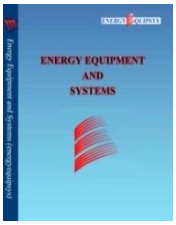
Energy Equip. Sys./ Vol. 12/No. 4/December 2024/ 423-453

ENERGY EQUIPSYS

Energy Equipment and Systems

www.energyequipsys.com



Thermal design of fire tube boiler with superheater and estimation of temperature increase in the superheater based on machine learning methods

Manuscript Type
Research Paper

Authors

Ebrahim Pilali[a,b]
Safiye Shafiei[a,b]
Ramin Kouhikamali[c]
Mohsen Salimi[d]
Majid Amidpour[a,b *]

[a]Department of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran

[b]Department of Energy System Engineering, Faculty of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran

[c]Department of Mechanical Engineering, Isfahan University of Technology, Isfahan, Iran

[d]Renewable Energy Research Department, Niroo Research Institute (NRI), Tehran, Iran

Article history:
Received : 24 May 2024
Revised : 24 June 2024
Accepted :30 June 2024

ABSTRACT

This research utilizes MATLAB and Python coding to optimize the thermal design of an industrial shell and tube steam boiler with an internal superheater. The paper outlines a systematic approach to steam boiler design, including heat transfer dynamics analysis, superheater configuration optimization, and implementation. They take action to enhance the performance of the third pass. The shell and tube steam boiler specifications, including an internal superheater, have been determined, with a steam capacity of 5 tons/hour and operating at a working pressure of 10 bar. According to the results, the opt substantially impacted 71 tubes in the second pass, each with a diameter of 5 cm, and an additional 82 tubes of identical size in the third pass (which includes revisions). To achieve a desired temperature increase of 15 ℃ in the superheater, incorporating the superheater section into the fire tube resulted in a 23.72% increase in the third pass level compared to the scenario without a superheater. For every 5 ℃ temperature increase in the superheater, the steam velocity in the third pass tubes decreases by approximately 1m/s. Adding the superheater to the end of the third pass reduces the temperature of this area from 525 ℃ to 500 ℃. Leveraging machine learning algorithms enabled the identification of parameters influencing the rise in superheater temperature. Linear regression emerged as the best predictor of superheater temperature increase among the eight models considered.

Keywords: Heat Transfer, Heat Exchanger Design, Fire Tube Boiler, Superheater, Machine Learning.

1. Introduction

Boilers are crucial components in various industrial applications, serving as primary sources for steam or hot water generation in processes such as manufacturing, heating, ventilation, air conditioning (HVAC), power generation, and other industries [1], [2], [3], [4], [5].

* Corresponding author: Majid Amidpour
Department of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran; Department of Energy System Engineering, Faculty of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran
Email: amidpour@kntu.ac.ir

Two steam boilers, fire-tube and water-tube boilers, are available to meet industrial and commercial steam needs. Choosing the best steam boiler is crucial for the long-term success of any project. The appropriate choice can save hundreds of thousands of dollars over the boiler's lifespan for the end user [6], [7], [8].

Fire-tube boilers are among the oldest types of boilers. They initially used this technology to heat water in the food and beverage industry. In these boilers, fuel (usually coal, wood, gas, or oil) is burned inside a fire tube, transferring heat to the water. Over time, with technological advancements, the design and performance of fire-tube boilers improved, and they are now

utilized in various shapes and sizes across different industries [9], [10].

Superheaters have also been used in fire tube boilers. They helped develop and deploy this technology in the food and beverage industry and were demonstrated in large power plants to achieve this. Still, they were expensive and occupied much space [11]. However, recent advances have led to new designs for superheaters. These new designs include vertically arranged heat-exchanging pipes with equal distances between them and U-shaped superheater tubes internally connected to inlet and outlet headers.

Additionally, suspended superheaters have been introduced, which are connected to inlet and outlet manifolds through a ceiling wall, reducing the building height and construction costs [12]. Other improvements include using a communication coil pipe with fins for better heat conduction and height adjustment capabilities. Furthermore, refractory components and sealing boxes have allowed unrestricted expansion displacement paths, reducing stress levels and prolonging the service life of the superheater.

Adding superheaters to boilers enhances efficiency. They demonstrated that steam temperatures improve industrial processes' performance, contributing to the development and enhancement of various industries [13].

Here, we discuss recent articles in the field of Firetube, helping them develop and deploy it in the food and beverage industry. They demonstrated the boiler and implemented a model-predictive control strategy to optimize its operation. This led to improved steam pressure regulation, reduced fuel consumption, and higher overall efficiency. This research provides valuable insights for industrial operators and researchers working on similar thermal systems.

Tognoli et al. [14] developed accurate and computationally efficient dynamic models for fire-tube hot-water boilers, which they validated against experimental data. They also presented a data-driven simulation approach using machine learning techniques, which provided valuable tools for analyzing, optimizing, and controlling these boilers.

Tognoli et al. [15] proposed a multi-setpoint control by helping them develop and deploy it in the food and beverage industry. They demonstrated significant fuel savings of up to 10-15% compared to traditional single-setpoint control. This approach has practical implications for boiler operators, plant managers, and energy efficiency specialists, contributing to industrial facilities' sustainability and environmental performance.

Morelli et al. [6] proposed a reduced finite volume model for accurately simulating the loss can improve fire tube Heat Recovery Steam Generators (HRSGs) efficiency, based on a pre-generated CFD database, demonstrated good agreement with experimental data and can be beneficial for the design and optimization of fire tube boilers and similar heat exchanger systems.

Seyed Mohammad javadi et al. [16] investigated the thermal performance and optimization of city gas station heaters equipped with turbulators in the fire tube section. The study demonstrated the potential for significant performance improvements using optimized turbulator designs, leading to substantial energy savings and reduced greenhouse gas emissions.

Habib and Nemitallah [17] investigated the design and potential benefits of integrating an Ion Transport Membrane (ITM) reactor into fire-tube boilers. The researchers developed a novel ITM reactor concept using selective ceramic membrane materials, which demonstrated high CO2 capture rates of up to 90% while maintaining acceptable levels of oxygen enrichment. Integrating ITM reactors into fire-tube boilers has practical implications for decarbonizing industrial heat and steam generation, reducing carbon footprints, and improving boiler efficiency.

According to Ozer Aydin and Y. Erhan Boke [18], optimizing the air-to-fuel ratio is crucial in controlling CO emissions from fire-tube water heaters, resulting in up to 50% reductions. The authors also explored other combustion-related modifications to improve CO emission performance further. These findings can benefit manufacturers and operators of fire-tube water heaters in developing and deploying heating solutions with lower environmental impact and enhanced safety, contributing to the broader knowledge of combustion process optimization in industrial and commercial heating equipment.

In another study by Mohammad Rofiqul Islam et al. [19], the properties of bio-oils derived from sugarcane waste through fixed-bed fire-tube heating pyrolysis were investigated. The researchers utilized a cost-

effective pyrolysis technology to convert the waste into bio-oils and comprehensively characterized their physical and chemical properties. These bio-oils could be suitable as alternative fuel sources for boilers, engines, and other thermal systems, aligning with the principles of a sustainable and circular economy. This study contributes to the understanding of biomass-based energy and chemical systems.

Bisetto et al. [20] examine the impact of tabulators on the operating conditions and heat transfer in a three-pass fire tube heat generator fueled by natural gas. The findings demonstrate that using tabulators reduces the emission temperatures of flue gases, resulting in improved efficiency. The researchers also developed a dynamic model to predict the system's behavior, validated through testing at various thermal power levels.

Marco Tognoli et al. [21] found that increasing the size of the boiler only slightly improves efficiency. They also noted that substantial boiler layouts are not economically beneficial and do not significantly reduce pressure variances.

Wim Beyne et al. [22] present a comprehensive thermal model for fire tube boilers, including the thermal influence of turn boxes. The study compares different models for tube passes and integrates radiation in the NTURAD model, which shows reduced sensitivity to control volume changes. The study finds that submerged turn boxes contribute 7% of heat transmission, while non-submerged turn boxes could reduce efficiency by 12%. Additionally, the study examines the impact of the boiler's initial state and feed water mass flow rate on the peak load period.

According to F.J. Gutiérrez Ortiz [23], a comprehensive model of a full-scale fire-tube boiler has been developed, which includes both the fire/gas side and the water/steam side. The model can be used for simulating and constructing controllers for multiple variables and has been validated through a case study on an 800 HP boiler. The model's performance aligns with existing literature and can be used for comparisons and oxy-combustion testing.

According to Tognoli et al. [24], their research on fire-tube boilers showed that larger boilers did not significantly affect performance or controllability. Additionally, smaller boilers experienced a decrease in efficiency of 1.8% with less variable demand profiles. The study also found that using smaller boilers did not substantially impact steam pressure, indicating that concerns over steam quality are not a significant limitation when considering potential savings.

According to Bhargav Akkinepally et al. [25], adjusting the tangential flow of the PC burner can minimize yaw-induced tube rupture caused by non-uniform velocity distribution in a tangentially fired pulverized coal boiler.

Ali Behbahani-nia et al. [26]suggest that optimizing exergy destruction and loss can improve the efficiency of fire tube Heat Recovery Steam Generators (HRSGs) while minimizing capital cost and energy loss. They use multi-objective optimization to find the best design variables and find that the optimal temperature for the pinch point is 6.26 °C, indicating that manufacturers should focus on enhancing heat recovery in fire tube HRSGs.

According to J. Gañan et al. [27], increasing the injection pressure in two three-stage peritubular boilers using gasoil C results in higher losses from sensible enthalpy of fumes, carbon dioxide emissions, and nominal power of the boiler, but lower concentrations of carbon monoxide. This trade-off determines the range within which maximum efficiency is achieved, where the decrease in efficiency from reduced losses of unburned gases has a more significant impact than the increase in losses from sensible heat.

This article has been written because there is a need for articles on the thermal design of fire tubes with superheaters. The primary objective of this project is to optimize the thermal design for an industrial sample of fire-tube boilers accompanied by an internal superheater. By applying thermal engineering methods and optimization algorithms, we aim to enhance the performance and reduce energy consumption in these boilers. Matlab and Python coding techniques will be utilized to achieve this goal, enabling us to simulate and analyze thermal models and obtain optimal results.

Given boilers' extensive importance in various industries and the critical role of superheaters in their efficiency and productivity, this project aims to contribute to advancing technology in this field, leading to

cost savings and energy efficiency across industries.

In this article, we will first delve into the history of fire-tube boilers and then discuss the role and importance of using superheaters in these boilers. Finally, we will outline the objective of this project, which involves optimizing the thermal design for an industrial sample of fire-tube boilers accompanied by an internal superheater, utilizing coding techniques in MATLAB software. Ultimately, leveraging machine learning algorithms enabled the identification of the varying degrees of influence of different parameters on the temperature rise within the superheater.

Nomenclature

$\dot{m}_g$	The mass flow rate of combustion products
$\dot{m}_w$	The mass flow rate of water vapor
M	Molar
V_{max}	The highest fluid velocity
v	Special volume
$\dot{V}_{CH_4}$	Volumetric flow rate of methane gas
$T_{gin.1}$	The inlet temperature of the Furnace
$T_{gout.1}$	The outlet temperature of the Furnace
$T_{gin.2}$	The inlet temperature of the Second pass
$T_{gout.2}$	The outlet temperature of the Second pass
$T_{gin.3}$	The inlet temperature of the third pass
$T_{gout.3}$	The outlet temperature of the third pass
T_{win}	The temperature of water entering the boiler
T_{wout}	The temperature of water steam leaving the boiler
T_{wall}	Pipe wall temperature
LMTD	Logarithmic mean temperature difference
ΔT_e	Excess temperature (boiling)
σ	Stephen's number
Nu	Nusselt number
Re	Reynolds number
Pr	Prantel number
h_f	Enthalpy of matter
h_{fg}	Enthalpy of vaporization
h_i	Internal displacement heat transfer coefficient
$h_{Radiation}$	Radiant heat transfer coefficient equivalent to displacement
$\bar{h}$	Average heat transfer coefficient
LHV	Lower heating value
U	Overall heat transfer coefficient
R_W	Pipe wall resistance
c_p	Specific heat
$cp_{v_{aper}}$	Special heat of superheated water vapor
Q_f	Burner capacity
$Q_{required}$	Heat received by the superheater
ε	Visibility factor
N	Number of tubes
ρ	Density
v	Special volume
K	Conductive heat transfer coefficient
μ	Viscosity
A_F	Cross section of the furnace
L_F	Furnace length
D	Diameter pipe
$D_{superheater}$	Diameter of the superheater
R	Combustion reactants
P	Combustion products
ST	Transverse step
S_L	Longitudinal step
R2	R-squared
MSE	Mean squared error

2. Boilers and machine learning methods

Fire-tube and water-tube boilers are steel steam boilers. Although they produce steam, they are similar in many ways. However, attention to each unit's specific characteristics distinguishes its design and performance. Figure 1 shows the General Classification of steam boilers.

2.1. Water tube boiler

Water tube boilers are heat exchangers of the shell-and-tube type, where the cold fluid (water) is located in the tubes, and the hot fluid (combustion gases) is located in the shell. In this type of boiler, unlike the fire-tube boiler, water vapor can reach a temperature and pressure higher than its saturation limit and become superheated [28].

Since power plants require high-pressure and high-temperature steam, other boiler designs, not fire-tube boilers, have been proposed.

Stephen Wilcox and George Babcock can be considered the pioneers of the blue tube boiler in its present form. Their steam boiler, designed in 1877, is shown in Fig. 2.

Fig. 1. General Classification of steam boilers

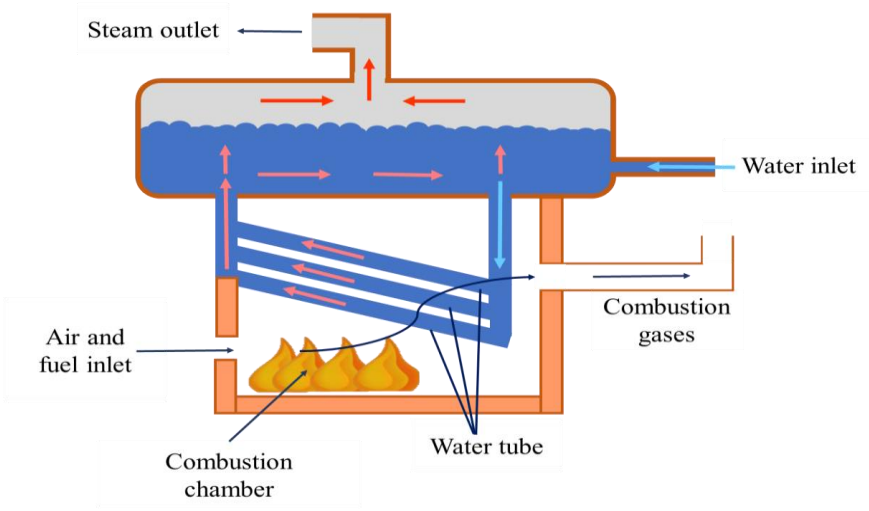


Fig. 2. A view of the blue tube boiler designed by Stephen Wilcox and George Babcock

As shown in Figure 2, the pipes containing water are inclined, and their two ends are connected to a cylindrical container called a drum by two conduits. The steam produced in the boiler pipes enters the drum before reaching the consumer. The heat exchange surface in this steam boiler consists of many tubes with an approximate diameter of 75 mm. Some of these pipes are directly exposed to the combustion flame, and the rest are exposed to hot gases from combustion. Several blades have been used to guide these gases around the pipes. By correcting the movement path of gases, these blades increase the level of heat transfer and, as a result, improve the boiler's efficiency. In this way, in this type of boiler, the heat of

combustion gases is transferred to water through a group of pipes whose cross section is relatively small compared to the shell diameter of the fire tube boiler. As a result, it makes it possible to increase steam production.

2.2. Fire tube boiler

The fire tube boiler contains pipes that enter it from one side and exit from the other; thus, the pipes occupy part of the boiler space, and the rest of the available space is considered for water [20].

The hot gases from combustion in the combustion chamber enter this set of pipes and pass through the boiler. Meanwhile, the heat transfer between the gases passing through the pipes and the water inside the boiler causes the water to heat up and produce steam [21].

In the fire tube boiler, the length of the furnace cannot be considered less than a specific limit Because, according to the dimensions of the cone-shaped flame, there will be limitations in the diameter and length of the furnace. On the other hand, the distance from the tip of the flame to the end of the stove has a specific limit to create a uniform heat transfer, avoid the concentration of thermal stress, and avoid the melting of the furnace wall [24].

These boilers are usually used for low steam capacities from 100 kg to 30 tons per hour or spa production. According to construction limitations, this type of steam boiler, made of thin-walled cylinders and considered a pressure vessel, cannot withstand pressures higher than 30 bar.

2.2.1. Classification of fire tube boilers

2.2.1.1. Classification based on the number of smoke passes

Fire tube boilers can be classified depending on the number of passes, which refers to the number of times hot combustion gases flow through the boiler. The combustion chamber, also known as the furnace, serves as the initial passage for combustion. Subsequently, fire-eating pipes are employed as the subsequent passage. The predominant kind of fire tube boiler in this category is the three-pass boiler, characterized by using two passes of fire tubes and the discharge of exhaust gases from either the rear or upper section of the boiler.

The following types of fire tube boilers can be mentioned:

- Single-pass fire tube boilers
- Two-pass fire tube boilers
- Fire-tube return steam boilers
- Three-pass fire tube boilers
- Four-pass fire tube boilers

Three-pass fire tube boilers are used with efficiencies of up to 86%. This type of boiler directs the fire through three reciprocating paths to transfer the energy from the hot gases to the water around the pipes. Figure 3 shows a view of a three-pass fire tube boiler.

Fig. 3. A view of three pass fire tube boiler

2.2.2. Classification based on the number of smoke passes

These steam boilers are divided into two categories: wet back and dry back.

2.2.2.1. Wetback

In the Wetback type, the hot gas circulation area is covered by water. Therefore, the hot exhaust gases in this case, unlike the dry back model, do not collide with the refractory materials installed at the end of the boiler. As a result, Wetback type boilers have fewer maintenance issues and offer higher efficiency.

Figure 4 shows wetback type fire tube boiler.

2.2.2.2. Dry back fire tube boilers

This type of boiler has a lower efficiency than the backfire tube boilers. Considering the performance of this model of steam boilers, their maintenance and repair costs are relatively higher than other types. It should be noted that the construction of this type of steam boiler is less popular due to the complexity of maintenance and repair and the need for sufficient efficiency. Fig. 5 shows a dry back-type fire tube boiler.

2.3. Fire tube boiler limitations

The use of solid fuels in fire tube boilers causes the formation of unwanted deposits due

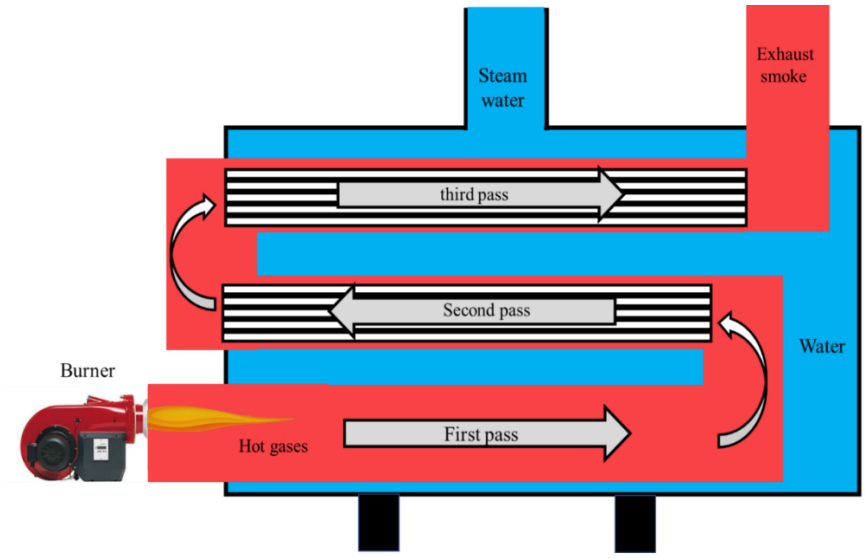


Fig. 4. Wetback type fire tube boiler

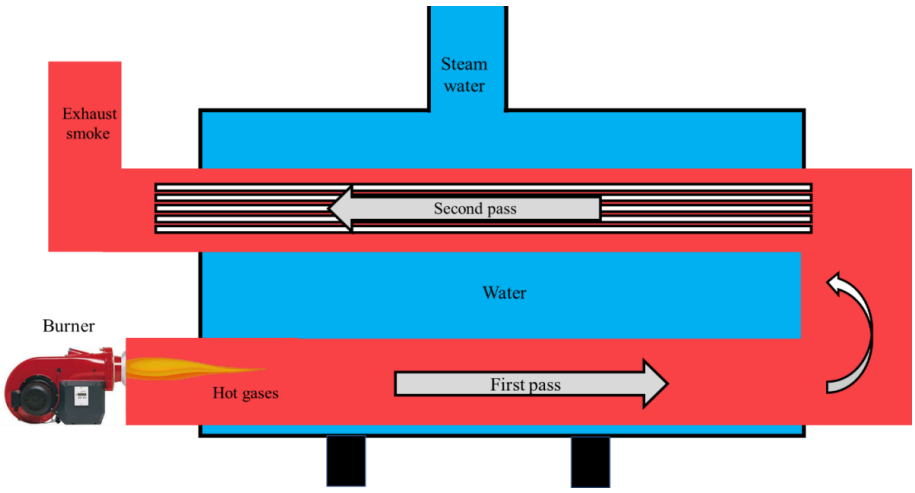


Fig. 5. Dry back type fire tube boiler

to combustion in the inner wall of the tubes, which brings many problems in the heat transfer performance. The coefficient of heat transfer in fire tube boilers where heat transfer occurs by gas flow inside the tubes is lower than when the gas flows outside the tubes. Therefore, fire tube boilers are more significant than water tube boilers in the same capacity. Due to the large volume of water around the fire tubes in these boilers, it takes more time to start them up and enter the line than water tube boilers.

Due to the lack of a drum (a container to separate water and steam) in these systems, the quality of steam produced could be better. Fire tube boilers are not used in industrial applications that require high-quality steam.

Water tube boilers are usually more suitable in industries that need to produce superheated steam. Although superheaters can also be installed in fire tube boilers, their available space is limited, and they can be installed only between the passes of fire tubes.

The thickness of the furnace wall (corrugated wall) at high pressures should be increased, as it becomes challenging to create waves on their body as this thickness increases. These waves control the thermal expansion between the furnace and the smaller tubes in the second and third passes. It should be noted that the plate containing the pipes is fixed at the end of the stove. Significant stresses are applied to the pipes and the plate containing them without this flexibility during contraction and expansion. Also, the number of pipe passes depends on the manufacturer's opinion according to the mentioned limitations (usually, three or four passes are used).

2.4. Superheater

Superheaters are used to raise the temperature of steam to a temperature higher than the saturation temperature of the fluid. The steam usually enters the superheater in a dry and saturated form. The saturated water vapor goes from the steam drum to the inlet valve of the superheater through the pipes [29].

For example, in a three-pass fire tube boiler, the steam collected from the boiler enters the converters, which include a group of U-shaped pipes, and on the other hand, the smoke from combustion enters the back area of the boiler from the end of the second pass. It transfers heat directly with the superheater, and the saturated steam turns into superheated steam. Superheaters are divided into two general categories based on the superheating process method. The first category is superheated by the flow of combustion gases of the steam boiler, and the other category is superheaters that perform the superheating process independently of the steam boiler. Also, superheaters are divided into displacement, radiation, and combined types regarding heat transfer. In combined superheaters, the constant temperature at different steam flow rates makes it more efficient in terms of working efficiency. This type of superheater is also called multi-stage. Figure 6 shows a schematic of the three-pass boiler with an internal superheater.

2. 5. Machine learning

This article discusses eight machine learning methods: Bayesian ridge regression, Gaussian process regressor, random forest, linear regression, support vector regression, decision trees, K-nearest neighbors, and gradient boosting. These methods can be broadly categorized into supervised and unsupervised learning. Supervised learning algorithms use labeled data to train a model for prediction, while unsupervised learning algorithms find patterns in unlabeled data. Linear regression is a fundamental supervised learning method that models the relationship between a dependent variable and one or more independent variables. It is a versatile technique that can be extended to handle more complex relationships. Support vector regression (SVR) is another supervised learning method that excels at handling nonlinear data. It achieves this by mapping the data into a higher-dimensional space and finding a hyperplane that fits the data with minimal error. Decision trees and K-nearest neighbors (KNN) are examples of algorithms that can be used for classification and regression tasks. Decision trees make predictions by recursively partitioning the feature space into smaller regions. KNN classifies or predicts the target variable by identifying the K nearest data points and aggregating their labels or values—ensemble methods like random forest and gradient boosting combine multiple weak learners to create a more robust model. A random forest

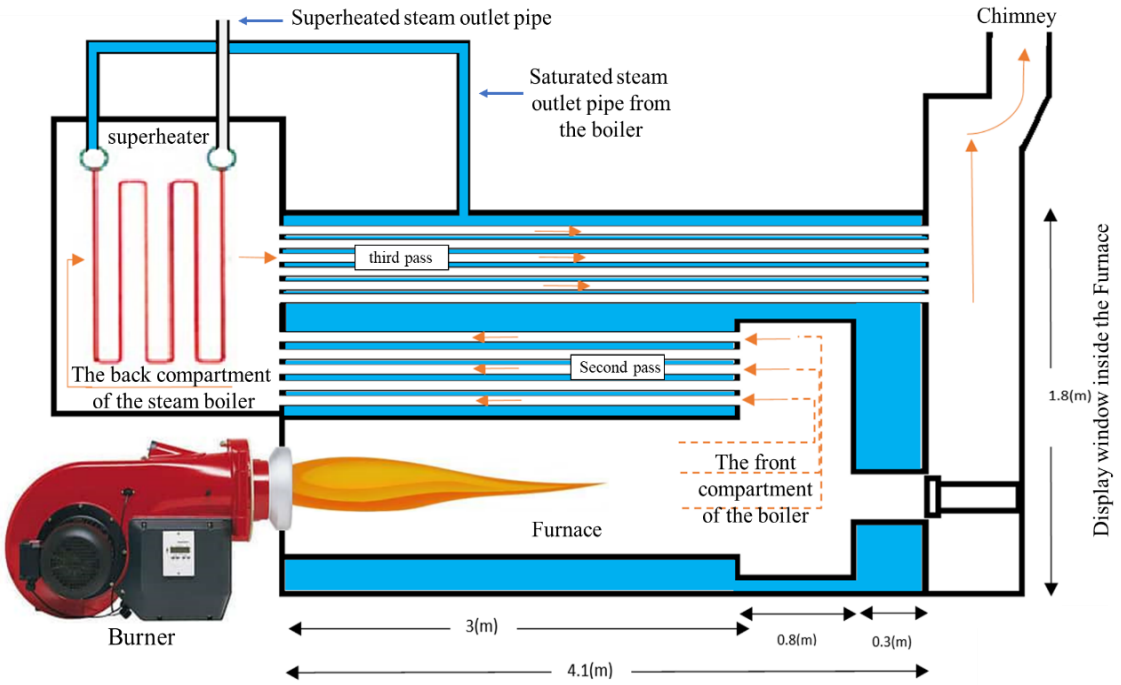


Fig. 6. Schematic of three-pass boiler with internal superheater

is a collection of decision trees trained on different bootstrap samples of the training data. Gradient boosting operates by incrementally enhancing the performance of a weak learner, concentrating on the errors made by the preceding model. These powerful tools can handle complex data and achieve high accuracy. The following section will discuss all these methods in detail.

3. Methodology

According to the explanations given in the first chapter about the fire tube boiler, this chapter examines the general design process in three sections: the furnace, the second and third passes, and the superheater.

3.1. Furnace

The burner provides controlled conditions to convert fuel energy into thermal energy continuously and safely in the boiler. The burner must control the mixture of fuel and air so that complete and stable combustion occurs. Burners are classified into single-stage, two-stage, and continuous types according to the kind of operation, and according to the type of fuel consumed, they are classified into gas, diesel, coal, and multi-fuel types [30].

In the boiler, the fire tube of the furnace is the first place for heat transfer of combustion products with water.

The furnace inlet temperature must first be calculated using fuel analysis to calculate the mass flow rate of combustion products ($\dot{m}_g$). The analysis of methane (CH4) municipal gas burning is shown in Eq. (1).

$$CH_4 + 2O_2 \rightarrow Co_2 + 2H_2O \quad (1)$$

The schematic of the methane fuel combustion equation is shown in Fig. 7.

Figure 8 is used for naming convenience.

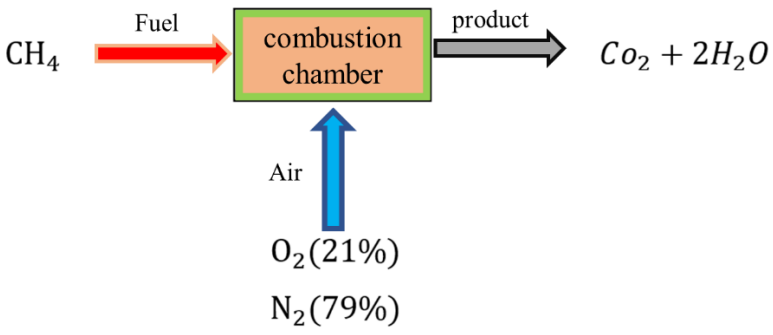


Fig. 7. Molecular schematic of combustion

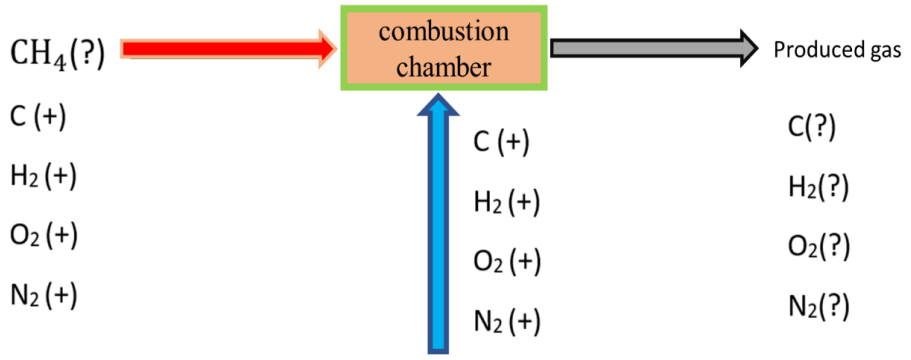


Fig. 8. Atomic schematic of combustion

Selecting one mole of CH4 as the basis reduces the number of unknowns to six parameters. For example, after the calculation, Table 1 and Table 2 present the amount of incoming air required for complete combustion, considering 30% additional air based on one mole of CH4 and two moles of O2.

Considering the 30% excess air, Eq. (2) will use the combustion equation.

$$CH_4 + 2.6O_2 + 9.78N_2 \rightarrow Co_2 + 2H_2O + 0.6O_2 + 9.78N_2 \quad (2)$$

To calculate the flame's adiabatic temperature, assuming the flow's stability and ignoring heat losses, the Equation for burning city gas with 30% excess air Eq.(3) is written. Then, by writing the enthalpy of the difference between reactants and products and interpolating from the thermodynamic tables, the output temperature of the combustion products from the burner is calculated, which is the same as the furnace inlet temperature ($(T_{gin.1})$).

R represents reactants, and P represents products.

$$\sum_{R}(nMh_f) = \sum_{p}(nMh_f) \quad (3)$$

Now, according to the above equality and interpolation from thermodynamic tables, the average temperature of the products will be equal to 1850 ℃.

The mass flow rate of combustion products ($\dot{m}_g$) is obtained directly by writing Eq. (4) for combustion products and water.

$$\dot{m}_w h_{fg} + \dot{m}_w cp_w (T_{win} - T_{woat}) = \dot{m}_g cp_g (T_{gin.1} - T_{gout.3}) \quad (4)$$

The temperature of the smoke coming out of the chimney must always be higher than the temperature of the steam coming out of the boiler.

Table 1. Number of moles of air in combustion

Description	Air	O_2	N_2
Required (mol)	9.25	2.00	7.52
Additional (mol)	2.856	0.6	2.26
Total (mol)	12.38	2.60	9.78

Table 2. The molar percentage of reactants and methane fuel combustion products.

Component	Mol	Percent
CH_4	0	0
CO_2	1	7.47
H_2O	2	14.95
O_2	0.6	4.48
N_2	$\binom{2.26}{7.52}$	73.10
Total	13.38	100

With the mass flow rate of combustion products, the mass of fuel and air input to the reaction is calculated from the methane burning Equation. The mass of air is obtained by knowing the moles of air, and the mass of consumed fuel is calculated by comparing the mass of combustion products with the mass of air. The fuel consumed is determined by dividing the fuel mass by the fuel density.

There are two solutions for calculating the burner capacity:

A) The volumetric flow rate of combustion products is calculated from Eqs. (5) to (7), and then the burner capacity is calculated according to Eq. (8).

$$\dot{m}_g = \dot{m}_{H2O} + \dot{m}_{CO2} + \dot{m}_{O2} + \dot{m}_{N2} \quad (5)$$

$$\dot{m}_{CH_4} = \dot{m}_g - \dot{m}_{O2} \quad (6)$$

$$\dot{V}_{CH_4} = \frac{\dot{m}_{CH_4}}{\rho_{CH_4}} \quad (7)$$

$$\text{Burner capacity} = \text{Calorific value of city gas} \times Volume\ of\ consumed\ fuel\ (cubic\ meters) \quad (8)$$

B) Water heat received in the boiler:

In boiler calculations, there are two efficiency models: the heat loss efficiency of the boiler body, which is about 2%, and the combustion efficiency, which is calculated according to the relations (9) to (11) below.

$$LHV = \left[\sum_{R} \left(nMh_f \right)_{25^0} - \sum_{p} \left(nMh_f \right)_{25^0} \right] \quad (9)$$

$$Q = \left[\sum_{p} \left(nMh_f \right)_{T.chimney=220^0} - \sum_{R} \left(nMh_f \right)_{25^0} \right] \quad (10)$$

$$\text{Combustion efficiency} = \frac{Q}{LHV} \quad (11)$$

The burner capacity is calculated from Eq. (12). Boiler steam loss is 2 percent.

$$\text{Burner capacity} = \frac{\dot{m}_w h_{fg} + \dot{m}_w cp_w \left(T_{win} - T_{woat} \right)}{Combustion\ efficiency \times Losses\ of\ boiler\ body} \quad (12)$$

According to the BS2790 standard and the net heat received by water, the minimum diameter of the furnace is determined, and the cross-sectional area of the furnace is calculated by Eq. (13).

$$A_F = \pi \frac{D^2}{4} \quad (13)$$

According to the EN676 standard, Eq. (14) below calculates the furnace's length.

$$L_F = 0.23 \sqrt{\frac{Q_f}{10}} \quad , \quad Q_f = (kilo\ watt)\ burner\ capacity \quad (14)$$

The relationship between the maximum flame length and the furnace length is calculated empirically from Eq. (15).

$$L_{Maximum\ pipe\ length} = \frac{3}{4} L_F \quad (15)$$

Suppose the maximum length of the flame exceeds 2/3 (maximum 3/4) of the size of the furnace. In that case, the resistance difference between the absorption surfaces between the pipes and the grid in front of the combustion return chamber has led to overheating of the pipe heads at the beginning of the 2nd pass, which will cause them to burn and eventually crack. The temperature at the end of the first pass $(T_{gout})_1$ is calculated using Eq. (16) and guess and error.

$$2\pi RL\varepsilon\sigma\left((0.5 \times \left(T_{outg.1} + T_{ing.1} \right))^4 - T_{wall}{}^4 \right) = 1000 \times \dot{m}_g cp_g \left(T_{gin.1} - T_{gout.1} \right) \quad (16)$$

3.2. Boiler passes

Heat transfer inside the desired fire tube boiler is done in three passes. The first pass is the furnace, and the second and third passes include a series of pipes. It should be noted that the heat transfer in all passes is conducted in all three ways: conduction, displacement, and radiation.

The parameters that are unknown in the design of the converter are the cross-section of the second and third passes and the temperature at the end of the second pass.

To calculate the displacement heat transfer in the second and third passes, first, some tubes are assumed for each pass, and then the mass flow rate in each tube is calculated by dividing the mass flow rate of smoke in each pass by the number of tubes. Then, according to the Reynolds number, the displacement heat transfer coefficient in that pass is determined.

The diameter of the pipes is equal in the second and third passes. Designers are looking for a suitable heat transfer coefficient and minimum pressure drop in the pipes. The size of the second and third pass levels are dependent on each other, so the following method is used for the correct calculation of three unknown parameters.

The speed of smoke in the two-way pass is considered to be 30 m/s. Now, according to the following relations, first, the Reynolds number and then the displacement heat transfer coefficient are calculated using relations (17) to (21).

$$\dot{m}_{Etch\ pipe} = \frac{\dot{m}_{\text{All pipe}}}{N} \tag{17}$$

$$A_{Etch\ pipe} = \frac{\dot{m}_{Etch\ pipe}}{\rho \times V} \tag{18}$$

$$D_{Etch\ pipe} = \sqrt{\frac{4A_{Etch\ pipe}}{\pi}} \tag{19}$$

$$\mathrm{Re} = \frac{V \times D}{v} \tag{20}$$

$$h_{conv.in} = \frac{N_u \times D}{K} \tag{21}$$

The heat transfer coefficient is calculated from Eqs. (22) and (23) below.

$$\mathrm{h}_{Radiation}\mathrm{A}\left(\mathrm{T}_{\mathrm{out.g}} - \mathrm{T}_{\mathrm{in.g}}\right) = \mathrm{A}\varepsilon\sigma\left((0.5\times\left(\mathrm{T}_{\mathrm{out.g}} + \mathrm{T}_{\mathrm{in.g}}\right))^4 - \mathrm{T}_{\mathrm{wall}}{}^4\right) \tag{22}$$

$$\mathrm{h}_{Radiation} = \frac{\varepsilon\sigma\left((0.5\times\left(\mathrm{T}_{\mathrm{out.g}} + \mathrm{T}_{\mathrm{in.g}}\right))^4 - \mathrm{T}_{\mathrm{wall}}{}^4\right)}{\left(\mathrm{T}_{\mathrm{out.g}} - \mathrm{T}_{\mathrm{in.g}}\right)} \tag{23}$$

The nuclear-boiling heat transfer coefficient is calculated using the relationship below (24).

$$h_{Nuclear\ fusion} = \frac{\mu_l \times h_{fg}}{\Delta T_e} \times \left[\frac{g\times\left(\rho_l - \rho_v\right)}{\sigma}\right]^{\frac{1}{2}} \times \left[\frac{C_{p.l}\times\Delta T_e}{C_{s.f}\times h_{fg}\times pr_l^{\,n}}\right]^3 \tag{24}$$

According to the water boiling curve for $\Delta T_e \approx 20$ (°C), the optimum and the highest possible external heat transfer coefficient will be available without approaching the danger zone.

The overall heat transfer coefficient is calculated using Eq. (25) below.

$$U = \frac{\frac{1}{A_0}}{\frac{1}{h_i A_i} + R_W + \frac{1}{\mathrm{h}_{Radiation} A_0}} \tag{25}$$

Since nuclear boiling has a substantial heat transfer coefficient and a slight resistance, it is not included in calculating the overall heat transfer coefficient.

The value of the logarithmic mean temperature difference is calculated using Eq. (26).

$$\mathrm{LMTD} = \frac{\Delta T_1 - \Delta T_2}{Ln\frac{\Delta T_1}{\Delta T_2}} \tag{26}$$

Due to the unknown temperature of the outlet of the second pass ($T_{gout.2} = T_{gin.3}$), which is the same as the inlet temperature of the third pass using the trial and error method for all temperatures between the furnace outlet temperature (first pass), the chimney output value of A_2 and A_3 is calculated. For each temperature considered for the second pass output, one number of pipes for the second pass and one for the third pass will be considered, and the guess value will be changed so that the value obtained for each level is equal to the NπDL value of the same pass. Finally, the minimum value of $A_2 + A_3$ obtained for different temperatures is considered the correct answer to the problem, and the temperature at which this minimum value is received is regarded as the output temperature of the second pass.

Relations (27) and (28) are used in calculations.

$$\mathrm{U}_2 \times A_2 \times LMTD = 1000 \times \dot{m}_g cp_{g.2}\left(T_{gin.2} - T_{gout.2}\right) \tag{27}$$

$$\mathrm{U}_3 \times A_3 \times LMTD = 1000 \times \dot{m}_g cp_{g.3}\left(T_{gin.3} - T_{gout.3}\right) \tag{28}$$

3.3. Superheater

First, the amount of heat received by the superheater is calculated by Eq. (29) according to the temperature of the superheated water vapor.

$$\mathrm{Q}_{\mathrm{required}} = \dot{m}_{water} \times cp_{v_{aper}} \times \left(\Delta T_{v_{aper}}\right) \tag{29}$$

Then, according to Eq. (30), the output temperature of the superheater is calculated.

$$\mathrm{Q}_{\mathrm{required}} = \dot{m}_g \times cp_g \times \left(\mathrm{T}_{\mathrm{outg.2}} - \mathrm{T}_{\mathrm{out-superheater}}\right) \tag{30}$$

There are two methods for calculating this part.

The first method:

To calculate the number of U-shaped coils, first, the cross-section of each superheater tube using Eq. (31) and then the gas mass flow rate inside the tubes using Eq. (32) and finally, the number of U-shaped coils is calculated by the Eq. (33).

$$A = \pi \frac{D_{\text{superheater}}{}^{2}}{4} \tag{31}$$

$$\dot{m}_{\text{Each pipe}} = \rho A V \tag{32}$$

$$N_{Number\ of\ rows} = \frac{\dot{m}_g}{\dot{m}_{\text{Each pipe}}} \tag{33}$$

According to the Eq. (34), the Reynolds number is calculated.

$$Re_{in} = \frac{4 \times \dot{m}_{\text{Each pipe}}}{\mu \times \pi \times D_{\text{superheater}}} \tag{34}$$

Equation (35) calculates the Nusselt number, and Eq. (36) calculates the internal displacement heat transfer coefficient.

$$\begin{aligned} Nu_{in} &= \frac{h_{in-\text{superheater}} \times D_{\text{superheater}}}{k_{vaperwater}} \\ &= 0.023 \times \text{Re}_{in}{}^{0.8} \times pr^{0.3} \end{aligned} \tag{35}$$

$$h_{in-\text{superheater}} = \frac{Nu_{in} \times k_g}{D_{\text{superheater}}} \tag{36}$$

Grimison's Eq. (37) calculates the external displacement heat transfer coefficient.

$$Nu_{out} = \overline{\text{Nu}}_D = C_1 \times Re_{D.max}{}^{m} \quad , \quad \begin{pmatrix} N_L \geq 10 \\ 2000 \leq Re \leq 40000 \\ pr = 0.7 \end{pmatrix} \tag{37}$$

Equation (38) calculates the Reynolds number of the external flow, and Eq. (39) calculates the maximum fluid velocity in a rectangular arrangement.

$$Re_{D.max} = \frac{V_{max} \times D_{\text{superheater}}}{v_{vaper.water}} \tag{38}$$

$$V_{max} = \left(\frac{ST}{ST - D} \right) \times V \tag{39}$$

The heat transfer coefficient of the displacement of the external flow is calculated from Eq. (40) below.

$$h_{out-\text{superheater}} = \frac{Nu_{out} \times k_{vaperwater}}{D_{\text{superheater}}} \tag{40}$$

The overall heat transfer coefficient is calculated from Eq. (41).

$$U = \frac{\frac{1}{A_0}}{\frac{1}{h_{in-\text{superheater}} A_i} + R_W + \frac{1}{h_{out-\text{superheater}} A_0}} \tag{41}$$

The logarithmic mean temperature difference value is calculated using Eq. (42).

$$\text{LMTD} = \frac{\Delta T_1 - \Delta T_2}{Ln \frac{\Delta T_1}{\Delta T_2}} \tag{42}$$

Then, according to Eq. (43), the overall cross-sectional area of the superheater is obtained.

$$A_{suerheater} = \frac{\text{Q}_{\text{required}}}{U \times LMTD} \tag{43}$$

The second method:
Calculating the value of the superheater's surface uses the Zukaskas relationship, which is directly calculated as follows without calculating the heat transfer coefficient of the average displacement of the superheater's cross-section.

$$\overline{\text{Nu}}_D = C \times Re_{D.max}{}^{m} \times pr^{0.36} \times \left(\frac{pr}{pr_s} \right)^{0.25} \tag{44}$$

$$\bar{\text{h}} = \frac{\overline{\text{Nu}}_D \times k_{\text{air}}}{D_{\text{superheater}}} \tag{45}$$

$$\Delta T_m = \frac{(T_s - T_i) - (T_s - T_o)}{Ln \left(\frac{T_s - T_i}{T_s - T_o} \right)} \tag{46}$$

$$\text{Q}_{\text{required}} = N_{\text{superheater}} \times \bar{\text{h}} \times \pi \times D_{\text{superheater}} \times \Delta T_m \tag{47}$$

$$N_{\text{superheater}} = \frac{\text{Q}_{\text{required}}}{\bar{\text{h}} \times \pi \times D_{\text{superheater}} \times \Delta T_m} \tag{48}$$

$$A_{\text{superheater}} = N_{\text{superheater}} \times \pi \times D_{\text{superheater}} \times L_{\text{superheater}} \tag{49}$$

3.4. Methods of machine learning

3.4.1. Bayesian ridge regression (BRR)

Linear regression is a workhorse for making predictions, but it doesn't account for the inherent uncertainty in those predictions. Enter Bayesian Ridge Regression (BRR). BRR injects probability theory (Bayesian) into the mix. It treats the influence of each variable as a range of possibilities (not just a single number) based on what we already know. This "prior knowledge"

gets refined using the actual data, resulting in a "posterior distribution" that reflects the true range of possibilities. This probabilistic approach lets us understand how reliable the predictions from the model actually are.

Traditional regression can struggle with data where features are too similar or when the model becomes overly specific to the training data (overfitting). BRR tackles these issues head-on. It introduces a penalty term that shrinks the influence of each variable, making the model less sensitive to specific data points and leading to more generalizable predictions. Additionally, BRR allows you to control the level of certainty in the initial probability ranges. This flexibility lets BRR adapt to complex datasets, making it a powerful tool for reliable predictions in fields like machine learning, finance, and medicine [31].

3.4.2. Gaussian Process Regressor (GPR):

GPR is a nonparametric, Bayesian approach to regression defining a probability distribution over functions rather than random variables, so Eq. (50) is solved by:

$$f(x) \sim GPR\left(m(x), k(x, x\text{^}')\right) \tag{50}$$

where m(x) is the mean and $k(\mathbf{x}, \mathbf{x}')$ is the covariance function. Please note that, as defined above, GPR does not require learning the parameters of the regression function f(x, θ) in a traditional sense. The mean and covariance are defined by:

$$m(\mathbf{x}) = E\left[f(\mathbf{x}) \right] \tag{51}$$

$$k(\mathbf{x}, \mathbf{x}') = E\left(f\left(\mathbf{x} - m(\mathbf{x})\right)\left(f\left(\mathbf{x}' - m(\mathbf{x}')\right)\right)\right] \tag{52}$$

GPR assigns a prior probability to every possible function, with higher probabilities given to functions that the algorithm considers more likely; for example, these may be smoother than other functions. As detailed in the ref [32], we implement the standard radial basis kernel, which also explains the algorithm mathematically. Other kernel options exist, but we do not explore the effect of kernel choice on algorithm performance here.

3.4.3. Random Forest (RF):

RF is a collection of constructed decision trees that sequentially conduct binary data splits to produce a homogeneous subset. For a comprehensive explanation of the algorithm, refer to ref. [33]. We adopt a bagging approach where the ensemble members are trained on different bootstrap samples of the training set, and we set the number of decision trees in the forest to 1,500. The variability of the predictions estimated by the RF has been investigated based on ref. [34], where the confidence interval's variance was obtained using the bootstrap replicates used to train the RF itself.

3.4.4. Linear regression:

Linear regression is a foundational methodology in machine learning used to model the relationship between a dependent variable (y) and one or more independent variables (x). The process typically begins with clearly formulating the problem at hand and collecting relevant data, ensuring its quality and representativeness. Preprocessing steps involve handling missing values, scaling features, and potentially engineering new ones to better capture underlying patterns.

Following data preparation, the dataset is split into training and testing sets. The training set is utilized to train the linear regression model, which involves initializing parameters and defining a cost function to measure the disparity between predicted and actual values. Optimization algorithms like Gradient Descent iteratively adjust parameters to minimize this cost function.

Subsequently, the trained model undergoes evaluation on both the training and testing sets to gauge its performance, utilizing metrics such as R-squared and Mean Absolute Error. Hyperparameters are fine-tuned based on evaluation results to enhance the model's predictive ability. Once satisfactory performance is achieved, the model can be deployed for making predictions on new data.

Continuous monitoring and maintenance are crucial post-deployment to ensure the model's ongoing accuracy and relevance. Linear regression's simplicity belies its versatility; it serves as a building block for more intricate models and can be extended to accommodate various complexities, such as polynomial features or regularization techniques like Ridge and Lasso regression. Overall, its methodology provides a structured approach to deriving insights and making predictions from data.

Linear regression is a type of supervised machine learning algorithm that predicts a continuous target variable based on one or more independent variables. It assumes a linear relationship between the dependent and independent variables and uses a linear equation to model this relationship. The general form of the linear Eq. (53) is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \grave{o} \tag{53}$$

where y is the dependent variable, $x_1,x_2,\ldots,x_n$ are the independent variables, $\beta_0,\beta_1,\beta_2,\ldots,\beta_n$ are the coefficients, and ϵ is the error term. The coefficients are the parameters that need to be learned from the data, and the error term represents the deviation of the observed value from the true value.

3.4.5. Support Vector Regression (SVR):

SVR is a robust machine-learning approach for regression tasks. It extends the principles of Support Vector Machines (SVM) to predict continuous outcomes.

Its methodology involves several key steps: preprocessing the data to handle missing values and scale features, selecting or engineering relevant features, choosing an appropriate kernel function (such as Linear, Polynomial, RBF, or Sigmoid), tuning hyperparameters like C and epsilon through techniques like grid search, training the model to identify support vectors that lie closest to the decision boundary, minimizing error with a loss function that balances prediction accuracy and margin maximization, and evaluating the model's performance using metrics like MSE or R2. Through these steps, SVR optimizes the margin between predicted and actual values while leveraging support vectors to capture complex relationships in the data. It is a versatile tool for regression tasks with nonlinear patterns, albeit requiring careful parameter tuning for optimal results.

SVR stands for support vector regression, a supervised learning algorithm that predicts a continuous output variable based on one or more input variables. It uses a kernel function to map the input variables to a higher-dimensional space, where it tries to find a hyperplane that fits the data with minimal error. The general form of the SVR equation is:

$$f(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b \tag{54}$$

where f(x) is the predicted output, x is the input vector, xi is the support vectors, are the coefficients, K(x, xi) is the kernel function, and b is the bias term. The coefficients and the bias term are the parameters that need to be learned from the data, and the kernel function determines the similarity between the input vectors. There are different types of kernel functions, such as linear, polynomial, radial basis function (RBF), or sigmoid. The choice of kernel function depends on the data's characteristics and the task's complexity.

3.4.6. Decision trees

Decision trees are a fundamental machine-learning technique for classification and regression tasks. The methodology of decision trees involves recursively partitioning the feature space into smaller regions, guided by the values of input features, to make decisions. The tree structure comprises nodes representing features, edges indicating decisions or outcomes, and leaves containing the final predictions. The algorithm selects the best feature and split point at each step based on specific criteria, commonly Gini impurity or entropy for classification and variance reduction for regression. This process continues until a stopping criterion is met, such as reaching a maximum depth, minimum samples per leaf, or no further reduction in impurity. Decision trees are interpretable, easily visualized, and capable of handling numerical and categorical data, making them versatile tools in machine learning.

Furthermore, decision trees can suffer from overfitting, especially when the tree grows too deep, or the dataset contains noise. Techniques like pruning, which involves removing branches of the tree that provide little predictive power, can mitigate overfitting and improve generalization. Ensemble methods like Random Forests and Gradient Boosting Machines leverage decision trees to build more robust models by aggregating predictions from multiple trees. These ensemble methods further enhance performance by reducing variance and improving predictive accuracy. Despite their simplicity, decision trees and their variants remain widely used in various applications due

to their interpretability, flexibility, and effectiveness in handling complex datasets.

3.4.7. K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a straightforward yet powerful algorithm for machine learning classification and regression tasks. The methodology of KNN revolves around the concept of similarity: it classifies or predicts the target variable by identifying the K nearest data points in the feature space and aggregating their labels or values. In classification, KNN assigns the most frequently occurring class among the K neighbors to the query point, while in regression, it computes the average or weighted average of the target values of those neighbors. The choice of K, the number of neighbors, is crucial and typically determined through cross-validation or domain knowledge, balancing between bias and variance. KNN operates assuming that similar instances will likely belong to the same class or exhibit similar behaviors, effectively handling nonlinear decision boundaries and locally varying data distributions.

However, the effectiveness of KNN can be influenced by the choice of distance metric and the curse of dimensionality, where the feature space becomes increasingly sparse in high-dimensional spaces, leading to degraded performance. Additionally, KNN's computational complexity grows linearly with the size of the training data, making it less efficient for large datasets. Despite these limitations, KNN remains famous for its simplicity, ease of implementation, and ability to handle nonparametric data without assuming a specific underlying distribution, making it a valuable tool in various machine-learning applications. Regularization techniques, such as feature scaling and dimensionality reduction, can alleviate some drawbacks and enhance performance in real-world scenarios.

K-Nearest Neighbors (KNN) is a supervised learning algorithm that predicts the output of a new data point based on the similarity of its features to the features of its k-closest neighbors in the training data. The formula of KNN depends on the distance metric used to measure the similarity between the data points. One of the most common distance metrics is the Euclidean distance, the straight-line distance between two points. The Euclidean distance between two points x and y with n features is given by:

$$d(x,y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \tag{55}$$

3.4.8. Gradient Boosting:

Gradient Boosting is a powerful ensemble learning technique used for classification and regression tasks in machine learning. The methodology of Gradient Boosting involves iteratively improving the performance of a weak learner, typically decision trees, by focusing on the errors made by the previous model. It works by sequentially adding new models to the ensemble, each trained to correct its predecessors' mistakes. During training, the algorithm calculates the gradients of a loss function concerning the predictions. Then, it fits a new model for these gradients, effectively minimizing the errors made by the ensemble. By combining multiple weak learners, Gradient Boosting produces a robust predictive model capable of capturing complex relationships within the data. Critical components of Gradient Boosting include the learning rate, which controls the contribution of each new model, and regularization techniques like tree depth and shrinkage, which prevent overfitting and improve generalization performance.

Despite its effectiveness, Gradient Boosting can be computationally expensive and sensitive to hyper parameters, requiring careful tuning to achieve optimal results. However, its ability to handle heterogeneous data types, deal with missing values, and provide feature importance estimates makes it a popular choice in various real-world applications. Furthermore, advancements such as XGBoost, LightGBM, and CatBoost have enhanced the efficiency and scalability of Gradient Boosting algorithms, enabling them to handle large-scale datasets and outperform other machine learning techniques regarding predictive accuracy. Gradient Boosting remains a versatile and widely used method in the machine learning toolkit, capable of producing highly accurate models with relatively little data preprocessing.

Gradient boosting is a supervised learning algorithm that combines several weak learners (such as shallow trees) into a strong learner by training each new model to minimize the loss function of the previous model using gradient

descent. The formula for gradient boosting depends on the type of problem (classification or regression), the loss function (such as mean squared error or cross-entropy), and the base learner (such as decision trees or linear models). One of the most common implementations of gradient boosting is gradient-boosted trees, where the base learner is a decision tree. The formula of gradient-boosted trees for regression problems is:

$$f(x)=\sum_{m=0}^{M}\gamma_m h_m(x) \tag{56}$$

where f(x) is the predicted output, x is the input vector, γ_m is the number of trees in the ensemble, $h_m(x)$ is the prediction of the m-th tree, and is the learning rate that controls the contribution of each tree. The learning rate is a hyper parameter that needs to be tuned to avoid overfitting or under fitting. The trees are trained sequentially, and each tree tries to fit the residual errors of the previous tree. The residual errors are computed as the negative gradient of the loss function concerning the predicted output. For example, if the loss function is the mean squared error, the residual errors are:

$$r_{im}=-\frac{\partial L\left(y_i, f(x_i)\right)}{\partial f(x_i)}=y_i-f(x_i) \tag{57}$$

where is the actual output, and f(xi) is the predicted output of the previous ensemble. The m-th tree is then fitted to the residual errors, and the tree predictions are added to the ensemble with a learning rate. The process is repeated until a stopping criterion is met, such as the maximum number of trees or the minimum improvement in the loss function.

3.5. R-squared (R2)

R-squared (R2) is a statistical metric that quantifies the proportion of the variability in the dependent variable that can be explained by the independent variables in a regression model. The metric quantifies the extent to which the independent factors account for the variability seen in the dependent variable [35]. R2 quantifies the degree to which the regression model accurately represents the observed data. The scale spans from zero to one, with zero being the lowest value and one representing the highest value. When the value of R2 is equal to zero, the regression model fails to account for any variation in the response data around its average value. When the value of R2 is equal to one, the regression model accurately accounts for the variation in the response data around its average value. Within the regression analysis framework, a more excellent R2 value signifies a superior alignment between the model and the data. Nevertheless, it is essential to exercise caution when interpreting R2 and utilize it in conjunction with other metrics to evaluate the overall performance and validity of the regression model [36], [37].

R2 is also known as the coefficient of determination. It is calculated as follows:

$$R^2=1-\frac{SS_{res}}{SS_{tot}} \tag{58}$$

SS_{res} is the sum of squares of residuals (the difference between the observed and predicted values), and SS_{tot} is the total sum of squares (the difference between the observed values and the mean of the dependent variable).

R2 measures the proportion of the total variation in the dependent variable that is explained by the independent variables in the regression model.

3.6. Mean Squared Error

Mean Squared Error (MSE) is a widely employed measure for assessing the effectiveness of regression models. The metric calculates the mean of the squared differences between the actual and anticipated values in the model [38]. Mathematically, MSE is calculated as the average of the squared differences between the predicted values (y_i) and the actual values ($\hat{y}_i$)

$$MSE=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2 \tag{59}$$

N is the number of samples or data points, y_i is the actual value of the dependent variable, and $\hat{y}_i$ is the predicted value of the dependent variable.

MSE, or mean squared error, measures the average of the squared differences between projected values and actual values. A lower mean squared error (MSE) signifies that the model's predictions are more closely aligned with the actual values, indicating superior performance of the model[39]. It should be emphasized that the Mean Squared Error (MSE) is not adjusted according to the scale of the dependent variable, which makes it challenging to evaluate on its own. Thus, MSE is commonly

employed alongside other measures to evaluate the overall effectiveness of regression models.

4. Results and discussion

This section presents the design results of an industrial example of a three-pass fire tube boiler with an internal superheater. The desired parameters have been calculated using MATLAB software.

The code is written in 5 sections to calculate the following parameters. These sections include the exit temperature from the furnace, finding the diameter and surface and the number of tubes of each pass, finding the nuclear-boiling coefficient, finding the surface of the superheater and the inlet temperature of the third pass, and correcting the level of the third pass. Table 3 shows information related to the design of a three-pass fire tube boiler (capacity of 5 tons per hour of steam and working pressure of 10 bar).

Table 4 shows Information about the superheater and third pass corrections for different superheated steam outlet temperatures.

Table 3. Information related to the design of a three-pass fire tube boiler (capacity of 5 tons per hour of steam and working pressure of 10 bar)

$\mathrm{T}_{\mathrm{in}_{\mathrm{water}}}(C^0)$	80
$\mathrm{T}_{\mathrm{ing.1}}(C^0)$	1850
$\mathrm{T}_{outg.1} = \mathrm{T}_{\mathrm{ing.2}}(C^0)$	1069
$\mathrm{T}_{outg.2}(C^0)$	527
$\mathrm{T}_{\mathrm{ing.3}}(C^0)$	499.82
$\mathrm{T}_{outg.3} = \mathrm{T}_{Chimney}(C^0)$	220
$\mathrm{T}_{in\_superheater}(C^0)$	180
$\mathrm{T}_{out\_superheater}(C^0)$	195
$\Delta T_{superheater}(C^0)$	15
$D_{Furnaces}\ (m)$	0.65
$D_{Tube}(m)$	0.05
$D_{superheater}(m)$	0.025
$D_{BOILER}(m)$	1.8
$L_1(m)$	4.1
$L_2(m)$	3
$L_3(m)$	4.1
$L_{superheater}(m)$	7
$L_{Rap}(m)$	0.8
$L_{Acc}(m)$	0.3
$L_{boiler}(m)$	5
N_2	71
$N_{3-N0-SUPERHEATER}$	67
$N_{3-WITH-SUPERHEATER}$	82
$N_{SUPERHEATER}$	17
$A_{Furnaces}(m^2)$	0.33
$A_2(m^2)$	33.87
$A_{3-N0-SUPERHEATER}(m^2)$	42.15
$A_{3-WITH-SUPERHEATER}(m^2)$	52.15
The percentage of increase in the level of the third pass	23.72
$A_{SUPERHEATER}(m^2)$	17
$V_{Furnaces}(\frac{m}{s})$	19.04
$V_2(\frac{m}{s})$	39.90
$V_{3-N0-SUPERHEATER}(\frac{m}{s})$	27.19
$V_{3-WITH-SUPERHEATER}(\frac{m}{s})$	22.21
$V_{SUPERHEATER}(\frac{m}{s})$	30
Burner capacity$(\frac{kcal}{h})$	4000000

Table 4. Information about the superheater and third pass corrections for different superheated steam outlet temperatures

$\Delta T_{superheater}(C^0)$	20	25	30	35	40
$T_{out\_superheater}(C^0)$	200	205	210	215	220
$T_{ing.3}(C^0)$	489	479	470	460	451
$N_{3-WITH-SUPERHEATER}$	89	96	103	112	121
$N_{SUPERHEATER}$	24	31	39	47	55
$A_{3-WITH-SUPERHEATER}(m^2)$	57	61	66	71	77
The percentage of increase in the level of the third pass	26	31	36	41	45
$A_{SUPERHEATER}(m^2)$	13	17	21	26	30
$V_{3-WITH-SUPERHEATER}(\frac{m}{s})$	21	19	18	16	15

Here, we show the graph of the results. Figure 9 shows the decrease in the inlet temperature in the third pass due to the increase in temperature in the superheater. Figure 9 compares the inlet temperature of the third pass with the Temperature increase in the superheater. Adding the superheater to the end of the third pass reduces the temperature of this area from 525 ℃ to 500 ℃.

Figure 10 shows the reduction of the steam speed in the third pass due to the increase in the superheater's temperature. According to industrial information on fire tube boilers, the velocity range of hot gases resulting from repass combustion should be between 20 m/s and 40 m/s.

According to the project's results, the boiler's optimal performance requires a temperature increase in the superheater of less than 22 ℃. However, according to the coding outputs, our final optimal value for the temperature increase in the superheater is 15 ℃. For every 5 ℃ temperature increase in the superheater, the steam velocity in the third pass tubes decreases by 1 m/s.

Figure 11 shows the decrease in the inlet temperature of the third pass in exchange for increasing the number of superheater tubes.

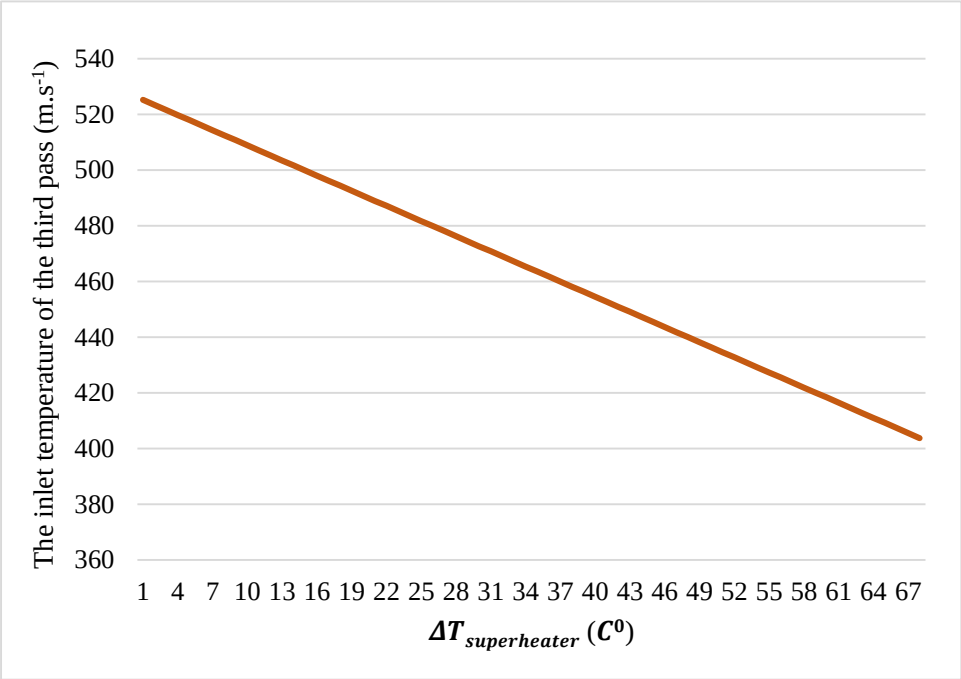


Fig. 9. Comparing the inlet temperature of the third pass with the Temperature increase in the superheater

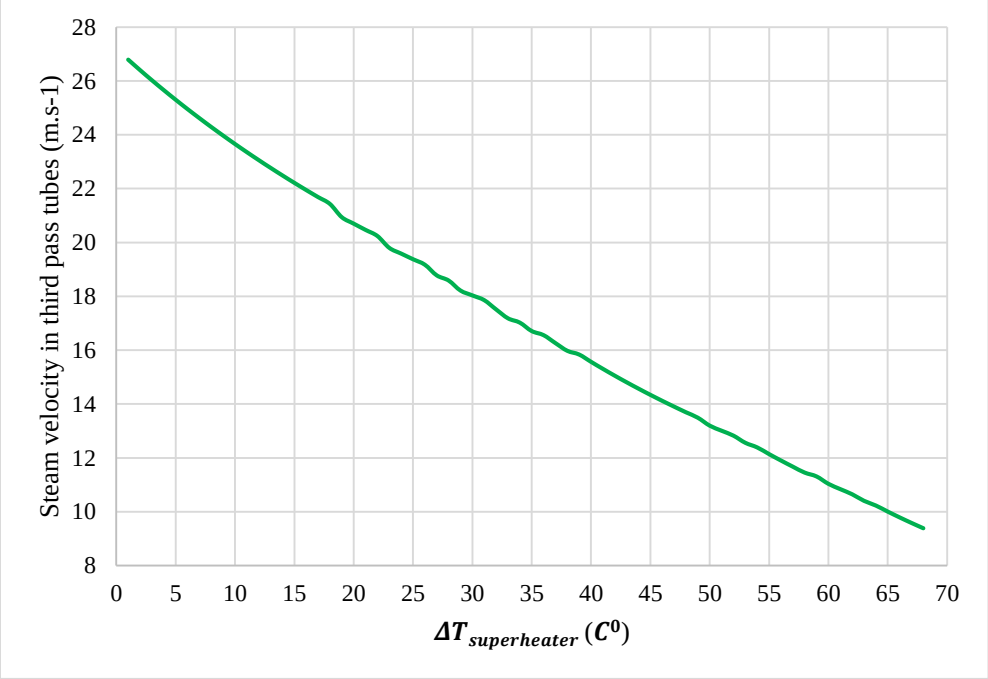


Fig. 10. Comparing the steam speed in the third pass with the increase in temperature in the superheater

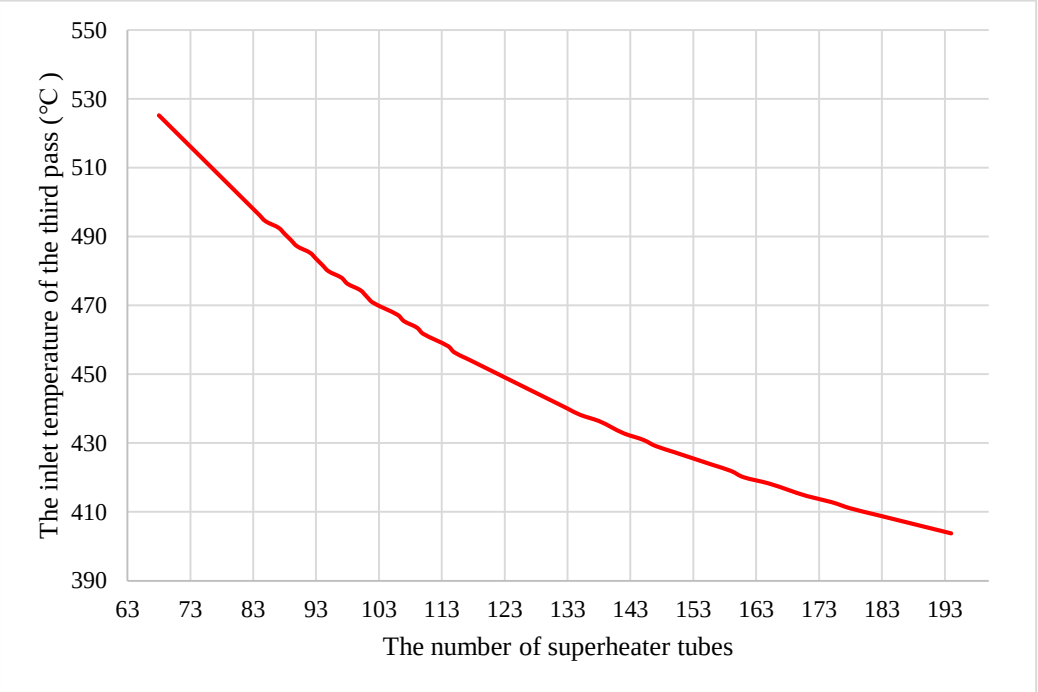


Fig. 11. Comparing the inlet temperature of the third pass with the number of superheater tubes

Figure 12 shows the graph of the increase in steam speed in the superheater by increasing the inlet temperature of the third pass. Since the optimal value of the speed in the fire tube passes is 20 to 40 m/s, according to the figure below, reducing the temperature to below 490 ℃ will cause the speed of the hot gases in the third pass to decrease to below 20 m/s, which causes a sharp decrease. The speed of hot gases in the boiler increases, and heat transfer is not done effectively, causing damage to the boiler.

Figure 13 shows the changes in some parameters by increasing the superheater temperature.

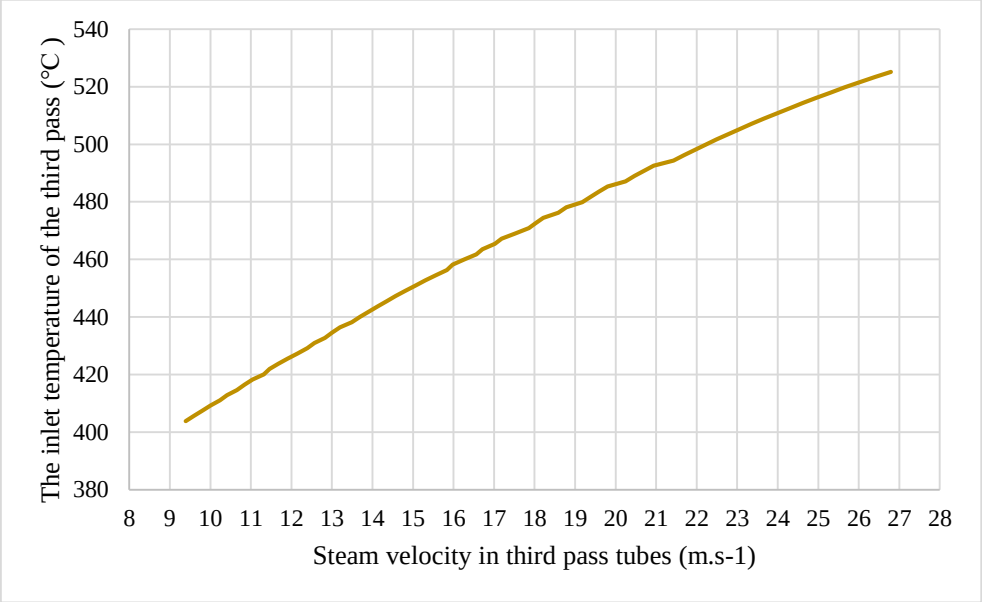


Fig. 12. Comparison of the steam speed in the superheater with the inlet temperature of the third pass

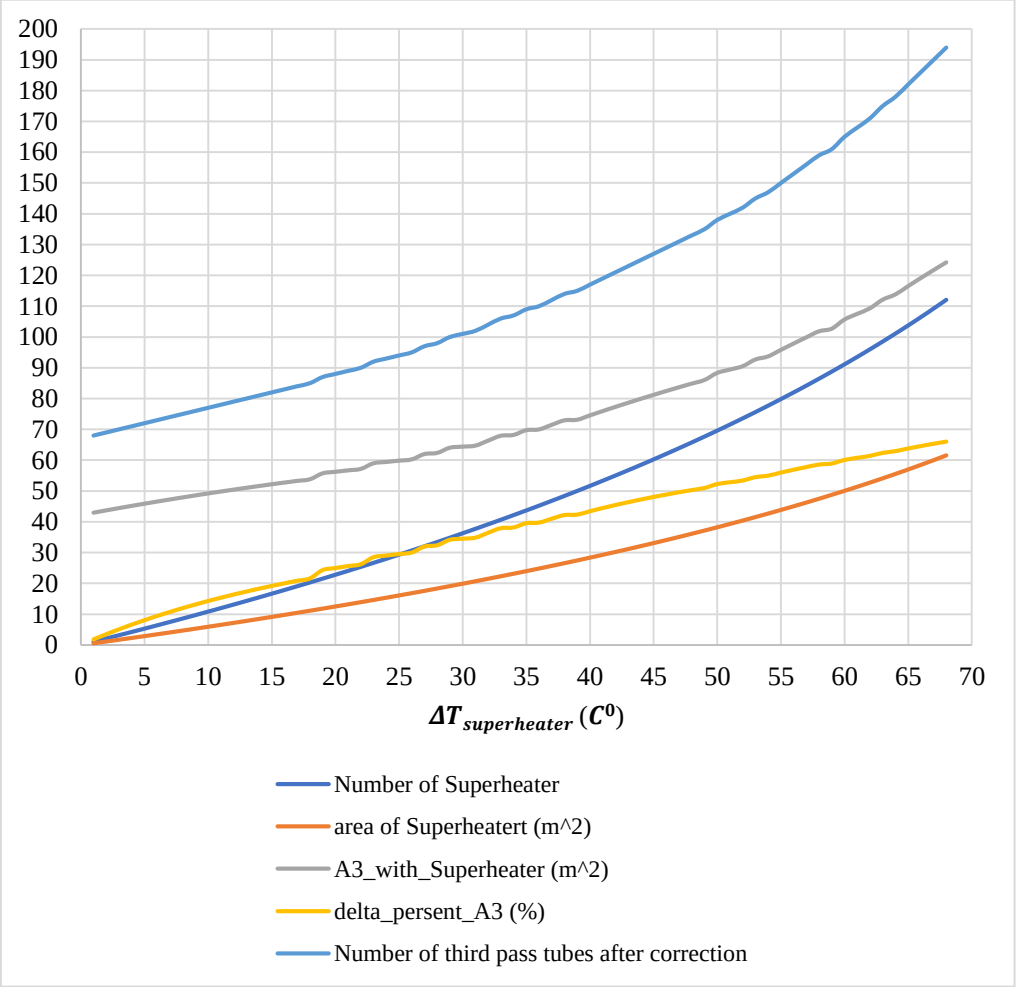


Fig. 13. Comparing the graph of changes in some parameters with the increase in superheater temperature

We consider the desired data whose dimensions are 68*8 according to the output of the superheater and boiler design code.

Now, by coding in Python software, we first select the most influential parameters among the effective parameters in calculating the temperature increase in the superheater—table 5 shows selected features with Percentage of impact coefficients.

In the machine learning stages, we consider 80% of the data as training, and we will consider the remaining 20% of the data as testing. The best model among eight machine learning models has been selected to predict the temperature increase in the superheater. Figures14 to 21 display graphs comparing the actual and expected values of the Temperature increase in the superheater ($\boldsymbol{\Delta T}$superheater).

Table 5. Selected features with Percentage of impact coefficients

Selected features	impact coefficients
$\boldsymbol{T_{gout.2}}$	51.12%
The percentage of changes in the level of the third pass	16.59%
Number of Superheater pipes	13.85%
Area of third pass with Superheater	10.82%
Area of Superheater	7.61%

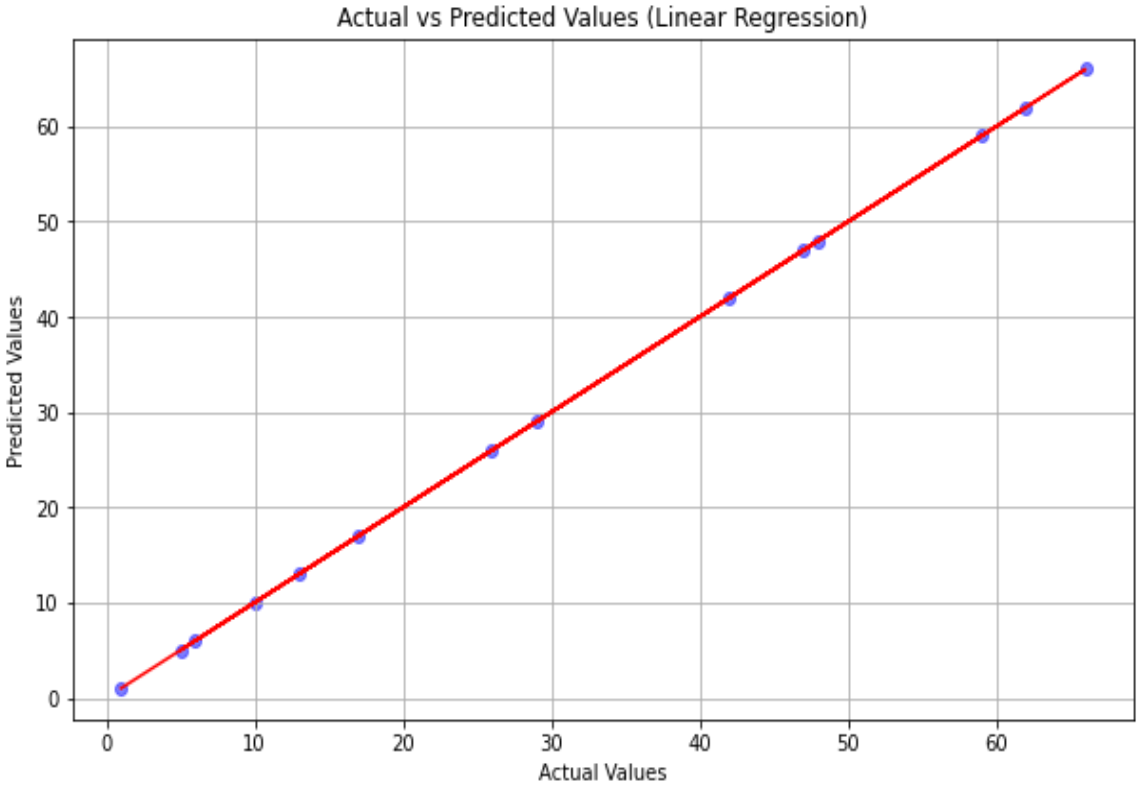


Fig. 14. Actual vs Predicted Values of $\boldsymbol{\Delta T}_{\text{superheater}}$ for Linear Regression.

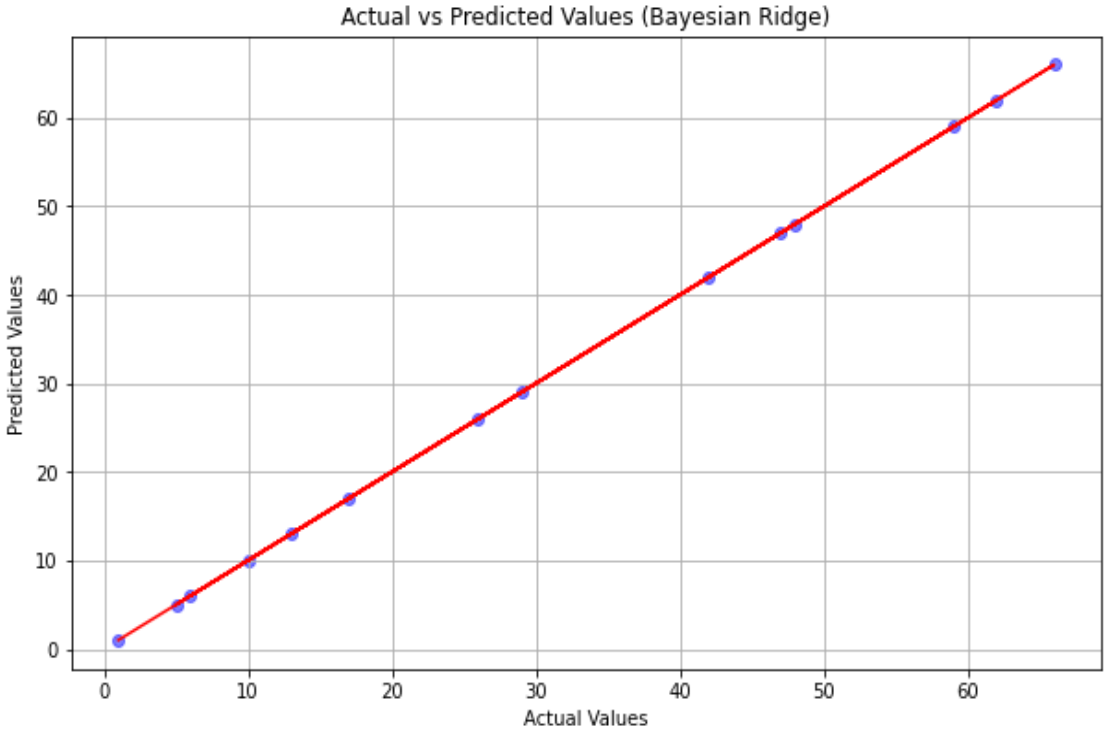


Fig. 15. Actual vs Predicted Values of $\Delta T_{\text{superheater}}$ for Bayesian Ridge.

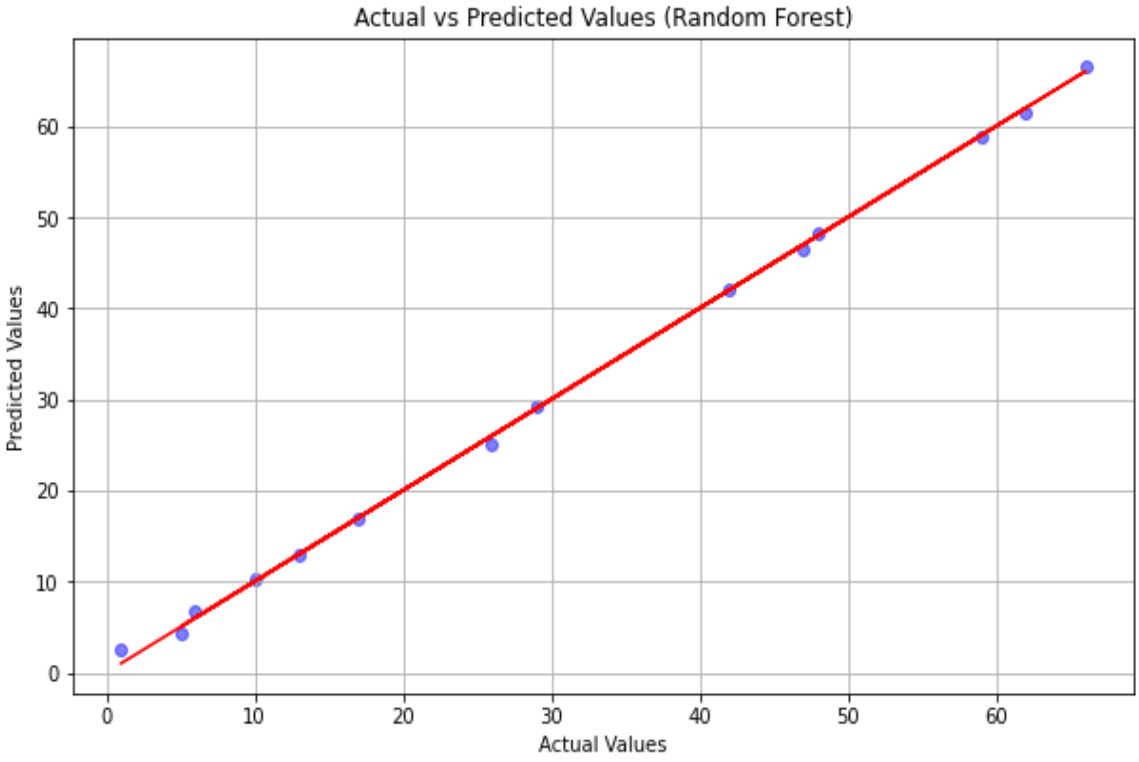


Fig. 16. Actual vs Predicted Values of $\Delta T_{\text{superheater}}$ for Random Forest.

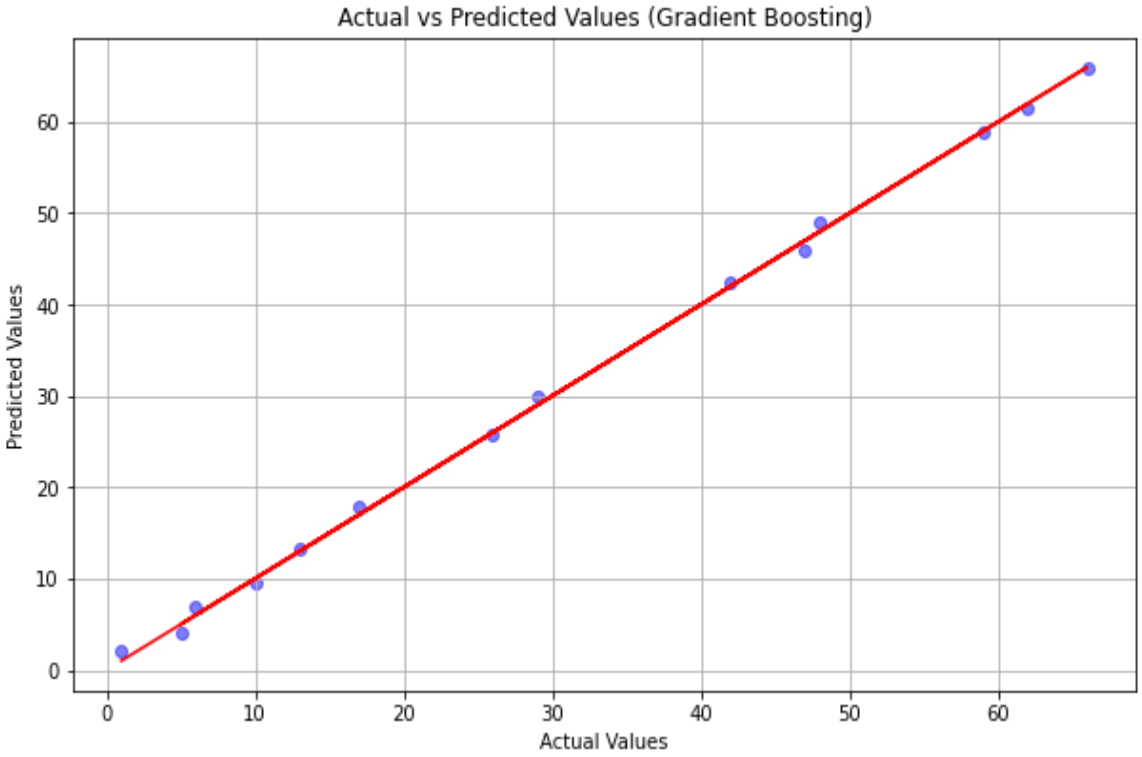


Fig. 17. Actual vs Predicted Values of $\boldsymbol{\Delta T}_{\text{superheater}}$ for Gradient Boosting.

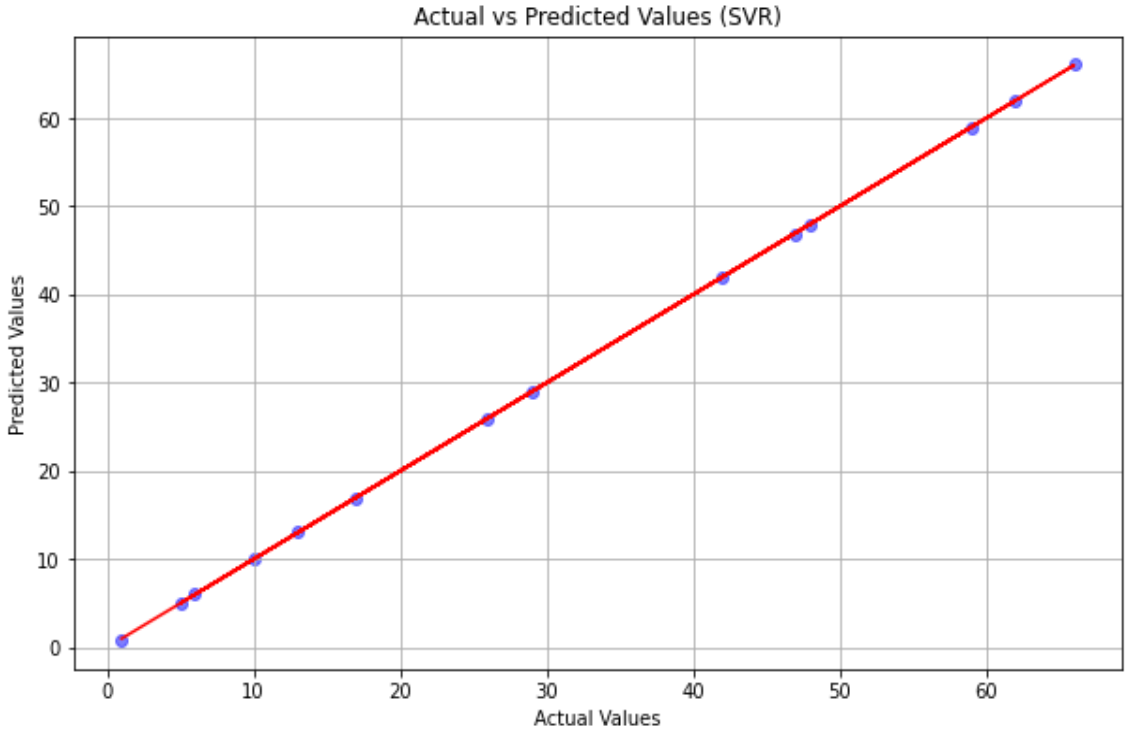


Fig. 18. Actual vs Predicted Values of $\boldsymbol{\Delta T}_{\text{superheater}}$ for SVR.

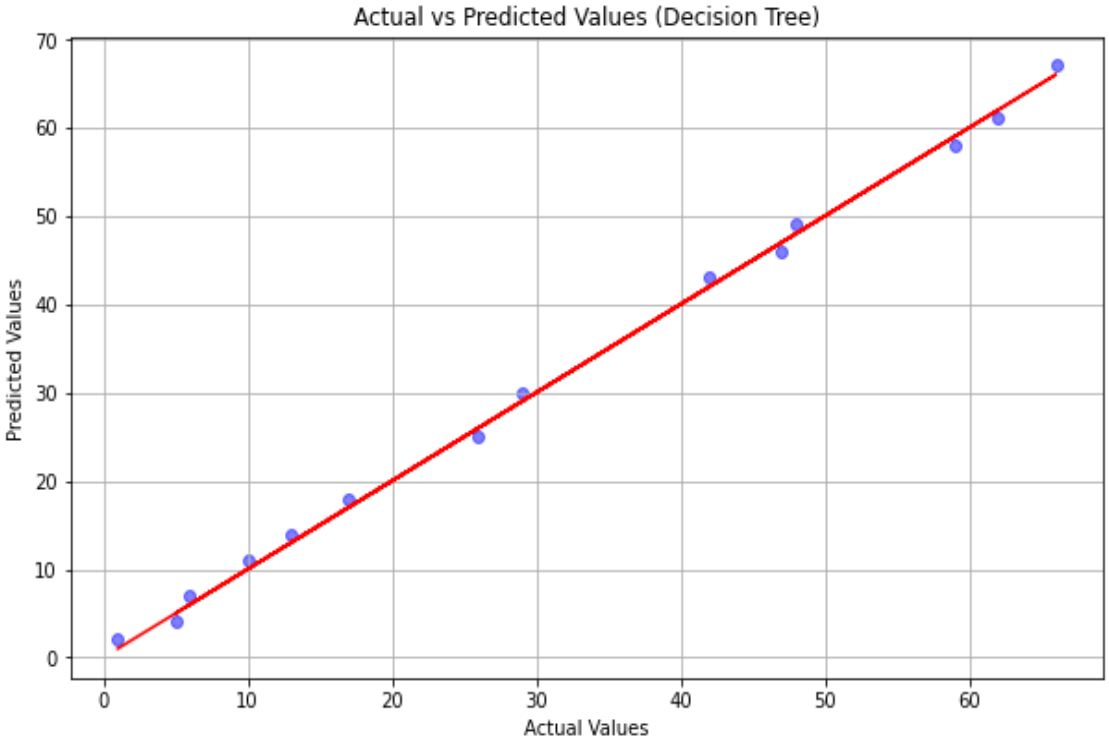


Fig. 19. Actual vs Predicted Values of $\Delta T_{superheater}$ for Decision Tree.

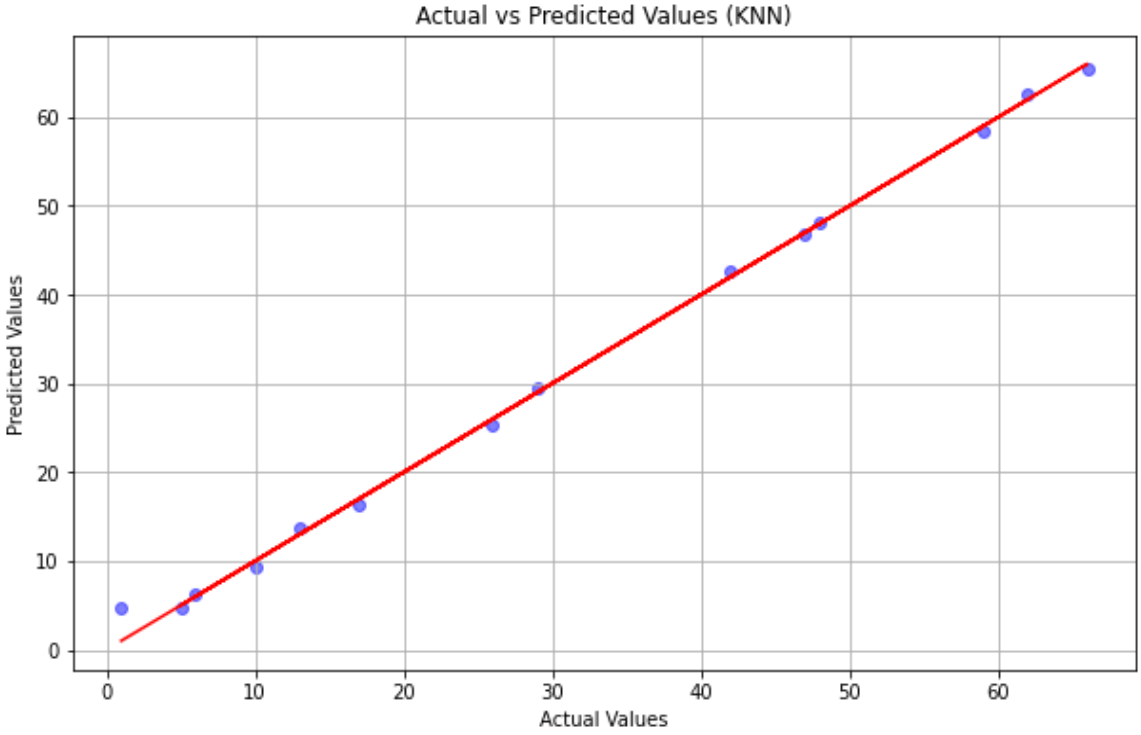


Fig. 20. Actual vs Predicted Values of $\Delta T_{superheater}$ for KNN.

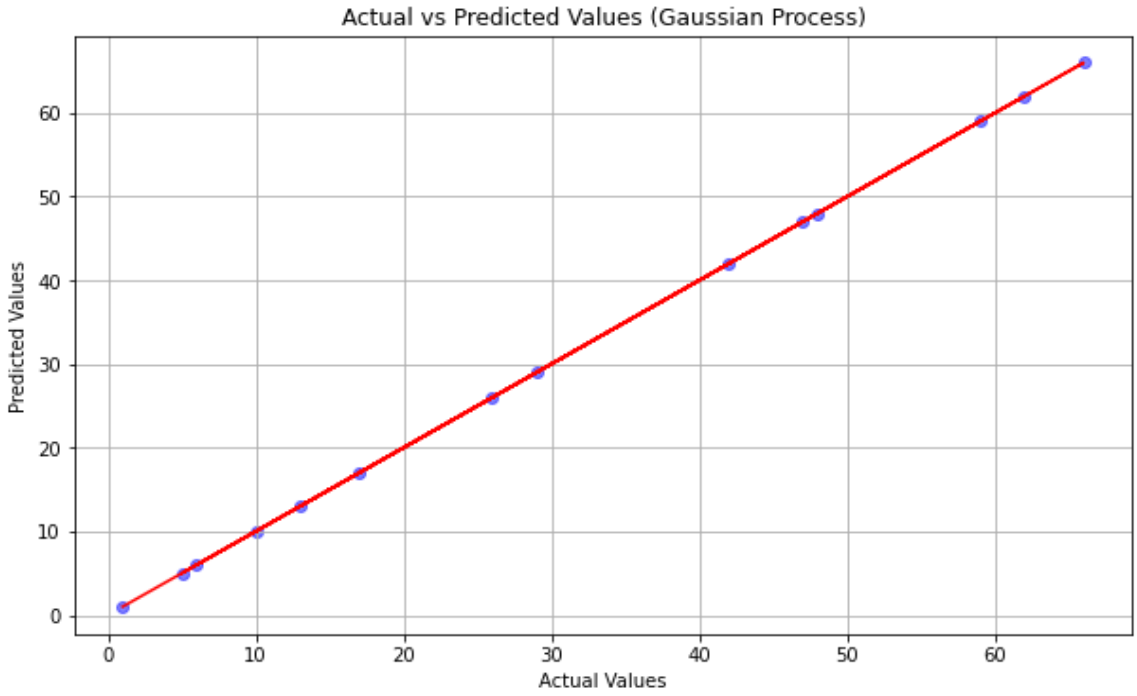
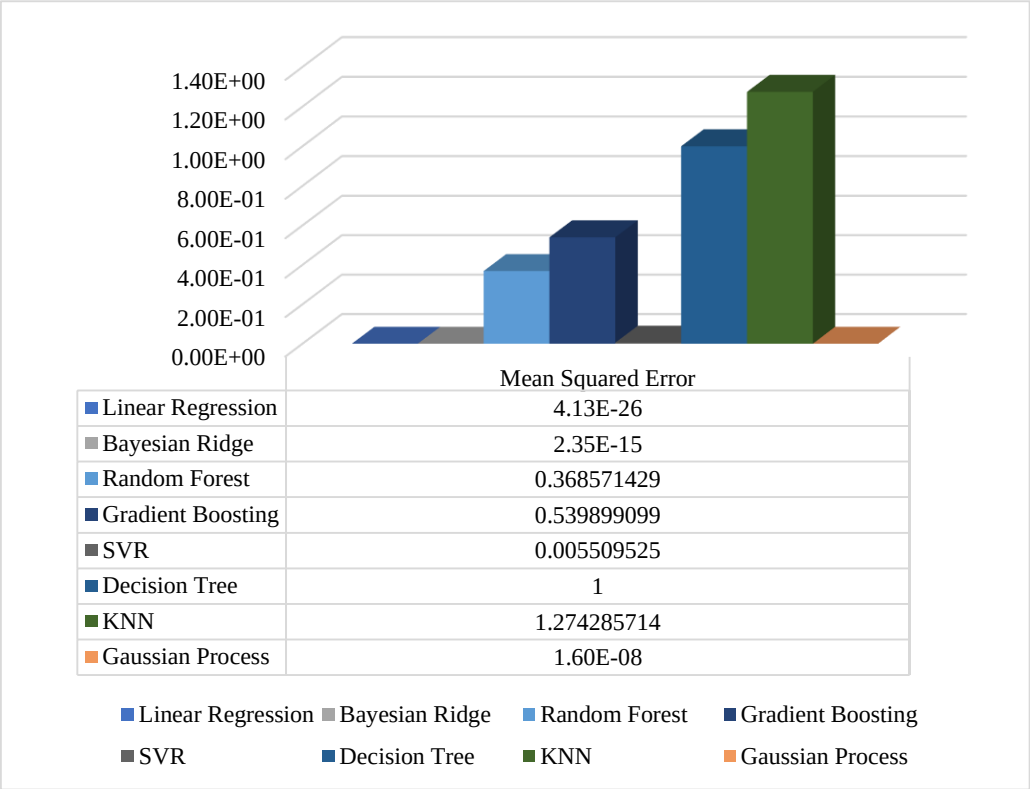


Fig. 21. Actual vs Predicted Values of $\Delta T_{superheater}$ for Gaussian Process.

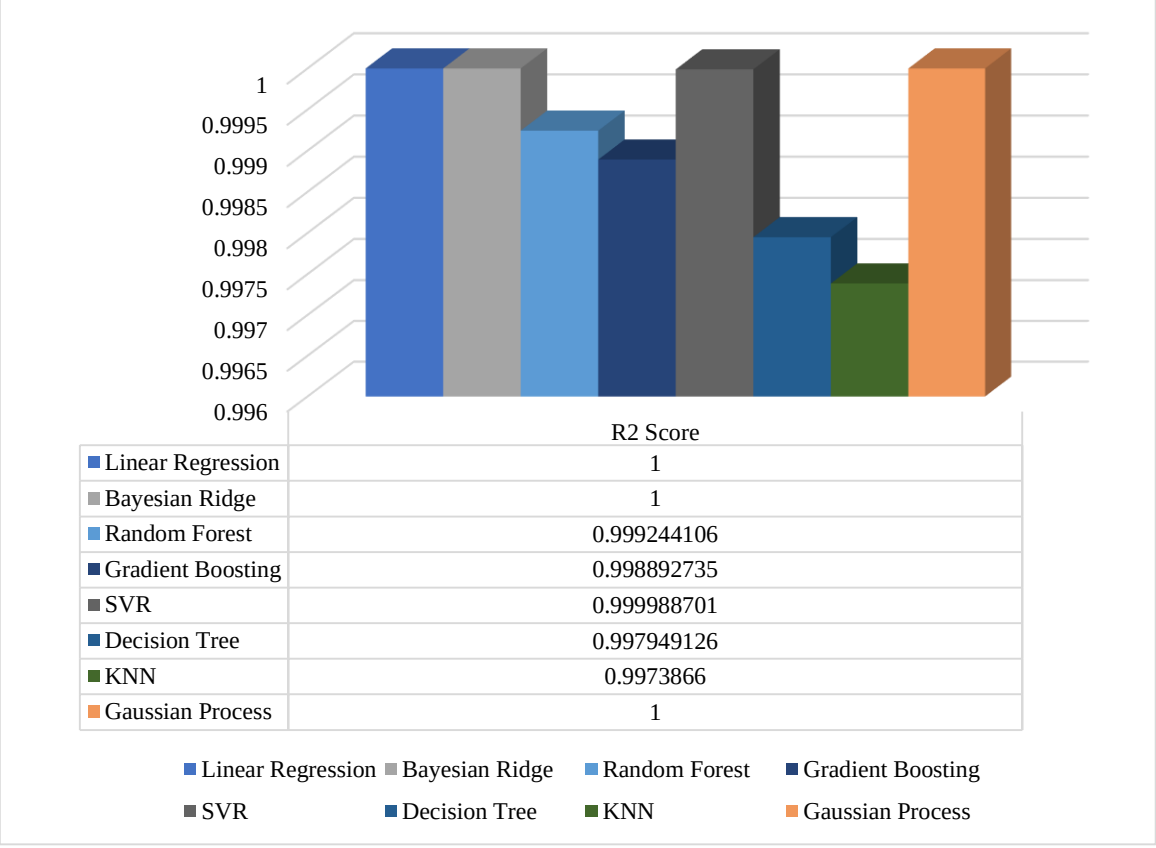
Figures 22 and 23 show the results of the parameters related to the accuracy of the machine learning methods.

According to Figs. 22 and 23, linear regression is the best prediction model for this project because it has the highest R2 score and the lowest mean squared error. Figure 24 shows the graph related to the fitting of the actual and predicted graph by the Bayesian Ridge model.



	Mean Squared Error
Linear Regression	4.13E-26
Bayesian Ridge	2.35E-15
Random Forest	0.368571429
Gradient Boosting	0.539899099
SVR	0.005509525
Decision Tree	1
KNN	1.274285714
Gaussian Process	1.60E-08

Fig. 22. Comparison of Mean squared error in different machine learning methods



	R2 Score
Linear Regression	1
Bayesian Ridge	1
Random Forest	0.999244106
Gradient Boosting	0.998892735
SVR	0.999988701
Decision Tree	0.997949126
KNN	0.9973866
Gaussian Process	1

Fig. 23. Comparison of R2 score in different machine learning methods.

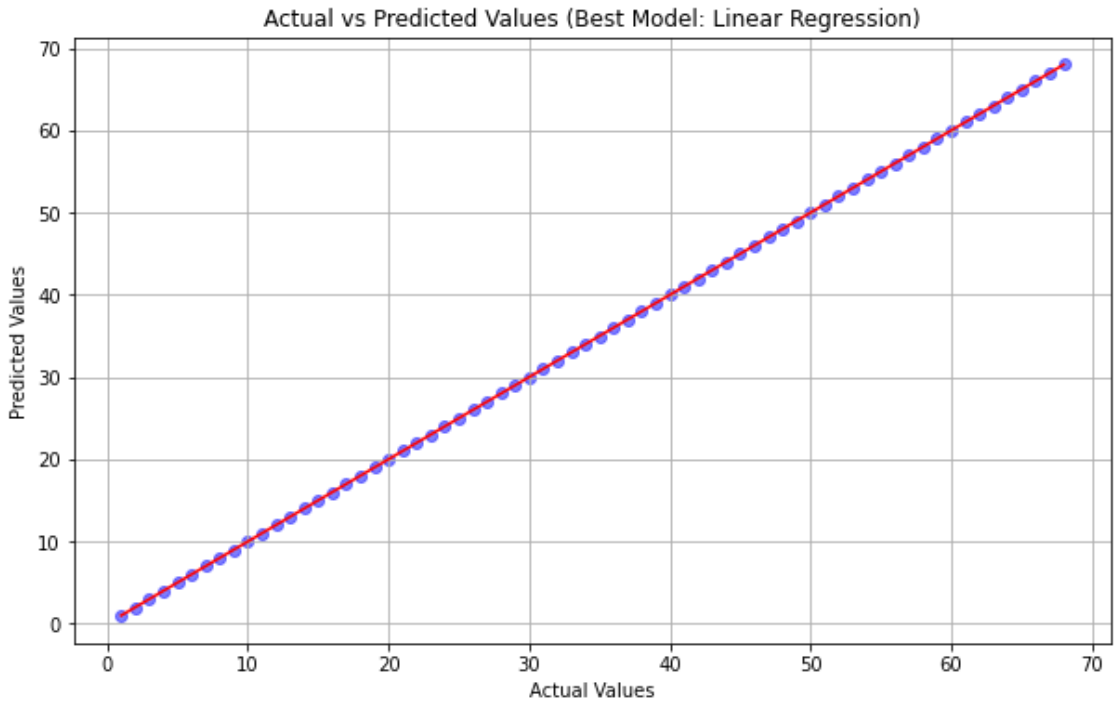


Fig. 24. Fitting of the actual and predicted graph by the Bayesian Ridge model

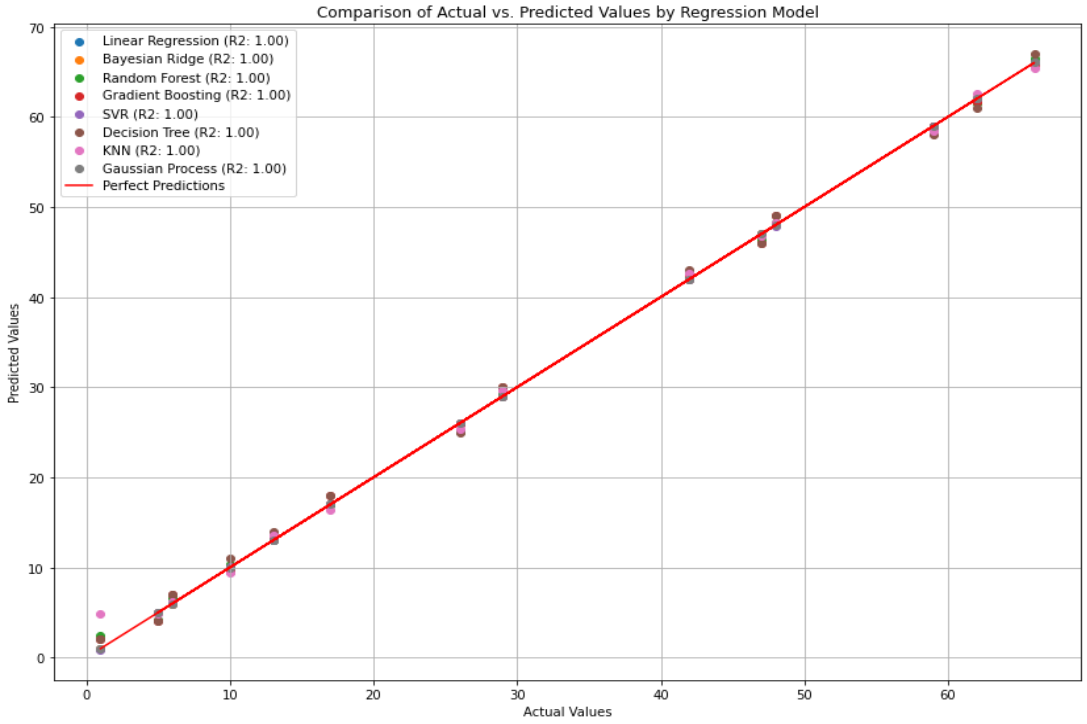


Fig. 25. Comparison of Actual vs. Predicted Values by Regression Model.

Figure 25 shows a comparison between all models and the actual value of $\Delta T_{\text{superheater}}$.

This optimized fire tube boiler design with a superheater has various applications across different industries, including power generation, chemical processing, and heating systems. Power generation can enhance efficiency and reliability, leading to increased output and lower operational costs. Improved heat transfer and steam quality are crucial for precise temperature control in chemical processing. This design can provide more efficient and reliable heating solutions for heating systems, reducing energy consumption and costs.

5. Conclusion

This project seeks to design a three-pass fire tube industrial boiler model with a steam capacity of 5 tons per hour and a working pressure of 10 bar so that the maximum useful heat transfer can be achieved by optimally increasing the number of firetube and superheater passes. The water vapor reaches the industrial consumer at a higher temperature. After stating the governing equations of heat transfer in the fire tube boiler and traditional calculations, the fire tube heat exchanger was designed using Matlab software and Python coding.Here, some important results of the codes' output are discussed. The optimally designed boiler has 71 tubes with a diameter of 5 cm in the second pass and 82 tubes of the same size in the third pass (revisions of the third pass). The speed of hot gases from combustion is between 20 m/s and 40 m/s.

The desired temperature increase in the superheater is 15 ℃; in this case, by adding the superheater section to the fire tube, the level of the third pass increases by 23.72% compared to the case without the superheater. When the superheater temperature increases by 5 ℃, the steam velocity in the third pass tubes reduces by approximately 1 m/s. Adding the superheater at the end of the third run decreases the temperature in this region from 525 ℃ to 500 ℃. This optimized fire tube boiler design with a superheater has various applications across different industries, including power generation, chemical processing, and heating systems. In the end, machine learning algorithms helped determine the percentage of influence of different parameters on the temperature increase in the superheater. Then, the best model among eight machine learning models was selected to predict the temperature increase in the superheater. In this design, the best prediction model is the linear regression model.

6. Future work

Based on the discoveries of this study, forthcoming research will give precedence to key domains aimed at improving the development and utilization of fire tube boilers equipped with superheaters. Key focuses include conducting extensive long-term performance evaluations in real-world industrial environments to validate the durability and efficiency of optimized designs over extended periods. Quantifying the environmental benefits of optimized designs, such as reduced emissions and enhanced energy efficiency, will guide efforts to minimize carbon footprints. Economic viability studies will comprehensively assess the feasibility of adopting optimized designs across industrial sectors, factoring in initial investments, maintenance costs, and energy savings. Furthermore, exploring integration possibilities with renewable energy sources such as solar thermal or biomass seeks to create hybrid systems maximizing energy efficiency and sustainability. Enhanced control algorithms and automation systems will optimize boiler operation precision, improving performance and energy conservation. Addressing these research areas promises to advance the development of more efficient, reliable, and environmentally friendly fire tube boilers with superheaters, pushing the boundaries of thermal engineering.

Contribution

Ebrahim Pil-Ali: Conceptualization; Formal Analysis; Methodology; Investigation; Software; Validation; Visualization; Writing - Original draft; Review & Editing.

Safiye Shafiei: Formal Analysis; Methodology; Investigation; Validation; Review & Editing.

Ramin Kouhikamali: Conceptualization; Methodology; Investigation; Supervision; Review & Editing.

Mohsen Salimi: Formal Analysis; Methodology; Investigation; Validation; Review & Editing.

Majid Amidpour: Formal Analysis; Methodology; Investigation; Supervision; Review & Editing.

References

[1] Barma MC, Saidur R, Rahman SMA, Allouhi A, Akash BA, Sait SM. A review on boilers energy use, energy savings, and emissions reductions. Renew Sustain Energy Rev. 2017 Nov;79:970–83. doi: 10.1016/j.rser.2017.05.187.

[2] Huang H, Liu J, Yang H, Gao Y, Ma X, Zhou J, et al. Biomass briquette fuel, boiler types and pollutant emissions of industrial biomass boiler: A review. Particuology. 2023;77. doi: 10.1016/j.partic.2022.08.016.

[3] Liu J, Zhao J, Zhu Q, Huo D, Li Y, Li W. Methanol-based fuel boiler: Design, process, emission, energy consumption, and techno-economic analysis. Case Stud Therm Eng. 2024 Feb;54:103885. doi: 10.1016/j.csite.2023.103885.

[4] Liu Z, Zhong W, Shao Y, Liu X. Conceptual design of a small-capacity supercritical CO2 coal-fired circulating fluidized bed boiler by an improved design calculation method. Energy. 2022 Sep;255:124534. doi: 10.1016/j.energy.2022.124534.

[5] Hashemi Beni M, Emami S, Meghdadi Isfahani AH, Shirneshan A, Kalbasi R. Transient simulation and exergy analysis of a D-type steam boiler with natural circulation during cold start: A case study. Appl Therm Eng. 2022 Jun;210:118367. doi:10.1016/j.applthermaleng.2022.118367

[6] Drobyazko S, Hilorme T. Methods for evaluating technical innovations in the implementation of energy-saving measures in enterprises. MethodsX. 2022;9.
doi: 10.1016/j.mex.2022.101658.

[7] Thong-On A, Boonruang C. Design of boiler welding for improvement of lifetime and cost control. Materials (Basel). 2016;9(11). doi: 10.3390/ma9110891.

[8] Ranjeeth R, Aditya D, Gautam A, Singh SP, Bhattacharya S. Failure investigation of a water-wall tube of fossil-fuel fired boiler. Eng Fail Anal. 2024 Jan;155:107727. doi: 10.1016/j.engfailanal.2023.107727.

[9] Bhatt A, Ravi V, Zhang Y, Heath G, Davis R, Tan ECD. Emission factors of industrial boilers burning biomass-derived fuels. J Air Waste Manag Assoc. 2023 Apr;73(4):241–57. doi: 10.1080/10962247.2023.2166158.

[10] Horak J, Machálek P, Kralik T, Fencl M, Suchomel J, Sáňka M, et al. PAH emissions

from old and new types of domestic hot water boilers. Environ Pollut. 2017 Jun;225:31–9.
doi: 10.1016/j.envpol.2017.03.034.

[11] Oh S, Seok Yun D, Kim J. Evaluation of high-temperature degradation of platen superheater tube in thermoelectric power plant using nonlinear surface ultrasonic waves. Ultrasonics. 2024 Jan;136:107162.
doi: 10.1016/j.ultras.2023.107162.

[12] Thapa S, Borquist E, Weiss L. Thermal energy recovery via integrated small scale boiler and superheater. Energy. 2018 Jan;142:765-772.
doi: 10.1016/j.energy.2017.10.063.

[13] Oh S-B, Kim J, Han S-W, Kim K-M, Yun D-S, Kim D-W. Analysis of Platen Superheater Tube Degradation in Thermal Power Plants via Destructive/Non-Destructive Characteristic Evaluation. Materials (Basel). 2022 Jan;15(2):581.
doi: 10.3390/ma15020581.

[14] Tognoli M, Keyvanmajd S, Najafi B, Rinaldi F. Simplified finite volume-based dynamic modeling, experimental validation, and data-driven simulation of a fire-tube hot-water boiler. Sustain. Energy Technol. Assessments. 2023 Aug;58:103321.
doi: 10.1016/j.seta.2023.103321.

[15] Tognoli M, Najafi B, Lucchini A, Colombo LPM, Rinaldi F. Implementation of a multi-setpoint strategy for fire-tube boilers utilized in food and beverage industry: Estimating the fuel saving potential. Sustain. Energy Technol. Assessments. 2022 Oct;53:102481.
doi: 10.1016/j.seta.2022.102481.

[16] Mohammad javadi S, Banihashemi S. Study of thermal performance and optimization of city gas station heaters equipped with turbulator in the fire tube section. Therm. Sci. Eng. Prog. 2023 Jan;37:101573.
doi: 10.1016/j.tsep.2022.101573.

[17] Habib MA, Nemitallah MA. Design of an ion transport membrane reactor for application in fire tube boilers. Energy. 2015 Mar;81:787-801.
doi: 10.1016/j.energy.2015.01.029.

[18] Aydin O, Boke YE. An experimental study on carbon monoxide emission reduction at a fire tube water heater. Appl. Therm. Eng. 2010 Dec;30(17-18):2658-2662.
doi: 10.1016/j.applthermaleng.2010.07.014

[19] Islam MR, Parveen M, Haniu H. Properties of sugarcane waste-derived bio-oils obtained by fixed-bed fire-tube heating pyrolysis. Bioresour. Technol. 2010 Jun;101(11):4162-4168.
doi: 10.1016/j.biortech.2009.12.137.

[20] Bisetto A, Del Col D, Schievano M. Fire tube heat generators: Experimental analysis and modeling. Appl. Therm. Eng. 2015 Mar;78:236-247.
doi: 10.1016/j.applthermaleng.2014.10.095

[21] Tognoli M, Najafi B, Marchesi R, Rinaldi F. Dynamic modelling, experimental validation, and thermo-economic analysis of industrial fire-tube boilers with stagnation point reverse flow combustor. Appl. Therm. Eng. 2019 Feb;149:1394-1407.
doi: 10.1016/j.applthermaleng.2018.12.087

[22] Beyne W, Lecompte S, Ameel B, Daenens D, Van Belleghem M, De Paepe M. Dynamic and steady state performance model of fire tube boilers with different turn boxes. Appl. Therm. Eng. 2019 Feb;149:1454-1462.
doi: 10.1016/j.applthermaleng.2018.09.103

[23] Gutiérrez Ortiz FJ. Modeling of fire-tube boilers. Appl. Therm. Eng. 2011 Nov;31(16):3463-3478.
doi: 10.1016/j.applthermaleng.2011.07.001

[24] Tognoli M, Najafi B, Rinaldi F. Dynamic modelling and optimal sizing of industrial fire-tube boilers for various demand profiles. Appl. Therm. Eng. 2018 Mar;132:341-351.
doi: 10.1016/j.applthermaleng.2017.12.082

[25] Akkinepally B, Shim J, Yoo K. Numerical and experimental study on biased tube temperature problem in tangential firing boiler. Appl. Therm. Eng. 2017 Nov;126:92-99.
doi: 10.1016/j.applthermaleng.2017.07.121

[26] Behbahani-nia A, Bagheri M, Bahrampoury R. Optimization of fire tube heat recovery steam generators for cogeneration plants through genetic algorithm. Appl. Therm. Eng. 2010 Nov;30(16):2378-2385.
doi: 10.1016/j.applthermaleng.2010.06.007

[27] Gañan J, Al-Kassir A, González JF, Turegano J, Miranda AB. Experimental study of fire tube boilers performance for public heating. Appl. Therm. Eng. 2005 Aug;25(11-12):1650-1656. doi: 10.1016/j.applthermaleng.2004.10.008

[28] Habib MA, Emara-Shabaik HE, Al-Zaharnah I, Ayinde T. A thermal nonlinear dynamic model for water tube drum boilers. Int. J. Energy Res. 2010 Jan;34(1):20-35. doi: 10.1002/er.1548.

[29] Modliński N, Janda T. Mathematical procedure for predicting tube metal temperature in the radiant superheaters of a tangentially and front fired utility boilers. Therm. Sci. Eng. Prog. 2023 May;40:101763. doi: 10.1016/j.tsep.2023.101763.

[30] Krzywanski J, Sztekler K, Szubel M, Siwek T, Nowak W, Mika Ł. A Comprehensive, Three-Dimensional Analysis of a Large-Scale, Multi-Fuel, CFB Boiler Burning Coal and Syngas. Part 2. Numerical Simulations of Coal and Syngas Co-Combustion. Entropy. 2020 Jul;22(8):856. doi: 10.3390/e22080856.

[31] Bishop CM. Pattern Recognition And Machine Learning. Springer; 2006.

[32] Seeger M. Gaussian processes for machine learning. Int. J. Neural Syst. 2004;14(2):1899-1915. doi: 10.1142/S0129065704001899.

[33] Breiman L. Random forests. Mach. Learn. 2001 Oct;45(1):5-32. doi: 10.1023/A:1010933404324.

[34] Wager S, Hastie T, Efron B. Confidence intervals for random forests: The jackknife and the infinitesimal jackknife. J. Mach. Learn. Res. 2014;15:1625-1651.

[35] Karch J. Improving on Adjusted R-Squared. Collabra Psychol. 2020 Jan;6(1):343. doi: 10.1525/collabra.343.

[36] Ricci L. Adjusted -squared type measure for exponential dispersion models. Stat. Probab. Lett. 2010 Sep;80(17-18):1365-1368. doi: 10.1016/j.spl.2010.04.019.

[37] Sterba SK, Rights JD. R-squared Measures for Multilevel Mixture Models with Random Effects. Struct. Equ. Model. A Multidiscip. J. 2022 Jul;29(4):489-506. doi: 10.1080/10705511.2021.1962325.

[38] Hodson TO, Over TM, Foks SS. Mean Squared Error, Deconstructed. J. Adv. Model. Earth Syst. 2021 Dec;13(12):e2021MS002681. doi: 10.1029/2021MS002681.

[39] Guo M, Ghosh M. Mean squared error of James–Stein estimators for measurement error models. Stat. Probab. Lett. 2012 Nov;82(11):2033-2043. doi: 10.1016/j.spl.2012.06.019.