

**AN APPROACH OF SOFT-COMPUTING IN OPTIMIZING
CONTROLLED RELEASE PRODUCTS**

Nam Phuong Nguyen, Nam Huu Bui, Duong Quang Do

University of Medicine and Pharmacy Ho Chi Minh City

ABSTRACT: *In the pharmaceutical market, all products have a life cycle. Out of date products should be replaced by new ones, which have better quality. For this reason, modelling and optimizing formulation are the regular demands. Traditional methods of design and optimization - such as statistics, simplex - can only be used for simple and linear data. In case of complicated or non-linear data, alternative methods that are able to deal with such data are needed.*

This paper presents a solution for optimizing controlled release product formulation using a combination of AI techniques (Soft-Computing): neural networks, fuzzy logic and genetic algorithms. This achievement will help to significantly reduce time and labour in R&D process thank to its good accuracy and high processing speed. The results obtained from this research indicate that the alternative approach can be considered as an effective and efficient method for modelling and optimising controlled release formulations.

Keywords: *Neural networks, Genetic Algorithms, Optimization, Soft computing, Controlled Release.*

1. INTRODUCTION

Formulation design is regular work of pharmacist because all products have a life cycle products quality need to be constantly improved. Out of date products should be replaced by better ones. For this reason, modelling and optimization of formulation are the regular demands^[1]. Traditional methods of design and optimization of formulation - such as *statistics, simplex,...* - are only used for simple and linear data. These methods are not suitable for complicated or non-linear data.

The formidable task of formulation research is to navigate multidimensional design

space to find the point, which has the optimum balance of properties. Nowadays, formulators can develop complex dosage forms by design and optimization way. Since product properties are affected not only by the ratio in which the ingredients are combined but also by the processing parameters, the ingredient levels and processing conditions should be taken into account in formulation design. Computer technology in the form of artificial intelligence provides an affordable means of improvement in product formulation and has more promising of solving an optimization of product formulation because it is not finite of ingredients (X) and can simultaneously

optimize many properties (Y) of the formulation and is suitable for the problems with complicated and non-linear data.

In this study, a combination of neural networks, fuzzy logic and genetic algorithms (GA) called Soft-Computing (SC) is employed with neural networks considered as a method for modelling whilst GA combined with fuzzy logic acted to optimisation process. Each of techniques has advantages and disadvantages, but if they are accurately combined all together, the disadvantages of this will be overcome by advantages of another [2, 3, 4, 5, 6] - for example, neural networks is difficult to extract knowledge, but fuzzy inference systems does it easily. The paper then reports the application of SC to two sets of published formulation data, one for a matrix tablet, and the other for controlled release microspheres and compares the results obtained with statistical analyses.

2. SOFT-COMPUTING CONCEPT

2.1. Modelling formulation data with Neural networks

Neural networks are complementary technologies in the design of adaptive intelligent systems. Artificial Neural Network (ANN) learns from scratch by adjusting the interconnections between layers. For over 60 years, ANNs have been applied to design a model of relationships between cause and

effect, particularly to nonlinear and complex data. A comparison can be observed for ANN with mammalian nerve connectivity. The mammalian nervous system is built up from biological neurons. Each neuron collects input stimuli and triggers an output to the next neurons in the system (see Figure 1). Similarly, artificial neural networks also involve connecting signal and nodes that collect mathematical inputs and produce the output signals that are passed to the next neurons [6, 7].

The units in the input layer only have one input signal assigned to them, while the nodes in the hidden layer are connected and assigned by many of the input signals. The output layer depends upon the structure of network in that there are only one or many output nodes with respectively many or a unique output signal. An artificial neural network is generally composed of several layers: input layer, hidden layers (one or many), and output layer. For example, the structure of a neural network with 4 inputs, 2 output, and 3 nodes in a single hidden layer is detailed in Figure 2. However, neural networks are often known as “black box” technologies in that the means of mapping inputs to output(s) is hidden within the network structure. It is also quite different from statistical methods in that a neural network does not produce a mathematical equation. Neural networks are often used to design predictive models.



Figure 1. Structure of biological neuron

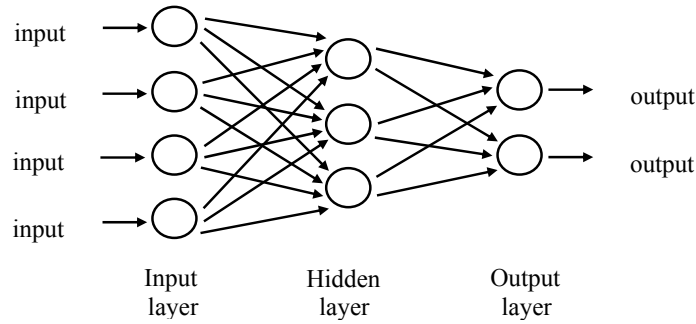


Figure 2. Structure of neural networks

2.2. The combination model of GA and fuzzy logic for optimization

Genetic Algorithms (GA) are derivative-free stochastic optimization methods based on the concepts of natural selection and

evolutionary processes (detailed in Figure 3). This step, genetic algorithms associate with fuzzy logic to optimise formulation - the fitness function based on cause-and-effect relationships^[8].

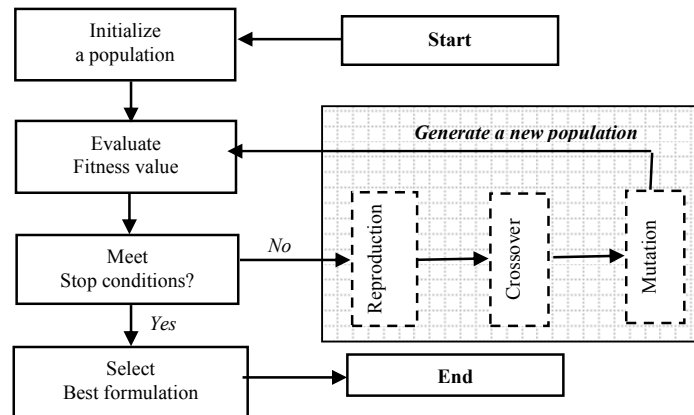


Figure 3. The cycle of selection and evolutionary processes with GA

Given a way or a method of encoding solution of a problem into the form of chromosomes and given an evaluation function that returns a measurement of the cost value of any chromosome in the context of the problem, the processing GA includes 6 steps [6, 9].

Step 1: Initialize a set of solutions (potential formulations) randomly - called population.

Step 2: Evaluate each formulation in the population

Step 3: Create new formulations by mating current formulations; apply mutation and

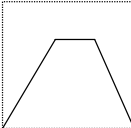
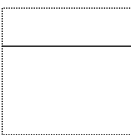
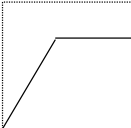
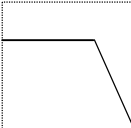
recombination as the “parent” formulations mate.

Step 4: Delete members of the population to make room for new formulations

Step 5: Evaluate the new formulations and insert them into the population

Step 6: If the stopping criterion is satisfied, then stop and return the optimum formulations; otherwise, go to Step 3

The detailed membership functions from the fuzzy logic, applied to optimization with GA, are as follows:

<p><i>Flat-Tent function</i> (a): desirability drops linearly between Mid1 and the minimum, and between Mid2 and the maximum, but between Mid1 and Mid2, the values are perfectly acceptable. That is, their membership function in the set of acceptable values is 1.</p>	 <p style="text-align: center;">(a)</p>	<p><i>Flat function</i> (d): any value is acceptable; its membership function in the set of acceptable values is 1.</p>	 <p style="text-align: center;">(d)</p>
<p><i>Up-Hill function</i> (b): any value between the mid-point (Mid1 = Mid2) and the maximum is completely acceptable; its membership function in the set of acceptable values is 1. Any value from minimum to mid-point, the desirability decreases linearly until it is zero at the minimum point.</p>	 <p style="text-align: center;">(b)</p>	<p><i>Down-Hill function</i> (c): any value between the mid-point (Mid1 = Mid2) and the minimum is completely acceptable; its from mid-point to maximum, the desirability membership function in the set of acceptable values is 1. Any value decreases linearly until it is zero at the maximum point.</p>	 <p style="text-align: center;">(c)</p>

2.3. Solving the optimization problem with neural networks, fuzzy logic and genetic algorithms

A fusion of neural networks, fuzzy logic and genetic algorithms to deal with an optimization of product formulation problem is illustrated in Figure 4.

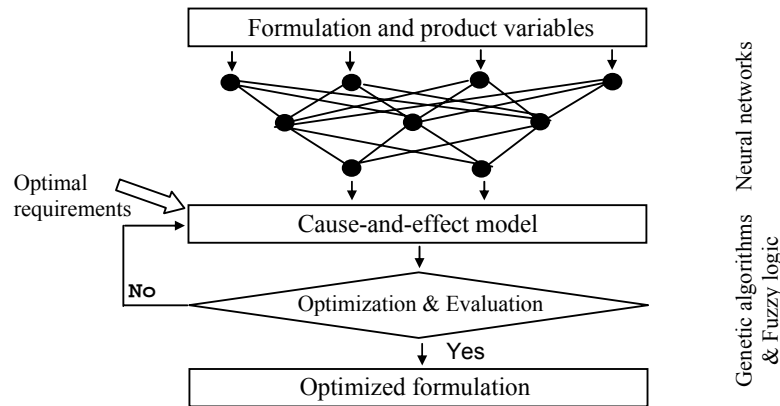


Figure 4. The cycle of modelling and optimization

The detailed processing of optimization is as follows:

Step 1: establish cause-and-effect relationship by using neuro-fuzzy system or neural networks.

Step 2: determine optimal requirements defined by user.

Step 3: optimize ingredients corresponding to optimal condition of properties by using genetic algorithms combined to fuzzy logic, the fitness function of GA is cause-and-effect relationship model determined from Step 1. Repeat Step 3 until a stopping criterion is met or optimal condition is reached.

2.4. Software tool

The software was used in this research is BCPharSoft OPT. This is a software tool, which is built in C#.net programming language. It was a modified form of that described previously – INForm (www.intelligensys.co.uk), but with additional functionalities in order to improve the quality of predictive models and the optimum formulation.

In order to evaluate the quality of a predictive model generated by ANN, the correlation coefficient R -squared (R^2) was computed, with higher values of R^2 indicating the improved quality of the model ^[10].

$$R^2 = \left(1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right) \times 100$$

where \bar{y} : the mean of the dependent variable; \hat{y} : the predicted value from the model; n: number of records.

3. EXPERIMENTAL DATA

The formulation database of the matrix tablet taken from the literature (Bodea and Leucuta, 1997) ^[11], consisted of 14 experimental records, and involved varying percentages of two hydrophilic polymers (hydroxypropylmethylcellulose, HPMC - X₁, sodium carboxymethylcellulose, CMCNa - X₂) and propranolol HCL - X₃. The measured outputs were the cumulative percentages of drug released after 1, 6, and 12h sampling intervals (Y₁, Y₂, and Y₃, respectively). These data were modelled and optimised in the original study ^[11] by statistical methods using a D-optimal quadratic model. In the present study, 11 records were used for training and 2 records used as unseen data for testing the predictive models. Another formulation database for controlled release diclofenac sodium microspheres containing 27 experimental records taken from a published paper (Gohel and Amin, 1998) ^[12] was used for validating the capability of SC for such of formulation as well. In this study, microspheres

were prepared using sodium alginate as a polymer and CaCl₂ as a cross-linking agent. A 3³ full factorial design was used to investigate the joint influences of three variables - the stirring speed during preparation of the microspheres (X₁), concentration of CaCl₂ (X₂) and % of heavy liquid paraffin in a blend of heavy and light liquid paraffin in the dispersion medium (X₃) - on the time for 80% drug dissolution (t₈₀). In addition, in the published study ^[12], the % drug released after 60 (Y₆₀), 360 (Y₃₆₀), and 480 min (Y₄₈₀) was also considered as outputs that were analysed. 25 records were used as training data, and 2 records used as unseen data to test predictive power.

4. EXPERIMENTAL RESULTS

4.1. Matrix tablet formulation

By selecting suitable values of control parameters, SC generated satisfactory models for all responses of the matrix tablet formulation. The correlation coefficient R² values of the predictive models generated from SC were showed in Table 1.

Table 1. R^2 values of the predictive models generated from SC and statistical method ^[11]

Method	Y_1	Y_2	Y_3
<i>Soft-Computing</i>	R^2 Train= 0.98 R^2 Test= 0.99 $R^2 = 0.98$	R^2 Train= 0.99 R^2 Test= 0.9 $R^2 = 0.97$	R^2 Train = 0.99 R^2 Test = 0.99 $R^2 = 0.99$
<i>Statistical</i>	$R^2 = 0.96$	$R^2 = 0.88$	$R^2 = 0.91$

Compared with a published study ^[11], the present study gave improved models for all responses. The analyses in Table 1 showed that

for the models of the cumulative percentage release after 1h (Y_1), 6h (Y_2) and 12h (Y_3), the quality of the models was improved with significantly higher R^2 values.

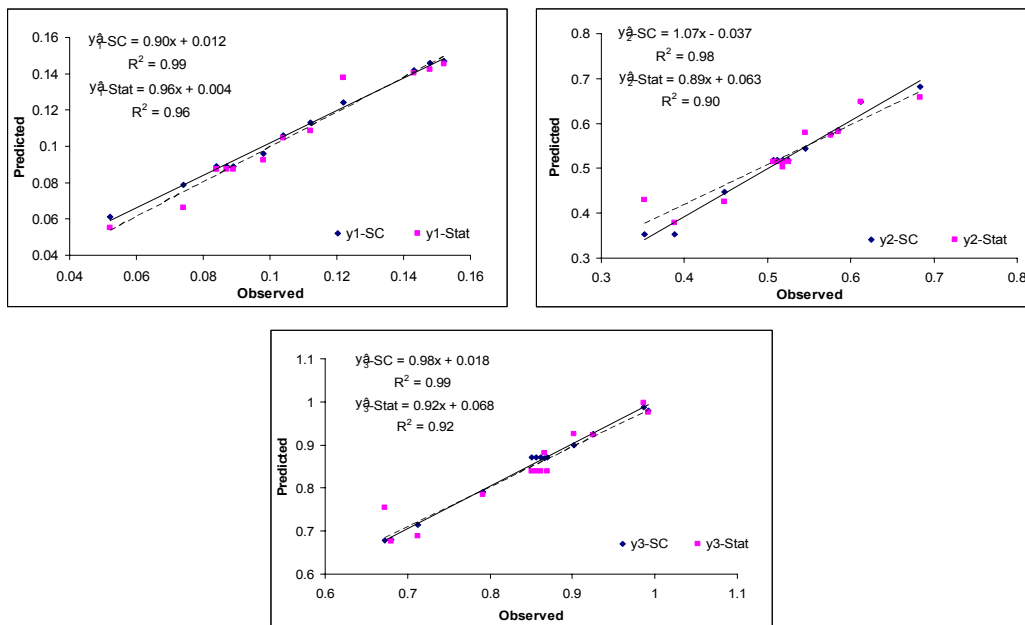


Figure 5. Scatter plots, linear equations and R^2 values for the observed data points from SC and statistical methods for Y_1 , Y_2 and Y_3 .

In comparison with the statistical result reported in the literature ^[11] showed in Figure 5, the linear R^2 values for all observed responses were significantly higher to those from the statistical models. Moreover, for the outputs Y_2 and Y_3 , the slope and the intercept coefficients

from the SC models were much improved compared to those from the statistical models. All of these results proved that overall the predictive models generated from SC were superior when compared to the results presented in the literature ^[11].

For the optimisation of this product, the constraints of optimum formulation used in this study were also taken from the literature that was as follows:

$X_2+X_3 \leq 0.8$	$0.1 \leq Y_1 \leq 0.2$
$X_3 \geq 0.34$	$0.45 \leq Y_2 \leq 0.55$
	$0.8 \leq Y_3$

Table 2. Optimum formulations generated from SC

	X_1	X_2	X_3	Y_1	Y_2	Y_3
(1)	0.453	0.007	0.519	0.145	0.546	0.843
(2)	0.334	0.101	0.508	0.113	0.550	0.907
(3)	0.316	0.101	0.498	0.104	0.548	0.908

From Table 2, it also demonstrated that though SC generated 3 different optimum formulations generated, they still met the required constraint. The first formulation showed the maximum value for Y_1 , the second formulation showed the maximum value for Y_2 , while Y_3 obtained the maximum value with the third formulation. For these formulations the formulators could get more selections for their different purposes, for example if they want to maximize the % of drug dissolved in 6h (Y_2) and optimize the formulation of this drug following the constrains showed above, they could consider the second formulation as the optimum one by themselves.

As showed in Table 2, SC generated several optimum formulations for this product that met all optimum conditions mentioned above. In addition, when compared to a single outcome optimised from statistical method ^[11], this approach is definitely superior because of its multiple formulations optimised.

4.2. Controlled release diclofenac sodium microspheres formulation

It is similar to the first data, by selecting suitable values of control parameters, the correlation coefficient R^2 values of the predictive models for the diclofenac sodium microspheres formulation generated from SC were showed in Table 3. The results in Table 3 showed that for this product SC achieved significantly higher quality predictive models for all responses. In particular, SC predicted a model with $R^2 = 0.93$ for Y_{60} whilst statistical method gave $R^2 = 0.74$ only for this property.

Table 3. R^2 values of the predictive models generated from SC and statistical method ^[12]

Method	t_{80}	Y_{60}	Y_{360}	Y_{480}
<i>Soft-Computing</i>	R^2 Train= 0.99 R^2 Test= 0.95	R^2 Train= 0.93 R^2 Test= 0.84	R^2 Train= 0.97 R^2 Test= 0.99	R^2 Train= 0.97 R^2 Test= 0.96

	$R^2 = 0.99$	$R^2 = 0.93$	$R^2 = 0.98$	$R^2 = 0.97$
<i>Statistical</i>	<i>0.99</i>	<i>0.74</i>	<i>0.95</i>	<i>0.92</i>

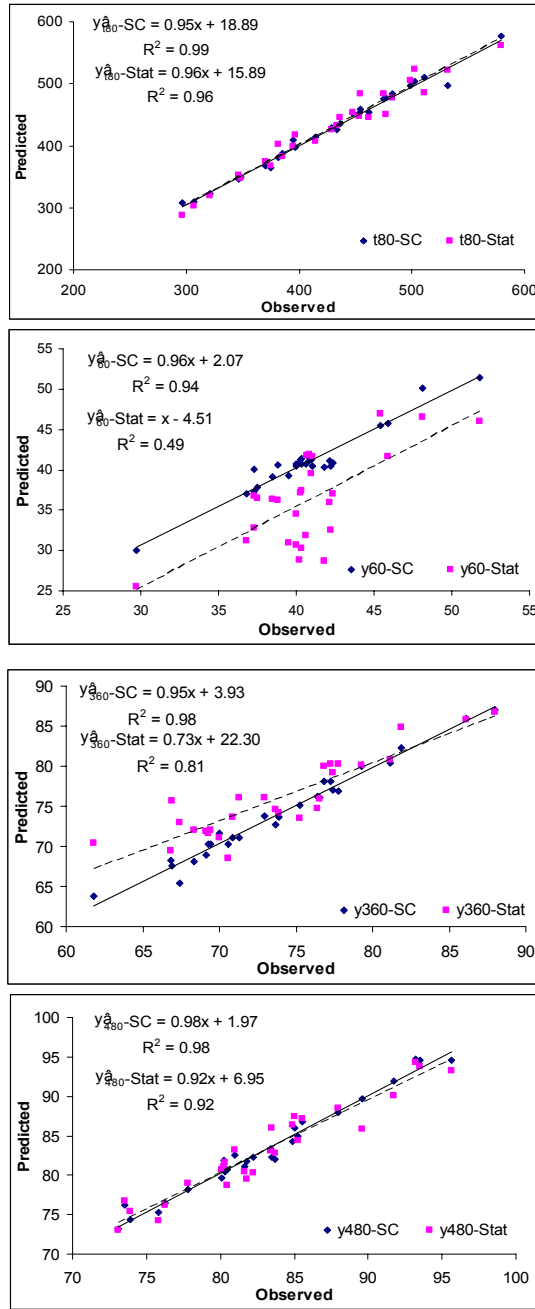


Figure 6. Scatter plots, linear equations and R^2 values for the observed data points from SC and statistical methods for t_{80} , Y_{60} , Y_{360} and Y_{480} .

From Figure 6, it is clear that the satisfactory predictive power of the SC models for the observed data can be seen. The linear R^2 values for all these responses were significantly high and the slope and the intercept coefficients from the SC models were acceptable as well. In general in comparison with the statistical method, SC produced satisfactory models for all responses. Moreover for the Y_{60} response, the predictive model of SC for this formulation significantly overcame the result generated from statistical analysis.

For the optimisation of this product, the constraints of optimum formulation used in this study were also taken from the literature that was as follows: $20\% \leq Y_{60} \leq 40\%$, $50\% \leq Y_{360} \leq 70\%$, $65\% \leq Y_{480} \leq 80\%$ and X_1 : integer ^[12]. As showed in Table 4, SC generated several optimum formulations for this product that met all optimum conditions mentioned above. In addition, when compared to a single outcome optimised from statistical method ^[12], this approach is definitely superior because of its multiple formulations optimised.

Table 4. Optimum formulations generated from SC

	X_1	X_2	X_3	t_{80}	Y_{60}	Y_{360}	Y_{480}
(1)	500	14.13	32.79	560.31	39.99	67.77	77.17
(2)	1500	14.31	49.30	470.98	39.03	69.90	79.00
(3)	540	7.50	44.75	482.26	37.69	69.40	80.00

From Table 4, it also demonstrated that though SC generated 3 completely different optimum formulations generated, they also met the required constraint. For these formulations the formulators could get more selections for their own purposes.

4.3. General comments

When validating the capability of SC and comparing the predictive power of this method to the statistical methods for both controlled release products, it was recognised that the basis of the statistical approach is to use standard equations and procedures based on statistical theory to obtain the final equation considered as predictive model. The statistical

output is fixed and if a formulator wants to improve the quality of the final statistical equation, he must carry out further experiments to obtain a higher quality data set. However with SC, a formulator can obtain alternative outputs, with a selection of an appropriate training model. For example, by changing values of control parameters, the quality of the predictive equation can be improved. In other words, a formulator can perform SC in an iterative manner by directed change of control parameter values until the most appropriate and/or predictive model is obtained. Moreover, a single optimised formulation generated from statistical analysis is also a major

inconvenience of this method when compared to SC.

5. CONCLUSIONS

Although neural networks, fuzzy logic and genetic algorithms had been introduced for a long time, applications using theories of neural networks, fuzzy logic and genetic algorithm are still interested; the application using the neural networks, fuzzy logic and genetic algorithm for

solving an optimization of product formulation in pharmaceuticals is an example. This solution helps formulator reduce time and labor more than traditional methods do. In contrast to statistical approaches, Soft-computing, with its advantage of generating several optimum formulations and superior predictive models, has been shown to be an efficient method for modelling and optimising controlled release formulations.

ỨNG DỤNG KỸ THUẬT TÍNH TOÁN MỀM GIẢI QUYẾT BÀI TOÁN TỐI ƯU HÓA CÔNG THỨC VIÊN PHÓNG THÍCH CÓ KIỂM SOÁT

Nguyễn Phương Nam, Bùi Hữu Nam, Đỗ Quang Dương

Trường Đại học Y Dược Tp. Hồ Chí Minh

TÓM TẮT: Trong ngành dược, đối với một nhà sản xuất thuốc việc thành lập cũng như tối ưu hóa công thức là việc làm thường xuyên bởi vì mỗi sản phẩm đều có một vòng đời nhất định và nhu cầu cạnh tranh trên thị trường đòi hỏi phải không ngừng cải tiến sản phẩm hiện có hay thay thế sản phẩm mới. Chính vì lý do này, tối ưu hóa công thức dược đã được đề cập đến. Các phương pháp tối ưu hóa truyền thống (toán thống kê, đơn hình...) chỉ có thể áp dụng với các dữ liệu đơn giản và tuyến tính. Chúng không còn phù hợp với các dữ liệu phức tạp và phi tuyến. Ngoài ra, các phương pháp truyền thống không tối ưu hóa được đồng thời nhiều biến phụ thuộc trong khi mỗi sản phẩm thường có rất nhiều tính chất. Phương pháp tối ưu hóa thông minh có nhiều triển vọng thay thế các phương pháp truyền thống.

Bài báo này đưa ra một phương pháp tối ưu hóa thông minh. Đó là một sự kết hợp giữa mạng thần kinh, logic mờ và thuật toán di truyền. Phương pháp này đã giải quyết được những khó khăn mà các phương pháp truyền thống không thể thực hiện được. Các kết quả đã thu được từ nghiên cứu này chứng minh rằng đây là một phương pháp tối ưu hóa hiệu quả.

Từ khóa: mạng thần kinh, logic mờ, thuật toán di truyền, kỹ thuật tính toán mềm

REFERENCES

- [1]. Giap D.V. *Applying an Artificial Intelligence Software for Optimizing Formulation and Process*. Faculty of Pharmacy - University of Medicine and Pharmacy at Ho Chi Minh City, Viet Nam (2001).
- [2]. Nauck D., Kruse R. *Neuro_Fuzzy Systems for Function Approximation*. Faculty of Computer Science_ Neural and Fuzzy Systems, Otto_von_Guericke_University of Magdeburg, Universitaetsplatz 2, D-39106 Magdeburg, Germany (1997).
- [3]. Jantzen J. *Neurofuzzy Modelling*. Department of Automation, University of Denmark, report no 98-H-874 (Oct 1998).
- [4]. Jang J.R., Sun C., Mizutani E. *Neuro-fuzzy and Soft Computing: A Computation Approach To Learning And Machine Intelligence*. Upper Saddle River, Prentice-Hall (1997).
- [5]. Bonissone P.P. *Adaptive Neural Fuzzy Inference Systems (ANFIS): Analysis and Applications*. GE CRD, Schenectady, NY USA (1997).
- [6]. Lin C.T., George Lee C.S. *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent System*. International ed, ISBN 0-13-261413-8. Upper Saddle River, Prentice-Hall (1996).
- [7]. Kasabov N.K., Kim J.S., Gray A.R., Watts M.J. *FuNN - A Fuzzy Neural Network Architecture for Adaptive Learning and Knowledge Acquisition*. Department of INForm 3.0ation Science, University of Otago, P.O.Box 56, Dunedin, New Zealand.
- [8]. Sivanandam S.N., Deepa S.N. *Introduction to Genetic Algorithms*. ISBN 978-3-540-73189-4. Berlin, Heidelberg, New York Springer (2008).
- [9]. Kiem H., Thai L.H. *Genetic Algorithms: Solving the problems in computer naturally*. Viet Nam, Educational Publisher (2000).
- [10]. Shao Q., Rowe R.C., York P. *Comparison of neurofuzzy logic and neural networks in modelling experimental data of an immediate release tablet formulation*. Eur J Pharm Sci, Vol. 28, pp. 394-404. (2006).
- [11]. Bodea A., Leucuta S.E. *Optimization of hydrophilic matrix tablets using D-optimal design*. Int J Pharm, Vol. 153, pp. 247-255. (1997).
- [12]. Gohel M.C., Amin A.F. *Formulation optimization of controlled release diclofenac sodium microspheres using factorial design*. J Control Release, Vol. 51, pp. 115-122. (1998).