# The Use of Data Mining for the Optimization of the Control Model of the Reheating Furnace

# Využití data miningu k optimalizaci řídicího modelu ohřívací pece

**Mgr. Ing. Tomáš Tykva**[1]; **doc. Ing. Ivo Špička, Ph.D.**[1]; **Ing. Vendula Fialová**[2]

[1] VŠB – Technical University of Ostrava, Faculty of Metallurgy and Materials Engineering, Department of Automation and Computing in Metallurgy, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic
[2] VŠB – Technical University of Ostrava, Faculty of Metallurgy and Materials Engineering, Department of Economics and Management in Metallurgy, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic

*The paper presents basic data mining tools for optimizing the control model of the reheating furnace. It highlights data pre-processing as an important step in the data mining process mainly due to the frequent manifestations of technological indiscipline and problems with scanning the operating values of reheating furnace; its control system collects for the purpose of monitoring a number of operational data and its databases are as simple as possible; each trend is usually stored in a separate table. The aim of the paper is to demonstrate the possibilities of pre-processing data from the walking beam furnace in order to make them suitable for subsequent use for estimation parameters for the model chosen to refine the predictive properties of the selected model. These planned activities should also lead to the fulfillment of certain parts of the national program Industry 4.0 in the area of creating a "digital enterprise".*

***Key words**: data mining; genetic algorithms; reheating furnace*

*Příspěvek představuje základní nástroje data miningu pro optimalizaci řídicího modelu ohřívací pece. Akcentuje předzpracování dat jako důležitý krok v procesu data miningu hlavně kvůli častým projevům technologické nekázně a problémy se záznamem provozních hodnot ohřívací pece. Cílem tohoto příspěvku je demonstrovat možnosti předzpracování dat z krokové pece, aby byly vhodné pro následné použití odhadu parametrů zvoleného modelu ke zpřesnění prediktivních vlastností vybraného modelu.*

*Monitorovací a řídicí systémy shromažďují za účelem sledování řadu provozních dat. Tyto systémy budují své databáze tak snadno, jak je to možné; každý trend je obvykle uložen v samostatné tabulce. Údaje v tabulkách jsou ukládány podle dvou základních principů, a to buď pravidelně v nastaveném časovém intervalu, nebo po určité události, jako je například změna hodnoty proměnné. To ale vede k nadbytečnosti (duplicitě). Řešením by mohlo být uložení každé referenční proměnné v samostatné tabulce. Mnoho systémů používá své vlastní mechanismy pro ukládání dat, které kombinují oba přístupy, tj. pravidelné ukládání dat a ukládání hodnot proměnných jen při jejich změně.*

*Z výsledků modelování je pak jasně patrné, že koncept kombinující data mining z provozních dat a využití genetických algoritmů pro nastavování parametrů modelu přinesl očekávaný výsledek. V budoucnosti lze očekávat, že tento postup bude použit pro další technologie a jejich modely, s cílem vytvořit jejich digitální podobu tak, aby mohly být použity pro další optimalizační aktivity, resp. pro další funkce. Tyto plánované aktivity by rovněž měly vést k naplňování některých částí národního programu Průmysl 4.0 v oblasti vytvoření "digitálního podniku".*

***Klíčová slova**: data mining; genetické algoritmy; ohřívací pec*

Data mining is a general term for various methods of data analysis which frequently use artificial intelligence. The process of data mining is based on tabulated or alternatively arranged, usually like blind dates. It concludes by finding knowledge (patterns) contained in these data, which are expressed in a simpler and clearer form, and which can be used more easily in dealing with future situations than the original data.

Monitoring and control systems collect for the purpose of monitoring some operational data. These systems construct their databases as simply as possible; each trend is usually stored in a separate table. Data in the tables are stored according to two basic principles, either periodically after a preset time interval, or after a certain event - change of a variable value. This, however, leads to redundancy. A solution could be to store each reference variable in a separate table. Many systems use their mechanisms for storing data that combine both periodic data storage and storage management using a change of value. These data must be pre-processed before the deployment of data mining [1].

Hutnické listy č. 2/2017, roč. LXX
ISSN 0018-8069

## 1. Data Acquisition and Pre-Treatment of Data

From the scanned data it is necessary to select those that are important for further processing and possibly to derive additional data that describe the environment of a heated continuously cast slab (position, the temperature of the furnace environment in a given position). Furthermore, it is necessary to select the appropriate continuously cast slabs from the database to determine the times of moving the continuously cast slab from one position to the next position and to assign to these events the corresponding flow rates of the heating medium and the temperature..

The next step (using genetic algorithms) is to determine the constants that indicate the direct contribution of heat generated by the combustion of the heating medium and the contribution, which forms a heat transfer between the furnace lining and the heated continuously cast slab. The already mentioned use of genetic algorithms is described in the next chapter [2].

## 2. The Use of Genetic Algorithms

Genetic algorithms (GAs) represent powerful search algorithms that perform an exploration of the search space that evolves in analogy to the evolution in nature. The power of GAs consists in the fact that they need only objective function evaluations. So derivatives or other auxiliary knowledge are not used. Instead, probabilistic transition rules of deterministic rules and handling a population of candidate solutions (called individuals or chromosomes) that evolves iteratively are used. Each iteration of the algorithm is called a generation. The evolution of a species is simulated through a fitness function and some genetic operators such as reproduction, crossover and mutation [3].

The fittest individuals will survive generation after generation, while reproducing and generating offspring that might be stronger and stronger. At the same time, the weakest individuals disappear from each generation. Individuals must be encoded in some alphabet, with binary strings, real numbers, vectors and other. In a practical application of genetic algorithms a population pool of chromosomes has to be installed and they can be initially randomly set. In each cycle of genetic evolution subsequent generation is created from the chromosomes in the current population. The cycle of evolution is repeated until a termination criterion is reached. The number of evolution cycles or a pre-defined fitness value can set this criterion.

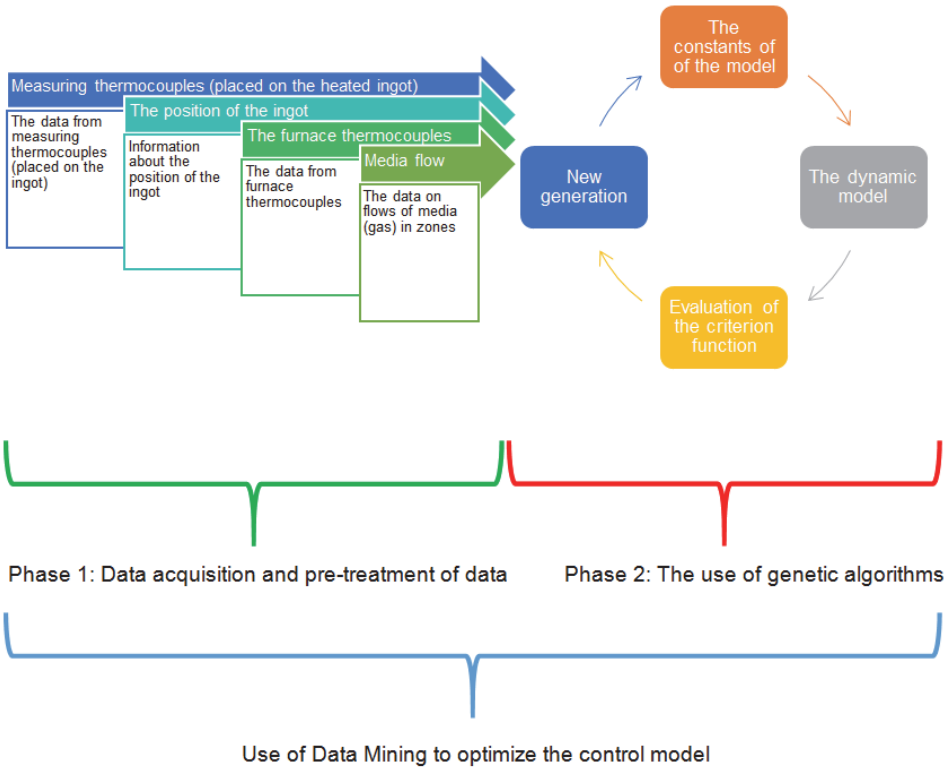


Fig. 1 Data preparation and application of genetic algorithms. (Source: our own processing)
Obr. 1 Příprava dat a aplikace genetických algoritmů. (Zdroj: vlastní zpracování)

The principle of genetic algorithms is that in the initial phase a default population (generation) is generated; each of these populations is characterized by a specific combination of genes that determine ultimately the behaviour of the examined model. They can, therefore, be considered as parameters of a general criterion function, which is later used to select the individuals of an appropriate generation. Choosing the most

appropriate individuals within the meaning of criterion values through crossbreeding methods, or mutations create a new generation containing the same number of individuals, however, with modified genes. These, in turn, are used to evaluate the criterion function, and the process is repeated until the set-point of the behaviour of the objective function (model) is reached. This shows the cycle on the right side of the figure (see the left side of the Fig. 1. [4].

The result (using genetic algorithms) is the determination of constants, which specify the direct contribution of heat generated by the combustion of the heating medium and the contribution, which forms a heat transfer between the furnace lining and the heated continuously cast slab. This part was solved using genetic algorithms, however, without the application of the principles of data mining genetic algorithms would be unable to assign the input data correctly at appropriate times to be input into the selected dynamic model.

The selected data (Fig. 2) about the flows of combustion media in each zone are computed in the block named gas and zones. The furnace temperature is read from the table temperature furnace; in the block model it is compared with the material temperature and the difference between these two values are put to the port labelled error. From this value, the–square roots are computed and they are integrated during the whole time of a simulation. At the end of the simulation, the value is stored into a variable named pow_sum_err.

This value represents the value of the criterion function for the next usage in the genetic algorithm program. Parameters for the block model are set by the genetic algorithm. In general, the block model implements the system for solving the heat transfer between the furnace atmosphere and the heated material, the heat generated by the combustion of gas and the heated material, and finally the transmission determining the relation between the quantity of gas in the zone and usable heat [5].

Criterion function is determined as the sum of squared deviations between the modelled temperature of the heated material and the actual measured temperature.

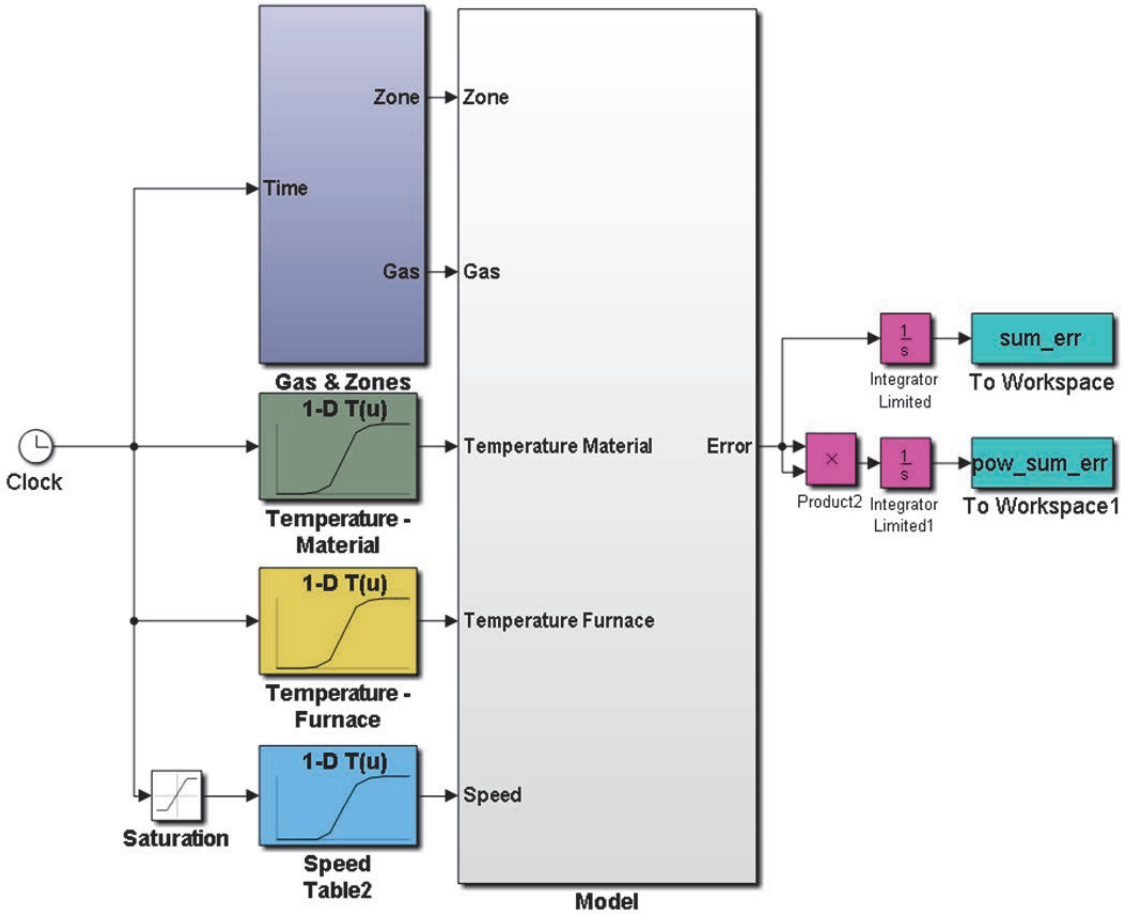


Fig. 2 The structure of the model of heating in a heating furnace. Source: (our own processing)
Obr. 2 Struktura modelu ohřívání v ohřívací peci. Zdroj: (vlastní zpracování)

Fig. 3 shows the result which was obtained by procedures outlined above; curve 5 was obtained by data mining methods. The temperature from the individual thermocouples in the furnace was converted to the temperature of the furnace environment above the appropriate continuously cast slab.

Hutnické listy č. 2/2017, roč. LXX
ISSN 0018-8069

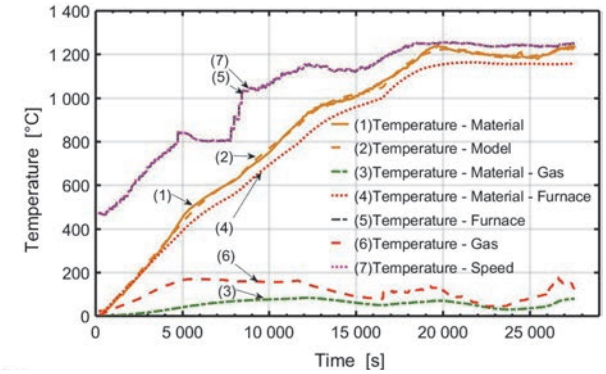


Fig. 3 The temperature curve. Source: [6]
Obr. 3 Teplotní křivka. Zdroj: [6]

At the time of 20 000 – 25 000 seconds, the temperature at the thermocouple of the continuously cast slab was decreasing even when the temperature of the furnace environment above the continuously cast slab was almost constant. Therefore, using genetic algorithms the system sets the parameters of the model which take into consideration not only the temperature of the furnace environment, but also the contribution resulting from burning of gas in the zone.

## Conclusions

It is evident from Figure 3 that offsetting the influence of the current input of the zone progress simulated and measured temperatures are nearly identical. From the results of modelling, it can be concluded that the concept of combining the data mining operational data and use of genetic algorithms for adjusting parameters of the model, which in this case was 28, brought the expected result. In the future, it may be expected that this procedure will be applied to other technologies and their models to create their digital models so that they can be used for the subsequent optimized solution of the problem with varying outlined criterion functions [7]. These planned activities should also lead to the fulfillment of certain parts of the national program Industry 4.0 in the area of creating a "digital enterprise."

***Acknowledgement***

*This article was supported by the specific university research No. SP2017/63.*

**Literature**

[1] ŠPIČKA, I., HEGER, M., ZIMNÝ, O. et al. Industrial Control Systems and Data Mining. In 20th *Annual International Conference on Metallurgy and Materials: Metal 2011*. Ostrava: Tanger, 2011, 1229-1234.

[2] HAN, J., KAMBER, M. *Data Mining: Concepts and Techniques*. 3rd ed. Burlington, MA: Elsevier, 2011.

[3] BENISIS, A. *Business Process Management: a Data Cube to Analyze Business Process Simulation Data for Decision Making*. Saarbrücken: VDM Verlag Dr. Müller Aktiengesellschaft & Co. KG, 2010.

[4] JIN, Y. *Multi-Objective Machine Learning* [online]. Berlin, Heidelberg: Springer, 2006 [cit. 2016-11-14]. ISBN 978-3-540-33019-6. Available from: http://dx.doi.org/10.1007/3-540-33019-4.

[5] KOŠTIAL, P., ŠPIČKA, I., JANČÍKOVÁ, Z., DAVID, J., VALÍČEK, J., HARNIČÁROVÁ, M., RUSNÁK, V. Lumped Capacitance Model in Thermal Analysis of Solid Materials. *Journal of Physics*, 588 (2015) 1, 1–7.

[6] ŠPICKA, I., HEGER, M., ZIMNÝ, O., JANČÍKOVÁ, Z., TYKVA, T. Optimizing the Model of Heating the Material in the Reheating Furnace in Metallurgy. *METALURGIJA*, 55 (2016) 4, 719–722.

[7] SENDLER, U., BAUM, G., BORCHERDING, H., BROY, M., EIGNER, M., HUBER A. S., KOHLER, H. K., RUSSWURM, S., STÜMPFLE, M. *Industrie 4.0: Beherrschung der industriellen Komplexität mit SysLM*. Berlin: Springer Verlag, 2013.

# Výrobci ploché oceli plánují zvýšení cen

*Stahl Aktuel* *09.01.2017*

Po dlouhé žíznivé cestě plánují brazilští výrobci ploché oceli Cia Siderúrgica Nacional, Usiminas a ArcelorMittal zvyšování cen. Oznamuje to agentura Reuters s odvoláním na znalce odvětví. V plánu je zvýšit ceny referenčních výrobků širokých pásů, válcovaných za tepla a jemného plechu ještě v lednu o 8 až 10 %. Ceny pozinkovaných plechů mají ale zůstat stejné.

# Přebytečné kapacity: likvidační plány by mohly být k dispozici před oslavou Nového roku

*Stahl Aktuell* *16.01.2017*

Přebytečné kapacity v čínském ocelářském průmyslu by měly v roce 2017 dále klesat. Tak čínská vláda podle údajů Commerzbank plánuje mimo jiné uzavření ilegálních výroben, které jsou zodpovědné za cca 40 mil. tun oceli ročně. Kromě toho by měly být do poloviny roku uzavřeny veškeré výrobny, které zpracovávají méně hodnotný ocelový šrot. Postiženy mají být rovněž podniky, které nedodržují státní předpisy o ochraně životního prostředí.