Highlights

An integrated data-driven surrogate model and thermofluid network-based model of a 620 MW_e utility-scale boiler using a fully connected mixture density network

B.T. Rawlins,Ryno Laubscher,Pieter Rousseau

- Data-driven surrogate model development using CFD solution data.
- Development of mixture density network using simulation data.
- Surrogate model integration with 1-D process model.

An integrated data-driven surrogate model and thermofluid network-based model of a 620 MW_e utility-scale boiler 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
Data-driven surrogate modelling
Computational fluid dynamics

ABSTRACT

An integrated data-driven surrogate model and one-dimensional (1D) process model of a 620 MW_e utility scale boiler is presented. A robust and computationally inexpensive computational fluid dynamic (CFD) model of the utility boiler was utilized to generate the solution dataset for surrogate model training and testing. Both a standard multi-layer perception (MLP) and MLP connected mixture density network (MDN) machine learning architectures are compared for use as a surrogate model to predict the furnace heat loads and the flue-gas inlet conditions to the convective pass. A hyperparameter search was performed to find the best MLP and MLP-MDN architecture. The MDN was selected for surrogate model integration since it showed comparable accuracy and provides the ability to predict the uncertainty in the results. Validation of the integrated model against plant data was performed for a wide range of loads and critical results were predicted within 5-8% of the measured results. The validated model was subsequently used to investigate the effects of using a poor-quality fuel for the 100% maximum continuous rating (MCR) load case. The uncertainties predicted by the surrogate model were propagated through the integrated model using the Monte Carlo technique, adding valuable insight into the operational limits of the power plant and the uncertainties associated with it.

1. Introduction

A significant portion of the developing world still rely on conventional coal fired power plants to meet local energy demands [1]. However, increased penetration of renewable energy sources onto electrical grids will force conventional base-load thermal power plants to operate on a mid-merit basis. Mid-merit operation requires improved flexibility with faster load following characteristics and start-up rates [2], as well as running at off-design and low load conditions for extended periods. This can result in an increase in thermal stresses on key operational components [3]. Therefore, it is important to be able to predict the integrated thermofluid response of boilers and the impact on components at off-design conditions.

Computational fluid dynamics (CFD) is well suited for the analysis of solid fuel combustion systems. It can adequately resolve the complex fluid flow and heat transfer phenomena, including the mechanical and thermal interaction between the flue-gas and the particles. CFD simulations are predominately used for detailed exploration of the flue-gas side flow and heat transfer phenomena of boilers to investigate pollutant production [4], do design optimization [5, 6] and study particle effects [7]. For utility-scale boilers the associated computational burden practically limits the use of CFD to steady-state analysis of specific areas of interest, such as the furnace and radiant heat exchangers, and does not allow for integrated whole-boiler modelling in other words modelling the water- and flue gas-sides in a single model

Investigating the integrated whole-boiler thermofluid response requires the use of 1-D process models to capture the response of the steam/water cycle integrated with suitable models of the flue-gas side in a co-simulation approach. Rousseau and Laubscher [1] used this approach to investigate the effects of high ash fuel on the heat absorption of

*Corresponding author
rwlbra001@myuct.ac.za (B.T. Rawlins)
ORCID(s):

radiant superheaters (SHs). As expected, this requires significant computational resources on the CFD side to allow for the integration of the two different models, and is therefore not a feasible approach for extensive studies of flexible operation. This has led to the development of reduced order multiphase modelling approaches [8, 9], which proved to reduce the computational effort by approximately 25% for a full-scale utility boiler. However, this is still not enough to allow direct integration of CFD into whole-boiler system analysis tools which cannot be used to quickly analyse different operating cases.

Data-driven surrogate models can potentially be used to reduce the computational burden introduced via the CFD model. Data driven surrogate models form part of statistical regression models, where the predicted objective function is fitted using true evaluations of the objective function [10]. Artificial neural networks (ANNs) are one of the most popular machine learning algorithms/techniques to facilitate the learning process. Other procedures include Kriging models [11], gradient boosting models [12, 13] and support vector machines [14].

The works of Singh and Abbassi [15] incorporated the use of an ANN and CFD model to investigate the transient thermal modelling of an off-highway machinery cabin. The ANN training data was obtained from a 1-D process model for the refrigeration cycle. The methodology was used successfully to integrate the ANN capabilities with a 3-D CFD simulation model thereby negating the need for a coupled simulation procedure. Similarly, Fei et al [11] made use of a fast set of reduced order models (ROMs), using the Kriging method, based on CFD data to investigate the retrofitting of a coal fired power plant under oxy-fuel conditions. Using an integrated ROM with a process model, a range of air-coal conditions were investigated, with the model illustrating sufficiently accurate results. However, no operational uncertainty was incorporated into the input parameters to account for real world operational effects (disturbances, local atmospheric conditions etc.).

Acknowledging and predicting the impact of uncertainties in conventional coal fired power plants are paramount for the safe operation and maintenance of these plants [16]. The uncertainties associated with the operational parameters of thermal power plants, such as fuel quality, combustion processes, and burner biasing, can lead to significant changes in the thermofluid response [17]. The design space load uncertainty of a combined cooling, heating and power system was investigated by Lu et al [18]. Using a multi-objective optimization model, factors such as pollutant emission, economy and the system reliability were incorporated to improve the operational performance. The model illustrated its future use for designers in providing the best configuration scheme under system load uncertainties.

It is noted that CFD cannot meet the near real time requirement for on line performance monitoring, due to the complex solution mechanisms and time-consuming calculations. However, it is a feasible alternative to generate training and testing datasets for ANN implementation, since using CFD does not interfere with normal operation of the plant. In addition to this a wider range of appropriate operating conditions can be simulated.

In the present work, an integrated data-driven surrogate model and 1-D process model is proposed to investigate the thermal response of a 620 MW_e utility scale boiler. An existing steady-state multiphase CFD model [8, 9] of the 620 MW_e utility scale boiler under investigation was used to generate a simulation dataset for 180 different operating states. To determine the required complexity of the data-driven surrogate model, two different ANN frameworks were considered, namely a standard multi-layer perception (MLP) and a mixture density network MDN, where the latter allows the introduction of uncertainty into the model predictions. A hyper-parameter search was conducted to establish the best performing model, with the lowest mean absolute error (MAE) being the essential criterion. Validation of the integrated surrogate and process model was performed for maximum continuous ratings (MCR) of 100%, 80% and 60% load for which measured data is available. Following this, the validated model was used to investigate the impact of poor fuel quality at 100% MCR, where the ash and moisture contents were varied for two separate cases.

2. Case study boiler

This section provides an overview of the case study boiler layout as well as a description of the integrated model.

2.1. Boiler layout

The overall layout of the case study boiler is shown schematically in Figure 1 [1].

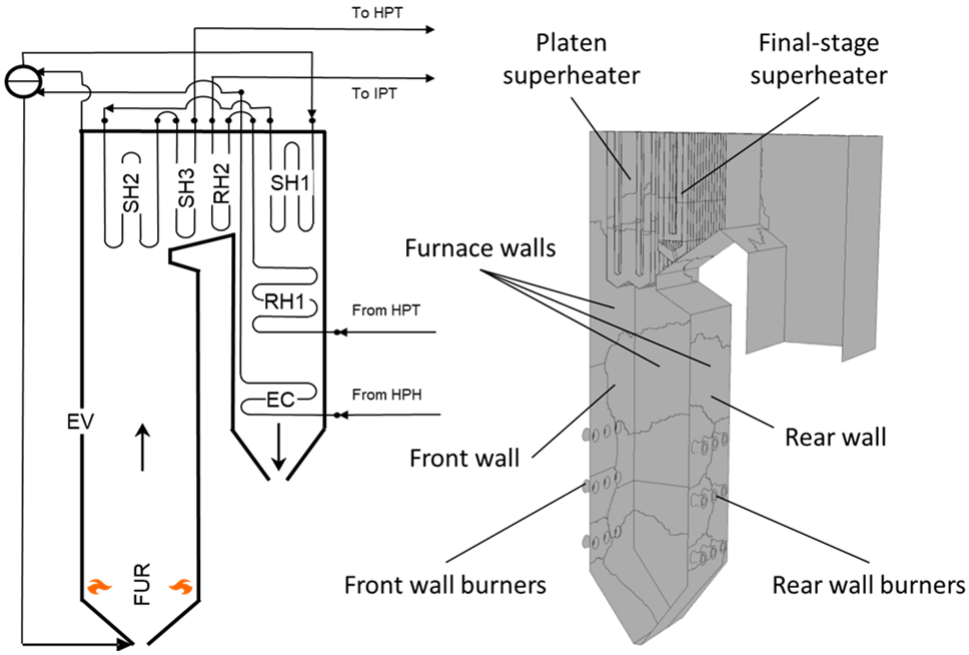


Figure 1: Case study boiler layout [1].

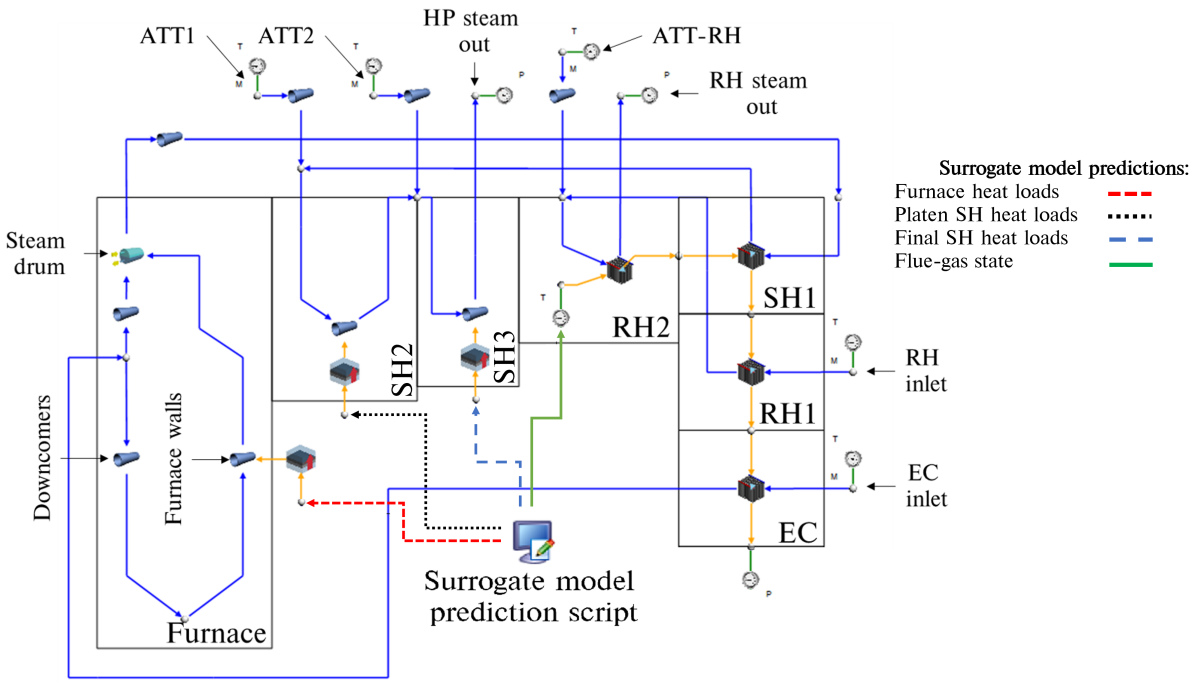
The boiler forms part of a 620 MWe power plant and consists of a furnace with water wall evaporator (EV), primary superheater (SH1), platen superheater (SH2) and final superheater (SH3), as well as a primary reheater (RH1), secondary reheater (RH2) and economizer (EC). The furnace height, width and depth are 64 m, 24 m, and 13.7 m respectively. The combustion chamber is fitted with three rows of six swirl burners each on the front and rear walls. Primary air is injected via the inner annulus of the burners, and secondary air via the outer annulus (refer to Figure 5).

2.2. Integrated model layout

The validation and the subsequent case-studies presented here, make use of an integrated surrogate and network based 1D process model. The process model is used to capture the thermodynamic response of the water/steam side of the utility boiler under investigation, with the developed data-driven surrogate model providing predictions of the gas-side thermal characteristics (i.e. combustion product species, temperature and incident radiation flux) and heat-exchanger heat loads to the EV, SH2 and SH3 walls.

The network model is set up in the process modelling software, Flownex SE® 2021 [19]. The software can solve both the steady state and transient forms of the fundamental conservation equations of mass, energy, and momentum together with built in fluid property relations and component characteristics representative of all the different types of components. It is also possible to add control elements to obtain a complete integrated dynamic system simulation model of a plant, sub system or component.

In the current work a C# script is used to access the Python application programming interface (API). This allows for predictions to be made using the trained MDN model developed in Python code. The predictions are retrieved using the script and transferred to the respective process model components as inputs. A schematic of the process model and its integration with the surrogate model is provided in Figure 2.



S

Figure 2: Schematic of network based process model integrated with the surrogate model.

The boiler can be split into three sections namely, the furnace consisting of EV waterwalls, the radiative pass consisting of SH2 and SH3, and the convective pass consisting of RH2, SH1, RH1 and the EC. The furnace heat load predictions to the EV walls determine the evaporation rate. The high pressure (HP) steam outlet flow rate is calculated as the sum of the attemperation flow rates (ATT1 and ATT2) and the evaporation flow rate. The attemperators are used to control the steam temperature to ensure an exit temperature of $808K$ at both the HP outlet and RH outlet. In addition to the furnace heat loads the surrogate model predicts the SH2 and SH3 heat loads. Lastly the surrogate model is able to predict the flue-gas conditions at the inlet to the convective pass (i.e. RH2). The EV, SH2 and SH3 walls are each represented by a single lumped parameter pipe flow component, interconnected with nodes that introduce the attemperation flows to the water/steam side. The convective pass components are modelled using an in-house heat exchanger model that accounts for the radiative (gas and direct), convective and conduction heat transfer mechanisms to and from the gas-side and steam-side control volumes, as well as between the up- and downstream heat exchangers on the gas side.

3. 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 MDN models.

3.1. Multivariate linear regression

A multivariate linear regression model was established as a base model for comparative purposes in the validation case study of Section 6.1. The main assumption of a multiple linear regression model is that the output/s can be calculated from a linear combination of the input variables. In other words, a linear regression models aims to determine the quantitative linear relationship between the dependent and independent variables [20]. The representation of the i-th dependent variable (y_i) can be written as follows for m-number of independent variables (x_{mi});

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{1i} + ... + \beta_m x_{mi} \\ i &= 1, 2, 3...n \end{aligned} \tag{1}$$

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. Multivariate linear regression problems typically utilize the mean square error (MSE) between the desired and estimated ($\hat{y}_i$) values [10], which is given in Equation 2.

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

In the present work the gradient descent algorithm [20] was utilised to minimize the cost function. The gradient descent algorithm 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.

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

where η is known as the learning rate.

In most cases the relationship between the dependent and independent variables are not linear. Special non-linear basis models, such as polynomial, sinusoidal, and radial can be used to optimize the training results [21]. For the current work a multiple linear regression model is utilized as a base model for comparison in Section 6.1.

3.2. Multilayer perception networks

Artificial neural networks (ANN) are machine learning systems inspired by biological neural activity [21]. There are many classifications of ANNs, with multilayer perception networks (MLP) being the standard representation [22]. 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 ($\overline{w}$) and biases ($\overline{b}$) 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 3, illustrating the common features of a MLP, that being the input, hidden and output layers.

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 until an output is determined. Each network layer output is calculated using two steps, the first being the calculation of the summed signal $\overline{z}_l$ and secondly the use of an activation function to generate the output signal $\overline{h}_l$. Equation 4 highlights the first step, where $\overline{h}_{l-1}$ is the output signal from the previous layer.

$$\overline{z}_l = \overline{h}_{l-1} \cdot \overline{w}_l + \overline{b}_l \tag{4}$$

The result of equation 4 is subsequently passed to an activation function ($\overline{h}_l = \sigma_l(\overline{z}_l)$). There are various activation functions that can be utilized, such as linear, ReLu, Elu, and the hyperbolic tangent [22], with the final layer activation function usually being linear for regression models. The current work makes use of ReLu for the hidden layers, since the ReLu are simple to implement and fast to compute [10], and a linear activation function for the output layer/s. The ReLu and linear activation functions are shown in Equation 5.

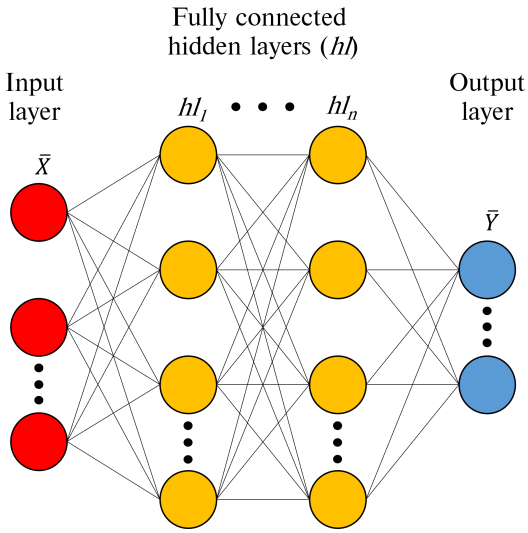


Figure 3: Traditional MLP schematic.

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

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 [23]. 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 [22] alternative to the gradient descent algorithm, which 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.

$$\overline{m} \leftarrow \beta_1 \overline{m} + (1-\beta_1)\nabla_\theta \boldsymbol{J}_{MSE}(\overline{\theta})$$
$$\overline{s} \leftarrow \beta_1 \overline{s} + (1-\beta_2)\nabla_\theta \boldsymbol{J}_{MSE}(\overline{\theta}) \otimes \quad \nabla_\theta \boldsymbol{J}_{MSE}(\overline{\theta})$$
$$\overline{m} \leftarrow \frac{\overline{m}}{1-\beta_1^t} \tag{6}$$
$$\overline{s} \leftarrow \frac{\overline{s}}{1-\beta_2^t}$$
$$\overline{\theta} \leftarrow \overline{\theta} - \eta \overline{m} \otimes (\sqrt{\overline{s}+\epsilon})^{-1}$$

The variable $\bar{\theta}$ in Equation 6 represents the model weights and biases for each layer. Scaling ($\bar{s}$) and the momentum ($\bar{m}$) matrices are initialized to zero when beginning the training phase. t is the iteration counter, while 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}. In the present work various learning rates (ν) were investigated during the hyper-parameter search for both the MLP and MDN neural architectures.

3.3. Mixture density networks (MDNs)

Mixture density networks are fundamentally built from two components, namely an ANN and a mixture model. This allows for multi-modal predictions. The ANN can be a standard feed-forward MLP or a recurrent neural network (RNN), with RNNs being used in transient applications with at least one feedback loop [24].

MDNs are used to predict the parameters of a probability distribution ($P(\overline{X} \mid \overline{Y})$), allowing for non-Gaussian distributions to be modelled, thus making MDNs a probabilistic machine learning framework. MDNs estimate the conditional probability distribution as a mixture of Gaussian distributions where the mixing coefficients ($\overline{\pi}_k$) and component densities are flexible functions of the input data ($\overline{X}$). Equation 7 illustrates the conditional probability function.

$$P(\overline{X} \mid \overline{Y}) = \sum_{k=1}^{K} \overline{\pi}_k(\overline{X}) \cdot \mathbb{N}(\overline{Y} \mid \overline{\mu}_k(\overline{X}), \overline{\sigma}_k^2(\overline{X})) \tag{7}$$

where K represents the number of selected normal distributions, while $\overline{\mu}_k$ and $\overline{\sigma}_k^2$ are the predicted means and variances for each distribution k given the input data $\overline{X}$ respectively.

A schematic of a simple MDN network is given in Figure 4, highlighting the mixing coefficients, predicted means and deviations. It is shown that modifications are made to the output layer by splitting the network output into three parts to calculate the $\overline{\pi}_k$, $\overline{\mu}_k$ and $\overline{\sigma}_k^2$ for each k distribution. This enables the MDN network to learn the $P(\overline{X} \mid \overline{Y})$.

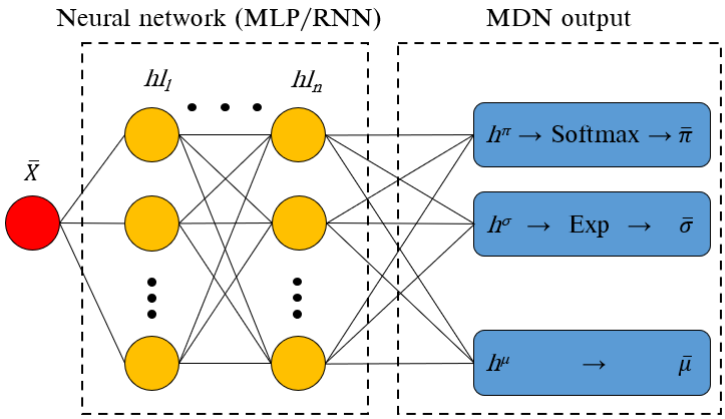


Figure 4: Simple MDN network.

Bishop [25] proposed the following restrictions for the mixing coefficients and variance components of the MDN output layer. Since the mixing coefficients contain the discrete probabilities of an output belonging to each K normal distribution for all observations, the mixing coefficients must satisfy the constraints listed in Equation 8.

$$\begin{aligned} &\sum_{k=1}^{K} \overline{\pi}_k^n = 1 \\ &0 \leq \overline{\pi}_k^n \leq 1 \end{aligned} \tag{8}$$

These constraints are met by using a Softmax function on the output of the hidden layer h^{π}. The mixing coefficient is calculated using Equation 9.

$$\overline{\pi}_k^n(\overline{x}^n) = \frac{exp(h_k^{\pi,n})}{\sum_{k=1}^{K} exp(h_k^{\pi,n})} \tag{9}$$

Similarly, a constraint is applied to the standard deviation values ensuring a positive value, this is achieved using an exponential function being applied to the standard deviation leg of the MDN output layer, namely h^{σ}. Equations 10 shows the imposed constraint for the standard deviation leg.

$$
\begin{aligned}
&(\overline{\sigma}_k^n(\overline{x}^n))^2 \geq 0 \\
&\sigma_k^n(\overline{x}^n) = exp(h_k^{\sigma,n})
\end{aligned}
\tag{10}
$$

The MDN weights and biases are optimized by minimizing the negative log-likelihood for all observations in a batch. This is shown in Equation 11.

$$
J_{NLL}(\overline{Y}, \overline{\pi}, \overline{\sigma}, \overline{\mu}) = -\sum_{n=1}^{N} ln \left\{ \sum_{k=1}^{K} \overline{\pi}_k(\overline{X}^n, \overline{\theta}) \cdot \mathbb{N}(\overline{Y}^n \mid \overline{\mu}_k(\bar{X}^n, \bar{\theta}), \overline{\sigma}_k^2(\overline{X}^n, \overline{\theta})) \right\} \tag{11}
$$

4. Data generation

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

4.1. CFD model setup

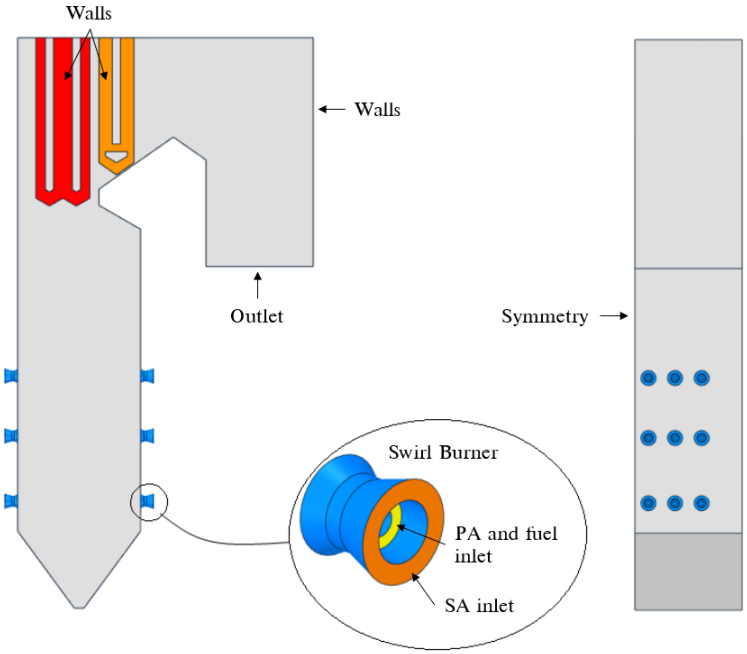
The current study makes use of an existing steady-state multiphase CFD model [8, 9] of the boiler under investigation using the commercial CFD software package ANSYS® Fluent 2019 R3 to resolve the fluid flow, heat transfer and combustion processes. This section provides a short overview of the model for the sake of completeness.

The modelled boiler consisted of a furnace with a depth of 13.77 (m), a width of 24.01 (m) and a maximum height of 64.00 (m). The CFD computational domain was created for half of the furnace width and has a symmetry boundary condition loacted at 12.01 (m). This was done to reduce the total cell count of the numerical mesh. Figure 5 highlights the computational domain and defines the important boundary conditions. The general conservation equations of mass, momentum, energy, and species were solved using a Eulerian approach. These are provided in Equation (12).

$$
\begin{aligned}
&\frac{\partial}{\partial x_i}(\rho \bar{u}_i) = S \\
&\frac{\partial}{\partial x_i}(\rho_{eff} u_i u_j) + \frac{\partial \overline{p}}{\partial x_j} = \\
&\frac{\partial}{\partial x_i}\left[\mu \left\{ \frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} - \frac{2}{3}\delta_{ij}\frac{\partial u_i}{\partial x_i} \right\}\right] + \frac{\partial}{\partial x_i}(-\rho \overline{u_i' u_j'}) + S_m \\
&\frac{\partial}{\partial x_i}(u_i[\rho E + p]) = \frac{\partial}{\partial x_j}\left[\lambda \frac{\partial T_g}{\partial x_j}\right] + S_h \\
&\frac{\partial}{\partial x_i}(\rho u_j Y_k) = -\frac{\partial}{\partial x_j}(\vec{J}_k) + \sum_r R_{j,r} + S_k
\end{aligned}
\tag{12}
$$

The solution of the Reynolds stress term found in the momentum equation ($-\rho \overline{u_i' u_j'}$) was approximated using the Boussineq equation [26]. In the present study the realizable k-ε turbulence model was utilized to address the turbulence closure problem and was selected for its applicability in modelling the effects of coal-fired swirl burners [27].

The P1 radiation model was used to resolve the radiative field in the domain. Particle transport was modelled using a multiphase Eulerian-Eulerian approach, further details of which are provided in the validation study of Rawlins et al [8, 28]. Modelling of combustion follows a four-step sequential process, beginning with the heating and evaporation



Boundary conditions		
Name	Type	Specify
Outlet	Pressure outlet	-
Symmetry	Symmetry	-
Walls (Grey, red and orange)	Convection [31]	Free stream, temperature and heat transfer coefficient
PA and fuel inlet	Velocity	Axial velocity
SA inlet	Velocity	Axial & tangential velocity

Figure 5: 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
Devolatilization		
Single rate kinetic model	$\frac{dm_{vol}}{dt} = R_{vol}(m_{0,vol} - m_{vol})$,	$A_{vol} = 2 \times 10^5 [s^{-1}]$,
	$R_{vol} = A_{vol} exp\left(\frac{E_{a,vol}}{RT_p}\right)$	$E_{a,vol} = 6.7{\times}10^7 [J/kmol]$ - [29]
Char oxidation		
Diffusion/kinetic - [30]	$\frac{dm_{char}}{dt} = -A_p p_{O_2} \frac{R_{diff} R_c}{R_{diff}+R_c}$,	$A_c = 0.0053 [kg/m^2 s Pa]$,
	$R_c = A_c exp\left(\frac{E_{a,c}}{RT_p}\right)$,	$E_{a,c} = 8.37{\times}10^7 [J/kmol]$ - [29]
	$R_{diff} = \frac{5\times10^{-12}}{d_p}\left(\frac{T_g+T_p}{2}\right)^{0.75}$	
Gaseous reactions of volatiles and CO		
Eddy dissipation model - [31]	$R_{k,r,P} = \vartheta_{k,r} M_{w,k} A B \rho \frac{\varepsilon}{k} min\left(\frac{\sum_p Y_p}{\sum_j \vartheta_{j,r} M_{w,j}}\right)$,	$A = 4.0,\ B = 0.5$
	$R_{k,r,R} = \vartheta_{k,r} M_{w,k} A \rho \frac{\varepsilon}{k} min\left(\frac{Y_R}{\vartheta_{R,r} M_{w,R}}\right)$	

of the moisture present in the fuel, followed by the devolatilization process where the volatiles are liberated from the solid particle, followed by the phenomena of char burnout, and finally the gas phase reactions would commence. The char oxidation 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 is 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 approximately 6 million cells. The

convergence criteria for the simulation model were 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. Convergence criteria was enforced for all the various simulation runs.

4.2. Simulated dataset

The inputs to the surrogate model include 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 (which is a function of the fuel composition). Thus the input field has a dimensionality of $d_{inputs} = 14$.

To obtain a representative set of results for training the surrogate model, a design of experiments (DOE) was conducted to generate a set of 180 simulation cases. The various model input ranges used in the DOE are provided in Table 2. The ranges where selected to cover a wide range of operational loads with maximum continuous ratings (MCR) between 100% and 30%. The DOE matrix was populated using the Latin Hypercube Sampling (LHS) method provided in the Python pyDOE library.

Table 2
Design of experiments input ranges for simulations

Input variables	Min	Max	Units
Total fuel flow rate for mills 1 to 6	39.5	120.2	kg/s
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^2K/W
Final SH fouling thermal resistance, R_{final}	0.01	0.017	m^2K/W
Input dependent variables	**Function of**		**Units**
Higher heating value	Fuel constituents Y_{H_2O} and Y_{ash}		J/kg
Excess air	Total fuel flow rate for mills		$\%$
Platen SH internal steam temperature	R_{platen}		K
Final SH internal steam temperature	R_{final}		K

Once the CFD simulations achieved convergence, the target data, comprising of the discretized heat loads to the furnace, platen SH and final SH, the exit flue-gas temperatures and flue-gas composition, was stored for each case. A total of 27 output target values ($d_{targets}$) are extracted from the results for each CFD simulation case. Due to the inherently unsteady nature of the CFD simulations even when convergence has been achieved, the output target values were extracted after every 50 iterations for an additional 2500 iterations once convergence was achieved. This results in each CFD simulation case having a solution data matrix size of ($\bar{Y} \in \mathbb{R}^{50\times27}$).

The target dataset was split with 80% for training purposes and the remainder for testing. To reduce the training time of the neural networks a min-max normalization was utilized to scale all the dataset entries to values between $0 \rightarrow 1$.

5. Surrogate model development

The present work makes use of two types of machine learning models, namely a standard artificial neural network (ANN/MLP) and a mixture density designated model connected to a standard ANN (MDN). 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.

5.1. Model configuration

The goal of the surrogate model is to predict the heat load distributions to the furnace water walls (EV), the platen superheater (SH2) tubes, and the final superheater (SH3) tubes, as well as the combustion characteristics and flue-gas temperatures at the exit, using the high level inputs for the ranges provided in Table 2.

The total CFD solution data matrix size will consist of $180 \times \overline{Y} = 243000$ data points. Using the total solution data matrix for training and testing would typically result in a slower convergence [22]. However, the use of subset data packages, referred to as batches or mini-batches, allows for a faster convergence. The batch size is a term used in machine learning that specifies a number of training/testing examples utilized in one iteration [10]. Thus, for a data batch size, m_b, the output tensor for the MLP model will be $\hat{Y} \in \mathbb{R}^{m_b \times 27}$ since $d_{targets} = 27$. However, the output data for the MDN model will consist of three parts, namely; the mixing coefficients tensor of shape $\overline{\pi} \in \mathbb{R}^{m_b \times K}$, the output standard deviation tensor of shape $\overline{\sigma} \in \mathbb{R}^{K \times m_b \times 27}$, and the predicted means of tensor shape $\overline{\mu} \in \mathbb{R}^{K \times m_b \times 27}$, where K is the number of distributions. The input data fed into both the MLP and MDN will have shape $\overline{X} \in \mathbb{R}^{m_b \times d_{inputs}}$. The input features will be varied based on the DOE, to account for burner mill biassing, fuel quality and the superheaters fouling thickness.

5.2. Hyper-parameter tuning & final model selection

Table 3 highlights the hyper-parameter search spaces for both the MLP and MDN model. The MDN has an added parameter, namely the number of additional distributions that the MDN would need to fit to the output data. The hyper-parameter tuning of both the MLP and MDN models made use of 1000 epochs. This was deemed adequate to train/test the models for the various hyper-parameter search spaces and was utilized to reduce the computational effort/time.

Table 3
Hyper-parameter search space for fully connected MLP and MDN models

Parameter	MLP search space	MDN search space
Number of distributions	-	1,2,3,4
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-6	1e-3, 1e-4, 1e-5, 1e-6
Mini batch sizes	32, 64, 128, 256	32, 64, 128, 256

The hyper-parameter search was conducted in a sequential manner with the MAEs and root mean square errors (RMSE) being taken as important performance indicators for each trained case. Firstly, the hidden layer architecture of both models (MLP & MDN) were varied by considering the number of hidden layers and neurons per layer. Figure 6 (a) illustrates the MLP model performance for the various hidden layer architectures. The neuron capacity reaches a minimum MAE at 80 neurons per layer for a 4-layer architecture. Further increasing the neuron capacity results in an increase in the MAE, possibly indicating an over fitting of data. Secondly, the learning rates were varied for the best performing architecture of the first hyper-parameter tuning step. Considering Figure 6 (b), a decrease in the learning rate shows an improvement in the MAE, with a learning rate of 1×10^{-5} being the best for the current epoch size. For comparative purposes, learning rates of 1×10^{-5} and 1×10^{-6} were used in the final step, where the mini-batch sizes are varied. Figure 6 (c) highlights the comparison for a fixed epoch, with a mini-batch size of 32 and a corresponding learning rate of 1×10^{-5} showing the best MAE improvement.

The MDN hyper-parameter tuning was conducted in a similar manner except that an additional step was required. This was to consider the number of distributions the MDN would use to capture the probabilistic characteristics. Figure 7 shows the MAE for the various distributions, with a distribution of one representing the best MLP model. An increase in the number of distributions tends to improve the MAE, however it is evident that a threshold of 3 distributions results in the best improvement.

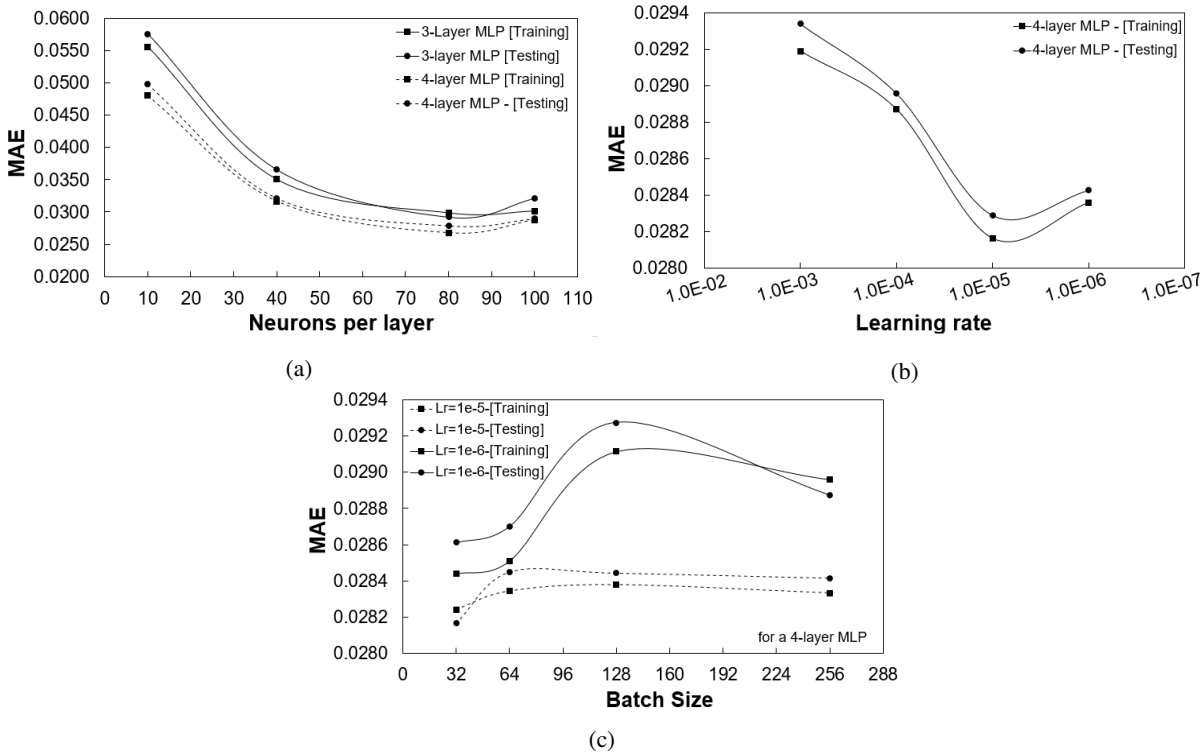


Figure 6: MLP model performance for; (a) hidden layer architecture, (b) varying learning rates, and (c) varying mini-batch sizes.

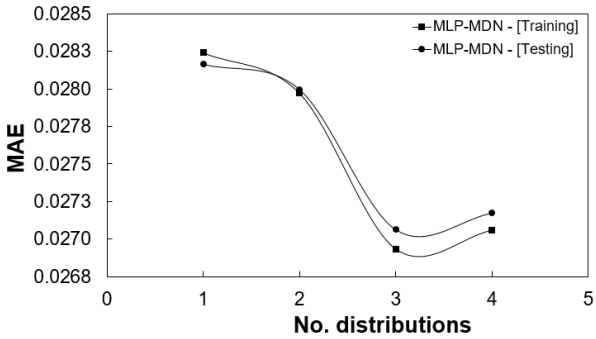
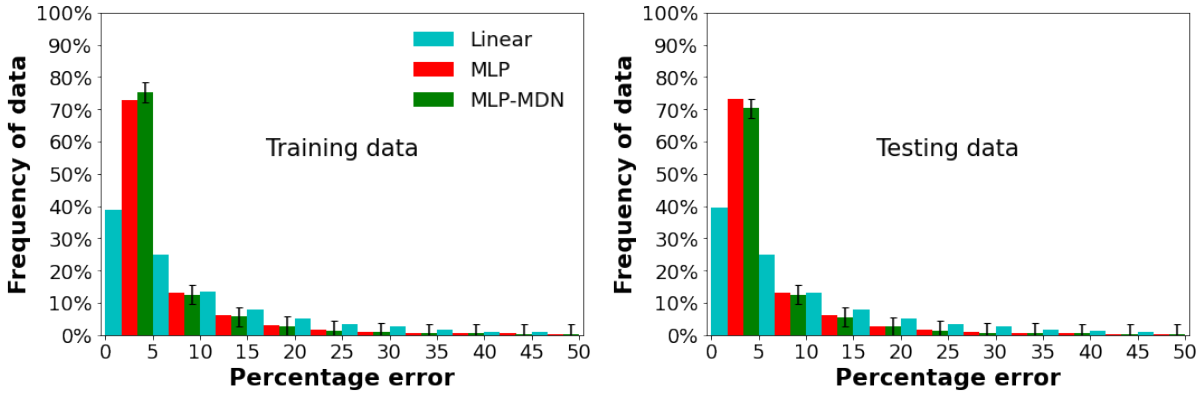


Figure 7: MDN model performance for number of distributions.

The results of hyper-parameter search are given in Table 4. Unlike the MLP, which can only produce a single set of predictions for an input, the MDN is able to produce a distribution of predictions based on the most probable mixing coefficient. This highlights one of the main advantages of the using the MDN model, which is its ability to learn the uncertainty that comes from using a probabilistic model that enables the modelling approach to take into account the effect information density in the training data. This is shown in the error distributions graphs of Figures 8 (a) and (b). A simple linear regression model was used as a base model for comparative purposes. For the MDN model it is seen that approximately 80-90% of the training data has mean absolute percentage errors (MAPEs) below 10%, with the MLP model showing a similar trend. Both models show considerable improvement in comparison to the linear model.

Based on the above analysis, the MDN model was selected as the best model to be used for the surrogate model implementation. The benefit of using the MDN surrogate model is the ability to provide not only the expected mean (most-probable) value of each output parameter, but also the associated standard deviation. Having the mean value of



(a) (b)

Figure 8: Overall error distributions for the selected MLP and MDN model; (a) training data and (b) testing data.

Table 4
Hyper-parameter search results

Parameter	MLP	MDN
Number of distributions	-	3
Number of layers	4	4
Number of neurons per layer	80	80
Learning rate	1e-5	1e-5
Mini batch size	32	32
Errors		
RMSE	0.0213	0.0211
MAE	0.0282	0.0263

each of the furnace and heat exchanger heat loads together with its uncertainty in the form of a standard deviation, allows the uncertainty to be propagated through the entire integrated model. It is therefore possible to obtain the resultant mean values of the overall plant performance parameters together with their uncertainties emanating from the surrogate model prediction.

6. Results and discussion

This section provides the results of the validation study of the integrated model, as well as an application case study that investigates the effects of fuel quality on the integrated boiler performance.

6.1. Validation

The results predicted using the surrogate model were validated for the 100%, 80% and 60% MCR load cases. Measured data was made available for these load cases, which allow the mean and standard deviation to be determined. The mean and standard deviation of each output can then be compared with the predicted mean and standard deviation obtained with the integrated model.

In the current work the integrated model is applied in conjunction with the Monte Carlo method [32]. Firstly, random sampling of the input variables to the integrated model are performed based on the mean and standard deviation predicted by the surrogate model. Following this, simulations are run for each sample and the outputs are evaluated to determine the predicted mean and standard deviation of each overall plant parameter.

Table 5
MDN input vectors ($\overline{X}$) for the validation and varied fuel case studies

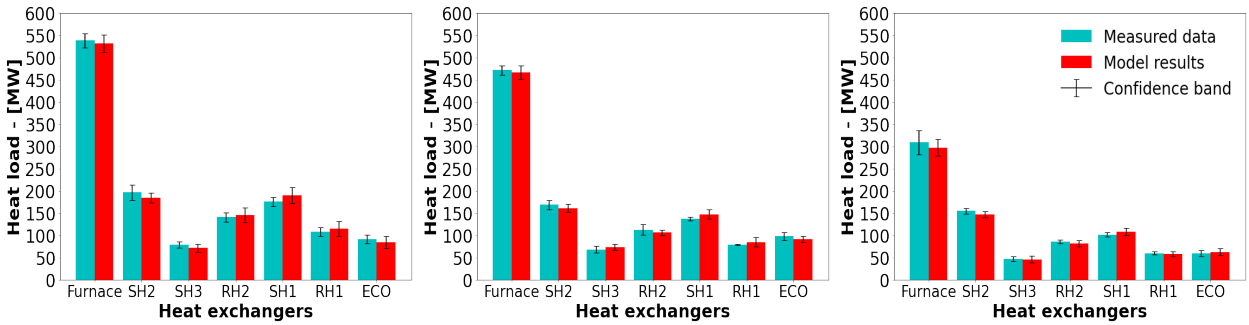
	Validation load cases			Varied fuel load cases	
Inputs [**mean** - ($\overline{\sigma}$)]	100%	80%	60%	High-ash	High-moisture
Excess air ratio - [-]	1.155	1.209	1.263	1.155	1.155
Y_{ASH} - $[kg/kg]$	0.409	0.409	0.409	0.5	0.409
Y_{H_2O} - $[kg/kg]$	0.055	0.055	0.055	0.055	0.146
HHV $= f(Y_{ASH}, Y_{H_2O})$ - $[MJ/kg]$	15.07	15.07	15.07	12.68	12.68
Platen SH fouling - $[K/W]$	0.012	0.012	0.012	0.012	0.012
Final SH fouling - $[K/W]$	0.0067	0.0067	0.0067	0.0067	0.0067
Total fuel flowrate - $[kg/s]$	108.6	90.72	60.12	108.6	108.6
Platen SH mean steam temp. - $[K]$	698	697	781	698	698
Final SH mean steam temp. - $[K]$	787	793	781	787	787

Table 6
Mill fuel flowrate biasing for the validation case study.

MCR load	Mill 1	Mill 2	Mill 3	Mill 4	Mill 5	Mill 6	Unit
100%	19.14	20.22	19.92	19.98	19.41	0.00	kg/s
80%	18.23	19.72	19.02	17.82	18.02	0.00	kg/s
60%	15.46	16.13	0.00	15.11	15.32	0.00	kg/s

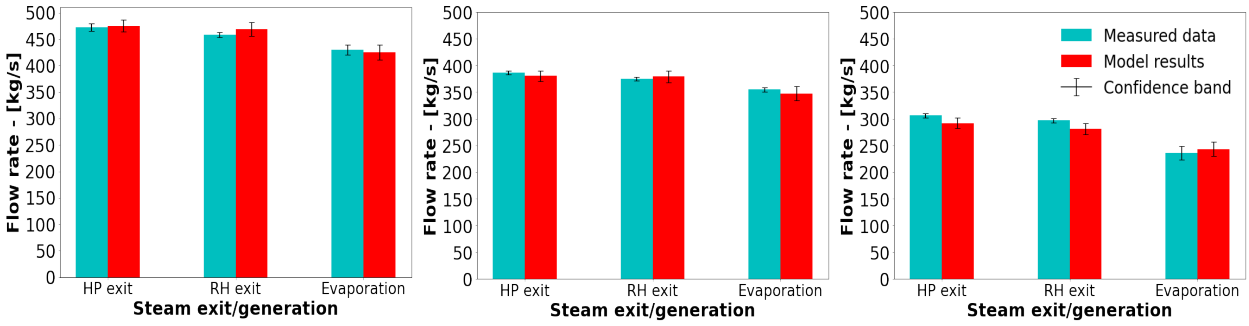
Fortunately, Flownex SE® 2021, provides a built-in Monte Carlo facility. It requires as input the mean and standard deviation values for each model input [19]. The Monte Carlo model input used in this study are the predicted mean and standard deviation results obtained with the surrogate model, which utilized the MDN input vectors ($\overline{X}$) provided in Table 5. These inputs were obtained from the works of Laubscher and Rousseau [33]. For this study $N = 1000$ simulations were performed for each validation case with the mean and standard deviation being used as measures to determine convergence. Table 6 provides the total mill fuel flow rate for the validation cases. Biasing of the mill fuel flow rates was introduced to provide a realistic operation of the combustion system.

Figures 9 - 11 provide comparative results of the measured data and integrated model responses for all three load cases. It can be seen for a range of MCR loads that the integrated model is able to capture the heat loads to the various heat exchangers (Figures 9 (a-c)) with sufficient accuracy. In general, the model under-predicts the mean of the heat loads in the furnace and radiative pass, with slight over-predictions seen in the convective pass. However, the associated uncertainty of the integrated model overlaps with the measured data uncertainty with minimal outliers, therefore, inferring that the model predictions are sufficiently accurate.



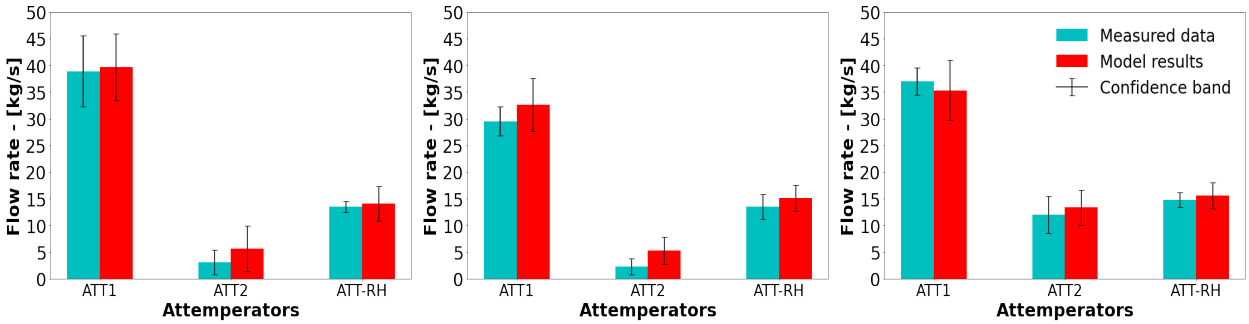
(a) 100% MCR (b) 80% MCR (c) 60% MCR

Figure 9: Load validation result comparison of the measured data and MDN model results for the various heat exchangers



(a) 100% MCR (b) 80% MCR (c) 60% MCR

Figure 10: Load validation result comparison of the measured data and MDN model results for the steam exit and steam generation flowrates



(a) 100% MCR (b) 80% MCR (c) 60% MCR

Figure 11: Load validation result comparison of the measured data and MDN model results for the required attemperator flowrates to maintain operational integrity

Considering the steam generation and exit flowrates of Figure 10 (a-c), the integrated model can sufficiently capture the hydraulic response of the case study boiler for a wide range of loads. More uncertainty is associated with the integrated model, but this is deemed acceptable since the response overlaps with the measured data in all but one case. The attemperators of a utility scale boiler are an integral part of the control systems put in place to ensure safe operating conditions and minimal wear on the downstream components, namely the steam HP and LP steam turbines [34]. Figure 11 (a-c) highlight the attemperator flowrates across a wide range of loads. The integrated model is shown to adequately resolve the flowrates with a slightly higher mean being predicted for the 80% load case.

It has been shown that the integrated model is able to resolve the thermal response based on the predictions of the MDN model for a wide range of loads with sufficient accuracy. The heat loads are within a 5% tolerance of the measured mean values, while the steam generation rates exhibit a maximum difference of 7%, which is seen in the 60% MCR load case (refer to Figure 10 (c)). The attemperator flow rate results show the largest uncertainty/confidence band of all the results (Figure 11), however these overlap with the measured values in all but one case.

6.2. Application

This section demonstrates the application of the validated integrated model by investigating the impact of fuel quality on the overall plant performance. The fuel quality parameters that are varied are the energy content of the fuel (HHV), together with the ash (Y_{ASH}) and moisture (Y_{H_2O}) contents. Two case studies are considered to determine the impact of the parameters on the boiler efficiency, heat pick-up and thermal response. The input vectors for the two fuel cases are provided in Table 5. The inputs are the same as those of the 100% load case except for the ash, moisture, and higher heating value of the respective fuels.

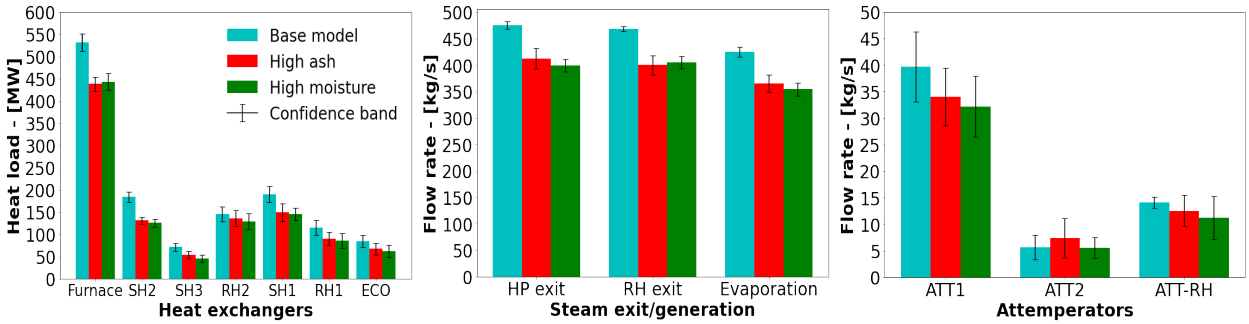
Table 7
Model results for poor quality fuel characteristics

Variable (**mean** (*min*,*max*))	Base	High-ash	High-moisture
Mean boiler efficiency [%]	**88.2** (*75.6,91.2*)	**77.6** (*73.4,80.9*)	**75.9** (*69.8,78.3*)
Convective pass inlet flue-gas temp. [K]	**1352** (*1339,1370*)	**1368** (*1354,1378*)	**1311** (*1295,1327*)
Radiative heat transfer percentage (Convective pass)			
RH2	**52.3** (*50.1,54.0*	**56.2** (*52.9,58.5*	**50.3** (*47.8,53.9*
SH1	**46.3** (*43.0,49.2*)	**49.8** (*44.8,52.3*	**47.2** - (*43.2,50.8*
RH1	**21.3** (*16.1,24.6*)	**26.7** (*20.6,30.1*	**18.2** (*13.7,23.5*
ECO	**6.7** (*5.1,9.8*)	**8.3** (*6.3,12.6*	**5.4** - (*2.8,9.7*

As with the validation case studies of Section 6.1, the Monte Carlo method was employed to ascertain the mean and standard deviation values predicted via the integrated model. The base case in Figure 12 (a-c) refers to 100% load case of Figures 9-11. Figure 12 (a) highlights the effects that poor quality fuel has on the heat load to the various heat exchangers, with a 20% drop being observed on average. This can reasonably be expected since the energy content of the fuel is lower than that of the base case. To maintain 100% load capabilities using the poor quality fuel would require an increase in the fuel flowrate, which would provide more energy content. However, the mill operational capacities are limited. The operational protocol for the 100% load case makes use of 5 mills operating 30 of the 36 burners at full capacity, with a standby mill and burner arrangement placed in reserve to help mitigate operational risks, such as maintenance schedules. Using the the standby mill would allow the fuel flowrate to be increased by a maximum of 17%. However, this would in turn decrease the operational integrity and increase the associated risks. This situation would therefore imply an unwanted load loss. When burning the high-ash content fuel the likelihood of ash deposition/fouling of the heat exchanger components also increases.

Figure 12 (b-c) similarly illustrate a decrease in the water/steam flowrates for the steam exit and attemperator conditions. The generation rates exhibit similar characteristics to the the 80% load case of Figure 10 (b), illustrating a significant decrease in the operational capabilities when using a poor quality fuel. The predicted flue-gas temperatures entering the the convective pass are shown in Table 7. It is shown that for an increase in the ash content of the fuel, the flue-gas inlet temperature increases, which results in a higher radiative heat transfer percentage for the convective pass components. It is also interesting to note that a higher ash content results in the largest uncertainties in the prediction of the HP and RH exit steam and steam generation flow rates. The presence of more moisture in the fuel results in a lower inlet temperature, primarily due to the fact that the extra moisture requires a higher heat rate of evaporation, leading to lower radiative heat transfer percentages.

The boiler efficiency is the measure used to convey how well the combustion heat is transferred to the working fluid, and is defined as the ratio of total amount of heat absorbed by the heat exchangers (i.e. sum of the furnace, radiative and convective pass heat loads) and the combustion energy released ($m_{fuel}HHV$). A significant decrease in the boiler efficiency is highlighted in Table 7 for both fuel cases with the high ash fuel showing a slightly better result. The higher efficiency of the high-ash fuel is attributed to the fact that in general the heat loads to various heat exchangers are higher than that of the high-moisture fuel, which can be seen in Figure 12 (a).



(a) (b) (c)

Figure 12: Integrated base case, high-ash and high-moisture case study comparison for; (a) the various heat exchangers, (b) the steam exit and generation flowrates, and (c) the required attemperator flowrates to maintain operational integrity

7. Conclusion

The present work establishes the basis for utilizing an integrated data-driven surrogate and 1D process model to investigate the overall operational response for a 620 MW_e utility scale boiler.

The data-driven surrogate modelling approach using CFD simulations can predict the various parameters needed to capture the combustion and heat transfer characteristics. These include the heat loads to the furnace evaporator walls (EV) and radiative superheaters (SH2 and SH3), as well as the flue-gas characteristics entering the convective pass, which include the temperatures, species mass fractions and radiation intensity. Training and testing datasets were generated using the CFD model of the 620 MWe. A total of 180 different CFD simulation cases were performed, where the input variables of Table 2 were varied.

It was found that the best surrogate model architecture was the use of a MDN configuration, which also allows prediction of the associated uncertainties. The configuration comprised of a MDN network consisting of 4 hidden layers each consisting of 80 neurons and 3 normal output distributions, while using a learning rate of 1×10^{-5} and batch size of 32. The MDN network illustrated that approximately 80-85% of the model predictions had MAPEs below 10%.

A validation case study was performed using the integrated model for a wide range of operational loads. The model results were compared to the measured thermal response for all load cases. The resolution of the heat loads, steam flowrates and attemperation flowrates resulted in a maximum difference of 7%. The uncertainties predicted by the surrogate model were propagated through the integrated model using the Monte Carlo technique, adding valuable insight into the operational limits of the power plant and the uncertainties associated with it.

The application case studies considered the impact of poor quality fuels at 100% MCR operation. The results highlight the drop in performance due to the fuel quality resulting in a substantial decrease in boiler efficiency. Using a high-ash fuel can result in a higher exit flue-gas temperatures resulting in a marginal increase in performance of the convective pass heat exchangers when compared to the high moisture fuel. However, it results in greater uncertainty in the predicted steam exit and generation flow rates and increases the likelihood of ash deposition occurring.

The current work has shown that it is possible to adequately resolve the thermofluid response of a utility scale boiler using a data-driven surrogate model trained on CFD simulation results. The integration of the surrogate and 1-D process model is paramount in modelling the thermal response and provides valuable operational insight for various load cases, including the uncertainty in the predicted results. The methodology used in the present work can be applied by engineers and other researchers to investigate various effects on the whole boiler operation.

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.

References

[1] P. Rousseau, R. Laubscher, Analysis of the impact of coal quality on the heat transfer distribution in a high-ash pulverized coal boiler using co-simulation, Energy 198 (2020) 117343. doi:10.1016/j.energy.2020.117343.
URL https://doi.org/10.1016/j.energy.2020.117343

[2] S. M. Safdarnejad, J. F. Tuttle, K. M. Powell, Dynamic modeling and optimization of a coal-fired utility boiler to forecast and minimize NOx and CO emissions simultaneously, Computers and Chemical Engineering 124 (2019) 62–79. doi:10.1016/j.compchemeng.2019.02.001.
URL https://doi.org/10.1016/j.compchemeng.2019.02.001

[3] N. Modliński, K. Szczepanek, D. Nabagło, P. Madejski, Z. Modliński, Mathematical procedure for predicting tube metal temperature in the second stage reheater of the operating flexibly steam boiler, Applied Thermal Engineering 146 (October 2018) (2019) 854–865. doi:10.1016/j.applthermaleng.2018.10.063.

[4] X. Liu, J. Zhang, H. Tan, Q. Mo, X. Wang, Y. Wang, Numerical and experimental study on co-firing of low volatile coal in a 330 MW tangentially fired boiler, Journal of the Energy Institute 96 (x) (2021) 242–250. doi:10.1016/j.joei.2021.03.015.

[5] A. Dugum, K. Hanjalić, Numerical simulation of coal-air mixture flow in a real double-swirl burner and implications on combustion anomalies in a utility boiler, Energy 170 (2019) 942–953. doi:10.1016/j.energy.2018.12.121.

[6] J. Gu, Q. Liu, W. Zhong, A. Yu, Study on scale-up characteristics of oxy-fuel combustion in circulating fluidized bed boiler by 3D CFD simulation, Advanced Powder Technology 31 (5) (2020) 2136–2151. doi:10.1016/j.apt.2020.03.007.
URL https://doi.org/10.1016/j.apt.2020.03.007

[7] R. Laubscher, P. Rousseau, Numerical investigation on the impact of variable particle radiation properties on the heat transfer in high ash pulverized coal boiler through co-simulation, Energy 195 (2020) 117006. doi:10.1016/j.energy.2020.117006.
URL https://doi.org/10.1016/j.energy.2020.117006

[8] B. T. Rawlins, R. Laubscher, P. Rousseau, Validation of a thermal non-equilibrium Eulerian-Eulerian multiphase model of a 620 MWe pulverized fuel power boiler., in: S. Skatulla (Ed.), 12th South African Conference on Computational and Applied Mechanics (SACAM2020), MATEC Web of Conferences, Cape Town, 2021. doi:https://doi.org/10.1051/matecconf/202134700004.

[9] B. T. Rawlins, R. Laubscher, P. G. Rousseau, A fast thermal non-equilibrium Eulerian-Eulerian numerical simulation methodology of a pulverized fuel combustor, in: V. Scherer, N. Fricker, A. Reis (Eds.), 13th European Conference on Industrial Furnaces and Boilers, 2022.

[10] M. Kochenderfer, T. Wheeler, Algorithms for optimization, 1st Edition, The MIT Press, Cambridge, Massachusetts, 2019.

[11] Y. Fei, S. Black, J. Szuhánszki, L. Ma, D. B. Ingham, P. J. Stanger, M. Pourkashanian, Evaluation of the potential of retrofitting a coal power plant to oxy-firing using CFD and process co-simulation, Fuel Processing Technology 131 (2015) 45–58. doi:10.1016/j.fuproc.2014.10.042.

[12] M. Si, K. Du, Development of a predictive emissions model using a gradient boosting machine learning method, Environmental Technology and Innovation 20 (x). doi:10.1016/j.eti.2020.101028.

[13] P. Qin, L. Zhao, Z. Liu, State of health prediction for lithium-ion battery using a gradient boosting-based data-driven method, Journal of Energy Storage 47 (November 2021) (2022) 103644. doi:10.1016/j.est.2021.103644.
URL https://doi.org/10.1016/j.est.2021.103644

[14] Y. Lv, F. Hong, T. Yang, F. Fang, J. Liu, A dynamic model for the bed temperature prediction of circulating fluidized bed boilers based on least squares support vector machine with real operational data, Energy 124 (2017) 284–294. doi:10.1016/j.energy.2017.02.031.
URL http://dx.doi.org/10.1016/j.energy.2017.02.031

[15] S. Singh, H. Abbassi, 1D/3D transient HVAC thermal modeling of an off-highway machinery cabin using CFD-ANN hybrid method, Applied Thermal Engineering 135 (January) (2018) 406–417. doi:10.1016/j.applthermaleng.2018.02.054.

[16] P. Madejski, Thermal Power Plants - New Trends and Recent Developments, IntechOpen, London, United Kingdom, 2018. doi:10.5772/intechopen.69696.
URL https://doi.org/10.5772/intechopen.69696

[17] D. Sarkar, THERMAL POWER PLANT Design and Operation, Elsevier Inc., Amsterdam, 2015.

[18] C. Lu, J. Wang, R. Yan, Multi-objective optimization of combined cooling, heating and power system considering the collaboration of thermal energy storage with load uncertainties, Journal of Energy Storage 40 (May) (2021) 102819. doi:10.1016/j.est.2021.102819.
URL https://doi.org/10.1016/j.est.2021.102819

[19] MTech, Flownex Theory Guide, MTech inc., 2021.

[20] D. Wen, Y. Pan, X. Chen, M. Aziz, Q. Zhou, N. Li, 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 (2022) 101137. doi:10.1016/j.tsep.2021.101137.

[21] C. Rasmussen, C. Williams, Gaussian Processes for Machine Learning, 1st Edition, The MIT Press, Cambridge, Massachusetts, 2006.

[22] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, first edit Edition, MIT Press, Chennai, 2017.

[23] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors, Nature 323 (6088) (1986) 533–536. doi:10.1038/323533a0.

[24] E. Oko, M. Wang, J. Zhang, Neural network approach for predicting drum pressure and level in coal-fired subcritical power plant, Fuel 151 (2015) 139–145. doi:10.1016/j.fuel.2015.01.091.

URL http://dx.doi.org/10.1016/j.fuel.2015.01.091
[25] C. Bishop, Mixture Density Networks, Tech. rep., Aston University (1994). doi:10.1007/978-3-322-81570-5_8.
[26] H. Versteeg, W. Malalasekera, Introduction to Computational Fluid Dynamics, The finite volume method, 2nd Edition, Pearson Prentice Hall, 2007. doi:10.1002/9781119369189.
[27] N. Modlinski, Computational modeling of a utility boiler tangentially-fired furnace retrofitted with swirl burners, Fuel Processing Technology 91 (11) (2010) 1601–1608. doi:10.1016/j.fuproc.2010.06.008.
[28] B. Rawlins, R. Laubscher, P. Rousseau, Application of computational fluid dynamics and process modelling to investigate low-load operation of a subcritical utility-scale boiler., ASME Journal of Thermal Science and Engineering Applications (2022) 1–29arXiv:https://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4054010/6856445/tsea-21-1636.pdf, doi:10.1115/1.4054010.
URL https://doi.org/10.1115/1.4054010
[29] C. Sheng, B. Moghtaderi, R. Gupta, T. F. Wall, A computational fluid dynamics based study of the combustion characteristics of coal blends in pulverised coal-fired furnace, Fuel 83 (11-12) (2004) 1543–1552. doi:10.1016/j.fuel.2004.02.011.
[30] M. Baum, P. Street, Predicting the combustion behaviour of coal particles, Combustion science and Technolodgy 3 (43) (1971) 231.
[31] ANSYS, ANSYS Fluent Theory Guide, 20th Edition, Ansys Inc., 2021.
[32] N. T. Thomopoulos, Essentials of Monte Carlo simulation: Statistical methods for building simulation models, Vol. 9781461460220, 2013. doi:10.1007/978-1-4614-6022-0.
[33] R. Laubscher, P. Rousseau, CFD study of pulverized coal-fired boiler evaporator and radiant superheaters at varying loads, Applied Thermal Engineering 160. doi:10.1016/j.applthermaleng.2019.114057.
[34] S. Kakac, Boilers, Evaporators and Condensers, Wiley Interscience, 1991.