

Highlights

Validation of an integrated data-driven surrogate model and a thermo-hydraulic network based model to determine boiler operational loads using a fully connected mixture density network

B.T. Rawlins, Ryno Laubscher, Pieter Rousseau

- Development of mixture density network using simulation data.
- Model based on validated CFD model of a 620 MW_e sub-critical boiler.
- Surrogate model prediction errors are below 10%.

Validation of an integrated data-driven surrogate model and a thermo-hydraulic network based model to determine boiler operational loads using a fully connected mixture density network

B.T. Rawlins^{a,*}, Ryno Laubscher^b and Pieter Rousseau^a

^a*Department of Mechanical Engineering, Applied Thermal-Fluid Process Modelling Research Unit, University of Cape Town, Library Road, Rondebosch, Cape Town, 7701, South Africa*

^b*Department of Mechanical Engineering, Stellenbosch University, Banghoek Road, Stellenbosch, 7600, South Africa*

ARTICLE INFO

Keywords:

Mixture density network

Surrogate modelling

Boiler operation


ABSTRACT

A data-driven surrogate model is proposed for a $620MW_e$ sub-critical power boiler. The surrogate model was developed using computational fluid dynamic (CFD) simulation data. The simulation data covered a varied range of inputs.

1. Introduction

The use of neural networks for the modelling of energy systems has been awesome. Optimization of a plant is extremely fun. Using the surrogate model development the model can be developed

*Corresponding author

 rw1bra001@myuct.ac.za (B.T. Rawlins)

ORCID(s):

Nomenclature

λ [W/mK]	Thermal conductivity	S_h [W/m ³]	Energy source term
μ [Pa.s]	Viscosity	S_k [kg/m ³]	Species source term
<i>abbreviation</i>	explanation for the abbreviation	S_m [N/m ³]	Momentum source term
<i>CFD</i>	Computational Fluid Dynamics	T_g [K]	Gas temperature
E [J/kg]	Total energy	u [m/s]	Directional velocity
p [Pa]	Pressure	Y_k [kg/kg]	Mass fraction of species k
S [kg/m ³]	Mass source term		

2. Applicable machine learning theory

The present work makes use of various machine learning architectures to develop a surrogate model that can predict the thermal and combustion characteristics of a utility scale boiler using high level inputs. The current section discusses the details of the respective architectures, namely linear regression, ANN and ANN-MDN models.

2.1. Multiple linear regression

The present work makes use of a multiple linear regression model as a base model for comparative purposes. The main assumption of a multiple linear regression model is that the output/s can be calculated from a linear combination of the variable inputs. In other words a linear regression models aims to determine the quantitative relationship between the dependent and independent variables [10]. The representation of the i -th dependent variable (y_i) can be written as follows for m -number of independent variables (x_{mi}):

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_m x_{mi} \quad (1)$$

$i = 1, 2, 3 \dots n$

where β_0 is a constant term, β_m is the m -th coefficient, and n is the total number of observations.

The optimal solution can be estimated by minimizing the cost function (J). A cost function usually calculates the difference between the estimated and the desired values and is reported as a single number. Multiple linear regression problems namely utilize the mean square error (MSE) between the desired and estimated (\hat{y}_i) values [4], which is given in equation 2.

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

To minimize the cost function, the gradient descent algorithm is the preferred method [10], which is an iterative procedure used to find the local minimum/maximum. Considering the cost function as a function of the weight the algorithm can be written as shown in equation 3.

$$\beta_m = \beta_m - \eta \frac{\partial}{\partial \beta_m} J(\beta_m) \quad (3)$$

where η is known as the learning rate.

In most cases the relationship between the dependent and independent variables are not always linear. Special linear basis models, such as polynomial, sinusoidal and radial can be used to optimize the training results [6]. For the current work a multiple linear progression model is utilized as a base model for comparison.

2.2. Artificial neural networks (ANNs)

There are many classifications of ANNs, with multilayer perceptron networks (MLP) being the standard representation [3]. Typically MLPs are adapted for supervised learning problems where the input variables are mapped to labelled output variables. The relationship between the input and output variables are learned by optimizing the weights (\bar{w}) and biases (\bar{b}) to minimize a selected cost function, which for most cases is the MSE given in equation 2. A standard MLP schematic is given in figure 1, illustrating the common features of a MLP, that being the input, hidden and output layers.

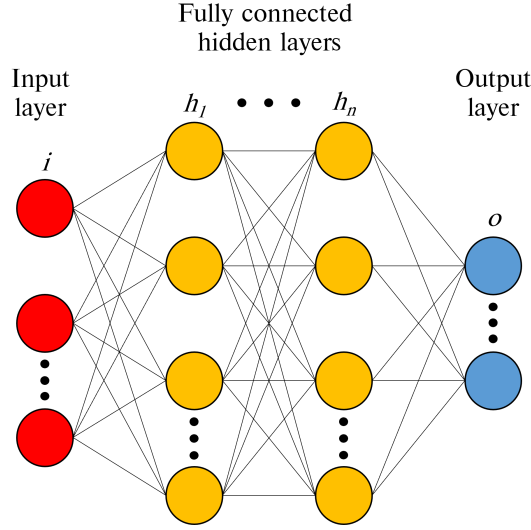


Figure 1: Traditional MLP schematic

To calculate the output values (\hat{y}_i) the forward propagation algorithm is utilized, which calculates the output for each layer and moves sequentially through the network till an output is determined. Each network layer output is calculated using two steps, the first being the calculation of the summed signal \bar{z}_l and secondly the use of an activation function to generate the output signal \bar{h}_l . Equation 4 highlights the first step, where \bar{h}_{l-1} is the output signal from the previous layer.

$$\bar{z}_l = \bar{h}_{l-1} \cdot \bar{w}_l + \bar{b}_l \quad (4)$$

The result of the equation 4 is subsequently passed to an activation function ($\bar{h}_l = \sigma_l(\bar{z}_l)$). There are various activation functions that can be utilized, such as linear, ReLu, Elu, and the hyperbolic-tangent [3], with the final layer activation function usually being linear for regression models. The current work makes use of ReLu and linear activation functions for the hidden layers. The ReLu and linear activation functions are shown in Equation 5.

$$\bar{h}_l = \sigma_{ReLU}(\bar{z}_l) = \bar{z}_l = \begin{cases} \bar{h}_{l-1} \cdot \bar{w}_l + \bar{b}_l & \text{if } \bar{z}_l > 0 \\ 0 & \text{if } \bar{z}_l < 0 \end{cases} \quad (5)$$

$$\bar{h}_l = \sigma_{linear}(\bar{z}_l) = \bar{z}_l = \bar{h}_{l-1} \cdot \bar{w}_l + \bar{b}_l$$

When the forward propagation step is complete, the network weights and biases can be updated to minimize the cost function (refer to Equation 2) using the backward propagation method [3]. The methodology calculates the gradient of the cost function with respect to the weights and biases for each layer. Once the gradients have been calculated the weights and biases are updated using the gradient descent algorithm. The current work makes use of the Adam [3] alternative to the gradient descent algorithm and is illustrated in Equation 6. Forward- and backward-propagation algorithms would be iteratively conducted until the cost function is reduced to below a desired threshold.

Algorithm 1.

$$\begin{aligned}
\bar{m} &\leftarrow \beta_1 \bar{m} + (1 - \beta_1) \nabla_{\theta} J_{MSE}(\bar{\theta}) \\
\bar{s} &\leftarrow \beta_1 \bar{s} + (1 - \beta_2) \nabla_{\theta} J_{MSE}(\bar{\theta}) \otimes \nabla_{\theta} J_{MSE}(\bar{\theta}) \\
\bar{m} &\leftarrow \frac{\bar{m}}{1 - \beta_1^t} \\
\bar{s} &\leftarrow \frac{\bar{s}}{1 - \beta_2^t} \\
\bar{\theta} &\leftarrow \bar{\theta} - \eta \bar{m} \otimes (\sqrt{\bar{s}} + \epsilon)^{-1}
\end{aligned} \tag{6}$$

The variable $\bar{\theta}$, of Equation 6, represents the model parameters weights (\bar{w}) and biases (\bar{b}) for each layer. Scaling (\bar{s}) and momentum (\bar{m}) matrices are initialized to zero when beginning the training phase, t is the iterative counter, β_1 and β_2 are the momentum and scaling decay hyper-parameters set to values of 0.9 and 0.999 respectively. Lastly ϵ is a smoothing term set to a value of 10^{-8} .

2.3. Mixture density networks (MDNs)

MDNs are simply a MLP with an additional layer that allows the prediction of the Gaussian distribution.

3. Data generation

A steady-state multiphase non-thermal equilibrium CFD model was used to generate the target data, which was subsequently used for training/development of an appropriate surrogate model.

3.1. CFD model setup

The current study makes use of the commercial CFD software package ANSYS[®] Fluent 2019 R3 to resolve the fluid flow, heat transfer and combustion processes for a 620 MW_e utility scale coal fired boiler. The computational domain is modelled on a symmetry plane half way through the depth of the boiler. Figure 2 highlights the computational domain and defines the important boundary conditions.

The general conservation equations, which include, continuity, momentum, energy and species, were solved using a Eulerian approach. The subsequent equations are stated in Equation (7).

$$\begin{aligned}
\frac{\partial}{\partial x_i} (\rho \bar{u}_i) &= S \\
\frac{\partial}{\partial x_i} (\rho_{eff} u_i u_j) + \frac{\partial \bar{p}}{\partial x_j} &= \frac{\partial}{\partial x_i} \left[\mu \left\{ \frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} - \frac{2}{3} \delta_{ij} \frac{\partial u_i}{\partial x_i} \right\} \right] + \frac{\partial}{\partial x_i} (-\rho \overline{u'_i u'_j}) + S_m \\
\frac{\partial}{\partial x_i} (u_i [\rho E + p]) &= \frac{\partial}{\partial x_j} \left[\lambda \frac{\partial T_g}{\partial x_j} \right] + S_h \\
\frac{\partial}{\partial x_i} (\rho u_j Y_k) &= -\frac{\partial}{\partial x_j} (\bar{J}_k) + \sum_r R_{j,r} + S_k
\end{aligned} \tag{7}$$

The resolution of the Reynolds stress term found in the momentum equation, $-\rho \overline{u'_i u'_j}$, was approximated using the Boussineq equation [9]. In the present study the realizable k- ϵ turbulence model was utilized to address the turbulence closure problem, this model was selected for its applicability in modelling the effects of coal-fired swirl burners [5].

The P1 radiation model was used to resolve the radiative field in the domain. Particle transport was modelled using a multiphase approach, further details on the approach are provided in the validation study of Rawlins et al [7]. The combustion follows a four step sequential process, beginning with the heating and evaporation of the moisture present in the fuel, followed by the devolatilization process where the volatiles are liberated from the solid particle, succeeded by the phenomena of char burnout, and finally the gas phase reactions would commence. The char oxidation

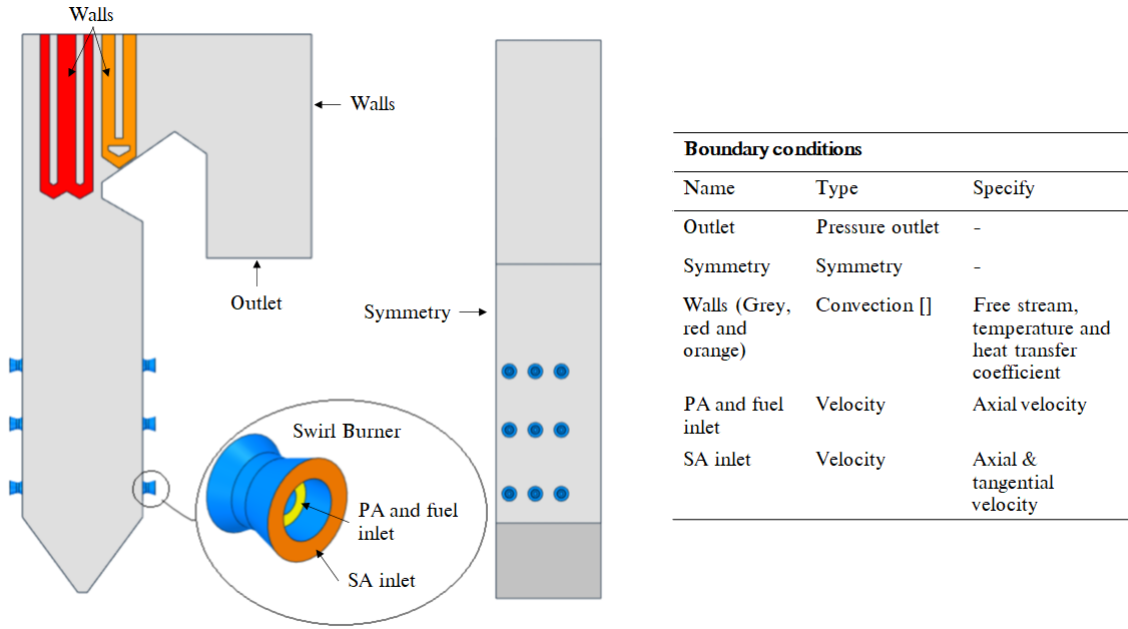


Figure 2: CFD model geometry and boundary condition descriptions.

Table 1
Summary of combustion models and constants used in the CFD model

Model	Equation/s	Constant/s
<i>Devolatilization</i>		
Single rate kinetic model	$\frac{dm_{vol}}{dt} = R_{vol}(m_{0,vol} - m_{vol}),$ $R_{vol} = A_{vol} \exp\left(\frac{E_{a,vol}}{RT_p}\right)$	$A_{vol} = 2 \times 10^5 [s^{-1}],$ $E_{a,vol} = 6.7 \times 10^7 [J/kmol] - [8]$
<i>Char oxidation</i>		
Diffusion/kinetic - [2]	$\frac{dm_{char}}{dt} = -A_p p O_2 \frac{R_{diff} R_c}{R_{diff} + R_c},$ $R_c = A_c \exp\left(\frac{E_{a,c}}{RT_p}\right),$ $R_{diff} = \frac{5 \times 10^{-12}}{d_p} \left(\frac{T_g + T_p}{2}\right)^{0.75}$	$A_c = 0.0053 [kg/m^2 s Pa],$ $E_{a,c} = 8.37 \times 10^7 [J/kmol] - [8]$
<i>Gaseous reactions of volatiles and CO</i>		
Eddy dissipation model - [1]	$R_{k,r,P} = \vartheta_{k,r} M_{w,k} A B \rho_k^\epsilon \min\left(\frac{\sum_j Y_p}{\sum_j \vartheta_{j,r} M_{w,j}}\right),$ $R_{k,r,R} = \vartheta_{k,r} M_{w,k} A \rho_k^\epsilon \min\left(\frac{Y_R}{\epsilon_{R,r} M_{w,R}}\right)$	$A = 4.0, B = 0.5$

reaction product species was set to that of carbon monoxide (CO). For the gas-phase reactions the turbulence-chemistry interaction was approximated using the eddy dissipation model (EDM). A summary of the combustion equations and constants are provided in Table 1.

The simulations were solved using the SIMPLE pressure-velocity coupling scheme. The pressure term was discretized using the PRESTO! scheme. Momentum, species and energy equations were discretized using the second-order upwind scheme and the turbulent kinetic energy and dissipation rate using the first-order upwind scheme. The numerical mesh was generated using quadrilateral elements consisting of 6 million cells. The convergence criteria for the simulation model was set to 1×10^{-3} for the continuity equation, 1×10^{-4} for the velocity equations, 1×10^{-6} for the remaining transport equations, and 1×10^{-4} for monitored key parameters.

Table 2

Design of experiments input ranges for simulations

Input variable	Min	Max	Units
Total fuel flow rate for mills 1 to 6	39.5	120.2	kg/s
Excess air	1.155	1.401	%
Fuel proximate analysis moisture mass fraction, Y_{H_2O}	0.025	0.085	kg/kg
Fuel proximate analysis ash mass fraction, Y_{ash}	0.259	0.559	kg/kg
Platen SH fouling thermal resistance, R_{platen}	0.004	0.007	$m^2 K/W$
Final SH fouling thermal resistance, R_{final}	0.01	0.017	$m^2 K/W$

3.2. Simulated dataset

As previously mentioned the aim this study is to illustrate the use of a data-driven surrogate model, integrated with a 1D process model, to predict the heat loads to the various heat exchanging components, the flue-gas composition and exit gas temperatures for a utility scale boiler using various high-level inputs. The inputs include the the following, the excess air ratio per burner, the total mill flowrate for the six mills in operation, the average steam temperatures for the platen and final superheaters, the fouling resistance for the platen and final superheaters, the composition of ash and moisture of the fuel and the gross calorific value of the fuel. Thus the input field has a dimensionality of $d_{input} = 14$.

A design of experiments (DOE) was conducted to generate a set of 180 simulation cases to obtain a representative set of results. The various model input ranges used in the DOE are given in Table 2. The ranges were selected to cover a wide range of operational loads with maximum continuous ratings (MCR) between 100% and 30%.

4. Model development

The present work makes use of two types of machine learning models, namely a standard artificial neural network (ANN) and a mixture density designated model connected to a standard ANN (MDN-ANN). The following section will discuss the hyper parameter tuning and final selected model configuration. The programming language Python 3.7.8 and the Tensorflow machine learning libraries were utilized in the present study.

4.1. Model configuration

4.2. Hyper parameter tuning & final model selection

table of NN and MDN data comparison for tuning

Table 3

Hyper-parameter search space for fully connected NN and MDN models

Parameter	NN search space	MDN search space
Number of distributions	-	2,3,4
Number of layers	2,3,4	2,3,4
Number of neurons per layer	10, 40, 80, 100	10, 40, 80, 100
Learning rates	1e-3, 1e-4, 1e-5	1e-3, 1e-4, 1e-5
Mini batch sizes	16, 32, 64	16, 32, 64

Table 4

ANN model selection results

Parameter	NN search space	MDN search space
Number of distributions	-	2,3,4
Number of layers	2,3,4	2,3,4
Number of neurons per layer	10, 40, 80, 100	10, 40, 80, 100
Learning rates	1e-3, 1e-4, 1e-5	1e-3, 1e-4, 1e-5
Mini batch sizes	16, 32, 64	16, 32, 64

Table 4 highlights the hyper-parameter search spaces for both the ANN and ANN-MDN model. The ANN-MDN has an added parameter namely the number of additional distributions the ANN-MDN would fit to the output data.

The hyper-parameter search was conducted in a sequential manner with the number of layers and neurons per layer being the initial step, after which the learning rates were varied, followed by the batch sizes. The ANN-MDN hyper-parameter tuning was conducted in a similar manner only that an additional step is required to establish the best amount of distributions.

5. Results and discussion

5.1. Multiple load validation

Validation includes the use of integrating the surrogate model with a network based-process model of the water/steam side of the utility boiler under investigation. A C# script is used to access the Python API available in the process modelling software, Flownex SE® 2021. This allows for predictions to be made using the trained MDN model. The most probable predictions are retrieved using the surrogate model script and transferred to the respective process model components. A schematic of the surrogate and process models integration is provided in Figure 3.

Measured plant data for a 100%, 81% and 60% MCR load ratings was made available and is to be used compare the accuracy of the integrated model.

5.2. Utility boiler response to poor fuel combustion

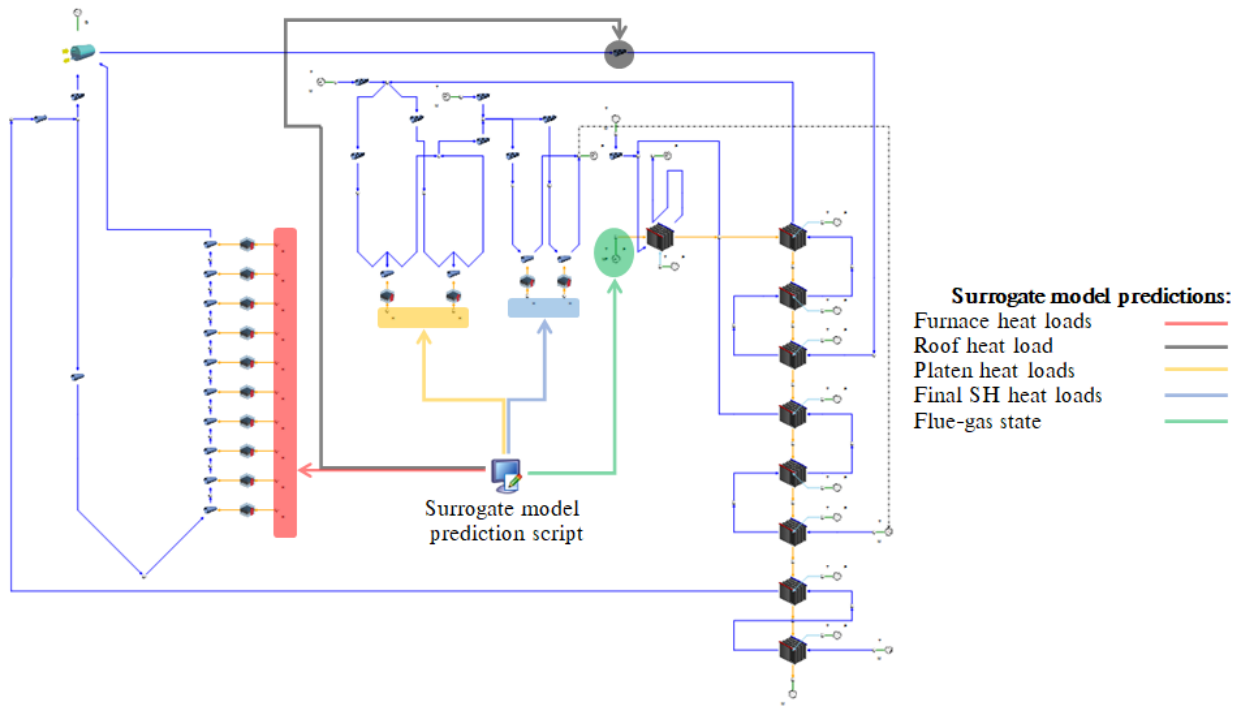
The fuel quality is an important factor when determining the Fuels with total moisture contents exceeding The following results illustrate the effects poor quality fuel has on the case studies boiler operational. The study made use of the developed surrogate model to investigate the steady state operation of the boiler burning poor quality coal. In this section the results of a study conducted

6. Conclusion

The present work has shown it is possible

CRedit authorship contribution statement

B.T. Rawlins: Methodology, Software, Validation, Formal analysis, Investigation, Writing original draft, Visualization.. **Ryno Laubscher:** Writing review & editing, Methodology, Resources, Conceptualization.. **Pieter Rousseau:** Writing review & editing, Resources, Conceptualization.



S

Figure 3: Schematic of integrated surrogate and network based process model

References

- [1] , 2021. ANSYS Fluent Theory Guide. 20 ed., Ansys Inc.
- [2] Baum, M., Street, P., 1971. Predicting the combustion behaviour of coal particles. *Combustion science and Technology* 3, 231.
- [3] Goodfellow, I., Bengio, Y., Courville, A., 2017. *Deep Learning*. First edit ed., MIT Press, Chennai.
- [4] Kochenderfer, M., Wheeler, T., 2019. *Algorithms for optimization*. 1 ed., The MIT Press, Cambridge, Massachusetts.
- [5] Modlinski, N., 2010. Computational modeling of a utility boiler tangentially-fired furnace retrofitted with swirl burners. *Fuel Processing Technology* 91, 1601–1608. doi:10.1016/j.fuproc.2010.06.008.
- [6] Rasmussen, C., Williams, C., 2006. *Gaussian Processes for Machine Learning*. 1 ed., The MIT Press, Cambridge, Massachusetts.
- [7] Rawlins, B.T., Laubscher, R., Rousseau, P., 2021. Validation of a thermal non-equilibrium Eulerian-Eulerian multiphase model of a 620 MWe pulverized fuel power boiler., in: Skatulla, S. (Ed.), 12th South African Conference on Computational and Applied Mechanics (SACAM2020), MATEC Web of Conferences, Cape Town. doi:<https://doi.org/10.1051/mateconf/202134700004>.
- [8] Sheng, C., Moghtaderi, B., Gupta, R., Wall, T.F., 2004. A computational fluid dynamics based study of the combustion characteristics of coal blends in pulverised coal-fired furnace. *Fuel* 83, 1543–1552. doi:10.1016/j.fuel.2004.02.011.
- [9] Versteeg, H., Malalasekera, W., 2007. *Introduction to Computational Fluid Dynamics, The finite volume method*. Second ed., Pearson Prentice Hall. doi:10.1002/9781119369189.
- [10] Wen, D., Pan, Y., Chen, X., Aziz, M., Zhou, Q., Li, N., 2022. Analysis and prediction of thermal stress distribution on the membrane wall in the arch-fired boiler based on machine learning technology. *Thermal Science and Engineering Progress* 28, 101137. doi:10.1016/j.tsep.2021.101137.