Highlights

An integrated data-driven surrogate model and thermofluid network-based model of a 620 MW_e utility-scale boiler

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.

- Off-design fuel quality investigation.

An integrated data-driven surrogate model and thermofluid network-based model of a 620 MW_e utility-scale boiler

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

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 perceptron (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 dynamic (CFD) models are typically utilised in modelling the fire-side interactions of pulverised coal fired boilers [], since the modelling approach can adequately resolve the complex fluid flow and heat transfer phenomena, including the mechanical and thermal interactions 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], perform 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, utilising a 1-D process model of the water-side and CFD model of the fire-side, to investigate the effects high ash fuel has on the heat absorption of the 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

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

Nomenclature

Variables

$\bar{h}_l$	Neural network signal for layer l	$-$
$\bar{P}_r$	Surrogate model prediction vector	$-$
$\bar{X}$	High-level input vector	$-$
$\bar{Y}$	CFD solution output vector	$-$
$\bar{z}_l$	Summed signal to network neural for layer l	$-$
A	Area	m^2
A_c	Char pre-exponential factor	$kg/m^2 sPa$
A_{vol}	Volatile pre-exponential factor	s^{-1}
d_p	Particle diameter	m
E	Total energy	J/kg
$E_{a,c}$	Char activation energy	J/mol
$E_{a,vol}$	Volatile matter activation energy	J/mol
h_M	Mixture enthalpy	J/kg
J_k	Diffusion flux of species k	kg/m^3
m_b	Batch size	$-$
m_c	Char mass	kg
m_c	Volatile mass	kg
P	Perimeter	m
p	Pressure	Pa
p_{O_2}	Oxygen partial pressure	Pa
R_c	Chemical reaction rate	$kg/m^2 sPa$
R_{diff}	Diffusion reaction rate coefficient	$kg/m^2 sPa$
R_{vol}	Volatile reaction rate	s^{-1}
S	Mass source term	kg/m^3
S_h	Energy source term	W/m^3
S_k	Species source term	kg/m^3
S_m	Momentum source term	N/m^3
T_g	Gas temperature	K
T_p	Particle temperature	K
u	Directional velocity	m/s
x	Quality	$-$
Y_k	Mass fraction of species k	kg/kg

Greek variables

α_M	Mixture fraction	$-$
$\bar{\mu}_k$	MDN predicted mean for distribution k	$-$
$\bar{\pi}_k$	MDN mixing coefficient for distribution k	$-$
$\bar{\theta}_k$	MDN standard deviation for distribution k	$-$
η	Learning rate	$-$
λ	Thermal conductivity	W/mK
μ	Viscosity	$Pa.s$
ϕ_{mp}	Pseudo particle mass fraction	kg/kg_{fuel}
ρ	Gas density	kg/m^3
ρ_{eff}	Effective density	kg/m^3
ρ_g	Gas phase density	kg/m^3
ρ_l	Liquid phase density	kg/m^3
ρ_M	Mixture density	kg/m^3

Abbreviations

ANN	Artificial Neural Network
ATT-RH	Reheater Attemperator
$ATT1$	Attemperator 2
$ATT2$	Attemperator 1
CFD	Computational Fluid Dynamics
DOE	Design of Experiments
EC	Economiser
EDM	Eddy Dissipation Model
EV	Evaporator
HHV	Higher Heating Value
J_{MSE}	Mean Square Error cost function
J_{NLL}	Negative Log Likelihood cost function
MAE	Mean Absolute Error
MDN	Mixture Density Network
MLP	Multi-layer Perceptron
MSE	Mean Square Error
$RH1$	Primary Reheater
$RH2$	Secondary Reheater
$RMSE$	Root Mean Square
RNN	Recurrent Neural Network
ROM	Reduced Order Model
$SH1$	Primary Superheater
$SH2$	Platen Superheater
$SH3$	Final Superheater

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, a wider range of appropriate operating conditions can be simulated. The works of Shi et al. [19] developed an optimisation method, which employed an ANN to predict the thermal efficiency and emission properties, for a 660 $[MW_e]$ ultra-supercritical coal fired power plant. Training and testing data was obtained from historical operating data, as well as using the solution data of a 100 unique CFD simulation cases covering a wide range of operating cases. The additional CFD simulation data strengthened the generalization of the ANN model, mainly because the CFD simulation provided different kinds of operating condition making the model adapt to different working conditions. Good agreement between the predictions and measured data was realised with a maximum mean error of 0.04%. Researchers Dal Secco et al. [20], conducted a comparable study to identify boiler operational configurations to minimise the NO_x emissions in a 600 $[MW]$ tangentially fired pulverised coal boiler. The study made use of a genetic algorithm to correlate the operating parameters to the output data produced from a large CFD simulation dataset. The genetic algorithm helped identify the optimal firing configurations from the CFD dataset that establish a good compromise between low NO_x emissions and the reduction of fire-side corrosion.

In the present work, a modelling methodology is employed that incorporates the use of a CFD generated solution database in the training and testing of data-driven surrogate model that captures the fire-side of a 620 $[MW_e]$ utility scale boiler. Subsequently, the data-driven surrogate model is integrated with a 1-D process model of the water-/steam-side. This integrated model is proposed to investigate the thermal response of the 620 $[MW_e]$ utility scale boiler under investigation. An existing validated steady-state multiphase CFD model [8, 9] was used to generate a CFD 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 perceptron (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 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. Materials and methods

The following section provides an overview of the case study boiler, the CFD modelling methodology utilised for the generation of testing and training data, the 1-D process model, and a description of the integrated model.

2.1. Case study boiler

The overall layout of the case study boiler is shown schematically in Figure 1 [1]. The boiler is defined as a two-pass sub-critical boiler, which is designed to deliver steam at a pressure of 16 [MPa] and an exit temperature of 535 [$°C$] at full load.
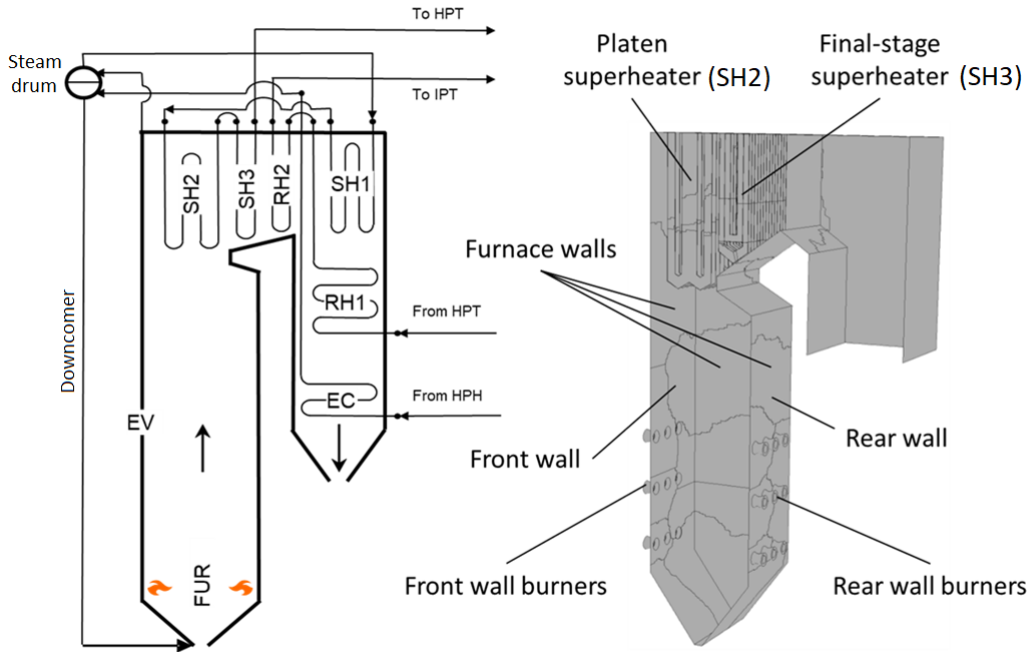

Figure 1: Case study boiler layout [1].

The boiler forms part of a 620 [MW_e] 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 economiser (EC). The furnace height, width and depth are 64.0 [m], 24.0 [m], and 13.7 [m] respectively. The boiler furnace is fed by six mills, each supplying a pulverised fuel and primary air (PA) mixture to a burner row consisting of six opposing wall-mounted swirl burners. The fuel and PA mixture is injected through the inner annulus of the burner while the SA air is fed through the outer annulus as indicated in Figure 2.

2.2. CFD model

The current study makes use of an existing steady-state multiphase CFD model [9]. The CFD model utilises an Eulerian-Eulerian approach to capture the solid and gaseous phase interactions. The model was implemented in the commercial CFD software package ANSYS Fluent® v19.5 to resolve the fluid flow, heat transfer and combustion processes. The primary features of developed multiphase CFD model is the solution of two energy equations, one for the gas phase and another for the solid phase. This approach allows for the adequate resolution of the temperature-dependent combustion processes and accurate prediction of the particulate radiative effects. The model has been validated and verified for a 2.165 [MW_{th}] lab scale pulverised fuel burner by Rawlins et al. [9]. Furthermore, the CFD model was subsequently validated for the full-scale utility boiler of the current study operating for a wide range of loads [8]. Thus, this section provides a summary of the CFD models features and governing equations.

The CFD computational domain was created for half of the furnace width with a symmetry boundary condition located at 12.01 [m]. This was done to reduce the total cell count of the numerical mesh. Figure 2 highlights the computational domain and boiler geometry.
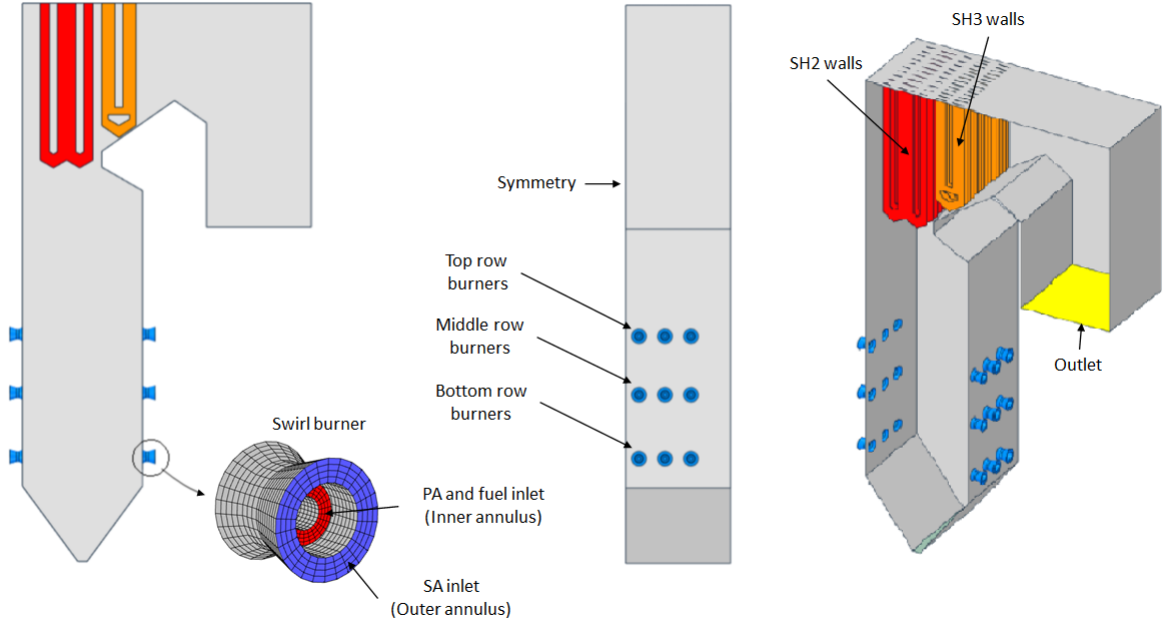
Figure 2: Overview of CFD model geometry and the computational domain.

The general conservation equations of mass, momentum, energy, and species were solved using an Eulerian approach. These are provided in Table 1.

Table 1
Summary of CFD models governing equations

Conservation entity	Conservation equation
Mass	$\frac{\partial}{\partial x_i}(\rho \bar{u}_i) = S$
Momentum	$\frac{\partial}{\partial x_i}(\rho_{eff} u_i u_j) + \frac{\partial \bar{p}}{\partial x_j} = \frac{\partial}{\partial x_i}\left[\mu\left\{\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} - \frac{2}{3}\delta_{ij}\frac{\partial u_i}{\partial x_i}\right\}\right] + \frac{\partial}{\partial x_i}(-\overline{\rho u_i' u_j'}) + S_m$
Energy	$\frac{\partial}{\partial x_i}(u_i[\rho E + p]) = \frac{\partial}{\partial x_j}\left[\lambda \frac{\partial T_g}{\partial x_j}\right] + S_h$
Species	$\frac{\partial}{\partial x_i}(\rho u_j Y_k) = -\frac{\partial}{\partial x_j}(\vec{J}_k) + \sum_r R_{j,r} + S_k$

To account for the inertial effects the particles exert on the gas phase, an effective density (ρ_{eff}) was defined as follows:

$$\rho_{eff} = \frac{\rho \rho_p (\phi_{mp+1})}{\rho \phi_{mp} + \rho_p} \tag{1}$$

where ρ, ρ_p and ϕ_{mp} are the gas density, particulate density and the mass fraction of the pseudo particles present in a cell [kg/kg_{fuel}].

The solution of the Reynolds stress term found in the momentum equation ($-\overline{\rho u'_i u'_j}$) was approximated using the Boussineq equation [21]. 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 [22]. The particle transport was modelled using general scalar transport equations of the pseudo-particle mass concentrations. The particulate scalar fields define the injected fuel based on the proximate fuel composition, namely the moisture, volatile matter, fixed carbon, and ash content.

The P-1 radiation model was used to resolve the radiative field in the domain. User-defined sources and functions were used to configure the Eulerian representation of the particle laden flow P-1 transport equation. In addition, the particles emissivity and scattering factor were modelled using variable correlations as applied by Lockwood et al. [23] and Yin [24], respectively. Modelling of combustion follows a four-step sequential process, beginning with the heating and evaporation of the moisture present in the fuel, followed by the devolatilisation 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 are provided in Table 2.

Table 2
Summary of combustion models and constants used in the CFD model

Model	Equation/s	Constant/s
Devolatilization		
Single rate kinetic model	$\frac{dm_{vol}}{dt} = R_{vol}(m_{0,vol} - m_{vol}),$	$A_{vol} = 2 \times 10^5 [s^{-1}],$
	$R_{vol} = A_{vol} exp\left(\frac{E_{a,vol}}{RT_p}\right)$	$E_{a,vol} = 6.7 \times 10^7 [J/mol]$ - [25]
Char oxidation		
Diffusion/kinetic - [26]	$\frac{dm_{char}}{dt} = -A_p p_{O_2} \frac{R_{diff} R_c}{R_{diff}+R_c},$	$A_c = 0.0053 [kg/m^2 s Pa],$
	$R_c = A_c exp\left(\frac{E_{a,c}}{RT_p}\right),$	$E_{a,c} = 8.37 \times 10^7 [J/mol]$ - [25]
	$R_{diff} = \frac{5\times10^{-12}}{d_p}\left(\frac{T_g+T_p}{2}\right)^{0.75}$	
Gaseous reactions of volatiles and CO		
Eddy dissipation model - [27]	$R_{k,r,P} = \vartheta_{k,r} M_{w,k} AB\rho\frac{\varepsilon}{k} min\left(\frac{\sum_P Y_P}{\sum_j \vartheta_{j,r} M_{w,j}}\right),$	$A = 4.0, \; B = 0.5$
	$R_{k,r,R} = \vartheta_{k,r} M_{w,k} A\rho\frac{\varepsilon}{k} min\left(\frac{Y_R}{\vartheta_{R,r} M_{w,R}}\right)$	

The interested reader is directed to the works of Rawlins et al. [8, 9, 28] for a more detailed description of the developed CFD multiphase modelling approach.

To ensure that the results are grid independent, simulations were performed for mesh sizes of 3.5 million, 6 million and 10 million cells. The numerical mesh consisting of approximately 6 million cells was selected as the most efficient grid and was subsequently used for all the CFD simulations for generating the CFD solution data needed for the training and testing of the surrogate model. To ensure numerical stability during the simulation runs, the aspect ratio and mesh orthogonality quality were kept below 20 and above 0.2, respectively. The simulations were solved using the SIMPLE pressure–velocity coupling scheme. The pressure term was discretised using the PRESTO! scheme. Momentum, species, and energy equations were discretised using the second order upwind scheme and the turbulent kinetic energy and dissipation rate using the first-order upwind scheme. The convergence criteria for the simulation model were set to 1×10^{-3} for the continuity equation, 1×10^{-4} for the velocity monitors, 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.
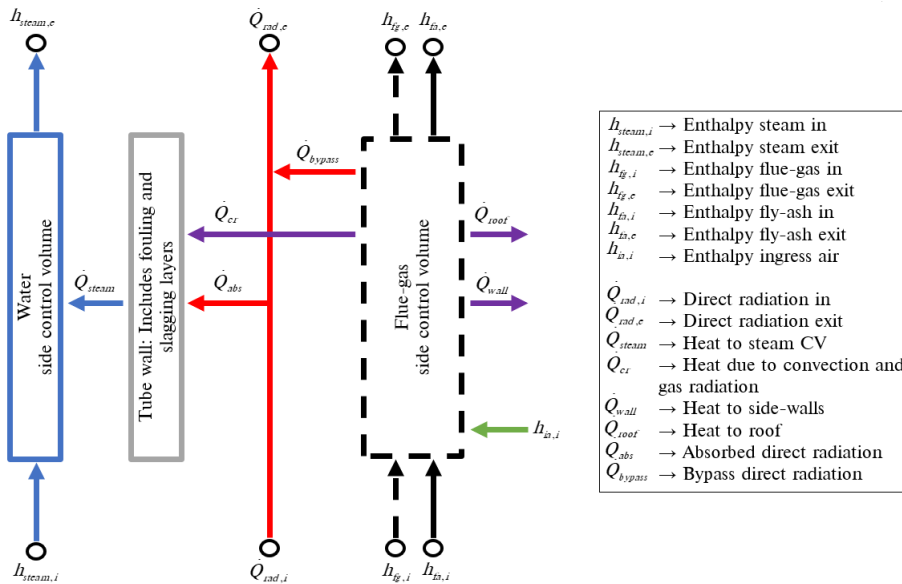
2.3. Process model

A 1-D discretised model of the furnace evaporator (EV), SH2, SH3, and subsequent downstream heat exchangers was developed using FlownexSE® 2021 [29]. 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.

The homogeneous two-phase mixture approach was used in modelling the heat exchanger water-side components of the entire boiler, whereby the fluid properties, phase velocities and temperatures are uniform per cross-sectional area. Table 3 provides a summary 1-D models homogeneous property equations and governing transport equations.

Table 3
Summary of 1-D process models properties and governing equations

Mixture property	Equation
Mixture fraction	$\alpha_H = \frac{\rho_l x}{\rho_l x + \rho_g (1-x)}$
Mixture density	$\rho_M = (1-\alpha_H)\rho_l + \alpha_H \rho_g$
Conservation entity	**Conservation equation**
Mass	$\frac{\partial}{\partial t}(\rho_M A) + \frac{\partial}{\partial s}(\rho_M A u) = 0$
Momentum	$\frac{1}{A}\frac{\partial}{\partial t}(\rho_M A u) + \frac{1}{A}\frac{\partial}{\partial s}(\rho_M A u^2) = -\frac{\partial p}{\partial s} - \frac{\tau_W P}{A} - \rho_M g \frac{\partial z}{\partial s}$
Energy	$\frac{\partial}{\partial t}(\rho_M h_M) + \frac{1}{A}\frac{\partial}{\partial s}(\rho_M A u h_M) + \frac{1}{2}\frac{\partial}{\partial s}(\rho_M u^2) + \frac{1}{2A}\frac{\partial}{\partial s}(\rho_M A u^3) = \frac{\partial p}{\partial t} + \frac{\dot{Q}_w}{V} - g\rho_M u \frac{\partial z}{\partial s}$

The process model includes all the heat exchangers up to and downstream of SH3, which includes heat exchangers RH2, SH1, RH1 and the EC. In addition, the model includes all the relevant attemperators (ATT1, ATT2, and ATT-RH), inlets, and outlets (refer to Figure 5). The process model is used to determine the required attemperation flow rates and the overall water-/steam-side thermal response. The furnace EV, SH2 and SH3 walls are represented by single lumped pipe flow components.
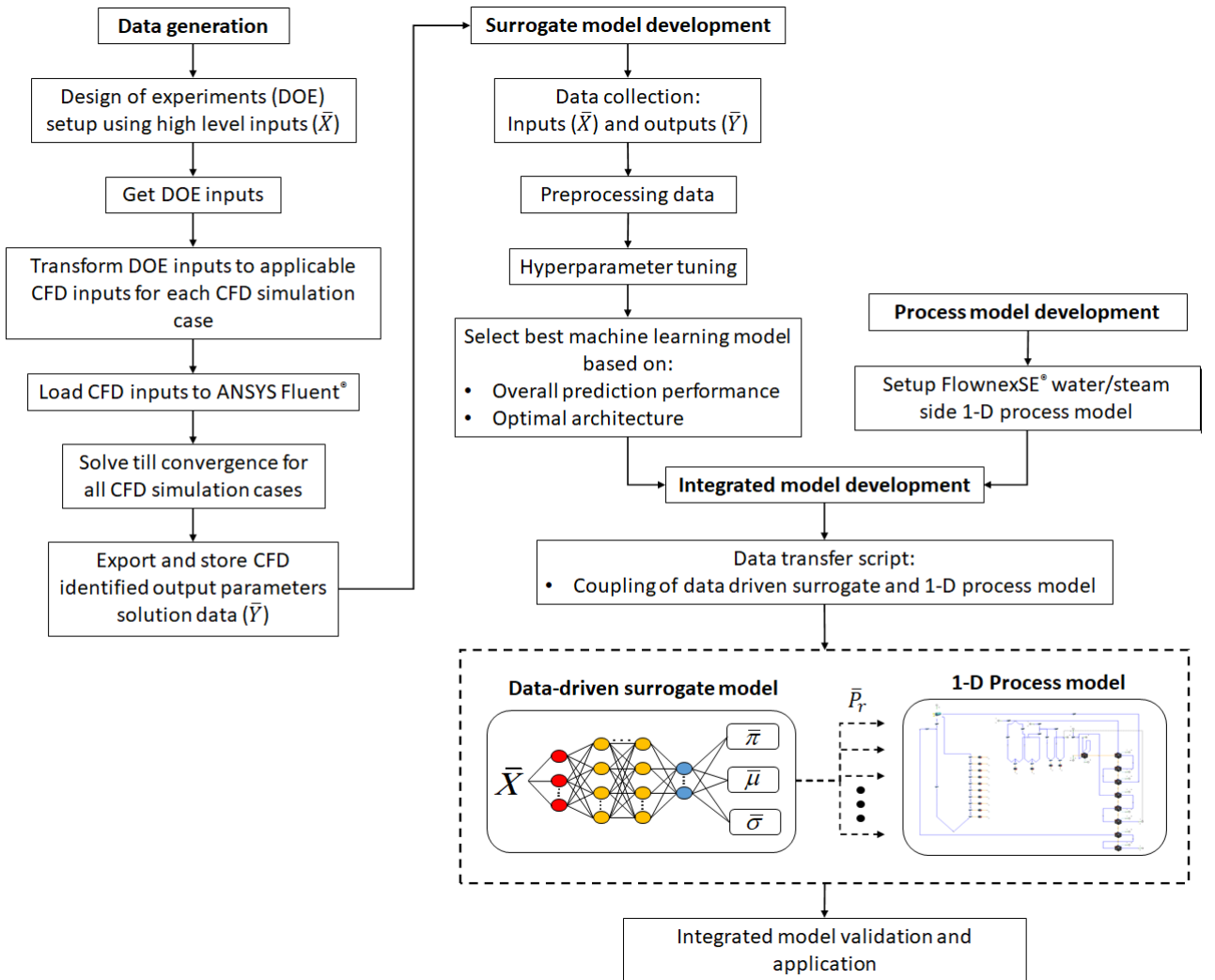

Figure 3: Heat exchanger component model [28]

The heat exchanger downstream of SH3 are modelled using a heat exchanger process model, which has successfully utilised in the works of Rawlins et al. [28]. Figure 3 illustrates the utilised heat exchanger process model. The component accounts for the radiative (gas and direct), convective, and conduction heat transfer mechanisms to and from the flue gas-side and water-/steam-side control volumes. The total heat transferred to the water-/steam-side control volume can be written as the sum of the absorbed direct radiation and the combined convective and gas radiative heat transfer

2.4. Integrated model

The integrated model is primarily the integration of a data-driven surrogate model and a 1-D 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. In the current study the development of the surrogate model follows two steps, namely data generation, utilising the CFD model presented in Section 2.2 to generate a training/testing dataset, and secondly the selection of the best machine learning model, based on the results of a hyperparameter search.

Figure 4: Summary of the model development process from data generation and collection to the final integrated model

Figure 4 provides a flow chart that indicates the development process from the data generation and collection to the final integrated model development. The data generation process is further described in Section 4, which provides an evaluation of the 14 high level inputs ($\bar{X}$) and CFD output parameters ($\bar{Y}$) that are to be utilised in the surrogate model development. The surrogate model development is primarily covered by the machine learning theory discussed in Section 3 and the results of Section 5.1 which highlights the selection of the best machine learning model. Section 2.3 provides an overview of the developed 1-D process model. Thus, the remainder of this section will discuss the integrated models coupling and layout development.

In the current work a C# script was used to access the Python application programming interface (API). This allows for the predictions ($\bar{P}_r$) to be made using the selected surrogate 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 5. The process model accepts the predicted heat loads for the EV, SH2 and SH3 at interconnected boundary nodes, the same is illustrated for the flue gas state.

Figure 5: Schematic of 1-D process model integrated with the data-driven surrogate model.

The boiler can be split into three sections namely, the furnace consisting of EV water walls, 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 808 [K] 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).

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 MLP and MDN machine learning models.

3.1. Multilayer perceptron networks

Artificial neural networks (ANN) are machine learning systems inspired by biological neural activity [30]. There are many classifications of ANNs, with multilayer perceptron networks (MLP) being the standard representation [31]. 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 cost function usually calculates the difference between the estimated ($\hat{y}_i$) and the desired values and is reported as a single number [10].

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

A standard MLP schematic is given in figure 6, illustrating the common features of a MLP, that being the input, hidden and output layers.
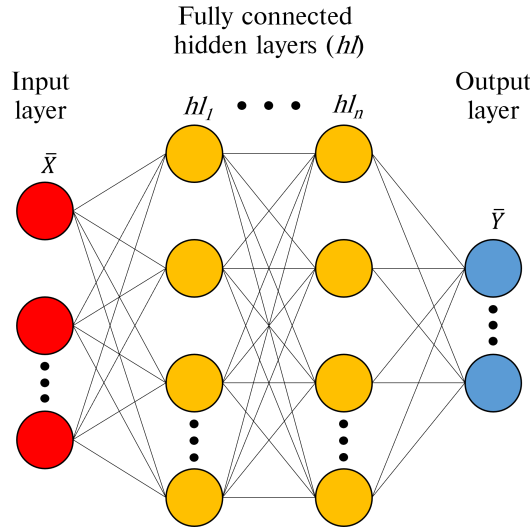
Figure 6: Traditional MLP schematic.

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 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 (3) highlights the first step, where $\overline{h}_{l-1}$ is the output signal from the previous layer.

$$\overline{z}_l = \overline{h}_{l-1} \cdot \overline{w}_l + \overline{b}_l \tag{3}$$

The result of equation 3 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 [31], 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 4.

$$\overline{h}_l = \sigma_{ReLu}(\overline{z}_l) = \overline{z}_l = \begin{cases} \overline{h}_{l-1} \cdot \overline{w}_l + \overline{b}_l & if \ \ \overline{z}_l > 0 \\ 0 & if \ \ \overline{z}_l < 0 \end{cases} \tag{4}$$

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

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 [32]. 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 [31] alternative to the gradient descent algorithm, which is illustrated in Equation (5). Forward- and backward-propagation algorithms would be iteratively conducted until the cost function is reduced to below a desired threshold.

Algorithm 1.

$$\overline{m} \leftarrow \beta_1 \overline{m} + (1 - \beta_1) \nabla_\theta J_{MSE}(\overline{\theta})$$

$$\overline{s} \leftarrow \beta_1 \overline{s} + (1 - \beta_2) \nabla_\theta J_{MSE}(\overline{\theta}) \otimes \quad \nabla_\theta J_{MSE}(\overline{\theta})$$

$$\overline{m} \leftarrow \frac{\overline{m}}{1 - \beta_1^t} \tag{5}$$

$$\overline{s} \leftarrow \frac{\overline{s}}{1 - \beta_2^t}$$

$$\overline{\theta} \leftarrow \overline{\theta} - \eta \overline{m} \otimes (\sqrt{\overline{s} + \epsilon})^{-1}$$

The variable $\overline{\theta}$ in Equation (5) represents the model weights and biases for each layer. Scaling ($\overline{s}$) and the momentum ($\overline{m}$) matrices are initialized to zero when beginning the training phase. The variable 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 (η) were investigated during the hyper-parameter search for both the MLP and MDN neural architectures.

3.2. 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 [33].

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 (6) illustrates the conditional probability function.

$$P(\overline{X} \mid \overline{Y}) = \sum_{k=1}^{K} \overline{\pi}_k(\overline{X}) \cdot \mathbb{N}(\overline{Y} \mid \overline{\mu}_k(\overline{X}), \overline{\sigma}_k(\overline{X})) \tag{6}$$

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

A schematic of a simple MDN network is given in Figure 7, 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 $\overline{\sigma}_k$ for each k distribution. This enables the MDN network to learn the $P(\overline{X} \mid \overline{Y})$.
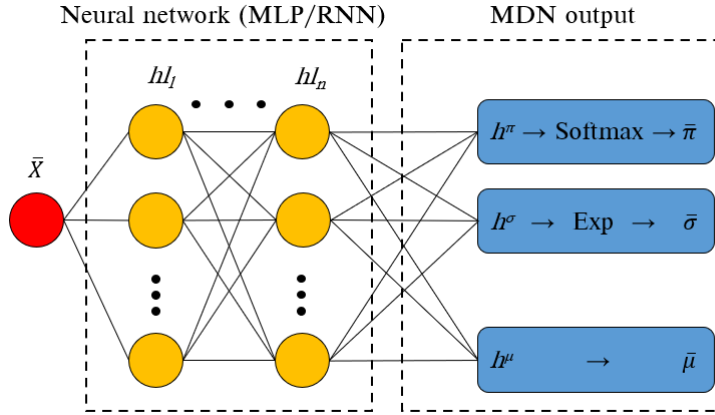
Figure 7: Simple MDN network.

Bishop [34] 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 (7).

$$\sum_{k=1}^{K} \overline{\pi}_k^n = 1$$

$$0 \leq \overline{\pi}_k^n \leq 1 \tag{7}$$

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 (8).

$$\overline{\pi}_k^n(\overline{x}^n) = \frac{exp(h_k^{\pi,n})}{\sum_{k=1}^{K} exp(h_k^{\pi,n})} \tag{8}$$

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 (9) shows the imposed constraint for the standard deviation leg.

$$(\overline{\sigma}_k^n(\overline{x}^n))^2 \geq 0$$

$$\sigma_k^n(\overline{x}^n) = exp(h_k^{\sigma,n}) \tag{9}$$

The MDN weights and biases are optimized by minimizing the negative log-likelihood for all observations in a batch. This is shown in Equation (10).

$$J_{NLL}(\overline{Y}, \overline{\pi}, \overline{\sigma}, \overline{\mu}) = -\sum_{n=1}^{N} ln \left\{ \sum_{k=1}^{K} \overline{\pi}_k(\overline{X}^n, \overline{\theta}) \cdot \mathbb{N}(\overline{Y}^n \mid \overline{\mu}_k(\bar{X}^n, \overline{\theta}), \overline{\sigma}_k^2(\overline{X}^n, \overline{\theta})) \right\} \tag{10}$$

4. Data generation

The steady-state multiphase non-thermal equilibrium CFD model, discussed in Section 2.2 was used to generate the training and testing data, which was subsequently used for training/testing of an appropriate data-driven surrogate model.

4.1. Simulated dataset

The high-level inputs to the surrogate model include the following: the excess air ratio, the fuel flow rate for the six mills in operation, the average steam temperatures for SH2 and SH3, the fouling resistance for SH2 and SH3, the proximate composition of ash and moisture of the fuel and the higher heating value (HHV) of the fuel, which is a function of the fuel composition determined using the Dulong equation [35]. Thus, the input field has a dimensionality of $d_{inputs} = 14$. To obtain a representative set of results for training and testing the surrogate model, a design of experiments (DOE) was conducted to generate a set of 180 unique simulation cases. The various high-level input ranges used in the DOE are provided in Table 4. 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 method provided in the Python *pyDOE* 0.3.8 library.

Table 4
Design of experiments input ranges for CFD simulations

Input variables	Mean	Min	Max	Units
Total fuel flow rate*, $(\sum_{i=1}^{6} \dot{m}_{fuel,i})$	68.9	39.5	120.2	kg/s
Fuel moisture content, (Y_{H_2O})	0.056	0.025	0.085	kg/kg
Fuel ash content, (Y_{ash})	0.418	0.259	0.559	kg/kg
SH2 fouling resistance, (R_{SH2})	0.006	0.004	0.007	K/W
SH3 fouling resistance, (R_{SH3})	0.014	0.008	0.017	K/W
Higher heating value, (HHV)	14.987	12.535	19.105	MJ/kg
Excess air, (γ_{air})	1.145	1.101	1.272	%
SH2 steam temperature, $(\overline{T}_{SH2})$	701	686	717	K
SH3 steam temperature, $(\overline{T}_{SH3})$	781	768	796	K

*Sum of the 6 mill fuel flow rates

The resultant matrix of 180 simulation cases would then be transformed into appropriate CFD inputs, as indicated in Figure 4. These 180 unique CFD cases were subsequently run using 216 Intel® Xeon® 2.6 GHz cores, made available by the Centre for High-Performance Computing, South Africa. Once the CFD simulations achieved convergence, the target data, comprising of the discretised 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 schematic of the output data reporting surfaces is given in Figure 8, and highlights the 10 discretised EV wall sections, the two tube bank legs making up both SH2 and SH3 superheaters, as well as the SH3 exit plane where the mass weighted flue gas state variables are reported.

A total of 22 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 monitored values, 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 ($\bar{Y} \in \mathbb{R}(50 \times d_{targets})$). Table 5 provides a statistical summary of the target data that was generated from the 180 CFD simulations, subsequently the data was pre-processed and cleaned to eliminate any unrealistic operational conditions.
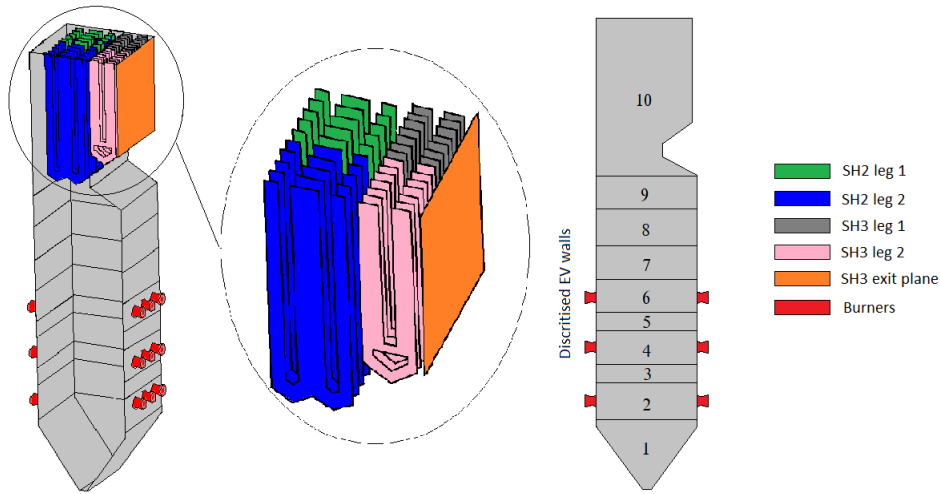
Figure 8: CFD solution data reporting surfaces, including the ten discretised EV wall sections, the SH2 and SH3 tube bank split and the SH3 exit plane.

Table 5
Summary of the CFD generated output data

Output variables	Mean	Std. deviation	Min	Max	Units
		Heat loads			
Wall - 1	30.76	25.85	1.01	98.12	MW
Wall - 2	21.56	7.68	0.84	43.24	MW
Wall - 3	20.85	8.03	1.17	38.37	MW
Wall - 4	29.17	8.82	5.20	46.99	MW
Wall - 5	22.77	8.61	4.70	43.24	MW
Wall - 6	30.97	10.19	4.95	52.99	MW
Wall - 7	39.92	14.26	10.98	84.82	MW
Wall - 8	38.70	11.06	13.66	67.89	MW
Wall - 9	30.62	8.22	11.78	52.06	MW
Wall - 10	78.06	20.70	30.04	148.02	MW
SH2 leg - 1	65.96	25.47	13.70	124.01	MW
SH2 leg - 2	78.98	26.13	15.56	132.30	MW
SH3 leg - 1	23.65	7.97	2.63	45.12	MW
SH3 leg - 2	24.46	7.58	3.80	43.75	MW
Roof	12.14	3.34	3.15	22.81	MW
		Flue gas composition and exit state			
Y_{CO_2}	0.188	0.025	0.126	0.235	kg/kg
Y_{H_2O}	0.041	0.005	0.031	0.051	kg/kg
Y_{O_2}	0.062	0.020	0.023	0.117	kg/kg
Y_{SO_2}	2.47×10^{-3}	3.2×10^{-3}	1.73×10^{-3}	3.02×10^{-3}	kg/kg
Temperature, $(T_{fg,exit})$	1143	109	894	1417	K
Mass flow rate, $(\dot{m}_{fg,exit})$	514.6	102.9	298.3	791.8	kg/s
Incident radiation, (G_{exit})	399.0	126.5	126.7	750.8	kW/m^2

5. Results and discussion

The following section provides the results of the hyperparameter tuning conducted in the development of the of the data-driven surrogate model using the generated CFD solution dataset of Section 4. Utilising the best machine learning model, the integrated model, as discussed in Section 2.4, was developed via the coupling of the 1-D process model of the water-/steam-side and the developed data-driven surrogate model. In addition, the results of a validation case study using experimental data was conducted for a wide range of loads. Finally, the results of an application case study, utilising the integrated model, was conducted that investigates the effects of fuel quality have on the integrated boiler performance.

5.1. Hyper-parameter tuning & final model selection

The present work makes use of two types of machine learning models, namely a standard MLP and a mixture density designated model connected to a standard MLP (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* 2.6.1 machine learning libraries were utilised in the present study.

The total CFD solution data matrix size will consist of $180 \times \overline{Y} = 198000$ data points. Using the total solution data matrix for training and testing would typically result in a slower convergence [31]. 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 $\hat{Y} \in \mathbb{R}(m_b \times d_{targets})$. However, the output data for the MDN model will consist of three parts, namely; the mixing coefficients tensor of shape $\bar{\pi} \in \mathbb{R}(m_b \times K)$, the output standard deviation tensor of shape $\bar{\sigma} \in \mathbb{R}(K \times m_b \times d_{targets})$, and the predicted means of tensor shape $\bar{\mu} \in \mathbb{R}(K \times m_b \times d_{targets})$, where K is the number of distributions. The input data fed into both the MLP and MDN will have shape $\bar{X} \in \mathbb{R}(m_b \times d_{inputs})$. The target dataset was split with 80% of the data being used 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 $0 \rightarrow 1$.

Table 6 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 6

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) was varied by considering the number of hidden layers and neurons per layer. Figure 9 (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 9 (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 9 (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.

(a)   (b)

(c)
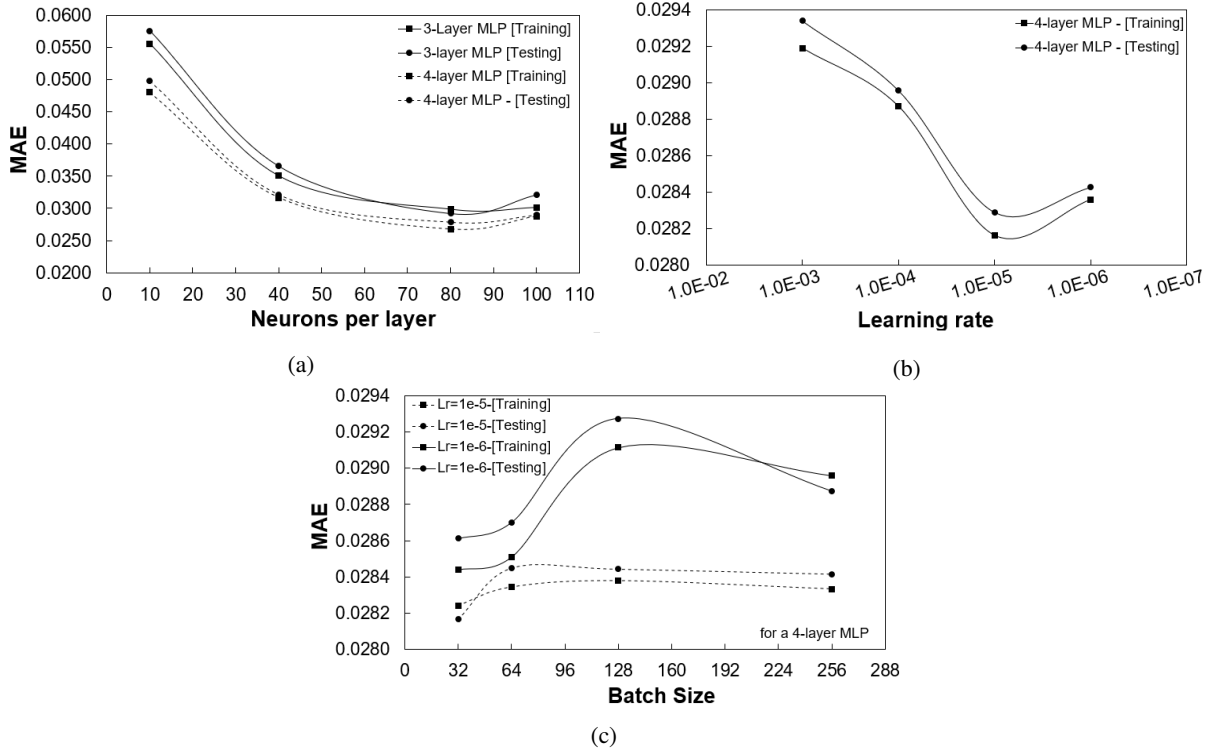
Figure 9: MLP model performance for; (a) hidden layer architecture, (b) varying learning rates, and (c) varying mini-batch sizes.

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 10 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.
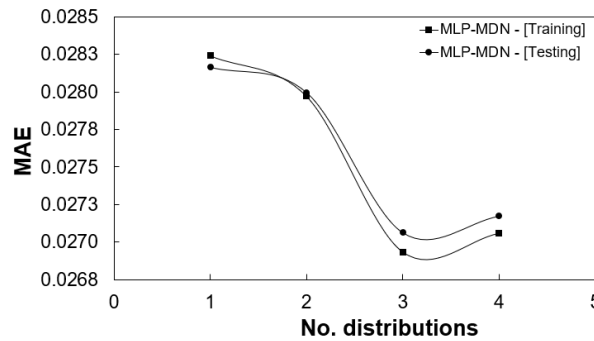

Figure 10: MDN model performance for number of distributions.

Table 7
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

The results of hyper-parameter search are given in Table 7. 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 11 (a) and (b). 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.
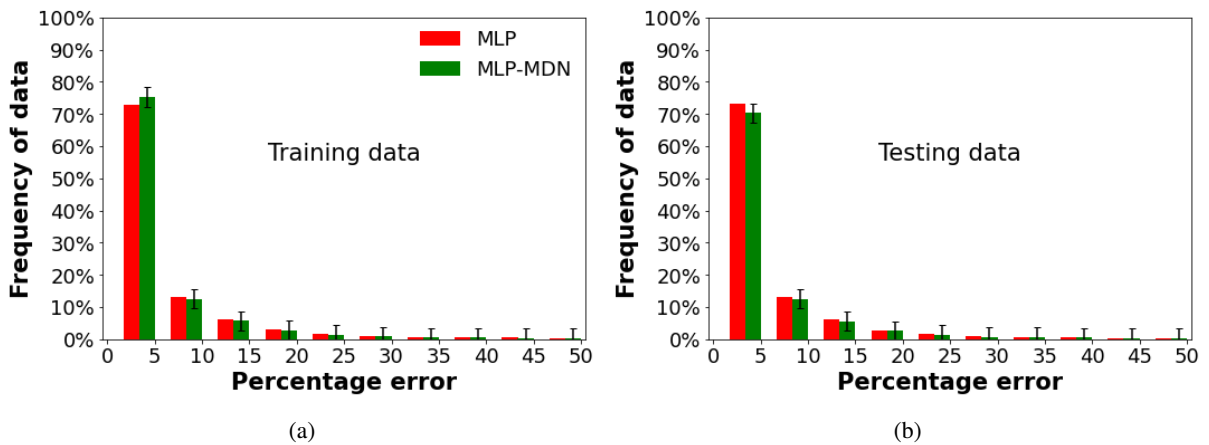
Figure 11: Overall error distributions for the selected MLP and MDN model; (a) training data and (b) testing data.

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 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.

5.2. Validation of integrated model using experimental data

The results predicted using the surrogate model were validated for the 100%, 80% and 60% MCR load cases. Measured data were 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.

The most probable predictions of the surrogate model can only be used to estimate the mean and standard deviation values for the EV, SH2 and SH3 wall heat loads. However, to propagate the uncertainty of the most probable predictions and ascertain the summary statistics for the convective section heat exchangers (RH2, SH1, RH1 and EC), as well as the steam generation rates and the attemperator flow rates, the Monte Carlo method was utilised. The Monte Carlo method is a mathematical technique used to estimate the outcome of a given stochastic process [36]. In the current work the integrated model was applied in conjunction with the Monte Carlo method [37]. 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.

Fortunately, FlownexSE® 2021, provides a built-in Monte Carlo facility. It requires as input the mean and standard deviation values for each model input [29]. The Monte Carlo model inputs used in this study are the predicted mean and standard deviation results obtained with the data-driven surrogate model. Table 8 illustrates the high-level MDN input vectors $(\overline{X})$ utilised for the validation and off-design case studies. For this study $N = 1000$ Monte Carlo simulations were performed for each validation load.

Table 8
MDN input vectors $(\overline{X})$ for the validation and the off-design fuel case studies

		Validation load cases			Varied fuel load cases	
Input variables		100%	81%	60%	High-ash	High-moisture
Excess air content		1.155	1.209	1.263	1.155	1.155
Y_{ASH} - $[kg/kg]$		0.409	0.409	0.409	0.501	0.409
Y_{H_2O} - $[kg/kg]$		0.055	0.055	0.055	0.055	0.150
HHV - $[MJ/kg]$		15.07	15.07	15.07	13.21	13.09
R_{SH2} - $[K/W]$		0.012	0.012	0.012	0.012	0.012
R_{SH3} - $[K/W]$		0.0067	0.0067	0.0067	0.0067	0.0067
SH2 steam temperature - $[K]$		698	697	781	698	698
SH3 steam temperature - $[K]$		787	793	781	787	787
		Mill flow rates - $[kg/s]$				
MCR load	**Mill 1**	**Mill 2**	**Mill 3**	**Mill 4**	**Mill 5**	**Mill 6**
100%	19.14	20.22	19.92	19.98	19.41	0.00
81%	18.23	19.72	19.02	17.82	18.02	0.00
60%	15.46	16.13	0.00	15.11	15.32	0.00

Figures 12 - 14 provide comparative results of the measured data and integrated model responses for all three MCR load cases. Included on all the plots are confidence bands which represent the range/spread of data for each variable. Figures 12(a) to (c) show the integrated model's response for the total heat loads to the various heat exchangers for the 100%, 80% and 60% MCR loads.
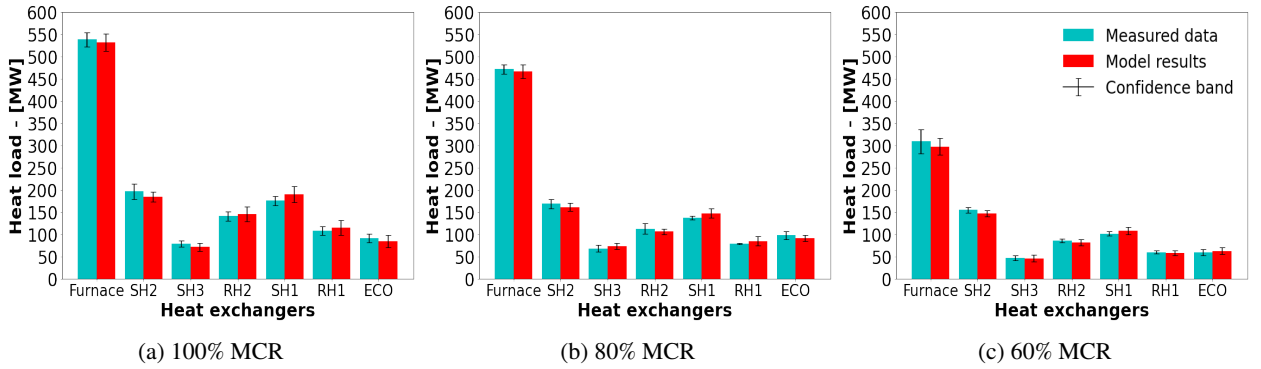
Figure 12: Load validation result comparison of the measured data and MDN model results for the various heat exchangers

For the 100% MCR load, the integrated model under predicts the mean results of the heat loads in the radiative section (EV, SH2 and SH3) by a maximum of 4%, which is realised in SH2 (approximately 8 [MW] for a mean SH2 heat load of 196 [MW]). Furthermore, the model over predicts the mean heat loads to the convective section (RH2 through EC) with a maximum deviation of 7%, which is realised in SH1 (approximately 10 [MW] for a mean SH1 heat load of 175 [MW]). A similar trend is realised for 80% and 60% MCR loads, with the former showing an over prediction of 4% of the mean heat load to SH3. 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, this is illustrated by comparing the confidence band intervals of the measured and model results.
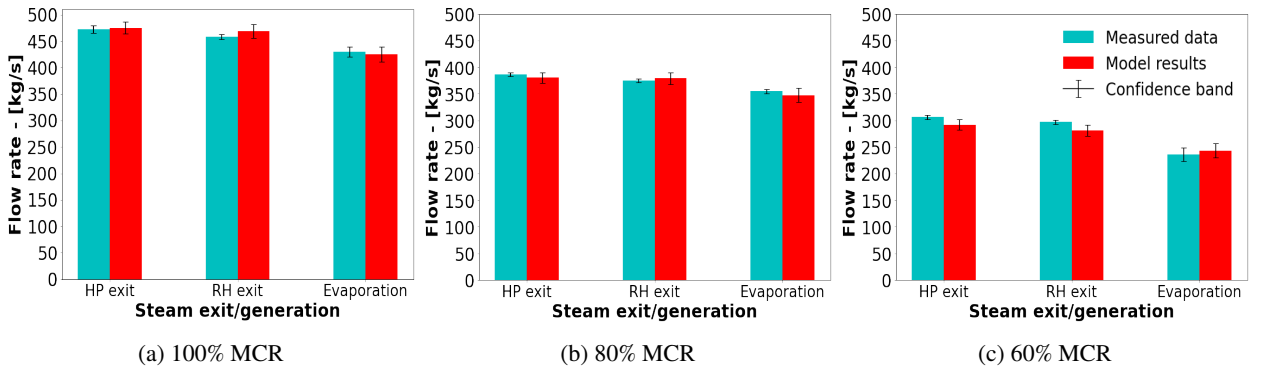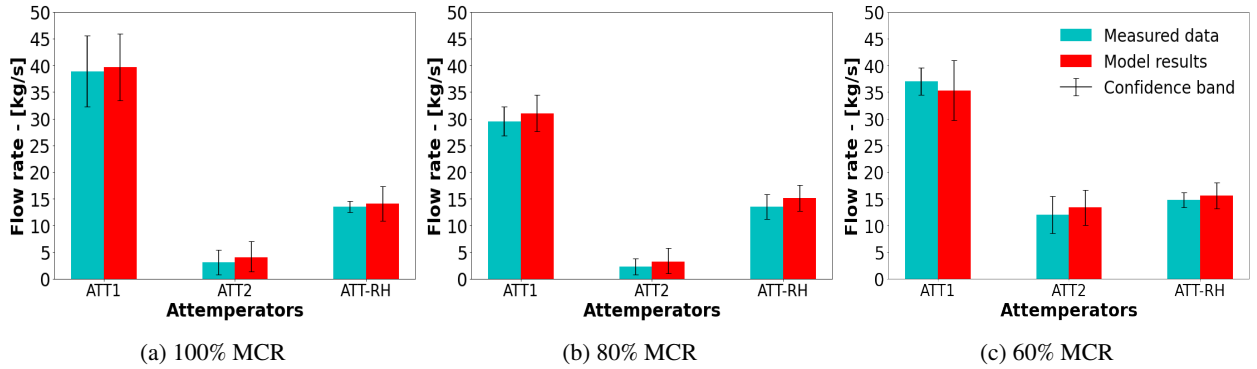

Figure 13: Load validation result comparison of the measured data and MDN model results for the steam exit and steam generation flowrates

Considering the steam generation and exit flow rates of Figure 13 (a-c), the integrated model is shown to sufficiently capture the hydraulic response of the case study boiler for a wide range of loads. A maximum under prediction of 8% is found in the RH exit of the 60% MCR load, and an over prediction of 4% is realised in the RH exit ow rate of the 100% MCR load. 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, namely the 60% MCR cases high-pressure and RH exit flow rates.

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 high-pressure and low-pressure steam turbines [38]. Figure 14 (a-c) highlight the attemperator flow rates for the various MCR load cases. The results of ATT1 show that the integrated model can capture the attemperator flow rate entering between SH1 and SH2 with adequate accuracy, however the 80% MCR load tends to over predict the mean ATT1 flow rate by 8%. Furthermore, the ATT1 flow rates demonstrate the largest uncertainty of the integrated model results, which all

overlap with the measured data confidence bands. The attemperator ATT2 flow rates are adequately resolved for the MCR loads, with a maximum over prediction of 8%. The attemperator ATT-RH, demonstrates sufficiently accurate results across all validation MCR loads.

Figure 14: Load validation result comparison of the measured data and MDN model results for the required attemperator flowrates to maintain operational integrity

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 10% tolerance of the measured mean values, while the steam generation rates exhibit a maximum difference of 8%, which is seen in the 60% MCR load case (refer to Figure 13 (c)). The attemperator flow rate results show the largest uncertainty/confidence band of all the results (Figure 14), however these overlap with the measured values.

5.3. Off-design analysis using the integrated model

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 ash (Y_{ASH}) and moisture (Y_{H_2O}) contents, which in turn affect the HHV value of the respective fuel. Two case studies are considered, one where the ash content is significantly increased, and a second case whereby the moisture content is only increased. The boiler efficiency, heat pick-up and thermal response are investigated to determine the impact these abnormal off-design conditions can have on the boiler's performance. The input vectors for the two fuel cases are provided in Table 8. 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.

As with the validation case studies of Section 5.2, the Monte Carlo method was employed to ascertain the mean and standard deviation values predicted via the integrated model. The base case in Figure 15 (a-c) refers to 100% load case of Figures 12-14.

Figure 15 (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 flow rate, 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 standby mill would allow the fuel flow rate 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.
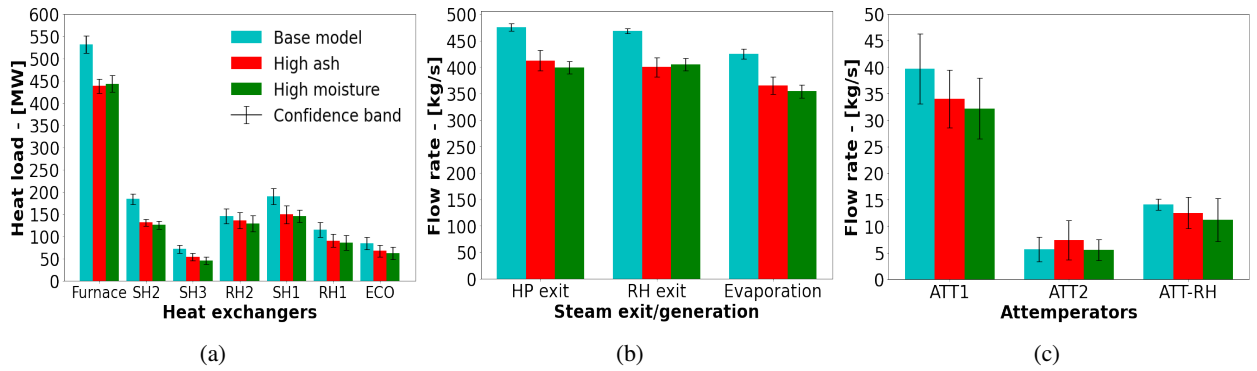
Figure 15: Integrated base case, high-ash and high-moisture case study comparison for; (a) the various heat exchangers, (b) the steam exit and generation flowrates, and (c) the required attemperator flowrates to maintain operational integrity

Figure 15 (b-c) similarly illustrate a decrease in the water/steam flow rates for the steam exit and attemperator conditions. The generation rates exhibit similar characteristics to the 80% load case of Figure 13 (b), illustrating a significant decrease in the operational capabilities when using a poor quality fuel. The predicted flue gas temperatures entering the convective pass are shown in Table 9. 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 high-pressure and RH exit steam and steam generation flow rates. The presence of more moisture in the fuel results in a lower inlet temperature, primarily since 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 ($\dot{m}_{fuel}HHV$). A significant decrease in the boiler efficiency is highlighted in Table 9 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 15 (a).

Table 9

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*

6. Conclusion

The present work establishes the basis for utilising an integrated data-driven surrogate and 1-D 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 $[MW_e]$. A total of 180 different CFD simulation cases were performed, where the input variables of Table 4 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 flow rates and attemperation flow rates 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.

Acknowledgements

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] Y. Shi, W. Zhong, X. Chen, A. B. Yu, J. Li, Combustion optimization of ultra supercritical boiler based on artificial intelligence, Energy 170 (2019) 804–817. `doi:10.1016/j.energy.2018.12.172`.

[20] S. Dal Secco, O. Juan, M. Louis-Louisy, J. Y. Lucas, P. Plion, L. Porcheron, Using a genetic algorithm and CFD to identify low NOx configurations in an industrial boiler, Fuel 158 (x) (2015) 672–683. `doi:10.1016/j.fuel.2015.06.021`.
URL `http://dx.doi.org/10.1016/j.fuel.2015.06.021`

[21] H. Versteeg, W. Malalasekera, Introduction to Computational Fluid Dynamics, The finite volume method, 2nd Edition, Pearson Prentice Hall, 2007. `doi:10.1002/9781119369189`.

[22] 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`.

[23] F. Lockwood, S. Rizvi, N. Shah, Comparative predictive experience of coal firing, Proceedings of the Institution of Mechanical Engineers 200 (1986) 79–87. `doi:10.1243/PIME`.

[24] C. Yin, On gas and particle radiation in pulverized fuel combustion furnaces, Applied Energy 157 (2015) 554–561. `doi:10.1016/j.apenergy.2015.01.142`.
URL `http://dx.doi.org/10.1016/j.apenergy.2015.01.142`

[25] 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`.

[26] M. Baum, P. Street, Predicting the combustion behaviour of coal particles, Combustion science and Technolodgy 3 (43) (1971) 231.

[27] ANSYS, ANSYS Fluent Theory Guide, 20th Edition, Ansys Inc., 2021.

[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) 1–29`arXiv: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] MTech, Flownex Theory Guide, MTech inc., 2021.

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

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

[32] 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`.

[33] 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`

[34] C. Bishop, Mixture Density Networks, Tech. rep., Aston University (1994). `doi:10.1007/978-3-322-81570-5_8`.

[35] V. Ganapathy, Steam generators and waste heat boilers: For process and plant engineers, 2014. `doi:10.1201/b17519`.

[36] D. Chen, J. Chen (Eds.), Monte-Carlo Simulation-Based Statistical Modeling, Springer, Singapore, 2017. `doi:10.1007/978-981-10-3307-0`.

[37] 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`.

[38] S. Kakac, Boilers, Evaporators and Condensers, Wiley Interscience, 1991.