Energy Proceedings
Vol 55, 2025
ISSN 2004-2965

# Utilizing Transfer Learning for Enhanced Prediction of Supercritical $CO_2$ Critical Flows#

Haifan Liao[1], Zhicheng Liang[1], Hongfei Hu[1], Haijun Wang[1*]

1 State Key Laboratory of Multiphase Flow in Power Engineering, Xi'an Jiaotong University, Xi'an 710049, China.

* Corresponding author：whj@mail.xjtu.edu.cn

**ABSTRACT**

With the widespread adoption of Carbon Capture and Storage (CCS) technologies and investigations into advanced power cycles such as the supercritical $CO_2$ ($sCO_2$) Brayton cycle, it is particularly crucial to focus on the risks and dynamics associated with accidental leaks under supercritical conditions. One key aspect is the accurate prediction of critical flow rates of supercritical fluids during leakage events. Owing to the scarcity of experimental data on critical flow in $sCO_2$ microchannels, this study harnesses the abundant experimental data from critical flow studies of water, employing machine learning algorithms enhanced by transfer learning to predict the critical flow rates of $sCO_2$. The approach involves pre-training a neural network on critical flow data from water, followed by fine-tuning with $sCO_2$ data, thereby bridging the gap in data availability and enhancing the model's generality across various parameters. The principal research findings indicate that transfer learning can improve prediction accuracy and adaptability, suggesting that transferring knowledge from extensively studied fluids like water can effectively enhance the predictive performance of less-documented supercritical fluids. This research provides a valuable tool for designing safer CCS and energy systems.

**Keywords:** Carbon Capture and Storage; Supercritical Carbon Dioxide; Critical Flow; Neural Networks; Transfer Learning.

**NONMENCLATURE**

| | |
|---|---|
| *Abbreviations* | |
| CCS | Carbon Capture and Storage |
| HEM | Homogeneous Equilibrium Model |
| LBB | Leak Before Break |
| ML | Machine Learning |
| SFM | Separated Flow Model |
| SHAP | SHapley Additive exPlanations |
| TL | Transfer Learning |
| *Symbols* | |
| $A$ | Cross-section Area / $m^2$ |
| $\boldsymbol{b}_{\text{net}}^{i}$ | Biases for Each Layer |
| $D$ | Hydraulic Diameter/m |
| $E_k$ | Interfacial Energy Exchange/ $kg{\cdot}m{\cdot}s^{-3}$ |
| $f$ | Friction Coefficient |
| $G$ | Critical Mass Flux /$kg{\cdot}m^{-2}{\cdot}s^{-1}$ |
| $h$ | Enthalpy /$J{\cdot}kg^{-1}$ |
| $L$ | Channel Length /m |
| $M$ | Interfacial Momentum Exchange Term /$kg{\cdot}s^{-2}$ |
| $p$ | Pressure /Pa |
| $p_r$ | Reduced Pressure/Pa |
| $q_w$ | Heat Flux of Liquid-wall Interface / $kg{\cdot}m^{-1}{\cdot}s^{-3}$ |
| $T$ | Temperature/K |
| $T_r$ | Reduced Temperature |
| $u$ | Velocity/$m{\cdot}s^{-1}$ |
| $\boldsymbol{W}_{\text{net}}^{i}$ | Weight Matrices |
| $\varepsilon$ | Surface Roughness /m |
| $\alpha$ | Void Fraction |
| $\rho$ | Density /$kg{\cdot}m^{-3}$ |
| $\sigma_{\text{net}}$ | Nonlinear Activation Function |

## 1. INTRODUCTION

Carbon dioxide ($CO_2$), as one of the primary sources of greenhouse gases and a working fluid in new power cycles, is increasingly attracting the attention of researchers and professionals in the field of engineering and technology, particularly concerning its safety and reliability during transportation and operation.

Carbon Capture and Storage (CCS) is considered one of the most cost-effective and efficient methods to mitigate $CO_2$ emissions from power plants and other carbon-intensive sources, thereby slowing global warming [1, 2]. The CCS system primarily involves the following processes: capturing $CO_2$ and compressing it to a supercritical state (or dense phase), and then transporting the supercritical $CO_2$ ($sCO_2$) to be stored in geological formations. Additionally, in response to the

# This is a paper for the 16th International Conference on Applied Energy (ICAE2024), Sep. 1-5, 2024, Niigata, Japan.

challenges of energy crises and environmental pollution, countries worldwide are developing advanced power cycles such as the $sCO_2$ Brayton cycle to achieve higher levels of energy sustainability, reliability, and economic efficiency. The $sCO_2$ Brayton cycle has been extensively studied over the past decade, achieving considerable thermal efficiency and reducing the overall system size. Currently, $sCO_2$ power cycles are considered an efficient and economical choice for various applications including solar thermal systems, next-generation nuclear reactors, and waste heat recovery systems [3-5]. In any of these applications, accidental $CO_2$ leaks pose a potential risk that could threaten people, equipment, and the environment.

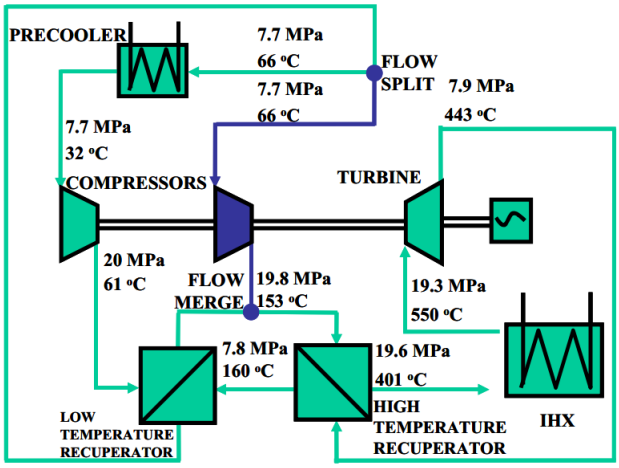



Fig. 1. Schematic of the supercritical carbon dioxide cycle [6].

In industrial applications, $sCO_2$ is typically transported and stored in high-pressure pipelines and containers. Over time, potential corrosion can lead to the development of small fissures in the pipelines, seals, and valves, resulting in leaks. When the $sCO_2$ system is compromised, $CO_2$ can escape through any cracks in the pipelines and containers, leading to hazardous situations characterized by low temperatures, asphyxiation, and damage. Therefore, it is essential to assess the safety of these systems and develop early leak detection and monitoring technologies. These advancements provide crucial support for the establishment of a warning, prevention, and control mechanism for micro-crack channel leakage incidents.

In this process, a critical aspect is understanding the leakage behavior and release rates of $sCO_2$ through compromised systems. The release of $sCO_2$ through microchannel fissures occurs in a critical state of maximum flow, commonly referred to as critical flow or choked flow. This flow is influenced by factors such as upstream pressure, upstream temperature, inlet morphology, and aspect ratio, while being unaffected by downstream pressure.

Despite extensive research on the flow dynamics of $sCO_2$ leaks, the critical flow of $sCO_2$ remains complex. The paucity of experimental data for $sCO_2$ in microchannels poses significant challenges to developing accurate predictive models. Existing models for $sCO_2$ critical flow are largely based on critical water flow models. However, due to differences in the physical properties of $CO_2$ and water and distortions in the pseudo-critical region, models based on water critical flow do not adequately describe $CO_2$ critical flow.

With the advancement of machine learning techniques, utilizing these methods to discover patterns in experimental data for accurate prediction has become a promising direction. Current research has shown promising reductions in prediction errors. Lahiri et al. [7] combined artificial neural networks with genetic algorithms to effectively integrate neural network hyperparameters for predicting fluid critical velocities. Zhang et al. [8] employed a genetic neural network (GNN) to predict leakage prior to rupture (LBB) under various conditions, reporting a relative error of 22.7% and suggesting an associated methodology. Xu et al. [9] proposed a method based on genetic neural networks to predict critical flow, considering not only the critical mass flux but also the critical pressure and mass. Yuan et al. [10] established a $sCO_2$ critical flow prediction model using a recurrent neural network (RNN) approach, applying K-fold cross-validation and L2 regularization, and utilized genetic algorithms to determine the optimal hyperparameters, reporting an average error of 4.88%.

Typically, traditional machine learning relies heavily on large amounts of training data. It operates under a critical assumption [11]: the training and testing data are drawn from the exact same distribution. However, this assumption often does not hold true in many real-world scenarios. Consequently, most conventional ML algorithms face three primary challenges: insufficient data, incompatible computational power, and distribution mismatch. This is particularly true for the issue of critical flow in $sCO_2$, which lacks experimental data across a broad range of parameters.

Given that experimental data on critical flow in microchannels based on water are far more abundant than those for $sCO_2$, transfer learning (TL) emerges as an effective means to enhance the generalization capabilities of current machine learning models [12]. The primary objective of TL is to leverage knowledge learned from source tasks in different domains to address target tasks, thereby obviating the need to start learning from

scratch in the data [13, 14]. This approach initially addresses the crucial issue of a lack of well-labeled training data. Moreover, it can significantly reduce the time and computational resources needed to train models because it allows for the reuse of pre-learned knowledge from other domains and tasks. Additionally, mismatches in distribution can lead to significant performance degradation in ML models. TL can also address this issue by integrating knowledge from one or more different domains.

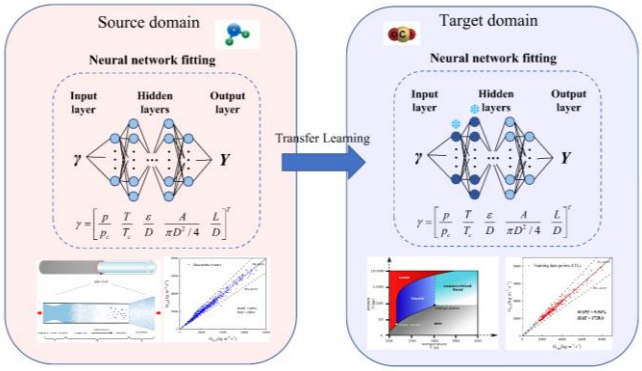



Fig. 2. Schematic of the transfer learning.

This paper first conducts a deep neural network fitting using experimental data on water critical flow, followed by transfer learning based on experimental data for $sCO_2$ critical flow. This approach enhances the accuracy and generalizability of the model predictions. The predictive method established in this study provides a valuable reference for developing critical flow models across a wide range of parameters.

## 2. DATASET AND PARAMETER DESCRIPTION

### *2.1 Dataset based on water critical flow*

Data from seven different experimenters on subcooled water microchannel critical flow have been collected to serve as the source training data for transfer learning. This dataset, used for initial deep neural network fitting, encompasses a wide range of microchannel geometrical dimensions and inlet conditions. The specific parameters are summarized in Table 1.

Table 1 Summary of Experimental Parameters(water)

| Source | Inlet fluid pressure /MPa | Inlet fluid temperature /K | Length /mm | Hydraulic diameter /mm | Flow channel area /mm$^2$ | Number of data |
|---|---|---|---|---|---|---|
| Collier et al.[15] | 4.04 - 9.11 | 493.1 - 555.4 | 20 | 0.09 - 0.35 | 0.07 - 4.96 | 29 |
| Huang et al. [16] | 3.92 - 11.12 | 459.1 - 538.1 | 80 | 1.13 | 6 | 19 |
| Abdollahian et al. [17] | 3.26 - 11.53 | 451.6 - 540.9 | 57.2 | 0.4 - 2.2 | 12.70 - 71.12 | 27 |
| Amos et al. [18] | 2.73 - 15.92 | 465.8 - 646.8 | 63.5 | 0.16 - 0.75 | 0.03 - 7.81 | 71 |
| John et al. [19] | 3.91 - 12.16 | 462.9 - 595.9 | 46 | 0.5 - 1.3 | 20 - 52 | 458 |
| Revankar et al. [20] | 6.6 - 6.87 | 508.6 - 531.7 | 0.32 | 0.59 - 0.71 | 0.83 - 1.74 | 19 |
| Mignot et al. [21] | 24.43-25.09 | 480 - 514 | 280 | 1.59 | 1.98 | 7 |
| Total | | | | | | 630 |

### *2.2 Dataset based on $sCO_2$ critical flow*

As the target task for transfer learning, critical flow data for $sCO_2$ has been collected from various researchers, with specific parameters summarized in Table 2. Additionally, the experimental dataset from Fan et al. [22] has been selected as the validation dataset, which does not participate in the training process. This selection is intended to test the generalization ability of the model.

Table 2 Summary of Experimental Parameters($sCO_2$)

| Source | Inlet fluid pressure /MPa | Inlet fluid temperature /K | Length /mm | Hydraulic diameter /mm | Flow channel area /mm$^2$ | Number of data |
|---|---|---|---|---|---|---|
| Liu et al. [23] | 7.52 - 9.03 | 308.1 - 318.1 | 8.02 -25.42 | 0.83 - 1.53 | 0.99 - 1.43 | 46 |
| LI et al. [24] | 7.5 - 9.2 | 288.4 - 329.5 | 2-20 | 2 | 3.14 | 76 |
| Wang et al. [25] | 8 - 11 | 307.2 - 368.8 | 1-100 | 1 | 0.785 | 39 |
| Mignot et al. [26] | 10.1 | 312 - 378.9 | 338.1 | 2 - 3.175 | 3.14 - 7.92 | 59 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Fan et al. [22] | 8 - 10.71 | 308 - 373.8 | 1 - 15 | 1.01 | 0.8 | 65 |
| Edlebeck et al. [27] | 7.7 - 11 | 304.6 - 331.7 | 3 - 20 | 1.01 | 0.79 - 0.81 | 32 |
| Total | | | | | | 317 |

## 3. ESTABLISHMENT OF THE DEEP LEARNING MODEL

### *3.1 Determination of input parameters*

Traditional methods for solving critical flow primarily include simplified analytical models, such as the Moody and Henry-Fauske models [28, 29], along with numerical computation models that solve sets of two-phase control equations. These include the Homogeneous Equilibrium Model (HEM) and the Two-fluid model/Separated Flow Model (SFM) [30-35].

In this context, Richter [36] developed a detailed SFM that describes the interphase transfer characteristics of subcooled water gas-liquid critical flow. This model assumes one-dimensional, steady-state flow within the tube where the pressure is uniform across any cross-section. It also disregards the friction between the vapor and the tube walls. The conservation equations for mass, momentum, and energy in this model are as follows:

$$\frac{\partial}{dz}\left(\rho_k u_k \alpha_k A\right) = \Gamma_k \tag{1}$$

$$\frac{\partial}{\partial z}\left(\alpha_k A \rho_k u_k^2\right) + \alpha_k A \frac{\partial p}{\partial z} = -A f_{\mathrm{wk}} + M_k - \alpha_k A \rho_k g \sin\theta \tag{2}$$

$$\frac{\partial}{\partial z}\left[A \alpha_k \rho_k u_k \left(h_k + \frac{u_k^2}{2}\right)\right] = -A q_{\mathrm{wk}} + E_k - \alpha_k A \rho_k u_k g \sin\theta \tag{3}$$

where α, ρ, A, u, p, g, θ and z indicate the void fraction, fluid density, cross-section area of the channel, fluid velocity, fluid pressure, acceleration of gravity, inclination angle and the coordinate along the flow channel, respectively. The terms Γ , M , E , fw , qw indicate the rate of mass transfer, momentum transfer, and energy transfer per unit channel length due to evaporation along the vapor-liquid interface, wall force and interfacial heat transfer. The subscript k=g is for the vapor phase and k=l is for the liquid phase.

From the classical phase flow equations (1)-(3), it is apparent that the main parameters affecting critical flow include inlet pressure, inlet temperature (or inlet enthalpy), the geometry of the inlet cross-section, channel length, and the effects of friction.

In summary, inlet pressure and inlet temperature are considered as inputs to characterize the influence of thermal parameters on mass flow rate. Additionally, the impact of the breach shape on mass flow is taken into account, incorporating inlet cross-sectional area, hydraulic diameter of the inlet cross-section, channel length, and channel roughness as input parameters. Considering the requirements for generalizability and ease of training, these input parameters are dimensionless. The final selected parameters include Inlet reduced pressure $p/p_c$, Inlet reduced temperature $T/T_c$, inlet area ratio $4A/\pi D^2$ (ratio of the inlet cross-sectional area to that of a circular tube), channel L/D ratio, and channel relative roughness $\varepsilon/D$.

### *3.2 Neural network structure*

A basic Feedforward Neural Network (FNN), or Multilayer Perceptron (MLP), is defined as a fusion of linear and nonlinear transformations. A deep neural network is a type of feedforward neural network with multiple hidden layers. Compared to traditional feedforward neural networks, deep neural networks have more layers and a more complex structure, enabling them to capture and learn more complex patterns and features. The architecture of such networks typically includes an input layer, several hidden layers, and an output layer, defined as follows:

$$\begin{aligned} &\text{input layer: } N^0(\boldsymbol{\gamma}) = \boldsymbol{\gamma}, \\ &\text{hidden layers: } N^i(\boldsymbol{\gamma}) = \sigma_{\mathrm{net}}\left(\boldsymbol{W}_{\mathrm{net}}^i N^{i-1}(\boldsymbol{\gamma}) + \boldsymbol{b}_{\mathrm{net}}^i\right), 1 \le i \le L-1 \\ &\text{output layer: } N^L(\boldsymbol{\gamma}) = \boldsymbol{W}_{\mathrm{net}}^L N^{L-1}(\boldsymbol{\gamma}) + \boldsymbol{b}_{\mathrm{net}}^L \end{aligned} \tag{4}$$

Its composition includes the weight matrices $\boldsymbol{W}_{\mathrm{net}}^i$ and biases $\boldsymbol{b}_{\mathrm{net}}^i$ for each layer, coupled with a nonlinear activation function $\sigma_{\mathrm{net}}(\cdot)$. In the pursuit of computing the active subspace for the dataset ($\boldsymbol{\gamma}$, $g(\boldsymbol{\gamma})$), the neural network $N(\boldsymbol{\gamma})$ is utilized for fitting, optimizing the network's weights $\boldsymbol{W}_{\mathrm{net}}^i$ and biases $\boldsymbol{b}_{\mathrm{net}}^i$.

The network parameters are updated using the AdamW [37], which combines weight decay techniques with the Adam optimizer. Weight decay is a technique that biases the optimization towards solutions with smaller norms. It has long been a standard technique for improving the generalization capability of machine learning models and continues to be widely employed in the training of modern deep neural networks [38, 39].

The performance of the two models was evaluated using the mean absolute percentage error (MAPE), mean absolute error (MAE), defined as follows:

$$MAPE = \frac{100\%}{N} \times \sum_{i=1}^{N} \frac{\left| y_{\text{pre,i}} - y_{\text{true,i}} \right|}{y_{\text{true,i}}}$$
$$MAE = \frac{1}{N} \times \sum_{i=1}^{N} \left| y_{\text{pre,i}} - y_{\text{true,i}} \right| \quad (5)$$

Here, the subscripts pre and true represent the network's predicted and true values from the dataset.

### *3.3 Transfer learning*

Transfer learning is a machine learning strategy that allows the use of models trained on an original task to be applied to a different but related task. This method leverages the model libraries already trained on the original task, reducing the amount of training data required for the target task, enhancing learning efficiency, and often improving model performance. The goal of transfer learning is to utilize knowledge obtained from the source task to enhance the learning performance on the target task. The main steps involved include:

1. Pre-trained Model: Train a model $f_S(\mathbf{x}_i^S; \theta_S)$ on the source task, optimizing the loss function $\mathcal{L}_S$:

$$\theta_S^* = \arg\min_{\theta_S} \mathbb{E}_{(\mathbf{x}_i^S, y_i^S) \sim \mathcal{D}_S} [\mathcal{L}_S(f_S(\mathbf{x}_i^S; \theta_S), y_i^S)]$$

2. Transfer Model: Use the pre-trained model parameters $\theta_S^*$ as initial values and fix some parameters to construct the target model $f_T(\mathbf{x}; \theta_T)$:

$$f_T(\mathbf{x}; \theta_T) = g(f_S(\mathbf{x}; \theta_S^*), \theta_T)$$

Here, $g$ represents the model structure that relates the pre-trained model $f_S$ to the new task.

3. Fine-tune Model: Optimize the target model $f_T$ parameters $\theta_T^*$ on the target domain $\mathcal{D}_T$:

$$\theta_T^* = \arg\min_{\theta_T} \mathbb{E}_{(\mathbf{x}_i^T, y_i^T) \sim \mathcal{D}_T} [\mathcal{L}_T(f_T(\mathbf{x}_i^T; \theta_T), y_i^T)]$$

Here the subscript T denotes 'target' and S denotes 'source'.

## 4. RESULTS AND DISCUSSION

### *4.1 Neural network predictive performance and generalization verification*

The neural network is initially trained on the source task, configured with nodes arranged in the layers as [5, 8, 16, 16, 8, 1]. Batch normalization layers are incorporated, bringing the total number of layers in the deep neural network to 9. The network is trained using the AdamW optimizer over 4000 steps with a learning rate of 1e-3.

The training results, as shown in Figure 3, indicate that the neural network accurately predicts the critical flow in the source dataset, which involves the critical flow of water as the inlet medium. The model achieves an average relative error of 5.87% and an average absolute error of 1313.8.

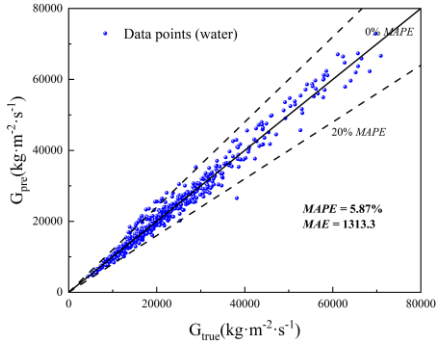
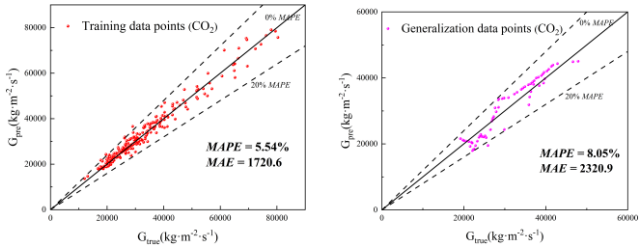



Fig. 3. Neural network fitting of critical flow mass flux (water, maximum nodes 16)

In the transfer learning phase, it is common to freeze the initial layers (which typically learn more generic features) and retrain the later layers (which learn features more specific to the new task). Thus, the last three layers of the network are retrained while the first six layers are frozen. The network is retrained on the basis of the previously trained model to test its fitting effectiveness.

In Figure 4, the neural network's fit on the training dataset for $sCO_2$ shows an average relative error of 5.54% and an average absolute error of 1720.6. On the generalization dataset, it achieves an average relative error of 8.05% and an average absolute error of 2320.9. The fitting results are favorable, with the neural network requiring only a few additional training steps (400 steps) to effectively transfer from the source dataset (water) to the target dataset ($CO_2$), ensuring good generalizability. This generalizability is validated by a dataset that was not used in the parameter training process, comprising 65 data points from Fan's [22] experiments on $sCO_2$ critical flow.

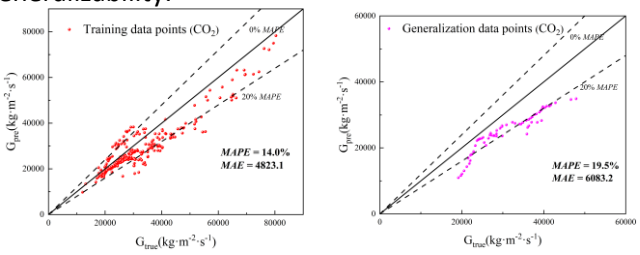



(a) Training set (b) Generalization set

Fig. 4. Neural network fitting of critical flow mass flux (with TL, $CO_2$, maximum nodes 16)

The results of training directly from the target data source ($CO_2$) are shown in Figure 5. Direct training fails to

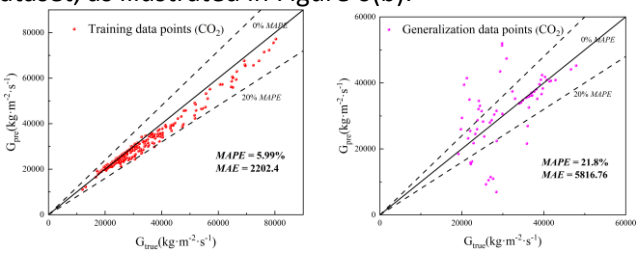
ensure high accuracy on the training dataset, with an average relative error of 14.0% and an average absolute error of 4823.1. This highlights that transfer learning not only enables the neural network to learn the characteristics of $sCO_2$ critical flow more effectively and quickly but also better ensures the requirements for generalizability.




(a) Training set (b) Generalization set

Fig. 5. Neural network fitting of critical flow mass flux (no TL, $CO_2$, maximum nodes 16)

Increasing the number of nodes in the neural network can improve accuracy on the training dataset, with an average relative error of 5.99% and an average absolute error of 2202.4, but it risks overfitting, as shown by a deterioration in performance on the generalization dataset, as illustrated in Figure 6(b).




(a) Training set (b) Generalization set

Fig. 6. Neural network fitting of critical flow mass flux (no TL, $CO_2$, maximum nodes 32)

### *4.2 Analysis of transfer learning results*

The study examined the impact of the number of neurons on transfer learning, as illustrated in Figure 7. It was found that increasing the number of nodes tends to reduce errors on the training dataset regardless of whether transfer learning is used; however, it increases errors on the generalization dataset, making the model more prone to overfitting. In the context of transfer learning, too few nodes hinder successful transfer, while too many lead to overfitting.

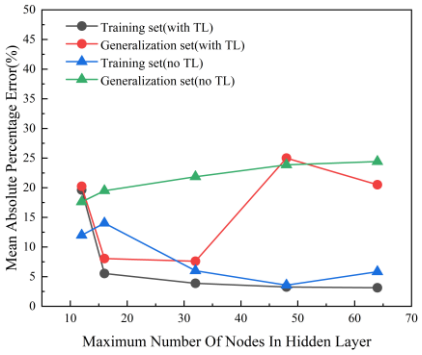



Fig. 7. The impact of the maximum number of nodes in neural network on transfer learning

Another critical parameter in transfer learning is the number of layers frozen. Fewer frozen layers mean more parameters need retraining, and less information from the source database is retained. As shown in Table 3, if too many layers are frozen within a limited number of retraining steps, the transfer effect cannot be achieved. Conversely, freezing too few layers retains insufficient information from the source database, which can increase errors on the generalization dataset.

Table 3 Error after freezing different layers in transfer learning.

| Num of frozen layers | Training set | | Generalization set | |
|---|---|---|---|---|
| | Mean percentage error (%) | Max percentage error (%) | Mean percentage error (%) | Max percentage error (%) |
| 7 | 17.1 | 130.6 | 31.5 | 149.9 |
| 6 | 5.5 | 21.6 | 8.0 | 24.6 |
| 5 | 3.2 | 11.9 | 9.5 | 40.8 |
| 4 | 2.6 | 11.7 | 7.2 | 24.7 |
| 3 | 3.3 | 13.6 | 9.5 | 40.3 |

### *4.3 SHAP values of the model*

To reveal the impact of each variable on the model more specifically and to compare differences after training with data from different media, SHAP (SHapley Additive exPlanations) values were introduced. SHAP, proposed by Lundberg et al. [40, 41] in 2017, is a unified framework for explaining predictions that integrates ideas from game theory [42] and local explanations [43]. It quantifies the contribution of each feature in the model to the final prediction. Currently, SHAP values are widely used to explain machine learning models.

SHAP values were calculated for both the source model before transfer learning and the model after transfer learning. The results, as illustrated in the Figure 8, indicate that after transferring the critical flow prediction model trained on water to $sCO_2$, the influences of inlet pressure, inlet temperature, and L/D ratio increased, while the impact of roughness decreased. Since all collected $sCO_2$ datasets had circular inlet cross-sections, the inlet area ratio had almost no impact.

It must be noted that due to the local interpretive nature of SHAP values, the explanations of importance for the data-poor $sCO_2$ dataset are significantly limited by the dataset's richness.

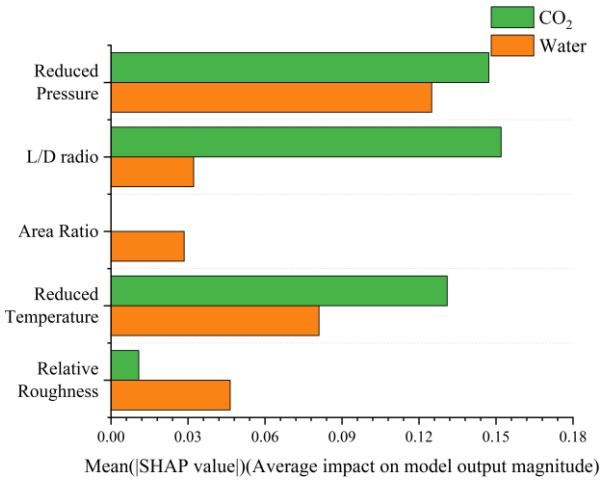



Fig. 8. Contribution of each feature to the model.

## 5. CONCLUSIONS

In summary, this study established a microchannel $sCO_2$ critical flow prediction model using neural networks and transfer learning. The main efforts and conclusions drawn are as follows:

(1) Feature selection for the neural network was grounded in traditional numerical models, utilizing a source database comprised of experimental data on critical flow in water microchannels. Through transfer learning, the model was successfully adapted to $sCO_2$ critical flow scenarios, exhibiting enhanced accuracy and generalizability.

(2) The effects of the number of neurons and the number of frozen layers on the transfer learning outcome were studied, revealing that while more neurons and fewer frozen layers decrease errors on the training dataset, they do not guarantee reduced errors on the generalization dataset.

(3) The comparison of SHAP values before and after transfer learning on the collected dataset revealed that the importance of inlet pressure, inlet temperature, and L/D ratio in determining the model's predictions increased after the model was transferred from water to $sCO_2$.

## ACKNOWLEDGEMENT

This work was supported by National Key R&D Program of China (Grant No.2023YFB4102205), Innovation Capability Support Program of Shaanxi (Program No.2023-CX-TD-18) and Innovative Scientific Program of CNNC.

## REFERENCE

[1] Lund H, Mathiesen BV. The role of carbon capture and storage in a future sustainable energy system. Energy. 2012;44:469-76.
[2] Sathre R, Chester M, Cain J, Masanet E. A framework for environmental assessment of CO2 capture and storage systems. Energy. 2012;37:540-8.
[3] Bowen F. Barriers to carbon capture and storage may not be obvious. Nature. 2010;464:160-.
[4] Iribarren D, Petrakopoulou F, Dufour J. Environmental and thermodynamic evaluation of CO2 capture, transport and storage with and without enhanced resource recovery. Energy. 2013;50:477-85.
[5] Wang J, Huang Y, Zang J, Liu G. Research activities on supercritical carbon dioxide power conversion technology in China. Turbo Expo: Power for Land, Sea, and Air: American Society of Mechanical Engineers; 2014. p. V03BT36A009.
[6] Davis CB, Marshall TD, Weaver K. Modeling the GFR with RELAP5-3D. Idaho National Lab.(INL), Idaho Falls, ID (United States); 2005.
[7] Lahiri S, Ghanta KC. Artificial neural network model with parameter tuning assisted by genetic algorithm technique: study of critical velocity of slurry flow in pipeline. Asia - Pacific Journal of Chemical Engineering. 2010;5:763-77.
[8] Zhang J, Chen R, Wang M, Tian W, Su G, Qiu S. Prediction of LBB leakage for various conditions by genetic neural network and genetic algorithms. Nuclear Engineering and Design. 2017;325:33-43.
[9] Xu H, Tang T, Zhang B, Liu Y. Application of artificial neural network for the critical flow prediction of discharge nozzle. Nuclear Engineering and Technology. 2022;54:834-41.
[10] Yuan Y, Chen T, Zhou Y, Feng H, Wang J, Zhai H, et al. Supercritical carbon dioxide critical flow model based on deep learning. Progress in Nuclear Energy. 2024;170:105121.
[11] Niu S, Liu Y, Wang J, Song H. A decade survey of transfer learning (2010–2020). IEEE Transactions on Artificial Intelligence. 2020;1:151-66.
[12] Inubushi M, Goto S. Transfer learning for nonlinear dynamics and its application to fluid turbulence. Physical Review E. 2020;102:043301.
[13] Maria Carlucci F, Porzi L, Caputo B, Ricci E, Rota Bulo S. Autodial: Automatic domain alignment layers. Proceedings of the IEEE international conference on computer vision2017. p. 5067-75.
[14] Pan SJ, Tsang IW, Kwok JT, Yang Q. Domain adaptation via transfer component analysis. IEEE transactions on neural networks. 2010;22:199-210.
[15] Collier RP, Stulen FB, Mayfield ME, Pape DB, Scott PM. Two-phase flow through intergranular stress corrosion cracks and resulting acoustic emission. Batelle, Columbus, Ohio, United States. 1984.
[16] Huang b, ZHANG B, SUN Y. Investigation on the Critical Flow Through a Rectangular Narrow Slit. JOURNAL OF XIANJIAOTONG UNIVERSITY. 2020;54:26-31.
[17] Abdollahian D, Chexal B. Calculation of leak rates through cracks in pipes and tubes. Final report. Levy (S.), Inc., Campbell, CA (USA); 1983.

[18] Amos CN, Schrock VE. Critical discharge of initially subcooled water through slits.[PWR; BWR]. Lawrence Berkeley Lab., CA (USA); 1983.
[19] John H, Reimann J, Eiseie G. Kritische Leckströmung aus rauhen Rissen in Druckwasserbehältern. Report KfK. 1987;4192.
[20] Revankar S, Wolf B, Riznic JR. INVESTIGATION OF subscriptCOOLEDWATER DISCHARGE THROUGH SIMULATED STEAM GENERATOR TUBE CRACKS. Multiphase Science and Technology. 2013;25.
[21] Mignot G, Anderson M, Corradini M. Critical Flow Experiment and Analysis for Supercritical Fluid. Nuclear Engineering and Technology. 2008;40:133-8.
[22] Fan X, Wang Y, Zhou Y, Chen J, Huang Y, Wang J. Experimental study of supercritical CO2 leakage behavior from pressurized vessels. Energy. 2018;150:342-50.
[23] Liu JP, Niu YM, Chen JP, Chen ZJ, Feng X. Experimentation and correlation of R744 two-phase flow through short tubes. Experimental Thermal and Fluid Science. 2004;28:565-73.
[24] LI W, ZHANG D, ZHAO M. Experimental Study on Critical Flow of Supercritical $CO_2$ at Steady State and Model Verification. Atomic Energy Science and Technology. 2022.
[25] Wang Y. Experimental Study on Critical Flow of Supercritical CO2. Nuclear Science and Technology. 2018;06:61-8.
[26] Mignot GP, Anderson MH, Corradini ML. Measurement of supercritical CO2 critical flow: Effects of L/D and surface roughness. Nuclear Engineering and Design. 2009;239:949-55.
[27] Edlebeck J, Nellis GF, Klein SA, Anderson MH, Wolf M. Measurements of the flow of supercritical carbon dioxide through short orifices. The Journal of Supercritical Fluids. 2014;88:17-25.
[28] Henry RE, Fauske HK. The two-phase critical flow of one-component mixtures in nozzles, orifices, and short tubes. 1971.
[29] Moody FJ. Maximum flow rate of a single component, two-phase mixture. 1965.
[30] Ardron K. A two-fluid model for critical vapour-liquid flow. International Journal of Multiphase Flow. 1978;4:323-37.
[31] Benhmidene A, Chaouachi B, Gabsi S, Bourouis M. Modeling of boiling two-phase flow in the bubble pump of diffusion-absorption refrigeration cycles. Chemical engineering communications. 2015;202:15-24.
[32] Liao H, Liu Q, Gao Y, Zhang S, Yang K, Wang H. A simplified two-fluid model for more stable microchannel two-phase critical flow prediction. Chemical Engineering Science. 2024:119885.
[33] Seixlack A, Barbazelli M. Numerical analysis of refrigerant flow along non-adiabatic capillary tubes using a two-fluid model. Applied Thermal Engineering. 2009;29:523-31.
[34] Xu H, Badea AF, Cheng X. Analysis of two phase critical flow with a non-equilibrium model. Nuclear Engineering and Design. 2021;372:110998.
[35] Yin S, Wang H, Xu B, Yang C, Gu H. Critical flow leakage of a vapour-liquid mixture from sub-cooled water: Nucleation boiling study. International Journal of Heat and Mass Transfer. 2020;146:118807.
[36] Richter HJ. Separated two-phase flow model: application to critical two-phase flow. International Journal of Multiphase Flow. 1983;9:511-30.
[37] Loshchilov I, Hutter F. Decoupled weight decay regularization. arXiv preprint arXiv:171105101. 2017.
[38] Devlin J, Chang M-W, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:181004805. 2018.
[39] Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. International conference on machine learning: PMLR; 2019. p. 6105-14.
[40] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Advances in neural information processing systems. 2017;30.
[41] Lundberg SM, Erion GG, Lee S-I. Consistent individualized feature attribution for tree ensembles. arXiv preprint arXiv:180203888. 2018.
[42] Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. Knowledge and information systems. 2014;41:647-65.
[43] Ribeiro MT, Singh S, Guestrin C. " Why should i trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining2016. p. 1135-44.