INTELLIGENT PRODUCTION DATA ANALYTICS FOR METAL INDUSTRY 4.0

Marcin PERZYK, Jacek KOZLOWSKI

Warsaw University of Technology, Warszawa, Poland, EU
M.Perzyk@wip.pw.edu.pl, jkozlows76@gmail.com

Abstract

The fundamental role of data analytics in the era of Industry 4.0 is characterized. The main types of data-driven advanced models, including learning systems, are presented. Potential applications in process control and fault diagnosis in metal industry are reviewed. Examples related to metal casting processes, based on the authors' expertise are presented. Potential chances and threads of intelligent data analytics in metal industry are discussed.

Keywords: Metal industry 4.0, data-driven modeling, process control, fault diagnosis

1. INTRODUCTION

In all industries, including the metal industry, the 4th revolution has come. A combination of modern technologies leading to a 'smart factory' is the essence of the Industry 4.0. Big Data analytics and machine learning technologies are doubtless the key issues, driving greater efficiency, higher production and faster fulfillment giving new opportunities [1]. One of the main challenges is development of methodologies capable of extracting useful knowledge from the huge amounts of data now available in most of manufacturing companies and further to utilize it in control of the processes, predicting process behavior or equipment failures and diagnosing the process disturbances and the products' defects.

The aim of a process control is always to limit its variability, irrespectively of its nature and manifestation. There are two main strategies to accomplish this objective: Engineering Process Control (EPC), based on the regulation of the process and Statistical Process Control (SPC), based on process monitoring [2]. EPC originates from process industry such as chemical, metallurgical, food, textile etc. whereas SPC was first applied in the parts industry, such as automotive, electrical and electronic, machine-building, aircraft, shipbuilding etc. EPC actively counteracts the process disturbances by making adjustments to process variables in order to keep the output quality parameter on target. The three classic types of control systems are utilized also in the manufacturing process control, i.e. open-loop, feed-forward and feedback. SPC assumes that the process output can be described by statistically independent observations fluctuating around a constant mean and is intended to detect signals which represent the special (assignable) causes of external disturbances increasing the process variation and subsequently remove them. There are two main types of SPC and the corresponding types of control charts: for continuous variables (e.g. product dimensions) and for attributes, e.g. defining product as acceptable or non-acceptable (e.g. defective). Typically, samples consisting of several products are checked in regular time intervals, however, in many metal industry branches e.g. foundry production, single measurements are common. The extreme case is that all products are checked which is typical for some automated systems, becoming common in contemporary manufacturing industry.

The present paper explains and systemizes possibilities for advanced data analytics in application to control of manufacturing processes, including fault diagnosis and predictions. The most advanced data-driven models are learning systems, often associated with the discipline named Computational Intelligence, have proved to be very successful in many areas of human activity. They have found various and valuable applications also in control and fault diagnosis of complex manufacturing processes, including such branches like metallurgy and metal casting [3-16]. The characteristic applications of that kind of tools, illustrated with examples from metal industry, mostly based on the authors experience in metallurgy and metal casting areas, are presented in the paper.

2. DATA-DRIVEN EPC IN MANUFACTURING

Probably the most widespread type of advanced data-driven (soft) models used in process control are artificial neural networks often utilizing imprecisely expressed (fuzzy) data in the form of neural or neuro-fuzzy controllers. Typically they utilize the data recorded in the course of normal production, however, sometimes artificially obtained data from process simulations are also helpful.

The models used for the process control can be in various forms, including hardware components working in the automated systems but also they can be the computer programs acting as virtual assistants for the operators or engineering staff. An example of that kind of control utility was presented in [17]. A neural model of the complex heat treatment process necessary to obtain Austempered Ductile Iron was build on the basis of a large amount of data collected in many different companies and laboratories in the world and then successfully implemented in a dedicated software, illustrated in **Figure 1**.

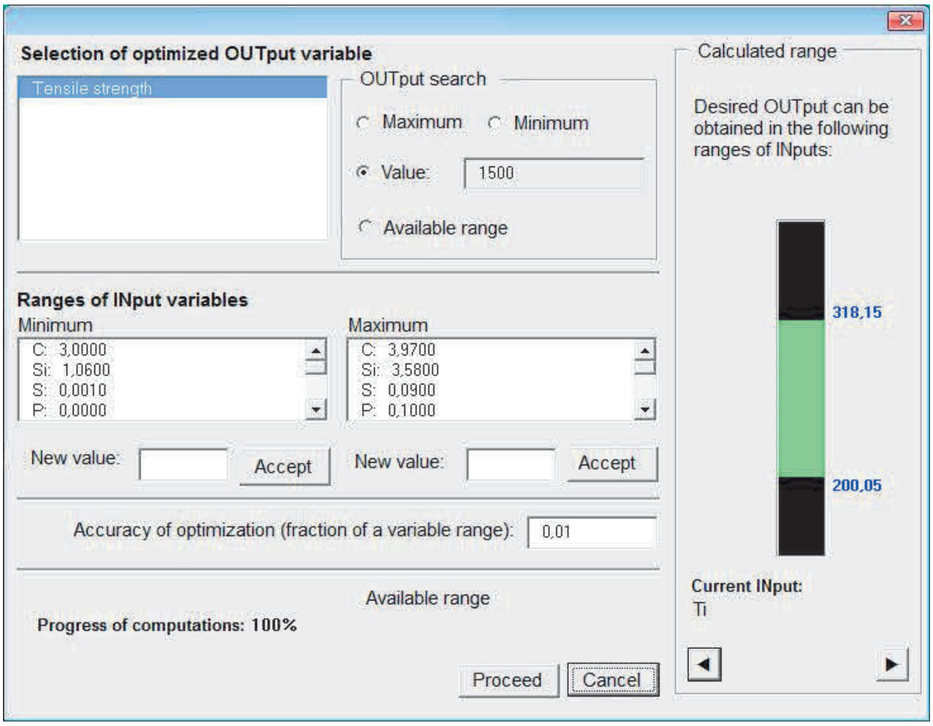

Figure 1 A dialog box facilitating the optimization of ADI heat treatment parameters utilizing neural modeling of the process [17]

Another type of the data-driven models utilized in control of manufacturing process are the logic rule systems, based on automated knowledge extraction from the recorded industrial data. The logic rules of the type: 'IF [conditions] THEN [decision class]' can be obtained with several types of tools. For manufacturing problems the decision trees are probably the most frequently used whereas the methods based on Rough Sets Theory seem to be their newer alternative [18-21]. In **Figure 2** an example of decision tree, obtained from the data described above, is shown. Here, the output variable was the grade of ADI, classified on the basis of its tensile strength, elongation and Brinell hardness. From the fragment of the tree shown in **Figure 2** the following recommendation can be obtained: the high strength cast iron of the grade 1,400 can be obtained by applying the isothermal austempering temperature below 337 °C and Si content above 2.9 %, irrespective the other heat treatment parameters and contents of other elements [17].

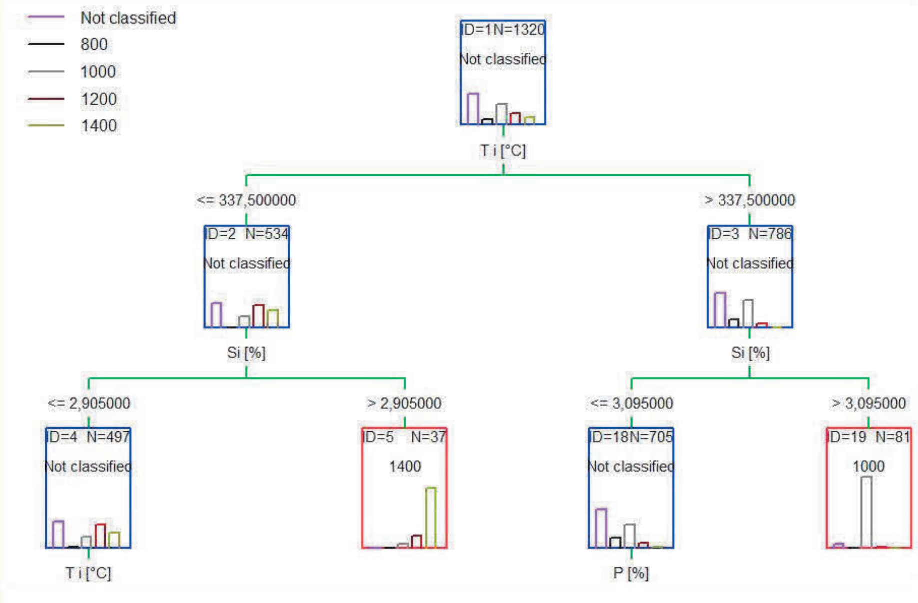

Figure 2 Fragment of decision (classification) tree utilized for obtaining ADI grades (see the chart legend) [17]; red rectangles mark end nodes showing the results of classification, blue rectangles are splitting (intermediate) nodes, the bars in all rectangles show distributions of the ADI grades (classes) in the node

3. STATISTICAL PROCESS CONTROL AND DIAGNOSIS

In general, three type of process variations can be observed on the process run charts: random normal variations around a constant mean, abnormal sequences (patterns) of points being signals of external disturbances and autocorrelations, i.e. dependencies of the current values on the preceding ones. SPC assumes the absence of the autocorrelations.

A general approach in application of the data-driven modeling to process fault diagnosis is that a process model linking the process parameters as inputs with the process outputs enables to make process simulations leading to identification of the cause of the process fault (e.g. product quality or abnormal sequence of any process parameter). Several examples of applications of process modeling leading to successful identification of causes of metallic products defects, taken from the present authors experience, illustrate this approach:

- Gas porosity appearing in steel castings was attributed by the present authors to the combination of molding sand properties and current atmospheric conditions based on neural modeling of the process
- Various data-driven models were successfully applied in order to find the root causes of heavy oscillation marks appearing on surfaces of steel billets produced by continuous casting process
- The ductile iron melting process was modeled using classification trees in order to find the most probable causes of missed melts, i.e. the combination of mechanical properties of the cast iron incompatible with the target iron grade, treated here as the process output

In [22] a neural model was applied to find the most likely causes of abnormal patterns appearing on Shewhart charts. Like in [23] the ductile iron melting process was analyzed, however, the process output was the iron tensile strength with all the values in the tolerance range. In **Figure 3** a characteristic example is shown. The abnormal sequence of points defined as *9 consecutive points on one side of central line* was observed on the output chart, but it could not be easily attributed to any of the process inputs (the chemical contents of the main chemical elements). However, it can be seen that the neural model imitates the real run chart quite well. It gives an opportunity to find the most likely root causes of the process disturbance by making suitable simulations, for example assuming some of the inputs at their levels from before the appearance of the abnormal pattern.

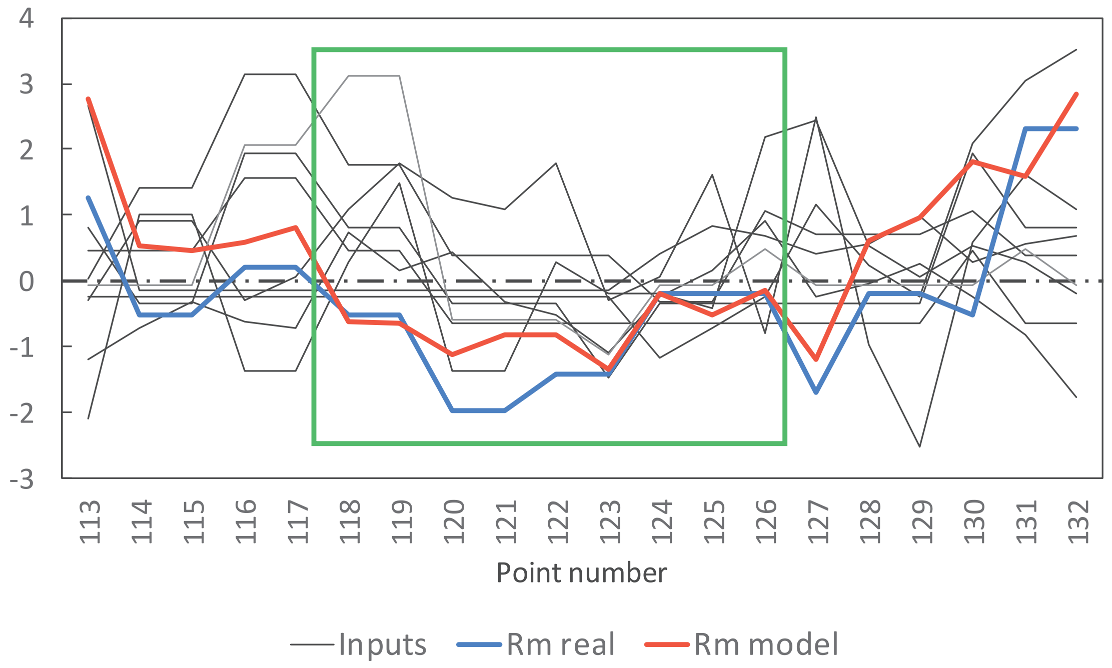

Figure 3 Example of the possible applications of neural regression modeling in finding the root causes of abnormal patterns on Shewhart charts; the green rectangle marks the abnormal sequence of points on the run chart [22]

4. AUTOCORRELATIONS IN PROCESSES

Autocorrelations appearing in runs of process variables enable making predictions of the future values using the methods of time-series analysis. However, they can also be a sign of hidden imperfections in the process and they can be misinterpreted as SPC abnormal patterns. The main types of the autocorrelations include variable means' trend and periodicity (seasonality). Surprisingly, autocorrelations may appear quite frequently in the real processes, including metallurgy and metal casting areas [23, 24, 25].

An illustrative example of detection of hidden imperfections in the grey cast iron melting process is given in [23]. Average Si content in the last 3 measurements of a day was the process variable in which a significant periodicity equal 7 was detected. It could indicate that the stock of the additives containing silicon is suitable just for a period of about 7 working days and the practice of the operators is to spend it in full in that period (increasing or decreasing the amounts added at the end of the period).

The main purpose of the predictions of the future values in a production process is to enable the staff to take appropriate actions anticipating the expected changes in the process. It was found [23] that the prediction accuracy for both current and daily-based averaged values of the chemical components of grey cast iron appeared to be about 20 % of the whole ranges of their concentrations, which can be accepted for undertaking suitable anticipated actions in most of the practical situations.

Removing the autocorrelations components from the original data allows to avoid false signals in SPC. The resulting Special Cause Control (SCC) charts utilize the residual data obtained from the time-series analysis. In **Figure 4** a characteristic example is shown. Here the main autocorrelation component found in the data was a significant periodicity, unnoticeable on the original run charts and also unexpected by the operators and the engineering staff. The results of application of SCC charts to a green sand processing system in a grey cast iron foundry showed that application of the SCC charts reduces numbers of the signals in almost all cases [24] which can be helpful in avoiding false signals, i.e. resulting from predictable factors.

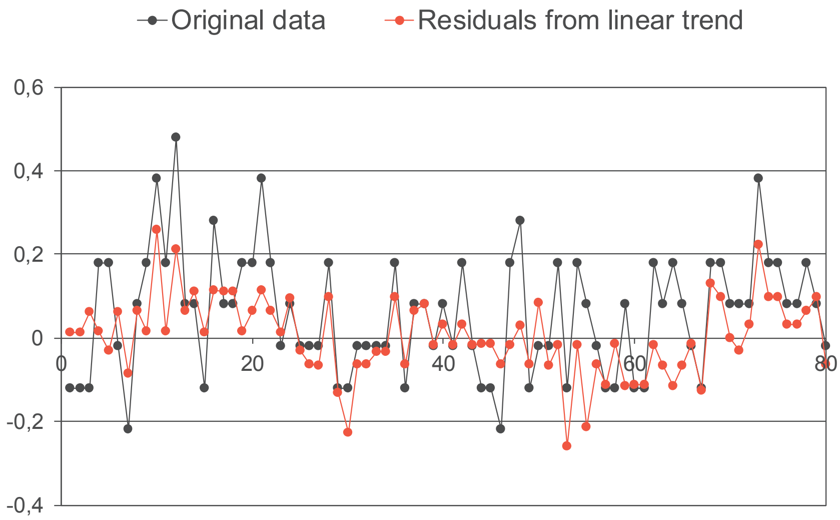

Figure 4 Example of the original data transformation and residuals obtained from time-series analysis (used for SCC charts) for moisture of green molding sand in a medium-size grey cast iron foundry [24]

5. CONCLUSION

The present paper presents possibilities and examples of applications of advanced data-driven modeling in metal industry. They become remarkably more real and in the era of the 4th industrial evolution, mainly due to the unprecedented availability of the automatically gained and recorded production data.

Of course, these chances are accompanied by some significant threats. One of them is certainly a strongly limited availability of some important data, for example those obtained from destructive tests of costly products which are typical for many branches of metal industry. Also the cases of equipment failure are usually seldom. It seems to be a more general observation that in the era of Industry 4.0 some important limitations due to the physical nature of products and processes may play a significant role.

A successful application of the big data analytics and machine learning technologies in metal industry presents a challenge to analysts and engineers. Without good understanding the nature of the data-driven models engineers will not be able to recognize their possibilities and successfully apply them in industrial practice. The role of updating all forms of education is therefore essential.

REFERENCES

[1] ZAVALISHINA, Jane. Manufacturing and the fourth revolution. *Control Engineering* [online]. 2016. [viewed 11.12.2017]. Available from: https://internetofbusiness.com/manufacturings-4th-revolution.

[2] MONTGOMERY, Douglas C., KEATS, Bert J., RUNGER, George C. and MESSINA, William S. Integrating statistical process control and engineering process control. *Journal of Quality Technology*. 1994. vol. 26, iss. 2, pp. 79-87.

[3] CHIU, Chih C., SHAO, Yuehjen E., LEE, Tian S. and LEE, Ker M. Identification of process disturbance using SPC/EPC and neural networks. *Journal of Intelligent Manufacturing*. 2003. vol. 14, iss. 3-4, pp. 379-388.

[4] HARDING, Jenny A., SHAHBAZ, Muhammad, SRINIVAS and KUSIAK, Andrew. Data mining in manufacturing: a review. *ASME, Journal of Manufacturing Science and Engineering*. 2006. vol. 128, iss. 4, pp. 969-976.

[5] JACOBS, Dereya A. and LUKE, Stephen R. Training Artificial Neural Networks For Statistical Process Control, In *Proceedings of the Tenth Biennial University/Government/Industry Microelectronics Symposium*. Research Triangle Park, NC, USA, 1993, pp. 235-239.

[6] KAMAL, Amr M.M. A data mining approach for improving manufacturing processes quality control. In *Proceedings of 2nd International Conference on Next Generation Information Technology, ICNIT 2011*. Hong Kong, China: IEEE, 2011, pp. 207-212.

[7] KOONCE, David, FANG, Cheng H. and TSAI, Shi C. Data mining tool for learning from manufacturing systems. *Computers and Industrial Engineering*. 1997. vol. 33, no. 1-2, pp. 27-30.

[8] KUSIAK, Andrew. Data mining: manufacturing and service applications. *International Journal of Production Research*. 2006. vol. 44, iss. 18-19, pp. 4175-4191.

[9] OZTURK, Atakan, KAJALIGIL, Sinan and OZDEMIREL, Nur E. Manufacturing lead time estimation using data mining. *European Journal of Operational Research*. 2006. vol, 173, iss. 2, pp. 683-700.

[10] SHAHBAZ, Muhammad, SRINIVAS, HARDING, Jenny A. and TURNER Marc. Product design and manufacturing process improvement using association rules. *Proceedings of the Institution of Mechanical Engineers, Journal of Engineering Manufacture*. 2006. vol. 220, iss. 2, pp. 243-254.

[11] SHEN, Lixiang, TAY, Francis E. H., QU, Liangsheng and SHEN, Yudi. Fault diagnosis using Rough Sets Theory. *Computers in Industry*. 2000. vol. 43, iss. 1, pp. 61-72.

[12] STITES, Robert, WARD, Bryan and WALTERS, Robert, Defect Prediction With Neural Networks. In *Proceedings of the Conference on Analysis of Neural Network Applications ANNA '91*. Fairfax, VA, USA: ACM, 1991, pp. 199-206.

[13] TSANG, Kay F., LAU, Henry C. W. and KWOK, Seung K. Development of a data mining system for continual process quality improvement. *Proceedings of the Institution of Mechanical Engineers, Journal of Engineering Manufacture*. 2007. Part B, vol. 221, pp. 179-193.

[14] TSENG, Tsu L., JOTHISHANKAR, WU, Tong, XING, Guangming and JIANG, Fuhua. Applying data mining approaches for defect diagnosis in manufacturing industry. In *IIE Annual Conference and Exhibition*. Houston: Institute of Industrial Engineers, 2004, pp. 1441-1447.

[15] VAZAN, Pavel, TANUSKA, Pavol, KEBISEK, Michal, MORAVCIK, Oliver and SCHREIBER, Peter. Data Mining Model Building as a Support for Decision Making in Production Management. In *Advances in Computer Science, Engineering & Applications, Proceedings of Second International Conference on Computer Science, Engineering and Applications (ICCSEA 2012)*. 2012. New Delhi, India: Springer-Verlag Berlin Heidelberg, vol. 1, pp. 695-701.

[16] WANG, Kesheng. Applying data mining to manufacturing: the nature and implications. *Journal of Intelligent Manufacturing*. 2007. vol. 18, iss. 4, pp. 487-495.

[17] PERZYK, Marcin A., KOCHANSKI, Andrzej, KOZLOWSKI, Jacek and MYSZKA, Dawid. Control of ductile iron austempering process by advanced data driven modeling. In *71st WFC*: *71st World Foundry Congress*. Bilbao, Spain: World Foundry Organization, 2014.

[18] KUSIAK, Andrew and KURASEK, Christian. Data mining of printed-circuit board defects. *IEEE Transactions on Robotics and Automation.* 2001. vol. 17, iss. 2, pp. 191-196.

[19] TSENG, Tsu L., JOTHISHANKAR, WU, Tong, XING, Guangming and JIANG, Fuhua. Applying data mining approaches for defect diagnosis in manufacturing industry. In *IIE Annual Conference and Exhibition*. Institute of Industrial Engineers. Houston: Institute of Industrial Engineers, 2004, pp. 1441-1447.

[20] SHEN, Lixiang, TAY, Francis E. H., QU, Liangsheng and SHEN, Yudi. Fault diagnosis using Rough Sets Theory. *Computers in Industry*. 2000. vol. 43, no. 1, pp. 61-72.

[21] SADOYAN, Hovhannes, ZAKARIAN, Armen and MOHANTY, Pravansu, Data mining algorithm for manufacturing process control. *Journal of Advanced Manufacturing Technology*. 2006. vol. 28, no. 3-4, pp. 342-350.

[22] PERZYK, Marcin and KOZLOWSKI, Jacek. Methodology of Fault Diagnosis in Ductile Iron Melting Process. *Archives of Foundry Engineering*. 2016. vol. 16, iss. 4, pp. 101-108.

[23] PERZYK, Marcin and RODZIEWICZ, Agnieszka. Application of Time-series Analysis in Control of Chemical Composition of Grey Cast Iron. *Archives of Foundry Engineering*. 2012. vol. 12, no. 4, pp. 171-175.

[24] PERZYK, Marcin and RODZIEWICZ, Agnieszka. Application of Special Cause Control Charts to Green Sand Process. *Archives of Foundry Engineering*. 2015. vol. 15, iss. 4, pp. 55-60.

[25] RODZIEWICZ, Agnieszka and PERZYK, Marcin. Application of Time-Series Analysis for Predicting Defects in Continuous Steel Casting Process. *Archives of Foundry Engineering*. 2016. vol. 16, iss. 4, pp. 125-130.