Use of the Circle-Segments Method as a Data Visualization Tool for an Artificial Neural Network

Shir Li Wang   **Chee Peng Lim***   **Zalina Abdul Aziz**

School of Electrical and Electronic Engineering
University of Science Malaysia, Engineering Campus
14300 Nibong Tebal, Penang, MALAYSIA
Email (corresponding author*): cplim@eng.usm.my

Abstract

Process modeling and prediction is one of the major tasks in many industrial applications. The Multi-Layered Perceptron (MLP) neural network has been a popular approach in process modeling and prediction, and has produced good results. One of the disadvantages of the MLP is that it is unable to provide a visualization effect of the underlying relationships between the input and output data. In this paper, we propose to use the circle-segments method as a visualization tool for the MLP. The applicability of the hybrid MLP and circle-segments approach is demonstrated using a case study on a closed-loop disk drive head system. The performance is compared with that from the Response Surface Methodology (RSM). From the results obtained, the MLP network shows a better prediction capability as compared with the RSM. In terms of visualization, the circle-segments method is able to overcome the limitations of contour plots in RSM in disclosing the relationships of the input-output data.

Keywords

Artificial neural networks, Multi-Layer Perceptron, circle segments, design of experiments, data visualization

1. Introduction

Process modeling and prediction is one of the major tasks that has applications in many fields [1-6]. Various methods have been applied to solve this problem, as the solution provides a way on understanding how the process works. Artificial neural networks (ANNs) are one of the popular methods used in process modeling and prediction owing to its ability to learn from data samples. With the given experimental data, an ANN is able to learn and form the relationship between the pairs of inputs and outputs with minimum understanding of the underlying process. Another classical method known as response surface methodology (RSM) is often used in unison with ANNs [2-6]. As the ANN is deficient in providing a way to visualize the relationship of outputs and inputs, the RSM can provide two-dimensional (contour plot) or three-dimensional (response surface plot) for visualizing the changes of response(s) against factors. The RSM, however, is limited to produce a graphical plot for only two factors. If a process involves multi-responses and multi-factors (exceed 2 factors), the RSM is not able to provide a visualization effect simultaneously in the same graph. In order to solve this problem, the circle-segments method is proposed as a visualization tool for the ANN.

In this paper, the Multilayer Perceptron network is used to demonstrate the applicability of ANNs integrated with the circle-segments approach in process modeling and prediction. In section 2, the MLP and circle-segments method are briefly described. A case study on a closed-loop control of a disk drive head system is presented in Section 3. The RSM is used to provide a performance comparison in the aspect of data visualization. The results obtained are analyzed and discussed in Section 4. A summary of the work is presented in Section 5.

2. Approach and Methods

As shown in Figure 1, the MLP consists of three main layers, i.e., input layer, hidden layer and output layer. Each layer contains neurons which are linked to neurons in other layers through the weight and bias values. The network learns the relationship between the pairs of inputs-outputs (factors-responses) vectors by altering the weight and bias values.The circle-segments method has been described in [9]. It is used to display the history of the data. The data samples are arranged in such a way that the oldest data samples are at the middle of the circle and the most recent samples are at the outside of the circle. However, in this paper, the circle-segments method is used in a different way, whereby it is used to correlate the inputs (factors) and outputs (responses).

Assume that a process involves k-dimensional data samples which consist of multi-factors and multi-responses. If n dimensions represent the number of responses, the remaining $k-n$ dimensions represent the factors. Figure 2 shows the circle-segments method to represent the data samples as mentioned

above. Each dimension is represented by a segment. The maximum and minimum values of the dimensions are scaled to the upper and lower values of the color map, and the values within it are mapped linearly.
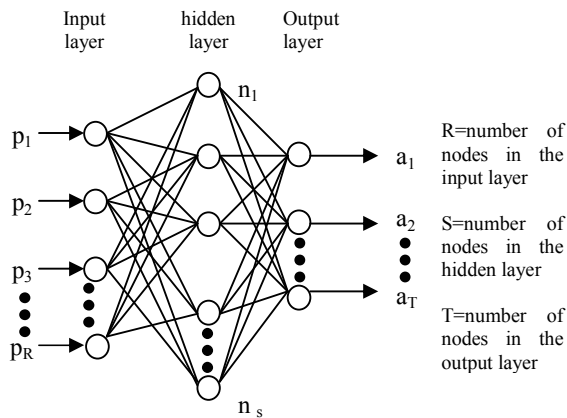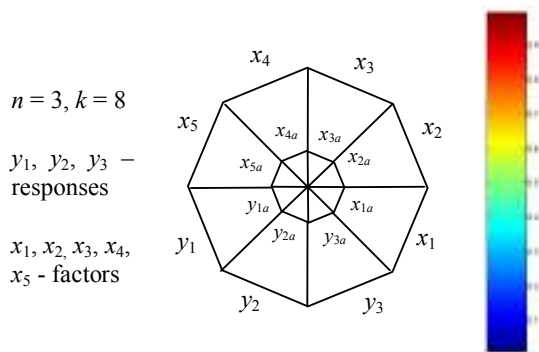

Figure 1 – The Multilayer Perceptron Network

R=number of nodes in the input layer

S=number of nodes in the hidden layer

T=number of nodes in the output layer

$n = 3, k = 8$

$y_1, y_2, y_3 -$ responses

x_1, x_2, x_3, x_4, x_5 - factors

Figure 2 – Circle-segments with 3 responses and 5 factors

3. Experiment

An engineering design case study in control systems is employed to evaluate the models of the RSM and MLP-circle-segments in process modeling and prediction. The goal of control engineering design is often to obtain configuration, specification and identification of the key parameters of a proposed system to meet an actual need [10]. As mentioned in [10], the process of control system design can be summarized as per flow chart in Figure 3. The proposed methods is used in steps 6–7, whereby the key parameters of the controller will be selected as factors, while the performance of the system is the response(s).

A closed-loop disk drive head system with an optional velocity feedback is adopted from [10]. The goal is to find the combination of key controller parameters (K_1 and K_a) to meet requirements of performance. Figure 4 shows the overall system.
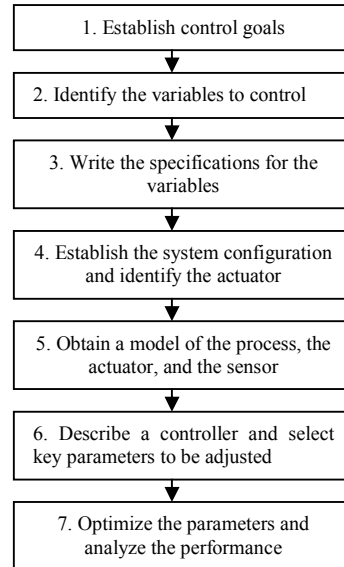

Figure 3 – The Control System Design Process

For the system shown in Figure 4, K_a and K_1 are the factors while the overshoot (*OS*), settling time (*ST*), and maximum response to a unit disturbance (*RD*) are the responses. As stated in [6], it is difficult to find the combination of factors which fulfill all of the responses simultaneously because varying a factor may not cause desired effect to all of the responses. Therefore, the constrained-optimization approach is applied to find the region which fulfills certain constraints.

The closed-loop disk drive head with an optional velocity feedback system was simulated using MATLAB. The range of K_a and K_1 were divided equally into ten levels for 0–200 and 0–0.10 respectively. Since each factor had ten levels, the total number of factor combinations were 100 (10^m, m=number of factors). The combinations of factors were simulated using MATLAB and the responses (*OS*, *ST* and *RD*) were recorded. Then, the pairs of factors-responses were used as the inputs-outputs data to train the MLP. Among the 100 samples collected, 70 samples were used as the training set while the remaining samples were used as the validation set. The validation set was used to monitor the generalization of the MLP. Another 10 samples were simulated to be used as the testing set. The selection of MLP parameters was done using the trial and error approach.

By using the same range of K_a and K_1, the RSM was applied. Table 1 shows the factors involved and their corresponding values. By using the model formed by the RSM, the responses (*OS*, *ST*, and *RD*) for the same set of training set, validation set and testing set of the MLP were determined. The performances of the MLP and RSM were evaluated using the mean squared error (MSE) and R-squared prediction ($R^2_{prediction}$), as shown in Table 2.
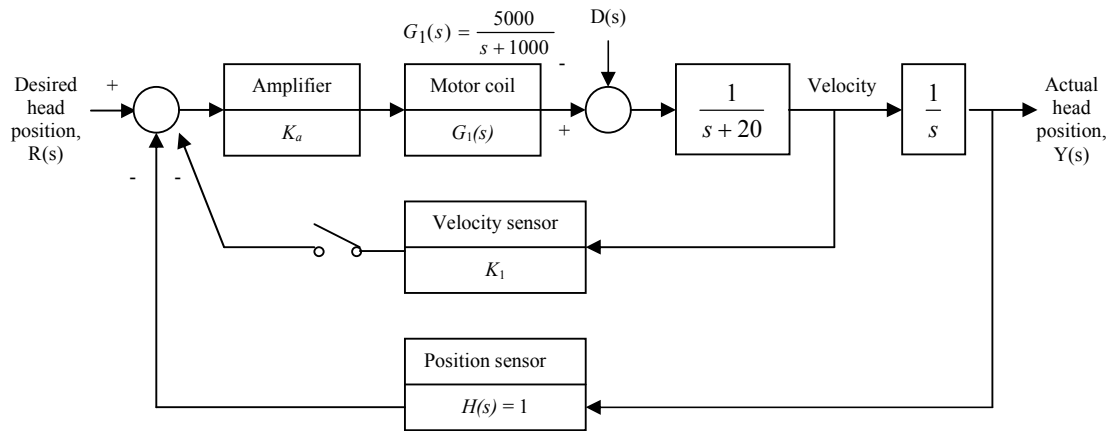
$$G_1(s) = \frac{5000}{s + 1000}$$

Figure 4 – The Closed –Loop Disk Drive With An Optional Velocity Feedback

Table 1 – Factors in Coded and Actual Values and Relevant Responses

Factors		Coded	Values			Responses	Response spec
	-1.41421	-1	0	1	1.41421		
K_a	20	46	110	174	200	Overshoot (*OS*)	< 0.05
K_1	0.010	0.023	0.055	0.087	0.100	Settling Time (*ST*)	< 0.25ms
						Maximum response to a unit disturbance (*RD*)	< 0.002

Table 2 – Statistical Measurement for the MLP and RSM

	Responses			Average
	OS	*ST*	*RD*	
$R^2_{prediction} - MLP(\%)$	99.14	97.80	87.21	94.72
$R^2_{prediction} - RSM(\%)$	72.83	83.28	89.07	81.73
Improvement (%)	26.31	14.52	-1.86	12.99
MSE_tes (MLP)	0.000038	0.004800	1.00×10^{-6}	0.001613
MSE_tes (RSM)	0.000693	0.005769	3.51×10^{-7}	0.002154
Improvement (%)	-	-	-	25.12

The $R^2_{prediction}$ value provides an indication of the predictive capability of the model formed. It is calculated using equation (1), where the SSE is the error sum of squares and SST is the total sum of squares.

$$R^2_{prediction} = 1 - \frac{SSE}{SST} \quad (1)$$

$$SSE = \sum_{i=1}^{N}(t_i - y_i)^2 \quad (2)$$

$$SST = \sum_{i=1}^{N} t_i^2 - \frac{\left(\sum_{i=1}^{N} t_i\right)^2}{N} \quad (3)$$

where t_i is the target/actual output, y_i is the predicted output, and N is the data sample size.

4. Result and Discussion

The RSM was used to model the responses based on the details in Table 1. The fitted equation for *OS*, *ST* and *RD* are shown in equations (4) – (6).

$$OS = 0.00614K_a - 0.03292K_1 - 0.00438K_a^2 \\ + 0.03014K_1^2 - 0.01225K_aK_1 \qquad (4)$$

$$ST = 0.2850 - 0.1406K_a + 0.1089K_1 \\ + 0.1116K_a^2 + 0.0191K_1^2 - 0.0337K_aK_1 \qquad (5)$$

$$RD = -0.00182 + 0.002391K_a - 0.001612K_a^2 \\ + 0.000227K_1^2 \qquad (6)$$

For the MLP, it was found, by trial and error, that network with 13 hidden neurons and 18200 epochs was suitable to model the process. Both of the RSM and MLP had been tested using the data set. Table 2 presents the results of $R^2_{prediction}$ and MSE of the testing set (MSE_tes) for both the models, thus verifying the performance of the models.

The results in Table 2 indicate that both the RSM and MLP had successfully modeled the process, with $R^2_{prediction}$ above 70% for each response. It is found that the result of $R^2_{prediction}$ of the MLP improved a lot as compared with that of the RSM, except for *RD*. However, the difference of $R^2_{prediction}$ for *RD* between the RSM and MLP was small (2%), while the improvements for *OS* and *ST* achieved by the MLP were, respectively, 26% and 15%. Based on the MSE_tes, the MLP had a smaller error in predicting responses except for *RD*. However, the MSE_tes for *RD* was so small for both of the models (in $10^{-6} \sim 10^{-7}$), that it didn't make a significance difference. In general, the MLP outperformed the RSM, with higher $R^2_{prediction}$ (94.7%) and lower MSE_tes (0.0016), on average.

In Figure 5, the data samples were arranged in such a way that *OS* was minimum at the center of the circle and maximum at the outside of the circle. Then, other factors (K_1 and K_a) and responses (*ST* and *RD*) were arranged accordingly along the perimeter from inside to outside of the circle. Therefore, a combination of colours along the perimeter represents a combination of factors and responses. The values of colours are shown as per color map on the right side of the circle-segments.

One can see that the segments of *ST* and *RD* indicate some correlations between the two responses. When segment *ST* had high values of colors, segment *RD* would have low values of colours. This indicated that *ST* was inversely proportional to *RD*. As for the consequences, varying the combination of factors to improve *ST* will cause opposite effect on *RD*. Since we are interested in minimizing both the responses, it is difficult to find the combination of factors which fulfill both the responses simultaneously.
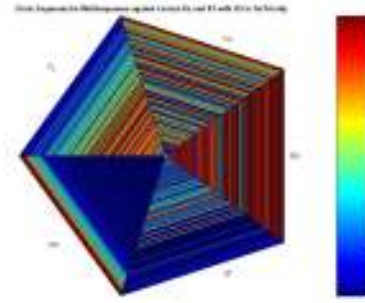

Figure 5 – Circle-Segments of Multi-Factors and Multi-Responses

Based on the segments of *ST* and *RD* that had high and low values of colours respectively, we observe there are some similarities among them. For these combinations, K_a had low value of color (0.1) while K_1 had colour values of equal to or above $0.5 \, (\geq 0.5)$. Therefore, it is found that holding K_a at 0.1 and decreasing the value of K_1 from 1 to 0.5 can increase *RD* and decrease *ST*.

As *OS* was arranged from minimum at the center to maximum at the outside of the circle, it is found the segment of K_1 had high values of colors at the center while low values of colors at the outside of the circle. This finding suggests that K_1 is inversely proportional to *OS*.

By referring to the segments of the three responses, it is difficult to find the combination which has low values of colors. The effort to adjust the combination of factors may not provide the same effect to all of the responses. A method known as constrained-optimization approach was applied to limit the region of responses into certain constraints. With the constrained-optimization approach, we formulated the problem as below statements.

$-0.001 < OS < 0.05$

$ST < 0.25$ ms

$RD < 0.002$

Again, the circle-segments method was used together to solve the problem. In Figure 6, the responses which met the above statements were mapped to the maximum value of colour (1.0 or equal to dark red). For the responses which didn't meet above statements, they were mapped to the minimum value of colour (0 or equal to dark blue). Based on the same circle-segments, we observe that the region which is near to the center of the circle can meet all the constraints. By zooming in the center of the circle as shown in Figure 7, we analyze the combination of factors that produce the desired responses. It is found K_1 had colour values of 0.3 – 0.5 while the colour values of K_a varied between 0.4 – 1.0. The area of desired responses can be minimized further by tightening the

requirements of constraints. Based the combination at the center of the circle (Figure 7), with the colour values of K_1=0.4, K_a=1.0 (K_1=0.04, K_a=200 in actual values), the step response of the system is shown in Figure 8.
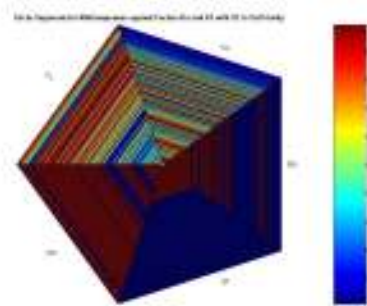

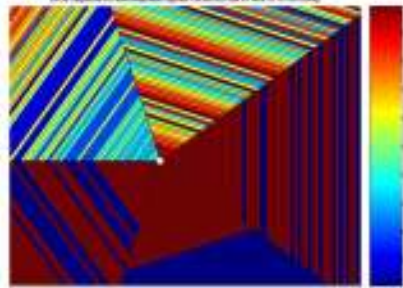*Figure 6 – Circle-Segments with Constrained-Optimization Approach*

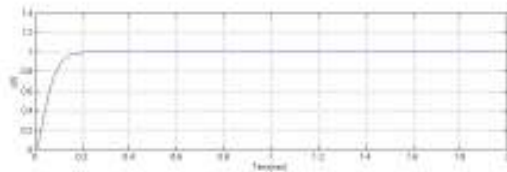*Figure 7 – Zoom in the Center of Circle-Segment*

Figure 8 -Step response of the control system

Figure 9 shows the overlaid contour plot of the responses, which is the visualization tool for the RSM. The white area is the area where all the responses met the statements. From the overlaid contour plot, it is difficult to visualize the possible correlations that exist between the responses. If the number of factors increases, the contour plot is able to display the changes of multi-responses against two factors only by fixing the remaining factors at a certain level. While for the circle-segments method, it is not limited by the number of factors to be displayed. In other words, the circle-segments method can be used as the visualization tool for the process which involves multi-factors and multi-

responses simultaneously. Furthermore, it provides a visualization effect to display the possible correlations that exist not only between factors and responses, but also between the responses as well.
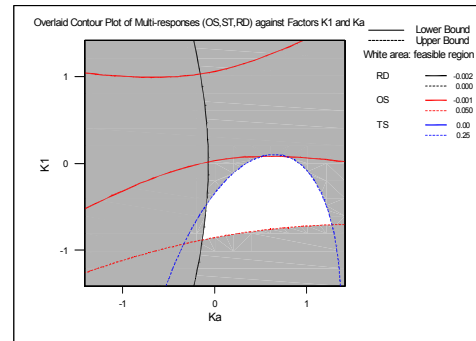

Figure 9 - Overlaid Contour Plot of Multi-Factors and Multi-Responses

5. Conclusion

In this paper, RSM and MLP have been applied to model and predict a control system. Even though both methods provided satisfactory result, the MLP outperformed as compared with the RSM, with higher $R^2_{prediction}$ and lower MSE_tes values. The circle-segments method is used as the visualization tool for the MLP. It has been used in two different ways: first to visualize the possible correlation among the factors and responses through colour mapping, and second to find the desired region based on the constrained-optimization approach. Unlike the contour plot, the circle-segments method is able to display multi-factors and multi-responses simultaneously. Furthermore, it provides insights into the possible correlations that exist not only between factors-responses but also among the responses.

References

[1] K.K. Peh, C.P. Lim, S.S. Quek, and K.H. Khoh, Use of artificial neural networks to predict drug dissolution profiles and evaluation of network performance using similarity factor, *Pharmaceutical Research*, vol. 17, 2000, pp. 1384-1388.

[2] C.P. Lim, S.S. Quek, and K.K. Peh, Prediction of drug release profiles using an intelligent learning system: an experimental study in transdermal iontophoresis, *Journal of Pharmaceutical and Biomedical Analysis*, vol. 31, 2003, pp. 159-168.

[3] J.R. Dutta, P.K. Dutta, and R. Banerjee, Optimization of culture parameters for extracellular protease production from a newly isolated Pseudomonas sp. using response

surface and artificial neural networks models, *Process Biochemistry*, vol.39, 2004, pp. 2193-2198.

[4] W.G. Lou and S. Nakai, Application of artificial neural networks for predicting the thermal inactivation of bacteria: a combined effect of temperature, pH and water activity, *Food Research International*, vol. 34, 2001, pp. 573-579, 2001.

[5] J. Bourquin, H. Schmidli, P. van Hoogevest, and H. Leuenberger, Advantages of artificial neural networks as alternative modeling technique for data sets showing non-linear relationships using data from a galenical study on a solid dosage form, *European Journal of Pharmaceutical Sciences*, vol. 7, 1998, pp. 5-16.

[7] U. Fayyad, G.G. Grinstein, and A. Wierse, *Information Visualization in Data Mining and Knowledge Discovery*, Morgan Kaufmann Publishers Inc, San Francisco, 2001.

[8] D.C. Montgomery, *Design and Analysis of Experiments*, 5th ed., John Wiley & Sons, New York, 2001.

[9] M. Ankerst, D.A. Keim, and H.P Kriegel, 'Circle Segments': A Technique for Visualizing Exploring Large Multidimensional Data Sets, *Proc Visualization '96, Hot Topics Session*, San Francisco, CA, 1996.

[10] Richard C. Dorf, Robert H. Bishop, *Modern Control System*, 10th ed., Prentice Hall, New Jersey, 2005.

[6] T.A Spedding, and Z.Q. Wang, Study on modeling of wire EDM process, *Journal of Materials Processing Technology*, vol. 69, 1997, pp. 18-28.