



Mathematical Representation of Vasorelaxant Effect of Benzopyranes Using Box-Jenkins Model

Yashar Sarbaz^{a,*}, Hossein Babaei^{b,c}, Shahriar Gharibzadeh^a, Farzad Towhidkhah^a,
Hamed Osuli^c, Somaieh Soltani^d, Abolghasem Jouyban^{b,e}

^a*Neuromuscular Systems Lab, Biomedical Engineering Faculty, Amirkabir University of Technology, Tehran, Iran*

^b*Drug Applied Research Center, Tabriz University of Medical Sciences, Tabriz 51664, Iran*

^c*Department of Electrical Engineering, Sahand University of Technology, Tabriz, Iran*

^d*Kimia Research Institute, P.O. Box 51665-171, Tabriz, Iran*

^e*Faculty of Pharmacy, Tabriz University of Medical Sciences, Tabriz 51664, Iran*

Abstract

Drug design is generally achieved through trial and error methods, which is a time and resource consuming process and novel methods are needed to improve it. Mathematical modelling is one of the possible solutions to speed up this process. In this study, we have presented Box-Jenkins model, to predict the vasorelaxant activity (pEC₅₀) of a set of benzopyrane compounds. We used the HyperChem software to extract the molecular features of these compounds (51 molecules); then these features were used to generate the Box-Jenkins model. The dataset was divided into three groups: 37 for training, 9 for prediction and 5 for validation. The absolute relative deviation of the predicted values was 4.8% and 4.1% for validation and prediction sets, respectively. The correlation coefficient between the predicted points and the experimental values was 0.9657, revealing that the proposed model is capable of representing pEC₅₀ of benzopyranes and could be used to speed up drug discovery processes.

Keywords: Benzopyrane; Box-Jenkins; Modeling; Vasorelaxant Effect.

Received: January 10, 2007; *Accepted:* March 2, 2007.

1. Introduction

The common method for investigating new bio-active compounds is serial synthesis of chemicals and evaluating their activities, and then choosing the best compound through trial and error methods, e.g. examining the

pharmacological effects of the different compounds on animals and determining the best one. This attempt is obviously time consuming and costly.

It seems that novel methods are needed to improve drug designing. Recently many efforts made to use mathematical modelling to predict the biological activity of a compound instead of experimental tests. Quantitative structure activity relationships (QSARs) have been

*Corresponding author: Yashar Sarbaz, biomedical Engineering Faculty, Amirkabir University of Technology, Tehran, Iran.
Tel: (+98) 914-3144373.
E-mail: yashar22c@yahoo.com

developed to facilitate the designing of new drugs [1, 2]. Most QSAR models use classical least square methods and the developed models correlated observed biological activities with a number of physico-chemical properties of drug compounds. However, the accuracy of this method must be improved with sophisticated mathematical methods.

In this study, we have presented a complete transfer function structure, Box-Jenkins (BJ) model, in order to improve the prediction ability of QSAR methods. As an example, we used this method to predict the vasorelaxant activity (pEC_{50}) of a drug family, benzopyrane compounds, which have vasorelaxant activity and have been used for decreasing cardiovascular disease (CVD) in menopause women recently. Such new drugs are used because of the need for serious efforts to develop new agents with low or no deleterious effects, as possible alternatives for conventional hormone replacement therapy [3].

In this paper, we will first explain the experimental method of data extracting. Then, BJ model will be described and implemented for predicting the pEC_{50} of various bezopyrane compounds. Finally, our results will be compared with experimental findings and advantages of BJ model will be discussed.

2. Materials and methods

2.1. Experimental data

The thoracic aorta of male rats (200-250 g) were dissected free from the surrounding connective tissues and cut into rings of 3 to 5 mm in length. The rings were then bathed in physiological salt solution. The rings were mounted on stainless-steel hooks connected to a force-displacement transducer joined to a recorder system, and a computer analyzer.

All of the rings were equilibrated for 60 min. under a resting tension of 2 g and then exposed repeatedly to KCl as a vasoconstrictor until the responses stabilized; then, a control contraction was produced by a vasoactive drug such as noradrenalin. After a

sustained tension was obtained, test solutions (vasorelaxant agents or equivalent concentrations of their vehicles) were added to the bath solution. Then relaxant effect was calculated as the percentage of the initial contraction induced by vasoactive drug and expressed as $mean \pm SEM$. Finally, pEC_{50} was calculated. More details of experimental data collection were given in the reference [3].

The structure of each compound was converted to 3D using HyperChem 7.0. Then, this software was used to extract 11 physicochemical features of each compound, i.e. grid and approximate surface areas, molar volume, hydration energy, molar refractivity, polarizability, logarithm of partition coefficient, molecular weight, total energy, dipole moment and RMS gradient. These 11 features will be the inputs of our designed model.

2.2. System and model description

A system is an organized collection that responds to its stimulators, which are called inputs; the responses are output [4]. In this work, the system is smooth muscle of rat aorta and the inputs are physicochemical properties of a set of drugs that are calculated by HyperChem software. The output is the negative log of median effective concentration (pEC_{50}). Any mathematical relationship which relates inputs to the output is called a model. The model could be presented either by relating output directly to inputs with ordinary mathematical equations or by using transfer functions (TF). TF is a special mathematical approach in which the output of any instance is related not only to the input, but also to the previous outputs and inputs of the system. In some cases, using a common mathematical equation to show the model leads to inappropriate results, especially when the system's behaviour is nonlinear. In addition, some environmental factors in the laboratory might affect the process of biological assay results. These are called the effects of noise. It is difficult to simulate the effect of noise

Table 1. Numerical values of optimized orders.

i	1	2	3	4	5	6	7	8	9	10	11
$Bi(q)$	2	2	2	2	2	2	2	2	1	2	2
$Fi(q)$	0	0	0	0	0	0	0	0	1	0	0
N_i	0	0	0	0	0	0	0	1	1	1	1

with ordinary mathematical equations, but the noise effect can be modelled by transfer functions. In this paper, we used transfer function in order to generate an accurate model and to simulate the noise effects.

2.3. Model selection

There are three different methods of system identification, black box, grey box, and white box. The main factor which makes them different from each other is the level of available information. Taking this criterion into account, for black box identification there is no available information from system. For white box, all of the relations between various parts of the system are known. Gray box identification is placed between two categories [5]. In this paper, the system is considered as a black box one, since there was almost no information about the system behaviour.

The most common way of estimating black box transfer function parameters is to use the least square method (minimizing sum of squared errors) which is a simple linear structure. The justification for using linear structure for vasorelaxant activity could be

justified by the range of pEC_{50} values. All data used in this study is obtained from a drug family, benzopyranes, which possess similarity in their chemical structure. Thus, their pEC_{50} values would differ slightly from each other and can be approximated as a linear system which operates in a narrow interval around an operating point.

There are different structures in least square system identification [6]. We adopted BJ model for minimizing the effect of white noise. The BJ model is:

$$y(t) = \left(\frac{B(q)}{F(q)} \right) u(t - nk) + \left(\frac{C(q)}{D(q)} \right) e(t) \quad (1)$$

Where $y(t)$ is the output of the system (pEC_{50}), $u(t - nk)$ denotes the input of the system (i.e. physicochemical properties of benzopyranes), and $e(t)$ is the effect of white noise. $y(t)$, $u(t - nk)$ and $e(t)$ are column vectors. For above equation, $u(t - nk)$ is a compact way of writing:

$$u_i(t - nk) \quad i = 1, 2, \dots, p \quad (2)$$

p is the number of the inputs of the system.

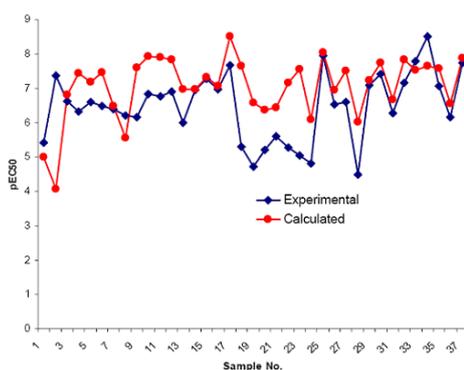


Figure 1. The fitness of experimental and calculated data for training set.

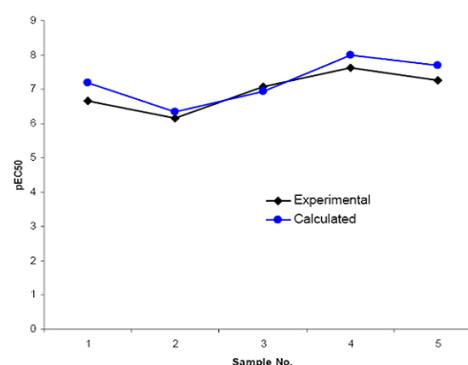


Figure 2. The fitness of experimental and calculated data for validation set.

Table 2. Numerical values of transfer functions.

i	Descriptors	Bi(q)	Fi(q)
1	Surface area (approximate)	$0.0009061 + 0.004541q^{-1}$	1
2	Surface area (grid)	$0.02863 + 0.02231q^{-1}$	1
3	Molar volume	$0.01747 + 0.03002q^{-1}$	1
4	Hydration energy	$20.52 + 47.66q^{-1}$	1
5	Molar refractivity	$0.01398 + 0.008336q^{-1}$	1
6	Polarizability	$-0.02523 - 0.008281q^{-1}$	1
7	Log P	$0.008855 - 0.01095q^{-1}$	1
8	Molecular weight	$-0.06002q^{-1}$	1
9	Total energy	$0.02067q^{-1} - 0.01242q^{-2}$	$1+0.6019q^{-1}$
10	Dipole moment	$-0.1854q^{-1} + 0.4571q^{-2}$	1
11	RMS gradient	$0.07928q^{-1} - 0.1467q^{-2}$	1

B(q), F(q), C(q) and D(q) are functions of shift operator, i.e. q^{-a} . The q operator shifts the sample α steps to past. For example:

$$q^{-1}y(t) = y(t-1) = y_{k-1} \quad (3)$$

It should be noted that the sequence of data (t) is the number of samples, rather than time; that is:

$$\text{if } y(t) = y_k \text{ then } y(t-1) = y_{k-1} \quad (4)$$

The shift operator (q) polynomials are compact ways of writing differential equations [7, 8]. For example, the B(q) polynomial in hand would be:

$$B(q) = b_{nk}q^{-nk} + b_{nk+1}q^{-nk-1} + \dots + b_{nk+nb-1}q^{-nk-nb+1} \quad (5)$$

Where nk is the shift which is determined through Ni by user. bi for $i=nk, nk+1, \dots, nk+nb-1$, are the constant coefficient. Our purpose is to calculate these constant coefficients. Needless to say, F(q), C(q) and D(q) are exactly like B(q).

We used system identification toolbox of MATLAB 7 software for generating the model.

2.4. Error criterion

The molecular features of all available benzopyrane drugs, extracted from HyperChem were divided into three groups: 37 for training the model, 5 for validation, and 9 for prediction.

In all three groups, the outputs of the model were compared with the experimental pEC₅₀s and the absolute relative deviation (%D) were

computed as an accuracy criterion using following equation:

$$\%D = 100 \left(\frac{|Calculated - Observed|}{Observed} \right) \quad (6)$$

3. Computational results

For the proposed model, two sets of parameters should be determined. 1) The coefficients of the transfer functions, i.e. B, F, C, and D in equation (1), and 2) The order of the model, i.e. the maximum power of q in equation (5). Although a system with higher orders for its TFs is more likely to fit with the real system, but its main disadvantage is that the model requires more data points in the training process.

The orders of the transfer functions are usually optimized using trial and error approach [7, 8]. These orders are summarized in Table 1.

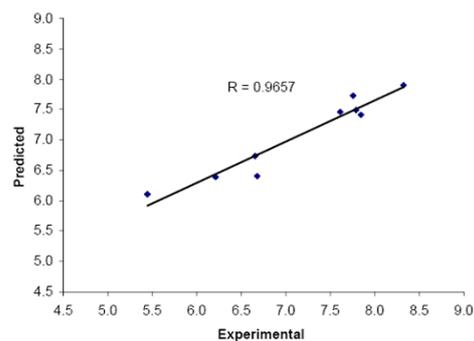


Figure 3. The correlation between experimental and prediction sets.

The numerical values of the transfer functions were calculated using MATLAB software. The parameters of optimized transfer functions are listed in Table 2. Figure 1 shows the good correlation of predicted pEC_{50s} with experimental data.

To check the validity of the trained model, features of another five benzopyrane compounds (validation set) were introduced to the trained model and the fitness of the data to the model was studied (Figure 2). The mean %D of validation set was 4.8%.

Now, the model is trained and ready to predict unmeasured output values. pEC_{50} of nine new compounds (prediction set) were predicted using the trained model. Figure 3 shows the predicted versus experimental pEC_{50} values.

The correlation coefficient of predicted versus experimental pEC_{50} values is 0.9657 which reveals excellent prediction capability of the proposed model. The mean %D was 4.1% which is an acceptable prediction error.

4. Discussion

The traditional trial and error experimental procedures for designing new drugs are time consuming and costly. Therefore, it seems that mathematical methods are needed to speed up these procedures.

The routine mathematical methods used in this regard include QSAR and related procedures. QSAR is a simple statistical method which can be used by researchers who are familiar with fundamental mathematics. However, this method is not so precise and may have greater error in predicting pEC_{50} of new drugs.

In this research, we used a classical linear mathematical method, Box- Jenkins model, which is more precise and applicable in predicting the biological activities of new drug entities.

As it is obvious in the Figure 1, the proposed model is capable of correlating the

pEC_{50} with physico-chemical properties of the drugs. In Figure 2, the fitness of experimental and calculated data for validation set shows that the model is valid.

Our model error in predicting pEC_{50} was less than 5%, which is an acceptable error range in drug designing studies. Using more inputs in this model may lessen the error. Our results indicated that the designed model may be used to design new compounds. The researchers in pharmacological fields may use this method as a preliminary route to limit the number of compounds which must be experimentally evaluated.

We showed the effectiveness of our model in predicting the characteristics of benzopyranes. This is just a typical example for multiple cases of drug designing in which using BJ can be advantageous.

References

- [1] Gonzalez MP, Teran C, Teijeira M, Helguera AM. Quantitative structure activity relationships as useful tools for the design of new adenosine receptor ligands. 1. Agonist. *Curr Med Chem* 2006; 13: 2253-66.
- [2] Kim RB. Transporters and drug discovery: Why, when, and how? *Mol Pharm* 2006; 3: 26-32.
- [3] Uhrig U, Holtje HD, Mannhold R, Weber H, Lemoine H. Molecular modelling and QSAR studies on K(ATP) channel openers of the benzopyran type. *J Mol Graph Model* 2002; 21: 37-45.
- [4] Oppenheim J, Willsky AS. *Signals and systems*. New Jersey: Prentice Hall, 1985.
- [5] Sjoberg J, Zhang Q, Ljung L, Benveniste A, Delyon B, Glorennec P, Hjalmarsson H, Juditsky A. Nonlinear black-box modelling in system identification: A unified overview. *Automatica* 1995; 31: 1691-724.
- [6] Ljung L, Wahlberg B. Asymptotic properties of least-squares method for estimating transfer function and disturbance spectra. *Adv Appl Prob* 1992; 24: 412-40.
- [7] Ljung L. *System identification theory for the user*. London: Prentice Hall International, 1999.
- [8] Soderstrom T, Stoica P. *System identification*. London: Prentice Hall International, 1989.