Modeling Performance of Response Surface Methodology and Artificial Neural Network

SADA, SO

Department of Mechanical & Production Engineering, Faculty of Engineering, Oleh Campus, Delta State University. Nigeria.
Email: samorosada@gmail.com; 08034085480

ABSTRACT: In recent years, response surface methodology (RSM) which is a statistical technique and artificial neural network (ANN) a soft computing technique have been highly used for modelling, simulation and optimization of several physical processes in engineering. Both RSM and ANN strategies have particular computational properties that makes them suitable for making predictions, but differ in their extrapolation and interpolation capabilities on complex non-linear processes, and thus potentially conflict in their predictive accuracy. This study models and compares the capabilities of RSM and ANN in predicting the tensile strength of a 6 mm thick mild steel gas tungsten arc welded plate based on the effects of input variables such as weld current, weld speed, gas flow rate and filler rod. The RSM and ANN based models for prediction were compared using the coefficient of determination criteria. With a higher value of 0.836, the ANN model proved to be a better modeling technique than the RSM model.

Modeling has been a useful tool for many engineering designs and analysis. Depending on its application, the definition of modeling may vary, but the basic concept remains unchanged. Kuhns and Johnson, 2013 defined predictive modelling as "the process of developing a mathematical tool or model that generates an accurate prediction". Welding processes are examples of complicated systems in which modelling and optimization have been extensively applied. Many attempts have been made in the last decade using different techniques, to understand and estimate the effect of process parameters on certain weld features. These techniques can be grouped into two (2), statistical techniques and soft computing techniques. Statistical techniques involve determining explicit information only. They do not involve any validation mechanism. However, data mining methods can discover implicit knowledge through data analysis. Among the most prominently used statistical techniques are Response Surface Methodology, Taguchi and Factorial Designs. On the other hand, Soft Computing (SC), a concept which was introduced in the early nineties by Dr. Zadeh, is an evolving collection of artificial intelligence (AI) methodologies aimed at exploiting the tolerance for imprecision and uncertainty inherent in human thinking and in real life problems, in order to deliver an efficient and optimal solutions and in addition explore valuable design knowledge. (Saridakis *et al*, 2008). Soft Computing

(SC) encompasses many techniques amongst which Fuzzy Logic (FL), Artificial Neural Networks (ANN), and Genetic Algorithms (GAs) are the core methodologies. Fuzzy logic (FL) reasoning systems, are based on knowledge-driven reasoning, while Artificial Neural Networks (Neuro Computing), and Genetic Algorithms (Evolutionary Computing) are data – driven search and optimization approaches. SC yields rich knowledge representation (symbol and pattern), flexible knowledge acquisition (machine learning), and flexible knowledge processing (inference by interfacing symbolic and pattern knowledge). Research has been deployed in the direction of applying SC to engineering design in the context of replacing existing analytical models with approximated models or meta-models. (Simpson *et al.,* 2001) investigated the potential of soft computing techniques by comparing them to the statistical techniques in meta-modeling and they provided some recommendations about their appropriate uses. Besides meta-modeling, SC techniques may be combined with expert and knowledge-based systems.

Recent studies, have shown that response surface methodology (RSM) which is a statistical technique and artificial neural network (ANN) a soft computing technique have been highly used for modelling, simulation and optimization of several physical processes in engineering. They both offer huge

advantages over the conventionally followed one factor-at-a-time approach, and are considered an effective modelling tool for solving complex nonlinear multivariable systems. They don't require the explicit expressions of the physical meaning of the system or process under investigation. They both develop or approximate the functional relationships between input and the output variables of the process applying experimental data. Both RSM and ANN strategies have particular computational properties that makes them suitable for making predictions, but differ in their extrapolation and interpolation capabilities on complex non-linear processes, and thus potentially conflict in their predictive accuracy. Hence the need for a study of their comparative performance. The response surface methodology (RSM) has been a widely used approach for modeling of welding processes (Sada, 2018). It encompasses a group of statistic based approaches which cuts across applications in model building, experimental designs, exploration of factor effects and searching for optimal conditions (Kalil *et al.,* 2000). The experimental responses in RSM are fitted to a quadratic function and one of its advantages is its ability to optimize a process and interpret the interactive effects of the process variable on the response using a lesser number of experiments. It requires good prior knowledge or extra preliminary experiments to fix the search criteria, and works only for a nonlinear quadratic correlation. On the other hand, ANN is one of the most widely used AI techniques and has been successfully employed by researchers in areas such as function approximation, classification, association, pattern recognition, time series analysis; signal processing, data compaction, non-linear system modeling, prediction, estimation, optimization and control (Joshi *et al.,* 2014). For manufacturing processes where no satisfactory analytic model exist or a low order empirical polynomial model is inappropriate, neural networks offers a good alternative approach. It has the ability to learn the mapping between a set of input and output values. ANN is more efficient than RSM based on certain features: (i) its ability to process highly nonlinear complex systems unlike RSM which is limited to quadratic approximations (Basheer and Hajmeer, 2000) (ii) its excellent ability in data fitting and prediction, (iii) it does not require a standard experimental design to develop a model or a prior description of proper fitting function and it has the ability of universal approximation, i.e. approximation of almost all kinds of nonlinear functions (iv) ANN is structured in nature and useful for getting more insight information, i.e. it has the ability to provide sensitivity analysis and to reveal the interactive effect of two factors on the system [Jev *et al.,* 2005, Gav *et al.,* 2006, Sew *et al.,* 2015]. (V) If the process under

analysis changes, new data can be added and the neural network can be retrained. This is much easier than determining new models or rules. This study models and compares the capabilities of RSM and ANN in predicting undercut weld defects in a Gas Tungsten Arc Welded mild steel rod.

MATERIALS AND METHODS

The gas tungsten arc welding process which is reputable for its quality of weld, and a mild steel plate of 6mm thickness was employed for the experiment. In addition, argon gas, was selected as the shielding gas. The mild steel sheet was cut into the required sizes (300mm x 150mm) using a hacksaw, thereafter cleaned and clamped for the welding experiment. In this study, tensile strength was selected as the response while weld current (240-300amp), weld speed (256-270mm), gas flow rate (8.5-10L/min) and filer rod diameter (3.2-4.0mm) were the independent variables. Thirty experiments was performed using the thirty experimental runs generated based on central composite design (CCD) of the RSM. The welded samples was tested for tensile strength in a universal testing machine, recorded and tabulated as shown in Table 1.

Table 1: Experimental Results

Exp No	Weld Current (amp)	Weld Speed (mm)	Gas Flow Rate (L/min)	Filler Rod Dia (mm)	Tensile Strength (N/mm2)
1	240	270	8.50	4.0	326.4
2	270	249	9.25	3.6	344.7
3	210	263	9.25	3.6	336.3
4	270	263	10.75	3.6	305.0
5	270	263	9.25	4.4	343.0
6	240	256	8.50	4.0	335.0
7	300	270	8.50	3.2	351.0
8	270	263	9.25	3.6	317.0
9	270	263	9.25	3.6	324.0
10	270	263	9.25	2.8	323.3
11	270	277	9.25	3.6	322.2
12	240	270	8.50	3.2	339.0
13	300	256	8.50	4.0	330.2
14	240	256	10.00	4.0	340.0
15	300	256	10.00	4.0	331.0
16	300	256	8.50	3.2	328.0
17	240	256	8.50	3.2	351.0
18	300	256	10.00	3.2	321.0
19	300	270	10.00	3.2	316.0
20	300	270	8.50	4.0	326.0
21	240	270	10.00	4.0	322.2
22	270	263	9.25	3.6	328.0
23	300	270	10.00	4.0	317.7
24	240	270	10.00	3.2	322.8
25	240	256	10.00	3.2	334.9
26	270	263	7.75	3.6	325.6
27	270	263	9.25	3.6	332.2
28	330	263	9.25	3.6	327.0
29	270	263	9.25	3.6	316.5
30	270	263	9.25	3.6	310.5

RSM Based Modeling: RSM is a collection of mathematical and statistically techniques effective for modeling and analysis of problems with several process variables, widely used in applications such as design, development, and formulation of new products, as well as improvement of existing product designs (Montgomery 2008). RSM is very effective in determining the main, quadratic and interactive effects of the operating variables upon the response or responses as the case may be. The most extensive applications of RSM are in situations in which some performance measure or quality characteristics of a product or process (response) is being influenced by several input variables (independent variable).

With no knowledge of the form of the relationship between the response and the independent variable, it's first approach is to find a suitable approximation for the true functional relationship between response (y) and the set of independent variables (x) (Thepsonthi and Ozel 2012). Usually a low-order polynomial in some relatively small region of the independent variable space provides a suitable approximation of the true form of the response function. In the case of curvature in the response surface, a higher degree polynomial can be used.

For this model, a second-order polynomial regression equation was used to fit the experimental data and to describe the relevant model terms. The regression analysis and process optimization were performed using the Statistical Software package "Design Expert". The model, which also includes the linear model, was described by the Equation 1.

$$y = \beta_o + \sum_{i=1}^{k} \beta_i x_i + \sum_{j=1}^{k} \beta_{ii} x_i^2 + \sum \sum_{ij} \beta_i x_i x_i + \varepsilon \quad for\ i < j \quad 1$$

Where y is the predicted response; β_o is the intercept constant; β_j, β_{jj} and β_{ij} are the interaction coefficients of the linear, quadratic and the second-order terms, respectively. k is the number of factors; x_i and x_j are variables (i and j range from 1 to k); ε is the error. The second-order model is widely used in RSM for several reasons. Amongst which are: flexibility, ability to take on a wide variety of functional forms and the suitability of using least square method in estimating the coefficient (β). Based on the experimental result in Table 1, the second order RSM model for the response was formulated using the estimated regression coefficients as shown in Equation 2.

$TS = 5065.611 - 5.645A - 33.196B + 138.049 - 76.282D + 0.015AB - 0.050BC + 0.068AD - 0.552BC - 0.8445BD + 14.073CD + 0.003A ** 2 + 0.069B ** 2 - 1.992C ** 2 + 20.889D ** 2$ (2)

Where TS = Tensile Strength

Analysis of variance (ANOVA) was performed using the design expert software, to evaluate the significance of the model based on the Fischer (p-value). A large F-value along with a P value less than 0.05, confirms that the model fits the experimental data significantly. Also variable with a large F-value and $p < 0.05$ are considered significant. The model F value of 3.47 and P value of 0.0113 as shown in Table 2, shows that the model is significant, haven obtained a P value less than 0.05. Also the following model terms, weld current, weld speed and gas flow rate were found to have had the most significant effect on the tensile strength judging from their F and P values as well. In addition, the goodness of fit statistics was used in validating the model. A coefficient of determination (R^2), Adjusted (R^2), and predicted (R^2) values of 0.764, 0.750 and 0.8110 respectively was obtained. Signifying that 76.4% of the variability was accounted for by the model, and that the model is 81.1% suitable for making prediction.

Table 2: Analysis of Variance (ANOVA)

Sources	Sum of Square	df	Mean Square	F-value	p-value	
Model	2716.23	14	194.02	3.47	0.0113	significant
A-Weld Current	198.09	1	198.09	3.54	0.0794	
B-Weld Speed	376.44	1	376.44	6.73	0.0203	
C-Gas Flow Rate	622.71	1	622.71	11.13	0.0045	
D-Filler Rod Dia.	0.7526	1	0.7526	0.0135	0.9092	
AB	162.24	1	162.24	2.90	0.1092	
AC	20.36	1	20.36	0.3639	0.5553	
AD	10.64	1	10.64	0.1902	0.6689	
BC	134.27	1	134.27	2.40	0.1422	
BD	89.54	1	89.54	1.60	0.2252	
CD	285.19	1	285.19	5.10	0.0393	
A²	241.49	1	241.49	4.32	0.0553	
B²	320.29	1	320.29	4.32	0.0303	
C²	34.43	1	34.43	5.72	0.4450	
D²	306.38	1	34.43	0.6153	0.0335	
Residual	839.28	15	55.95			
Lack of Fit	510.15	10	51.01	0.7750	0.6589	Not significant
Pure Error	329.13	5	65.83			
Cor Total	3555.51	29				

Artificial Neural Networks Model: Artificial Neural Network (ANN) are defined as a functional demonstration of the biological structures of the human central nervous system. They are commonly used for prediction, matching, identification, pattern recognition, classification, and optimization of problems (Radhakrishnan and Mohamed 2010). ANN comprises of interconnection of simple, non-linear computational elements which are called nodes or neurons. A neuron typically consists of three components: (i) a group of weights, (ii) a summation function, and (iii) a non-linear activation function f(R). ANN can contain a large number of neurons, neurons linked with variable weights (Prasad and Prasanna

2008). The computational behavior of an ANN model is determined by the values of the weight associated with each of the neurons. This behavior can be derived usually on trained sample data, normally represented in the form of an input matrix, X, and the corresponding output vector, Y. The job of the ANN is to learn the input-output relations embedded in X→Y through a learning process. The operations of the processing units consist of a number of steps. First, each input field xi is multiplied by a corresponding weight ω_t, $\omega_t \in W = \{ \omega_1, \omega_2…, \omega_n\}$. The product of each input field and its corresponding weight are then summed to produce the cumulative weighted combination R, as shown in Equation 4:

$$R = \omega_1 x_1 + \omega_2 x_2 … … . . \omega n x n = \sum_{t=1}^{n} \omega t x t = W.X$$
(3)

In order to adjust the behavior of the neurons, a quantity called the bias can be used as threshold.

$$R = \omega_1 x_1 + \omega_2 x_2 … … . . \omega n x n + \theta = \sum_{t=1}^{n} \omega t x t + \theta$$
(4)

Depending on the behaviour of the system being modelled, the function f (;) can take many forms, some of which are linear, sigmoid, exponential, etc. The computed value of y can serve as input to other neurons or as an output of the neural network depending on its position in the network configuration. The training process usually involves minimizing the sum of square error between actual and predicted output. The behaviour in the available training data is captured by continuous adjustment and by finally determining the weight connecting neurons in adjacent layers. Backward error propagation (also called the back propagation or back-prop) algorithm is the most commonly used learning algorithm. This algorithm uses the gradient descent method in its implementation (Anaraki *et al*., 2008). Theoretically, a limited amount of training data points does not guarantee that a neural network will generalize the "true" behaviour desired. Cross-validation process is used in order to verify the result of generalization. It involves the sectioning of the parent database into three subsets: training, test, and validation (Basheer and Hajmeer, 2000). Training is done by feeding teaching patterns to the network and letting the network to change its weighting function based on some previously defined learning rules. There are two types of learning: supervised and unsupervised. In supervised learning the network under study is trained by giving it inputs and output patterns during supervised learning whereas for the unsupervised learning the output of the network is trained to respond to input patterns. The training subset usually includes all the data belonging to the problem domain and is used in the training phase to update the weights of the network. During the learning process, the test subset uses data distinct from those used in the training, in checking the network response for untrained data. Based on the performance of the ANN on the test subset, the architecture may be changed and/or more training cycles applied. The validation subset is the third portion of the data, it usually includes sample data different from those in the other two subsets. This subset is used after selecting the best network to further examine the network or confirm its accuracy before being implemented in real-life systems.   Several ANN architectures, networks and algorithms have been developed, amongst which are the Perceptron, Hopfield and Hamming Network and the various types of algorithms used for training ANNs include Back propagation, Delta Learning Rule, Hebb Learning Rule and Bayesian Regularization Algorithm (Du and Swamy 2013). There are also several transfer functions such as Hardlim, Tansig, Purelin and Logsig function utilized in ANN models.
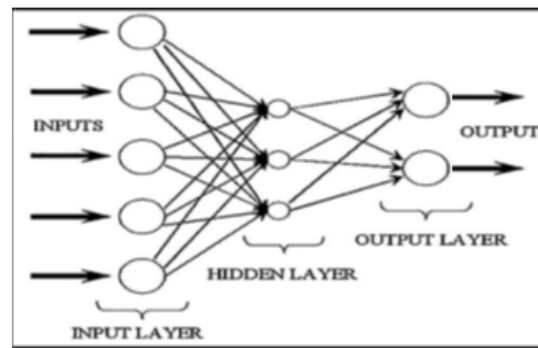

Fig 1. Multi-layered perceptron architecture

Among the several architectures are: A single layer perception (SLP) network, which also includes are single layer of output nodes and inputs, will suffice for the simplest form of ANN. These are directly fed to the outputs through a series of weights, but a Multi-layer Perceptron (MLP) also known as feed forward neural networks (FF network) which is a three-layered network as shown in Figure 1, is perhaps the most commonly and widely used models in many engineering applications. The three layered network comprises of the input layer which is the first layer that establishes the first contact points to the data, the hidden layer, followed by the last which is the output layer responsible for presenting the result of the ANN to the outside world. The output of the last hidden layer neurons are fed into the input of the output layer neurons. Succeeding layers in the network sums the inputs of previous layers, adds a bias to the sum and apply the activation function to produce its own output.

In this study, the multilayer perceptron (MLP) or feed-forward neural network (FF) along with Trainlm, a training function that updates weights and bias values based on Levenberg-Marquardt back propagation (LMB) algorithm, was used for training the network. Neural Network Toolbox 8.0 of MATLAB mathematical software was used for simulation. The same experimental data, earlier used for the RSM design, were also employed in designing the artificial neural network. The data were divided into three groups, in the training set 70%, in the validation set 15% and in the test set 15%. The externally normalized input values were normalized between 0 and 1 for the reduction of network error and higher homogeneous results and then forwarded from the input layer to the hidden layer and then to the output layer to predict the response. Evaluating the network performance, ANN output for test input data are compared with experimentally obtained data. As mentioned earlier in step two, If the results are not satisfactory, the network is re-trained. If the test results are good enough training parameters is saved. In assessing the developed model, the mean squared error (MSE), an error function which measures the performance of the ANN model and the correlation coefficient (R^2) on the unseen validation data were used as the performance criterion to show the effectiveness of the trained network. As Figure 2 illustrates, the variation of the MSE during the training was achieved in 5 epochs and the training was terminated.
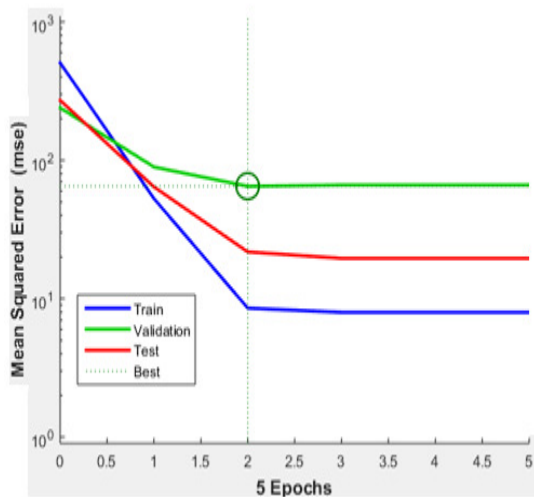
The best training performance is 64.9174 at epoch 3 which is acceptable. The similar characteristic curve for the test and the validation were also observed, suggesting no significant over-fitting.
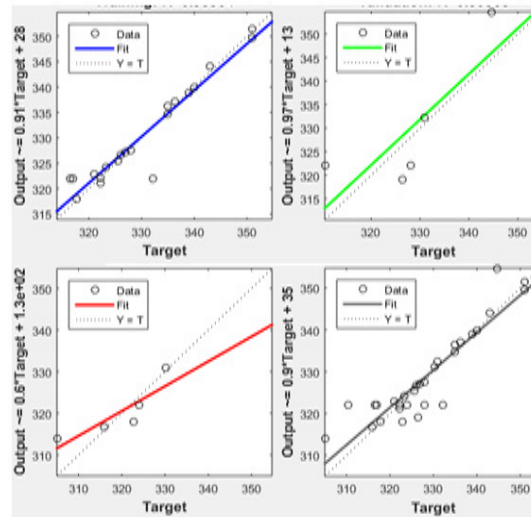

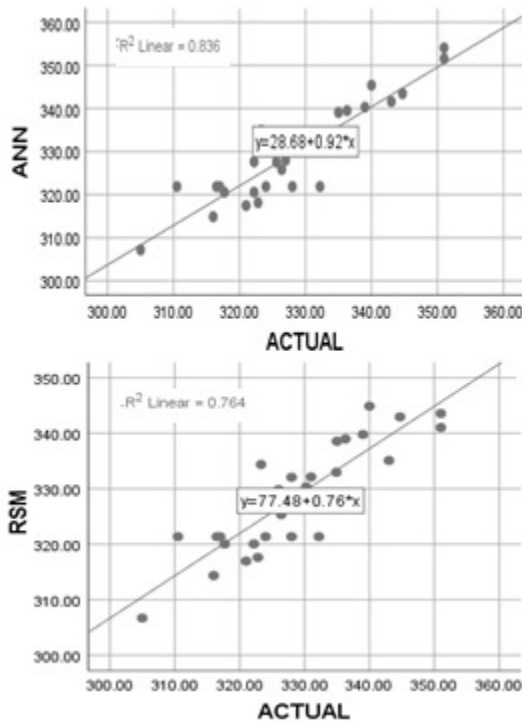
Fig 3 – ANN Model Showing the Regression Plots

Fig 2: MSE for training, validation, and test dataset computed using LMB

Fig 4: Reliability Plot Showing Actual Versus Predicted Values using both models
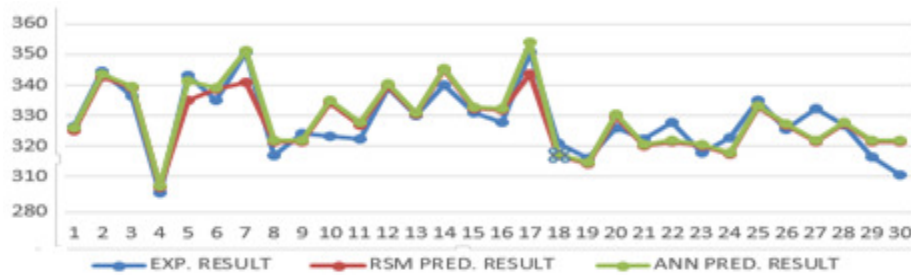
Fig 5. Comparison of the RSM and ANN models response with experimental results

Figure 3 describes the ANN regression plot for training, validation, testing, and overall prediction set in the form of network output versus experimental. The correlation coefficients 'R' for training, validation, and testing, were 0.959, 0.804, 0.860, respectively, whereas the overall prediction set was 0.915, which confirms that the ANN model is satisfactory for interpolating the experimental data.

Comparison of The RSM and ANN Models: In this work, a comparison of the capabilities of both the techniques (ANN and RSM) was made, and the estimation was examined on the basis of their root-mean-squared error (RMSE) and coefficient of determination R^2. The predictive models developed by RSM and ANN were compared on the basis of their prediction accuracy. As shown in Figure 4, a comparison between the two models was made according to their evaluation method, results shows that the coefficient of determination denoted by R^2, of 0.764 and 0.836 was obtained for RSM and ANN, respectively. Thus, higher R^2 and lower RMSE values for ANN, demonstrates that ANN is much accurate in making prediction than RSM.

Figure 5 shows a comparison of the results of RSM and ANN prediction performances with experimental results. Although each model has similar tracking ability ANN model is closer to the real values. This amount represents the superiority of ANN over RSM.

Conclusion: An attempt has been made to apply the RSM and ANN in predicting the tensile strength of a 6mm
Mild steel gas tungsten arc welded plate. A RSM model and an ANN model based on the Levenberg-Marquardt algorithms was developed and compared with the experimental results obtained earlier to evaluate the performance of both models. The model results are compared with each other in terms of the performance criteria regression coefficient (R^2) and root mean squared error (RMSE). Result showed that the ANN model, had the highest

coefficient of determination (R^2): 0.836, which proves its superiority over the RSM model.

REFERENCES

Anaraki, MT; Sanjari, M; Akbarzadeh, A (2008). Modeling of High Temperature Rheological Behavior of AZ61 Mg-alloy using Inverse Method and ANN. *J. Mat. & Design* 29: 1701-1706

Basheer, IA; Hajmeer, A (2000). Artificial Neural Networks; Fundamentals, Computing, Design and Application. *J. of microbiological methods*, 43(1), 3-31.

Du, KL; Swamy, MN (2013). Neural Networks and Statistical Learning. Springer Science & Media.

Bhanu, P; Prasana SRM (2008). Speech, Audio, Image and Biomedical Signal Processing using Neural Networks. Studies in computational Intelligence, vol 83, pp. 169-188.

Dutta, P; Kumar, D; (2007). Modelling of GTAW process using Conventional Regression and Artificial Neural Network-based approaches, *J. of Mat. Processing Tech.* 184, 56-68.

Montgomery, DC (2008). Design and analysis of experiments: Wiley.

Gevrey, M; Dimopoulos, I; Lek, S (2006). Two-way interaction of input variables in the sensitivity analysis of neural network models. Ecol Model. ; 195:43–50.

Jaiswal, S; Benson, E; Bernard, J (2005). Neural Network Modelling and Sensitivity Analysis of a Mechanical Poultry Catching System. *Biosys Eng.* 2005; 92:59–68.

Joshi, SR; Ganjigatti, JP; Kulkarni, S (2014). Application of Statistical and Soft Computing based Modeling and Optimization Techniques for Various Welding Processes. *Int. J. of Latest*

Trends in Engineering and Technology, Vol. 3 Issue 4, pp 375-385.

Kalil, S; Maugeri, F; Rodrigues, M (2000). Response Surface Analysis and Simulation as a Tool for Bioprocess Design and Optimization. *Process Biochem.* 35:539–550.

Kuhn, M; Johnson K (2013). Applied Predictive Modeling, Springer Science & Media; 2013.

Radhakrishnan, VR; Mohamed, AR (2010). Neural Networks for the Identification and Control of Blast Furnace Hot Metal Quality. *Journal of Process Control*; 10: 509-524.

Sada, SO 2018 "Optimization of Weld Strength Properties of TIG mild steel welds using RSM" Nigerian J. of Tech. vol 37, no 2, pp 407-415.

Saridakis, KM; Dentsoras, AJ (2008) Soft Computing in Engineering Design – A Review/ Advanced Engineering Informatics 22, 202–221

Sewsynker, Y; Kana, EB; Lateef, A (2015). Modelling of Biohydrogen Generation in Microbial Electrolysis cells using Artificial Neural Networks (ANNs). Biotechnol Equip. ;29:1208–1215.

Simpson, TW; Peplinski, JD; Koch, PN; Allen, JK (2001). Metamodels for computer-based engineering design: Engineering with Computers, vol 17, pp. 129-150.

Thepsonthi, T; Özel, T (2012). Multi-objective Process Optimization for Micro-end Milling of Ti-6Al-4V titanium alloy, *Int J Adv Manuf Techl* 63, 903–914.

Witek-Krowiak, A; Chojnacka, K; Podstawczyk, D; Pokomeda, K (2014) App of RSM and ANN Methods in Modelling and Optimization of Bio Process. *J. of Bio Tech.* vol 160, pp150-160

Yang, LJ; Chandel, RS; Bibby, MJ (1992). The Effects of Process Variables on the Bead Width of Submerged-arc Weld Deposits. *J. of Mat. Proc Tech*, 29, 133–144.