Operational guide to stabilize, standardize and increase power plant efficiency

Lara Werncke Vieira[a], Augusto Delavald Marques[a], Jéssica Duarte[a], Rafael Petri Zanardo[a], Paulo Smith Schneider[a], Felipe Antonio Chegury Viana[b], Antônio José da Silva Neto[c], Felipe Roman Centeno[a], Julian David Hunt[d], Julio Cezar Mairesse Siluk[e]

[a] *Federal University of Rio Grande do Sul, Porto Alegre, Brazil*
[b] *University of Central Florida, Orlando, UCF, United States*
[c] *State University of Rio de Janeiro - UERJ, Nova Friburgo, Brazil*
[d] *International Institute for Applied Systems Analysis - IIASA, Vienna, Austria*
[e] *Federal University of Santa Maria - UFSM, Santa Maria, Brazil*

Abstract

Complex engineering systems, such as power plants, deliver their best performance when operating along a designed range of some priority parameters. However, plant field operation may deviate from design conditions, and new references must be identified. Actions towards high-quality operation can be supported by fine modeling, which helps building decision support tools. The present work proposes a standardization strategy for the operation of an actual coal-fired power plant based on a Design of Experiment approach, partially tested onsite and finally accomplished with surrogate models built upon a 2 year long database. Artificial Neural Networks (ANNs) and Mass and Energy balances (M&Es) are used to represent the plant´s steam generator and its mills subset, which is the core of an operational guide to increase system efficiency under actual operation. Primary and secondary air flows, pulverized coal outlet temperature, speed of the dynamic classifier, primary air flow, excess O_2, primary and secondary air pressures are the seven controllable factors selected as the most relevant ones among an extensive set of parameters, able to perform effective maneuvers. The application of the operational guide indicates combinations of ranges of the seven controllable parameters that allow for achieving steam generator efficiency within the 84.0% to 88.92% range. The proposed methodology aims as well to improve safe and stable conditions to a system that undergoes operation different than the one prescribed by the original design. The study case results show an opportunity to raise efficiency by up to 2.28% during operation, which represents a reduction in coal consumption by 3.1 t/h and above 6% on CO_2

emissions.

Keywords: Coal-fired power plant, Surrogate model, Optimal operating strategy, Plant operation

1. Introduction

The need to produce electricity while respecting environmental restrictions pushed engineering to enhance power plant efficiency. A unity gain in efficiency on conventional pulverized coal power plants can lead to a 2-3% reduction in CO_2 emissions. Complex engineering systems, such as power plants, deliver their best performance when operating along a designed range of some priority parameters. The plant control system handles systemic instabilities, leaving the operator to manage controllable losses [1, 2]. Equipment wear, external limitations and eventual disturbances can prevent to reach the project design performance [3, 4], creating a gap between that original target and current operating conditions. The operational team must know how to act in these situations, which leads to the need for strategies and decision support [5].

The influence of external factors and disturbances must be considered when looking for a condition of higher performance and stability of the power plant, which goes beyond the search for a fixed optimum point. Optimal conditions are seldom feasible and achieved with the aid of automatic control, but some manned intervention can impact to attenuate the magnitude of losses [1, 2].

The realistic balance between efficiency and performance also becomes meaningful. Efficiency is a project condition pursued by automatic control, while performance is the best response to be achieved with the current system conditions.

Decision support tools can help improving the system performance by reducing its variability by means of process standardization. These tools can be based on computational representations able to simulate the system behavior in a broad range of conditions while delivering a quick response.

Some well known tools can help to build that approach, which aims to represent the system behavior by means of surrogate models. The Design of Experiments (DoE) techniques and Response Surface Methodologies (RSM) allow for scrutinizing actual systems and to deliver easy to use models. A systematic review on both approaches to produce surrogate models was performed in the context of coal-fired power plants. The why and how of the systematic review was synthesized in this research question: how surrogate modeling techniques supported by DoE and RSM can help to enhance coal-fired power plant efficiency? The extensive search in Scopus and Web of Science databases returned 434 relevant works, but only 13 studies were selected as relevant to answer the proposed review question.

The analysed studies were applied to multiple power plant systems, from the design of labyrinth seal leakage in steam turbines to the CO2 capture process, maintenance decision support and optimization. Papers were mainly connected to the energy theme.

Literature review indicated that surrogate models built with the aid of DoE or RSM can help testing power plants without actually subjecting them to adverse or unsafe conditions. Nine of the selected studies applied RSM and DoE, two applied RSM without addressing DoE, and the last two applied DoE solely, as presented in Figure 1.

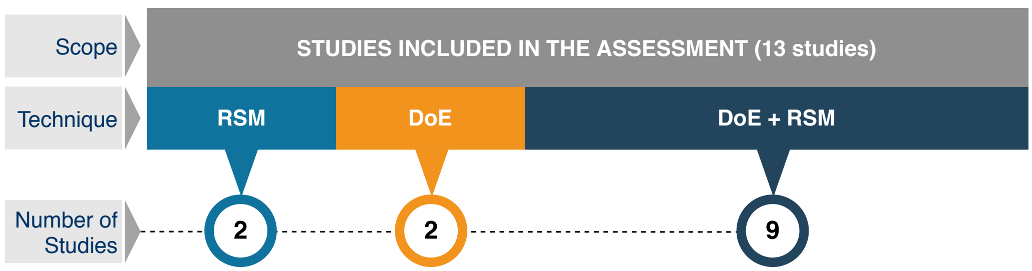


Figure 1: Overview of the techniques used to build surrogate models found in the literature review.

The literature research indicated that surrogate models built with the aid of DoE or RSM can help enhancing coal-fired power plant efficiency by identifying significant operation factors, establishing the functional relationship between them and building a strategy to optimize operation. However, these studies did not explore the full potential of integrated tools. Each study explored one of its advantages, such as the system optimization or the parameter selection, but none of them included the interaction between parameters, ranking the parameters by order of importance, and construction of a surrogate model for quick response. Not only the full potential of the tools was not explored simultaneously, but studies were not applied to coal-fired power plants or were applied at different conditions.

Besides, the main question of this study is about the identification and recognition that systems can deviate from their design operational range, which also evidence the role of manned interference, with on-field experience. Those deviations cause the system to loose performance in respect to its primary or designed efficiency state, and the present work proposes a continuous improvement tool to help process control and decision making.

The goal is to organize a comprehensive guideline designed to assist the system staff to improve performance by recovering efficiency losses, through a standard procedure in the form of an Operational Guide. The steam generator systems of the EDP PECEM 2x360MW coal-fired power plant[1] were used to develop and test the tool, to be presented along this paper.

2. Modeling approach

The methodology to regain plant efficiency and therefore improve performance was based on the construction of surrogate models to simulate the actual system, as displayed inf Fig. 2.

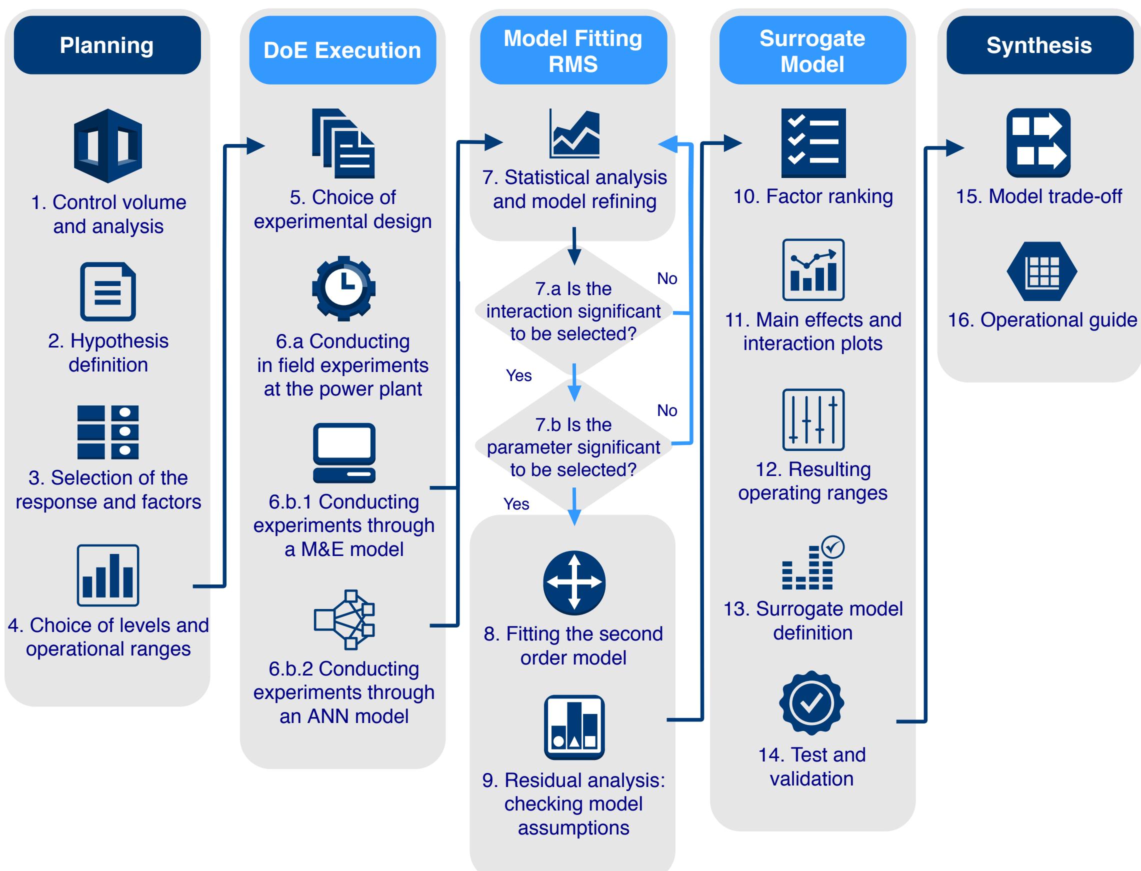


Figure 2: Methodology for the construction of a standard power plant operation procedure

[1]https://pecem.brasil.edp.com/pt-br/power-plant

The Planning phase concerns all basic and structural definitions, like bounding the study object (Step 1), followed by the application of the well-known DoE procedures (Steps 2 to 4) that provide the elements to elect a specific experimental design (Step 5). That step balances the designed number of experiments with the available time and resources, by taking into account the factor types and nature, replication, and blocking. The resulting design matrix contains the controlled factors, their levels, and the experiment running order. Step 6 deals with procedures for conducting the experiments at the power plant (step 6.a) or through simulation models. Two simulation models were proposed in this paper, a Mass and Energy model (step 6.b.1) and an Artificial Neural Network model (step 6.b.2). The simulation models proposed in this paper are a Mass and Energy model (step 6.b.1) and an Artificial Neural Network model (step 6.b.2). A detailed study of Artificial Neural Networks and Design of Experiments approach has already been reported by [6].

The model fitting RMS phase builds a Response Surface Model (RSM) from the collected data. Step 7 performs an Analysis of Variance (ANOVA) hypothesis test, based on the definition of a confidence interval and its complementary significance level (α). Hierarchical factor interactions are tested in step 7.a. The null hypothesis H_0 is rejected for (p-value $< \alpha$), meaning that the interaction is significant to be selected, otherwise (p-value $> \alpha$) it is removed from the model. Process is restarted from step 7, and it is repeated until all remaining model interactions are identified as significant. Step 7.b tests the significance of individual factors.

The fitting of the second-order model (Step 8) is followed by the checking of model assumptions in Step 9, normality, constant variance, and independence. Four residual plots are generated in this step (normal probability, residual histogram, residual versus fitted values, and residual versus observation order). The simplest option to produce random residuals is a good candidate for a relatively precise and unbiased model. A new analysis would become necessary only if some of the model assumptions could not be verified, which can be caused by the missing of variables, higher-order terms of a variable in the model to explain the curvature or a missing interaction between terms already in the model [7, 8].

The Surrogate model phase aims to build mathematical expressions to standardize the operation. Step 10 performs the controlled parameter ranking in descending order of importance in respect to the response, in order to find the optimal settings that minimizes variability. Key parameters are identified and ranked on a Pareto plot. The main effect and interaction plot is built in the 11^{th} step, which helps identifying the settings that yield to improve steam generator efficiency. Regions of operating ranges are settled in Step 12 based on the seek of identifying the best possible responses. Lines of constant yield are connected to form response contours using contour plots, which are projections on the interest regions [8].

Step 13 defines the surrogate model which assures that only the significant terms are considered, whose test and validation is performed in Step 14. The model trade-off is carried out in Step 15, by comparing all the available ones to select the more appropriate. These former steps must assure the competence of the chosen surrogate model to represent the system behavior.

Finally, an operational guide is then synthesized in Step 16, which is the main product of the methodology for standardizing power plant operations, providing a sequence of maneuvers to the operator focused on increasing efficiency along with the plant stability.

The operational guide is applied based on the observation of a necessity or opportunity for efficiency improvement in respect to the current operating condition. Figure 3 presents the cyclic scheme of the proposed methodology.

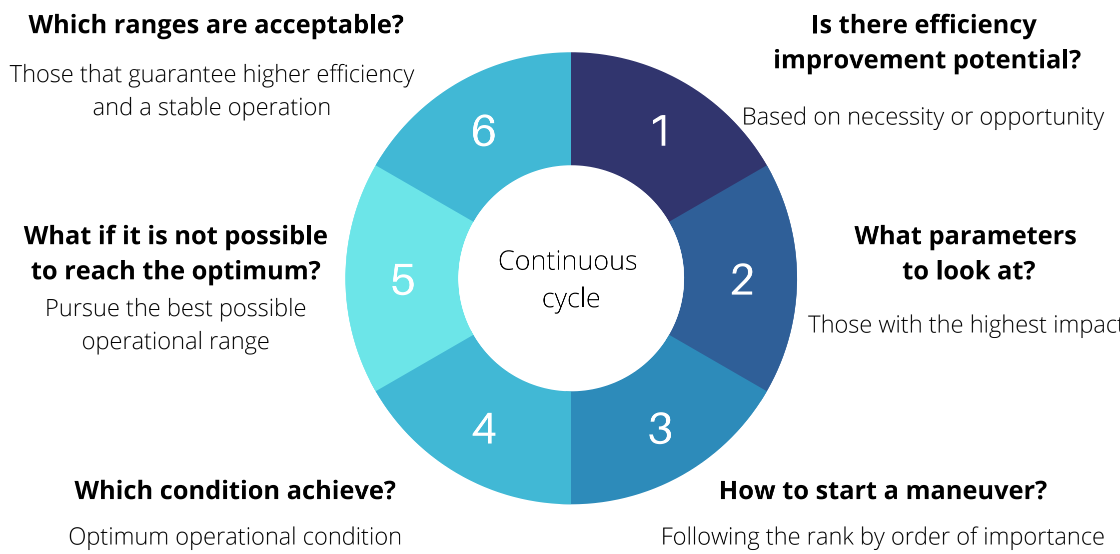


Figure 3: Operational guide to increase efficiency by standardizing the operation

3. Pecem power plant: a logbook to build a surrogate model

PECEM I [2] is composed by two independent sub-critical coal-fired power units of 360MW electric power output each. The identical steam generators are equipped with heat exchangers such as superheaters, reheaters, economizers and air heaters, arranged to efficiently absorb heat released by fuel combustion and deliver steam at rated temperature, pressure and capacity. These last parameters determine the steam generator configuration [9, 10]. Independent mills feed each steam generator with dry pulverised coal.

The modeling approach presented in section 2 was applied to the power group B running at the base load of 360 MW. The schematic layout of the subset steam generator and its mills is presented in Figure 4.

The natural choice of control volume (CV) is around the steam generator, but it was enlarged to contain the coal mills, as they are directly related to the system performance. Coal consumption and granulometry, air flow rate, flame stability, among others, are all influenced by the mill activity. Three out of four mills are connected to the steam generator at base load operation to rationalize costs and maintenance. Mill operation can be assessed individually or by their average. The former option was adopted as it allows to notice and detail each equipment condition. Burners are arranged in four rows placed on the furnace front and rear walls, with six burners each. One mill feeds a burner line of six pulverized coal burners, placed in independent wind boxes.

It is worth recalling that this paper research goal is to standardize the operation of the steam generator subset in order to improve its performance and thereafter impact the plant overall behaviour. The significance of the system controllable parameters and their interactions are the hypothesis to be tested to reduce process variability related to the operator actions.

3.1. Planning

The classification was based on controllable and uncontrollable parameters. The formers are the ones that can be directly impacted by the actions of the unit control operator [1]. Uncontrollable parameters or noise variables are those which are difficult or expensive to control or that cannot be controlled, but they can be monitored and included in models. Parameters were identified based on both literature and technical advisoring from PECEM engineering staff, and the data were recovered through their KKS codes (identification system for power stations). Eleven controllable parameters

[2] https://pecem.brasil.edp.com/en/power-plant.

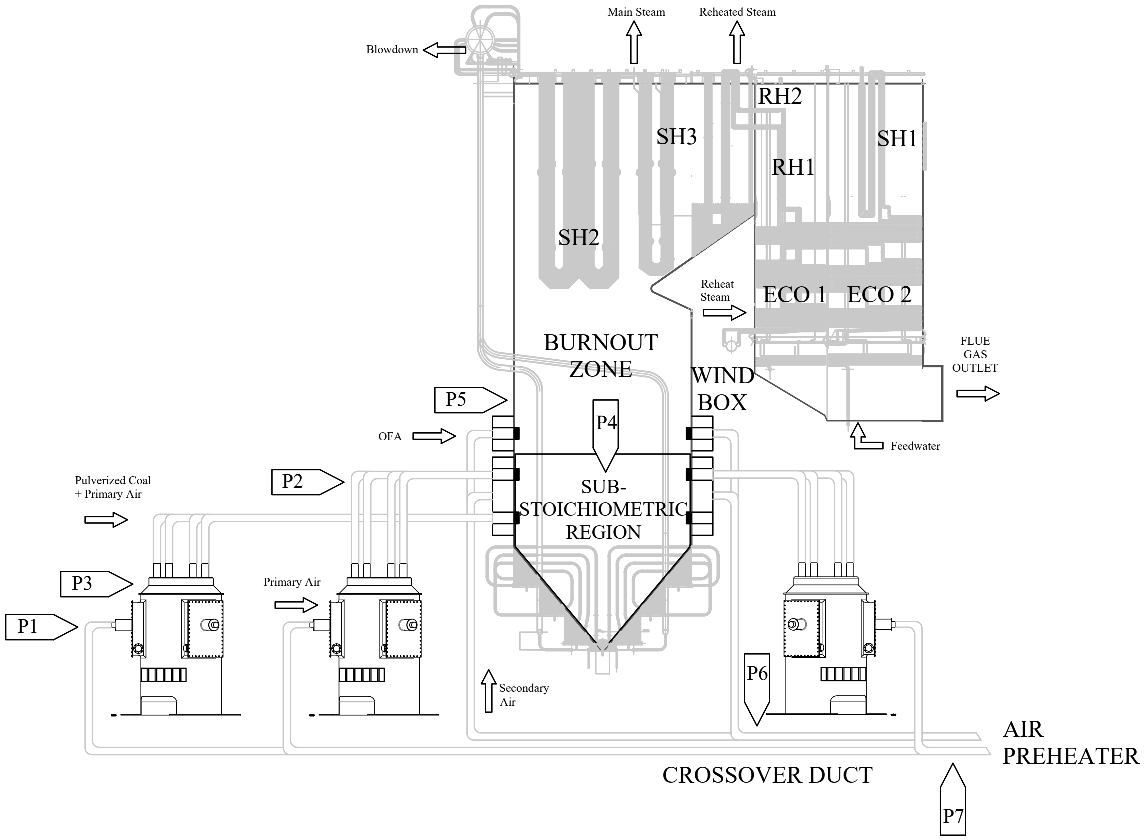


Figure 4: Schematic layout of the steam generator and mills of PECEM power plant, indicating the selected parameters.

were identified but only seven of them were kept due to the actual possibility of action on plant maneuvers. These 7 parameters were labeled from P1 to P7 and the response as S1 (see Table 1).

Controlable factors P1 to P3 concerned the mills and the remaining ones were related to the steam generator. The efficiency S1 was selected as the response since it represents the performance of the steam generator subset in a single parameter. The primary air flow P1 performs coal drying and transport it to the burners, already pulverized, and this parameter is directly related to coal granulometry [11]. Pulverized coal outlet temperature (P2) at the mill outlet is related to the coal drying process, and operates around 80°C . The speed of the dynamic classifier (P3) is the last parameter related to the mill, and it impacts directly the coal granulometry, along with P1.

The combustion air flow demand is coupled to the system steam production and its fuel flow rate. Combustion takes place in the furnace sub-stoichiometric region and completed at the burnout zone, with the injection of Over Fired Air (OFA). Stoichiometry (P4) refers to the sub-stoichiometric region, whose value is set below 1.0. Coal and air are rapidly mixed and burned in the furnace under such sub-stoichiometric conditions, to be completed with extra oxygen from the over fire air (OFA) ports at the burnout zone.

The secondary air stream is directly connected to the wind box, pressure-balanced, and admitted in the combustor by the OFAs, arranged in two rows with six injectors each above the top rows of the pulverized fuel burners. The excess of O_2 (P5) refers to the burnout zone and it defines the global stoichiometry of the combustion process and commands the OFA operation. The combustion total air is the summation of the primary, secondary and over-firing air flows, and its global stoichiometry is kept approximately constant around 1.2.

The last two controllable parameters are the primary and secondary air pressures (P6 and P7). Preheated air from a common heating device at approximately 300°C (air preheater) is splited into the two feeding paths via independent crossover ducts. The primary crossover duct supplies hot primary air to the mills, while the secondary crossover duct delivers hot air to the burner windboxes [11] and OFA . This arrangement maintains the correct secondary air pressure for all firing conditions. The crossover duct pressure acts on flame stability and distribution.

From the steam generator side, feedwater is admitted in countercurrent to the flue gases at the economizers (ECO1 and ECO2), to evaporate at the furnace water walls and superheated on three superheaters (SH1, SH2, and SH3). Both the main steam stream and the reheated stream feed the cycle turbine at 180 bara, and 36 bara, respectively, both at 540°C.

The operating range of the controllable parameters are determined according to the plant history to provide safe and stable conditions. Experiments must not cause additional stresses to the power plant, but to standardize operation ensuring safety.

The historical data does not enable the knowledge of the real conditions of the power plant at a given moment, as they do not allow to conclude if the power plant is under normal operation. For instance, coal moisture due to the rain, or unusual equipment behavior, cannot be observed with this approach. Thus, the performance of experiments through DoE is essential due to its ability of covering wide operational ranges.

Table 1 summarizes the main values collected for the controllable parameters for group 2 operating on the 340 to 360 MW range.

Table 1: Summary of factors (controllable parameters) operation range and respective levels

Factor	**Lower Level**	**Medium Level**	**Upper Level**	**Description**
P1 (kg/s)	24.0	26.0	28.0	Primary air mass flow rate
P2 (°C)	65	75	85	Pulverized coal outlet temperature
P3 (rpm)	90	100	110	Speed of the dynamic classifier
P4 (kg/s)	55	66	77	Secondary air mass flow rate
P5 (%)	1.5	2.3	3.0	O_2 excess
P6 (mbar)	18	21	23	Secondary air pressure
P7 (mbar)	70	78	85	Primary air pressure

The operational range of each parameter was defined with the assistance of the PECEM technical team, and limits were changed according to their experience and recommendations

3.2. Design of Experiments Execution

The choice of the experimental design corresponds to the 5th step on Figure 2 and involves the sample size, number of replicates, selection of the randomized order of experimentation, need for blocking and constrain analysis.

The number of required experiments by the full factorial 3^k is expressively bigger than the Box-Behnken Design (BBD) and Central Composite Design (CCD). The advantages of BBD become more evident with the increased number of factors. BBD stood out with 62 experiments for 7 parameters in the present study case whereas 3^k indicated 2187 experiments. On the top of it, BBD does not need to perform experiments at the range limits or extremes, whose advantage is to avoid expensive or not viable situations [12, 8].

3.2.1. Experiments on site

Experiment setting is prescribed in the sixth and last step of the planning and execution phase (Figure 2), whose major concerns are discussed in this section.

Samples of five different types of coal were collected from the plant coal stockyard, based on the [13] methodology to verify if the incoming raw coal composition was in agreement with the one declared by the supplier. Results discarded the eventual blending of the delivered coal shipments, excluding the need to consider coal type as a DoE factor. Water content can highly impact the steam generator operation and it is not a controllable parameter. Pulverized coal outlet temperature (P2) is expected to be around 85°C, but it can drop down to 65°C on rainy days. That knowleadge helped to restrict the experiments to the drought season, from August to December. Coal water content was measured twice a day, colected at the storage silo due to safety reasons.

The impacts of the primary air flow (P1) and the dynamic classifier speed (P3) on coal granulometry and ash carbon content were checked by collecting both pulverized and unburned coal samples out of all level combinations of P1 and P3.

The execution schedule is part of the design matrix and must be discussed with the operators previously to the conduction of the experiments. The schedule includes the adjustment of the start time to lounch the first factor, whose calling order is not relevant. The prescribed values for factors P1 to P7 are hardly reached, but they must be in accordance with their corresponding uncertainties. The adjustment end-time refers to the moment when factors effectively reached their prescribed values, under stable condition. External and non-controllable factors must be monitored in order to avoid interferences in the steam generator operation, like changes in the condensing system or soot blowers effects. Attention must be paid whenever a different stockpile is selected to feed the combustion system since experimental data is acquired for a given type of coal. Sootblowing was also stopped for around 30 min. Actual maneuvers were slowly conducted in order to avoid plant destabilization. The controllable parameters were set one at a time, allowing to observe the development of the operation and to ensure safe control. Results of the performed experiments are presented in Table 2.

The total amount of coal flow is the summation of individual mill flows. Experiments were conducted at mill A, and coal flow was fixed to 45 ton/h. The survey of data quality was carried out for all KKS directly related to the experiment.

Table 2: Performed experiments at the PECEM power plant following a DoE planning

							Factors (controllable parameters)								
Date	Experiment number	Responsible operator	Adjustments start time	Adjustments end time	Experiment end time	Coal Stockpile	P1 (KKS)	P2	P3 (KSS)	P4 (KKS)	P5 (KSS)	P6 (KSS)	P7 (KSS)	Sampling	S1 (KKS)
8/12/2019	1	Operator A	11:36	13:05	14:20	2D	26.0	65	90	0.88	2.3	23	78		91.20
8/12/2019	2	Operator A	14:20	15:23	16:24	2D	24.0	75	100	0.88	2.3	18	70	proceed sampling	91.10
8/12/2019	3	Operator B	16:30	17:46	18:43	3A	26.0	75	100	0.80	3.0	23	78		90.90
8/12/2019	4	Operator B	18:47	20:05	21:05	3A	26.0	75	100	0.80	3.0	18	78		90.30
8/12/2019	5	Operator B	22:25	23:10	0:10	3A	26.0	75	100	0.88	2.3	21	78		90.30
8/13/2019	6	Operator C	9:28	10:44	11:15	3A	24.0	75	110	0.88	3.0	21	78	proceed sampling	91.00
8/13/2019	7	Operator C	11:26	12:34	13:04	3A	26.0	65	100	0.88	3.0	21	85		90.30
8/13/2019	8	Operator C	13:58	15:22	15:38	3A	28.0	75	100	0.88	2.3	23	85		90.40
8/13/2019	9	Operator C	15:51	17:08	17:14	3A	26.0	85	100	0.88	3.0	21	85		90.10
8/13/2019	10	Operator C	17:31	18:31	18:48	3A	24.0	75	100	0.88	2.3	18	85		90.00
8/13/2019	11	Operator A	20:22	21:50	22:10	3A	24.0	85	100	0.95	2.3	21	78		90.40
8/13/2019	12	Operator A	22:10				28.0	75	110	0.88	3.0	21	78		

Only 11 experiments out of the 62 planned ones were performed due to a current overload on a mill inverter during the execution of the 12^{th} experiment. Generation output dropped from 360 to 190 MW, with the primary air flow (P1) and the speed of the dynamic classifier (P3) at their upper limits.This experiment brought relevant information to a new limitation that was not within the system alarms, but must be included. Historical data showed that this current value had only been reached once before, which also caused failure and consequently stopped the mill. An investigative process was carried out, and it was found a 20 A deviation from the supervisory reading (72 A) and the actual value (92 A), acquired on the field. The inverter was replaced but still mills presented an extra current failure, causing the experiment planning to be aborted.

It was noticed a 1.2% variation in the steam generator efficiency (S1), calculated under steady state regime. Although that variation may seem small, it can represent about 10.000 tons on coal consumption per year [3].

These results from 11 runs are not conclusive and should be followed by a complete set of experiments, repetitions, the inclusion of all mills, evaluation of external factors, among other measures. It is worth highlighting that performing experiments on a plant the size of PECEM is a hard task to accomplish, despite all the strictness and attention given by the operational team.

3.3. System Representation

Experiments following the DoE planning could not be concluded at the PECEM power plant and were substituted by both a mass and energy simulation model and an artificial neural network representation.

3.3.1. Mass and Energy (M&E) model

The steam generator and mill subset modeling with the commertial software EBSILON is depicted in Figure 5. EBSILON solves non-linear mass and energy balances to simulate a variety of thermodynamic cycles, and it disposes of extensive component, fluid and fuel libraries [14].

The steam generator subset was modeled with 149 components, concerning its equipments, fluid streams and fuel. The subsystems mill (1) and steam generator furnace (2) were highlighted in the figure to expose their working parameters. Subsystem 1 input parameters are the hot and cold primary air, coal flow and sealing

[3]The calculation considered 2018 as the base year. The power plant operated 40,6% of the time in baseload.

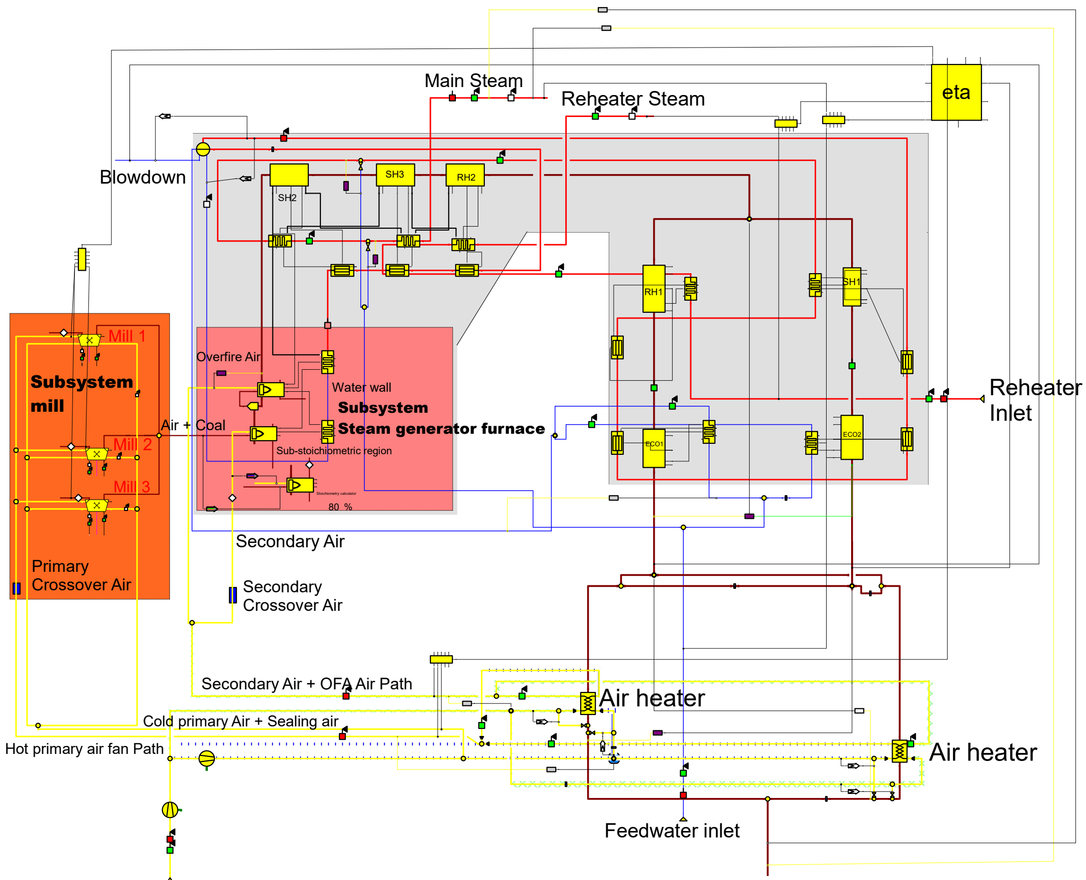


Figure 5: PECEM steam generator subset modeled by EBSILON.

air, coal and air mixture outlet temperature, and its outputs are the fuel air mixture and moisture.

Subsystem 2 segregates two volumes to represent the sub-stoichiometric combustion and the burnout zones. Inputs are the secondary air and the over fire air (OFA) flow rates, and outputs are the flue gases sent to the heat exchangers and the steam generation flow rates. Subset efficiency is calculated for design and off-design conditions, based on the direct method (15). Results for steam generator efficiency based on experimental and simulated data are presented in Table 3.

The relative deviation was calculated by the ratio between the efficiency difference of the PECEM plant and the simulation model in relation to the PECEM plant efficiency. The maximum relative deviation values for each given case reached 1.21 and the MSE is 0.3390.

Table 3: Relative deviation of real experiments at the PECEM power plant and the simulation model.

Steam Generator Efficiency (S1)			
Experiment number	PECEM power plant	Simulation model	Relative deviation
1	84.19%	83.52%	0.80
2	84.37%	83.53%	1.00
3	84.02%	83.00%	1.21
4	82.89%	83.00%	-0.13
5	83.90%	83.52%	0.45
6	83.61%	82.81%	0.96
7	83.19%	82.80%	0.47
8	83.76%	83.52%	0.29
9	82.92%	82.79%	0.16
10	82.82%	83.49%	-0.82
11	83.71%	83.46%	0.30

3.3.2. Artificial Neural Network - ANN

Data from the steam generator subset came from the Distributed Control System (DCS), on a half-hour mean value base, from August 2018 to October 2019, and concerned the 360 MW baseload ranges. That database was randomized and divided into 70% for training and 30% for testing and validation [16]. Parameters were standardized with respect to their correspondent standard deviation. ANN performance was evaluated in respect to the MSE and the MAE error metrics for training and testing and validation. ANNs were built using the Keras programming interface [17] running on top of the Tensorflow machine learning library [18].

The activation functions included ReLU (Rectified Linear Unit) and Tanh (hyperbolic tangent). The investigation process aimed at the simplest ANN capable of representing the system behavior. The chosen topology for the ANN is presented in Figure 6, and detailed in Table 4.

Table 4: Chosen ANN - Backpropagation learning algorithm and Multi-Layer Perceptron network type for 200 epochs with a batch size of 256

Input layer	**Hidden neurons**	**Hidden layers**	**Activation function**	**Output layer**	**MAE training**	**MSE training**	**MAE testing**	**MSE testing**
7	4 - 8 - 8	3	tanh-tanh-relu	1	0.0130	0.0003	0.0135	0.0003

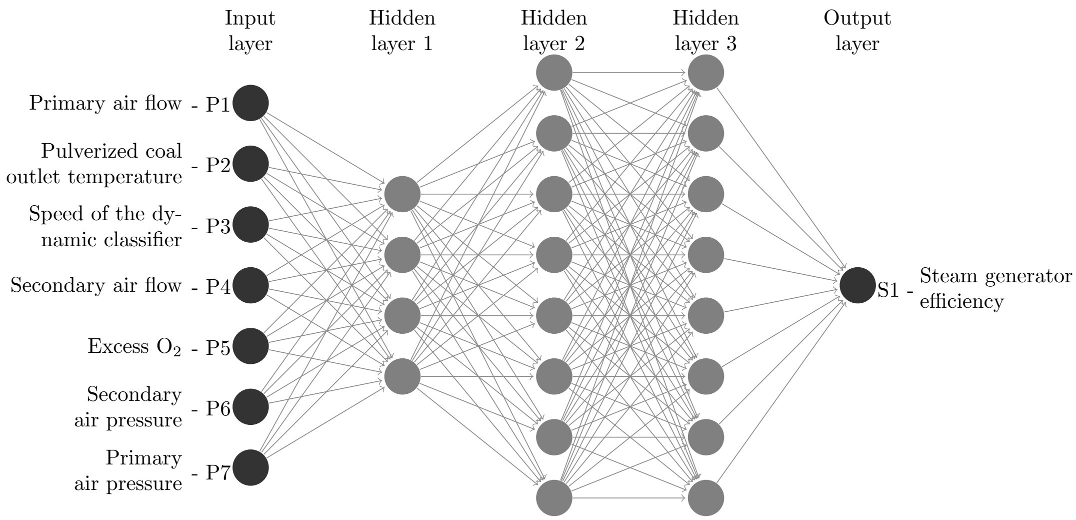


Figure 6: ANN steam generation subset model topology

3.4. DoE assessment with the simulation models

The Mass and Energy model (M&E) was able to take into account all controllable parameters but the speed of the dynamic classifier (P3), according to Table 1. The ANN model was able to take all the original parameters into account. The initial 64 number of experiments prescribed by the Box-Behnken (BBD) design approach was downsized to 54 experiments excluding the speed of the dynamic classifier for the M&E model, for the same operational range.

3.5. Model Fitting RSM

Response Surface modeling starts by appling Analysis of Variance (ANOVA) for both the M&E and ANN models (step 7). The importance of each factor and their effects were not a priori known, although inferred due to the process knowledge. The eighth step is dedicated to the second-order fitting of the steam generator efficiency (S1), whose expression intends to replace the original models proposed in steps 6.b.1 and 6.b.2 (Figure 2). Equations 1 and 2 present the RSM expressions for the M&E and ANN models, based on the 1st and 2nd order terms with statistical significance.

$$\begin{aligned} S1_{M\&E} = 38.77 + \ 0.1341\ P1 + \ 1.22355\ P2 - \ 0.05048\ P4 - \ 0.4101\ P5 \\ - \ 0.1280\ P6 - \ 0.0228\ P7 - \ 0.002193\ P1^2 - \ 0.007614\ P2^2 \\ + \ 0.000390\ P4^2 - \ 0.04651\ P5^2 + \ 0.003144\ P6^2 + \ 0.000147\ P7^2 \\ - \ 0.000422\ P1P4 - \ 0.01172\ P1P5 \end{aligned} \tag{1}$$

$$
\begin{aligned}
S1_{ANN} = & -8.57 + \ 0.2314P1 + \ 0.06444P2 + \ 0.04946P3 + 0.01863 \ P4 \\
& + \ 0.4439P5 - \ 0.01851P6 + \ 0.01097P7 - \ 0.001293P1^2 \\
& - \ 0.000098P2^2 - \ 0.000114P3^2 - \ 0.00759P5^2 - \ 0.000065P7^2 \\
& - 0.000632P1P2 - \ 0.000547P1P3 - \ 0.00825P1P5 - 0.000505P1P7 \\
& - \ 0.000163P2P3 - \ 0.000208P2P4 - \ 0.000071P3P4 - 0.001365P3P5 \\
& + \ 0.000187P3P6 + \ 0.000061P3P7 - \ 0.000825P4P5 + \ 0.000069P4P7
\end{aligned}
\tag{2}
$$

Error metrics for both RSM models were found as: adjusted R^2 of 99.98% and a predicted R^2 of 99.95% with M&E and adjusted R^2 of 74.30% and a predicted R^2 of 64.00% with ANN which can be considered suitable to calculate S1 [19]. M&E errors displayed better results compared to ANN, although the former model does not account for noise and variability, such as the latter does.

The objective of the 9th step is to check the model assumptions of normality, constant variance, and independence through residual plots. Both approaches showed symmetrical data distribution, the residual versus fitted values showed a random distribution around the zero line with constant variance, and no recognizable patterns or trends on the residuals. The present step concludes the second phase model fitting - RSM.

The algebraic RSM expressions (Eq. 1 and 2) were used to rank the factors by order of importance (step 10) in respect to the system efficiency S1. The coeffients are described in Table 5.

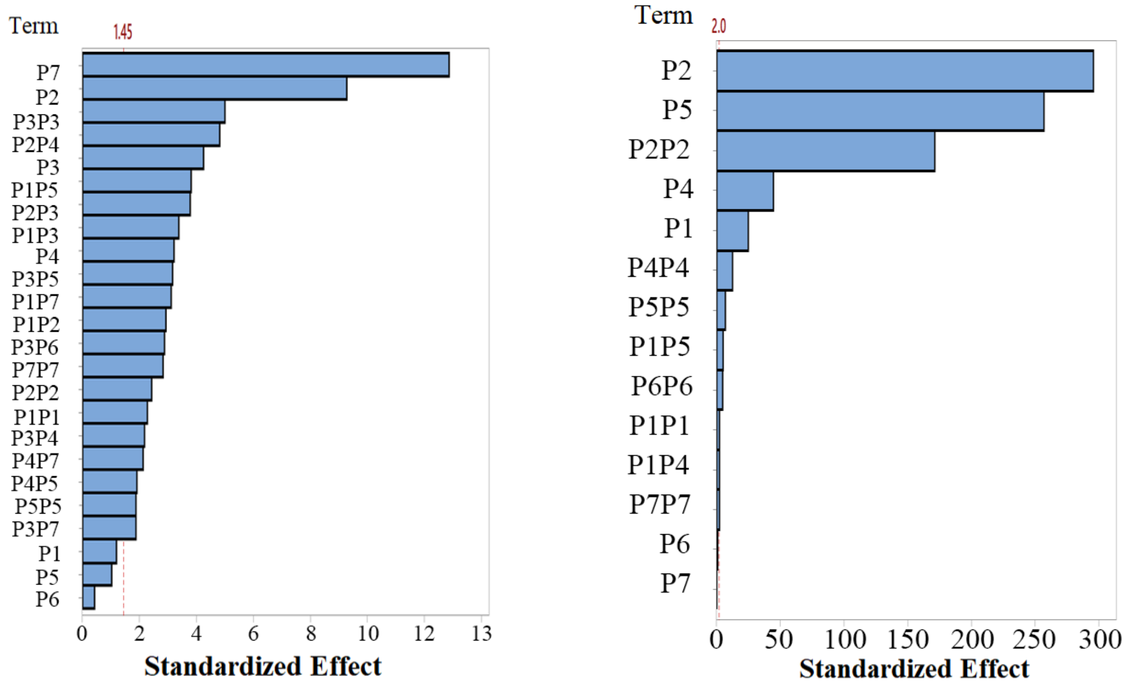
Table 5: Regression coefficients of the second-order models in terms of coded and uncoded coefficients from the M&E and the ANN model. P1 (Primary air flow), P2 (Pulverized coal outlet), P3 (Speed of the dynamic classifier),P4 (Secondary air flow), P5 (O_2 excess), P6 (Secondary air pressure), P7 (Primary air pressure)

	Coded coefficient		**Uncoded coefficient**	
	M&E model	ANN model	M&E model	ANN model
Constant	83.3262	0.81582	43.44	-8.57
Linear				
P1	-0.0684	-0.00223	0.1604	0.2314
P2	0.8141	0.01735	1.22355	0.06444
P3	N.A.	0.00796	N.A.	0.04946
P4	-0.1095	-0.00602	-14.54	0.01863
P5	-0.693	0.00191	-0.4101	0.4439
P6	0.0022	0.0008	-0.128	- 0.01851
P7	-0.0002	-0.02407	0.0228	0.01097
Square				
P1*P1	-0.0088	-0.00517	-0.002193	- 0.001293
P2*P2	-0.7614	-0.00552	-0.007614	- 0.000098
P3*P3	N.A.	-0.0114	N.A.	- 0.000114
P4*P4	0.0472	-	8.394	-
P5*P5	-0.0262	-0.00427	-0.04651	- 0.00759
P6*P6	0.0197	-	0.003144	-
P7*P7	0.0083	-0.00646	0.000147	- 0.000065
2 - Way interaction				
P1*P2	-	-0.00948	-	- 0.000632
P1*P3	N.A.	-0.01094	N.A.	- 0.000547
P1*P4	-0.0093	-	-0.062	-
P1*P5	-0.0176	-0.01238	-0.01172	- 0.00825
P1*P6	-	-	-	-
P1*P7	-	-0.01011	-	-0.000505
P2*P3	N.A.	-0.01225	N.A.	- 0.000163
P2*P4	-	-0.01562	-	- 0.000208
P3*P4	N.A.	-0.00708	N.A.	- 0.000071
P3*P5	N.A.	-0.01023	N.A.	-0.001365
P3*P6	N.A.	0.00934	N.A.	0.000187
P3*P7	N.A.	0.00607	N.A.	0.000061
P4*P5	-	-0.00619	-	- 0.000825
P4*P6	-	-	-	-
P4*P7	-	0.00687	-	0.000069

N.A. - Not applicable

The coefficient signal indicates direct or indirect proportionality with respect to the equation response. Uncoded coefficients allow expressing the equation response in a more meaningful way as they express the original operational scales and ranges as in Equation 1 and 2. Ranking was performed using coded coefficients, calculated as if all the factors were to be varied in the same range. Therefore, different orders of magnitude of the factors won't impact results. Similar results to the coded coefficients can be observed graphically through the Pareto chart presented in Figure 7.

Factors by order of importance in respect to the system efficiency (S1) for the M&E simulation were identifyied as the pulverized coal outlet temperature (P2), excess O_2 (P5), secondary air flow (P4), primary air flow (P1), secondary air pressure (P6), and primary air pressure (P7), while the ones in respecto to the ANN were the primary air pressure (P7), pulverized coal outlet temperature (P2), speed of the



(a) Pareto plot of the most important parameters of the ANN

(b) Pareto plot of the most important parameters of the M&E model

Figure 7: Pareto chart of the standardized effects (response S1, α=0.05). P1 (Primary air flow), P2 (Pulverized coal outlet), P3 (Speed of the dynamic classifier),P4 (Secondary air flow), P5 (O_2 excess), P6 (Secondary air pressure), P7 (Primary air pressure)

dynamic classifier (P3), secondary air flow (P4), primary air flow (P1), excess O_2 (P5), secondary air pressure (P6). Single effect on the steam generator efficiency (S1) in respect to the selected factorsare displayed in Figure 8.

The slope is proportional to the effect. Factors P2 and P5 displayed a significant impact on S1 variation, specially for the M&E model, also pointed out by the Pareto chart (Figure 7). S1 presented a wide range of variation in respect to P7 for the ANN model. Primary air pressure (P7) in the M&E model show negletible impact in air enthalpy and amount of energy.

The steam generator efficiency (S1) increases as the primary air flow (P1), stoichiometry (P4) and excess O_2 (P5) decrease. In the burning process, the more air is presented the greater the energy is used to promote the combustion. It is worth remembering that the Excess O_2 (P5) is related to the burnout zone (Figure 4). Regarding the pulverized coal outlet temperature (P2) the higher the temperature

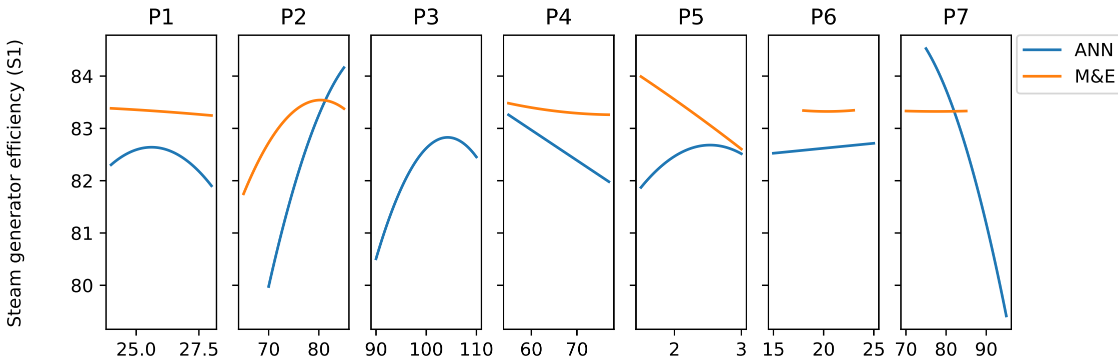


Figure 8: Main effects plot for the response steam generator efficiency (S1). P1 (Primary air flow), P2 (Pulverized coal outlet), P3 (Speed of the dynamic classifier),P4 (Secondary air flow), P5 (O_2 excess), P6 (Secondary air pressure), P7 (Primary air pressure)

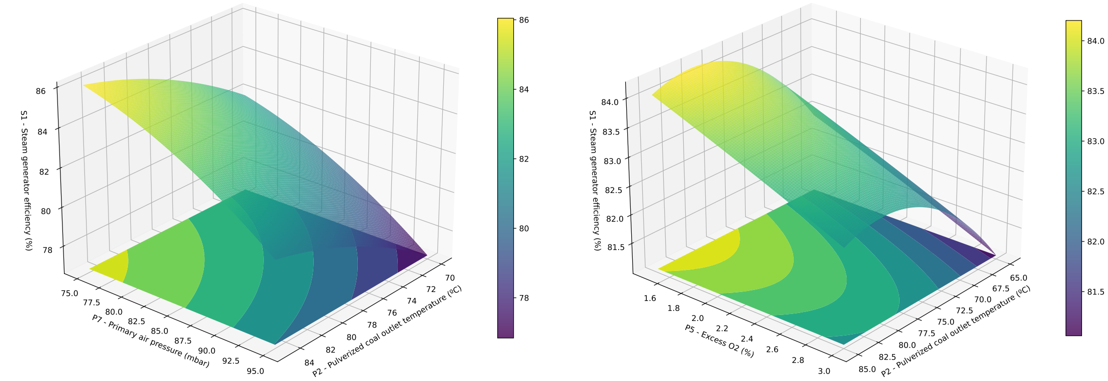
of the pulverized coal the better for the burning process. This temperature must be high enough to remove coal moisture, however, it cannot be so high as to cause the auto-ignition process. The condition of higher efficiency is around 80°C, corresponding to the nominal operating point of the mills for the M&E model. Analyzing the M&E model the least impact on the steam generator efficiency (S1) are the secondary and primary air pressure (P6 and P7).

Contour plots display response surfaces as a two-dimensional plane with response isolines, and indicate the parameter ranges. Graphs are assembled by pairs of factors, while all others parameters are hold at their average values. Factors P2 and P5 showed to be the more relevant in respect to the system efficiency S1 for the M&E model while for the ANN the P2 and P7. These contour plots are presented in Figure 9.

The highest steam generator efficiency (S1) ranges are related to high pulverized coal outlet temperatures (P2), low values for the excess O_2, up to around 1.6%, and low primary air pressures (P7). The higher the pulverized coal outlet temperature the better the process, because the energy required to burn coal will be less. On the other hand, the higher the excess O_2 more energy will be required to burn the air, decreasing the efficiency of the process.

3.6. Choice of model

Both models were able to satisfactorily represent the system. The M&E model presented a MSE of 0.03390 and the ANN model a MSE of 0.0003, which is significantly lower than the latter. The ANN proved to be superior in capturing the complete set of relevant factors, their nonlinear behavior and interconnections, as



(a) Contour plot of the most important parameters of the ANN

(b) Contour plot of the most important parameters of the simulation model

Figure 9: The two contour plots represent the most important parameters in efficiency impact for the two models considered

it took into account the speed of the dynamic classifier (P3), and individualize the activity of each mill. Therefore, the ANN model was chosen as the surrogate model to be used.

3.7. Summary and Result Discussions

This section aims to identify how to improve the steam generator performance through the previous results, answering the operational guide main question (Fig. 3): "How to increase efficiency by standardizing the operation?". The surrogate model was applied to determine operational maneuvers to assure best-operating conditions following pre-defined operating ranges.

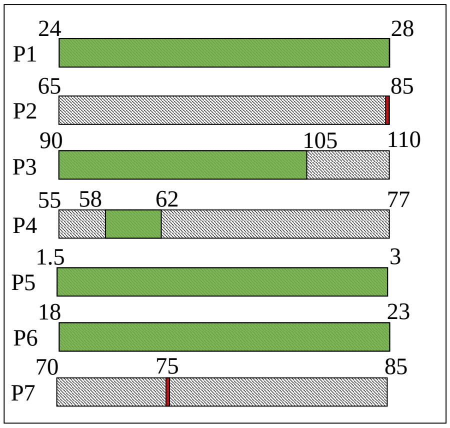
Historical data showed that the steam generator efficiency varies from 80.80 to 88.92% provided that the controllable parameters were kept within the ranges presented in Table 1. A narrower efficiency range was defined with 84.00% and 88,92% as its lower and upper limits, based on the factor frequency. The upper limit was achieved with the unique combination displayed in Table 6, while any other state within that efficiency domain can result from multiple combinations of the controllable parameters.

Figure 10 presents four scenarios whose parameter combination lead to steam generator efficiency within the 84.00% to 88,92% range. The pulverized coal outlet temperature (P2) and the primary air pressure (P7) were identified as the most important factors for the ANN model (Fig. 7a), and their values were randomly

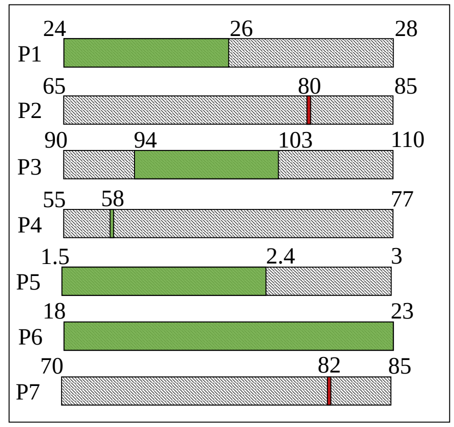
Table 6: Reference state to achieve 88.92% steam generator efficiency (S1)

S1 = 88.92%						
P1 (kg/s)	P2 (°C)	P3 (rpm)	P4 (kg/s)	P5 (%)	P6 (mbar)	P7 (mbar)
24.8	85	92	58.0	3.0	15	75

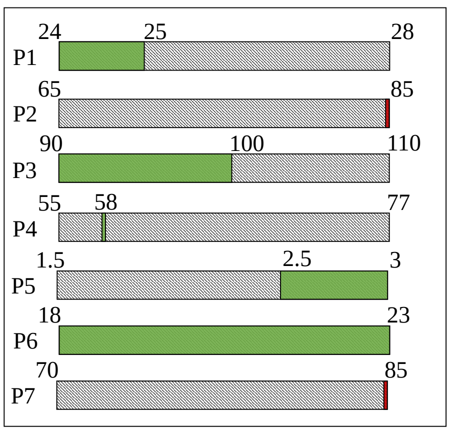
preset at constant values, represented by red narrow bars. Green intervals delineate the recommended operational range within the allowable parameter range spam.



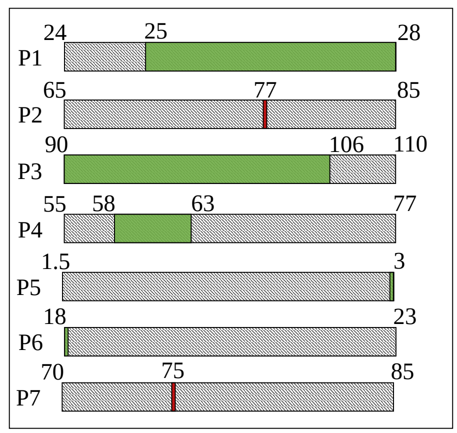
(a) Possible range values with parameters P2=85°C and P7=75 mbar



(b) Possible range values with parameters P2=80°C and P7=82 mbar



(c) Possible range values with parameters P2=85°C and P7=85 mbar



(d) Possible range values with parameters P2=77°C and P7=75 mbar

Figure 10: Operation maneuvers to assure best-operating conditions. P1 (Primary air flow), P2 (Pulverized coal outlet), P3 (Speed of the dynamic classifier),P4 (Secondary air flow), P5 (O_2 excess), P6 (Secondary air pressure), P7 (Primary air pressure)

Scenario Fig. 10a allows for more flexibility concerning factors P1, P5 and P6, who displayed a wide exploitation along their ranges. For all other scenarios, the secondary airflow (P4) displayed narrow possible spam. On the other hand, the speed of the dynamic classifier (P3) allowed to operate within a wide spam.

The plant's historical average steam generator efficiency at 360 MW base load was reported as 82.45%, which could be raised to at least 84.73% by following the operational guide. That 2.28% difference could reduce coal consumption by 3.1 t/h, with the same coal LHV of 25,750 kJ/kg. Reduction in fuel consumption can reach an annual saving of \$368.000 US and above 6% reduction on CO_2 emissions [20]. The estimate was calculated using the direct method approach [21].

4. Conclusions

The methodology proposed in this paper to regain plant efficiency and therefore improve performance was based on the construction of surrogate models and summarized in the for of an operational guide. The operational guide is dedicated to those that operate complex systems submitted to effects of aging, maintenance issues, together with a variety of operational choices of the technical staff that can bias or decrease the plant performance. It proposes operational maneuvers to assure best-operating conditions following pre-defined operating ranges developed for the steam generator and coal mills subset of a 2x360MW coal fired power plant.

A two year long database allowed to identify the subset main parameters and their levels of integration and dependency, based on a Design of Experiment (DoE) approach with on-field experience. The indicated maneuvers from the DoE matrix were first actually performed onsite and finished by ANNs and Mass and Energy balance models. Both models were able to satisfactorily represent the system. The M&E model presented a MSE of 0.03390 and the ANN model a MSE of 0.0003. The simulation of the steam generator and its mills subset following a DoE matrix generated different algebraic surrogate models, whose main output was the subset efficiency based on seven controllable input parameters: the mills' primary air flow, pulverized coal outlet, and speed of the dynamic classifier, steam generator's secondary air flow, excess O_2, and primary and secondary air pressures.

The actual application of the operational guide as a continuous improvement tool to help process control and decision making on the PECEM power plant would confirm the expected 2.28% efficiency raise during operation, which represents a reduction in coal consumption by 3.1 t/h and a consequent 6% reduction on CO_2 emissions.

5. Acknowledgments

Authors acknowledge Energy of Portugal EDP for the financial and technical support to this project; L.W.Vieira acknowledges the INCT-GD and the financial support from CAPES 23038.000776/2017-54 for her Ph.D. grant; A.D.Marques acknowledges the financial support from CNPq 132422/2020-4 for his MSc grant; J.Duarte acknowledges the financial support from CNPq 154147/2020-6 for her undergraduate scholarship; P.S. Schneider acknowledges CNPq for his research grant

(PQ 305357/2013-1). A. J. Silva Neto acknowledges FAPERJ (E-26/202.878/2017), CNPq (308958/2019-5) and CAPES (88887.311757/2018-00) for the financial support.

References

[1] GP Strategies Corporation, Heat Rate Awareness (2013).

[2] IEA, Statistics & Data (2019).
URL https://www.iea.org/

[3] X. Li, N. Wang, L. Wang, Y. Yang, F. Maréchal, Identification of optimal operating strategy of direct air-cooling condenser for Rankine cycle based power plants, Applied Energy 209 (November 2017) (2018) 153–166. doi:10.1016/j.apenergy.2017.10.081.
URL https://doi.org/10.1016/j.apenergy.2017.10.081

[4] A. Stoppato, A. Mirandola, G. Meneghetti, E. Lo Casto, On the operation strategy of steam power plants working at variable load: Technical and economic issues, Energy 37 (1) (2012) 228–236. doi:10.1016/j.energy.2011.11.042.
URL http://dx.doi.org/10.1016/j.energy.2011.11.042

[5] G. Manente, A. Toffolo, A. Lazzaretto, M. Paci, An Organic Rankine Cycle off-design model for the search of the optimal control strategy, Energy 58 (2013) 97–106. doi:10.1016/j.energy.2012.12.035.
URL http://dx.doi.org/10.1016/j.energy.2012.12.035

[6] L. W. Vieira, A. D. Marques, P. S. Schneider, A. J. da Silva Neto, F. A. C. Viana, M. Abdel-jawad, J. D. Hunt, J. C. M. Siluk, Methodology for ranking controllable parameters to enhance operation of a steam generator with a combined artificial neural network and design of experiments approach, Energy and AI 3 (2021) 100040.

[7] P. G. Mathews, Design of Experiments with MINITAB, ASQ Quality Press Milwaukee ˆ eWisconsin Wisconsin, 2005.

[8] D. C. Montgomery, Design and Analysis of Experiments, 8th Edition, John Wiley & Sons, Inc., 2013.

[9] D. Annaratone, Steam generators: Description and design (2008).

[10] The Babcock & Wilcox Company, Steam: its generation and use, 42nd Edition, The Babcock & Wilcox Company, Charlott, 2015.

[11] Doosan Babcock Energy, Pecem Units 1 & 2 Boiler Operating & Maintenance Manual, Doosan Babcock Energy, 2011.

[12] R. H. Myers, D. C. Montgomery, C. M. Anderson-Cook, Response Surface Methodology: process and product optimization using designed experiments, John Wiley & Sons, Inc., Hoboken, 2016.

[13] ASTM International, ASTM D2234 / D2234M-19 - Standard Practice for Collection of a Gross Sample of Coal (2019).
URL `www.astm.org`

[14] STEAG, Design and engineer your plants with EBSILON®Professional – STEAG's innovative, world-leading planning tool (2020).
URL `https://www.ebsilon.com/en/#header`

[15] DIN 1942, Acceptance Testing of Steam Generators (1994).

[16] A. Burkov, The hundred-page machine learning book, Vol. 1, Andriy Burkov Canada, 2019.

[17] F. Chollet, et al., Keras, `https://keras.io` (2015).

[18] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, software available from tensorflow.org (2015).
URL `https://www.tensorflow.org/`

[19] N. J. Salkind, Statistics for people who (think they) hate statistics, 6th Edition, SAGE, London, 2017.

[20] K. Burnard, S. Bhattacharya, Power generation from coal (2011).
URL `https://www.oecd-ilibrary.org/content/paper/5kg3n27ts06b-en`

[21] S. Shah, D. Adhyaru, Boiler efficiency analysis using direct method, in: 2011 Nirma University International Conference on Engineering, IEEE, 2011, pp. 1–5.

[22] A. Dresch, D. Lacerda, J. Antunes, Design Science Research: A Method for Science and Technology Advancement, Springer, 2014.
URL https://books.google.com.br/books?id=-SQAswEACAAJ

[23] M. K. Linnenluecke, M. Marrone, A. K. Singh, Conducting systematic literature reviews and bibliometric analyses, Australian Journal of Management 45 (2) (2020) 175–194.

[24] K. S. Khan, R. Kunz, J. Kleijnen, G. Antes, Five steps to conducting a systematic review, Journal of the royal society of medicine 96 (3) (2003) 118–121.

[25] S. Chandrasekharan, R. C. Panda, B. N. Swaminathan, Statistical modeling of an integrated boiler for coal fired thermal power plant, Heliyon 3 (6) (2017) e00322.

[26] S. Chandrasekharan, R. C. Panda, B. N. Swaminathan, Predictive modeling of integrated boiler unit for a coal fired thermal power plant, International Journal of Emerging Electric Power Systems 18 (3) (2017).

[27] K. Cremanns, D. Roos, S. Hecker, P. Dumstorff, H. Almstedt, C. Musch, Efficient multi-objective optimization of labyrinth seal leakage in steam turbines based on hybrid surrogate models, in: Turbo Expo: Power for Land, Sea, and Air, Vol. 49712, American Society of Mechanical Engineers, 2016, p. V02CT45A024.

[28] V. Seetharama-Yadiyal, G. Brighenti, P. Zachos, Advanced gas turbine performance modelling using response surface methods, The Aeronautical Journal 122 (1258) (2018) 1871–1883.

[29] M. Amiri, S. Shahhosseini, A. Ghaemi, Optimization of co2 capture process from simulated flue gas by dry regenerable alkali metal carbonate based adsorbent using response surface methodology, Energy & Fuels 31 (5) (2017) 5286–5296.

[30] V. S. Chandane, A. P. Rathod, K. L. Wasewar, S. S. Sonawane, Efficient cenosphere supported catalyst for the esterification of n-octanol with acetic acid, Comptes rendus chimie 20 (8) (2017) 818–826.

[31] V. S. Chandane, A. P. Rathod, K. L. Wasewar, S. S. Sonawane, Synthesis of cenosphere supported heterogeneous catalyst and its performance in esterification reaction, Chemical Engineering Communications 205 (2) (2018) 238–248.

[32] M. Kumar, L. Goswami, A. K. Singh, M. Sikandar, Valorization of coal fired-fly ash for potential heavy metal removal from the single and multi-contaminated system, Heliyon 5 (10) (2019) e02562.

[33] S. Mahanta, M. Chandrasekaran, S. Samanta, R. Arunachalam, Multi-response ann modelling and analysis on sliding wear behavior of al7075/b4c/fly ash hybrid nanocomposites, Materials Research Express 6 (8) (2019) 0850h4.

[34] L. Remenárová, M. Pipíška, E. Florková, M. Horník, M. Rozložník, J. Augustín, Zeolites from coal fly ash as efficient sorbents for cadmium ions, Clean Technologies and Environmental Policy 16 (8) (2014) 1551–1564.

[35] A. Verma, A. Rao, G. Samuelsen, Sensitivity analysis of a vision 21 coal based zero emission power plant, Journal of Power Sources 158 (1) (2006) 417–427.

[36] J. M. Wakiru, L. Pintelon, P. N. Muchiri, P. K. Chemweno, A simulation-based optimization approach evaluating maintenance and spare parts demand interaction effects, International journal of production economics 208 (2019) 329–342.

[37] Y. Xu, Y. Sun, Z. Ma, R. Wang, X. Wang, J. Wang, L. Wang, X. Gao, J. Gao, Response surface modeling and optimization of electrodialysis for reclamation of ro concentrates in coal-fired power plants, Separation Science and Technology 55 (14) (2020) 2593–2603.

[38] N. J. Van Eck, L. Waltman, Software survey: Vosviewer, a computer program for bibliometric mapping, scientometrics 84 (2) (2010) 523–538.

APPENDIX A - Systematic Literature Review

Figure 1 presents the step by step procedure for the conduction of the systematic review.

Figure 1: Steps of the systematic literature review applied in the present study, based on [22, 23, 24].

A brief description of each selected study is presented. [25, 26] used RSM supported by DoE to optimize the operating parameters of an integrated boiler unit for a coal-fired power-plant. These papers proposed the optimization of operational parameters considering as outputs the pressure and temperatures at the economizer, drum, superheater, and the integrated boiler unit. The input parameters considered by [25] were coal feed, feed water, and air, while [26] considered specific heat transfer rate of flue gas, the flow rate of feed water and enthalpy of the feedwater.

[27] proposed DoE combined to surrogate modeling to obtain an optimized design of labyrinth seal leakage in steam turbines through multi-dimensional target

functions. It is a well-developed work, which presents all the steps of creating a CFD-based surrogate model.

[28] focused on the development and application of RSM to capture the performance of a complex power system through a surrogate model. The generated surrogate model becomes part of a wider computational platform and enabled to system optimization.

[29] applied DoE and RSM on a CO_2 capture process from the flue gases of fossil fuel power plants. DoE and RSM were also applied by [30], which used a byproduct of coal-fired thermal power plants, the cenospheres, to develop a heterogeneous acid catalyst. RSM was used to optimize the various process parameters for the synthesis through Box-Behnken design. Polynomial model equations were developed to predict the esterification conversion and yield [31].

[32] performed experiments in a thermosyphon integrated thermoelectric generator using coal-fired fly ash (CFFA) collected from a local thermal power plant. The experiments were conducted through DoE and RSM to select process parameters for optimization, but there was no development of a surrogate model to represent the system.

[33] used boron carbide along with fly ash from a thermal power plant as reinforcement particles in the aluminum matrix to fabricate a new class of composite. DoE and RSM were employed for investigation of the response variables. Multiple linear regression models were obtained to establish the functional relationship between response variables and process parameters.

[34] verified the applicability of fly ash from the combustion of brown-coal in the ENO Novaky power plant (Slovak Republic) for the synthesis of zeolitic materials. A RSM using Box–Behnken design was applied for investigation of interaction and competitive effects in a binary metal system. Second-order polynomial models were obtained. In addition, Pareto graphs were used to present the effects of observed factors and their combined impacts.

[35] approached the development and demonstration of a technology for ultra-clean 21^{st} century energy plants that could effectively remove environmental concerns associated with the use of fossil fuels for producing electricity. The DoE methodology was applied to identify the significant factors that affected the system performance, ranked them to further propose system design modifications. The authors did not apply RSM nor developed surrogate models of the system.

[36] established, and selected significant optimization parameters that affected equipment performance with the objective of offering maintenance decision support on a thermal power plant. The DoE approach was utilized to quantify the effects and interactions of the variables on equipment availability and total repair time. The

paper did not stress DoE capabilities nor proposed any RSM or surrogate model. The optimization stage was conducted through simulation models.

[37] employed RSM to model and optimize the electrodialysis process for Reverse Osmosis Concentrate (ROC) reclamation in coal-fired power plants. They applied DoE combined with RSM to develop a surrogate model.

[26, 25] papers stood out for the development of individual regression models for efficient calculation of boiler performance using RSM supported by DoE with Box-Behnken design. The statistics part was well described, but none of them mentioned parameter selection, rank, or stabilization. [34] applied RSM through Box-Behnken design using multiple regression analysis to develop second-order polynomial models to a chemical process. However, they did not explain the individual impact of each parameter, their interactions or ranking.

[30, 31], [37], [29] and [32] applied RSM supported by Box-Behnken design for parameter optimization. However, just [30, 31], and [37] employed polynomial equations to describe the system of interest. Similarly, [33] applied RSM supported by central composite design. The process parameter contribution and their interactions were studied to develop equations to describe the system.

[27] developed a surrogate model based on DoE applied to simulation models, but they did not include polynomial equations, RSM, parameter raking, or the discussion of operational parameters. [28] identified the key design parameters and their impact on the system performance using RSM to build a surrogate model. The main objective was optimization. Finally, [35] and [36] applied DoE to assist the determination of critical system parameters to be optimized. [36] applied a full-factorial 2^k in a simulation model but focused on maintenance. [35] presented the ranking of the parameters as an advantage.

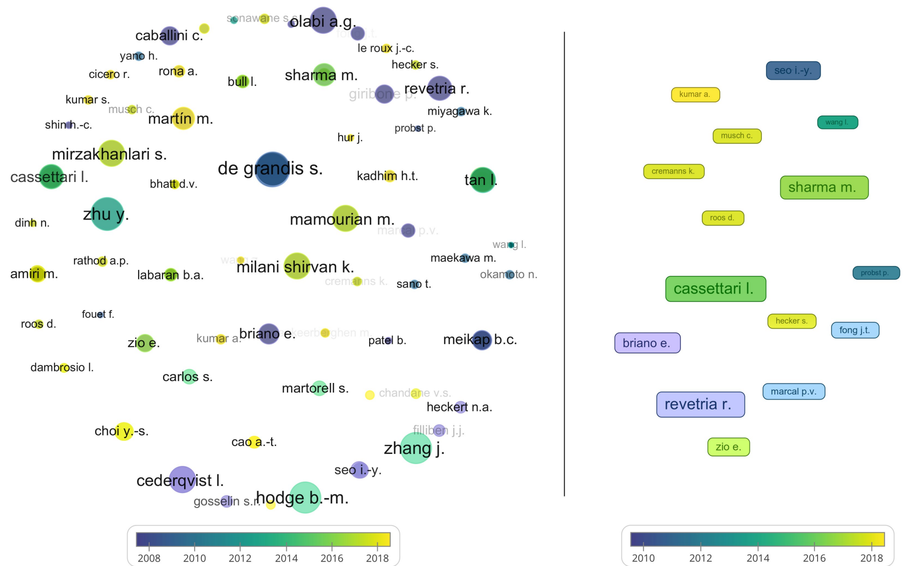
APPENDIX B - Bibliometric analysis

Bibliometric analysis was chosen in the present work to organize the sources. Results are presented hereafter based on the original 423 papers identified at the beginning of the present work, in order to provide the reader a broad view about the field. Networks maps were developed based on citations and common keywords. Papers were mainly connected to the energy theme.

The analysis was conducted by network maps with the aid of the software VOSviewer [4] [38]. The motivation was to look for trends in the topics of interest in the research literature. Color map indicates the publication year and circle size represents the number of citations. The citation network maps were developed based on the number of documents and citations per author.

Figure 2(a) displays authors with a minimum of two published papers with at least one citation per paper. Restricting the analysis, Figure 2(b) presents authors with at least three published papers and with at least one citation per paper.

[4] https://www.vosviewer.com



(a) (b)

Figure 2: Citation network map. (a) Minimum of two published papers per author with at least one citation (b) Minimum of three published papers per author with at least one citation. Circle sizes represent the number of citations.

From a universe of 780 authors, the restriction of a minimum of two or three documents per author returned 97 matches and the second one drop it down to 15 authors. The reduced number of publications indicates that there are no featured authors in this specific area of surrogate modeling applied to coal power plants.

The network map was limited to the keywords related to the study theme and is presented in Figure **??**.

The first feature that stands out is the strong connection among keywords, with an emphasis on the four most cited ones: *Design of Experiments*, *Response Surface Methodology*, *Optimization*. Keywords related to *coal-fired power plants* and *surrogate models* are connected to more recent studies, and it can be identified keywords related to *performance assessment* and *process control*. It is worth mentioning the strong presence of the keyword *optimization*, which was never mentioned during the search.