CHEMICAL ENGINEERING TRANSACTIONS
VOL. 103, 2023
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: Petar S. Varbanov, Panos Seferlis, Yee Van Fan, Athanasios Papadopoulos

ISBN 979-12-81206-02-1; **ISSN** 2283-9216
DOI: 10.3303/CET23103117

Using a Neural Network Approach to Predict Deposits on the Surfaces of Heat Exchange Equipment

Oleg Rudenko[a], Oleksandr Bezsonov[a], Oleg Ilyunin[a,*], Olexiy Demirskiy[b], Nataliia Serdiuk[a], Olga Arsenyeva[c], Oksana Semenenko[c]

[a]Kharkiv National University of Radioelectronics, Department of Computer Intelligent Technologies and Systems, 14 Nauky ave., 61166 Kharkiv, Ukraine
[b]AO SODRUGESTVO-T LLC, 2 Krasnoznamenny Per., 61002 Kharkiv, Ukraine
[c]O.M. Beketov National University of Urban Economy in Kharkiv, Department of Automatic Control and Computer Integrated Engineering, 17 Marshal Bazhanov Street, 61002 Kharkiv, Ukraine
oleh.iliunin@nure.ua

This work proposes a neural network (NN) approach for predicting the following values: the heat transfer coefficient at the point of interest in the operational period of plate heat exchangers (PHEs), and the time-point to reach the lower allowable limit of the heat transfer coefficient. In this approach, neural network models replace complex mathematical modelling that used systems of differential equations and matrices of heuristic coefficients to calculate the flow rate of deposits on PHE plates, which required the involvement of serious computing resources. Training a feed-forward neural network (FFNN) on a small dataset simulated in the vicinity of reference points obtained by industrial measurements showed the proper coefficient of determination R^2 = 0.99 (accuracy) of the short-term prediction forecasts and for operational evaluation of the heat transfer coefficient due to the static type of NN.

1. Introduction

The significance of the complexity of problems of world energy consumption is constantly growing, while the optimization of energy consumption, energy saving, reduction of pollution emissions is inextricably linked with sustainable development. The achievement of the plan concerning the climate-neutral economy by 2050 requires reducing global CO_2 emissions to zero. This goal requires the implementation of CO_2 and energy-reducing technologies in all industries (Lameh et al., 2021). The use of efficient compact PHEs with increased heat transfer is one of the most promising directions for solving this set of problems (Klemeš et al., 2015).
Deposits formed on the plates of PHEs are inevitable in technological processes. Deposits adversely affect production efficiency, shorten the turnaround time and the life of equipment in heat recovery systems. When deposits accumulate on the surfaces of heat exchangers, they reduce the rate of heat transfer flow, resulting in increased energy consumption and reduced energy efficiency. This, in turn, increases fuel consumption, which is not only economically unprofitable, but also increases the environmental burden on the environment and the cost of cleaning. The practical solution to the problem of automating the operational forecasting of the start of routine maintenance for cleaning the surfaces of PHE plates is a serious task on the way to sustainable development and is consistent with the Industry 4.0 conception.
At present, to solve the problem of predicting the period of scheduled dismantling and cleaning of PHEs, rather cumbersome mathematical models of differential equations are used (Bansal et al., 2008) that use heuristic coefficients, which require the involvement of serious computing resources.
Efficient energy management has always attracted the attention of both industrial manufacturers and the scientific community. Recently, machine learning (ML) has been introduced as a methodology for solving identification and control problems in the energy sector (Wu and Wang, 2018). One of the examples is the use of the long short-term memory (LSTM) methodology as a predictive approach to prognose the state of charge of batteries at battery charging stations, as proposed by Malek et al. (2019).

Paper Received: 11 June 2023; Revised: 30 June 2023; Accepted: 23 July 2023
Please cite this article as: Rudenko O., Bezsonov O., Ilyunin O., Demirskiy O., Serdiuk N., Arsenyeva O., Semenenko O., 2023, Using a Neural Network Approach to Predict Deposits on the Surfaces of Heat Exchange Equipment, Chemical Engineering Transactions, 103, 697-702 DOI:10.3303/CET23103117

A generalized and scalable statistical model with flow physics (Sundar et al., 2020) is presented to accurately predict fouling resistance using commonly measured industrial heat exchanger parameters. The 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 a large number of training samples. The effectiveness of this modeling approach has been demonstrated for fouling prediction in an analytically simulated cross-flow heat exchanger designed to recover flue gas waste heat using room temperature water. The results of the trained models show that the coefficients of determination (R^2), which characterize the accuracy of the correspondence between forecasts and observed data, exceed 99%. NN model (see Kashani et al., 2012) was applied for online short-term prediction and monitoring of crude oil fouling in heat exchangers. The model is automatically updated online at any time, when a new data set can be achieved. The results revealed, that the model could well predict (multi-step-ahead) the ascending processes. The generalization of the model was proved by the proximity of errors for training and prediction subsets. The mean relative errors (MRE) of the training and prediction subsets were about 6.61 % and 8.06 %, respectively. Trzcinski and Markowski (2018) presented an NN-based identification method of the fouling influence on heat losses due to growing deposits in a heat exchanger. Jradi et al. (2020) proposed the use of the dimensionality reduction method and machine learning of artificial neural networks (ANNs) by partial least squares (PLS) algorithm, which use available process data to determine the level of contamination of models. Although these approaches provide promising results, they, unfortunately, require specialized research engineers familiar with these models to periodically train and tune these algorithms for application to different processes, which reduces their generalizability and practical applicability. The listed studies were carried out only for tubular heat exchangers and battery charging stations. ANN, tuned by deep machine learning methods, is an efficient, high-speed non-linear computational structure, created by analogy with a biological neural system and consisting of simple and highly interconnected elementary units called nodes. The accuracy of the trained ANN is guaranteed to reach the accuracy of cumbersome regression models (Murdoch et al., 2019), which use significant computing resources during operation. The mentioned advantages, and the proven applicability of ANN for tubular heat exchangers, make it promising to create a network model for predicting time to reach the lower allowable limit of the heat transfer coefficient value for PHE, which is the subject of the current work.

2. Data extraction

The flowsheet of the five-effect evaporator, which is used in the process of sugar production to heat the thin juice and is examined in the work, is presented in Figure 1.

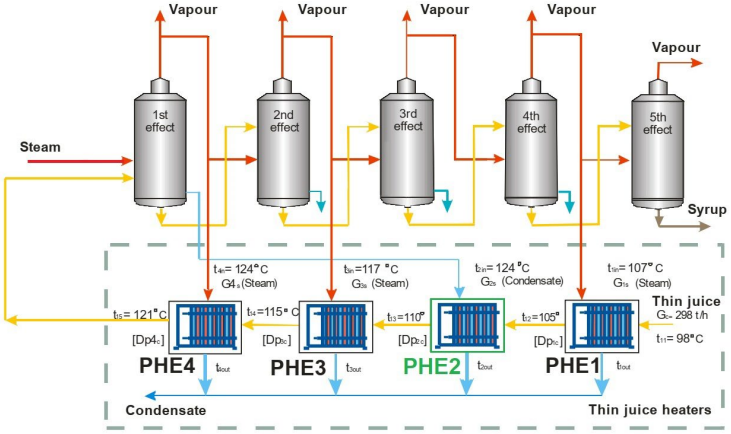


Figure 1: Flowsheet of 5-effect thin juice evaporation station

For efficient energy consumption and process intensification, the juice is preheated before evaporation. PHE2 (highlighted in green in Figure 1) heats up the thin juice by the condensate from the 1st evaporation effect. For the PHE2 position, the Alfa Laval plate-and-frame heat exchanger of M15M type with 150 heat transfer plates was selected and installed at the sugar plant. From the start of the operation, the heat and hydraulic characteristics of this heat exchanger were monitored. The data from on-site monitoring were given in the paper by Demirskyy et al. (2016).

The heat exchanger was operating 13 days after the last cleaning. The first measurement took place after the start-up, which in this case, took a long time because of several stops of the equipment. And the stable operating conditions started after 96 h of work. The operating parameters were calculated according to equations

presented by Arsenyeva et al. (2013). The values of overall heat transfer coefficients for clean and contaminated surface of PHE2 for different time periods during 13 days of operation are presented in Table 1.

Table 1: The operating parameters and calculated values of PHE2 heat exchanger

Parameters and calculated values / N	1	2	3	4	5
Parameters:					
Time point of process τ, hours	96	144	216	264	312
The flowrate of thin juice G_2, m^3/h	265	260	270	277	265
Inlet temperature of thin juice T_{21}, ºC	103	101	100.5	102	101.7
Outlet temperature of thin juice T_{22}, ºC	108	105	106	107	106
Pressure loss of thin juice flow ΔP_2, Pa	0.5	0.5	0.6	0.6	0.6
Condensate flowrate G_1, m^3/h	65	63	61	66	64
Inlet temperature of condensate T_{11}, ºC	123.5	123.5	123.5	123.5	123.5
Outlet temperature of condensate T_{12}, ºC	105	102.8	104.8	106.1	104.8
Calculated values:					
Fouling thermal resistance ×10^4 R_f, m^2 K / W	0.27	1.10	1.55	1.67	1.9
Heat transfer coefficient for clean surface K, W/(m^2·K)	2.673	2.220	2.668	2.686	2.382
Heat transfer coefficient with fouling K_f, W/(m^2·K)	2.493	1.784	1.887	1.853	1.640
Relation $K_R = K_f/K$×100, %	93.3	80.4	70.7	69.0	68.8

Assuming that the changes in the parameter values in the time intervals between iterations are linear (similar to Galčíková et al., 2022), the parameters $\Delta P_2(\tau), G_i(\tau), T_{ij}(\tau)$ were emulated with a random deviation of up to 5 % of the estimated parameter value, and the parameters $K_f(\tau), K(\tau)$ were calculated similarly to the ones in Table 1. Based on five reference points and "black-box" approach, production rules of the form $IF(Parameters)\ THEN(Calculated\ value)$ were formed, on the basis of which the hypersurface $F(\tau, G_1, T_{11}, T_{12}, G_2, T_{21}, T_{22}, \Delta P_2, K_f, K)$ was simulated on the time interval $\Delta\tau = 312 - 96 = 216$ h. As a result, data set of 427 points with uniform time increment $\tau_{j+1} = \tau_j + 0.5$ h for feed-forward neural network (FFNN) training was obtained. Data set of 200 points for testing $\{\Delta P_2(\tau), G_i(\tau), T_{ij}(\tau)\}$ with a random deviation of up to 12 % was obtained in the same way. Further, the already developed mathematical models are replaced by ANN models.

3. Neural networks for predicting PHE fouling indicators

This section briefly discusses the construction technique and testing results of the simplest FFNN for real-time prediction of the heat transfer coefficient using the data from the previous section. Below is a method for constructing RNNs with LSTM that allows one to study long-term dependencies, process sequential input data, and use feedback loops to transfer information from one-time step to another. As a result, the LSTM net considers the influence of trends (time gradients) of input parameters more adequately at the net`s outlet.

3.1 Feed-Forward neural network for real-time prediction of the heat transfer coefficient

The ability of neural networks to approximate unknown areas of the "input-output" mapping is widely used to identify, control and predict the behavior of objects. The FFNN properties are completely determined by the activation functions Φ used in the nodes of the hidden layer and forming a certain basis for the input vector image x. Complex objects are modeled using a multidimensional Gaussian radial basis (GRB) activation functions having a peak at the center of c and monotonically decreasing with distance from the center:

$$\varphi(x) = \Phi(\|x - c\|, \Sigma) = exp\left(-(x - c)^T \Sigma^{-1}(x - c)\right) = \exp\left(-\|x - c\|^2_{\Sigma^{-1}}\right), \quad (1)$$

where the covariance matrix Σ determines the shape, size and orientation of the so-called receptor field of the radial-basis function. At $\Sigma = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$ – is a hyper ellipsoid whose axes coincide with the axes of the input space and have a length $2\sigma_i$ of the i-th axis.

The task of learning of the approximating neural network is to find a function $F(x)$ so close $f(x)$ to that:

$$\|F(x) - f(x)\| \le \varepsilon, \forall x(k): k = 1,2, \ldots N, \quad (2)$$

where $F(x)$ – the mapping realized by the network, ε – is a small positive number, that determines the accuracy of the approximation.

The process of a model building is divided into two stages - structural and parametric identification, and the application of the ANN also requires solving two problems: determining the network structure, its setting parameters and training. Usually, a change in the network structure is made by its gradual complication of adding new nodes, performed each time when an additional identification error $e = d - y$ occurs when a new input signal appears, exceeding the permissible one. Training (parametric identification) consists in determining the network parameters and reduces to minimizing the identification error – as a rule, a quadratic error functional:

$$J(k) = \|\boldsymbol{\varepsilon}(k)\|^2 = \|\boldsymbol{d}(k) - \boldsymbol{y}(k)\|^2. \quad (3)$$

In practice, the most common are discrete learning algorithms of the form:

$$w_{ji}(k+1) = w_{ji}(k) + \eta(k) e_j(k) x_i(k), \quad (4)$$

The speed of the learning process using algorithms (3), (4) is completely determined by the choice of the parameter η_k that determines the step of the displacement in the space of the tunable parameters. It is natural to choose this parameter so that the rate of convergence of the current values of synaptic weights of NN hidden layer $w_j(k)$ to the optimal hypothetical weights will be maximal. Introducing into consideration that the optimal value of the step parameter may be obtained in the form:

$$\eta(k) = \|x(k)\|^{-2}, \quad (5)$$

That leads to a one-step learning the known Kaczmarz – Widrow – Hoff (KWH) algorithm (Kaczmarz, 1993):

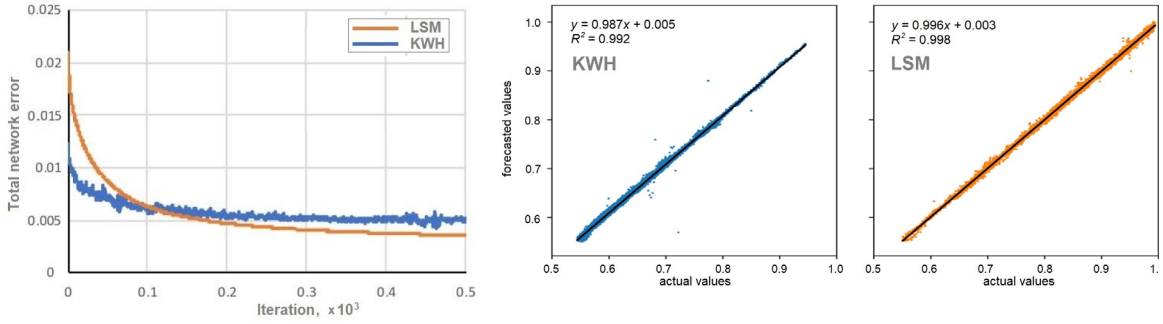
$$w_j(k+1) = w_j(k) + \frac{e(k)x(k)}{\|x(k)\|^2}. \quad (6)$$

FFNN architecture implements the mapping $R^{10} \longrightarrow R^1$: $K_R(\tau) = F(\tau, G_1, T_{11}, T_{12}, G_2, T_{21}, T_{22}, \Delta P_2, K_f, K)$.

Table 2: FFNN $K_R(\tau) = F(\tau, G_1, T_{11}, T_{12}, G_2, T_{21}, T_{22}, \Delta P_2, K_f, K)$ Structure

	Input Layer	Hidden Layer	Output Layer	Activation Function
Nodes	10	96	1	GBR

FFNN was built in the NeuroPh package (Perry J.S., 2018), its tuning was performed using the least squares method (LSM) and accelerated KWH algorithms. Figure 2b shows two scatterplots representing the prediction accuracy of each algorithm. Actual values are displayed on the horizontal axis, and predicted values are displayed on the vertical axis. The coefficient of determination (R^2) is less in KWH case because one-step algorithm computes steeper weight gradients when tuned, resulting in a local overfitting effect. And although the total network calculation error already at the 200th training iteration was set at 0.005 (Figure 2a), which exceeds the value of 0.021 achieved by Arsenyeva et.al. (2022), it should be noted that the presented model does not take into account trends in temporal changes in the input parameters. The small amount and the short period of extraction (216 h) of real current PHE parameters significantly reduce the adequacy of the model (Demirskiy et al., 2016). FFNN, designed as a static model, can be effective for the operative evaluation of $K_R(\tau)$.



a *b*

Figure 2: FFNN learning and testing results: a) total network calculation error; b) KWH and LSM algorithms` accuracy

3.2 Recurrent neural network with a long short-term memory

Unlike the one proposed by Trigkas et.al. (2022), Time Delayed NN with two hidden layers for battery state of charge prediction stores input states from previous $4n$ steps in memory to make a prediction for n future steps,

whereas LSTM selectively stores states of NN inputs. A variant of a recurrent NN (RNN) such as LSTM is designed to solve the problem of the decay of the temporal gradient of the input parameter (the parameter changes little over time) and provide more efficient learning on long-term sequences. LSTM does not have a fundamentally different architecture from RNN, but it uses a different function to calculate the hidden state. LSTM cells can be thought of as black boxes that take as input the previous state $h(t-1)$ and the current input $x(t)$. Inside these cells, they decide which state of memory to keep, and from what moment $(t-k)$ to erase the previous states $h(t-k)$. Cells then concatenate their previous state, current memory, and input parameter. These elements are very effective in storing long-term dependencies and analyzing the temporal gradients (trends) of these dependencies.

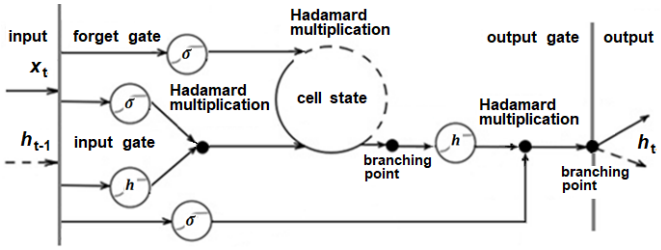


Figure 3: Scheme of the sequence of computational processes at each time step in LSTM

An LSTM cell can be schematically represented (Figure 3) as a sequence of processes that occur at each step:

- The input data $x(t)$ and the previous output state $h(t-1)$ are fed to the input gate, the forget gate, and the output gate. The input data and the previous output state are passed through various weights and biases before being fed to the input of the activation functions.
- The input gate decides what information from the current input should be retained. This is done by elementary multiplication of the output of the sigmoid function and the hyperbolic tangent.
- The forget gate decides which information from the current internal state should be forgotten. This is done by elementary multiplication of the output of the sigmoid function and the current internal state.
- The internal state is updated by combining the results of the input gate and the forget gate. The updated internal state can take new information into account and forget the old.
- The output gate decides what information from the updated internal state should be passed to the output. This is done by elementary multiplication of the output of the sigmoid function and the hyperbolic tangent of the updated internal state.
- The output state $h(t)$ is passed to the next time step or used as the final output of the LSTM cell.

Thanks to this structure, LSTM can analyse and remember long-term dependencies in data and train efficiently on sequences of different lengths. Since the FFNN has 10 input nodes $\tau, G_1, T_{11}, T_{12}, G_2, T_{21}, T_{22}, \Delta P_2, K_f, K$ and 1 output node - $\tau(K_f, K)$, a similar structure can be used to create an LSTM network with 96 nodes in one hidden layer. Due to the small data set for training the LSTM network, the coefficient of determination R^2 = 0.8918 in testing data set was achieved. The activation function (AF) ReLu: $A(x) = \max(0, x)$ is less computationally demanding other AFs performs simple mathematical operations, like an activator allows only some neurons be activated, which makes activations sparse and efficient. Using the most common functions like hyperbolic tangent or sigmoid as AF entails activating all neurons in the output layer. Hence, almost all activations are involved in describing the network output: the activation is dense, and this is inefficient in the computing resources used.

Table 3: LSTM $\tau(K_f, K) = F(\tau, G_1, T_{11}, T_{12}, G_2, T_{21}, T_{22}, \Delta P_2, K_f, K)$ Structure

	Input Layer	Hidden Layer	Output Layer	Activation Function
Nodes	10	96	1	ReLu

4. Conclusions

The proposed static FFNN makes it possible to determine the PHE heat transfer coefficient only in short-term forecasts and for operational evaluation. The coefficient of determination (accuracy of the model in the test data set) of FFNN trained by both algorithms LSM and KWH reached 0.99. An RNN architecture with LSTM is also proposed for long-term prediction of the time-point to reach the limiting heat transfer coefficient, as the accuracy of the model on the test data set reached 0.89. It is noted that in order to obtain adequate predictive models, a

huge amount of data on the current parameters associated with the process of deposit formation on the PHE plates is required. The presented models, after training on real data for specific instances of the PHEs, can be used in pairs for rapid evaluation and long-term prediction of these parameters in sugar and other industries.

Acknowledgments

This research was supported by the Ministry of Education and Science of Ukraine, project No. 0123U102775. Olga Arsenyeva is grateful to the Sustainable Process Integration Laboratory – SPIL, NETME Centre, Faculty of Mechanical Engineering, Brno University of Technology - VUT Brno, Technická 2896/2, 616 69 Brno, Czech Republic for the support and collaboration.

References

Arsenyeva O., Matsegora O., Kapustenko P., Yuzbashyan A., Klemeš J.J., 2022, The water fouling development in plate heat exchangers with plates of different corrugations geometry, Thermal Science and Engineering Progress, 32, 101310.

Arsenyeva O.P., Crittenden B., Yang M., Kapustenko P.O., 2013, Accounting for the thermal resistance of cooling water fouling in plate heat exchangers, Applied Thermal Engineering, 61(1), 53-59.

Bansal B., Chen X.D., Müller-Steinhagen H., 2008, Analysis of “classical” deposition rate law for crystallisation fouling, Chemical Engineering Process: Process Intensif, 47(8), 1201-10.

Demirskyy O.V., Kapustenko P., Khavin G., Arsenyeva O., Matsegora O., Kusakov S., Bocharnikov I., Tovazhnianskyi V., 2016, Investigation of fouling in plate heat exchangers at sugar factory, Chemical Engineering Transactions, 52, 583-588.

Galčíková L., Horváthová M., Oravec J., Bakošová M., 2022, Self-Tunable Approximated Explicit Model Predictive Control of a Heat Exchanger, Chemical Engineering Transactions, 94, 1015-1020.

Jradi R., Marvillet C., Jeday M.R., 2020, Modeling and comparative study of heat exchangers fouling in phosphoric acid concentration plant using experimental data, Journal of Heat and Mass Transfer, 56, 2653-2666.

Kaczmarz S., 1993, Approximate solution of systems of linear equations, International Journal of Control, 53,1269-1271.

Klemes J.J., Arsenyeva O, Kapustenko P, Tovazhnyanskyy L., 2015, Compact heat exchangers for energy transfer intensification: low grade heat and fouling mitigation, CRC Press, Boca Raton, FL, USA.

Lameh M., Dhabia M.A., Linke P., 2021, Cost Analysis for CO2 Reduction Pathways, Chemical Engineering Transactions, 88, 583–588.

Kashani M.N., Aminian J., Shahhosseini S., Farrokhi M., 2012, Dynamic crude oil fouling prediction in industrial preheaters using optimized ANN based moving window technique, Chemical Engineering Research and Design, 90(7), 938-949.

Murdoch W.J., Singh C., Kumbier K., Abbasi-Asl R., Yu B., 2019, Definitions, Methods, and Applications in Interpretable Machine Learning, Proceedings of the National Academy of Sciences of the United States of America, 116, 22071–22080.

Nait Malek Y., Najib M., Bakhouya M., Essaaidi M., 2019, Forecasting the State-of-Charge of Batteries in Micro-Grid Systems, 4th World Conference on Complex Systems (WCCS), 1-6.

Perry J.S., 2018, Create an artificial neural network using the Neuroph Java framework, <developer.ibm.com/tutorials/cc-ann-neuroph-machine-learning/ > accessed 10.05.2023.

Sundar S., Rajagopal M.C., Zhao H., Kuntumalla G., Meng Y., Chang H.C., Shao C., Ferreira P., Miljkovic N., Sinha S., Salapaka S., 2020, Fouling modeling and prediction approach for heat exchangers using deep learning, International Journal of Heat and Mass Transfer, 159, 120112.

Trigkas D., Gravanis G., Diamantaras K., Voutetakis S., Papadopoulou S., 2022, Energy Management in Microgrids Using Model Predictive Control Empowered with Artificial Intelligence, Chemical Engineering Transactions, 94, 961-966.

Trzcinski P., Markowski M., 2018, Diagnosis of the fouling effects in a shell and tube heat exchanger using artificial neural network, Chemical Engineering Transactions, 70, 355-360.

Wu N., Wang H., 2018, Deep learning adaptive dynamic programming for real time energy management and control strategy of micro-grid, Journal of cleaner production, 204, 1169-1177.