APPLICATION OF MODELING TOOLS IN MANUFACTURING TO IMPROVE QUALITY AND PRODUCTIVITY WITH CASE STUDY

Mahesh B. PARAPPAGOUDAR[1,*], Pandu R. VUNDAVILLI[2]

[1] Principal & Ph.D., Professor in Mechanical Engineering Dept, CSIT, Durg (C.G), India
[2] Ph.D, Professor & Head, Mechanical Engineering Dept, MIC College of Technology, Kanchikacherla, AP, India

Abstract: *The productivity improvement, especially in the manufacturing industries helps to reduce the cost and further it assists the firms to be competitive in the market. One of the tools which help the industries to reduce cost, manufacturing lead time and defective products is process modeling. Hence, the present day's need of the market to achieve the quality with the reduction in cost is possible with the effective utilization of modeling tools. Many modeling and simulation tools are available to simulate manufacturing processes. In the present paper, both conventional regression tools (Design of Experiments, Response Surface Methodology) as well as un-conventional modeling tools based on Neural Networks (Back Propagation Neural Network and Genetic-Neural Network) are discussed. All manufacturing processes can be represented as multi input multi output system. It is important to understand the influence of process variables and establish input-output relationships. Design of Experiments combined with Response Surface Methodology is a powerful statistical tool to analyze the influence of process variables and establish accurate input-output relationships. The requirement of manufacturing processes is to know a set of process variables, which will result in desired output (Reverse mapping). However, most of the times the statistical tools fail in reverse mapping. Hence, the use of soft computing based Neural Network, Genetic Algorithm, Fuzzy Logic and their different combinations can be successfully used in reverse mappings. The application of these tools is illustrated with the help of a case study.*

Key words: *Modeling, Manufacturing Process, DOE, RSM, NN, GA, FLC.*

1. INTRODUCTION

The main objective of any manufacturing process is to produce good quality product at a low cost and at the right time. This objective can be achieved by minimizing the incidence of the product defects. To reduce the defects, one should have sound knowledge of the process through which the products are made. The manufacturing process involves a number of stages, depending on the type of manufacturing process. For example the casting process involves pattern making, mould preparation, melting and pouring. Many a times the control of manufacturing process is difficult due to the large number of variables associated with it. Hence, modeling tools can be successfully applied to improve the quality and reduce the cost of manufacturing.

Modeling of the physical system means identifying, establishing and analyzing the input-output relationships of the system. In other words, once the model is established with the help of planned experiments, it is possible to know the quantitative change in the response values when the independent variables are changed from one set of values to another set without conducting the real experiments. Regression analysis is the most common

method of establishing the relationships. This will help to identify the sensitivity of process variables and establish better control over the process.

Design of experiment with response surface methodology (one of the regression methods using statistical tools) provides with a very good analysis and precise input-output relationships. In a multi-input multi-output system, there may be either the linear or non-linear input-output relationship for the responses. The major requirement of the modeling approach is to have well-defined, quantifiable input variables and the well-defined measurable response (output).

In conventional statistical regression analysis of a system, only one response can be determined at a time, as the function of input process parameters. However, in a multi-input and multi-output system, all the outputs are to be determined simultaneously and there might be some dependency among themselves. Thus, the above approaches are unable to capture the physics of the system fully. These approaches can solve the problems of forward mapping (that is, to predict the responses for a set of input process parameters). However, they may not be able to solve the problems of reverse mapping (that is, to determine the process parameters to ensure a set of desired mould properties) always efficiently. These problems may be solved using soft computing-based modeling techniques. Soft computing-based approaches, namely Genetic Algorithms (GAs), Fuzzy Logic Controller (FLC), Neural Networks (NNs) and their different com-

---
* Corresponding author: Mechanical Engineering Dept, CSIT, Durg (C.G) - 491001 – India
Tel.: +91 788 2320884
Fax: +91 788 2334808
E-mail addresses: *maheshpg@gmail.com* (M. Parappagoudar), *panduvundavilli@gmail.com* (P. Vundavilli)

binations have been used by many researchers in modeling manufacturing processes.

Many researchers are working on the application of above discussed modeling tools to analyze and establish input-output relations in manufacturing processes. The literature shows the application of the above discussed modeling tools in manufacturing processes, namely machining, forming, casting etc [1–3].

2. MODELING APPROACHES

Conventional modeling tools-based on Design of Experiments, Response Surface Methodology and Unconventional modeling tools based on Neural Network (BPNN and GA-NN) have been discussed in the following subsections

2.1. Conventional regression analysis method (design of experiments and response surface methodology)

Statistically designed experiments are outstanding in comparison with the conventional engineering experimental approach of varying one variable at a time. The most important advantage of statistical Design of Experiment (DOE) lies in the fact that several variables are simultaneously studied for a more complete insight into the combined effects of the factors on the response under investigation. The experiments usually involve a large number of variables. The well-planned statistically designed experiment requires a less number of experiments compared to the conventional engineering experimental approach.

Regression analysis is the most common method of establishing the relationships. Design of experiment with response surface methodology (one of the regression methods using statistical tools) provides with a very good analysis and precise input-output relationships. The conventional statistical regression analysis based on a Design of Experiments (DOE) provides a very good understanding of the effect of input variables on the responses of a physical system. Moreover, the input-output relationships developed in this method yield accurate results and the regression models can be used to predict the responses for a set of input variables.

Statistical Design of Experiment (DOE) refers to the process of planning the experiment, so that an appropriate set of data can be collected and analyzed using the statistical methods for drawing inferences on the input-output relationships. It is an efficient technique to estimate the effects of several variables simultaneously. Each experimental design will contain a group of experimental runs. Optimizing a product or process design means determining the best architecture, levels of control factors and tolerances. The DOE combined with the Response Surface Methodology (RSM) is a powerful statistical tool to develop the input-output relationships. The RSM is an empirical modeling approach using polynomials as the local approximations to the true input-output relationships. The statistical regression analysis based on DOE consists of the following steps [4–5]:

1. Identify the important process parameters (input variables) and their feasible upper and lower limits. Decide on the number of replicates. Replicates make the test of significance possible by providing an estimate of experimental error. It also makes the statistical test more sensitive and increases the precision.
2. Develop the design matrix based on the number of variables and their levels chosen.
3. Conduct the experiments with the input variable combinations as per the design matrix and record the corresponding response values.
4. Use response surface methodology to develop linear or non-linear input output relationships (depending on the number of levels of the input variables and design matrix).
5. Check the statistical adequacy of the model through significance test, ANOVA test and coefficient of determination (that is, R^2 value).
6. Analyze the effect of variables on the responses with the help of effects plot (linear models) and surface plots (non-linear models).
7. Test the performance (practical suitability) of the developed models using some test cases.

The more important practical requirement is to obtain the combination of input variables that will produce the desired output, that is, the back prediction. The back prediction using statistical regression was attempted by Lee and Rhee [6] to predict the process parameters in welding. The use of statistical regression method in back prediction requires the transformation matrix to be a square one, so that its inverse can be determined. However, it is possible, when we consider only the linear terms leaving all interaction terms in the regression equation. In most of the situations, leaving all the interaction terms may lead to imprecise input-output relationships. In short, a statistical regression model provides with a scope to perform very good analysis. However, there has to be a separate model for each of the outputs as a function of input variables. Therefore, the dependency among the outputs, if any, will be lost. Moreover, the back prediction may be difficult to obtain in statistical regression analysis. These problems may be solved using soft computing-based modeling techniques

2.2. Neural network-based modeling approaches

Artificial neural networks are excellent tools in modeling the complex manufacturing processes. Most of the times, these techniques have proved better than the conventional modeling tools. An Artificial Neural Network (ANN) is an interconnected network of many simple processing units, which are analogous to the biological neurons in the human brain. ANN is an information-processing paradigm inspired by the way the densely interconnected, parallel structure of the mammalian brain processes information. It is composed of a large number of highly interconnected processing elements that are analogous to neurons and are tied together with weighted connections that are analogous to synapses.

2.2.1. Back-propagation neural network. With the fixed topology, specifying the weights of a neural network can be seen as an optimization process with the goal to find a set of weights that minimizes the network's error on the training set. Usually, the problem's error surface will be multi-dimensional and contains many local minima. The most widely used algorithm for this

problem is the back-propagation algorithm, which is a local gradient search method. The back-propagation is prone to get struck at the local minima, as it needs gradient information. The success of back-propagation method depends on good, problem specific parameter settings. In general, there are two different schemes for the training of neural network, namely incremental and batch training, depending on feeding of the training data. In incremental training, the weights are updated for each training data set, whereas in case of batch mode of training, the weights will be updated only after the entire set of training data is passed through the network. Incremental training can be used for the online training, whereas off-line training uses the batch mode of training. BPNN algorithm is based on steepest descent method to minimize error. The performance of BPNN is based on nature of input-output data. Further, huge input-output data is required for better training of Neural Network.

Figure 1 shows a feed-forward neural network. The network is iteratively trained to reduce the mean squared error. The error function is the mean squared error, as given below.

$$\text{Minimize } E = \frac{1}{R \times N} \sum_{i=1}^{R} \sum_{j=1}^{N} \frac{1}{2}(T_{ij} - O_{ij})^2, \qquad (1)$$

where T_{ij} and O_{ij} represent the target and predicted values of the responses, respectively, R indicates the number of responses and N represents the number of training scenarios.

Back-propagation algorithm (which works based on a steepest descent method) with a momentum term α will be used to update the weights of the neural network, as given below.

$$\Delta W_{jk}(t) = -\eta \frac{\partial E}{\partial W_{jk}}(t) + \alpha.\Delta W_{jk}(t-1), \qquad (2)$$

where η indicates the learning rate, α represents the momentum constant, t indicates the iteration number and $\frac{\partial E}{\partial W_{jk}}$ can be determined using the chain rule of differentiation as given below.
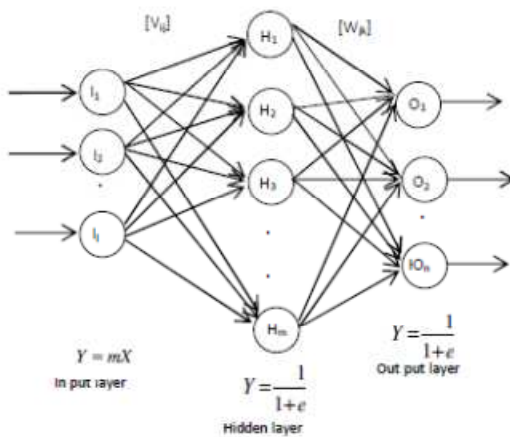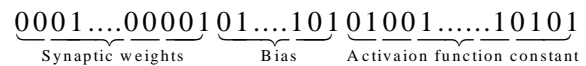

Fig. 1. Structure of neural network.

$$\frac{\partial E}{\partial W_{jk}} = \frac{\partial E}{\partial Y_k} \cdot \frac{\partial Y_k}{\partial U_k} \cdot \frac{\partial U_k}{\partial W_{jk}}, \qquad (3)$$

where Y_k and U_k represent the output and input, respectively, of kth neuron lying on the output layer. Similarly, the change in $[V_{ij}]$ values can also be determined [7].

2.2.2. Genetic-Neural (GA-NN) System. The back-propagation algorithm works based on gradient-based approach (i.e., steepest descent) and its solutions may get trapped into the local minima. Hence, some good solutions may be left out in the process. On the other hand, a Genetic Algorithm (GA) carries out its search in a huge space. Hence, the possibility of its solutions for being trapped into the local minima is less. Discontinuous functions could be handled easily using a GA, whereas a back-propagation algorithm may not train an NN, if the threshold functions are not continuous because it requires a derivative of the function. In GA-NN, the GA replaces the back-propagation mechanism of learning. The error, which is caused due to the difference between the desired output and actual output obtained from the NN, is fed back to GA in the form of its fitness and the weights of the network are optimized to obtain the minimum error. The complete set of weights for the network, activation function constants and bias values are coded in the binary string. An error function is used to evaluate the fitness of each string. In successive generations of the GA, reproduction, crossover and mutation are performed to minimize the error.

The GA-string will supply the weights, constants of activation functions and bias values, whereas the neural network will compute the expected output. A typical GA-string is shown below:

$$\underbrace{0\,0\,0\,1\,....\,0\,0\,0\,0\,1}_{\text{Synaptic weights}}\,\underbrace{0\,1\,....\,1\,0\,1}_{\text{Bias}}\,\underbrace{0\,1\,0\,0\,1\,......\,1\,0\,1\,0\,1}_{\text{Activaion function constant}}.$$

3. CASE STUDY

Modeling and analysis of clay-bonded moulding sand system (green sand moulding system) is considered in this case study. Conventional statistical regression analyses as well as neural network-based models have been developed to establish the input-output relationships.

The quality of castings in a green sand mould is influenced significantly by its properties, such as green compression strength, permeability, mould hardness, and others, which depend on input parameters. The relationships of these properties with the input parameters, like sand grain size and shape, binder, water etc. are complex in nature.

The clay-bonded moulding sand system is represented as an input-output system as shown in Fig. 2. The process parameters and their levels for the clay-bonded system are shown in Table 1.

3.1. Determination of permeability, green compression strength, mould hardness and bulk density

The experiments were carried out to measure permeability number, green compression strength, mould hardness and bulk density. The standard test specimens (of cross-sectional area a_c, and height h) were prepared with
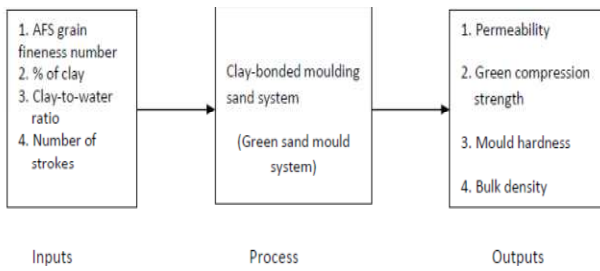
Fig. 2. Clay-bonded moulding sand system.

Table 1

Process parameters and their levels

Sl. No	Parameters (Description)	Notation		Levels		
		Coded	Uncoded	High (+1)	Middle (0)	Low (-1)
1	AFS Grain fineness no.	X_1	A	94	73	52
2	% of Clay	X_2	B	12	10	8
3	Clay: Water Ratio	X_3	C	3	2.25	1.5
4	No. of strokes	X_4	D	5	4	3

different combinations of the variables and their levels as per the 2-level full-factorial design, Central composite design of experiments.

3.2. Conventional statistical regression models, analyses and testing

The experimental data collected as per the DOE was utilized to develop the linear as well as non-linear regression models. The statistical adequacy of the developed models was tested with the help of significance and ANOVA tests. Moreover, the performance of the developed models was tested with the help of test cases [8, 9].

3.2.1. Permeability. Linear regression model based on 2-level full-factorial design and non-linear regression models based on central composite design have been developed for the response – permeability.

Linear model based on 2-level full-factorial design. The linear regression model for the response – permeability with the parameters expressed in coded value is given below.

$$P_{fact-c} = 1075 - 4627X_1 - 1658X_2 + 282X_3 - 2223X_4 + 4.83X_1X_2 + 3.15X_1X_3 + 10.96X_1X_4 + 1.79X_2X_3 + \quad (4)$$
$$3.97X_2X_4 - 2.32X_3X_4 - 2.01X_1X_2X_3 + 1.28X_1X_2X_4 - 0.97X_1X_3X_4 - 5.91X_2X_3X_4 + 5.79X_1X_2X_3X_4$$
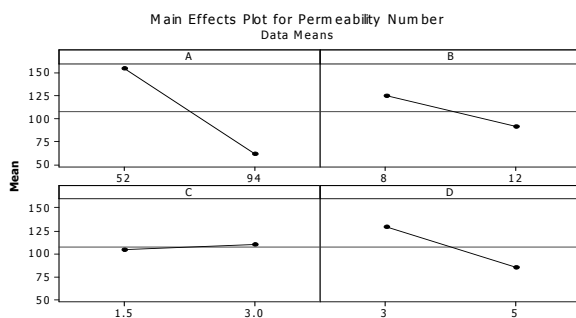

Fig. 3. Main effect plot – permeability.

Figure 3 shows the graphical representation of the effects of main factors. From these figures, it is observed that among the main factors GFN, percentage of clay and number of strokes is found to have negative effect, whereas clay-to-water ratio has a small positive contribution.

The input parameters are coded using the following relationships:

$$X_1 = \frac{A-73}{21}, \quad X_2 = \frac{B-10}{2}, \quad X_3 = \frac{C-2.25}{0.75}, \quad X_4 = \frac{D-4}{1}.$$

where X_1, X_2, X_3 and X_4 represent the input parameters, such as grain fineness number A, % of clay B, clay-to-water ratio C and number of strokes D, respectively, in the coded form. The response equation in the real form can be written as follows:

$$P_{fact} = 2398.35 - 23.216A - 185.165B - 769.43C - 451.795D + 1.79201AB + 8.44065AC +$$
$$4.49427AD + 75.2983BC + 38.83BD + 175.023CD - 0.799412ABC - 0.38335ABD -$$
$$1.90059ACD - 17.362BCD + 0.183894ABCD \quad (5)$$

The advantage of using real form of the equation is that the input parameters need not be converted to their coded form to obtain the predicted value of the response. However, the coefficients of the main and interaction terms will not provide the true picture of their contributions. ANOVA test was conducted to check the adequacy of the developed linear model. The developed regression model is found to be statistically adequate. Statistical adequacy is tested using T-test, ANOVA and R^2 value.

Non-linear model based on central composite design. Permeability of the green sand mould was expressed as the non-linear function of the input process parameters (in coded form) as follows:

$$P_{ccd} = 73.874 - 47.148X_1 - 15.29X_2 + 3.76X_3 - 20.903X_4 + 30.357007X_1^2 + 1.801X_2^2 - 2.325X_3^2 + 4.22X_4^2 + \quad (6)$$
$$4.829X_1X_2 + 3.149X_1X_3 + 10.963X_1X_4 + 1.785X_2X_3 + 3.971X_2X_4 - 2.323X_3X_4$$

Significance test was conducted to examine the effect of different process parameters and their interaction terms on the said response.

The response equation can be written in the uncoded form as follows:

$$P_{ccd} = 11925.1 - 1598A - 3566B + 9.51C - 10566D + 0.07A^2 + 0.45B^2 - 4.13C^2 + 4.22D^2 + \quad (7)$$
$$0.11AB + 0.2AC + 0.52AD + 1.19BC + 1.99BD - 3.1CD.$$

where A, B, C, and D represent the input process parameters - grain fineness number, % of clay, clay: water ratio and no. of strokes, respectively.

To determine input-output relationships in green sand mould system, full-factorial design was adopted for the linear modeling and two other techniques, namely central composite design and Box-Behnken design were used for developing the non-linear models. The adequacy of the developed models has been checked statistically. To validate the results of the developed models, twenty randomly-generated test cases were passed and deviations in prediction were determined. For each response, the predicted values using three different models have been compared separately with their respective target values. Moreover, three models were compared, in terms of % deviation in prediction of different responses.

In the neural network based forward mapping, mould properties are expressed as the functions of input parameters, whereas attempts can also be made to determine an appropriate set of input parameters, to ensure a set of desired properties, in reverse mapping. In the pre-

sent work, the problems related to both the forward as well as reverse mappings in green sand mould system were tackled by using a back–propagation neural network (BPNN) and a genetic–neural network (GA-NN). Batch mode of training had been provided to both the networks with the help of one thousand data, generated artificially from the regression equations obtained earlier by the authors. The performances of the developed models had been compared among themselves for twenty randomly-generated test cases. The results show that GA-NN outperforms the BPNN and that both the NN approaches are able to carry out the reverse mapping effectively.

Fig. 4 compares the performances of two NN-based approaches among themselves and with that of the conventional regression analysis in terms of average absolute % deviation in prediction of the responses (Forward Mapping).

The performances of NN-based approaches are found to be comparable with that of the conventional statistical regression analysis. It is to be noted that GA-NN has performed better than the BPNN for all the responses. It could be due to the fact that the chance of BPNN algorithm for being trapped into the local minima is more, as it works based on the principle of steepest descent algorithm, whereas the chance of the GA-solutions for getting struck at the local minima is less. It is also important to mention that the performances of BPNN and GA-NN are dependent on the nature of error surface to be minimized during the training of the network. Further, the results of reverse mapping showed that the performance of GA-NN was better than BPNN in predicting process parameters.

Fig. 5 compares the performances of BPNN and GA-NN, in terms of average of absolute % deviation in prediction of different process parameters (Reverse Mapping). It is important to note that GA-NN outperformed the BPNN in predicting all the parameters except the number of strokes. The performances of the BPNN and GA-NN might be dependent on the nature of error surface. The aim of reverse mapping is to determine a suitable set of input process parameters, such that a desired set of outputs can be obtained.

4. CONCLUSIONS

The important modeling tools, namely, Statistical regression models and Neural Net work-based models have been discussed in the present paper. The application of these tools is illustrated with the help of case study. The followings conclusions have been made from the present work.

- Design of experiment combined with Response Surface Methodology can be successfully used to analyze the effect of independent variables and their sensitivity on the responses. Further the regression models can be developed, which can be used to predict the response (output) from the known input process parameters.

- More information can be obtained on the input-output relationships with less number of experiments. This will help to reduce manufacturing lead time and the cost of trial runs.
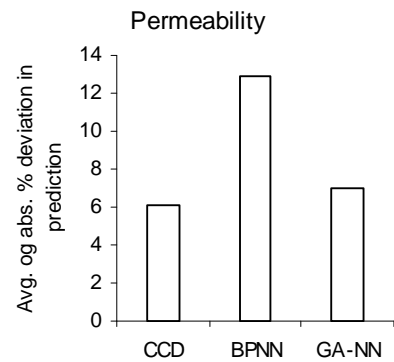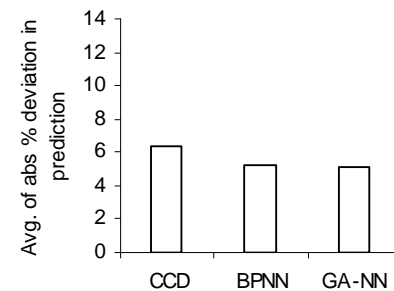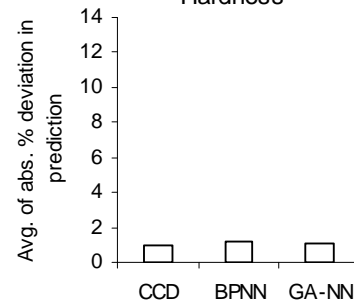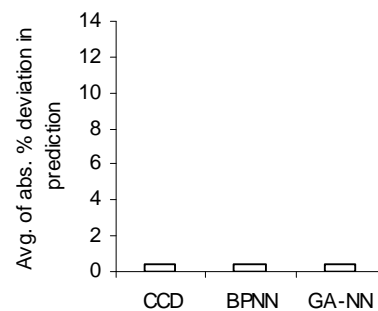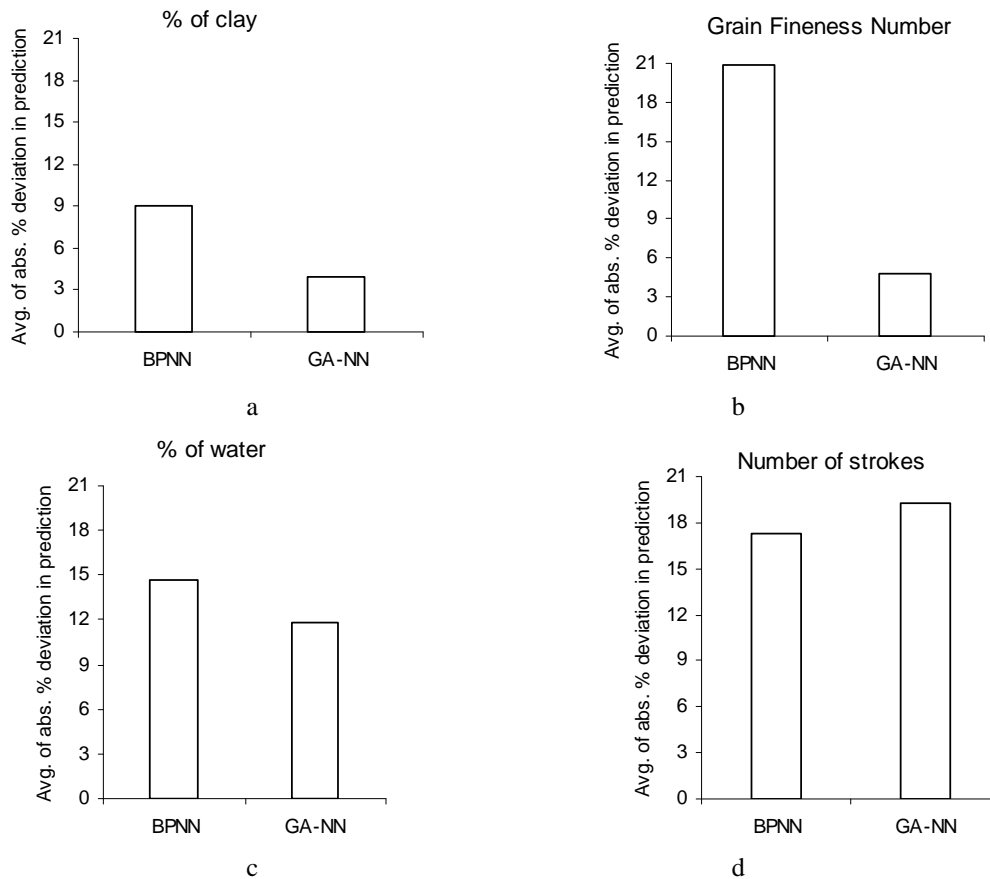

Fig. 4. Comparison of the models in terms of average absolute % deviation in prediction of the responses: *a* – permeability; *b* – green compression strength; *c* – mould hardness; *d* – bulk density.

Fig. 5. Comparison of the models in terms of average absolute % deviation in prediction of the process parameters: *a* − grain fineness number; *b* − % of clay; *c* − % of water; *d* − number of strokes.

- The important requirement of manufacturers is to know the setting of input parameters for a set of desired output Reverse modeling). This can be successfully achieved using soft-computing based modeling. NN-based approaches can tackle the problems of both forward as well as reverse mappings efficiently.

- The literature shows there is a lot of scope in developing the above discussed models in manufacturing. The application of these tools will help the industries to manufacture the product with higher efficiency. The reverse mapping will help to establish on-line control of the process where quick change in the parameters can be made to get the desired output.

REFERENCES

[1] U. Caydas and A. Hascalik, *A study on surface roughness in abrasive water jet machining process using artificial neural networks and regression analysis method*, Journal of Materials Processing Technology, Vol. 202, No. 1−3, 2007, pp. 574−582,.

[2] C. Ko, D.H. Kim and B.M. Kim, *Application of artificial neural network and taguchi method to preform design in metal forming considering workability*, International Jour-

nal of Machine Tools and Manufacture, Vol. 39, No. 5, 1999, pp. 771−785.

[3] A. Krimpenis, P. G. Benardos, G. C. Vosniakos, A. Koukouvitaki, *Simulation-based selection of optimum pressure die casting process parameters using neural nets and genetic algorithms*, International Journal of Advanced Manufacturing Technology, 27 (2006) 509−517.

[4] D.C. Montgomery, *Design and Analysis of Experiments*, 2001, (John Wiley & Sons, New York).

[5] C.F.J. Wu, M. Hamada, *Experiments Planning*, Analysis, and Parameter Design Optimization, 2000, (John Wiley & Sons, New York).

[6] J.E. Lee, S. Rhee, *Prediction of Process Parameters for Gas Metal Arc Welding by Multiple Regression Analysis*, Proc. I MechE Part B Journal of Engineering Manufacture, Vol. 214, 2000, pp. 443−449.

[7] D.K. Partihar, *Soft Computing*, Narosa Publishing House, New Delhi, India, 2008.

[8] M.B. Parappagoudar, D.K. Pratihar, G.L. Datta, *Linear and non-linear statistical modelling of green sand mould system, International Journal of Cast Metals Research*, Maney Publications, Vol. 20, No. 1, 2007, pp. 1−13.

[9] M.B. Parappagoudar, D.K. Pratihar, G.L Datta, *Forward and reverse mappings in green sand mould system using neural networks*, Elsevier Publications, Applied Soft Computing, Vol. 8, No. 1, 2008, pp. 239−260.