# Improvement of the Critical Heat Flux Correlation using Symbolic Regression

Yeonha Lee[a] and Jeong Ik Lee[a*]

[a]Department of Nuclear and Quantum Engineering, N7-1 KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, Korea 34141
*Corresponding author: jeongiklee@kaist.ac.kr

***Keywords :*** Critical Heat Flux, W-3 correlation, Symbolic Regression

## 1. Introduction

In the design and safety analysis of pressurized water reactors (PWRs), the critical heat flux (CHF) is a key parameter for ensuring fuel integrity. In particular, the departure from nucleate boiling (DNB) causes a sharp decrease in the local heat transfer coefficient. This phenomenon can lead to a rapid rise in cladding temperature, potentially resulting in fuel damage. Therefore, accurately predicting DNB is essential to maintain sufficient thermal margins in nuclear power plants.

The W-3 correlation[1] has been a standard tool for predicting local DNB heat flux in PWRs for several decades. The W-3 correlation is expressed as a product of five distinct terms, as shown in Eq. (1). Each individual term is defined from Eq. (2) to Eq. (6), respectively:

$$q''_{W3} = F_1 \times F_2 \times F_3 \times F_4 \times F_5 \tag{1}$$

$$F_1 = (2.022 - 0.06238P) + (0.1722 - 0.01427P)\ \exp((18.177 - 0.5987P)X) \tag{2}$$

$$F_2 = [2.326(0.1484 - 1.596X + 0.1729X|X|)G + 3271] \tag{3}$$

$$F_3 = 1.157 - 0.869X \tag{4}$$

$$F_4 = 0.2664 + 0.8357\exp(-124.1D_h) \tag{5}$$

$$F_5 = 0.8258 + 0.0003413(h_{sat} - h_{in}) \tag{6}$$

where the variables and their respective units are defined as follows:
P: System pressure [MPa]
X: Local steam quality [-]
G: Mass flux [kg/m²s]
$D_h$: Hydraulic diameter [m]
$h_{sat}$: Saturation enthalpy [kJ/kg]
$h_{in}$: Inlet enthalpy [kJ/kg]

The W-3 correlation was optimized based on experimental data within specific operating ranges, as shown in Table I. However, the W-3 correlation has a structural weakness due to the exponential function in its first term (shaded part of Eq. (2)), which amplifies small variations in pressure and quality and leads to non-physical divergence outside the original calibration range. This causes the predicted values to diverge unnaturally when the conditions go beyond the original validity range. Fig. 2 illustrates how the W-3 correlation exhibits non-physical explosion depending on the variations in pressure and steam quality.

Table I: Comparison of operational parameters between normal PWR operation, accident conditions, and W-3 validity ranges

| | Normal PWR Operation | Accident Condition (e.g., LOFA/SBLOCA) | W-3 strict range | W-3 Expanded range |
|---|---|---|---|---|
| Pressure [MPa] | – 15.5 | 5.5 – 10 | 6.9 – 15.9 | 5.5 – 20 |
| Mass flux [Mg/m2s] | – 3.5 | < 1.0 | 1.36 – 6.78 | 1.0 – 8.0 |
| Local steam quality | Subcooled | Up to 0.15 | -0.15 – 0.15 | -0.15 – 0.15 |

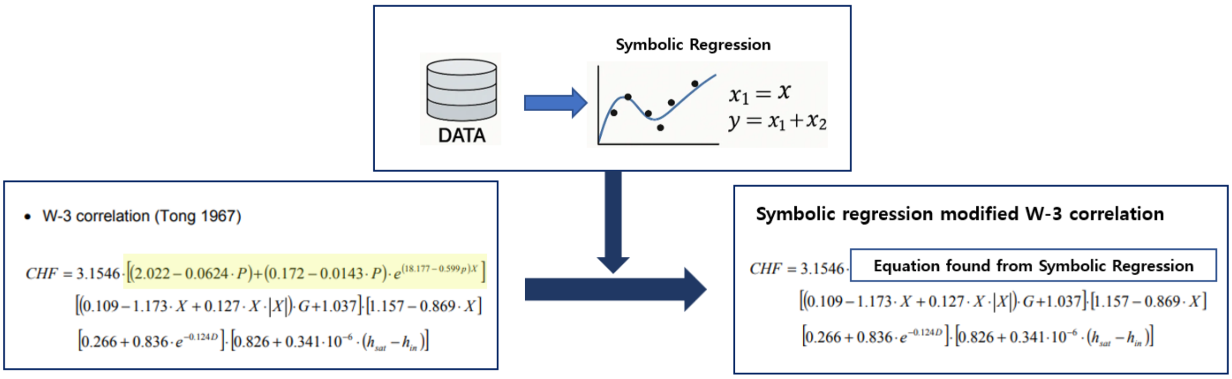


Fig. 1. Hybrid modeling framework of the W-3 correlation using Symbolic Regression.

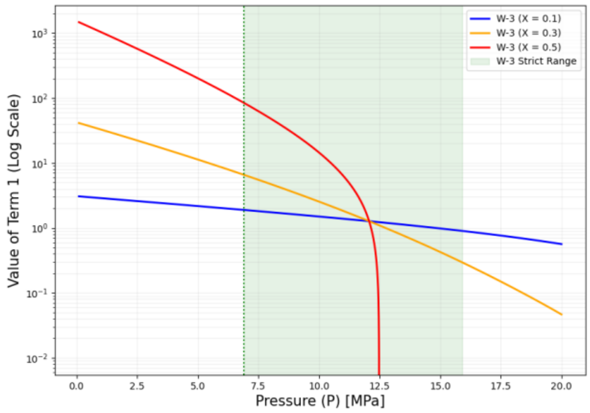


Fig. 2. W-3 Divergence Sensitivity to Steam Quality

During actual plant operation, transient or accident scenarios like a loss-of-flow accident (LOFA) can occur. In these cases, system pressure and flow rates often drop below the range where the W-3 correlation is reliable. The expanded range of interest is presented in Table I. A preliminary analysis using experimental data in the expanded range indicates that the original. As shown in Fig. 3, W-3 correlation exhibits a high prediction error, with a relative RMSE of approximately 32.3%. It shows a strong tendency to over-predict CHF, especially under low-pressure conditions.

In this study, symbolic regression (SR) is applied to improve the accuracy and robustness of the W-3 correlation while maintaining its physical interpretability. By optimizing only the unstable exponential term through SR and preserving the rest of the original framework, a hybrid W-3 correlation is developed. The performance of the proposed model is verified through a detailed comparison with experimental data.

## 2. Methodology

Section 2 describes the composition of the experimental dataset and provides a detailed explanation of the symbolic regression (SR) method and its training conditions.

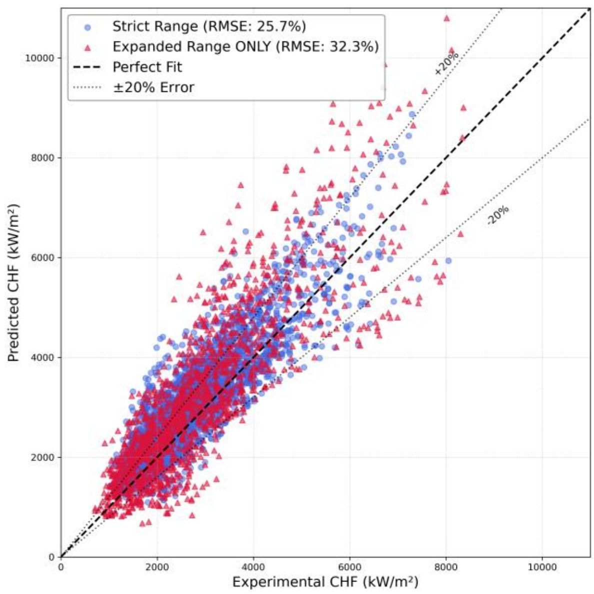


Fig. 3. CHF predictability of the W-3 correlation for strict and expanded operational ranges.

### *2.1 Experimental Dataset*

The CHF experimental data used in this study were obtained from the NRC dataset [2], consisting of a total of 24,579 data points. Among these, 2,994 points fall within the strict validity range of the original W-3 correlation, while 1,638 points are located in the expanded range. In total, 4,632 data points were utilized for the training and evaluation of the proposed model.

*2.2 Symbolic Regression*

Symbolic regression (SR) is a genetic programming-based method that discovers the optimal mathematical expression from a given dataset. In SR, equations are represented as tree structures composed of operators, variables, and constants. The model evolves through mutation and crossover processes to find the best-fit expression. To prevent overfitting, the total number of nodes in the tree is defined as the complexity, which limits the structural depth of the equation. The final expression is selected by calculating a score, which represents the rate of loss reduction relative to the increase in complexity.

In this study, the basic operators used were addition, subtraction, multiplication, division, power, and absolute value. To ensure physical consistency and prevent overly stiff functional forms, the power operator was limited to a single occurrence. Since the original Term 1 of the W-3 correlation only considered pressure and quality, these two variables were selected as input features. The target value for SR training was set as the experimental CHF divided by the product of the remaining W-3 terms. This formulation ensures that the SR model exclusively captures the pressure–quality dependency while preserving the physical structure and calibration of the remaining W-3 terms.

## 3. Results

As a result of the SR training, the final expression was determined at a complexity of 10. As shown in Fig. 4,an analysis of the Pareto front revealed that the rate of error reduction became marginal beyond this complexity level, suggesting that a complexity of 10 provides an optimal balance between accuracy and simplicity.

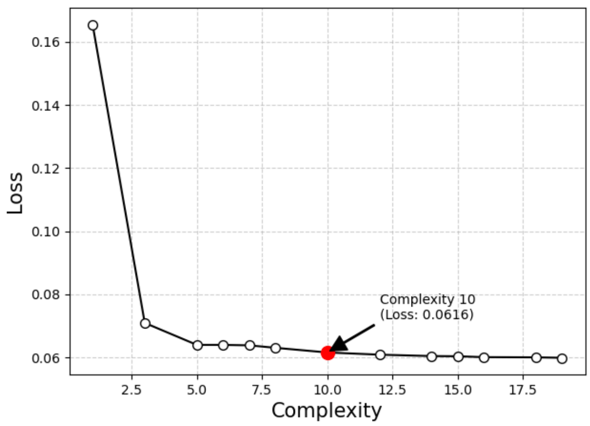


Fig. 4. CHF predictability of the W-3 correlation for strict and expanded operational ranges.

The derived expression for Term 1 is presented in Eq. (7):

$$F_{1\,SR} = \left(\frac{30182}{P} - 1.1658\right)^{|X|+0.38164} \tag{7}$$

where the variables and their respective units are defined as follows:

P: System pressure [MPa]

X: Local steam quality [-]

This equation maintains a monotonic power-law relationship with quality (X), while preventing the numerical divergence observed in the original W-3 correlation. Furthermore, the SR-derived term is structurally simpler than the original. While a direct physical interpretation of each constant may be limited due to the hybrid training approach—where only one term was optimized relative to the existing framework—the model functions as a robust empirical correction for the pressure-quality effect.

The predictive performance was significantly improved by incorporating the SR-derived term into the modified correlation. Within the original validity range (strict range), the relative RMSE was reduced from 25.7% to 19.8%. More importantly, in the expanded range (outside the original limits), the relative RMSE decreased from 32.3% to 26.6%.

The parity plots for the expanded range are displayed in Fig. 5. It is clearly observed that the number of data points falling outside the ±20% error bands is substantially reduced in the modified correlation compared to the original W-3. This demonstrates that the proposed model provides a more stable and reliable CHF prediction.

Table II: Comparison of Relative RMSE (%) between Original W-3 and SR-Modified Model

| | W-3 correlation | SR correlation |
|---|---|---|
| Strict range | 25.7 | 19.8 |
| Outside range | 32.3 | 26.6 |

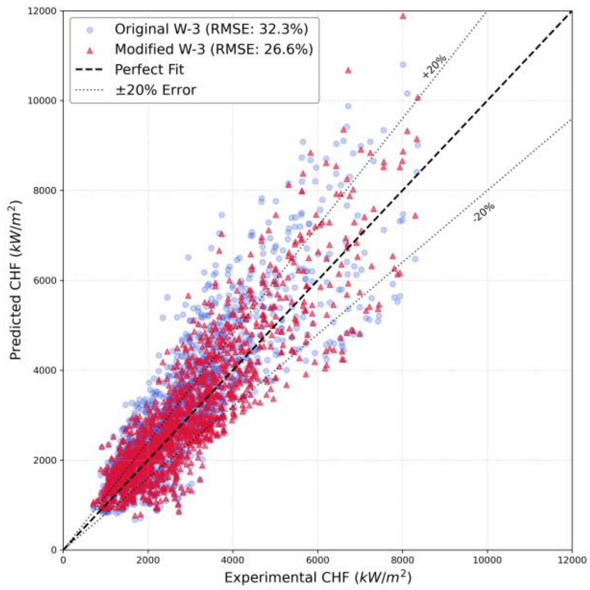


Fig. 5. CHF predictability of the W-3 correlation and modified SR correlation for outside ranges.

## 4. Summary and Conclusions

In this study, a hybrid CHF correlation was developed by integrating symbolic regression (SR) into the existing W-3 framework to improve predictive performance in expanded operational ranges. By replacing the unstable exponential term of the original W-3 correlation with an SR-derived expression, the structural vulnerability leading to non-physical divergence under low-pressure conditions was effectively resolved. The proposed hybrid model demonstrated superior accuracy across all datasets. Specifically, the relative RMSE was reduced from 25.7% to 19.8% in the strict validity range and from 32.3% to 26.6% in the expanded range. The modified correlation showed a stable prediction trend within the ±20% error band, particularly in transient regions such as loss-of-flow scenarios where the original correlation often fails. The results suggest that the SR-based hybrid approach is a powerful tool for developing robust thermal-hydraulic correlations that maintain both physical interpretability and high predictive accuracy. The proposed model is expected to contribute to enhancing the safety margins and reliability of thermal-hydraulic system codes used for PWR safety analysis.

## ACKNLOGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT)(No. : RS-2024-00436693).

## REFERENCES

[1] Tong, L. S, Prediction of departure from nucleate boiling for an axially non-uniform heat flux distribution, Journal of Nuclear Energy, 21(3), 241-248, 1967.

[2] Groeneveld, D. C, Critical heat flux data used to generate the 2006 Groeneveld critical heat flux lookup tables, United States Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, 2019.

[3] Cranmer, Miles, Interpretable machine learning for science with PySR and SymbolicRegression. jl, arXiv preprint arXiv:2305.01582, 2023.