Prediction of Productivity and Energy Consumption in a Consteel Furnace Using Data-Science Models

Panagiotis Sismanis(✉)

Sidenor Steel Industry SA, 15125 Athens, Greece
psismanis@sidenor.vionet.gr

Abstract. The potential to predict the productivity and the specific electric-energy furnace consumption is very important for the economic operation and performance of a Consteel electric-arc furnace. In this work, these two variables were predicted based on specific operating parameters with the use of machine learning. Actually, three different algorithms were tested for this study: the BRF method of support vector machine (SVM), the light gradient boosting method (lightGBM), and the Keras system with TensorFlow as backend. The results appear to be good enough for production scheduling, and are presented and discussed in this work.

Keywords: Gradient boosting method · Support vector machine · Keras · TensorFlow · Productivity · Energy · Consteel

1 Introduction

The need for the potential of predicting the productivity, and specific electrical-energy consumption based on process parameters has been always the desire in steelmaking plants worldwide. It facilitates procurement and results in a more rational approach in the schedule of actions. Budget development is then performed in a faster and more reliable basis. In a monumental work, Koehle [1] had come up with a very compact formula based on regression analysis from data collected from a relatively large number of electric-arc furnaces (EAF) worldwide. This equation had taken under consideration the specific electric-energy reduction from installations with scrap preheating, like Consteel. On the other hand, in a more recent work, Memoli [2] elaborated data from Consteel installations and pointed out the effect of liquid heel upon productivity in addition to the installed and operating electric-arc power. However, each Consteel installation has its own peculiarities depending upon scrap quality, furnace condition, and operating personnel skills. In this study, process parameters that are recorded daily per heat were selected and an attempt to predict the expected productivity and specific electric-furnace energy consumption was carried out. Machine learning was applied in order to derive supervised models that could be used in off-line scheduling from the scrap yard till the secondary metallurgy treatment of liquid steel.

© Springer Nature Switzerland AG 2019
W. Abramowicz and R. Corchuelo (Eds.): BIS 2019, LNBIP 353, pp. 85–99, 2019.
https://doi.org/10.1007/978-3-030-20485-3_7

2 Preparation for the Computations

2.1 EAF Installation

The SOVEL plant, which is part of the VIOHALCO/SIDENOR group of companies, is located at a seacoast area of Almyros next to the city of Volos in the middle of Greece. In 2006, SOVEL decided to convert to a Consteel operation in order to increase production capacity and decrease electrical energy consumption. The furnace is a 3-phase EBT EAF with a 120 MVA transformer, and 600-mm-diameter electrodes.

Table 1. Considered operating parameters (independent variables).

No.	Symbol	Description
1	*YIELD*	Yield (t of good billet per t of scrap)
2	*POW_ON*	Power on time (min)
3	*TAP_TAP*	Tap-to-tap time (min)
4	*POWER_AVG*	Average power per heat (MW)
5	*SPEC_OX*	Specific oxygen consumption (Nm^3/t)
6	*SPEC_LIME*	Specific lime consumption (kg/t)
7	*SPEC_MGO*	Specific MgO consumption (kg/t)
8	*SPEC_CHRG_CARBON*	Specific charged carbon (kg/t)
9	*SPEC_INJ_CARBON*	Specific injected carbon (kg/t)
10	*TEMP*	Tapping temperature (°C)
11	*ppmO*	Oxygen content (ppm)
12	*PC_HMS_1*	Scrap type HMS #1 (percentage)
13	*PC_HMS_2*	Scrap type HMS #2 (percentage)
14	*PC_SHREDDED*	Shredded scrap (percentage)
15	*PC_BUSHELLING*	Busheling scrap (percentage)
16	*PC_RETURNS*	Returns (percentage)
17	*PC_TURNINGS*	Turnings (percentage)
18	*PC_PIG_IRON*	Pig iron (percentage)

It has the necessary modules to supply chemical energy as well as to inject carbon units in order to retain the appropriate volume of foaming slag. The meltshop has a ladle furnace (LF) for the secondary metallurgy treatment of liquid steel, and a 6-strand continuous caster (CCM) that mainly casts 140 × 140 mm × mm billets. The furnace-tapped weight is 130 t and the annual capacity is around 1 million tons of steel. The interested reader may refer to [3] for a more detailed description of the furnace installation.

2.2 Parameters Considered in the Computations

Two dependent parameters were analyzed: productivity in t/h (*TON_PER_HR*), and specific electrical-energy consumption in kWh/t good billet (*SPEC_ENRGY_BILLET*).

However, since these two parameters are very critical with respect to the economics policy of a company, only scaled data/results were presented in this study. Nevertheless, the degree of correlation achieved in the analysis is mostly important to the data scientist and not the actual values.

Table 1 presents the 18 independent operating parameters considered as the important factors that were supposed to have an influence upon the two dependent variables. The main parameters related to the electrical energy consumption are the average power of the heat (*POWER_AVG*), and the time duration in which the energy was supplied (*POW_ON*). The tap-to-tap time (*TAP_TAP*) influences productivity and electrical energy consumption; it reflects the proper condition of a furnace not only with respect to operations but with respect to maintenance, as well. The specific oxygen consumption (*SPEC_OX*) plays an important role not only to the selected dependent variables but also to the chemistry of the tapped liquid steel. The specific addition of lime (*SPEC_LIME*) and magnesia (*SPEC_MGO*) play a paramount role in the proper chemical composition of the slag for foaming. In practice, we aim for $FeO \approx 20\%$, $MgO \approx 6\%$, and $B_3 \approx 2.0$; the index B_3 is given by the following equation [4]:

$$B_3 = \frac{CaO}{SiO_2 + Al_2O_3} \tag{1}$$

We wanted to keep the derived supervised results as close as possible to the operating parameters that we controlled more easily in practice. For that reason, we decided to employ only the specific additions of the slag fluxes than including the slag chemical analyses into the computations. The liquid steel temperature (*TEMP*) and oxygen (*ppmO*) as measured by the CELOX probe [5] have an effect on energy consumption and liquid steel cleanliness; the yield (*YIELD*) plays a great role on specific consumptions and cost-effective meltshop operation. Finally, seven scrap parameters related to scrap mix were taken under consideration (Table 1, 12–18). One may realize that scrap mix is the top factor that greatly influences the two selected dependent parameters, so it was impossible to proceed in this type of study without considering it. Nevertheless, a meltshop facility may not have the desired scrap mix all year around. On this basis, proper management may help eliminate extreme cases and still make production at desired costs.

2.3 Computational Approach

Pieces of software were developed in order to tackle the process of data, and tune the selected machine-learning algorithms. Python (version 3.5.4, 64-bit, Anaconda installation [6]) was the deployed language. The total number of cases (heats) that were selected for the analysis were 9148 in total, and they belonged in the production period from January 2016 until July 2018, that is about 1.2 million tons of produced liquid steel. In fact, the initial number of heats for that period was more but almost one-third of heats were filtered out in order to compensate for ambiguous or missing data. The library of pandas [7] was used for data manipulation, and the scikit-learn library [8] was used for validation and tuning of the models. Furthermore, the kernel RBF (radial basis function) from the support vector machine (SVM) library of the scikit-learn package was deployed as one of the three selected algorithms for machine learning. SVM was selected, as it is a powerful and widely used machine-learning algorithm. It generally constructs a hyperplane or set of hyperplanes in a multi-dimensional space which can be used for regression by selecting the hyperplane that has the largest distance to the nearest training-data point of any class (margin), since in general the larger the margin, the lower the generalization error of the derived supervised model. The second selected algorithm was lightGBM [9] (gradient boosting method) that appears to be reliable, simple, and flexible. LightGBM is a gradient boosting framework that uses histogram-based algorithms, which bucket continuous feature (attribute) values into secret bins. This speeds up training and reduces memory usage. Finally, the last algorithm was Keras [10, 11] with TensorFlow [12] as backend. Keras uses neural networks for the derivation of the supervised models. Keras is an open-source neural-network library written in Python. Designed to enable fast experimentation with deep neural networks, it focuses on being user-friendly, modular, and extensible. It supports standard, convolutional, and recurrent neural networks. An attempt to apply machine learning in a Python environment was the triggering mechanism to work with the three aforementioned machine-learning algorithms after having published another study in an environment under R [13]. The software was run in a DELL Alienware laptop with the Intel i7-6700HQ CPU (8 cores) @2.6 GHz, 16 GB RAM, running under a 64-bit Windows 10 Professional O/S. The data were scaled before further processing. Finally, two important functions from the scikit-learn package were deployed: the StandardScaler function was used in order to scale the data so that errors coming from computations amongst too big and too low values to be minimized; furthermore, the GridSearchCV function was called in order to fine-tune the machine learning algorithms.

3 Results and Discussion

3.1 Principal Component Analysis

At first, a search was held in order to verify whether the independent variables exhibited some patterns. For this purpose, the KMeans method from scikit-learn was applied after scaling in order to identify the optimum number of clusters in which the independent variables might be included; the average-silhouette-width [14] criterion was put into effect and proved that the optimum number of clusters was two. Furthermore, applying a principal component analysis (PCA) after filtering the scaled data through the GradientBoostingRegressor (all scikit-learn [8] functions) for the two optimum clusters, Fig. 1 was created showing the results.

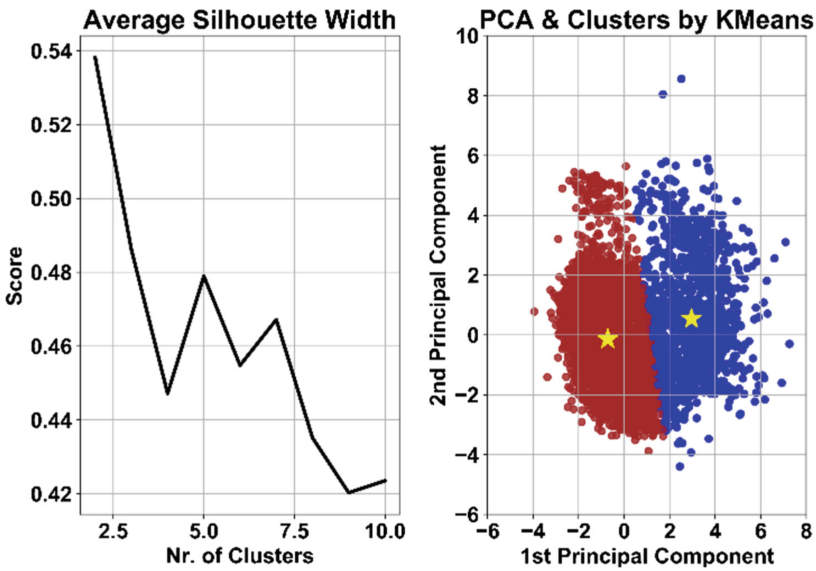

Fig. 1. Clusters and PCA analysis.

The two stars in Fig. 1 illustrate the centroids for the two clusters. It seems that the two clusters are related to the heats that were produced at larger average-power values (left-hand-side centroid), and smaller average-power values (right-hand-side centroid).

3.2 Productivity Analysis-SVM

Figure 2 depicts actual versus predicted scaled-productivity values. One may realize that there is a relatively good prediction upon productivity. In fact, the analysis of variance (ANOVA) between actual and predicted values showed that the root-mean-squared-error (RMSE) was 0.150, with a correlation coefficient (R2_squared) equal to 0.9778. However, in order to deduce these relatively good results with the RBF kernel of the support vector machine we had to tune the model. Two tuning parameters were required [15, 16] in order to improve the accuracy of the model: C, and gamma; actually, after tuning using a grid search (GridSearchCV), the optimum values were C = 100, and gamma = 0.01.

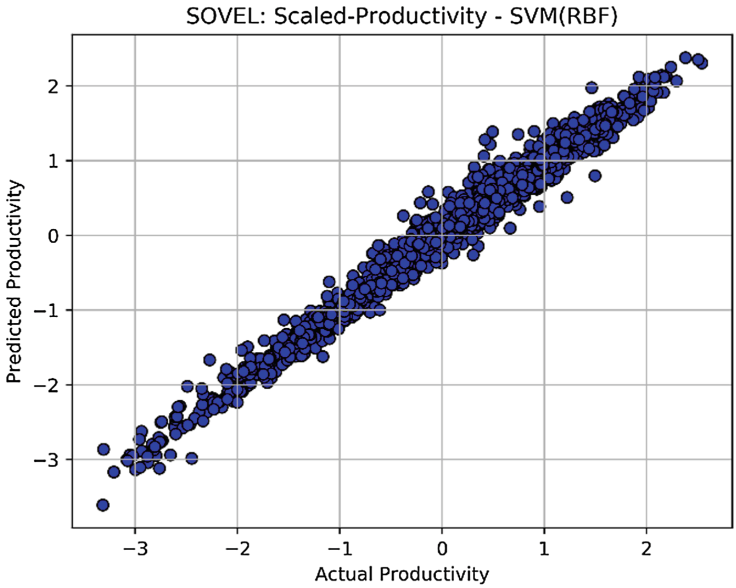

Fig. 2. Actual vs. predicted scaled-productivity values using the RBF kernel of the support vector machine.

3.3 Productivity Analysis-LightGBM

Figure 3 illustrates actual versus predicted scaled-productivity values by the lightGBM model. Here, the ANOVA showed that the RMSE value was 0.098, with R2_squared equal to 0.9905. Furthermore, a grid search was held in order to tune the appropriate parameters. In this case, the optimum parameters were learning_rate = 0.01; n_estimators = 10000.

3.4 Productivity Analysis-Keras

Keras was found more difficult to tune and the attained optimum parameters gave adequate values for the RMSE = 0.156, and R2_squared = 0.9760. Figure 4 illustrates the attained training and validation accuracy versus epochs. One may notice that the error minimizes flattening at epochs equal to 700. Actually, the optimum parameters were found to be batch_size = 150, and epochs = 700. Figure 5 depicts the actual versus predicted productivity values by Keras. Two main sequential models were applied the first with 64 units, and the second 18 units; the activation selected was of 'relu' type that is a good choice for regression models, with a 2% dropout percentage, and an 'l2' regularizer kernel equal to 0.1%. The 'rmsprop' type was selected as the routine for the mean-squared-error optimizer.

As a concluding remark, it seems that productivity can be predicted to a reasonable extent by the operating parameters from all three selected models.

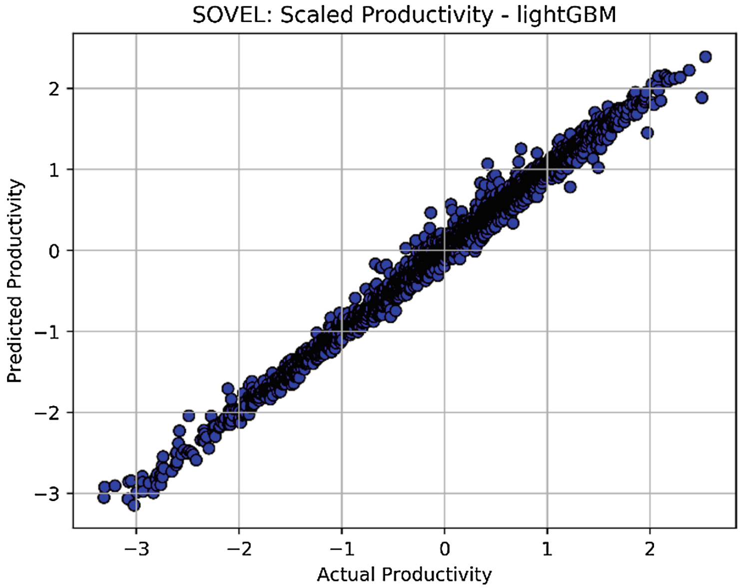

Fig. 3. Actual vs. predicted scaled-productivity values using lightGBM.

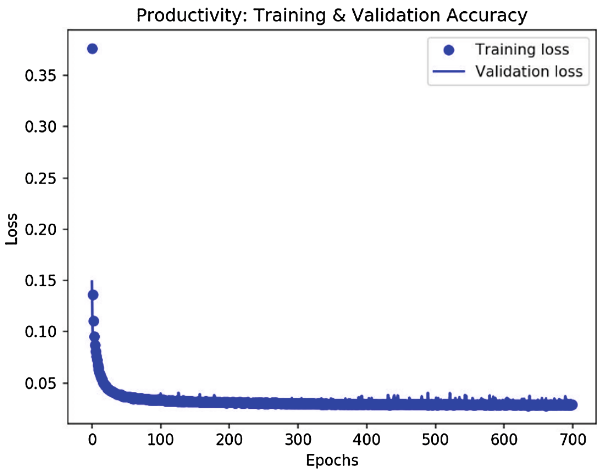

Fig. 4. Training and validation accuracy vs. epochs during grid-search tuning of the Keras model for productivity.

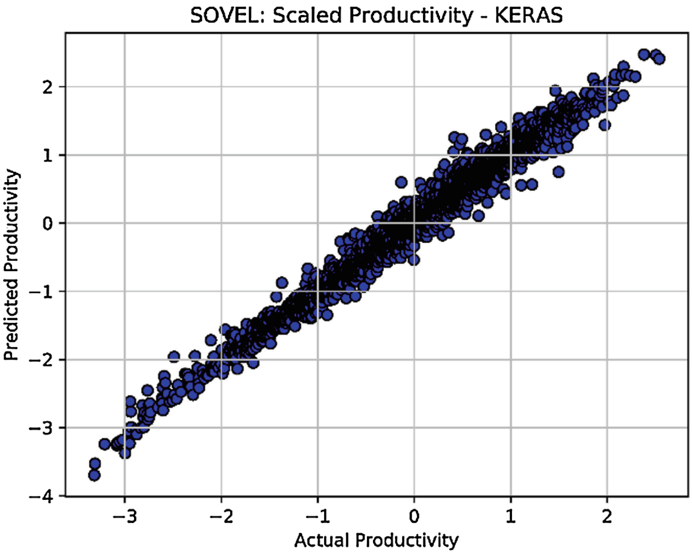

Fig. 5. Actual vs. predicted scaled-productivity values using Keras.

3.5 Energy Analysis-SVM

The scaled specific electrical-energy consumption (*SPEC_ENRGY_BILLET*) was predicted by the RBF kernel of the support vector machine; a grid search was applied in order to quantify the tuning parameters (C, gamma) and optimize the mean-squared-error. Figure 6 depicts the actual and predicted values for the scaled specific-energy consumption. The ANOVA on the actual and predicted values computed a RMSE value equal to 0.240, and a R2-squared value equal to 0.9421; these values were calculated for the optimum values of C = 1000, and gamma = 0.01.

3.6 Energy Analysis-LightGBM

Figure 7 illustrates the actual and predicted values for the specific electrical energy consumption. A grid search was also performed calculating as optimum values a learning_rate = 0.010, and n_estimators = 10000.

For Fig. 7 under the optimum parameters the ANOVA resulted in a RMSE value equal to 0.1858 with a correlation coefficient R2_squared = 0.9654. Nevertheless, what appears to be interesting is the prediction of the relative importance of the independent parameters upon the two dependent parameters under study. Figure 8 shows the variable importance of the selected independent parameters upon the specific energy consumption; it is worth saying that the same relative importance of the independent parameters appears true for productivity, as well. One may notice that the independent parameters related to power and time play a paramount importance upon energy consumption, as well as productivity. Indeed, *POWER_AVG*, *POW_ON*, and *TAP_TAP* seem to be very important, as expected from real practice experience. Yield (*YIELD*) cannot be underestimated and energy-inducing factors like oxygen (*SPEC_OX*) and carbon (*SPEC_CHRG_CARBON, SPEC_INJ_CARBON*) are important as well.

We should notify that we operate the furnace at low natural-gas consumptions; actually, we use natural gas only for the proper operation of modules and not as burners. Lime is an important slag constituent (*SPEC_LIME*), and tapping temperature (*TEMP*) and oxygen content (*ppmO*) could not be excluded from the estimating picture. The influence of the scrap mix is more-or-less expected according to the experience from our practice. However, the low effect from pig iron (*PC_PIG_IRON*) comes from the fact that we do not add more than 3% in most cases, for the heats that we do add it.

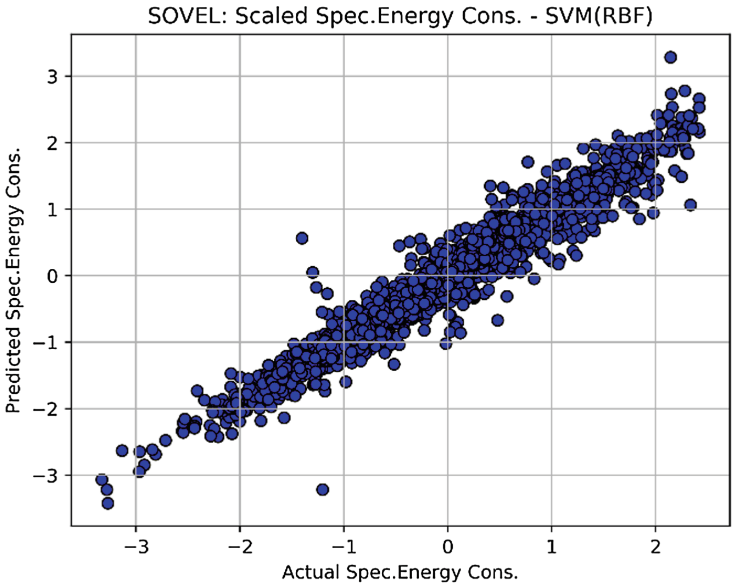

Fig. 6. Actual vs. predicted scaled specific-energy-consumption values using the RBF kernel of the support vector machine.

3.7 Energy Analysis-Keras

Figure 9 illustrates the training and validation accuracy in the tuning process.

Applying a grid search to tune the Keras model it was found that the batch_size = 100, at epochs = 600 gave rise to the best possible optimization of the mean-squared-error. In fact, the ANOVA gave the value of RMSE = 0.238, at a multiple correlation coefficient R2_squared = 0.9432. Neural networks seem to be hard to tune in order to attain the best desirable values. Again in this case, two main sequential units were applied with the first being 64 dense units, and 18 the second; the optimizer selected was the type 'rmsprop' at a 'relu' type of activation. A 2% dropout was applied, too.

Figure 10 illustrates the actual and predicted specific energy-consumption values from this analysis with Keras.

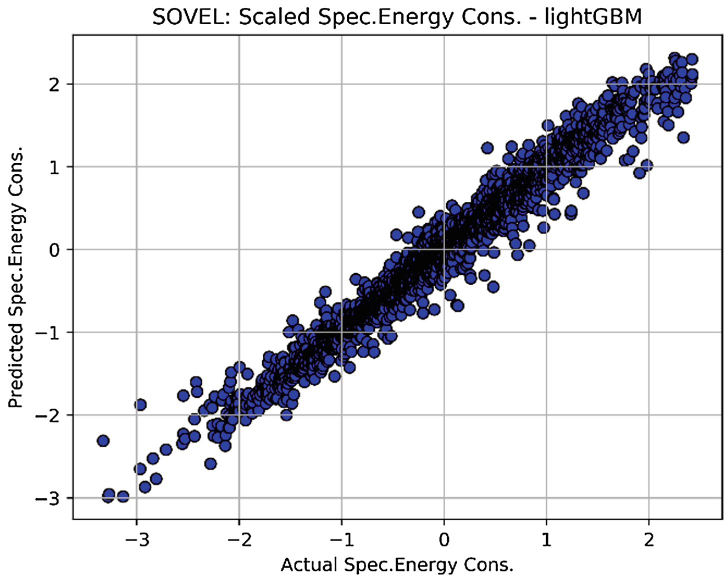

Fig. 7. Actual vs. predicted scaled specific-energy-consumption values using lightGBM.

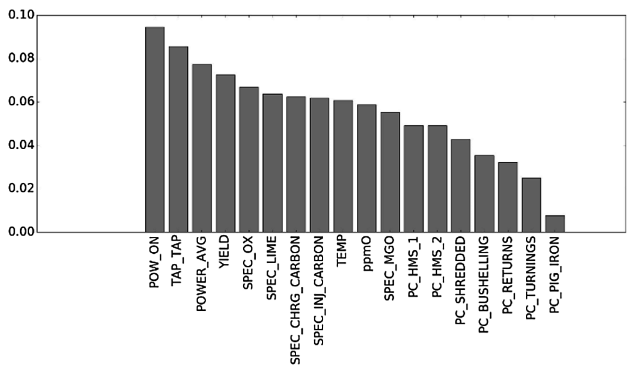

Fig. 8. Relative importance of the independent parameters upon the specific energy consumption (the same holds true for productivity).

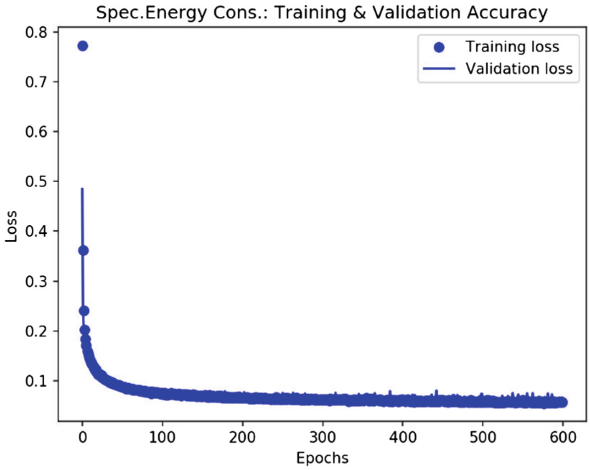

Fig. 9. Training and validation accuracy vs. epochs during grid-search tuning of the Keras model for specific energy consumption.

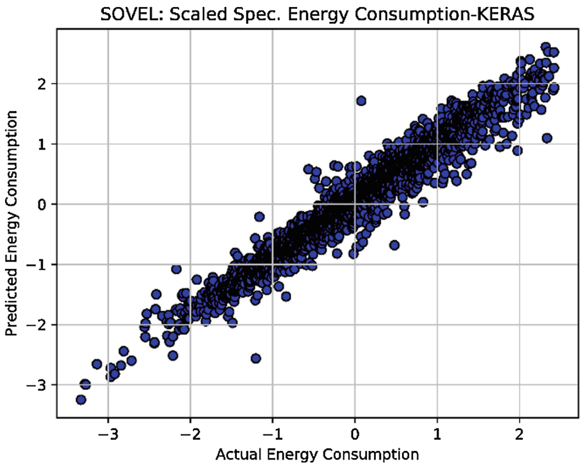

Fig. 10. Actual vs. predicted scaled specific-energy-consumption values using Keras.

3.8 Energy Analysis-Verification by Practice

After the summer-2018 maintenance period, the scrap yard collected enough quantity of relatively good scrap with minimal gangue content. Consequently, we experienced a good September-2018 month with high productivities and small specific electrical-energy consumptions. For this reason, testing the supervised model predictions for productivity and energy consumption based on the past set of data for the independent variables would be a good verification about the reliability of these models.

For illustration purposes, results from lightGBM model predictions were selected for presentation. Figure 11 shows actual and predicted scaled-productivity values for the “good” period Aug-Sep’18. The RMSE (standard deviation) was 0.303 with a multiple correlation coefficient R2_squared equal to 0.9108.

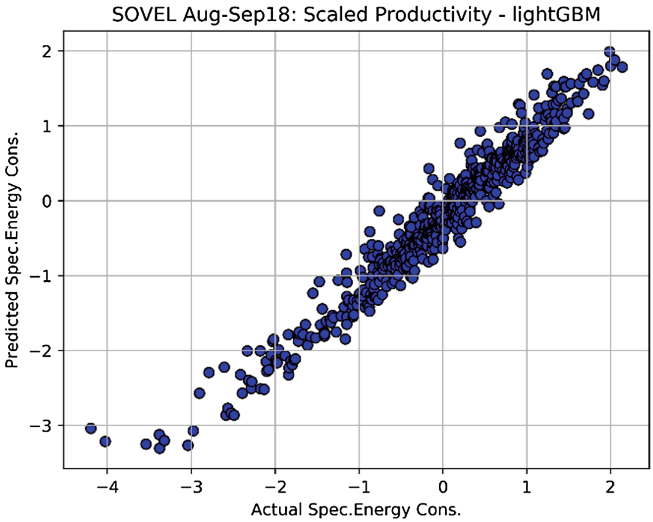

Fig. 11. Actual vs. predicted scaled-productivity values for the period Aug-Sep’18 using lightGBM.

Similarly, Fig. 12 depicts actual and predicted scaled specific electrical-energy consumption values for the same period. The standard deviation was in this case 0.699 with a multiple correlation coefficient equal to 0.4463. Since the model had been tuned with a very small percentage of heats with very low scaled-specific electrical-energy consumption values, one may conclude that the poor correlation results were to be expected. However, the salient features of the high productivity and low energy consumption were remarkably achieved albeit to a less reliable base. Nevertheless, including the Aug-Sep’18 data in the tuning database of the models it was expected that new and better-supervised models could be computed that will behave much better in similar cases in the future. Since statistical analysis has exhibited much better results for the prediction of productivity for the limited data of the Aug-Sep’18 period, emphasis was given to energy analysis for that period by encompassing these data into the initial database and deriving a new lightGBM supervised model in order to check any potential improvement.

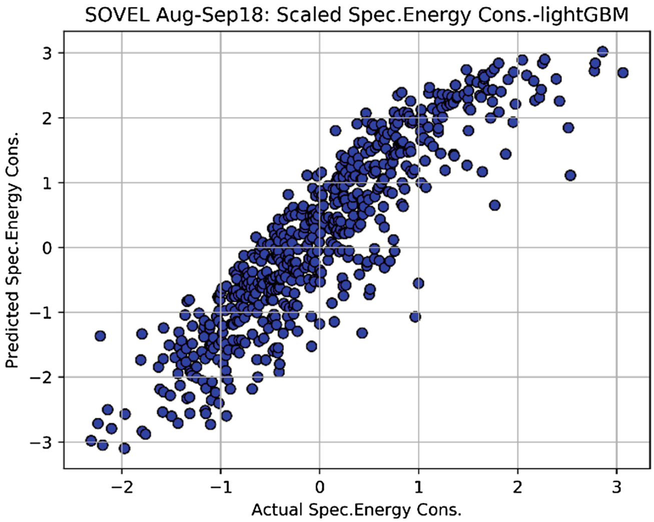

Fig. 12. Actual vs. predicted scaled specific-energy-consumption values for the period Aug-Sep’18 using lightGBM

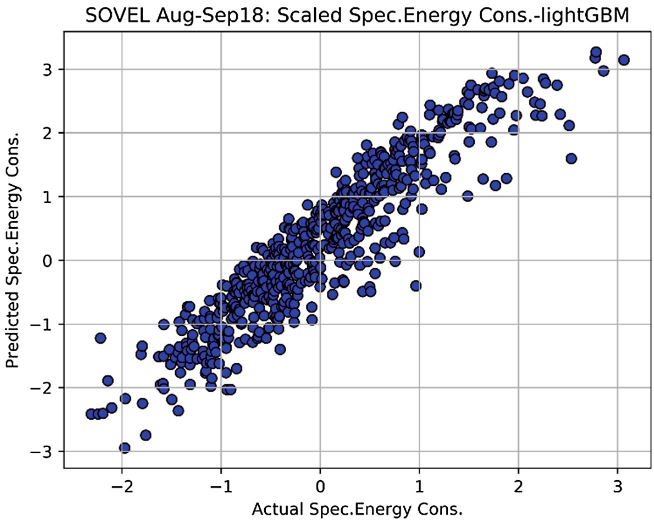

Fig. 13. Actual vs. predicted scaled specific-energy-consumption values for the period Aug-Sep’18 using the new derived lightGBM-model by including the data of this period

Figure 13 depicts actual and predicted scaled specific electrical-energy consumption values for the new derived model. The model was tuned with a learning_rate = 0.01, and n_estimators = 10000; the ANOVA showed improved model values, RMSE = 0.1266, and R2_squared = 0.9833. Furthermore, the predictions for the Aug-Sep'18 data exhibited further improved values, RMSE = 0.573, and R2_squared = 0.628. We would like to point out that this is the 'heart' of machine learning: the system is trained to behave better by getting fresh data by time.

4 Conclusion

Supervised models were developed that predict productivity and specific electrical-energy consumption for the Consteel furnace at SOVEL. SVM and lightGBM based models predicted adequately well the dependent parameters with the latter exhibiting the best correlation characteristics, while Keras was found a bit more difficult to tune. As a further work, one may propose that a potentially online model can be deduced by incorporating Level 2 automation data.

Acknowledgment. The author is grateful to the top management for the continuous support on this type of studies.

References

1. Koehle, S.: Improvements in EAF operating practices over the last decade. In: Electric Furnace Conference Proceedings, Iron & Steel Society, Pittsburgh, PA, vol. 57, pp. 3–14 (1999)
2. Memoli, F., Guzzon, M., Giavani, C.: The evolution of preheating and the importance of hot heel in supersized Consteel® systems. In: AISTech 2011 Proceedings, Indianapolis, IN, vol. I, pp. 823–832 (2011)
3. Bouganosopoulos, B., Papantoniou, V., Sismanis, P.: Start-up experience and results of Consteel® at the SOVEL meltshop. Iron Steel Technol. (2), 38–46 (2009)
4. Pretorius, E.B., Carlisle, R.C.: Foamy slag fundamentals and their practical application to electric furnace steelmaking. In: Electric Furnace Conference Proceedings, Iron & Steel Society, New Orleans, LA, vol. 56, pp. 275–292 (1998)
5. Maes, R.: Celox® for on-line process control in modern steelmaking, (brochure), Heraeus Electro-Nite (2012)
6. Continuum/Anaconda Homepage. https://www.anaconda.com/. Accessed Sept 2017
7. McKinney, W.: pandas: a foundational Python library for data analysis and statistics. http://pandas.pydata.org/. Accessed 20 Aug 2018
8. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
9. Ke, G., et al.: LightGBM: a highly efficient gradient boosting decision tree. In: 31st Conference on Neural Information Processing Systems (NIPS 2017), Advances in Neural Information Processing Systems 30, Microsoft Research (2017)
10. Chollet, F.: Deep Learning with Python. Manning Publications, Shelter Island (2018)
11. Brownlee, J.: Develop Deep Learning Models on Theano and TensorFlow Using Keras. Deep Learning with Python. Jason Brownlee, Melbourne (2018)

12. TensorFlow. https://www.tensorflow.org/. Accessed Aug 2018
13. Sismanis, P.: Analysis of rolled-plates' mechanical properties with a machine-learning software. In: AISTech 2018 Proceedings, Philadelphia, PA, AIST, pp. 2273–2286 (2018)
14. Zumel, N., Mount, J.: Practical Data Science with R. Manning Publications, Shelter Island (2014)
15. Avila, J., Hauck, T.: scikit-learn Cookbook, 2nd edn. Packt Publishing, Birmingham (2017)
16. Garreta, R., Moncecchi, G., Hauck, T., Hackeling, G.: scikit-learn: Machine Learning Simplified. Packt Publishing, Birmingham (2017)