Intelligence Analysis Method of Automation Control System Archive Database for controlling Hot Blast Stove Block

Oleksiy Koyfman [1 [0000-0003-2075-7417]], Oleksandr Simkin [2 [0000-0002-9939-7866]]
Kseniia Serdiuk [3 [0000-0002-4598-0562]]

Pryazovskyi State Technical University, str. Universytets'ka 7, Mariupol, 87555, Ukraine,

[1] koifman_o_o@pstu.edu, [2] simkin@pstu.edu
[3] serdiuk.xen@gmail.com

Abstract. In automated process control systems the data from the archive database are used in most cases in a graphic form for analyzing the technological process history. Frequently, the controlling of a hot blast stove block occurs in the conditions of incomplete information. The method of data mining of the archive database was developed and implemented in the form of computer programs to obtain new interconnections between the main technological parameters of the blast air heating process in a hot blast stove block. The developed algorithm allows to determine the functional dependence of the base parameter on two other parameters with fixing the third parameter in a certain range. The obtained analytical dependencies allow us to make decisions about changes in control algorithms for a hot blast stove block.

Keywords: intelligence analysis, data mining, method, approximation, hot blast stove, HBS, flow-chart, computer application

1 Introduction

Nowadays, all the basic technological processes of metallurgical production are automated by Automated Control Systems (ACS) of various levels of complexity, in which the technological parameters measured by the sensors, the values of the calculated variables enter the visualization systems and are stored in current databases. This information is used both for process control and for informing personnel about the process condition. Current databases are stored forming an archive database of the technological process (ADB). Unfortunately, in most cases ADB information is used to display the process flow in a graph form and, if necessary, to evaluate the operation of the unit in the past.

The software of ACS includes programs that implement mathematical models – informational or controlling. During the operation of the technological facility, such tuning coefficients of these programs must be adjusted due to aging or a change in the state of equipment and structural elements.

To solve this problem, some ACS are used data mining in recent years.

The goal of the research: the development and application of new methods of data mining to obtain new interconnections between the technological parameters of blast air heating in hot blast stove block (HBSB) in order to configure control algorithms for HBSB.

2 Formal problem statement

In many cases HBSB consists of three or four hot blast stoves (HBS). The block is designed to heat blast air (air enriched with oxygen up to 25-30%) to a temperature of 1000 - 1200 °C, which enters a blast furnace (BF).

HBS is a unit of discontinuous operation. The main two modes of operation are the checkerwork heating (on-gas period) and the blast heating (on-blast period). In on-gas period in the combustion chamber a fuel gas is burned (blast furnace gas (BFG) or a mixture of BFG and natural or coke gas). Then the hot combustion products heat the block refractory checkerwork, giving it the heat. In on-blast period, BF air passes through the hot checkerwork in the direction opposite to the combustion products, taking the heat from the checkerwork. In case of operation of HBSB, which includes 4 units, in sequential mode, three HBS operate in on-gas period, and one operates in on-blast period. As soon as one of HBS ends on-blast period, it is switched to the on-gas period, and the heated HBS is switched to on-blast period.

HBSB is controlled by an automation system that also includes a subsystem for maintaining a database of process parameters for each HBS. In modern ACS, there is a system for controlling the dome temperature, the fuel-air ratio and a monitoring system for a fuel gas flow, an air for combustion flow, a flue gases temperature, chemical analysis of the flue gases composition and fuel value.

In the studied system, there is a system for controlling the dome temperature and a system for monitoring the fuel gas flow and the temperature of the flue gases. Such number of parameters does not allow a clear understanding of process quality, i.e. in practice, ACS operates in the conditions of incomplete information.

Thus, the objective of the research aims at by using ADB that contains the above-mentioned minimum set of parameter values for the long-term operation of HBSB to find the interconnections between these parameters in order to their further use to correct the existing mathematical models of HBS operation and in block control algorithms.

3 Literature review

The article [1] describes various methods used in data analysis and approaches to the implementation of these methods.

Liu Y. et al. [2] use accumulated production data to compensate for the lack of technological information for the gas flow and conclude that the obtained sufficient accuracy of reproducing the missing data provides powerful support for the operation of gas blast furnace facilities.

The quality and accuracy of the technological information stored in the current and archive databases are very important for the decision-making process in order to save resources. Unfortunately, due to design errors, instability and vulnerability of the Supervisory Control and Data Acquisition (SCADA), these anomalies usually exist in practice. This usually leads to a negative impact on the accuracy and stability of data-driven methodologies. The authors of [3] proposed two methods for searching for anomalies in the data presented in the current database.

Methodology, problems and solutions in the field of using multivariate polynomial regression are described in [4]. According to the authors, descriptive data analysis is a process of evaluating values based on a given data set, which can be achieved by regression analysis on this data set. In real time, data can include various data sets with different properties. Predictive (intelligent) Data Mining is the process of evaluating or predicting future values from an available set of values, uses the concept of regression to evaluate future values.

An example of the use of quadratic multiparametric polynomial regression is [5], where the authors performed an analysis to determine the relationships between the parameters (gas temperature, wall thickness) of heat exchange during drying with infrared convection. By analyzing the dispersion, the authors evaluate the order of influence of the significance of all factors on the temperature of hot air.

As indicated above, both the blast furnace itself and the hot blast stoves are units operating in conditions of incomplete information. For such objects, according to various authors, the only real opportunity for implementing management functions is data analysis.

The standard approach to the analysis of BF process is based on a detailed study of its individual phenomena (heat and mass transfer, gas and hydrodynamics, etc.) with the subsequent use of the identified patterns for determination of possible and desired trends in the development of processes and management. The authors of [6] proposed a thermochemical model of BF smelting, allowing the operator to apply corrective actions, based on an analysis of data from an archive database for 6 months. The calculation results obtained on the basis of the model were compared with real data.

The success of cluster data analysis primarily depends on the controllability of BF processes based on existing sets of measured parameters and control factors [7]. The uncertainty of the control of BF process is determined by uncontrolled parameters, which determine the spread of the values of the output parameters of the process. An analysis of the controllability of BF processes, carried out in this work on the basis of existing statistics of BF smelting, shows that setting specific values of control factors determine the productivity of BF with a certain accuracy. This fact justifies the management technique proposed in this paper. It is based on real data as applied to the calculation of the theoretical combustion temperature.

According to the authors of [8], measurements of technological parameters of BF smelting usually have non-Gaussian characteristics. This problem can be solved if we do not stick to the analytical solution, but use the empirical method to determine the control limit. In the general case, if the distribution of the measurement data is unimodal and the detection index value is constant for the data under normal conditions. To diagnose disturbances in the operation of BF, the authors of the article used multivariate statistical analysis, the purpose of which was to monitor disturbances associated with switching HBS to detect anomalies.

To monitor the process of cast iron production in BF, the authors of [9] use the Principal Component Analysis (PCA) method of analysis of the main components - a statistical procedure that orthogonally converts the initial *n* coordinates of the data set into a new set of *n* coordinates called the main components. The work shows that the distribution is not normal for the following reasons: peak-like interference caused by HBS switching; periodic loading of coke and tapping of cast iron, etc. Before detecting abnormalities, outliers are removed from the original database, the size of the training sample and the number of main components are selected.

In article [10] an approach is proposed to control BF process to increase the efficiency of process, based on modeling the operation of BF and on-line stabilization of parameters on the basis of the obtained mathematical models. Using machine learning methods, dependencies are determined to specify the controlled parameters at the output and coke speed. Intelligent data analysis for BF operation was performed to identify models of distribution and subsequent melting of materials. Based on the developed models an expert system was developed and implemented.

In practice, operators of BF make decisions on controlling the unit based on a visual assessment of the graphic profiles of key variables, and not through a complex calculation of the exact numerical values. Zhou, B. et al. [11] proposed a new qualitative trend analysis algorithm (QTA) based on piecewise linear polynomial approximation. Using the first and second derivatives, the rate of change and the direction of the trend are determined. The effectiveness of the algorithm is shown by testing on various models and real data of blast furnaces.

An analysis of the existing quality control systems performed by the author [12] showed that currently there are practically no systems that could be used for operational control of the dynamics of the quality of cast iron during the smelting process in real time. In this regard, the article presents a system for collecting and analyzing data on the controlled parameters of the BF process with the aim of real-time monitoring of the quality of cast iron during smelting. Data on the current values of the technological parameters are sent to a server from the controllers every 3 seconds and are smoothed using an exponential algorithm with the choice of the optimal parameter.

The work [13] describes the operation of a BF automation system based on the use of an archive database in the mode of an operator advisor. The author outlines a methodology for analyzing and optimizing the volume of the operational database in order to predict the course of BF smelting in real time and offers a method for analyzing data of the technological parameters, which allows you to obtain additional information about the smelting process. When designing the system, hierarchical clustering (the Lance-Williams method) was used, which allows them to be found in the process of performing the clustering procedure.

The main goal of [14] is to show the potential of modern intelligent technologies based on self-organizing neural networks for clustering operating modes and model-predictive control of the BF process. The method uses a prediction model of control technology. Moreover, the construction of a model of the BF process includes data obtained in real time. The authors propose to use two approaches to clustering the effective values of parameters using the criteria for the efficiency of a blast furnace: using elliptical surfaces and a self-organizing Kohonen network.

A sufficient amount of research also applies to the use of modeling, solving optimization problems, multivariate statistical analysis, cluster analysis and data mining for the operation of automation systems for HBS.

For example, Nose, K. et al. [15] formulated and solved using modeling the non-linear optimization problem by using the generalized reduced gradient (GRG) method.

For a HBS control system the authors of [16] proposed a sequential procedure for improving process control based on multivariate methods of statistical analysis. In order to take advantage of a large amount of historical data, a combination of the hierarchical clustering method and statistical process control methods is used to detect and analyze key factors that significantly affect process performance. This technique consists of four successive stages: data collection and multivariate statistical analysis, hierarchical clustering and detection of the operating mode, the selection of dominant variables, a new operational benchmark and its verification.

The authors of [17] confirm that with intelligent data processing the Six Sigma Methodology increases the processing speed and the quality of the “gained” data.

4 Statement of Basic Materials

The procedure for obtaining the desired information from the archive database is shown in Fig. 1.

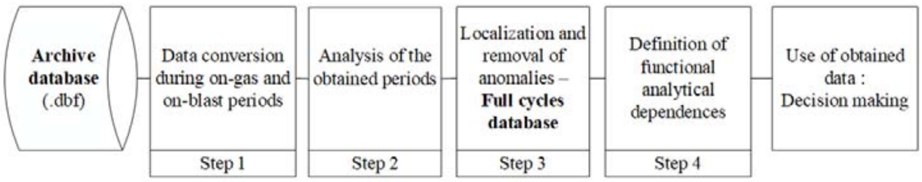

Fig. 1. Stages of information conversion

At one of the iron and steel plants of Ukraine, the values of 194 technological parameters of HBS in ACS for cast iron smelting are recorded in ADB (file with the extension * .dbf) with a frequency of 5 seconds, and the file is generated for every 24 hours and, accordingly, one file contains 3352320 records.

The ADB file structure is shown in Fig. 2, where Date and Time columns correspond to the date and time the parameter was written to the database, TagIndex is the index corresponding to the technological parameter, Value is the value of the technological parameter; Tagname – the name of the technological parameter in the database.

The studied ADB has a form that is inconvenient for its further analysis and use due to its structure (outdated recording format) and accordingly due to its size.

	A	B	C	D
1	Date	Time	TagIndex	Value
19	26.02.2019	00:00:11	17	0,000
24	26.02.2019	00:00:11	22	1,000
26	26.02.2019	00:00:11	24	0,000
29	26.02.2019	00:00:11	27	23982,777
42	26.02.2019	00:00:11	40	122,300
49	26.02.2019	00:00:11	47	1240,300
52	26.02.2019	00:00:11	50	0,000

a)

	A	B
1	Tagname	TTagIndex
19	VN1\D	17
24	VN1\N	22
26	VN1\O	24
29	VN1\RSG	27
42	VN1\TD	40
49	VN1\TKF	47
52	VN2\D	50

b)

Fig. 2. ADB structure: a) file with values, b) tag table

Therefore, firstly, it is necessary to transform the database into a form convenient for further analysis. The authors developed and implemented an ADB processing algorithm taking into account the following provisions:

1. The operation of each HBS will be described with the characteristics (values of technological parameters) of two periods – on-gas and on-blast including:

a) for on-gas period: time of the beginning and end of the period (H_G_{beg}, M_G_{beg}, S_G_{beg}, H_G_{end}, M_G_{end}; S_G_{end}); the duration of the period; the beginning and end values of the dome temperature (TDG_HBS1_{beg}, TDG_HBS1_{end}); the beginning and end values of the flue gases temperature (TFG_HBS1_{beg}, TFG_HBS1_{end}); the average integral value of the dome temperature (TDG_HBS1_{avg}); the average integral value of the flue gases temperature (TFG_HBS1_{avg}); dome heating rate; the rate of heating of the flue gases; the beginning and end consumption of the BFG (FG_HBS1_{beg}, FG_HBS1_{end}); volume of the consumed BFG (VG_HBS1); the average integral flow rate of the BFG during the period (FG_HBS1_{end});

b) for on-blast period: time of the beginning and end of the period (H_B_{beg}, M_B_{beg}, S_B_{beg}, H_B_{end}, M_B_{end}; S_B_{end}); the duration of the period; the beginning and end values of the dome temperature (TDB_HBS1_{beg}, TDB_HBS1_{end}); the beginning and end values of the temperature of the blast air (TBB_HBS1_{beg}, TBB_HBS1_{end}); the average integral value of the dome temperature (TDB_HBS1_{avg}); the average integral value of the blast temperature at the inlet to HBS (TBB_HBS1_{avg}); the cooling rate of the dome temperature during the period.

2. Divide ADB into 8 parts (files), each of which will contain data on on-gas and on-blast periods for each HBS of the block.

Let us consider in details the proposed algorithm (Fig. 3) for converting the ADB into a form convenient for analysis of HBS #1 in particular.

When the program is initialized, a directory in which the ADB files are located and a directory in which the processing results will be written are selected.

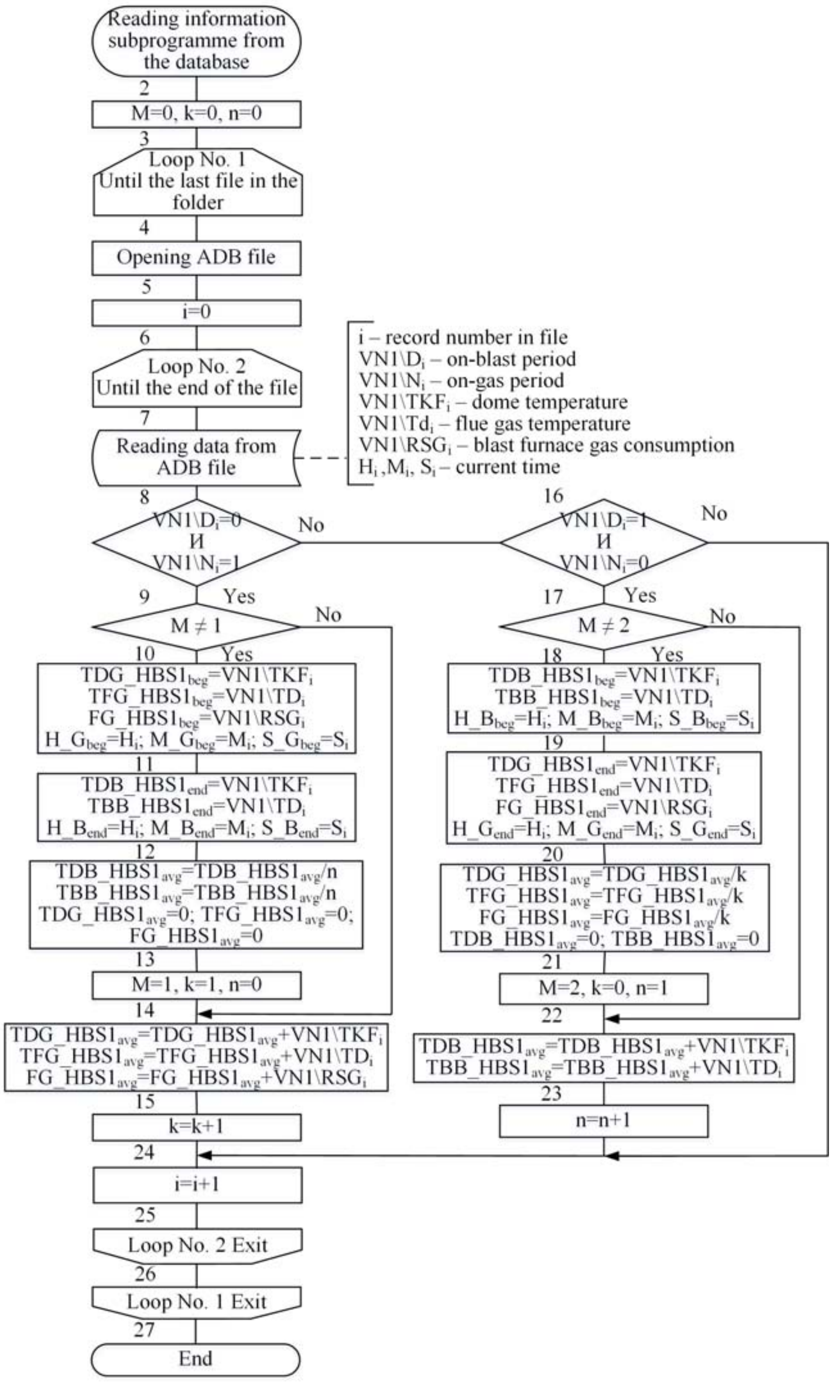

Fig. 3. Flow-chart of algorithm of ADB conversion in a form convenient for analysis

Step 1. Beginning of reading of information subprogramme from the database.

Step 2. Resetting variables: M – mode of operation of HBS, $M = 1$ – on-gas period, $M = 2$ – on-blast period, n – number of parameters values recorded over on-gas period; k – number of parameters values recorded over on-blast period.

Step 3. Loop No. 1 starting, it operates until all ADB files from the folder are processed.

Step 4. Opening the next ADB file.

Step 5. Resetting variable: i – number of records in a file.

Step 6. Loop No. 2 starting, it operates until the end of a file is reached.

Step 7. Reading data from ADB file using SQL-queries: i ; $VN1\backslash D_i$ – on-blast period; $VN1\backslash N_i$ – on-gas period; $VN1\backslash TKF_i$ – dome temperature; $VN1\backslash TD_i$ – flue gas temperature; $VN1\backslash RSG_i$ – BFG flow.

Step 8. Mode checking: $VN1\backslash D_i = 0$ AND $VN1\backslash N_i = 1$? If YES – on-gas period, transition to step 9, if NOT – transition to step 16.

Step 9. Mode change checking: $M \neq 1$? If YES – mode changing from on-blast period to on-gas period, transition to step 10; if NOT – transition to step 14.

Step 10. Assigning initial values for on-gas period beginning: TDG_HBS1_{beg} , TFG_HBS1_{beg} , FG_HBS1_{beg} , H_G_{beg} , M_G_{beg} , S_G_{beg} .

Step 11: Assigning finite values for on-blast period end: TDB_HBS1_{end} , TBB_HBS1_{end} , H_B_{end} , M_B_{end} , S_B_{end}

Step 12. Calculating for on-blast period: TDB_HBS1_{avg} , TBB_HBS1_{avg} .

Resetting for on-gas period: TDG_HBS1_{avg} , TFG_HBS1_{avg} , FG_HBS1_{avg} , VG_HBS1 .

Step 13: Assigning initial values: $M = 1$ – on-gas period; $k = 1$ – beginning of new on-gas period; $n = 0$ – end of on-blast period.

Step 14. Calculating: TDG_HBS1_{avg} , TFG_HBS1_{avg} , FG_HBS1_{avg} , VG_HBS1 .

Step 15. Incrementing number of parameters values $k = k + 1$.

Step 16. Mode checking: $VN1\backslash D_i = 1$ AND $VN1\backslash D_i = 0$? If YES – on-blast period, transition to step 17, if NOT – transition to step 24.

Step 17. Mode change checking: $M \neq 2$? If YES – mode changing from on-gas period to on-blast period, transition to step 18; if NOT – transition to step 22.

Step 18. Assigning initial values for on-blast period beginning: TDB_HBS1_{beg} , TBB_HBS1_{beg} , H_B_{beg} , M_B_{beg} , S_B_{beg} .

Step 19. Assigning finite values for on-gas period ending: TDG_HBS1_{end} , TFG_HBS1_{end} , FG_HBS1_{end} , H_G_{end} , M_G_{end} ; S_G_{end} .

Step 20. Calculating for on-gas period: TDG_HBS1_{avg} , TFG_HBS1_{avg} , FG_HBS1_{avg} .

Resetting for on-gas period: TDB_HBS1_{avg} , TBB_HBS1_{avg} .

Step 21. Assigning initial values: $M = 2$ – on-blast period; $k = 0$ – ending of new on-gas period; $n = 1$ – beginning of on-blast period.

Step 22. Calculating: TDB_HBS1_{avg}, TBB_HBS1_{avg}.

Step 23. Incrementing number of parameters values $n = n + 1$.

Step 24. Incrementing number of records $i = i + 1$.

Step 25. Exiting from the loop No. 2.

Step 26. Exiting from the loop No. 1.

Step 27. End.

The results of processing ADB are written to an Excel file (Fig. 4). Each row of eight tables represents one on-gas period or one on-blast period.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Period Start Date	Period End Date	Period Duration	Period Duration (Mins)	Beg Dome Temp	End Dome Temp	Avg Dome Temp	Dome Temp Change Rate	End BFG Flow Rate	Avg BFG Flow Rate	Vol BFG Flow Rate	Beg FlueG Temp	End FlueG Temp	FlueG Temp Change Rate
137	1.3.20 14:30:53	1.3.20 20:06:39	6:35:46	335,8	1206	1333	1297	23	44150	35240	163543	94	401	55
138	1.3.20 22:40:40	1.4.20 6:49:02	8:08:22	488,4	1242	1339	1292	12	43301	34291	212925	123	401	34
139	1.4.20 9:37:35	1.4.20 14:43:57	5:06:22	306,4	1222	1274	1248	10	41793	31169	93918	128	401	54

Fig. 4. ADB conversion results

The next second processing step is the analysis of the obtained periods and the comparison of on-gas periods with the corresponding subsequent on-blast periods.

The following ADB processing results were obtained, which included 7 months of operation of HBSB in on-gas periods / on-blast periods: for HBS1 696 / 761; for HBS2 773 / 763; for HBS3 306 / 718; HBS4 207 / 237. For the next 3 months: for HBS1 155 / 152; for HBS2 166 / 167; for HBS3 59 / 81; for HBS4 173 / 173. For example, inconsistencies in the number of periods can be in the case when the recording of the technological parameters values in the ADB is incorrect (absence or failure of the sensor) and there is a mismatch between the values of the tags responsible for the periods and these periods.

In addition the operating values of time of on-gas and on-blast periods are known from the operating map of HBS block, so periods that are known to be small or long in duration are initially removed.

Then a search is made for on-blast periods, which correspond to the current on-gas period. After processing the following number of full cycles was obtained: for HBS1 – 641, for HBS2 – 736, for HBS3 – 135, for HBS4 – 201. Thus, for example, for HBS1 there were 851 entries in ADB, and only 641 full cycles fell into the table, i.e. approximately 25% of records are filtered.

At the third processing step, localization and removal of anomalies of the obtained values for each of the columns is performed using descriptive statistics: calculation of the maximum and minimum values, average values, standard deviation (σ): elements that do not belong to the range $[-3\sigma;3\sigma]$ are removed from the selection. It can be seen from Fig. 5 that the initial distribution of the parameter values is processed in such a way that the resulting distribution is close to normal.

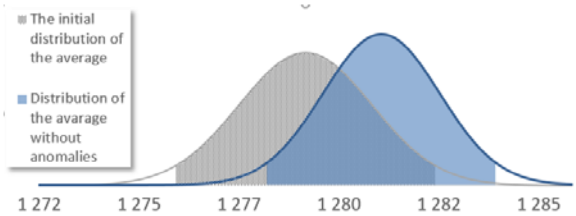

Fig. 5. Result of processing in steps 3

As a result, the final version of the database was obtained with a breakdown into full HBS operation cycles database (hereinafter referred to as FCDB).

In the fourth processing step we get the data from FCDB. At this stage, the opportunity is realized using FCDB and the algorithm developed by the authors (Fig.6) to determine the functional dependence of the base parameter on two other parameters with fixing the third parameter in a certain range.

Below is an algorithm that implements this possibility, using the example of obtaining the dependence of on-gas duration τ_{dur} (Z) on the average dome temperature TDG_HBS_{avg} (X) and the volume of consumed BFG VG_HBS (Y) within the range of BFG flow FG_HBS (F) from $FG\_HBS1_{\min}$ to $FG\_HBS1_{\max}$ ($F_{\min}$ to $F_{\max}$):

$$Z = f(X,Y) \; within \; F \in [F_{\min}, F_{\max}]. \tag{1}$$

Let us consider in details the proposed algorithm (Fig. 6):

Step 1. Beginning of the subprogram.

Step 2. Opening FCDB file.

Step 3. Reading k - the number of records in FCDB file.

Step 4. Reading data from FCDB file to X, Y, Z, F.

Step 5. Calculating the values of average, minimum, maximum, standard deviation for X and Y arrays: $X_{\min}$, $X_{\max}$, X_{avg}, X_{ds}, $Y_{\min}$, $Y_{\max}$, Y_{avg}, Y_{ds}.

Step 6. Input $F_{\min}$, $F_{\max}$.

Step 7. Input range values in which the X and Y data will be processed: σ_{beg}, σ_{end}, σ_{step}. By default : $\sigma_{beg} = -2$, $\sigma_{end} = 2$, $\sigma_{step} = 1$.

Step 8. Assigning initial values $\sigma = \sigma_{beg}$, $m = 0$.

Step 9. Loop No. 1 starting, it operates until $\sigma \le \sigma_{end}$.

Step 10. Calculating values of $convX_m$ and $convY_m$ arrays.

Step 11. Incrementing σ and m.

Step 12. Exiting from the loop No. 2

Step 13. Resetting $Xrez$ array.

Step 14. Loop No. 2 starting, it operates until $i \le m-1$.

Step 15. Resetting $Yrez$ array elements to zero.

Step 16. Loop No. 3 starting, it operates until $j \le m-1$.

Step 17. Resetting $Zrez$ array elements to zero; n=0.

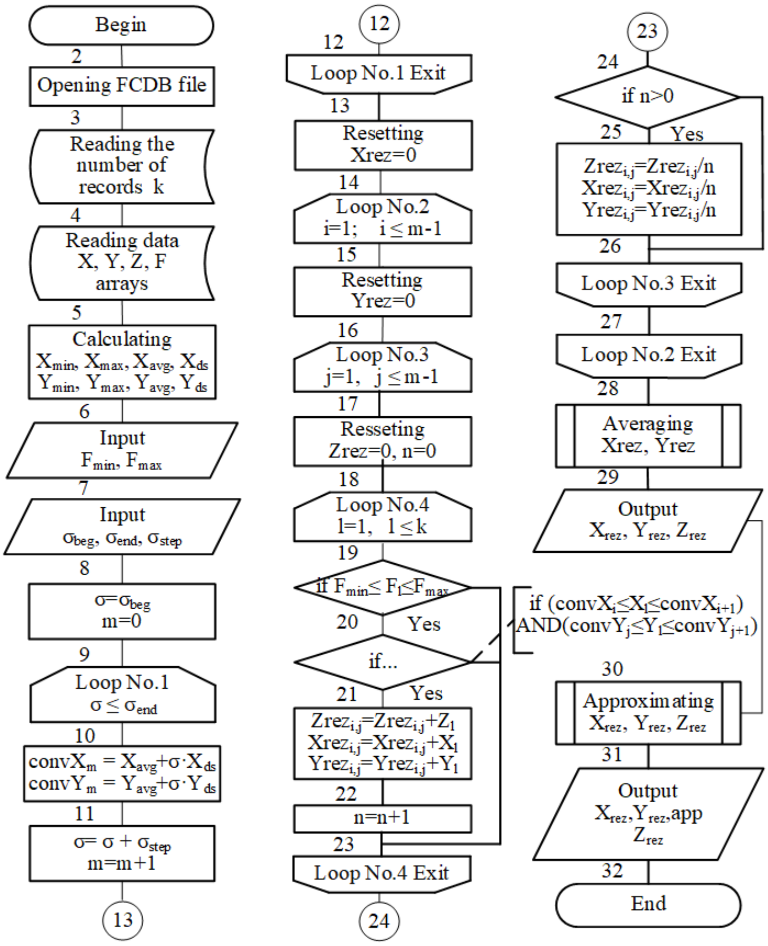

Fig. 6. Flow-chart of algorithm for determining the functional dependence

Step 18. Loop No. 4 starting, it operates until $l \le k$.

Step 19. Checking condition: $F_{\min} \le F_l \le F_{\max}$? If YES – transition to step 20, if NOT – transition to step 23.

Step 20. Checking condition:

if $\left(\left(convX_i \le X_l \le convX_{i+1}\right)AND\left(convY_j \le Y_l \le convY_{j+1}\right)\right)$? If YES – transition to step 21, if NOT – transition to step 23.

Step 21. Calculating: $Zrez_{i,j}$, $Xrez_{i,j}$, $Yrez_{i,j}$.

Step 22. Incrementing n .

Step 23. Exiting from the loop No. 4

Step 24. Checking condition: if $n > 0$? If YES – transition to step 25, if NOT – transition to step 26.

Step 25. Calculating of average values: $Zrez_{i,j}$, $Xrez_{i,j}$, $Yrez_{i,j}$.

Step 26. Exiting from the loop No. 3

Step 27. Exiting from the loop No. 2

Step 28. Averaging *Xrez* and *Yrez* values – getting a one-dimensional array from a two-dimensional array.

Step 29. Output *Xrez* , *Yrez* , *Zrez* .

Step 30. Regression analysis of the obtained results *Xrez* , *Yrez* , *Zrez* with multivariate approximation.

Step 31. Output *Xrez* , *Yrez* , *appZrez* .

Step 32. End.

5 Assessment of the research results

Based on the above algorithms, the following programs have been developed: for the implementation of the first and second stages (Fig. 1); for the implementation of the third and fourth stages (Fig. 1).

Below are the results of the programs.

To analyze the operation of HBS #4, the dependence (1) was obtained (Fig.7) with the following samples: $[-3\cdot\sigma; 3\cdot\sigma]$, $step\,\sigma$ (Fig. 7, a, b); $[-2{,}5\cdot\sigma; 2{,}5\cdot\sigma]$, $step\ 0{,}5\sigma$ (Fig. 7, c, d, e); $[-2{,}5\cdot\sigma; 2{,}5\cdot\sigma]$, $step\ 0{,}25\sigma$ (Fig. 7, f, g, h).

It can be seen from the graphs that when using a step of $0{,}5\sigma$ and especially $0{,}25\sigma$ (Fig. 7, c, f), the problem of the lack of input data in FCDB that simultaneously satisfy the set of specified parameters values (dome temperature and gas volume) arises. This is determined by the presence of zero values of heating time. In this case, the program replaces these points using an approximation based on nonzero values, which is visually seen by the presence of flat sections on the surface of the graphs (Fig. 7, d, g).

Thus using an interval equal to σ eliminates the problems that arise in the absence of values in the database (Fig. 7, a).

It should be noted that the most optimal sampling range is the range from -2σ to 2σ with an interval equal to σ .

The obtained results (Fig 7, a, d, g) in terms of the totality of values are unsuitable for subsequent use in control algorithms of HBS. Using the multi-parameter approximation, we obtain these dependencies in the form of analytical expressions ($R^2 > 0.8$), the graphs of which are shown in Fig. 7 (b, e, h). Examples of such expressions are given in Table 1.

Earlier [18], the authors developed and implemented a mathematical model of the thermal operation of HBS of BF, which was identified on the basis of ADB used in this research. For this purpose, the averaged basic operation parameters of each HBS of the block ([18], Fig. 4) were used. Thus, for example, for HBS4, the average dome temperature corresponds to 1275 °C, and the average on-gas period duration is 230 minutes, which fully corresponds to the calculated value in accordance with the analytical expression (Tab. 1).

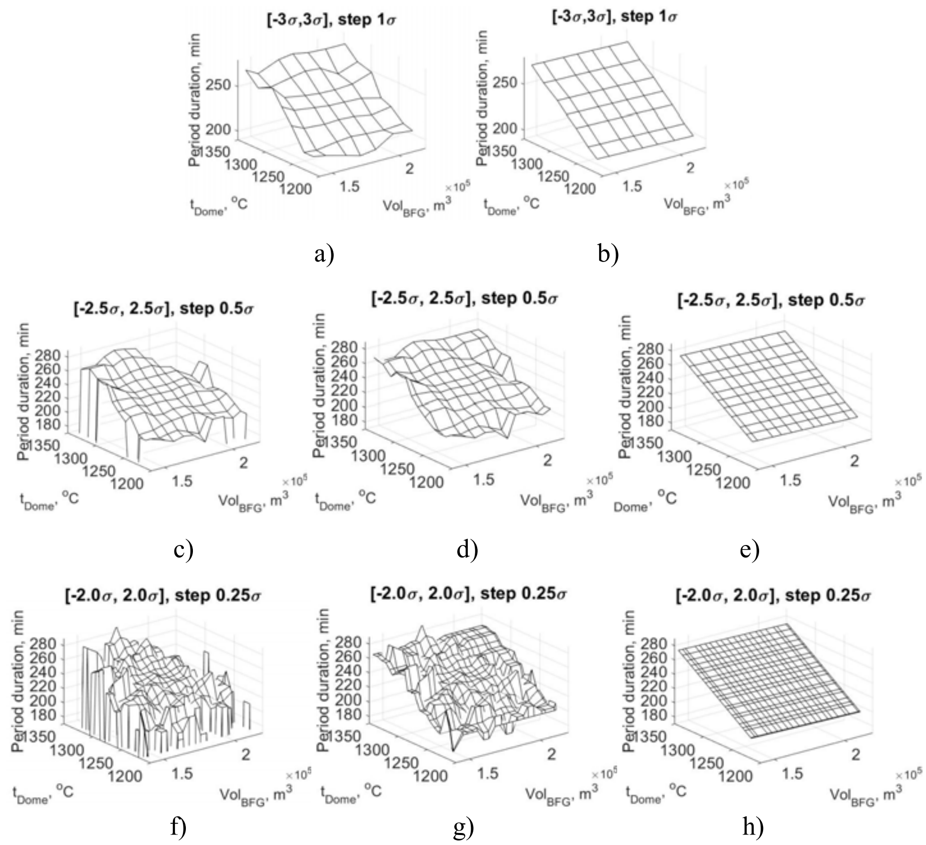

Fig. 7. The dependence (1) for HBS #4 for different range of σ

Table 1. Analytical expressions

Range	Analytical expressions	R^2
$[-3\cdot\sigma;3\cdot\sigma]$, *step* σ	$\tau=-0.0001\cdot V_{BFG}+0.5\cdot t_{Dome}-471.3$	0.8939
$[-2{,}5\cdot\sigma;2{,}5\cdot\sigma]$, *step* $0{,}5\sigma$	$\tau=-0.0001\cdot V_{BFG}+0.6\cdot t_{Dome}-511.6$	0.8578
$[-2{,}5\cdot\sigma;2{,}5\cdot\sigma]$, *step* $0{,}25\sigma$	$\tau=-0.00005\cdot V_{BFG}+0.6\cdot t_{Dome}-534.3$	0.8281

Fig. 8 graphically presents the dependences of on-gas duration on the average dome temperature and the BFG volume for various average BFG flow for each HBS of the block obtained as a result of the research.

Such data presentation (Fig.8) makes it possible to change in real time the parameters of the mode map of the blast heating of HBSB in case of production need or significant changes in the production situation.

6 Conclusion

1. The methods of data mining and their use, including in control systems for blast furnace production, which allows one to extract new opportunities from a minimum set of values of technological parameters, have been analyzed. The authors concluded that there is no information on the use of data mining in automation systems of HBSB.

2. For the first time, a method and algorithm for processing an archive database which includes 4 stages is proposed.

3. The specified algorithm is implemented in the form of two software modules developed using Visual Studio 2019.

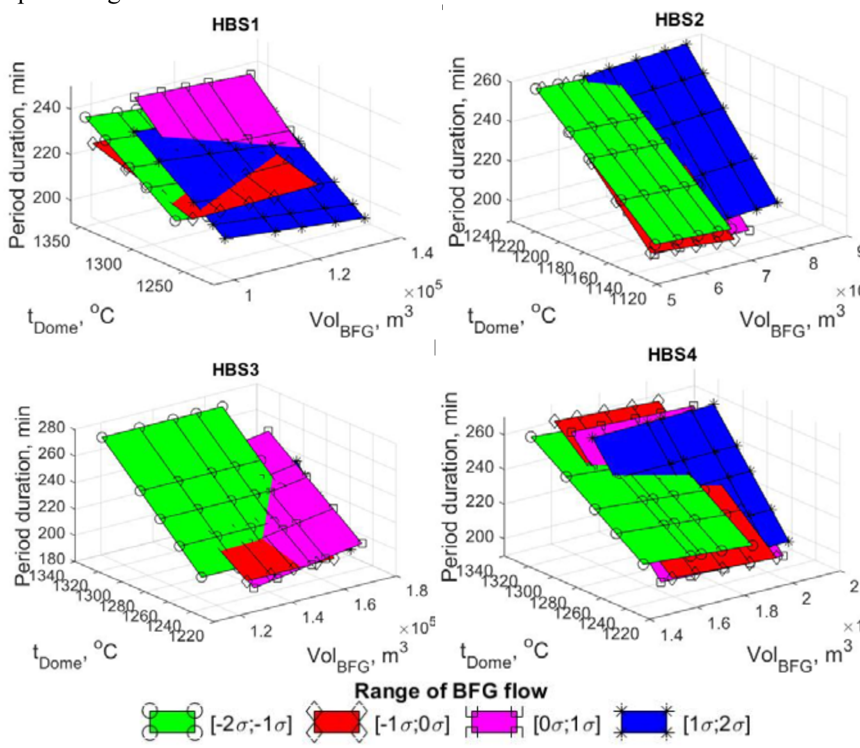

Fig. 8. Dependence of on-gas duration on the average dome temperature and the volume of consumed BFG within the range of BFG flow for each HBS of block

4. As a result of the research, analytical dependencies have been obtained, they allow to make decisions about changes in the heating flow chart and implement control algorithms for HBSB and each HBS of the block.

References

1. Qin, S. J.: Process data analytics in the era of big data. AIChE Journal, Vol. 60(9): pp. 3092-3100 (2014). doi: 10.1002/aic.14523
2. Liu, Y., Lv, Z., Wang, W.: An Improved Generalized-Trend-Diffusion-Based Data Imputation for Steel Industry. Mathematical Problems in Engineering, Vol. 2013: pp. 1-10 (2013). doi: 10.1155/2013/136241
3. Zhao, J., Liu, K., Wang, W., Liu, Y.: Adaptive fuzzy clustering based anomaly data detection in energy system of steel industry. Information Sciences, Vol. 259: pp. 335-345 (2014). doi: 10.1016/j.ins.2013.05.018
4. Sinha, P. Multivariate polynomial regression in data mining: methodology, problems and solutions. International Journal of Scientific and Engineering Research, Vol. 4(12): pp. 962-965 (2013).
5. Liu, C., Chen, S., Hai, Y., et al.: Experimental study and analysis on heat transfer hot air temperature of far infrared convection combined drying. In 2018 International Conference on Robots & Intelligent System (ICRIS): pp. 552-555. IEEE (2018). doi: 10.1109/ICRIS.2018.00142
6. Gasparini, V. M., de Castro, L. F. A., et. al.: Thermo-chemical model for blast furnace process control with the prediction of carbon consumption. Journal of Materials Research and Technology. Vol. 6(3): pp. 220-225 (2017). doi:10.1016/j.jmrt.2016.12.001
7. Kazarinov, L.S., Barbasova, T.A., Mohsen, H.M.: An efficient control method for furnace process based on cluster analysis. Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics. Vol. 16 (2): pp. 164–169 (2016). (In Russian) doi: 10.14529/ctcr160220
8. Shang, J., Chen, M., Zhang, et al.: Increment-based recursive transformed component statistical analysis for monitoring blast furnace iron-making processes: An index-switching scheme. Control Engineering Practice. Vol. 77, pp. 190–200 (2018). doi: 10.1016/j.conengprac.2018.05.012
9. Zhou, B., Ye, H., Zhang, H., Li, M.: Process monitoring of iron-making process in a blast furnace with PCA-based methods. Control Engineering Practice. Vol. 47, pp. 1–14. (2016). doi: 10.1016/j.conengprac.2015.11.006
10. Barbasova, T., Zagoskina, E.: Blast-Furnace Melting Blast Control. 2019 IEEE Russian Workshop on Power Engineering and Automation of Metallurgy Industry: Research & Practice (PEAMI), Magnitogorsk, Russia, pp. 13-17. (2019). doi: 10.1109/peami.2019.8915284
11. Zhou, B., Ye, H., Zhang, H., Li, M.: A new qualitative trend analysis algorithm based on global polynomial fit. AIChE Journal, Vol. 63(8): pp. 3374-3383. (2017). doi: 10.1002/aic.15706
12. Vorotnikova, Z.J., Novoselzev, M.A.: Cast iron fusion quality control functional monitoring. Reporter of Priazovskyi State Technical University: Collection of scientific papers. Vol. 26: pp. 221-227 (2013). (In Russian)
13. Vorotnikova, Z.Ye.: Forming and use of an archived database in the system "Adviser of a blast furnace operator". Scientific works of Donetsk National Technical University. Vol. 1 (26): pp.14-23 (2014) (In Russian).
14. Shnayder, D. A., Kazarinov, L. S., Barbasova, T. A., Lipatnikov, A. V.: Data mining and model-predictive approach for blast furnace thermal control. In 2017 Intelligent Systems Conference (IntelliSys), pp. 653-660 (2017). doi: 10.1109/IntelliSys.2017.8324364

15. Nose, K., Takemura, M., Morita, T.: Systems engineering approach to the optimization of hot blast stoves. IFAC Proceedings Volumes. Vol.: 16(15), pp. 47-54. (1983). doi: 10.1016/S1474-6670(17)64255-1/
16. Lee, Y.-H., Min, K. G., Han, C., Chang, K. S., Choi, T. H.: Process improvement methodology based on multivariate statistical analysis methods. Control Engineering Practice. Vol. 12(8), pp. 945–961 (2004). doi:10.1016/j.conengprac.2003.11.004
17. Jang, G.-S., Jeon, J.-H.: A Six Sigma Methodology Using Data Mining: A Case Study on Six Sigma Project for Heat Efficiency Improvement of a Hot Stove System in a Korean Steel Manufacturing Company. Cutting-Edge Research Topics on Multiple Criteria Decision Making. MCDM 2009. Communications in Computer and Information Science. Vol. 35: pp. 72-80 (2009). doi: 10.1007/978-3-642-02298-2_12
18. Koifman, A., Simkin, A.: Development and Software Implementation of the Hot Blast Stove Computer Model. CEUR Workshop Proceedings, Vol. 2353: pp. 440-454 (2019). Online http://ceur-ws.org/Vol-2353/paper35.pdf.