Methodology of Fault Diagnosis in Ductile Iron Melting Process

M. Perzyk *, J. Kozlowski

Institute of Manufacturing Technologies, Warsaw University of Technology, Narbutta 85, 02-524 Warszawa, Poland
*Corresponding author. E-mail address: M.Perzyk@wip.pw.edu.pl

Abstract

Statistical Process Control (SPC) based on the Shewhart's type control charts, is widely used in contemporary manufacturing industry, including many foundries. The main steps include process monitoring, detection the out-of-control signals, identification and removal of their causes. Finding the root causes of the process faults is often a difficult task and can be supported by various tools, including data-driven mathematical models. In the present paper a novel approach to statistical control of ductile iron melting process is proposed. It is aimed at development of methodologies suitable for effective finding the causes of the out-of-control signals in the process outputs, defined as ultimate tensile strength (Rm) and elongation (A5), based mainly on chemical composition of the alloy. The methodologies are tested and presented using several real foundry data sets. First, correlations between standard abnormal output patterns (i.e. out-of-control signals) and corresponding inputs patterns are found, basing on the detection of similar patterns and similar shapes of the run charts of the chemical elements contents. It was found that in a significant number of cases there was no clear indication of the correlation, which can be attributed either to the complex, simultaneous action of several chemical elements or to the causes related to other process variables, including melting, inoculation, spheroidization and pouring parameters as well as the human errors. A conception of the methodology based on simulation of the process using advanced input - output regression modelling is presented. The preliminary tests have showed that it can be a useful tool in the process control and is worth further development. The results obtained in the present study may not only be applied to the ductile iron process but they can be also utilized in statistical quality control of a wide range of different discrete processes.

Keywords: Quality management, Application of information technology to the foundry industry, Process fault diagnosis, Ductile iron, Data-driven models

1. Introduction

In modern industry the Statistical Process Control (SPC) based on the Shewhart's control charts is widely used. The SPC methods assume 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. The main steps include process monitoring, detection the out-of-control signals, finding and removal of their causes.

Identifying the root causes of the process instability and providing the means for optimum control of the process is often a difficult task. Quality engineers and operational staff can be supported by various types of models linking the potential causes with the process outputs, particularly the product characteristics. Qualitative models, such as the Cause and Effect Diagrams (also known as Ishikawa or "Fishbone" diagrams) are widely used. However, because of the obviously limited capabilities of the qualitative models, various advanced data-driven models,

including Computational Intelligence methods, become more common in industrial practice (see e.g. [1] and publications cited there).

The classic SPC methods are commonly used also in foundry industry. The foundry technology covers a wide range of highly diversified processes, among which the melting process is one of the key issues deciding about quality of castings. One of the most advanced and demanding cast materials produced on a large scale is ductile iron.

The ductile iron properties are a result of complex, multistage process, in which melting plays a crucial role as it is responsible for chemistry of the base iron. The other main stages include inoculation, an spheroidizing treatment and pouring. The main control parameters in the melting process are: charge (weight of each component), temperature, heating time and chemistry of the base iron. The metallic charge usually include returns, steel scrap and high purity iron units, additionally recarburizers, ferroalloys and other silicon additions. Spectrometric analysis is the most widely used for checking the chemical composition, but the carbon and sulphur contents can only be taken as indicative. Thermal analysis and chill wedge samples may be also used, as complementary tests. A comprehensive characterization of ductile iron production process can be found in [2, 3].

In the present paper a novel approach to statistical control of ductile iron melting process is presented. It is aimed at development of methodologies suitable for effective finding the causes of the out-of-control signals in the process outputs, defined as ultimate tensile strength (Rm) and elongation (A5).

2. Research methodology and characteristic of data sets

2.1. Out-of-control signals definition and detection

The sequences of points, representing the out-of-control signals, included the 8 types of standard abnormal patterns, defined in Table 1. Some of them are defined using the notion of the three zones above and below the chart centerline. If samples of a certain size are taken from the production process and the points denote the sample means, the zones limits are usually expressed in terms of the standard deviation of the points from the centerline in a stationary process. However, in the present work the data were single measurements, therefore the borders of the zones must be calculated using the following quantity [6]:

$$S = \frac{\overline{MR}}{1.128} \qquad (1)$$

where $\overline{MR}$ is the moving range of two successive observations.

These three zones are typically denoted as: Zone A – the area between $2S$ and $3S$ above and below the center line; Zone B – the area between S and $2S$, and Zone C – the area between the center line and S.

For finding the standard patterns of points the especially programmed Excel spreadsheet, developed in a previous work [7], was used. It detects and counts all the occurrences of the patterns in a given data set, assuming that any adjacent appearances of point sequences of this same type is indicated and counted only once.

Table 1.
Standard abnormal patterns of points, representing out-of-control signals in a process (after [4, 5])

Pattern type	Definition	Comments
1	1 point beyond Zone A (outside control lines)	The most frequently studied case
2	9 consecutive points on one side of central line	The process average has probably changed
3	6 points in a row steadily increasing or decreasing	A signal of a drift in the process average
4	14 points in a row alternating up and down	Two systematically alternating causes are producing different results
5	2 out of 3 points in a row in Zone A or beyond	An 'early warning' of a process shift
6	4 out of 5 points in a row in Zone B or beyond	An 'early warning' of a process shift
7	15 points in a row in Zone C (above and below the center line)	A reduced variability of the process
8	8 consecutive points on both sides of the centerline with no points falling in zone C	Different points are affected by different factors, resulting in a bimodal distribution of measurements

2.2. Ductile iron melting process data sets

The main input variables for cast iron melting process are contents of the chemical components of the base iron. They are usually controlled in all foundries and in the present study it is assumed that they are the only available and recorded inputs of the process. In the cooperating foundry, the following 9 elements were tested: carbon (C), manganese (Mn), silicon (Si), phosphorus (P), sulphur (S), chromium (Cr), nickel (Ni), copper (Cu) and magnesium (Mg).

Three grades of ductile iron were produced: 400/18 (low strength and high elongation), 500/07 (high strength and lower elongation) and 500/07 special (with increased hardness). They were obtained by adjustment of the matrix structure through controlling the pearlite promoting chemical elements contents, mainly copper. However, it is worth noticing, that also other elements such as Ni, Mn, Cr, Si, P have this characteristic [2, 3] and therefore disturbances in their concentrations can also be a potential source of the process abnormalities.

In Table 2 the basic statistical characteristics of the five data sets utilized in the present study are presented. It is worth noticing that the conventional approach to detecting the abnormal patterns in SPC assumes that the observations are normally distributed around a constant mean. The Kolmogorov's normality test revealed that for the typical significance level equal 0.05 this assumption was not true for the data sets no 1, 4 and 5. However, as indicated in [5], also for highly skewed distributions all the usually considered abnormal patterns of points have small probabilities (less than 1%) and therefore it is reasonable to make decisions based on the appearance of such patterns also for non-normal distributions.

Table 2.
Characteristic of ductile iron production data sets used in the study

Set no.	Iron grade	No of records	Output variable	Normality test for output (p-value from Kolmogorov's test)
1	400-18	285	Rm	0,025
2	400-18	194	Rm	0,335
3	400-18	194	A5	0,076
4	500-07	161	Rm	0,010
5	500-07 spec	351	Rm	0,000

In Figs. 1 and 2 typical fragments of the run charts obtained from two exemplary data sets are shown. It can be seen that the variability of all variables (outputs and inputs) is different in character and magnitude and also depend on the data set (iron grade). The measurements exhibit various types of deviations from expected or desired values, resulting from the real foundry process in which various type of errors (human, equipment etc.) are inherent and unavoidable. For example, the copper contents in the low strength iron occasionally reaches 0.6%. The shape of the runs confirm a non-normal distribution of tensile strength of the 500-07 special grade and also indicate that some of the inputs (alloying components) seem to have bimodal rather than normal distributions. A characteristic shift in the average contents of some alloying elements with change of the iron grade can be also observed, particularly the pearlite promoting elements such as copper, manganese, nickel and chromium. It is also worth noticing, that the increase of the average contents of these elements for the high strength ductile iron is accompanied by a significant increase of their variabilities.

Finding the causes of abnormal behaviour of the process outputs can be often facilitated if relative significances of all outputs are known. The most significant input variables, i.e. those with the greatest influences on a given output, are usually the first candidates for being responsible for appearance of the out-of-control signal. Application of the significance analysis in production processes was a subject of some previous works [8,9]. The relative significance of the process input can be understood, defined and calculated in different ways. For the purpose of the present study the methodology based on one-way ANOVA (analysis of variance) was utilized [8], giving reasonable accuracies for the regression type relationships.

Fig. 2. Examples of typical run charts for the process inputs (chemical components contents of ductile iron) extracted from the data sets No 1 and 5
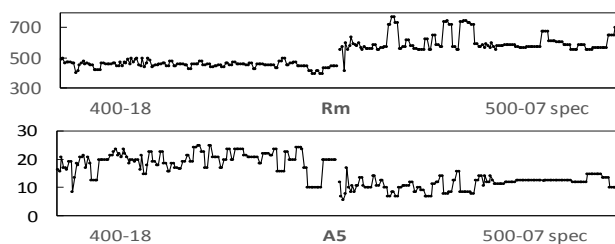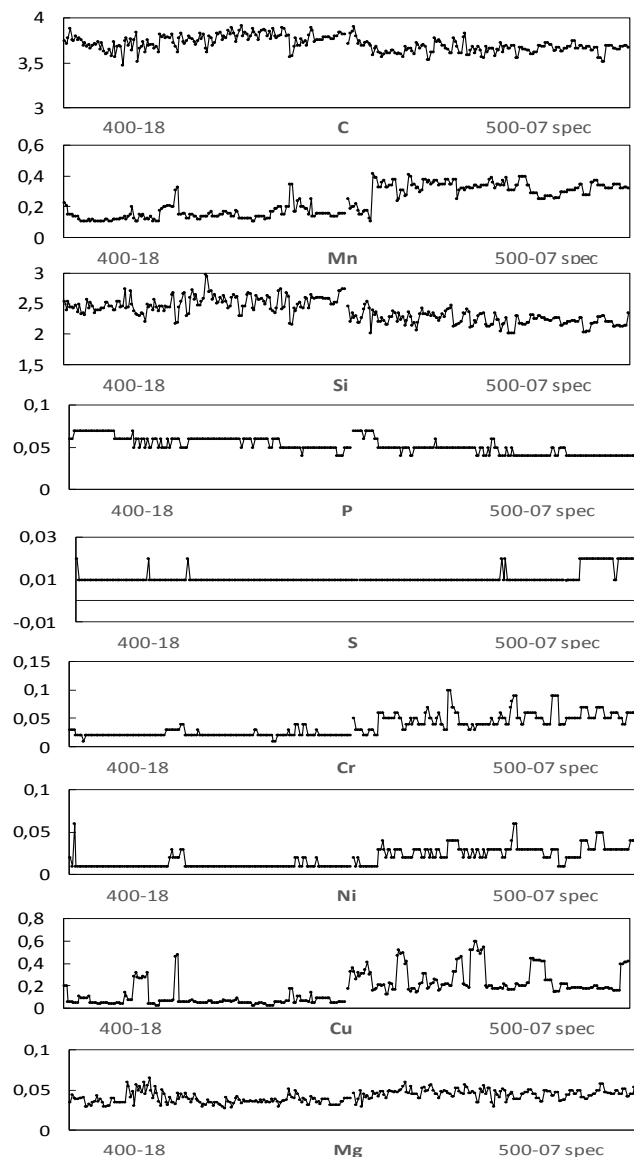

Fig. 1. Examples of typical run charts for the two process output variables (ductile iron ultimate tensile strength Rm and elongation A5) extracted from the data sets No 1, 3 and 5

The results are presented in Fig. 3. It can be observed that significances of individual inputs (chemical elements) are very different in different data sets. However, some of them seem to be generally low, e.g. for carbon and some seem to be generally important, e.g. for the pearlite promoting copper and chromium.
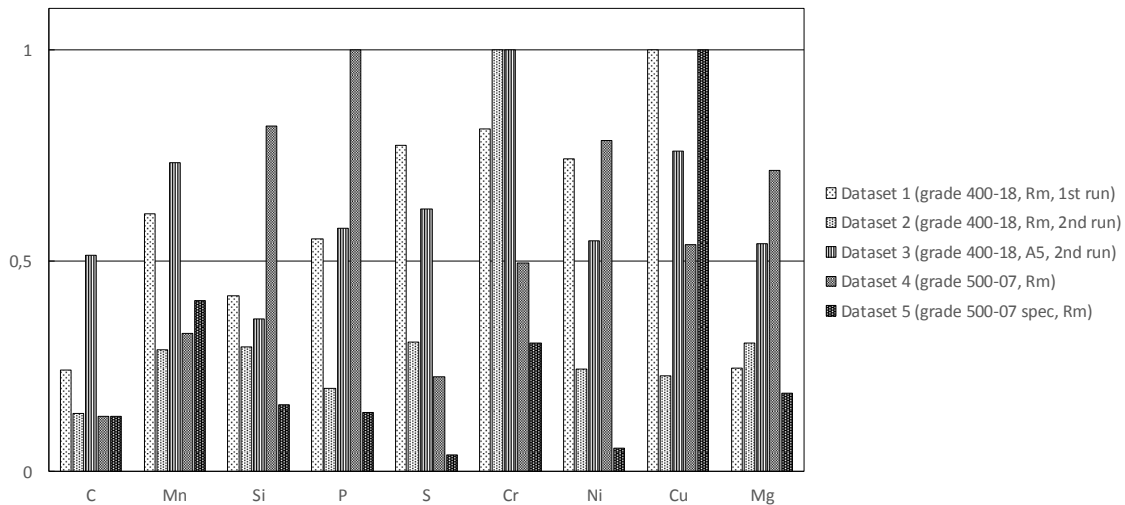

Fig. 3. Relative significances of the chemical components of ductile iron obtained for the five data sets defined in Table 2, using the ANOVA based methodology

It is worth noticing, that the significance of a given variable is not only a result of its physical potential but also depends on its variability range in a given data set.

3. Results of computations and analysis

3.1. Correlation between abnormal output patterns and inputs runs

A natural way of finding the root causes of an out-of-control signals in the process, appearing as one of the abnormal patterns of the points, is checking all the monitored inputs for appearance of similar patterns. In the present study all abnormal patterns of the types described in Table 1 were detected for the process outputs (ductile iron properties) as well as for all inputs (chemical components), for all data sets listed in Table 2. The results are given in Table 3 in the form of numbers of occurrences of the abnormal patterns. For the inputs only those occurrences are included which appeared in this same records as for the given output, i.e. which may be considered as an indication of the cause of the abnormal behaviour of the process output.

It can be seen that only abnormal patterns of the types 3 and 4 were not found in the data. The total number of the abnormal patterns appeared in the process output variables was 111, whereas only in 31 cases one or more inputs revealed the same type of pattern. It means, that in most cases finding the cause of the out-of-control signal in the process is not straightforward.

In Fig. 4 two examples of identification of probable causes of abnormal patterns in the process, based on numerically detected similar patterns of inputs, are shown. Here and in the following

graphs, the values shown on the vertical axis are the normalized values of the variables using the sample mean and deviation. The black rectangles mark the ranges of actual abnormal patterns. It should be noticed that those inputs which can be easily identified as responsible for the output abnormal signals (marked by broken lines) may have different overall significances in the corresponding data sets (copper has the maximum attainable value 1 whereas phosphorus only 0,55).

It is worth noticing that in most of such cases also shapes of the run charts confirm that the given inputs could be probable causes of the out-of-control signals. However, in some cases the abnormal pattern of the process output cannot be assigned to any input, in spite of the presence of the same type abnormal pattern detected for some inputs. An illustrative example of such situation is presented in Fig. 5 where the contents of the significant pearlite promoting elements are on opposite sides of the centreline.

Table 3.
Numbers of abnormal patterns in data sets used in the study

| Set no. | No of cases of abnormal patterns of points for: output \| output & arbitrary input \| output & neural model output | | | | | |
	Type 1	Type 2	Type 5	Type 6	Type 7	Type 8
1	0|0|0	5|1|0	4|1|0	9|4|1	2|0|0	1|0|0
2	0|0|0	6|2|1	1|0|0	1|0|0	5|0|1	0|0|0
3	2|0|1	10|5|1	2|0|0	3|1|1	6|1|0	0|0|0
4	3|2|1	5|2|1	2|0|1	0|0|0	1|0|0	0|0|0
5	7|1|0	7|3|2	6|2|0	16|6|1	5|0|0	2|0|0
Σ	12|3|2	33|13|2	15|3|1	29|11|3	19|1|1	3|0|0

Notes:
a) pattern types are defined in Table 1,
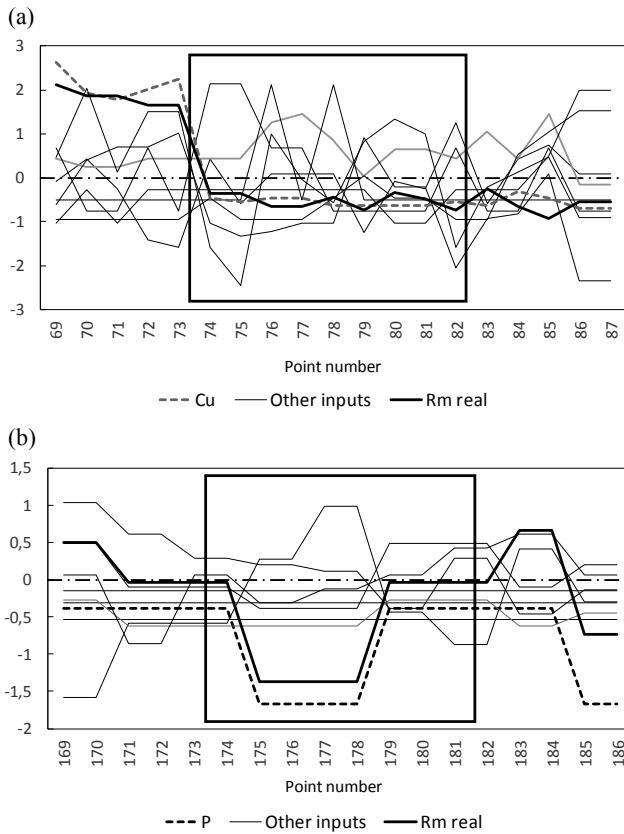b) neural model is described in Section 3.2

(a)

(b)

Fig. 4. Examples of run charts with abnormal patterns of the process output and inputs exhibiting the same types of patterns with clear correlations: (a) dataset no. 5, pattern type 2; (b) dataset no. 1, pattern type 6
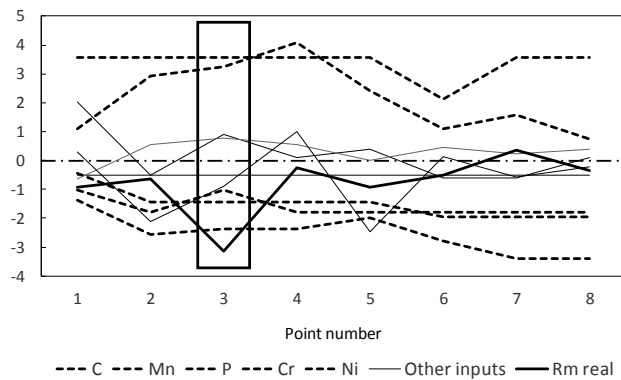

Fig. 5. Example of run charts with abnormal patterns of the process output and inputs exhibiting the same type of pattern without discernible correlation: dataset no. 5, pattern type 1 (the pearlite promoting elements are marked with dotted lines)

As shown in Table 3, most of the out-of-control signals in process outputs are not accompanied by the same type of abnormal patterns in an input. However, it was found that in many of such cases identification of the probable cause is easy by a visual inspection of the output and the inputs run charts. In Fig. 6

an example of that possibility is presented. In most cases the inputs designated as the probable causes of the out-of-control signals appearing in the process outputs are the inputs with high relative significances, e.g. the pearlite promoting chemical elements like copper, manganese and chromium.
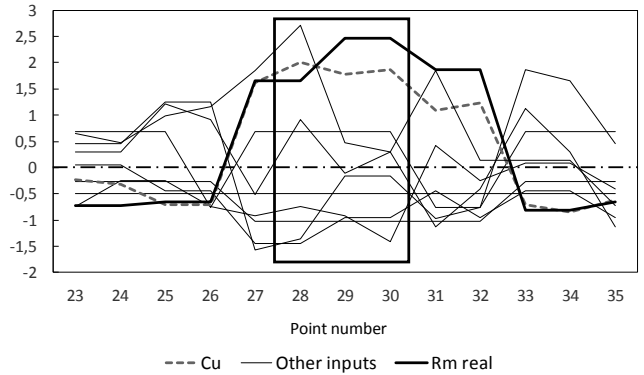

Fig. 6. Example of a run chart with abnormal pattern of the process output only and an input (copper) indicating noticeable correlation; dataset no. 5, pattern type 5

An interesting case is shown in Fig. 7, where the numerical analysis of the carbon contents reveals the abnormal pattern but there is another input (phosphorus) which is more likely to be the actual cause. It is also worth noticing that the relative significances of these two inputs are small (see Fig. 3).
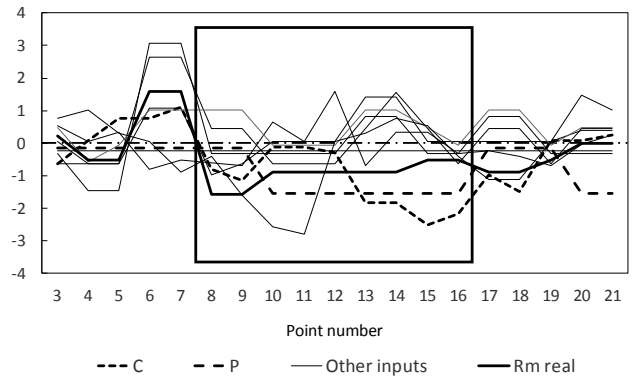

Fig. 7. Example of run charts with abnormal patterns of the process output, an input (carbon) exhibiting the same type of pattern but with questionable correlation with the output and another input (phosphorus) indicating better correlation; dataset no. 2, pattern type 2

A significant number of cases were found in which none of the inputs exhibits the abnormal pattern, neither determined numerically (as in Fig. 4) nor in a visual manner (as in Fig. 6). Some examples of such cases are presented in Fig.8. The exact number of such cases cannot be determined because of the obvious subjectivity in qualifying the input variations recognised in the run chart as correlated with the output. Examples of such dubious conclusions are illustrated in Fig. 7.
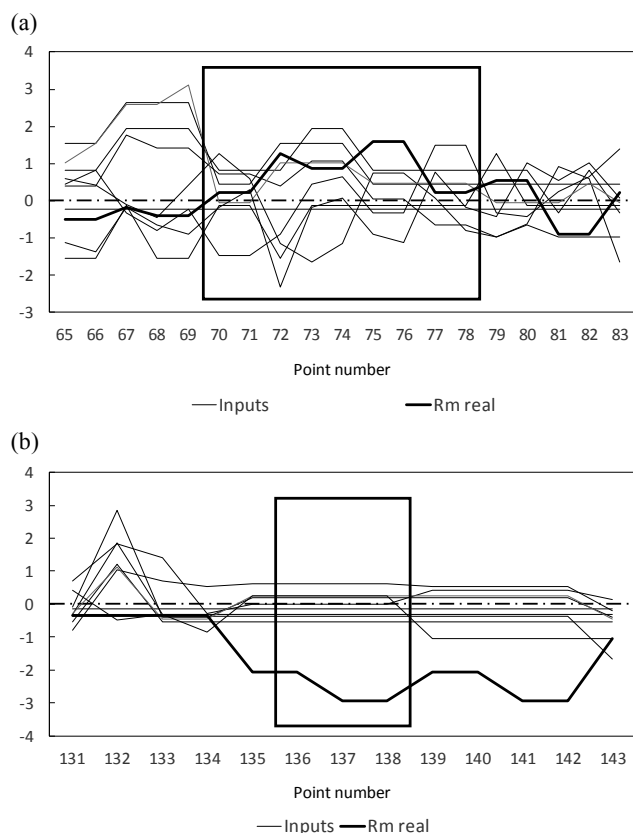
(a)

(b)

Fig. 8. Examples of run charts with abnormal patterns of the process output only and with absence of noticeable correlations: (a) dataset no. 2, pattern type 2; (b) dataset no. 1, pattern type 5

In Fig. 8 (a) the variability of all inputs within the output's abnormal pattern range is large and highly diversified. Here the diagnosis of the appearance of the output out-of-control signal is difficult due to the complexity of the inputs' effects, though it would be possible. In Fig. 8 (b) the situation is different: all inputs are very stable. In these cases assigning the out-of-control signals appearing in the process output to one or more process inputs would be definitely unjustified. It is very likely that the causes of the out-of-control signal in the process output are not bound up with the iron chemistry.

The situations like those presented in Figs. 7 and 8 (a), when the variations of the process inputs don't clearly point at the cause of the out-of-control signal appearing in the process output, seem to be quite frequent in industrial practice. Hence, it would be desirable to develop a mathematical tool which could help the operational staff and quality engineers in finding the cause. A novel conception of such tool is described in the next section.

3.2. Process diagnosis based on regression modelling

The idea of the proposed tool intended for supporting the diagnosis of the out-of-control signals in a manufacturing

processes is based on simulations of the process output variations due to changes of the process outputs, carried out using a multivariate regression model. It is built on the basis of the process inputs and output values recorded in a longer time interval. The general methodology of finding the abnormal pattern causes includes the following three steps.

Step 1. Calculate the process outputs' values from the model in the region of appearance of the abnormal pattern.

Step 2. Compare the model values with the real ones. If the model values also exhibit the same type of abnormal pattern or their run charts are at least similar, the conclusion can be made that the monitored process inputs are the possible causes of the out-of-control signal.

Step 3. Test the influence of the inputs on the output using the model and find those of them which could bring about the appearance of abnormal pattern of the output in the analysed case.
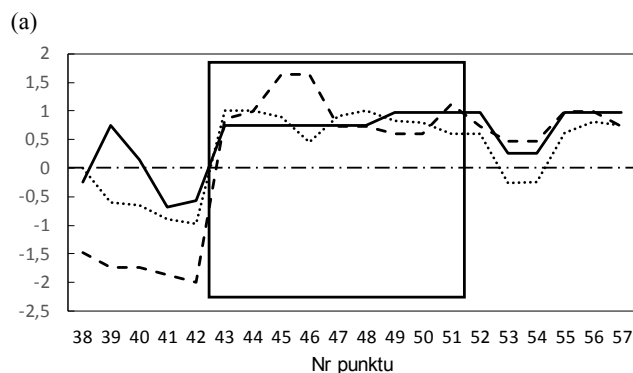
Obviously, the proposed methodology described in the above three steps is very general and requires a systematic development and refinement; some main issues are indicated it the final chapter of the paper. In the present work only some preliminary testing of the presented idea was carried out.

It should be noticed that the idea of application a regression model capable of predicting the expected values of the process output was also proposed and implemented by Zhang [10]. He used that type of modelling in situations when a product characteristic is achieved in several consecutive processes; the model inputs were values of the product characteristic obtained in the previous stages.
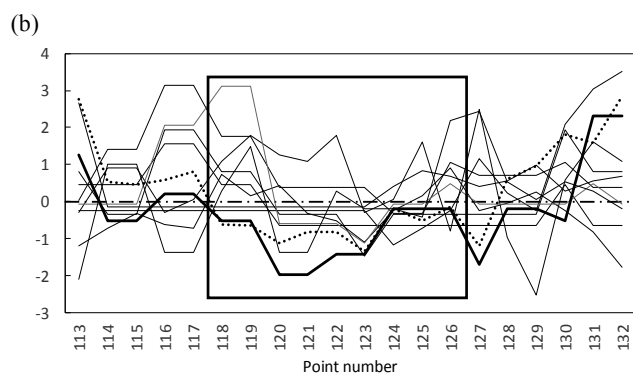
The regression neural models for all data sets listed in Table 2 were obtained using the Statistica ver. 12 package. One hundred MLP-type networks were built for each data set, using randomly selected number of hidden neurons and splitting the sets into training (70% of records), validating (15% of records) and testing (15% of records) subsets. The best neural model was chosen on the basis of the value of product of the three network quality indices obtained for these three subsets.

In Figs. 9 and 10 some examples of characteristic results obtained from the regression models are presented. In the two cases shown in Fig. 9 the abnormal patterns were observed also on the run charts for the values calculated from the models. As shown in Table 3, such cases are rather rare. However, in many cases the model values display the run charts shapes similar to those of the real values, though the abnormal patterns are not detected numerically. This fact is promising for successful implementations of the proposed methodology of finding the abnormal pattern causes described earlier in this section. Two such cases are shown in Fig. 10.

It is also worth noticing that in some cases, like those presented in Figs. 9 (a) and 10 (a), the corresponding abnormal pattern are observed also on the run charts for some process input. Such cases could be used for testing the model behaviour and the proposed methodology.
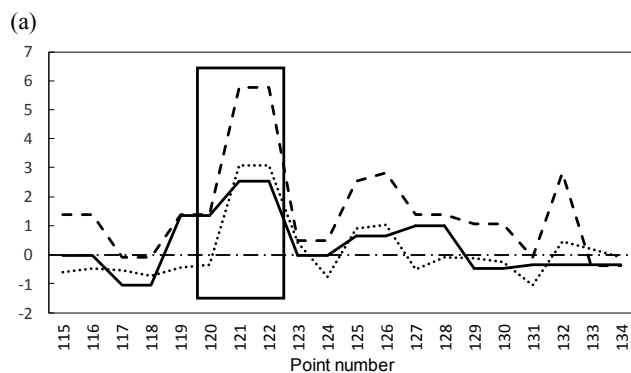
(a)

(b)

Fig. 9. Examples of run charts with abnormal patterns of the process output and the corresponding model values: (a) dataset no. 4, pattern type 2, observed also on one input (Mn); (b) dataset no. 2, pattern type 2, not observed on any input

(a)

(b)

Fig. 10. Examples of run charts with abnormal patterns of the process output, not appearing on the model run charts: (a) dataset no. 1, pattern type 5, observed also on one input (Mn); (b) dataset no. 1, pattern type 7, not observed on any input
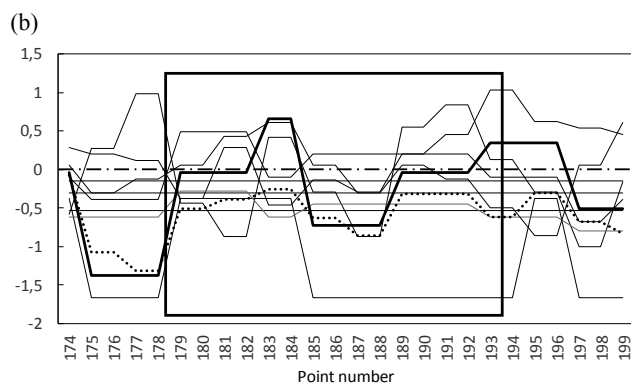
4. Summary, conclusions and future work

The present study revealed some new possibilities of identification of the root causes of manufacturing process faults resulting in appearance of the out-of-control signals on the process output run charts. Real production data collected in a cooperating iron foundry were used in the study and the conclusions from the analysis can be particularly useful in the ductile iron melting process fault diagnosis.

A limited usefulness of a visual inspection of the complex run charts including also process inputs runs was demonstrated and discussed. The main observation was that in many cases there is no clear indication which input is responsible for the out-of-control signal in the process output. It can be ascribed to the influence of the process variables which are not monitored but also to the simultaneous actions of several monitored input variables. This inspired the author to propose a novel methodology of finding abnormal pattern causes appearing on the process output run charts, utilizing an advanced regression model linking the process output with all the monitored inputs.

The preliminary testing of this methodology, in which neural models were used, revealed its substantial potential. However, in the current stage it is just a general conception and requires a systematic development. Especially the testing procedure mentioned in the third step (see Chapter 3) must be devised. Also, the criteria of similarity between the model and real values, required in the second step, should be developed.

References

[1] Perzyk, M., Kozlowski, J. & Zarzycki, K. (2013). Application of computational intelligence methods in control and diagnosis of production processes. In Kazmierczak J. (Ed.). S*ystems supporting production engineering*. 2013, 104-125.

[2] Sobczak J.J. (Ed.). (2013). Foundryman's Handbook (Poradnik Odlewnika). Kraków, Poland: Wydawnictwo Stowarzyszenia Technicznego Odlewników Polskich (in Polish).

[3] (2004). *The Sorelmetal Book of Ductile Iron*. Montreal (Quebec) Canada: Rio Tinto Iron & Titanium Inc.

[4] StatSoft, Inc. (2014). *STATISTICA (data analysis software system), version 12*. Retrieved April 13, 2016, from www.statsoft.com.

[5] Hoyer, R.W. & Ellis W.C. (1996). A Graphical Exploration of SPC Part 2: The probability structure of rules for interpreting control charts. *Quality progress.* 29(5), 57-64.

[6] Montgomery, D.C. (2010). *Statistical Quality Control: A Modern Introduction. (6th ed.)*. John Wiley & Sons, Inc.

[7] Perzyk, M. & Rodziewicz, A. (2015). Application of Special Cause Control charts to green sand process. *Archives of Foundry Engineering.* 15(4), 55-60.

[8] Perzyk, M., Biernacki, R. & Kozłowski, J. (2008). Data mining in manufacturing: significance analysis of process parameters. *Journal of Engineering Manufacture.* 222, 1503-1516.

[9] Perzyk, M. & Kochanski, A. (2003). Detection of causes of casting defects assisted by artificial neural networks. *Journal of Engineering Manufacture.* 217, 1279-1284.

[10] Zhang G. (1990). A new diagnosis theory with two kinds of quality. *Total Quality Management.* 1(2), 249-257.