Metamodelling and optimization of copper flash smelting process

Marcin Gulik
AGH University of Science and Technology
Faculty of Metals Engineering
and Industrial Computer Science
al. Mickiewicza 30, 30-059 Kraków
POLAND
mgulik21@gmail.com

Piotr Jarosz
AGH University of Science and Technology
Faculty of Non-Ferrous Metals
al. Mickiewicza 30, 30-059 Kraków
POLAND
pijar@agh.edu.pl

Jan Kusiak
AGH University of Science and Technology
Faculty of Metals Engineering
and Industrial Computer Science
al. Mickiewicza 30, 30-059 Kraków
POLAND
kusiak@agh.edu.pl

Stanisław Małecki
AGH University of Science and Technology
Faculty of Non-Ferrous Metals
al. Mickiewicza 30, 30-059 Kraków
POLAND
stanmal@agh.edu.pl

Paweł Morkisz
AGH University of Science and Technology
Faculty of Applied Mathematics
al. Mickiewicza 30, 30-059 Kraków
POLAND
morkiszp@agh.edu.pl

Piotr Oprocha
AGH University of Science and Technology
Faculty of Applied Mathematics
al. Mickiewicza 30, 30-059 Kraków
POLAND
oprocha@agh.edu.pl

Wojciech Pietrucha
AGH University of Science and Technology
Faculty of Applied Mathematics
al. Mickiewicza 30, 30-059 Kraków
POLAND
wojtekpietrucha@gmail.com

Łukasz Sztangret
AGH University of Science and Technology
Faculty of Metals Engineering
and Industrial Computer Science
al. Mickiewicza 30, 30-059 Kraków
POLAND
szt@agh.edu.pl

Abstract: The aim of the paper is metamodelling of copper flash smelting process and an attempt of its optimization. Flash smelting process is an important step in the metallurgical process of production of copper blister (pure copper). Three different modelling techniques: Artificial Neural Networks (ANN), Support Vector Machines (SVM) and Random Forest (RF) were used in construction of metamodels and the obtained results were compared. The metamodels are used in finding a Pareto-optimal solution (and control parameters) for the considered industrial process.

Key–Words: Metamodelling, Artificial Neural Networks, Random Forest, Support Vector Machines, optimization, flash copper smelting process

1 Introduction

Industrial processes are usually very complex, nonlinear and depend on numerous parameters. Moreover, not all parameters of a real industrial process can be measured, while some of process parameters (signals) depend on many other elements, such as technological condition (wear) of tools, quality of materials or semi-products used in the process, etc. Therefore, we often have to rely only on averaged data, experience of engineers, etc. The models of such industrial processes are usually very complex and their simulations are costly with regard to the computing time. The computation time increases rapidly in case of optimization calculations, requiring numerous iterations (simulation calls). Hence, the motivation for the current research was a replacement of computationally costly model of the analyzed process by a metamodel. Developed metamodels should describe the analyzed process with required accuracy and decrease the computation time of the optimization procedure, as well, can be an important aid in the control of the production.

In the present paper we will test three different methods of construction of metamodels: Artificial Neural Networks (see [5, 6]), Support Vector Machines (see [1, 12]) and Random Forests [2, 5]. Our

ISBN: 978-1-61804-347-4

testing ground will be simplified model of Outotec type copper flash smelting process. We will use data generated by theoretical model of the process (see [3]).

2 Copper flash smelting process

We focus on a copper flash smelting process (Outotec type) which is an important step in production of copper blister (99 % Cu). The process takes place in a flash smelting furnace, which schematically is presented in Figure 1. Main feedstock of the process is

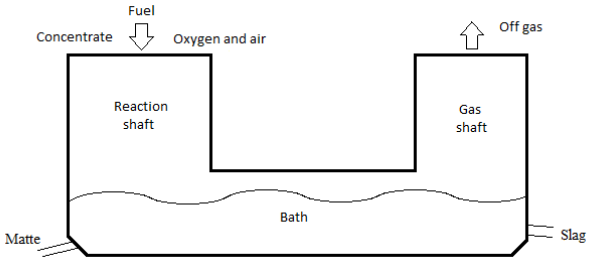


Figure 1: Schematic structure of outotec copper flash smelting furnace

a dry Cu-Fe-S concentrate. Very roughly speaking, dry concentrate is blasted with O_2 in reaction shaft. Furnace is equipped with hydrocarbon fuel burners located concentrically in a reaction shaft. Burners are used to adjust temperature, so that the reaction in the shaft is performed properly, that is a concentrate smelts at specified rate.

Main factor of the process is a composition of dry concentrate which is a mixture of Cu_2S and $CuFeS_2$. Let us assume that the furnace is loaded with 1 000 kg of concentrate. Then the amount (on average) of the Cu, Fe and S in the concentrate can be evaluated using the the following relations (see [3]):

$$\begin{aligned} Cu &= 346x + 800(1-x), \\ Fe &= 304x, \\ S &= 350x + 200(1-x), \end{aligned}$$

where the parameter $x \in [0, 1]$ depends on the contents/origin of the concentrate.

Parameter x cannot be controlled, since it depends on the composition of the material that is delivered to the plant. There are three parameters of the process that can be controlled, and have an essential impact on the quality of the final product of smelting

- *fuel* - fossil fuel burned in the furnace. In the process are used between 0 and 40 kg of fuel.
- O_2 - mass of O_2 in blast,
- N_2 - mass of N_2 in blast.

As a result of smelting in furnace a molten matte is produced.

The metamodels have the following four input signals: *fuel*, O_2, N_2 and x. Constructed metamodels predict (separately) two output parameters:

- Q_1 - concentration of Cu in molten Cu-Fe-S matte (between 45% and 75% of Cu) - a semi-product which goes to the further production stage - oxidation converting into metallic copper,
- Q_2 - concentration of FeO in molten Fe silicate slag.

The aim of the further optimization procedure is obtaining the required values of Q_1^{opt} and Q_2^{opt}.

3 Metamodelling of copper flash smelting process

One of our main goals was to build and to compare three different approaches to metamodels: ANN, SVM and RFs. Two independent metamodels, separately for each quality parameter, i.e. separate model for Q_1 and Q_2 were constructed using each of these methods.

We had 46 741 records (observations) as starting point for building our metamodels. We used 98% of the records (45 707 observations) as training samples, while 2% (934 elements) were used at the testing stage.

The first metamodel was built on the basis of Artificial Neural Network[1] (Multi-Layer Perceptron type). The Levenberg-Marquardt backpropagation method was employed in training process of the networks. Since the ANN's accuracy depends on the topology of the network, we trained 50 networks using different topologies for each quality criterion and the best were chosen: 3-layer network of 4-28-1 structure with *tan-sigmoid*, *tan-sigmoid* and *linear* activation functions for metamodel of Q_1; the 4-layer network of 4-22-2-1 structure with *log-sigmoid*, *tan-sigmoid*, *tan-sigmoid* and *linear* activation functions for the metamodel describing the Q_2 quality criterion.

In the second approach, Support Vector Machines[2] were used as a regression approach in metamodelling of considered process. The learning process was executed using the *radial basis kernel* and *eps-regression* type as parameters.

The last methodology used for metamodelling was Random Forest. Again, our models were implemented in R, using "randomForest" package. We constructed two different forests consisting of 500 trees. We used standard entries of the function `randomForest` in the first case of regression mode

[1] Matlab 2011b (Neural Network Toolbox) was used.

[2] Package "e1071' in R program was used.

ISBN: 978-1-61804-347-4

and we additionally set parameter[3] `nodesize` $= 5$ in the second attempt.

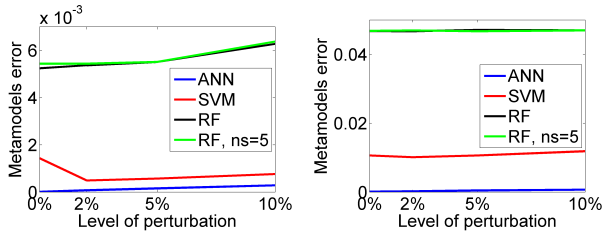
To test fitting of our metamodels on verification data we calculated mean error separately for each quality function (here Q^i is value in metamodel, and $\overline{Q^i}$ is value in the testing dataset):

$$E = \frac{1}{n}\sum_{i=1}^{n}|Q^i - \overline{Q^i}|. \quad (1)$$

When constructing our metamodel, we considered 4 modifications of data adding random perturbation, which should reflect problems in evaluation of quality parameters. We considered 4 different levels of perturbation of real (accurate value) of Q^i, that is: 0% (accurate measurement), 2%, 5% and 10%. The accuracy of each metamodel was computed using equation (1) for all levels of training data perturbation. The obtained errors are presented in Table 1 and Figure 2.

Table 1: Errors of fitting (1) of considered metamodels.

Pert.		Error			
		RF	RF ns=5	SVM	ANN
0%	Q_1	0,005245	0,005442	0,001432	0,000012
	Q_2	0,046774	0,046840	0,010633	0,000123
2%	Q_1	0,005374	0,005440	0,000492	0,000077
	Q_2	0,046738	0,046967	0,010128	0,000215
5%	Q_1	0,005515	0,005515	0,000573	0,000157
	Q_2	0,047113	0,046743	0,010628	0,000463
10%	Q_1	0,006281	0,006375	0,000763	0,000286
	Q_2	0,047022	0,046974	0,011860	0,000701



(a) Error of fitting of Q_1 (b) Error of fitting of Q_2

Figure 2: Graphical illustration of fitting errors (1) of metamodels.

4 Optimization

As we mentioned above, two different quality functions were considered. Therefore, it is possible to get only Pareto-optimal solutions. We decided to apply Global Criterion Method (see [8] or [4]), with goals $Q_1^{opt} = 0.65$ and $Q_2^{opt} = 0.54$, which are technologically motivated optimal values of quality functions. The multi-objectives problem was transformed into single-objective one, and each of the quality functions was associated with a weight. Recapitulating, the considered objective function had the following form:

$$F = w\frac{|Q_1 - Q_1^{opt}|}{Q_1^{opt}} + (1-w)\frac{|Q_2 - Q_2^{opt}|}{Q_2^{opt}}. \quad (2)$$

The assumed feasible decision space of control parameters was technologically limited by the following ranges: *fuel:* 0–40, O_2: 122–467, N_2: 0–1634. All optimization calculations were performed for $x = 0.5$, which means that the input concentrate was composed of 500 kg of Cu_2S and 500 kg of $CuFeS_2$.

In the first approach we performed single-objective optimization assuming the equal importance of both quality functions, that is $w = 0.5$. The optimization process was performed using Monte-Carlo method with 50 000 objective function tests for given metamodel (in this approach we used metamodels trained with no perturbations). In each case we performed optimization 100 times, obtaining 100 different candidates for optimal solution.The best (B), average (A) and worst (W) values of considered optimization criteria (F, Q_1 and Q_2)[4] corresponding to the found optimal solutions of the parameters (*fuel*, O_2 and N_2) calculated using each metamodel are presented in Table 2.

Table 2: Optimal values of criteria and corresponding parameters.

		Criteria			Parameters		
		F^*	Q_1^*	Q_2^*	$fuel^*$	O_2^*	N_2^*
ANN	B	8.78e-4	0.649	0.538	31.5	184	142
	A	0,011	0.654	0.537	-	-	-
	W	0.0273	0.67	0.517	39	226	430
SVM	B	0.0021	0.648	0.539	31.9	187	161
	A	0.019	0.653	0.538	-	-	-
	W	0.047	0.664	0.482	39	221	421
RF	B	4.29e-4	0.65	0.539	27.8	173	60
	A	0.004	0.649	0.537	-	-	-
	W	0.0081	0.649	0.537	28.6	167	56
RF *ns*=5	B	1.67e-4	0.65	0.538	27.5	174	61
	A	0.005	0.649	0.537	-	-	-
	W	0.0096	0.652	0.526	27.9	136	48

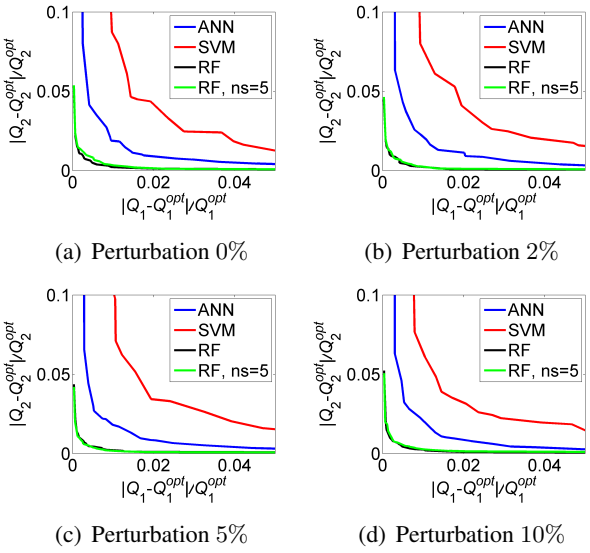
We can see, that optimization performed based on RF and RF ns=5, metamodels give much better solutions than in case of optimization based on two other metamodels. However, it must be notice that RF and RF ns=5 metamodels are characterized by the biggest metamodelling errors (see Figure 2).

[3] In what follows, we will write RF ns=5 to emphasize that Random Forest with this extra parameter was applied

[4] We are looking for the minimum of F, while Q_1, Q_2 should be close to the given, required values $Q_1^{opt} = 0.65$ and $Q_2^{opt} = 0.54$

ISBN: 978-1-61804-347-4

In second approach, our goal was to obtain Pareto fronts. Therefore, for each metamodel and each level of perturbation we set different values of weight w (in equation (2)) to $0, 0.05, \ldots, 0.95, 1$. As previously, in each case we run optimization procedure 100 times. To create Pareto fronts (presented in Figure 3) we used average values of obtained results.



(a) Perturbation 0% (b) Perturbation 2%

(c) Perturbation 5% (d) Perturbation 10%

Figure 3: Pareto front for different metamodels and perturbation ranges

Static fitting evaluation (1) comparing metamodels responses on test data is not the only criterion to be taken into account. Applied methodology may have strong influence on the topology of the model and as consequently different metamodels may behave differently when searching for optimal solution of quality function (2). For example, Random Forest always extrapolate its answers in the max-min range provided by data set in learning process, while ANN or SVM may respond with values outside desired range and so penalty functions have to be used. Clearly, there are many other tiny details that have to be taken into account in one model but does not exist in the other approach.

5 Conclusion

We presented three different methodologies of metamodeling of copper flash smelting process. Analyzing average errors of fitting models into data (see Table 1 and Figure 2) we see that the best results were obtained in modeling by multi-layer neural network. Especially random forest methodology related to Q_2 was very far from desired. Surprisingly, fitting results of random forest with parameter $ns = 5$ were only slightly better. One of the reasons can be that some values of control parameters, while in acceptable range, can be unacceptable from technological point of view. For these parameters we probably will not have records in the data set (failure of the process). There is no chance to avoid these regions, since RS methodology is a kind of averaged classification method, hence is extrapolating values between min-max. It seems that in practice, to improve performance of metamodel built using this methodology, some additional records for non-acceptable parameters set should be provided and set with some unacceptable value (e.g. -1 as value of Q_1 which is concentration of Cu in molten Cu-Fe-S matte). It suggests further investigation on the structure of forest (and learning methodology) in modeling of the process of flash smelting may be necessary.

It is also interesting that performance of SVM increased when data became perturbed (random noise was added). It also behaved much better than ANN in the optimization process when modelling using perturbed data. It suggests that in the case of usually very noisy industrial data, SVM can be good choice. This observation also motivates further investigation in this direction.

Acknowledgements: The research was supported by National Science Centre in Poland (NCN), Grant no. DEC-2013/11/B/ST8/00352.

References:

[1] B. E. Boser, I. M. Guyon, V. N. Vapnik, *A Training Algorithm for Optimal Margin Classifiers*. Proceedings of the Fifth Annual Workshop on Computational Learning Theory 5, Pittsburgh, 1992, pp. 144-152.

[2] L. Breiman, Random Forests, *Machine Learning* 45, 2001, pp. 5-32

[3] W. G. Davenport, M. D. Jones, M. J. King, E. H. Partelpoeg, *Flash Smelting: Analysis, Control and Optimization*, The Minerals, Metals & Materials Society, Warrendale 2001.

[4] K. Deb, *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons, Chichester 2001.

[5] T. Hastie, R. Tibshirani, J. Friedman, Jerome. *The elements of statistical learning. Data mining, inference, and prediction*. Second edition. Springer Series in Statistics. Springer, New York, 2009.

[6] J. Kusiak, D. Szeliga, Ł.Sztangret, Modelling techniques for optimizing metal forming processes, *in: Microstructure evolution in metal forming processes*, Woodhead Publishing Limited, 2012.

[7] J. Kusiak, Ł.Sztangret, Ł. Rauch, M. Pietrzyk, Metamodel driven optimization of thermomechanical industrial processes, *Computer Methods in Materials Science : quarterly* 14, 2014, pp. 2026.

[8] S. S. Rao, *Engineering Optimization: Theory and Practice*, John Wiley & Sons, New York, 1996.

[9] M. E. Schlesinger, M. J. King, K. C. Sole, W. G. Davenport, *Extractive Metallurgy of Copper*, Fifth Edition. Elsevier Ltd., Amsterdam 2011.

[10] Ł Sztangret, D. Szeliga, J. Kusiak, M. Pietrzyk, Application of inverse analysis with metamodelling for identification of metal flow stress, *Canadian Metallurgical Quarterly* 51, 2012, pp. 440446.

[11] Ł. Sztangret, Ł. Rauch, J. Kusiak, P. Jarosz, S. Małecki, Modeling of the oxidizing roasting process of zinc sulphide concentrates using artificial neural networks, *Computer Methods in Materials Science : quarterly* 11, 2011, pp. 122127

[12] V. N. Vapnik. *The Nature of Statistical Learning Theory*. 2nd edition, Springer, 1999.

ISBN: 978-1-61804-347-4