역산문제 방법을 적용한 제네릭 의약품 개발 프로세스의 강건 설계

Nguyen Khoa Viet Truong* · 신상문*† · 정성훈**

* 인제대학교 시스템경영공학과
** 부산대학교 약학대학

Integrating inverse problem to robust design for a generic drug development process

Nguyen Khoa Viet Truong* · Sangmun Shin*† · Seong Hoon Jeong**

* Department of Systems Management & Engineering, Inje University
** College of Pharmacy, Pusan National University, 30 Geumjeong-gu, Busan 609-735, South Korea

Key Words : Inverse Problem, Robust Design, Response Surface Methodology, Pharmaceutical Formulation, Mixture Experiment

Abstract

Robust design (RD) has emerged as a key feature in process design and development for more than twenty years. Many researchers and industrial engineers around the world have invested their intensive efforts to develop and apply RD in many fields in order to improve quality of output products. However, there is also room for improvement. The primary objective of this research is to determine "robust formulation" of a medicine by checking its gelation index. In order to achieve this target, based on the nature of problem, at first, a customized experimental format is designed for obtaining data. Second, time-depended responses based models are developed by the proposed inverse problem (IP) methodology. Third, an RD model based on mean square error (MSE) concept is introduced for time-depended responses. Finally, the proposed approach is illustrated by a case study while comparing obtained results to the response surface methodology (RSM) approach.

1. Introduction

Even though our global economy is in bad situation recent years, health care industry still is one of the world's largest and fastest-growing industries. In such competition environment, pharmaceutical manufacturers are making much effort to keep and increase their market shares. Besides, under the pressure of strict regulations for new drugs development processes as well as the tight control of the Food and Drug Administration (FDA) (Peterson et al. 2009), quality improvement becomes one of key issues of researchers in the pharmaceutical field. In order to raise output product quality, researchers must carefully consider all possible problems from design stage. This is also mentioned as quality by design in which robust design (RD) concept has played important role. In pharmaceutical studies, robust design can be im-

† 교신저자 sshin@inje.ac.kr

※ This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (20100104).

plemented throughout following main steps: (1) Det ermine control factors and interested output responses; (2) Design experiments to collect data; (3) Develop empirical models to represent relationships between inputs and outputs (for both mean and variance responses) and then (4) Optimize selected RD model to obtain optimal solutions.

The situation in which this paper confronted is to determine "robust formulation" of one kind of drug by investigating its time-depended gelation index. In RD literature of model building, the associated parameters of empirical models have to be determined by a number of estimation methodologies, such as least squares, maximum likelihood, and Bayesian estimations. Even though RD along with response surface methodology (RSM) provides significant advantages in mathematical foundation and practicability, there is room for impro vement. Based on the least squares method, RSM requires several basic assumptions, such that all errors of observations must be mutual independent, normally distributed with the constant variances and zero expectations (Myers and Montgomery, 2002). Another aspect, it may not reasonable to assume that all observed data have the same variance and zero expectation in RD context because the variance values must be different at all mean values of output characteristics based on RD philosophy. On the other hand, responses measured from phar ma ceutical study in this paper are time- depended rather than fixed values. To handle this situation, inverse problem (IP) can be a potential alternative because of its advantages. First, IP can relax basic assumptions on that RSM must base. Second, for variance model, IP can overcome the weak point as discussed. Finally, IP approach considers both data and model parameters as random variables so that it can exploit more information from observed data than conventional RSM approach. From this point of view, the primary motivations of this paper can be identified as follows:

- A customized experimental format for a phar maceutical experimental design is identified.
- Estimation methods to handle time-depended responses are also developed by using both RSM and IP.
- Based on the mean square error (MSE) con cept, a new inverse problem based robust de sign (IPRD) model for time-depended responses is proposed.
- A case study on a drug development is per formed for verification purpose, and a compar ative study for RSM and IP based RD methods is also conducted.

2. Experimental framework

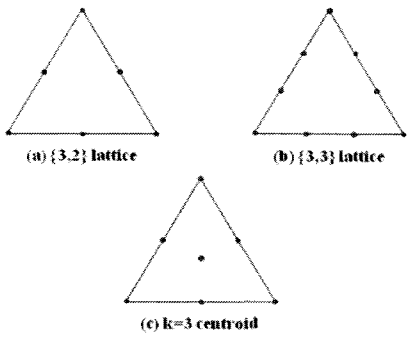
In researches, data are commonly collected through implementing experiments. The quality of obtained information is based on experiment's design (Montgomery, D.C.(2001). Determining a suitable formulation of drug is fundamental in pharmaceutical researches. So, it is required to handle mixture problems in which the proportion of ingredients within the mixture, rather than their amounts must be considered. In such problems, ingredients of formulations or control factors are inherently dependent upon others. Therefore, experimental design used to be implemented in common optimization studies cannot be utilized. Instead, a special kind of designed experiment referred as a mixture experiment is used. In the mixture experiment, factors are ingredients of a mixture and the quality characteristics or responses are based on the proportionality of those ingredients. In the early stages of drug development, mixture experiments are of great interest to pharmaceutical companies and experts because they can provide reliable tools for optimizing and accelerating the move ment of a drug from the R&D stages towards its introduction into the marketplace.

Assume that there are k ingredients x1,x2,⋯,xk required to form an target formulation. It is required that the proportions of k ingredients in the mixture must be under the constraint as $\sum_{i=1}^{k} x_i = 1$,

where $x_i \geq 0, i = 1, 2, ..., k$. In another word, the degree of freedom for mixture will be k-1. Ther efore, the measured responses depend only on the relative proportions of the ingredients or components in the mixture, not on the amount of the mix ture. The mixture experiments base on the same assumptions as those are made for factorial experiments in which the errors are independent and identical distributed with zero mean and common variance.

In the situation that there is no additional constraint, standard mixture designs are available for fitting standard models, such as Simplex-Lattice designs and Simplex-Centroid designs. When mixture components are subject to additional constraints, such as a maximum and/or minimum value for each component, designs other than the standard mixture designs, referred to as constrained mixture designs or Extreme-Vertices designs, are appropriate. In the Figure 1, visual comparison of simplex-lattice designs for (a) k = 3 m= 2, (b) k = 3 m= 3, (c) and simplex-centroid design for k = 3 are displayed below. Further information on mixture experiments and the simplex-centroid designs can be referred to Scheffé (1958, 1963).

As discussed, data are often collected overtime naturally in pharmaceutical studies in order to investigate the effect of interested drug properties on human being. In such studies, responses are collected as time series data at all runs in designed experiments instead of static responses at a given time. At a given time ti, data are collected as shown in Table 1, the well-known designed experimental format of n runs appeared frequently in literature. With the vector of control factors x, a suitable mixture designed experimental format is selected while considering capability of research group. In order to utilize robust design concept, n runs of experiments are performed with r replicati ons at each run. Then, the vector of mean $\overline{y_i}$ and vector of variance s_i^2 can be calculated.



<Figure 1> Simplex-lattice design

The response y_{uv} denotes the v^{th} observation in r replications at the u^{th} experimental run, $\overline{y_{iu}}$ and s_{iu}^2 represents the uth element in vector of mean responses and vector of variance responses, respec tively. When responses are obtained overtime from t_1 to t_m, the above experimental format can be extended as shown in below Table 2 in which original data are not showed.

<Table 1> Experimental format for a given time ti

Runs	Input factors**x**	y_i (Replications)					$\overline{y_i}$	si2
1	Mixture design	y_{11}	y_{12}	⋯	y_{1v}	⋯	y_{1r}	$\overline{y}_{i1}$	s_{i1}^2
2		y_{21}	y_{22}	⋯	y_{2v}	⋯	y_{2r}	$\overline{y}_{i2}$	s_{i2}^2
⋮		⋮	⋮	⋱	⋮	⋮	⋮	⋮	⋮
u		y_{u1}	y_{u2}	⋯	y_{uv}	⋯	y_{ur}	$\overline{y}_{iu}$	s_{iu}^2
⋮		⋮	⋮	⋯	⋮	⋱	⋮	⋮	⋮
n		y_{n1}	y_{n2}	⋯	y_{nv}	⋯	y_{nr}	$\overline{y}_{in}$	s_{in}^2

<Table 2> Proposed experimental format for time-oriented responses

Runs	Input factors x	t1		t2		···		ti		···		tm	
		$\bar{y}_1$	s12	$\bar{y}_2$	s22	···	···	$\bar{y}_i$	si2	···	···	$\bar{y}_m$	sm2
1		$\bar{y}_{11}$	s112	$\bar{y}_{21}$	s212	···	···	$\bar{y}_{i1}$	si12	···	···	$\bar{y}_{m1}$	sm12
2		$\bar{y}_{12}$	s122	$\bar{y}_{22}$	s222	···	···	$\bar{y}_{i2}$	si22	···	···	$\bar{y}_{m2}$	sm22
···	Mixture design	···	···	···	···	···	···	···	···	···	···	···	···
u		$\bar{y}_{1u}$	s1u2	$\bar{y}_{2u}$	s2u2	···	···	$\bar{y}_{iu}$	siu2	···	···	$\bar{y}_{mu}$	smu2
···		···	···	···	···	···	···	···	···	···	···	···	···
n		$\bar{y}_{1n}$	s1n2	$\bar{y}_{2n}$	s2n2	···	···	$\bar{y}_{in}$	sin2	···	···	$\bar{y}_{mn}$	smn2
	Targets	Tt1		Tt2				Tti				Ttn	

From time t_1 to t_m, the responses are collected as the format shown in Table 1, the replications are not displayed in this table. At the given time t_i, $\overline{y_{iu}}$ denotes the u^{th} element in vector of mean responses and s_{iu}^2represents u^{th} element in vector of variance responses, T_{ti} represents corresponding target value for output quality.

3. Response surface methodology (RSM)and inverse problem (IP) based robust design

3.1 Response surface methodology based model development for time-depended responses

This section is summarized results of model development for time-depended responses by using RSM approach proposed by Park et al. (2010) and Shin et al. (2011). From available data in above Table 2, mean and variance models can be built for RD procedure. For developing model of mean function, let Y be the mean response matrix obtained from Table 2. Vertically, mean response matrix Y can be considered as Y = [Yc1 Yc2 ···Yci ··· Ycm] where Yci denotes mean column vector of $\bar{y}_i$.

Considering the whole mean response matrix Y, the general vertical form of the relationship between Y and x can be expressed as

$$\hat{\mathbf{Y}}(\mathbf{x}) = \breve{\mathbf{x}}^T \times \mathbf{M}_{\mu c} \tag{1}$$

where $\mathbf{M}_{\mu c}$ denotes a matrix of parameters in vertical analysis given as follows:

$$\mathbf{M}_{\mu c} = \left[\hat{\boldsymbol{\beta}}_{\mu 1} \quad \hat{\boldsymbol{\beta}}_{\mu 2} \quad \cdots \quad \hat{\boldsymbol{\beta}}_{\mu m}\right] = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y} \tag{2}$$

By analyzing vertically, the relationship between mean response matrix Y and input control factors x can be expressed. In horizontal direction, the relationship between mean response matrix Y and time t can be proposed. Horizontally, the mean response matrix Y can be considered as Y = [Yr1 Yr2 ···Yri ··· Yrm] where Yru represents uth row vector of the mean response matrix in Table 2. Let $\boldsymbol{t}^T$ represents a polynomial form of time t to express the empirical relationship between row responses over time and T denotes as the designed matrix of t, by using least squares method for the whole matrix Y:

$$\hat{\mathbf{Y}}(t) = \mathbf{M}_{\mu r} \times \mathbf{t}^T \tag{3}$$

where $\mathbf{M}_{\mu r} = [\gamma_{\mu 1}, \gamma_{\mu 2}, ..., \gamma_{\mu n}]$ is the transposed matrix of parameters represents relationship be-

tween responses and time.

From the results of Equations (1), (2) and (3), it is reasonable to integrate two separated relationships into following general Equation:

$$\hat{\mathbf{Y}}(\mathbf{x},t)=\breve{\mathbf{x}}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{M}_{\mu r}\mathbf{t}^T \quad (4)$$

The Equation (4) describes mean response matrix Y as a function of control factors x and time t. In similar way, let s2 be the variance response matrix obtained from Table 2 for developing model of variance function. Hence, the empirical relationship between variance response matrix and input control factors over time can be obtained as

$$\hat{\mathbf{S}}^2(\mathbf{x},t)=\breve{\mathbf{x}}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{M}_{\sigma r}\mathbf{t}^T \quad (5)$$

where $\mathbf{M}_{\sigma\mathrm{T}}$ is the transposed matrix of parameters for horizontal analysis with variance response matrix.

The developed Equations (4) and (5) can describe the relationships between mean, variance responses and control factors at any given time that are necessary for handling the interested RD problem which based on time-depended responses.

3.2 Inverse problem based model development for time-depended responses

By using IP, mean and variance should be estimated in different way because they are based on different basic assumptions. Considering the responses are sampled from populations those are of normal distributions, mean values will belong to normal distributions and variance values will belong to Chi-square distributions.

For mean model, assume that each row vector Yru can be a function of t as follows:

$$\hat{\mu}_{ruIP}(t)=\boldsymbol{\gamma}_{\mu uIP}\times\mathbf{t}^T \quad (6)$$

where tT represents a polynomial form of time t to explain the empirical relationship between responses and time. The information of variances should be identified before estimating model parameters by using IP approach. Because of the nature of the observations, the repetitions of responses will provide the variance information for analyzing relationship between responses and control factors, not time. Hence, the relationship between responses and time can be estimated based on the assumption that there is no variance information and well-known normal equation can be utilized to estimate the model parameter vector

$$\boldsymbol{\gamma}_{\mu uIP}=Y_{ru}\mathbf{T}\left(\mathbf{T}^T\mathbf{T}\right)^{-1} \quad (7)$$

Extending this result to the whole matrix, the horizontal form of the relationship between mean response matrix Y and time can be derived as follows:

$$\hat{\mathbf{Y}}_{IP}(t)=\boldsymbol{\Gamma}_{\mu IP}\times\mathbf{t}^T \quad (8)$$

where $\Gamma_{\mu IP}=[\Gamma_{\mu 1IP};\Gamma_{\mu 2IP};\cdots;\Gamma_{\mu nIP}]$ is the matrix of parameters for horizontal analysis. Embedding Equation (7) on $\Gamma_{\mu IP}$, the Equation (8) will become:

$$\hat{\mathbf{Y}}_{IP}(t)=\mathbf{YT}\left(\mathbf{T}^T\mathbf{T}\right)^{-1}\times\mathbf{t}^T \quad (9)$$

In order to analysis the relationship between responses and control factors, mean response matrix Y can be analyzed vertically as Y = [Yc1 Yc2 ···Yci ··· Ycm] where Yci denote mean column vectors of $\overline{y}_i$. Individually, each column vector Yci can be a function of control factors as

$$\hat{Y}_{ciIP}(\mathbf{x})=\breve{\mathbf{x}}^T\hat{\boldsymbol{\beta}}_{\mu iIP} \quad (10)$$

where $\hat{\beta}_{\mu iIP}$ is the model parameter vectors that can be estimated by using IP approach

$$\hat{\boldsymbol{\beta}}_{\mu iIP}=\left(\mathbf{X}^T\mathbf{C}_{\mathbf{Di}}^{-1}\mathbf{X}\right)^{-1}\mathbf{C}_{\mathbf{Di}}^{-1}\mathbf{X}^T Y_{ci} \quad (11)$$

Generally, consider the whole mean response matrix Y, the vertical form of the relationship between Y and control factors x can be expressed as

$$\hat{\mathbf{Y}}_{IP}(\mathbf{x}) = \breve{\mathbf{x}}^T \times \mathbf{B}_{\mu cIP} \tag{12}$$

where $\hat{\beta}_{\mu iIP}$ denotes a matrix of parameters in vertical analysis given as follows:

$$\mathbf{B}_{\mu cIP} = \left[\hat{\boldsymbol{\beta}}_{\mu 1IP} \quad \hat{\boldsymbol{\beta}}_{\mu 2IP} \quad \cdots \quad \hat{\boldsymbol{\beta}}_{\mu mIP}\right] \tag{13}$$

From the results of Equations (9)-(13), the functional form of empirical relationship between mean response matrix Y and control factors x over time t using IP approach as can be derived as

$$\hat{\mathbf{Y}}_{IP}(\mathbf{x},t) = \breve{\mathbf{x}}^T \mathbf{B}_{\mu cIP} \mathbf{T}\left(\mathbf{T}^T\mathbf{T}\right)^{-1} \mathbf{t}^T \tag{14}$$

In the case of variance, its empirical model can be developed in similar way of mean model but with different basic assumptions. Let V be the variance response matrix obtained from Table 2. In the horizontal direction approach, the variance response matrix can be considered horizontally as V = [Vr1 Vr2 ···Vru ··· Vrn] where Vru represents uth row vector of the variance response matrix. Similar to mean model, the horizontal form of the relationship between variance response matrix V and time can be expressed as follows:

$$\hat{\mathbf{V}}_{IP}(t) = \mathbf{V}\mathbf{T}\left(\mathbf{T}^T\mathbf{T}\right)^{-1} \times \mathbf{t}^T \tag{15}$$

In vertical direction approach, variance response matrix V can be analyzed as V = [Vc1 Vc2 ···Vcu ··· Vcn] where Vci denote variance column vectors of si2. Individually, each column vector Vci can be a function of control factors as

$$\hat{V}_{ciIP}(\mathbf{x}) = \breve{\mathbf{x}}^T \hat{\boldsymbol{\beta}}_{\sigma iIP} \tag{16}$$

where $\hat{\beta}_{\mu iIP}$ is the model parameter vectors that can be estimated based on the assumption that observed variances belong to Chi-square distributions (Walpole et al., 2007)

$$\hat{\boldsymbol{\beta}}_{\sigma iIP} = \left(\mathbf{K}_i^T \mathbf{K}_i\right)^{-1} \mathbf{K}_i^T \mathbf{F} \tag{17}$$

where Ki is the modified form of designed matrix X for ith vector of observed variances si2 and F is the vector of degree of freedom as addressed in following Equation (18), (19).

$$\mathbf{K} = \left[\frac{(r-1)\mathbf{X}_{r1}}{s_1^2} \quad \frac{(r-1)\mathbf{X}_{r2}}{s_2^2} \quad \cdots \quad \frac{(r-1)\mathbf{X}_{rn}}{s_n^2}\right]^T \tag{18}$$

$$\mathbf{F} = \left[r-1 \quad r-1 \quad \cdots \quad r-1\right]^T \tag{19}$$

Generally, consider the whole variance response matrix V, the relationship between variance responses and control factors x in vertical analysis can be derived as

$$\hat{\mathbf{V}}_{IP}(\mathbf{x}) = \breve{\mathbf{x}}^T \times \mathbf{B}_{\sigma cIP} \tag{20}$$

where $\mathbf{B}_{\mu c}$ denotes a matrix of parameters in vertical analysis given as follows:

$$\mathbf{B}_{\sigma cIP} = \left[\hat{\boldsymbol{\beta}}_{\sigma 1IP} \quad \hat{\boldsymbol{\beta}}_{\sigma 2IP} \quad \cdots \quad \hat{\boldsymbol{\beta}}_{\sigma mIP}\right] \tag{21}$$

From the results of Equations (15)-(21), the functional form of empirical relationship between mean response matrix V and control factors x over time t as can be derived as

$$\hat{\mathbf{V}}_{IP}(\mathbf{x},t) = \breve{\mathbf{x}}^T \mathbf{B}_{\sigma cIP} \mathbf{T}\left(\mathbf{T}^T\mathbf{T}\right)^{-1} \mathbf{t}^T \tag{22}$$

3.3 Proposed robust design models

Once time-depended responses based mean and variance empirical models are developed, a suitable RD model can be selected for optimizing in order to obtain optimal settings of control factors. Considering the situation of pharmaceutical study needed to handle in this study, the RD based on MSE concepts can be selected. Let Tti be the target response at the time ti. The proposed time-depended response surface methodology based robust design (RSMRD) and inverse problem based robust design (IPRD) optimization model based on MSE concepts can be formulated as follows:

RSMRD model:

$$\text{Minimize} \quad \sum_{i=1}^{m}\left[\left(\hat{\mathbf{Y}}(\mathbf{x},t_i)-T_{t_i}\right)^2+\hat{\mathbf{S}}^2(\mathbf{x},t_i)\right] \qquad (23)$$

Subject to: $\mathbf{x}\in\Omega$

IPRD model:

$$\text{Minimize} \quad \sum_{i=1}^{m}\left[\left(\hat{\mathbf{Y}}_{IP}(\mathbf{x},t_i)-T_{ti}\right)^2+\hat{\mathbf{V}}_{IP}(\mathbf{x},t_i)\right] \qquad (24)$$

Subject to: $\mathbf{x}\in\Omega$

in which Ω represents the feasible region of control factors.

Comparison of two RD procedures can be summarized as following steps:

- In order to conduct the RSM approach, mean and variance models can be estimated by Equations (4) and (5) where the matrix of parameters can be obtained by using Equation (2). For variance model, the matrix of variance responses S is used instead of Y.

- To the proposed IP approach, mean model can be estimated by Equation (14) similar to the RSM approach. However, for variance model, the procedure of estimating parameters is different from the RSM approach as identified in Equations (17) to (21).

- Once mean and variance models are estimated by using both RSM and IP approaches, optimizations are performed by utilizing RSMRD and IPRD models as shown in Equations (23) and (24). Finally, the optimal solutions can be obtained as demonstrated in Tables 6 and 7.

4. Case study

The aim of this pharmaceutical study is to determine mixture of a tablet so that it can give optimal solution. The formulations of each tablet are shown in designed experimental format as Table 3. The total weight of each tablet was 234.31mg. Gelation index might be a useful tool to represent the portion of a tablet that has undergone gelation on time.

$$\textit{Gelation Index}\,(G,\%) = \left\{1-\frac{(D_{obs})^3}{(D_{ini})^3}\right\}\times 100$$

Dobs : diameter of the portion not gelled after the test

Dini : diameter of the tablet before the test

The mixture experiment is designed with the constraint of that x1 is not less than 40 percent and simple-lattice design is chosen for ten ingredients. Gelation indexes are measured overtime from 0.5 hour to 5 hours. At any given time, responses are replicated 4 times at all runs so that mean and variance values can be calculated. Data profile and their target values are displayed in following Table 4 and Figure 2.

Once data are available, mean and variance models for RD utilization can be developed using results of Equations (4), (5) for RSM based model and (14), (22) for IP based model. Based on the curves of data profile in Figure 2, second order models of time are chosen for both vertical and horizontal analysis. For examining the fitted models to ensure that they provide an adequate approximation to the true system, following tables show R-square value of separated regression models for 8 columns and 21 rows in data profile at first.

Also, after checking the normal probability plot overtime, mean absolute error criterion is used to show the adequacy of models.

$$MAE = \frac{1}{runs\times tn}\sum_{j=1}^{runs}\sum_{i=1}^{tn}\left|\overline{y}_{ij}-\hat{y}_{ij}\right| \qquad (23)$$

where runs, tn are respectively the number of experimental runs and the number of given time, in this case study, runs = 21 and tn = 8. For RSM model, mean absolute error MAE_RSM = 4.9260 and for IP model, MAE_RSM = 4.7401. Then, for evaluating model predictability, results of RSMRD and IPRD model are compared in term of MSE (mean squared error) defined as below:

$$MSE_{RSM} = \frac{1}{tn}\sum_{i=1}^{tn}\left(y_i^{RSM\_opt} - T_i\right)^2 \quad (24)$$

$$MSE_{IP} = \frac{1}{tn}\sum_{i=1}^{tn}\left(y_i^{IP\_opt} - T_i\right)^2 \quad (25)$$

For RSM model, mean absolute error MSE_RSM = 1.821 and for IP model, MSE_IP = 1.766.

Then, RD models as discussed in Equations (23), (24) can be used in order to obtain optimal setti ngs. The solutions of two approaches can be compared in below Table 6.

Based on these solutions, optimal time-depended responses can be estimated. For comparative purpose, Table 7 is showed in which absolute biases and bias percents are also calculated as below. For illustration, the obtained data are plotted in Figures 3-5.

$$Absolute\ bias_i = \left|Optimal\ drug\ release\ rate_i - Target_i\right|$$

$$\%bias_i = \frac{Absolute\ bias_i}{Target_i} \times 100\%$$

<Table 3> Experimental format for gelation study

Std order	Run order	Input factors									
		x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
8	1	93.71	0	0	0	0	0	0	140.6	0	0
19	2	100.74	7.03	7.03	7.03	7.03	7.03	7.03	77.33	7.03	7.03
3	3	93.71	0	140.6	0	0	0	0	0	0	0
15	4	100.74	7.03	7.03	77.33	7.03	7.03	7.03	7.03	7.03	7.03
6	5	93.71	0	0	0	0	140.6	0	0	0	0
10	6	93.71	0	0	0	0	0	0	0	0	140.6
17	7	100.74	7.03	7.03	7.03	7.03	77.33	7.03	7.03	7.03	7.03
14	8	100.74	7.03	77.33	7.03	7.03	7.03	7.03	7.03	7.03	7.03
11	9	107.77	14.06	14.06	14.06	14.06	14.06	14.06	14.06	14.06	14.06
16	10	100.74	7.03	7.03	7.03	77.33	7.03	7.03	7.03	7.03	7.03
18	11	100.74	7.03	7.03	7.03	7.03	7.03	77.33	7.03	7.03	7.03
4	12	93.71	0	0	140.6	0	0	0	0	0	0
20	13	100.74	7.03	7.03	7.03	7.03	7.03	7.03	7.03	77.33	7.03
9	14	93.71	0	0	0	0	0	0	0	140.6	0
12	15	171.04	7.03	7.03	7.03	7.03	7.03	7.03	7.03	7.03	7.03
2	16	93.71	140.6	0	0	0	0	0	0	0	0
13	17	100.74	77.33	7.03	7.03	7.03	7.03	7.03	7.03	7.03	7.03
1	18	234.31	0	0	0	0	0	0	0	0	0
7	19	93.71	0	0	0	0	0	140.6	0	0	0
21	20	100.74	7.03	7.03	7.03	7.03	7.03	7.03	7.03	7.03	77.33
5	21	93.71	0	0	0	140.6	0	0	0	0	0

<Table 4> Obtained data for the gelation study based on the designed experimental format with target values

Runs	0.5h		1h		1.5h		2h		2.5h		3h		4h		5h	
	$\bar{y}$	s12	$\bar{y}$	s22	$\bar{y}$	s32	$\bar{y}$	s42	$\bar{y}$	s52	$\bar{y}$	s62	$\bar{y}$	s72	$\bar{y}$	s82
1	49.61	2.67	62.06	1.14	77.19	1.19	84.62	1.14	87.72	1.13	91.64	1.28	94.73	0.53	96.09	0.11
2	48.46	1.86	66.35	0.62	77.24	1.67	82.12	0.57	86.14	1.64	87.40	1.45	93.57	0.63	94.97	1.09
3	14.56	1.17	22.00	4.24	27.63	4.57	31.32	4.78	36.50	1.81	37.25	0.83	54.48	1.86	64.74	1.36
4	33.41	3.53	19.33	2.59	49.04	4.13	62.12	4.26	65.84	2.93	60.27	2.02	59.69	1.75	59.85	2.44
5	33.50	2.38	53.13	1.80	61.85	0.67	70.57	1.45	74.24	1.62	83.18	0.67	89.47	1.18	95.40	0.48
6	41.13	1.34	56.48	1.23	66.64	2.08	77.61	0.41	81.04	0.82	85.02	1.04	91.67	0.83	95.67	0.68
7	40.83	1.98	56.56	2.04	69.34	1.51	78.12	1.84	81.09	0.90	87.71	1.01	94.87	0.81	97.94	1.03
8	10.04	2.91	8.90	4.89	41.43	0.97	40.94	4.52	46.97	5.86	56.43	3.01	59.99	3.47	68.16	3.98
9	37.11	2.22	55.18	0.64	63.17	7.66	77.97	1.86	82.88	1.35	90.11	1.76	96.48	0.40	98.61	0.25
10	37.45	3.48	54.34	3.20	68.03	2.83	74.78	2.86	83.35	1.53	86.90	0.40	94.80	0.58	96.24	0.88
11	34.05	3.16	47.94	3.24	58.73	1.52	64.07	1.15	73.24	0.11	77.73	0.47	85.64	1.31	92.99	0.96
12	22.82	2.18	35.46	1.46	50.78	1.92	61.46	2.38	70.03	1.01	79.11	2.75	86.35	1.29	97.16	0.25
13	34.74	3.18	45.80	1.81	58.03	1.82	66.09	1.49	75.26	1.69	83.93	0.78	90.95	0.86	96.43	0.74
14	22.82	2.01	37.16	2.77	48.81	2.86	53.43	1.95	65.32	3.58	75.69	1.50	82.78	2.60	91.85	1.59
15	41.84	0.6	58.87	0.9	66.38	0.6	76.01	0.8	81.03	0.7	84.58	1.1	92.78	0.4	96.90	0.3
16	14.99	2.47	25.57	3.12	40.31	1.58	47.40	0.72	55.51	1.15	63.77	1.65	67.70	0.61	80.32	2.07
17	35.96	0.82	48.02	2.68	55.70	3.33	70.98	2.55	78.17	1.94	83.29	1.19	93.38	1.18	94.85	0.44
18	37.42	1.79	52.18	0.90	59.83	1.31	68.09	0.64	75.22	0.57	81.01	0.67	87.58	1.11	93.10	0.46
19	43.33	1.40	58.05	2.75	65.47	2.99	71.65	1.92	81.72	2.39	88.40	1.24	91.15	1.77	92.73	2.48
20	42.28	2.09	56.91	0.73	64.41	1.26	72.38	1.51	81.62	1.22	84.53	1.02	91.53	0.36	97.11	0.79
21	37.13	1.00	56.06	2.80	64.24	4.56	76.47	2.17	81.72	0.50	87.04	1.04	95.17	0.96	98.55	0.47
Targets	37.75		47.61		56.71		65.54		69.32		77.55		88.42		88.81	

<Table 5> R-square of models for both analysis directions

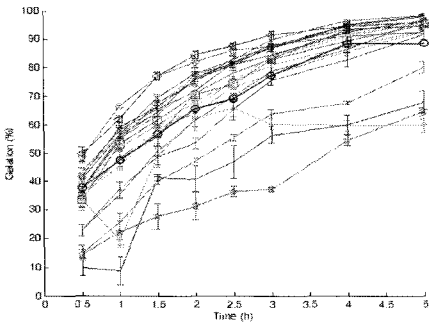
R-square for vertical analysis			R-square for horizontal analysis		
Columns	RSM	IP	Rows	RSM	IP
1	0.78	0.88	1	0.98	0.99
2	0.72	0.84	2	0.96	0.98
3	0.89	0.94	3	0.99	0.99
4	0.83	0.91	4	0.73	0.78
5	0.84	0.91	5	0.98	0.99
6	0.81	0.89	6	0.99	0.99
7	0.66	0.80	7	0.99	0.99
8	0.55	0.71	8	0.92	0.95
			9	0.99	0.99
			10	0.99	0.99
			11	0.99	1.00
			12	1.00	1.00
			13	1.00	1.00
			14	0.99	1.00
			15	0.99	0.99
			16	0.99	0.99
			17	0.99	0.99
			18	1.00	1.00
			19	0.99	0.99
			20	0.99	1.00
			21	0.99	0.99

<Table 6> Optimal settings obtained from two comparative approaches

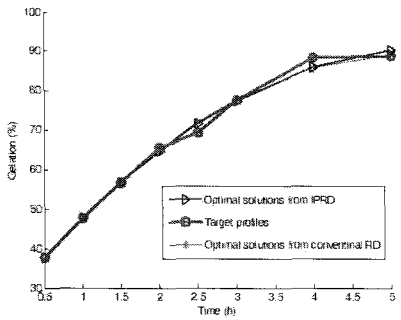
	Optimal setting x*									
Ingredients	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
RSM based RD	206.51	0.44	21.40	0.00	0.00	0.00	0.00	0.00	0.00	5.97
IP based RD	180.08	0.00	29.09	0.00	25.13	0.00	0.00	0.00	0.00	0.00

<Table 7> Optimal time-depended responses

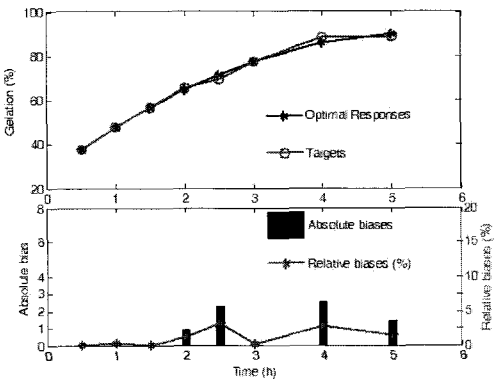
Gelation rates (%) at observed times (hours)								
Time	0.5h	1.0h	1.5h	2.0h	2.5h	3.0h	4.0h	5.0h
Targets	37.75	47.61	56.71	65.54	69.32	77.55	88.42	88.81
RSM based RD								
Optimal solutions	37.21	47.65	56.92	65.04	71.99	77.77	85.86	89.31
Absolute biases	0.54	0.04	0.21	0.50	2.67	0.22	2.56	0.50
% Biases	1.43	0.08	0.37	0.76	3.85	0.28	2.90	0.56
IP based RD								
Optimal solutions	37.73	47.76	56.74	64.67	71.55	77.39	85.91	90.24
Absolute biases	0.02	0.15	0.03	0.87	2.23	0.16	2.51	1.43
% Biases	0.06	0.31	0.05	1.33	3.22	0.21	2.84	1.61



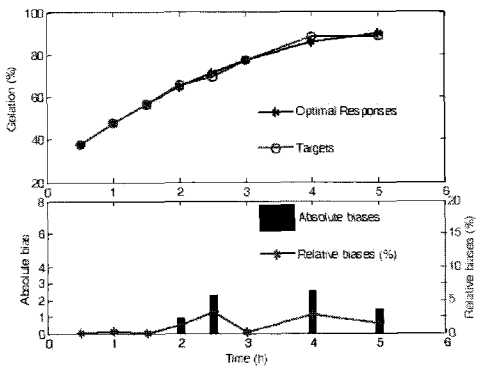
<Figure 2> Data profile. Each curve represents one set of time-depended responses at one experimental run. Along the curves are error bars represent variance data.



<Figure 3> Optimal solutions and their target values



<Figure 4> IP based optimal solutions



<Figure 5> RSM based optimal solutions

5. Conclusions

By implementing comparative research, this paper successfully illustrates how to apply proposed IPRD methodology to handle the situation of time-depended responses based RD problem. Then, the obtained results are compared with those came from conventional RSM based approach in term of absolute biases and relative biases. Comparative study showed that IP approach can dominate RSM approach in term of MSE (MSE_RSM = 1.821 and MSE_IP = 1.766). The results of study confirm that by using proposed IP approach in the RD procedure, robust solutions can give less biases than using RSM based approach. The reason can attribute to that IP approach can exploit the variance information better than conventional one. Further study will explore the larger data set so that some basic assumptions can be relaxed and the empirical models can be considered more complicated.

References

[1] Montgomery, D.C.(2001), *Design and Analysis of Experiments*, 5th Edition, John Wiley and Sons, New York.

[2] Myers, R.H. and Montgomery, D.C.(2002), *Response Surface Methodology*, John Wiley and Sons, New York.

[3] Park J. S., Shim J. Y., Truong N. K. V., Park J. S., Shin S.M., Choi Y. W., Lee J. H., Yoon J. H., Jeong S. H. (2010), "A pharma-robust design method to investigate the effect of PEG and PEO on matrix tablets. *International Journal of Pharmaceutics*, Vol. 393, No. 1-2, 30, pp. 80-88.

[4] Peterson, J.J., Snee, R.D., McAllIster, P.R., Schofield, T.L., Carella, A.J., (2009), "Statistics in Pharmaceutical Development and Manufacturing". *Journal of Quality Technology*, Vol. 41, pp. 111-132.

[5] Scheffé, H. (1958), "Experiments with Mixtures," *Journal of the Royal Statistical Society*. Series B, Vol. 20, No. 2, pp. 344-360.

[6] Scheffé, H.(1963), "The Simplex-Centroid Design for Experiments with Mixture," *Journal of the Royal Statistical Society*. Series B, Vol.25, No. 2, pp. 235-263.

[7] Shin S.M., Choi D. H., Truong N. K. V., Kim N. A., Chu K. R., Jeong S. H. (2011), "Time-oriented experimental design method to optimize hydrophilic matrix formulations with gelation kinetics and drug release profiles". *International Journal of Pharmaceutics*, Vol. 407, No. 1-2, 4, pp. 53-62.

[8] Walpole, R.E., Myers, R.H., Myers, S.L., Ye, K., (2007). *Probability and Statistics for Engineers and Scientists*, 4th Edition. Prentice Hall International Paperback Editions.

2011년 6월 9일 접수, 2011년 3월28일 수정, 2011년 8월 9일 채택