

Article

Using Statistical Modeling to Predict the Electrical Energy Consumption of an Electric Arc Furnace Producing Stainless Steel

Leo S. Carlsson ^{*}, Peter B. Samuelsson  and Pär G. Jönsson

Royal Institute of Technology; Brinellvägen 23, 114 28 Stockholm, Sweden; petersam@kth.se (P.B.S.); parj@kth.se (P.G.J.)

^{*} Correspondence: leoc@kth.se

Received: 29 November 2019; Accepted: 21 December 2019; Published: 24 December 2019



Abstract: The non-linearity of the Electric Arc Furnace (EAF) process and the correlative behavior between the process variables impose challenges that have to be considered if one aims to create a statistical model that is relevant and useful in practice. In this regard, both the statistical modeling framework and the statistical tools used in the modeling pipeline must be selected with the aim of handling these challenges. To achieve this, a non-linear statistical modeling framework known as Artificial Neural Networks (ANN) has been used to predict the Electrical Energy (EE) consumption of an EAF producing stainless steel. The statistical tools Feature Importance (FI), Distance Correlation (dCor) and Kolmogorov–Smirnov (KS) tests are applied to investigate the most influencing input variables as well as reasons behind model performance differences when predicting the EE consumption on future heats. The performance, measured as kWh per heat, of the best model was comparable to the performance of the best model reported in the literature while requiring substantially fewer input variables.

Keywords: Electric Arc Furnace; electrical energy consumption; statistical modeling; machine learning; predictive modeling

1. Introduction

In the light of the increased use of Electric Arc Furnaces (EAF) to produce steel, it has become of more interest to study ways of reducing the raw material and energy consumption of the process. Not only will successful attempts reduce the environmental impact, but also increase the financial result for the company producing the steel. One common tool to evaluate new production strategies is through the act of modeling. Modeling is a valuable tool because it enables process engineers to evaluate proposed changes without interfering with the process, which may be more or less harmful to the steel plant supply chain. These modeling approaches can be, for example, physicochemical models using established relationships such as the mass- and energy balance equations [1–19]. Another approach is through Computational Fluid Dynamics (CFD) modeling [20]. While these models are mainly based on physicochemical equations, there is another type of model approach that is purely based on data. These models are known as statistical models and have frequently been used to predict the Electrical Energy (EE) consumption, per heat [21–24], or per ton of tapped steel, from the EAF [25–42].

Several gaps in previous attempts to model the EE consumption using statistical modeling have recently been highlighted [43]. One such gap is related to the black-box behavior of non-linear statistical models, such as Artificial Neural Networks (ANN), which imposes lack of trust in the model in the eyes of the process engineers. Other gaps are related to the statistical modeling framework

itself. For example, lack of a detailed description of the data- and modeling pipeline, which includes variable selection, data collection, data cleaning, parameter-search, modelling, and evaluation of model robustness. Another gap is the lack of connection between metallurgical knowledge and statistical modeling, both of which must be carefully considered if one aims to create a statistical model that is of practical use in the steel plant. In some cases, this has led to a lack of relevant variables and in some other cases an overuse of the number of variables. The aim of this article is to address these gaps in the context of the aforementioned statistical modeling of an EAF producing stainless steel.

ANN will be used to predict the EE consumption of an EAF producing stainless steel. This is a suitable algorithm for this prediction problem since some important physicochemical relations governing the EE consumption are non-linear. The choice of variables to model the EE consumption will be motivated using both metallurgical reasoning and statistical reasoning. First, the energy balance equation will be used to highlight the physicochemical relations between the variables governing the EE consumption [43]. Second, the expected correlative relation between the variables with respect to the EE consumption will be highlighted. The common thread in both approaches is the use of process knowledge in the reasoning.

The data- and modeling pipeline will be explained in detail for reproducibility and for reference to future studies aiming to use statistical models to predict the EE consumption of the EAF. To validate the models and to address the black-box behavior of the ANN models, statistical analysis tools such as Feature Importance (FI), Kolmogorov–Smirnov (KS) test, and Distance correlation (dCor), will be used. These methods have not previously been used in the context of statistical modeling of the EE consumption in the EAF. The consequences and aspects of this approach will also be discussed.

The results of the modeling shows that it is not necessary to use a high amount of input variables to create the best performing model. It is enough to select a subset of a set of well-chosen variables using process knowledge. Using dCor, KS tests, and FI, as complementary tools can aid the process engineers in both finding the most important variables with respect to the model performance and the reasons why a model performs differently on previously unseen heats (test data).

2. Background

2.1. EAF Process

The EAF is the main melting process in the mini-mill, type of steel plant. It uses raw material such as steel scrap and alloys to create molten steel for further processing in downstream processes in the steel plant.

The EAF process begins with the charging phase during which raw materials are added to the furnace. The melting phase starts when the electrodes are powered on and bored down into the raw material. This phase lasts until enough of the raw material has melted to make room for the second basket of raw material. This is followed by yet another melting phase. During the melting phases, burners are activated to remove cold spots, which facilitates an even melting behavior of the charged scrap. After most of the raw material is melted, the refining phase starts during which the steel is adjusted to a pre-specified composition. Additional raw material such as carbon and silicon are added in combination with oxygen lancing to facilitate exothermic chemical reactions. This reduces the amount of EE needed to heat the steel. Lastly, the steel is tapped into a ladle for further processing in the steel plant. Any necessary preparations are then made for the furnace, such as fettling of the refractories, before the next heat starts. This generalized EAF process is illustrated in Figure 1.

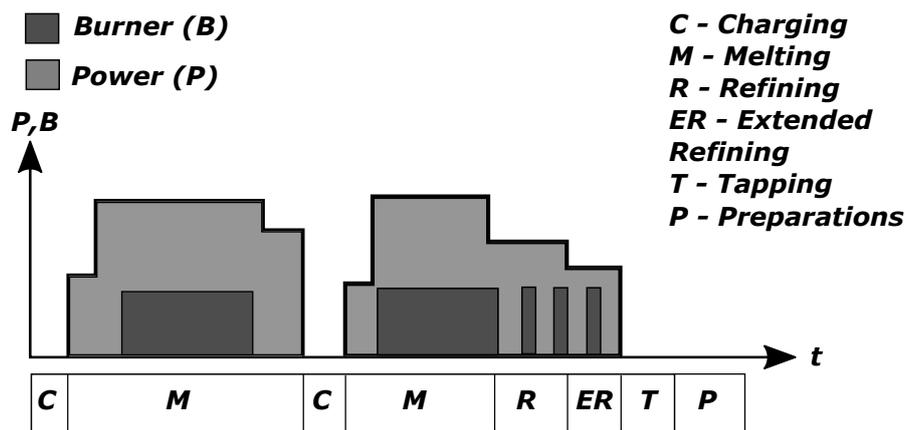


Figure 1. An idealized timeline over the EAF process for one heat. The Extended Refining is unique to the steel plant of study.

2.2. Energy Balance Equation

The energy balance equation of the EAF process makes it possible to express EE consumption, E_{El} , as a sum of ingoing and outgoing energy factors (Equations (1)–(3)).

$$E_{tot,in} = E_{El} + E_{Chem} + E_{Bu} \quad (1)$$

$$E_{tot,out} = E_{Steel} + E_{Slag} + E_{Gas} + E_{Dust} + E_{Cooling} + E_{Rad} + E_{Conv} + E_{El,loss} \quad (2)$$

$$E_{tot,in} - E_{tot,out} = 0 \quad (3)$$

Each of the energy factors are related to physical and chemical entities (see Table 1).

Table 1. Ingoing and outgoing energy terms governing the energy balance equation of the Electric Arc Furnace (EAF) process.

Energy Factor	Description	Equation	Proportionality
In E_{El}	Total Electrical Energy (EE) output from transformer	$E_{El} = \eta_{Arc} \cdot \eta_{El} \cdot P_{Avg} \cdot t_{PON}$	$E_{El} \propto t_{PON}$
E_{Chem}	Total energy from chemical reactions in steel and slag	$E_{Chem} = \sum \Delta H_{298}^0(\text{products}) - \sum \Delta H_{298}^0(\text{reactants}) + \int_{T_s}^{T_{Tap}} \left[\sum c_p(\text{products}) - \sum c_p(\text{reactants}) \right] dT$	$E_{Chem} \propto T_s; T_{Tap}; \text{composition}$
E_{Bu}	Total energy input from burner	$E_{Bu} = \eta_{Fuel} h_{Fuel} V_{Fuel}$	$E_{Bu} \propto V_{Fuel}; \eta_{Fuel}$
Out E_{Steel}	Total energy output to steel	$E_{Steel} = H_{Melt,steel} + m_{Steel} \cdot c_{Steel} \int_{T_s}^{T_{Tap}} dT$	$E_{Steel} \propto T_s; T_{Tap}; \text{composition}$
E_{Slag}	Total energy lost in slag	$E_{Slag} = H_{Melt,slag} + m_{Slag} \cdot c_{Slag} \int_{T_s}^{T_{Tap}} dT$	$E_{Slag} \propto T_s; T_{Tap}; \text{composition}$

Table 1. Cont.

Energy Factor	Description	Equation	Proportionality
Out E_{Gas}	Total energy lost in gas	$E_{Gas} = \frac{PM\dot{V}_{Gas}}{R} \cdot t_{TTT} \cdot c_{Gas} \int_{T_s}^{T_{Offgas}} \frac{dT}{T}$	$E_{Gas} \propto \ln(T_s); \ln(T_{Tap}); t_{TTT}; \dot{V}_{Gas}$
E_{Dust}	Total energy lost in dust	$E_{Dust} = \dot{m}_{Dust} \cdot t_{TTT} \cdot c_{Dust} \int_{T_s}^{T_{Offgas}} dT$	$E_{Dust} \propto T_s; T_{Tap}; t_{TTT}; \dot{m}_{Dust}$
$E_{Cooling}$	Total energy lost in cooling water	$E_{Cooling} = kA_{CS}(T_{CS} - T_{H2O}) \cdot t_{TTT}$	$E_{Cooling} \propto T_{CS}; t_{TTT}$
E_{Rad}	Total energy lost through radiation	$E_{Rad} = \epsilon\sigma A_{EAF} T_{EAF}^4 \cdot t_{TTT}$	$E_{Rad} \propto T_{EAF}^4; t_{TTT}$
E_{Conv}	Total energy lost through convection	$E_{Conv} = hA_H(T_H - T_{Amb}) \cdot t_{TTT}$	$E_{Conv} \propto T_H; T_{Amb}; t_{TTT}$
$E_{El,loss}$	Energy lost in electrical system and arc transfer	$E_{El,loss} = (1 - \eta_{Arc})(1 - \eta_{El}) \cdot E_{El}$	$E_{El,loss} \propto t_{PON}$

Calculations of ingoing and outgoing energy in the EAF process have previously been calculated [9,25,44–46]. This is compiled in Table 2.

Table 2. Ingoing and outgoing energy in the EAF process as percentages of energy sources and energy sinks, respectively. Values are computed from a synthesis of reported values [9,25,44–46].

	Energy Factor	Percentage of in and out Energy Balance
In	Electric	40–66%
	Oxidation	20–50%
	Burner/fuel	2–11%
Out	Liquid steel	45–60%
	Slag and dust	4–10%
	Off-gas	11–35%
	Cooling	8–29%
	Radiation and electrical losses	2–6%

The ingoing and outgoing energy of the EAF, as percentage of each total, provide guidance in choosing input variables to a statistical model predicting the EE consumption.

2.3. Non-Linearity

Some of the terms in the energy balance equation listed in Table 1 are related to non-linear physical phenomena in the process. Examples include, E_{Gas} and E_{Rad} . Breaking down the Tap-to-Tap time (TTT) into its smaller components Charging, Melting, Refining, Extended Refining, Tapping, and Preparation, the process evidently becomes considerably more non-linear. The reason is because the energy loss through convection, conduction, and radiation are very different from these sub-processes. For example, the energy loss through radiation is a lot higher when the steel is molten, i.e., during refining, compared to the energy loss during the first charging phase. Adding the dynamic timelines of each sub-process due to varying amounts of scrap, refining times, and various delays between heats, the physicochemical relation to the EE consumption becomes even more complex.

Because of the non-linear and non-normal distributed outcome of the process, linear statistical models, such as Multivariate Linear Regression (MLR), and linear statistical metrics, such as the Pearson correlation, are sub-optimal as tools in the context of predicting the EE consumption.

2.4. Statistical Modeling

Statistical modeling differs from other types of mathematical modeling such as those based on physical and chemical relations. See the energy balance equation in Section 2.2 for an example of a physicochemical model. Physical models act directly upon the values of the input variables to predict the output variable. Statistical models, on the other hand, act on the value with respect to previous values of the input variable it has observed in the context of the values of the other input variables. The value of the output variable is always a probability or a value with the highest probability. This means that statistical models do not adhere to the physicochemical relations between the variables. The connection between the physicochemical relations and the model is, by and large, dependent on the data quality, variations in the data, and correlative behavior between the variables. Hence, statistical models can seem counter-intuitive in the eyes of a practitioner of physicochemical-based modeling.

Data quality: A bad data quality will negatively impact the performance of a statistical model since uncertainties are imposed. There are many different ways to get good data quality. The variables must be correctly defined with respect to what they represent. This may not always be the case because everything that can go wrong is not possible to account for. For example, if the furnace is considered closed when the roof is closed, then the charging time of the furnace could still be counted if the roof is not completely shut even though the furnace is in the melting stage. Connecting plant events to logged data is not always trivial and fault free. Other examples of where errors are prevalent are from online measurements and estimations of energies based on those measurements as well as manual logging of data.

Data variability: Near constant variables add little value to a statistical model because the variations in the data are what the statistical model learns and then uses to make predictions on test data. Statistical models require variations in the data to function properly. This is perhaps one of the most counter-intuitive traits of statistical modeling since a constant term in an arbitrary physicochemical model is important for its prediction. For example, a statistical model for an EAF where the total amount of ingoing lance oxygen is constant will not receive any benefit from including that variable. This is in contrast to a physical model where the amount of added oxygen leads to exothermic chemical reactions.

Correlation: Strongly correlated variables are in a statistical sense redundant variables. Weakly correlated variables with respect to the output variable may be redundant. The degree of redundancy is closely related to the correlation between the input variable and other input variables. It may be that the input variable in question does not add any further information to the statistical model with respect to lowering the prediction error. Because of this, some variables that are considered important for predicting EE consumption in a physical model may appear as not important for the statistical model. On one hand, correlation does not imply causation, on the other hand, statistical models cannot distinguish correlation from causation. However, correlation can point to areas where causation may exist. This stresses the importance of possessing domain-specific knowledge about the process to distinguish between causative and non-causative relations between the variables governing the statistical model. Three cases of correlation related to the EAF process with respect to the EE consumption are presented, which highlight the complex relations between the different entities in the EAF process governing the EE consumption. These cases are also illustrated in Figure 2.

First, the composition of the ingoing material affects the EE consumption exothermic reactions. Here, the composition together with oxygen lancing give rise to exothermic reactions which in turn reduce the EE consumption.

Second, the various sub-processes of the EAF affect the EE consumption in various rates. For example, during the melting process, EE is added in a more intense rate than, for example, during

the refining process. Furthermore, the energy loss through radiation, convection, and conduction is more intense during the refining process when most of the steel is melted. Adding all sub-processes together gives us the total process time. The delays in the process are added to the sub-process where the delays occur. Thus, the time variables are both interrelated and have a complex relation to the EE consumption.

Third, the raw material types that are added to the EAF have a more complicated relation to EE consumption. Raw material types correlate positively to the EE consumption since they directly add to the total weight of the ingoing material to be melted. The raw material types are also correlated with one another since, given a pre-determined charge weight, the raw material types combined must make up to that specific weight. If one raw material type is not available, another is used as replacement. Raw material types can also contribute negatively to the EE consumption if the raw material type contains C, Si, Al, Fe, and Cr, all of which react exothermic with oxygen.

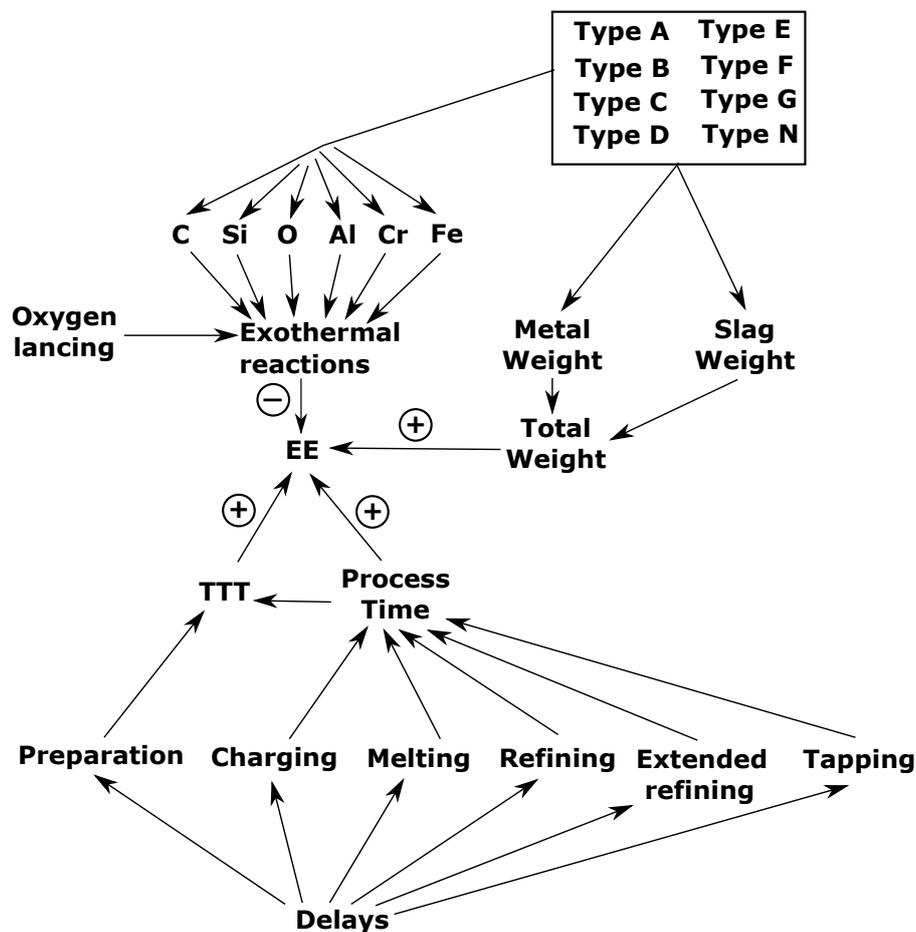


Figure 2. Graph-model describing the expected correlations between the input variables and the output variable in a statistical model predicting the EE consumption. The arrows illustrate the correlative relation between the variables. The circled signs illustrate the positive or negative contributions imposed on EE.

Variable selection comes down to how much additional useful information, with respect to predictive performance, the model gains after adding another input variable. The three points previously mentioned also tie closely to the specific EAF where the data comes from. Even though the interrelations between the input variables to the EE consumption are similar between EAF furnaces, it is not possible to know, ad hoc, the specific impact on the predictive performance by a setup of input variables.

In the supervised statistical modeling framework, which will be used in this paper, each row of input data has a corresponding output value during the training phase and test phase. The supervised statistical modeling framework can be abstracted into the following steps.

1. Select statistical model and values for the model-specific parameters.
2. Train the model using training data until the model accuracy converges, i.e., stops improving.
3. Test the model on previously unseen data.
4. Record the accuracy on test data and evaluate its practical applicability.
5. If the accuracy on the test data is satisfactory, deploy the model into production.
6. Re-train the model if the model accuracy has deteriorated. This deterioration is bound to happen over time in a production setting, partly due to changes in the process.

To find the optimal combination of model parameters, multiple models are trained for each combination of parameters. This is known as grid-search or parameter-search.

There are also subcategories of supervised statistical models. In the scope of this paper, these are divided into linear and non-linear statistical models. Linear statistical models can learn linear relationships between the input variables and the output variables. One such example is Multivariate Linear Regression (MLR). On the other hand, non-linear statistical models can learn both linear and non-linear relationships between the input variables and the output variable. One such example is ANN. Both ANN and MLR have been commonly used to predict the EE consumption in the EAF [43]. While an ANN model can learn non-linear relations in the data, the model is almost impossible to interpret. This is commonly referred to as a black-box model and is one of the main reasons why such models are not widely accepted in practice in the context of steel process modeling. However, by using statistical analysis tools, it is possible for the process engineers to get insights on which input variables are the most important for the model to make accurate predictions as well as reasons behind performance differences on test data.

Furthermore, it is always possible to create a more complex statistical model to any given problem. Nevertheless, one should always strive for model parsimony, which is achieved when picking the simplest of a group of models achieving similar performance. The model complexity is related to both the number of input variables as well as the chosen type of statistical model.

3. Method

The workflow explained in this section can be followed in Figure 3.

The software used for all parts of the experiments was Python provided by Anaconda distribution. The hardware and software specifics are shown in Appendix A.1.

3.1. Furnace Information

The data used in this study comes from an Alternate Current (AC) EAF producing stainless steel. The furnace nominal capacity is 80 tons of molten steel and the electrical system has a maximum power of 80 MW. The charged raw material is 100% scrap, and alloys. A preheater, which uses part of the furnace off-gas enthalpy as energy source and heat transfer medium, is employed to remove entrained moist and to burn off residual oil and grease in the scrap. Conventional oxy-fuel burners are used to facilitate an even temperature distribution during melting. Oxygen lancing is used to facilitate oxidation of elements such as C and Si. The number of heats produced per year is approximately 5000.

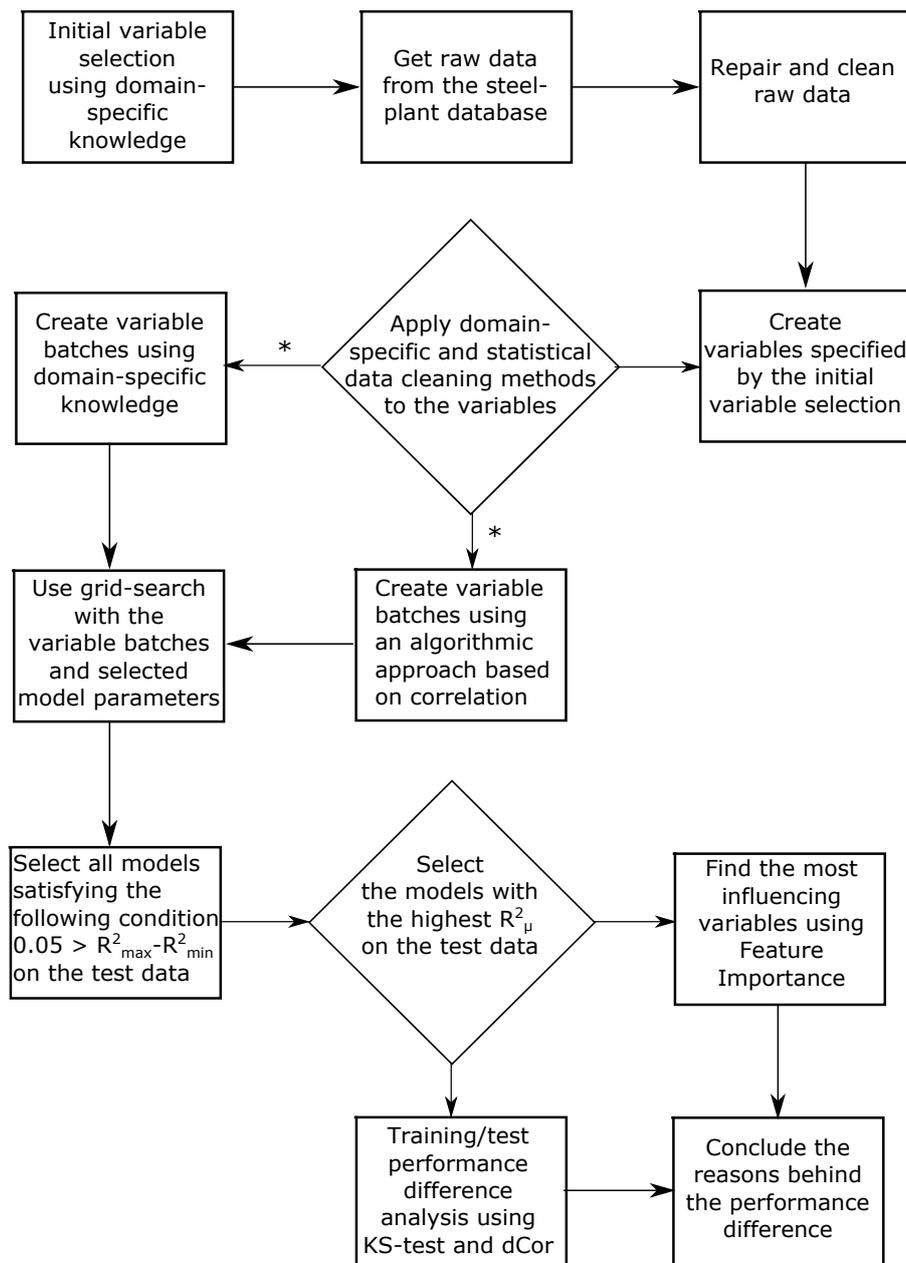


Figure 3. Flowchart explaining the data and statistical modeling pipeline. * The flowchart is split to emphasize the two different approaches used to select the variable batches for the grid-search.

3.2. Data

3.2.1. Variable Selection

The variable selection process is partly based on the discussion regarding the energy balance equation, non-linearity, and the reasoning about the correlative relations of the EAF process with respect to the EE consumption, (see Sections 2.2–2.4). The variables can be divided into four categories:

1. **Time:** The logged times for the whole process as well as the various EAF sub-processes are important for predicting the EE demand. Each process sub-stage, Charging, Melting, Extended refining, Refining, and Tapping, contribute differently to the heat loss. For example, the heat loss is higher when molten steel is present compared to when the first bucket of scrap is charged. Both the TTT and the Process Time were included. The total time imposed by delays, defined as the sum of the deviation from nominal time of each EAF sub-process, was also included.

Some obvious correlations can be expected with regards to these time variables. As Charging, Melting, Extended refining, Refining, and Tapping makes up the Process Time and the majority of TTT, these variables are expected to be relatively highly correlated.

2. **Chemical:** Oxidation and the oxyfuel-burners can account for as much as 50% and 11% of the total ingoing energy, respectively (see Table 2). The contribution by the burners was accounted for by one variable propane. The oxygen gas injected through the burners was not included due to a near stoichiometric relationship between oxygen gas and propane (5:1). The wt.% of C, Si, Al, O of the total charged metallic material and the oxygen lancing were also included to account for exothermic chemical reactions occurring during the process.

The wt.% Fe, Cr, and Ni of the total charged metallic material were included to account for possible deviations in the process due to the steel grade being produced.

The contents of the charged oxide bearing raw material are, of course, expected to impact the thermodynamics and kinetics of each reaction. However, there are more complicated factors in play between the metallic elements in the steel melt and the oxides in the slag. To account for these complex effects, the following oxides, in wt.% of total charged oxide bearing raw material, were included: Cr_2O_3 , MgO , CaO , FeO , SiO_2 , and Al_2O_3 .

It is known, by experience, that the specific heat per volume unit of oxygen is inversely proportional to total amount of lanced oxygen gas.

3. **Charged material:** The different material types are expected to contribute differently to the melting behavior and heat transfer of the scrap. Hence, all 8 of the available Material Types were included in the list of input variables. The Total Weight of the ingoing material was included because it is closely connected to the required energy to melt the steel. It was also divided into two separate variables, Metal Weight and Slag Weight, which represent the total weight of metallic material and total weight of oxide bearing raw material, respectively.

The sum of all Material Types equals the Total Weight. Hence, the Material Types are expected to be correlated with the Total Weight. Furthermore, the sum of the Slag Weight and Metal Weight equals the Total Weight and are also expected to be correlated.

4. **Energy:** There are numerous energy variables that can be included in the list of input variables. However, most are in one way or another connected to the already mentioned variables. For example, heat loss is linearly related to time and amount of ingoing scrap is linearly related to the heat required to heat and melt the scrap.

The preheating energy was included because the actual preheating is conducted before the start of the EAF process itself. In the steel plant of study, the scrap is preheated with partial off-gas from the EAF operation. Hydrocarbons such as grease and oil, and moisture entrenched in the scrap are burned off. Due to the varying amounts of moisture and hydrocarbons in the scrap, the preheating energy as ingoing variable is therefore motivated.

EE is well defined in the transformer system and is subject to a negligible error. This means that the logged EE value can be trusted and is therefore also taken as the true value when training the statistical model.

The selected variables are shown in Table 3.

Table 3. The variables used in the models.

Variables	Unit	Definition
Delays	<i>min</i>	The sum of all delays, which is defined as all deviations from the nominal time of each sub-process.
Tap-to-Tap time (TTT)	<i>min</i>	The between the end of the tapping from the previous heat to the end of tapping of the current heat
Charging	<i>min</i>	Total time needed to charge all baskets
Melting	<i>min</i>	Total melting time for all scrap baskets
Refining	<i>min</i>	Total refining time
Extended refining	<i>min</i>	Total Extended refining time.
Tapping	<i>min</i>	Total tapping time
Total Weight	<i>kg</i>	Total weight of all materials added during the EAF process
Propane	<i>m³</i>	Total amount of propane gas added by burners
O ₂ -lance	<i>m³</i>	Total amount of oxygen added by lance
Preheater energy	<i>kWh</i>	Estimated thermal energy added to the scrap baskets by the preheater
EE consumption	<i>kWh</i>	Total EE consumption for the heat. This is the output variable in the models.
Process Time	<i>min</i>	Defined as the sum of Charging, Melting, Refining, Extended refining, and Tapping.
C	<i>wt%</i>	The weight percent with respect to the total charged metallic material during the heat. Hence, added dolomite, lime, and carbon by lance are included
Si	<i>wt%</i>	
Cr	<i>wt%</i>	
Fe	<i>wt%</i>	
Ni	<i>wt%</i>	
O	<i>wt%</i>	
Al	<i>wt%</i>	
Cr ₂ O ₃	<i>wt%</i>	The weight percent with respect to the total charged oxide bearing raw material during the heat.
MgO	<i>wt%</i>	
CaO	<i>wt%</i>	
FeO	<i>wt%</i>	
SiO ₂	<i>wt%</i>	
Al ₂ O ₃	<i>wt%</i>	
Metal Weight	<i>kg</i>	Total weight of metallic material
Slag Weight	<i>kg</i>	Total weight of oxide bearing raw material
Type A	<i>kg</i>	All raw material types as defined by the plant engineers.
Type B	<i>kg</i>	
Type C	<i>kg</i>	
Type D	<i>kg</i>	
Type E	<i>kg</i>	
Type F	<i>kg</i>	
Type G	<i>kg</i>	
Type N	<i>kg</i>	

Correlations between the variables in the 4 categories are also expected. For example, EE demand should correlate with Total Weight. Another example is Material Types and chemical compositions as these material categories are partly based on chemical analyses.

The reason correlation is frequently referred to is because it is one of the main underlying factors that connects the input variables to the output variable by the trained statistical model. Furthermore, input variables that are correlated “soaks up” parts of the “correlation potential” with respect to the output variable. This is especially the case when the correlated input variables are correlated with the

output variable. A simple example is the following correlative relationship between input variables A and B, and output variable C: $A \xrightarrow{\text{corr}} B \xrightarrow{\text{corr}} C$.

3.2.2. Variable Batches

The variables in each variable batch in Table 4 are explicitly shown in Table 5.

Table 4. The variable batches governing the domain-specific variable selection approach. The variables contained within each variable group are shown in Table 5.

Variable Batch	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Base	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Process Time	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Sub-processes	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Metallic elements	x	x	x	x	x	x	x	x								
Oxide compounds	x	x	x	x					x	x	x	x				
Met-Slag weight	x	x			x	x			x	x			x	x		
Material Types	x		x		x		x		x		x		x		x	
Variable Batch	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Base	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Process Time																
Sub-processes	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Metallic elements	x	x	x	x	x	x	x	x								
Oxide compounds	x	x	x	x					x	x	x	x				
Met-Slag weight	x	x			x	x			x	x			x	x		
Material Types	x		x		x		x		x		x		x		x	
Variable Batch	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Base	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Process Time	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Sub-processes																
Metallic elements	x	x	x	x	x	x	x	x								
Oxide compounds	x	x	x	x					x	x	x	x				
Met-Slag weight	x	x			x	x			x	x			x	x		
Material Types	x		x		x		x		x		x		x		x	
Variable Batch	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
Base	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Process Time																
Sub-processes																
Metallic elements	x	x	x	x	x	x	x	x								
Oxide compounds	x	x	x	x					x	x	x	x				
Met-Slag weight	x	x			x	x			x	x			x	x		
Material Types	x		x		x		x		x		x		x		x	

The motivation behind using a setup of variable batches is to investigate the true impact of each variable type. The impact of a variable on the model outcome can be diminished if that variable is partly represented by another variable, i.e., correlated, (see Section 2.4). One clear example is the

percentage of *Cr* and a Material Type containing high amounts of *Cr*. Likewise, the total process time is partly represented by the melting time because the melting time adds to the total process time.

The decision to divide the variables into the specific variable batches shown in Table 4 was two-fold. First, there are multiple variables that represents the same effect of a physical phenomenon on the EE consumption, as motivated previously. Second, the number of model types to create given the number of combinations using each variable is very large ($2^{35} \approx 3.44 \cdot 10^9$) and it is not possible to model all variants within a reasonable time frame. Hence, bundling together variables that are of the same type, for example Material Types, will reduce the number of unique variable batch combinations to 64 (2^6).

The base variables were included in all variable batches because they are believed to be strongly related to the EE consumption. The specific variables in the base variable group are shown in Table 5.

Table 5. Input variables for each variable group. The metallic elements and oxide compounds are the composition of the total charged metallic material and oxide bearing raw material, respectively.

Variable Group	Variables	No. Variables	Variable Group	Variables	No. Variables	
Base	Delays	6	Sub-processes	Charging	5	
	TTT			Melting		
	Total Weight			Refining		
	Propane			Extended refining		
	O ₂ -lance			Tapping		
	Preheater energy		Metallic elements	C	7	
Met-Slag weight	Metal Weight	2		Si		
	Slag Weight		Cr			
Material types	Type A	8		Fe		
	Type B		Ni			
	Type C		O			
	Type D		Al			
	Type E		Oxide compounds	Cr ₂ O ₃		6
	Type F		MgO			
	Type G		CaO			
Type N	FeO					
Process Time	Process Time	1		SiO ₂		
				Al ₂ O ₃		

3.2.3. Selection of Test Data

The test data is commonly selected from a random sample of the complete data set. However, in this regard, the training data and the test data become chronologically intertwined. This shortcoming has been discussed in previous research [43]. From a practical process perspective, a statistical model will predict on data from heats that are produced after the training of model. This means that these heats will always be from a future point in time. To consider this in a model evaluation, the test data should be selected in chronological order with respect to the training data. In the experiments, the test data will be from heats produced within 30 days after the last heat in the training data.

3.3. Data Treatment

3.3.1. Purpose

The purpose of data treatment is to clean, and possibly repair, erroneous data or remove data that is not part of regular production. Data treatment is a double-edged sword. On one hand, the aim of any given model is to predict well on as many heats as possible. On the other hand, including unrepresentative data makes it harder to optimize the model enough to make it practically useful.

There are two different categories for data treatment. One is domain-specific, where the values in the data are assessed to what is physically possible within the application of the model. For example,

if the EAF has a maximum capacity of 80t, then any value above that value should be viewed as an erroneous value. The second strategy category is statistical outlier detection. For example, removing all values that are above 3σ from the mean value of a variable that is normal distributed. However, it is important to consider that statistical outlier detection methods do not take the application domain into account. These methods must be used with caution and carefully assessed on a case-to-case basis.

The data used in the modeling was aggregated from two disparate data flows. The first data flow consists of raw material compositions, raw material types, total charge weight, metal and oxide bearing raw material weights, energy data, and added propane through burners and oxygen gas through lance. The second data flow consists of process time, times for different sub-processes during the EAF process, and delay times.

All data from the two data flows was gathered after the 31 March 2016 and before 18 October 2018.

3.3.2. Domain-Specific Methods

Data flow 1: 12.1% of the raw material entries lacked oxide compositions. These entries were repaired using previously recorded oxide compositions for the same material entries. This is not expected to dramatically affect the performance of the models. An average of 176 kg of added material per heat, amounting to an approximate 0.2% of the total weight, had not available (NA) values for Rawtype, oxide, and metal composition. These raw material types were categorized into a variable called Type N. The total number of heats removed due to a raw material input with zero or negative weight was 59. The resulting number of heats at the end of the flow was 12,805.

Data flow 2: Only heats with logged events “heat started” and “heat ended” were removed. This was to safeguard against any complications such an error would give rise to in the rest of the entries. No other data treatment was commenced in this stage. A total of 12,649 heats were gathered from this data flow.

Aggregated data: Aggregating the two data flows resulted in a total of 12,587 heats. The resulting number of heats are less than from both data flows due to mismatching overlap of heats. After the aggregation of the two data flows, domain-specific data treatment rules were applied and are shown in Table 6.

Table 6. The domain-specific cleaning rules that were applied to the data. The rules were applied to the training data and test data separately to simulate an applied data filtering system. Rules applied to the training data must be applied to the test data to safeguard against data points outside the scope of the statistical model.

Filter	Motivation
Removal of all Trial heats	Trial heats are not part of regular production since the aim is primarily to investigate the properties for new scrap types
Heats with EE above 60 MWh	Identified as abnormal EE by the process engineers
Heats with Total charged weight above 110,000 kg	Physically impossible weight due to furnace size limitations
45 min < TTT < 180 min	45 min are considered unusually short and above 180 min are likely due to a longer delay in the process or a scheduled stop. Usually, the TTT is aimed at 60–70 min
Delays < 180 min	Heats with delays over 3 h are because of longer stops due to, for example, broken equipment.

3.3.3. Statistical Methods

Statistical data cleaning methods were not applied to the data set in this study. There are three reasons for this:

1. A more robust outlier detection algorithm was used in an attempt to remove outliers [47]. However, the resulting data set became too small after applying the algorithm on just a few of the variables.
2. It makes sense to not apply statistical outlier detection methods on some variables. For example, the wt% content in the charged material can vary tremendously depending on what scrap types are available. Since multiple steel grades are produced using different charging strategies, the wt% of elements will also vary. Some stainless steel types need more wt% Cr and wt% Ni than others.
3. Most conventional outlier detection methods assume that the data is normal distributed, or normal distributed with various levels of skewness and kurtosis. This is not always the case in EAF production data, see Figure 4.

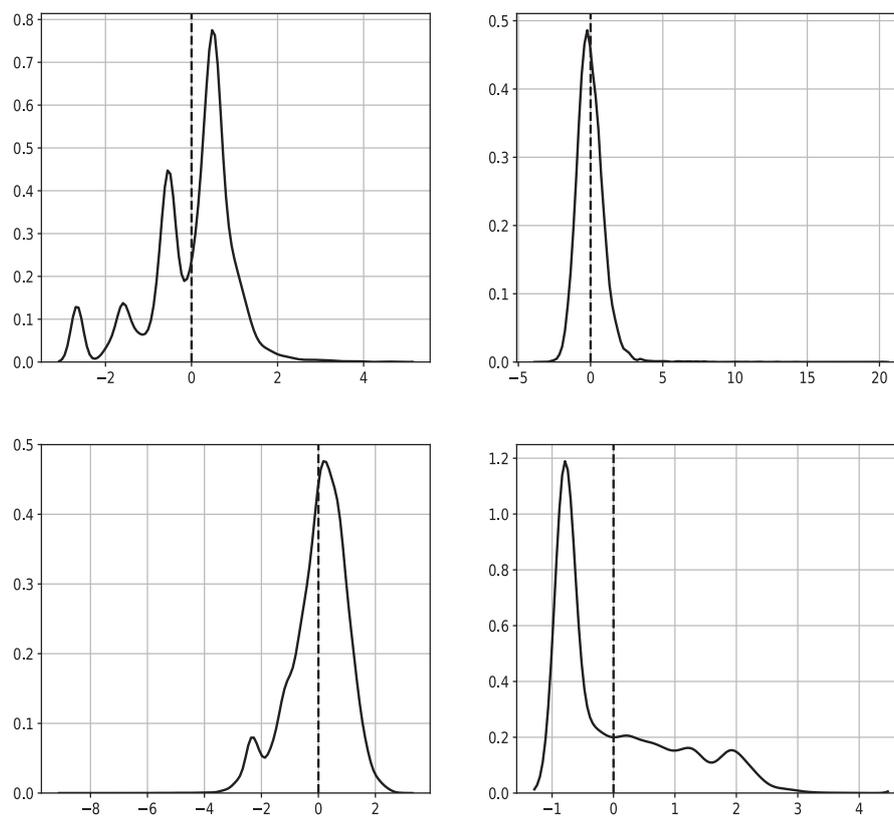


Figure 4. Examples of distribution of variables from real EAF production. All variables are normalized.
Top left: Propane. **Top right:** Refining time. **Bottom left:** Total weight. **Bottom right:** Material Type B.

3.4. Modeling

3.4.1. Artificial Neural Networks (ANN)

ANN, as mentioned earlier, is one type of non-linear statistical model framework. The idea is to connect the input variables to the output variable using a network of nodes that are fully connected [48]. The first layer is the input layer and the last layer is the output layer. The intermediate layers are called hidden layers. See Figure 5 for an illustration of a simple ANN.

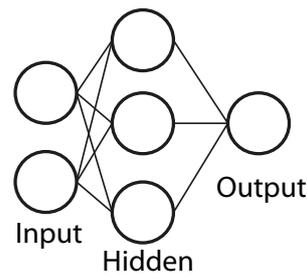


Figure 5. An Artificial Neural Network (ANN) with two input nodes, one hidden layer with three nodes, and one output node. Each line drawn between the nodes illustrates the forward flow of calculations in the network.

The number of hidden layers and the number of nodes in each hidden layer determine the complexity of the model. In each of the hidden layers, and in the output layer, each node multiplies a weight with each of the values from the nodes in the previous layer. The resulting values are summed together. This process can be mathematically expressed as:

$$s_j = \sum_{i=1}^P w_i \cdot x_i \quad (4)$$

where P is the number of nodes in the previous layer and j is the j th node in the layer.

A function, known as activation function, is then applied on s_j resulting in the value that the current node sends forward in the network. The hyperbolic tangent (tanh) and the logistic sigmoid functions are two commonly used activation functions.

During the training phase, the training data is sequentially fed into the network upon which the network weights are updated in the direction of minimizing the loss function. The loss function can be, for example, the mean squared error (MSE) or R-square (R^2). As the output value is a function of all the weights in the network, it is also possible to express the loss function as a function of the values of those weights. The updating of the weights is also known as backpropagation, because the errors are propagated back through the network and the weights updated. Since the loss space is a function of all weights in the network, finding a good enough local minima requires a well-engineered algorithm. These types of algorithms are known as gradient-descent algorithms because their aim is to descent to the most optimal local minima in loss space [48].

The sequentially feeding of training data, and update of the weights, is done numerous iterations until the accuracy improvement vanishes. One also uses a validation set, a subsample of the training data, during the training phase to ensure that the model does not overfit on the training data. Overfitting means that the model has learned the training data so well that it fails to predict well on unseen data. This is one of the drawbacks with neural network models because of its ability to learn complex relations even though the relations are not of value to solving the prediction problem. The validation set is the data set that the model is benchmarked against during the training phase to account for this drawback. After the training phase is completed, the test data is predicted by the model and model performance metrics are calculated.

3.4.2. Model Performance Metrics

To compare the performance of the models, R^2 and the regular error metric will be used.

One should use the adjusted- R^2 formula if one aims to compare R^2 between models with different number of input variables. This is because each added predictor slightly increases the R^2 -value given that the number of data points is fixed [49]. The formula for the adjusted- R^2 can be written as follows

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n - 1}{n - p - 1} \quad (5)$$

where R^2 is the standard R-square, n is the number of data points, and p is the number of input variables.

The regular error metric is advisable to use over the absolute error metric since an overestimated prediction is vastly different than an underestimated prediction in a practical steel plant context. The regular error metric for mean values is written as follows

$$E_{\mu} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (6)$$

where y_i is the true value, \hat{y}_i as the predicted value, and $i \in 1, 2, \dots, n$. The standard deviation, E_{σ} , is derived from E_{μ} the ordinary way.

3.4.3. Hyperparameter Optimization

Hyperparameter optimization aims to find the combination of parameters that creates the best model with respect to minimizing the model error [50]. A common method is to use a pre-specified grid of these parameters where the framework trains one, or more, models for each combination of the parameters. This specific method is also known as grid-search. For the experiments in this paper, parameters include model-specific parameters as well as application-/domain-specific parameters both of which together totals 36864 parameter combinations. See Table 7.

Table 7. The parameters used in the grid-search of which $2 \cdot 3 \cdot 48 = 288$ are model-specific and $64 \cdot 2 = 128$ are domain-specific. A total of $288 \cdot 128 = 36,864$ parameter combinations.

Model-Specific			
Parameter	Description	Values	#
Activation function	Activation functions have different gradient intensity in the updating step of each iteration and different upper and lower bounds. This influences the training phase.	[<i>tanh</i> , <i>logistic</i>]	2
Learning rate	A larger learning rate means that the gradient-descent algorithm takes a larger step in error space and may therefore miss more optimal local minimum points. Smaller steps, i.e., smaller learning rates, are preferable to find more optimal local minimum points.	[0.001, 0.01, 0.1]	3
Hidden node topology	Increasing number of layers and nodes means increased model complexity. Here, the goal is to investigate whether or not more hidden layers leads to a better model. One should always strive for model simplicity when two models are statistically significant equally good.	(z) or (z, z) where $z \in$ 1, 2, ..., 24	48
Domain-Specific			
Parameter	Description	Values	#
Input variables	It is not possible to know, a priori, which combination of input variables is the most relevant for a statistical model to predict the output variable with high precision and accuracy. See Section 2.4 for reasoning.	See Table 4.	64
Validation set	The effect of a randomized validation set and ordered validation set with respect to the rest of the training data will be investigated.	<i>ordered</i> , <i>randomized</i>	2

Parameters which were the same for all model combinations are explained below:

- Validation fraction: 0.2, which specifies the fraction of the training data used as validation set.
- Gradient-descent algorithm : An adaptive learning rate optimization algorithm known as Adam. This algorithm has been shown to outperform other gradient-descent algorithms in a variety of models and datasets. The algorithm-specific parameter values were selected as recommended in the paper [51].
- Early stopping set to True. Specified by the following items. **1.** Number of iterations with no change = 20. Specifies how many iterations of increasing or equal performance are required before

the training phase stops. 2. Tolerance = 10^{-7} . The tolerance of improvement to reset the number of iterations with no change.

Each combination of parameters represents one trained model type. Furthermore, each model type will be instanced 10 times to investigate the stability of the parameter selection and to reduce the impact of randomness. The metrics used to evaluate the model types based on the 10 instances are explained in Table 8.

Table 8. The metrics used to evaluate the performance of each model type.

Symbol	Definition
\bar{R}_μ^2	Mean adjusted R-square of the 10 model instances on the test data
\bar{R}_σ^2	Standard deviation of adjusted R-square of the 10 model instances on the test data
\bar{R}_{min}^2	Minimum adjusted R-square of the 10 model instances on the test data
\bar{R}_{max}^2	Maximum adjusted R-square of the 10 model instances on the test data
Δ_μ	Mean error of the mean error of the 10 model instances on the test data
Δ_σ	Standard deviation of the mean error of the 10 model instances on the test data.
Δ_{min}	Minimum error of the mean error of the 10 model instances on the test data.
Δ_{max}	Maximum error of the mean error of the 10 model instances on the test data.

To determine stability, the idea was to keep the \bar{R}_{min}^2 and \bar{R}_{max}^2 as close to the \bar{R}_μ^2 as possible. Hence, for filter out the best model type of each variable batch, the following algorithm was used.

1. Filter out any model that does not pass the following condition: $\bar{R}_{max}^2 - \bar{R}_{min}^2 \leq 0.05$. The motivation using this specific condition was to ensure that only stable models were selected while still enabling at least one model type from each variable batch to pass the condition.
2. Sort the models on decreasing \bar{R}_μ^2 .
3. Pick the first model in the list.

The reason adjusted R-square was used for determining stability is because the metric indicates goodness of fit for the model. Mean model error and standard deviation of model error were included to determine the best model out of models who have close to equal R-square.

3.4.4. Algorithmic Approach

In the algorithmic approach, process knowledge will not be the basis of choosing variable batches. By contrast, the variables in each batch will be chosen based on their one-to-one correlation with the EE consumption. In this manner, the approach will be mostly algorithmic except for the chosen setup of available variables, which are the same as in the domain-specific approach. The proposed algorithm is described below:

1. Calculate the pair-wise correlation values between the input variables and the output variable (EE consumption).
2. Use all input variables in the first model.
3. For each subsequent model, remove the input variable with lowest correlation value with respect to the output variable. Save the performance metric results from all models for further analysis.

In the experiments, dCor will be used as correlation metric due to its ability to detect non-linear correlative relations between variables. dCor is explained in Section 3.4.5.

The rest of the model-specific and domain-specific parameters will be the same as in the domain-specific approach. This results in a total of $35 \cdot 2 = 70$ domain-specific parameter combinations and a total of $2 \cdot 3 \cdot 48 = 288$ model-specific parameter combinations. Using the algorithmic approach, a total of $70 \cdot 288 = 20,160$ model types were trained. To filter out the best model for each variable batch, 10 instances of each parameter combination were created and the same algorithm as in the domain-specific approach was used.

3.4.5. Model Analysis

To ensure the reliability of a model, its transparency must be highlighted. To achieve this, three different statistical methods will be used. Two of these methods, KS test and dCor, are model independent and point to reasons behind model performance deviation from the training and test data. The third method, FI, is model dependent and explains how important each input variable from the model's viewpoint. By using FI in combination with dCor and KS tests, the goal is to home in on the variables that are the strongest reason behind model performance deviation between the training and test data.

dCor: Correlation metrics are used to investigate whether a variable is correlated with another variable. Correlated variables may or may not be causative. Nevertheless, two correlated variables may provide a hint to a possible causative relationship. Using domain-specific knowledge it is possible to assess the relative strength of the causative relation based on the correlation value, if there is evidence of such from physical deliberations.

The commonly used Pearson correlation metric only detects linear and monotonic relationships between two variables [52]. This severely limits the relevance of the correlation metric within the scope of this study since it is well known that parameters governing the EAF process are non-linear with respect to the EE consumption, (see Sections 2.2 and 2.3).

Alternative correlation metrics, such as dCor, can detect both linear and non-linear relationships between variables [53]. The resulting mathematical expression for dCor is similar to the Pearson correlation coefficient:

$$dCor(V_1, V_2) = \frac{dCov(V_1, V_2)}{\sqrt{dVar(V_1)dVar(V_2)}} \quad (7)$$

where $dCov(V_1, V_2)$ is the distance covariance and $dVar(V_1)$, $dVar(V_2)$ is the distance variance of the random variables V_1 and V_2 , respectively. The square root of the latter is the distance standard deviations of V_1 and V_2 .

To calculate dCor, the task is first to calculate the $n \times n$ distance matrices of each random variable, V_1 and V_2 :

$$a_{j,k} = \|V_{1,j} - V_{1,k}\| \quad (8)$$

$$b_{j,k} = \|V_{2,j} - V_{2,k}\| \quad (9)$$

where $j, k = 1, 2, \dots, n$ and $\|\cdot\|$ is the Euclidean norm. Using the distance matrices, the doubly centered distances are calculated as:

$$A_{j,k} = a_{j,k} - \bar{a}_j - \bar{a}_k + \bar{a} \quad (10)$$

$$B_{j,k} = b_{j,k} - \bar{b}_j - \bar{b}_k + \bar{b} \quad (11)$$

where \bar{a}_j is the row mean of the distance matrix, \bar{a}_k is the column mean of the distance matrix, and \bar{a} is the grand mean of the distance matrix for random variable V_1 . Analogous for random variable V_2 . The distance covariance is then calculated as the following arithmetic average:

$$dCov^2(V_1, V_2) = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n A_{j,k} B_{j,k} \quad (12)$$

Analogous, the distance variance for V_1 and V_2 are:

$$dVar^2(V_1) = dCov^2(V_1, V_1) = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n A_{j,k}^2 \quad (13)$$

$$dVar^2(V_2) = dCov^2(V_2, V_2) = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n B_{j,k}^2 \quad (14)$$

The dCor metric assumes values between 0 and 1, where 0 implies that the variables are independent and 1 implies that the variables are equal.

It has also been theoretically proven that dCor asymptotically detect deviations from independence. This implies that dCor is, at least in theory, a solid tool to detect any relationship between two variables given that enough data is provided. Some alternative correlation metrics with the similar characteristics as dCor are Hoeffding's D measure, HHG (Heller, Heller and Gorfine) measure and MI (Mutual Information) [52]. The use of any particular correlation metric depends in large on the number of observations and the type of relationship one intends to identify and the "nature" of the variables. This implicitly justifies that any correlation metric is not necessarily the "holy grail" of determining variable dependence [52]. It also means that a correlation metric may give low values if the pattern one intends to identify is challenging for the metric to detect. However, due to the limitations of the current study to the statistical modeling of the EE consumption in the EAF, dCor is chosen as the preferred metric due to its ability to detect both linear and non-linear relationships. Furthermore, for large number of observations, dCor has proven to be reliable for many different types of dependencies [52].

dCor will be calculated for each input variable to the output variable for both the training and the test data. A change in dCor between the training and test data indicates that the relation between the input variable and the output variable has changed. The relationship between the input variables and the output variable is what a statistical model learns using the training data. Thus, the change in dCor between the training and test data highlights the variable as a reason behind the model performance deviation.

In the analysis, input variable pairs with dCor values at or above 0.1 will be referred to as having relatively high correlation values. The reasoning behind this limit is two-fold. First, most of the dCor values between the input variables pairs were shown to be lower than 0.1 and most of the expected correlations between the input variable pairs was higher than 0.1. Second, to the authors knowledge, a clear guidance in the literature to what is considered a low or high dCor value does not exist.

KS tests: The KS-value is the maximum distance between two cumulative distribution functions (CDF). The CDF are either two samples of the same variable or one sample and one ideal CDF, such as the CDF for the normal distribution. Furthermore, the KS test is a non-parametric statistical test which means that it does not make any presumptions about the distributions governing the samples [54]. Variables from the EAF production are often from varying classes and superpositions of distributions, see Figure 4. The KS test is conducted by calculating the confidence, the p -value, of the KS-value under the null hypothesis, H_0 , that the two samples are from the same distribution. The KS-value takes values between 0 and 1 where 0 indicates the distributions are identical and 1 indicates that the distributions are completely different. The p -value is the probability that the two samples are in fact from the same distribution. Hence, a high KS-value with a low p -value strongly indicates that the two distributions underlying the two samples are different.

In the experiments, the two-sample KS test is of particular interest because it enables the study of the difference in two sampled distributions of the same variable. The two-sample KS test equations follows:

$$D_{n_1, n_2} = \sup_x |F_{n_1}(x) - G_{n_2}(x)| \quad (15)$$

where F_{n_1} and G_{n_2} are the two distribution functions with n_1 and n_2 are the number of samples from the two distributions, respectively. x is the total sample space. \sup is the supremum. The two-sample KS test is illustrated in Figure 6.

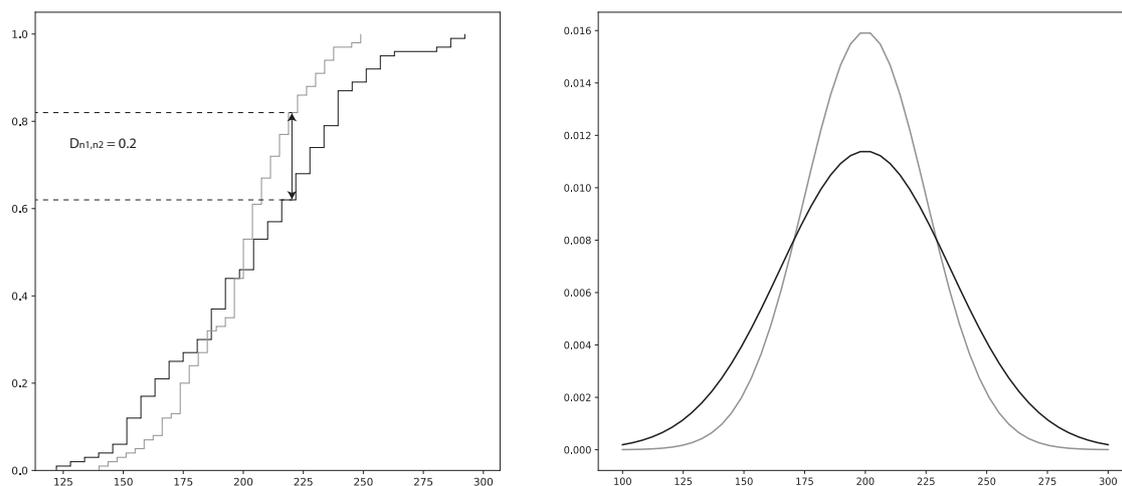


Figure 6. An illustration of the two-sample Kolmogorov–Smirnov (KS)-test for the random variables X_1 and X_2 , where $X_1 \sim Norm(200, 25)$ and $X_2 \sim Norm(200, 35)$ **Left:** The cumulative distribution functions (CDF) of X_1 and X_2 with the D -value as the difference between the upper and lower dashed lines. 100 samples were drawn from each distribution. **Right:** The probability density functions of X_1 and X_2 .

The null hypothesis, which is that the samples come from the same distribution, is rejected if:

$$D_{n_1, n_2} > c(\alpha) \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} \quad (16)$$

where α is the significance level and c is the threshold value calculated using α and the cumulative KS distribution [55].

One drawback of KS test is that it is sensitive to any difference between the two distributions. A large, but local, difference between the distributions will be captured by the KS test, but may not necessarily be representative over the complete distribution space.

The two-sample KS test will be used to investigate the distribution difference between the training data and the test data for both the input variables and the output variable. Since the neural network weights are adapted to the training data, any significant difference in distribution between the training and the test data for any variable (more affect on the most important variables) will affect the performance of the model on previously unseen data (test data). The aim is to pinpoint variables that are a probable cause to the performance difference rather than to prove specific performance changes by each variable. The latter approach is difficult, if at all doable, partly due to the complex interrelations between the variables in the data.

KS-values above or equal to 0.2 with p -values at or below 0.05 will be considered to be variables of interest in the analysis. The main reason is to avoid capturing all variables while at the same time capturing variables that change significantly enough between the training and test data. The implications of this limit on the KS test will be discussed. Distributions plots for variables above the KS-value threshold can further help to identify if the difference between the training and test data is significant enough.

FI: It is important to investigate how much each input variable affects the output variable. A useful way to do this is to use interpretable machine learning algorithms. One such algorithm is called FI which ranks each input feature, e.g., variable, in order of importance with respect to all predictions. Hence, FI is known as a global interpretable machine learning model. Each variable is singly broken and the recorded error from the model on this new data set is compared with the error from the original, unbroken, data set. If the error is higher, then the input variable is of some predictive importance with

respect to the output variable. If the error is statistically indifferent, then one can conclude that the input variable is of little importance to the model [56]. The algorithm is described as follows:

1. Train a model and record its error $L(X)$.
2. Permute one of the input variables, \bar{x}_j , in the input matrix X .
3. Apply the permuted input matrix, X_j , to the trained model and record the error, $L(X_j)$.
4. Repeat steps 2 and 3 for all input variables $j \in 1, 2, \dots, m$.
5. Order all variables in the order of decreasing $L(X_j)$.
6. Normalize all $L(X_j)$ on $\max(L(X_j))$ (optional).

where L is the model loss function, X is the complete and non-permuted input data matrix, and m is the total number of variables.

FI will be applied to both the training data and the test data. FI on training data tells us how much the model relies on each input variable for making predictions on new data. On the other hand, FI on test data reveals how important each input variable is to the actual model performance on unseen data [57]. Hence, both approaches can be viewed as complementary tools for model transparency. If the training and test data for a given variable come from the same distribution, then the FI value will be the same given that the number of data points is large enough. A larger deviation in FI indicates that the underlying relations between the input variables and the output variable have changed. FI will be calculated 20 times for each variable and data set to account for randomness in the permutations. The presented values will be the mean FI-values for each variable.

4. Results

4.1. Modeling

See Table 9 for the domain-specific cleaning results. Only 10% of the original data was removed after all cleaning steps.

Table 9. The remaining data from the domain-specific cleaning rules that were applied to the data. The training and test data were selected as described in Section 3.2.3.

Filter	Training Data	Test Data	Total Data
Before cleaning	12,183	404	12,587
Removal of all Trial heats	11,530	386	11,916
Heats with EE above 60 MWh	11,530	386	11,916
Heats with Total charged weight above 110,000 kg	11,530	386	11,916
45 min < TTT < 180 min	10,990	362	11,352
Delays < 180 min	10,966	362	11,328
After cleaning (No.)	10,966	362	11,328
(% data loss)	10.0%	10.4%	10.0%

In a practical application context, only the best model will be selected. However, some interesting observations were made when analyzing the results from both the domain approach and the algorithmic approach. Therefore, a total of 6 models were selected for further analysis. D1 was selected because it uses all input variables, D64 was selected because it uses only the base variables in the domain approach. D15, D17, and D31 were selected because they are the top 3 best performing models on the test data from the domain approach. A21 was selected because it is the best performing model from the algorithmic approach. Hence, we refer to D15, D17, D31, and A21, as the best models. The true EE consumption is plotted against the predicted EE consumption by models D15 and A21 in Figure 7. The performance of the selected models are presented in Table 10. The performance of the

best model from each variable batch is presented in Figures A1 and A2 for the domain approach and algorithmic approach, respectively.

Table 10. Performance of the selected models on the training and test data. Models called D are from the Domain approach and models called A are from the algorithmic approach. TR and TE are short for training and test, respectively.

Model	D1		D15		D17		D31		D64		A21	
No. Variables	35		20		34		19		6		16	
Dataset	TR	TE	TR	TE	TR	TE	TR	TE	TR	TE	TR	TE
\bar{R}_μ^2	0.790	0.698	0.717	0.731	0.766	0.706	0.660	0.722	0.601	0.697	0.725	0.706
Δ_μ	1	-331	-2	-554	-23	-266	2	-710	-8	-686	-5	-566
Δ_σ	1145	1167	1392	1126	1235	1176	1562	1105	1691	1216	1378	1185
Δ_{min}	-5679	-4249	-6606	-3819	-5275	-3662	-7526	-4089	-8149	-4805	-6519	-4333
Δ_{max}	5310	5442	5564	5735	5657	6072	6462	6020	8520	6087	6038	6610
$\bar{R}_{\mu,train}^2 - \bar{R}_{\mu,test}^2$	0.092		-0.014		0.060		-0.062		-0.096		0.019	
$\Delta_{\mu,train} - \Delta_{\mu,test}$	332		552		243		712		676		561	

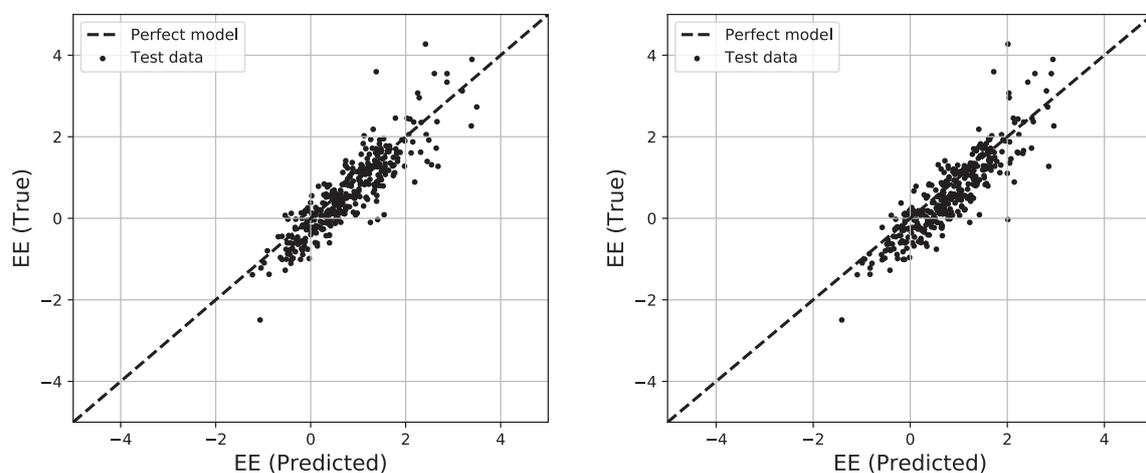


Figure 7. The predicted and true EE consumption. The dashed lines represent a perfect model where all predicted EE consumption values are equal the true EE consumption values. All values are normalized. **Left:** Model D15. **Right:** Model A21.

A significant negative mean error on the test data compared to the training data can be observed for all selected models. This is also consistent for a majority of the best models from the variable batches, as can be seen in Figures A1 and A2. The minimum and maximum errors are in the ranges of -3.7 MWh to -4.8 MWh and 5.4 MWh to 6.6 MWh on the test data, for the best models.

Comparing the performance from the best model to reported ANN models in the literature is shown in Table 11. The standard deviation of error is better than all reported models evaluated on test data, while the mean error is worse. The minimum and maximum error are comparable. However, the number of variables is significantly lower for the best model from the experiments done in this paper; 20 compared to 82 to 100 in the literature. Furthermore, the results from the models in the present paper are the average over 10 iterations. No such approach to combat model instability was done on the models in the literature. In addition, one has to keep in mind that the reported ANN models, evaluated on test data, in the literature lack R^2 for all models and % cleaned data for all but one model [43].

Table 11. The best model compared with ANN models created by Baumert et al. [21–23]. Only models that were tested on test data and presented the number of test heats are reported. Dashed entries indicate missing information.

Model	Baumert C1	Baumert D1	Baumert E1	Baumert E2	D15
No. Hidden layers	1	10	10	11	1
Total No. Hidden nodes	7	50	50	58	11
No. Variables	100	84	82	95	20
Train/test split	375/20	668/150	707/62	921/1011	10,966/362
% cleaned data	-	2.6%	-	-	10%
R^2	-	-	-	-	0.731
Mean error (kWh/heat)	~0	−100	~0	−300	−554
Standard deviation of error (kWh/heat)	1300	3300	1800	1900	1126
Minimum error (kWh/heat)	−3500	-	-	−8000	−3819
Maximum error (kWh/heat)	6000	-	-	5000	5735

The best model from the algorithmic approach is almost as good as the best model from the domain approach with a difference in \bar{R}_μ^2 of only 0.025.

The model with the least amount of input variables, D64, has a \bar{R}_μ^2 that is only 0.001 higher than the \bar{R}_μ^2 for the model with the highest number of input variables, D1.

The performance on the test data is better than on the training data for D15, D31, D64, and D21. Significantly so for D31 and D64 where a difference in \bar{R}_μ^2 0.062 and 0.096 are observed. This is somewhat unexpected since a model usually performs worse on test data.

4.2. Model Analysis

Assuming a threshold p -value of 0.05, all variables except Delays, Type B, Type C, and Type F, are of particular interest for further analysis of the KS-values, see Table 12. MgO shows an unusually large KS-value of 0.85. Al, Metal Weight, Slag Weight, and Cr_2O_3 have KS-values of 0.34, 0.41, 0.5, and 0.36, respectively. Then there is a plethora of variables that have KS-values between 0.20 and 0.30. From the Base variables these are, Total Weight, Propane, O_2 -lance, Preheater. From the time variables, Refining is the only variable. From the oxide composition, these variables are CaO, FeO, Al_2O_3 . For the raw material types, Type N satisfies the criteria. EE consumption has a KS-value of 0.25. The distributions for the training and test data for EE consumption are shown in Figure 8.

The dCor between the input variables and the EE consumption are shown in Table 12. Changes in dCor above 0.1 between the training and test data are observed for Delays, TTT, Total Weight, Propane, Process Time, Extended refining, C, Si, O, Metal Weight, Type C, Type E, and Type N, raw materials.

Expected correlations between variables, as discussed in Section 2.4, can be observed in Table A4. dCor calculations between time variables and time variables have higher dCor values compared to dCor calculations between time variables against other variable types. The intra-correlation for metal composition, oxide compositions, and raw material types, are expected since the sum of the compositions are equal to one and the sum of the raw material types are bound by the total weight. The inter-correlation between metal compositions, oxide compositions, and the raw material types, indicates the complex relationship between these variables. Furthermore, the differences in dCor between the training and test data for all variables are explicitly shown in Table A3. While some larger changes in dCor can be observed for the time variables, the variables with the most number of larger dCor changes are the weight variables, metal composition, oxide composition, and raw material type variables.

Table 12. Two-sample KS tests for the training data against the test data for all input variables and the EE consumption. Distance correlation (dCor) for training and test data sets are calculated as a one-to-one correlation between the input variables and EE consumption. The difference in dCor is defined as $dCor(training) - dCor(test)$.

Variable	KS-Value	p-Value	dCor EE Train	dCor EE Test	dCor EE Diff
Delays	0.04	0.49	0.135	0.270	−0.135
TTT	0.13	0.16	0.262	0.396	−0.134
Total Weight	0.24	0.0	0.325	0.116	0.209
Propane	0.22	0.0	0.082	0.217	−0.135
O ₂ -lance	0.24	0.0	0.123	0.114	0.009
Preheater energy	0.23	0.0	0.060	0.074	−0.014
Process Time	0.18	0.0	0.399	0.534	−0.135
Charging	0.08	0.01	0.025	0.095	−0.07
Melting	0.08	0.02	0.107	0.087	0.02
Refining	0.24	0.0	0.448	0.499	−0.051
Extended refining	0.08	0.03	0.237	0.463	−0.226
Tapping	0.18	0.0	0.152	0.129	0.023
C	0.07	0.04	0.226	0.096	0.129
Si	0.16	0.0	0.105	0.216	−0.111
Cr	0.19	0.0	0.131	0.139	−0.008
Fe	0.13	0.0	0.095	0.181	−0.086
Ni	0.16	0.0	0.097	0.187	−0.09
O	0.11	0.0004	0.032	0.257	−0.225
Al	0.34	0.0	0.098	0.140	−0.042
Cr ₂ O ₃	0.36	0.0	0.095	0.099	−0.004
MgO	0.85	0.0	0.193	0.241	−0.048
CaO	0.24	0.0	0.057	0.131	−0.074
FeO	0.20	0.0	0.032	0.123	−0.091
SiO ₂	0.13	0.0	0.079	0.111	−0.032
Al ₂ O ₃	0.26	0.0	0.130	0.210	−0.08
Metal Weight	0.41	0.0	0.344	0.135	0.208
Slag Weight	0.5	0.0	0.108	0.179	−0.071
Type A	0.18	0.0	0.064	0.094	−0.03
Type B	0.07	0.06	0.111	0.180	−0.069
Type C	0.07	0.08	0.150	0.291	−0.141
Type D	0.14	0.0	0.187	0.271	−0.084
Type E	0.18	0.0	0.128	0.300	−0.172
Type F	0.02	0.98	0.041	0.046	−0.005
Type G	0.19	0.0	0.151	0.155	−0.004
Type N	0.23	0.0	0.098	0.303	−0.205
EE consumption	0.25	0.0	-	-	-

FI for the 6 selected models are presented in Table 13. Delays, TTT, and Total Weight are the three most important variables for all models, except for model D15, with respect to the training data. Some significant changes in FI from the training and test data sets are also observed. These are marked with bold and underlined numbers in Table 13.

Table 13. Feature importance (FI) for the selected models. Models called D are from the Domain approach and models called A are from the algorithmic approach. TR and TE are short for training and test, respectively. Significant changes in FI are indicated in underlined bold font.

	D1		D15		D17		D31		D64		A21	
	TR	TE										
Delays	0.85	0.75	1.37	1.30	0.65	0.62	1.21	1.13	<u>1.90</u>	<u>1.71</u>	<u>0.71</u>	<u>0.58</u>
TTT	<u>1.15</u>	<u>1.34</u>	<u>1.35</u>	<u>1.58</u>	<u>0.94</u>	<u>1.14</u>	<u>1.71</u>	<u>1.96</u>	<u>2.58</u>	<u>2.91</u>	0.82	0.88
TotWeight	<u>1.15</u>	<u>0.70</u>	<u>0.16</u>	<u>0.02</u>	<u>1.16</u>	<u>0.63</u>	<u>0.15</u>	<u>0.04</u>	<u>0.13</u>	<u>0.03</u>	0.43	0.40
Propane	0.07	0.01	0.05	0.00	0.03	0.00	0.02	0.0	0.02	0.0	-	-
O ₂ -lance	0.03	0.0	0.02	0.0	0.02	0.0	0.0	0.0	0.02	0.0	-	-
Preheater	0.03	0.02	0.03	0.01	0.02	0.00	0.02	0.01	0.02	0.01	-	-
Process Time	0.08	0.05	<u>0.23</u>	<u>0.49</u>	-	-	-	-	-	-	<u>0.20</u>	<u>0.30</u>
Charging	0.01	0.0	0.01	0.01	0.01	0.0	0.01	0.0	-	-	-	-
Melting	0.0	0.01	0.0	0.0	0.0	0.0	0.0	0.0	-	-	-	-
Refining	0.09	0.06	0.05	0.04	0.09	0.02	0.05	0.05	-	-	0.03	0.03
Ext. ref.	0.03	0.06	0.03	0.07	<u>0.03</u>	<u>0.15</u>	<u>0.03</u>	<u>0.12</u>	-	-	0.03	0.06
Tapping	0.03	0.0	0.05	0.04	0.00	0.00	0.0	0.0	-	-	0.07	0.06
C	0.01	0.0	-	-	0.01	0.00	-	-	-	-	0.01	0.0
Si	0.09	0.14	-	-	0.08	0.09	-	-	-	-	-	-
Cr	0.02	0.0	-	-	0.01	0.0	-	-	-	-	0.01	0.01
Fe	<u>0.13</u>	<u>0.27</u>	-	-	0.02	-0.02	-	-	-	-	-	-
Ni	0.07	0.05	-	-	0.01	0.04	-	-	-	-	-	-
O	0.08	0.11	-	-	0.07	0.07	-	-	-	-	-	-
Al	0.04	0.03	-	-	0.03	0.01	-	-	-	-	-	-
Cr ₂ O ₃	0.03	-0.01	-	-	0.01	0.00	-	-	-	-	-	-
MgO	<u>0.15</u>	<u>0.0</u>	-	-	<u>0.13</u>	<u>0.01</u>	-	-	-	-	0.07	0.0
CaO	<u>0.44</u>	<u>0.82</u>	-	-	<u>0.32</u>	<u>0.61</u>	-	-	-	-	-	-
FeO	<u>0.05</u>	<u>0.14</u>	-	-	<u>0.03</u>	<u>0.10</u>	-	-	-	-	-	-
SiO ₂	<u>0.24</u>	<u>0.46</u>	-	-	<u>0.19</u>	<u>0.38</u>	-	-	-	-	-	-
Al ₂ O ₃	0.06	0.05	-	-	0.07	0.03	-	-	-	-	0.01	-0.01
Metal	<u>0.44</u>	<u>0.29</u>	-	-	<u>0.49</u>	<u>0.25</u>	-	-	-	-	0.15	0.16
Slag	0.06	0.01	-	-	0.04	0.00	-	-	-	-	-	-
Type A	0.03	0.0	0.02	0.00	0.01	0.01	0.01	0.01	-	-	-	-
Type B	0.02	0.0	0.04	0.02	0.01	0.01	0.03	0.02	-	-	-	-
Type C	0.04	0.02	0.07	0.01	0.06	0.0	0.01	0.01	-	-	0.05	0.04
Type D	0.02	0.04	0.05	0.08	0.02	0.01	0.03	0.07	-	-	0.01	0.02
Type E	<u>0.05</u>	<u>0.14</u>	0.02	0.0	0.02	0.04	0.0	0.0	-	-	0.01	0.04
Type F	0.0	0.0	0.01	0.0	0.0	0.0	0.0	0.0	-	-	-	-
Type G	0.10	0.13	0.05	0.04	0.09	0.08	0.04	0.02	-	-	0.08	0.12
Type N	0.02	0.0	0.03	0.01	0.02	0.0	0.0	0.0	-	-	-	-

Training and test data distributions for some of the most important variables for the 6 selected models are shown in Figure 8. These are Delays, TTT, Process Time, Total Weight, and Metal Weight. EE is also plotted because of its KS-value of 0.25. Metal compositions, oxide compositions, and raw material types were not included because of their complex relationships shown by the relatively high inter-correlation and intra-correlation, (see Tables A4 and A5). Analyzing their combined impact on the EE consumption is too extensive for this article.

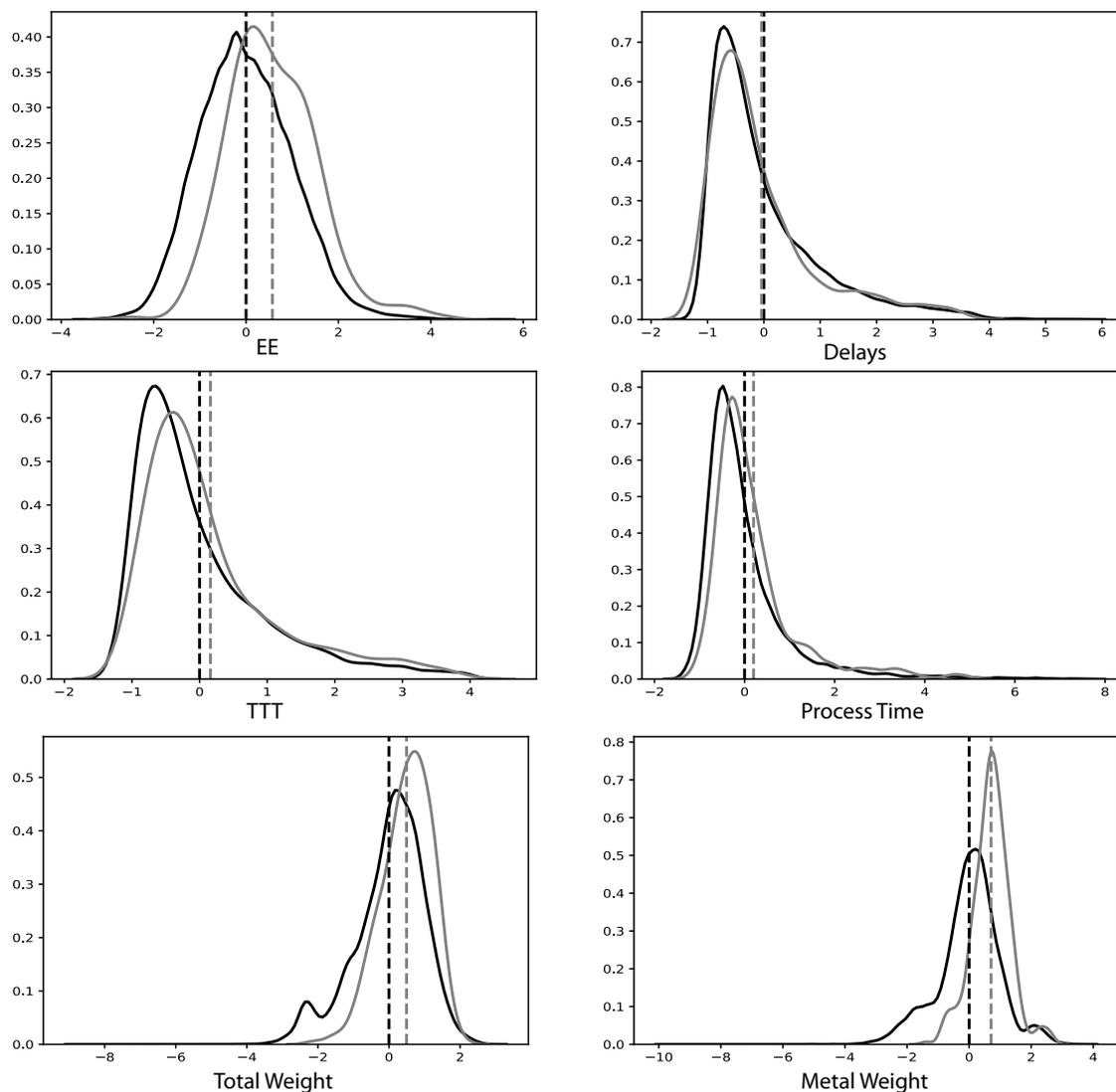


Figure 8. Distributions for the training data (black) and test data (gray) for EE and the 5 of the variables with higher FI for the 6 selected models. The vertical dashed lines show the mean of each distribution. Both the training and test data are normalized with the mean error and standard deviation of error of the training data.

The EE consumption have increased in the test data set. So have the TTT, Process Time, Total Weight, and Metal Weight. The delay is about equal on average between the training and test data sets.

4.3. Grid-Search Metadata

The parameters from the models passing the $\bar{R}_{max}^2 - \bar{R}_{min}^2 \leq 0.05$ criterion (see Section 3.4.3) and the parameters from the best models are summarized in Tables 14 and 15, respectively. Validating on chronologically ordered data during training is not always beneficial as the ordered is only present in 27% of the best models. Number of hidden layers, which is a proxy for model complexity, is equal to one for 87% of the best models. A similar reasoning can be made for the algorithmic approach, which metadata is presented in the same tables.

Table 14. Parameter metadata for the models passing the $\bar{R}_{max}^2 - \bar{R}_{min}^2 \leq 0.05$ criterion (see Section 3.4.3). The upper row is for the domain-specific approach and the bottom row is for the algorithmic approach.

Learning rate	0.1 2%	0.01 19%	0.001 79%	Activation function	L 68%	T 32%	Ordered	Yes 39%	No 61%	Hidden layers	1 60%	2 40%
Learning rate	0.1 9%	0.01 32%	0.001 59%	Activation function	L 65%	T 35%	Ordered	Yes 48%	No 52%	Hidden layers	1 55%	2 45%

Table 15. Parameter metadata for the best model from each variable batch. The upper row is for the domain-specific approach and the bottom row is for the algorithmic approach.

Learning rate	0.1 0%	0.01 6%	0.001 94%	Activation function	L 69%	T 31%	Ordered	Yes 19%	No 81%	Hidden layers	1 88%	2 12%
Learning rate	0.1 5%	0.01 40%	0.001 54%	Activation function	L 63%	T 37%	Ordered	Yes 23%	No 77%	Hidden layers	1 69%	2 31%

5. Discussion

5.1. Modeling

The best model from the domain approach has a \bar{R}_{μ}^2 of 0.706 while the best model from the algorithmic approach has a \bar{R}_{μ}^2 of 0.731. These values were calculated from the average of 10 model instances, and the difference is not large. This means that the algorithmic approach can be used as a tool to select variables that governs a model that will have a top-tier accuracy. However, it is important to consider the domain-specific reasoning behind the initial selection of the 35 available input variables. It is, therefore, not suggested that one should blindly select variables stored in an arbitrary EAF-database and then perform the algorithmic variable selection explained in this paper. While this could work, it is not what is intended using the algorithmic approach.

Using a large scale grid-search with over 57,000 model types, using both a domain-specific approach and an algorithmic approach, the best model type achieved an R_{μ}^2 of 0.731, a mean error of -554 kWh/heat, a standard deviation of error of 1126 kWh, a minimum error of -3819 kWh/heat, and a maximum error of 5735 kWh/heat. The best model in the literature had a mean error of approximately zero, a standard deviation of error of 1300 kWh/heat, a minimum error of -3500 kWh/heat, and a maximum error of 6000 kWh/heat. The errors from both models are similar. However, the literature model did not report averages of 10 model iterations. Therefore, it is not possible to know if the performance is stable or purely based on luck. The models in this paper were selected after a filtering criterion, $\bar{R}_{max}^2 - \bar{R}_{min}^2 \leq 0.05$ (see Section 3.4.3). After this filtering, the best models based on the highest \bar{R}_{μ}^2 on the test data were selected. In addition, the model from the literature did not report R_{μ}^2 and was only evaluated on 20 test data points compared to 362 test data points for the best model produced in this paper. Assuming a production of 20 heats per day, this represents a model evaluation difference of approximately 17 production days. Furthermore, neither the data treatment specifics nor the percentage of cleaned data were specified for that model. Model reporting shortfalls, of which some are mentioned above, are frequently occurring in the available literature on the subject of statistical modeling to predict the EE consumption in the EAF. It is important to clearly describe all steps in the data- and modeling pipeline that impose changes to the end result. In the context of this paper, the end result is the best model. The best model presented in this paper is also less complex than the best reported model in the literature even though both models perform similarly. While the number of hidden layers are equal between both models, the model reported in the literature has 100 input variables compared to 20 input variables for the best model produced in this paper. In the interest of model parsimony, simplicity should be followed when creating any model. On the same note, the model using 6 of the base variables has an R_{μ}^2 that is not significantly better (a difference of only 0.001) than the model using all 36 input variables, and the best model uses only 20 variables. Hence,

adding a lot of input variables does not necessarily create a model with better performance. One has to search for the optimal combination of variables using either a purely domain-specific approach or an algorithmic approach, such as the one presented in this paper. Tables 14 and 15 further support model simplicity since models with 2 layers are not preferred over models with 1 layer with respect to model stability (the $\bar{R}_{max}^2 - \bar{R}_{min}^2 \leq 0.05$ criterion, see Section 3.4.3) or with respect to the best model from each variable batch.

3 of the 6 selected models perform better on the test data than on the training data. This is somewhat contradictory since models usually perform worse on test data. However, the selection of the best models from each variable batch was based on the performance on the test data. The performance on the training data was not considered since, in a practical environment, only the performance on test data is important.

5.2. Model Analysis

Investigating the dCor-matrix for the training data, Table A4, one can observe many relatively high 1-to-1 correlation among the metal composition, oxide composition, and raw material type variables. This is in line with what is expected from a process-metallurgical perspective. Because of this trait, these variables share information between one another to a large extent. For example, Type D is relatively highly correlated with all metal composition and oxide composition variables. Simplified, Type D has the potential to account for the metal and oxide composition which renders those variables less important to the models performance. This is the reason adding a lot of input variables does not necessarily create a model with better performance, and the reason variables that are important in a physical model may be next to worthless in a statistical model. Changes in dCor from the training data to the test data, see Table A3, indicate that the underlying relation between the variables have changed. This is not as frequently occurring for the time variables as for the metal composition, oxide composition, and raw material types where changes to both lower and higher dCor are present. Due to changes in dCor, changes in the model performance are also expected. However, analyzing a 36×36 correlation matrix in depth is both impractical and does not show which variables are the most important for the performance of the model.

Table 12 shows the results from the KS tests performed on each variable on training and test data sets. KS-values greater than or equal to 0.20 were considered of significant change given that the p -value is 0.05, or less. In conjunction with FI for the selected models, see Table 13, it is possible to identify reasons behind the performance differences from training to test. Delays, TTT, Process Time, Total Weight, and Metal Weight are some of the most important variables for the 6 selected models. Out of these variables, Total Weight and Metal Weight are variables with KS-values above 0.20. Process Time, TTT, and Delays, had KS-values of 0.18, 0.13, and 0.04, respectively. Furthermore, the KS-value for EE consumption is 0.25, which indicates that the EE consumption has also changed from the training to test data. However, KS-values do only provide a difference in the CDF and do not indicate whether the variable has increased or decreased. Observing the variable distributions for the training and test data for these variables in Figure 8, it is clear that all of these variables, except Delays, have increased from the training data to test data. Since the EE consumption increases with the amount of charged materials and longer process times, and because these input variables are important for the models, it is highly probable that these variables give rise to the decrease in Δ_{μ} and the change in \bar{R}_{μ}^2 for the 6 selected models.

Please note that the limitation of demarcating “interesting” variables to only those at or above 0.2 KS-value becomes clear when observing the clear and consistent increase in Process Time and TTT in Figure 8. Nevertheless, using KS test in conjunction with FI and the additional distribution plots enabled the discovery of the implications on the model performance due to the changes in these variables.

Δ_{μ} is defined as the average error for all predictions across all 10 model instances. Each error is in turn defined as the true value minus the predicted value. Since all 6 selected models get a significant

negative Δ_{μ} on the test data, the models overestimate the EE consumption due to the changes in the variables with the highest FI. In particular, the Total Weight, Metal Weight, TTT and Process Time. This behavior is not unexpected since any statistical model will predict the value of the highest probability with respect to the data it has used to adapt its parameters. In this case, the 10,966 training data points have adapted the ANN weights. Any change to the distribution of the variables with high FI will affect the models prediction on data sampled from this aggregated distribution.

6. Conclusions

One of the aims with the present work was to address the previously reported shortcomings in the statistical modeling approach to predict the EE consumption of the EAF [43]. In addition to addressing these shortcomings with a detailed data- and modeling pipeline, the choice of statistical modeling framework and statistical analysis tools was adapted to the non-linearity of the EAF process and the complex correlative relations between the selected process variables. The main conclusions of the modeling approach, and suggested future work, may be summarized as follows:

6.1. Modeling

- In the interest of model parsimony, increasing the number of input variables does not automatically increase model performance. This is both observed when comparing the best model from the experiments to the best model reported in the literature and when comparing model D64, which uses 6 variables, and model D1, which uses 35 variables. It is sufficient to choose several variables from an initial selection demarcated by the use of process knowledge.
- Similar performance was retrieved by the best models from the algorithmic approach and the domain approach, respectively. Given a setup of variables that are selected with process expertise, an algorithmic approach can find an optimal selection of variables. This is important from a practical applicability standpoint since less time need to be invested in searching for the optimal number of variables and the optimal subset of variables.
- Neural networks with many layers, increasing the model complexity, are not necessary to produce state-of-the-art model performance. This was observed both in the grid-search performed in the experiments and in comparing the best model from the experiments with the best models reported in the literature.
- The mean error is slightly higher, and the standard deviation of error is better for the best model compared to the best model reported in the literature. The minimum and maximum errors are approximately the same as the best model reported in the literature. However, these are -3.8 MWh to 5.7 MWh, respectively, which are quite high from a practical application context.

6.2. Model Analysis

- Using KS test and FI as complementary tools, it was possible to identify an increase in Total Weight, Metal Weight, and Process Times as a highly probable cause behind the performance change from the training to test data for the 6 selected models.
- An analysis of the Δ_{μ} error from the selected models on the test data indicates that the models overestimate the EE consumption with regards to the change in the most important variables to the models. In particular, the Total Weight, Metal Weight, TTT, and Process Time variables.
- High intra-correlation between the metal composition, oxide composition, and raw material type variables was found using dCor on the training data. This explains why some of these variables get a high FI in some of the models. The changes in dCor from the training to test data were also prominent for these variables and can explain part of the performance changes for the selected models attributing higher FI to some of these metal composition, oxide composition, and raw material type variables.

6.3. Future Work

Based on the present work carried out and the conclusions drawn, it is suggested to further investigate the following:

- Use upsampling on the training data to get more data points as outliers. Could potentially solve the problem with large maximum and minimum errors.
- Develop a more advanced variable selection algorithm based on dCor that takes into account one-to-many correlations for each input variable. The current algorithm only considers the correlation between the EE and each input variable.
- Apply dCor to groups of variables and the EE consumption instead of variables in a 1-to-1 fashion. This feature of dCor should be explored further.
- Use other non-linear correlation metrics to determine appropriate variable selection. These include Hoeffding's, MI, and HHG.
- Investigate the effects of different types of scrap on the EE consumption using new classifications based on density, and heat and melting behavior.
- The effect on models by normalized variables that are not normally distributed is not known. See Figure 4 for a few examples.

Author Contributions: Conceptualization, L.S.C., P.B.S. and P.G.J.; Methodology, L.S.C. and P.B.S.; Validation, L.S.C., P.B.S. and P.G.J.; Formal analysis, L.S.C.; Investigation, L.S.C. and P.B.S.; Resources, P.B.S. and P.G.J.; Data curation, L.S.C.; Writing—original draft preparation, L.S.C.; Writing—review and editing, L.S.C., P.B.S. and P.G.J.; Visualization, L.S.C. and P.B.S.; Software, L.S.C.; Supervision, P.B.S. and P.G.J.; Project administration, P.G.J.; Funding acquisition, P.G.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by “Hugo Carlssons Stiftelse för vetenskaplig forskning” in the form of a scholarship granted to the corresponding author.

Acknowledgments: We want to thank the plant engineers Pär Ljungqvist, Christoffer Schmidt, and Jesper Janis at the Outokumpu Stainless Avesta mill for their support and data provisioning during this project.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

$E_{tot,in}$	Total ingoing energy
E_{El}	Total Electrical Energy (EE) output from transformer
E_{Chem}	Total energy from chemical reactions in steel and slag
E_{Bu}	Total energy input from burner
$E_{tot,out}$	Total outgoing energy
E_{Steel}	Total energy output into steel
E_{Slag}	Total energy lost in slag
E_{Gas}	Total energy lost in gas
E_{Dust}	Total energy lost in dust
$E_{Cooling}$	Total energy lost in cooling water
E_{Rad}	Total energy lost through radiation
E_{Conv}	Total energy lost through convection
$E_{El,loss}$	Total energy lost in electrical system and arc transfer
T_{CS}	The temperature of the cooling panels
T_{EAF}	The temperature of the surface area subject to radiation losses
T_s	Temperature of ingoing material and gas at the start of the EAF process
T_{Tap}	Temperature of the steel at tapping
T_{Offgas}	Temperature of the off-gas leaving the EAF through the off-gas system
T_{H2O}	Temperature of the cooling water
T_H	The temperature of the surface area subject to convection losses
T_{Amb}	The temperature of the air surrounding the EAF
m_{Steel}	Mass of ingoing metallic material
m_{Slag}	Mass of ingoing oxidic material

\dot{m}_{Dust}	Mass flow of dust in the off-gas system
c_{Steel}	The heat capacity of steel at constant pressure
c_{Slag}	The heat capacity of slag at constant pressure
c_{Dust}	The heat capacity of dust at constant pressure
c_{Gas}	The heat capacity of EAF ambient gas at constant pressure
$c_p(reactants)$	The heat capacity of reactants at constant pressure
$c_p(products)$	The heat capacity of products at constant pressure
$\Delta H_{298}^0(reactants)$	Standard heat of formation for reactants at 298K
$\Delta H_{298}^0(products)$	Standard heat of formation for products at 298K
$\Delta H_{Melt,steel}$	Heat of fusion for steel
$\Delta H_{Melt,slag}$	Heat of fusion for slag
k	Conductivity of the cooling panels
h	Heat transfer coefficient of the EAF ambient gas
ϵ	Emissivity factor of the radiating surface area of the EAF
σ	Stefan-Boltzmann constant
η_{El}	Efficiency factor for the transformer system
η_{Arc}	Efficiency factor for the energy transferred from the arcs
η_{Fuel}	Efficiency factor for burning the fuel in the burners
h_{Fuel}	Heat generated per volume unit of fuel
V_{Fuel}	Volume of the fuel consumed by the burners
A_H	Surface area of the EAF subject to convection losses
A_{AC}	Surface area of the cooling panels
A_{EAF}	Surface area of the EAF subject to radiation losses
P_{Avg}	Average power of the transformer system
P	Furnace pressure
M	Molar mass of the furnace gas
R	Universal gas constant
\dot{V}_{Gas}	Volume flow of gas in the off-gas system
t_{TTT}	Tap-to-tap time
t_{PON}	Power-on time
R^2	Coefficient of determination
\bar{R}^2	Coefficient of determination adjusted for number of data points and variables
n	Number of data points
p	Number of input variables
y_i	True value of the output variable for data point i
\hat{y}_i	Predicted value of the output variable for data point i
P	Number of nodes in the previous layer.
s_j	Summation of the input values for j th node in the current layer.
w_i	Weight of node i in the previous layer.
x_i	Value of node i in the previous layer.
Δ_μ	Mean error of the mean error of the 10 model instances on the test data
Δ_σ	Standard deviation of the mean error of the 10 model instances on the test data
Δ_{min}	Minimum of the mean error of the 10 model instances on the test data
Δ_{max}	Maximum of the mean error of the 10 model instances on the test data
\bar{R}_μ^2	Mean adjusted R-square of the 10 model instances
\bar{R}_σ^2	Standard deviation of adjusted R-square of the 10 model instances
\bar{R}_{min}^2	Minimum of adjusted R-square of the 10 model instances
\bar{R}_{max}^2	Maximum of adjusted R-square of the 10 model instances
V_1	Random variable
V_2	Random variable
$dCor(V_1, V_2)$	Distance correlation between V_1 and V_2
$dCov(V_1, V_2)$	Distance covariance between V_1 and V_2
$dVar(V_1)$	Distance standard deviation for V_1
$dVar(V_2)$	Distance standard deviation for V_2

$dVar^2(V_1)$	Distance variance for V_1
$dVar^2(V_2)$	Distance variance for V_2
$a_{j,k}$	Distance between values j and k for random variable V_1
$b_{j,k}$	Distance between values j and k for random variable V_2
$V_{1,j}$	Value j for random variable V_1
$V_{2,j}$	Value j for random variable V_2
$A_{j,k}$	Doubly centered distance for value j and k for random variable V_1
$B_{j,k}$	Doubly centered distance for value j and k for random variable V_2
$a_{j,k}$	
\bar{a}_j	Row mean of the distance matrix for the random variable V_1
\bar{a}_k	Column mean of the distance matrix for the random variable V_1
\bar{a}	Grand mean of the distance matrix for random variable V_1
$b_{j,k}$	
\bar{b}_j	Row mean of the distance matrix for the random variable V_2
\bar{b}_k	Column mean of the distance matrix for the random variable V_2
\bar{b}	Grand mean of the distance matrix for random variable V_2
n_1	Number of samples from the first distribution in the KS test
n_2	Number of samples from the second distribution in the KS test
D_{n_1,n_2}	KS-value from the KS test between F_{n_1} and G_{n_2}
x	The total sample space in the KS test
F_{n_1}	First distribution function in the two-sample KS test
G_{n_2}	Second distribution function in the two-sample KS test
H_0	Null hypothesis
α	Significance level
c	Threshold value calculated from the cumulative KS distribution
L	Model error function
X	Input matrix
\bar{x}_j	Input variable j
X_j	Input matrix with permuted variable j
E_μ	Mean error
E_σ	Standard deviation of error

Abbreviations

The following abbreviations are used in this manuscript:

CFD	Computational Fluid Dynamics
EE	Electrical Energy
EAF	Electric Arc Furnace
AC	Alternate Current
MLR	Multivariate Linear Regression
ANN	Artificial Neural Network
MSE	Mean Squared Error
FI	Feature Importance
KS	Kolmogorov–Smirnov
CDF	Cumulative distribution function
dCor	Distance correlation
MI	Mutual Information
HHG	Heller, Heller and Gorfine
TTT	Tap-to-Tap Time
TR	Training
TE	Test
NA	Not Available

Appendix A

Appendix A.1. Hardware and Software

Table A1. Hardware specifications.

Computer model	Dell Latitude E5570
CPU	Intel Core i7 2376 MHz
RAM	16,203 MB

Table A2. Software specifications for the experiments.

Purpose	Software/Package	Version
Operating system	Microsoft Windows 7 Professional	6.1.7601 Service Pack 1 Build 7601
Programming language	Python 3	3.7.1
Python distribution	Anaconda 3	4.6.7
Data handling	Pandas	0.23.4
	Numpy	1.17.4
Statistical modeling	Scikit-learn	0.20.1
Feature importance	eli5	0.8.1
Distance correlation	dcor	0.3
KS test	scipy	0.15.0
Visualization	Matplotlib	3.0.2

Appendix A.2. Variable Batch Performance

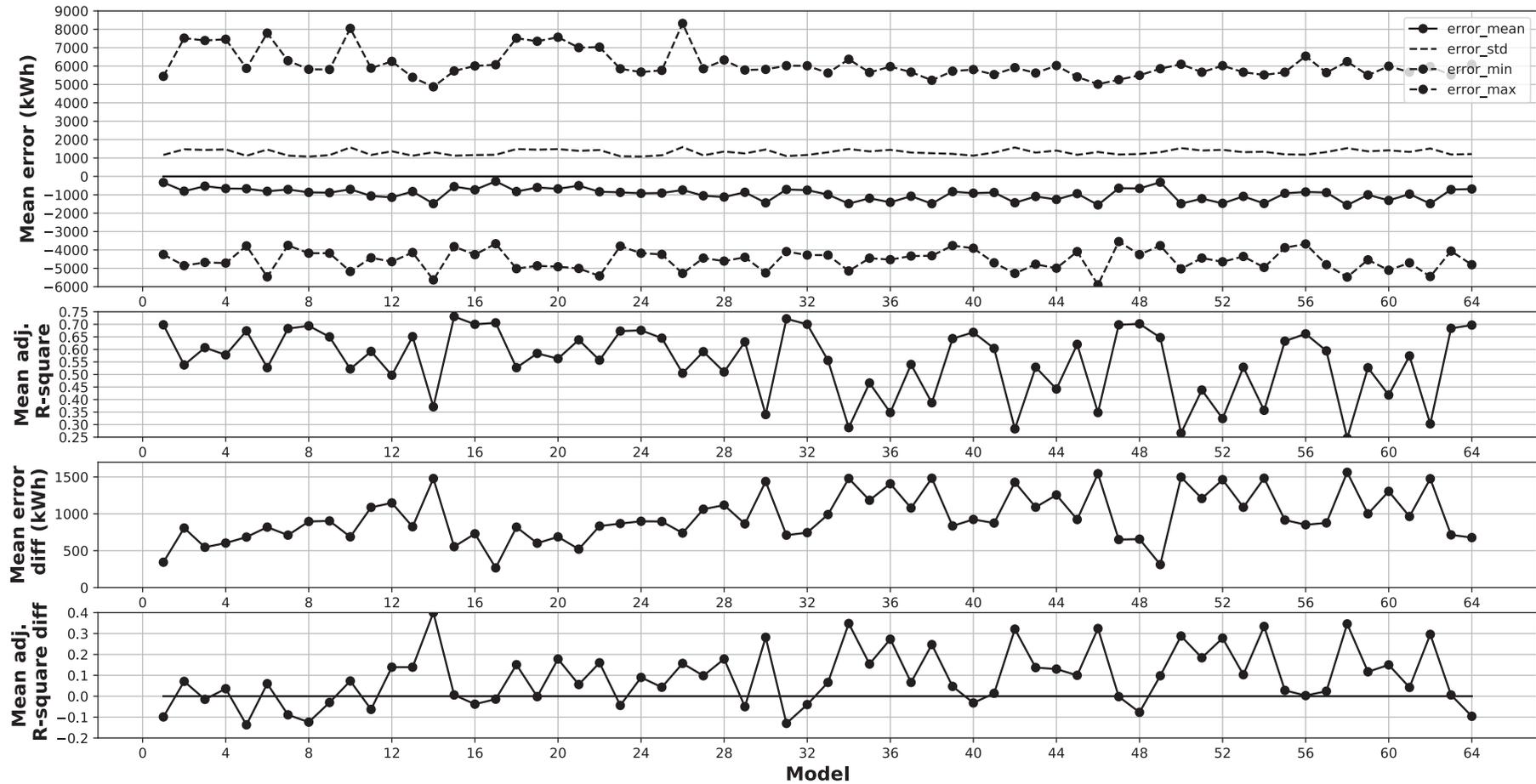


Figure A1. The modeling results from the domain approach. **Top panels:** Mean error plots for the best models of the 64 variable batches. The solid line with circles, Δ_{μ} . The dashed line without circles, Δ_{σ} . The dashed lines with circles, Δ_{min} and Δ_{max} , respectively. The solid line represents zero error and is used as reference for clarity purposes. The \bar{R}_{μ}^2 of each model. The model with the highest \bar{R}_{μ}^2 is model 15 with $\bar{R}_{\mu}^2 = 0.731$. **Bottom panels:** The difference in Δ_{μ} and \bar{R}_{μ}^2 between the training and the test data.

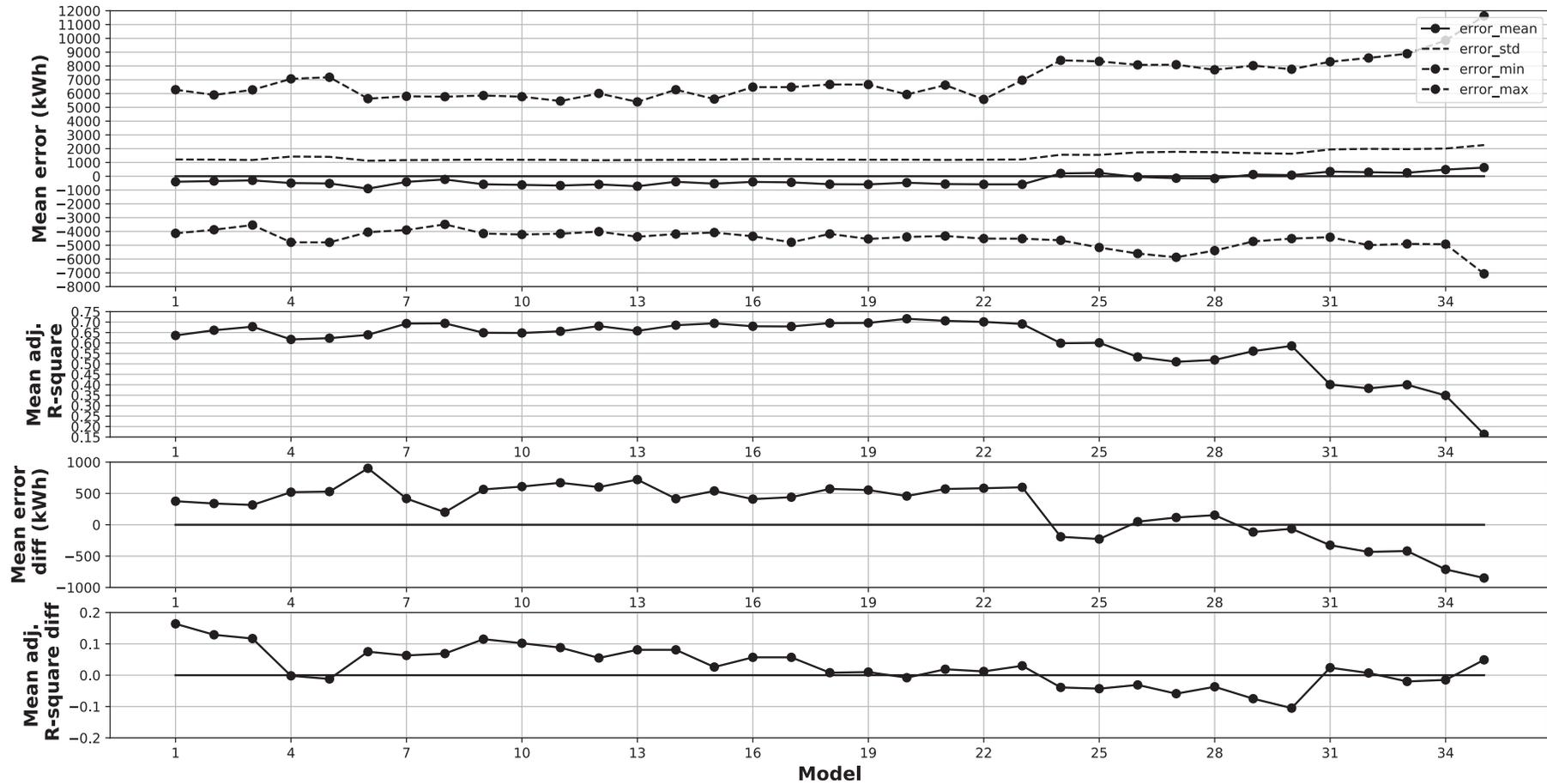


Figure A2. The modeling results from the algorithmic approach. **Top panels:** Mean error plots for the best models for each of the 36 models in the algorithmic approach. The solid line with circles, Δ_μ . The dashed line without circles, Δ_σ . The dashed lines with circles, Δ_{min} and Δ_{max} , respectively. The solid line represents zero error and is used as reference for clarity purposes. The \bar{R}_μ^2 of each model. The model with the highest \bar{R}_μ^2 is model 21 with $\bar{R}_\mu^2 = 0.716$. **Bottom panels:** The difference in Δ_μ and \bar{R}_μ^2 between the training and the test data.

Appendix A.3. Distance Correlation (dCor) Matrices

Table A3. Difference in Variable-variable distance correlation (dCor) between the training and test data. dCor values at ± 0.1 , or above, is bold and underlined.

	Delays	TTT	Process	Charging	Ext. Ref.	Melting	Refining	Tapping
Delays	0.0	-0.035	-0.054	0.018	<u>-0.112</u>	-0.041	-0.063	-0.048
TTT	-0.035	0.0	-0.046	0.036	<u>-0.121</u>	-0.022	-0.076	-0.028
Process	-0.054	-0.046	0.0	0.074	<u>-0.212</u>	-0.009	-0.059	-0.009
Charging	0.018	0.036	0.074	0.0	-0.058	-0.088	-0.078	-0.059
Ext. ref.	<u>-0.112</u>	<u>-0.121</u>	<u>-0.212</u>	-0.058	0.0	<u>-0.103</u>	-0.015	0.007
Melting	-0.041	-0.022	-0.009	-0.088	<u>-0.103</u>	0.0	-0.017	-0.067
Ext. ref.	-0.063	-0.076	-0.059	-0.078	-0.015	-0.017	0.0	-0.002
Tapping	-0.048	-0.028	-0.009	-0.059	0.007	-0.067	-0.002	0.0
Propane	<u>-0.181</u>	<u>-0.188</u>	<u>-0.241</u>	-0.073	<u>-0.121</u>	-0.068	-0.017	<u>-0.142</u>
O ₂	-0.044	-0.056	-0.044	-0.081	-0.06	-0.031	-0.037	-0.024
TotWeight	-0.018	0.007	0.081	-0.003	-0.054	-0.007	-0.078	0.016
Metal	-0.047	-0.005	0.087	-0.031	-0.055	-0.016	-0.037	0.015
Slag	-0.027	-0.032	-0.014	-0.042	-0.067	-0.016	0.013	<u>-0.112</u>
%Fe	-0.043	-0.034	-0.052	-0.008	<u>-0.125</u>	-0.056	-0.092	-0.028
%O	-0.052	-0.055	-0.065	-0.045	-0.06	-0.06	-0.039	-0.073
%Al	-0.055	-0.039	-0.031	-0.053	-0.053	-0.042	<u>-0.115</u>	-0.021
%Cr	-0.056	-0.06	-0.021	-0.036	-0.068	-0.056	0.022	-0.049
%Si	-0.048	-0.055	-0.077	-0.058	-0.058	-0.052	<u>-0.113</u>	-0.093
%C	-0.042	-0.015	-0.031	-0.063	<u>-0.162</u>	-0.073	0.079	-0.018
%Ni	-0.067	-0.059	-0.09	-0.028	<u>-0.112</u>	-0.06	-0.078	-0.039
%FeO	-0.066	-0.046	-0.035	-0.08	-0.035	-0.062	-0.067	-0.039
%SiO ₂	-0.044	-0.041	-0.031	-0.041	-0.068	-0.057	<u>-0.116</u>	-0.058
%Al ₂ O ₃	-0.089	-0.076	-0.056	<u>-0.134</u>	-0.088	-0.043	<u>-0.166</u>	-0.041
%Cr ₂ O ₃	<u>-0.116</u>	<u>-0.128</u>	-0.072	<u>-0.112</u>	<u>-0.114</u>	-0.05	<u>-0.119</u>	-0.031
%Mgo	-0.074	-0.082	-0.087	-0.044	-0.056	-0.026	-0.038	-0.021
%CaO	-0.039	-0.029	-0.026	-0.048	-0.065	-0.045	-0.084	-0.054
Type A	-0.061	-0.062	-0.054	-0.051	-0.059	-0.073	<u>-0.229</u>	-0.041
Type B	-0.033	-0.03	-0.055	-0.053	-0.088	-0.038	-0.016	-0.057
Type C	-0.006	-0.016	-0.049	-0.012	-0.035	-0.078	<u>-0.192</u>	-0.054
Type D	-0.036	-0.026	-0.038	-0.033	-0.063	-0.077	<u>-0.137</u>	-0.03
Type E	-0.049	-0.054	-0.056	-0.047	-0.058	-0.062	-0.015	-0.088
Type F	-0.034	-0.022	-0.023	-0.043	-0.02	-0.042	-0.071	-0.096
Type G	-0.06	-0.07	<u>-0.161</u>	-0.038	<u>-0.112</u>	-0.081	<u>-0.202</u>	-0.076
Type N	<u>-0.133</u>	<u>-0.164</u>	<u>-0.127</u>	-0.048	<u>-0.126</u>	-0.087	-0.091	-0.031
PreHeater	-0.055	-0.055	-0.029	-0.069	-0.066	-0.052	<u>-0.151</u>	-0.066
EE	<u>-0.135</u>	<u>-0.134</u>	<u>-0.135</u>	-0.07	<u>-0.226</u>	0.02	-0.051	0.023

Table A3. Cont.

	Propane	O ₂	TotWeight	Metal	Slag	%Fe	%O	%Al
Delays	<u>-0.181</u>	-0.044	-0.018	-0.047	-0.027	-0.043	-0.052	-0.055
TTT	<u>-0.188</u>	-0.056	0.007	-0.005	-0.032	-0.034	-0.055	-0.039
Process	<u>-0.241</u>	-0.044	0.081	0.087	-0.014	-0.052	-0.065	-0.031
Charging	-0.073	-0.081	-0.003	-0.031	-0.042	-0.008	-0.045	-0.053
Remelting	<u>-0.121</u>	-0.06	-0.054	-0.055	-0.067	<u>-0.125</u>	-0.06	-0.053
Melting	-0.068	-0.031	-0.007	-0.016	-0.016	-0.056	-0.06	-0.042
Ext. ref.	-0.017	-0.037	-0.078	-0.037	0.013	-0.092	-0.039	<u>-0.115</u>
Tapping	<u>-0.142</u>	-0.024	0.016	0.015	<u>-0.112</u>	-0.028	-0.073	-0.021
Propane	0.0	<u>0.174</u>	<u>0.201</u>	<u>0.197</u>	-0.057	-0.077	<u>-0.144</u>	0.012
O ₂	<u>0.174</u>	0.0	<u>0.135</u>	<u>0.141</u>	-0.025	-0.048	-0.049	-0.022
TotWeight	<u>0.201</u>	<u>0.135</u>	0.0	0.001	<u>-0.219</u>	-0.08	<u>-0.209</u>	0.076
Metal	<u>0.197</u>	<u>0.141</u>	0.001	0.0	-0.058	<u>-0.186</u>	<u>-0.154</u>	-0.096
Slag	-0.057	-0.025	<u>-0.219</u>	-0.058	0.0	0.098	-0.066	<u>0.155</u>
%Fe	-0.077	-0.048	-0.08	<u>-0.186</u>	0.098	0.0	-0.075	0.025
%O	<u>-0.144</u>	-0.049	<u>-0.209</u>	<u>-0.154</u>	-0.066	-0.075	0.0	-0.089
%Al	0.012	-0.022	0.076	-0.096	<u>0.155</u>	0.025	-0.089	0.0
%Cr	-0.038	-0.02	0.044	<u>-0.125</u>	0.018	<u>0.133</u>	-0.038	-0.066
%Si	-0.025	0.045	<u>-0.219</u>	-0.041	-0.094	<u>-0.183</u>	<u>-0.326</u>	-0.059
%C	0.056	<u>-0.126</u>	-0.012	-0.034	-0.046	-0.007	-0.068	-0.004
%Ni	-0.058	-0.055	<u>-0.137</u>	<u>-0.254</u>	<u>0.118</u>	-0.058	-0.081	0.033
%FeO	0.062	0.049	0.09	-0.066	<u>0.141</u>	-0.055	<u>-0.139</u>	-0.05
%SiO ₂	0.017	-0.003	<u>0.143</u>	<u>-0.131</u>	<u>0.222</u>	<u>-0.122</u>	<u>-0.145</u>	-0.046
%Al ₂ O ₃	<u>-0.168</u>	-0.059	-0.011	<u>-0.125</u>	-0.054	<u>-0.125</u>	<u>-0.296</u>	<u>-0.133</u>
%Cr ₂ O ₃	<u>-0.127</u>	-0.098	0.056	<u>-0.194</u>	<u>0.193</u>	<u>-0.182</u>	-0.088	<u>-0.108</u>
%MgO	-0.035	0.065	-0.018	-0.079	<u>0.117</u>	<u>-0.180</u>	<u>-0.236</u>	0.030
%CaO	0.038	0.014	<u>0.123</u>	<u>-0.145</u>	<u>0.264</u>	<u>-0.119</u>	<u>-0.223</u>	-0.052
Type A	-0.036	-0.042	<u>-0.211</u>	<u>-0.292</u>	-0.025	<u>-0.171</u>	-0.093	<u>-0.224</u>
Type B	-0.057	-0.047	-0.089	-0.099	-0.036	<u>-0.168</u>	<u>-0.214</u>	-0.095
Type C	-0.02	0.003	<u>-0.110</u>	<u>-0.119</u>	-0.085	-0.054	<u>-0.15</u>	-0.071
Type D	-0.047	-0.019	<u>-0.140</u>	<u>-0.144</u>	0.01	0.0	-0.094	<u>-0.109</u>
Type E	<u>-0.123</u>	-0.02	<u>-0.160</u>	<u>-0.130</u>	0.043	<u>-0.117</u>	<u>-0.189</u>	-0.027
Type F	-0.025	0.018	-0.02	-0.036	-0.034	-0.096	<u>-0.114</u>	-0.038
Type G	-0.028	<u>0.150</u>	<u>-0.205</u>	0.006	<u>0.101</u>	<u>-0.366</u>	<u>-0.344</u>	<u>-0.241</u>
Type N	<u>0.164</u>	<u>0.244</u>	<u>0.106</u>	<u>0.135</u>	-0.089	<u>-0.108</u>	<u>-0.104</u>	-0.036
PreHeater	-0.055	-0.035	-0.066	-0.043	<u>-0.137</u>	-0.095	-0.06	-0.086
EE	<u>-0.135</u>	0.009	<u>0.209</u>	<u>0.208</u>	-0.071	-0.086	<u>-0.225</u>	-0.042

Table A3. Cont.

	%Cr	%Si	%C	%Ni	%FeO	%SiO ₂	%Al ₂ O ₃
Delays	−0.056	−0.048	−0.042	−0.067	−0.066	−0.044	−0.089
TTT	−0.06	−0.055	−0.015	−0.059	−0.046	−0.041	−0.076
Process	−0.021	−0.077	−0.031	−0.09	−0.035	−0.031	−0.056
Charging	−0.036	−0.058	−0.063	−0.028	−0.08	−0.041	<u>−0.134</u>
Ext. ref.	−0.068	−0.058	<u>−0.162</u>	<u>−0.112</u>	−0.035	−0.068	−0.088
Melting	−0.056	−0.052	−0.073	−0.06	−0.062	−0.057	−0.043
Refining	0.022	<u>−0.113</u>	0.079	−0.078	−0.067	<u>−0.116</u>	<u>−0.166</u>
TT	−0.049	−0.093	−0.018	−0.039	−0.039	−0.058	−0.041
Propane	−0.038	−0.025	0.056	−0.058	0.062	0.017	<u>−0.168</u>
O ₂	−0.02	0.045	<u>−0.126</u>	−0.055	0.049	−0.003	−0.059
TotWeight	0.044	<u>−0.219</u>	−0.012	<u>−0.137</u>	0.09	<u>0.143</u>	−0.011
Metal	<u>−0.125</u>	−0.041	−0.034	<u>−0.254</u>	−0.066	<u>−0.131</u>	<u>−0.125</u>
Slag	0.018	−0.094	−0.046	<u>0.118</u>	<u>0.141</u>	<u>0.222</u>	−0.054
%Fe	<u>0.133</u>	<u>−0.183</u>	−0.007	−0.058	−0.055	<u>−0.122</u>	<u>−0.125</u>
%O	−0.038	<u>−0.326</u>	−0.068	−0.081	<u>−0.139</u>	<u>−0.145</u>	<u>−0.296</u>
%Al	−0.066	−0.059	−0.004	0.033	−0.05	−0.046	<u>−0.133</u>
%Cr	0.0	−0.066	0.028	0.016	<u>−0.136</u>	−0.092	<u>−0.135</u>
%Si	−0.066	0.0	<u>−0.146</u>	<u>−0.184</u>	<u>−0.101</u>	<u>−0.225</u>	<u>−0.337</u>
%C	0.028	<u>−0.146</u>	0.0	−0.04	0.024	0.035	−0.069
%Ni	0.016	<u>−0.184</u>	−0.04	0.0	−0.047	−0.065	<u>−0.130</u>
%FeO	<u>−0.136</u>	<u>−0.101</u>	0.024	−0.047	0.0	−0.04	0.025
%SiO ₂	−0.092	<u>−0.225</u>	0.035	−0.065	−0.04	0.0	−0.082
%Al ₂ O ₃	<u>−0.135</u>	<u>−0.337</u>	−0.069	<u>−0.130</u>	0.025	−0.082	0.0
%Cr ₂ O ₃	−0.009	<u>−0.258</u>	−0.006	<u>−0.118</u>	−0.087	−0.091	<u>0.116</u>
%MgO	−0.058	<u>−0.196</u>	0.025	<u>−0.158</u>	0.053	−0.016	<u>−0.351</u>
%CaO	−0.077	<u>−0.238</u>	0.045	−0.053	<u>−0.110</u>	−0.071	<u>−0.117</u>
Type A	<u>−0.143</u>	<u>−0.174</u>	−0.065	<u>−0.185</u>	<u>−0.244</u>	−0.049	<u>0.129</u>
Type B	−0.084	<u>−0.213</u>	−0.053	<u>−0.106</u>	−0.089	<u>−0.165</u>	<u>−0.177</u>
Type C	0.003	<u>−0.155</u>	0.085	−0.058	0.027	−0.06	<u>−0.261</u>
Type D	−0.004	<u>−0.131</u>	<u>0.183</u>	−0.017	−0.05	−0.039	<u>−0.194</u>
Type E	−0.015	<u>−0.217</u>	−0.029	−0.087	<u>−0.165</u>	<u>−0.197</u>	<u>−0.238</u>
Type F	−0.021	<u>−0.165</u>	−0.041	−0.065	−0.018	−0.056	−0.036
Type G	<u>−0.212</u>	<u>−0.231</u>	<u>−0.154</u>	<u>−0.291</u>	<u>−0.112</u>	<u>−0.339</u>	<u>−0.122</u>
Type N	−0.043	0.068	0.021	<u>−0.126</u>	−0.028	−0.056	−0.094
PreHeater	−0.092	−0.04	<u>−0.112</u>	−0.096	0.043	−0.036	<u>−0.119</u>
EE	−0.008	<u>−0.111</u>	<u>0.129</u>	−0.09	−0.091	−0.032	−0.08

Table A3. Cont.

	%Cr ₂ O ₃	%MgO	%CaO	Type A	Type B	Type C	Type D	Type E
Delays	<u>-0.116</u>	-0.074	-0.039	-0.061	-0.033	-0.006	-0.036	-0.049
TTT	<u>-0.128</u>	-0.082	-0.029	-0.062	-0.03	-0.016	-0.026	-0.054
Process	-0.072	-0.087	-0.026	-0.054	-0.055	-0.049	-0.038	-0.056
Charging	<u>-0.112</u>	-0.044	-0.048	-0.051	-0.053	-0.012	-0.033	-0.047
Ext. ref.	<u>-0.114</u>	-0.056	-0.065	-0.059	-0.088	-0.035	-0.063	-0.058
Melting	-0.05	-0.026	-0.045	-0.073	-0.038	-0.078	-0.077	-0.062
Refining	<u>-0.119</u>	-0.038	-0.084	<u>-0.229</u>	-0.016	<u>-0.192</u>	<u>-0.137</u>	-0.015
Tapping	-0.031	-0.021	-0.054	-0.041	-0.057	-0.054	-0.03	-0.088
Propane	<u>-0.127</u>	-0.035	0.038	-0.036	-0.057	-0.02	-0.047	<u>-0.123</u>
O ₂	-0.098	0.065	0.014	-0.042	-0.047	0.003	-0.019	-0.02
TotWeight	0.056	-0.018	<u>0.123</u>	<u>-0.211</u>	-0.089	<u>-0.110</u>	<u>-0.140</u>	<u>-0.160</u>
Metal	<u>-0.194</u>	-0.079	<u>-0.145</u>	<u>-0.292</u>	-0.099	<u>-0.119</u>	<u>-0.144</u>	<u>-0.130</u>
Slag	<u>0.193</u>	<u>0.117</u>	<u>0.264</u>	-0.025	-0.036	-0.085	0.01	0.043
%Fe	<u>-0.182</u>	<u>-0.180</u>	<u>-0.119</u>	<u>-0.171</u>	<u>-0.168</u>	-0.054	0.0	<u>-0.117</u>
%O	-0.088	<u>-0.236</u>	<u>-0.223</u>	-0.093	<u>-0.214</u>	<u>-0.150</u>	-0.094	<u>-0.189</u>
%Al	<u>-0.108</u>	0.03	-0.052	<u>-0.224</u>	-0.095	-0.071	<u>-0.109</u>	-0.027
%Cr	-0.009	-0.058	-0.077	<u>-0.143</u>	-0.084	0.003	-0.004	-0.015
%Si	<u>-0.258</u>	<u>-0.196</u>	<u>-0.238</u>	<u>-0.174</u>	<u>-0.213</u>	<u>-0.155</u>	<u>-0.131</u>	<u>-0.217</u>
%C	-0.006	0.025	0.045	-0.065	-0.053	0.085	<u>0.183</u>	-0.029
%Ni	<u>-0.118</u>	<u>-0.158</u>	-0.053	<u>-0.185</u>	<u>-0.106</u>	-0.058	-0.017	-0.087
%FeO	-0.087	0.053	<u>-0.110</u>	<u>-0.244</u>	-0.089	0.027	-0.05	<u>-0.165</u>
%SiO ₂	-0.091	-0.016	-0.071	-0.049	<u>-0.165</u>	-0.06	-0.039	<u>-0.197</u>
%Al ₂ O ₃	<u>0.116</u>	<u>-0.351</u>	<u>-0.117</u>	<u>0.129</u>	<u>-0.177</u>	<u>-0.261</u>	<u>-0.194</u>	<u>-0.238</u>
%Cr ₂ O ₃	0.0	<u>-0.226</u>	-0.052	0.03	<u>-0.141</u>	<u>-0.136</u>	-0.017	<u>-0.173</u>
%Mgo	<u>-0.226</u>	0.0	<u>-0.254</u>	<u>-0.126</u>	<u>-0.206</u>	-0.033	-0.044	<u>-0.164</u>
%CaO	-0.052	<u>-0.254</u>	0.0	-0.076	<u>-0.236</u>	-0.003	-0.008	<u>-0.243</u>
Type A	0.03	<u>-0.126</u>	-0.076	0.0	<u>-0.106</u>	<u>-0.225</u>	<u>-0.226</u>	<u>-0.127</u>
Type B	<u>-0.141</u>	<u>-0.206</u>	<u>-0.236</u>	<u>-0.106</u>	0.0	0.011	-0.025	<u>-0.210</u>
Type C	<u>-0.136</u>	-0.033	-0.003	<u>-0.225</u>	0.011	0.0	<u>-0.122</u>	-0.065
Type D	-0.017	-0.044	-0.008	<u>-0.226</u>	-0.025	<u>-0.122</u>	0.0	-0.01
Type E	<u>-0.173</u>	<u>-0.164</u>	<u>-0.243</u>	<u>-0.127</u>	<u>-0.210</u>	-0.065	-0.01	0.0
Type F	0.017	-0.034	-0.058	-0.065	-0.028	0.006	-0.084	<u>-0.118</u>
Type G	<u>-0.398</u>	-0.048	<u>-0.302</u>	<u>-0.256</u>	-0.072	<u>-0.209</u>	<u>-0.281</u>	<u>-0.230</u>
Type N	-0.088	-0.086	-0.054	<u>-0.107</u>	-0.054	-0.039	-0.063	-0.078
PreHeater	<u>-0.179</u>	0.027	-0.05	<u>-0.112</u>	-0.062	-0.091	<u>-0.124</u>	-0.015
EE	-0.004	-0.048	-0.074	-0.03	-0.069	<u>-0.141</u>	-0.084	<u>-0.172</u>

Table A3. Cont.

	Type F	Type G	Type N	PreHeater	EE
Delays	−0.034	−0.06	−0.133	−0.055	−0.135
TTT	−0.022	−0.07	−0.164	−0.055	−0.134
Process	−0.023	−0.161	−0.127	−0.029	−0.135
Charging	−0.043	−0.038	−0.048	−0.069	−0.07
Ext. ref.	−0.02	−0.112	−0.126	−0.066	−0.226
Melting	−0.042	−0.081	−0.087	−0.052	0.02
Refining	−0.071	−0.202	−0.091	−0.151	−0.051
Tapping	−0.096	−0.076	−0.031	−0.066	0.023
Propane	−0.025	−0.028	0.164	−0.055	−0.135
O ₂	0.018	0.150	0.244	−0.035	0.009
TotWeight	−0.02	−0.205	0.106	−0.066	0.209
Metal	−0.036	0.006	0.135	−0.043	0.208
Slag	−0.034	0.101	−0.089	−0.137	−0.071
%Fe	−0.096	−0.366	−0.108	−0.095	−0.086
%O	−0.114	−0.344	−0.104	−0.06	−0.225
%Al	−0.038	−0.241	−0.036	−0.086	−0.042
%Cr	−0.021	−0.212	−0.043	−0.092	−0.008
%Si	−0.165	−0.231	0.068	−0.04	−0.111
%C	−0.041	−0.154	0.021	−0.112	0.129
%Ni	−0.065	−0.291	−0.126	−0.096	−0.09
%FeO	−0.018	−0.112	−0.028	0.043	−0.091
%SiO ₂	−0.056	−0.339	−0.056	−0.036	−0.032
%Al ₂ O ₃	−0.036	−0.122	−0.094	−0.119	−0.08
%Cr ₂ O ₃	0.017	−0.398	−0.088	−0.179	−0.004
%MgO	−0.034	−0.048	−0.086	0.027	−0.048
%CaO	−0.058	−0.302	−0.054	−0.05	−0.074
Type A	−0.065	−0.256	−0.107	−0.112	−0.03
Type B	−0.028	−0.072	−0.054	−0.062	−0.069
Type C	0.006	−0.209	−0.039	−0.091	−0.141
Type D	−0.084	−0.281	−0.063	−0.124	−0.084
Type E	−0.118	−0.230	−0.078	−0.015	−0.172
Type F	0.0	−0.195	−0.02	−0.028	−0.005
Type G	−0.195	0.0	0.116	−0.068	−0.004
Type N	−0.02	0.116	0.0	−0.126	−0.205
PreHeater	−0.028	−0.068	−0.126	0.0	−0.014
EE	−0.005	−0.004	−0.205	−0.014	0.0

Table A4. Variable-variable dCor for the training data. dCor values at ± 0.1 , or above, is bold and underlined.

	Delays	TTT	Process	Charging	Ext. Ref.	Melting	Refining	Tapping	Propane
Delays	1.0	<u>0.944</u>	<u>0.488</u>	<u>0.557</u>	<u>0.148</u>	<u>0.131</u>	0.054	<u>0.199</u>	0.09
TTT	<u>0.944</u>	1.0	<u>0.558</u>	<u>0.557</u>	<u>0.189</u>	<u>0.146</u>	<u>0.139</u>	<u>0.218</u>	<u>0.102</u>
Process	<u>0.488</u>	<u>0.558</u>	1.0	<u>0.351</u>	<u>0.29</u>	<u>0.203</u>	<u>0.309</u>	<u>0.44</u>	<u>0.212</u>
Charging	<u>0.557</u>	<u>0.557</u>	<u>0.351</u>	1.0	0.013	0.029	0.024	0.025	0.072
Ext. ref.	<u>0.148</u>	<u>0.189</u>	<u>0.29</u>	0.013	1.0	0.022	0.048	0.072	0.051
Melting	<u>0.131</u>	<u>0.146</u>	<u>0.203</u>	0.029	0.022	1.0	<u>0.103</u>	0.018	0.066
Refining	0.054	<u>0.139</u>	<u>0.309</u>	0.024	0.048	<u>0.103</u>	1.0	<u>0.115</u>	<u>0.144</u>
Tapping	<u>0.199</u>	<u>0.218</u>	<u>0.440</u>	0.025	0.072	0.018	<u>0.115</u>	1.0	<u>0.118</u>
Propane	0.09	<u>0.102</u>	<u>0.212</u>	0.072	0.051	0.066	<u>0.144</u>	<u>0.118</u>	1.0
O ₂	0.042	0.029	0.044	0.052	0.033	0.046	0.098	0.057	<u>0.322</u>
TotalWeight	0.057	0.082	<u>0.206</u>	0.07	0.014	<u>0.110</u>	0.078	<u>0.131</u>	<u>0.301</u>
Met	0.043	0.082	<u>0.186</u>	0.053	0.011	<u>0.106</u>	<u>0.164</u>	<u>0.112</u>	<u>0.298</u>
Slag	0.035	0.042	0.068	0.034	0.018	0.067	<u>0.208</u>	0.054	0.089
%Fe	0.043	0.06	0.066	0.072	0.045	0.059	<u>0.151</u>	0.052	0.066
%O	0.033	0.041	0.036	0.037	0.013	0.029	<u>0.134</u>	0.08	0.045
%Al	0.046	0.051	0.063	0.055	0.033	0.055	<u>0.153</u>	0.062	<u>0.124</u>
%Cr	0.05	0.045	0.071	0.086	0.039	0.062	<u>0.198</u>	0.051	0.059
%Si	0.046	0.034	0.04	0.031	0.031	0.063	0.067	0.049	<u>0.192</u>
%C	0.032	0.065	0.074	0.026	0.03	0.04	<u>0.193</u>	0.057	<u>0.159</u>
%Ni	0.034	0.05	0.039	0.052	0.039	0.059	<u>0.181</u>	0.044	0.052
%FeO	0.038	0.055	0.042	0.057	0.04	0.038	<u>0.124</u>	0.036	<u>0.166</u>
%SiO ₂	0.046	0.06	0.066	0.057	0.031	0.062	<u>0.187</u>	0.038	<u>0.105</u>
%Al ₂ O ₃	0.021	0.029	0.051	0.024	0.018	0.047	0.049	0.037	0.059
%Cr ₂ O ₃	0.027	0.03	0.055	0.03	0.016	0.061	<u>0.127</u>	0.082	0.056
%MgO	0.026	0.033	0.037	0.059	0.017	0.054	<u>0.129</u>	0.071	0.086
%CaO	0.047	0.062	0.067	0.054	0.029	0.069	<u>0.185</u>	0.041	<u>0.127</u>
Type A	0.017	0.026	0.029	0.033	0.014	0.035	<u>0.107</u>	0.05	0.073
Type B	0.031	0.03	0.029	0.015	0.013	0.025	0.068	0.038	0.04
Type C	0.054	0.079	0.062	0.063	0.073	0.035	<u>0.181</u>	0.053	0.096
Type D	0.047	0.087	0.085	0.049	0.06	0.036	<u>0.235</u>	0.067	0.068
Type E	0.041	0.043	0.045	0.035	0.014	0.047	<u>0.134</u>	0.047	0.059
Type F	0.017	0.022	0.025	0.016	0.014	0.02	0.045	0.024	0.031
Type G	0.019	0.026	0.032	0.036	0.028	0.037	0.06	0.071	0.092
Type N	0.025	0.034	0.07	0.042	0.048	0.043	<u>0.108</u>	<u>0.118</u>	<u>0.278</u>
PreHeater	0.021	0.032	0.051	0.022	0.025	0.029	0.055	0.052	<u>0.132</u>
EE	<u>0.135</u>	<u>0.262</u>	<u>0.399</u>	0.025	<u>0.237</u>	<u>0.107</u>	<u>0.448</u>	<u>0.152</u>	0.082

Table 4A. Cont.

	O ₂	TotWeight	Metal	Slag	%Fe	%O	%Al	%Cr	%Si
Delays	0.042	0.057	0.043	0.035	0.043	0.033	0.046	0.05	0.046
TTT	0.029	0.082	0.082	0.042	0.06	0.041	0.051	0.045	0.034
Process	0.044	0.206	0.186	0.068	0.066	0.036	0.063	0.071	0.04
Charging	0.052	0.07	0.053	0.034	0.072	0.037	0.055	0.086	0.031
Ext. ref.	0.033	0.014	0.011	0.018	0.045	0.013	0.033	0.039	0.031
Melting	0.046	0.110	0.106	0.067	0.059	0.029	0.055	0.062	0.063
Refining	0.098	0.078	0.164	0.208	0.151	0.134	0.153	0.198	0.067
Tapping	0.057	0.131	0.112	0.054	0.052	0.08	0.062	0.051	0.049
Propane	0.322	0.301	0.298	0.089	0.066	0.045	0.124	0.059	0.192
O ₂	1.0	0.235	0.262	0.067	0.056	0.047	0.071	0.055	0.137
TotWeight	0.235	1.0	0.793	0.342	0.273	0.151	0.274	0.260	0.154
Metal	0.262	0.793	1.0	0.170	0.309	0.173	0.132	0.220	0.274
Slag	0.067	0.342	0.170	1.0	0.341	0.379	0.500	0.354	0.287
%Fe	0.056	0.273	0.309	0.341	1.0	0.533	0.336	0.571	0.426
%O	0.047	0.151	0.173	0.379	0.533	1.0	0.206	0.266	0.603
%Al	0.071	0.274	0.132	0.500	0.336	0.206	1.0	0.479	0.174
%Cr	0.055	0.260	0.220	0.354	0.571	0.266	0.479	1.0	0.233
%Si	0.137	0.154	0.274	0.287	0.426	0.603	0.174	0.233	1.0
%C	0.042	0.112	0.151	0.186	0.274	0.129	0.200	0.269	0.071
%Ni	0.051	0.246	0.249	0.387	0.822	0.377	0.450	0.488	0.241
%FeO	0.134	0.228	0.184	0.356	0.356	0.200	0.611	0.333	0.225
%SiO ₂	0.086	0.358	0.203	0.570	0.429	0.255	0.612	0.405	0.205
%Al ₂ O ₃	0.048	0.222	0.091	0.243	0.269	0.177	0.118	0.126	0.154
%Cr ₂ O ₃	0.04	0.318	0.116	0.447	0.318	0.248	0.271	0.308	0.096
%MgO	0.138	0.184	0.150	0.328	0.132	0.158	0.212	0.137	0.151
%CaO	0.099	0.321	0.205	0.577	0.434	0.239	0.585	0.411	0.263
Type A	0.03	0.245	0.109	0.426	0.318	0.231	0.253	0.184	0.100
Type B	0.038	0.085	0.101	0.115	0.296	0.140	0.099	0.146	0.165
Type C	0.127	0.255	0.214	0.389	0.450	0.259	0.484	0.424	0.176
Type D	0.079	0.226	0.230	0.547	0.548	0.308	0.508	0.533	0.196
Type E	0.07	0.100	0.179	0.361	0.509	0.658	0.254	0.354	0.655
Type F	0.059	0.034	0.057	0.074	0.082	0.081	0.04	0.056	0.046
Type G	0.255	0.086	0.238	0.389	0.111	0.235	0.063	0.061	0.388
Type N	0.334	0.281	0.286	0.049	0.051	0.044	0.076	0.055	0.203
PreHeater	0.054	0.063	0.091	0.059	0.035	0.048	0.063	0.029	0.07
EE	0.123	0.325	0.344	0.108	0.095	0.032	0.098	0.131	0.105

Table 4A. Cont.

	%C	%Ni	%FeO	%SiO ₂	%Al ₂ O ₃	%Cr ₂ O ₃	%Mgo	%CaO
Delays	0.032	0.034	0.038	0.046	0.021	0.027	0.026	0.047
TTT	0.065	0.05	0.055	0.06	0.029	0.03	0.033	0.062
Process	0.074	0.039	0.042	0.066	0.051	0.055	0.037	0.067
Charging	0.026	0.052	0.057	0.057	0.024	0.03	0.059	0.054
Ext. ref.	0.03	0.039	0.04	0.031	0.018	0.016	0.017	0.029
Melting	0.04	0.059	0.038	0.062	0.047	0.061	0.054	0.069
Refining	0.193	0.181	0.124	0.187	0.049	0.127	0.129	0.185
Tapping	0.057	0.044	0.036	0.038	0.037	0.082	0.071	0.041
Propane	0.159	0.052	0.166	0.105	0.059	0.056	0.086	0.127
O ₂	0.042	0.051	0.134	0.086	0.048	0.04	0.138	0.099
TotWeight	0.112	0.246	0.228	0.358	0.222	0.318	0.184	0.321
Metal	0.151	0.249	0.184	0.203	0.091	0.116	0.150	0.205
Slag	0.186	0.387	0.356	0.570	0.243	0.447	0.328	0.577
%Fe	0.274	0.822	0.356	0.429	0.269	0.318	0.132	0.434
%O	0.129	0.377	0.200	0.255	0.177	0.248	0.158	0.239
%Al	0.200	0.450	0.611	0.612	0.118	0.271	0.212	0.585
%Cr	0.269	0.488	0.333	0.405	0.126	0.308	0.137	0.411
%Si	0.071	0.241	0.225	0.205	0.154	0.096	0.151	0.263
%C	1.0	0.237	0.125	0.195	0.068	0.159	0.130	0.209
%Ni	0.237	1.0	0.360	0.434	0.236	0.349	0.152	0.426
%FeO	0.125	0.360	1.0	0.719	0.243	0.174	0.251	0.667
%SiO ₂	0.195	0.434	0.719	1.0	0.311	0.530	0.334	0.891
%Al ₂ O ₃	0.068	0.236	0.243	0.311	1.0	0.631	0.121	0.343
%Cr ₂ O ₃	0.159	0.349	0.174	0.530	0.631	1.0	0.145	0.524
%MgO	0.130	0.152	0.251	0.334	0.121	0.145	1.0	0.177
%CaO	0.209	0.426	0.667	0.891	0.343	0.524	0.177	1.0
Type A	0.141	0.322	0.227	0.495	0.448	0.508	0.225	0.450
Type B	0.066	0.263	0.212	0.191	0.084	0.082	0.119	0.179
Type C	0.288	0.440	0.492	0.496	0.092	0.263	0.183	0.470
Type D	0.355	0.527	0.524	0.603	0.138	0.385	0.212	0.576
Type E	0.146	0.373	0.220	0.264	0.243	0.217	0.177	0.296
Type F	0.031	0.062	0.069	0.084	0.056	0.105	0.047	0.076
Type G	0.106	0.089	0.225	0.175	0.099	0.067	0.219	0.180
Type N	0.126	0.04	0.110	0.073	0.039	0.036	0.118	0.064
PreHeater	0.094	0.035	0.128	0.106	0.035	0.026	0.103	0.089
EE	0.226	0.097	0.032	0.079	0.130	0.095	0.193	0.057

Table 4A. Cont.

	Type A	Type B	Type C	Type D	Type E	Type F	Type G	Type N	PreHeater
Delays	0.017	0.031	0.054	0.047	0.041	0.017	0.019	0.025	0.021
TTT	0.026	0.03	0.079	0.087	0.043	0.022	0.026	0.034	0.032
Process	0.029	0.029	0.062	0.085	0.045	0.025	0.032	0.07	0.051
Charging	0.033	0.015	0.063	0.049	0.035	0.016	0.036	0.042	0.022
Ext. ref.	0.014	0.013	0.073	0.06	0.014	0.014	0.028	0.048	0.025
Melting	0.035	0.025	0.035	0.036	0.047	0.02	0.037	0.043	0.029
Refining	0.107	0.068	0.181	0.235	0.134	0.045	0.06	0.108	0.055
Tapping	0.05	0.038	0.053	0.067	0.047	0.024	0.071	0.118	0.052
Propane	0.073	0.04	0.096	0.068	0.059	0.031	0.092	0.278	0.132
O ₂	0.03	0.038	0.127	0.079	0.07	0.059	0.255	0.334	0.054
TotWeight	0.245	0.085	0.255	0.226	0.100	0.034	0.086	0.281	0.063
Metal	0.109	0.101	0.214	0.230	0.179	0.057	0.238	0.286	0.091
Slag	0.426	0.115	0.389	0.547	0.361	0.074	0.389	0.049	0.059
%Fe	0.318	0.296	0.450	0.548	0.509	0.082	0.111	0.051	0.035
%O	0.231	0.140	0.259	0.308	0.658	0.081	0.235	0.044	0.048
%Al	0.253	0.099	0.484	0.508	0.254	0.04	0.063	0.076	0.063
%Cr	0.184	0.146	0.424	0.533	0.354	0.056	0.061	0.055	0.029
%Si	0.100	0.165	0.176	0.196	0.655	0.046	0.388	0.203	0.07
%C	0.141	0.066	0.288	0.355	0.146	0.031	0.106	0.126	0.094
%Ni	0.322	0.263	0.440	0.527	0.373	0.062	0.089	0.04	0.035
%FeO	0.227	0.212	0.492	0.524	0.220	0.069	0.225	0.110	0.128
%SiO ₂	0.495	0.191	0.496	0.603	0.264	0.084	0.175	0.073	0.106
%Al ₂ O ₃	0.448	0.084	0.092	0.138	0.243	0.056	0.099	0.039	0.035
%Cr ₂ O ₃	0.508	0.082	0.263	0.385	0.217	0.105	0.067	0.036	0.026
%MgO	0.225	0.119	0.183	0.212	0.177	0.047	0.219	0.118	0.103
%CaO	0.450	0.179	0.470	0.576	0.296	0.076	0.180	0.064	0.089
Type A	1.0	0.075	0.329	0.414	0.194	0.054	0.05	0.042	0.035
Type B	0.075	1.0	0.252	0.198	0.205	0.056	0.158	0.034	0.039
Type C	0.329	0.252	1.0	0.807	0.294	0.116	0.112	0.084	0.055
Type D	0.414	0.198	0.807	1.0	0.372	0.043	0.07	0.043	0.041
Type E	0.194	0.205	0.294	0.372	1.0	0.059	0.257	0.107	0.075
Type F	0.054	0.056	0.116	0.043	0.059	1.0	0.026	0.026	0.03
Type G	0.05	0.158	0.112	0.07	0.257	0.026	1.0	0.238	0.032
Type N	0.042	0.034	0.084	0.043	0.107	0.026	0.238	1.0	0.062
PreHeater	0.035	0.039	0.055	0.041	0.075	0.03	0.032	0.062	1.0
EE	0.064	0.111	0.150	0.187	0.128	0.041	0.151	0.098	0.06

Table 4A. Cont.

	EE
Delays	0.135
TTT	0.262
Process	0.399
Charging	0.025
Ext. ref.	0.237
Melting	0.107
Refining	0.448
Tapping	0.152
Propane	0.082
O ₂	0.123
TotWeight	0.325
Metal	0.344
Slag	0.108
%Fe	0.095
%O	0.032
%Al	0.098
%Cr	0.131
%Si	0.105
%C	0.226
%Ni	0.097
%FeO	0.032
%SiO ₂	0.079
%Al ₂ O ₃	0.130
%Cr ₂ O ₃	0.095
%MgO	0.193
%CaO	0.057
Type A	0.064
Type B	0.111
Type C	0.150
Type D	0.187
Type E	0.128
Type F	0.041
Type G	0.151
Type N	0.098
PreHeater	0.06
EE	1.0

Table A5. Variable-variable distance correlation for the test data.

	Delays	TTT	Process	Charging	Ext. Ref.	Melting	Refining	Tapping
Delays	1.0	0.979	0.542	0.539	0.26	0.172	0.117	0.247
TTT	0.979	1.0	0.604	0.521	0.31	0.168	0.215	0.246
Process	0.542	0.604	1.0	0.277	0.502	0.212	0.368	0.449
Charging	0.539	0.521	0.277	1.0	0.071	0.117	0.102	0.084
Ext. ref.	0.26	0.31	0.502	0.071	1.0	0.125	0.063	0.065
Melting	0.172	0.168	0.212	0.117	0.125	1.0	0.12	0.085
Refining	0.117	0.215	0.368	0.102	0.063	0.12	1.0	0.117
Tapping	0.247	0.246	0.449	0.084	0.065	0.085	0.117	1.0
Propane	0.271	0.29	0.453	0.145	0.172	0.134	0.161	0.26
O ₂	0.086	0.085	0.088	0.133	0.093	0.077	0.135	0.081
TotWeight	0.075	0.075	0.125	0.073	0.068	0.117	0.156	0.115
Metal	0.09	0.087	0.099	0.084	0.066	0.122	0.201	0.097
Slag	0.062	0.074	0.082	0.076	0.085	0.083	0.195	0.166
%Fe	0.086	0.094	0.118	0.08	0.17	0.115	0.243	0.08
%O	0.085	0.096	0.101	0.082	0.073	0.089	0.173	0.153
%Al	0.101	0.09	0.094	0.108	0.086	0.097	0.268	0.083
%Cr	0.106	0.105	0.092	0.122	0.107	0.118	0.176	0.1
%Si	0.094	0.089	0.117	0.089	0.089	0.115	0.18	0.142
%C	0.074	0.08	0.105	0.089	0.192	0.113	0.114	0.075
%Ni	0.101	0.109	0.129	0.08	0.151	0.119	0.259	0.083
%FeO	0.104	0.101	0.077	0.137	0.075	0.1	0.191	0.075
%SiO ₂	0.09	0.101	0.097	0.098	0.099	0.119	0.303	0.096
%Al ₂ O ₃	0.11	0.105	0.107	0.158	0.106	0.09	0.215	0.078
%Cr ₂ O ₃	0.143	0.158	0.127	0.142	0.13	0.111	0.246	0.113
%MgO	0.1	0.115	0.124	0.103	0.073	0.08	0.167	0.092
%CaO	0.086	0.091	0.093	0.102	0.094	0.114	0.269	0.095
Type A	0.078	0.088	0.083	0.084	0.073	0.108	0.336	0.091
Type B	0.064	0.06	0.084	0.068	0.101	0.063	0.084	0.095
Type C	0.06	0.095	0.111	0.075	0.108	0.113	0.373	0.107
Type D	0.083	0.113	0.123	0.082	0.123	0.113	0.372	0.097
Type E	0.09	0.097	0.101	0.082	0.072	0.109	0.149	0.135
Type F	0.051	0.044	0.048	0.059	0.034	0.062	0.116	0.12
Type G	0.079	0.096	0.193	0.074	0.14	0.118	0.262	0.147
Type N	0.158	0.198	0.197	0.09	0.174	0.13	0.199	0.149
PreHeater	0.076	0.087	0.08	0.091	0.091	0.081	0.206	0.118
EE	0.27	0.396	0.534	0.095	0.463	0.087	0.499	0.129

Table 5A. Cont.

	Propane	O ₂	TotWeight	Metal	Slag	%Fe	%O	%Al
Delays	0.271	0.086	0.075	0.09	0.062	0.086	0.085	0.101
TTT	0.29	0.085	0.075	0.087	0.074	0.094	0.096	0.09
Process	0.453	0.088	0.125	0.099	0.082	0.118	0.101	0.094
Charging	0.145	0.133	0.073	0.084	0.076	0.08	0.082	0.108
Ext. ref.	0.172	0.093	0.068	0.066	0.085	0.17	0.073	0.086
Melting	0.134	0.077	0.117	0.122	0.083	0.115	0.089	0.097
Refining	0.161	0.135	0.156	0.201	0.195	0.243	0.173	0.268
Tapping	0.26	0.081	0.115	0.097	0.166	0.08	0.153	0.083
Propane	1.0	0.148	0.1	0.101	0.146	0.143	0.189	0.112
O ₂	0.148	1.0	0.1	0.121	0.092	0.104	0.096	0.093
TotWeight	0.1	0.1	1.0	0.792	0.561	0.353	0.36	0.198
Metal	0.101	0.121	0.792	1.0	0.228	0.495	0.327	0.228
Slag	0.146	0.092	0.561	0.228	1.0	0.243	0.445	0.345
%Fe	0.143	0.104	0.353	0.495	0.243	1.0	0.608	0.311
%O	0.189	0.096	0.36	0.327	0.445	0.608	1.0	0.295
%Al	0.112	0.093	0.198	0.228	0.345	0.311	0.295	1.0
%Cr	0.097	0.075	0.216	0.345	0.336	0.438	0.304	0.545
%Si	0.217	0.092	0.373	0.315	0.381	0.609	0.929	0.233
%C	0.103	0.168	0.124	0.185	0.232	0.281	0.197	0.204
%Ni	0.11	0.106	0.383	0.503	0.269	0.88	0.458	0.417
%FeO	0.104	0.085	0.138	0.25	0.215	0.411	0.339	0.661
%SiO ₂	0.088	0.089	0.215	0.334	0.348	0.551	0.4	0.658
%Al ₂ O ₃	0.227	0.107	0.233	0.216	0.297	0.394	0.473	0.251
%Cr ₂ O ₃	0.183	0.138	0.262	0.31	0.254	0.5	0.336	0.379
%MgO	0.121	0.073	0.202	0.229	0.211	0.312	0.394	0.182
%CaO	0.089	0.085	0.198	0.35	0.313	0.553	0.462	0.637
Type A	0.109	0.072	0.456	0.401	0.451	0.489	0.324	0.477
Type B	0.097	0.085	0.174	0.2	0.151	0.464	0.354	0.194
Type C	0.116	0.124	0.365	0.333	0.474	0.504	0.409	0.555
Type D	0.115	0.098	0.366	0.374	0.537	0.548	0.402	0.617
Type E	0.182	0.09	0.26	0.309	0.318	0.626	0.847	0.281
Type F	0.056	0.041	0.054	0.093	0.108	0.178	0.195	0.078
Type G	0.12	0.105	0.291	0.232	0.288	0.477	0.579	0.304
Type N	0.114	0.09	0.175	0.151	0.138	0.159	0.148	0.112
PreHeater	0.187	0.089	0.129	0.134	0.196	0.13	0.108	0.149
EE	0.217	0.114	0.116	0.136	0.179	0.181	0.257	0.14

Table 5A. Cont.

	%Cr	%Si	%C	%Ni	%Feo	%Sio2	%Al2o3
Delays	0.106	0.094	0.074	0.101	0.104	0.09	0.11
TTT	0.105	0.089	0.08	0.109	0.101	0.101	0.105
Process	0.092	0.117	0.105	0.129	0.077	0.097	0.107
Charging	0.122	0.089	0.089	0.08	0.137	0.098	0.158
Ext. ref.	0.107	0.089	0.192	0.151	0.075	0.099	0.106
Melting	0.118	0.115	0.113	0.119	0.1	0.119	0.09
Refining	0.176	0.18	0.114	0.259	0.191	0.303	0.215
Tapping	0.1	0.142	0.075	0.083	0.075	0.096	0.078
Propane	0.097	0.217	0.103	0.11	0.104	0.088	0.227
O ₂	0.075	0.092	0.168	0.106	0.085	0.089	0.107
TotWeight	0.216	0.373	0.124	0.383	0.138	0.215	0.233
Metal	0.345	0.315	0.185	0.503	0.25	0.334	0.216
Slag	0.336	0.381	0.232	0.269	0.215	0.348	0.297
%Fe	0.438	0.609	0.281	0.88	0.411	0.551	0.394
%O	0.304	0.929	0.197	0.458	0.339	0.4	0.473
%Al	0.545	0.233	0.204	0.417	0.661	0.658	0.251
%Cr	1.0	0.299	0.241	0.472	0.469	0.497	0.261
%Si	0.299	1.0	0.217	0.425	0.326	0.43	0.491
%C	0.241	0.217	1.0	0.277	0.101	0.16	0.137
%Ni	0.472	0.425	0.277	1.0	0.407	0.499	0.366
%FeO	0.469	0.326	0.101	0.407	1.0	0.759	0.218
%SiO ₂	0.497	0.43	0.16	0.499	0.759	1.0	0.393
%Al ₂ O ₃	0.261	0.491	0.137	0.366	0.218	0.393	1.0
%Cr ₂ O ₃	0.317	0.354	0.165	0.467	0.261	0.621	0.515
%MgO	0.195	0.347	0.105	0.31	0.198	0.35	0.472
%CaO	0.488	0.501	0.164	0.479	0.777	0.962	0.46
Type A	0.327	0.274	0.206	0.507	0.471	0.544	0.319
Type B	0.23	0.378	0.119	0.369	0.301	0.356	0.261
Type C	0.421	0.331	0.203	0.498	0.465	0.556	0.353
Type D	0.537	0.327	0.172	0.544	0.574	0.642	0.332
Type E	0.369	0.872	0.175	0.46	0.385	0.461	0.481
Type F	0.077	0.211	0.072	0.127	0.087	0.14	0.092
Type G	0.273	0.619	0.26	0.38	0.337	0.514	0.221
Type N	0.098	0.135	0.105	0.166	0.138	0.129	0.133
PreHeater	0.121	0.11	0.206	0.131	0.085	0.142	0.154
EE	0.139	0.216	0.097	0.187	0.123	0.111	0.21

Table 5A. Cont.

	%Cr ₂ O ₃	%Mgo	%CaO	Type A	Type B	Type C	Type D	Type E
Delays	0.143	0.1	0.086	0.078	0.064	0.06	0.083	0.09
TTT	0.158	0.115	0.091	0.088	0.06	0.095	0.113	0.097
Process	0.127	0.124	0.093	0.083	0.084	0.111	0.123	0.101
Charging	0.142	0.103	0.102	0.084	0.068	0.075	0.082	0.082
Ext. ref.	0.13	0.073	0.094	0.073	0.101	0.108	0.123	0.072
Melting	0.111	0.08	0.114	0.108	0.063	0.113	0.113	0.109
Refining	0.246	0.167	0.269	0.336	0.084	0.373	0.372	0.149
Tapping	0.113	0.092	0.095	0.091	0.095	0.107	0.097	0.135
Propane	0.183	0.121	0.089	0.109	0.097	0.116	0.115	0.182
O ₂	0.138	0.073	0.085	0.072	0.085	0.124	0.098	0.09
TotWeight	0.262	0.202	0.198	0.456	0.174	0.365	0.366	0.26
Metal	0.31	0.229	0.35	0.401	0.2	0.333	0.374	0.309
Slag	0.254	0.211	0.313	0.451	0.151	0.474	0.537	0.318
%Fe	0.5	0.312	0.553	0.489	0.464	0.504	0.548	0.626
%O	0.336	0.394	0.462	0.324	0.354	0.409	0.402	0.847
%Al	0.379	0.182	0.637	0.477	0.194	0.555	0.617	0.281
%Cr	0.317	0.195	0.488	0.327	0.23	0.421	0.537	0.369
%Si	0.354	0.347	0.501	0.274	0.378	0.331	0.327	0.872
%C	0.165	0.105	0.164	0.206	0.119	0.203	0.172	0.175
%Ni	0.467	0.31	0.479	0.507	0.369	0.498	0.544	0.46
%FeO	0.261	0.198	0.777	0.471	0.301	0.465	0.574	0.385
%SiO ₂	0.621	0.35	0.962	0.544	0.356	0.556	0.642	0.461
%Al ₂ O ₃	0.515	0.472	0.46	0.319	0.261	0.353	0.332	0.481
%Cr ₂ O ₃	1.0	0.371	0.576	0.478	0.223	0.399	0.402	0.39
%MgO	0.371	1.0	0.431	0.351	0.325	0.216	0.256	0.341
%CaO	0.576	0.431	1.0	0.526	0.415	0.473	0.584	0.539
Type A	0.478	0.351	0.526	1.0	0.181	0.554	0.64	0.321
Type B	0.223	0.325	0.415	0.181	1.0	0.241	0.223	0.415
Type C	0.399	0.216	0.473	0.554	0.241	1.0	0.929	0.359
Type D	0.402	0.256	0.584	0.64	0.223	0.929	1.0	0.382
Type E	0.39	0.341	0.539	0.321	0.415	0.359	0.382	1.0
Type F	0.088	0.081	0.134	0.119	0.084	0.11	0.127	0.177
Type G	0.465	0.267	0.482	0.306	0.23	0.321	0.351	0.487
Type N	0.124	0.204	0.118	0.149	0.088	0.123	0.106	0.185
PreHeater	0.205	0.076	0.139	0.147	0.101	0.146	0.165	0.09
EE	0.099	0.241	0.131	0.094	0.18	0.291	0.271	0.3

Table 5A. Cont.

	Type F	Type G	Type N	PreHeater	EE
Delays	0.051	0.079	0.158	0.076	0.27
TTT	0.044	0.096	0.198	0.087	0.396
Process	0.048	0.193	0.197	0.08	0.534
Charging	0.059	0.074	0.09	0.091	0.095
Ext. ref.	0.034	0.14	0.174	0.091	0.463
Melting	0.062	0.118	0.13	0.081	0.087
Refining	0.116	0.262	0.199	0.206	0.499
Tapping	0.12	0.147	0.149	0.118	0.129
Propane	0.056	0.12	0.114	0.187	0.217
O ₂	0.041	0.105	0.09	0.089	0.114
TotWeight	0.054	0.291	0.175	0.129	0.116
Metal	0.093	0.232	0.151	0.134	0.136
Slag	0.108	0.288	0.138	0.196	0.179
%Fe	0.178	0.477	0.159	0.13	0.181
%O	0.195	0.579	0.148	0.108	0.257
%Al	0.078	0.304	0.112	0.149	0.14
%Cr	0.077	0.273	0.098	0.121	0.139
%Si	0.211	0.619	0.135	0.11	0.216
%C	0.072	0.26	0.105	0.206	0.097
%Ni	0.127	0.38	0.166	0.131	0.187
%FeO	0.087	0.337	0.138	0.085	0.123
%SiO ₂	0.14	0.514	0.129	0.142	0.111
%Al ₂ O ₃	0.092	0.221	0.133	0.154	0.21
%Cr ₂ O ₃	0.088	0.465	0.124	0.205	0.099
%MgO	0.081	0.267	0.204	0.076	0.241
%CaO	0.134	0.482	0.118	0.139	0.131
Type A	0.119	0.306	0.149	0.147	0.094
Type B	0.084	0.23	0.088	0.101	0.18
Type C	0.11	0.321	0.123	0.146	0.291
Type D	0.127	0.351	0.106	0.165	0.271
Type E	0.177	0.487	0.185	0.09	0.3
Type F	1.0	0.221	0.046	0.058	0.046
Type G	0.221	1.0	0.122	0.1	0.155
Type N	0.046	0.122	1.0	0.188	0.303
PreHeater	0.058	0.1	0.188	1.0	0.074
EE	0.046	0.155	0.303	0.074	1.0

References

1. MacRosty, R.; Swartz, C. Dynamics Optimization of Electric Arc Furnace Operation. *Inst. Chem. Eng.* **2007**, *53*, 640–653. [[CrossRef](#)]
2. Ledesma-Carrión, D. Energy Optimization of Steel in Electric Arc Furnace. *Glob. J. Technol. Optim.* **2016**, *7*, 1–10.
3. Gerardi, D.; Marlin, T.; Swartz, C. Optimization of Primary Steelmaking Purchasing and Operation under Raw Material Uncertainty. *Ind. Eng. Chem. Res.* **2013**, *52*, 12383–12398. [[CrossRef](#)]
4. Morales, R.; Rodríguez-Hernández, H.; Conejo, A. A Mathematical Simulator for the EAF Steelmaking Process Using Direct Reduced Iron. *ISIJ Int.* **2005**, *41*, 426–435. [[CrossRef](#)]
5. Nyssen, P.; Colin, R.; Junqué, J.-L.; Knoops, S. Application of a dynamic metallurgical model to the electric arc furnace. *La Revue de Métallurgie* **2004**, *10*, 317–326. [[CrossRef](#)]
6. Çamdali, Ü. Determination of the Optimum Production Parameters by Using Linear Programming in the AC Electric Arc Furnace. *Can. J. Met. Mater. Sci.* **2013**, *44*, 103–110. [[CrossRef](#)]

7. MacRosty, R.; Swartz, C. Dynamic Modeling of an Industrial Electric Arc Furnace. *Ind. Eng. Chem. Res.* **2005**, *44*, 8067–8083. [[CrossRef](#)]
8. Mapelli, C.; Baragiola, S. Evaluation of energy and exergy performances in EAF during melting and refining period. *Ironmak. Steelmak.* **2006**, *33*, 379–388. [[CrossRef](#)]
9. Kirschen, M.; Badr, K.; Pfeifer, H. Influence of Direct Reduced Iron on the Energy Balance of the Electric Arc Furnace in Steel Industry. *Energy* **2011**, *36*, 6146–6155. [[CrossRef](#)]
10. Logar, V.; Dovžan, D.; Škrjanc, I. Modeling and Validation of an Electric Arc Furnace: Part 1, Heat and Mass Transfer. *ISIJ Int.* **2012**, *52*, 402–412. [[CrossRef](#)]
11. Logar, V.; Dovžan, D.; Škrjanc, I. Modeling and Validation of an Electric Arc Furnace: Part 2, Thermo-chemistry. *ISIJ Int.* **2012**, *52*, 413–423. [[CrossRef](#)]
12. Çamdalı, Ü.; Tunç, M. Modelling of Electric Energy Consumption in the AC Electric Arc Furnace. *Int. J. Energy Res.* **2002**, *26*, 935–947. [[CrossRef](#)]
13. Opitz, F.; Treffinger, P. Physics-Based Modeling of Electric Operation, Heat Transfer, and Scrap Melting in an AC Electric Arc Furnace. *Met. Mater. Trans. B* **2016**, *47*, 1489–1503. [[CrossRef](#)]
14. Morales, R.; Conejo, A.; Rodríguez, H. Process Dynamics of Electric Arc Furnace during Direct Reduced Iron Melting. *Met. Mater. Trans. B* **2002**, *33*, 187–199. [[CrossRef](#)]
15. Prakash, S.; Mukherjee, K.; Singh, S.; Mehrotra, S.P. Simulation of energy dynamics of electric furnace steelmaking using DRI. *Ironmak. Steelmak.* **2007**, *34*, 61–70. [[CrossRef](#)]
16. Kho, T.S.; Swinbourne, D.R.; Blanpain, B.; Arnout, S.; Langberg, D. Understanding stainless steelmaking through computational thermodynamics Part 1: Electric Arc Furnace Melting. *Miner. Process. Extr. Met.* **2010**, *119*, 1–8. [[CrossRef](#)]
17. Kirschen, M.; Risonarta, V.; Pfeifer, H.; Energy efficiency and the influence of gas burners to the energy related carbon dioxide emissions of electric arc furnaces in steel industry. *Energy* **2009**, *34*, 1065–1072. [[CrossRef](#)]
18. Trejo, E.; Martell, F.; Micheloud, O.; Teng, L.; Llamasa, A.; Montesinos-Castellanosa, A. A novel estimation of electrical and cooling losses in electric arc furnaces. *Energy* **2012**, *42*, 446–456. [[CrossRef](#)]
19. Fathi, A.; Saboohi, Y.; Škrjanc, I.; Logar, V. Comprehensive Electric Arc Furnace Model for Simulation Purposes and Model-Based Control. *Steel Res. Int.* **2017**, *88*, 1600083. [[CrossRef](#)]
20. Odenthal, H.-J.; Kemminger, A.; Krause, F.; Sankowski, L.; Uebber, N.; Vogl, N. Review of Modeling and Simulation of the Electric Arc Furnace (EAF). *Steel Res. Int.* **2018**, *89*, 1–36. [[CrossRef](#)]
21. Baumert, J.-C.; Engel, R.; Weiler, C. Dynamic modelling of the electric arc furnace process using artificial neural networks. *Revue de Métallurgie* **2002**, *99*, 839–849. [[CrossRef](#)]
22. Baumert, J.-C.; Vigil, J.R.; Nyssen, P.; Schaefer, J.; Schutz, G.; Gillé, S. *Improved Control of Electric arc Furnace Operations by Process Modelling*; European Commission: Luxembourg, 2005; ISBN 92-894-9789-0.
23. Mathy, C.; Terho, K.; Chouvet, M.; Coq, X.L.; Baumert, J.; Engel, R.; Hoffmann, J. *Production of Steel at Lower Operating Costs in EAF*; European Commission: Luxembourg, 2003; ISBN 92-894-6377-5.
24. Chen, C.; Liu, Y.; Kumar, M.; Qin, J. Energy Consumption Modelling Using Deep Learning Technique—A Case Study of EAF. In Proceedings of the 51st CIRP Conference on Manufacturing Systems, Stockholm, Sweden, 16–18 May 2018.
25. Sandberg, E. Energy and Scrap Optimisation of Electric Arc Furnaces by Statistical Analysis of Process Data. Ph.D. Thesis, Luleå University of Technology, Luleå, Sweden, 2005.
26. Köhle, S.; Lichtenbeck, R.; Paura, G. *Verbesserung der energetischen Betriebsführung von Drehstrom-Lichtbogenöfen*; European Commission: Brussels, Belgium, 1996; ISBN 92-827-6467-2.
27. Köhle, S. Effects on the Electric Energy Consumption of Arc Furnace Steelmaking. In Proceedings of the 4th European Electric Steel Congress, Madrid, Spain, 3–6 November 1992.
28. Köhle, S. Variables influencing electric energy and electrode consumption in electric arc furnaces. *Met. Plant Technol. Int.* **1992**, *6*, 48–53.
29. Bowman, B. Performance comparison update—AC vs DC furnaces. *Iron Steel Eng.* **1995**, *72*, 26–29.
30. Kleimt, B.; Köhle, S. Power consumption of electric arc furnaces with post-combustion. *Met. Plant Technol. Int.* **1997**, *3*, 56–57.

31. Köhle, S. Improvements in EAF operating practices over the last decade. In Proceedings of the Electric Furnace Conference, Pittsburgh, PA, USA, 14–16 November 1999.
32. Köhle, S. Recent improvements in modelling energy consumption of electric arc furnaces. In Proceedings of the 7th European Electric Steelmaking Conference, Venice, Italy, 26–29 May 2002.
33. Köhle, S.; Hoffmann, J.; Baumert, J.; Picco, M.; Nyssen, P.; Filippini, E. *Improving the Productivity of Electric Arc Furnaces*; European Commission: Luxembourg, 2003; ISBN 92-894-6136-5.
34. Kleimt, B.; Köhle, S.; Kühn, R.; Zisser, S. Application of models for electrical energy consumption to improve EAF operation and dynamic control. In Proceedings of the 8th European Electric Steelmaking Congress, Birmingham, UK, 9–11 May 2005; pp. 183–197.
35. Kirschen, M.; Zettl, K.-M.; Echterhof, T.; Pfeifer, H. Models for EAF energy efficiency. *Steel Times Int.* **2017**, *44*, 1–6.
36. Conejo, A.; Cárdenas, J. Energy Consumption in the EAF with 100% DRI. In Proceedings of the Iron & Steel Technology Conference, Cleveland, OH, USA, 1–4 May 2006; Volume 1.
37. Czapla, M.; Karbowniczek, M.; Michaliszyn, A. The Optimisation of Electric Energy Consumption in the Electric Arc Furnace. *Arch. Met. Mater.* **2008**, *53*, 559–565.
38. Sandberg, E.; Lennox, B.; Undvall, P. Multivariate Prediction of End Conditions for Electric Arc Furnaces. In Proceedings of the 2nd International Conference on Process Development in Iron and Steelmaking, Luleå, Sweden, 6–9 June 2004.
39. Sandberg, E.; Lennox, B.; Marjanovic, O.; Smith, K. Multivariate process monitoring of EAFs. *Ironmak. Steelmak.* **2005**, *32*, 221–226. [[CrossRef](#)]
40. Sandberg, E.; Lennox, B.; Undvall, P. Scrap management by statistical evaluation of EAF process data. *Control Eng. Pract.* **2007**, *15*, 1063–1075. [[CrossRef](#)]
41. Gajic, D.; Savic-Gajic, I.; Savic, I.; Georgieva, O.; Gennaro, S.D. Modelling of electrical energy consumption in an electric arc furnace using artificial neural networks. *Energy* **2016**, *108*, 132–139. [[CrossRef](#)]
42. Haupt, M.; Vadenbo, C.; Zeltner, C.; Hellweg, S. Influence of Input-Scrap Quality on the Environmental Impact of Secondary Steel Production. *J. Ind. Ecol.* **2016**, *21*, 391–401. [[CrossRef](#)]
43. Carlsson, L.S.; Samuelsson, P.B.; Jönsson, P.G. Predicting the Electrical Energy Consumption of Electric Arc Furnaces Using Statistical Modeling. *Metals* **2019**, *9*, 959. [[CrossRef](#)]
44. Pfeifer, H.; Kirschen, M. Thermodynamic analysis of EAF electrical energy demand. In Proceedings of the 7th European Electric Steelmaking Conference, Venice, Italy, 26–29 May 2002.
45. Steinparzer, T.; Haider, M.; Zauner, F.; Enickl, G.; Naussed, M.M.; Horn, A.C. Electric Arc Furnace Off-Gas Heat Recovery and Experience with a Testing Plant. *Steel Res. Int.* **2014**, *85*, 519–526. [[CrossRef](#)]
46. Keplinger, T.; Haider, M.; Steinparzer, T.; Trunner, P.; Patrejko, A.; Haselgrübler, M. Modeling, Simulation, and Validation with Measurements of a Heat Recovery Hot Gas Cooling Line for Electric Arc Furnaces. *Steel Res. Int.* **2018**, *89*, 1800009. [[CrossRef](#)]
47. Carling, K. Resistant outlier rules and the non-Gaussian case. *Comp. Stat. Data Anal.* **2000**, *33*, 249–258. [[CrossRef](#)]
48. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; The MIT Press: Cambridge, MA, USA, 2016; ISBN 9780262035613.
49. Ajossou, A.; Palm, R. Impact of Data Structure on the Estimators R-square and Adjusted R-square in Linear Regression. *Int. J. Math. Comput.* **2013**, *20*, 84–93.
50. Claesen, M.; De Moor, B. Hyperparameter Search in Machine Learning. *arXiv* **2015**, arXiv:1502.02127.
51. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference for Learning Representations, San Diego, CA, USA, 7–9 May 2015.
52. de Siqueira Santos, S.; Yasumasa Takahashi, D.; Nakata, A.; Fujita, A. A comparative study of statistical methods used to identify dependencies between gene expression signals. *Brief. Bioinform.* **2014**, *1*, 906–918. [[CrossRef](#)]
53. Székely, G.J.; Rizzo, M.L. Brownian distance covariance. *Ann. Appl. Stat.* **2009**, *3*, 1236–1265. [[CrossRef](#)]
54. Dodge, Y. Kolmogorov–Smirnov Test. In *The Concise Encyclopedia of Statistics*; Springer: New York, NY, USA, 2009; pp. 283–287, ISBN 978-0-387-32833-1.
55. Pratt, J.W.; Gibbons, J.D. Kolmogorov–Smirnov Two-Sample Tests. In *Concepts of Nonparametric Theory*; Springer: New York, NY, USA, 1981; pp. 318–344, ISBN 978-1-4612-5931-2.

56. Fisher, A.; Rudin, C.; Dominici, F. All Models are Wrong but Many are Useful: Variable Importance for Black-Box, Proprietary, or Misspecified Prediction Models, using Model Class Reliance. *arXiv* **2018**, arXiv:1801.01489v3.
57. Molnar, C. *Interpretable Machine Learning—A Guide for Making Black Box Models Explainable*. 2019. Available online: <https://christophm.github.io/interpretable-ml-book/> (accessed on 19 October 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).