JOM, Vol. 68, No. 8, 2016
DOI: 10.1007/s11837-016-2002-2

Optimization of Sinter Plant Operating Conditions Using Advanced Multivariate Statistics: Intelligent Data Processing

DANIEL FERNÁNDEZ-GONZÁLEZ,[1,4] RAMÓN MARTÍN-DUARTE,[2,5] ÍÑIGO RUIZ-BUSTINZA,[2,6] JAVIER MOCHÓN,[2,7] CARMEN GONZÁLEZ-GASCA,[3,8] and LUIS FELIPE VERDEJA[1,9]

1.—Research Group in Siderurgy, Metallurgy and Materials, Department of Materials Science and Metallurgical Engineering, Oviedo School of Mines, University of Oviedo, Independencia Street, 13, 33004 Oviedo, Asturias, Spain. 2.—Research Group in Processes, Materials and Energy in Ecological and Sustainable Metallurgy (PROMESS), Department of Primary Metallurgy and Recycling, National Centre of Metallurgical Researches, National Council of Scientific Researches (CENIM-CSIC), Gregorio del Amo Avenue, 8, 28001 Madrid, Spain. 3.—Universidad Europea – Laureate International Universities, Calle Tajo s/n, Villaviciosa de Odón, 28670 Madrid, Spain. 4.—e-mail: fernandezgdaniel@uniovi.es. 5.—e-mail: ramon.martinduarte@gmail.com. 6.—e-mail: irbustinza@cenim.csic.es. 7.—e-mail: jmochon@cenim.csic.es. 8.—e-mail: mcarmen.gonzalez@uem.es. 9.—e-mail: lfv@uniovi.es

Blast furnace operators expect to get sinter with homogenous and regular properties (chemical and mechanical), necessary to ensure regular blast furnace operation. Blends for sintering also include several iron by-products and other wastes that are obtained in different processes inside the steelworks. Due to their source, the availability of such materials is not always consistent, but their total production should be consumed in the sintering process, to both save money and recycle wastes. The main scope of this paper is to obtain the least expensive iron ore blend for the sintering process, which will provide suitable chemical and mechanical features for the homogeneous and regular operation of the blast furnace. The systematic use of statistical tools was employed to analyze historical data, including linear and partial correlations applied to the data and fuzzy clustering based on the Sugeno Fuzzy Inference System to establish relationships among the available variables.

INTRODUCTION

The optimization of ore mixtures that are employed in the sintering process plays a key role in sinter quality and production levels. In the sintering process,[1–4] the physical, chemical, and mineralogical properties of raw materials have a relevant influence that should be considered, along with their availability and cost.

Sinter productivity and quality requirements are a function of the blast furnace requirements,[5–14] which impose restrictions on sinter features such as basicity, reducibility, strength, etc.

In daily plant operations, it is necessary to address the actual availability of raw materials and requirements for the blast furnace to reach an optimal performance point. Iron ore features do not always fit precise values; they may change for the same ore deposit and even for the same ore stock.[15] On the other hand, quality requirements can be modified inside certain limits without impairing blast furnace performance.

The main objective of this research was to obtain the least expensive iron ore blend for the sintering process (with a more effective use of raw materials and energy), with homogenous properties and with suitable chemical and mechanical properties (to ensure a soft and regular blast furnace performance).

Blends for sintering include several iron by-products and other wastes (mill scale, dust catcher powder, etc.) that are obtained in ironmaking and steelmaking factories. The supply of these materials is not always consistent, but they should be consumed in the sintering process (to save money and recycle wastes). These materials change the chemical and mechanical properties of the blends, so they must be considered as restrictions in the optimization of an iron ore blend.

METHODOLOGY

Optimization techniques are used to find a set of design parameters that can be defined as optimal in a particular way (maximization or minimization of a certain characteristic subjected to restrictions).[16]

An efficient and accurate solution to this problem is not only dependent on the size of the problems in terms of the number of constraints and design variables but also on the characteristics of the objective function and constraints. For the present case, the optimization problem to be tackled should be considered as a nonlinear programing problem,[17,18] particularly because some of the constraints involved are nonlinear functions. A solution to a nonlinear programing problem generally requires an iterative procedure to establish a search direction at each major iteration, which was the approach to the problem that was used in this work.

Methods of Data Analysis: Subtractive Clustering

The systematic use of statistical tools was employed to analyze historical data. These tools included linear correlation and partial correlation applied to the data and clustering based on the Sugeno Fuzzy Inference System (FIS)[19–21] to establish relationships among the available variables (121,[15]). Sugeno method of fuzzy inference (also known as Takagi–Sugeno–Kang) was introduced in 1985.[15] It was developed as systematic approach to generating fuzzy rules from a given input–output dataset. A rule in a Sugeno fuzzy model is:[6,15]

$$\text{if } x_1 \text{ is } A \text{ and } x_2 \text{ is} B \text{ then } z = f(x_1, x_2) \tag{1}$$

where A and B are fuzzy sets in the antecedent, while $z = f(x_1, x_2)$ is a numerical function in the consequent [typically $f(x_1, x_2)$ is a polynomial function of the entry variable]. The Sugeno Inference System was used because it is effective for optimization problems and is computationally efficient.[6]

Once main relationships among iron ore and sinter were established, a non-linear optimization algorithm based on step-descent methods was developed. Applied restrictions were related with the conditions imposed on the sinter properties and the use of iron ore and by-products.

The clustering of numerical data forms the basis of many classification and system modeling algorithms. The purpose of clustering is to identify natural groupings of data from a large dataset to produce a concise representation of a system's behavior,[6] which can use the cluster information to generate a Sugeno-type FIS that best models the data behavior using a minimum number of rules.

The advantage of using the subtractive clustering algorithm is that the number of clusters does not need to be a priori specified; instead, the method can be used to determine the number of clusters and their values.

The Subtractive Clustering method[22,23] assumes that each data point $z_j = (x_j, y_j)$ has assigned a potential, P_j, according to its location to all other data points. The potential, P_i^*, at data point x_i is defined as:

$$P_i^* = \sum_{i=1}^{n} \exp\left(-\frac{\|x_i - x_j\|^2}{(r_a/2)^2}\right) \tag{2}$$

where P_i^* is the potential-value i-data as a cluster center; r_a is a positive constant called cluster radius (a neighbor radius); and x is the data point. Hence, the potential of a data point to be a cluster center is higher when there are more neighboring data points. The data point with the highest potential (P_k^*) is considered as the first cluster center (x_{c1}). The potential is then recalculated for all other points (x_i) excluding the influence of the first cluster center as follows:

$$P_i^* = P_i^* - P_k^* \cdot \exp\left(-\frac{\|x_i - x_{c1}\|^2}{(r_b/2)^2}\right) \tag{3}$$

where r_b is a positive constant which defines a neighborhood that has measurable reductions in potential value. After revising the potential value, the next cluster center is selected as the point having the greatest potential value. The process continues until a sufficient number of clusters are defined.

The generation of the FIS matrix, which is a MATLAB object that contains all the FIS information (including variable names, membership function definitions, etc.),[24] is accomplished through previous training, which is conducted prior to building this matrix. Relationships between inputs and outputs a priori are searched, and the data with similar behavior are clustered, such that the number of rules has been reduced, being equal to the number of clusters. Therefore, the FIS matrix has an equal number of membership functions for each input, as clusters have been found.[25,26]

The best cluster parameters for this system were obtained using the trial and error method. Thus, the cluster radius[27] that was utilized is a function of the amount of variables that were used. Higher number of variables resulted in longer time required in the calculus process. All processes (data introduction in the software, the estimation process and the results presentation) should not be longer than 60 min (minimum time required for the iron ore blend for reaching the blast furnace hopper feeders). Longer calculus times (higher number of variables) suppose a better accuracy, although the problem would be that the estimated values could not be relevant in the factory control (because the results would be obtained once the sinter was fed to the blast furnace), and the quality control of the sinter that feeds the blast furnace would definitely not be achieved.

Optimization of Ore Blends for Sintering

Data Collection

Two different sets of data were collected from a European sinter factory plant. The first set covers a period of 6 years and a total amount of 216 stockpiles (15,000–30,000 tons of mineral).

A second set of data was collected during the early months of the project, covering a period of 10 months and a total amount of 35 ore stockpiles.

To establish suitable relationships among iron ore and the chemical properties of blends, the European factory also supplied the results of the standard chemical analysis that are employed on the raw materials, which are used as other variables in the optimization process.[15]

Data Analysis

Three steps were employed to perform the data analysis: (1) simple data inspection and data pretreatment, including the removal of outliers and the rejection of non-significant parameters; (2) a search for a linear relationship among variables; and (3) the study of partial correlation to elucidate the mutual influence among variables.

Linear Correlation

Obtaining a simple correlation between variables is the second step. This correlation was obtained by means of the correlation coefficient calculated in a standard way:

$$r_{xy} = \frac{\sum_N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\left[\sum_N (x_i - \bar{x})^2 \cdot \sum_N (y_i - \bar{y})^2\right]^{1/2}} \tag{4}$$

where x_i is the value of x for observation i, $\bar{x}$ is the mean x value, y_i is the value of y for observation i, and $\bar{y}$ is the mean y value. The following description will be focused on the most significant results on linear correlation. Therefore, the effort will be focused on the correlation between plausible input variables and plausible output variables. In the cases in which the existence of a correlation between input variables could be noted, the existence of co-linearity will also be indicated.

The correlation coefficient between iron content in ore blend and sinter is approximately 0.55. Similar results were achieved for FeO, although this variable is known to be largely altered during the sintering process. In fact, the FeO content in sinter also presents some degree of negative correlation with the sinter strand bed height and positive correlation with the sinter strand speed. Moreover, it has a negative correlation with coke and the total amount of Fe that is present in sinter.

For CaO, there is not a significant correlation between ore blend and sinter content, but a correlation of 0.5 is found between CaO sinter content and the additions of limestone and lime employed. MgO content in sinter present a reasonable (~ 0.55) correlation with MgO content in ore blends and with the additions of dunite. Al_2O_3 content in sinter is correlated with coke, which provides an appreciable amount of Al_2O_3 to the sinter, and with its initial amount in the ore blend. No correlations are found for SiO_2, possibly because there is not significant variation of SiO_2 content during the analyzed period.

Tumbler index [ISO 3271 (2015)] is a sinter quality index used in the ironmaking industry that provides a measure of the resistance of iron oxides to breakage or degradation by impact and abrasion.[12] The Tumbler index (%> 6.3 mm) is positively correlated with the bed height and negatively correlated with strand speed, air temperature and lime addition.

RDI (Reduction Degradation Index) [ISO 4696-2 (2015)] is a sinter quality index used in the ironmaking industry that provides a measure of the degradation of the sinter that could occur in the upper section of the blast furnace after some reduction (a high degree of reduction disintegration generates fines in the top of the furnace that affects the flow distribution within the blast furnace).[12] RDI (Reduction Degradation Index) is positively correlated with bed height and coke consumption, while being negatively correlated with strand speed, lime addition, CaO content and Mn content.

Apart from the stochastic aspects, there are thermodynamic aspects that induce the use of a partial correlation between variables.

Partial Correlation

The combined effect of input variables and their possible interrelations cause problems when attempting to carry out a conclusion from the simple correlation obtained by the method described above. Therefore, it is necessary to obtain the correlation between two variables while avoiding the effect of the rest,[28] which is a known partial correlation. To perform this, a new partial correlation coefficient was employed as described below:

$$\hat{r}_{x_i x_j, R} = \frac{t_j}{\sqrt{t_j^2 + n - k - 1}} \tag{5}$$

where n is the sample size, k is the degree of freedom, and t_j is defined as:

$$t_j = \frac{\hat{b}_j}{\hat{s}_R \sqrt{q_{jj}}} \tag{6}$$

$\hat{b}_j$ represents the coefficient of variable x_j in the linear regression model of x_i, $\hat{S}_R$ would be the variance non-explained with the linear regression model and q_{jj} will be the diagonal element 'j' of the covariance matrix.

Although partial correlation was studied for every variable, only the results concerning sinter properties will be presented here.

Some remarkable correlations are:

Tumbler index is positively correlated (among others) with: variables that represent the ignition furnace parameters, an index of gas consumption in the ignition furnace, the temperature in the windbox at the end of the strand, and the coke content.

Tumbler index is negatively correlated with: Fe content in the blend, MgO content in the blend and strand speed.

In the RDI case, the strong influence of Alkalis ($K_2O + Na_2O$) and Al_2O_3 that is observed is remarkable.

As is well known, the influence of alkalis activates the coke gasification kinetics (Boudouard Reaction):[29–31]

$$\mathrm{Coke} + CO_2 \Leftrightarrow 2CO(g)$$

Therefore, there is a maximum calorific power of the coke reduction (total oxidation in excess of oxygen to CO_2) and the possibility of reaching maximum temperatures in the sintering front. Consequently, the Tumbler and RDI indexes decrease.

Fuzzy Inference Systems for RDI and Tumbler Index Estimation

RDI and Tumbler indexes depend on a complex method on the chemical properties of iron ores, as well as the performance of the sinter strand.

To estimate the values for these sinter properties, a FIS was developed. The employed model was a Sugeno-type FIS, which was obtained by means of the subtractive clustering algorithm.[22] This algorithm permits the modelling of data behavior, which clusters the experimental data around some values that are obtained during the modelling process.[22] The model can be fitted by employing the cluster radius.[27] This parameter indicates the range of influence of a cluster and must be specified beforehand.[27] Specifying a small cluster radius will typically yield many small clusters in the data.[27] Specifying a large cluster radius will typically yield a few large clusters in the data.[27]

In the case of RDI estimation, the input variables were selected according to the results obtained in the previous study using partial correlation.[15] The input variables selected were sinter Al_2O_3 content, sinter Alkalis content, coke content, air temperature and bed height. A collection of 200 samples of historical data was divided into two sets of 100 samples. The first set was employed for training, and the second set was employed for validation. The model was developed with a rather large value for the cluster radius. Because RDI measurements are often imprecise, the model should show only the tendency of RDI.

The Tumbler index was modeled in a very similar way.[15] The input variables selected in this case were: sinter SiO_2 content, sinter CaO content, air flow, air temperature, and coke content.

To evaluate the adjustment between the model and real measures, a correlation coefficient was used. The performance index in the *training period* (this period helps to select the model and estimate its parameters) was 0.934 for RDI and 0.702 for the Tumbler index. The performance index in the *validation period* (where the forecasted model is tested to know if it functions properly) was 0.916 for RDI and 0.832 for the Tumbler index.

The models for the RDI and Tumbler indexes are included in the optimization process as two quality elements, to ensure a material with the blast furnace operator quality parameters.

Table I. Restrictions in the optimization process for sinter

	Value		
Variable	**Maximum**	**Minimum**	**Exact**
% Fe		56	
% MgO			1.65
Basicity Index			1.70
% SiO_2	5.40	5.30	
% Al_2O_3	1.35		
% Alkalis	0.11		
% Phosphorus	0.04		
Tumbler index (%)		74	
RDI index (%)	33		
% < 0.125 mm	15.00		
0.2 mm < % < 0.7 mm	18.00		

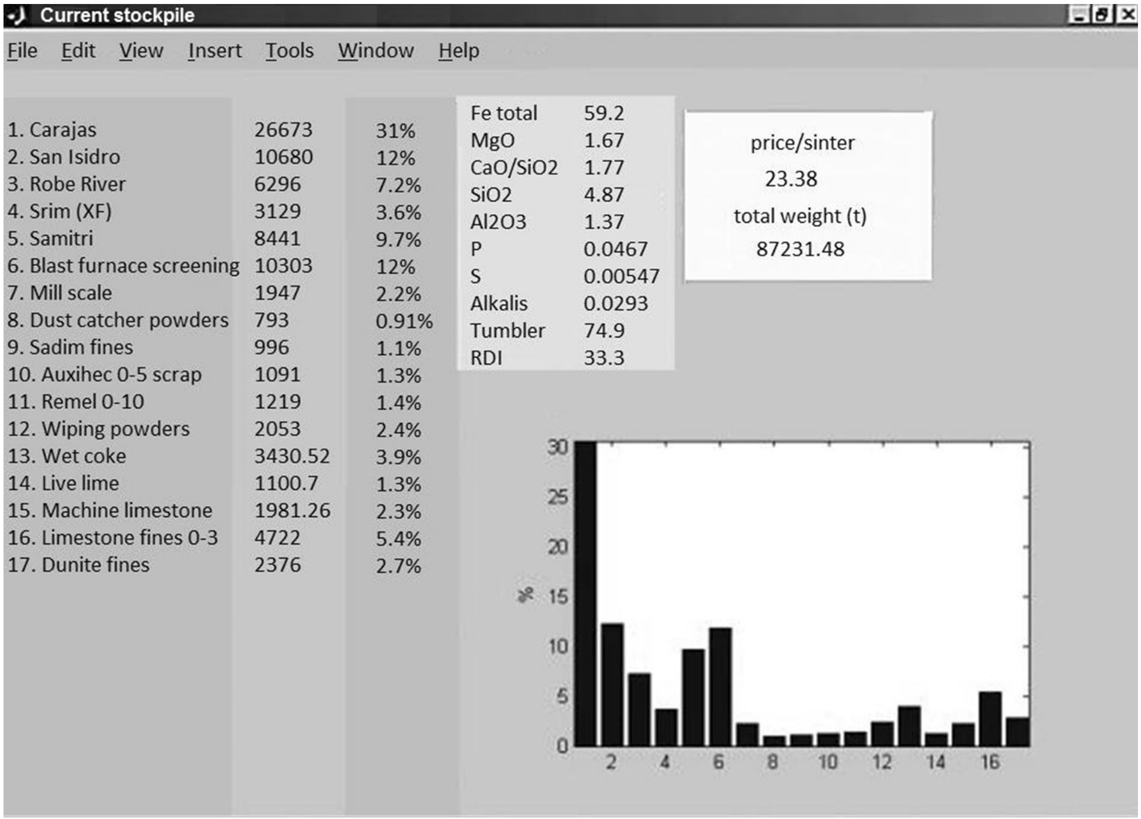

Fig. 1. Features of the optimum blend obtained from stockpile number 30.

Optimization Problem Definition

As a final goal, the optimization process, must minimize the price of the ore blend to ensure the minimum quality requirements (restrictions in the optimization process can be read in Table I). To reach a proper definition of the problem, it is necessary to establish the relationship between these restrictions and the ore blend properties and to set their values.

The content in the sinter of the variables considered as restrictions in the European factory practice is calculated as follows:

$$\%X(\mathrm{S}) = \frac{\sum_i \%X(i) \cdot M(i)}{\sum_i \mathrm{RS}(i) \cdot M(i)} \cdot 100 \qquad (7)$$

where $X(\mathrm{S})$ is the X phase content in the sinter, $X(i)$ is the X phase content in the ore i, $M(I)$ is the mass of ore (i) that should be employed in the blend and $\mathrm{RS}(i)$ is the ore-sinter yield for ore (i). This last variable is obtained by subtracting the losses due to humidity, calcinations, de-sulfuration and de-alkalization.

The results are:

- Iron content ($>56\%$): The results of operation show that the work point is typically far beyond this limit, mainly because Fe content has a negative influence on the sinter strength. Therefore, its content may not be far beyond the lower limit imposed.
- MgO (1.65%): Data obtained from the last 6 months show that the results are close to this value (mean: 1.65, deviation 0.15).
- Basicity Index (1.7): The results that were obtained show a mean value of 1.69 with a standard deviation of 0.10.
- SiO_2 content (5.3–5.4%): The results that were obtained show a mean value of 5.03 with a standard deviation of 0.24. These results lead to a review of the way in which this restriction is imposed, not only because the mean value is out of range but also because the allowed interval width is narrower than the actual standard deviation of data.
- Al_2O_3 ($<1.35\%$): The results show a mean value of 1.12 with a deviation of 0.06. In fact, the upper bound is hardly ever crossed.
- Alkalis ($<0.11\%$): To estimate the alkalis content in the sinter, the de-alkalization process (%dealk) must be taken into account. Therefore, the equation in this case is:

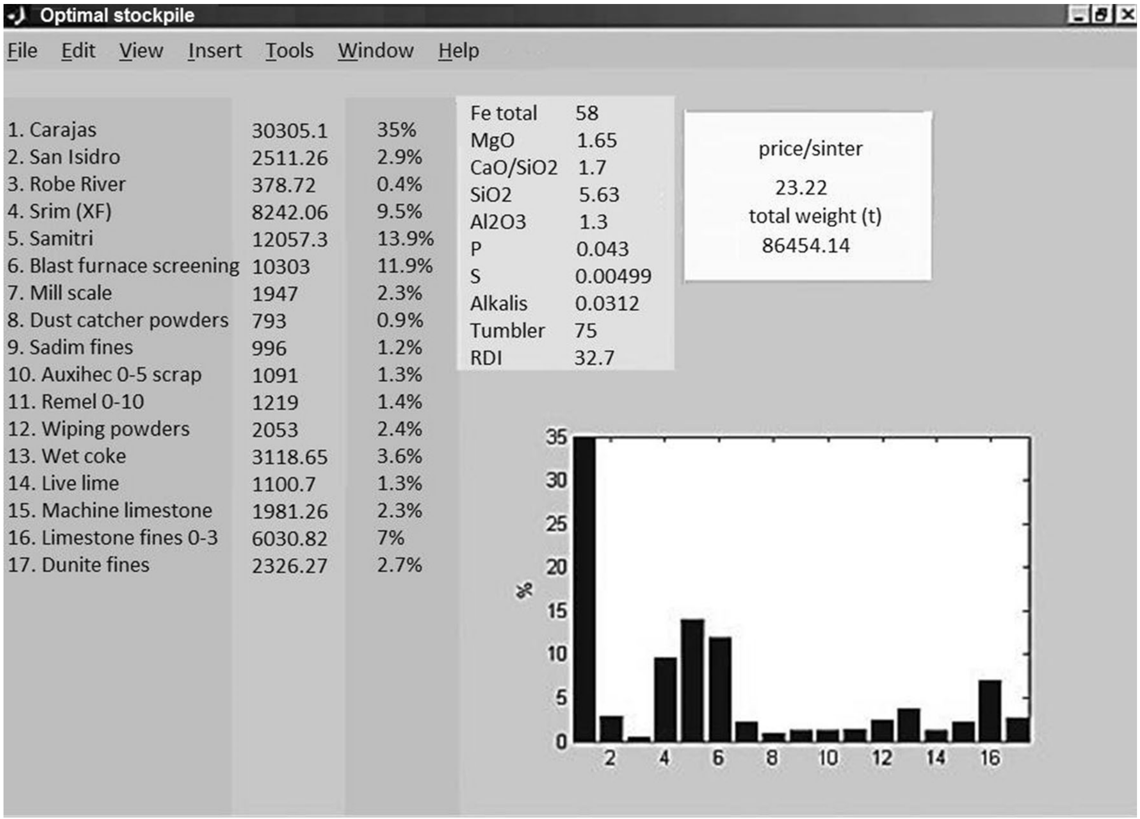

Fig. 2. Features of the optimum blend obtained from stockpile number 30 fixing set points to their actual limits.

$$\%\mathrm{Alk}(\mathrm{S}) = \frac{\sum_i \%\mathrm{Alk}(i)\cdot M(i)\left(1-\frac{\%\mathrm{dealk}}{100}\right)}{\sum_i \mathrm{RS}(i)\cdot M(i)}\cdot 100 \quad (8)$$

%dealk has been established at 40%. Data on alkali content show a mean value of 0.022 with a standard deviation of 0.0035 and, thus, do not appear to be an important restriction.

– Phosphorus ($<0.04\%$): Data for P content show a mean value of 0.046 with a standard deviation of 0.003.

MATLAB Software Tool

With all that has been previously mentioned, a program was developed in MATLAB to conduct price and quality optimization for the iron ore blends that are used in a European sinter plant. To analyze the performance of the optimizer, blends employed in the past were reproduced.

A reduction in the price per ton of sinter of approximately 0.15 € [original 23.38 €/*t* sinter (Fig. 1), optimizer tool 23.22 €/*t* sinter (Fig. 2), in the case of the stockpile 30] was predicted. In a sinter plant of 2.5 Mt, the cost savings would be approximately 375,000 €. The software shows the optimum mix of iron ore, recycled products, slag-forming elements, fluxes, etc. that satisfies the restrictions described in previous sections.

CONCLUSION

During this work, data from a European sinter plant were collected and studied by means of statistical tools to establish the relationship among variables (chemical and mechanical properties of both blends and sintered products, and sintering process variables). Non-linear relationships were found; therefore, non-linear optimization was proposed.

The RDI and Tumbler indexes were considered as restrictions of the process. As a consequence of their relationship with other variables, fuzzy programing (Sugeno's inference model) was used with the purpose of predicting them.

The goal was to obtain a mixture that fulfilled the conditions imposed on sinter quality, while also obtaining the least expensive mixture. This objective was reached by means of the development of a MATLAB optimizer, which obtained an iron ore blend that was less expensive than that previously used and contained quality parameters that ensured homogeneous and consistent operation in the blast furnace.

ACKNOWLEDGEMENTS

The authors wish to extend their appreciation to the Spanish MICYT (MAT 2001-4435-E) for their financial support. The research was also supported by the Spanish Ministry of Education, Culture and Sports via an FPU (Formación del Profesorado Universitario) grant to Daniel Fernández González (FPU014/02436).

REFERENCES

1. P.R. Dawson, *Ironmak. Steelmak.* 20, 137 (1993).
2. J. Mitterlehner, G. Loeffler, F. Winter, H. Hofbauer, H. Schmid, E. Zwittag, T.H. Buergler, O. Pammer, and H. Stiasny, *ISIJ Int.* 44, 11 (2004).
3. W. Yang, R. Changkook, S. Choi, E. Choi, D. Lee, and W. Huh, *ISIJ Int.* 44, 492 (2004).
4. G. Szatvanyi, C. Duchesne, and G. Bartolacci, *Ind. Eng. Chem. Res.* 45, 4706 (2006).
5. L. Wang, *Adaptive Fuzzy System and Control Design and Stability Analysis* (New York: Prentice Hall, 1994), p. 232.
6. N. Gulley and J.R. Jang, *Fuzzy Logic Toolbox User's Guide* (Natick: The Math Works, 2005).
7. J.A. Freeman and D.M. Skapura, *Neuronal Network—Algorithms, Application and Programming Techniques* (New York: Addison Wesley Publishing Company, 1991), p. 401.
8. R.K. Sarma, A. Gupta, and S. Vadhavkar, *Neural Network-Based Data Mining Techniques for Steel Making*, ed. H. Bozdogan (Florida: CRC Press LLC, 2004), p. 401.
9. J. Hinnelä and H. Saxén, *ISIJ Int.* 41, 142 (2001).
10. P.J. Laitinen and H. Saxén, *Ironmak. Steelmak.* 34, 109 (2007).
11. A. Cores, L.F. Verdeja, S. Ferreira, I. Ruiz-Bustinza, and J. Mochón, *Dyna-Colombia* 80, 152 (2013).
12. J. Mochón, A. Cores, I. Ruiz-Bustinza, L.F. Verdeja, J.I. Robla, and F. García-Carcedo, *Dyna-Colombia* 81, 168 (2014).
13. A. Cores, L.F. Verdeja, S. Ferreira, I. Ruiz-Bustinza, J. Mochón, J.I. Robla, and C. González-Gasca, *Dyna-Colombia* 82, 227 (2015).
14. J. Iglesias, J. Sancho, M. Ordiales, D. Fernández, J. Mochón, I. Ruiz-Bustinza, A. Fuentes, and L.F. Verdeja, *Ironmak. Steelmak.* 42, 618 (2015).
15. M.R. Martín, *Optimización de las condiciones de operación de una planta de sínter y localización del parque de minerales por medio de estadísticas multivariantes avanzadas y lógica difusa* (Oviedo, ES: Universidad de Oviedo, PhD. Thesis, 2015).
16. R. Fletcher, *Practical Methods of Optimisation: Constrained Optimisation*, vol. 2 (Chichester: Wiley, 1981), p. 224.
17. R. Barea, J. Mochón, A. Cores, and R. Martín, *ISIJ Int.* 46, 687 (2006).
18. R. Martín, J. Mochón, L.F. Verdeja, R. Barea, P. Rusek, and J. Jiménez, *Steel Res. Int.* 80, 185 (2009).
19. J.S.R. Jang, C.T. Sun, and E. Mizutani, *IEEE Trans. Autom. Control* 42, 1482 (1997).
20. J.S. Bendat and A.G. Piersol, *Random Data. Analysis and Measurements Procedures*, (New York: Wiley, 1987), p. 566.
21. J.M. Mendel, *IEEE Proc.* 83, 345 (1995).
22. J.S.R. Jang, C.T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing* (New York: Prentice Hall, 1996), p. 614.
23. A. Priyono, M. Ridwan, A.J. Alias, R. Atiq, O.K. Rahmat, A. Hassan, and A. Mohd, *J. Teknol.* 43, 143 (2005).
24. A. Lofti, H.C. Andersen, and A.C. Tsoi, *IEEE Trans. Syst. Man Cybern. B* 26, 332 (1996).
25. R.R. Yager and D.P. Filev, *J. Intell. Fuzzy Syst.* 2, 209 (1994).
26. M.V. Ramos, E. Kasai, J. Kano, and T. Nakamura, *ISIJ Int.* 40, 448 (2000).
27. A.K. Lohani, N.K. Goel, and K.K.S. Bhatia, *J. Hydrol.* 331, 146 (2006).
28. L.S. Meyers, G. Gamst, and A.J. Guarino, *Applied Multivariate Research. Design and Interpretation*, 2nd ed. (California, CA: SAGE Publications, Inc. 2013), pp. 337–338.
29. J.P. Sancho, L.F. Verdeja, and A. Ballester, *Metalurgia Extractiva. Volumen II: Procesos de obtención*, (Madrid, ES: Síntesis, 2000), pp. 22–27.
30. A. Babich, D. Senk, H.W. Gudenau, and TTh Mavrommatis, *Ironmaking, (Aachen, DE: Department of Ferreous Metallurgy* (Aachen: RWTH Aachen University, 2008), pp. 50–64.
31. E. Lwamba and A.M. Garbers-Craig, *J. S. Afr. I Min. Metall.* 108, 293 (2008).