metals

MDPI

Article

Data-Driven Modelling and Optimization of Energy Consumption in EAF

Simon Tomažič *, Goran Andonovski, Igor Škrjanc and Vito Logar

Faculty of Electrical Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia; goran.andonovski@fe.uni-lj.si (G.A.); igor.skrjanc@fe.uni-lj.si (I.Š.); vito.logar@fe.uni-lj.si (V.L.)
* Correspondence: simon.tomazic@fe.uni-lj.si; Tel.: +386-1-4768-760

Abstract: In the steel industry, the optimization of production processes has become increasingly important in recent years. Large amounts of historical data and various machine learning methods can be used to reduce energy consumption and increase overall time efficiency. Using data from more than two thousand electric arc furnace (EAF) batches produced in SIJ Acroni steelworks, the consumption of electrical energy during melting was analysed. Information on the consumed energy in each step of the electric arc process is essential to increase the efficiency of the EAF. In the paper, four different modelling approaches for predicting electrical energy consumption during EAF operation are presented: linear regression, k-NN modelling, evolving and conventional fuzzy modelling. In the learning phase, from a set of more than ten regressors, only those that have the greatest impact on energy consumption were selected. The obtained models that can accurately predict the energy consumption are used to determine the optimal duration of the transformer profile during melting. The models can predict the optimal energy consumption by selecting pre-processed training data, where the main steps are to find and remove outlier batches with the highest energy consumption and identify the influencing variables that contribute most to the increased energy consumption. It should be emphasised that the electrical energy consumption was too high in most batches only because the melting time was unnecessarily prolonged. Using the proposed models, EAF operators can obtain information on the estimated energy consumption before batch processing depending on the scrap weight in each basket and the added additives, as well as information on the optimal melting time for a given EAF batch. All models were validated and compared using 30% of all data, with the fuzzy model in particular providing accurate prediction results. It is expected that the use of the developed models will lead to a reduction in energy consumption as well as an increase in EAF efficiency.

Keywords: electric arc furnace; energy consumption; profile optimization; modelling; machine learning; steelmaking; regression; fuzzy modelling; evolving modelling

check for updates

Citation: Tomažič, S.; Andonovski, G.; Škrjanc, I.; Logar, V. Data-Driven Modelling and Optimization of Energy Consumption in EAF. *Metals* **2022**, *12*, 816. https://doi.org/10.3390/met12050816

Academic Editors: Thomas Echterhof, Ko-Ichiro Ohno and Ville-Valtteri Visuri

Received: 30 March 2022
Accepted: 6 May 2022
Published: 9 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Current market demands for steel quality, price and production times require the introduction of several technological innovations in electric arc furnace (EAF) steelmaking. Electric Arc Furnaces (EAF) are improving very rapidly. Twenty years ago, the performance of today's EAFs would have seemed impossible. Thanks to an impressive number of innovations, the melting time in the most efficient furnaces (with a capacity of 100–130 t) has been reduced to 30–40 min. Electrical energy consumption was decreased by 1.8 times, from 630 to 340 kWh/t and hourly productivity increased by six times, from 40 to 240 t/h. The share of electrical energy in total energy consumption per melt fell to 50%. Electrode consumption was reduced by about six times [1,2]. It can be assumed that such performance should be normal for most steelworks in the near future.

In modern furnaces, the fundamental processes are melting the solid scrap and heating the liquid bath. The productivity of today's furnaces therefore depends mainly on these

high-energy processes. To set these processes in motion, heat must be obtained from electrical or chemical energy and then passed to the regions of the solid charge or liquid bath [3,4]. The heating technology, furnace designs and other EAF equipment are evolving very fast. Every year, new technical solutions are offered and widely advertised. Steelmakers are struggling to find their way through the flood of innovations. According to the latest trends, modern steelworks should meet four essential requirements in the following way:

- By producing different types of steel in the desired quality, the specified process requirements are met.
- By reducing the manufacturing costs, the specified economic requirements are met, which means that the profitability and competitiveness of the products can be increased.
- By limiting excessive pollution, which is regulated by government regulations, the specified environmental requirements are met.
- By limiting physically and mentally demanding work that is unacceptable for the population of a given country above a certain level of social development, the specified health and safety requirements are met.

The total costs of the EAF can be divided into the cost of scrap and ferroalloys, which account for about 70%, and the so-called operating costs, which account for the remaining 30% of the total cost. The operating costs can be further divided into the costs of electrical energy, fuel and electrodes, which account for about 40% of the operating costs [1,5,6]. The total costs can be reduced in the following ways:

- By reducing the consumption of loaded materials, refractory materials, energy sources, etc. per ton of product;
- By speeding up and increasing production and thus reducing the costs of maintenance, personnel and other specific production costs;
- By finding cheaper input materials and energy sources.

Over the last fifty years, the main objective of EAF development has been to increase productivity. During this period, almost all innovations introduced were dedicated to this problem. Apart from the cost of scrap, productivity represents a crucial factor on which the overall steelmaking economy depends to a large extent [7]. When productivity increases, labour and maintenance costs usually decrease, as do the costs of electrodes, energy sources, refractory materials, electrical energy and other operating costs [8,9]. The proposed EAF innovations, in addition to their positive contributions, also bring some drawbacks. For example, the use of oxygen-gas burners and the introduction of carbon injection for slag foaming enable a drastic reduction in electrical energy consumption, but, on the other hand, increase carbon dioxide emissions [5,6]. Due to environmental protection, the use of biomass (and biofuel produced from renewable biomass) as a renewable energy source in the electric arc furnace is also becoming increasingly important [10,11].

The electrical energy consumption can be controlled by the electrical mode, which is determined by the programme for changing the electrical parameters (current, voltage, arc power, etc.) of the EAF's circuit during the melting process. These parameters can be changed over a wide range due to the special design of the furnace transformer. The control of the transformer voltage levels during the melting process ("on-load") can be done either manually by the operator or fully automatically. The biggest challenge in EAF operation, i.e., determining the optimum melting programmes, times and batch quantities, is thus still left to the operator and his experience. Since the control of the melting process is based on indirect measurements (e.g., arc stability, energy consumption, power-on time, etc.) and not on the actual conditions in the EAF (e.g., bath temperature, melting stage, bath composition), EAF operation is suboptimal (lower raw material and energy efficiency, lower steel quality and increased CO_2 emissions), which consequently means higher operating costs [2,11,12].

With extensive use of oxygen and carbon during the melting process, chemical heat plays a major role in reducing electrical energy consumption and increasing EAF productiv-

ity as the bath absorbs a large amount of chemical heat, which is released during oxidation of carbon, iron and its alloys such as Mn, Si, etc [2,6].

Higher oxygen consumption usually occurs during bath blowing, as it depends on the use of carbon powder, which is added into the bath at the same time as oxygen. The impressive results achieved by the additional oxygen consumption cannot be achieved without the carbon injection. The latter reduces the iron oxides and thus prevents an undesirable reduction in yield. Otherwise, the amount of oxidised iron increases drastically the more oxygen is blown into the bath. In addition, the injected carbon leads to the release of CO and CO_2, which causes the slag to foam. Immersing the arc in foamy slag provides a large increase in efficiency in the use of electrical energy [2,13].

This study addresses the optimization of electric arc furnace (EAF) to increase its efficiency and thus reduce electrical energy consumption. This can be achieved by defining optimal control profiles for the EAF, i.e., transformer power, oxygen balancing, and carbon addition [13,14]. The optimization is based on a data-driven approach where different models (from linear models to evolving fuzzy models) [15–17] and statistical analyses [12] have been performed. The models can be run online in parallel with the actual EAF process and help the operator to control the EAF. Many authors have shown through simulations that optimised operating profiles allow significant reductions in production times and operating costs [2,13,18,19].

Advanced technological solutions such as post-combustion, off-gas [20] and slag [21] heat recovery, oxygen lancing [22], gas burners [23], bottom stirring, adjustable alternating current transmission systems and high-power transformers [24], have reduced energy consumption. Nevertheless, additional process improvements can be made in EAF through heat recovery systems [25], various additives [5,26] and optimal control with operating profiles [2].

The melting profiles are usually selected in advance by the operator based on the maximum energy input. The predefined profiles have the disadvantage that they do not take into account the variations in EAF conditions. Therefore, adaptive control of the EAF (via oxygen and carbon input) is required to achieve suitable conditions and also slag properties. The latter enables to protect the water-cooled panels and walls, reduce energy consumption and contribute to the correct steel composition [2]. Due to the lack of measurements, the operator has limited insight into the EAF process. Consequently, the predefined timed inputs (charging, oxygen lancing and carbon injection) may differ from the optimal times that ensure higher EAF efficiency. Many authors [2,13,18] have conducted studies to investigate EAF efficiency through optimised control. However, very few of them have considered the optimisation of energy sources over the entire tap-to-tap interval. The reason for this could be insufficiently defined optimisation objectives and rough EAF models that are not accurate enough to be used in the optimisation procedure.

The aim of this study is to find key influential factors from which energy consumption in EAF is estimated using the proposed predictive models. These can be used in a simulator to improve the EAF process in such a way that less electrical energy is consumed and the production of a certain type of steel is possible in a shorter time than with the existing process. The total energy (electrical and chemical) consumed in the EAF process is distributed between the three products (steel, slag and off-gas) and the various losses. Only the energy that is delivered to the steel bath can be considered as useful energy.

The paper is organised as follows: Section 2 describes, first, the dataset used and the preprocessing steps applied on it; second, the procedure for selecting the key input variables; and, third, four different modelling approaches (based on machine learning and fuzzy methods) for predicting the electrical energy consumption of EAF. Section 3 discusses the experimental results, comparing all the developed models. A discussion and concluding remarks are given in Sections 4 and 5.

2. Materials and Methods

This section presents the methods needed to build models for predicting electrical energy consumption. These models will be used as part of the operator advisory system to assist the operator in managing the EAF. This prevents the operator from frequently selecting suboptimal settings in the semi-automatic furnace control mode that result in lower steel yield and quality and higher energy and material consumption.

2.1. Data Description and Pre-Processing

The operation of the EAF is monitored by measuring all variables and parameters that could affect energy consumption and overall efficiency. All parameters and variables are stored and organised separately for each batch. Some of the measurements are recorded event-based at specific times, while others are recorded continuously. All important variables from the charging and melting phases are listed in Table 1. The charging recipe is determined by the scrap weight in each basket and the hotheel at the beginning. These data are aggregated for all baskets used. In the melting phase, there are several parameters that affect the total energy consumption and the overall efficiency of the process. The most important criterion and the focus of this article is the electrical energy consumption per total weight of scrap (kWh/t), which is presented in this article as a percentage of the maximum electrical energy consumption per total weight of scrap.

Table 1. List of input variables at the charging and melting phase of the EAF process.

Charging		**Melting**	
Description	**Unit**	**Description**	**Unit**
Total scrap weight	[kg]	Melting time	[s]
Hotheel start	[kg]	Delays	[s]
Scrap weight in basket 1	[kg]	Temperature	[°C]
Scrap weight in basket 2	[kg]	Total oxygen	[Nm^3]
Scrap weight in basket 3	[kg]	Total carbon	[kg]
Type of charged scrap		Hotheel end	[kg]
		Slag weight	[kg]

In the development of the electrical energy consumption prediction models, the first required step is the preprocessing and filtering of data (removal of a part of the data). Since the data are stored in different databases and with different sampling times during the operation of the EAF, it is necessary to resample and synchronise the data before starting the analysis phase. Since historical data from completed batches are often incomplete, these batches must be removed from the modelling process during filtering. The data cleaning procedure to eliminate all corrupted data should also be applied to efficiently identify and remove outliers (e.g., unusually long tap-to-tap time spans, i.e., more than four hours, extremely high power consumption, etc.). The steps of data pre-processing cannot be performed completely automatically, since in some special cases the knowledge and experience of the staff (especially the EAF operator) must also be taken into the account. Each batch may consist of two, three, or four baskets of raw material. Since in the available database the melting process was most frequently performed with three baskets, only these batches were used in all further analyses. After filtering the data, the first, the second, and the third baskets have an average capacities of 46 t, 36 t, and 18 t, respectively. Each individual charge takes about three minutes. The melting of the scrap after charging with the first, second, and third baskets takes about 17 min, 11 min, and 20 min on average, and the average delay per batch is 13 min.

2.2. Selection of the Key Input Variables

The operation of the EAF is a subject to several factors that affect the final product quality and energy consumption. Deviations from potentially optimal performance can

be influenced by all parameters and settings during the charging and melting phases. Therefore, the most influential independent variables must be identified from historical data, as this information is necessary for the development of the models for energy consumption prediction. In the study of Glavan et al. [27], it has been shown that the input variable selection (IVS) approach can efficiently find the most important input variables from a big database for modelling and prediction purposes. The IVS approach is based on the analysis of historical data and combines a data mining approach with various selection criteria [28]. The selection of input variables has a great impact on the prediction performance, the effectiveness of the model and the better understanding of the system. Therefore, the IVS represents an important step for model identification. The authors in [27] tested and compared different methods from the literature for variable selection. They also evaluated each method to find out the most suitable methods for model-based prediction problems. Finally, the authors selected the following methods as the most effective: partial correlation measure (Pcorr) [29], partial mutual information (PMI), linear-in-the-parameters (LIP) [30], non-negative Garrote (NNGarr) [31], variable importance in projection (PLS VIP) [32], distance correlation (dCorr) [33], and least absolute shrinkage and selection operator (LASSO) [34]. All of these methods, briefly discussed in [27], were used in this study. The influential factors from all methods were averaged to determine the order of the most influential variables. In the following, all the machine learning methods that were used to obtain the predictive models for estimation of the electrical energy consumption are briefly described.

2.3. Machine Learning Methods

2.3.1. Linear Regression

The linear model has been a mainstay of statistics for the past 30 years and remains one of the most important tools [35,36]. Linear regression is a linear approach for modelling the relationship between a scalar response and one or more explanatory variables (also known as dependent and independent variables). In linear regression, the relationships are modelled using linear predictor functions, whose unknown model parameters are estimated from the data. Such models are called linear models. Using linear regression, the single output of the model $\hat{y}_j$ can be determined in the following way:

$$\hat{y}_j = \hat{\beta}_0 + \sum_{i=1}^{p} x_i \hat{\beta}_i, \tag{1}$$

where $x_j^T = [x_1, x_2, \ldots, x_p]_j$ is a regression vector (where $j = 1, \ldots, m$; m is the total number of test samples and p is the total number of independent variables) and $\hat{\beta}^T = [\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p]$ is the vector of linear coefficients. The term $\hat{\beta}_0$ is the intercept, which in machine learning is also called the bias [35]. It is often convenient to include the constant variable 1 in the vector x_j and include $\hat{\beta}_0$ in the vector of coefficients $\hat{\beta}$, and then write the linear model in vector form as an inner product:

$$\hat{y}_j = x_j^T \hat{\beta}. \tag{2}$$

There are many different methods of fitting the linear model to a set of training data. By far the most popular is the least squares method. In this approach, the coefficients β are chosen to minimise the residual sum of squares:

$$RSS(\beta) = \sum_{i=1}^{n} (y_i - x_i^T \beta)^2, \tag{3}$$

where n is the total number of training samples. $RSS(\beta)$ is a quadratic function of the parameters, and hence its minimum always exists but may not be unique. The solution is most easily characterised in matrix notation:

$$RSS(\beta) = (y - X\beta)^T (y - X\beta), \tag{4}$$

where X is an $n \times p$ matrix with each row an input vector x_i^T, and y is an n-vector of the outputs in the training set. Differentiating w.r.t. β the normal equations can be written as follows:

$$X^T(y - X\beta) = 0. \tag{5}$$

If $X^T X$ is non-singular, then the unique solution is given by:

$$\hat{\beta} = (X^T X)^{-1} X^T y, \tag{6}$$

and the fitted value at the i-th input x_i is $\hat{y}_i = x_i^T \hat{\beta}$.

2.3.2. K-Nearest Neighbour Method

K-nearest neighbours (k-NN) algorithms [37,38] are non-parametric supervised machine learning algorithms commonly used in the field of pattern recognition. The k-NN algorithms can be used for both classification and regression. In both cases, the input to the algorithm consists of the labelled training dataset:

$$D = \{(x_i, y_i)\} \text{ for } i = 1, \ldots, n, \tag{7}$$

where n is the number of samples in the dataset, x_i is the regression vector and y_i is the class label or a continuous output variable. To make a prediction (class label or continuous target variable), the k-NN algorithms find the k nearest neighbours of a query point $\hat{x}_j$ and compute the class label (i.e., classification) or continuous target variable (i.e., regression) based on the k nearest (most "similar") points. Since the prediction is based on a comparison of a query point with data points (regression vectors) in the training dataset, k-NN is also categorised as an instance-based (or "memory-based") method.

In k-NN regression, the output prediction is based on the labels of the k nearest neighbours. The output value $\hat{y}$ is usually the average of the values of k nearest neighbours:

$$\hat{y}_j = \frac{1}{k} \sum_{i=1}^{k} y_i. \tag{8}$$

For both classification and regression, a distance-weighted k-NN algorithm [38] can also be used, which assigns weights to the contributions of the neighbours, so that the closer neighbours contribute more to the average than the more distant ones. For example, a common weighting scheme is to assign a weight of $w_i = 1/d_i$ to each neighbour, where d_i is the distance to the i-th nearest neighbour.

$$\hat{y}_j = \frac{\sum_{i=1}^{k} w_i y_i}{\sum_{i=1}^{k} w_i}. \tag{9}$$

The best choice of k depends upon the data. In general, larger values of k reduce the effects of noise on classification but make the boundaries between classes less clear. A good k can be selected by various heuristic techniques. By changing the value of k, the complexity of a k-NN model is affected. In practise, a good trade-off must be found between high bias (the model is not complex enough to fit the data well when k is too large) and high variance (the model fits the training data too closely when k is too small).

For k-NN algorithms, many distance metrics or measures can be used to select k nearest neighbours. There is no "best" distance measure, and the choice is highly context- or problem-dependent. For continuous features, the most common distance metric is the Euclidean distance. Another popular choice is the Manhattan distance, which puts less emphasis on the differences between "distant" feature vectors or outliers than the Euclidean distance. The Mahalanobis distance would be another good choice for a distance metric, as it takes into account the variance of the different feature vectors as well as the covariance among them.

One of the main advantages of k-NN is that it is relatively easy to implement and interpret. Moreover, with its approach to approximate complex global functions locally, it can be a powerful predictive "model". Another advantage is that k-NN has some strong consistency results. As the amount of data approaches infinity, the two-class k-NN algorithm is guaranteed to yield an error rate no worse than twice the Bayes error rate (the minimum achievable error rate given the distribution of the data). The drawback is that k-NN is very sensitive to the curse of dimensionality [38] and is expensive to compute with an $O(n)$ prediction step. Therefore, various data structures have been developed to improve the computational performance of k-NN in prediction. In particular, the idea is to identify the k nearest neighbours more intelligently. Instead of matching each training sample in the training set to a given query point vector, various approaches have been developed to partition the search space as efficiently as possible and reduce the number of distance evaluations actually performed. Data structures such as KD-trees and Ball-trees are often used for this purpose, as they can make k-NN substantially more efficient.

2.4. Takagi–Sugeno Fuzzy Modeling

Fuzzy logic was developed in 1965 as an extension of the classical (Boolean) logic. The classical logic assigns to a variable or a statement the value of 1 for "true" or the value of 0 for "false", fuzzy logic allows the value assignment at an interval between $[0, 1]$. The reason for this can be found in the observation of the way of human thinking when deciding on the very approximate estimates of various facts that they present to themselves in the form of rules. To address such a concept, a mechanism for recording knowledge based on rules in the form of approximate reasoning based on fuzzy logic has been introduced. First, some basic concepts of fuzzy logic and approximative reasoning, which are necessary for the understanding of fuzzy models, are introduced. Fuzzy logic records relationships, knowledge and decisions in the form of rules. For conjunction of the linguistic statements, the conjunction operator (t-norm) "min" is used. The combination of the affiliation of all linguistic expressions determines degree of rule fulfillment or rule firing strength because it expresses how well the premise matches the given values of input variables. For the entire fuzzy system, only fulfillment degrees greater than zero are important. It must be guaranteed that the rules complete the entire possible input space to avoid situations where no rule gets activated for certain input values. In the case of non-explicit local affiliation functions, this problem does not exist because all rules are always fulfilled, although with very small values. After the degree of fulfillment of an individual rule is calculated, the contributions of individual consequent parts have to be determined and assembled to obtain the output of the fuzzy system. This is called accumulation. Usually, the output of the fuzzy system is a fuzzy set that needs to be transformed into a sharp form for further work. This is called the process of defuzzification. Of course, this is not necessary if a sharp value is chosen for the consequent part, or if the result is used for qualitative estimations. In general, three basic types of fuzzy systems exist, i.e., a linguistic or Mamdani, a special fuzzy system or a singleton, and a Takagi–Sugeno Fuzzy System. In our case, the focus is on the Takagi–Sugeno (TS) Fuzzy approach, which provides excellent interpretability and the best fuzzy modelling results [39]. The first step of fuzzy modelling is the fuzzification, where the degree of membership for all linguistic statements $\mu_{ij}(x_j)$ ($i = 1, \ldots, M$ and $j = 1, \ldots, p$) is calculated, where M is the number of fuzzy system rules and p is the number of inputs x_j. The TS fuzzy rule $\mathcal{R}_i$ can be written as follows:

$$\mathcal{R}_i: \text{ IF } x_1 = \mathsf{A}_{i1} \text{ AND } \ldots \text{ AND } x_p = \mathsf{A}_{ip} \text{ THEN } y = f_i(x_1, x_2, \ldots, x_p) \tag{10}$$

where A_{ij} represents a fuzzy set for the variable x_j and y is the output. By aggregation, the individual linguistic statements into the level of activation of the rule (with respect to the operators between them) are composed. The output of the TS fuzzy model is defined as:

$$\hat{y} = \frac{\sum_{i=1}^{M} f_i(x)\,\mu_i(x)}{\sum_{i=1}^{M} \mu_i(x)} \tag{11}$$

where x is the input vector and $f_i(x) = w_{i0} + w_{i1}x_1 + w_{i2}x_2 + \ldots + w_{ip}x_p$ is linear regression function. If the fuzzy model is written in a conjunctive form and min function is used for the t-norm, then the degree of fulfillment of the rule is:

$$\mu_i(x) = \min(\mu_{ij}(x_j))\,. \tag{12}$$

In this study, Gaussian membership functions were used:

$$\mu_{ij}(x_j) = \exp\left(-\frac{1}{2}\frac{(x_j - c_{ij})^2}{\sigma_{ij}^2}\right), \tag{13}$$

where σ_{ij}^2 is variance and c_{ij} is expected value of the Gaussian function (belonging to A_{ij} fuzzy set).

In fuzzy models, the nonlinear parameters in the premise (i.e., the parameters in the causal part of the rule that define the membership functions, their positions and widths) and the linear parameters (w_{ip}) in the consequent part of the rules can be optimized. The latter can be easily estimated using the least squares method. The parameters in the causal part of the rule correspond to the parameters on a hidden layer of neural networks and are nonlinear. The optimization of the rule structure is a combination problem that can be solved by a selection of linear subsets or by a nonlinear global optimization, for example, by an optimization with a genetic algorithm (GA) or a particle swarm optimization (PSO) [40].

2.5. Evolving the Cloud-Based Prediction Model

Due to the refinement of the technological process of melting in the EAF, the data collected from the new batches are increasingly different and the consequently developed models are predicting electrical energy consumption worse and worse. The evolving modelling approach is appropriate for the purpose of constantly updating models also during the process of melting. In this paper, an online evolving fuzzy identification method (based on data clouds) [41], which represents an upgrade according to the Takagi–Sugeno fuzzy modelling, is used. By upgrade means the ability to evolve the structure of the model online and to adapt the parameters of each local model during the process.

In evolving modelling, the structure of the fuzzy model is identified online using the evolving mechanisms, i.e., principles for adding and removing fuzzy rules. The rule-based form of i-th rule is defined as:

$$\mathcal{R}^i : \text{IF} \quad (x_f(k) \sim X^i) \quad \text{THEN} \quad y_i(k) = f_i(x_f(k)) \tag{14}$$

where $x_f(k) = \left[u_1(k), u_2(k) \ldots, u_p(k)\right]$ represents the input (regression) vector, X^i stands for the i-th data-cloud, $y_i(k)$ represents the output of that fuzzy rule, and $f_i(x_f(k))$ represents an arbitrary function. In our case, the NARX model is used and therefore the output function is defined as:

$$f_i(k) = \theta_i^T \psi(k) \tag{15}$$

where $\psi(k) = \left[x_f(k),\, 1\right]^T$ stands for the extended regressor and θ_i^T is vector of local parameters of i-th fuzzy rule, which are calculated using the recursive Weighted Least Squares method (rWLS) as presented in [41]. The final value of the output is calculated as follows:

$$y(k) = \sum_{i=1}^{c} \beta_i \theta_i^T \psi_k \tag{16}$$

where c is the number of data-clouds (fuzzy rules), and β_i stands for normalized relative density, which is defined as relation between the current data sample $x_f(k)$ and the i-th fuzzy rule X^i. Normalized relative density is calculated as follows:

$$\beta_i(k) = \frac{\gamma_i(k)}{\sum_{j=1}^{c} \gamma_j(k)}, \qquad i = 1, \ldots, c \tag{17}$$

where $\gamma_i(k)$ stands for the local density of the data $x_f(k)$ and is calculated as:

$$\gamma_i(k) = \frac{1}{1 + \|x_f(k) - \mu_i(k)\|^2 + \sigma_i(k) - \|\mu_i(k)\|^2} \tag{18}$$

In Equation (18), $\mu_i(k)$ and $\sigma_i(k)$ denote mean value vector and mean-square length of the data vector from i-th cloud, respectively. Please refer to [41] for more details about the whole evolving algorithm including the evolving mechanisms of adding and removing (data-clouds) fuzzy rules.

3. Results

The data used within the methods for key input variables selection and for the validation of the developed models were collected from the actual EAF in the SIJ Acroni company. From the collected database, 577 different batches were selected with the filtering. In the stage of predictive models development, the whole dataset was divided into training (404 batches) and testing (173 batches) subsets (70% of the data for training and 30% for testing). For each batch, 13 input variables were recorded, which are listed in Table 1. For each batch, the loading recipe (marked from 1 to 12) and melting program (marked from 1 to 15) are also selected according to the required properties of the steel produced.

3.1. Results of the Selection of Key Input Variables

Using the methods presented in Section 2.2 (Pcorr, PMI, LIP, NNGarr, PLS VIP, and LASSO), the most influential variables for predictive models were found. Since the results vary widely from one method to another, average influential factors were calculated to be more generally usable regardless of modelling method. Figure 1 shows the sorted results of finding the most influential variables considering all data in the database. The boxes in the figure show the average values, the median values and the intervals within the 25th and 75th percentiles.

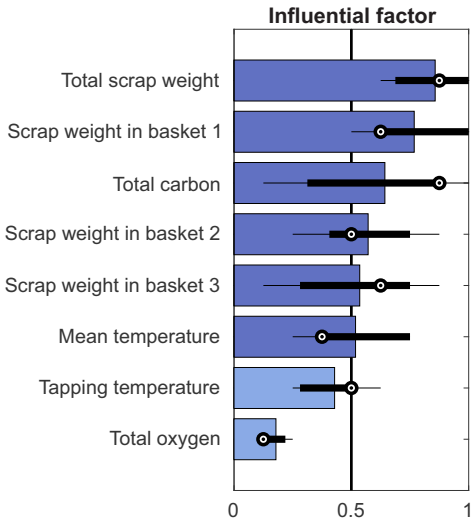


Figure 1. The average influential factors for all independent input variables.

Table 2 shows the average values for each influential factor. According to the obtained results, it is reasonable to include the following variables in further consideration: total scrap weight, scrap weight in individual baskets, total carbon, average temperature during melting, tapping temperature and total oxygen. The significance of the individual input variables can also be partially inferred from Figure 2, where linear models describe the relationships between the various input variables and the total electrical energy consumption (as a percentage of the maximum value (kWh/t)). In determining the most influential variables, the dispersion or data distribution plays a major role. Figure 2 shows one of the most influential variables and one of the least influential variables in each case. The simultaneous use of multiple independent variables to predict electrical energy consumption can change the influential factor of a single variable (due to the interconnectedness of the variables). Therefore, it is difficult to conclude from Figure 2 why total scrap weight is more important than the total carbon variable.

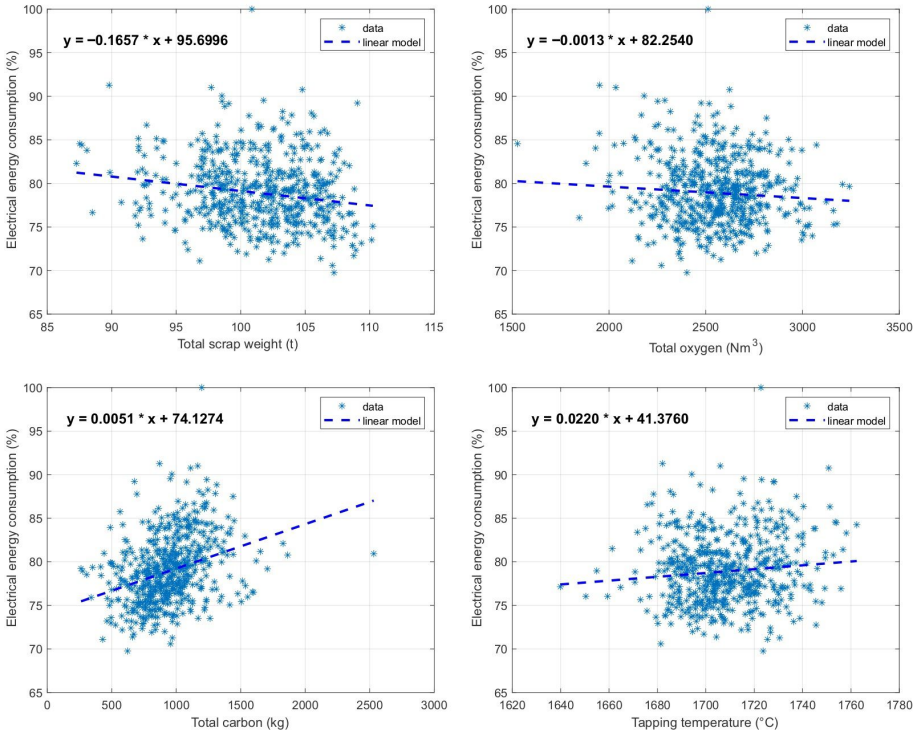


Figure 2. Linear models of electrical energy consumption (as a percentage of the maximum value) as a function of total scrap weight (**top left**), total oxygen (**top right**), total carbon (**bottom left**), and tapping temperature (**bottom right**), respectively.

Table 2. The average influential factors for all independent input variables.

Variable	Influential Factor
Total scrap weight	0.8571
Scrap weight in basket 1	0.7679
Total carbon	0.6429
Scrap weight in basket 2	0.5714
Scrap weight in basket 3	0.5357
Mean temperature	0.5179
Tapping temperature	0.4286
Total oxygen	0.1786

When modelling electrical energy consumption, reducing the dimensionality of the input space is also very important; otherwise, the (fuzzy) model structure may become too complex and the large number of model parameters may be difficult to determine.

If the modelling method also includes an optimization phase of the model parameters, the modelling process can become very slow and inefficient. On the other hand, considering only a limited number of the most influential variables can lead to worse prediction results as some of the information is lost. Therefore, different variations of combined input variables were also considered. Using the methods presented in Section 2.2, the following combined input variables (Figure 3) were selected as the most influential: the quotient of tapping temperature and total scrap weight, the quotient of mean temperature and scrap weight in the first two baskets, chemical energy (calculated from total carbon and total oxygen as proposed in [42]), the quotient of total oxygen and total carbon, scrap weight in the third basket. The average influential factors for all combined input variables are listed in Table 3. Figure 4 shows that the use of only one combined input variable does not drastically improve the prediction of electrical energy consumption, but as mentioned earlier, the main advantage of selecting the most influential variables is shown only when all input variables are used together in the exact combination.

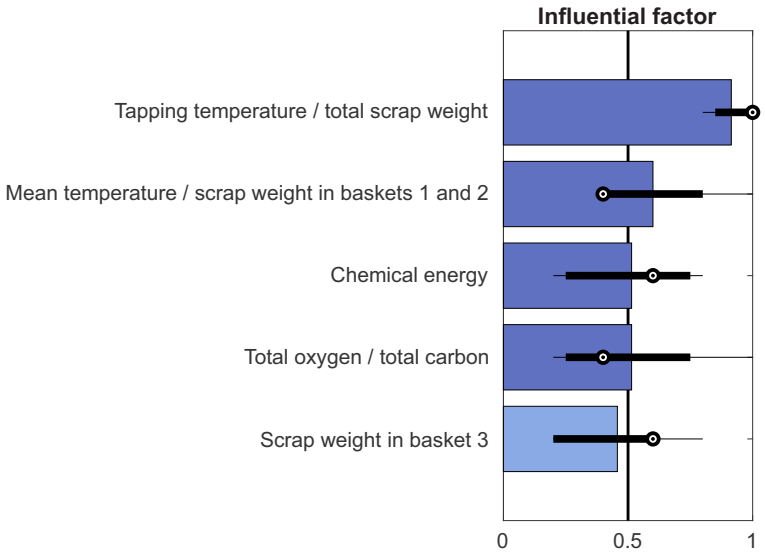


Figure 3. The average influential factors for the five most influential combined input variables.

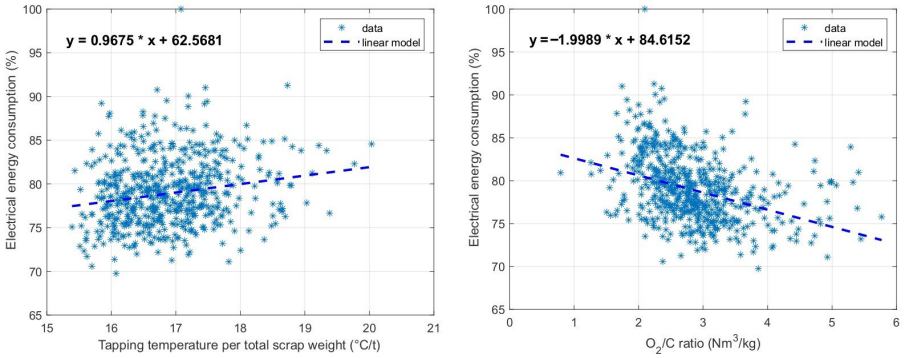


Figure 4. Linear models of electrical energy consumption as a function of the quotient of tapping temperature and total scrap weight (**left**) and the quotient of total oxygen and total carbon (**right**).

Table 3. The average influential factors for the five most influential combined input variables.

Variable	Influential Factor
Tapping temperature/total scrap weight	0.9143
Mean temperature/scrap weight in baskets 1 and 2	0.6000
Chemical energy	0.5143
Total oxygen/total carbon	0.5143
Scrap weight in baskets 3	0.4571

3.2. Analysis of Models for Energy Consumption Prediction

This subsection presents the comparative results of predicting electrical energy consumption with the static models explained in Section 2. Each model is used to predict the total electrical energy consumption of the current batch as a function of the key input variables listed in Table 3. All models are compared using the root–mean–square error (RMSE), which is a measure of the differences between the values (electrical energy consumption in percentages) predicted by a model $\hat{y}_i$ and the observed values y_i:

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2}, \tag{19}$$

where m is the number of all test batches.

Figure 5 shows the results of predicting electrical energy consumption with the k-NN model (left) and the linear regression model (right) compared to the electrical energy consumption measurements. In the figure, the line shows the ideal (completely accurate) prediction of electrical energy consumption according to the test samples. The k-NN model was constructed to consider Mahalanobis distance and the six nearest neighbours. The output of the k-NN model is calculated according to Equation (9), which means that the nearer neighbour has more influence on the output than the farther neighbour. Compared to the prediction results of the k-NN model, the linear regression model achieves slightly better results (see Table 4) in terms of R^2 (coefficient of determination) and RMSE, although this model is simpler.

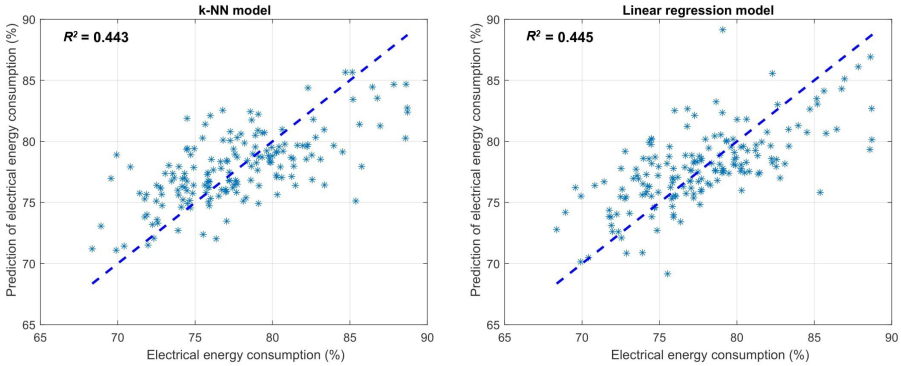


Figure 5. Prediction of electrical energy consumption with the k-NN model (**left**) and the linear regression model (**right**) compared to measurements of electrical energy consumption.

Artificial intelligence algorithms, i.e., evolving and fuzzy modelling approaches proposed in this work, achieve better prediction results than machine learning methods (k-NN and linear regression), as expected. Figure 6 shows the results of predicting electrical energy consumption with the evolving model (left) and the fuzzy model (right) compared to the electrical energy consumption measurements. When looking at Figures 5 and 6, it is difficult to decide which model is the best because the differences are quite small. Therefore, all RMSE and R^2 results for each method are presented in Table 4. From the table, it can be concluded that the best results were obtained with the conventional fuzzy method and the evolving method proposed in this paper. In the conventional fuzzy modelling, the PSO optimization method was used to determine the optimal structure (number, distribution and width of Gaussian membership functions) of the fuzzy logic system that gives the best prediction results.

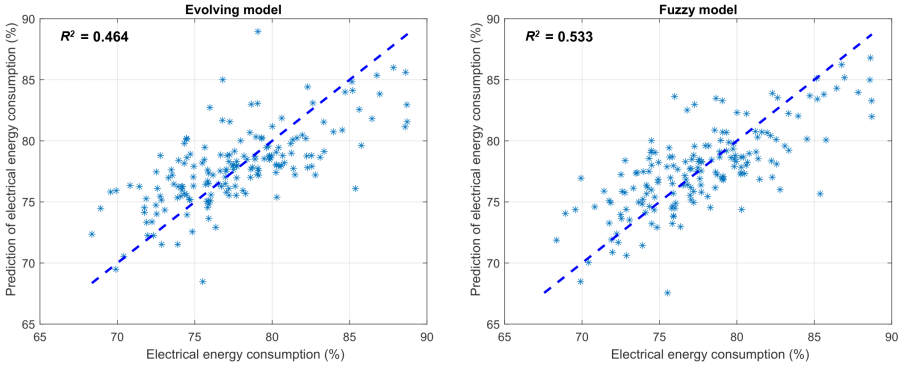


Figure 6. Prediction of electrical energy consumption with the evolving model (**left**) and fuzzy model (**right**) compared to measurements of electrical energy consumption.

Table 4. Comparison of RMSE results for the prediction models.

Method	RMSE (%)	R^2
k-NN method	3.177	0.443
Linear regression	3.171	0.445
Evolving model	3.118	0.464
Fuzzy model	2.910	0.533

All the developed models can also be compared with the calculation of the cumulative distribution functions, which are shown in Figure 7. From this graph, for example, it is easy to see that 90% of all errors are less than 5% (of the maximum electrical energy consumption) when the k-NN model is used. Thus, a steeper curve represents a better model.

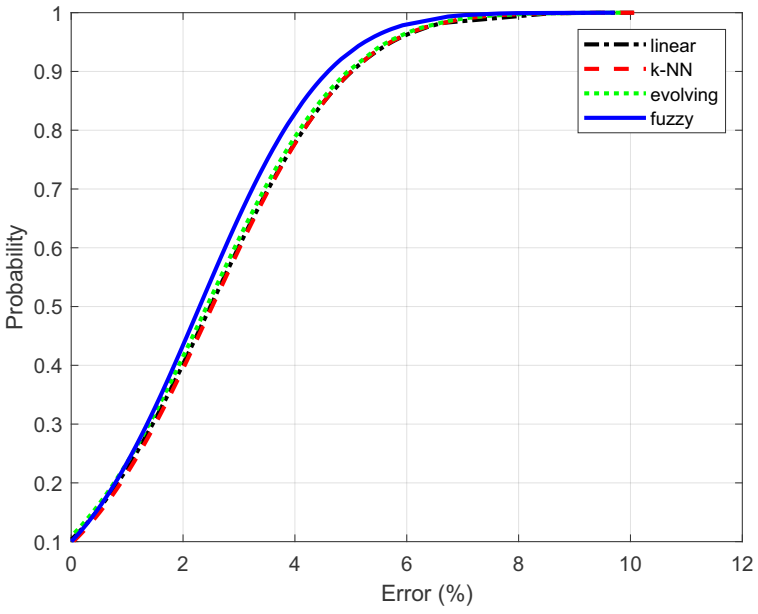


Figure 7. Cumulative distribution functions for the electrical energy prediction errors.

The comparison between the results considering all input variables and only the most influential variables (selected variables) shows that reducing the independent variables can improve the fitting results by at least 20% according to the RMSE of linear regression. The effect of reducing the input space is even more evident when evolving or a fuzzy modelling approach is used, since in these cases model complexity translates into more challenging optimization conditions due to the large number of input variables. The larger number of optimization parameters slows down the training process and may lead to

suboptimal results. Fuzzy membership functions may not be optimally defined and distributed, and consequently the model may be over-fitted to the training dataset. However, over-fitted models fail quickly when applied to new batches that differ slightly from those in the past.

The prediction results of all developed models can be drastically improved (by at least 20% according to the RMSE of linear regression) if melting time is also used as an independent input variable. Although up to 80% of authors of all published papers dealing with the EAF energy consumption prediction have used melting time as an input variable to achieve better results, this approach is completely incorrect as the melting time is not known in advance. If melting time was known in advance, advanced models would actually be unnecessary because the melting time is almost entirely proportional to electrical energy consumption (see Figure 8). In Figure 8, two different linear models are shown according to the maximum transformer tap level (in the profile), which is either seven or eight for all melting programmes. The linear models show that the melting programmes with the maximum transformer tap level eight have a slightly higher energy consumption than the melting programmes with the maximum transformer tap level seven, but the slope is almost the same in both cases. The obtained models, shown in Figure 8, are used to predict the melting time from the electrical energy consumption prediction (obtained with the fuzzy model). This information is essential for the EAF operator as he can try different scenarios in the simulator and determine the optimal time to complete the batch. This is one of the possible ways to partially reduce the electrical energy consumption without intervening in the EAF itself because until now, in most cases, electrical energy consumption was only too high due to an unnecessary prolongation of the melting time. This is because it is difficult for the operator to determine exactly when the material is completely melted.

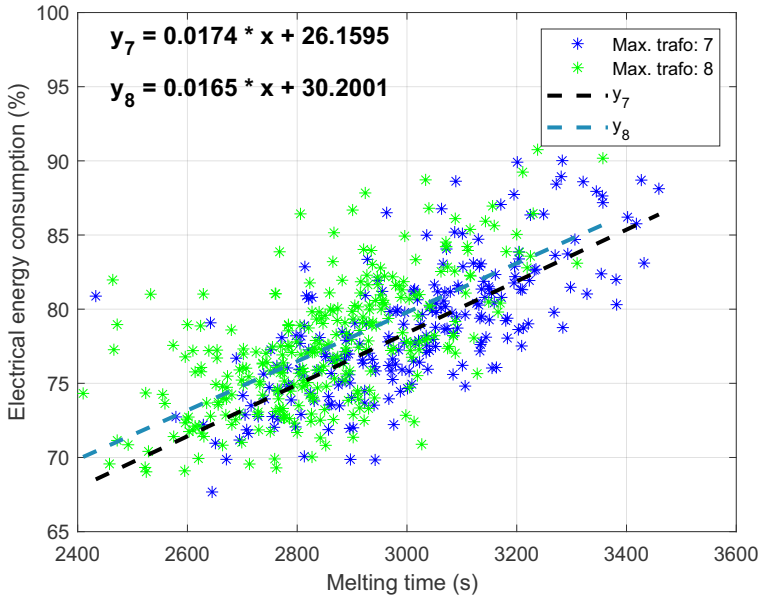


Figure 8. Electrical energy consumption as a function of melting time for two different maximum transformer tap levels in the profile.

4. Discussion

Technological processes in the steel industry have improved greatly in recent decades. Further optimization of the processes is possible by introducing digital tools that advise operators on setting parameters and help control production (also in terms of equipment maintenance). In this study, the focus is on the optimization of electrical energy consumption through the analysis of existing historical data and the construction of prediction models. The latter allows the operator to perform preliminary simulations through an advisory tool that determines the electrical energy consumption according to the selected conditions.

The operator can thus test the optimal values for the materials added, the amount of carbon and oxygen added, the melting temperature and, above all, the final melting time. From all of the influential variables, total melting time is the one that total energy consumption depends the most on it but should not be considered as an input variable, which is a common mistake. Although the transformer profiles (in historical data) that define the EAF electrical parameters (current, voltage, arc power) have two different final values for the power levels in the existing melting programs (7 and 8), these values have an almost negligible impact on the final consumption compared to the final melting time (which is also defined with the transformer profiles). The developed models can predict the electrical energy consumption quite accurately since the error is less than 5% (of the maximum energy consumption) for 90% of all errors. Converting the electrical energy consumption to the final melting time is also very straightforward since consumption and melting time are proportional to each other. The choice of input variables is critical to developing applicable models, especially when a large number of variables are available. Without algorithms to analyze influential factors, the types of charged materials would certainly be chosen as input variables, as well as the amount of slag or delays during the process. As the results show, the total mass has the greatest influence on the prediction of energy consumption, although the consumption is normalized with respect to the total mass (in kWh/t). When determining the key variables, their simultaneous consideration is crucial because the influential factors are distributed differently than when only one input variable is considered at a time. For successful model construction, it is also critical to eliminate bad measurements (outlier filtering) that occur in batches with many interruptions and extended melting time due to faults at the EAF. Poor measurements are also possible due to incorrectly recorded charged materials (quantities and types), but not all such anomalies in the measurements can be detected. All the developed models are comparable with each other in terms of the prediction error (RMSE) and the coefficient of determination R^2, which means that, if the combined variables are chosen appropriately, the linear methods also work effectively. Each of the methods has its advantages and disadvantages. For example, the evolving method, although it does not give the best results, may be best suited for online updating of models during the process itself, which may improve the prediction for the current batch. The conventional fuzzy method is computationally the most demanding because it involves a PSO optimization, but it provides the best prediction results. On the other hand, linear regression is the simplest since it does not require parameter adjustments, while the k-NN method is the fastest since it does not require a training phase.

5. Conclusions

This paper presents the results of a study in which preprocessed historical data from the real EAF process were used to identify the influential variables that have the greatest impact on electrical energy consumption during melting. The results show that the root mean square error in predicting electrical energy consumption can be reduced by at least 20% with proper selection of the influential variables. Four different prediction models were constructed from the filtered data, using linear regression, k-NN, evolving, and fuzzy modelling methods. When comparing the errors in the prediction of electrical energy consumption, the fuzzy model was found to be the most accurate, as the root mean square error has the lowest value and the coefficient of determination has the highest value. The developed models will be used within the advisory tool, which will help the EAF operator to adjust the parameters correctly during the melting process and, in this way, improve the efficiency of the EAF.

Author Contributions: Conceptualization, S.T., I.Š. and V.L.; Data curation, V.L.; Formal analysis, S.T. and G.A.; Methodology, S.T., G.A., I.Š. and V.L.; Project administration, I.Š.; Software, S.T., G.A. and V.L.; Visualization, S.T. and V.L.; Writing—original draft, S.T., G.A., I.Š. and V.L.; Writing—review & editing, V.L. All authors have read and agreed to the published version of the manuscript.

Funding: The work presented in this paper is funded by the European Union's Horizon 2020 research and innovation programme, the SPIRE initiative, under Grant No. 869815, the INEVITABLE project ("Optimization and performance improving in metal industry by digital technologies") and by the Slovenian Research Agency Programme: Modelling, simulation and control of processes (P2-0219).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Toulouevski, Y.N.; Zinurov, I.Y. Modern Steelmaking in Electric Arc Furnaces: History and Development. In *Innovation in Electric Arc Furnaces: Scientific Basis for Selection*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 1–24. [CrossRef]
2. Saboohi, Y.; Fathi, A.; Škrjanc, I.; Logar, V. Optimization of the Electric Arc Furnace Process. *IEEE Trans. Ind. Electron.* **2019**, *66*, 8030–8039. [CrossRef]
3. 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]
4. Kovačič, M.; Stopar, K.; Vertnik, R.; Šarler, B. Comprehensive Electric Arc Furnace Electric Energy Consumption Modeling: A Pilot Study. *Energies* **2019**, *12*, 2142. [CrossRef]
5. Sung, Y.; Lee, S.; Han, K.; Koo, J.; Lee, S.; Jang, D.; Oh, C.; Jang, B. Improvement of Energy Efficiency and Productivity in an Electric Arc Furnace through the Modification of Side-Wall Injector Systems. *Processes* **2020**, *8*, 1202. [CrossRef]
6. Echterhof, T. Review on the Use of Alternative Carbon Sources in EAF Steelmaking. *Metals* **2021**, *11*, 222. [CrossRef]
7. Ahmed, W.; Moazzam, M.; Sarkar, B.; Ur Rehman, S. Synergic Effect of Reworking for Imperfect Quality Items with the Integration of Multi-Period Delay-in-Payment and Partial Backordering in Global Supply Chains. *Engineering* **2021**, *7*, 260–271. [CrossRef]
8. Mahapatra, A.S.; N Soni, H.; Mahapatra, M.S.; Sarkar, B.; Majumder, S. A Continuous Review Production-Inventory System with a Variable Preparation Time in a Fuzzy Random Environment. *Mathematics* **2021**, *9*, 747. [CrossRef]
9. Bhuniya, S.; Pareek, S.; Sarkar, B. A supply chain model with service level constraints and strategies under uncertainty. *Alex. Eng. J.* **2021**, *60*, 6035–6052. [CrossRef]
10. Sarkar, B.; Mridha, B.; Pareek, S. A sustainable smart multi-type biofuel manufacturing with the optimum energy utilization under flexible production. *J. Clean. Prod.* **2022**, *332*, 129869. [CrossRef]
11. Yadav, D.; Kumari, R.; Kumar, N.; Sarkar, B. Reduction of waste and carbon emission through the selection of items with cross-price elasticity of demand to form a sustainable supply chain with preservation technology. *J. Clean. Prod.* **2021**, *297*, 126298. [CrossRef]
12. Carlsson, L.S.; Samuelsson, P.B.; Jönsson, P.G. Using Statistical Modeling to Predict the Electrical Energy Consumption of an Electric Arc Furnace Producing Stainless Steel. *Metals* **2020**, *10*, 36. [CrossRef]
13. Logar, V.; Fathi, A.; Škrjanc, I. A Computational Model for Heat Transfer Coefficient Estimation in Electric Arc Furnace. *Steel Res. Int.* **2016**, *87*, 330–338. [CrossRef]
14. Meier, T.; Logar, V.; Echterhof, T.; Škrjanc, I.; Pfeifer, H. Modelling and Simulation of the Melting Process in Electric Arc Furnaces—Influence of Numerical Solution Methods. *Steel Res. Int.* **2016**, *87*, 581–588. [CrossRef]
15. Núñez, A.; De Schutter, B.; Sáez, D.; Škrjanc, I. Hybrid-fuzzy modeling and identification. *Appl. Soft Comput.* **2014**, *17*, 67–78. [CrossRef]
16. Dovžan, D.; Logar, V.; Škrjanc, I. Implementation of an Evolving Fuzzy Model (eFuMo) in a Monitoring System for a Waste-Water Treatment Process. *IEEE Trans. Fuzzy Syst.* **2015**, *23*, 1761–1776. [CrossRef]
17. Škrjanc, I.; Iglesias, J.A.; Sanchis, A.; Leite, D.; Lughofer, E.; Gomide, F. Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: A Survey. *Inf. Sci.* **2019**, *490*, 344–368. [CrossRef]
18. 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]
19. Hay, T.; Visuri, V.V.; Aula, M.; Echterhof, T. A Review of Mathematical Process Models for the Electric Arc Furnace Process. *Steel Res. Int.* **2021**, *92*, 2000395. [CrossRef]
20. Lee, B.; Sohn, I. Review of Innovative Energy Savings Technology for the Electric Arc Furnace. *JOM* **2014**, *66*, 1581–1594. [CrossRef]
21. Barati, M.; Esfahani, S.; Utigard, T. Energy recovery from high temperature slags. *Energy* **2011**, *36*, 5440–5449. [CrossRef]
22. Lee, B.; Ryu, J.W.; Sohn, I. Effect of Hot Metal Utilization on the Steelmaking Process Parameters in the Electric Arc Furnace. *Steel Res. Int.* **2015**, *86*, 302–309. [CrossRef]
23. 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]
24. Bisio, G.; Rubatto, G.; Martini, R. Heat transfer, energy saving and pollution control in UHP electric-arc furnaces. *Energy* **2000**, *25*, 1047–1066. [CrossRef]

25. Meier, T.; Hay, T.; Echterhof, T.; Pfeifer, H.; Rekersdrees, T.; Schlinge, L.; Elsabagh, S.; Schliephake, H. Process Modeling and Simulation of Biochar Usage in an Electric Arc Furnace as a Substitute for Fossil Coal. *Steel Res. Int.* **2017**, *88*, 1600458. [CrossRef]
26. Gandt, K.; Meier, T.; Echterhof, T.; Pfeifer, H. Heat recovery from EAF off-gas for steam generation: Analytical exergy study of a sample EAF batch. *Ironmak. Steelmak.* **2016**, *43*, 581–587. [CrossRef]
27. Glavan, M.; Gradišar, D.; Atanasijević-Kunc, M.; Strmčnik, S.; Mušič, G. Input variable selection for model-based production control and optimisation. *Int. J. Adv. Manuf. Technol.* **2013**, *68*, 2743–2759. [CrossRef]
28. Van De Wal, M.; De Jager, B. Review of methods for input/output selection. *Automatica* **2001**, *37*, 487–510. [CrossRef]
29. May, R.; Dandy, G.; Maier, H. Review of Input Variable Selection Methods for Artificial Neural Networks. In *Artificial Neural Networks-Methodological Advances and Biomedical Applications*; BoD—Books on Demand: Norderstedt, Germany, 2011.
30. Li, K.; Peng, J.X. Neural input selection-A fast model-based approach. *Neurocomputing* **2007**, *70*, 762–769. [CrossRef]
31. Breiman, L. Better subset regression using the nonnegative garrote. *Technometrics* **1995**, *37*, 373–384. [CrossRef]
32. Chong, I.G.; Jun, C.H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **2005**, *78*, 103–112. [CrossRef]
33. Székely, G.J.; Rizzo, M.L.; Bakirov, N.K. Measuring and testing dependence by correlation of distances. *Ann. Stat.* **2007**, *35*, 2769–2794. [CrossRef]
34. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc.* **1996**, *58*, 267–288. [CrossRef]
35. Freedman, D. *Statistical Models: Theory and Practice*; Cambridge University Press: Cambridge, UK, 2009.
36. Montgomery, D.C.; Peck, E.A.; Vining, G.G. *Introduction to Linear Regression Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
37. Friedman, J.H.; Bentley, J.L.; Finkel, R.A. An Algorithm for Finding Best Matches in Logarithmic Expected Time. *ACM Trans. Math. Softw.* **1977**, *3*, 209–226. [CrossRef]
38. Chen, G.H.; Shah, D. Explaining the Success of Nearest Neighbor Methods in Prediction. *Found. Trends Mach. Learn.* **2018**, *10*, 1–250. [CrossRef]
39. Takagi, T.; Sugeno, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybern.* **1985**, *SMC-15*, 116–132. [CrossRef]
40. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948. [CrossRef]
41. Andonovski, G.; Mušič, G.; Blažič, S.; Škrjanc, I. On-line Evolving Cloud-based Model Identification for Production Control. *IFAC-PapersOnLine* **2016**, *49*, 79–84. [CrossRef]
42. Blažič, A.; Škrjanc, I.; Logar, V. Soft sensor of bath temperature in an electric arc furnace based on a data-driven Takagi–Sugeno fuzzy model. *Appl. Soft Comput.* **2021**, *113*, 107949. [CrossRef]