International Journal of Heat and Mass Transfer 159 (2020) 120112

Contents lists available at ScienceDirect

International Journal of Heat and Mass Transfer

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

Fouling modeling and prediction approach for heat exchangers using deep learning

Sreenath Sundar[a], Manjunath C. Rajagopal[a], Hanyang Zhao[a], Gowtham Kuntumalla[a], Yuquan Meng[a], Ho Chan Chang[a], Chenhui Shao[a], Placid Ferreira[a,d], Nenad Miljkovic[a,e,f,g], Sanjiv Sinha[a,d], Srinivasa Salapaka[a,b,c,*]

[a] Department of Mechanical Science and Engineering, University of Illinois, Urbana, IL 61801, USA
[b] Department of Aerospace Engineering, University of Illinois, Urbana, IL 61801, USA
[c] Department of Industrial and Enterprise Systems Engineering, University of Illinois, Urbana, IL 61801, USA
[d] Micro and Nano Technology Laboratory, University of Illinois, Urbana, IL 61801, USA
[e] Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801, USA
[f] Materials Research Laboratory, Urbana, IL 61801, USA
[g] International Institute for Carbon Neutral Energy Research (WPI-12CNER), Kyushu Univeristy, 744 Moto-oka Nishi-ku, Fukuoka 819-0395, Japan

ARTICLE INFO

Article history:
Received 16 February 2020
Revised 4 June 2020
Accepted 21 June 2020

Keywords:
Heat exchanger fouling
Deep learning
Cross-flow heat exchanger
Waste heat recovery
Fouling modeling and prediction
Ensemble model
Neural networks
Bagging

ABSTRACT

In this article, we develop a generalized and scalable statistical model for accurate prediction of fouling resistance using commonly measured parameters of industrial heat exchangers. This prediction model is based on deep learning where a scalable algorithmic architecture learns non-linear functional relationships between a set of target and predictor variables from large number of training samples. The efficacy of this modeling approach is demonstrated for predicting fouling in an analytically modeled cross-flow heat exchanger, designed for waste heat recovery from flue-gas using room temperature water. The performance results of the trained models demonstrate that the mean absolute prediction errors are under $10^{-4} KW^{-1}$ for flue-gas side, water side and overall fouling resistances. The coefficients of determination (R^2), which characterize the goodness of fit between the predictions and observed data, are over 99%. Even under varying levels of measurement noise in the inputs, we demonstrate that predictions over an ensemble of multiple neural networks achieves better accuracy and robustness to noise. We find that the proposed deep-learning fouling prediction framework learns to follow heat exchanger flow and heat transfer physics, which we confirm using locally interpretable model agnostic explanations around randomly selected operating points. Overall, we provide a robust algorithmic framework for fouling prediction that can be generalized and scaled to various types of industrial heat exchangers.

1. Introduction

Fouling is the process of continuous deposition and accumulation of unwanted materials and sediments such as scale, algae, suspended solid and insoluble salts on the interior or exterior surface of the heat exchanger [1]. Fouling on process equipment and heat exchanger surfaces often have a significant, detrimental impact on the working efficiency and operation of the heat exchangers. Specifically, fouling reduces heat transfer rate, impedes fluid flow, corrodes material surface and contaminates the working fluid. These effects result in financial strain on all major heat and process industries today, typically in terms of installation of additional extended surfaces, increased fuel consumption, production losses from unplanned fouling related shutdowns, and maintenance costs for removal of fouling deposits with chemicals and mechanical anti-fouling devices [2]. Various studies estimate the total fouling related costs for major industrialized nations to be over $4.4 billion with the corresponding fouling related economic losses amounting to about 0.25% to 0.30% of their overall GDP [3,4]. The fouling related cleaning costs range between $40,000-50,000 per heat exchanger per cleaning [5]. Further, fouling can reduce the overall effective wall thermal conductivity of hybrid metal-polymer pipes to below 1 $Wm^{-1}K^{-1}$, which is required for profitable waste heat recovery [6].

The existing research on fouling mitigation that spans several centuries can be primarily categorized as - (1) *fouling prevention, or mitigation techniques*: These include adding anti-foulant or hydrophobic chemicals [7,8] to the operating fluid for interrupting

* Corresponding author.
E-mail address: salapaka@illinois.edu (S. Salapaka).

https://doi.org/10.1016/j.ijheatmasstransfer.2020.120112

Nomenclature

D_o	Pipe outer diameter, m
D_i	Pipe inner diameter, m
t	Pipe wall thickness, m
L	Pipe length, m
s	Spacing between pipes(constant spacing along row and column), m
k_{tube}	Thermal conductivity of pipe material, $Wm^{-1}K^{-1}$
n_r	Number of tubes along flue gas flow direction
n_t	Number of tubes in each row
H	Height of flue gas flow duct, m
W	Width of flue gas flow duct, m
T_{fi}	Flue gas inlet temperature, $^{\circ}C$
T_{fo}	Flue gas outlet temperature, $^{\circ}C$
T_{wi}	Water inlet temperature, $^{\circ}C$
T_{wo}	Water outlet temperature, $^{\circ}C$
$\dot{M}_{flue}$	Mass flow rate of flue gas, $kg.s^{-1}$
$\dot{M}_w$	Mass flow rate of water, $kg.s^{-1}$
f_1	Flue gas side fouling factor, m^2KW^{-1}
f_2	Water side fouling factor, m^2KW^{-1}
R_1	Flue gas side fouling resistance, KW^{-1}
R_2	Water Side fouling resistance, KW^{-1}
$R_1 + R_2$	Overall fouling resistance, KW^{-1}
V	Water flow velocity in pipes, $m.s^{-1}$
$V_{\max}$	Maximum flue gas flow velocity, $m.s^{-1}$
Re_w	Reynolds number for water flow
Re_{flue}	Reynolds number for flue flow
Nu_w	Water-side Nusselt number
Nu_{flue}	Flue-side Nusselt number
h_w	Water-side heat transfer co-efficient, $Wm^{-2}K^{-1}$
h_{flue}	Flue-side heat transfer co-efficient, $Wm^{-2}K^{-1}$
U	Overall heat transfer co-efficient, $Wm^{-2}K^{-1}$
f	Fanning friction factor for pipe flow
Eu	Euler number for flue flow
$C_{\min}$	Minimum fluid heat capacity rate, WK^{-1}
$C_{\max}$	Maximum fluid heat capacity rate, WK^{-1}
C_r	Fluid heat capacity ratio
IQR	Inter-Quartile Range

Subscript

f	physical variable evaluated under fouled-circumstances

fouling-inducing mechanisms, real-time control of heat-exchanger environment and fluid composition for minimizing empirical fouling risks [9], engineering the surfaces on heat transfer equipment [10,11], and regular cleaning and optimal maintenance scheduling of heat exchangers during the operation cycle [12,13]. While these research techniques [14] can control the rate of fouling to a considerable extent, they can significantly be improved upon with better fouling modeling and prediction algorithms that lead to better planning and maintenance control strategies. Fouling prediction can potentially save billions of dollars lost annually on fouling related losses through significant reductions in heat exchanger downtime. (2) *Fouling modeling and prediction algorithms*: These algorithms have predominantly been based on empirical approaches or specific heat-exchanger dynamical models. Previous analytical models of fouling processes are based on rate equations [15] where the rate of fouling deposition growth is estimated to be a deterministic function (linear [16,17], asymptotic [18,19], non-linear [20]) or a stochastic approximation [21]. A non-linear model is presented in Ref. [20] that predicts the threshold fouling conditions as a function of Reynolds and Prandtl numbers. The functional parameters are then estimated from experimental coking data of preheat train and process heaters used in crude-distillation units. In Ref. [21], a heat-exchanger maintenance strategy is developed based on the scatter parameter of time to fouling distribution. Model based heat-exchanger fouling estimation methods have centered around the use of Kalman filters [22] and development of fuzzy polynomial observers for state-estimation [23] based on non-linear heat exchanger fouling dynamical models [24]. The limitations of empirical and model-based fouling estimation techniques have led to the development of data-driven prediction approaches in the recent literature.

The process of fouling formation and material deposition happens through crystallization of salts, particulate deposit from fluid stream, deposition and growth of biological matter, chemical reaction by-products from various reactants in the working fluids and corrosion. Each of these individual phenomena occurs at different rates (time scales) which makes the creation of an all-encompassing dynamic model for fouling monitoring and prediction infeasible. These challenges make it difficult to develop accurate models for predicting fouling from first principles. This has resulted in a surge of data-driven approaches for fouling prediction from easily measurable heat-exchanger parameters. Fouling prediction and detection algorithms based on support vector machines (SVM) [25], autoassociative kernel regression (AAKR) [26] and artificial neural networks [22,27–30] have been successfully developed and deployed in various process applications. The extensive use of artificial neural networks for fouling prediction and detection is inspired by their utility in solving inverse problems pertaining to modeling and evaluation of heat transfer coefficients for fin-tube heat exchangers [31] and fluid-particle systems [32]. These data-driven algorithms for fouling prediction, despite achieving good fouling predictive accuracies for their respective applications, face issues of scalability and generalization. These approaches that fall under the category of classical machine learning techniques are often limited by the size of data and require extensive feature reduction either by visual identification or by mathematical methods (such as principal component analysis (PCA) [28]) to reduce model complexity. These challenges often become significant impediments to their deployment in large-scale industrial applications, where copious amounts of process data is generated.

In this article, we present a generalized approach for modeling and predicting fouling in heat-exchangers using deep-learning based feed-forward ensemble neural network architectures. Deep learning, the primary driver of our model, utilizes neural network often consisting of a few thousand parameters to learn the non-linear functional relationship existing between fouling resistances (target variable) and heat-exchanger operational data (causal variables) [33]. Our framework ensures that the fouling model can be powered by large datasets with very little pre-processing [33]. We tested our fouling prediction algorithm on an analytical model of cross-flow heat exchanger designed for waste-heat recovery. The choice of application and selection of heat-exchanger type for the simulated system is motivated by United States's annual industrial energy consumption, which is approximately 32 quadrillion Btu (10^{15} Btu), of which a further 1478 trillion Btu of energy from exhaust gas thermal emissions remain unrecovered [34,35]. Acidic condensation, increased chemical activity and fouling are often cited as significant barriers affecting the economy and effectiveness of waste-heat recovery units (WHRU) which results in their long-payback periods [34–36]. Further, industrial cross-flow heat exchangers are widely used for transferring heat between two gas streams or between a liquid and gas such as in stack gas heat recovery, convective recuperators or air preheaters in WHR [37]. The high fouling rates typically experienced by these waste heat recovery units underline the need for an accurate fouling prediction algorithm, where its real-time deployment can create maximum working impact and cost reduction.

The article is organized as follows. In Section 2, we develop a robust data-generation scheme for analytically modelling the chosen heat-exchanger system under fouled circumstances by incorporating state of the art models for both hydrodynamic and thermal fouling effects. Details of the neural network architecture used in our fouling prediction algorithm are laid out in Section 3. We address historically faced training challenges in fouling prediction networks such as lower precision, slower network convergence and unstable model performance [25] through our choice of network activation functions and output data transformation techniques. Further, we demonstrate the versatility of the proposed deep-learning framework by adapting the network architecture to predict individual flue-gas side and water-side fouling resistances in addition to the overall fouling resistance, using same operating input samples. We also outline the deep neural network training and hyper-parameter tuning strategies for the optimization algorithm used for updating network weights. In Section 4, we present the training of our deep neural network using the data samples derived from the data-generation scheme described in Section 2. The accuracy of fouling predictions is quantified using three test datasets, each with its unique sample distribution. We obtained over 99% model fit accuracy (R^2) between predicted and actual fouling resistance in each of these data-sets. In Section 5, we highlight the important features of our deep-learning model. First, we quantify the predictions from our module under varying levels of practically relevant input noise. Finally, we analyze our fouling prediction module using local interpretable model-agnostic explanations (LIME) [38]. We show that the predictions agree with results from high fidelity heat exchanger flow and heat-transfer fouling physics models around randomly chosen operating points. With the rapid advances in high-speed computing, the proposed fouling module is a generalized yet scalable framework that can be adapted to any heat-exchanger environment to make fouling predictions in real-time.

2. Problem set-up

For a clear exposition, we demonstrate our fouling prediction framework on a cross-flow heat exchanger model. In this section, we discuss the analytical heat exchanger model that is used for data generation. We use a modified effectiveness-NTU ($\epsilon - NTU$) scheme that also accounts for hydrodynamical and thermal effects of fouling.

Table 1
Range of input flow-rate and temperature parameter values in simulation motivated by [34,39].

Heat Exchanger Variable	Operating Range
Unfouled flue gas flow rate ($\dot{M}_{flue}$)	$4\ kg.s^{-1} \le \dot{M}_{flue} \le 8.5\ kg.s^{-1}$
Unfouled water flow rate ($\dot{M}_w$)	$0.75\ kg.s^{-1} \le \dot{M}_w \le 10.65\ kg.s^{-1}$
Flue inlet temperature (T_{fi})	$170^\circ C \le T_{fi} \le 250^\circ C$
Water inlet temperature (T_{wi})	$20^\circ C \le T_{wi} \le 30^\circ C$
Flue-gas side fouling factor (f_1)	$0\ m^2KW^{-1} \le f_1 \le 0.00175\ m^2KW^{-1}$
Water side fouling factor (f_2)	$0\ m^2KW^{-1} \le f_2 \le 0.00053\ m^2KW^{-1}$

2.1. Heat exchanger model description

The cross-flow heat exchanger that we model consists of banks of water carrying tubes placed in cross-flow to a flue gas stream flowing in a stack. A series of water tubes, each with outer diameter D_o, thickness t, length L, wall thermal conductivity k_{tube}, is placed in an inline square arrangement with a equal spacing s, along lateral and flow directions. Corbels are attached to the stack walls to regulate the flow of flue-gas towards the tubes bank (Fig. 1). The hot flue-gas enters the duct with temperature T_{fi}, flow rate $\dot{M}_{flue}$ and reaches a peak velocity of $V_{\max}$ as it flows over the tubes bank. The hot flue-gas loses heat to cold water flowing in the tubes to ultimately exit the duct at a lower temperature T_{fo}. Heat that is lost from the flue-gas is gained by water flowing in the tubes at a flow-rate $\dot{M}_w$. This increases its temperature from T_{wi} at the pipe inlet to T_{wo} at its exit.

Range of Input Operating Conditions: The inlet temperatures and flow-rates for the heat exchanger are chosen for a typical waste heat recovery application. A range of input parameter values for the cross-flow heat exchanger is chosen for a maximal rating of about 0.5 MW (Table 1). The mass flow rates and temperatures in Table 1 correspond to clean (unfouled) conditions. Further, the maximum fouling factor values were fixed assuming the operating fluid to be coal flue gas (mixed stream) and river water (unmixed stream) [1]. The choice of parameter values based on Table 1 ensures that the training data-set for the proposed fouling model covers a wide-range of operating points and the corresponding predictions from the fouling model can be generalized across the entire operating region of the heat exchanger.

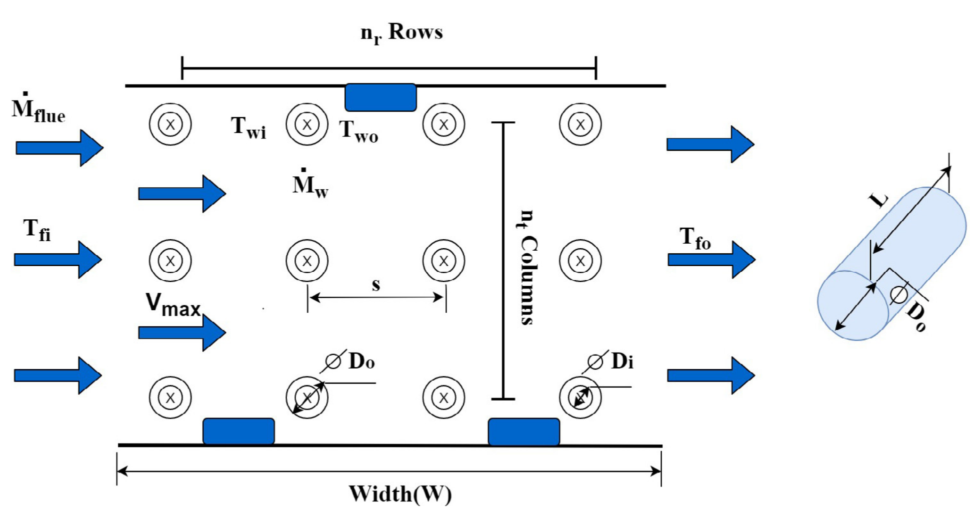


Fig. 1. Geometry of cross-flow tube arrays: The numerical values of heat-exchanger parameters used in simulations are: Number of tubes, $N = 150$, tube outer diameter $D_o = 3\,cm$, thickness $t = 5\ mm$, length $L = 1.5\ m$, wall thermal conductivity $k_{tube} = 10\ Wm^{-1}K^{-1}$, spacing $s = 4.5\ cm$, flue duct width and height $W, H = 0.5\ m$, duct length $L = 1.5\ m$, number of tubes in flow direction $n_r = 15$, number of tubes normal to flow direction $n_t = 10$.

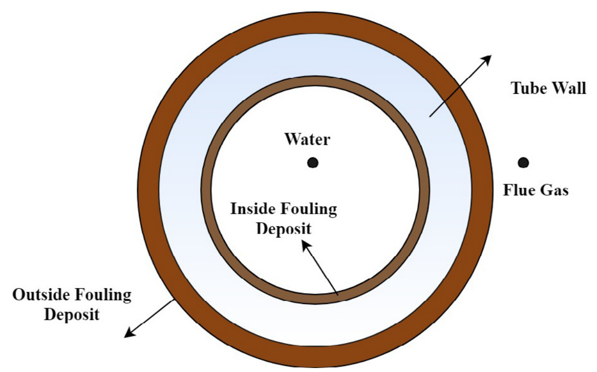


Fig. 2. Schematic of fouling layer deposits on heat exchanger pipe walls [40]. We show a cross section with fouling assumed on either sides of the pipe wall.

2.2. Fouling models for flue-gas and water side heat transfer coefficient

Water-side fouling and heat transfer model: The fouling deposit on interior of the pipes (Fig. 2) can occur due to crystallization of unwanted salts present in treated river-water that flows through these pipes. The frictional pressure drop in the tube for a single-phase flow is given by [41]:

$$\triangle p = 4f\left(\frac{L}{D_i}\right)\frac{\rho V_m^2}{2}, \tag{1}$$

where f, L, D_i, ρ and V_m in Eqn. 1 respectively represent the fanning friction factor, length, pipe inner diameter, density and mean flow velocity of water. The fouling layer causes the inner diameter of the pipe to reduce (Fig. 2) while increasing the roughness of the pipe surface. This causes an increase in the pressure drop according to Eqn. 1. The ratio of frictional pressure drops under fouled and clean pipe conditions is expressed as:

$$\frac{\triangle p_f}{\triangle p_c} = \frac{f_f}{f_c}\left(\frac{D_i}{D_{f,i}}\right)\left(\frac{V_f}{V}\right)^2, \tag{2}$$

where subscripts f and c represent properties under fouled and clean (unfouled) flow conditions. Further, it is reasonable to assume that the Fanning friction factor under fouled and clean conditions do not differ greatly [40]. If the pressure drop under fouled and clean conditions is held constant by maintaining the same pump head under both operating circumstances, the ratio of fouled inner pipe diameter ($D_{f,i}$) to the clean inner pipe diameter (D_i) is related to the ratio of the fouled water flow velocity (V_f) to the water flow velocity under clean conditions (V) as [40]:

$$\frac{D_i}{D_{f,i}} = \left(\frac{V}{V_f}\right)^2, \tag{3}$$

where D_i and $D_{f,i}$ are also related through the fouling layer thermal conductivity(k_f) and the water-side fouling factor(f_2) as:

$$ln\left(\frac{D_i}{D_{f,i}}\right) = \frac{2k_f f_2}{D_i} \qquad \Rightarrow D_{f,i} = D_i e^{-\frac{2k_f f_2}{D_i}}. \tag{4}$$

Eqn. 4 brings out how the pipe diameter decreases exponentially with an increase in fouling conductivity (k_f) and water-side fouling factor (f_2). The fouling thermal conductivity (k_f) depends on the nature of material deposit on the pipe walls. In our simulations, we chose $k_f = 2.3\,Wm^{-1}K^{-1}$, to reflect deposits due to calcium and magnesium salts that are widely present in river water [40]. For fouled conditions, the Reynolds number and mass flow rate, based on the fouled inner pipe diameter ($D_{f,i}$) and the fouled mass flow rate ($\dot{M}_{w,f}$), are given by:

$$Re_{w,f} = \frac{V_f D_{f,i}}{\nu}; \text{ and } \dot{M}_{w,f} = \rho\left(\frac{\pi D_{f,i}^2}{4}\right)V_f. \tag{5}$$

The friction factor [42,43], Nusselt number [44] and water-side heat transfer coefficient (h_w) [41,45] are estimated from the following empirical correlations:

$$f\left(Re_{w,f}\right) = \begin{cases} \left(1.58.ln\left(Re_{w,f} - 3.28\right)\right)^{-2} & \text{if } \; Re_{w,f} > 2300 \\ \dfrac{16}{Re_{w,f}} & \text{otherwise,} \end{cases} \tag{6}$$

and

$$Nu_w(D_{f,i}, L, Re_{w,f}, Pr) = \begin{cases} \left(\frac{f(Re_{w,f})}{2}\right).\dfrac{(Re_{w,f}-100).Pr}{1+12.7.\left(\frac{f(Re_{w,f})}{2}\right)^{0.5}.(Pr^{\frac{2}{3}}-1)} & \text{if } Re_{w,f} > 2300 \\ 1.86.\left(\frac{D_{f,i}.Re_{w,f}.Pr}{L}\right)^{\frac{1}{3}} & \text{otherwise.} \end{cases} \tag{7}$$

Flue gas-side fouling and heat transfer model: A fouling layer can develop on the outside of pipes due to contaminants in flue-gas exhaust streams. The rate of fouling depends on the composition and the phase of waste heat streams [35]. The extent and rate of fouling tends to be more severe on the flue-gas side as seen in Table 1. The mean pressure drop across a single row of pipes ($\triangle \overline{p}_{row}$) in a tube bank is often expressed in terms of an Euler number(Eu) [46] as:

$$\triangle\, \overline{p}_{row} = Eu.\frac{\rho V_{\max}^2}{2}; \qquad and \qquad V_{\max} = \frac{\dot{M}_{flue}}{\rho.n_t.(s - D_o)L}, \tag{8}$$

where $V_{\max}$ is the maximum flue gas flow-velocity calculated for the minimum flow area between tubes normal to the local flow direction [47]. An analytical expression for Eu is obtained empirically from approximation of several Euler number (Eu) vs Reynolds number (Re) experimental curves [48,49] for heat-exchangers with different parameters and flow configurations. This facilitates to calculate pressure drop under clean and fouled tube conditions as given by:

$$\frac{\triangle p_{flue,f}}{\triangle p_{flue,c}} = \left(\frac{Eu}{Eu_f}\right)\left(\frac{V_{max}}{V_{max,f}}\right)^2, \tag{9}$$

where $Eu = \sum_{i=0}^{i=4} \frac{c_i}{Re_{flue}^i}$ and the coefficients c_i are obtained from Ref [49]. for the chosen cross-flow configuration. Therefore Eqn. 9 is rewritten as:

$$\begin{aligned} c_0 x^4 &+ \left(\frac{c_1 \nu_{flue}}{D_{f,o}}\right)x^3 \\ &+ \left[\frac{c_2 \nu_{flue}^2}{D_{f,o}^2} - \left(c_0 + \frac{c_1}{Re_{flue}} + \frac{c_2}{Re_{flue}^2} + \frac{c_3}{Re_{flue}^3} + \frac{c_4}{Re_{flue}^4}\right)\right]x^2 \\ &+ \left(\frac{c_3 \nu_{flue}^3}{D_{f,o}^3}\right)x + \frac{c_4 \nu_{flue}^4}{D_{f,o}^4} = 0 \end{aligned} \tag{10}$$

where,

$$Re_{flue} = \frac{V_{\max} D_o}{\nu_{flue}}; \qquad \text{and} \qquad D_{f,o} = D_o e^{\frac{2k_f f_1}{D_o}}. \tag{11}$$

Further, an analytical expression (Eqn. 11) similar to Eqn. 4 relating external pipe diameter (D_o) to fouled external pipe diameter ($D_{f,o}$) in terms of fouling layer thermal conductivity (k_f), is obtained to create a bi-quadratic equation (Eq. 10). One of the positive roots of the bi-quadratic equation, whose magnitude is greater than maximum flow velocity ($V_{\max}$) is the fouled maximum flow velocity

Table 2
Algorithmic overview for data-generation scheme.

(1) Randomly select inlet flow-rates (clean conditions), inlet temperatures and fouling factors for flue-gas and water from multi-variate random distribution within working range (Table 1)
(2) Guess a feasible solution for flue-gas exit temperature ($T_{fo,guess} \leq T_{fi}$) and initialize another variable T_{fo} to be equal to T_{fi}
(3) Repeat (a)-(g) till $\left|T_{fo,guess} - T_{fo}\right| < 10^{-6}$ is satisfied
(a) Evaluate the flue gas fluid properties at the flue bulk mean-temperature
(b) Compute flue-side fouling variables $D_{f,o}, V_{max,f}, \dot{M}_{flue,f}$ and h_{flue} using equations (10), (11), (12), (13)
(c) Evaluate $T_{wo,\,guess}$ from energy balance and evaluate water properties at water bulk-mean temperature
(d) Determine water-side fouling variables $D_{f,i}, V_f, \dot{M}_{w,f}$ and h_w using equations (4), (5), (6), (7)
(e) Calculate overall heat transfer co-efficient($U = f(h_{flue}, h_w, R_1, R_2)$), dimensionless number of transfer units ($NTU = f(U, (\dot{M}C_p)_{\min})$) and heat exchanger efficiency ($\epsilon = f(U, NTU, C_r)$)
(f) Obtain T_{fo} and T_{wo} from ϵ computations in previous step (g) Update $T_{fo,guess}$ as $\frac{T_{fo,guess}+T_{fo}}{2}$ and $T_{wo,\,guess}$ as $\frac{T_{wo,guess}+T_{wo}}{2}$
(4) Record the generated heat-exchanger operating data sample along with their corresponding fouling resistances (R_1, R_2)

($V_{max,f}$). The flue gas-side Reynolds number ($Re_{flue,f}$) and mass flow rate ($\dot{M}_{flue,f}$) under fouled conditions, are evaluated based on $V_{max,f}$ as:

$$Re_{flue,f} = \frac{V_{\max,f} D_{f,o}}{\nu_{flue}}; \quad \text{and} \quad \dot{M}_{flue,f} = \rho . n_t . \left(s - D_{f,o}\right) . L . V_{\max,f}. \tag{12}$$

These expressions (Eqn. 12) are then used in suitable empirical correlations for determining the heat transfer coefficient on the flue gas side (h_{flue}) [1,50–52] given by:

$$Nu_{flue}\left(D_{f,o}, Re_{flue,f}, L, Pr_{flue}\right) = \frac{h_{flue}.D_{f,o}}{k_{flue}} = 0.211\left(Re_{flue,f}\right)^{0.651}\left(Pr_{flue}\right)^{0.34}. \tag{13}$$

2.3. Outlet temperature computation using ϵ-NTU

We use the ϵ-NTU method for computing the outlet fluid temperatures on the flue-gas and water side. The iterative scheme used for generating the heat exchanger operation data at different fouling levels is summarized in Table 2.

The analytical expression for overall heat transfer co-efficient (U) of the cross-flow heat exchanger used in our scheme, is obtained as:

$$\frac{1}{UA} = \frac{1}{h_w A_{f,i}} + \frac{\log(\frac{D_o}{D_i})}{2\pi L k_{tube}} + \frac{1}{h_{flue} A_{f,o}} + R_1 + R_2 \tag{14}$$

This expression (Eqn. 14) incorporates the thermal effects of fouling through the addition of both the thermal resistance terms (R_1, R_2) while the hydrodynamic fouling effects is accounted for in the heat transfer coefficient designs on the flue gas and water sides (h_{flue}, h_w). The overall heat transfer coefficient is then used to calculate the number of transfer units (NTU) and the effectiveness (ϵ) of the heat exchanger as:

$$NTU = \frac{UA}{C_{\min}}; \quad \text{and} \quad \epsilon = 1 - e^{-\frac{\left(1 - e^{NTU.C_r}\right)}{C_r}}, \tag{15}$$

where $C_r = \frac{C_{\min}}{C_{\max}} = \frac{min(\dot{M}_{flue,f}C_{p,flue}, \dot{M}_{w,f}C_{p,w})}{max(\dot{M}_{flue,f}C_{p,flue}, \dot{M}_{w,f}C_{p,w})}$.

2.4. Model convergence and fouling effects

The convergence of iterations in the effectiveness-NTU method was verified through 15,000 test runs with different combinations of input fluid flow rates, temperatures and fouling resistances. We ensured that the guess temperature, $T_{fo,guess}$, converged to the exit temperature, T_{fo}. The method was further verified by checking the energy balance - that is, ensuring that the energy lost by the flue gas is equal to the energy gained by the water. This 2-step approach ensured the accuracy and consistency of the fouling data generation scheme. The data-sample quality evaluation results are briefly summarized in the Table 3 for randomly selected 5 data samples.

Table 3 demonstrates that our data generation approach is accurate and robust. Therefore, our $\epsilon - NTU$ model incorporated with fouling is fairly representative of a cross-flow heat exchanger used in waste heat recovery systems. The graphs in Fig. 3 show the variation in flue gas outlet temperature and the quantity of heat exchanged for $T_{fi} = 170°C$, $\dot{M}_{flue} = 6.2\,kg.s^{-1}$, $T_{wi} = 25°C$, $\dot{M}_w = 3.54\,kg.s^{-1}$. While the exit temperature of the flue gas decreases from about 135°C to 111.5°C (Fig. 3(b)), the flue gas flow rate was greatly reduced from 6.2 $kg.s^{-1}$ to 2.5 $kg.s^{-1}$ (Fig. 3(c)) due to high fouling effects. Consequently, there is a reduction in the quantity of heat exchanged which drops by about 0.07 MW (Fig. 3(a)). These inefficiencies in heat exchanger's operations and its reduced energy performance induced by high fouling rates, present a strong case for monitoring fouling in real-time. This objective can be achieved through an algorithm that makes real-time fouling resistance predictions based on operating conditions of the heat exchanger. Such an algorithm can be used to substantially reduce operating costs through timely scheduling of heat-exchanger cleaning and maintenance operations.

Table 3
Input-output data from fouling data generation algorithm.

$\dot{M}_{flue,f}$ $kg.s^{-1}$	T_{fi} °C	$\dot{M}_{w,f}$ $kg.s^{-1}$	T_{wi} °C	$f_1 * 10^5$ m^2KW^{-1}	$f_2 * 10^5$ m^2KW^{-1}	T_{fo} °C	T_{wo} °C	$\dot{Q}$ MW	$e_T * 10^{15}$ °C	$e_{\dot{Q}} * 10^{11}$ MW
4.23	198.08	4.69	25.13	8.69	1.32	136.20	39.66	0.285	0	5.82
4.87	242.22	5.53	29.45	68.2	2.47	165.23	47.25	0.412	7.11	11.6
3.70	189.34	1.65	24.04	61.8	5.48	142.03	51.69	0.19	0	0
3.99	235.7	1.09	20.15	51.9	12.8	182.35	71.46	0.23	14.2	0
2.63	229.58	6.66	21.89	135	21.7	126.59	32.54	0.29	14.2	5.82

The error in convergence of the temperature variables ($\mathbf{e_T}$) is defined as: $e_T = \max(\left|T_{fo,guess} - T_{fo}\right|, |T_{wo,guess} - T_{wo}|)$. The variable $\mathbf{e_{\dot{Q}}}$ measures the difference in energies of the flue gas side and water side which can be defined as: $e_{\dot{Q}} = \left|[\dot{M}_{flue,f}C_{p,flue}(T_{fo} - T_{fi})] - [\dot{M}_{w,f}C_{p,w}(T_{wo} - T_{wi})]\right|$. The data-generation scheme satisfies both energy balance and temperature convergence criteria as illustrated by their negligible error magnitudes. The errors are even lower than the system's minimum floating point standard which results in their magnitudes to be zero in certain cases. Once we ascertain the temperature convergence and energy balance, the heat exchanger rating ($\dot{\mathbf{Q}}$) is computed as: $\dot{Q} = \dot{M}_{flue,f}C_{p,flue}(T_{fo} - T_{fi})$.

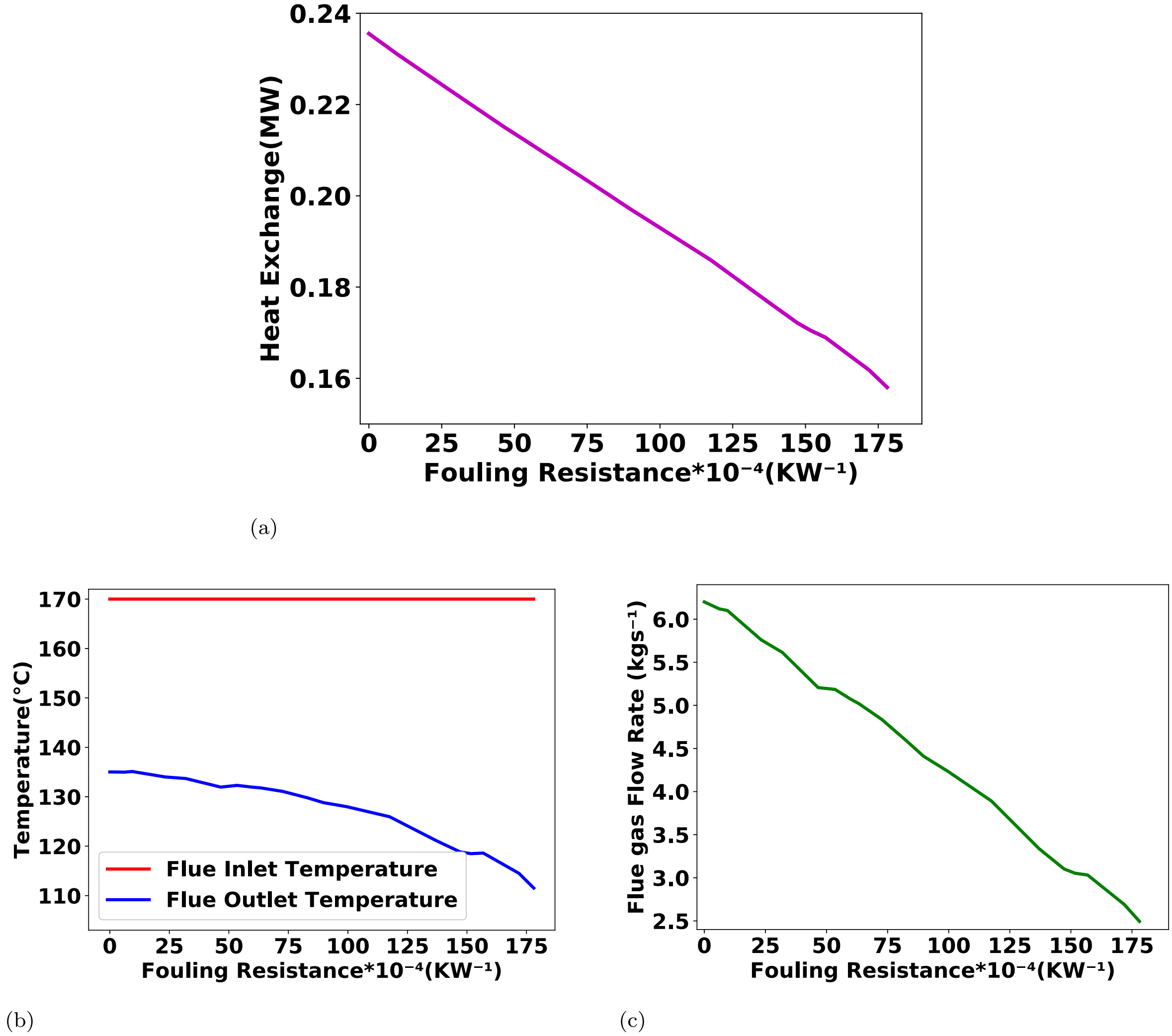


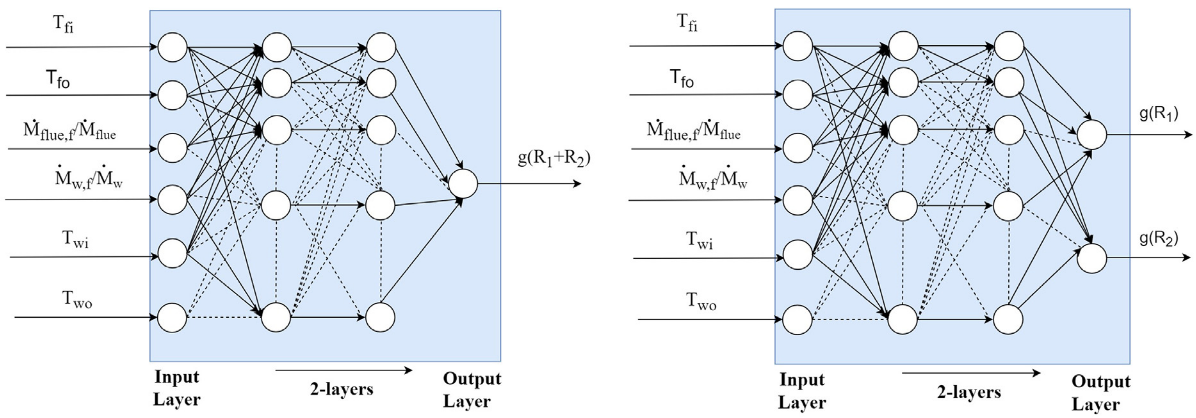
Fig. 3. Effect of increasing overall fouling resistance ($\mathbf{R_1} + \mathbf{R_2}$) on: (a) heat exchanged ($\dot{Q}$), (b) flue gas outlet temperature ($\mathbf{T_{fo}}$), and (c) flue gas flow rate ($\dot{\mathbf{M}}_{\mathbf{flue,f}}$).

3. Solution approach

Core Algorithmic Idea. The key component of our algorithm is a neural network architecture that takes in operating conditions data of cross-flow heat exchanger as its input and predicts fouling resistances. This general approach can be applied to any heat-exchanger module. The inputs to the network are inlet fluid temperatures (T_{fi}, T_{wi}), ratio of fouled fluid flow-rates to flow-rates under clean circumstances ($\frac{\dot{M}_{w,f}}{\dot{M}_w}$, $\frac{\dot{M}_{flue,f}}{\dot{M}_{flue}}$) and outlet fluid temperatures (T_{fo}, T_{wo}). All of these predictors are easy to measure in an industrial setting through fluid flow-rate and temperature sensors at the inlet and exit of the duct and pipes, respectively. Alternatively, pressure drop data can be used instead of fluid fouling flow ratio if the heat exchanger is operated under constant flow-rate conditions.

We use techniques from deep-learning [33,53] for training the neural network architecture of the fouling resistance module. In essence, a deep neural net is a specifically structured and parameterized function, often with millions of parameters, which can approximate complex nonlinear maps by determining appropriate parameters. These parameters are determined or *learned* by using *training data*; that is, the function parameters or *weights* are identified by fitting the observed values of function arguments - the *predictor variables* (e.g. operating conditions data such as measured temperatures, flow rates) and the corresponding function values - the target variables (e.g. fouling resistance). Such a fitting to learn fouling requires a large training set - a large dataset comprising the operating conditions variables with known values of fouling resistances on flue-gas and water sides. This data-set was generated using the scheme discussed in Table 2.

More specifically, a neural network is structured as a weighted directed graph where nodes are arranged in layers as shown in Fig 4. Each node, a *neuron*, is a computing unit that outputs $z = f(\sum w_i x_i + b_i)$, where $\Sigma w_i x_i$ is a weighted sum of its inputs, b_i is the bias and f is a specific *activation* function. That is, all the edges going out from a neuron have this output, each of which is an input to a neuron of the next layer. The first layer termed as in the input layer (Fig. 4) takes in the inputs to the network and the number of neurons in this layer corresponds to size of functional inputs, e.g., the number of operating conditions in our case. The main task of *learning* or *training* a neural network is to determine the *weights* w_i for each neuron in the network. During the training process, the weight parameters of the neural network (mathematical function) are initially randomly initialized. For every data sample in the dataset, the predicted value of target variable (fouling resistance) is compared with the observed value and the weights are updated to minimize the prediction error as $(y - \hat{y})^2$ where $\hat{y}$ is the network prediction and y is the observed value. This process is repeated for every data sample in the training dataset several



(a) Target: $g(R_1 + R_2)$ (b) Target: $g(R_1) + g(R_2)$

Fig. 4. Neural network architecture for predicting: (a) overall fouling resistance ($\mathbf{R_1} + \mathbf{R_2}$) and (b) flue gas-side fouling resistance ($\mathbf{R_1}$) and water side fouling resistance ($\mathbf{R_2}$).

times until the network *learns* to accurately predict the fouling resistance for all data samples in the training dataset. With the right training data-set that encompasses the entire range of operation of the heat-exchanger, the neural network can now be used to predict fouling resistance for new data samples. In our work, we segmented the dataset into training and validation datasets, where we trained the neural network on the training dataset and verified the performance of the trained network on the validation dataset.

Problem Formulation. The proposed fouling prediction module has been set-up to solve two prediction problems - the first is to predict the overall-fouling resistance ($R_1 + R_2$) as a function of heat-exchanger operating conditions (Fig. 4(a)), while the second is to similarly predict each of the individual fouling resistances (R_1 and R_2) (Fig. 4(b)). The $R_1 + R_2$ prediction quantifies the cumulative fouling effects, while predicting individually the resistances R_1 and R_2 gives a finer scale information of relative fouling levels on the inside and outside sides of the tubes. This approach can be used to get more specific (finer) information by increasing the number of neurons in the output layer. For instance, if data is available from different sections of a heat-exchanger, fouling resistances at each of those sections can be predicted. The accuracy of prediction in all these cases will depend on the data size and quality.

Fouling Prediction Network Architecture. The fouling resistance architecture used in this study (Fig. 4) is a feed-forward neural network with six neurons in the input layer (corresponding to input data-size), two hidden layers with twelve neurons and eight neurons respectively and an output layer with one neuron and two neurons pertaining to prediction of overall fouling resistance ($R_1 + R_2$) and each of the individual fouling resistance (R_1 and R_2) respectively (Fig. 4(a), 4(b)). We used an activation function, the Rectified Linear Unit (ReLU) function given by

$$\textbf{RELU:}\ f(x) = max(0, x). \tag{16}$$

RELU is easily implementable by thresholding a matrix of activations at zero which results in creation of a sparse matrix. RELU greatly accelerates the convergence of optimization algorithm used for network parameter update by a factor of six or more in certain cases, when compared to other activation functions such as sigmoid or hyperbolic tangent functions [54]. A linear activation function was used at the output layer of the network.

Data pre-processing. The inputs to the network are subjected to unity normalization (range: [0,1]) for achieving faster training rates [55]. The target variables (outputs from the network) is transformed to the logarithmic domain. The advantages of training the network to predict logarithmic fouling resistances in neural architectures shown in Fig. 4(a) and 4(b) are multifold and they follow. The kernel density estimation plots [56] (Fig. 5(a)) show that the sample probability distribution of overall fouling, flue-gas side and water-side fouling resistances ($R_1 + R_2$, R_1, R_2) in the training dataset are left-skewed with a long right tail. Further, the orders of magnitude of fouling resistances that have to be estimated are small, which slows down the neural training process consequently, and the process is besought with floating point underflow issues. We instead use logarithmic transformation of fouling resistance variables, which increases the prediction module's sensitivity, avoids floating point precision error and speeds up the training process[57,58].

Module Training and Evaluation. The objective of training the fouling prediction module is to learn the optimal network weights based on minimization of the prediction error. Accordingly, we pose and solve the following optimization problem:

$$W^* = argmin_{W \in \mathbb{W}} \sum_{i=1}^{i=n} L(y_i, \widehat{y_i});$$

$$\text{with} \qquad L(y_i, \widehat{y_i}) = (y_i - \widehat{y_i})^2, \tag{17}$$

where W represents the weights, $\mathbb{W}$ is the weight-parameter space, L represents the *loss* function between the true value y and the estimated value $\hat{y}$, and n is the number of training samples. Here the estimate $\hat{y}_i = f(x_i, W)$ as given by the neural network with input data x_i and network weights W. This minimization problem (Eqn. 17) is non-convex and difficult to solve. Deterministic approaches such as the gradient descent algorithm (refer Supplementary Information) are sensitive to initialization, computationally intensive and offer no theoretical performance guarantees for convergence to global minima [59]. We overcome this challenge by using distributed stochastic mini-batch gradient descent algorithm [60–62] with Nesterov momentum [63] (refer Supplementary Information). The mini-batch gradient is an unbiased stochastic estimator of the full gradient and is given by:

Mini-batch Stochastic Gradient:

$$\nabla_W L(W; x^{(i:i+n)}; y^{(i:i+n)}) = \frac{1}{n} \sum_{i=1}^{i=n} \nabla_W L(y_i, f(x_i; W)), \tag{18a}$$

(Full) Gradient Descent:

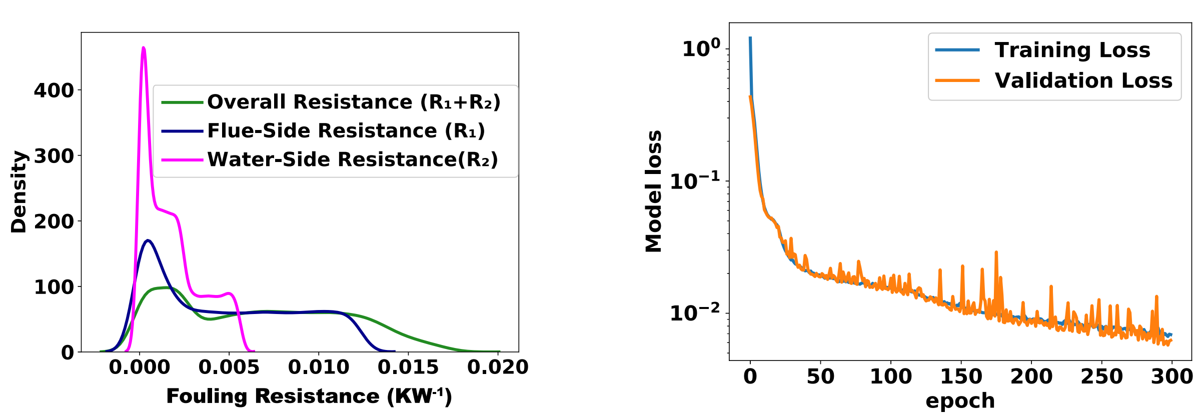
$$\nabla_W L(W; X; Y)) = \frac{1}{N} \sum_{i=1}^{i=N} \nabla_W L(y_i, f(x_i; W)), \tag{18b}$$

where n samples are chosen randomly without replacement from a dataset of size N ($n \leq N$). The gradient of the loss function in Eqn. 18a, with respect to the parameters of the prediction network is efficiently computed using back-propogation algorithm [64]. Back-propagation calculates the gradient of the network parameters using chain rule, starting from the final layer of neurons and iteratively progressing backwards to the intermediate layers of the network to avoid redundant computations (refer Supplementary Information). The expression obtained for the loss function gradient with respect to the parameters of each layer is a function of prediction network's output ($\hat{y}$) (refer Supplementary Information). This creates the need for a forward pass (computation) to determine the network output $\hat{y}$ prior to the backward computation of loss function gradient for every sample (x_i, y_i).

The iterative training procedure is summarized as follows. The fouling dataset generated through the algorithm in Table 2 is split into two parts before the commencement of training. (1) Training data-set: about 80% of the data-set is randomly selected to train the network parameters (2) validation data-set: about 20% of data is used to assess model performance, generalization apart from aiding the selection of network hyper-parameters such as optimization algorithm's step-size, batch-size and stopping criterion. The network parameters of the fouling prediction module are randomly initialized at the start of the first training epoch. An epoch is completed when the forward and backward pass computations through the prediction happen once for every sample in the training dataset. An epoch is structured into several iterations where the parameters of the network are adaptively tuned based on an optimization update rule (refer Supplementary Information). This rule involves three variables: the *magnitude* of local stochastic gradient estimated from a batch of n randomly chosen training samples using back-propagation, *size* of the update step controlled through a learning rate parameter (α) and *rate* of gradient descent regulated by momentum factor (γ). These hyper-parameters are model characteristics that cannot be estimated directly from training data-set but play a prominent role in training rate and model accuracy. With an optimal selection of these hyper-parameters and network structure, the fouling resistance prediction loss's magnitude decreases as training progresses through several epochs, as illustrated in Fig. 5(b). As the parameters of the network are updated through several iterations, the *size* of the gradient update needs to be reduced periodically to achieve convergence to the loss function's local minima. Accordingly, we incorporate a time-based step-size (α) decay scheduler: $\alpha_{t+1} = \frac{\alpha_t}{1+\beta t}$, where t is the iteration number and β is the decay rate parameter. The training process of the fouling resistance prediction network is stopped around 100 epochs when it is observed that the validation loss fails to improve on prolonged training. Training the prediction module beyond this limit leads to model over-fitting when the network learns to conform to the *noise* in the training dataset at the expense of marginally higher validation loss. An overfit network fails to accurately learn the critical functional maps between the heat-exchanger operating variables and fouling resistance. Consequently, their performance on *unseen* data samples is sub-optimal.

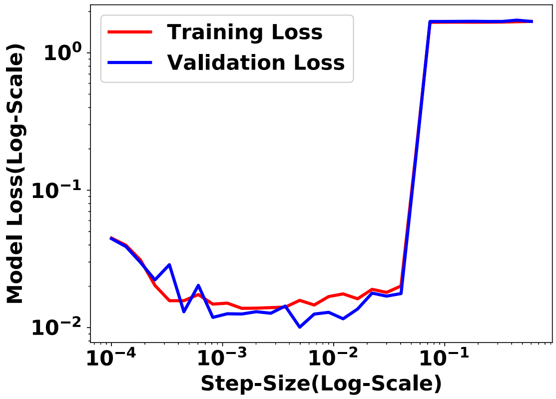
The two major hyper-parameters of the optimization update rule (refer Supplementary Information), the learning rate (α) and the batch-size (n), are tuned using grid-search algorithm [65]. The network is then trained for every tuple in this set with every training run lasting 50 epochs. Model loss is computed after each run and the tuple combination of learning rate and batch-size that resulted in low training and validation loss are taken to be the optimal hyper-parameter choice for subsequent algorithmic studies (Fig. 6(a), 6(b)). Finally, an ensemble of 5 neural networks [66], each trained on randomly selected samples amounting to about 80% of the training data (bagging process) was used for both types of prediction networks and the final predictions are computed by averaging over the predictions of each network in the ensemble. The predictive performance of the ensemble network trained through bagging process (bagged network) is compared against the performance of an individual network in the subsequent section.

The fouling prediction module for each of the two problem settings (Fig. 4(a), 4(b)) returns an output corresponding to the logarithm of fouling resistance. The performance of our trained fouling resistance prediction module is evaluated using absolute prediction errors and the co-efficient of determination(R^2) [67]. Let $y_1, y_2, y_3, \ldots y_n$ be the n values of target variable with mean $\bar{y} := \frac{1}{n}\sum_{i=1}^{i=n} y_i$ and $\hat{y}_1, \hat{y}_2, \hat{y}_3, \ldots \hat{y}_n$ be the corresponding model predictions, we define the coefficient of determination $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$, which represents the proportion of variance in the dependent variable (fouling resistance) that is explained by its predictor variables (heat-exchanger operation data) or the reduction in error achieved over a mean-only model. A higher model R^2 is indicative of better model prediction power as the model is able to explain a significant fraction of variation observed in fouling resistance values. R^2 values typically range between [0,1] though it may be negative in certain cases when the learning model performs worse than mean-only model. We quantified the performance of the prediction networks by computing the coefficient of determination (R^2) of our log-transformed model and the actual predictions, obtained by taking the exponential of the outputs returned by the log-transformed model. It is essential that the coefficient of determination (R^2) is evaluated for both these predictions as a higher coefficient of determination (R^2) of log-transformed model

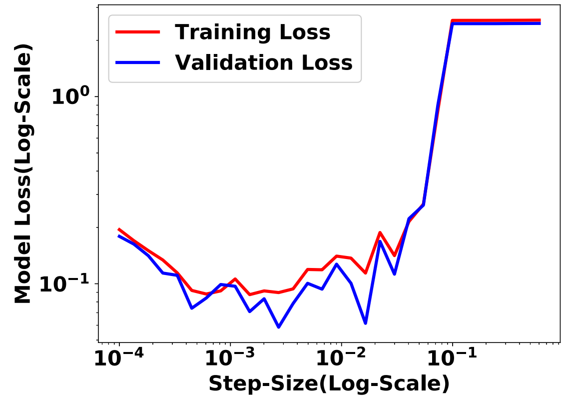


(a) (b)

Fig. 5. (a) Fouling resistance sample distribution plot: The density of fouling resistance values present in the training dataset and (b) Training-validation model loss curve: The loss curve shows the decline in fouling resistance prediction error as the weights of the network are tuned based on the training samples.



(a)



(b)

Fig. 6. Hyper-parameter tuning for: (a) overall fouling resistance ($R_1 + R_2$) prediction, (b) flue-gas and water side fouling resistance (R_1 & R_2). We used a step-size of 0.001 with a batchsize of 32 for the network in Fig. 4(a) and a step size of 0.002 with a batchsize of 32 for the network in Fig. 4(b). Further, the momentum factor (γ) and the learning rate decay parameter (β) is taken to be 0.9 and 10^{-6} respectively for both these networks..

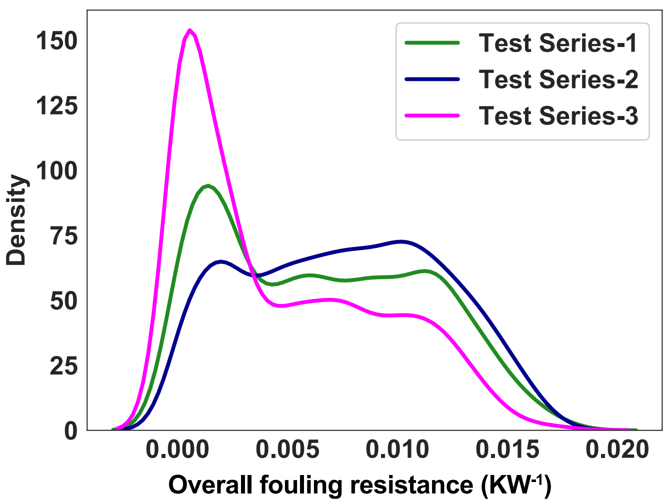


Fig. 7. Sample distribution of overall fouling resistance ($R_1 + R_2$) in three test data-sets.

does not necessarily imply a high coefficient for actual predictions [68,69]. Further, we proceed to evaluate the performance of both single and ensemble prediction networks on three test data-sets comprising of 2500 data samples each. A test data-set is one that is generated by our scheme (Table 2), on which the prediction model has not been trained on. The performance of the fouling prediction module on these 3 test data-sets, each with its signature sample distribution (Fig. 7) is indicative of prediction algorithm's generalization and utility. In the subsequent section, we report additional studies to quantify the effects of data-set size and measurement noise on fouling prediction model's performance.

4. Results

Effect of training data-set size on model performance. The neural network based model is trained on 5 different training data-sets with their sample sizes varying between 1000 and 30000. The model accuracy is then evaluated on three independent test data-sets with different sample distributions (Fig. 7). The model predictive accuracy curve for overall fouling resistance ($R_1 + R_2$) plotted against the number of training dataset samples, increases with an increase in sample size and achieves saturation after a sample-size of 10,000 (Fig. 8(a)). Similarly, the accuracy metric plots for flue-gas side and water-side fouling resistances (Fig. 8(b), 8(c)) saturate at a training data-set size of about 15,000 samples.

From Figs. 8(a), 8(b), 8(c), and Supplementary Information Table 1, we infer that the test data-set accuracies are over 97% for training data-set size of over 5000 points for both the prediction models. The relatively lower coefficient of determination observed for models trained on datasets with samples fewer than 5000, is due to lack of adequate information available to the models. Thus, they fail to learn the best generalized functional representation between the heat-exchanger variables and the fouling resistance. The higher accuracies observed on models trained on large datasets come at an increased computational cost. We chose a training dataset with 15,625 samples for both our prediction models to balance this trade-off.

It is also observed that the fouling resistance predictions on the flue-gas side is nominally better than the water-side on each of the test data-sets for the same training data-set sample size. This can perhaps be attributed to the fact that the water-side fouling resistances are smaller than flue-gas side fouling resistances which makes their accurate predictions more challenging. It should also be noted that the model predictive accuracies are significantly high partly due to the absence of measurement noise in test data-sets. Overall, the predicted fouling resistance values are in good agreement with the actual fouling resistance.

Optimizing prediction network configuration. The choice of optimal number of hidden layers and number of corresponding neurons in each hidden layer is dependent on the complexity of the non-linear functional relationship that exists between the heat exchanger's input variables and the fouling resistance. While a network with a single, huge hidden layer is a universal approximator of Borel measurable functions [70,71], single hidden layer solution is inefficient and sub-optimal compared to solutions with more hidden layers for approximating some functions [72]. There is a body of work which highlights the superior performance of deep network models over their shallow counterparts on a variety of tasks and datasets [73,74]. The higher representational power of deep networks with RELU activation units over their shallow counterparts stems from their ability to separate their input space into exponentially higher linear response units than shallow networks [75].

In Fig. 9, we perform a parametric study to capture the effect of number of hidden layers and number of neurons in each hidden layer on the performance of the fouling resistance prediction network. We compare the performance of models with different network architectures on the basis of their 5-fold cross validation loss

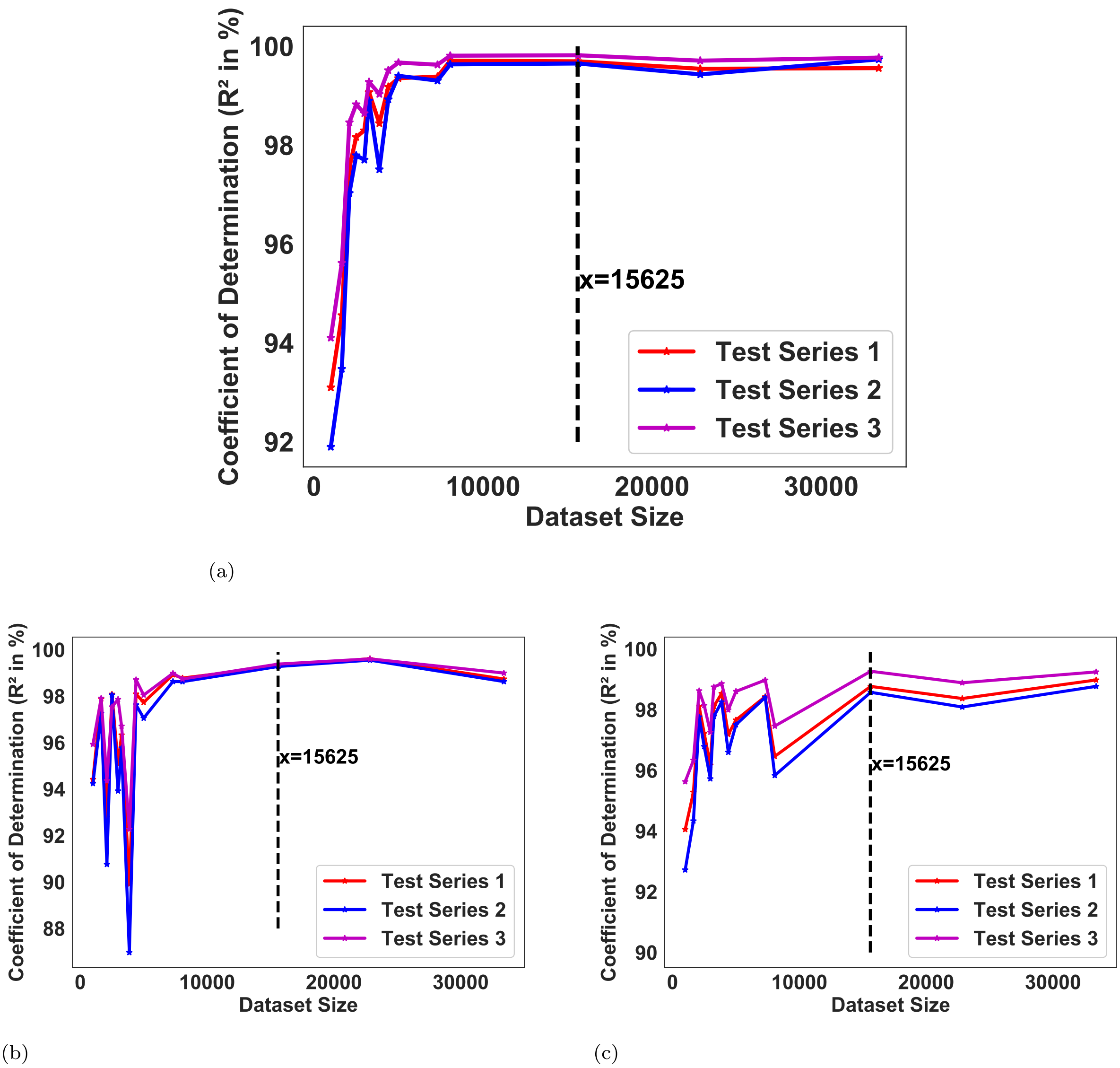


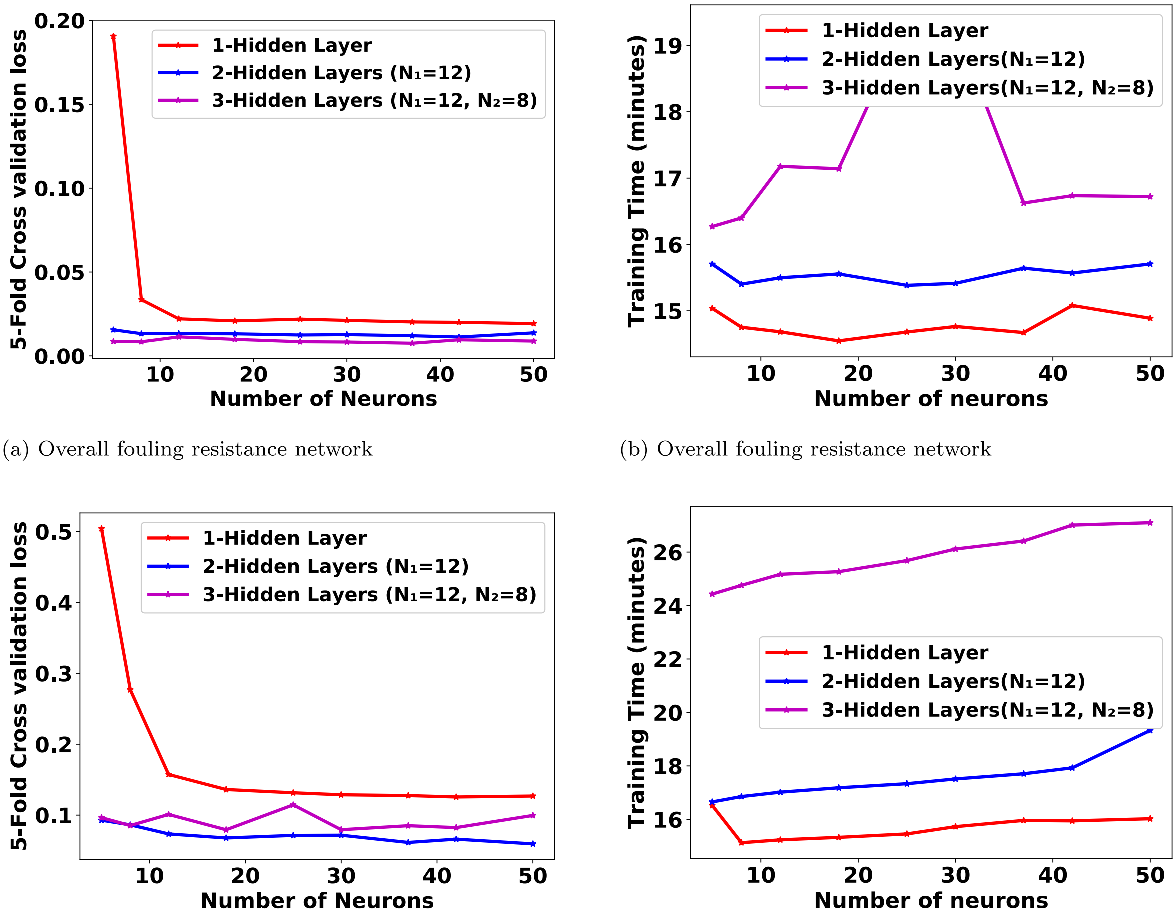
Fig. 8. Data-set size influence on R^2 for: (a) overall fouling resistance ($R_1 + R_2$) model, (b) flue-gas side fouling resistance (R_1), (c) water side fouling resistance model (R_2).

scores. For every configuration, we compute the cross-validation loss scores as follows: the training dataset of 15,600 samples is split into 5 equal parts. The network is trained for 100 epochs on 4 of these parts at a time and its predictive loss (Eqn. 17) is evaluated on the 5th partition during each run, for a total of 5 runs per configuration. The averaged root mean squared error between the network's predictions and actual values over the 5 partitions is the network configuration's 5-fold cross validation score, which is indicative of a trained model's ability to generalize to an independent dataset.

The cross-validation loss curves for both overall fouling resistance (Fig. 9(a)) and individual fouling resistances prediction networks (Fig. 9(c)) with one hidden layer has an observable elbow at around 12 neurons beyond which there isn't a signification reduction in the computed loss. Thus, subsequent network configurations with 2 and 3 hidden layers are built with 12 neurons in the first hidden layer (reduction in parametric search space). The prediction loss curve for network models with 2 hidden layers saturate at around 8 neurons. Further, the plots indicate that there is no significant improvement in the network predictions with the addition of a third hidden layer of neurons; however, we see an increase in training times (Fig. 9(b), 9(d)). Therefore, the choice of an optimal network configuration stems from a trade-off between the model's predictive power and training times. We achieve this balance with the selection of a network architecture with 2 hidden layers comprising of 12 neurons in the first hidden layer and 8 neurons in the second layer for our configuration shown in Fig 4(a),4(b).

Single and Ensemble Model Prediction Comparison. We compare the model prediction accuracy metrics for a trained single network and an ensemble (bagged model) of 5 neural networks on each of the three test data-sets. These studies were performed for both the prediction problems- overall fouling resistance ($R_1 + R_2$), flue gas and water-side fouling resistances (R_1, R_2) (network architecture in Fig. 4(a), 4(b)). We use Gaussian kernel density estimation plots for obtaining the spread of absolute prediction errors and assessing the fit between the predicted and actual sample distributions on each of the three test datasets. Gaussian kernel density is a non-parametric technique to estimate the probability density distribution of the chosen parameter based on the observed data samples.

The kernel-density estimation plot for absolute prediction errors of overall fouling resistance ($R_1 + R_2$) (Fig. 10(a)) shows that a significant proportion of errors is less than $10^{-4}KW^{-1}$ for both



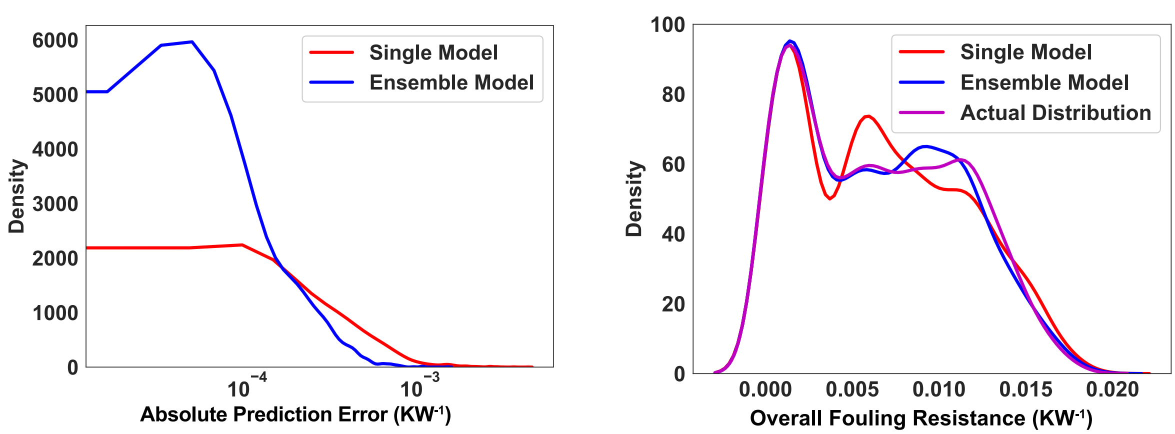
(a) Overall fouling resistance network

(b) Overall fouling resistance network

(c) Individual fouling resistances network

(d) Individual fouling resistances network

Fig. 9. (a) Cross-validation loss and (b) training time for different configurations of overall fouling resistance ($R_1 + R_2$) predictions. (c) Cross-validation loss and (d) training time for individual fouling resistance (R_1, R_2) predictions.



(a)

(b)

Fig. 10. Gaussian kernel density estimation plots of: (a) absolute model prediction errors, (b) predicted and actual sample density distributions of overall fouling resistance $(\mathbf{R_1} + \mathbf{R_2})(\mathbf{KW^{-1}})$ in test dataset-1.

single network and ensemble network based prediction models (Fig. 10(a)). However, the presence of higher sample densities at lower error magnitudes for the ensemble network curve along with the presence of a long right distribution tail for the single-network curve (Fig. 10(a)) make a solid case for the fouling resistance predictions of the ensemble model of five networks to be more robust and accurate compared to the single-network model. This finding is further reinforced through the kernel density plots shown in Fig. 10(b). The sample density distribution based on ensemble model predictions conforms to the actual sample distribution of overall fouling resistance values on test dataset-1. The predictions from a single network model, however, appears to be

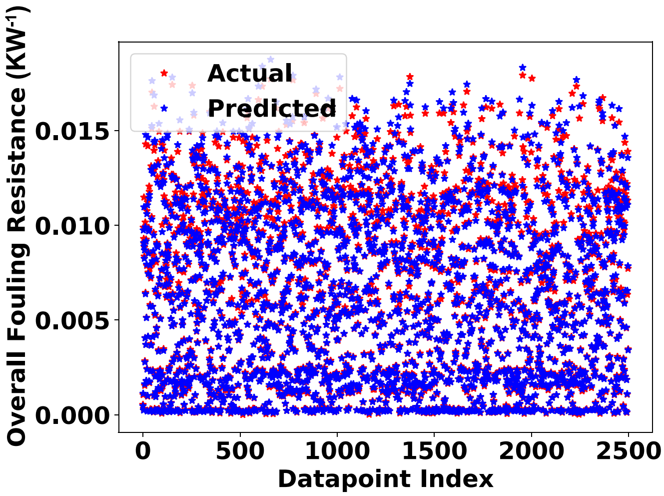


Fig. 11. Scatter plot of predicted and actual values of overall fouling resistance $(\mathbf{R_1+R_2})(\mathbf{KW^{-1}})$. This plot depicts the ensemble-network predictions and the corresponding "ground-truth" values of overall fouling resistance for each of the 2500 samples in test dataset-1.

Table 4
Ensemble network's model predictive accuracies (R^2).

Dataset	Overall Fouling (R_1+R_2)	Flue-Gas Side Fouling (R_1)	Water Side Fouling (R_2)
Test Dataset 1	99.86%	99.21%	99.20%
Test Dataset 2	99.80%	98.86%	99.05%
Test Dataset 3	99.92%	99.34%	99.47%

slightly discordant with the actual fouling resistance distribution in the range of $[0.005, 0.015]\,KW^{-1}$, accounting for its higher absolute prediction error rates. The scatter plot for ensemble model predictions with actual fouling resistance values in test dataset-1 (Fig. 11) illustrates that while both these values largely overlap, there is a slight mismatch seen in the range of $[0.01, 0.02]\,KW^{-1}$. These observations agree with the inferences from the distribution plots in Fig. 10.

The kernel density estimation plots for test data-sets-2 and 3 also point to the superior performance of an ensemble network over single network in predicting the overall fouling resistance (R_1+R_2) (Fig. 12). These observations are consistent with our earlier observations pertaining to the single and ensemble network's predictive performance on test dataset-1. The trained ensemble network model seems to slightly under-predict the overall fouling resistance in the distribution's central zone (roughly between $0.01\,KW^{-1}$ and $0.015\,KW^{-1}$) and over-predicts the fouling resistance towards the distribution's tail. The model predictive performance of the ensemble network has been tabulated in Table 4. The ensemble model's predictive performance is the highest in test data-series 3 that has a high proportion of low overall fouling resistance values and least in test data-series 2 that has a high density in the distribution's central zone. The coefficient of determination (R^2) for the trained ensemble network model is very close to 1 for both overall fouling resistance and individual flue-gas and water side predictions. These are indicative of the algorithm's ability to accurately learn the generalized functional representation between the operating conditions and fouling resistances from a training dataset.

Fig. 13 depicts the spread of absolute prediction errors for flue-side and water-side fouling resistances (R_1 and R_2) prediction model. We infer that the ensemble network model for flue-side and water side fouling resistances has lesser mean absolute prediction errors and lesser distribution outliers compared to a single network model (Fig. 13(a),13(b)). Consequently, the ensemble network model exhibits higher flue gas and water-side fouling resistance prediction accuracies (R^2) in all the test datasets compared to the single network model (Table 4). Further, the ensemble model delivers similar prediction accuracies for both flue-gas side and water side fouling resistances, unlike a single network model that has a marginally lower model fit on the water-side fouling resistance series.

The analysis of single and ensemble network model for overall and individual resistance predictions indicate that while a single network's predictions may have some disagreements in conforming to actual test-dataset's distribution in certain zones, the ensemble model comprising of 5 networks flattens out these individual mispredictions to predict a distribution that fits the actual distribution better. A combination of outputs from several networks is favorable only if they disagree on some inputs to give varied predictions and this scenario occurs in this present problem setting. It can be shown mathematically that learning of continuous valued function using ensemble network gives improved accuracy and lesser prediction inconsistencies (variance) under the discussed scenario [76] and our observations are in line with the theoretically proven results.

5. Discussion

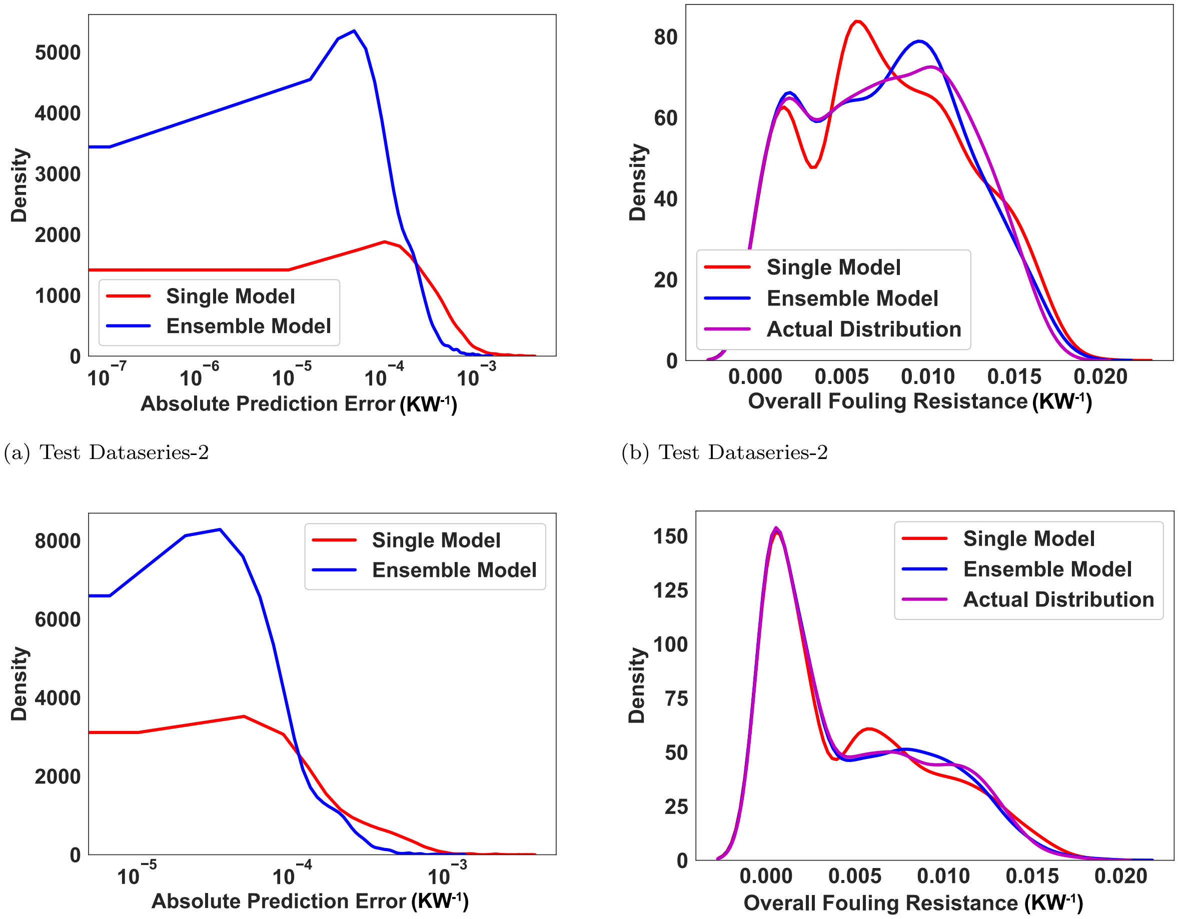
In this section, we explore the suitability of the model and assess its predictive performance on test datasets whose samples are augmented with simulated measurement noise. Further, we explain the working of the fouling resistance prediction network around select operating points through the technique of local interpretable model-agnostic explanations (LIME). This approach explains the working of the prediction network by locally approximating it with a linear model that is easy to comprehend. Consequently, this section brings out the practical applicability of our prediction model for fouling resistance in heat-exchangers.

Quantifying Prediction Model's performance under input measurement noise Sensor noise, measurement errors and other physical uncertainties introduce noise into the inputs of the fouling prediction network. These noisy real-time inputs to the prediction algorithm has a parameter distribution which deviates from the corresponding input distribution on which the algorithm is trained on. The developed algorithm is robust and practical if the loss in predictive accuracy is minimal under small levels of input noise.

We simulate the noisy datasets by adding varying degrees of simulated noise to a base test dataset, which is sampled from the same multi-variate distribution as the training dataset (Table 1). The developed module's performance is quantified on these datasets where the input temperature and flow-rate variables are augmented with a noise. This noise in measurement variables is modeled as:

$$I(t) = \bar{I}(t) + \triangle I, \tag{19}$$

where $\triangle I$ is a zero mean Gaussian noise with variance $\sigma(t)^2$. The degree of noise in a test data-set is controlled by adjusting the 3σ measurement noise levels. Noise is added to each nominal value ($I(t)$ in Eqn. 19) in these test datasets by augmenting it with a random value ($\triangle I$) which is sampled from a normal distribution with a mean 0 and $3 \times$ standard deviation (σ) level corresponding to a certain percentage of the nominal measurement. For instance, every flow-rate and temperature measurement in a test dataset with 5% 3σ noise will be augmented with a random value sampled from a normal distribution whose $3 \times$ standard-deviation (σ) corresponds to 5% of the measured value.



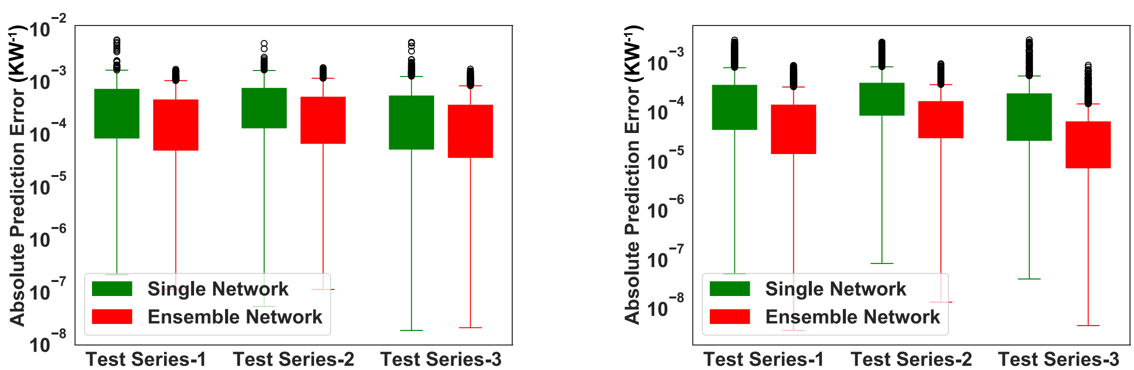
(a) Test Dataseries-2

(b) Test Dataseries-2

(c) Test Dataseries-3

(d) Test Dataseries-3

Fig. 12. Gaussian kernel density estimation plots: (a), (c) error plots comparing the sample distributions of absolute predictions errors of single and ensemble network prediction models on test datasets 2 and 3. (c), (d) The distributions of "ground truth" values and network predictions based on observed samples of test datasets 2 and 3.



(a) Flue-Side Fouling Resistance

(b) Water-Side Fouling Resistance

Fig. 13. Ensemble and single network prediction error box plots. These plots graphically represent model absolute prediction errors on each of the three datasets through their quartiles. The box in these error plots correspond to the 2nd and 3rd quartiles (middle 50% of the sample population) with vertical line extensions within 1.5*IQR on either side of the box. The black rings are the outliers that lie outside this statistical configuration.

The box plot in Fig. 14 shows the absolute prediction error spread for overall fouling resistance ($R_1 + R_2$) network under varying levels of input measurement noise.

The error bars of both single network and ensemble prediction models (Fig. 14) progressively shift upwards with an increase in 3σ measurement noise levels. They are indicative of rising mean absolute prediction errors with an increase in 3σ measurement noise levels. The mean absolute prediction errors are significantly lower than 0.01 KW^{-1} even at 15% 3σ measurement noise levels. This demonstrates the high noise resilience of both the single and ensemble network prediction modules for overall fouling resistance. The plot (Fig. 14) also confirms that the absolute prediction error bands for ensemble network is lower than the single network model with fewer distribution outliers. Thus, we infer that the

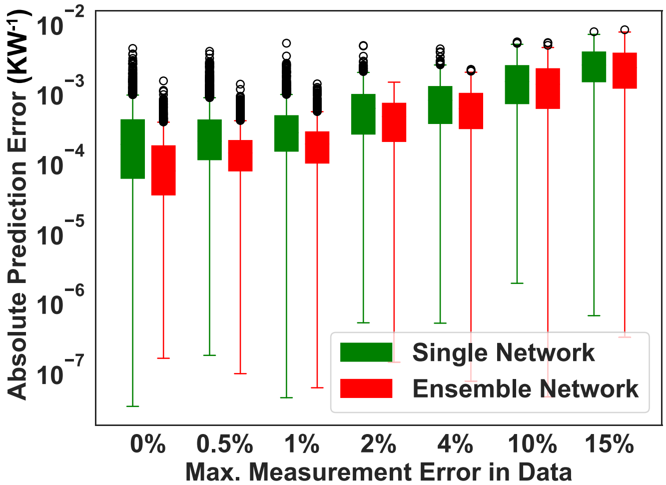


Fig. 14. Box plot of ensemble and single network model predictions of overall fouling resistance $(\mathbf{R_1} + \mathbf{R_2})(\mathbf{KW^{-1}})$ under varying levels of input measurement noise.

Table 5
Ensemble model's log-transformed and actual predictive performance in the presence of measurement noise.

Test Dataset 3σ Noise levels	Network Output ($log(R_1 + R_2)$)	Overall Fouling ($R_1 + R_2$)
0%	99.31%	99.85%
0.5%	97.63%	99.82%
1%	95%	99.73%
2%	79.2%	98.62%
4%	71.26%	97.26%
10%	43.13%	86.03%
15%	12.88%	61.47%

ensemble network predictions are more accurate than that of a single network, for overall fouling resistance.

The ensemble model's predictive performance on noisy test datasets is summarized in Table 5. The key observations from the model predictive statistics is that the coefficient of determination (R^2) of log-transformed ensemble model deteriorates rapidly with an increase in 3σ measurement noise levels. The accuracy of ensemble predictions for overall fouling resistance ($R_1 + R_2$) also decreases with an increase in measurement noise levels but the rate of decline is not as steep as that of the log-transformed model. The R^2 of ensemble model's overall fouling resistance is still as high as 86.03% at 10% measurement noise levels and 61.47% at 15% noise levels which brings out the practicality of the proposed ensemble model.

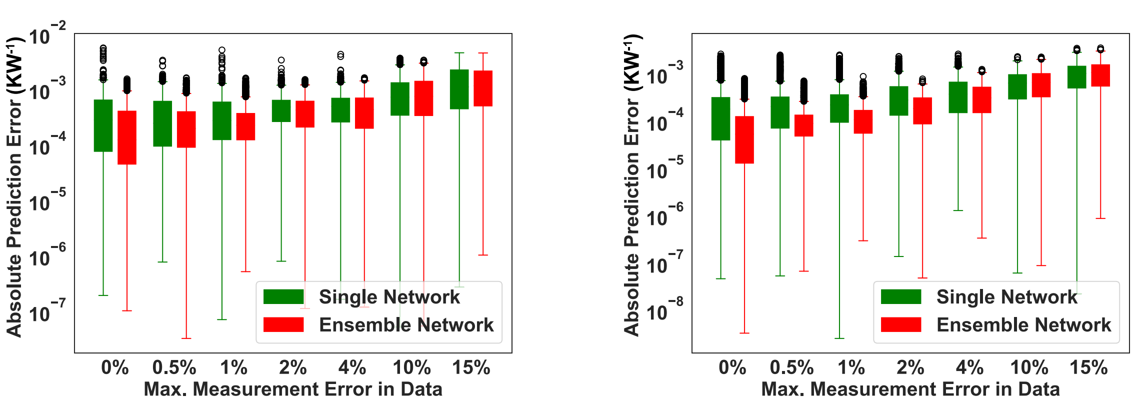
The error bars for flue-gas side (R_1) and water-side (R_2) fouling prediction model (Fig. 15) also display similar trends to that of overall fouling resistance prediction model ($R_1 + R_2$), where the mean absolute errors in the predictions of individual fouling resistances increase steadily with an increase in the test dataset's 3σ input noise levels. Table 6 quantifies the ensemble model's flue gas-side and water-side fouling resistance predictive performance on noisy input data. The rate of drop in model accuracies with an increase in input measurement noise is much faster in predictions of water-side fouling resistance compared to flue-side. Interestingly, it is seen that the co-efficient of determination (R^2) of log-transformed model becomes negative at 15% 3σ measurement noise levels. Thus, the model starts behaving worse than a mean-only prediction model at around 15% input noise levels.

Table 6
Ensemble model's log-transformed and actual predictive performance in the presence of input measurement noise.

3σ Noise levels	Flue-gas side Fouling (R_1) Log-Output ($\log(R_1)$)	Predicted (R_1)	Water Side Fouling (R_2) Log-Output ($\log(R_2)$)	Predicted (R_2)
0%	97.25%	99.20%	98.32%	99.19%
0.5%	94.1%	99.12%	91.93%	99.21%
1%	90.62%	99.15%	88.25%	99.12%
2%	70.39%	98.36%	69%	97.63%
4%	65.93%	97.86%	47.75%	93.28%
10%	39.75%	91.63%	15.08%	73.95%
15%	17.77%	80.86%	-16.93%	-

Table 7
Input variable values and neural network LIME model weights for test sample-1.

Test Sample-1 Input Variable	Variable Value	LIME Weights Single Network	Ensemble Network
Fouled Flue Flow-rate ratio ($\frac{\dot{M}_{flue,f}}{\dot{M}_{flue}}$)	0.34	0.47	1.77
Flue Inlet Temperature (T_{fi})	200.69°C	0.1	0.13
Flue Outlet Temperature (T_{fo})	129.20°C	-0.06	-0.04
Fouled Water Flow-rate ratio ($\frac{\dot{M}_{w,f}}{\dot{M}_w}$)	0.93	-0.14	-0.29
Water Inlet Temperature (T_{wi})	25.81°C	-0.07	0.26
Water Outlet Temperature (T_{wo})	35.99°C	0.09	-0.004
Coefficient of Determination (R^2)		82.87%	97.16%



(a) Flue-Side Fouling Resistance (b) Water-Side Fouling Resistance

Fig. 15. Box plot of ensemble and single network model predictions for: (a) flue-side and (b) water-side fouling resistance in the presence of varying levels of input measurement noise.

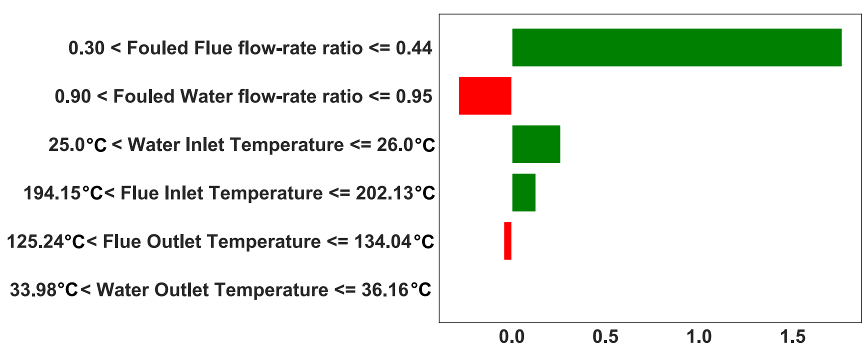
Table 8
Input variable values and neural network LIME model weights for test sample-2 .

Test Sample-2		LIME Weights	
Input Variable	Variable Value	Single Network	Ensemble Network
Fouled Flue Flow-rate ratio ($\frac{\dot{M}_{flue,f}}{\dot{M}_{flue}}$)	0.92	-1.32	-0.89
Flue Inlet Temperature (T_{fi})	216.13$^\circ C$	0.12	-0.17
Flue Outlet Temperature (T_{fo})	134.82$^\circ C$	-0.2	0.41
Fouled Water Flow-rate ratio ($\frac{\dot{M}_{w,f}}{\dot{M}_{w}}$)	0.65	0.31	0.39
Water Inlet Temperature (T_{wi})	21.01$^\circ C$	0.08	0.11
Water Outlet Temperature (T_{wo})	30.16$^\circ C$	0.34	0.05
Coefficient of Determination (R^2)		94.11%	97%

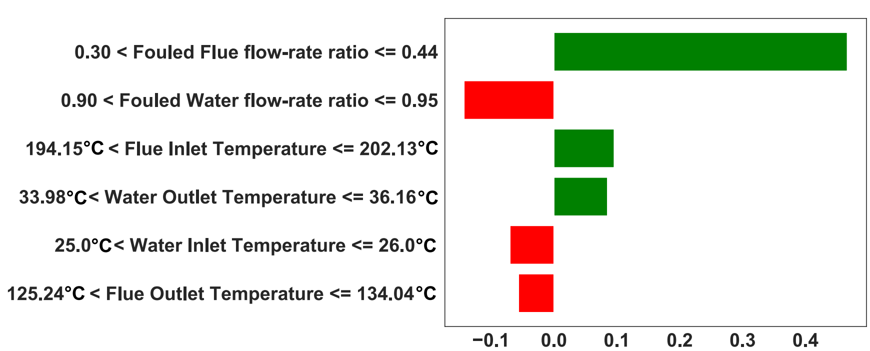
This study has helped us identify the noise threshold for predictions from the proposed fouling prediction network (Fig. 4(b)) for water-side and flue-side fouling resistances. We cap the maximum 3σ measurement noise levels for ensemble model predictors at 10% beyond which the flue-gas side and water-side fouling resistance network loses its predictive capabilities and utility. The water-side fouling resistance (R_2) samples have lower magnitudes and spread compared to their flue gas side (R_1) and overall fouling resistance ($R_1 + R_2$) counterparts, due to the smaller water-side fouling factors (f_2). This makes the water-side fouling resistance (R_2) predictions highly sensitive to their input predictors leading to their lower noise resiliency as observed (Table 6). Overall, the choice of using an ensemble (bagged) network over a single network in our prediction model greatly increases the prediction accuracy and makes it resilient to input measurement noise. The proposed fouling prediction model can operate effectively even at 10% input measurement noise levels which reinforces its suitability for practical applications.

Interpreting Fouling Prediction Model Results. Lack of mathematical models that account for generalized and holistic working of neural networks makes interpretation of the relationship between input features and output target a tedious exercise. Here, we attempt to explain the working of our fouling resistance prediction network not by model specific approaches but by using local interpretable model-agnostic explanations (LIME) [38].

LIME is based on the theory of linearization of non-linear functions. The proposed neural network module for fouling resistance prediction can be thought of a non-linear curve fitting technique based on input-output predictions. Thus, the predictions of our fouling resistance module, locally around an input sample, can be understood heuristically by fitting weighted linear regression model (refer Supplementary Information). The inputs to this weighted interpretable surrogate model for our fouling resistance network is obtained by perturbing the network around a point of interest and the corresponding outputs are the fouling resistance module predictions around the perturbed input. The weights for the generated new samples, close to the point of interest, were generated using proximity-measure (distance) evaluated between the sample and the point of interest. The weighted interpretable linear model is trained based on the generated input-output instances. This model acts as a local approximation of our network model around the point of interest. The feature coefficients of this local surrogate model is a measure of that particular input variable's importance in our fouling resistance module's local predictions.

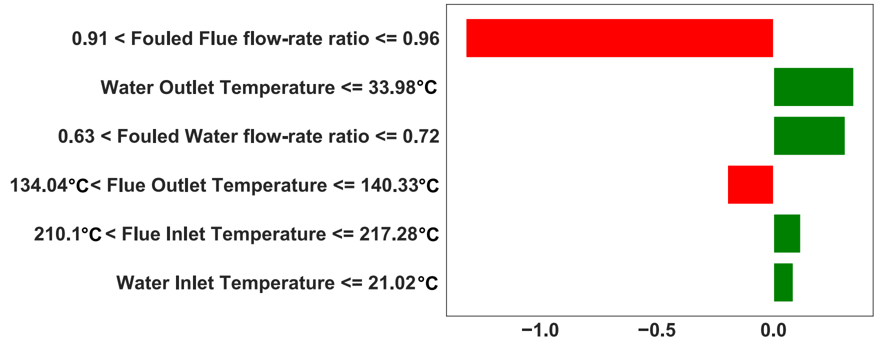


(a) Ensemble Network

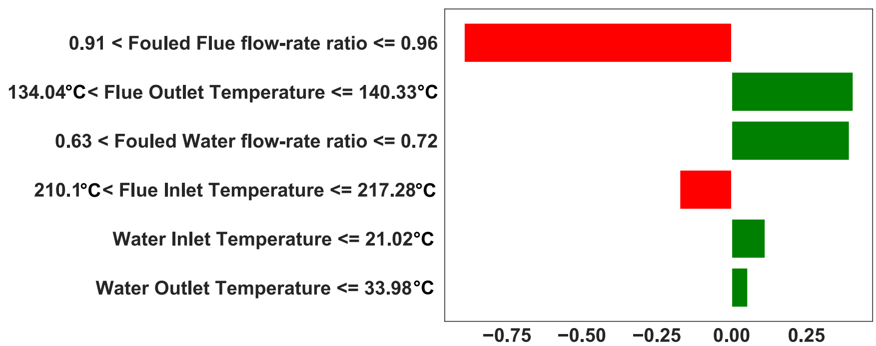


(b) Single Network

Fig. 16. LIME model weights for single and ensemble networks for test sample-1: the inputs of the training dataset are partitioned into groups of 10 (decile) based on their sample distribution. The local representative models for the networks were trained on 5000 samples in the neighborhood of the chosen operating point. The flow-rates and temperature variables were sampled from their respective deciles. The relevance of the generated sample to the chosen operating point was quantified using distance based similarity measure. This similarity measure was incorporated through the Gaussian kernel (refer Supplementary Information) with a kernel width of 0.3 in the error function of local surrogate models.



(a) Ensemble Network



(b) Single Network

Fig. 17. Model weights for LIME models for single and ensemble networks for test sample-2.

We explain the local working of the proposed overall fouling resistance (R_1+R_2) prediction module by considering two test instances (model inputs, actual and predicted model outputs in Tables 7 and 8). The order of importance of heat exchanger operating variables are similar in the developed local linear models for ensemble (Fig. 16(a)) and single (Fig. 16(b)) prediction networks. However, the two local linear models differ in how they weigh the individual predictor variables. Both the LIME models accord high positive weights to fouled flue flow ratio ($\frac{\dot{M}_{flue,f}}{\dot{M}_{flue}}$) that takes a value of 0.34 for the chosen data sample. The high fouled water flow ratio ($\frac{\dot{M}_{w,f}}{\dot{M}_{w}}$) of 0.93 has a negative effect on overall fouling resistance predictions in the local models as expected. The flue-gas and water temperature variables have lower weights than their corresponding flow rate variables in the local representative model. The contributions of the temperature variables in the local models need to be understood in the larger context of their range in training data (Table 1). The flue gas inlet temperature for the chosen sample is $200.69^{\circ}C$, which is between 30%-40% of the considered range in our data-generation scheme. This is seen as favorable conditions for fouling by the local surrogate models, however this influence is partly countered by the observed low flue outlet temperature of $129^{\circ}C$ in this sample. The water inlet and outlet temperature variable weights of local surrogate models are seen to be negligible.

The validity of linear models as an approximation to the proposed single and ensemble fouling resistance prediction networks is assessed by evaluating their coefficient of determination in the vicinity of the chosen sample. Both of these LIME models to our prediction networks display high accuracies in the considered local region (Table 7). Thus, these linear models encapsulate the local working of our fouling resistance prediction networks and simplify their analysis.

The magnitude of weights of predictors in local surrogate models for single and ensemble networks are different (Fig. 17). The local models for both these networks accord large negative weights to fouled flue flow-rate ratio ($\frac{\dot{\mathbf{M}}_{\mathbf{flue,f}}}{\dot{\mathbf{M}}_{\mathbf{flue}}}$) which lies between (0.91,0.96] for the chosen sample. The temperature variables, water outlet temperature (T_{wo}) and flue outlet temperature (T_{fo}), follow the flue flow-rate ratio variable in the precedence order with high positive weights in the linear surrogates of ensemble and single prediction networks respectively. These weight distributions for local representative models are different from that seen in the case of sample-1 (Fig. 16) where the fluid fouled flow-rate variables had a higher influence than the temperature variables. The local models around test sample-2 also differed in how they assessed the effect of flue temperature variables. The representative linear model for ensemble prediction network gave positive weights to flue-inlet temperature and negative weights to flue-outlet temperature with the opposite seen in the case of single network's local model. These local effects account for the differences in predictions of single and ensemble network models. These local-linear models have high co-efficient of determination in the vicinity of test-sample-2 (Table 8) with the prediction errors due to local non-linear effects between the input variables and overall fouling resistance.

The local exposition for proposed fouling network obtained through LIME model assigns higher weights to the hydrodynamic indicators like fouled flue-flow rate ratio ($\frac{\dot{\mathbf{M}}_{\mathbf{flue.f}}}{\dot{\mathbf{M}}_{\mathbf{flue}}}$) and fouled water flow ($\frac{\dot{\mathbf{M}}_{\mathbf{w.f}}}{\dot{\mathbf{M}}_{\mathbf{w}}}$) rate ratio variables over thermal indicators like fluid inlet and outlet temperatures. This can be explained by the fact that outlet fluid temperatures (T_{fo}, T_{wo}) are inherently dependent on fouled fluid flow-rates ($\dot{M}_{flue,f}$, $\dot{M}_{w,f}$) and fluid inlet temperatures (T_{fi}, T_{wi}). Hence, the hydrodynamic indicators like fouled flue-flow rate ratio and fouled water flow already account for some variations observed in thermal indicators like the fluid outlet temperatures and the additional information gained from the thermal variables with respect to fouling resistance predictions are only marginal.

The hydrodynamic and thermal effects dominated the heat exchanger's fouling physics, as evident from the data generation scheme used for generating the training dataset for the prediction models. Fig. 3 shows the extent of decrease in heat-exchanged from 0.24 *MW* under un-fouled heat exchanger conditions to 0.16 *MW* under maximal heat-exchanger fouling for an input variable combination of $T_{fi} = 170^{\circ}C$, $\dot{M}_{flue} = 6.2\, kg.s^{-1}$, $T_{wi} = 25^{\circ}C$, $\dot{M}_w = 3.54\, kg.s^{-1}$. We assessed the contribution of hydrodynamic effects of fouling in quantity of heat-exchanged without the thermal fouling effects (Table 2) for theoretical purposes. Interestingly, the quantity of energy exchanged by the heat-exchanger dropped from about 0.24 *MW* under un-fouled conditions to about 0.18 *MW* under fully fouled conditions which is nearly 71% of the energy drop witnessed with both hydrodynamic and thermal effects incorporated.

This presents a strong evidence that the hydrodynamic fouling effects dominate the local heat transfer and fluid flow physics of the considered cross-flow waste heat recovery heat-exchanger. The proposed fouling resistance prediction network also seems to operate in line with local heat transfer and flow physics by giving higher weights to hydrodynamic fouling variables such as fouled flue-flow rate ratio ($\frac{\dot{\mathbf{M}}_{\mathbf{flue.f}}}{\dot{\mathbf{M}}_{\mathbf{flue}}}$) and fouled water flow ($\frac{\dot{\mathbf{M}}_{\mathbf{w.f}}}{\dot{\mathbf{M}}_{\mathbf{w}}}$) over thermal variables in the local explanatory models for overall fouling resistance predictions. Thus, the local-linear surrogate model approach proposed for explaining the working of fouling prediction algorithm has successfully brought out the practical utility and physics behind the operation of the model.

6. Conclusion

We developed an accurate and generalized ensemble deep neural network framework capable of predicting overall fouling resistance and individual flue-gas side and water-side fouling resistances of a cross-flow heat exchanger used in waste heat recovery. The model predictions has an average co-efficient of determination (R^2) of 99.86%, 99.14% and 99.24% for overall fouling resistance, flue-gas side fouling resistance and water-side fouling resistance respectively on the 3 randomly generated test datasets. The ensemble based fouling prediction module was robust to input measurement noise with the model predictive accuracy (R^2) of 86.03%, 91.63% and 73.95% for overall fouling resistance, flue-gas side fouling resistance and water-side fouling resistance respectively at a 10% 3σ sensor noise cap. The developed fouling resistance prediction model's generalization and ability to learn heat-exchanger's fouling flow and heat transfer physics was demonstrated through an analysis of a local surrogate model of the network on a couple of test samples. Hence, an adequately trained model of the proposed fouling resistance prediction module can be deployed to make real-time predictions in an industrial setting with the inputs to the module obtained from various flow-rate and temperature sensors installed at the fluid inlets and outlets of the heat-exchanger. The proposed ensemble network framework for fouling resistance prediction can also be appropriately adapted to any heat exchanger problem-setting by suitably modifying the number of neurons in the input, inter-mediate and output layers of the network while also adjusting the number of networks in the prediction ensemble. The modularity of the proposed solution also provides opportunities for feeding in region specific temperature and flow-rate data of any chosen heat exchanger module to the network to identify and ascertain regions of the heat-exchanger most affected by fouling at that instant of time. This information will be particularly useful for undertaking region-specific corrective actions in large-scale heat exchangers. Moreover, this ensemble network fouling resistance prediction module can also be recast into a recurrent neural network with long-short term memory (LSTM) framework for learning heat-exchanger fouling dynamics and optimizing fouling cleaning and maintenance schedules. These discussed approaches will broadly be the subject of our future work.

Declaration of Competing Interests

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.

CRediT authorship contribution statement

Sreenath Sundar: Conceptualization, Data curation, Formal analysis, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing. **Manjunath C. Rajagopal:** Data curation, Formal analysis, Validation, Visualization, Writing - original draft, Writing - review & editing. **Hanyang Zhao:** Data curation. **Yuquan Meng:** Data curation. **Ho Chan Chang:** Data curation. **Chenhui Shao:** Conceptualization, Funding acquisition. **Placid Ferreira:** Conceptualization, Funding acquisition. **Nenad Miljkovic:** Conceptualization, Funding acquisition, Resources. **Sanjiv Sinha:** Conceptualization, Funding acquisition, Project administration, Resources. **Srinivasa Salapaka:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - review & editing.

Acknowledgement

We acknowledge support from the Advanced Manufacturing office (AMO) of the office of Energy Efficiency and Renewable Energy (EERE) under the U.S. Department of Energy (DOE) through the contract DE-EE0008312.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijheatmasstransfer.2020.120112.

References

[1] G.F. Hewitt, G.L. Shires, T.R. Bott, Process heat transfer, 113, CRC press Boca Raton, 1994.
[2] A. Pritchard, The Economics of Fouling, in: Fouling Science and Technology, Springer, 1988, pp. 31–45.
[3] B.I. Master, K.S. Chunangad, V. Pushpanathan, Fouling mitigation using helixchanger heat exchangers (2003).
[4] H. Müller-Steinhagen, M. Malayeri, A. Watkinson, Fouling of heat exchangers-new approaches to solve an old problem, Heat Transfer Eng. 26 (1) (2005) 1–4.
[5] T.R. Bott, Fouling of heat exchangers, Elsevier, 1995.
[6] M.C. Rajagopal, H.C. Chang, T. Man, G. Kuntumalla, Y. Meng, S. Sundar, H. Zhao, S. Salapaka, C. Shao, P. Ferreira, et al., Materials-to-device design of hybrid metal-polymer heat exchanger tubes for low temperature waste heat recovery, Int. J. Heat Mass Transf. 143 (2019) 118497.
[7] A. Watkinson, Critical review of organic fluid fouling, Technical Report, Argonne National Lab., IL (USA); British Columbia Univ., Vancouver, BC Q, 1988.
[8] A.P. Watkinson, Chemical reaction fouling of organic fluids, Chemical Engineering & Technology: Industrial Chemistry-Plant Equipment-Process Engineering-Biotechnology 15 (2) (1992) 82–90.
[9] Z.S. Saleh, R. Sheikholeslami, A. Watkinson, Fouling characteristics of a light australian crude oil, Heat Transfer Eng. 26 (1) (2005) 15–22.
[10] D. Klaren, End bailie re, Hydrocarbon Processing 3 (1989) 48.
[11] H. Zhao, C.A. Deshpande, L. Li, X. Yan, M.J. Hoque, G. Kuntumalla, M.C. Rajagopal, H.C. Chang, Y. Meng, S. Sundar, P. Ferreira, C. Shao, S. Salapaka, S. Sinha, N. Miljkovic, Extreme antiscaling performance of slippery omniphobic

covalently attached liquids, ACS Applied Materials & Interfaces 12 (10) (2020) 12054–12067, doi:10.1021/acsami.9b22145. PMID: 32045210
[12] F. Smaïli, V. Vassiliadis, D. Wilson, Mitigation of fouling in refinery heat exchanger networks by optimal management of cleaning, Energy & fuels 15 (5) (2001) 1038–1056.
[13] M. Markowski, K. Urbaniec, Optimal cleaning schedule for heat exchangers in a heat exchanger network, Appl. Therm. Eng. 25 (7) (2005) 1019–1032.
[14] P. Heggs, Heat exchanger fouling-mitigation and cleaning technologies by hans muller-steinhager, TCE CHEMICAL ENGINEER (2001). 49–49
[15] N. Epstein, General Thermal Fouling Models, in: Fouling Science and Technology, Springer, 1988, pp. 15–30.
[16] R.B. Ritter, Crystalline fouling studies, J Heat Transfer 105 (2) (1983) 374–378.
[17] B. Reitzer, Rate of scale formation in tubular heat exchangers. mathematical analysis of factors influencing rate of decline of over-all heat transfer coefficients, Industrial & Engineering Chemistry Process Design and Development 3 (4) (1964) 345–348.
[18] A. Watkinson, O. Martinez, Scaling of heat exchanger tubes by calcium carbonate, J. Heat Transfer 97 (4) (1975) 504–508.
[19] W. Augustin, Fouling on heat transfer surfaces; Verkrustung (Fouling) von Waermeuebertragungsflaechen, In institut fur Verfahrens- und kerntechnik., Technische Universitat Braunschweig: Germany, 1992.
[20] W. Ebert, C. Panchal, Analysis of Exxon crude-oil-slip stream coking data, Technical Report, Argonne National Lab., IL (United States), 1995.
[21] S. Zubair, A. Sheikh, M. Budair, M. Badar, A maintenance strategy for heat transfer equipment subject to fouling: a probabilistic approach, J. Heat Transfer 119 (3) (1997) 575–580.
[22] S. Lalot, O. Palsson, G. Jonsson, B. Desmet, Comparison of neural networks and kalman filters performances for fouling detection in a heat exchanger, Int. J. Heat Exch. 8 (1) (2007) 151.
[23] F. Delmotte, M. Dambrine, S. Delrot, S. Lalot, Fouling detection in a heat exchanger: a polynomial fuzzy observer approach, Control Eng. Pract. 21 (10) (2013) 1386–1395.
[24] G. Jonsson, Parameter estimation in models of heat exchangers and geothermal reservoirs., Department of Mathematical Statistics, Lund Institute of Technology, Sweden., 1990 Ph.D. thesis.
[25] L. Sun, Y. Zhang, X. Zheng, S. Yang, Y. Qin, Research on the fouling prediction of heat exchanger based on support vector machine, in: 2008 International Conference on Intelligent Computation Technology and Automation (ICICTA), 1, IEEE, 2008, pp. 240–244.
[26] T. Ardsomang, J.W. Hines, B.R. Upadhyaya, Heat exchanger fouling and estimation of remaining useful life, in: Annual Conference of Prognostics and Health Management Society, Estados Unidos eKnoxville Knoxville, 2013, pp. 1–9.
[27] C. Riverol, V. Napolitano, Estimation of fouling in a plate heat exchanger through the application of neural networks, Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology 80 (5) (2005) 594–600.
[28] V. Radhakrishnan, M. Ramasamy, H. Zabiri, V. Do Thanh, N. Tahir, H. Mukhtar, M. Hamdi, N. Ramli, Heat exchanger fouling model and preventive maintenance scheduling tool, Appl. Therm. Eng. 27 (17–18) (2007) 2791–2802.
[29] J. Zhang, A. Morris, E. Martin, C. Kiparissides, Estimation of impurity and fouling in batch polymerisation reactors through the application of neural networks, Computers & chemical engineering 23 (3) (1999) 301–314.
[30] S. Lalot, S. Lecoeuche, Online fouling detection in electrical circulation heaters using neural networks, Int. J. Heat Mass Transf. 46 (13) (2003) 2445–2457.
[31] A. Pacheco-Vega, M. Sen, K. Yang, R.L. McClain, Neural network analysis of fin–tube refrigerating heat exchanger with limited experimental data, Int J Heat Mass Transf 44 (4) (2001) 763–770.
[32] S.S. Sablani, A neural network approach for non-iterative calculation of heat transfer coefficient in fluid–particle systems, Chem. Eng. Process. 40 (4) (2001) 363–369.
[33] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436.
[34] A. Thekdi, S.U. Nimbalkar, Industrial waste heat recovery-potential applications, available technologies and crosscutting R&D opportunities, Technical Report, Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States), 2015.
[35] I. Johnson, W.T. Choate, A. Davidson, Waste heat recovery. Technology and opportunities in US industry, Technical Report, BCS, Inc., Laurel, MD (United States), 2008.
[36] H. Jouhara, N. Khordehgah, S. Almahmoud, B. Delpech, A. Chauhan, S.A. Tassou, Waste heat recovery technologies and applications, Thermal Science and Engineering Progress 6 (2018) 268–289.
[37] S.M. Jeter, Effectiveness and lmtd correction factor of the cross flow exchanger: a simplified and unified treatment, in: ASEE Southeast Section Conference, Citeseer, 2006, pp. 1–10.
[38] M.T. Ribeiro, S. Singh, C. Guestrin, Model-agnostic interpretability of machine learning, arXiv:1606.05386 (2016).
[39] N. Zhou, X. Wang, Z. Chen, Z. Wang, Experimental study on organic rankine cycle for waste heat recovery from low-temperature flue gas, Energy 55 (2013) 216–225.
[40] S. Kakaç, H. Liu, A. Pramuanjaroenkij, Heat exchangers: Selection, rating, and thermal design, CRC press, 2002.
[41] F.P. Incropera, D.P. DeWitt, Fundamentals of heat and mass transfer, 4, Wiley New York, 1996.
[42] L.F. Moody, Friction factors for pipe flow, Trans. Asme 66 (1944) 671–684.
[43] G. Filonenko, Hydraulic resistance in pipes. heat exchanger design handbook. teploenergetica, 1954.
[44] E.N. Sieder, G.E. Tate, Heat transfer and pressure drop of liquids in tubes, Industrial & Engineering Chemistry 28 (12) (1936) 1429–1435.
[45] W.M. Kays, A.L. London, Compact heat exchangers (1984).
[46] C.K. Batchelor, G. Batchelor, An introduction to fluid dynamics, Cambridge university press, 2000.
[47] A. Zukauskas, R. Ulinskas, Heat transfer in tube banks in crossflow (1988).
[48] J.P. Hartnett, W.M. Rohsenow, E. Ganic, Y. Cho, Handbook of heat transfer, McGraw-Hill, 1973.
[49] E.U. Schlunder, Heat exchanger design handbook (1983).
[50] J. Biery, Prediction of heat transfer coefficients in gas flow normal to finned and smooth tube banks, J. Heat Transfer 103 (4) (1981) 705–714.
[51] A.P. Colburn, E.d.P. de, Mean temperature difference and heat transfer coefficient in liquid heat exchangers, Industrial & Engineering Chemistry 25 (8) (1933) 873–877.
[52] A. Zukauskas, Convective heat transfer in cross flow, Handbook of single-phase convective heat transfer (1987).
[53] J. Schmidhuber, Deep learning in neural networks: an overview, Neural networks 61 (2015) 85–117.
[54] D.-A. Clevert, T. Unterthiner, S. Hochreiter, Fast and accurate deep network learning by exponential linear units (elus), arXiv:1511.07289 (2015).
[55] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv:1502.03167 (2015).
[56] V.A. Epanechnikov, Non-parametric estimation of a multivariate probability density, Theory of Probability & Its Applications 14 (1) (1969) 153–158.
[57] J. Pek, O. Wong, C. Wong, Data transformations for inference with linear regression: clarifications and recommendations., Practical Assessment, Research & Evaluation 22 (2017).
[58] J.W. Osbourne, Notes on the use of data transformation., Practical Assessment, Research & Evaluation 8 (6) (2002) n6.
[59] I.J. Goodfellow, O. Vinyals, A.M. Saxe, Qualitatively characterizing neural network optimization problems, arXiv:1412.6544 (2014).
[60] L. Bottou, Large-scale Machine Learning with Stochastic Gradient Descent, in: Proceedings of COMPSTAT'2010, Springer, 2010, pp. 177–186.
[61] N. Qian, On the momentum term in gradient descent learning algorithms, Neural networks 12 (1) (1999) 145–151.
[62] F. Chollet, et al., Keras, 2015 (https://keras.io).
[63] Y. Nesterov, A method for unconstrained convex minimization problem with the rate of convergence o (1/k 2), in: Doklady AN USSR, 269, 1983, pp. 543–547.
[64] R. Hecht-Nielsen, Theory of the Backpropagation Neural Network, in: Neural Networks for Perception, Elsevier, 1992, pp. 65–93.
[65] R. Bardenet, M. Brendel, B. Kégl, M. Sebag, Collaborative hyperparameter tuning, in: International conference on machine learning, 2013, pp. 199–207.
[66] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140.
[67] N.J. Nagelkerke, et al., A note on a general definition of the coefficient of determination, Biometrika 78 (3) (1991) 691–692.
[68] T.O. Kvålseth, Cautionary note about r 2, Am. Stat. 39 (4) (1985) 279–285.
[69] A. Scott, C. Wild, Transformations and r 2, Am. Stat. 45 (2) (1991) 127–129.
[70] K. Hornik, M. Stinchcombe, H. White, et al., Multilayer feedforward networks are universal approximators., Neural networks 2 (5) (1989) 359–366.
[71] G. Cybenko, Approximation by superpositions of a sigmoidal function, Mathematics of control, signals and systems 2 (4) (1989) 303–314.
[72] R. Reed, R.J. MarksII, Neural smithing: Supervised learning in feedforward artificial neural networks, Mit Press, 1999.
[73] I.J. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, Y. Bengio, Maxout networks, arXiv:1302.4389 (2013).
[74] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, et al., Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups, IEEE Signal Process. Mag. 29 (6) (2012) 82–97.
[75] G.F. Montufar, R. Pascanu, K. Cho, Y. Bengio, On the number of linear regions of deep neural networks, in: Advances in neural information processing systems, 2014, pp. 2924–2932.
[76] A. Krogh, J. Vedelsby, Neural network ensembles, cross validation, and active learning, in: Advances in neural information processing systems, 1995, pp. 231–238.