Artificial Intelligence enabled efficient power generation and emissions reduction underpinning net-zero goal from the coal-based power plants

Waqar Muhammad Ashraf [1,2*], Ghulam Moeen Uddin [2], Hassan Afroze Ahmad [3], Muhammad Ahmad Jamil [4], Rasikh Tariq[5], Muhammad Wakil Shahzad [4], Vivek Dua [1]

[1]*Centre for Process Systems Engineering, Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*

[2]*Department of Mechanical Engineering, University of Engineering & Technology, Lahore, Punjab 54890, Pakistan*

[3]*Huaneng Shandong Ruyi (Pakistan) Energy Private Limited, Sahiwal Thermal Power Plant, Sahiwal, Punjab 57000, Pakistan*

[4]*Mechanical & Construction Engineering Department, Northumbria University, Newcastle Upon Tyne, UK*

[5]*Facultad de Ingeniería, Universidad Autónoma de Yucatán, Av. Industrias No Contaminantes por Anillo Periférico Norte, Apdo. Postal 150, Cordemex, Mérida, 97203, Yucatán, México*

* Corresponding Author: *waqar.ashraf.21@ucl.ac.uk*

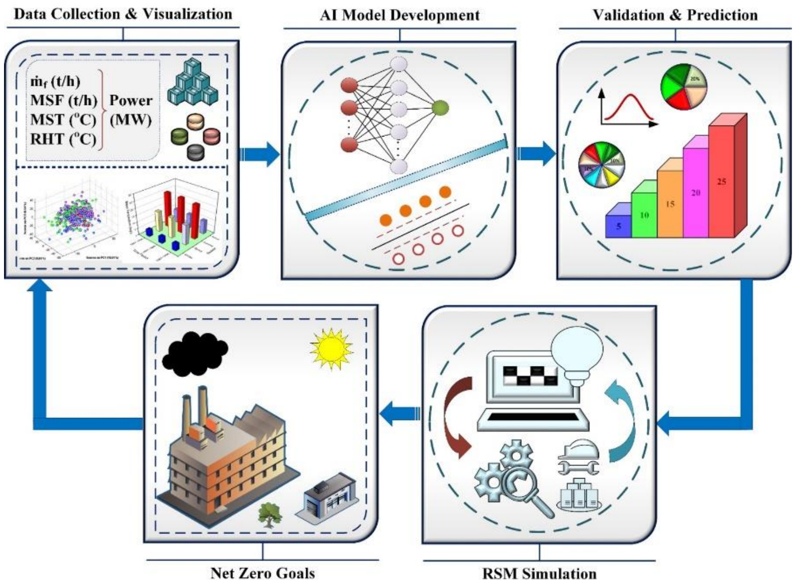
Highlights

- AI-RSM framework for real operational data of a 660 MW power plant is presented.
- AI-RSM enabled efficient power production of the power plant is realized.
- Annual reduction in CO_2, CH_4 and Hg emissions, i.e., 210.2 kt/y is estimated.
- AI based techno-environmental performance enhancement contributes to net-zero goal.

Abstract

A large power generation facility is a complex multi-criteria system associated with multivariate couplings, high dependency, and non-linearity among the operating variables which present a major challenge to ensure efficient power production. In this research, an integrated artificial intelligence (AI) and response surface methodology (AI-RSM) framework to achieve the efficient power production operation of a 660 MW coal power plant is presented. Two AI algorithms, i.e., extreme learning machine (ELM) and support vector machine (SVM) are trained comprehensively on the power plant's operational data and are validated as well. Full factorial design of experiments on the three levels of the operating parameters are constructed and simulated from the better performing AI model which is an effective non-linear representation of the complex power plant operation. RSM analysis is carried out under three power generation scenarios to simulate the effective values of the operating variables which are tested on the power plant's operation and a reasonable agreement is found with the experimental observations. The notable improvement in fuel consumption rate, thermal efficiency, and heat rate of the power plant under Half Load, Mid Load, and Full Load capacity of the power plant is achieved by the AI-RSM framework enabled analyses. It is estimated that annual reduction in CO_2, CH_4 and Hg emissions measuring 210 kilogram tons per year (kt/y), 23.8 t/y and 2.7 kg/y, respectively can be obtained corresponding to Mid Load operating state of the power plant. The research presents the reliable and robust utilization of AI-RSM framework for simulating the effective operating conditions for the fossil-based power plants' operation with an eventual goal to improve the techno-environmental performance which is expected to contribute to net-zero emissions goal from the energy sector.

Graphical Abstract



Keywords

Smart energy; CO_2 reduction; Net-Zero Emissions; Fossil plants; GHG emissions; Artificial Intelligence

Nomenclature

AI	Artificial Intelligence	MST	Main Steam Temperature (°C)
ELM	Extreme Learning Machine	N	Number of samples
IEA	International Energy Agency	RHT	Reheat Steam Temperature (°C)
kt/y	kilogram tons per year	RSM	Response Surface Methodology
$\dot{m}_f$	Coal Flow Rate (t/h)	SLFNs	Single-hidden-Layer Feedforward Networks
MSF	Main Steam Flow Rate (t/h)	SVM	Support Vector Machine

1. Introduction

The recovering global economic activity after the COVID-19 crisis is expected to push the demand for coal up by 4.5%, raising it to above the levels in the year 2019. Overall, electricity demand is expected to grow by 4 % in 2022, most of which will be catered for by coal and gas-fired power plants [1]. In this situation, it is becoming ever more difficult to achieve the targets set by the Paris Agreement and Conference of Parties-26, which stipulates that emissions must drop 7.6 per cent per year from 2020 to 2030 to keep temperatures from exceeding 1.5 °C and 2.7 per cent per year to stay below 2 °C [2].

It is evident that fossil fuel based power generation is one among the greatest contributor to global greenhouse gas emissions and is likely to remain so for the next decade [2]. The International Energy Agency (IEA) report on the energy sector roadmap to carbon neutrality underlines key technologies to achieve net-zero emissions target from the energy sector [3]. However, these technologies are still under the development stage and their reliable utilization across the communities and financial markets require cost competitiveness and innovation in their performance indicators to achieve a sustainable level. Therefore, the IEA report recommends that the existing fossil based power plants need to be operated more efficiently (improved energy efficiency and reduced emissions load) by adopting effective operational practices and strategies until the clean energy transition is robust and sustainable to achieve carbon neutrality [3].

At the same time, the development of effective operational strategies for the energy-efficient power production from large-scale coal power plants is quite challenging. The power production operation is sustained under the integration of multiple energy conversion processes which have a large number of highly non-linear and interacting control variables. The relative significance of these hundreds of variables, extent, nature of non-linearity and interactions among them, and the causal relationship with the critical performance indicator of energy processes is very important and needs to be included in conducting the performance analytics [4, 5]. The first-principle models on the interdisciplinary interacting domains and for the complex industrial processes do not exist [6, 7]. Furthermore, a large number of operational constraints and degradation in the system performance are not incorporated in the existing physical models [6]. In addition to that, the human experience and professional credentials are domain-specific, and cognitive abilities are limited to analyzing and visualizing high dimensional features. Thus, conventional operating strategies developed for complex and integrated industrial processes are built on human wisdom, experience and hit and trial approach which cannot yield effective results for such hyper dimensional problems.

Artificial intelligence (AI) based modeling approach is another paradigm to develop the data-driven models for the complex and multi-variate systems. Our research group has reported various AI based modeling studies for large-scale industrial systems and believes AI is one of the reliable tools for conducting the performance enhancement analytics for large-scale industrial systems [7-18]. However, conducting the AI-based performance enhancement analytics have many associated challenges which are mentioned in the supplementary material. It is the high need of the time to highlight the true potential of AI modeling algorithms to the industrial community especially to beginner and mid-level engineering managers to help them develop efficient techno-economic and environmental policies in order to achieve industrial competitiveness, and subsequently, contribute to industry 4.0 vision [18, 19] and net-zero emissions commitment [20].

Plethora of literature is available on deploying AI tools for modeling the power production from various power generation capacities of the power complexes. Lee et al. [21] developed an artificial neural network (ANN) model using a large number of features taken from various power-generating processes. The data was taken from the simulator of a 500 MW supercritical power plant. The model presented a promising performance in simulating the load variation from 250 to 500 MW [21]. Naveen Kumar [22] modeled the power production operation of a 660 MW by ANN using eight operating parameters to minimize the energy consumption and maximize the power production. Yasin Tunckaya [23] modeled the power production operation of a 660 MW power plant on thirty seven features taken from the operation of boiler, turbine and generator. Ravinder Kumar [24] also modeled the power production rate of a 660 MW power plant on a large number of features taken from various energy devices. Ashraf et al. modeled the apparent power of a 660 MW power plant using the least square support vector machine and other AI techniques. The authors presented the impact of the operating parameters on the power production [10]. In another study, Ashraf et al. constructed the power generation curve for a 660 MW coal-fired power plant by an AI technique [7]. Haddad et al. applied genetic algorithm and particle swarm optimization techniques for the parametric identification from a 660 MW power plant. The simulations proved the feasibility for cleaner and flexible power generation [25].

Liu et al. modeled the 1000 MW ultra-supercritical power plant operation by the significant operating parameters using fuzzy neural network methods. The model presented the merit of efficacy in modeling the complex and non-linear power production operation over the recursive least squares technique [26]. Zhang et al. developed a deep neural network with a stacked auto-encoder modeling approach to model power production from a 1000 MW power plant [27]. Cui et al. constructed a deep neural network in addition to a deep belief network for modeling and control of a 1000 MW ultra-supercritical power generation process [28]. Mikulandri et al. proposed the AI based control for optimizing the combustion process and emissions discharge from a thermal power plant [29]. Lei et al. presented the web-based digital twin network for the monitoring of a physical power plant [30]. In some other studies, machine learning tools were exploited to achieve cleaner power production and increase the thermal efficiency of the coal power plants [31-33].

In the research studies reported in the literature, researchers have deployed a large number of operating parameters for modeling the complex power generation operation of large-scale power complexes. A great number of features with sizable data potentially increase the computational resource, reduction in the prediction efficacy and generalization ability of the developed model [4, 34, 35]. Therefore, modeling the power production operation of a large-scale power plant under the reduced number of system's relevant variables is lacking in the literature which should be investigated. In certain scenarios, selecting a small number of critically relevant operating variables for modeling the power production operation would further reduce the complexity in conducting the performance enhancement analysis and subsequently the implementation of the simulated solutions on the actual power systems would be convenient that should be analyzed from the perspective of industrial applications. This would especially be very useful when the critical segments of power generation systems under investigation are covered by the interdisciplinary domain knowledge and no first principle models are available. Furthermore, an integrated framework comprising on AI based modeling algorithm and the structured analytical technique with which the industrial community is quite versed with for making multi-criteria decision-making should be developed. The potential of AI and process analysis techniques towards net-zero from the energy sector should be investigated to contribute to its pool of knowledge. Thus, the improvement in the performance indices of the power plant and the reduction in emissions discharge to the environment should be evaluated as the result of AI enabled efficient power production analytics so that meaningful and viable contributions to the net-zero goals are made from the energy sector [3, 36].

In this research, two advanced artificial intelligence algorithms are deployed for modeling the complex power generation operation of a 660 MW power plant under the reduced number of operating parameters. ELM is a variant of ANN and is becoming increasingly popular among the scientific community to conduct the modeling and optimization analytics for lab-scale and industrial engineering systems [37-40]. This brings the motivation to investigate the modeling performance of ELM for a 660 MW power plant's power production operation which has not been conducted previously and its application on coal power plant is relatively new. Similarly, SVM has demonstrated excellent performance in various engineering applications in terms of superior modeling capacity and generalization potential [41-44]. Another motivation to apply such data-driven AI modeling methods is to comprehend the behavior of the actual power plant's operation which would contribute to the existing body of knowledge on the effectiveness of the AI algorithms. Response surface methodology (RSM) is a reliable and structured data-analysis technique and is used in academia and industry for the analysis of hyper-dimensional and non-linear complex problems. Therefore, RSM is integrated with the developed AI model (having effectively learnt the complex and non-linear behavior of the power generation operation). The AI-RSM framework built on real industrial data is deployed for simulating the values of the operating parameters for efficient power generation under Half Load, Mid Load and Full Load capacity of the power plant. The simulated results are tested on the power plant's operation and the performance indices like fuel consumption rate, thermal efficiency and power plant heat rate are investigated under three power generation modes. In addition to that, the annual reduction in CO_2, CH_4 and Hg emissions discharge is also estimated. The research presents the effective utilization of AI-RSM framework for the complex operation of a 660 MW power plant that improved its techno-environmental performance and thus contributes to net-zero goals from the energy sector as highlighted by IEA. To the best of the authors' knowledge, the study presents a hybrid framework built on AI modeling algorithm and the statistical technique (RSM) for the power generation operation of a 660 MW capacity coal power plant that is not reported in the literature. Further, the current research work contributes to the body of knowledge of net-zero through AI, as proposed by IEA, and would also be beneficial for the industrial community to adopt the proposed AI-RSM framework for enhancing the performance of their industrial systems. The novelty of this work is summarized as follows:

1- The modeling of 660 MW power plant's power generation operation under reduced number of operating parameters;
2- Evaluating and comparing the modeling performance of ELM and SVM on the complex power generation operation of the power plant;
3- AI-RSM framework for the complex and large-scale power generation system is proposed. The benefits of the AI model (an effective representation of the complex and non-linear system) and RSM (structured and empirical approach to derive the solutions) are unified to estimate the effective values of the operating parameters. The developed solutions are tested on the power plant operation and a good agreement between the experimental and simulated observations is found; and
4- The contribution to the net-zero goal from the energy sector by improved thermal efficiency and reduced emissions load to the environment as suggested by IEA.

Furthermore, the AI-RSM hybrid framework contributes to the pool of knowledge on net-zero by AI and process analysis techniques.

The structure of the paper is partitioned into methodology (section 2), AI model development (section 3), result and discussion (section 4) and conclusion (section 5). Section 2 describes the research methodology followed in the study. Furthermore, the details associated with the data-collection, visualization and processing steps are provided therein. Section 3 provides the comprehensive information on the AI models development along with the hyperparameters tuning. Whereas, section 4 contains the results on the validation of the developed AI models, estimation of the effective values of the operating parameters by RSM under three power generation modes. Moreover, the validation of the developed solutions to achieve the efficient power generation from the power plant, and the improvement in the performance indices for the power generation operation are presented. Finally, the conclusions of the study are mentioned in section 5.

2. Methodology

The research framework developed for efficient power production from a 660 MW super critical power plant [4] is presented in Figure 1. Data collection is the first step in the AI-based data analytics studies and, in this work, is taken from the supervisory information system of the power plant. Furthermore, the generic schematic diagram of the coal power plant and the corresponding property points of the operating parameters are marked on the temperature-entropy diagram of Rankine Cycle as mentioned in the supplementary material. The advanced data-visualization techniques are utilized to confirm the reasonable distribution of data in the variables space. Data-normalization technique is deployed on the data which is subsequently exploited for the development of extreme learning machine (ELM) and support vector machine (SVM) models; the two advanced AI algorithms. The two models can extract useful engineering knowledge from the operational data in reasonable timeframes and with low margins of error thus helping in the establishment of robust process models [41, 45]. The recent publications and scientific activities from our research group have also presented the excellent ability of the AI models in creating value for engineering systems [9, 10, 46].

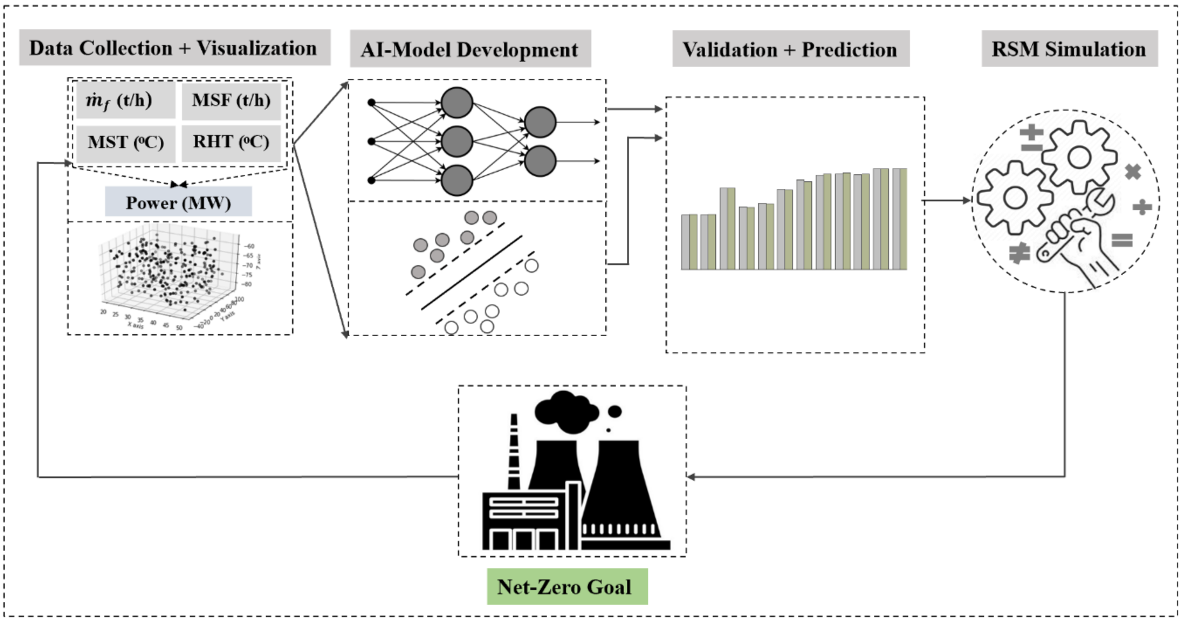


Figure 1. The proposed research framework for the efficient power production operation of a power complex underpinning the net-zero goal from the coal-fired power plants.

The two models are carefully trained by the tuning of the hyperparameters involved in their algorithms to achieve the effective prediction flexibility towards the operating conditions. The prediction performance of the developed AI models is evaluated before deploying them to conduct performance analysis. For this purpose, the models are externally validated by the dataset which complies with two fundamental conditions; (1) the variables should have nearly the same operating ranges comparable with that of training data, and (2) the dataset should be unseen to the models [9]. Four statistical measures namely correlation coefficient (R), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) are introduced to measure the deviation between the actual and model simulated responses and thereby selecting a better performing AI model. The full factorial design of experiment (DOE) technique is utilized for constructing the experiments based on the three levels of the operating parameters and are predicted from the selected AI model. Subsequently, response surface methodology (RSM) analysis is conducted to mine the relationships among the input-output variables. Furthermore, RSM analysis is further extended to determine the values of the operating variables for the efficient power production under three conditions since the power generation is generally maintained around them. These three conditions refer to Half Load, Mid Load and Full Load power generation corresponding to 660 MW capacity which is taken as 356 MW, 496 MW, and 655 MW respectively. The results simulated by RSM are tested on the power plant operation. The key performance indices of the power complex, i.e., fuel flow rate, thermal efficiency, power plant heat rate and the emissions discharge (CO_2, CH_4 and Hg) are investigated corresponding to the efficient power production.

The improvement in energy efficiency and reduction in emissions of fossil-based power plants are the key suggestions given by IEA and other organizations to contribute to net-zero from energy sector until the cost effective and reliable carbon capture technologies are available [3, 47, 48]. Therefore, the potential of advanced AI algorithms and structured RSM technique is applied to ensure the efficient power production, higher thermal efficiency and reduced environmental impact from the investigated coal power plant to contribute to net-zero target.

2.1. Data Collection, Visualization & Processing

In this paper, four operating variables are deployed to model the power production from a 660 MW supercritical coal power plant. The variables are selected based on extensive discussion with the various teams of the power plant, parametric interdependencies and relevancy, and the literature review [6, 10, 23] so that the true representative variables integrated with the power generation operation are considered. Therefore, coal flow rate ($\dot{m}_f$), main steam flow (MSF), main steam temperature (MST) and reheat steam temperature (RHT) are taken for modeling the power production from the power plant. $\dot{m}_f$ indicates the fuel consumption rate in the boiler and the thermal energy received as the result of fuel combustion is utilized corresponding to the sustained power production, main steam conditions, condition of feed water and air (primary and secondary) entering the boiler. Furthermore, $\dot{m}_f$ is also included for evaluating the techno-environment performance of the power plant like thermal efficiency, power plant heat rate and emissions discharge to the environment. Smrekar et al. [6] incorporated the properties of feed water at the inlet of the boiler like temperature and pressure for modeling the steam conditions exiting the boiler. However, the power plant's operational control experience as well as the strong correlation [49] among the conditions of feed water with $\dot{m}_f$ compels us to avoid incorporating these interdependent variables for the AI model development. Similarly, the other three parameters, i.e., MSF, MST, and RHT are the key control variables to be maintained carefully for the power production. MST and RHT indicate the work potential of steam entering the high-pressure and intermediate pressure turbine respectively and thus, are critical to ensure the smooth and energy-efficient power production operation of Rankine Cycle [50]. Whereas, MSF represents the quantity of steam entering the high-pressure turbine and thus is critically synchronized to maintain the sustained power production [50]. Generally, the steam turbines operation is influenced by the condition of steam entering the turbine and thus, the parameters associated with the sub-systems synchronized with the steam turbines operation like feed water regenerative heating system, deaerator functioning, condenser vacuum system etc., are adjusted accordingly [4, 9]. Therefore, the selected variables out of the large number of operating parameters of the power plant, are the ones that are to be critically controlled to sustain smooth power production. It is important to mention here that the end-state variables of a system are enough to develop generalized AI models without incorporating the variables associated with the underlying process physics [51]. At the same time, the selection of the input variables is influenced by the objectives of model development and the availability of the variables. However, the domain knowledge and insight of the system are the crucial requirements for structuring the relevant operating variables with the objective function in order to achieve the targeted results [6].

Numerous state-of-the art sensors are implanted at various points along the various power generation processes to measure the operating values of the operating parameters. The data measured by the sensors is stored in the centralized supervisory information system (SIS). In this study a dataset comprising on hourly-averaged observations of the operating parameters is obtained from the SIS considering the continuous power generation operation of the power plant, and each observation contains the values of four operating variables and the corresponding output variable. The outliers in the observations of the operating parameters are removed and the procedure is described in the supplementary material. Subsequently, a cleaned dataset consisting of 2017 observations is split into 80% and 20% by size for training and testing of the model under development. The operating variables have continuous values and the descriptive statistics of the dataset deployed for the AI model development are mentioned in Table 1. Furthermore, the make of the measuring sensors is also provided. It is pertinent to mention here that the values of the operating variables are carefully taken with reference to the variation in power production, i.e., from Half Load to Full Load capacity of the power plant. The range of the operating variables is fairly wide referring to the operating windows of the variables. Moreover, the data spread is also reasonable as measured by the standard deviation values for the variables.

Table 1. Descriptive statistics of dataset deployed for model development.

Variables	**Sensor**	**Unit**	**Min**	**Mean**	**Max**	**Standard Deviation**
Input Variables						
Coal Flow Rate ($\dot{m}_f$)	Vishay Precision Group (USA)	t/h	126	190	252	48.19
Main Steam Flow Rate (MSF)	Siemens (Germany)	t/h	997	1541	2135	434.96
Main Steam Temperature (MST)	Anhui Tiankang China Thermocouple	°C	550	562	569	5.04
Reheat Steam Temperature (RHT)	Anhui Tiankang China Thermocouple	°C	552	563	569	4.06
Output Variable						
Power	Nanjing Suatak Measurement and Control System	MW	354	517	660	135.20

Figure 2(a,b) show the data distribution of real process data of the input and output variables by box plots. The box plots present a viable way of visualizing the data in the distribution space of the variables. The straight line through the box represents the mean values of the variables while the box itself covers the range of the variables from 25% to 75%. Healthy data distribution is evident from the box plots of $\dot{m}_f$, Power and MSF. However, the box width for MST and RHT is appeared to be relatively tight. It is because the two variables are constrained to be critically controlled between a narrow operating window (min – max value mentioned in Table 1). In general, the data visualization of the variables within the operating limits is reasonable which is essentially required for the development of flexible AI models as discussed in the next section.

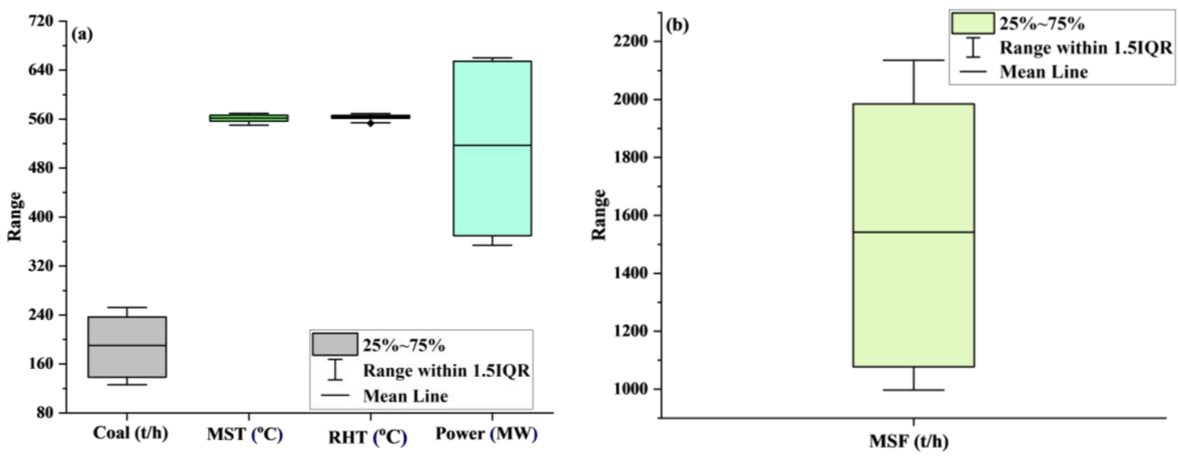


Figure 2. Data visualization of operating variables. a) and b) present the box plots of input ($\dot{m}_f$, MSF, MST and RHT) and output (Power) variables.

An important step in the training of AI models is the normalization of data. It can be observed that the data of the variables do not lie in a uniform range and thus normalization is essential to scale the input variables into equal ranges. This provides a fair chance for developing the association of the input variables with the output variable without biasing it towards some particular input variable. Thus, data normalization ensures effective construction of AI models based on the variables which can have large operating ranges. Min-max normalization method is adopted in this study whose mathematical expression is given as:

$$X_i^* = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{1}$$

where X_i is the actual value of the variable which is normalized to X_i^*. X_{max} and X_{min} are the maximum and minimum values of the variables respectively. After the normalized data set has been constructed, it is fed into the AI algorithm for construction of the model.

3. AI Model Development

In this work, two reliable and advanced AI algorithms have been used, namely support vector machine and extreme learning machine. Both algorithms are investigated for their suitability and performance analyses for function approximation problems [52, 53]. A basic theoretical background of both techniques is presented here:

3.1. Support Vector Machine

Support vector machine (SVM) is one of the popular AI algorithms which is widely used for regression and classification problems in recent years [54]. SVM makes use of the structured risk minimization principle which allows it to possess good generalization performance [54]. SVM transforms the input data onto a higher dimensional feature space which enables it to solve the non-linear problems in a linear pattern [55]. SVM incorporates Vapnik's ε-insensitive loss function, formulates the approximation problem as an inequality constrained optimization problem and learns the non-linear interactions by solving a convex quadratic programming problem. Further details about the SVM algorithm can be found from the supplementary material.

3.2. Extreme Learning Machine

Extreme Learning Machine (ELM) is an evolution of the popular single hidden layer feed forward network (SLFN). ELM was developed to improve the efficiency of SLFN, which faces difficulties in the manual tuning of control parameters (such as epochs, learning rate etc.) and can end up being confined to a local minimum [56]. On the contrary, ELM does not require any intervention by the user as it is implemented automatically with no iterative tuning and has faster learning speeds compared to more traditional SLFN. Moreover, ELM can also achieve a good generalization performance compared to SLFN with gradient descent and back propagation [57]. In ELM, the learning parameters of hidden nodes are assigned randomly and independently, and the output weights of the network are determined analytically by the use of a simple Moore-Penrose inverse operation [57, 58]. A brief description of the ELM development is provided in the supplementary material.

3.3. Hyper parameters tuning for SVM and ELM

A wide range of hyper parameters for the SVM, namely epsilon, kernel scale and box constraint are tested. These hyper parameters are tuned to ensure good generalization performance of the SVM. In this study, the range of hyper parameters, i.e., kernel scale, epsilon, and box constraint is varied between 0.001 ~ 1000, 0.20859 ~ 20859.15 and 0.001 ~ 1000 respectively. After selection of operating ranges, Bayesian optimizer along with the expected improvement per second plus acquisition function is deployed for selecting the different combination of hyper parameters in 30 epochs under a 5-fold cross validation training scheme [10]. The mean square error (MSE) between the modeled outputs and the actual observations is calculated corresponding to a set of values of the hyperparameters, and the MSE comparison corresponding to each set of the hyperparameters values is made for thirty epochs. The minimum MSE of 4.30 is achieved for kernel scale = 37.91, epsilon = 1.14 and box constraint of 997.9.

In the case of ELM, one only needs to specify the type of activation function and the number of neurons in the hidden layer, as ELM itself has good generalization performance for randomly assigned, small values of the input weights, and hidden layer biases of the network [10]. For this study, a sigmoidal activation function ($g(x) = 1/(1+\exp(-x))$) is deployed since it performs better nonlinear transformation and feature mapping compared with other activation functions [59]. Huang et al. [57] proved theoretically that ELM can approximate any continuous function with any degree of accuracy provided it has as many as possible hidden layer neurons included in its architecture. However, in practical applications, a higher number of hidden layer neurons cannot guarantee the best ELM performance [59]. It is recommended to apply an incremental approach in selecting the hidden layer neurons and subsequently, prune the unimportant ones to develop a better ELM model [59]. The hidden layer neurons are varied from 10% - 90% of the training dataset and the root mean square error (RMSE) between the predicted and actual outputs is compared. Hidden layer pruning is executed and a better ELM model comprising of seventy hidden layer neurons is developed for which the least RMSE has been the least. Subsequently, the developed ELM model is forwarded for external validation test as mentioned in section 4.2. Matlab 2020b software is used for the development of SVM and ELM network in this study, and the performance of SVM and ELM in the training phase is provided in the supplementary material.

4. Result and Discussion

4.1. Evaluation Criteria

Four statistical parameters are selected to assess the prediction performance of the trained AI models, namely correlation coefficient (R), root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) [10]. Their mathematical expressions are written as:

$$R = \frac{\sum_{i=1}^{N} (y_i - \bar{y}_i)\left(\hat{y}_i - \bar{\hat{y}}_i\right)}{\sqrt{\sum_{i=1}^{N} (y_i - \bar{y}_i)^2 \sum_{i=1}^{N} \left(\hat{y}_i - \bar{\hat{y}}_i\right)^2}} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \quad (3)$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \quad (4)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i-\hat{y}_i}{y_i}\right| * 100 \quad (5)$$

where, N is the sample size, $\hat{y}_i$ and y_i are the predicted and actual values; $\bar{y}_i$and $\bar{\hat{y}}_i$ are the mean of actual and predicted values. The range of R lies from 0 to 1. R = 0 describes a condition that no correlation exists between the actual and network predicted values. Whereas, R = 1 signifies a perfect correlation between actual and network predicted values. The other merits of performance like RMSE, MAE and MAPE account for the errors computed between the actual and network predicted responses and are measured in relevant measuring units.

4.2. *External Validation of AI models*

The developed AI models are deployed to predict the external validation dataset in order to evaluate their prediction efficacy towards the unseen dataset of the operating variables. The dataset is taken from the power plant for various power generation values and consists of 134 observations. The operating conditions are simulated by SVM and ELM models, the model's simulated responses are compared with the actual power values, and performance matrix built on the chosen statistical parameters is computed.

Figure 3 (a,b) present the graphical representation of actual vs SVM, and actual vs ELM predicted power values for the external validation dataset respectively. Comparing the models' responses with the actual values, it appears that SVM has well predicted the dataset confirming the good generalization ability of the model. Moreover, its predicted responses are far closer to the actual values compared with that of ELM. Although, the R value for ELM predictions is high measuring 0.99 which is comparable with that of SVM (R = 0.99). However, the RMSE, MAE, and MAPE for the ELM model, i.e., 7.81 MW, 5.10 MW, 0.95 % are significantly larger than that of SVM (2.96 MW, 2.59 MW, 0.51 %) respectively. It is explained on the basis of the difference in learning algorithms of the two models. SVM tends to get closer to the actual values through the iterative cost function minimization process which reduces the error between the actual and model predictive responses without sticking to the local extrema, and enables SVM to improve its prediction efficacy. On the other hand, ELM chooses a few of the learning parameters from the input space randomly and based on them, the remaining parameters required for the model development are determined. The iterative error reduction loop is missing in the training algorithm of ELM which could explain the existing predictive performance of the model. Similar results confirming the comparable or superior performance of SVM over ELM are reported in the literature [60, 61]. The performance matrix built for SVM provides confidence in the model's efficacy to predict the unseen operating scenarios which appears to be a little weak in the case of ELM. Therefore, SVM turns out to be a comparatively better model in conducting the further analysis as presented in the subsequent section.

It is important to note that correlation coefficient of the SVM model, constructed on four operating parameters in this work, is comparable with the ANN model (R = 0.99 on eight operating parameters) reported in literature [22]. It confirms that power generation operation can be modeled on fewer operationally and financially significant operating parameters. Subsequently, the estimation of the effective operating values of the operating parameters by RSM and the testing of the developed solutions on the power plant operation would be convenient as described in the subsequent sections 4.3-4.4.

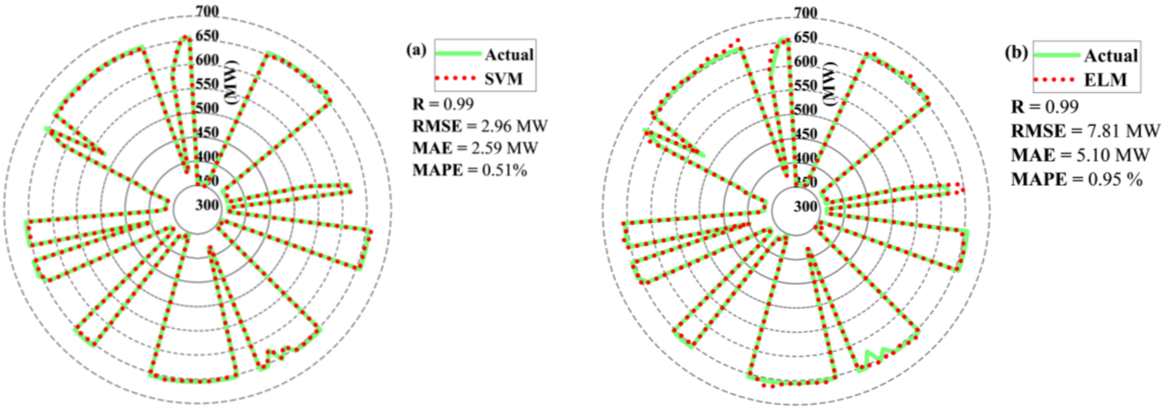


Figure 3. Comparison of actual power with the AI models' predicted responses; a) SVM, and b) ELM. SVM has presented a better performance in simulating the external validation dataset deployed for evaluating the prediction efficacy of the developed models.

4.3. *Response Surface Methodology*

Response surface methodology (RSM) is a statistical technique that mines the relationship among the system's driving variables and objective function. It incorporates linear / square / polynomial or interaction terms to develop the relations among the input and output variables. This technique is well-established and popular among the scientific and research community for deriving optimized solutions for engineering systems [62].

Full factorial design of experiment technique is deployed in this study to generate all possible combination of experiments using the operating variables. Three levels of the variables, i.e., minimum, average and maximum values are utilized for constructing the experiments ($3^4 = 81$). RSM is introduced in combination with the SVM model to develop the causal relationships among the input and output variables through systematically constructed experiments. The prediction of the constructed experiments from the SVM model and the subsequent fitting of RSM model provides the basis for conducting the value-creating analytics for efficient power production.

A second-order regression model along with the interaction terms is fitted through the constructed experiments and the simulated responses (output). The equation of the RSM model is provided in the supplementary material. The model has demonstrated excellent goodness of fit measuring coefficient of determination equal to 0.999. Moreover, the Pareto chart showing the significance (standardized effect) of variables and combinations on the output is presented in Figure 4. A line of statistical significance (95% confidence interval or $\alpha = 0.05$) is drawn on the Pareto chart. The bars represent linear, quadratic, or interaction terms, and the ones lying above the statistical significance line are significant for the output. It is observed that MSF terms out to be the most significant variable followed by $\dot{m}_f$, RHT and MST. Similarly, the interaction among the operating variables is also significant like MSF * RHT, MST * MSF etc. It is estimated that the accumulated contribution of the linear, interaction and square terms on the output variable is 85%, 12%, and 3% respectively representing the significance of operating variables on the power production.

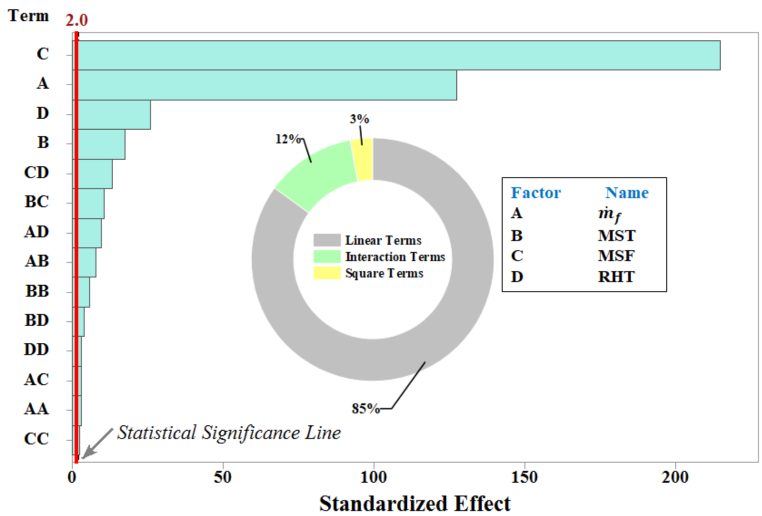


Figure 4. Pareto chart of the standardized effects for the operating variables on the output.

The main effect plot of the operating variables is presented in Figure 5. The plot is constructed on the complete range of a particular variable and keeping other variables at their average values. Thus, it provides the separate and marginal effect of the variable on the output function. The locally estimated scatterplot smoothing is performed for the ease of interpretation of the main effect plot for the selected variable. Referring to Figure 5(a), a rising trend in power production is observed with the increase in $\dot{m}_f$. The increase in $\dot{m}_f$ provides more thermal energy input in the boiler and the subsequent increased heat transfer through heating surfaces for producing high-quality steam which is exploited in power production. The result is supported by the research conducted on the thermal performance evaluation of a coal power plant by Kumar et al. . In Figure 5(b), a linear positive trend is observed between MSF and the power. It is explained by the power plants' operational physics that the increase in main steam flow and its subsequent expansion in the series of turbines produces large torque in the main shaft of the steam turbine system which is utilized in increasing the power production [50, 63].

MST and RHT refer to the steam temperatures before the high pressure and intermediate pressure turbines respectively. The increase in these two temperatures provides more work potential available to the respective turbines. The work potential is observed to be utilized for the power production as evident from Figure 5(c) (MST) and Figure 5(d) (RHT) respectively. The parametric response of MST and RHT on the power generation also complies with the explanation of the Rankine Cycle for steam turbine that the enthalpy of the steam before entering the turbines is high, and thus the net-work produced by the turbines is increased [50].

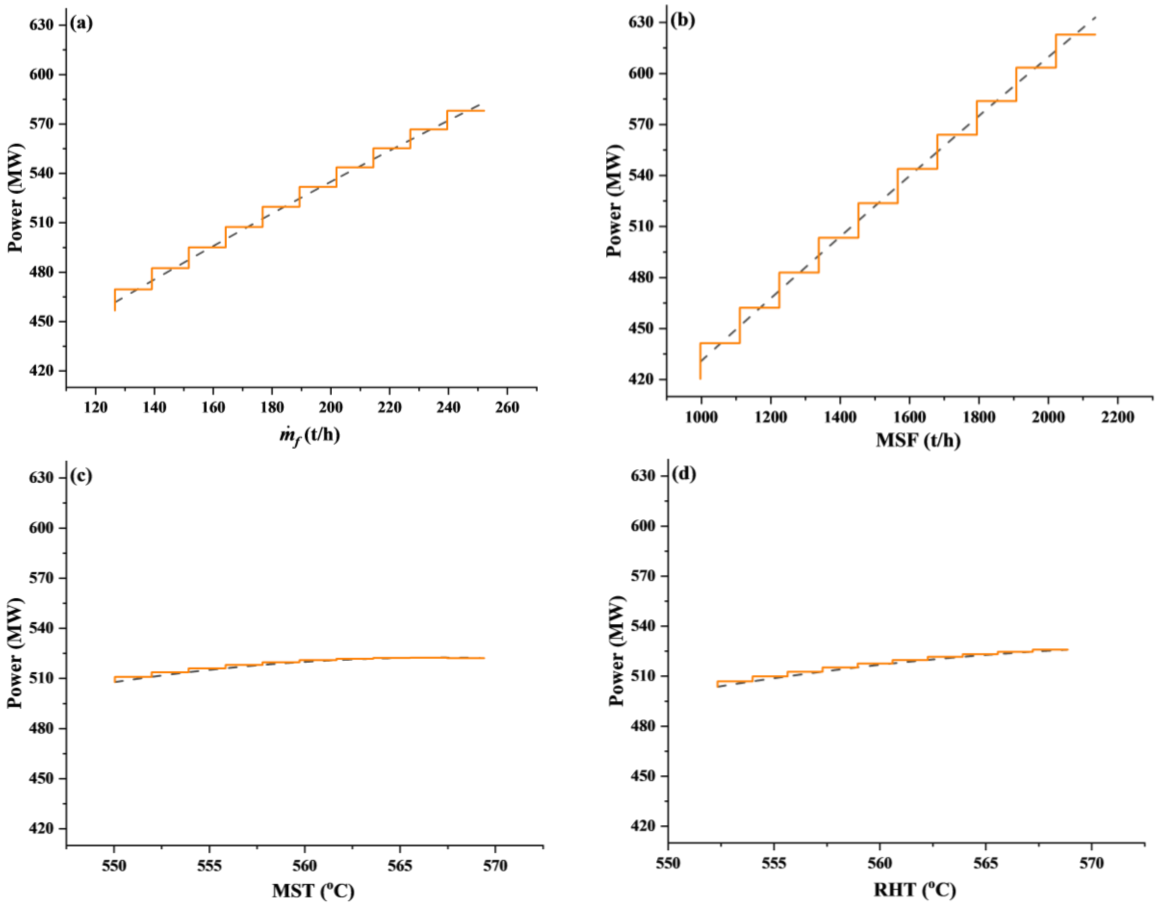


Figure 5. Main effect plot of operating parameters a) $\dot{m}_f$, b) MSF, c) MST, and d) RHT.

The combined effect of two variables on the power production in the form of a contour plot is presented in Figure 6. The contour plot represents the projection of the output function present in space on to a two-dimensional plane defined by the selected variables. Figure 6(a) presents the effect of $\dot{m}_f$ and MSF on the power production under the full variation in their operating ranges, i.e., 126 t/h to 252 t/h and 997 t/h to 2134 t/h respectively. Whereas, MST and RHT are kept at their average values. The power production under the parametric variation is observed to vary from lower than 400 MW to greater than 650 MW which complies with the reported study in literature [22, 63, 64]It is apparent that $\dot{m}_f$ and MSF have a positive impact on the power production and various power production brackets as indicated in the right side of the plot, are identified against the operating windows of the two parameters. Generally, power plants are designed to be operated at their rated full power production capacity. Therefore, constraining $\dot{m}_f$ and MSF in the operating limits, i.e., from 228 t/h to 252 t/h and 2004 t/h to 2132 t/h respectively, power production from the power plant can be sustained > 650 MW.

MST and RHT are the two critically controlled operating parameters during the power plant operational control and their influence on the power production is presented in Figure 6(b). In simulating the effect of MST and RHT on power production, $\dot{m}_f$ and MSF are kept at their average values. It is found that an increase in MST and RHT in their operating limits drives the power production system to enhance its output. It is explained by the fact that the work potential of the steam in Rankine Cycle is effectively increased resulting in higher power production [50]. However, the upper limit of the two temperatures is carefully controlled given the thermal stresses and material metallurgy is to be kept in the safe operating region [65]. For the given operating values of $\dot{m}_f$ and MSF as incorporated in the construct of the MST & RHT contour plot, MST and RHT can be maintained between 561 °C to 569 °C & 567 °C to 569 °C for which the power production would be more than 526 MW.

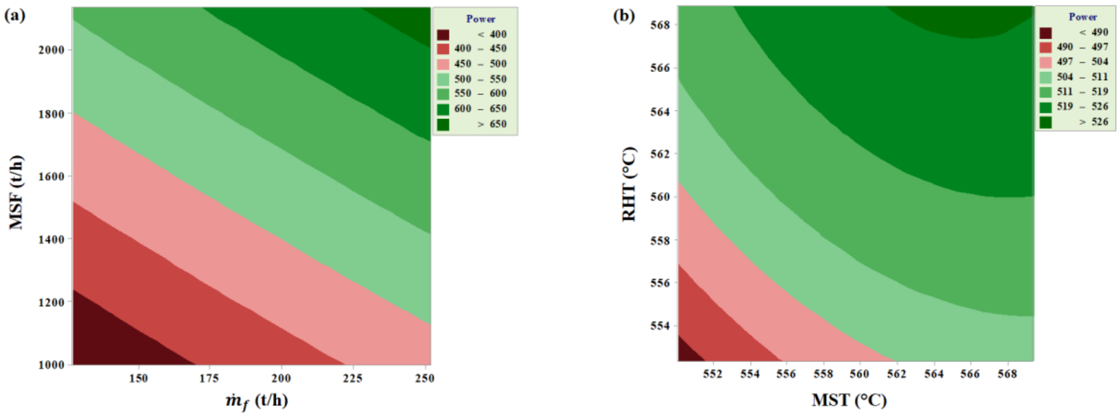


Figure 6. Contour plots of the operating variables a) Effect of $\dot{m}_f$ & MSF on the power production, and b) Effect of MST & RHT on the power production.

RSM analysis is further extended for estimating the values of the operating parameters for which the power production from the power plant is efficient. Various research studies reported in the literature have reported the optimum engineering solutions using the RSM based analysis [66, 67].

In this work, the operating values of the operating variables are determined for the set value of power production by the RSM technique. The limiting range of $\dot{m}_f$, MSF, MST and RHT deployed in the analysis is 126 ~ 252 t/h, 997 ~ 2135 t/h, 550 ~ 569 °C and 552 ~ 569 °C respectively. It is important to note that the operating range of power, i.e., 350~ 360 MW, 490 ~ 500 MW, and 650 ~ 660 MW is considered as Half load, Mid Load and Full Load operating zone for 660 MW power generation capacity of the power plant. Therefore, 356 MW, 496 MW, and 655 MW are taken as Half Load, Mid Load and Full Load set values, and are deployed for finding out the corresponding operating values of $\dot{m}_f$, MSF, MST, and RHT respectively.

Figure 7(a-c) depict the RSM-simulated values for $\dot{m}_f$, MSF, MST and RHT for Half Load, Mid Load and Full Load respectively. The operating values are determined with a desirability factor almost equal to one indicating the strong confidence in the estimated values of the operating variables. Moreover, the target values of power are quite close to the RSM-predicted responses. The operating values of $\dot{m}_f$ and MSF for Half Load, Mid Load, and Full Load are estimated to be 132 t/h, 185 t/h, 236 t/h and 1033 t/h, 1450 t/h, 1980 t/h respectively. Whereas, it is recommended to keep MST and RHT at 569 °C under three power production scenarios. Keeping the MST and RHT at the safe upper control limits is beneficial to ensure the efficient power generation [50].

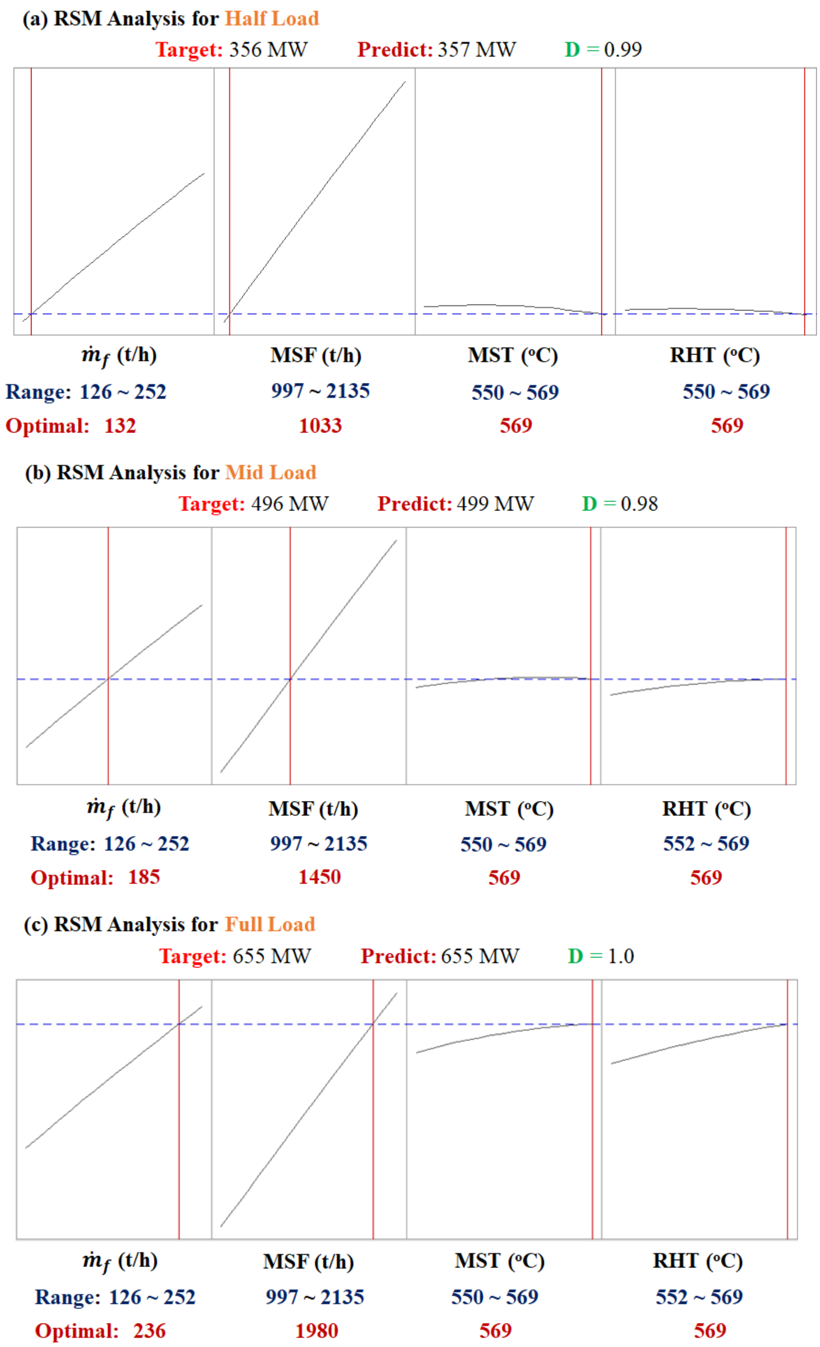


Figure 7. RSM-driven operating values of $\dot{m}_f$, MSF, MST and RHT for a) Half Load, b) Mid Load, and c) Full Load.

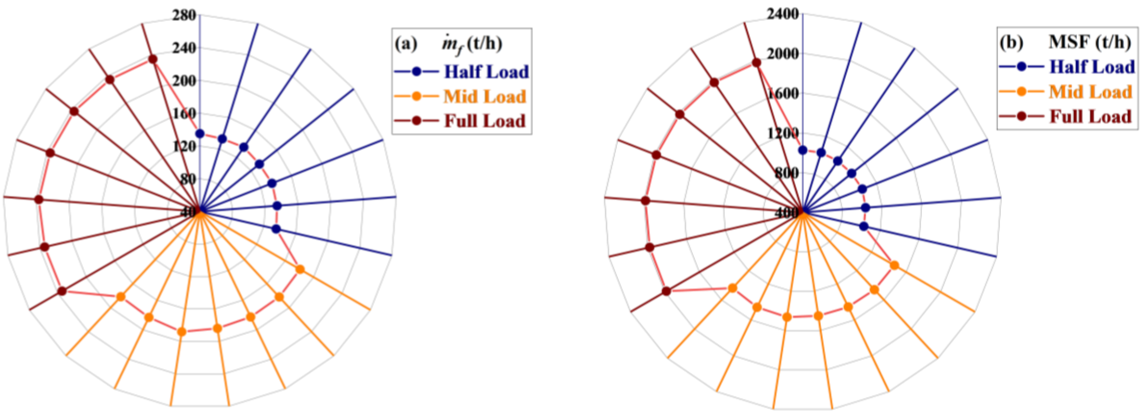
4.4. *Testing of the RSM-driven solutions on the power plant operation*

RSM-based operating values of $\dot{m}_f$, MSF, MST and RHT are tested on the power production operation of the power plant. The power production operation under the three power generation modes, i.e., Half Load, Mid Load, and Full Load is carefully sustained during which each operating parameter is maintained at an operating value for one hour (steady-state operation) with the tendency to keep it around the RSM-driven recommended values for the operating parameters. Subsequently, the average values of the operating variables and the power production observations maintained for one-hour duration are taken. During the test, seven hourly-averaged observations corresponding to the Half Load, Mid Load and Full Load are recorded and the corresponding values of $\dot{m}_f$, MSF, MST and RHT are plotted in Figure 8(a-d) respectively.

Figure 8(a) shows $\dot{m}_f$ observations recorded corresponding to three operating modes of the power plant. The mean value of $\dot{m}_f$ for the three power generation modes maintained during the Half Load, Mid Load, and Full Load is 134 t/h, 183 t/h and 235 t/h respectively. The mean percentage deviation of $\dot{m}_f$ observations with the corresponding RSM-driven values, i.e., 132 t/h, 185 t/h, and 236 t/h for Half Load, Mid Load and Full Load are 1.69%, -0.92% and -0.36% respectively. Similarly, MSF is maintained around 1031 t/h, 1456 t/h, and 1981 t/h (shown in Figure 8(b)) with the mean percentage deviation of -0.20%, 0.39% and 0.07% respectively from the RSM-driven operating values of the operating parameters for the three power generation modes.

Referring to Figure 8(c) and Figure 8(d) plotted for MST and RHT respectively, their average values maintained during the test corresponding to Half Load, Mid Load and Full Load are 568 ^{o}C & 566 ^{o}C, 563 ^{o}C & 565 ^{o}C and 565 ^{o}C & 566 ^{o}C respectively. Some fluctuations in the steam temperatures are observed which are attributed to different adjustments in the set-values of the various operating parameters like attemperation water flow rate, superheat degree, burners angles and soot accumulation on the heating surfaces etc. during three power generation modes of the power plant. However, the mean percentage deviation measured among the MST and RHT observations with the RSM based simulated values of the operating parameters for three power generation modes of the power plant are -0.09% & -0.45%, -1.13% & -0.76% and -0.62% & -0.54% respectively.

Figure 8(e) refers to the power produced under the influence of the operating parameters for three operating scenarios termed as Half Load, Mid Load and Full Load. The average power production for the three power generation modes is 356 MW, 497 MW, and 657 MW (shown by the purple color bar in Figure 8(e)) with the mean percentage deviation of -0.14%, -0.36% and 0.38% from the RSM-driven predicted values. The experimental results for the power production under three power generation modes of the power plant appear to be in good agreement with the RSM-driven solutions. It confirms the ability of the RSM technique to determine the efficient solutions which are not only viable to be implemented on the real engineering systems but could also enhance the performance of industrial systems.



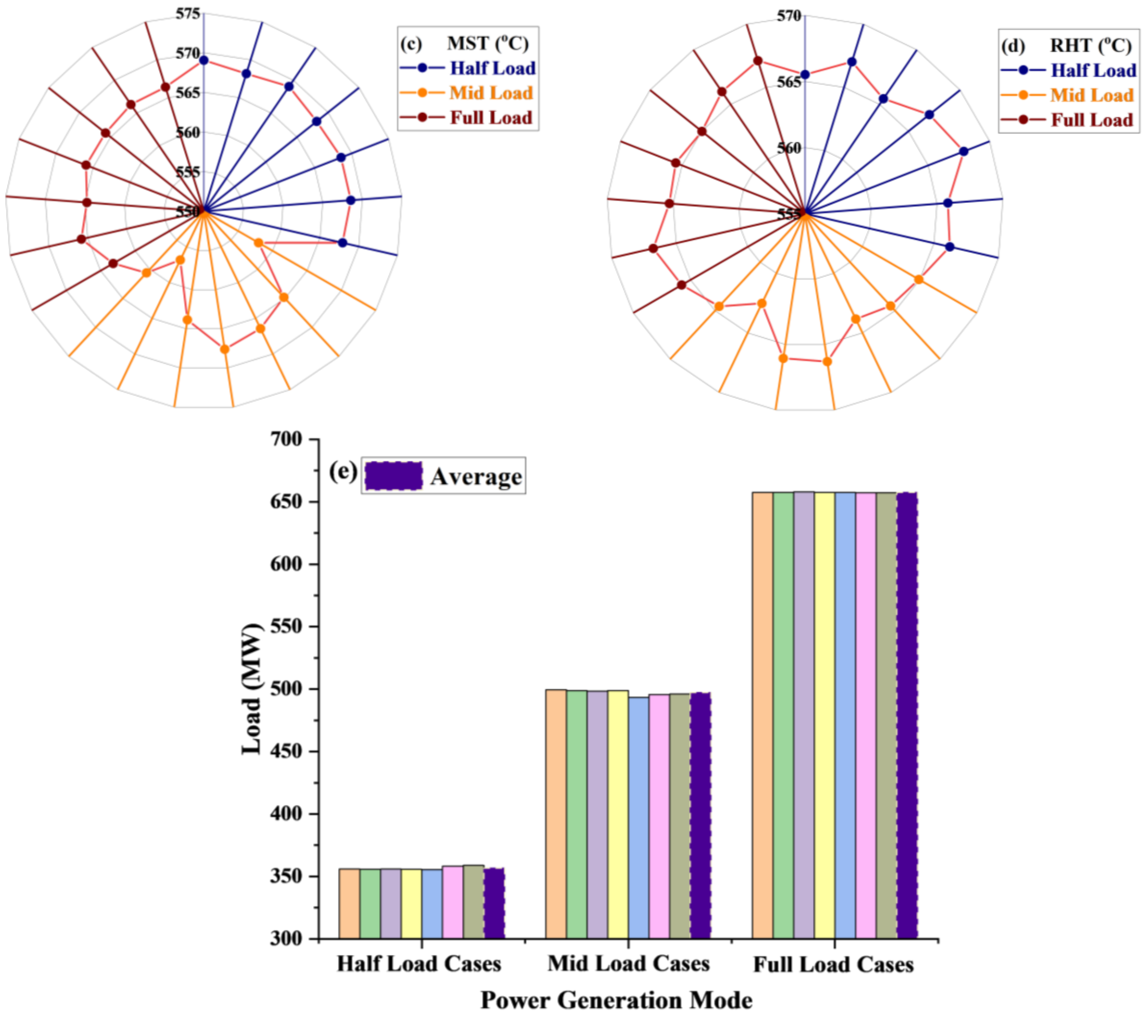


Figure 8. The operating values of the operating parameters maintained during testing of RSM-driven solutions for a) $\dot{m}_f$, b) MST, c) MSF, and d) RHT. (e) The sustained power production under the impact of $\dot{m}_f$, MSF, MST and RHT during Half Load, Mid Load and Full Load. The purple bar shows the average of the power production under the given operating mode of the power plant.

4.5. Performance Enhancement in the Power Plant Operation

The tested results obtained by the RSM technique are compared with the power plant operation. The mean values of $\dot{m}_f$, power, thermal efficiency and heat rate of the power plant under Half Load, Mid Load and Full Load are taken from the power plant operation. These values are taken corresponding to the business as usual (BAU) approach (the power plant's operation control by the operators influenced by their experience and instructions from the management), and are presented in Table 2. Similarly, the RSM-simulated values (RSM-SV) of $\dot{m}_f$, power, thermal efficiency and heat rate are also mentioned in Table 2.

Comparing the BAU observations with the RSM-SV, it is evident that the fuel consumption rate for the power production is less than that of BAU approach measuring 4 t/h, 12 t/h and 6 t/h fuel savings for three power production modes respectively. The thermal efficiency of the power plant is also improved estimating 1.29 percent point (pp), 2.63 pp and 0.94 pp improvement corresponding to RSM-SV based operational control of the power plant respectively. Moreover, the heat rate of the power plant is decreased by 289 kJ/kWh, 586 kJ/kWh and 188 kJ/kWh as the result of RSM-SV for operating parameters settings under Half Load, Mid Load and Full Load operating scenario respectively. It confirms that RSM-simulated operating values of the operating parameters ensured the efficient power production from the power plant.

Table 2. Comparison of test results with power plant operation.

Power Mode	**State**	$\dot{m}_f$ **(t/h)**	**Thermal Efficiency (%)**	**Heat Rate (kJ/kWh)**
Half Load	BAU	138	39.41	9135
	RSM-SV	134	40.70	8846
Mid Load	BAU	195	38.97	9239
	RSM-SV	183	41.60	8653
Full Load	BAU	241	41.89	8594
	RSM-SV	235	42.83	8406

The fuel savings are converted into annual reduction into CO_2, CH_4 and Hg emissions for three operating modes of the power plant, i.e., Half Load, Mid Load, and Full Load as shown in Figure 9. The formula for calculating the emissions is mentioned in the supplementary material. The schedule overhaul period and the allowed shutdown hours as committed with the power purchasing agency are removed for estimating the annual reduction in emissions. It is noted that significant reduction in CO_2 emissions is calculated measuring 70 kilo tons / year (kt/y), 210 kt/y and 105 kt/y corresponding to Half Load, Mid Load and Full Load respectively. Similarly, the savings in CH_4 and Hg emissions discharge to the environment is 7.9 t/y, 23.8 t/y, 11.9 t/y and 1.6 kg/y, 2.7 kg/y, 1.6 kg/y for the three power generation modes respectively. The accumulated reduction in emissions is largest for Mid Load operating mode of the power plant summing 210.2 kt/y. CO_2 and CH_4 are the major contributor to the global warming [68] whereas Hg has a hazardous and toxic impact on the environment [69]. Therefore, the reduction in these pollutants as the result of efficient power generation and improved thermal efficiency not only supports environmental protection but also contributes to global efforts and commitments like Paris Accord and net-zero emissions targets by the energy sector.

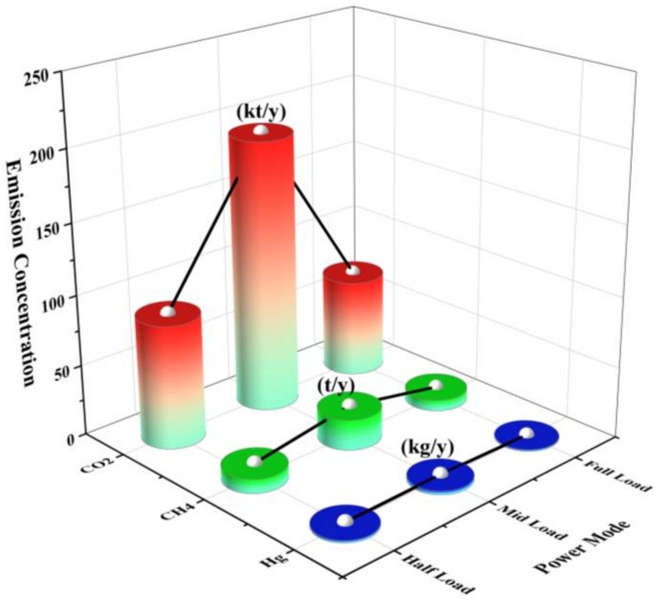


Figure 9. Annual reduction in CO_2, CH_4 and Hg emissions corresponding to Half Load, Mid Load and Full Load.

5. Conclusions

Efficient power production from a coal-based large-scale power facility is a significant challenge for the industrial community to ensure fuel economy, improved thermal efficiency and minimal emissions discharge to the environment. In this paper, AI-RSM framework to achieve the efficient power production from a 660 MW supercritical power plant is presented. Moreover, response surface methodology (RSM) technique coupled with the AI model is deployed to determine the effective operating values of the operating parameters for the efficient power production from the power plant.

The training data of the operating variables are carefully collected, visualized in the input and output space, and normalized. Subsequently, advanced AI models like ELM and SVM are developed, validated and a better-performing model evaluated on a performance matrix is retained, i.e., SVM.

Full factorial design of experiments based on the three levels of operating variables is constructed, simulated by the effective SVM model (representation of the complex and non-linear power generation operation of the power plant) and RSM model is constructed. The Pareto chart shows that MSF is the most significant operating variable to the power production followed by $\dot{m}_f$. RSM analysis is extended to determine the operating values of $\dot{m}_f$, MSF, MST, and RHT under three power generation modes of the power plant, i.e., Half Load, Mid Load and Full Load. The simulated values of the operating variables are tested on the power plant operation under three operating modes, and a reasonable agreement is found with the experimental observations.

The AI-RSM framework enabled simulated results have further improved the fuel consumption, thermal efficiency, and heat rate of the power plant. 12 t/h savings in $\dot{m}_f$, 2.63 pp improvement in thermal efficiency and 586 kJ/kWh reduction in power plant heat rate are achieved corresponding to Mid Load power production mode. Moreover, the annual reduction in CO_2, CH_4 and Hg are found to be the largest corresponding to Mid Load power generation measuring 210 kt/y, 23.8 t/y and 2.7 kg/y respectively. The improvement in the energy efficiency and reduction in the emissions from the coal power plant contribute to net-zero goal from the energy sector through efficient operation of the existing fossil-fuel based energy assets as highlighted by IEA.

The AI-RSM methodology is tested on a 660 MW coal power plant. However, the proposed framework should be tested on large-scale power plants (multi-capacity and different power generation technologies) and process industries to verify its effectiveness. The detailed insight for power generation operation can be investigated with the first-principle models and can be compared with the AI-RSM framework enabled results.

In the future work, ANN-PSO and ANN-fuzzy framework along with other optimization techniques would be investigated considering significant causal variables for the multi-objective optimization problem incorporating thermal efficiency and emissions for multi-capacity coal power plants. It is expected that contribution to the pool of knowledge on the carbon neutrality from the energy sector by AI would be continued to made in the future research studies.

Declaration of Competing Interest

The authors declare no competing financial benefits and personal relationships have influenced the findings of the paper.

Acknowledgment

The authors are thankful to the management of Sahiwal Coal Power Plant, Sahiwal, Pakistan. The technical support and access to the power plant resources made available by the power plant management are highly appreciated.

Funding

WMA declares to have received the funding from Punjab Education Endowment Fund (PEEF) to pursue his PhD at University College London.

References

1. IEA, *Global Energy Review 2021, IEA, Paris.* https://www.iea.org/reports/global-energy-review-2021. 2021.
2. (2020), U., *Emissions Gap Report 2020. Nairobi.* https://www.unep.org/emissions-gap-report-2020.
3. IEA, *An Energy Sector Roadmap to Carbon Neutrality in China;*https://www.iea.org/reports/an-energy-sector-roadmap-to-carbon-neutrality-in-china*.* 2021.
4. Ashraf, W.M., et al., *Optimization of a 660 MW e Supercritical Power Plant Performance—A Case of Industry 4.0 in the Data-Driven Operational Management Part 1. Thermal Efficiency.* Energies, 2020. **13**(21): p. 5592.
5. Ashraf, W.M., et al., *Artificial intelligence based operational strategy development and implementation for vibration reduction of a supercritical steam turbine shaft bearing.* Alexandria Engineering Journal, 2022. **61**(3): p. 1864-1880.
6. Smrekar, J., et al., *Prediction of power output of a coal-fired power plant by artificial neural network.* Neural Computing and Applications, 2010. **19**(5): p. 725-740.
7. Ashraf, W.M., et al., *Construction of operational data-driven power curve of a generator by industry 4.0 data analytics.* Energies, 2021. **14**(5): p. 1227.
8. Ashraf, W.M., et al., *Artificial intelligence based operational strategy development and implementation for vibration reduction of a supercritical steam turbine shaft bearing.* Alexandria Engineering Journal, 2021.
9. Ashraf, W.M., et al., *Strategic-level performance enhancement of a 660 MWe supercritical power plant and emissions reduction by AI approach.* Energy Conversion and Management, 2021. **250**: p. 114913.
10. Muhammad Ashraf, W., et al., *Optimization of a 660 MWe Supercritical Power Plant Performance—A Case of Industry 4.0 in the Data-Driven Operational Management. Part 2. Power Generation.* Energies, 2020. **13**(21): p. 5619.
11. Uddin, G.M., et al., *Artificial intelligence-based emission reduction strategy for limestone forced oxidation flue gas desulfurization system.* Journal of Energy Resources Technology, 2020. **142**(9): p. 092103.
12. Uddin, G.M., et al., *Neural networks assisted computational aero-acoustic analysis of an isolated tire.* Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2020. **234**(10-11): p. 2561-2577.
13. Uddin, G.M., et al., *Monte Carlo study of the high temperature hydrogen cleaning process of 6H-silicon carbide for subsequent growth of nano scale metal oxide films.* International Journal of Nanomanufacturing, 2013. **9**(5-6): p. 407-430.

14. Li, T., P. Liu, and Z. Li, *Operation Data based Modelling and Optimization of Thermal Power Units under Full Working Conditions*, in *Computer Aided Chemical Engineering*, A. Espuña, M. Graells, and L. Puigjaner, Editors. 2017, Elsevier. p. 2455-2460.
15. Yu, J., P. Liu, and Z. Li, *A Hybrid Modelling Approach to Developing Digital Twins of an Ultra-supercritical Steam Turbine Control Stage*, in *Computer Aided Chemical Engineering*, S. Pierucci, et al., Editors. 2020, Elsevier. p. 115-120.
16. Venkatasubramanian, V., *The promise of artificial intelligence in chemical engineering: Is it here, finally.* AIChE J, 2019. **65**(2): p. 466-478.
17. Suvarna, M., T.P. Araújo, and J. Pérez-Ramírez, *A generalized machine learning framework to predict the space-time yield of methanol from thermocatalytic CO2 hydrogenation.* Applied Catalysis B: Environmental, 2022: p. 121530.
18. Suvarna, M., et al., *Smart manufacturing for smart cities—overview, insights, and future directions.* Advanced Intelligent Systems, 2020. **2**(10): p. 2000043.
19. Suvarna, M., et al., *Cyber-physical production systems for data-driven, decentralized, and secure manufacturing—A perspective.* Engineering, 2021.
20. MacDougall, A.H., et al., *Is there warming in the pipeline? A multi-model analysis of the Zero Emissions Commitment from CO 2.* Biogeosciences, 2020. **17**(11): p. 2987-3016.
21. Lee, K.Y., et al. *Neural network-based modeling for a large-scale power plant*. in *2007 IEEE Power Engineering Society General Meeting*. 2007. IEEE.
22. Naveen Kumar, G. and E. Gundabattini, *Optimization and Analysis of Design Parameters, Excess Air Ratio, and Coal Consumption in the Supercritical 660 MW Power Plant Performance using Artificial Neural Network.* Journal of The Institution of Engineers (India): Series C, 2021: p. 1-13.
23. Tunckaya, Y. and E. Koklukaya, *Comparative prediction analysis of 600 MWe coal-fired power plant production rate using statistical and neural-based models.* Journal of the Energy Institute, 2015. **88**(1): p. 11-18.
24. Kumar, R., K. Nikam, and R. Jilte, *A simulation model to predict coal-fired power plant production rate using artificial neural network tool*, in *Applied Computer Vision and Image Processing*. 2020, Springer. p. 150-160.
25. Haddad, A., et al., *Parameter identification of a highly promising cleaner coal power station.* Journal of Cleaner Production, 2021. **326**: p. 129323.
26. Liu, X., et al., *Modeling of a 1000 MW power plant ultra super-critical boiler system using fuzzy-neural network methods.* Energy Conversion and Management, 2013. **65**: p. 518-527.
27. Zhang, H., et al., *Stacked auto-encoder modeling of an ultra-supercritical boiler-turbine system.* Energies, 2019. **12**(21): p. 4035.
28. Cui, J., T. Chai, and X. Liu, *Deep-Neural-Network-Based Economic Model Predictive Control for Ultrasupercritical Power Plant.* IEEE Transactions on Industrial Informatics, 2020. **16**(9): p. 5905-5913.
29. Mikulandric, R.A., et al., *Improvement of environmental aspects of thermal power plant operation by advanced control concepts.* Thermal science, 2012. **16**(3): p. 759-772.
30. Lei, Z., et al., *Toward a Web-Based Digital Twin Thermal Power Plant.* IEEE Transactions on Industrial Informatics, 2021. **18**(3): p. 1716-1725.
31. Strušnik, D., et al., *Optimisation of an old 200 MW coal-fired boiler with urea injection through the use of supervised machine learning algorithms to achieve cleaner power generation.* Journal of Cleaner Production, 2021. **290**: p. 125200.
32. Strušnik, D., *Integration of machine learning to increase steam turbine condenser vacuum and efficiency through gasket resealing and higher heat extraction into the atmosphere.* International Journal of Energy Research, 2022. **46**(3): p. 3189-3212.
33. Strušnik, D., M. Golob, and J. Avsec, *Artificial neural networking model for the prediction of high efficiency boiler steam generation and distribution.* Simulation Modelling Practice and Theory, 2015. **57**: p. 58-70.

34. Li, J., et al., *Machine learning aided supercritical water gasification for H2-rich syngas production with process optimization and catalyst screening.* Chemical Engineering Journal, 2021. **426**: p. 131285.
35. Sarkis, M., et al., *Decision support tools for next-generation vaccines and advanced therapy medicinal products: present and future.* Current Opinion in Chemical Engineering, 2021. **32**: p. 100689.
36. Liebreich, M. *The Quest for Resilience – What Could Possibly Go Wrong? BloombergNEF (Available:https://about.bnef.com/blog/liebreich-the-quest-for-resilience-what-could-possibly-go-wrong/)*. 2022.
37. Wang, M., et al., *Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses.* Neurocomputing, 2017. **267**: p. 69-84.
38. Naz, A., et al., *Short-term electric load and price forecasting using enhanced extreme learning machine optimization in smart grids.* Energies, 2019. **12**(5): p. 866.
39. Sahani, M. and P.K. Dash, *Automatic power quality events recognition based on Hilbert Huang transform and weighted bidirectional extreme learning machine.* IEEE Transactions on Industrial Informatics, 2018. **14**(9): p. 3849-3858.
40. Hua, L., et al., *Integrated framework of extreme learning machine (ELM) based on improved atom search optimization for short-term wind speed prediction.* Energy Conversion and Management, 2022. **252**: p. 115102.
41. Pisner, D.A. and D.M. Schnyer, *Support vector machine*, in *Machine Learning*. 2020, Elsevier. p. 101-121.
42. Maddipati, S., et al., *Learning patterns in combinatorial protein libraries by Support Vector Machines.* Computers & chemical engineering, 2011. **35**(6): p. 1143-1151.
43. Ji, C., et al., *Multi-objective optimization of operating parameters for a gasoline Wankel rotary engine by hydrogen enrichment.* Energy Conversion and Management, 2021. **229**: p. 113732.
44. Tong, J., et al., *Experiment analysis and computational optimization of the Atkinson cycle gasoline engine through NSGA Ⅱ algorithm using machine learning.* Energy Conversion and Management, 2021. **238**: p. 113871.
45. Li, L.-L., et al., *Using enhanced crow search algorithm optimization-extreme learning machine model to forecast short-term wind power.* Expert Systems with Applications, 2021. **184**: p. 115579.
46. Uddin, G.M., et al., *Artificial intelligence-based Monte-Carlo numerical simulation of aerodynamics of tire grooves using computational fluid dynamics.* AI EDAM, 2019. **33**(3): p. 302-316.
47. IEA, *Net Zero by 2050. 2021, IEA: Paris. p. 224*.
48. Rajat Gupta, J.W. *Asia's net-zero transition: Opportunity and risk amid climate action,Mckinsey & Company* 2022 [cited 2022 05-26]; Available from: https://www.mckinsey.com/featured-insights/future-of-asia/asias-net-zero-transition-opportunity-and-risk-amid-climate-action?cid=other-eml-alt-mip-mck&hdpid=49c93cdf-a569-464c-9a47-a0f6da5560bb&hctky=13407127&hlkid=867bd78c5f49494aba7ac81586cb6bf9.
49. Tang, Q., et al., *Machine learning prediction of pyrolytic gas yield and compositions with feature reduction methods: Effects of pyrolysis conditions and biomass characteristics.* Bioresource Technology, 2021. **339**: p. 125581.
50. Cengel, Y.A., M.A. Boles, and M. Kanoğlu, *Thermodynamics: an engineering approach*. Vol. 5. 2011: McGraw-hill New York.
51. Pistikopoulos, E.N., et al., *PAROC—An integrated framework and software platform for the optimisation and advanced model-based control of process systems.* Chemical Engineering Science, 2015. **136**: p. 115-138.

52. Han, F. and D.-S. Huang, *Improved extreme learning machine for function approximation by encoding a priori information.* Neurocomputing, 2006. **69**(16-18): p. 2369-2373.
53. Xiaofang, Y. and W. Yaonan, *Parameter selection of support vector machine for function approximation based on chaos optimization.* Journal of Systems Engineering and Electronics, 2008. **19**(1): p. 191-197.
54. Yaohua, T., G. Weimin, and G. Jinghuai. *Efficient model selection for Support Vector Machine with Gaussian kernel function*. in *2009 IEEE Symposium on Computational Intelligence and Data Mining*. 2009.
55. Lv, Y., et al., *Adaptive selective catalytic reduction model development using typical operating data in coal-fired power plants.* Energy, 2020. **192**: p. 116589.
56. Ding, S., X. Xu, and R. Nie, *Extreme learning machine and its applications.* Neural Computing and Applications, 2013. **25**: p. 549-556.
57. Huang, G., Q.-Y. Zhu, and C.K. Siew, *Extreme learning machine: a new learning scheme of feedforward neural networks.* 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), 2004. **2**: p. 985-990 vol.2.
58. Serre, D. *Matrices: Theory and Applications*. 2002.
59. Cao, W., et al. *Some tricks in parameter selection for extreme learning machine*. in *IOP conference series: materials science and engineering*. 2017. IOP Publishing.
60. Fikret Kurnaz, T. and Y. Kaya, *The comparison of the performance of ELM, BRNN, and SVM methods for the prediction of compression index of clays.* Arabian Journal of Geosciences, 2018. **11**(24): p. 770.
61. Cao, J. and Z. Lin, *Extreme Learning Machines on High Dimensional and Large Data Applications: A Survey.* Mathematical Problems in Engineering, 2015. **2015**: p. 103796.
62. Tacias-Pascacio, V.G., et al., *Comparison of acid, basic and enzymatic catalysis on the production of biodiesel after RSM optimization.* Renewable Energy, 2019. **135**: p. 1-9.
63. Kumar, R., et al., *A simulation model for thermal performance prediction of a coal-fired power plant.* International Journal of Low-Carbon Technologies, 2019. **14**(2): p. 122-134.
64. Nikam, K.C., R. Kumar, and R. Jilte, *Exergy and exergo-environmental analysis of a 660 MW supercritical coal-fired power plant.* Journal of Thermal Analysis and Calorimetry, 2021. **145**(3): p. 1005-1018.
65. Muhammad Ashraf, W., et al., *Optimization of a 660 MWe supercritical power plant performance—a case of Industry 4.0 in the data-driven operational management part 1. Thermal efficiency.* Energies, 2020. **13**(21): p. 5592.
66. Saldana-Robles, A., et al., *Optimization of the operating parameters using RSM for the Fenton oxidation process and adsorption on vegetal carbon of MO solutions.* Journal of Industrial and Engineering Chemistry, 2014. **20**(3): p. 848-857.
67. Pashaei, H., et al., *Experimental modeling and optimization of CO2 absorption into piperazine solutions using RSM-CCD methodology.* ACS omega, 2020. **5**(15): p. 8432-8448.
68. Peter, S.C., *Reduction of CO2 to chemicals and fuels: a solution to global warming and energy crisis.* ACS Energy Letters, 2018. **3**(7): p. 1557-1561.
69. Okpala, C.O.R., et al., *Hazardous properties and toxicological update of mercury: From fish food to human health safety perspective.* Critical reviews in food science and nutrition, 2018. **58**(12): p. 1986-2001.