Chemical Engineering Research and Design 227 (2026) 23–34

Contents lists available at ScienceDirect

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

Attention-based neural network fusion for fouling prediction in Ethylene-Vinyl Acetate heat exchangers

Yellam Naidu Kottavalasa [a,*], Andrea Battaglia [b], Giovanni Bevilacqua [b], Gianni Marchetti [b], Andrea Salfinger [a], Lauro Snidaro [a]

[a] Department of Mathematics, Computer Science and Physics, University of Udine, Udine, 33100, Italy
[b] Research Center, Versalis SpA, Via Giuseppe Taliercio, 14, Mantova, 46100, Italy

ARTICLE INFO

Keywords:
Fouling forecasting
EVA reactor
Deep learning
Heat exchanger system
Industrial polymer process

ABSTRACT

Fouling, a phenomenon in which materials originating from the process fluid settle onto heat-exchange surfaces, significantly reduces thermal efficiency, increases energy consumption, and raises maintenance costs, particularly in high-pressure tubular reactors used for Ethylene-Vinyl Acetate (EVA) polymerization. Accurate, real-time prediction of fouling factor is therefore essential for optimizing operational efficiency, maintaining product quality, and preventing unplanned downtime. This paper proposes a novel attention-based neural network that integrates parallel Bidirectional Gated Recurrent Unit branches with a Multi-Head Attention mechanism to enhance temporal feature extraction and focus on the most informative time steps. In addition to the neural architecture, the framework incorporates Mutual Information-based feature selection stage to retain highly relevant process variables, derived from temperature, pressure, and flow rate measurements collected through sensors across the reactor system. The model was trained on six years of industrial EVA reactor data from Versalis. Experimental results demonstrate that the proposed model consistently outperforms baseline architectures, achieving the lowest test MSE (3.48×10^{-3}), RMSE (4.17×10^{-2}), and highest R^2 (0.82) on normalized data. These improvements highlight the model ability to capture complex temporal dependencies and generalize under varying operational conditions. The proposed approach offers a scalable and effective solution for predictive fouling monitoring in polymerization heat exchangers, with potential applicability across other energy-intensive chemical manufacturing processes.

1. Introduction

Heat exchangers are essential components used across a wide range of sectors, including energy production (Liang et al., 2024), chemical processing (Jradi et al., 2024; Sansana et al., 2024), environmental engineering, electronics, and the food industry (Yin et al., 2024; Krüger et al., 2023). They play a critical role in regulating fluid temperatures either by heating or cooling, and contribute markedly to the overall efficiency and optimization of industrial processes.

Heat exchangers are designed in numerous configurations and operate under diverse thermal environments, each tailored to distinct functional requirements and design constraints. In aerospace and power generation systems, advanced geometries such as film cooled turbine guide vanes and squealer tip turbine rotors are engineered to withstand extreme gas temperatures while maintaining structural integrity through internal convection and film cooling (Jia et al., 2025; Tao et al., 2024). In nuclear safety applications, hybrid air–water passive residual heat removal systems integrate natural convection and evaporative cooling to ensure post-shutdown heat dissipation (Bi et al., 2025). Similarly, thermoelectric generator assemblies operate as coupled conduction–convection exchangers, optimizing heat-to-power conversion across broad temperature gradients (Zhu et al., 2025).

In industrial practice, the most common exchanger types include shell-and-tube configurations frequently used in refineries and petrochemical units (Ikram et al., 2023), cross-flow exchangers typical of acid concentration and chemical recovery plants (Jradi et al., 2022a), air-cooled units in aircraft environmental control systems (Wang et al., 2023a), and tubular reactors with cooling jackets in polymerization processes such as low-density polyethylene or ethylene-vinyl acetate production (Sholahudin Rohman et al., 2022; Kottavalasa et al., 2025). Regardless of their configuration or operating scale, all heat exchangers are susceptible to fouling, which incrementally reduces operational efficiency.

* Corresponding author.
E-mail addresses: kottavalasa.yellamnaidu@spes.uniud.it (Y.N. Kottavalasa), lauro.snidaro@uniud.it (L. Snidaro).

https://doi.org/10.1016/j.cherd.2026.01.043
Received 29 September 2025; Received in revised form 23 December 2025; Accepted 20 January 2026
Available online 21 January 2026

Fouling refers to the gradual buildup of unwanted materials such as solid particles, biological matter, mineral scales, or reaction byproducts on the surfaces of heat exchangers. These deposits may form on either the internal or external surfaces, depending on the properties of the process fluids and the operating environment. As fouling develops, it substantially impairs heat transfer efficiency, increases flow resistance, and may lead to corrosion or contamination of the fluid streams (Davoudi and Vaferi, 2018). The cumulative effect of these issues includes elevated energy consumption, reduced system reliability, and increased operational costs due to the need for frequent maintenance, unplanned shutdowns, and chemical or mechanical cleaning processes (Jradi et al., 2019). Thus, fouling represents one of the most severe operational challenges encountered in industrial heat-exchange systems.

In the context of global energy challenges, the consequences of fouling extend beyond economic losses. Fouling contributes to higher energy consumption and increases the carbon footprint of industrial processes. In a world aiming for sustainable and resilient energy systems, minimizing fouling is essential to improve process efficiency and reduce environmental impact (Kapustenko et al., 2023). This is especially relevant in energy-intensive chemical manufacturing processes, such as polymerization, where reliable heat transfer performance is crucial to maintaining thermal stability, product quality, and operational safety.

Among chemical manufacturing processes, Ethylene-Vinyl Acetate (EVA) polymerization is particularly vulnerable to fouling due to the exothermic characteristics of the reaction and the complex thermal behavior within tubular reactors. The EVA process requires carefully regulated heating and cooling cycles to maintain optimal operating conditions. These thermal variations, along with phase separation between residual monomers and growing polymer chains, facilitate the formation of insulating fouling layers on the inner reactor surfaces. These deposits reduce the efficiency of heat transfer through the cooling jacket, which can cause localized temperature increases, lower production performance, and create potential safety concerns.

Despite being a persistent problem in EVA polymerization, real-time prediction of fouling behavior remains difficult due to the complex and dynamic nature of the process. This challenge arises from the interdependence of multiple process variables, including temperature, pressure, flow rate, and fluid composition, that continuously change during operation (Tang et al., 2020). In industrial practice, the fouling factor represents the additional thermal resistance caused by the formation of a fouling layer, and is widely used as a quantitative indicator of heat transfer degradation. Traditional approaches to estimating the fouling factor typically rely on first-principle models or empirical correlations. However, these methods often assume steady-state conditions and struggle to adapt to fluctuations commonly encountered during normal operation. As a result, they tend to be reactive and detect fouling only after it has already impacted performance, rather than being predictive. Additionally, their reliance on indirect indicators, such as rising temperatures in cooling zones, makes real-time assessment and proactive intervention difficult. These limitations underscore the need for more adaptive and data-driven solutions capable of capturing complex, time-dependent behaviors and providing accurate fouling predictions under varying operating conditions.

To tackle the growing challenge of predicting fouling behavior in industrial heat exchangers, researchers have explored a variety of models from statistical to machine learning using both synthetic and real-world datasets (Kottavalasa and Snidaro, 2025). Early studies focused on empirical methods, i.e., non-parametric, data-driven approaches such as Autoassociative Kernel Regression, which effectively captured fouling degradation trends and enabled Remaining Useful Life estimation using a General Path Model combined with Bayesian updates (Ardsomang et al., 2013). Subsequently, statistical models like Multiple Linear Regression, Sliced Inverse Regression, and Random Forest Regression were employed to estimate cleanliness factors under ash fouling conditions, achieving high precision in covariate selection based on metal tube temperatures (Kumari et al., 2023). In addition to these predictive approaches, detection-oriented strategies have also been investigated, such as impulse-response analysis in shell-and-tube exchangers, which showed high sensitivity to fouling deposits and provided an effective diagnostic framework (Al Hadad et al., 2019). Other approaches have combined Response Surface Methodology (RSM) and Artificial Neural Network (ANN) to estimate fouling resistance in acid concentration processes, where RSM exhibited strong predictive capabilities with a correlation coefficient of 0.99 (Jradi et al., 2022a).

More recent advancements have centered around machine learning techniques such as Support Vector Regression, Bagged Trees, and Gaussian Process Regression, the latter achieving an impressive R^2 of 0.98 in fouling factor prediction (Hosseini et al., 2022). Similarly, advanced neural network models, including Feedforward Neural Network (FFNN), Nonlinear Autoregressive Networks with Exogenous Inputs, and Long Short-term Memory (LSTM)-based frameworks, have demonstrated exceptional accuracy in capturing nonlinear patterns associated with fouling buildup (Ikram et al., 2023; Wang et al., 2023a), while hybrid architectures that merge recurrent networks with attention mechanisms further improved forecasting accuracy (Kottavalasa et al., 2025). Specific architectures have been tailored for domain-specific applications, such as phosphoric acid concentration processing using ANN with Broyden–Fletcher–Goldfarb–Shanno (BFGS) training, Low-Density Polyethylene (LDPE) tubular reactors using FFNN trained via the Levenberg–Marquardt algorithm, and membrane filtration systems using Autoregressive Integrated Moving-Average and LSTM time series models by leveraging temperature, flow rate, density, and pressure as key input features for accurate fouling estimation (Jradi et al., 2022b; Sholahudin Rohman et al., 2022; Krüger et al., 2023). Likewise, Davoudi and Moghadas (2022) investigated multiple ANN paradigms and Adaptive Neural-based Fuzzy Inference System for fouling factor prediction in hot wire probes, with cascade feed-forward networks yielding the best performance. In addition, Yin et al. (2024) proposed an approach in which an Encoder–Decoder LSTM network was used to predict output temperature in heat exchangers, followed by a Random Forest model to estimate fouling. Similarly, in the case of ethylene oxide production, a hybrid framework integrating knowledge-based feature engineering with data-driven models was used to forecast fouling indicators one month in advance, achieving a predictive accuracy of $R^2 = 0.7$ (Sansana et al., 2024).

Across many of these studies, ANNs remain a popular modeling choice due to their adaptability and high predictive accuracy. Although deep networks can learn relevant representations directly from raw data, several works have reported that applying feature selection techniques, such as correlation analysis, principal component analysis, or sensitivity-based ranking, can help reduce redundancy and stabilize model performance in industrial environments (Jradi et al., 2023; Davoudi and Vaferi, 2018). Despite promising results, challenges persist in capturing complex temporal dependencies and ensuring robustness under non-stationary conditions typical of continuous industrial operations. These non-stationary effects result from shifts in production grade, variations in feed composition, fluctuations in hot-water flow within the jacket, and sensor drift developing over extended operating periods. Although notable progress has been made, research on fusion-based deep learning models that combine sequential learning with attention mechanisms for time-series fouling factor prediction remains limited (Kottavalasa et al., 2025). This limitation is especially evident in high-risk chemical processes where accurate forecasting is fundamental.

To address this gap, the present study introduces a novel deep learning framework that adopts a Bidirectional GRU-Bidirectional GRU (BiGRU-BiGRU) architecture with a Multi-Head Attention (MHA) mechanism. This hybrid design enables the model to capture both past and future temporal dependencies while dynamically attending to the

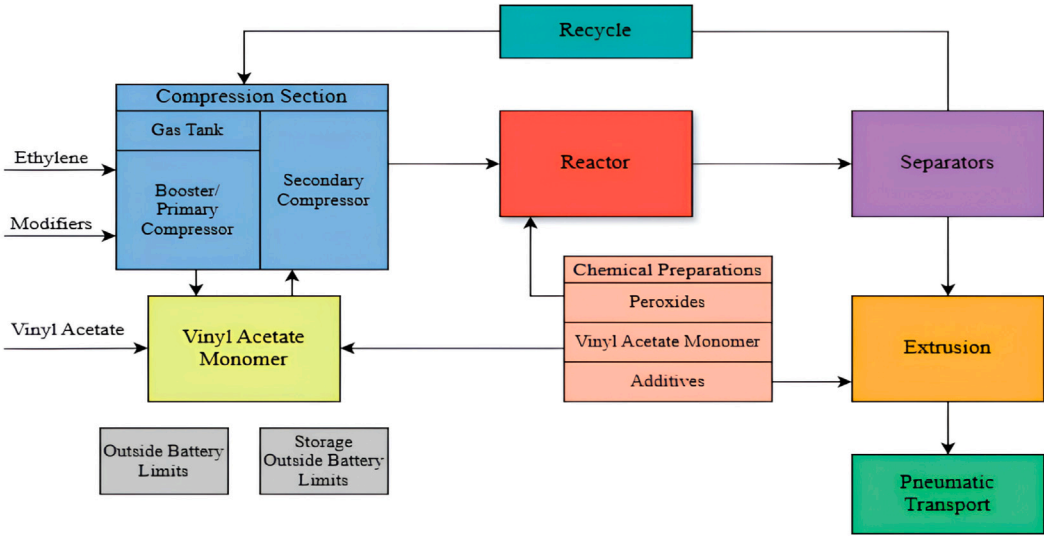


Fig. 1. Process flow diagram of the EVA polymerization line at the Versalis plant. The process begins at the raw material inlet (indicated by the starting arrow) and proceeds through the main units and auxiliaries.

most informative time steps. In addition, we employ Mutual Information (MI) for feature selection to retain only the most relevant input features and enhance generalization. By combining the strengths of sequence modeling and attention-based time-step weighting, the proposed approach offers a scalable and practical solution for real-time fouling prediction in EVA polymer reactors. Experimental results using real industrial data from Versalis demonstrate that the proposed model achieves better performance than other neural network architectures in terms of prediction accuracy.

2. Industrial system description and data preparation pipeline

This section describes the industrial heat exchanger system and the methodology used to derive the fouling factor, which serves as the target variable for model training. The dataset collected from on-site monitoring is introduced, followed by the data preprocessing workflow, feature selection method, and normalization procedure applied to prepare the input features for the predictive model.

2.1. Description of heat exchanger system

The EVA polymerization process at Versalis is carried out in a high-pressure tubular reactor system designed to operate under extreme thermal and mechanical conditions. Specifically, the tubular reactor functions at pressures ranging from 2000 to 2600 bar and temperatures between 180 °C and 280 °C, as required for the free radical polymerization of ethylene and vinyl acetate. The reactor is structurally designed as a long cylindrical tube, divided into multiple reaction zones that support different stages of the process: preheating, initiator injection, polymerization, and cooling. Surrounding this tubular reactor is a jacketed shell that serves as a pressurized hot water-based heat exchanger system, which dynamically regulates temperature along the reactors length. This configuration ensures that excess heat generated during the exothermic polymerization process is efficiently removed and that zone-specific thermal profiles are maintained to support precise conversion rates and safe operation. Reactants and peroxide initiators are strategically injected at multiple points to manage the initiation and progression of polymer chains (Versalis, 2022). The overall process layout of the EVA polymerization line at the Versalis plant, including the compression section, reactor, separators, extrusion, and auxiliary units, is illustrated in Fig. 1. This schematic outlines the industrial context in which the heat exchanger system operates.

The heat exchanger system is valuable not only for maintaining thermal stability and product quality but also for preventing localized overheating, which could otherwise degrade reactor performance and safety. The integration of this thermal control infrastructure enables real-time heat load balancing across varying operational conditions (Versalis, 2022).

2.2. Fouling factor calculation

In order to quantify the extent of deposit buildup and its impact on heat transfer performance, the fouling factor (FFR_i) for each reactor zone i is computed from the difference between the overall heat transfer coefficients under fouled and clean operating conditions. This calculation, based on on-site process and utility measurements, provides a direct indicator of the additional thermal resistance caused by fouling (Versalis, 2022).

The clean overall heat transfer coefficient is determined from the process-side and utility-side convection coefficients, $h_{r,i}$ and $h_{j,i}$, as Sholahudin Rohman et al. (2022), Jradi et al. (2022b):

$$U_{\text{clean},i} = \left(\frac{1}{h_{r,i}} + \frac{1}{h_{p,i}} + \frac{D_{r,i}}{D_{ci,i}} \cdot \frac{1}{h_{j,i}} \right)^{-1} \quad \left[\frac{W}{\text{m}^2 \cdot K} \right] \tag{1}$$

where $D_{r,i}$ and $D_{ci,i}$ are the reactor and channel diameters (m), respectively.

The fouled overall heat transfer coefficient is calculated from the measured heat duty and the log mean temperature difference (ΔT_{mli}), obtained as Jradi et al. (2022b), Sansana et al. (2024):

$$\Delta T_{mli} = \frac{(T_{r,iA} - T_{j,iA}) - (T_{r,iB} - T_{j,iB})}{\ln\left(\frac{T_{r,iA} - T_{j,iA}}{T_{r,iB} - T_{j,iB}} \right)} \quad [K] \tag{2}$$

$$U_{\text{fouled},i} = \frac{Q_{\text{jack},i} \cdot 4187}{3600 \cdot S_i \cdot \Delta T_{mli}} \quad \left[\frac{W}{\text{m}^2 \cdot K} \right] \tag{3}$$

where $Q_{\text{jack},i}$ is the heat exchanged in the jacket (kW), S_i is the heat transfer surface area (m^2), and 4187 is the specific heat capacity of water (J/kg· K). T_r and T_j denote the reactor and jacket temperatures at points A and B.

Finally, the fouling factor is obtained as Jradi et al. (2022b):

$$FFR_i = \frac{1}{U_{\text{fouled},i}} - \frac{1}{U_{\text{clean},i}} \quad \left[\frac{\text{m}^2 \cdot K}{W} \right] \tag{4}$$

which expresses the additional thermal resistance introduced by fouling layers. This parameter is computed for each zone and used as the target variable for subsequent model training.

2.3. Data from on-site monitoring

The dataset used in this study was collected from the EVA polymerization reactor system at the Versalis plant, covering the operational period from 2018 to 2024. It consists of 43,768 time-series records with 26 process variables and one target variable (i.e., fouling factor), which was computed using the heat-transfer formula described in Eq. (4). These process variables include key parameters such as temperature, pressure, and flow rates, measured by industrial-grade sensors positioned along the reactor. All measurements were recorded at a fixed sampling frequency of 15 min throughout the operating period. This real-time operational data reflects the dynamic behavior of the process under varying conditions and serves as the foundation for training and evaluating the proposed predictive model.

2.4. Data preprocessing

Several data cleaning steps were performed to enhance the reliability of the dataset and support stable model development. The preprocessing workflow was carefully designed to ensure accurate representation of operational behavior and includes:

- **Filtering of Inactive States:** Records corresponding to plant shutdowns or non-operational conditions were removed by applying threshold-based filters on important process variables such as system pressure, feed flow, and reactor temperature.
- **Handling Missing Values:** An Iterative Imputer strategy was employed to estimate missing entries. This approach allowed for the preservation of complex inter-feature relationships and minimized data loss.
- **Outlier Removal:** Unrealistic spikes in the computed fouling values, often due to sensor drift or noise, were addressed using the Interquartile Range (IQR) method. This statistical approach removed values falling outside the typical distribution range, improving data integrity while preserving genuine operational trends.

2.5. Feature selection using mutual information

Mutual information (MI) was employed for feature selection to reduce model complexity and improve generalization. MI is a widely used statistical measure that quantifies both linear and non-linear dependencies between input features and the target variable. Unlike correlation coefficients, which only capture linear relationships, MI estimates the shared information between random variables without assuming any specific distribution (Jain and Murthy, 2015; Ryu et al., 2018).

Mathematically, the MI between two discrete random variables X and Y is defined as:

$$I(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right) \tag{5}$$

where $p(x)$ and $p(y)$ are the marginal probability distributions of X and Y, and $p(x,y)$ is their joint probability distribution. If X and Y are independent, $I(X;Y) = 0$, whereas higher values indicate stronger dependencies.

In this study, MI scores were computed between each input feature and the target variable using the raw data. Based on the recommendations of domain experts from the plant, a threshold of 0.45 was applied to select features with significant influence on the fouling factor. This process resulted in a refined subset of 22 variables that were subsequently used to train the proposed predictive model. Although only four variables were excluded, these features exhibited very low MI scores and were known to be noise-prone, and their removal contributed to more stable model training. The selected features are primarily associated with reactor- and utility-side operational conditions, representing variations in thermal and flow behavior that govern heat-transfer performance within the system. These variables collectively capture both direct and indirect effects on fouling evolution.

2.6. Normalization

Normalization is the process of rescaling numerical features to a common range to ensure consistent input magnitudes for model training. After the preprocessing steps, Min–Max normalization was applied to transform each feature into a fixed range of [0, 1], helping the model converge more efficiently. The transformation is computed using the following equation:

$$x_{\text{normalized}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{6}$$

Where x is the original feature value, and $x_{\min}$, $x_{\max}$ are the minimum and maximum values of that feature, respectively.

3. Proposed model architecture

The objective of the proposed architecture is to improve the predictive capability for time series fouling factor forecasting by integrating complementary temporal features extracted from parallel Bidirectional GRU layers with MHA. This section provides a detailed explanation with the mathematical formulations of the BiGRU and MHA components, followed by a high-level overview of the proposed fusion architecture.

3.1. Bidirectional gated recurrent unit

The Gated Recurrent Unit (GRU) (Bahdanau et al., 2014) is a simplified variant of the LSTM (Hochreiter and Schmidhuber, 1997) designed to capture sequential dependencies in time series data using two gating mechanisms for information update and reset. Its compact setup reduces computational complexity while mitigating vanishing gradient issues and preserving the ability to learn temporal patterns (Li et al., 2020). The resulting hidden state update can be expressed as:

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \tag{7}$$

To further enhance contextual understanding of sequential data, a Bidirectional GRU (BiGRU) is adopted (Li et al., 2021). BiGRU processes the input in both forward and backward directions, allowing the model to consider both past and future temporal dependencies. At each time step, the forward GRU computes a hidden state $\vec{h}_t$, and the backward GRU computes $\overleftarrow{h}_t$. These are then concatenated as follows:

$$\vec{h}_t = \text{GRU}_{\text{forward}}(x_t) \tag{8}$$

$$\overleftarrow{h}_t = \text{GRU}_{\text{backward}}(x_t) \tag{9}$$

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \tag{10}$$

Fig. 2 compares the internal structures of a GRU cell (a) and a BiGRU layer (b), highlighting the differences in temporal context processing.

In this study, we employ a parallel BiGRU-BiGRU architecture, in which two BiGRU blocks are run independently on the same input sequence to extract diverse representations. The outputs of these parallel BiGRUs are subsequently fused using a Multi-Head Attention mechanism (detailed in Section 3.2) to generate a more richer and context-aware representation for final fouling prediction.

3.2. Multi-head attention mechanism

Multi-Head Attention (MHA) enables the network to attend to different temporal relationships in parallel by projecting an input sequence into multiple query Q, key K, and value V spaces and computing attention weights across them (Wang et al., 2023b). The core operation applied within each head is the scaled dot-product attention:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{11}$$

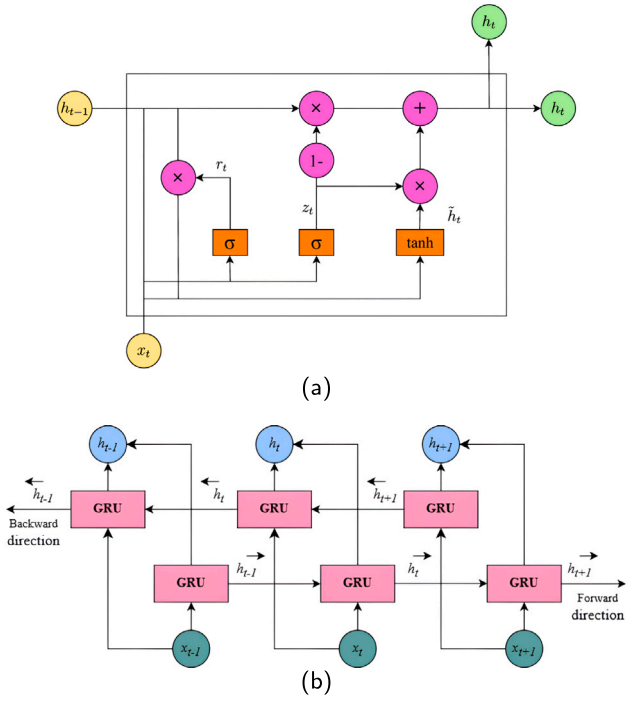


Fig. 2. Architecture of (a) a GRU cell, and (b) a BiGRU layer, consisting of a pair of forward and backward GRU cells at each time step.

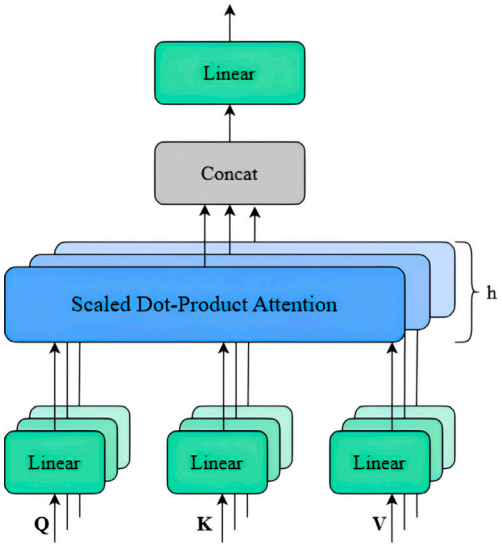


Fig. 3. Schematic diagram of the Multi-Head Attention mechanism. Each head performs scaled dot-product attention, and the outputs are concatenated and linearly transformed.

where d_k denotes the dimensionality of the key vectors. This operation allows the model to focus on the most relevant parts of the sequence when forming internal representations (Liu et al., 2024).

Outputs from individual heads are then concatenated and linearly transformed to produce the aggregated attention representation used in our model (Fig. 3).

3.3. Parallel BiGRU-BiGRU with MHA

The final forecasting architecture of our study fuses the strengths of recurrent sequence modeling and attention mechanisms to enhance the predictive performance on fouling factor time series. As shown in Fig. 4, the proposed model is composed of two parallel BiGRU branches followed by a Multi-Head Attention module and a compact feedforward output head.

The input sequence is fed simultaneously into two independent BiGRU encoders. Each branch consists of stacked BiGRU layers, enabling the network to capture rich bidirectional time-evolving dependencies from different representational perspectives. Batch normalization layers are inserted between BiGRU stacks to stabilize training and reduce internal covariate shift, promoting faster convergence.

The outputs from both BiGRU branches retain the full temporal resolution and are denoted as H_1 and H_2, respectively. These outputs are passed to the MHA block, where H_1 acts as the query matrix and H_2 serves as both the key and value matrices. This asymmetric configuration encourages the model to perform cross-branch attention and learn dependencies between independently learned sequential representations across time steps:

$$\mathrm{MHA}(Q,K,V) = \mathrm{MultiHead}(Q = H_1, K = H_2, V = H_2) \tag{12}$$

After computing scaled dot-product attention across multiple heads, the resulting attended representation is projected through a linear layer and subsequently reduced using a Global Average Pooling operation. This compresses the sequential dimension into a fixed-length feature vector that serves as the final sequence encoding.

The pooled vector is passed through a dense layer with ReLU activation, followed by a dropout layer for regularization. Finally, a single-unit output layer produces the predicted fouling factor value. The architectural design enables the model to combine mutually reinforcing chronological features with attention over time steps, and its impact on prediction accuracy and error distribution is examined in the experimental analysis (Table 1, Figs. 6 and 10). All experiments were implemented in Python using TensorFlow and Keras, while data preprocessing and analysis were conducted with Pandas, NumPy, and Scikit-learn.

4. Experimental results

The following results present the outcome of training and evaluating the proposed model alongside baseline architectures. Hyperparameter settings, evaluation metrics, comparative performance analysis, and ablation study are reported to showcase the effectiveness of the model under the given operational data.

4.1. Hyperparameter settings

The hyperparameters of the proposed fusion architecture were determined through iterative experimentation and comparative evaluation. Key parameters such as the number of BiGRU units, dropout rate, number of attention heads, learning rate, batch size, and sequence window length were systematically varied to assess their influence on model prediction accuracy. The time sequence window was tested with lengths of 15, 30, 45, and 90 time steps, with 30 providing the best trade-off between capturing short-term and long-term dependencies. In this configuration, the model performs single-step forecasting, where each time step corresponds to a 15-minute sampling interval; thus, the preceding 30 time steps represent 7.5 h of operational history used to predict the fouling factor at the subsequent time step. The final configuration employed two parallel BiGRU branches with 64 and 32 hidden units in the first and second layers, respectively, a dropout rate of 0.1, and the Adam optimizer with a learning rate of 0.001. The model was trained using a batch size of 64, with Mean Squared Error (MSE) as the loss function. The dataset is time-series data and was split in chronological order into 80% training and 20% testing, corresponding to 34,990 and 8784 samples, respectively.

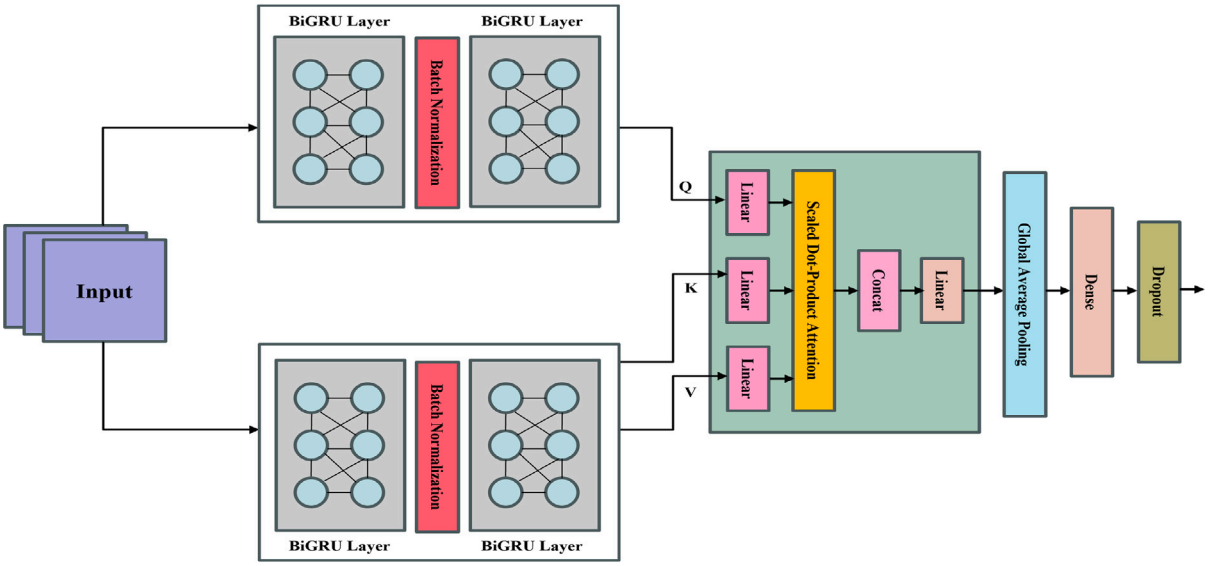


Fig. 4. Proposed attention-based fusion neural architecture for fouling prediction in EVA heat exchangers.

Table 1
Comparison of model performance on train and test sets using MSE, RMSE, R^2, and parameter counts.

Model	Train Set (Normalized)					Test Set (Normalized)					Params (k)
	MSE ($\times 10^{-3}$)	MSE σ ($\times 10^{-2}$)	RMSE ($\times 10^{-2}$)	RMSE σ ($\times 10^{-2}$)	R^2	MSE ($\times 10^{-3}$)	MSE σ ($\times 10^{-2}$)	RMSE ($\times 10^{-2}$)	RMSE σ ($\times 10^{-2}$)	R^2	
LSTM	6.00	1.57	5.69	5.26	0.80	6.51	1.69	5.66	5.75	0.67	210.6
BiLSTM	6.11	1.66	5.47	5.59	0.79	7.14	1.79	6.07	5.88	0.64	203.9
GRU	4.66	1.29	4.92	4.73	0.84	6.11	1.41	5.69	5.35	0.69	216.8
BiGRU	4.26	1.47	4.81	4.41	0.86	5.17	1.69	4.72	5.43	0.74	179.2
LSTM-GRU	5.66	1.73	4.92	5.69	0.81	7.70	2.11	5.73	6.64	0.61	155.0
BiLSTM-BiGRU	5.41	1.88	5.55	4.82	0.82	4.54	1.21	5.02	4.49	0.77	380.9
Parallel BiLSTM	3.15	1.39	3.63	4.28	0.89	5.58	1.52	5.28	5.29	0.72	193.5
Parallel BiGRU	3.79	1.27	4.32	4.39	0.87	4.73	1.26	4.84	4.89	0.76	182.2
Proposed Model	**2.88**	**1.19**	**3.50**	**4.07**	**0.90**	**3.48**	**1.07**	**4.17**	**4.18**	**0.82**	**151.1**

4.2. Evaluation metrics

Evaluating the prediction accuracy of the proposed model for fouling factor forecasting is a crucial step in identifying the optimal configuration and benchmarking it against alternative time series approaches. In this study, three widely used statistical error metrics are employed to quantify the deviation between predicted outputs and observed values, as defined below:

Mean Squared Error (MSE): Measures the mean of the squared differences between the predicted and actual fouling factor values. This metric penalizes larger errors more heavily, making it useful for identifying cases where the model significantly deviates from observed values:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Root Mean Squared Error (RMSE): Obtained by taking the square root of the Mean Squared Error, RMSE restores the error metric to the same units as the target variable (fouling values). It offers a direct interpretation of the average magnitude of prediction errors, giving a clearer sense of the practical scale of deviations:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Coefficient of Determination (R^2): Indicates the proportion of variance in the actual fouling factor values explained by the model. A higher R^2 value reflects a better fit and stronger predictive capability. This metric is particularly useful for assessing how effectively the model captures variability arising from different operating conditions and process fluctuations:

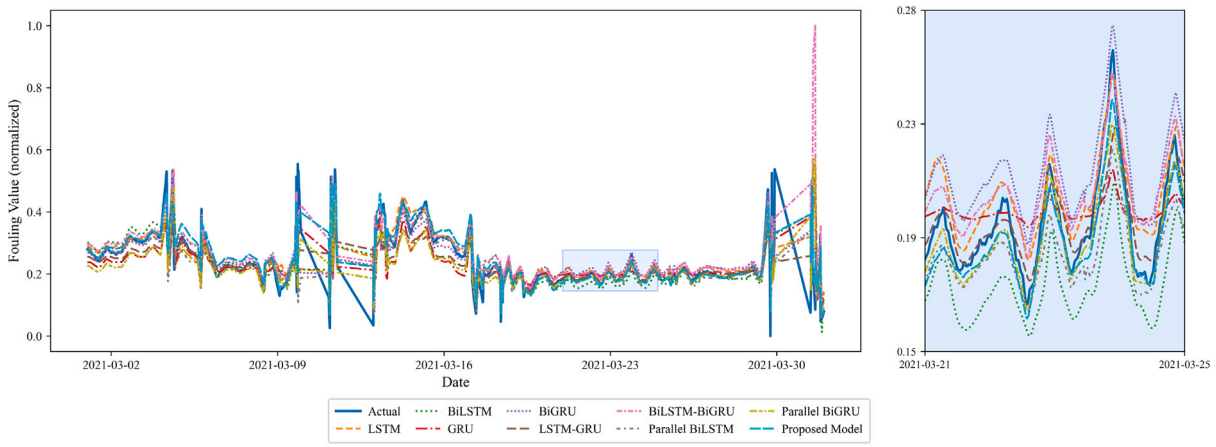
$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Here, y_i and $\hat{y}_i$ represent the actual and predicted fouling factor values, respectively, n is the total number of samples, and $\bar{y}$ is the mean of the actual values.

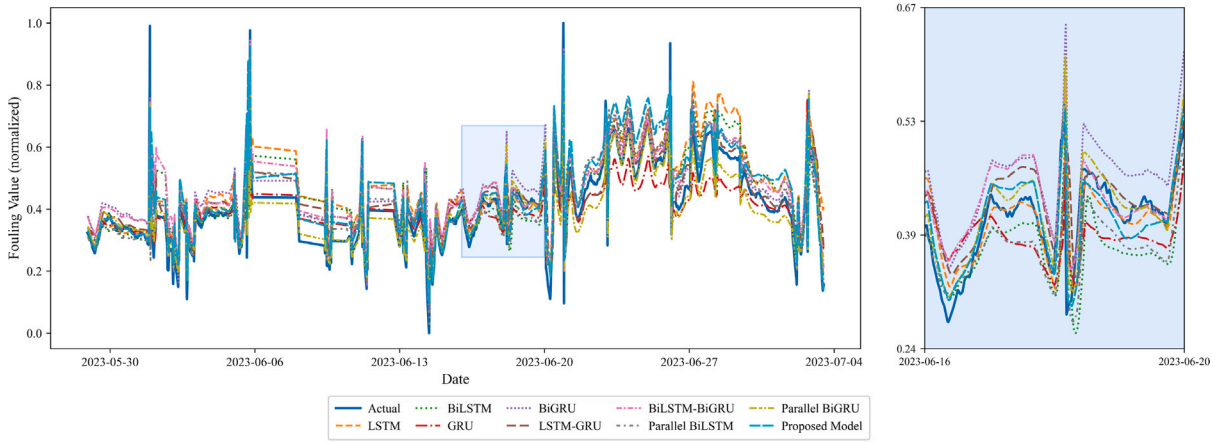
4.3. Model performance comparison

Table 1 presents a comparative analysis of the proposed fusion-based BiGRU-BiGRU with MHA architecture against a diverse set of recurrent baselines, including single-layer (LSTM, GRU), bidirectional (BiLSTM, BiGRU), hybrid (LSTM-GRU, BiLSTM-BiGRU), and parallel variants (Parallel BiLSTM, Parallel BiGRU). In the hybrid models, the two distinct recurrent layers operate in parallel on the same input sequence, and their respective feature maps are concatenated before the pooling stage to combine heterogeneous temporal representations. The parallel variants comprise two identical bidirectional branches that learn different aspects of sequential behavior from the same input data, with their outputs merged at the feature level through concatenation. In the developed BiGRU-BiGRU fusion model, temporal feature representations learned by the two parallel BiGRU branches are fused through an attention-based integration strategy, as detailed in Section 3. The comparison is conducted on both the training and test sets using normalized data, in compliance with plant confidentiality constraints.

Across all metrics, the proposed novel model consistently outperforms the baseline architectures. On the test set, it achieves the lowest



(a) Train set



(b) Test set

Fig. 5. Actual versus predicted fouling factor values for the (a) train set (March 2021) and (b) test set (June 2023) across various recurrent models and the proposed architecture.

MSE (3.48×10^{-3}) and RMSE (4.17×10^{-2}), alongside the highest R^2 score (0.82), indicating improved predictive accuracy and variance explanation compared to the alternatives. Relative to the standard LSTM baseline, the proposed model reduces RMSE by approximately 26.3% and improves R^2 by 22.4%, revealing its superior ability to capture complex sequential dependencies in fouling progression. In addition to its high accuracy, the model exhibits the lowest variability, with MSE standard deviation (1.07×10^{-2}) and RMSE standard deviation (4.18×10^{-2}), confirming its stability against fluctuations in the data.

The performance gain is also evident in the training set results, where the proposed model attains the lowest MSE (2.88×10^{-3}), RMSE (3.50×10^{-2}) and highest R^2 (0.90), suggesting an effective balance between fitting the training data and forecasting coherence on unseen data. The sustained reduction in error metrics across both datasets indicates that the proposed fusion strategy, combining supplementary bidirectional dynamic representations with attention-driven feature weighting, enhances the model resilience to unseen operational conditions. The significance of the attention mechanism in the predictive process is visualized in the attention weight matrix (Fig. 6), which represents the average distribution learned over the test set. Each element in this heatmap quantifies the relative importance assigned by the model to the contextual time steps when predicting the current output step. Higher intensity regions denote stronger dependencies, indicating that the model focuses more on recent and temporally relevant observations when estimating the response variable (i.e., fouling factor). The structure of the attention weights reveals that the most recent segments exert a dominant influence, an expected behavior in industrial processes where fouling is closely tied to immediate operational history. At the same time, non-zero attention assigned to earlier time steps suggests that the model leverages longer-term contextual information, reflecting the cumulative and history dependent nature of fouling trajectory. Besides accuracy, Table 1 also reports the number of trainable parameters, offering insight into model complexity. The proposed model achieves superior results with only 151k parameters, fewer than all compared methods. This is because each BiGRU branch uses smaller hidden dimensions than the single stream baselines, confirming that the improvements stem from the fusion and attention design rather than from a larger model size, as supported by the ablation analysis (Section 4.4).

In addition to the quantitative results summarized in Table 1, Fig. 5 provides a visual comparison between the actual and predicted fouling factor values for both the training and test sets. The proposed model tracks the real trend more closely than the other recurrent architectures, capturing both local fluctuations and broader progression patterns with reduced deviation. Moreover, Fig. 7 shows a strong alignment of data points along the equality line ($y = x$). This close correspondence between measured and estimated values indicates minimal bias and high predictive consistency across datasets. This visual evidence further substantiates the enhancements in forecasting precision and steadiness across conditions achieved by the proposed approach.

It is important to note that the apparent sudden variations visible in both the training (Fig. 5(a)) and test (Fig. 5(b)) curves reflect the

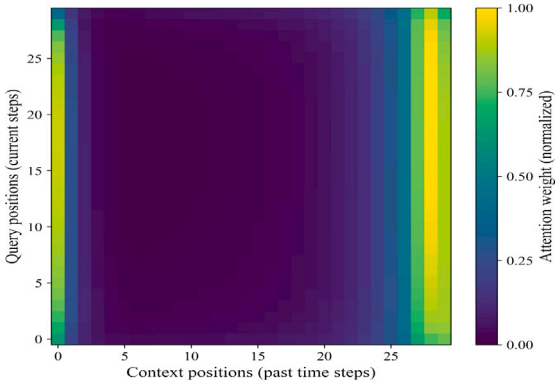


Fig. 6. Attention weight heatmap showing how each query time step attends to past context positions. The model assigns higher weights to boundary positions, with relatively low attention to intermediate time steps.

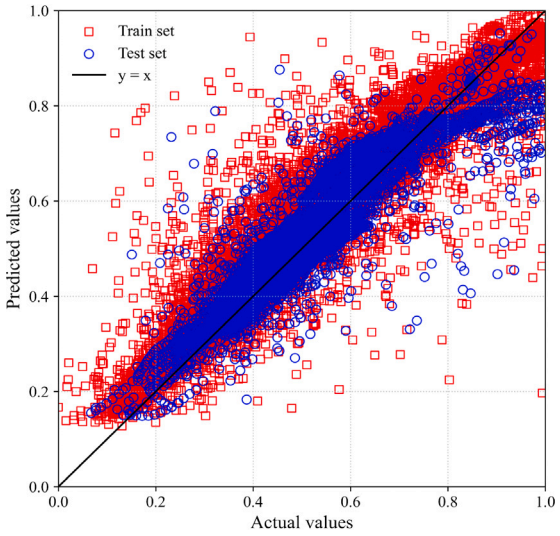


Fig. 7. Scatter representation of predicted and historical fouling values (normalized) for training and test data.

inherent nature of industrial data. In large-scale EVA polymerization plants, process sensors such as thermocouples and pressure transmitters are subject to calibration drift, flow disturbances, and transient control actions, all of which can momentarily alter measured heat-transfer behavior. Similarly, periods of gradual stabilization or sharp spikes often coincide with operational transitions such as temporary steady-state phases or utility adjustments. These prominent fluctuations are characteristic of routine process transients rather than modeling artifacts. The fact that all recurrent models follow the trend indicates that the fouling signal is intrinsically embedded within the process variables, while the proposed architecture reduces the deviation around these disturbance regions.

Fig. 8 illustrates the variation of the predicted fouling factor across different operating ranges of selected process variables. For each variable on the x-axis, the data are partitioned into value intervals, and the curves report the mean fouling factor observed within each range, with shaded regions indicating the associated $\pm\sigma$ variability for training and test sets. These variables include both directly and indirectly related operational parameters that influence heat-transfer performance. The observed trends manifest the model capacity to reproduce realistic fouling dynamics consistent with overall process behavior. Across all cases, the close agreement between the actual and forecast mean curves, together with the narrow $\pm\sigma$ confidence bands, indicates that the model can follow both gradual fouling evolution and rapid fluctuations observed during the process.

Table 2
Performance of alternative architectures on the test set (normalized).

Model	MSE ($\times10^{-3}$)	MSE σ ($\times10^{-2}$)	RMSE ($\times10^{-2}$)	RMSE σ ($\times10^{-2}$)	R^2
CNN	8.17	1.71	9.04	5.93	0.59
TCN	5.45	1.28	7.38	5.05	0.72
Transformer	6.02	1.59	7.76	5.50	0.70
xLSTM	6.52	1.63	5.72	5.70	0.67
Proposed Model	**3.48**	**1.07**	**4.17**	**4.18**	**0.82**

Following the extended evaluation, Table 2 reports the performance of the additional deep learning architectures considered in this work. Overall, the ranking observed from these experiments positions the Temporal Convolutional Network (TCN) (Pang et al., 2025) as the strongest among these models, yet still not at the level of the proposed fusion model, which achieved the lowest error values and highest coefficient of determination on the test set. The xLSTM (Alharthi and Mahmood, 2024) model also shows moderate performance; however, its design is tailored for long-horizon sequence modeling, and the relatively short input window used in our experiments limits its ability to fully exploit its extended memory mechanisms. Unlike the Convolutional Neural Network (CNN), Transformer, and xLSTM architectures, the proposed design benefits from attention-guided integration of recurrent representations, allowing the model to retain temporal information across successive operating periods, which appears essential for modeling progressive fouling in industrial heat-exchange environments.

4.4. Ablation study of proposed model

To analyze the contribution of each architectural component, we conducted an ablation study comparing three variants: a baseline Single BiGRU, Proposed_without_MHA, and the complete Proposed model on the test set. As shown in Fig. 9, the Single BiGRU baseline exhibits a higher median error and more pronounced extreme deviations, reflecting limited capacity to capture complex dynamics. The Proposed_without_MHA variant narrows this distribution, suggesting that parallel BiGRUs enrich the learned sequence features. The full Proposed model further reduces both the median error and the presence of large deviations, attesting to the stabilizing effect of incorporating the attention mechanism.

We use Gaussian Kernel Density Estimation (KDE) plots to analyze the spread of absolute prediction errors and to evaluate how well the predicted values match the actual sample distribution in the test set. KDE is a non-parametric method for estimating the probability density function of a variable, providing a smooth approximation of the underlying distribution from the available data samples.

The KDE plots provide additional insights into the behavior of the ablation variants. Fig. 10(a) shows the distribution of absolute prediction errors for the three models. The Single BiGRU baseline produces a peak at comparatively higher error values, reflecting the presence of larger errors, whereas the Proposed_without_MHA variant shifts the density toward lower magnitudes. The complete Proposed model exhibits the highest concentration of errors near zero, accompanied by a shorter right tail, demonstrating a clear reduction in both error magnitude and variability.

A supportive perspective is given in Fig. 10(b), which compares the predicted sample density distributions with the actual fouling factor distribution. The peaks of the Proposed_without_MHA and Proposed models lie close to one another, revealing similar estimation of the most frequent fouling levels. However, differences appear in the shoulder regions: the Proposed model spreads probability more smoothly across

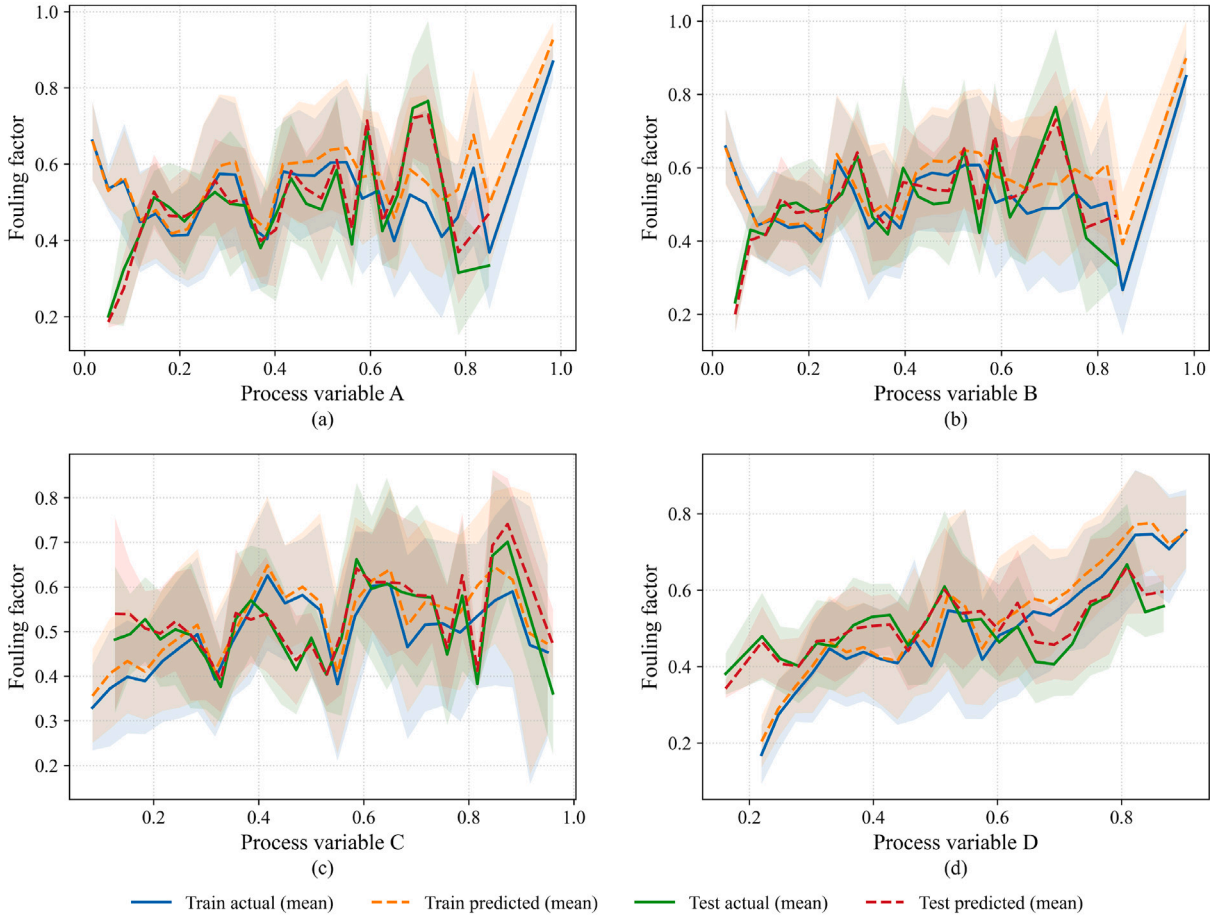


Fig. 8. Variation of the fouling factor with representative process variables for predicted and measured values.

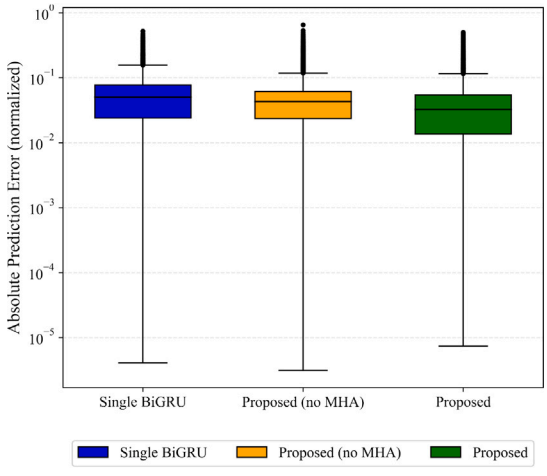


Fig. 9. The boxes represent the interquartile range (IQR), covering the middle 50% of errors, with whiskers extending to 1.5×IQR. The Proposed model achieves a lower median error compared to Single BiGRU and the no-MHA variant, indicating improved prediction accuracy, while exhibiting a comparable spread of errors.

moderate fouling ranges, whereas the Proposed_without_MHA shows sharper transitions. This suggests that the attention mechanism influences the representation of off-peak regions, rather than shifting the dominant density region.

5. Discussion

The proposed fusion-based BiGRU-BiGRU with MHA architecture reliably improves predictive accuracy and consistency to process variability in fouling factor forecasting. This improvement is evident across both training and test sets, as well as in the ablation results and error distribution analyses. When contrasted with prior research, several distinctions emerge. In Wang et al. (2023a), the proposed framework was validated primarily on experimental data from controlled test environments rather than on long-horizon industrial plant records. Consequently, its robustness under highly non-stationary operating conditions, such as those observed in EVA reactors, remains unverified and thus would need to be checked against real-world data. In Ikram et al. (2023), machine learning models for refinery shell-and-tube exchangers reported high accuracy. However, these models mainly relied on static operating variables and did not explicitly capture temporal dependencies or sequential correlations across time, limiting their applicability to predictive maintenance requiring timely forecasting. Other comparative studies, such as Yin et al. (2024), Krüger et al. (2023), achieved good short-horizon forecasts with univariate models (e.g., ARIMA, Prophet, LSTM). Furthermore, none of these works integrated attention mechanisms to explicitly pinpoint informative time steps, an aspect shown in our results to enhance stability and reduce large deviation errors. Expanding on the preceding discussion, Table 3 compares representative fouling prediction studies reported in the literature. Although many achieved high accuracy in controlled environments, they were predominantly trained on small and static datasets and therefore lack temporal scalability; hence, their reported accuracies should be interpreted as contextual rather than directly comparable benchmarks. It is also worth noting that our dataset is characterized by

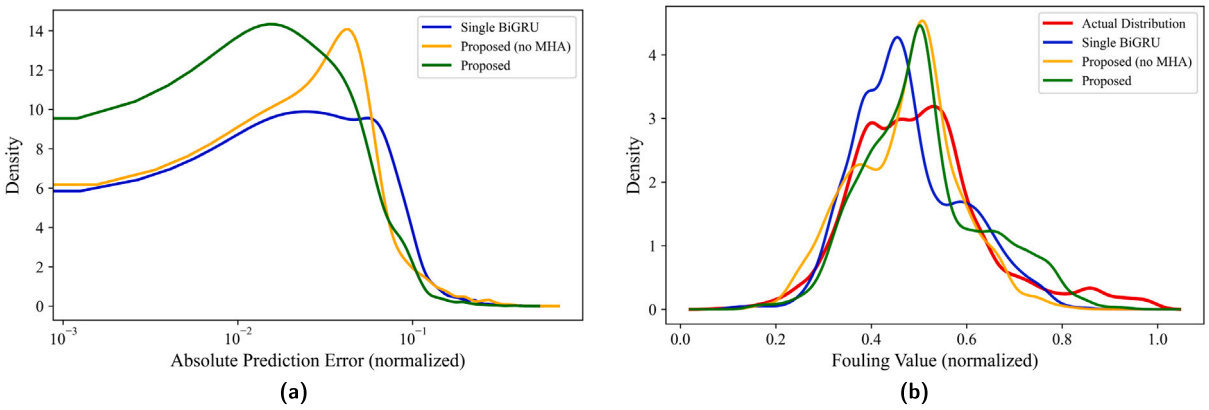


Fig. 10. Gaussian kernel density estimation plots from the ablation study: (a) absolute prediction error distributions, and (b) sample density distributions of predicted fouling factors compared with actual values in the test set.

Table 3
Comparison of representative fouling prediction studies based on dataset characteristics, model variety, accuracy, and application.

Reference	Application	Dataset	Data Type	Inputs	Models Tested	Best Accuracy
Davoudi and Vaferi (2018)	Portable Fouling Research Unit and single-tube heat exchangers	11,626	Experimental data	Fluid density, surface temperature, fluid temperature, fluid velocity, hydraulic diameter, dissolved oxygen concentration, and time	Multiple ANN configurations	MLP ANN with Bayesian regulation back-propagation: R^2 = 0.97, RMSE = 0.0355, MSE = 0.0013
Hosseini et al. (2022)	General shell and tube heat exchangers	11,626 (train: 9268, test: 2358)	Experimental data (non-temporal)	Fluid density, Surface temperature, Fluid temperature, Fluid velocity, Operation time, Equivalent diameter, Oxygen content	GPR, Decision Tree, Bagged Tree, SVR, Linear Regression	GPR: R^2 = 0.98, MAPE = 13.89%, MSE = 7.02 × 10^{-4}
Jradi et al. (2022a)	Phosphoric acid concentration plant	361 observations	Experimental plant data	Time, acid inlet temperature, acid outlet temperature, steam temperature, acid density, acid volume flow rate	RSM and multiple ANN configurations	ANN (6-6-1): R^2 = 0.99, RMSE = 4.08 × 10^{-6}, MSE = 1.67 × 10^{-11}
Ikram et al. (2023)	E101 CBA shell-and-tube heat exchanger (atmospheric-distillation unit)	290 samples (train:232, test:58)	Experimental industrial data (static measurements)	Tube inlet temperature, tube outlet temperature, tube mass flux, shell inlet temperature, shell outlet temperature, shell mass flux	FNN-MLP, NARX, and SVM-RBF	FNN-MLP: R = 0.99, nRMSE = 1.0031%, nMAE = 0.7939%, SEP = 1.0027%
Wang et al. (2023a)	Aircraft air-conditioning heat exchangers (Boeing 737-800)	25 records (laboratory setup)	Short experimental data	Ambient temperature, air conditioning system inlet pressure, primary heat exchanger outlet pressure, secondary heat exchanger outlet pressure	BP neural network, LSTM, and TA-LSTM	TA-LSTM: MAPE = 0.00231, RMSE = 0.00168, MAE = 0.00136
Sholahudin Rohman et al. (2022)	LDPE tubular polymerization reactor	8000	Simulated dynamic data (non-temporal modeling)	Heat transfer coefficient (zone 4), terminal-point temperature (zone 4), polymer exit density	FFNN with different configurations	FFNN (3-8-1): R^2 = 0.98; MSE = 1.17 × 10^{-9}
Sansana et al. (2024)	Ethylene oxide production (Dow chemical company)	–	3 years of real industrial time-series data	Reynolds number (KPI), lagged KPI_{t-1}, cycle gas flow, cumulative EO production, inlet tube temperature	Hybrid approach (Mechanistic data-driven approach using Multiple Linear Regression)	R^2 = 0.70
Ours	EVA polymerization heat exchanger (Polymer production plant)	43,768 records (train:34,990, test:8,784)	6 years of real industrial time-series data	Key process parameters including reactor temperature, pressure, and flow rates (22 selected features based on Mutual Information)	LSTM, BiLSTM, GRU, BiGRU, LSTM-GRU, BiLSTM-BiGRU, Parallel BiLSTM, Parallel BiGRU, and Proposed Parallel BiGRU with MHA	Proposed Parallel BiGRU with MHA: R^2 = 0.82, RMSE = 0.0417, MSE = 3.48×10^{-2}

relatively short time windows and a moderate number of samples. Under these conditions, transformer-based architectures are generally less effective, not only because they tend to benefit from longer sequences and larger datasets (Wen et al., 2023), but also because they lack the inductive bias of RNNs towards processing recent time dependencies. This further justifies our focus on recurrent architectures.

The present work lies in leveraging six years of real industrial monitoring data from an EVA tubular reactor, enabling evaluation under authentic process variability rather than synthetic or laboratory-generated datasets. The proposed fusion approach not only captures bidirectional sequential patterns through parallel BiGRU branches but also exploits MHA to adaptively emphasize critical temporal dependencies. This design choice was confirmed by the ablation study, which demonstrated that both fusion and attention contribute additive benefits, yielding lower error variance and improved fidelity to the actual fouling distribution (see Table 1). The KDE plot (Fig. 10), and boxplot analyses (Fig. 9) further confirm that the inclusion of attention reduces large-deviation errors and aligns the predicted distribution more closely with the true fouling behavior. In addition, MI-based feature selection was applied to identify the most influential input variables, ensuring that the model focuses on physically meaningful factors driving fouling progression while avoiding redundancy and noise. Together, these methodological advances establish a robust and practically relevant predictive framework suitable for industrial monitoring and maintenance planning. In summary, this study contributes to the existing literature by applying a fusion-based, sequence-aware model to real plant data. This advancement supports improved operational decision making for fouling control in industrial polymerization reactors.

While the results imply promising forecasting capabilities, the proposed framework is subject to certain limitations. First, the model was trained and evaluated on data from a single industrial EVA polymerization reactor, and although the dataset spans six years of operation,

its generalizability to other plants, process configurations remains to be verified. The forecasting strategy employed is single-step and assumes that historical windows provide sufficiently stable temporal representation; however, longer prediction horizons or rapidly shifting operating regimes may introduce additional uncertainties. Moreover, the approach is fully data-driven and does not incorporate physical fouling knowledge or mechanistic constraints, which may limit reliability when the process moves outside previously observed operating ranges.

6. Conclusion

In this study, we developed a fusion-oriented neural architecture combining parallel BiGRU layers with a Multi-Head Attention mechanism to predict fouling factors in EVA tubular polymerization reactors. The framework was trained and tested on six years of real-world industrial operating data from Versalis, with preprocessing and Mutual Information-based feature selection applied to ensure meaningful and robust inputs. The evaluation revealed that the proposed model provides superior forecasting capability compared with a range of recurrent architectures, achieving the lowest error values and the highest R^2 on the test dataset. The ablation analysis further confirmed that both the fusion of parallel BiGRUs and the adoption of the attention mechanism play essential roles in reducing prediction variance and enhancing stability under time-varying operating conditions. Overall, the developed model has proven to be accurate and stable for monitoring fouling growth in EVA reactors and can be considered as a reliable data-driven soft-sensor solution for industrial applications, particularly in supporting proactive maintenance and improving process efficiency. To strengthen industrial applicability, future studies will evaluate the dependability of the model against overfitting and validate its effectiveness using actual data from additional sections of the plant, where fouling forecast remains vital.

CRediT authorship contribution statement

Yellam Naidu Kottavalasa: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Andrea Battaglia:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Giovanni Bevilacqua:** Writing – review & editing, Resources, Investigation. **Gianni Marchetti:** Conceptualization, Investigation, Resources, Supervision, Writing – review & editing. **Andrea Salfinger:** Writing – review & editing, Conceptualization, Formal analysis, Investigation. **Lauro Snidaro:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was funded by Versalis S.p.A. and by the European Union - NextGenerationEU, National Recovery and Resilience Plan (NRRP) M4C2 Inv. 3.3 D.M. 117/2023. The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

References

Al Hadad, W., Schick, V., Maillet, D., 2019. Fouling detection in a shell and tube heat exchanger using variation of its thermal impulse responses: Methodological approach and numerical verification. Appl. Therm. Eng. 155, 612–619.

Alharthi, M., Mahmood, A., 2024. xLSTMTime: Long-term time series forecasting with xLSTM. AI 5, 1482–1495.

Ardsomang, T., Hines, J., Upadhyaya, B., 2013. Heat exchanger fouling and estimation of remaining useful life. PHM 2013 - Proceedings of the Annual Conference of the Prognostics and Health Management Society 2013, PHM 2013 - Proceedings of the Annual Conference of the Prognostics and Health Management Society 2013, vol. 5.150–158,

Bahdanau, D., Bengio, Y., Cho, K., Rachmad, Y., Merrienboer, B., 2014. On the properties of neural machine translation: Encoder-decoder approaches.

Bi, Y., Xia, G., Wang, C., Peng, M., Xu, Y., Wu, J., 2025. Design and transient analysis of a novel type passive residual heat removal system. Nucl. Eng. Des. 445, 114446.

Davoudi, E., Moghadas, B.K., 2022. Modeling and estimation of fouling factor on the hot wire probe by smart paradigms. Chem. Eng. Res. Des. 188, 81–95.

Davoudi, E., Vaferi, B., 2018. Applying artificial neural networks for systematic estimation of degree of fouling in heat exchangers. Chem. Eng. Res. Des. 130, 138–153.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780.

Hosseini, S., Khandakar, A., Chowdhury, M.E., Ayari, M.A., Rahman, T., Chowdhury, M.H., Vaferi, B., 2022. Novel and robust machine learning approach for estimating the fouling factor in heat exchangers. Energy Rep. 8, 8767–8776.

Ikram, K., Djilali, K., Abdennasser, D., Al-Sabur, R., Ahmed, B., Sharkawy, A.-N., 2023. Comparative analysis of fouling resistance prediction in shell and tube heat exchangers using advanced machine learning techniques. Res. Eng. Struct. Mater.

Jain, N., Murthy, C.A., 2015. A new estimate of mutual information based measure of dependence between two variables: properties and fast implementation. Int. J. Mach. Learn. Cybern. 7, 857–875.

Jia, Y., Liu, Y., He, X., Meng, Z., Zhao, S., 2025. Arrangement guideline of film holes along conjugate temperature difference in turbine guide vanes. Chin. J. Aeronaut. 38, 103400.

Jradi, R., Fguiri, A., Marvillet, C., Jeday, M.R., 2019. Tubular heat exchanger fouling in phosphoric acid concentration process. In: Bhattacharya, S., Ardekani, M.M., Biswas, R., Mehta, R.C. (Eds.), Inverse Heat Conduction and Heat Exchangers. IntechOpen, Rijeka, http://dx.doi.org/10.5772/intechopen.88936.

Jradi, R., Marvillet, C., Jeday, M.-R., 2022a. Analysis and estimation of cross-flow heat exchanger fouling in phosphoric acid concentration plant using response surface methodology (RSM) and artificial neural network (ANN). Sci. Rep. 12.

Jradi, R., Marvillet, C., Jeday, M.-R., 2022b. Application of an artificial neural network method for the prediction of the tube-side fouling resistance in a shell-and-tube heat exchanger. Fluid Dyn. Mater. Process. 18, 1511–1519.

Jradi, R., Marvillet, C., Jeday, M., 2023. Estimation and sensitivity analysis of fouling resistance in phosphoric acid/steam heat exchanger using artificial neural networks and regression methods. Sci. Rep. 13.

Jradi, R., Marvillet, C., Jeday, M.-R., 2024. Multi-objective optimization and performance assessment of response surface methodology (RSM), artificial neural network (ANN) and adaptive neuro-fuzzy interfence system (ANFIS) for estimation of fouling in phosphoric acid/steam heat exchanger. Appl. Therm. Eng. 248, 123255.

Kapustenko, P., Klemeš, J.J., Arsenyeva, O., 2023. Plate heat exchangers fouling mitigation effects in heating of water solutions: A review. Renew. Sustain. Energy Rev. 179, 113283.

Kottavalasa, Y.N., Battaglia, A., Bevilacqua, G., Marchetti, G., Snidaro, L., 2025. Fusion-based LSTM-GRU-attention model for time series forecasting of fouling factor in polymer production reactor. In: 2025 28th International Conference on Information Fusion. FUSION, pp. 1–7. http://dx.doi.org/10.23919/FUSION65864.2025.11123885.

Kottavalasa, Y.N., Snidaro, L., 2025. Advancing chemical manufacturing processes through data-driven approaches: A survey. Chemometr. Intell. Lab. Syst. 267, 105553.

Krüger, M., Vogel-Heuser, B., Land, K., Brandstetter, J., Lorenzer, J., Grim, G., Franzreb, M., Berensmeier, S., 2023. Forecasting membrane fouling in filtration processes using univariate data-driven models. In: 2023 IEEE 19th International Conference on Automation Science and Engineering. CASE, pp. 1–6. http://dx.doi.org/10.1109/CASE56687.2023.10260505.

Kumari, S., A, V., Esther, B., Alexander, D., 2023. Statistical model identification and variable selection for prediction of heat exchanger fouling. Math. Probl. Eng. 2023, 1–9.

Li, L., Hu, M., Ren, F., Xu, H., 2021. Temporal attention based tcn-bigru model for energy time series forecasting. In: 2021 IEEE International Conference on Computer Science, Artificial Intelligence and Electronic Engineering. CSAIEE, pp. 187–193. http://dx.doi.org/10.1109/CSAIEE54046.2021.9543210.

Li, X., Ma, X., Xiao, F., Wang, F., Zhang, S., 2020. Application of gated recurrent unit (GRU) neural network for smart batch production prediction. Energies 13.

Liang, Y., Zhu, L., Wang, Y., Wu, H., Zhang, J., Guan, J., Wang, J., 2024. Fouling prediction of a heat exchanger based on wavelet neural network optimized by improved particle swarm optimization algorithm. Processes 12.

Liu, W., Bai, Y., Yue, X., Wang, R., Song, Q., 2024. A wind speed forcasting model based on rime optimization based VMD and multi-headed self-attention-LSTM. Energy 294, 130726.

Pang, S., Zou, L., Zhang, L., Wang, H., Wang, Y., Liu, X., Jiang, J., 2025. A hybrid TCN-BiLSTM short-term load forecasting model for ship electric propulsion systems combined with multi-step feature processing. Ocean Eng. 316, 119808.

Ryu, U., Wang, J., Kim, T., Kwak, S., J.U., 2018. Construction of traffic state vector using mutual information for short-term traffic flow prediction. Transp. Res. Part C: Emerg. Technol. 96, 55–71.

Sansana, J., Rendall, R., Castillo, I., de Bruijne, L., Huggins, J., Phillips, A., Reis, M.S., 2024. Hybrid approach for advanced monitoring and forecasting of fouling with application to an ethylene oxide plant. Ind. Eng. Chem. Res. 63, 10666–10676.

Sholahudin Rohman, F., Muhammad, D., Sudibyo, Nazri Murat, M., Azmi, A., 2022. Application of feed forward neural network for fouling thickness estimation in low density polyethylene tubular reactor. Mater. Today: Proc. 63, S95–S100, 2nd International Conference on Chemical Engineering and Applied Sciences.

Tang, S.-Z., Li, M.-J., Wang, F.-L., He, Y.-L., Tao, W.-Q., 2020. Fouling potential prediction and multi-objective optimization of a flue gas heat exchanger using neural networks and genetic algorithms. Int. J. Heat Mass Transfer 152, 119488.

Tao, Z., Li, W., Guo, Z., Chen, Y., Song, L., Li, J., 2024. Aerothermal optimization of a turbine rotor tip configuration based on free-form deformation approach. Int. J. Heat Fluid Flow 110, 109644.

Versalis, S.p.A., 2022. EVA technology and process. https://versalis.eni.com/assets/documents/versalis/it/documentazione/licensing/2022/EVA.pdf.

Wang, J., Sun, L., Li, H., Ding, R., Chen, N., 2023a. Prediction model of fouling thickness of heat exchanger based on TA-LSTM structure. Processes 11, 2594.

Wang, Y.-y., Wang, W.-c., Chau, K., Xu, D.-M., Zang, H.-f., Liu, C.-j., Ma, Q., 2023b. A new stable and interpretable flood forecasting model combining multi-head attention mechanism and multiple linear regression. J. Hydroinformatics 25, 2561–2588.

Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., Sun, L., 2023. Transformers in time series: a survey. In: Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence. IJCAI '23, http://dx.doi.org/10.24963/ijcai.2023/759.

Yin, J., Jarmatz, N., Mauermann, M., Augustin, W., Scholl, S., 2024. Predictive fouling detection in food production using machine learning models based on real data. In: Proceedings of the 15th International Conference on Heat Exchanger Fouling and Cleaning-2024 (Peer-Reviewed) April. pp. 21–26.

Zhu, X., Zuo, Z., Wang, W., Jia, B., Liu, R., 2025. Parameter interaction analysis and comprehensive performance optimization of a thermoelectric generator system integrating a wide temperature range of thermoelectric modules. Energy Convers. Manage. 342, 120027.