European Journal of Pharmaceutics and Biopharmaceutics 74 (2010) 316–323

Contents lists available at ScienceDirect

European Journal of Pharmaceutics and Biopharmaceutics

journal homepage: www.elsevier.com/locate/ejpb

Research paper

Artificial neural networks in the optimization of a nimodipine controlled release tablet formulation

Panagiotis Barmpalexis [a], Feras Imad Kanaze [b,1], Kyriakos Kachrimanis [a], Emanouil Georgarakis [a,*]

[a] Department of Pharmaceutical Technology, Aristotle University of Thessaloniki, Thessaloniki, Greece
[b] Pharmathen S.A., Pharmaceutical Industry, Athens, Greece

ARTICLE INFO

Article history:
Received 10 June 2009
Accepted in revised form 30 September 2009
Available online 6 October 2009

Keywords:
Artificial neural networks
Pruning
Controlled release
Euclidian distance
Nimodipine

ABSTRACT

Artificial neural networks (ANNs) were employed in the optimization of a nimodipine zero-order release matrix tablet formulation, and their efficiency was compared to that of multiple linear regression (MLR) on an external validation set. The amounts of PEG-4000, PVP K30, HPMC K100 and HPMC E50LV were used as independent variables following a statistical experimental design, and three dissolution parameters (time at which the 90% of the drug was dissolved, $t_{90\%}$, percentage of nimodipine released in 2 and 8 h, Y_{2h}, and Y_{8h}, respectively) were chosen as response variables. It was found that a feed-forward back-propagation ANN with eight hidden units showed better fit for all responses (R^2 of 0.96, 0.90 and 0.98 for $t_{90\%}$, Y_{2h} and Y_{8h}, respectively) compared to the MLR models (0.92, 0.87 and 0.92 for $t_{90\%}$, Y_{2h} and Y_{8h}, respectively). The ANN was further simplified by pruning, which preserved only PEG-4000 and HPMC K100 as inputs. Optimal formulations based on ANN and MLR predictions were identified by minimizing the standardized Euclidian distance between measured and theoretical (zero order) release parameters. The estimation of the similarity factor, f_2, confirmed ANNs increased prediction efficiency (81.98 and 79.46 for the original and pruned ANN, respectively, and 76.25 for the MLR).

1. Introduction

Nimodipine (NIM) is an active pharmaceutical ingredient (API) that belongs to the class of pharmacological agents known as calcium channel blockers [1]. It has cerebrovasodilatory and neuronal affects and is mainly used in the treatment of subarachnoid haemorrhage (SAH), focal or global ischemia, as well as epilepsy [2]. Oral administration of NIM is associated with certain problems such as frequent dosing (30–60 mg every 4–8 h) varying half life and fluctuating plasma concentrations [1]. The preparation of a NIM-controlled release formulation could be an appealing solution to these problems, and until now, attempts for the preparation of a controlled release formulation include the development of floating-sustained release tablets [3] and soft gelatine capsules [4].

Development of a controlled release tablet formulation is commonly achieved by the use of hydrophilic polymer blends. Hydrophilic matrices are widely accepted because of their biopharmaceutical and pharmacokinetics advantages over conventional dosage forms [5], offering precise modulation of drug release as a result of hydration and swelling. Commonly used polymers for the preparation of hydrophilic matrices include substituted celluloses such as hydroxypropylmethylcellulose (HPMC) and hydroxypropylcellulose (HPC) [6], different grades of polyvinylpyrrolidone (PVP) and polyethylene glycol (PEG).

Furthermore, the preparation of controlled release formulations involves handling and optimizing a large number of factors, such as polymer matrix composition and drug to matrix ratio. Conventional trial-and-error methods require increased amounts of experiments in order to identify an acceptable solution. As the number of factors increases, these types of methods become highly inefficient. Therefore, they have been merely superseded by statistical experimental design methodology known as Design of Experiments (DoE). DoE usually includes a screening and an optimization phase [7]. Screening designs involve experimentation at two extreme factor levels (high and low), in order to identify factor's significance (the insignificant factors are excluded from the next step of optimization). Response surface methodology has been widely used as an optimization technique for drug delivery systems [8–10], and it is proven to be efficient in terms of cost, time and effort. This methodology involves the use of various types of experimental designs involving three or more factor levels, thus enabling the fitting of not only just linear (as is the case for two-level screening designs), but also quadratic or cubic polynomial equations. Mapping of response(s) over the experimental domain allows for the determination of optimum solution(s) [8,9].

* Corresponding author. Department of Pharmaceutical Technology, School of Pharmacy, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece. Tel.: +30 2310 997641; fax: +30 2310 997652.
E-mail address: georgara@pharm.auth.gr (E. Georgarakis).
[1] Present address: Alapis A.B.E.E. Research and Development Centre, R&D Department, Pallini 15351, Athens, Greece.

0939-6411/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejpb.2009.09.011

Despite the advantages of DoE-based polynomial model fitting, often the developed models show bad fit resulting to a poor optimum estimation. An alternative approach that has been successfully applied in cases where conventional DoE methods prove inadequate is the use of feed-forward artificial neural networks (ANNs) [11–15]. ANNs are biologically inspired computer algorithms that act as universal function approximators, having the ability to model highly complex relationships where the response variables are non-linearly related to the independent variables. Two of the main advantages of the ANN are [16] (a) there is no need to assume an underlying data distribution and (b) ANNs are applicable to highly non-linear multivariate problems. However, a critical evaluation of the ANN can reveal several disadvantages including [16] (a) minimizing over-fitting requires a great deal of computational effort and (b) it is not so straightforward to investigate the individual relations between the input variables and the output variables, as the ANN models lack a sound theoretical background.

Therefore, the present work investigates the use of ANNs in the development of a NIM-controlled release matrix tablet formulation typically releasing 90% of the drug in 12 h that follows a zero-order profile. Hydrophilic polymers (HPMC, PVP and PEG) are used as matrix formers. Screening of significant formulation factors (amounts of various hydrophilic polymers) is performed according to a two-level full factorial design, and the levels of the identified significant factors are subsequently optimized by fitting feed-forward ANN models to experimental drug release data obtained according to a circumscribed central composite design (CCD). In order to simplify ANNs' structure, different pruning techniques are applied, and the efficiency of the ANNs is compared to that of conventional polynomial multi-linear regression (MLR) models.

2. Materials and methods

2.1. Materials

Micronized NIM, with mean particle size of 18.2 (±11.2) μm, measured as circle equivalent diameter by digital imaging analysis using a Quantimet 500 system (Leica, Cambridge UK) supplied by Union Quimico Farmaceutica S.A. (UQUIFA, Barcelona, Spain) was used as an API.

Polyethylene glycol (PEG-4000, CLARIANT, Sulzbach, Germany), polyvinylpyrrolidone (PVP K30, BASF Co, Ledgewood, NJ) and two grades of hydroxylpropylmethylcellulose (HPMC) differing in the degree of substitution and viscosity (K100M grade with methoxy content 19–24% w/w and viscosity 80,000–120,000 cps, and E50LV grade with methoxy content 28–30% w/w, and viscosity 40–60 cps), purchased from Dow Chemical Company (Midland, MI, USA), were used as hydrophilic matrix formers.

Magnesium stearate (Mg Stearate) purchased from Katayama (Osaka, Japan) was used as lubricant.

Distilled water and sodium lauryl sulfate (SLS) obtained from COGNIS (Fino Mornasco, Italy) were used for the preparation of dissolution medium.

All other materials and reagents were of analytical grade and used as received.

2.2. Preparation of NIM dispersions and tableting

Pre-weighted amounts of PEG-4000 were placed in aluminium dishes and heated in a water bath at 70 °C to complete melting. NIM and the remaining polymers were then sequentially dispersed in the PEG melt in 5 min intervals. This process leads to a dispersion of crystalline NIM microparticles in the polymer matrix, verified by the strong birefringence when observed microscopically under crossed polarizers, according to the European Pharmacopoeia method using a PriorluxPol petrological microscope (Prior, UK). The resultant dispersions were stored at −18 °C for 2 days and then pulverised using a mortar and a pestle. The 100–150 mesh size fraction was separated, and Mg Stearate (0.5% w/w) was added and mixed for 5 min with a spatula. The powders were stored in hermetically sealed dark-glass vials for further processing.

Appropriate amounts of sample containing 30 mg of NIM were compressed on a manually operated hydraulic press equipped with a 7 mm diameter flat-faced punch and die set pre-lubricated with Mg Stearate. A compression pressure of 3185 N/cm^2 was applied for 5 s (sufficient to obtain compacts of minimum attainable porosity). This procedure led to tablets of varying thickness (2.51–5.27 mm), leading to differences in the initial surface area. However, the fast and extensive swelling of the tablets leads to a rapid increase in the initial surface area, depending on the hydrophilic polymer content; therefore, these initial differences are expected to have a negligible influence on drug release.

In order to prevent the photodecomposition of NIM, all formulations were protected from light.

2.3. HPLC verification of drug content

The actual drug content of tablets was assayed using a validated HPLC method [17]. A Shimadzu Prominence HPLC system consisting of a degasser (Model DGU-20A5), a pump (Model LC-20AD), an auto sampler (Model SIL-20AC), a UV–Vis detector (Model SPD-20A) and a column oven (Model CTO-20AC) was used. Chromatographic analysis was performed on an Interchrom C8 analytical column (5 μm particle size, 250 × 4.6 mm I.D.). The mobile phase used was acetonitrile/water (67.5:32.5, v/v) with a flow rate of 0.9 ml/min, and the column temperature was 40 °C. Injection volume was set at 10 μl, and NIM was detected at 236 nm. The excipients used in this study did not interfere with the assay of NIM.

2.4. In vitro release studies

In vitro release of NIM from tablets was evaluated using apparatus II (rotating paddle method) of the USP29, on a Distek 2100C dissolution tester (Distek, North Brunswick, NJ). The dissolution medium consisted of 1000 ml of distilled water containing 0.5% w/w SLS. The test was performed at 37 ± 0.5 °C under stirring at 50 rpm. Sink conditions were maintained throughout the test. An aliquot of 4 ml of samples was collected at 5, 15, 30, 60, 120, 240, 480, 720, 960 and 1200 min using an automatic sampler (Distek Evolution 4300) and assayed for NIM content by HPLC using the apparatus and methodology described above. Each test was performed in triplicate.

The cumulative% drug release was calculated, and the drug release data were plotted using *SigmaPlot v.8.0* software package (Systat Software, Inc., San Jose, California, USA). Some common release kinetic models (first order, zero order, Higushi, Hixson–Crowell and Korsmeyer–Peppas) were fitted to the results by linear regression analysis in order to describe the release mechanism of the examined formulations.

2.5. Swelling and erosion studies

The swelling behavior of the tablets, described as water absorption capacity, was determined gravimetrically on a modified USP apparatus [18–22] consisting of a 200 mesh stainless steel basket carrying the tablet, mounted above the stirring paddle. The baskets were weighed and immersed in 1000 ml of dissolution medium, under stirring at 50 rpm, allowing the tablet to swell at 37 ± 0.5 °C. The baskets were periodically detached from the appa-

ratus, blotted with absorbent tissue to remove any excess dissolution medium on the surface and weighed. Degree of swelling (% water uptake) was calculated according to the following equation [18]:

$$\text{Degree of swelling (\% water uptake)} = \left[\frac{W_t - W_o}{W_o}\right] \times 100 \quad (1)$$

where W_o is the initial weight of the dry tablet, and W_t is the weight of the wet, swollen tablet.

Matrix erosion was determined on the same tablets, at the same time intervals. Specifically, after weighing, the hydrated matrices were dried in an oven at 60 °C for 24 h, and the remaining dry weight, W_r, was determined. Matrix erosion was calculated according to the formula:

$$\text{Erosion (\% mass loss)} = \left[\frac{W_o - W_r}{W_o}\right] \times 100 \quad (2)$$

The mass loss due to the dissolution of NIM was considered negligible.

2.6. Experimental design

A two-level full factorial design was used as a screening tool, in order to identify significant formulation factors. The amounts of PEG 4000 (X_1), PVP K30 (X_2), HPMC K100 (X_3) and HPMC E50LV (X_4) were examined as formulation factors. Preliminary studies provided the settings for the levels of each factor (Table 1). Three dissolution parameters, namely, the time at which the 90% of the drug was dissolved ($t_{90\%}$) and the % release of NIM in 2 h (Y_{2h}) and in 8 h (Y_{8h}), were used as responses since a single response optimization is thought to yield misleading results [23].

Table 1
Experimental domain and averaged measured responses of three replicate experiments, for the ANN and MLR models.

Run	Factors (mg)				Responses		
	X_1	X_2	X_3	X_4	$t_{90\%}$[b] (h)	Y_{2h}[c] (%)	Y_{8h}[d] (%)
1[a]	90	30	20	20	7.21	33.95	94.95
2	120	20	30	30	9.44	25.23	81.48
3	60	20	30	10	10.00	19.37	78.16
4[a]	90	30	20	40	8.21	25.00	93.74
5	60	40	30	30	10.50	21.51	74.89
6	60	0	30	30	10.96	26.21	70.85
7	60	20	30	50	9.69	15.18	78.48
8[a]	30	30	40	20	12.35	15.46	60.90
9	60	20	50	30	13.81	10.78	57.26
10[a]	90	30	40	20	11.33	22.62	68.94
11[a]	90	10	40	20	13.17	17.45	57.12
12[a]	30	30	20	20	9.54	24.48	79.98
13	60	20	30	30	10.79	18.23	74.99
14[a]	30	30	40	40	12.64	13.58	58.85
15	60	20	10	30	8.61	23.46	83.72
16[a]	30	10	40	40	12.08	11.08	63.33
17[a]	30	10	20	40	9.50	15.43	83.25
18[a]	90	30	40	40	12.16	18.79	69.70
19[a]	90	10	40	40	12.58	14.78	62.78
20[a]	30	10	40	20	13.50	12.41	56.28
21	0	20	30	30	10.18	15.01	72.79
22[a]	90	10	20	20	7.44	29.47	95.08
23[a]	30	30	20	40	9.08	18.52	83.34
24	60	20	30	30	10.89	17.50	76.56
25[a]	30	10	20	20	8.96	21.12	78.95
26[a]	90	10	20	40	8.58	21.18	86.87

[a] The 16 runs of full factorial design.
[b] SD from 0.17 to 1.93.
[c] SD from 0.16 to 3.13.
[d] SD from 0.04 to 5.80.

A linear multiple regression model was used in order to estimate the magnitude of main effects and two-way interactions (Eq. (3)):

$$Y_i = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{14}X_1X_4 + b_{23}X_2X_3 + b_24X_2X_4 + b_34X_3X_4 \quad (3)$$

where Y_i is the measured response, b_0 is an intercept term, b_i to b_{ij} are regression coefficients for the main effects and two-way interactions, respectively, and X_{1-4}, coded levels of formulation factors.

Following screening, optimization of significant factor levels was attempted according to a circumscribed CCD (26 runs, $2^4 + 8$ star points + 2 centre points, Table 1) response surface method. The CCD was constructed by the addition of appropriate data points to the two-level full factorial screening. The additional experiments required were also conducted in a randomized order and in triplicate. Desirable ranges for each response (11.5–12.5 h for $t_{90\%}$, 10–20% for Y_{2h} and 55–65% for Y_{8h}) were considered as constraints for optimization.

Linear (Eq. (3)) and quadratic (Eq. (4)) polynomial models were fitted to the experimental data.

$$Y_i = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{14}X_1X_4 + b_{23}X_2X_3 + b_24X_2X_4 + b_34X_3X_4 + b_{11}X_1^2 + b_{12}X_2^2 + b_{33}X_3^2 + b_{44}X_4 \quad (4)$$

where Y_i is the measured response (dependent factor); b_0 is an intercept; b_i to b_{ij} are regression coefficients for the main effects and two-way interactions, respectively; X_1, X_2, X_3, X_4 are coded levels of independent factors. Cubic models were not used in the present study because they were aliased (i.e., the independent estimation of some main effects and interactions was not possible using cubic models due to insufficient number of data points). The best-fitting mathematical model was selected based on several statistical parameters including the p-values of lack of fit and the predicted residual sum of square (PRESS). The model with a p-value smaller than 0.05 and PRESS values as small as possible was selected as the fittest.

The collected data were modeled using the Design Expert (vs. 6.0.4, Stat-Ease Inc., Minneapolis, MN) software package.

2.7. Artificial neural network modeling

Feed-forward back-propagation neural networks were fitted to the data obtained according to the CCD. The data set was split into a training and a validation subset by randomly selecting 22 formulations and assigning them to the training set, while the remaining 4 were used as a validation subset. Each subset contained one of the two centre points included in the design (runs 13 and 24, Table 1). Four inputs, corresponding to the examined factors, and three output units, corresponding to the selected responses, were used in the developed networks. After preliminary trial-and-error tests using networks with 1–4 hidden layers containing 2–19 hidden units, the appropriate network structure was determined. The logistic sigmoid activation function (Eq. (5)) was used in all units.

$$f(x) = [1 + \exp(-x)]^{-1} \quad (5)$$

The input and output data were scaled within the range from 0 to 1, and the network was trained using the 'vanilla' or standard back-propagation (StBack) algorithm. In StBack, information from the input layer is fed forward through the network, and the link weights between neurons are optimized by backward propagation of the error during the training phase [24], using the three-layer elementary back-propagation topology [25] in order to minimize the error sum of squares cost function. In the present study, learning rate was set at 0.2 and maximum-tolerated difference between

a target value and an output was set at 0. The number of training cycles was selected on the basis of the mean squared error (MSE) of the validation subset, according to the "early stopping" method [26], which prevents the network from over-training (memorization of noise [27]). Additionally, two training algorithms, Back-propagation momentum (BackMom) and Resilient back-propagation (Rprop), were also tested as to whether they can improve the predictive ability of the neural network. The BackMom uses an additional momentum term that introduces the old weight change as a parameter for the computation of new weight change, while the Rprop algorithm uses a local adaptive learning scheme to eliminate the harmful influence of the size of the partial derivative on the weight step [25]. The maximum update value and weight decay exponent were set at 30 and 5, respectively, for the Rprop algorithm, and the momentum term was set at 0.5 for the BackMom algorithm.

Causal and predictive importance (sensitivity and saliency) of the input variables was evaluated in order to identify significant effects on the examined responses [26]. Sensitivity shows the change of the output corresponding to a given change in an input variable. Saliency refers to the increase in the error function when an input is omitted from the network [28]. The modified link-weight-magnitude approach (suitable for non-linear relations) was selected for estimating sensitivity [27] while for saliency, setting the inputs at their mean value and retraining of the network was employed [28].

In order to simplify the optimum trained network, pruning (size reduction) was performed [26]. Pruning can reduce the cost of the net (runtime, memory and cost for hardware implementation), improve the generalization and eliminate unnecessary input or hidden nodes (and/or connections) [25]. Two different pruning algorithms were used, namely, magnitude-based pruning (MBP) that belongs to the so-called sensitivity methods and optimal brain damage (OBD) which is an error-based method [27]. MBP is the simplest weight pruning algorithm. After each training, the link with the smallest weight is removed [25]. OBD approximates the changes of the error function when pruning a certain weight by using Taylor series [25].

The Stuttgart Neural Network Simulator (SNNS version 4.2 for WIN32) software package was employed for the development, training and size reduction in the ANN.

2.8. Validation of multiple linear regression (MLR) and ANN models

The developed MLR quadratic models and the original and pruned ANNs were tested for predictive ability on a randomly selected external validation set (Table 2). Mixtures of composition corresponding to the selected test points were formulated into tablets, and their release parameters were determined. The regression correlation coefficient R^2 of the observed vs. the predicted values was used to compare the performance of the MLR, original and pruned ANN models.

Table 2
External validation set used in the comparison of the ANN and MLR models prediction performance.

Run	Factors (mg)				Responses		
	X_1	X_2	X_3	X_4	$t_{90\%}$[a] (h)	Y_{2h}[b] (%)	Y_{8h}[c] (%)
27	46	25	22	24	9.08	22.69	82.05
28	44	14	29	34	10.72	16.08	71.23
29	32	23	30	21	10.67	15.88	72.10
30	35	30	32	20	11.92	14.48	70.87
31	57	11	40	22	12.94	15.46	58.18
32	31	10	39	20	13.14	11.22	57.45

[a] SD from 0.15 to 1.85.
[b] SD from 1.05 to 3.06.
[c] SD from 2.94 to 4.50.

2.9. Global optimization technique

The ability of the developed MLR and ANN models to propose formulations that approximate a zero-order release profile with 90% of the drug released at 12 h was evaluated by the minimization of the standardized Euclidian distance (Eq. (6)) [11] of the measured from the theoretical responses (corresponding to a zero-order release profile).

$$S(X) = \left[\sum_{k}^{n}\left(\frac{O_k(X) - E_k(X)}{SD_k}\right)^2\right]^{1/2} \quad (6)$$

where $S(X)$ is the distance function generalized by the standard deviation (SD_k) of the observed values for each response (k), X is a matrix set of the independent factors examined, $O_k(X)$ and $E_k(X)$ are the optimum value and the estimated response k.

Subsequently, the selected optimal formulations predicted by the MLR and ANN models were evaluated on the basis of the dissolution similarity factor (f_2) of the corresponding dissolution profiles with the desirable zero-order profile, according to the equation:

$$f_2 = 50\left[\left(1 + \frac{1}{n}\sum w_t(R_t - T_t)^2\right)^{-0.5} \times 100\right] \quad (7)$$

where n is the number of observations, w_t is an optional weight parameter, R_t is the percentage drug dissolved from the reference formulation, and T_t is percentage drug dissolved from the test formulation.

3. Results and discussion

3.1. In vitro drug release

The drug content of all formulations included in the experimental design varied between 98.6% and 101.1%. Results of the dissolution parameters $t_{90\%}$, Y_{2h} and Y_{8h} which were used as responses for the formulations included in the full factorial and the CCD are listed in Table 1. Also, representative dissolution profiles of the formulations that showed the highest variation are given in Fig. 1. It is seen that all tested formulations achieve sustained release with absence of burst or dose dumping effects. The variations between the examined responses were high enough to justify the application of an optimization procedure. More specifically, $t_{90\%}$ varied between 7.21 and 13.81 h, Y_{2h} varied between 10.78% and 33.95% and Y_{8h} between 56.28% and 95.08%.

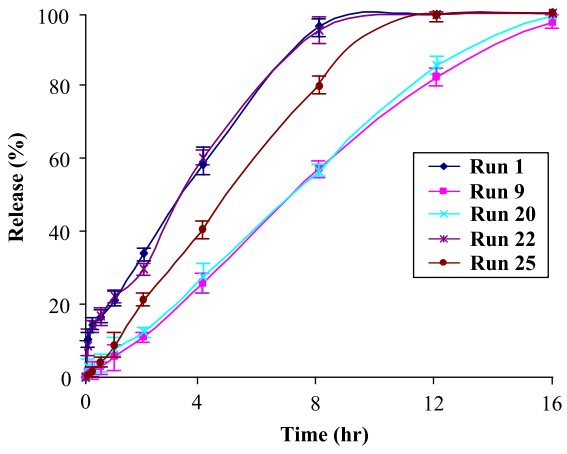

Fig. 1. Representative plots of formulations showing the highest variations of dissolution profiles (run numbers as in Table 1). (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

Table 3
Factor significance estimated by ANOVA (at 0.05 probability level) for the screening design (insignificant main effects and interactions are omitted).

Factors	$t_{90\%}$		Y_{2h}		Y_{8h}	
	F	p	F	p	F	p
X_1	83.30	<0.0001	1234.36	<0.0001	52.01	<0.0001
X_2	21.84	<0.0001	410.83	<0.0001	15.72	0.0003
X_3	1608.29	<0.0001	1915.37	<0.0001	558.58	<0.0001
X_4	–	–	729.23	<0.0001	–	–
X_1X_2	8.41	0.006	10.22	0.0028	4.19	0.0473
X_1X_3	33.57	<0.0001	39.48	<0.0001	9.59	0.0036
X_1X_4	18.16	0.0001	27.88	<0.0001	3.50	0.0688
X_2X_3	8.16	0.007	–	–	–	–
X_2X_4	3.97	0.05	–	–	–	–
X_3X_4	24.15	<0.0001	158.27	<0.0001	5.89	0.0199

Fitting of dissolution kinetic models showed that in general, the Korsmeyer–Peppas model had the best fit with R^2 = 0.991, followed by the zero-order model (R^2 = 0.987). The values of the exponent n ranged from 0.66 to 0.99, indicating that NIM is released by a combination of matrix swelling and erosion [29], in agreement with the widely accepted view that poorly soluble drugs such as NIM are expected to be released mainly by matrix erosion [30,31]. Furthermore, the high correlation coefficient of the zero-order release model along with the nearly-unity values of exponent n for some formulations, indicated that a controlled release formulation of NIM following zero-order kinetics is feasible.

3.2. Statistical experimental design

Table 3 lists the ANOVA results for the statistical significance of the examined formulation factors, according to the 2-level full factorial design. It is seen that three of the four examined factors, namely, PEG-4000 (X_1), PVP K30 (X_2) and HPMC K100 (X_3), exert a significant effect on all responses (p-level < 0.0003), while factor X_4 (HPMC E50LV) exerts significant effect only on Y_{2h} (with p-value < 0.0001). Increasing amounts of PEG-4000 (X_1) and PVP K30 (X_2) contribute to an increase in Y_{2h} and Y_{8h} and to a decrease in $t_{90\%}$. On the other hand, increasing amounts of HPMC K100M (X_3) decrease Y_{2h} and Y_{8h} and increase $t_{90\%}$. Finally, increasing amounts of HPMC E50LV (X_4) decrease Y_{2h}. The interactions of PEG-4000 with PVP-K30 (X_1X_2), HPMC K100M (X_1X_3) and HPMC E50LV (X_1X_4), as well as the interaction between HPMC K100M and E50LV (X_3X_4) have a significant effect on all examined responses (p-level 0.068, Table 3), while the interactions of PVP-K30 with HPMC K100M (X_2X_3) and E50LV (X_2X_4) have a significant synergistic effect only on $t_{90\%}$ (p-level < 0.05, Table 3).

Based on the results of the screening design none of the examined formulation factors was dropped from the CCD due to low statistical significance, since all factors are involved in significant main or interaction effects. Fitting of linear and quadratic models by MLR showed that the quadratic model has the best fit, with p-values lower than 0.0001 for all responses, and PRESS values of 28.11, 351.77, 1986.41 for $t_{90\%}$, Y_{2h} and Y_{8h} respectively. The developed regression models (after the removal of insignificant terms by a backward elimination process) are as follows:

$$t_{90\%} = 10.81 - 0.54X_1 - 0.19X_2 + 2.19X_3 - 0.24X_1^2 + 0.11X_3^2 - 0.23X_4^2 - 0.15X_1X_2 + 0.28X_1X_3 + 0.20X_1X_4 - 0.13X_2X_3 + 0.10X_2X_4 - 0.23X_3X_4$$

$$Y_{2h} = 17.40 + 3.43X_1 + 1.18X_2 - 3.17X_3 - 1.05X_4 + 0.67X_1^2 + 1.61X_2^2 + 0.29X_1X_2 - 0.57X_1X_3 - 0.48X_1X_4 + 1.15X_3X_4$$

$$Y_{8h} = 77.26 + 3.25X_1 + 1.73X_2 - 6.62X_3 - 1.03X_2^2 - 1.62X_3^2 + 1.08X_1X_2 - 1.63X_1X_3 - 0.98X_1X_4 + 1.28X_3X_4$$

Table 4
Correlation coefficient (R^2), coefficient of variation (% C.V.) and adequate precision of the MLR models fitted to the CCD data.

Response	R^2	% C.V.	Adequate precision
$t_{90\%}$	0.970	3.44	40.65
Y_{2h}	0.982	3.78	71.35
Y_{8h}	0.934	4.40	30.43

Table 4 summarizes the R^2, the % C.V. and the adequate precision (a measure of the range in predicted response relative to its associated error, or in other words a signal to noise ratio) obtained from the best-fitting MLR models.

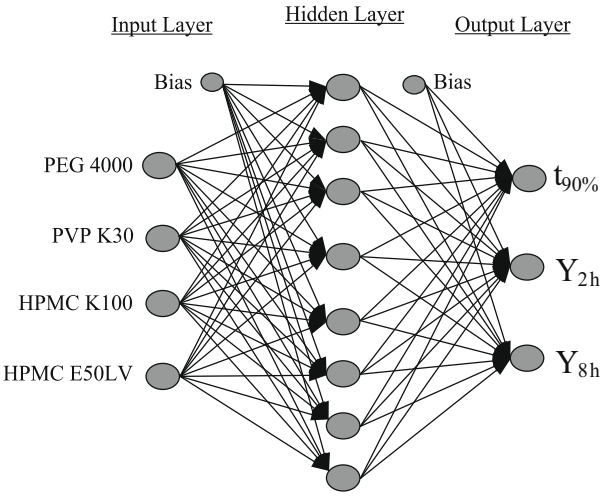

Fig. 2. Optimal neural network architecture employing PEG, PVP, HPMC K100 and HPMC E50LV as inputs and $t_{90\%}$, Y_{2h}, and Y_{8h} as output units.

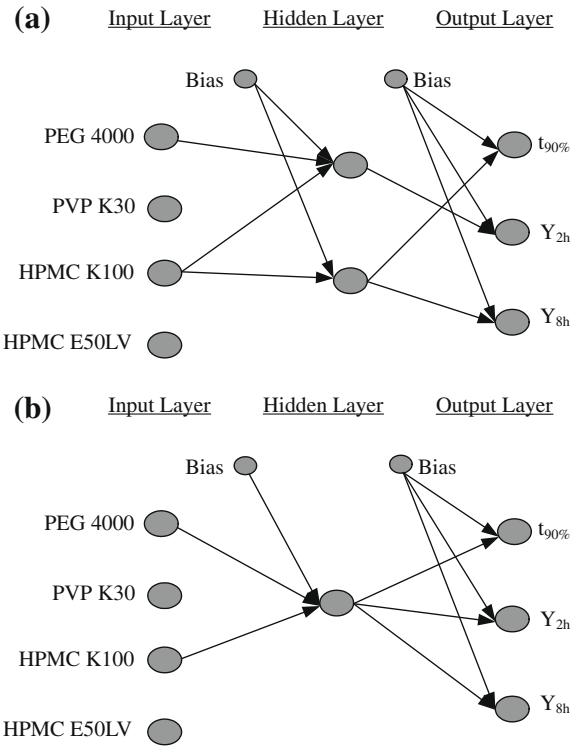

Fig. 3. Architecture of the ANN after pruning by the MBP (a) and OBD (b) algorithms.

Table 5
Causal importance (sensitivity) and predictive importance (saliency) of the examined input variables.

Input variable	Sensitivity	Saliency ($\times 10^2$)
PEG-4000	1.95	2.66
PVP K30	1.68	0.71
HPMC K100	6.82	9.60
HPMC E50LV	1.65	0.84

3.3. Development of artificial neural network models

While the complete CCD dataset was used for the MLR model fitting, in the case of the ANN, the dataset was partitioned to a training and a validation subset. In order to identify the optimum number of training cycles (iterations), a neural network having ten hidden nodes in a single hidden layer was trained using the StBack algorithm, and the mean squared error (MSE) of the validation subset was monitored. The results of validation MSE vs. the number of training cycles showed a minimum at 3000 training cycles, while the MSE of the training subset decreased asymptotically. Therefore, training was stopped at 3000 training cycles because at this point the network generalizes best, otherwise over-training occurs, and the performance of the network decreases [25].

In the next step, the optimal number of hidden layers and units was selected by trial and error, by training ANNs having 1–4 hidden layers containing 2–16 hidden units, for 3000 cycles with the StBack algorithm. Minimum validation MSE was achieved at 8 hidden units in a single layer. Hence, a network consisting of four input and three output units, with eight hidden units arranged in a single hidden layer was selected (Fig. 2).

Subsequently, the ANN was simplified applying two pruning algorithms (MBP and OBD). According to the MBP, after each training process, the link with the smallest weight is removed, while according to the OBD when pruning a certain weight, the change of the error function is approximated using a Taylor series [25], and if it is smaller than a certain threshold, the link or unit associated with that weight is permanently removed. Fig. 3 illustrates the two simplified networks suggested by the MBP (a) and OBD (b) pruning methods. The number of the hidden units was reduced to one by the OBD and two by the MBP, and the amounts of PVP K30 and HPMC E50LV were identified as insignificant from both pruning algorithms and hence excluded from the network. This is in agreement with the results of sensitivity (causal importance) and saliency (predictive importance) analysis listed in Table 5. Fig. 4 illustrates the contour plots of the pruned ANN for MBP (a) and OBD (b). In both cases, slightly curved contour lines were observed, indicating small but significant deviations from linearity in the relation between the release parameters and the retained formulation factors.

3.4. Comparison of ANN and MLR models

The performance of the ANN model was validated and compared to that of the MLR using an external test set (Table 2). Table 6 summarizes the R^2 of observed vs. predicted values, used to compare the predictive ability of the examined models. Results indicate that the original (unpruned) ANN, trained with the StBack algorithm, shows much better predictive ability than the MLR model, with R^2 of 0.955, 0.904 and 0.983 vs. 0.923, 0.873 and 0.924 of the MLR model for $t_{90\%}$, Y_{2h} and Y_{8h}, respectively. Moreover, the original StBack ANN outperforms all the rest of the ANN models (trained with Rprop and BackMom or trained with StBack and pruned ANN using the OBD and MBP algorithms). However, the close proximity of the R^2 values for the original and pruned by MBP network indicates that a good level of predictive performance can be preserved while greatly simplifying network's structure. Therefore, the original StBack, pruned by MBP neural networks and the MLR model were chosen for further evaluation.

3.5. Multi-objective optimization and experimental evaluation

The minimization of the standardized Euclidian distance (Eq. (6)) between the predicted value of each response and the optimum one (corresponding to a zero-order release profile) was selected as an appropriate global optimization tool. Composition

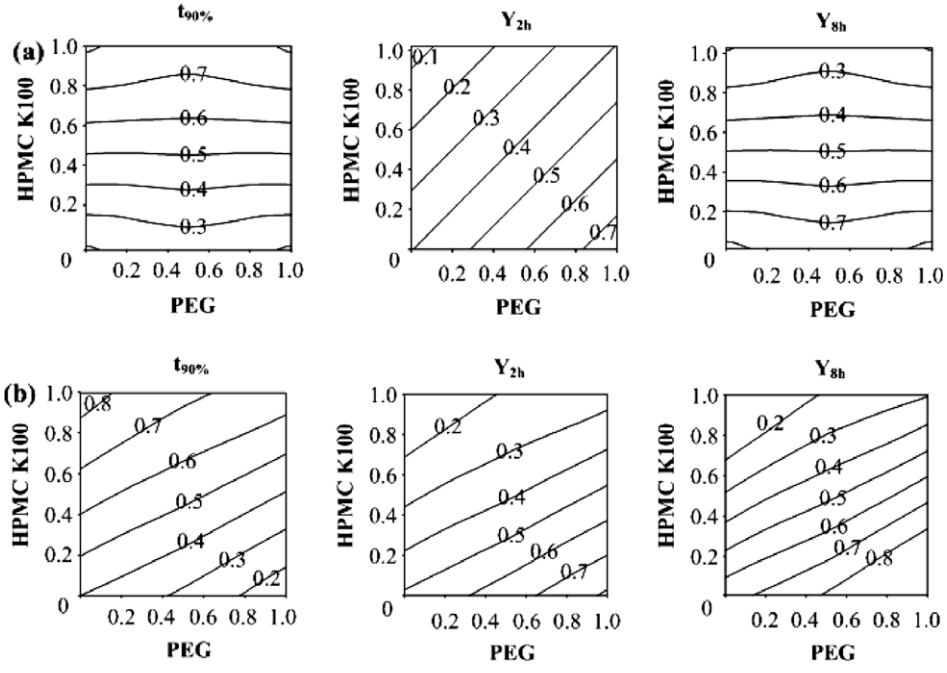

Fig. 4. Contour plots of retained formulation factors (normalized amounts of PEG and HPMC K100) vs. responses ($t_{90\%}$, Y_{2h}, and Y_{8h}) obtained using ANN models pruned by the MBP (a) and the OBD algorithms (b).

Table 6
Predictive performance (based on R^2 of observed vs. predicted values on the external validation subset) for the original ANN trained by the StBack, BackMom or Rprop algorithms, the pruned (by OBD and MBP) networks and the MLR model.

Response	R^2 achieved by model				
	StBack	BackMom	Rprop	OBD	MBP
$t_{90\%}$ (h)	0.957	0.923	0.921	0.934	0.947
Y_{2h} (%)	0.903	0.878	0.748	0.883	0.894
Y_{8h} (%)	0.981	0.938	0.927	0.899	0.971

Table 7
Composition of optimal formulations proposed according to the minimum standardized Euclidean distance for the MLR, the original StBack and the pruned ANN model.

Model	PEG (mg)	PVP (mg)	HPMC (mg)	
			K100	E50LV
MLR	30	30	37	36
ANN	40	23	38	20
ANN-MBP	30	–	40	–

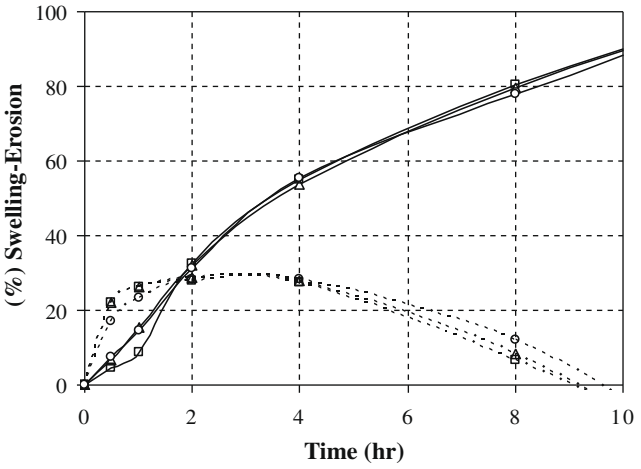

Fig. 6. Swelling (solid lines) and erosion (dashed lines) profiles of the optimum formulations proposed by the original StBack ANN (-Δ-), pruned with MBP (-○-) and the MLR (-□-) model.

results for the test formulations based on the three examined methodologies (MLR, original StBack ANN, and pruned ANN) are summarized in Table 7. It is seen that the composition proposed by the original ANN emphasizes the role of HPMC K100M (38 mg respectively), while keeping the amount of PVP at a middle level (23 mg) and HPMC E50LV and PEG close to the lowest factor levels used in the CCD (20 and 40 mg, respectively). MLR model suggested a composition with nearly equal amounts of all constituents. On the other hand, the pruned ANN, which lacks the PVP and HPMC E50LV inputs, suggests a higher amount of HPMC K100M than that of PEG (40 mg over 30 mg), outlining the predominant role of matrix swelling in determining the release kinetics of NIM. In general, both neural models used tend to emphasize the role of HPMC K100M in determining the release kinetics. The discrepancy observed between MLR and the neural models could be attributed to the non-linear nature of the release parameters–formulation factors relationship, and the inadequacy of MLR to describe non-linearities, contrary to the ANNs that are known to be excellent universal function approximators.

Fig. 5 illustrates the dissolution profiles of the proposed batches compared to the desirable zero-order profile. The similarity factor, f_2 (Eq. (7)), was used as a measure of convergence between the dissolution profiles of the obtained formulation and the desirable zero-order profile. Conventionally, values of f_2 between 50 and 100 indicate similarity between a test and a reference formulation. The lower acceptable value (i.e., 50) corresponds to 10% average absolute difference at each time-point, while values close to 100 indicate high similarity. All three models had f_2 values > 50, indicating a smaller than 10% average absolute difference between the reference and the test formulation at each time-point. The similarity factors of the original, the pruned ANN and the MLR model were 81.98, 79.46 and 76.25, respectively. The increased f_2 values of the neural networks indicated that the ANN models constitute better optimization tools in the development of a NIM-controlled release formulation with zero-order release profile, compared to conventional MLR models. This can be explained by the ANN's increased ability to approximate non-linear correlations.

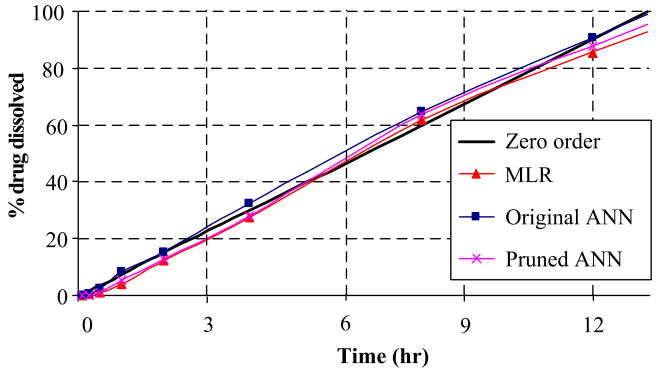

Fig. 5. Dissolution profiles of the optimal formulations proposed by the MLR and ANN models, together with the desirable zero-order profile (optimum). (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

3.6. Swelling and erosion of optimal formulations

In order to elucidate the role of the constituents of the proposed optimal formulations, the swelling and erosion of the matrix tablets was evaluated graphically (Fig. 6). From the plots, it can be concluded that the controlled release formulations undergo both swelling and erosion concomitantly, with a rather increased contribution of the erosion process. Specifically, the weight increase due to water uptake is very quickly compensated by the weight loss due to matrix erosion (after ca 3 h), and after that point, matrix erosion predominates. In such cases (when erosion is the main mechanism of release), constant release can be obtained by matrix increasing swelling. From Fig. 6, it is also evidenced that the main difference between the optimal tablet formulations proposed by the MLR and ANN models lies in the initial part of the matrix erosion phase, where a small hysteresis is seen for the formulations proposed by the MLR, which is absent from the formulations proposed by the ANN models. This means that there exists and initial acceleratory phase in matrix erosion behavior that probably results from the fast dissolution of PVP and HPMC E50LV, which are contained in increased amounts in that formulation.

4. Conclusion

The present work demonstrates the increased efficiency of the combination of neural networks and DoE, compared to the conventional MLR modeling techniques, in the systematic optimization of a NIM-controlled release tablet formulation having a zero-order profile. Furthermore, it is shown that quantitative and semi-quantitative information can be extracted about the relative importance

of the formulation factors by applying sensitivity analysis or network pruning, which additionally leads to more parsimonious ANN models without significantly impairing their generalizing ability.

References

[1] M.S. Langley, E. M Sorkin, Nimodipine a review of its pharmacodynamic and pharmacokinetic properties and therapeutic potential in cerebrovascular disease, Drugs 37 (1989) 669–699.
[2] A. Scriabine, T. Schuurman, J. Traber, Pharmacological basis for the use of nimodipine in central nervous system disorders, FASEB J. 3 (1989) 1799–1806.
[3] W. Wu, Q. Zhou, H.B. Zhang, G.D. Ma, C.D. Fu, Studies on nimodipine sustained-release tablet capable of floating on gastric fluid with prolonged gastric resident time, Acta Pharm. Sinica 32 (10) (1997) 786–790.
[4] S.J. Wie, J.P. Liu, R. Lu, W.J. Zheng, Release kinetics and in vitro-in vivo correlation of nimodipine sustained-release soft gelatin capsule, Chin. Pharm. J. 42 (2007) 852–856.
[5] S. Chopra, G.V. Patil, S.K. Motwani, Release modulating hydrophilic matrix systems of losartan potassium: optimization of formulation using statistical experimental design, Eur. J. Pharm. Biopharm. 66 (1) (2007) 73–82.
[6] M.L. Vueba, L.A.E. Batista de Carvalho, F. Veiga, J.J. Sousa, M.E. Pina, Influence of cellulose ether polymers on ketoprofen release from hydrophilic matrix tablets, Eur. J. Pharm. Biopharm. 58 (1) (2004) 51–59.
[7] B. Singh, R. Kumar, N. Ahuja, Optimization drug delivery systems using systematic "design of experiments". Part I: fundamental aspects, Crit. Rev. Ther. Drug Carrier Syst. 22 (1) (2005) 27–105.
[8] B. Singh, S.K. Chakkal, N. Ahuja, Formulation and optimization of controlled release mucoadhesive tablets of atenolol using response surface methodology, AAPS Pharm. Sci. Tech. 7 (1) (2006) E1–E10.
[9] Y.B. Huang, Y.H. Tsai, W.C. Yang, J.S. Chang, P.C. Wu, K. Takayama, Once-daily propranolol extended-release tablet dosage form: formulation design and in vitro/in vivo investigation, Eur. J. Pharm. Biopharm. 58 (3) (2004) 607–614.
[10] M. Kincl, S. Turk, F. Vrecer, Application of experimental design methodology in development and optimization of drug release method, Int. J. Pharm. 291 (1–2) (2005) 39–49.
[11] K. Takayama, M. Fujikawa, Y. Obata, M. Morishita, Neural network based optimization of drug formulation, Adv. Drug Deliv. Rev. 55 (9) (2003) 1217–1231.
[12] J. Bourquin, H. Schmidli, P. van Hoogevest, H. Leuenberger, Comparison of artificial neural networks (ANN) with classical modeling techniques using different experimental designs and data from a galenical study on a solid dosage form, Eur. J. Pharm. Sci. 6 (4) (1998) 287–301.
[13] J. Takahara, K. Takayama, T. Nagai, Multi-objective simultaneous optimization technique based on an artificial neural network in sustained release formulation, J. Control. Release 49 (1997) 11–20.
[14] K. Takayama, M. Fujikawa, T. Nagai, Artificial neural network as a novel method to optimize pharmaceutical formulations, Pharm. Res. 16 (1) (1999) 1–6.
[15] 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 (2) (2000) 175–186.
[16] T. Sando, R. Mussa, J. Sobanjo, L. Spainhour, Advantages and disadvantages of different crash modeling techniques, J. Safety Res. 36 (5) (2005) 485–487.
[17] P. Barmpalexis, F.I. Kanaze, E. Georgarakis, Developing and optimizing a validated isocratic reversed-phase high-performance liquid chromatography separation of nimodipine and impurities in tablets using experimental design methodology, J. Pharm. Biomed. Anal. 49 (2009) 1192–1202.
[18] H.Y. Karasulu, F. Taneri, E. Sanal, T. Guneri, G. Ertan, Sustained release bioadhesive effervescent ketoconazole microcapsules tabletted for vaginal delivery, J. Microencapsul. 19 (3) (2002) 357–362.
[19] L. Wang, X. Tang, A novel ketoconazole bioadhesive effervescent tablet for vaginal delivery: design in vitro and in vivo evaluation, Int. J. Pharm. 350 (1–2) (2008) 181–187.
[20] N. Kavanagh, O.I. Corrigan, Swelling and erosion properties of hydroxypropylmethyl-cellulose (Hypromellose) matrices: influence of agitation rate and dissolution medium composition, Int. J. Pharm. 279 (1–2) (2004) 141–152.
[21] L.S. Shenouda, K.A. Adams, G.J. Alcorn, M.A. Zoglio, Evaluation of a modified paddle method for dissolution testing, Drug Dev. Ind. Pharm. 12 (8–9) (1986) 1227–1239.
[22] K. Tahara, K. Yamamoto, T. Nishihata, Overall mechanism behind matrix sustained release (SR) tablets prepared with hydroxypropylmethylcellulose 2910, J. Control. Release 35 (1995) 59–66.
[23] C. Sanchez-Lafuente, S. Furlanetto, M. Fernandez-Arevalo, J. Alvarez-Fuentes, A.M. Rabasco, M.T. Faucci, S. Pinzauti, P. Mura, Didanosine extended-release matrix tablets: optimization of formulation variables using statistical experimental design, Int. J. Pharm. 237 (1–2) (2002) 107–118.
[24] S. Agatonovic-Kustrin, R. Beresford, Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research, J. Pharm. Biomed. Anal. 22 (5) (2000) 717–727.
[25] Stuttgart Neural Network Simulator (SNNS), User Manual, Version 4.2, <http://www.ra.cs.uni-tuebingen.de/downloads/SNNS/>.
[26] K. Kachrimanis, V. Karamyan, S. Malamataris, Artificial neural networks (ANNs) and modeling of powder flow, Int. J. Pharm. 250 (1) (2003) 13–23.
[27] I.V. Tetko, A.E. Villa, D.J. Livingstone, Neural network studies. 2. Variable selection, J. Chem. Inf. Comput. Sci. 36 (4) (1996) 794–803.
[28] W. Sarle, How to Measure the Importance of Inputs, SAS Institute, Cary, NC, USA, 2000, ftp://ftp.sas.com/pub/neural/importance.html.
[29] R.S. Langer, N.A. Peppas, Chemical and physical structure of polymers as carriers for controlled release of bioactive agents: a review, Rev. Macromol. Chem. Phys. 23 (1983) 61–126.
[30] J.L. Ford, M.H. Rubinstein, F. McCaul, J.E. Hogan, P.J. Edgar, Importance of drug type, tablet shape and added diluents on drug release kinetics from hydroxypropylmethylcellulose matrix tablets, Int. J. Pharm. 40 (1987) 223–234.
[31] K. Tahara, K. Yamamoto, T. Nishihata, Application of model-independent and model analysis for the investigation of effect of drug solubility on its release rate from hydroxypropyl methylcellulose sustained release tablets, Int. J. Pharm. 133 (1–2) (1996) 17–28.