Data-driven nonlinear modeling for superheating degree in organic Rankine cycle systems

Andrés Hernández[1*], Fredy Ruiz[2], Vincent Lemort[1]

[1] *Thermodynamics Laboratory / Université de Liège, Liège, Belgium*
[2] *Dipartimento di Elettronica, Informazione e Bioingegneria/ Politecnico di Milano, Milan, Italy*

* Corresponding Author: jahernandez@uliege.be

ABSTRACT

Organic Rankine Cycle (ORC) power systems have become a promising solution for improving energy efficiency, particularly in waste heat recovery (WHR) applications. These systems convert low-grade heat into electricity, contributing to energy savings and emission reductions. A key control objective is the regulation of the superheating degree, as it directly affects thermodynamic performance and system reliability. This work explores a data-driven modeling strategy using neural networks (NNs) to capture the nonlinear dynamics of the superheating process in a small-scale (11 kWel) ORC unit. To enhance generalization and interpretability, an automatic feature selection framework based on reinforcement learning is developed. The approach evaluates the relevance of multiple candidate input variables, selecting the most informative ones to optimize predictive accuracy. Experimental results show that the proposed model effectively reproduces system behavior while maintaining strong generalization to unseen operating conditions. This modeling framework lays the foundation for advanced control development and contributes to data-driven methodologies for energy systems optimization.

1. INTRODUCTION

Organic Rankine Cycle (ORC) power systems have gained increasing relevance due to their ability to convert low-grade thermal energy into useful electrical power, particularly in waste heat recovery (WHR) applications where conventional technologies are inefficient. By utilizing organic working fluids with favorable thermophysical properties, ORCs operate effectively in low-temperature regimes unsuitable for steam-based cycles (Tchanche *et al.*, 2011).

A key challenge in ORC deployment is the development of accurate dynamic models to support advanced control strategies. In particular, the regulation of the superheating degree at the expander inlet is essential for maximizing thermal efficiency and avoiding damage (Quoilin *et al.*, 2011; Schuster *et al.*, 2009). Traditional modeling approaches, such as empirical or physics-based methods, often fail to capture the strong nonlinearities in ORC dynamics, especially near optimal operating points.

To address these challenges, data-driven models based on neural networks (NNs) have emerged as powerful tools for approximating nonlinear system behavior without requiring explicit knowledge of internal dynamics (Zhou *et al.*, 2023). However, their performance critically depends on selecting relevant input features, especially in systems with many correlated variables and lagged effects.

Recent literature has applied NNs and reinforcement learning (RL) to ORC-related tasks, primarily focusing on control (Egan *et al.*, 2022; Lin *et al.*, 2024) or performance prediction with predefined input sets (Feng *et al.*, 2020). For example, (Lin *et al.*, 2024) investigates sim-to-real transfer for RL control, while (Egan *et al.*, 2022) and (Wang *et al.*, 2020) apply RL to optimize superheat regulation. Other works, such as (Wu *et al.*, 2024; Feng *et al.*, 2023; Oyekale and Oreko, 2023), explore machine learning for system design and optimization but rely on synthetic data and do not address input dimensionality or

dynamic feature selection.

In contrast, this contribution introduces a novel reinforcement learning-based framework for automatic feature selection in dynamic ORC modeling. By embedding policy-gradient RL into the training process, the model identifies the most informative subset of lagged sensor measurements to predict the superheating degree. The resulting neural network achieves high predictive accuracy with a significantly reduced number of inputs, as validated on experimental data. To the best of the authors' knowledge, this work presents the first experimental study of RL-based feature selection for dynamic ORC system identification.

2. ORC EXPERIMENTAL SYSTEM AND SETUP

This section describes the small-scale ORC pilot plant and the experimental procedures carried out to capture the system's nonlinear dynamic behavior.

2.1 ORC Experimental Facility Description

The experimental setup consists of a recuperative ORC system with R245fa as the working fluid, shown schematically in Figure. 1. The nominal electrical output of the system is 11 kW_{el}. The expander is a twin-screw machine mechanically coupled to a permanent magnet synchronous generator connected to the grid via a four-quadrant inverter. This allows flexible speed control, although for the experiments discussed here, the expander speed was fixed at 5,000 rpm for consistency, even though the system supports speeds up to 10,000 rpm. The working fluid is circulated by a variable-speed, 14-stage centrifugal pump with a nominal electrical power of 2.2 kW_{el} and a maximum discharge pressure of 14 bar.

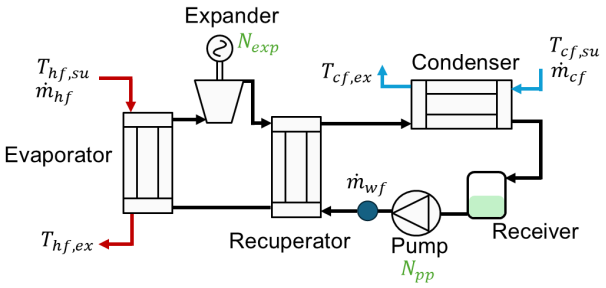


Figure 1: Schematic layout of the ORC pilot facility.

The ORC process starts at the condenser outlet, where the fluid is in saturated liquid state and is collected in a liquid receiver. It is then pressurized by the pump and directed through the cold side of the recuperator and the evaporator. The evaporator utilizes waste heat to bring the fluid to a superheated vapor state. Upon exiting the evaporator, the vapor is expanded in the twin-screw expander and passes through the hot side of the recuperator before returning to the condenser, thus completing the thermodynamic cycle.

The degree of superheating at the expander inlet is a critical parameter influencing both system performance and component safety (Quoilin *et al.*, 2011). It is defined as:

$$\Delta T_{sh} = T_{exp,su} - T_{sat}(P_{evap}) \tag{1}$$

where $T_{exp,su}$ is the measured temperature at the expander inlet, and $T_{sat}(P_{evap})$ is the saturation temperature corresponding to the evaporation pressure, assuming negligible pressure drop in the evaporator:

$$T_{sat}(P_{evap}) = f(P_{evap}) \tag{2}$$

Here, f denotes the nonlinear relationship for R245fa, typically derived from thermodynamic property libraries. Key performance indicators for the ORC include the net electrical output and thermal effi-

ciency:

$$\dot{W}_{net} = \dot{W}_{exp} - \dot{W}_{pump}, \qquad \eta_{cycle} = \frac{\dot{W}_{net}}{\dot{Q}_{in}} \tag{3}$$

where $\dot{W}_{exp}$ and $\dot{W}_{pump}$ are the electrical power outputs of the expander and pump, respectively, and $\dot{Q}_{in}$ is the thermal power supplied to the ORC via the evaporator.

To ensure repeatable experimental conditions, all tests are conducted under controlled settings, including closed-loop regulation of the heat source and sink using dedicated control strategies. Details on the instrumentation, data acquisition system, and secondary loop control design are available in (Desideri *et al.*, 2016; Hernandez *et al.*, 2017).

2.2 Nonlinear Dynamic Response of Superheating Degree

An experimental campaign was conducted to highlight the nonlinear and time-varying dynamics of the ORC system, acquiring data with a sampling time $T_s = 1\,\text{s}$. Figure 2 shows the response of the superheating degree ΔT_{sh} to a staircase variation in pump rotational speed ($N_{\text{pp}} = 50$ rpm), under two distinct heat source mass flow rates at a fixed temperature of $T_{\text{hf}} = 115\,^{\circ}\text{C}$. Although identical step inputs are applied to N_{pp} in both cases, the resulting changes in ΔT_{sh} differ markedly in both amplitude and dynamic profile. At lower heat source flow rate ($\dot{m}_{\text{hf}} = 1.32$ kg/s), the system displays stronger nonlinear behavior, with sharper and less predictable transitions. In contrast, the response at higher flow rate ($\dot{m}_{\text{hf}} = 2.1$ kg/s) is smoother and more gradual. These observations confirm that the system exhibits not only nonlinear dynamics but also time-varying characteristics, with response sensitivity dependent on the thermal boundary conditions which will be constantly changing due to fluctuating nature of the waste heat sources. This complexity underscores the need for flexible, data-driven modeling approaches capable of capturing such varying dynamics.

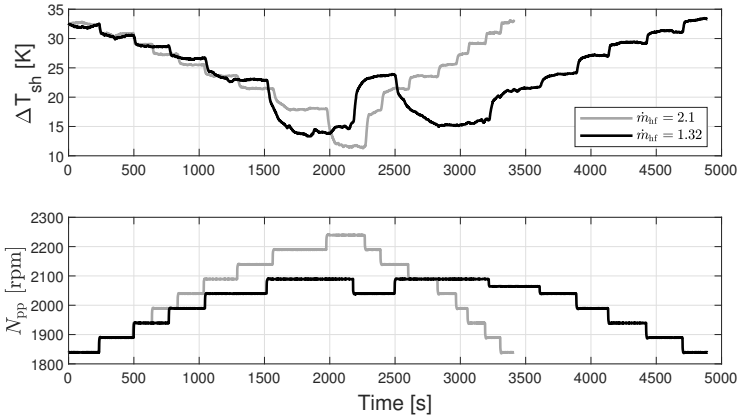


Figure 2: Dynamic response of superheating degree ΔT_{sh} to a staircase variation in pump rotational speed N_{pp}, under two different heat source mass flow rates.

It is worth noting that in the described experiment, the pump rotational speed N_{pp} serves as the manipulated variable, while the heat source conditions—namely the mass flow rate $\dot{m}_{\text{hf}}$ and temperature T_{hf}—are considered measured exogenous inputs. As demonstrated, these thermal conditions have a strong influence on the system dynamics and must be accounted for in any predictive modeling framework.

3. DATA-DRIVEN MODELING METHODOLOGY

This section describes the modeling methodology developed for predicting the superheating degree in an Organic Rankine Cycle (ORC) system using neural networks enhanced with reinforcement learning (RL) for automatic feature selection. The framework consists of three core stages: data preparation, neural network training, and feature selection using the reinforcement learning algorithm.

3.1 Data Preparation and Feature Engineering

Experimental data from a small-scale (11 kWel) ORC unit operating in a waste heat recovery configuration was used. The dataset includes multiple sensor measurements, such as pump speed, temperatures, pressures, mass flow rates, and electrical output. The superheating degree at the expander inlet, denoted $y(t)$, is the modeling target.

To capture system dynamics, a lagged representation was constructed. For a lag order L, the input vector at time t is defined as:

$$\mathbf{x}(t) = [u_1(t), u_1(t-1), \ldots, u_1(t-L), \ldots, u_N(t), \ldots, u_N(t-L), y(t-1), \ldots, y(t-L)]^\top, \tag{4}$$

where $u_i(t)$ are the N candidate input variables. This structure incorporates both control inputs and autoregressive terms from the output (Box *et al.*, 2015). Additionally, all variables were normalized using a Min-Max scaler to lie within $[0,1]$, improving convergence and stability during training (LeCun *et al.*, 2012).

3.2 Neural Network Model

The model used is a fully connected feedforward neural network with two hidden layers. Given an input vector $\mathbf{x}(t) \in \mathbb{R}^d$, the prediction $\hat{y}(t)$ is computed as:

$$\mathbf{h}_1 = \sigma_1(\mathbf{W}_1\mathbf{x}(t) + \mathbf{b}_1), \tag{5}$$

$$\mathbf{h}_2 = \sigma_2(\mathbf{W}_2\mathbf{h}_1 + \mathbf{b}_2), \tag{6}$$

$$\hat{y}(t) = \mathbf{W}_3\mathbf{h}_2 + b_3, \tag{7}$$

where $\mathbf{W}_i$ and $\mathbf{b}_i$ are weight matrices and biases for each layer, and σ_i are activation functions, here a rectified linear unit is employed (i.e., $f(x) = max(0,x)$), as commonly utilized for deep neural networks for nonlinear system identification (Han *et al.*, 2022).

The model is trained by minimizing the mean squared error (MSE):

$$\mathscr{L}_{\text{MSE}} = \frac{1}{T}\sum_{t=1}^{T}(\hat{y}(t) - y(t))^2. \tag{8}$$

3.3 Reinforcement Learning-Based Feature Selection

To identify the most informative subset of features from the candidate pool, a reinforcement learning strategy was employed using the reinforce algorithm (Agarwal *et al.*, 2021). A stochastic policy network is trained to generate binary masks $\mathbf{m} \in \{0,1\}^d$ that select which features are used for each training iteration.

Let $\mathbf{p} = \pi_\theta(\mathbf{x})$ be the output of the policy network, representing the Bernoulli selection probabilities for each feature. A binary mask $\mathbf{m}$ is sampled as:

$$\mathbf{m} \sim \text{Bernoulli}(\mathbf{p}). \tag{9}$$

The selected input becomes $\tilde{\mathbf{x}} = \mathbf{m} \odot \mathbf{x}$, where $\odot$ denotes element-wise multiplication. A neural network is trained using $\tilde{\mathbf{x}}$ and the corresponding reward is computed as the negative RMSE on validation data:

$$RMSE = -\sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{y}(t) - y(t))^2}. \tag{10}$$

The policy parameters θ are updated using the reinforce gradient:

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta}\left[RMSE \cdot \nabla_\theta \log \pi_\theta(\mathbf{m})\right]. \tag{11}$$

To reduce variance, a batch of episodes is averaged, and training proceeds over multiple iterations. At convergence, the mean selection probability for each feature is evaluated, and a probability threshold selection is applied to define the final input subset (Atashgahi *et al.*, 2023).

Let the averaged feature probabilities be denoted by $\mathbf{p} = [p_1, p_2, \ldots, p_d]$. To extract a deterministic feature mask for training or deployment, a selection threshold $\tau \in [0,1]$ is introduced such that:

$$m_i = \begin{cases} 1, & \text{if } p_i > \tau, \\ 0, & \text{otherwise}, \end{cases} \tag{12}$$

where m_i indicates whether the i-th feature is included in the final set. This allows controlling the trade-off between model complexity and accuracy by tuning τ: higher values yield more compact models, while lower values retain more features.

3.4 Model Validation and Evaluation

The final neural network model is trained using the selected features and evaluated on both in-sample and out-of-sample (unseen) experimental data. Performance is assessed using RMSE, MAE, and the coefficient of determination (R^2). The entire pipeline is implemented in PyTorch, allowing for reproducibility and scalability to other datasets or control-oriented applications (Paszke *et al.*, 2019).

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|, \qquad R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{13}$$

4. EXPERIMENTAL RESULTS AND DISCUSSION

This section presents the experimental validation of the proposed neural network-based approach for predicting the superheating degree.

4.1 Experimental Data

The training dataset of 4000 samples was obtained under conditions designed to excite the system dynamically, varying the heat source and manipulated variable (pump speed) over a broad range of amplitudes and frequencies. This is illustrated in Figure 3.

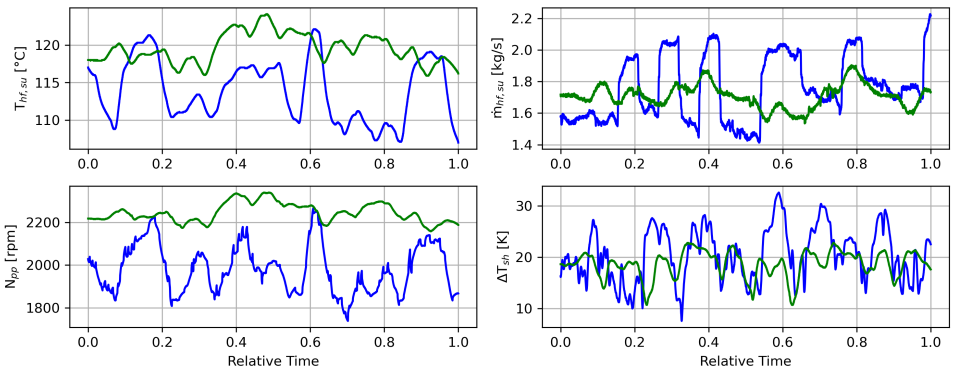


Figure 3: Experimental training (blue line) and unseen (green line) datasets showing system excitation through input perturbations

4.2 Features Candidates and Neural Network Architecture

Without prior knowledge of feature importance, thermodynamic understanding of the ORC system suggests including key sensor measurements for neural network training. These include temperature and mass flow rate from the heat source (indicating heat input), as well as temperature and pressure at the inlet and outlet of the evaporator where phase change and superheating occur. Measurements at the recuperator (inlet/outlet temperatures) are relevant for preheating effects, while pump speed and pressure

influence mass flow rate and evaporation pressure. The expander's fast dynamics make its temperature and pressure useful for predicting future superheating behavior. Lastly, condenser-side measurements (temperature and cooling mass flow) affect pressure and influence the superheating process.

Fifteen physical variables are considered as candidate inputs based on thermodynamic relevance, including heat source conditions ($T_{\mathrm{hf,su}}$, $T_{\mathrm{hf,ex}}$, $\dot{m}_{\mathrm{hf}}$), condenser inlet/outlet temperatures and flow rate ($T_{\mathrm{cf,su}}$, $T_{\mathrm{cf,ex}}$, $\dot{m}_{\mathrm{cf}}$), pump speed (N_{pp}), working fluid flow ($\dot{m}_{\mathrm{wf}}$), and temperatures and pressures at key points ($T_{\mathrm{ev,su}}$, $T_{\mathrm{exp,su}}$, $T_{\mathrm{exp,ex}}$, $p_{\mathrm{exp,su}}$, $p_{\mathrm{exp,ex}}$, $p_{\mathrm{pp,ex}}$, $p_{\mathrm{eva,su}}$). With a lag window of $L = 4$, each input is expanded to include its past four values, yielding $15 \times 4 = 60$ input features. Including 4 lagged values of the output ΔT_{sh} results in a total input vector of 64 features. The reinforcement learning algorithm is then used to automatically select the most informative features from this candidate set.

In general the NN model takes as many inputs as features available (from the 64 candidates described earlier), 64 neurons in the first and second hidden layers, and a single output (the predicted superheating value) with no activation function as this corresponds to a regression value, as depicted in Figure. 4, where four features are selected by the RL algorithm from the different candidates.

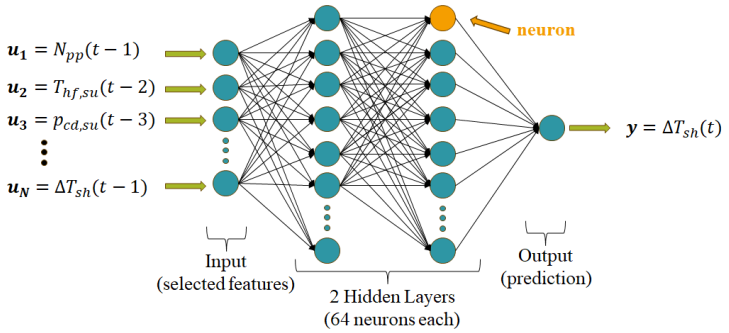


Figure 4: Illustrative example of general neural network architecture for superheating modeling

4.3 Reinforcement Learning-Based Feature Selection

To reduce model complexity while maintaining predictive performance, a reinforcement learning-based policy was employed to perform automatic feature selection. The algorithm evaluated the relevance of each feature by optimizing the prediction reward over multiple episodes. Figure 5 shows the average selection probabilities obtained for each feature across time lags.

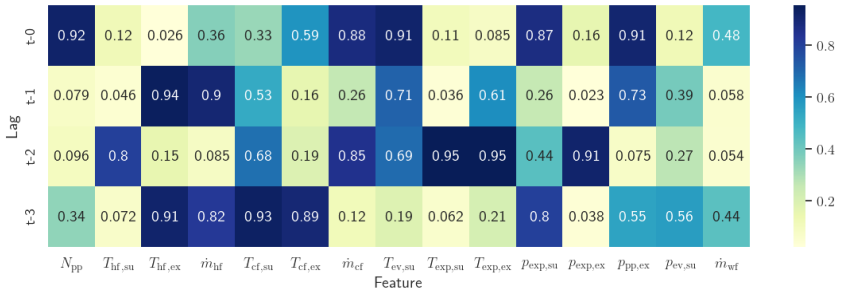


Figure 5: Feature selection probability heatmap from reinforcement learning policy

Based on these learned probabilities, a selection threshold $\tau = 0.6$ was applied to retain the most informative features. The resulting model, selected via reinforcement learning from the full candidate feature set, includes 24 inputs as summarized in Table. 1.

The neural network is trained on the reduced feature set, achieving accurate predictions as shown in Figure 6a. These results confirm the effectiveness of the reinforcement learning strategy in identifying a compact yet informative subset of inputs. From a thermodynamic perspective, the selected features closely align with the underlying physical processes that govern superheating in ORC systems.

Table 1: Selected input features and time lags for $\tau = 0.6$

Feature	Lag(s)	Feature	Lag(s)	Feature	Lag(s)	Feature	Lag(s)
N_{pp}	t	$\dot{m}_{cf}$	$t, t-2$	$T_{ev,su}$	$t, t-1, t-2$	$p_{exp,su}$	$t, t-3$
$p_{pp,ex}$	$t, t-1$	$T_{hf,ex}$	$t-1, t-3$	$T_{hf,su}$	$t-2$	$\dot{m}_{hf}$	$t-1, t-3$
$T_{cf,su}$	$t-2, t-3$	$T_{cf,ex}$	$t-3$	$T_{exp,su}$	$t-2$	$T_{exp,ex}$	$t-1, t-2$
$p_{exp,ex}$	$t-2$	ΔT_{sh}	$t-3, t-4$				

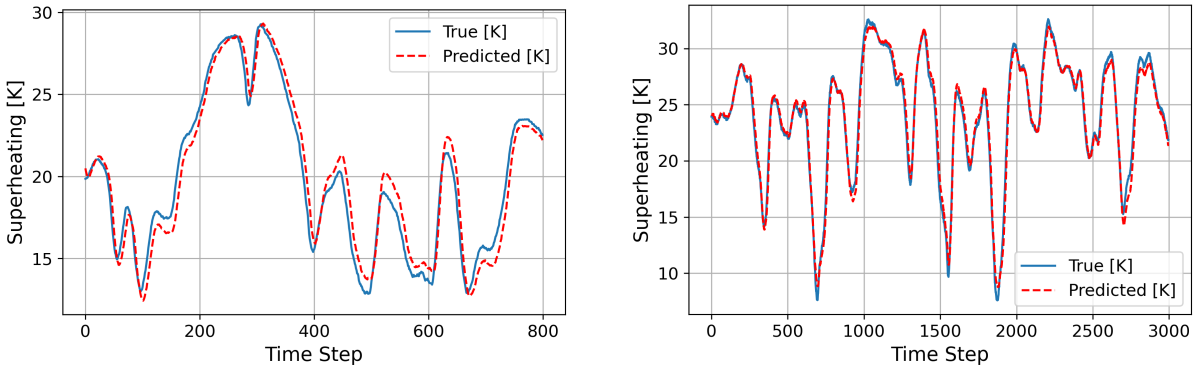
The chosen variables include key heat source characteristics such as evaporator oil inlet/outlet temperatures and mass flow rate, which influence the energy input into the system. Features such as evaporator and expander temperatures and pressures ($T_{ev,su}$, $p_{exp,su}$, $T_{exp,ex}$) are directly related to the phase change process and determine the degree of superheating. Pump speed (N_{pp}) impacts the mass flow rate and pressure of the working fluid entering the evaporator, thus influencing vaporization conditions. Cooling loop variables like condenser fluid inlet temperature and flow rate ($T_{cf,su}$, $\dot{m}_{cf}$) affect the condenser backpressure, which indirectly alters the evaporation pressure and hence the superheating degree.

Moreover, the inclusion of lagged input variables and delayed values of the output ΔT_{sh} reflects the system's thermal inertia and transport delays—dynamic effects that are critical for accurate modeling but are often difficult to capture using first-principles models alone. The reinforcement learning-based feature selection thus achieves a balance between physical interpretability and data-driven predictive power, enabling efficient modeling with reduced complexity.

4.4 Generalization to Unseen Experimental Data

To assess generalization, the trained model was evaluated on a completely independent dataset collected under different dynamic operating conditions but a similar amplitude range. Figure 3 presents in green line the new experimental dataset of 3000 samples used for this purpose.

The predictions on the unseen dataset are shown in Figure 6b. The model retains strong predictive accuracy, demonstrating robustness to shifts in system dynamics and operating regimes, even at very low superheating levels where nonlinearity is stronger (Hernandez *et al.*, 2021). It is worth noting that the prediction capabilities discussed so far are assessed in a one-step-ahead configuration, where at each time step, the lagged inputs of the model are updated using the actual measured values from the system. This setup reflects an idealized scenario and serves to evaluate the intrinsic accuracy of the model without the accumulation of prediction errors.



(a) Training dataset

(b) Unseen validation dataset

Figure 6: One sample a-head prediction performance using selected features ($\tau \geq 0.6$).

4.5 Accuracy vs. Model Complexity: A Trade-off Study

A key advantage of the proposed framework is the ability to balance model simplicity against predictive performance. By varying the probability threshold (τ) applied to the RL-derived feature selection

probabilities, different subsets of features are retained. Table 2 presents the resulting model metrics for selected threshold values.

It is observed that with as few as 18 features (at a threshold of 0.91), the model still achieves over 77% coefficient of determination (R^2) on the unseen dataset. This suggests that meaningful dimensionality reduction is possible without drastically compromising accuracy. However, overly aggressive feature pruning (e.g., probability threshold $\tau \geq 0.95$) leads to significant performance degradation, emphasizing the importance of careful threshold tuning.

Table 2: Performance metrics at selected feature selection thresholds

Threshold τ	**Features**	**Training Data**			**Unseen Data**		
		R^2	RMSE	MAE	R^2	RMSE	MAE
0.00	64	0.9885	0.504	0.407	0.9533	1.154	0.937
0.20	42	0.9786	0.686	0.549	0.9456	1.244	1.022
0.40	39	0.9716	0.790	0.637	0.9312	1.400	1.132
0.60	32	0.9762	0.724	0.582	0.9542	1.142	0.915
0.80	30	0.9522	1.025	0.811	0.8658	1.955	1.624
0.90	20	0.9000	1.483	1.214	0.7393	2.725	2.277
0.91	18	0.9214	1.315	1.050	0.7724	2.546	2.100
0.92	13	0.7789	2.205	1.736	0.3106	4.431	3.726
0.94	11	0.7400	2.392	1.925	0.2805	4.527	3.794
0.96	7	0.8083	2.053	1.668	0.6766	3.035	2.453
0.98	1	0.0684	4.527	3.709	-0.6622	6.881	5.813

4.6 Prediction Capabilities of the Neural Network Simulation Model

To evaluate the predictive performance of the developed neural network model, it is used in simulation mode. The model is initialized with measured data to provide the required lagged inputs (including manipulated and exogenous variables). Subsequent predictions are then based solely on exogenous inputs and previously estimated outputs, without further updates from real output measurements. This sequential prediction approach emulates real-time deployment and serves as a basis for potential applications in control strategies such as Model Predictive Control (MPC). The results, illustrated in Figure 7, demonstrate strong predictive accuracy.

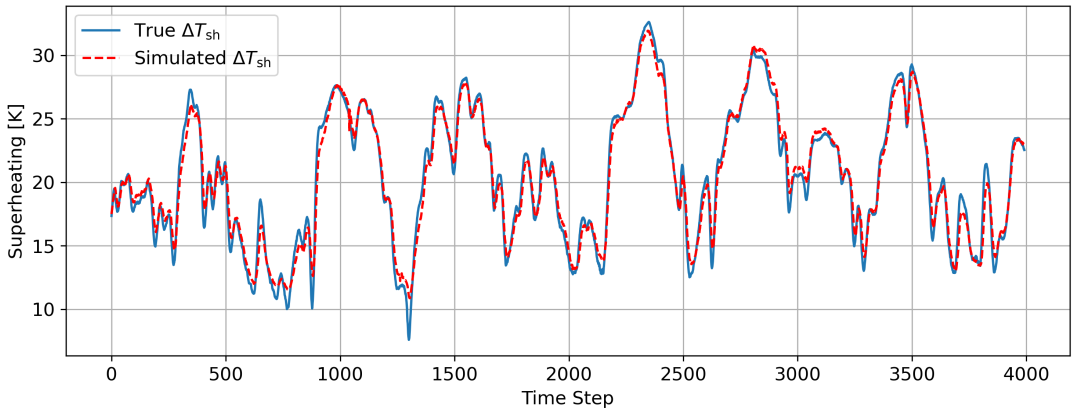


Figure 7: Simulation-based evaluation of the model's prediction capabilities.

Finally, by comparing the one-step ahead error metrics from Table 2—specifically for the 24 selected features in Table 1—with those obtained in simulation mode, it is possible to assess the model's robustness under accumulated prediction error. The one-step ahead evaluation yields $R^2 = 0.9703$, $RMSE = 0.809$, and $MAE = 0.609$, while the simulation mode results in $R^2 = 0.9662$, $RMSE = 0.9544$, and $MAE =$

0.7524. This corresponds to an increase of approximately 18% in RMSE and 24% in MAE. Despite these increases, the performance degradation remains within acceptable bounds, confirming the model's capability for multi-step prediction and potential for use in model-based control applications.

5. CONCLUSIONS

This contribution demonstrates the effectiveness of neural networks for modeling the superheating dynamics in organic Rankine cycle (ORC) systems using experimental data. The key findings are:

- Reinforcement learning-based feature selection successfully identified a compact subset of 24 out of 64 candidate features, reducing the input dimensionality by over 60% while maintaining high accuracy.
- The neural network model trained on the selected features achieved a one-step-ahead prediction performance of $R^2 = 0.9703$, $RMSE = 0.809$, and $MAE = 0.609$, demonstrating strong fit to the measured data with limited input complexity.
- In simulation (sequential prediction) mode, where the model uses only past predictions and inputs without access to measured outputs, the model still attained $R^2 = 0.9662$, $RMSE = 0.9544$, and $MAE = 0.7524$, confirming its robustness and suitability for control-oriented applications.
- The selected features were physically meaningful and consistent with ORC thermodynamics, including variables related to heat input, fluid flow, and lagged effects due to thermal inertia.

The developed framework supports flexible experimentation with neural network architectures and input configurations. Future work will explore its integration into model predictive control (MPC) schemes, where compact yet accurate dynamic models are essential for real-time implementation under varying operating conditions with long prediction horizon.

REFERENCES

Agarwal, A., Kakade, S. M., Lee, J. D., and Mahajan, G. (2021). On the theory of policy gradient methods: Optimality, approximation, and distribution shift.

Atashgahi, Z., Zhang, X., Kichler, N., Liu, S., Yin, L., Pechenizkiy, M., Veldhuis, R., and Mocanu, D. (2023). Supervised feature selection with neuron evolution in sparse neural networks. *Transactions on Machine Learning Research*, 2023(02).

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

Desideri, A., Hernandez, A., Gusev, S., Broek, M. V. D., Lemort, V., and Quoilin, S. (2016). Steady-state and dynamic validation of a small scale waste heat recovery system using the thermocycle modelica library. *Energy*, 115.

Egan, D., Xu, B., Zhu, Q., and Prucka, R. (2022). Reinforcement learning based control of an organic rankine cycle waste heat recovery system over a drive cycle for heavy-duty diesel engines. In *Proceedings of the ASME 2022 ICE Forward Conference.*, Indianapolis, Indiana, USA.

Feng, Y.-q., Liu, Y.-Z., Wang, X., He, Z.-X., Hung, T.-C., Wang, Q., and Xi, H. (2020). Performance prediction and optimization of an organic rankine cycle (orc) for waste heat recovery using back propagation neural network. *Energy Conversion and Management*, 226:113552.

Feng, Y.-q., Xu, K.-j., Zhang, Q., Hung, T.-C., He, Z.-x., Xi, H., and Rasheed, N. (2023). Experimental investigation and machine learning optimization of a small-scale organic rankine cycle. *Applied Thermal Engineering*, 224:120120.

Han, Y., Huang, G., Song, S., Yang, L., Wang, H., and Wang, Y. (2022). Dynamic Neural Networks: A Survey . *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 44(11):7436–7456.

Hernandez, A., Desideri, A., Gusev, S., Ionescu, C.-M., van den Broek, M., Quoilin, S., Lemort, V., and De Keyser, R. (2017). Design and experimental validation of an adaptive control law to maximize the power generation of a small-scale waste heat recovery system. *Applied Energy*, 203:549–559.

Hernandez, A., Ruiz, F., Gusev, S., De Keyser, R., Quoilin, S., and Lemort, V. (2021). Experimental validation of a multiple model predictive control for waste heat recovery organic rankine cycle systems. *Applied Thermal Engineering*, 193:116993.

LeCun, Y., Bottou, L., Orr, G. B., and Müller, K.-R. (2012). Efficient backprop. *Neural Networks: Tricks of the Trade*, pages 9–48.

Lin, R., Luo, Y., Wu, X., Chen, J., Huang, B., Su, H., and Xie, L. (2024). Surrogate empowered sim2real transfer of deep reinforcement learning for orc superheat control. *Applied Energy*, 356:122310.

Oyekale, J. and Oreko, B. (2023). Machine learning for design and optimization of organic rankine cycle plants: A review of current status and future perspectives. *WIREs Energy and Environment*, 12(4):e474.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., *et al.* (2019). Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, volume 32.

Quoilin, S., Aumann, R., Grill, A., Schuster, A., Lemort, V., and Spliethoff, H. (2011). Dynamic modeling and optimal control strategy of waste heat recovery organic rankine cycles. *Applied Energy*, 88(6):2183–2190.

Schuster, A., Karellas, S., Kakaras, E., and Spliethoff, H. (2009). Energetic and economic investigation of organic rankine cycle applications. *Applied Thermal Engineering*, 29(8):1809–1817.

Tchanche, B. F., Lambrinos, G., Frangoudakis, A., and Papadakis, G. (2011). Low-grade heat conversion into power using organic rankine cycles – a review of various applications. *Renewable and Sustainable Energy Reviews*, 15(8):3963–3979.

Wang, X., Wang, R., Jin, M., Shu, G., Tian, H., and Pan, J. (2020). Control of superheat of organic rankine cycle under transient heat source based on deep reinforcement learning. *Applied Energy*, 278:115637.

Wu, X., Lin, L., Xie, L., Chen, J., and Shan, L. (2024). Fast robust optimization of orc based on an artificial neural network for waste heat recovery. *Energy*, 301:131652.

Zhou, X., Fang, S., Zhang, H., Xu, Z., Yao, Y., Gan, H., Zhi, X., Qiu, L., and Wang, K. (2023). Dynamic modeling of a mechanically coupled organic rankine-vapor compression system for compression heat recovery based on an improved lumped parameter model. *Applied Thermal Engineering*, 234:121237.