REFEREED PAPER

APPLICATION OF MACHINE LEARNING ALGORITHMS IN BOILER PLANT ROOT CAUSE ANALYSIS: A CASE STUDY ON AN INDUSTRIAL SCALE BIOMASS UNIT CO-FIRING SUGARCANE BAGASSE AND FURFURAL RESIDUE AT EXCESSIVE FINAL STEAM TEMPERATURES

LAUBSCHER R[1], ENGELBRECHT Q[2], AND MARAIS CFP[2]

[1]University of Stellenbosch
[2]John Thompson a division of ACTOM

francoism@johnthompson.co.za

Abstract

The current work sets out to showcase the power of statistical learning algorithms to mine boiler operational data in an attempt to create a predictive model capable of capturing the plant specific behavior. The machine learning predictive model can be used to perform investigations such as boiler diagnostics, sensitivity analysis on operational parameters and root cause analysis to determine cause of upset/detrimental conditions. A data mining analysis was performed on an industrial scale biomass boiler co-firing sugarcane bagasse and furfural residue, which operated at excessive final steam temperatures (420-440°C) when compared to the design steam temperature (400°C). The goal of the analysis was to find the cause of the excessive final steam temperatures and propose remedial action. The analysis comprised of using artificial neural network, support vector regression and random forest machine learning algorithms to mine the operational data acquired from the boiler's distributed control system and generate a statistical predictive model.

A sensitivity analysis is performed on the boiler input parameters (fuel moisture, fuel density, fuel feeder speeds, induced draught fan speed, and forced draught fan damper position) using the machine learning model, to find the inputs which cause excessive temperature excursions. The model was able to accurately capture the boiler trends, and was used to find that it was the fuel moisture, density and upward flow velocity in the furnace which caused the flame to be positioned much higher in the furnace than intended. The higher flame position caused an increase in thermal radiation heat transfer to the radiant superheater above the design values which resulted in the higher final steam temperature.

Keywords: boilers, machine learning, artificial neural networks, diagnostics

Introduction

Upset or detrimental boiler conditions can lead to the damage of equipment, plant inefficiency, plant downtime and danger to staff members working on or around the unit. To find the cause of these conditions is mostly not an easy task as the amount of boiler input and output parameters creates a multidimensional problem to troubleshoot. Human beings excel at understanding 1 to 3 dimensional problems and the internal dependencies between the individual parameters, and struggle to comprehend the relationships between the various parameters in higher dimensional problems. Such was the case that will be discussed in the current work where a radiant superheater and internal de-superheater were installed in a pre-existing boiler unit. During operation the steam temperature increase in the new radiant superheater (ΔT=135°C) was consistently 35% higher than the design value

(ΔT=100°C). Excessive final steam temperatures can lead to thermal fatigue of the superheater material and in turn boiler downtime, continuous lifting of the superheater safety valve and superheater automatic vent valve which reduces plant efficiency or even trips turbines due to excessive inlet steam temperatures. The objective of the current work was to develop a predictive model which can be used to determine the cause(s) of the excessive final steam temperature by varying combinations of boiler input parameters to the machine learning model and predicting the final steam temperature response.

The machine learning algorithms chosen for the current work were limited to: regression artificial neural networks (RANN), random forests (RF) and support vector regression (SVR). The mentioned machine learning models have been previously used for various applications in the power generation sector. Neural networks were used by Tan *et al*. (2015) to model the correlation between operational parameters (boiler input parameters such as induced draught fan speed, O_2 concentration in the furnace, and secondary air fan speed) and NO_X emissions for a 700 MW coal-fired power plant. The model developed by Tan *et al*. (2015) had an absolute prediction error of 0.92%. The model was used along with various optimisation algorithms to minimise NO_X emissions within a set of constraints. Machine learning was used by Hu *et al.* (2017) to develop a fault detection model based on signal reconstruction for a combined-cycle power plant compressor. The trained machine learning model had a mean squared error (MSE) in the order of 0.01 for the reconstructed signal, and Tufekci (2014) utilised machine learning to develop a model capable of predicting electrical power output for a base load combined cycle power plant using ambient temperature, atmospheric temperature, relative humidity and exhaust steam pressure as the predictors for the model. Using similar approaches to the above mentioned work a machine learning model will be trained on a year's operational data, where boiler input parameters will act as predictors and final steam temperature as target/response variable for the model.

This paper is comprised of the following sections:

(i) Boiler and fuel information, where the boiler operational parameters will be presented and discussed along with the fuel characteristics and boiler history.
(ii) Data modelling, where a brief discussion of the theory regarding the implemented machine learning models will be presented, dataset exploratory analysis, data augmentation, comparison of models' in-sample and validation MSEs and final model selection.
(iii) Results and discussion, where the selected machine learning model is used to predict the final steam temperature and inference is made regarding the observations.

Boiler and fuel information

The boiler used for the case study has an evaporation rate of 140 000 kg/h and final team conditions of 30 bar(g) and 400°C, whilst co-firing sugarcane bagasse and furfural residue. The boiler has a three-pass evaporator bank configuration, with a two-stage superheater with interstage attemperation by means of an integral indirect contact de-superheater. The primary superheater is a screened drainable horizontal superheater and the secondary a radiant pendant superheater. Table 1 shows the boiler design parameters for the maximum continuous rating (MCR).

Table 1. Boiler design parameters.

Parameter	Units	MCR Value
Boiler evaporation rate	kg/h	140 000
Final steam pressure	kPa(g)	3000
Final steam temperature	°C	400
Final gas temperature	°C	175
Boiler efficiency (LHV)	%	85.6

Data were collected from the boiler distributed control system storage servers for the period April to December 2016. The data streams that were collected and used for the development of the predictive model can be seen in Table 2. Other data streams were collected such as superheater steam pressure, furnace temperature, oxygen content at evaporator outlet, final gas temperature and evaporator bank outlet gas temperature but were not utilised in the development of the machine learning model because the cause of the excessive final steam temperatures was related to the combustion dynamics in the furnace, which will be discussed later.

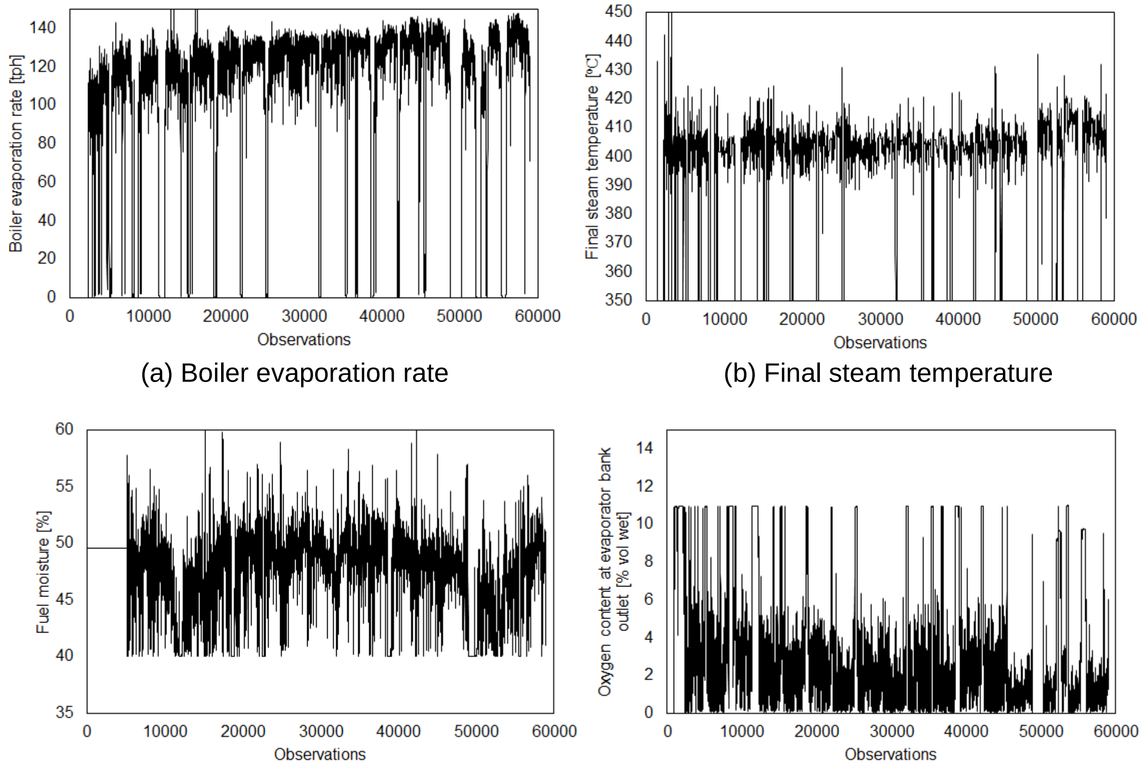
Table 2. Collected data streams from boiler DCS for predictive model creation.

Parameter	Units
Final steam temperature	°C
Boiler evaporation rate	tph
Induced draught (ID) fan speed	rpm
Feeder speeds (7 feeders)	rpm
Forced draught (FD) damper position	%
Front wall over-fire air pressure	Pa
Rear wall over-fire air pressure	Pa
Fuel densities (measured on 2 chutes)	%
Fuel moistures	%

The front wall over-fire air pressure reading is for a duct located behind the front wall of the boiler. The duct feeds the pneumatic fuel spreaders along with a front wall over-fire air injection above the feeder's entrance to the furnace. The rear wall has two sets of over-fire air nozzles located 1.2 m and 2.2 m above the grate where each set is comprised of 28 nozzles. The fuel densities are measured by nuclear density measurement devices located on two of the seven biomass feeding chutes. The fuel moisture is measured using an online moisture measurement device and has been calibrated using laboratory measurements of fuel moisture (error of roughly 2% on mass). Roughly 85% of required air is fed by the forced draught fan and is injected through the boiler grate. The amount of air introduced by forced draught fan is controlled by the fan's inlet damper position.

Figures 1a to 1d show plots of some of the boiler data streams for the mentioned period. The multiple boiler evaporation rate dips in Figure 1a are due to the monthly shuts the sugar mill has, otherwise the boiler is operated with an evaporation rate of between 110-130 tph for the majority of the period. Figure 1b shows that the boiler operates for the majority of the period with a final steam temperature above 400°C, with the attemperator valve completely open to the de-superheater device itself (not shown in figures). Therefore, the entire attemperator margin is dissipated by the high heat transfer rates to the secondary superheater. The original fuel moisture used for the design of the new radiant superheater and de-superheater was 51% on mass basis. From Figure 1c it is seen that the fuel moistures varies continuously between 45 and 51%. The oxygen content in the flue gas varies between 2 and 4% for the majority of the data points which are relatively close to the design value of 3.1%.

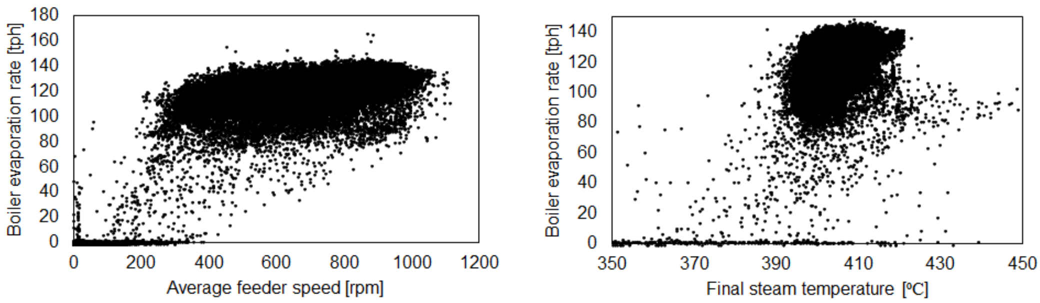
From figures 2a and 2b it is seen that for a given steaming rate the average feeder speed varies markedly (at 120 tph the feeder speeds vary between 400 and 1 000 rpm), pointing towards a substantial change in fuel composition or density. The biomass feeders for the boiler under consideration are volumetric feeding devices and changes in the fuel density may introduce more fuel to the furnace than required for the given load requirement, thus one would anticipate that the feeder speeds will vary significantly at a specific load due to the variation in fuel bulk density. Figure 2b shows that for a given evaporation rate the final steam temperature variation is notable, roughly 25°C at 130 000 kg/h. Looking at the large variations in final steam temperature and feeder speeds for a given boiler load points towards notable fuel variability and variation in combustion dynamics.



(a) Boiler evaporation rate (b) Final steam temperature

c) Fuel moisture (d) Flue gas oxygen

Figure 1. Boiler distributed control system data streams for the time period: 2016-04 to 2016-12.



(a) (b)

Figure 2. Boiler evaporation rate as a function of average biomass feeder speed and final steam temperature.

Sugarcane bagasse is a fibrous fuel stream which is generated during the processing of sugar from cane. Furfural in turn is produced from bagasse in large reactors. The furfural production process produces another fibrous stream namely furfural residue which can be burnt in a boiler for steam production. While bagasse and furfural residue are fibrous materials from the same feedstock, they have very different physical properties and combustion characteristics (Naude *et al*. 1993).

Table 3. Chemical and physical characteristics of furfural residue and bagasse (Naude *et al.* 1993)

Parameter	Units	Furfural residue	Bagasse
Proximate analysis			
Ash	%	2.9	2.0
Volatiles	%	37.2	40.2
Fixed carbon	%	8.0	5.8
Moisture	%	51	52
Ultimate analysis			
Carbon	%	56.2	48.6
Hydrogen	%	5.8	6.0
Oxygen	%	37.4	45.4
Nitrogen	%	0.5	-
Sulphur	%	0.1	-
Combustion			
Gross calorific value per kg	kJ/kg	9 800	8 740
Temperature at onset of combustion	°C	250	228
Apparent activation energy	kJ/mol	110	100
Physical properties			
Bulk density	kg/m^3	450	140

The texture of furfural residue indicates a size grading much finer than that of bagasse. Typical grading of furfural residue and bagasse are represented in Figure 3. The finer size grading could result in the furfural residue being suspended higher up in the furnace (due to upward air flow from the secondary air injections and primary air fed through the grate at the bottom of the furnace) causing the flame position to be closer to the radiant superheater than during normal bagasse only operation. In addition, the higher activation energy of furfural residue would also delay particle ignition and could add to a higher flame position. The higher flame position in turn increases the radiative heat transfer to the secondary superheater and leads to possible temperature excursions.

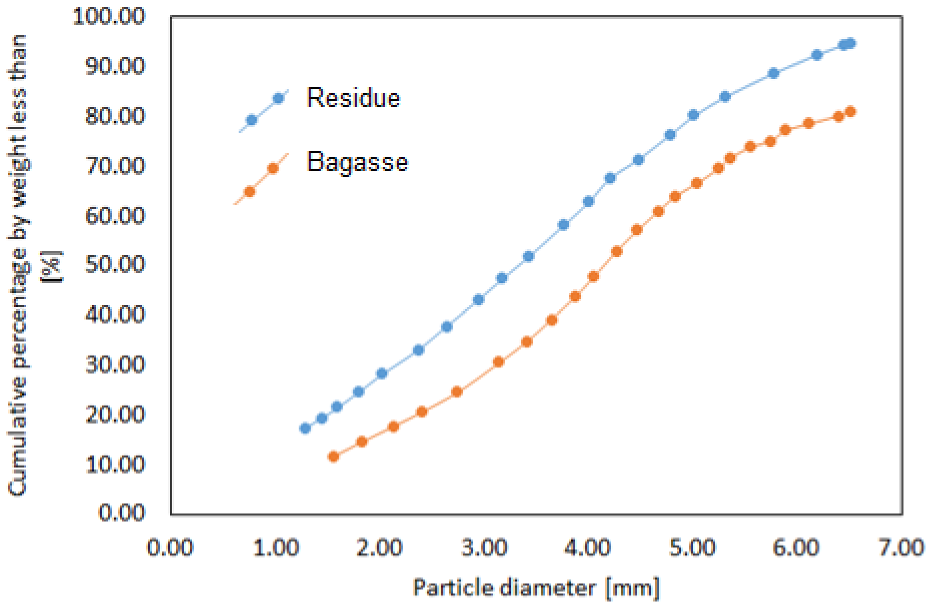


Figure 3. Particle size distribution of furfural residue and sugarcane bagasse.

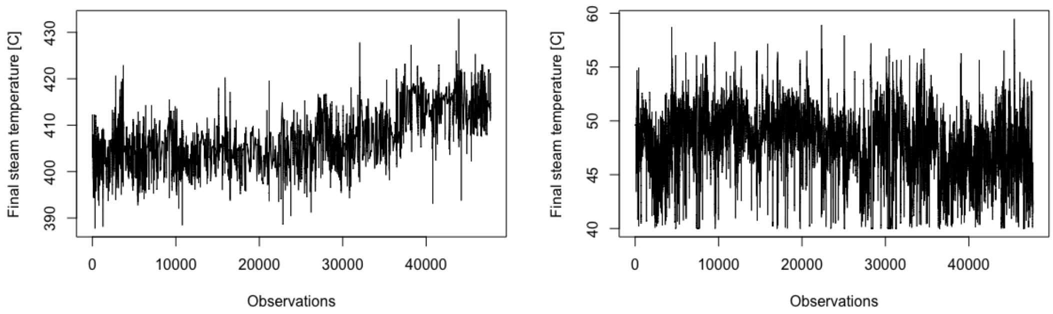
As mentioned previously, the design moisture used for the design of the new radiant superheater and de-superheater was 51% and a fuel blend of 90% furfural residue and 10% bagasse on a mass basis. The lower average moisture content of the fuel leads to a higher adiabatic flame temperature and can further increase radiative heat transfer to the new superheater along with the higher flame position. Other factors which influence the combustion dynamics in the furnace are the draught profile by means of enhancing the oxidizer and fuel mixing rates. The draught or air flow profile in the furnace is created by the 85% primary air introduced by the forced draught fan through the grate of the furnace vertically upward and the remaining fraction through the horizontally positioned over-fire air nozzles and pneumatic fuel spreaders which are fed by the secondary air fan. From the preliminary data and information on the fuel blend being fired in the boiler it can be ascertained that the combustion dynamics in the furnace plays a major role on the secondary superheater performance. The predictor values used to create a machine learning model able to capture the superheater performance should therefore be dependent on furnace inputs, which directly affect combustion dynamics. Additionally, to capture the required fuel and air amounts the boiler evaporation rate data stream can be used. The necessary predictor values used in creating the predictive model are listed in Table 2. In the next section the machine learning methods used to develop the predictive models from these data streams will be discussed.

Data modelling

Machine learning finds its origins in artificial intelligence, and the goal is to solve complex problems using machine learning methodologies as a collection of tools to learn complex relationships and then act or predict based on what has been learned. In other words machine learning takes complex relationships and regularities and learns from the observations through the use of computer algorithms. In this section the inner workings of the three machine learning models selected for this study will not be discussed, only data preparation will be discussed.

Data preparation

As the input space's dimension increases we encounter a variety of problems such as algorithmic complexity and storage space required to fit the target function increases significantly, complexity of target function increases and thus the risk of overfitting and the required sample size required to develop an accurate model increases. In the present work some simple feature selection and extraction techniques were utilised to reduce the feature space dimensionality and sample size. The original dataset of the operational parameters comprised of 28 parameters. Seeing as the goal of the investigation was to determine the cause of the excessive steam temperatures, feature selection was performed by selecting a smaller subset of features such as feeder speeds, FD damper position and ID fan speed. As discussed in a previous section only the features influencing the combustion dynamics in the furnace were selected, the selected features can be seen in Table 2. The subset selection reduced the input feature dimensionality from 28 to 15. To further reduce the input feature space's dimensionality some basic feature extraction was performed by combining existing features. The seven feeder speeds were combined into an averaged feeder speed data stream for all the feeders and the two chute density readings were also averaged into a single stream resulting in a feature space dimensionality of eight. The response variable or output was selected as final steam temperature only, therefore the data used to develop the machine learning models have eight input features and a single output feature. The time series data for the boiler load observed in figure 1a, shows that the unit for the majority of the observations generated steam at a load of >100 tph, with intermittent boiler shuts, where the boiler was brought offline for weekly maintenance, these observations should not form part of the data that will be used to model the boiler performance. Figure 2b shows that the high steam temperatures occurred at boiler loads above ±80 tph, therefore only observations where the boiler load was above 80 tph were used for model development. To remove the unwanted observations the data was sent through a high-pass filter. Therefore, all observations with a load below 80 tph were removed. Figures 4a and 4b below show the final steam temperature and fuel moisture data streams after it has been passed through the high-pass filter.



a) Final steam temperature (b) Fuel moisture

Figure 4. Some filtered data streams.

As mentioned one of the machine learning techniques used to model the data is artificial neural networks. The neural networks use gradient ascent to minimise the prediction error, but as the problem's dimensionality increases the minimisation worsens. A method used to assist the gradient ascent algorithm to converge faster is to scale the input features by

$$z_i = \frac{x_i - \mu_i}{\sigma_i}$$

where zi is the scaled observation value, xi is the original observation feature values, μi is the mean value of the feature and σi is the standard deviation of the feature.

Now that the feature space's dimensionality has been reduced and the data adjusted, the predictive models were developed and their performance accessed.

Results and Discussion

The major parameters which influences the combustion dynamics in the furnace are the fuel quality (particle size and moisture) and furnace draught (forced draught and over-fire air amounts). The results section is divided into two parts: firstly the combined effect of fuel density and moisture changes on the final superheater temperature was investigated and secondly the combined effect of upward draught and particle size on the final steam temperature was investigated.

The average boiler steam generation rate for the training dataset was 129 tph, therefore the investigation into the steam excursions using the machine learning models will be performed at this load only. For the first part where the fuel moisture and particle size effects are investigated the fuel density (chute density) was varied from 40→70% and for each density value the moisture was varied from 40→60%. The results for the three models is seen in figures 5 to 7.

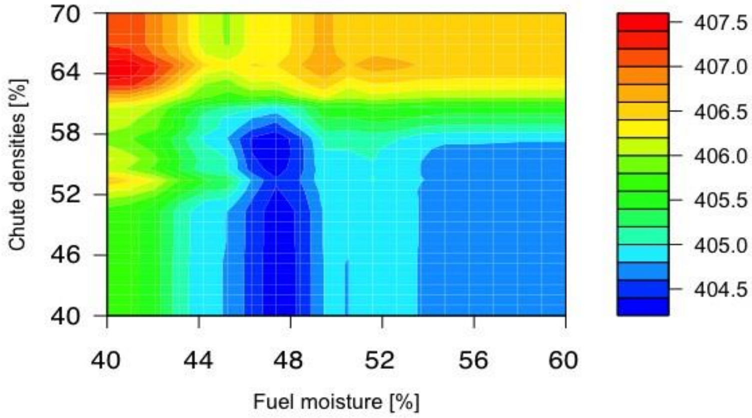


Figure 5. Random forest final steam temperature predictions for varying fuel moisture and density.

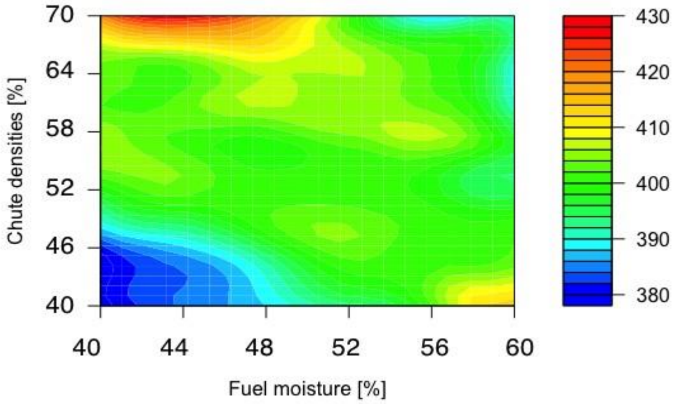


Figure 6. Artificial neural network final steam temperature predictions for varying fuel moisture and density.

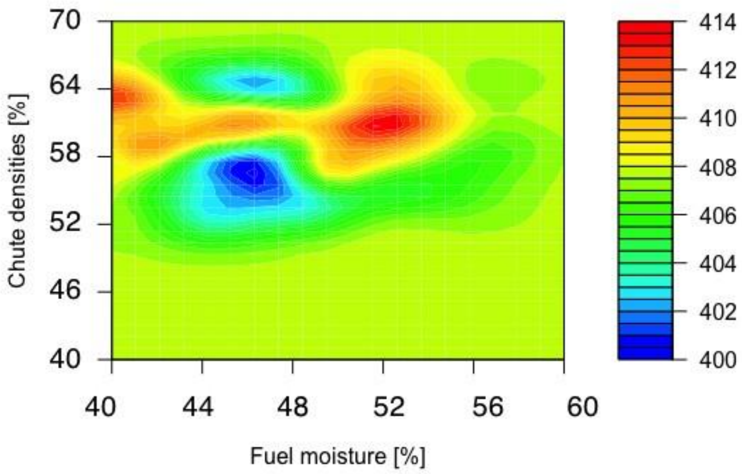


Figure 7. *E*-SVR final steam temperature predictions for varying fuel moisture and density.

From figures 5, 6 and 7 it is seen that the highest temperatures are found for fuel qualities that exhibit low moistures and high particle densities (small fuel particles). Lower moistures will increase the fuel's higher heating value thus causing the flame temperature to increase. The smaller fuel particles will cause the combustion zone/flame to be positioned higher up in the furnace (closer to the superheater) due to the increased relative drag force the particles experience. The combined effect of these two occurrences results in the heat transfer to the superheater to increase beyond the expected design values. From the results of the three models it can be seen that fuel moisture and particle size play a major role in the superheater performance and a clear temperature increase is seen as the fuel particle size reduces (chute density increases). For the low range of fuel moisture values; as the moisture increases the superheater temperature decreases due to the lower flame temperatures, but for the higher range of moistures as the moisture increases the superheater temperature slightly increases. The slight increase in superheater heat transfer at excessive moistures is due to the increased convective heat flux the superheaters experience which is due the increased mass flow rate of the fuel (the lower heating value of the fuel requires more mass of fuel to sustain boiler load).

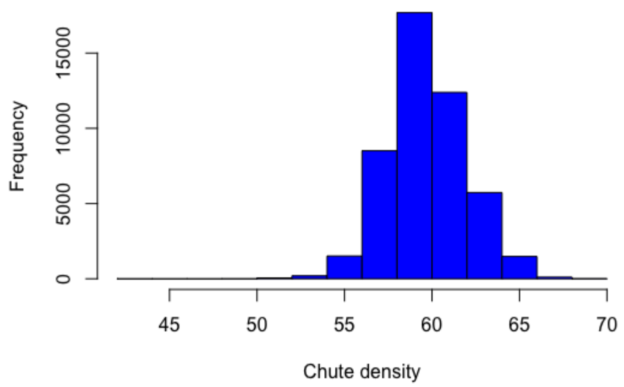


Figure 8. Chute density histogram plot.

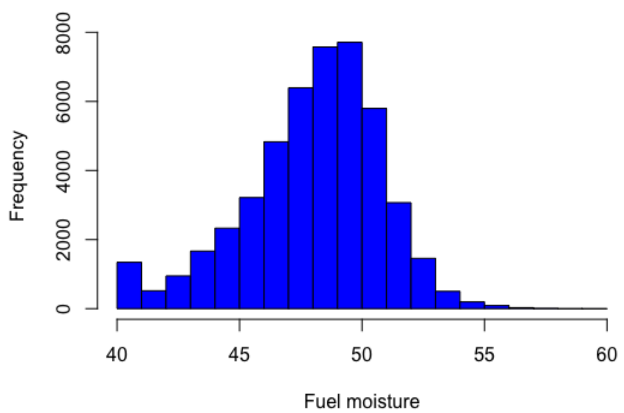


Figure 9. Fuel moisture histogram plot.

The average fuel moisture and chute density for the operation period, as mentioned, was 48% (with the majority of the variation being between 46-51%) and 60 % (with the majority of the variation being between 59-61%) respectively and the respective histogram plots can be seen in Figures 8 and 9. At these fuel conditions the neural network and E-SVR predicts temperatures above 410°C, whereas the Random forest model predicts 408°C. The Random forest model averages the output values that fall in a single partition, thus its prediction values is the averaged value from a wide range of temperatures, therefore the Random forest model predicts a slightly lower temperature compared to the other two models. In conclusion the major contributor to increased superheater heat transfer thus far is the fuel particle size, therefore the combined effect of upward draught and fuel density will be investigated next.

To investigate the combined effect of the upward furnace draught and chute densities, the forced draught fan damper position input to the machine learning models was varied between 20-100% for chute density inputs of 45, 55 and 65% and the average result of the three models was predicted and can be seen in Figure 10.

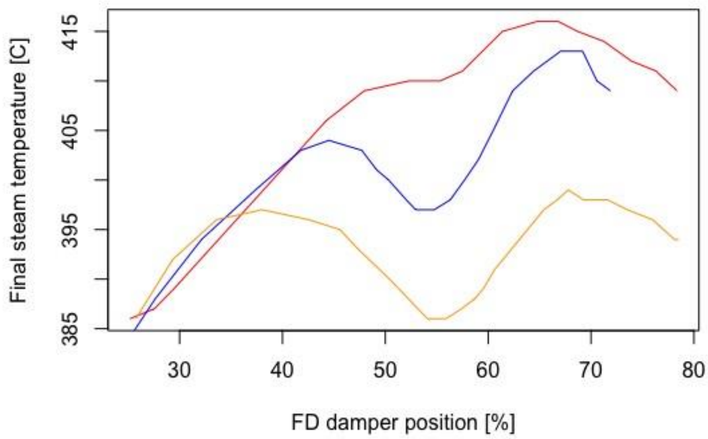


Figure 10. Effect of forced draught fan damper position and chute densities on final super-heater temperature. Red: 65%, Blue: 55% and Orange: 45% chute density.

For the majority of the operations period the forced draught damper position was varying between 50- 55% open. The histogram plot of the forced draught fan damper position is seen in Figure 11.

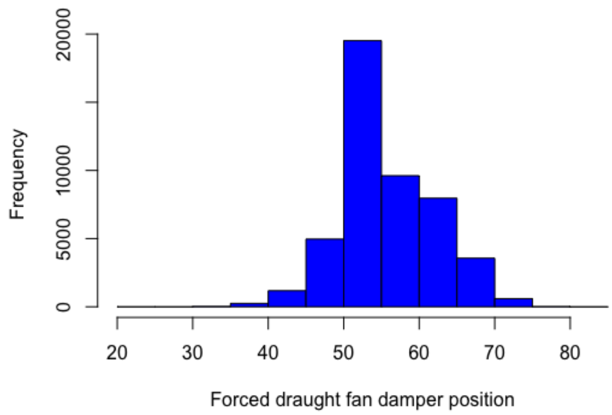


Figure 11. Forced draught fan damper position histogram.

Investigating Figure 11, we can conclude that there were not enough data points in the damper range between 20-40% for the machine learning algorithms to really capture the trend and thus the models predict similar superheater performance. Beyond 40% damper position the story is completely different, due to the fact that there is a multitude of data, the machine learning models are able to capture more accurately the trends in the data. We see from a damper position of 40% and upward that finer particles cause higher super-heater heat transfer and final steam temperatures. In addition, as the damper position is opened further the superheater temperatures climb, this effect is more pronounced the finer the fuel is.

Conclusions

The current work set out to find the cause of excessive final steam temperatures that are experienced in an industrial biomass boiler. Preliminary calculations and data pointed towards the combustion dynamics (flame temperature, position and height) as the cause of the higher radiant superheater heat transfer. To find the input parameters that cause the excessive steam temperatures, advanced machine learning algorithms were used to develop predictive models capable of calculating final superheater temperatures based on the boiler input parameters. These models were then used to pin point the cause of the over-performance. The random forest, neural network and E-SVR models all showed that the major contributor to the excessive performance of the super heater was the fuel particle size. The combination of fine particles and much lower moistures than the design values cause the superheater to over-perform. The upward draught which is influenced by the forced draught fan damper position also affects the superheater performance, especially the finer the fuel particles are.

The machine learning models were able to accurately capture the boiler superheater performance as a function of the boiler input parameters and were successfully used to find the root cause of the excessive final steam temperatures.

Future work

Future work includes developing a model capable of predicting all the boiler output parameters and using the model as a complete fault-finding tool and possibly a control system.

Use the forced draught damper and final superheater temperature data model (Figure 10) to adjust the control logic of the boiler to limit upward draught when the fuel particles' size reduces. The control logic will increase the over-fire air fraction at high chute densities.

REFERENCES

Hu Y, Palme T, and Fink O (2017). Fault detection based on signal reconstruction with auto-associative extreme learning machines. *Engineering Applications of Artificial Intelligence* 57: 105-117.

Naude D, McIntyre P, Field J (1993). The design and operation of boiler plant utilising furfural residue as fuel. *Proc S Afr Sug Technol Ass* 67:176-180

Tan P, Xia J, Zhang C, Fang Q and Chen G (2015). Modelling and reduction of nox emissions for a 700 mw coal-fired boiler with the advanced machine learning method. *Energy* 94: 672-679.

Tufekci P (2014). Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. *Electrical Power and Energy Systems* 60: 126-140.