Neural network based optimization of drug formulations

Devendra Singh

Advanced drug delivery …

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Advanced Drug Delivery Reviews 55 (2003) 1217–1231
Advanced DRUG DELIVERY Reviews
www.elsevier.com/locate/addr

Neural network based optimization of drug formulations

Kozo Takayama*, Mikito Fujikawa, Yasuko Obata, Mariko Morishita

Department of Pharmaceutics, Hoshi University, Ebara 2-4-41, Shinagawa, Tokyo 142-8501, Japan

Received 10 March 2003; accepted 19 May 2003

Abstract

A pharmaceutical formulation is composed of several formulation factors and process variables. Several responses relating to the effectiveness, usefulness, stability, as well as safety must be optimized simultaneously. Consequently, expertise and experience are required to design acceptable pharmaceutical formulations. A response surface method (RSM) has widely been used for selecting acceptable pharmaceutical formulations. However, prediction of pharmaceutical responses based on the second-order polynomial equation commonly used in an RSM, is often limited to low levels, resulting in poor estimations of optimal formulations. The purpose of this review is to describe the basic concept of the multi-objective simultaneous optimization technique, in which an artificial neural network (ANN) is incorporated. ANNs are being increasingly used in pharmaceutical research to predict the nonlinear relationship between causal factors and response variables. Superior function of the ANN approach was demonstrated by the optimization for typical numerical examples.

Keywords: Artificial neural networks; Response surface method; Multi-objective optimization; Polynomial equation; Pharmaceutical formulation

Contents

*Corresponding author. Tel.: +81-3-5498-5781; fax: +81-3-5498-5782.
E-mail address: takayama@hoshi.ac.jp (K. Takayama).

0169-409X/03/$ – see front matter
doi:10.1016/S0169-409X(03)00120-0

1. Introduction

A pharmaceutical formulation is composed of several formulation factors and process variables. Several responses relating to the effectiveness, usefulness, stability, as well as safety must be optimized simultaneously. One of the difficulties in the quantitative approach for formulation design is approximating the actual relationship between causal factors and pharmaceutical responses. Another difficulty is an optimal formulation for one property is not always desirable for the other characteristics. This is called a multi-objective optimization problem. Consequently, expertise and experience are required to design acceptable pharmaceutical formulations.

The response surface method (RSM) has widely been used for selecting acceptable pharmaceutical formulations [1–17]. The RSM includes (1) statistical factorial experimental designs, (2) modeling between causal factors and response variables, and (3) multi-objective optimization for seeking the best formulation under a set of pragmatic constraints. Composite experimental design can be applied for selecting rationale model formulations, which are composed of several formulation factors and process variables. Compared with a normal analysis based on a one-factor-at-a-time experiments, we can greatly reduce the number of experiments for the preparation of model formulations. Response variables of these model formulations are predicted quantitatively by the combination of causal factors. In a classical way, multiple regression analysis has been applied on the basis of a quadratic polynomial equation [18,19], since theoretical relationships between causal factors and response variables are not clear. Finally, multi-objective optimization algorithms are applied for predicting the best formulation [11,20–23].

Prediction of pharmaceutical responses based on the quadratic polynomial is often limited to low levels. It may result in the poor estimation of optimal formulations. To overcome the shortcomings of the poor estimation based on the quadratic polynomial, application of an artificial neural network (ANN) has been investigated [24–30]. The ANN is a learning system based on a computational technique, which can simulate the neurological processing ability of the human brain [24]. The ANN has been applied to solving various problems such as product development [25–36], QSARs (quantitative structure–activity relationships) [37–39], QSPRs (quantitative structure–pharmacokinetic relationships) [39–45], estimating diffusion coefficient [46], predicting the skin permeability [47–49], predicting the Caco-2 cell permeability [50], and predicting the mechanism of drug action [51].

The purpose of this review is to describe the basic concept of the multi-objective simultaneous optimization technique, in which an ANN is incorporated [26–31]. Superior function of the ANN approach was demonstrated by the optimization for typical numerical examples [30,31].

2. ANN modeling

Theoretical details of a hierarchical ANN have been given elsewhere. Briefly, the general ANN structure has one input layer, one or more hidden layers and one output layer. Each layer has some units corresponding to neurons. The units in neighboring layers are fully interconnected with links corresponding to synapses. The strengths of connections between two units are called ‘‘weights’’. In each hidden layer and output layer the processing unit sums its input from the previous layer and then applies the sigmoidal function to compute its output to the following layer according to the following equations:

$$y_q = \sum w_{pq} x_p \tag{1}$$

$$f(y_q) = \frac{1}{1 + \exp(-\alpha y_q)} \tag{2}$$

where w_{pq} is the weight of the connection between

unit q in the current layer to unit p in the previous layer, x_p is the output value from the previous layer, $f(y_q)$ is conducted to the following layer as an output value, and α is a parameter relating to the shape of the sigmoidal function. Nonlinearity of the sigmoidal function is strengthened with an increase in α. The ANN learns an approximate nonlinear relationship by a procedure called ''training'', which involves varying weight values. Training means a search process for the optimized set of weight values, which can minimize the squared error between the estimation and experimental data of units in the output layer. A back-propagation method has widely been applied for training ANNs [52]. Training is long iterative process, and an ANN often gets stuck in a local minima. Certain empirical techniques have been reported to improve the convergence of ANNs in the global minima [53]. Another essential approach is to use a simulated annealing technique along with an extended Kalman filter algorithm for ANN training [54–56]. We can greatly reduce the number of iterative training by using the extended Kalman filter algorithm, and also can avoid the ANN getting stuck in a local minima using the simulated annealing technique [57]. Although multiple layers can be set between the input layer and the output layer, many ANNs consist of only one hidden layer [52]. One layer is usually sufficient to provide adequate prediction even if continuous variables are adopted as the units in the output layer [58–60].

In order to enable reasonable prediction of each response variable by an ANN, Carpenter and Hoffman introduced an equation relating to the number of units in the input layer, the hidden layer and the output layer [61]:

$$n_s = \beta\{n_h(n_i + 1) + n_o(n_h + 1)\} \quad (3)$$

where n_h is the number of hidden units, n_i is the number of input units, n_o is the number of output units, and n_s is the number of training data pairs. The constant β is the parameter relating to the degree of over-determination.

The unknown parameters associated with an ANN are the weights of the network. Over-determined ($\beta > 1$), exact-determined ($\beta = 1$) and under-determined ($\beta < 1$) approximations have more, an equal number, or fewer training data pairs than the number of unknown parameters associated with the approximation. For example, $\beta = 1.5$ would give a 50% over-determined approximation. With an under-determined approximation ($\beta < 1$), each output data point is fitted perfectly by iterative training, but the approximation may vary wildly between the output data points; i.e., the over-training problem. Thus, the selection of $\beta > 1$ is usually recommended to enable reasonable prediction of each response variable adopted as the unit in the output layer. To ensure the optimality of an ANN structure, a cross-validation technique such as a ''leave-one-out (LOO) method'' should be applied [37,47,50]. In the LOO method, one data pair is systematically removed from the training data set, and the ANN is then trained by using the reduced data set. The trained ANN is adopted to predict the response variable on the removed data pair. This process is repeated so as to complete the prediction of the response variable in the all data set. As a judging standard to evaluate the optimality of ANN, Akaike's information criterion (AIC) can be applied [30]:

$$\mathrm{AIC} = n_s \times \ln(\mathrm{SS}) + 2 \times n_w \quad (4)$$

where n_s is the number of data pairs, n_w is the number of weights in the ANN, and SS is the residual sum of squares between observed and predicted response variables. From among possible ANN structures estimated based on Carpenter and Hoffman's concept (Eq. (3)), we may choose the optimal ANN, which gives the smallest value of AIC.

In the general structure of an ANN, the same units in the hidden layer are used for the prediction of different response variables. This may occasionally lead to poor estimation of some responses. To avoid this problem, we developed a partitioned ANN in which every response could be estimated by an independent set of units in the hidden layers (Fig. 1) [28–30]. This is equivalent to predicting each response variable independently by different ANN systems. In the optimization study for pharmaceuticals, model formulations are usually prepared according to statistical experimental designs to reduce the number of experiments. Hence the number of

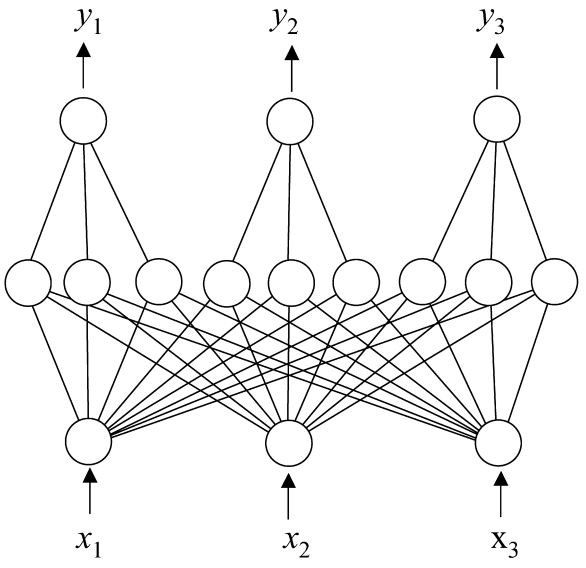

Fig. 1. Typical structure of a partitioned ANN composed of three input units, three hidden units and three output units. Every response (output unit) in the partitioned ANN can be estimated by the independent set of units in the hidden layer, although the same units in the hidden layer are used for the prediction of different responses in the case of a general hierarchical ANN.

data pairs available for ANN training is limited to low levels. This may often lead to the under-determined approximation in the general ANN structure composed of plural units in the input, hidden and output layers. On the other hand, it is easier for the partitioned ANN to avoid the under-determined approximation because $n_o = 1$ can be adopted in Eq. (3).

3. Mathematical concept for optimization

3.1. Single-objective optimization

In general, optimization problems are described mathematically so as to minimize the objective function, $F(X)$, under the following inequality and/or equality constraints:

$$G_i(X) \geq 0 \quad i = 1, 2, 3, \ldots, p \tag{5}$$

$$H_j(X) = 0 \quad j = 1, 2, 3, \ldots, q \tag{6}$$

where $G_i(X)$ is the inequality constraint and $H_j(X)$ is the equality constraint. It is difficult to solve the constrained optimization problem described above without any mathematical modifications. Thus, the constrained optimization problem is transformed to an unconstrained optimization problem by adding penalty functions as follows:

$$\begin{aligned} T(X,r) = F(X) &+ r^{-1}\sum_{i=1}^{p}\{\phi_i G_i(X)\} \\ &+ r^{-1}\sum_{j=1}^{q}\{H_i(X)\} \quad \text{when } G_i(X) < 0,\ \phi_i \\ &= 1; \quad \text{when } G_i(X) \geq 0,\ \phi_i = 0 \end{aligned} \tag{7}$$

where $T(X,r)$ is the transformed unconstrained objective function, r is a perturbation parameter of $T(X,r)$ and ϕ_i is a step function by which the objective function, $F(X)$, is penalized. Mathematical details of the other transformations have been well described in the literature [62,63]. The second and third terms in Eq. (7) act as penalty functions, because the values of these terms will increase abruptly when the values of $G_i(X)$ are negative or the values of $H_j(X)$ deviate from zero. The meaning of perturbation parameter, r, can be explained using a simple optimization problem as follows [9]:

$$F(X) = X_1^2 + X_2^2 \tag{8}$$

$$-2 \leq X_1 \leq 2 \tag{9}$$

$$-2 \leq X_2 \leq 2 \tag{10}$$

$$X_1 + X_2 = -1 \tag{11}$$

where Eq. (8) is an objective function to be minimized, Eqs. (9) and (10) are inequality constraints, and Eq. (11) is an equality constraint. Fig. 2 shows the effect of r values on the three-dimensional diagrams for the transformed objective function, $T(X,r)$, which is derived from Eqs. (8)–(11). The shape of $T(X,r)$ is gradually sharpened and the minimum points of $T(X,r)$ sequentially approach the accurate solution of the optimization problem with decreasing r values. The optimum solution is obtained as the point, $X(r)$, which gives the minimum

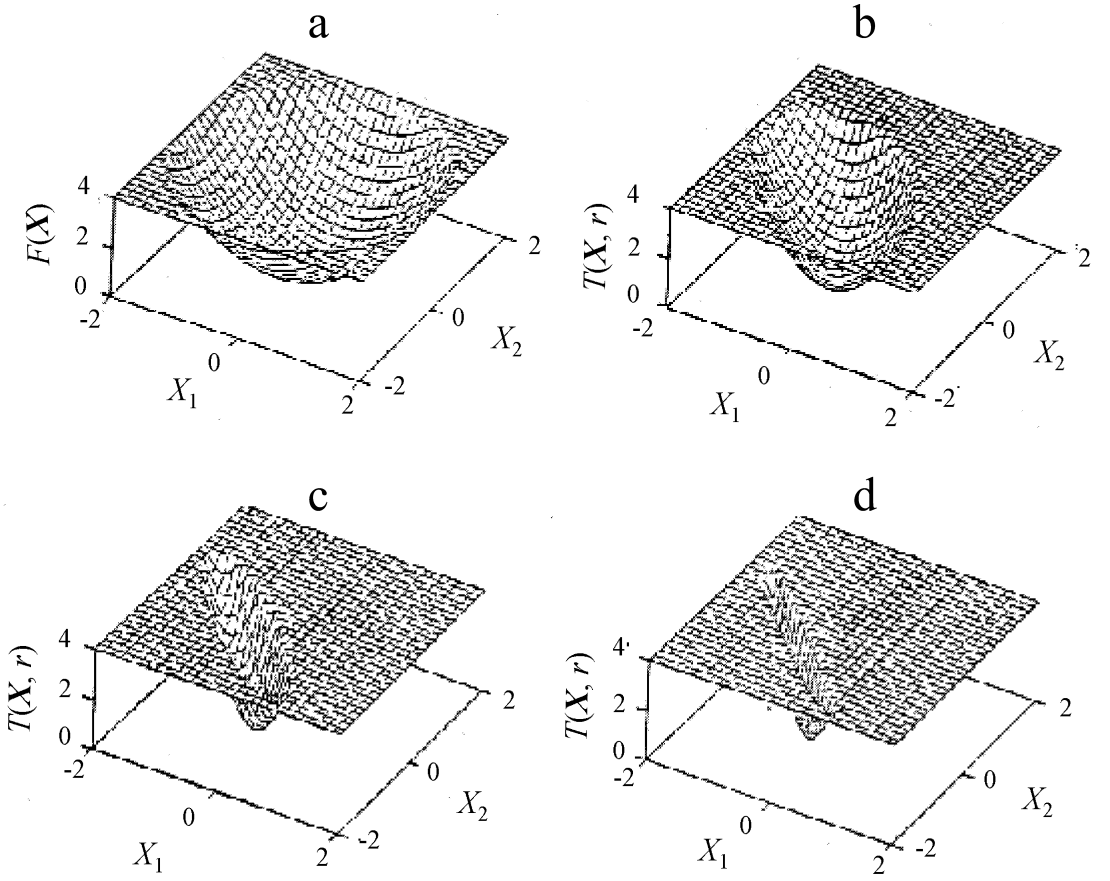

Fig. 2. Three-dimensional diagrams for the objective function, $F(X)$, and transformed function, $T(X,r)$, derived from Eqs. (9)–(11) as a function of X_1 and X_2. (a) $F(X)$; (b) $T(X,r)$ at $r=1$; (c) $T(X,r)$ at $r=0.1$; (d) $T(X,r)$ at $r=0.01$.

value of $T(X,r)$ when the value of r is sufficiently close to zero.

3.2. Multi-objective optimization

When the optimization problem includes several objectives, response variables [$F_k(X)$, $k=1, 2, 3, \ldots, n$] should be incorporated into a single function in order to consider all the responses simultaneously. Derringer and Suich [20] introduced general transformations based on the concept of desirability associated with a given response function. This transformation, a desirability function method, requires minimum and maximum acceptable values for every response. The individual response can be normalized to the desirability functions (d_k, $k=1, 2, 3, \ldots, n$) which have values inside the interval [0,1] using the distance between minimum and maximum acceptable values. The normalized functions are then combined into a multi-objective function, D_{total}, by means of the geometric mean of predicted values of each function:

$$D_{\text{total}} = (d_1 \times d_2 \times d_3 \times \ldots \times d_n)^{1/n} = \left(\prod_{k=1}^{n} d_k\right)^{1/n} \quad (12)$$

The desirability function defined in Eq. (12) allows simultaneous optimization taking into account the relative importance of each response. This method has widely been applied to the development of pharmaceutical products and the method has been useful for solving practical optimization problems [22,23]. However, one of the basic shortcomings of

this approach is the subjectivity in the selection of the minimum and maximum acceptable values for each response. That is, improper values of minima and/or maxima may lead to inaccurate solutions for the optimum formulation.

In order to avoid the problem of subjectivity in application of the desirability function method, we can employ another approach based on the standardized Euclidian distance between the predicted value of each response and the optimum one that was obtained individually [11,21]:

$$S(X) = \left(\sum_{k=1}^{n} \left\{ \frac{\mathrm{FD}_k(X) - \mathrm{FO}_k(X)}{\mathrm{S.D.}_k} \right\}^2 \right)^{1/2} \quad (13)$$

where $S(X)$ is the distance function generalized by the standard deviation, $\mathrm{S.D.}_k$, of the observed values for each response variable, $\mathrm{FD}_k(X)$ is the optimum value of each response variable optimized individually over the experimental region and $\mathrm{FO}_k(X)$ is the estimated value of all the responses given in the same set of causal factors, X. Substituting $F(X)$ in Eq. (7) with $S(X)$ in Eq. (13), the transformed function, $T(X,r)$, in the case of multi-objectives can be given as follows:

$$T(X,r) = \left(\sum_{k=1}^{n} \left\{ \frac{\mathrm{FD}_k(X) - \mathrm{FO}_k(X)}{\mathrm{S.D.}_k} \right\}^2 \right)^{1/2} + r^{-1} \sum_{i=1}^{p} \{\phi_i G_i(X)\}^2 + r^{-1} \sum_{j=1}^{q} \{H_i(X)\}^2 \quad (14)$$

The simultaneous optimum solution is estimated as the point, $X(r)$, which gives minimum value of $T(X,r)$ in Eq. (14).

3.3. Formulator's preference

Further improvement of the distance function given in Eq. (13) can be made as follows [11]:

$$S(X) = \left(\sum_{k=1}^{n} \left\{ \frac{\mathrm{FD}_k(X) - \mathrm{FO}_k(X)}{\mathrm{S.D.}_k} \right\}^p \right)^{1/p} \quad (15)$$

where p is a parameter relating to the impartiality among the response variables. Eq. (13) is the special case of Eq. (15) for equal weightings. Increasing the p value to greater than 2 leads to an enlargement in the importance of the response of which deviation from the optimum value was greater than that of the other responses. Thus, the formulator's preference can be incorporated into the multi-objective function to a certain extent as a function of the p values in Eq. (15).

4. Numerical examples

4.1. Controlled-release tablet of theophylline

The simultaneous optimization method incorporating ANN was applied to the development of a theophylline tablet, which has an optimized release behavior [30]. Plasma concentrations of theophylline were predicted on the basis of pharmacokinetic analysis to obtain desirable release profiles of theophylline from the tablets. Controse, the mixture of hydroxypropylmethylcellulose (HPMC) and lactose, was employed as a gel-forming material while cornstarch was used as a disintegrant.

4.1.1. Preparation of model formulations

As model formulations, 16 kinds of theophylline tablets composed of Controse and cornstarch were prepared according to the three-factor spherical second-order composite experimental design (Table 1). The amounts of Controse (X_1), cornstarch (X_2) and compression pressure (X_3) were selected as causal factors. The amount of theophylline was fixed at a constant value (100 mg/tablet). The three-factor spherical second-order composite experimental design shown in Table 1 requires a total of 16 experiments including a repetition of the center experiment. The first eight experiments represented a factorial design for three factors at two levels, resulting in 2^3 trials. Three additional levels were selected: zero represented the base level midway between the aforementioned levels and the positive and negative 1.732 values represented the extreme values. This experimental design has a unique characteristic that each experimental point was located at the same distance from the center experiment.

4.1.2. Analysis of release behaviors

Release profiles of theophylline from the tablets in Table 1 showed the fast release rate during the initial stage and consecutive slow release in the following

Table 1
Experimental design and model formulations of controlled-release tablets of theophylline

Formulation	X_1	Controse (mg)	X_2	Cornstarch (mg)	X_3	Pressure (kg F/cm^2)
1	−1	71	−1	11	−1	92
2	−1	71	−1	11	1	208
3	−1	71	1	39	−1	92
4	−1	71	1	39	1	208
5	1	129	−1	11	−1	92
6	1	129	−1	11	1	208
7	1	129	1	39	−1	92
8	1	129	1	39	1	208
9	−1.732	50	0	25	0	150
10	1.732	150	0	25	0	150
11	0	100	−1.732	0	0	150
12	0	100	1.732	50	0	150
13	0	100	0	25	−1.732	50
14	0	100	0	25	1.732	250
15	0	100	0	25	0	150
16	0	100	0	25	0	150

The amount of theophylline was fixed at 100 mg.

stage. Assuming first-order release kinetics, the weight fraction of theophylline remaining at time t is given:

$$W = W_0 \exp(-kt) \tag{16}$$

where W is the weight fraction of theophylline remaining at time t, W_0 is the initial weight fraction of theophylline and k is the first-order release rate constant. The fast and slow release fractions of theophylline in the tablets are:

$$\begin{aligned} W &= W_f + W_s = W_{f0}\exp(-k_f t) + W_{s0}\exp(-k_s t) \\ &= W_{f0}\exp(-k_f t) + (W_0 - W_{f0})\exp(-k_s t) \end{aligned} \tag{17}$$

where W_f and W_s are the fast and slow release fractions, W_{f0} and W_{s0} are the fast and slow release fractions at $t=0$, and k_f and k_s are the fast and slow release rate constants, respectively. The percentage of released theophylline (C, %) at time t is given as:

$$C = 100 - \frac{100\{W_{f0}\exp(-k_f t) + (W_0 - W_{f0})\exp(-k_s t)\}}{W_0} \tag{18}$$

A wide deviation in the release parameters was observed, indicating that they were greatly affected by changes in the level of causal factors such as the quantities of Controse and cornstarch as well as the compression pressure.

4.1.3. Prediction of release parameters

Three causal factors corresponding to different levels of Controse (X_1), cornstarch (X_2) and compression pressure (X_3) were used as each unit of the input layer in the partitioned ANN. The output layer was composed of three response variables, i.e., W_{f0}, k_f and k_s. A set of release parameters and causal factors were used as tutorial data for the ANN. To optimize the structure of an ANN, a simulated annealing technique can be applied employing AIC as judging standard as defined in Eq. (4).

As the optimal ANN structure, two units in the hidden layer for each response were selected. The degree of over-determination β in Eq. (3) was calculated to be 1.45, suggesting that a 45% over-determined approximation for each response can be attained as a function of three causal factors. Fig. 3 shows the three-dimensional diagrams of each response variable as a function of X_1 (amount of Controse) and X_2 (amount of cornstarch) at a constant value of X_3 (compression pressure, 150 kg F/cm^2). Increase in the amount of Controse (X_1) resulted in decrease in the values of W_{f0}, k_f and k_s while these parameters increased with increasing the quantities of cornstarch (X_2). This suggests that Controse acts as a controlling agent in the release of

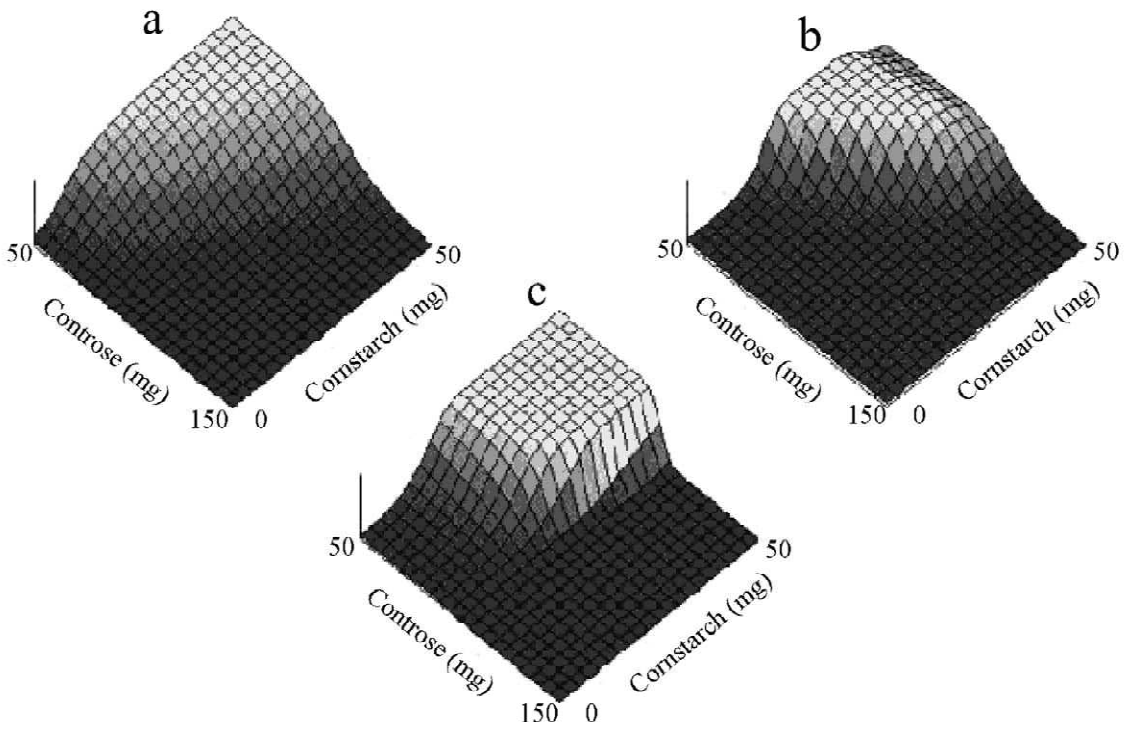

Fig. 3. Response surfaces of W_{f0}, k_f and k_s predicted using the partitioned ANN as a function of the amounts of Controse and cornstarch at a constant compression pressure (150 kg F/cm^2). (a) W_{f0} (6.96–61.0 mg); (b) k_f (5.71–37.3 h^{-1}); (c) k_s (0.0608–1.00 h^{-1}).

theophylline from the tablet while cornstarch acts as an accelerator. Nonlinear relationships between the causal factors and the release parameters were represented well with the response surfaces generated by the partitioned ANN. However, the second-order polynomial equation exhibited relatively plane surfaces for all release parameters (Fig. 4). In the multi-variable system, the quantitative relationships between causal factors and response variables are thought to be complex and nonlinear. The ANN appears to be more useful than polynomial equations in cases where approximations of such relationships are required.

4.1.4. Simultaneous optimization

The optimization of the controlled-release tablet containing theophylline was performed according to the distance function method defined in Eq. (14). The maximum and the minimum values of the individual release parameters were estimated under the restriction of experimental regions ($3 \geq X_1^2 + X_2^2 + X_3^2$; in coded form). Several cases composed of a set of these parameters were investigated in order to predict the simultaneous optimal formulations. Fig. 5 shows the release profiles of the respective optimal formulations predicted by a set of the minimum W_{f0}, the maximum k_f and the minimum k_s as individual optimal values. Release profiles predicted by the partitioned ANN coincided well with the experimental values while the approximations based on the polynomial equation were somewhat poorer. These findings demonstrate that a multi-objective optimization technique incorporating the partitioned ANN is more useful for optimizing pharmaceutical formulations, compared with a classical response surface approach based on the polynomial regression equation.

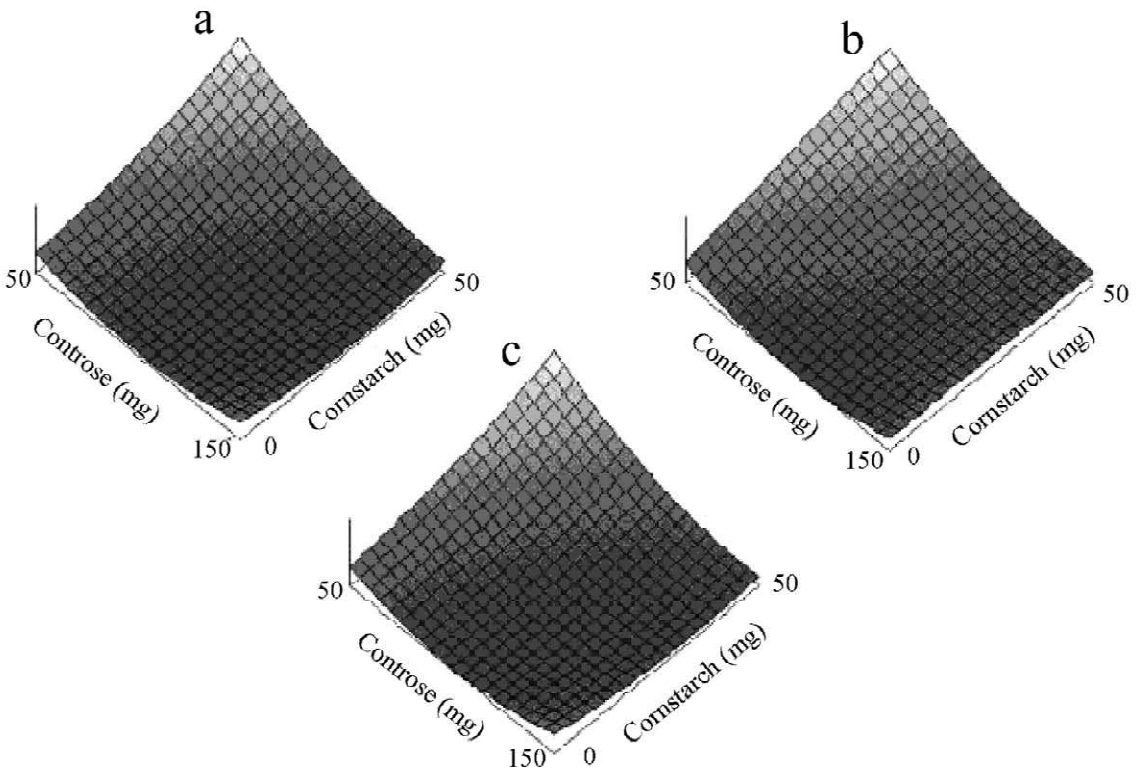

Fig. 4. Response surfaces of W_{f0}, k_f and k_s predicted using the second-order polynomial equation as a function of the amounts of Controse and cornstarch at a constant compression pressure (150 kg F/cm^2). (a) W_{f0} (1.55–71.6 mg); (b) k_f (2.87–32.4 h^{-1}); (c) k_s (−0.0125–0.982 h^{-1}).

4.2. Ketoprofen hydrogel containing chemical enhancers

A large number of chemical enhancers were synthesized by using L-menthol as a lead compound [64]. Structure–activity and structure–toxicity relationship studies revealed that the lipophilicity of molecules, the molecular size, and the affinity to the skin tissues were important factors for exhibiting enhancement activity as well as skin damage [65]. Among these chemicals, 1-*O*-ethyl-3-*n*-butylcyclohexanol (OEBC) was found to be a promising compound with potent enhancement activity and relatively low skin irritation [65]. Diisopropyl adipate (DIA) has also been reported to enhance the permeation of diclofenac [66], indomethacin [67], and hydrocortisone butyrate [68] through the skin. In addition, several studies have reported that the lipophilic penetration enhancer combination with a short-chain alkanol, such as isopropanol (IPA), could produce a synergistic enhancement and decrease the skin irritation [69–71]. The combined effect of penetration absorption as well as skin irritation was evaluated in the ketoprofen hydrogel formulation incorporating OEBC, DIA and IPA [31].

4.2.1. Preparation of model formulations

A three-factor spherical second-order composite experimental design was used. As shown in Table 2, 16 types of ketoprofen hydrogels composed of hydroxypropylcellulose (HPC) and hydroxyethylcellulose (HEC) as gel bases, OEBC and DIA as penetration enhancers, and IPA and water as solvents were prepared. The quantities of OEBC

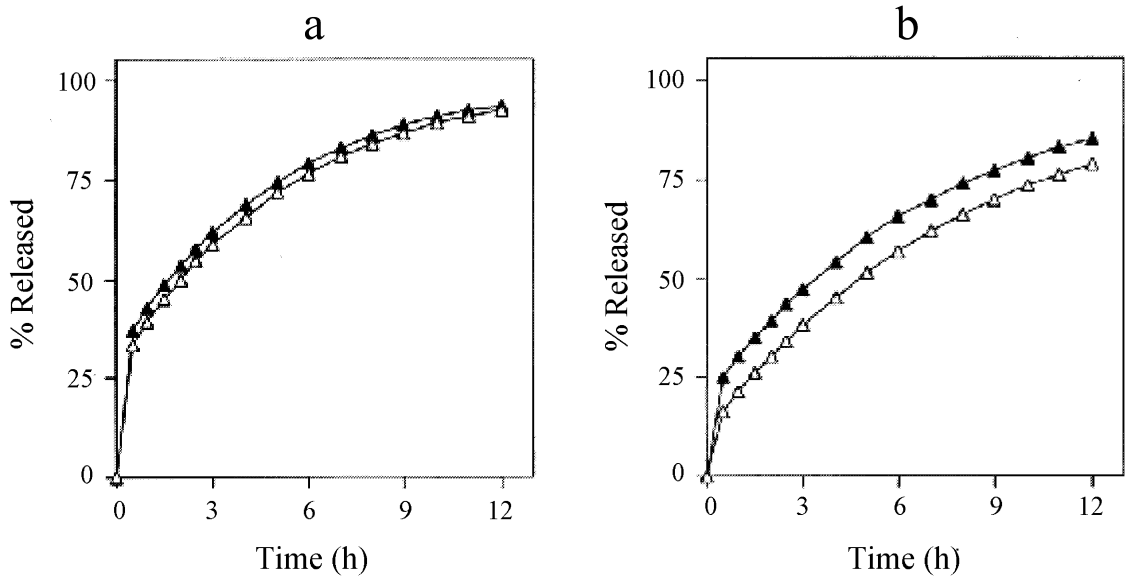

Fig. 5. Release profiles of theophylline from optimum formulations. (a) The formulation optimized using the partitioned ANN; (b) the formulation optimized using the second-order polynomial equation. △, Predicted values; ▲, experimental values (mean of six determinations).

Table 2
Experimental design and model formulations of ketoprofen hydrogels

Formulation	X_1	OEBC (%)	X_2	DIA (%)	X_3	IPA (%)
1	−1	0.42	−1	1.06	−1	24.23
2	−1	0.42	−1	1.06	1	35.77
3	−1	0.42	1	3.94	−1	24.23
4	−1	0.42	1	3.94	1	35.77
5	1	1.58	−1	1.06	−1	24.23
6	1	1.58	−1	1.06	1	35.77
7	1	1.58	1	3.94	−1	24.23
8	1	1.58	1	3.94	1	35.77
9	−1.732	0	0	2.50	0	30.00
10	1.732	2.00	0	2.50	0	30.00
11	0	1.00	−1.732	0	0	30.00
12	0	1.00	1.732	5.00	0	30.00
13	0	1.00	0	2.50	−1.732	20.00
14	0	1.00	0	2.50	1.732	40.00
15	0	1.00	0	2.50	0	30.00
16	0	1.00	0	2.50	0	30.00

The amounts of ketoprofen, HPC and HEC were fixed at 3, 1 and 1%, respectively. The total amount of each hydrogel was adjusted to 10.00 g by addition of water.

(X_1), DIA (X_2) and IPA (X_3) were selected as causal factors. The amounts of ketoprofen, HPC, and HEC were fixed at 3, 1, and 1%, respectively. The total amount of each hydrogel was adjusted to 100% by the addition of water. The gel bases, HPC and HEC, were dissolved in water. Ketoprofen was dissolved in IPA containing OEBC and DIA, separately. Then, the both components were mixed well, and the resulting hydrogels were stored in air-tight containers at room temperature prior to use.

4.2.2. Evaluation of percutaneous absorption and skin irritation

To evaluate the percutaneous absorption of experimental formulations, the test hydrogel was applied to the rat abdominal skin. Blood samples were taken via the jugular vein at 1–8 h after topical application. The concentration of ketoprofen in the blood was analyzed using a high-performance liquid chromatography. The rate of penetration (R_p) of ketoprofen was estimated from a pharmacokinetic model based on the assumption that the rate of penetration of ketoprofen absorbed from the hydrogel is constant after a lag time.

Irritation evoked by model formulations on rat skin was microscopically judged at the end of absorption experiment. The site of application of each formulation was excised and fixed. Tissues were divided into small pieces and stained with hematoxylin and eosin. All sections were examined by light microscopy. The microscopic findings were graded as five levels of irritation from no change to a marked change including liquefaction of epidermis, edema of subepidermis, collagen fiber swelling, and inflammatory cell infiltration in both the dermis and hypodermis, as well as degeneration of skin appendages. The total irritation score (TIS) was obtained by summation of each irritation score and used as an index of skin damage caused by the application of ketoprofen hydrogel.

4.2.3. Prediction of response variables

An ANN was applied to the prediction of response variables such as R_p and TIS as a function of causal factors. The second-order polynomial regression equation was used for comparing the prediction ability. Three causal factors corresponding to different levels of OEBC, DIA, and IPA were used as each unit of the input layer of ANN. Response variables, R_p and TIS, were predicted by the partitioned structure of ANN. To optimize the structure of ANN, the simulated annealing technique was applied, employing AIC defined in Eq. (4) as a judging standard. To achieve reasonable prediction of both R_p and TIS, three units were appropriate as the number of units in the hidden layer.

The response surfaces generated by the ANN and a second-order polynomial equation are shown in Figs. 6 and 7, respectively, as a function of X_1 (the amounts of OEBC) and background factors [sum of the amounts of DIA (X_2) and IPA (X_3)]. The response surfaces generated by ANN clearly demonstrated the combined effect of OEBC and the other factors on both R_p and TIS values. Furthermore, the nonlinear relationship between the factors and the

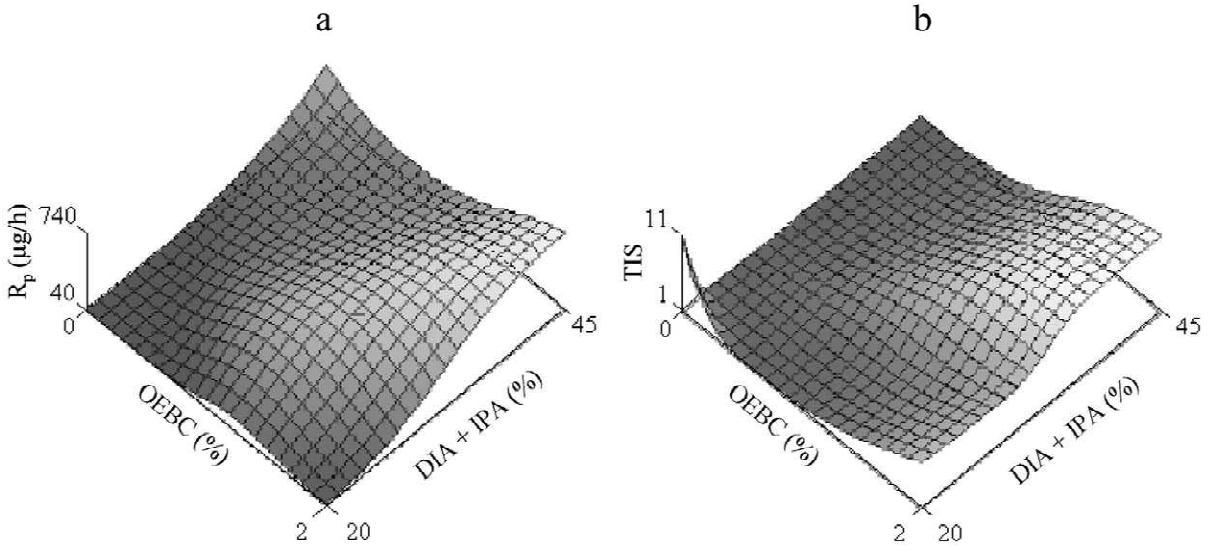

Fig. 6. Response surface of R_p and TIS predicted using the partitioned ANN as a function of the amounts of OEBC and background factors (DIA + IPA). (a) R_p; (b) TIS.

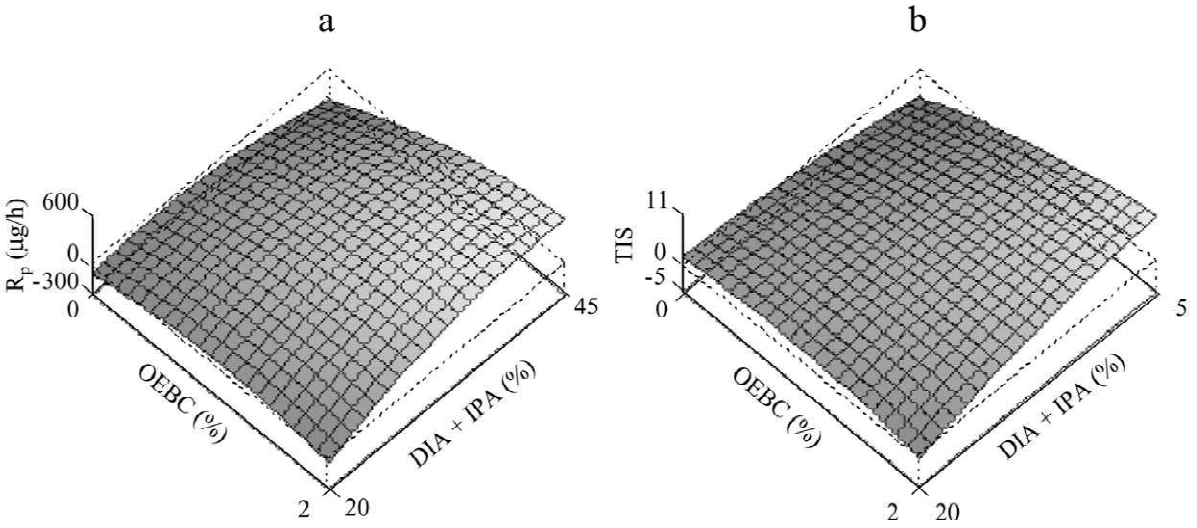

Fig. 7. Response surface of R_p and TIS predicted using the second-order polynomial equation as a function of the amounts of OEBC and background factors (DIA + IPA). (a) R_p; (b) TIS.

responses was represented well with the response surface predicted by ANN (Fig. 6). However, a second-order polynomial equation exhibited relatively plain surfaces (Fig. 7). Furthermore, polynomial equation analysis predicted negative values in the boundary region of the experimental limits that were out of physical reality. Generally, the quantitative relationships between causal factors and response variables in vivo are thought to be complex and nonlinear. This may suggest that ANN is more useful than the polynomial equation in cases when approximations of such relationships are required.

Table 3
Optimal formulations predicted by the ANN and the second-order polynomial regression equation

Factor	ANN	Polynomial equation
OEBC (%)	1.15	0.610
DIA (%)	1.15	0.923
IPA (%)	31.1	36.7

4.2.4. Simultaneous optimization

Simultaneous optimization was performed according to the distance function defined in Eq. (14). Results of the simultaneous optima are given in Table 3. The concentration of OEBC predicted by the ANN was twofold greater than that estimated by the polynomial equation. However, no significant difference was seen between the ANN and the polynomial equation in the optimal concentrations of DIA and IPA. The R_p and TIS values predicted by the ANN coincided well with the experimental values, whereas the approximations based on the polynomial equation were somewhat poorer (Table 4). Furthermore, the optimal formulation predicted

Table 4
Predicted and experimental values of response variables of optimal formulations based on the ANN and the second-order polynomial regression equation

Response	ANN		Polynomial equation	
	Predicted	Experimental[a]	Predicted	Experimental[a]
R_p (μg/h)	621	623±73	506	399±31
TIS	4	4±1	6	4±1

[a] Mean±S.D. for three determinations.

by ANN exhibited the most prominent activity on the percutaneous absorption of ketoprofen despite the finding that the skin damage was sufficiently low.

In general, the strong percutaneous absorption action can be obtained by increasing the concentrations of the chemical enhancers, but this increase often causes significant skin damage. However, the above findings clearly demonstrate that strong enhancement action as well as the low skin damage can be attained by seeking the optimal combination of enhancers and other additives in the formulations. It is obvious that it would be unable to reach the exact combination of causal factors using a normal analysis based on a one-factor-at-a-time experiment. Optimization of ketoprofen hydrogel containing OEBC, DIA and IPA was successfully accomplished by means of the optimization technique incorporating the partitioned ANN.

Acknowledgements

This work was supported by the Ministry of Education, Science, Sports and Culture of Japan.

References

[1] A.I. Khuri, J.A. Cornel, Response Surface, Design and Analysis, Marcel Dekker, New York, 1987.
[2] G.A. Lewis, D. Mathieu, R. Phan-Yan-Luu, Pharmaceutical Experimental Design, Marcel Dekker, New York, 1999.
[3] K. Takayama, N. Nambu, T. Nagai, Computer optimization of formulation of flufenamic acid/polyvinylpolypyrrolidone/methyl cellulose solid dispersions, Chem. Pharm. Bull. 31 (1983) 4496–4507.
[4] T. Takai, K. Takayama, N. Nambu, T. Nagai, Optimum formulation of griseofulvin/hydroxypropyl cellulose solid dispersions with desired dissolution properties, Chem. Pharm. Bull. 32 (1984) 1942–1947.
[5] K. Takayama, H. Imaizumi, N. Nambu, T. Nagai, Mathematical optimization of formulation of indomethacin/polyvinylpolypyrrolidone/methyl cellulose solid dispersions by the sequential unconstrained minimization technique, Chem. Pharm. Bull. 33 (1985) 292–300.
[6] Y. Akitoshi, K. Takayama, Y. Machida, T. Nagai, Computer optimization of the formulation of acrylic plaster, Chem. Pharm. Bull. 33 (1985) 4536–4543.
[7] J. Privorsky, K. Takayama, T. Nagai, D. Waitzova, J. Elis, Combination effect of penetration enhancers and propylene glycol on in vitro transdermal absorption of insulin, Drug Design and Delivery 2 (1987) 91–97.
[8] K. Arai, K. Takayama, Y. Nambu, N. Nambu, H. Ueda, Y. Machida, T. Nagai, Preparation of polylactic acid–polylipoic acid nanospheres as drug targeting carriers, Drug Design and Delivery 2 (1987) 109–120.
[9] K. Takayama, T. Nagai, Novel computer optimization methodology for pharmaceutical formulations investigated by using sustained-release granules of indomethacin, Chem. Pharm. Bull. 37 (1989) 160–167.
[10] K. Takayama, H. Okabe, Y. Obata, T. Nagai, Formulation design of indomethacin gel ointment containing D-limonene using computer optimization methodology, Int. J. Pharm. 61 (1990) 225–234.
[11] K. Takayama, T. Nagai, Simultaneous optimization for several characteristics concerning percutaneous absorption and skin damage of ketoprofen hydrogels containing D-limonene, Int. J. Pharm. 74 (1991) 115–126.
[12] M. Hirata, K. Takayama, T. Nagai, Formulation optimization of sustained-release tablet of chlorpheniramine maleate by means of extreme vertices design and simultaneous optimization technique, Chem. Pharm. Bull. 74 (1992) 741–746.
[13] K. Imasaka, K. Takayama, M. Yoshida, H. Omichi, T. Nagai, A new attempt to design erodible terpolymers using computer-optimization methodology, Eur. Polym. J. 7 (1992) 709–712.
[14] S. Ogawa, T. Kamijima, Y. Miyamoto, M. Miyajima, H. Sato, K. Takayama, T. Nagai, A new attempt to solve the scale-up problem for granulation using response surface methodology, J. Pharm. Sci. 83 (1994) 439–443.
[15] M. Matsumura, H. Nakagami, T. Yamao, K. Takayama, T. Nagai, Computer optimization for the formulation of controlled-release theophylline tablet made of micronized low-substituted hydroxypropylcellulose and methylcellulose, Chem. Pharm. Bull. 42 (1994) 1902–1908.
[16] K.K. Levison, K. Takayama, K. Isowa, K. Okabe, T. Nagai, Formulation optimization of indomethacin gels containing a combination of three kinds of cyclic monoterpenes as percutaneous penetration enhancers, J. Pharm. Sci. 83 (1994) 1367–1372.
[17] Y. Miyamoto, S. Ogawa, M. Miyajima, M. Matsui, H. Sato, K. Takayama, T. Nagai, An application of the computer optimization technique to wet granulation process involving explosive growth of particles, Int. J. Pharm. 149 (1997) 25–36.
[18] D.E. Fonner Jr., J.R. Buck, G.S. Banker, Mathematical optimization techniques in drug product design and process analysis, J. Pharm. Sci. 59 (1970) 1587–1596.
[19] J.B. Schwartz, J.R. Flamholz, R.H. Press, Computer optimization of pharmaceutical formulations I. General procedure, J. Pharm. Sci. 62 (1973) 1165–1170.
[20] G. Derringer, R. Suich, Simultaneous optimization of several response variables, J. Quality Tech. 12 (1980) 214–219.
[21] A.I. Khuri, M. Conlon, Simultaneous optimization of multiple responses represented by polynomial regression functions, Technometrics 23 (1981) 363–375.
[22] A.D. McLeod, F.C. Lam, P.K. Gupta, C.T. Hung, Optimized

synthesis of polyglutaraldehyde nanoparticles using central composite design, J. Pharm. Sci. 77 (1988) 704–710.
[23] B.G. Muller, H. Leuenberger, T. Kissel, Albumin nanospheres as carriers for passive drug targeting: An optimized manufacturing technique, Pharm. Res. 13 (1996) 32–37.
[24] A.S. Achanta, J.G. Kowalski, C.T. Rhodes, Artificial neural networks: Implications for pharmaceutical sciences, Drug Dev. Ind. Pharm. 21 (1995) 119–155.
[25] A.S. Hussain, X. Yu, R.D. Jornson, Application of neural computing in pharmaceutical product development, Pharm. Res. 8 (1991) 1248–1252.
[26] J. Takahara, K. Takayama, T. Nagai, Multi-objective simultaneous optimization technique based on an artificial neural network in sustained release formulations, J. Control. Release 49 (1997) 11–20.
[27] J. Takahara, K. Takayama, K. Isowa, T. Nagai, Multi-objective simultaneous optimization based on artificial neural network in a ketoprofen hydrogel formula containing *O*-ethylmenthol as a percutaneous absorption enhancer, Int. J. Pharm. 158 (1997) 203–210.
[28] K. Takayama, J. Takahara, M. Fujikawa, H. Ichikawa, T. Nagai, Formula optimization based on artificial neural networks in transdermal drug delivery, J. Control. Release 62 (1999) 161–170.
[29] K. Takayama, M. Fujikawa, T. Nagai, Artificial neural network as a novel method to optimize pharmaceutical formulations, Pharm. Res. 16 (1999) 1–6.
[30] K. Takayama, A. Morva, M. Fujikawa, Y. Hattori, Y. Obata, T. Nagai, Formula optimization of theophylline controlled-release tablet based on artificial neural networks, J. Control. Release 68 (2000) 175–186.
[31] P.C. Wu, Y. Obata, M. Fujikawa, C.J. Li, K. Higashiyama, K. Takayama, Simultaneous optimization based on artificial neural networks in ketoprofen hydrogel formula containing *O*-ethyl-3-butylcyclohexanol as percutaneous absorption enhancer, J. Pharm. Sci. 90 (2001) 1004–1014.
[32] C.P. Lim, S.S. Quek, K.K. Peh, Prediction of drug release profiles using an intelligent learning system: an experimental study in transdermal iontophoresis, J. Pharm. Biomed. Anal. 31 (2003) 159–168.
[33] K. Kachrimanis, V. Karamyan, S. Malamataris, Artificial neural networks (ANNs) and modeling of powder flow, Int. J. Pharm. 250 (2003) 13–23.
[34] A.P. Plumb, R.C. Rowe, P. York, C. Doherty, The effect of experimental design on the modeling of a tablet coating formulation using artificial neural networks, Eur. J. Pharm. Sci. 16 (2002) 281–288.
[35] S. Ibric, M. Jovanovic, Z. Djuric, J. Parojcic, L. Solomun, The application of generalized regression neural network in the modeling and optimization of aspirin extended release tablets with Eudragit RS PO as matrix substance, J. Control. Release 82 (2002) 213–222.
[36] S. Vaithiyalingam, M.A. Khan, Optimization and characterization of controlled release multi-particulate beads formulated with a customized cellulose acetate butyrate dispersion, Int. J. Pharm. 234 (2002) 179–193.
[37] T. Aoyama, Y. Suzuki, H. Ichikawa, Neural networks applied to quantitative structure–activity relationship analysis, J. Med. Chem. 33 (1990) 2583–2590.
[38] J. Huuskonen, QSAR modeling with the electrotopological state indices: predicting the toxicity of organic chemicals, Chemosphere 50 (2003) 949–953.
[39] D.K. Agrafiotis, W. Cedeno, V.S. Lobanov, On the use of neural network ensembles in QSAR and QSPR, J. Chem. Inf. Comput. Sci. 42 (2002) 903–911.
[40] A.S. Hussain, R.D. Jornson, N. Vachhrajani, W.A. Ritschel, Feasibility of developing a neural network for prediction of human pharmacokinetic parameters from animal data, Pharm. Res. 10 (1993) 466–469.
[41] E. Brier, J.M. Zurada, G.R. Aronoff, Neural network predicted peak and trough gentamicin concentrations, Pharm. Res. 12 (1995) 406–412.
[42] J.V.S. Gobburu, W.H. Shelver, Quantitative structure–pharmacokinetic relationship (QSPR) of beta blockers derived using neural networks, J. Pharm. Sci. 84 (1995) 862–865.
[43] B.P. Smith, M.E. Brier, Statistical approach to neural network model building for gentamicin peak predictions, J. Pharm. Sci. 85 (1995) 65–69.
[44] H. Gao, V. Shanmugasundaram, P. Lee, Estimation of aqueous solubility of organic compounds with QSPR approach, Pharm. Res. 19 (2002) 497–503.
[45] I.S. Nestorov, S.T. Hadjitodorov, I. Petrov, M. Rowland, Empirical versus mechanistic modeling: comparison of an artificial neural network to a mechanistically based model for quantitative structure pharmacokinetic relationship of a homologues serious of barbiturates, AAPS PharmSci. 1 (1999) E17.
[46] B.K. Jha, S.S. Tambe, B.D. Kulkarni, Estimating diffusion coefficients of a micellar system using an artificial neural network, J. Coll. Interf. Sci. 170 (1995) 392–398.
[47] C.W. Lim, S. Fujiwara, F. Yamashita, M. Hashida, Prediction of human skin permeability using a combination of molecular orbital calculations and artificial neural network, Biol. Pharm. Bull. 25 (2002) 361–366.
[48] T. Degim, J. Hadgraft, S. Ilbasmis, Y. Ozkan, Prediction of skin penetration using artificial neural network (ANN) modeling, J. Pharm. Sci. 92 (2003) 656–664.
[49] X.C. Fu, X.W. Ma, W.Q. Liang, Prediction of skin permeability using an artificial neural network, Pharmazie 57 (2002) 655–656.
[50] S. Fujiwara, F. Yamashita, M. Hashida, Prediction of Caco-2 cell permeability using a combination of MO-calculation and neural network, Int. J. Pharm. 237 (2002) 95–105.
[51] J.N. Weinstein, K.W. Kohn, M.R. Grever, V.N. Viswanadhan, L.V. Rubinstein, A.P. Monks, D.A. Scudiero, L. Welch, A.D. Koutsoukos, A.J. Chiausa, K.D. Paull, Neural computing in cancer drug development: predicting mechanism of action, Science 258 (1992) 447–451.
[52] J.L. McClelland, D.E. Rumelhart, Explorations in Parallel Distributed Processing, MIT Press, Cambridge, MA, 1988.
[53] R.J. Erb, Introduction to backpropagation neural network computation, Pharm. Res. 10 (1993) 165–170.
[54] H. Murase, S. Koyama, N. Honami, T. Kuwabara, Kalman filter neuron training, Bull. Univ. Osaka Pref., Ser. B 43 (1991) 91–101.

[55] T.B. Blank, S.D. Brown, Adaptive, global, extended Kalman filters for training feedforward neural networks, J. Chemometr. 8 (1994) 391–407.
[56] R. Simutis, I. Havlik, M. Dors, A. Luebbert, Training of artificial neural networks extended by linear dynamic subsystems, Process Control Qual. 4 (1993) 211–220.
[57] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 200 (1983) 671–680.
[58] R.P. Lippman, An introduction to computing with neural nets, IEEE ASSP Mag. April (1987) 4–22.
[59] D.G. Bounds, P.J. Lloyd, A multilayer perceptron network for the diagnosis of low back pain, in: Proceedings of the 2nd IEEE International Conference on Neural Networks, San Diego, CA, 24–27 July 1988, pp. II481–489.
[60] G. Cybenko, Approximations by superpositions of a sigmoidal function, Math. Control Signals Syst. 2 (1989) 303–314.
[61] W.C. Carpenter, M.E. Hoffman, Understanding neural network approximations and polynomial approximations helps neural network performance, AI Expert March (1995) 31–33.
[62] A.V. Fiacco, G.P. McCormik, Nonlinear Programming: Sequential Unconstrained Minimization Techniques, Wiley, New York, 1968.
[63] J. Kowalik, M.R. Osborne, Methods for Unconstrained Optimization Problems, Elsevier, New York, 1968.
[64] Y. Obata, H. Sato, C.J. Li, K. Takayama, K. Higashiyama, T. Nagai, K. Isowa, Effect of synthesized cyclohexanol derivatives using L-menthol as a lead compound on the percutaneous absorption of ketoprofen, Int. J. Pharm. 198 (2000) 191–200.
[65] Y. Obata, C.J. Li, M. Fujikawa, K. Takayama, H. Sato, K. Higashiyama, K. Isowa, T. Nagai, Evaluation and strucure–activity relationship of synthesized cyclohexanol derivatives on percutaneous absorption of ketoprofen using artificial neural network, Int. J. Pharm. 212 (2001) 223–231.
[66] H. Okuyama, Y. Ikeda, S. Kasai, K. Inamori, K. Takayama, T. Nagai, Influence of diisopropyl adipate on percutaneous absorption and subcutaneous tissue penetration of diclofenac from alcoholic gel ointment, Yakuzaigaku 59 (1999) 75–83.
[67] T. Inagi, T. Muramatsu, H. Terada, Interaction of indomethacin with the vehicle component diisopropyl adipate, Chem. Pharm. Bull. 34 (1986) 1228–1234.
[68] Y. Ozawa, T. Yamahira, H. Sunada, T. Nadai, Influence of fatty acid–alcohol esters on percutaneous absorption of hydrocortizone butyrate propionate, Chem. Pharm. Bull. 36 (1988) 2145–2151.
[69] M. Goodman, B.W. Barry, Lipid–protein-partitioning (LPP) theory of skin enhancer activity: finite dose technique, Int. J. Pharm. 57 (1989) 29–40.
[70] A. Arellano, S. Santoyo, C. Martin, O. Ygartua, Influence of propylene glycol and isopropyl myristate on the in vitro percutaneous penetration of diclofenac sodium from carbopol gels, Eur. J. Pharm. Sci. 7 (1998) 129–135.
[71] M. Goldberg-Cettina, P. Liu, J. Nightingale, T. Kurihara-Bergstrom, Enhanced transdermal delivery of estradiol in vitro using binary vesicles of isopropyl myristate and short-chain alkanols, Int. J. Pharm. 114 (1995) 237–245.