr. R. Pecnik, Dr.ir. J.W.R. Peeters and Dr. H. Caesar for their valuable guidance and support during the thesis. Further, I would like to thank Dr. Y. Zhang for sharing the data set used in this research.

# References


[1] May 1990. URL: https://en.wikipedia.org/wiki/File:Airplane_vortex_edit.jpg.

[2] Yoon-Yeong Bae, Hwan-Yeol Kim, and Deog-Ji Kang. “Forced and mixed convection heat transfer to supercritical CO2 vertically flowing in a uniformly-heated circular tube”. In: *Experimental Thermal and Fluid Science* 34.8 (2010), pp. 1295–1308.

[3] Yoon-Yeong Bae, Hwan-Yeol Kim, and Tae Ho Yoo. “Effect of a helical wire on mixed convection heat transfer to carbon dioxide in a vertical circular tube at supercritical pressures”. In: *International journal of heat and fluid flow* 32.1 (2011), pp. 340–351.

[4] Leo Breiman. *Arcing the edge*. Tech. rep. Citeseer, 1997.

[5] Wanli Chang et al. “Heat transfer prediction of supercritical water with artificial neural networks”. In: *Applied Thermal Engineering* 131 (2018), pp. 815–824.

[6] Junghui Chen, Kuan-Po Wang, and Ming-Tsai Liang. “Predictions of heat transfer coefficients of supercritical carbon dioxide using the overlapped type of local neural network”. In: *International journal of heat and mass transfer* 48.12 (2005), pp. 2483–2492.

[7] Xu Cheng and Thomas Schulenberg. *Heat transfer at supercritical pressures: literature review and application to an HPLWR*. Vol. 6609. FZKA Karlsruhe, Germany, 2001.

[8] Bong-Hyun Cho, Young-In Kim, and Yoon-Yeong Bae. “PREDICTION OF A HEAT TRANSFER TO CO□ FLOWING IN AN UPWARD PATH AT A SUPERCRITICAL PRESSURE”. In: *Nuclear Engineering and Technology* 41.7 (2009), pp. 907–920.

[9] Xu Chu et al. “A computationally light data-driven approach for heat transfer and hydraulic characteristics modeling of supercritical fluids: From DNS to DNN”. In: *International Journal of Heat and Mass Transfer* 123 (2018), pp. 629–636.

[10] *Convolutional neural networks for visual recognition*. Mar. 2023. URL: https://cs231n.github.io/neural-networks-1/#actfun.

[11] R Dhanuskodi et al. “Artificial Neural Networks model for predicting wall temperature of supercritical boilers”. In: *Applied Thermal Engineering* 90 (2015), pp. 749–753.

[12] Xue Du, Junsheng Hu, and Genglei Xia. “Operation characteristic of supercritical carbon dioxide-cooled reactor system under coordination control scheme”. In: *International Journal of Advanced Robotic Systems* 17.3 (2020), p. 1729881420933833.

[13] *EU legislation control f-gases*. URL: https://climate.ec.europa.eu/eu-action/fluorinated-greenhouse-gases/eu-legislation-control-f-gases_en.

[14] Gestione. “Turbulence models in CFD”. In: *IdealSimulations* (Apr. 2022). URL: https://www.idealsimulations.com/resources/turbulence-models-in-cfd/.

[15] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[16] Sahil Gupta et al. “Developing empirical heat-transfer correlations for supercritical CO2 flowing in vertical bare tubes”. In: *Nuclear Engineering and Design* 261 (2013), pp. 116–131.

[17] Matthew T. Hughes, Brian M. Fronk, and Srinivas Garimella. “Universal condensation heat transfer and pressure drop model and the role of machine learning techniques to improve predictive capabilities”. In: *International Journal of Heat and Mass Transfer* 179 (2021), p. 121712. ISSN: 0017-9310. DOI: https://doi.org/10.1016/j.ijheatmasstransfer.2021.121712. URL: https://www.sciencedirect.com/science/article/pii/S0017931021008188.

[18] JD Jackson. “Progress in Developing an Improved Empirical Heat Transfer Equation for Use in Connection With Advanced Nuclear Reactors Cooled by Water at Supercritical Pressure”. In: vol. Volume 3: Thermal Hydraulics; Current Advanced Reactors: Plant Design, Construction, Workforce and Public Acceptance. International Conference on Nuclear Engineering. July 2009, pp. 807–819. DOI: 10.1115/ICONE17-76022. eprint: https://asmedigitalcollection.asme.org/ICONE/proceedings-pdf/ICONE17/43536/807/2740039/807_1.pdf. URL: https://doi.org/10.1115/ICONE17-76022.

[19] Andrej Karpathy. *A Recipe For Training Neural Networks*. Apr. 2019. URL: https://karpathy.github.io/2019/04/25/recipe/.

[20] Dong Eok Kim and Moo-Hwan Kim. “Experimental investigation of heat transfer in vertical upward and downward supercritical CO2 flow in a circular tube”. In: *International Journal of Heat and Fluid Flow* 32.1 (2011), pp. 176–191.

[21] Hwan Yeol Kim et al. “Heat transfer test in a vertical tube using CO2 at supercritical pressures”. In: *Journal of Nuclear Science and Technology* 44.3 (2007), pp. 285–293.

[22] Hyungrae Kim et al. “Experimental investigation on the heat transfer characteristics in upward flow of supercritical carbon dioxide”. In: *Nuclear technology* 164.1 (2008), pp. 119–129.

[23] Diederik P Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: *arXiv preprint arXiv:1412.6980* (2014).

[24] Shenghui Liu et al. “Improvement of buoyancy and acceleration parameters for forced and mixed convective heat transfer to supercritical fluids flowing in vertical tubes”. In: *International Journal of Heat and Mass Transfer* 106 (2017), pp. 1144–1156. ISSN: 0017-9310. DOI: https://doi.org/10.1016/j.ijheatmasstransfer.2016.10.093. URL: https://www.sciencedirect.com/science/article/pii/S0017931016321615.

[25] Xiaoshu Lü et al. “Improving the energy efficiency of buildings based on fluid dynamics models: A critical review”. In: *Energies* 14.17 (2021), p. 5384.

[26] DM McEligot, CW Coon, and HC Perkins. “Relaminarization in tubes”. In: *International Journal of Heat and Mass Transfer* 13.2 (1970), pp. 431–433.

[27] Feng Nie et al. “A universal correlation for flow condensation heat transfer in horizontal tubes based on machine learning”. In: *International Journal of Thermal Sciences* 184 (2023), p. 107994. ISSN: 1290-0729. DOI: https://doi.org/10.1016/j.ijthermalsci.2022.107994. URL: https://www.sciencedirect.com/science/article/pii/S1290072922005221.

[28] Ashish Patel, Rafael Diez, Rene Pecnik, et al. “Turbulence modelling for flows with strong variations in thermo-physical properties”. In: *International Journal of Heat and Fluid Flow* 73 (2018), pp. 114–123.

[29] Jurriaan Willem Reinier Peeters. “Turbulence and turbulent heat transfer at supercritical pressure”. PhD thesis. 2016.

[30] JWR Peeters. “On the effect of pseudo-condensation on the design and performance of supercritical CO2 gas chillers”. In: *International Journal of Heat and Mass Transfer* 186 (2022), p. 122441.

[31] BS Petukhov and PL Kirillov. “About heat transfer at turbulent fluid flow in tubes”. In: *Thermal Engineering* 4 (1958), pp. 63–68.

[32] Igor Pioro, Sahil Gupta, and Sarah Mokry. “Heat-transfer correlations for supercritical-water and carbon dioxide flowing upward in vertical bare tubes”. In: *Heat Transfer Summer Conference*. Vol. 44779. American Society of Mechanical Engineers. 2012, pp. 421–432.

[33] Tiberiu Preda et al. “Development of a heat transfer correlation for supercritical CO2 based on multiple data sets”. In: *International Conference on Nuclear Engineering*. Vol. 44991. American Society of Mechanical Engineers. 2012, pp. 211–217.

[34] Martin Rohde et al. “A blind, numerical benchmark study on supercritical water heat transfer experiments in a 7-rod bundle”. In: *Journal of Nuclear Engineering and Radiation Science* 2.2 (2016), p. 021012.

[35] G Scalabrin and L Piazza. “Analysis of forced convection heat transfer to supercritical carbon dioxide inside tubes using neural networks”. In: *International Journal of Heat and Mass Transfer* 46.7 (2003), pp. 1139–1154.

[36] Xinhuan Shi et al. “Prediction of supercritical CO2 heat transfer behaviors by combining transfer learning and deep learning based on multi-fidelity data”. In: *International Journal of Heat and Mass Transfer* 218 (2024), p. 124802. ISSN: 0017-9310. DOI: https://doi.org/10.1016/j.ijheatmasstransfer.2023.124802. URL: https://www.sciencedirect.com/science/article/pii/S001793102300947X.

[37] JH Song et al. “Heat transfer characteristics of a supercritical fluid flow in a vertical pipe”. In: *The Journal of Supercritical Fluids* 44.2 (2008), pp. 164–171.

[38] *STEP Demo Benefits: Compact Power Plant Components*. https://www.gti.energy/step-demo/step-demo-benefits/compact-power-plant-components/. [Accessed: 11-04-2023].

[39] Robert Tibshirani. “Regression shrinkage and selection via the lasso”. In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 58.1 (1996), pp. 267–288.

[40] Lei Wang et al. “Convective heat transfer characteristics of supercritical carbon dioxide in vertical miniature tubes of a uniform heating experimental system”. In: *International Journal of Heat and Mass Transfer* 167 (2021), p. 120833.

[41] S Wang, LQ Yuan, and LKH Leung. “Assessment of supercritical heat-transfer correlations against AECL database for tubes”. In: (2010).

[42] K Yamagata et al. “Forced convective heat transfer to supercritical water flowing in tubes”. In: *International journal of heat and mass transfer* 15.12 (1972), pp. 2575–2593.

[43] SK Yang. *Heat transfer modes in supercritical fluids*. Tech. rep. Atomic Energy of Canada Limited, 2012.

[44] Kai Ye et al. “Modeling convective heat transfer of supercritical carbon dioxide using an artificial neural network”. In: *Applied Thermal Engineering* 150 (2019), pp. 686–695.

[45] H Zahlan, D Groeneveld, and S Tavoularis. “Measurements of convective heat transfer to vertical upward flows of CO2 in circular tubes at near-critical and supercritical pressures”. In: *Nuclear Engineering and Design* 289 (2015), pp. 92–107.

[46] Qian Zhang et al. “Special heat transfer characteristics of supercritical CO2 flowing in a vertically-upward tube with low mass flux”. In: *International journal of heat and mass transfer* 122 (2018), pp. 469–482.

[47] Liwei Zhou et al. “Machine learning algorithms to predict flow condensation heat transfer coefficient in mini/micro-channel utilizing universal data”. In: *International Journal of Heat and Mass Transfer* 162 (2020), p. 120351. ISSN: 0017-9310. DOI: https://doi.org/10.1016/j.ijheatmasstransfer.2020.120351. URL: https://www.sciencedirect.com/science/article/pii/S0017931020332877.

[48] Bingguo Zhu et al. “Heat transfer prediction of supercritical carbon dioxide in vertical tube based on artificial neural networks”. In: *Journal of Thermal Science* 30 (2021), pp. 1751–1767.

[49] Bingguo Zhu et al. “The general supercritical heat transfer correlation for vertical up-flow tubes: K number correlation”. In: *International Journal of Heat and Mass Transfer* 148 (2020), p. 119080. ISSN: 0017-9310. DOI: https://doi.org/10.1016/j.ijheatmasstransfer.2019.119080. URL: https://www.sciencedirect.com/science/article/pii/S0017931019342073.

[50] Guangya Zhu, Tao Wen, and Dalin Zhang. “Machine learning based approach for the prediction of flow boiling/condensation heat transfer performance in mini channels with serrated fins”. In: *International Journal of Heat and Mass Transfer* 166 (2021), p. 120783.

[51] Jun-Yan Zhu et al. “Unpaired image-to-image translation using cycle-consistent adversarial networks”. In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2223–2232.

# A

# Data set

**Table A.1:** The data set developed by Ye et al [44] is based on the studies described in this table. Their parameter ranges and data points are stated.

| Dataset | $D$ [mm] | $T_b$[°C] | $P$ [MPa] | $G$ [kg/m$^2$s] | $q_w$ [kW/m$^2$] | Data points |
|---|---|---|---|---|---|---|
| Kim et al. [20] | $4.5$ | $30-115$ | $7.53-9.23$ | $230-874$ | $52.8-231$ | $180$ |
| Bae et al. [3] | $6.32$ | $15-89$ | $7.75$ | $400$ | $30-50$ | $159$ |
| Kim et al. [21] | $4.4$ | $27-110$ | $7.75-8.85$ | $400-1200$ | $30-150$ | $740$ |
| Bae et al. [2] | $6.32$ | $15-79$ | $7.75-8.12$ | $285-600$ | $29.3-70$ | $865$ |
| Cho et al. [[8] | $4.4$ | $28-64$ | $8.12$ | $1200$ | $50-130$ | $72$ |
| Song et al. [37] | $4.4, 9$ | $7-194$ | $8.12$ | $400$ | $30-50$ | $395$ |
| Zahlan et al. [45] | $8$ | $7-82$ | $7.48-8.67$ | $402-2027$ | $19.6-436.4$ | $7485$ |
| total | $4.4-9$ | $7-116$ | $7.48-9.23$ | $230-2027$ | $19.6-436.4$ | $9956$ |

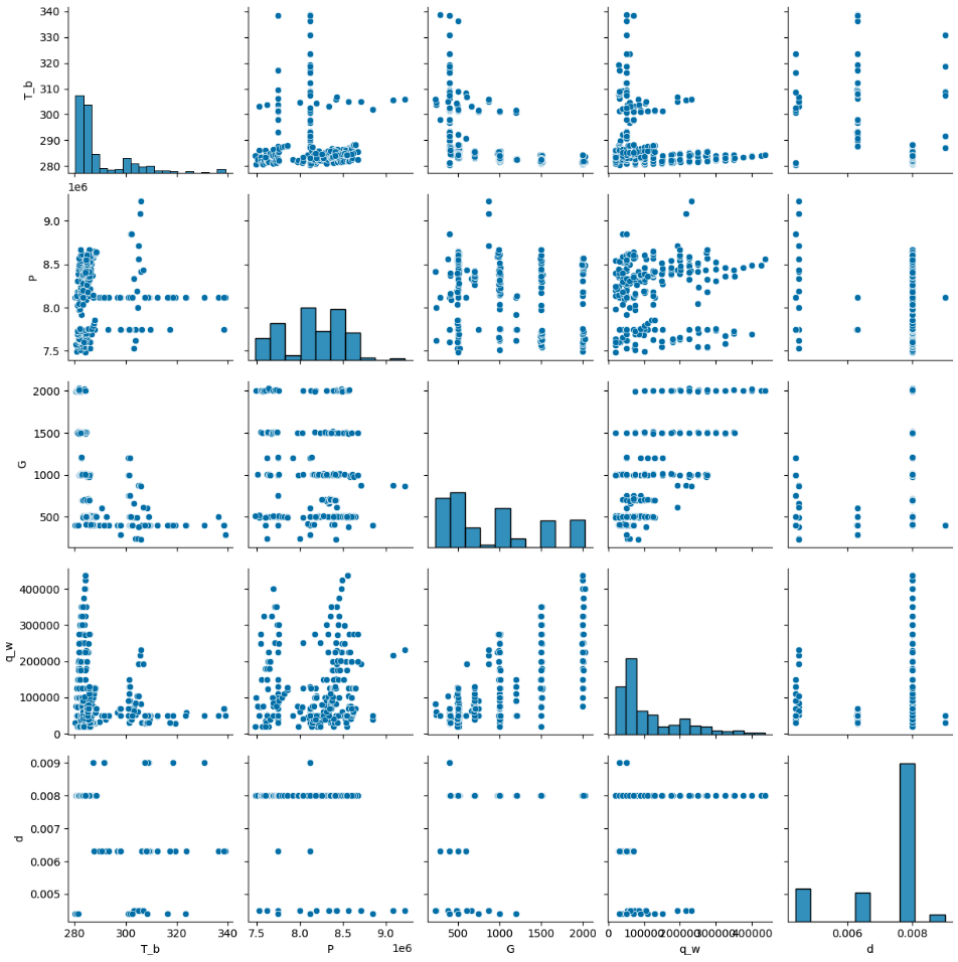



**Figure A.1:** Data distribution of the data set retrieved from the papers in table A.1

# B

# Nusselt correlations

**Table B.1:** Nusselt Correlations for heat transfer to supercritical $CO_2$ in an upward heated tube

| Author | Nusselt Correlation |
|---|---|
| Gupta [16] | $Nu_b = 0.01 Re_b^{0.89} \overline{Pr}_b^{-0.14} \left(\frac{\rho_w}{\rho_b}\right)^{0.93} \left(\frac{k_w}{k_b}\right)^{0.22} \left(\frac{\mu_w}{\mu_b}\right)^{-1.13}$ |
| Kim [22] | $Nu_b = 0.0182 Re_b^{0.824} \overline{Pr}_b^{0.515} \left(\frac{\rho_w}{\rho_b}\right)^{0.299}$ |
| Jackson [18] | $Nu_b = hD/\lambda_b = 0.021 Re_b^{0.8} Pr_b^{0.4} \left(\frac{\rho_w}{\rho_b}\right)^{0.3} \left(\frac{\overline{c_p}}{c_{p,b}}\right)^n$ |
| | $n = \begin{cases} 0.4 & \frac{T_b}{T_{pc}} < \frac{T_w}{T_{pc}} \leq 1, \text{ or }, 1.2 \leq \frac{T_b}{T_{pc}} < \frac{T_w}{T_{pc}} \\ 0.4 + 0.2\left(\frac{T_w}{T_b} - 1\right) & \frac{T_b}{T_{pc}} \leq 1 \leq \frac{T_w}{T_{pc}} \\ 0.4 + 0.2\left(\frac{T_w}{T_{pc}} - 1\right)\left[1 - 5\left(\frac{T_b}{T_{pc}} - 1\right)\right] & 1 < \frac{T_b}{T_{pc}} < 1.2, \text{ and }, \frac{T_b}{T_{pc}} < \frac{T_w}{T_{pc}} \end{cases}$ |
| Preda [33] | $Nu_w = 0.0015 Re_w^{1.03} \overline{Pr}_w^{0.76} \left(\frac{\rho_w}{\rho_b}\right)^{0.46} \left(\frac{\mu_w}{\mu_b}\right)^{0.53} \left(\frac{\kappa_w}{\kappa_b}\right)^{-0.43}$ |
| Pioro [32] | $Nu_w = 0.0038 Re_w^{0.96} \overline{Pr}_w^{-0.14} \left(\frac{\rho_w}{\rho_b}\right)^{0.84} \left(\frac{k_w}{k_b}\right)^{-0.75} \left(\frac{\mu_w}{\mu_b}\right)^{-0.22}$ |
| Wang [41] | $Nu_b = 0.0324 Re_b^{0.79} Pr_b^{0.66} \left(\frac{\rho_w}{\rho_b}\right)^{0.38} \left(\frac{\overline{c_p}}{c_{p_b}}\right)^n$ |
| | $n = \begin{cases} 0.66 \text{ for } T_b < T_w < T_{\mathrm{pc}}, \text{ or } 1.2T_{\mathrm{pc}} < T_b < T_w \\ 0.66 + 0.2\left(\frac{T_w}{T_{\mathrm{pc}}} - 1\right) \text{ for } T_b < T_{\mathrm{pc}} < T_w \\ 0.66 + 0.2\left(\frac{T_w}{T_{\mathrm{pc}}} - 1\right)\left[1 - 5\left(\frac{T_b}{T_{\mathrm{pc}}} - 1\right)\right] \text{ for } T_{\mathrm{pc}} < T_b < 1.2T_{\mathrm{pc}} \text{ and } T_b < T_w \end{cases}$ |
| Yang [43] | NHT: $K_c \leq 0.27$, $Nu_b = 0.41179 Nu_0^{1.10223} \left(\frac{P}{P_c}\right)^{-0.43274} \left(\frac{T_b}{T_{\mathrm{pc}}}\right)^{1.84087} \left(10^4 \frac{q_w}{GH_b}\right)^{0.13205} \times \left(\frac{\mu_b}{\mu_w}\right)^{-0.92839} \left(\frac{k_b}{k_w}\right)^{0.16801} \left(\frac{\overline{c_p}}{c_{p_b}}\right)^{0.72487}$ |
| | HTD: $K_c > 0.27$, $Nu_b = 1.7065 Nu_0^{0.94871} \left(\frac{P}{P_c}\right)^{-0.53838} \left(\frac{T_b}{T_{\mathrm{pc}}}\right)^{2.46823} \left(10^4 \frac{q_w}{GH_b}\right)^{-0.32562} \times \left(\frac{\mu_b}{\mu_w}\right)^{0.50388} \left(\frac{k_b}{k_w}\right)^{-0.54941} \left(\frac{\overline{c_p}}{c_{p_b}}\right)^{0.57156}$ |
| Zhang [46] | $Nu_{\mathrm{b}} = \begin{cases} 0.00672 \cdot Re_{\mathrm{b}}^{1.414} \overline{Pr}_{\mathrm{b}}^{-0.005} \left(\frac{\rho_{\mathrm{w}}}{\rho_{\mathrm{b}}}\right)^{0.448} \left(\frac{\overline{c_p}}{c_{p,\ \mathrm{b}}}\right)^{0.218} \cdot Bu^{0.586}, & h_{\mathrm{b}} < 0.9 h_{\mathrm{pc}} \\ 0.056 \cdot Re_{\mathrm{b}}^{0.829} \overline{Pr}_{\mathrm{b}}^{0.35} \left(\frac{\rho_{\mathrm{w}}}{\rho_{\mathrm{b}}}\right)^{-0.095} \left(\frac{\bar{c}_{\mathrm{p}}}{c_{\mathrm{p,b}}}\right)^{0.214} \cdot Bu^{0.142}, & h_{\mathrm{b}} \geqslant 0.9 h_{\mathrm{pc}} \end{cases}$ |
| Zhu [49] | $Nu_b = 0.0012 Re_b^{0.9484} \overline{Pr}_b^{0.718} K^{-0.0313}$ |

# C

# Experiments

## C.1. Experiment 1: Baseline model

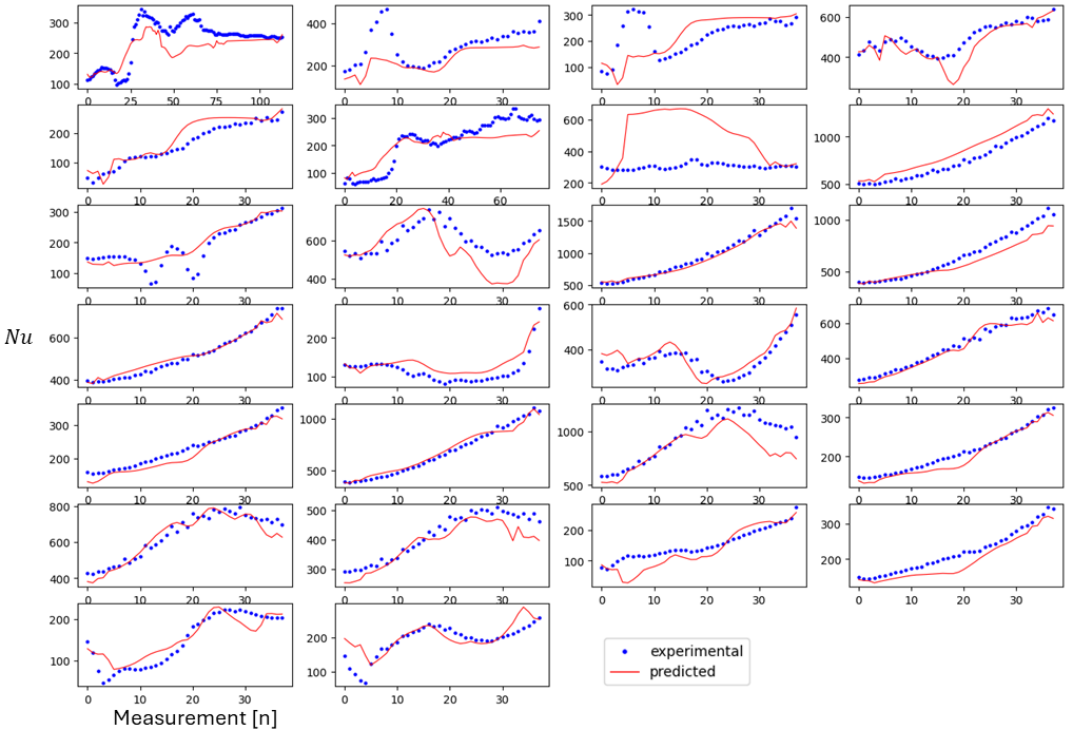



**Figure C.1:** Measurement location [n] plotted against predicted Nusselt number [-] per heat transfer trajectory for baseline model

## C.2. Experiment 2: Comparison of data processing methods

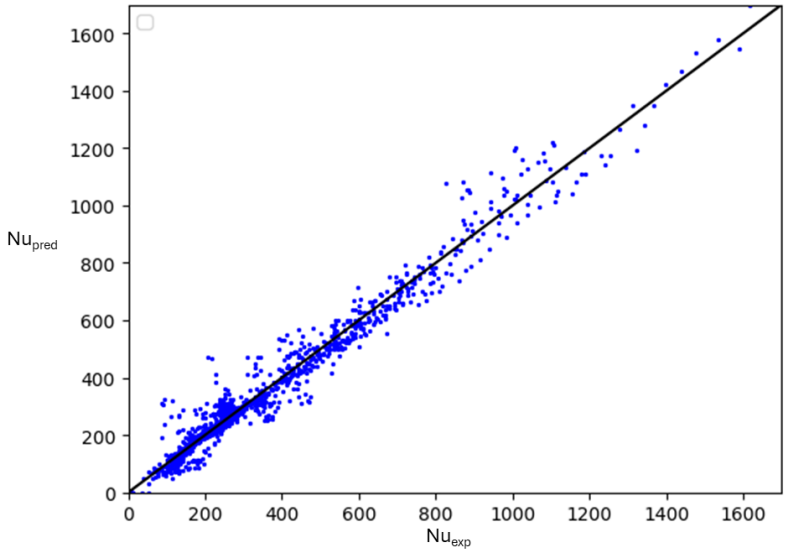



**Figure C.2:** Experimental Nusselt number vs. predicted Nusselt number for segmentation and padding

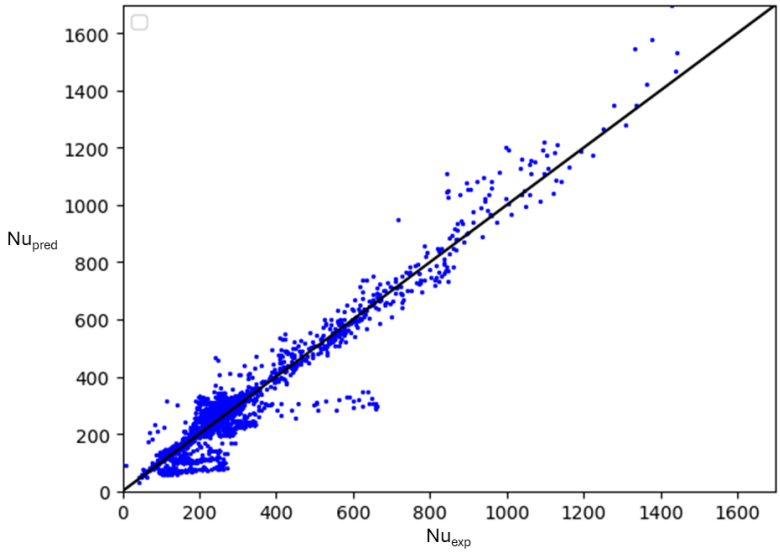



**Figure C.3:** Experimental Nusselt number vs. predicted Nusselt number for overlapped segments

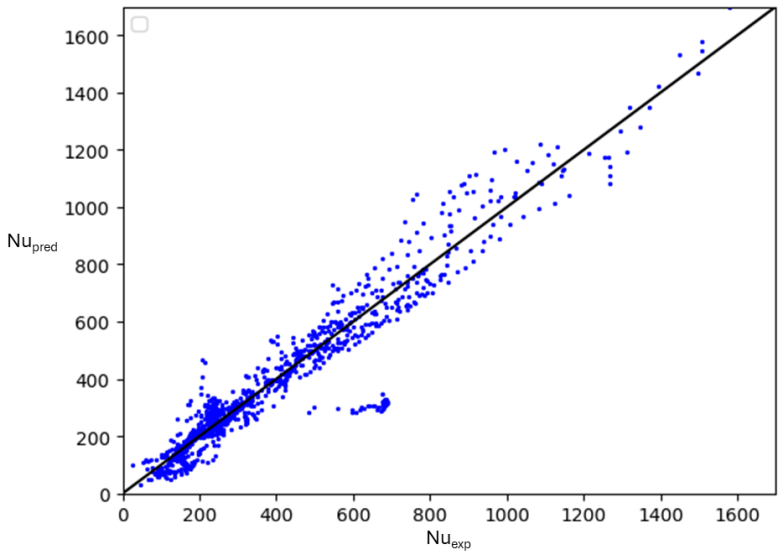



**Figure C.4:** Experimental Nusselt number vs. predicted Nusselt number for segmentaton and interpolation

## C.3. Experiment 3: Feature importance

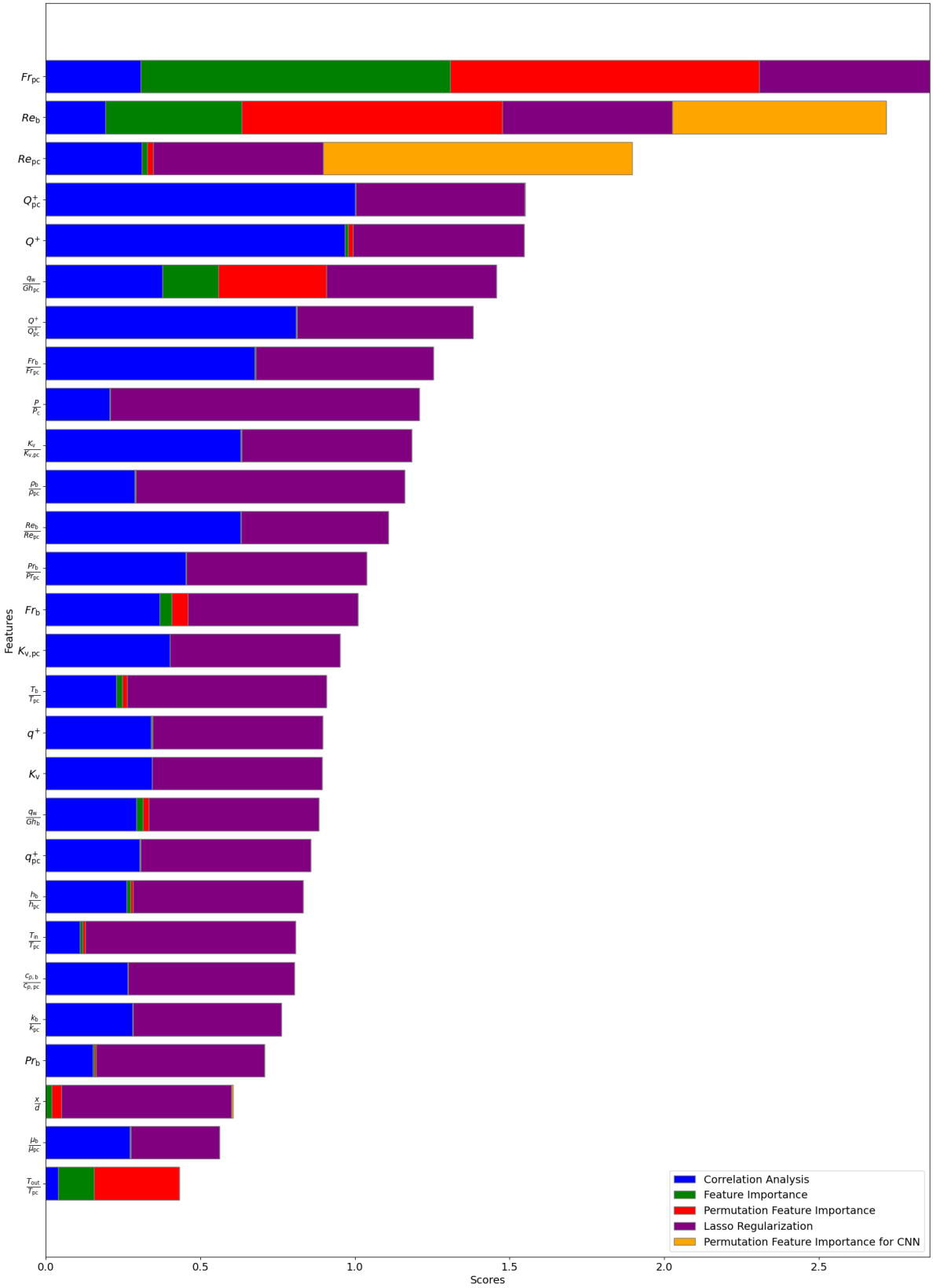



**Figure C.5:** Bar chart of feature analysis

## C.4. Experiment 6: Benchmarking CNN against ANN and Nusselt correlations

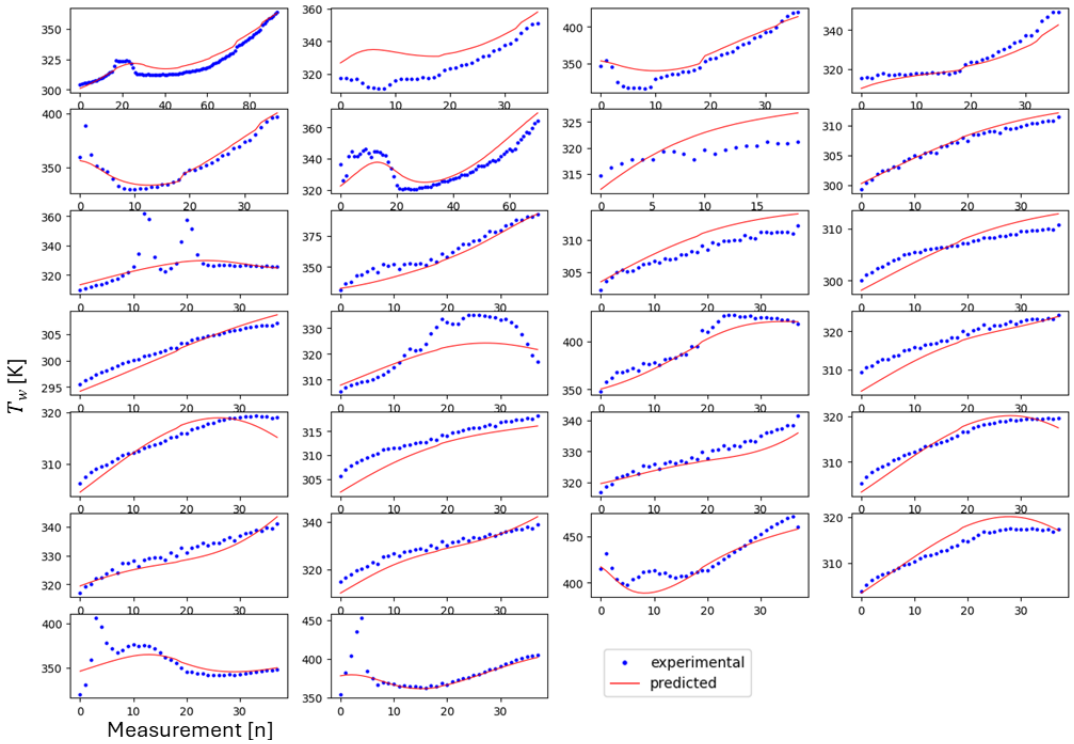



**Figure C.6:** Comparison between predicted and experimental wall temperature using the ANN model from section 4.6

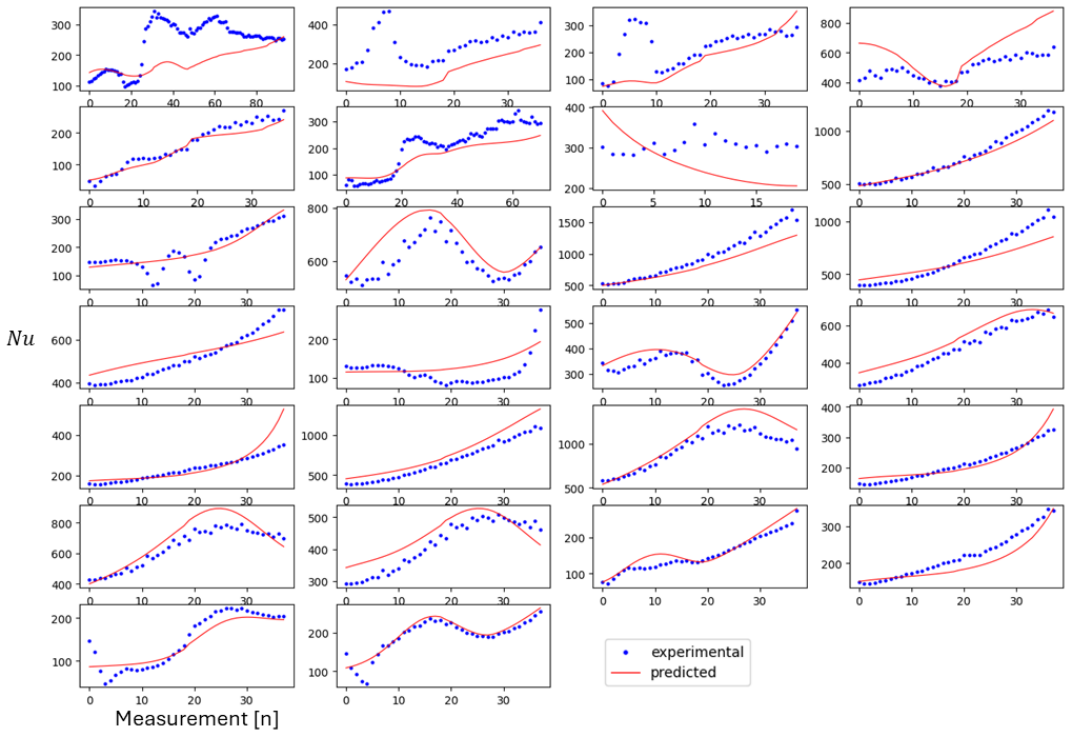



**Figure C.7:** Comparison between experimental Nusselt number and the calculated Nusselt number using the wall temperature predicted by the ANN from section 4.6

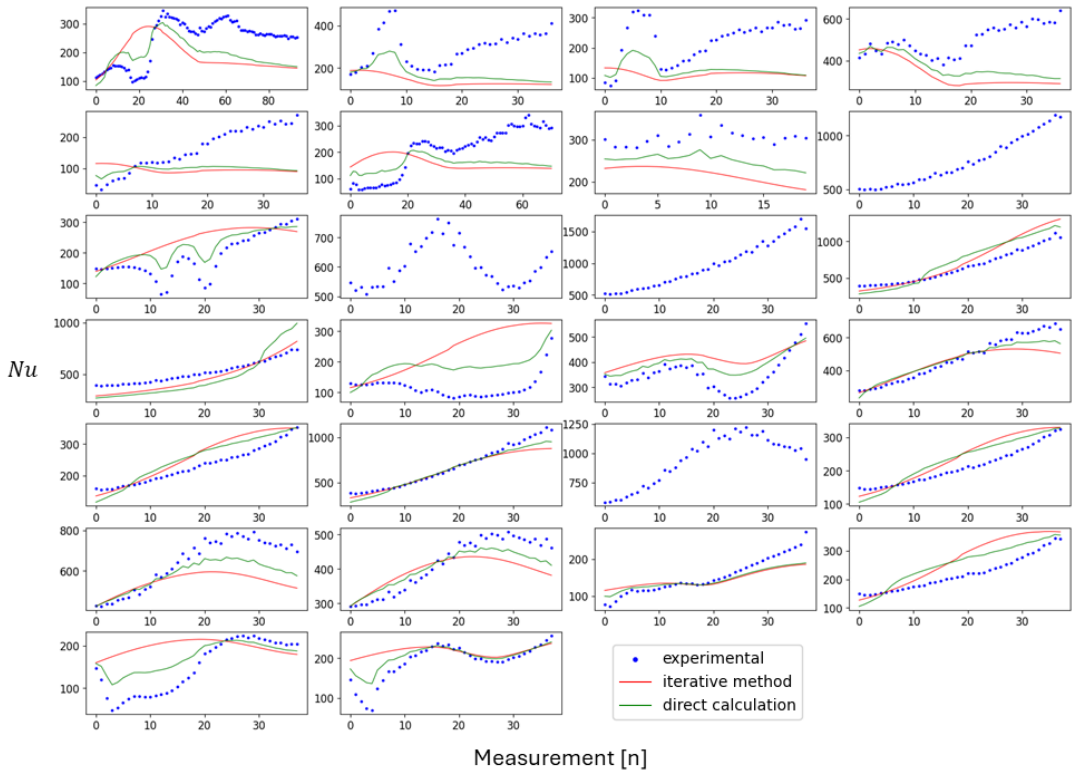



**Figure C.8:** Comparison between experimental and calculated Nusselt number. It is calculated using a iterative method for the Nusselt correlation and using the wall temperature directly to calculate the Nusselt number. The calculation is done using the correlation of Zhu et al [48]

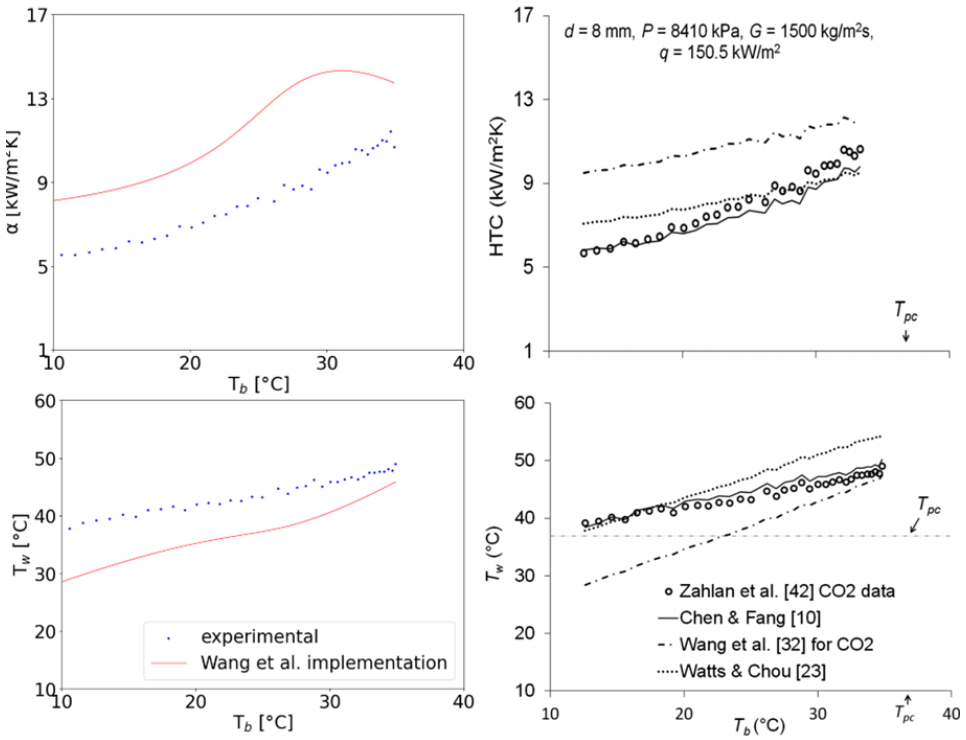



**Figure C.9:** The calculation for the Nusselt correlation of Wang et al. [41] is compared to the calculation done by Zahlan et al. [45] to validate that the method used in this paper is correct.

## C.5. Experiment 7: Impact of feature normalization on model performance

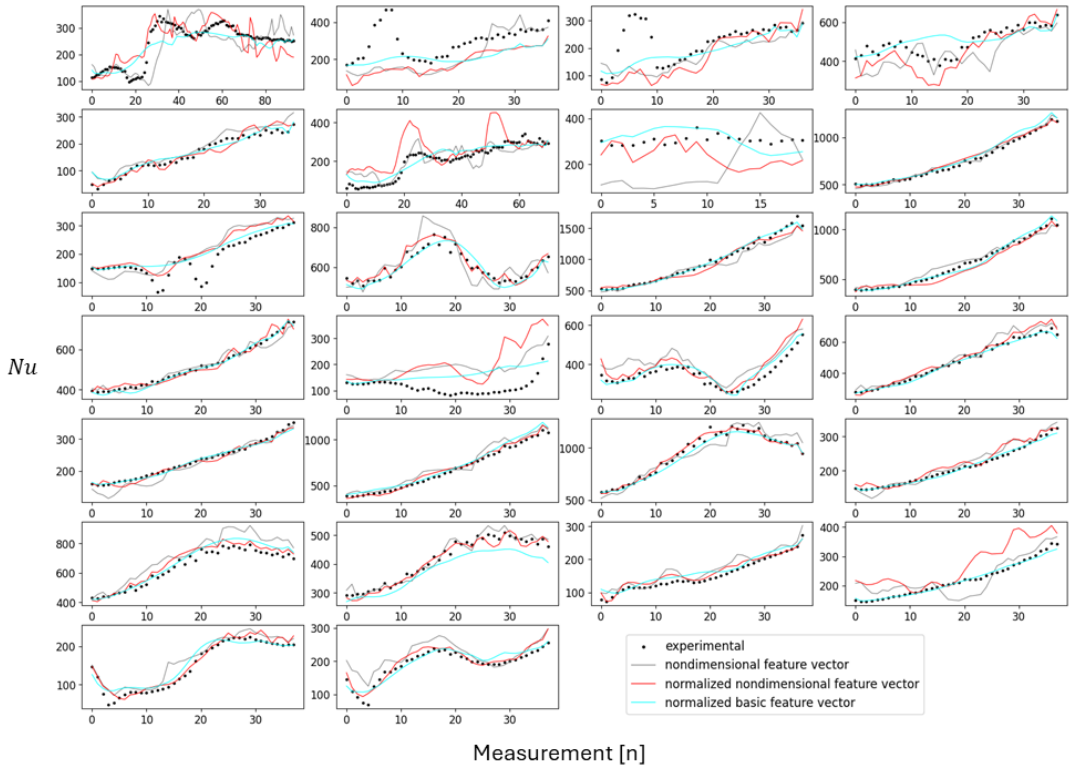



**Figure C.10:** Comparison between prediction and experimental Nusselt number for all trajectories

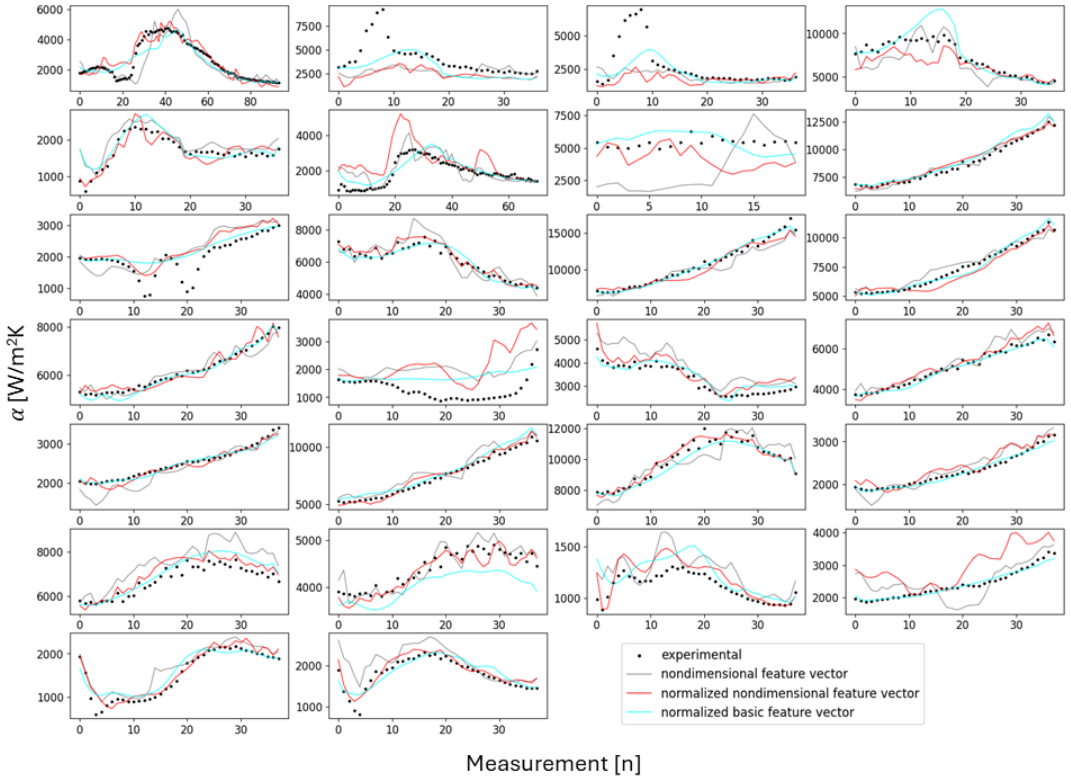



**Figure C.11:** Comparison between experimental heat transfer coefficient and heat transfer coefficient calculated using the predicted Nusselt number for all trajectories

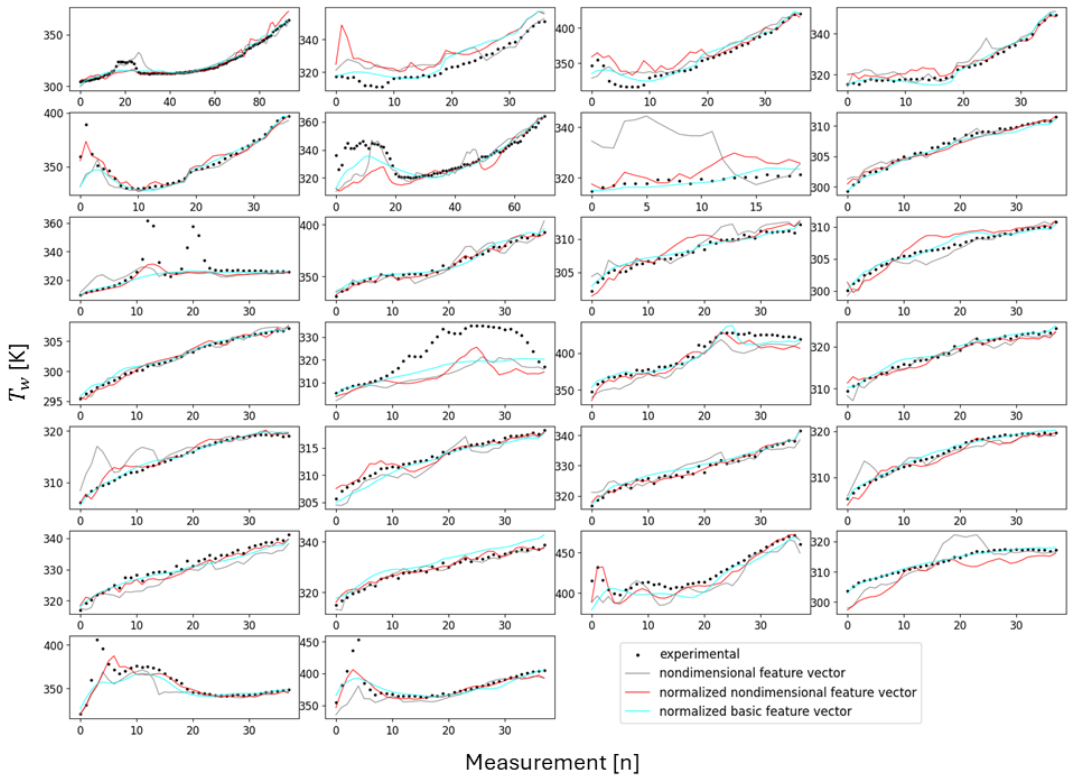



**Figure C.12:** Comparison between experimental wall temperature and wall temperature calculated using the predicted Nusselt number for all trajectories