Revista Colombiana de Estadística
Junio 2013, volumen 36, no. 1, pp. 153 a 176

Response Surface Optimization in Growth Curves Through Multivariate Analysis

Optimización de superficies de respuesta en curvas de crecimiento a través de análisis multivariado

Felipe Ortiz[1,a], Juan C. Rivera[2,b], Oscar O. Melo[2,c]

[1]Facultad de Estadística, Universidad Santo Tomás, Bogotá, Colombia

[2]Departamento de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia

Abstract

A methodology is proposed to jointly model treatments with quantitative levels measured throughout time by combining the response surface and growth curve techniques. The model parameters, which measure the effect throughout time of the factors related to the second-order response surface model, are estimated. These estimates are made through a suitable transformation that allows to express the model as a classic MANOVA model, so the traditional hypotheses are formulated and tested. In addition, the optimality conditions throughout time are established as a set of specific combination factors by the fitted model. As a final step, two applications are analyzed using our proposed model: the first was previously analyzed with growth curves in another paper, and the second involves two factors that are optimized over time.

Key words: Growth curves, Multiple optimization, Response surfaces, Second order models.

Resumen

En este artículo se propone una metodología para modelar conjuntamente tratamientos con niveles cuantitativos medidos en el tiempo, mediante la combinación de técnicas de superficies de respuesta con curvas de crecimiento. Se estiman los parámetros del modelo, los cuales miden el efecto en el tiempo de los factores relacionados con el modelo de superficie de respuesta de segundo orden. Estas estimaciones se realizan a través de una transformación que permite expresar el modelo como un modelo clásico de MANOVA; de esta manera, se expresan y juzgan las hipótesis tradicionales.

[a]Lecturer. E-mail: andresortiz@usantotomas.edu.co
[b]MsC in Statistics. E-mail: jcriverar@unal.edu.co
[c]Associate professor. E-mail: oomelom@unal.edu.co

Además, las condiciones de optimización a través del tiempo son establecidas para un conjunto de factores específicos por medio del modelo ajustado. Como paso final, se analizan dos aplicaciones utilizando el modelo propuesto: la primera fue analizada mediante curvas de crecimiento en otro artículo, y la segunda consiste en dos factores que son optimizados a lo largo del tiempo.

***Palabras clave*:** curvas de crecimiento, optimización múltiple, superficies de respuesta, modelos de segundo orden.

1. Introduction

Sometimes in experimentation, researchers interest focuses on analyzing data over time to know the tendencies of an individual or groups of individuals. In other cases, the goal is not only the trend but also to know what combination of factors can optimize the process over time. This latter context is the starting point for analysis of growth curves and response surface methodology (RSM). Response surface and growth curves are statistical methods frequently used in the analysis of experiments. The purpose of the first is to determine the optimum operating conditions of a process, whereas the latter method is used to model the effect of treatments throughout time.

Two applications of the above hybrid model approach are analyzed in this paper. The first is an experiment to analyze the effect of dietary ingestion of sodium Zeolite A (SZA) on the growth and physiology of sixty horses reported by Frey, Potter, Odom, Senor, Reagan, Weir, Elsslander, Webb, Morris, Smith & Weigand (1992). The horses were randomly assigned to four treatments: control and three levels of dietary SZA (0.66%, 1.32% and 2%). In addition, plasma silicon concentration was measured at the times: $t = 0, 1, 3, 6, 9$ hours after ingestion on eighty four days into the diet. The second study is an experiment about the waste-water treatment, in which is common adding inhibitory agents to reduce the negative environmental impact generated by these substances discharged into the receiving water bodies. In such cases, we study the biological oxygen demand (BOD) as a water pollution measure. Montoya & Gallego (2012) performed a central composite rotatable design adding combinations of detergent (D) and animal fat (AF) to the residual water. The BOD, biomass growth and substrate consumption at $t = 24, 48, 72, 96, 120$ hours after of mixture were observed.

In both experiments, we are interested in studying the optimum combination of factors throughout time that optimizes our response variable. Therefore, in these kinds studies, we want to observe if the growth curves can be represented by a cubic, quadratic or linear polynomial in time, and if the response surface can be expressed by a quadratic or linear polynomial in the treatments. Furthermore, we want to obtain the confidence band(s) for the expected combination of factors over time (response surface throughout growth curves).

A growth curve is a model of the evolution of a quantity over time. Growth curves are widely used in biology for quantities such as population size, body height or biomass. Growth curve experiments have been considered from various angles by Rao (1959), Potthoff & Roy (1964), Khatri (1966), Khatri (1973), Verbyla

& Venables (1988), Kshirsagar & Boyce (1995), Srivastava (2002), Pan & Fang (2002), Chiou, Müller, Wang & Carey (2003) and Kahm, Hasenbrink, Lichtenberg-Fraté, Ludwig & Kschischo (2010). All of these growth curve studies involve successive and correlated measurements on the same individuals which are divided into two or more groups of treatments. We use in this paper treatments that are combinations of quantitative factors which are based on polynomial models in the response surface.

RSM uses statistical models and therefore practitioners need to be aware that even the best statistical model is an approximation to reality. In this way, if researches are interested in modeling and analyzing situations to determine optimum operating conditions for a process; this particular analysis is performed through the RSM. It is widely applicable in the biological sciences, chemistry, social experimentation agriculture, engineering, food sciences, quality control and economics, among others. The RSM has been developed in experimental and industrial production by Box & Wilson (1951), Hill & Hunter (1966), Mead & Pike (1975), Lucas (1976), Box & Draper (1982), Draper & Ying (1994), Chiou, Müller & Wang (2004) and Box & Draper (2007). These authors discussed some first-order and second-order response surface designs from the point of view of their ability to detect certain likely kinds of lack of fit for a higher's degree polynomial than has been fitted.

The two previous approaches to growth curve and RSM problems are now mixed to give a solution to our two applications because we need to know what is the combination of factors over time that best works in the optimization process. Our methodology is derived from the theory of multivariate normal analysis of variance, and it is based on polynomial models for both growth curve and response surface. Moreover, we provide both confidence bands and the over-all tests of significance for various kinds of compound hypotheses that involve the parameters of the proposed model. Furthermore, we find the optimal operating conditions over time.

This kind of problem was previously studied by Guerrero & Melo (2008) providing a solution where they combined the response surface and the growth curve techniques using an univariate analysis. In this paper, the same is done to obtain the functional relationship that exists between the treatment and time in order to predict its effect in any future time. Although, there are several phenomena of this kind where these two techniques may be used simultaneously, a procedure that combines them at the same time is not known using multivariate analysis. This analysis works better than the univariate approximation presented by Guerrero & Melo (2008) because the different statistics for hypothesis testing are exact, which does not always happen in the univariate approach.

The experiments to be considered are characterized by the presence of k fixed quantitative factors, $\zeta_1, \zeta_2, \dots, \zeta_k$, associated with a continuous variable of interest $\boldsymbol{Y}$, where the observed levels of each factor are equally spaced and the response variable is measured on the same experimental units in several moments.

The plan of the paper is the following: Section 2 presents the response surfaces model in growth curves. Then in Section 3, parameter estimation, hypotheses testing and test statistics are presented. Section 4 is dedicated to locating the optimum; at first the model is reparametrized (Section 4.1), and then the optimal point is found (Section 4.2). Finally, two applications of our procedure are showed in Section 5, and the conclusions are exposed in Section 6.

2. Response Surfaces Model in Growth Curves

The growth curve model implies that there are g different groups or treatments and a single growth variable y, which is measured at q time points $t_1, t_2, \dots, t_q$ on n_j individuals chosen at random from the j-th group ($j = 1, 2, \dots, g$). A polynomial regression of degree $(p-1)$ for y on the time variable t is defined. Thus,

$$E(y_t) = \phi_{j0} t^0 + \phi_{j1} t^1 + \cdots + \phi_{j(p-1)} t^{p-1} \tag{1}$$

$t = t_1, t_2, \dots, t_q, q > p-1, j = 1, 2, \dots, g$. The observations $y_{t_1}, \dots, y_{t_q}$ on the same individual are correlated, and come from a multivariate normal distribution with unknown variance-covariance matrix $\boldsymbol{\Sigma}$, equal for all the individuals. Let $\boldsymbol{Y}_j$ denote the $n_j \times q$ matrix of the observations for the j-th group, and let

$$\mathbf{Y}' = [\boldsymbol{Y}'_1, \boldsymbol{Y}'_2, \dots, \boldsymbol{Y}'_g]$$

denote the $q \times n$ matrix for all the $n = n_1 + n_2 + \cdots + n_g$ individuals. Then from (1)

$$E(\boldsymbol{Y}_j) = \begin{pmatrix} \boldsymbol{\phi}_j \mathbf{G} \\ \boldsymbol{\phi}_j \mathbf{G} \\ \vdots \\ \boldsymbol{\phi}_j \mathbf{G} \end{pmatrix} = \mathbf{J}_{n_j 1} \boldsymbol{\phi}_j \mathbf{G}, \quad j = 1, 2, \dots, g \tag{2}$$

where $\boldsymbol{\phi}_j = [\phi_{j0}, \phi_{j1}, \dots, \phi_{j(p-1)}]'$ denotes the vector of the regression or growth curve coefficients for the j-th group, and

$$\mathbf{G} = \begin{pmatrix} t_1^0 & t_2^0 & \dots & t_q^0 \\ t_1^1 & t_2^1 & \dots & t_q^1 \\ \vdots & \vdots & \ddots & \vdots \\ t_1^{p-1} & t_2^{p-1} & \dots & t_q^{p-1} \end{pmatrix}$$

and $\mathbf{J}_{a \times b}$ denotes, in general, an $a \times b$ matrix with all unit elements. Furthermore, the matrix $\mathbf{G}_{p \times q}$ relates the parameters of the curve with the corresponding polynomial degree.

Combining (2) for all g groups, we now have

$$E(\mathbf{Y}) = \begin{pmatrix} \mathbf{J}_{n_1 1}\boldsymbol{\phi}_1\mathbf{G} \\ \mathbf{J}_{n_2 1}\boldsymbol{\phi}_2\mathbf{G} \\ \vdots \\ \mathbf{J}_{n_g 1}\boldsymbol{\phi}_g\mathbf{G} \end{pmatrix} = \mathbf{A}\boldsymbol{\Phi}\mathbf{G} \tag{3}$$

where

$$\boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \\ \vdots \\ \boldsymbol{\phi}_g \end{pmatrix}$$

is the $g \times p$ matrix of the growth curve coefficients, and $\boldsymbol{A} = diag[\mathbf{J}_{n_1 1}, \mathbf{J}_{n_2 1}, \dots \mathbf{J}_{n_g 1}]$ is a block diagonal matrix of order $n \times g$ 'group indicator'. Therefore, assuming independence between individuals, we have that

$$Var(Vec(\mathbf{Y})) = \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_q \tag{4}$$

where $\otimes$ denotes the Kronecker product of two matrices (see Magnus (1988)).

The equations (3) and (4) conform to the growth curve model introduced by Potthoff & Roy (1964), and later analyzed by Khatri (1966), Grizzle & Allen (1969), Kabe (1974), and Khatri (1988), among many others.

2.1. Construction of Proposed Model

With the idea of making a joint modeling of growth curves and response surfaces, a couple of aspects were considered:

1. The matrix $\mathbf{A}_{n\times g}$, whose columns contain information about treatments, was changed by a new matrix $\mathbf{X}_{n\times s}$, whose columns register the levels of a factor and their interactions for each of the n individuals, just like in second order response surfaces designs with k quantitative fixed factors and $s = 1 + k + (k + \binom{k}{2})$ parameters in the surface.

2. For relating the parameters from the response surface with each of the groups, a new matrix of parameters $\boldsymbol{\theta}_{s\times g}$ was included in the model where θ_{lj} measures the effect of the l-th parameter in the surface for the $j-th$ group. Let $\boldsymbol{\Phi}_{g\times p}$ be the matrix that relates the groups with the growth curve coefficients, i.e., ϕ_{jm} is the parameter associated to the degree coefficient m in the growth curve for the j-th group ($l = 1, 2, \dots, s; j = 1, 2, \dots, g; m = 0, 1, \dots, p-1$).

Under the usual assumptions described above and maintaining the same structure and interpretation for the matrices $\mathbf{Y}_{n\times q}$ and $\mathbf{G}_{p\times q}$, the proposed model is given by

$$E(\mathbf{Y}_{n\times q}) = \mathbf{X}_{n\times s}\,\boldsymbol{\theta}_{s\times g}\,\boldsymbol{\Phi}_{g\times p}\,\mathbf{G}_{p\times q} \tag{5}$$

Notice that the model (5) is a classic model Potthoff & Roy (1964) adaptation in which the matrix $\boldsymbol{\xi}_{s\times p} = \boldsymbol{\theta}_{s\times g}\boldsymbol{\Phi}_{g\times p}$, whose components are given by

$$\xi_l^m = \sum_{j=1}^{g} \theta_{lj}\phi_{jm}$$

which is the parameter associated with the l-th component of the surface in the m-th growth curve degree ($l = 1, 2, \dots, s$ and $m = 0, 1, \dots, p-1$). This allows to write (5) as

$$E(\mathbf{Y}_{n\times q}) = \mathbf{X}_{n\times s}\,\boldsymbol{\xi}_{s\times p}\,\mathbf{G}_{p\times q} \tag{6}$$

Another form for writing this model is

$$E(y_{ia}) = \sum_{m=0}^{p-1}\left(\xi_{\,0}^{\,m} + \sum_{r=1}^{k}\xi_{\,r}^{\,m} x_{ir} + \sum_{r=1}^{k}\sum_{r'=1}^{k}\xi_{\,rr'}^{\,m}\, x_{ir}x_{ir'}\right) t_a^m \tag{7}$$

or equivalently the model (6) can be written as

$$E(y_{ia}) = \left(\sum_{m=0}^{p-1}\xi_0^m t_a^m\right) + \sum_{r=1}^{k} x_{ir}\left(\sum_{m=0}^{p-1}\xi_r^m t_a^m\right) + \sum_{r=1}^{k}\sum_{r'=1}^{k} x_{ir}x_{ir'}\left(\sum_{m=0}^{p-1}\xi_{rr'}^m t_a^m\right) \tag{8}$$

with $a = 1, 2, \dots, q$ and $i = 1, \dots, n$, and where $\xi_{\,rr'}^{\,m}$ is the parameter that denotes the effect of the interaction between the factors r and r' in the m-th growth curve degree ($r, r' = 1, 2, \dots, k$ and $m = 0, 1, \dots, p-1$), x_{ir} and $x_{ir'}$ are encoded explanatory variables associated to the factors r-th and r'-th, respectively, and y_{ia} is the response variable associated to the i-th individual in the a-th time.

Note that the model (7) is in fact a growth curve whose coefficients are themselves a response surface of order two, and the model (8) is a response surface whose parameters are growth curves. Moreover in (7), it is necessary to point out that for a fixed m, all the parameters of the form ξ_0^m, ξ_r^m, $\xi_{rr'}^m$ $(r, r' = 1, 2, \dots, k)$ belong to the m-th column of $\boldsymbol{\xi}$. Similarly, in (8), each set of parameters of the form ξ_0^m, ξ_r^m, $\xi_{rr'}^m$ with $m = 0, 1, \dots, p-1$ and fixed r, r', conforms the rows of $\boldsymbol{\xi}$. The remarks above are of great utility in section 3.2 for building the hypotheses of interest on the model parameters.

3. Inference on the Model

3.1. Parameter Estimation

Parameter estimation is achieved by expressing the model (6) as a MANOVA classic model, using the following transformation

$$\mathbf{Y}^{\triangle} = \mathbf{Y}\mathbf{P}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}')^{-1} \tag{9}$$

with $\mathbf{P}$ any symmetric positive definite matrix, such that $(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}')^{-1}$ exists.

By applying the transformation (9) in (6), the next expression is obtained

$$\begin{aligned} E(\mathbf{Y}^{\triangle}_{n\times p}) &= \mathbf{X}_{n\times s}\,\boldsymbol{\xi}_{s\times p} \\ Var(\mathbf{Y}^{\triangle}_{i}) &= (\mathbf{G}\mathbf{P}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{P}^{-1}\boldsymbol{\Sigma}\mathbf{P}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}')^{-1} \\ &= \boldsymbol{\Sigma}^{\triangle}_{p}, \qquad\qquad i = 1, 2, \dots, n \end{aligned} \tag{10}$$

Potthoff & Roy (1964) found that taking $\mathbf{P} = \boldsymbol{\Sigma}$ produces the minimum variance estimator for $\boldsymbol{\xi}$; however, since $\boldsymbol{\Sigma}$ is unknown, in practice $\mathbf{P}$ is given by

$$\mathbf{P} = \mathbf{S} = \mathbf{Y}'\{\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\}\mathbf{Y} \tag{11}$$

Note that $\mathbf{P}$ can take different forms which depend of the data correlation structure; a complete discussion about $\mathbf{P}$ can be found in Davis (2002), Molenberghs & Verbeke (2005), and Davidian (2005).

Then, for model (10), the parameter estimators obtained with the maximum likelihood method are given by

$$\widehat{\boldsymbol{\xi}} = (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}^{\triangle} \tag{12}$$

From a slight extension of the result given by Rao (1967) in equation 50, we can find that the unconditional covariance matrix of the elements of $\widehat{\boldsymbol{\xi}}$ can be expressed as

$$Var\left(\widehat{\boldsymbol{\xi}}\,'\right) = \tfrac{n-s-1}{n-s-q+p-1}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \otimes \boldsymbol{\Sigma}^{\triangle} \tag{13}$$

where $\otimes$ is the Kronocker product, and $Var(\widehat{\boldsymbol{\xi}}\,')$ denotes the covariance matrix of the elements of $\widehat{\boldsymbol{\xi}}$ taken in a columnwise manner.

It is easily shown that $E(\widehat{\boldsymbol{\xi}}) = \boldsymbol{\xi}$, and using a result given by Grizzle & Allen (1969), we find that $E((\mathbf{G}\mathbf{S}^{-\mathbf{1}}\mathbf{G}')^{-1}) = (n-s-q+p)\boldsymbol{\Sigma}^{\triangle}$. From this last equation and equation (13), it follows that an unbiased estimator of the variance of $\widehat{\boldsymbol{\xi}}$ is

$$\widehat{Var}\left(\widehat{\boldsymbol{\xi}}\,'\right) = \tfrac{n-s-1}{n-s-q+p-1}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \otimes \widehat{\boldsymbol{\Sigma}}^{\triangle} \tag{14}$$

where

$$\widehat{\boldsymbol{\Sigma}}^{\triangle} = \frac{1}{n-s-q+p}(\mathbf{G}\mathbf{S}^{-\mathbf{1}}\mathbf{G}')^{-\mathbf{1}}$$

In next Subsection, we will present a classic technique for testing a hypothesis of the form $\mathbf{C}\,\boldsymbol{\xi}\,\mathbf{U} = \mathbf{0}$ under the generalized expectation model (6), and also we will obtain related confidence bounds.

3.2. Hypothesis of Interest and Test Statistics

As shown in the section 2.1, the model (6) can be written by expressions (7) and (8), where it can be observed that the hypotheses of interest lie mainly on the rows or the columns of the matrix $\boldsymbol{\xi}$. These and many other can be written in the conventional general linear hypothesis form

$$H_0: \quad \mathbf{C}\,\boldsymbol{\xi}\,\mathbf{U} = \mathbf{0} \qquad vs \qquad H_1: \quad \mathbf{C}\,\boldsymbol{\xi}\,\mathbf{U} \neq \mathbf{0} \tag{15}$$

where $\mathbf{C}_{c\times s}$ and $\mathbf{U}_{p\times u}$ are known matrices of ranges c ($\leq s$) and u ($\leq p$), respectively. The matrices that define the main hypotheses, together with their corresponding interpretation, are shown in Table 1.

TABLE 1: Hypotheses more common over treatments and times.

H_0	Interpretation	$\mathbf{C}$	$\mathbf{U}$
$\boldsymbol{\xi} = \mathbf{0}$	The time-parameter interaction adjusted by the intercepts is not significant.	$\begin{pmatrix} 0 & \mathbf{0}_{1\times s-1} \\ \mathbf{0}_{s-1\times 1} & \mathbf{I}_{s-1} \end{pmatrix}$	$\begin{pmatrix} 0 & \mathbf{0}_{1\times p-1} \\ \mathbf{0}_{p-1\times 1} & \mathbf{I}_{p-1} \end{pmatrix}$
$\boldsymbol{\xi}^{(m)} = \mathbf{0}$	The m-th column of ξ is zero, indicating that the degree m coefficient is not important in the growth curve.	$\mathbf{I}_s$	$(0,\ldots,\underset{m-th}{\overset{1}{\downarrow}},\ldots,0)'_{p\times 1}$
$\boldsymbol{\xi}_{(l)} = \mathbf{0}$	The l-th row of ξ is zero, indicating that the parameter of the surface is not significant.	$(0,\ldots,\underset{l-th}{\overset{1}{\downarrow}},\ldots,0)_{1\times s}$	$\mathbf{I}_p$
$\xi_l^m = 0$	The l-th component of the surface does not exercise influence in the m-th degree of the curve.	$(0,\ldots,\underset{l-th}{\overset{1}{\downarrow}},\ldots,0)_{1\times s}$	$(0,\ldots,\underset{m-th}{\overset{1}{\downarrow}},\ldots,0)'_{p\times 1}$

For the construction of the test statistics, the following two matrices should be kept in mind

$$\mathbf{H} = \mathbf{U}'\widehat{\boldsymbol{\xi}}'\mathbf{C}'[\mathbf{C}\mathbf{R_1}\mathbf{C}']^{-\mathbf{1}}\mathbf{C}\widehat{\boldsymbol{\xi}}\mathbf{U}$$
$$\mathbf{E} = \mathbf{U}'(\mathbf{G}\mathbf{S}^{-\mathbf{1}}\mathbf{G}')^{-\mathbf{1}}\mathbf{U}$$

where

$$\mathbf{R_1} = \left\{\mathbf{I} + (\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}'\mathbf{Y}\mathbf{S}^{-\mathbf{1}}\left[\mathbf{I} - \mathbf{G}'(\mathbf{G}\mathbf{S}^{-\mathbf{1}}\mathbf{G}')^{-\mathbf{1}}\mathbf{G}\mathbf{S}^{-\mathbf{1}}\right]\mathbf{Y}'\mathbf{X}\right\}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}$$

H and **E** play a decisive role in building the four classic multivariate test statistics used in testing hypothesis (15) under the model (10): the Roy's test uses the largest characteristic root of $(\mathbf{H}\mathbf{E}^{-\mathbf{1}})$, the Lawley-Hotelling $T^2 = tr(\mathbf{H}\mathbf{E}^{-\mathbf{1}})$, the trace of Bartlett-Nanda-Pillai $V = tr(\mathbf{H}(\mathbf{H}+\mathbf{E})^{-\mathbf{1}})$, and the statistic proposed by Wilks (1932), $|\mathbf{E}|/|\mathbf{H}+\mathbf{E}| \sim \Lambda_{(u,c,m)}$ with $m = n-[s+(q-p)]$, which is known as the λ-criterion.

The hypotheses in Table 1 involves one row vector or one column vector. Therefore, we now state the general rule of rejecting a null hypothesis based on Wilks's Λ, using a level of significance α. To test the null hypothesis $\boldsymbol{\xi}^{(m)} = \mathbf{0}$, we use the following test presented in Kshirsagar & Boyce (1995)

$$u_1 = \frac{1-\Lambda}{\Lambda}\frac{ddf}{ndf}$$

where ddf is the denominator degree of freedom and ndf is the numerator degrees of freedom. Then, the null hypothesis $\boldsymbol{\xi}^{(m)} = \mathbf{0}$ is rejected if $u_1 > F_{(\alpha,ndf,ddf)}$.

To test the null hypothesis $\boldsymbol{\xi}_{(l)} = \mathbf{0}$, we use the following test presented in Kshirsagar & Boyce (1995)

$$c_1 = \frac{1-\Lambda}{\Lambda}\left(\frac{c+ddf-u}{u}\right)$$

Then, the null hypothesis $\boldsymbol{\xi}_{(l)} = \mathbf{0}$ is rejected if $c_1 > F_{(\alpha,u,c+ddf-u)}$.

On the other hand, simultaneous 100(1-α)% confidence bounds for the function $\mathbf{b}'\mathbf{C}\boldsymbol{\xi}\mathbf{U}\mathbf{f}$, $\forall\ \mathbf{b}_{(c\times 1)}$ and $\mathbf{f}_{(u\times 1)}$, are given by

$$\mathbf{b}'\mathbf{C}\widehat{\boldsymbol{\xi}}\mathbf{U}\mathbf{f} \pm \left\{\left(\tfrac{h_\alpha}{1-h_\alpha}\right)(\mathbf{b}'\mathbf{C}\mathbf{R}_1\mathbf{C}'\mathbf{b})(\mathbf{f}'\mathbf{E}\mathbf{f})\right\}^{1/2} \tag{16}$$

where the prediction is, of course, the first term of the equation (16) and h_α stands for the α fractile of the distribution for the Roy's largest characteristic root statistic tabulated by Heck (1960) with its three parameters (denoted by s, m and n in Heck's notation, but here denoted, respectively, by s^*, m^* and n^*) equal to $s^* = \min(c,u)$, $m^* = (|c-u|-1)/2$ and $n^* = (n-s-(q-p)-u-1)/2$.

Other test statistics are presented in some works; for instance, Grizzle and Allen's statistic (1969) which considers a variant for the matrix associated the hypothesis (relating to the herein presented). Singer & Andrade (1994) remarked on the appropriate selection of error terms and presented a test statistic that follows an exact F distribution (under H_0). This was also used in the application of Section 5, since it yielded the same decisions as the test statistics exposed there.

4. Location of the Optimum

The crucial goal of the response surface methodology is to find the optimal operating conditions for the variable of interest, and in this scenario, their behavior throughout time is added.

4.1. Reparameterization of the Model

In order to find the optimal operating conditions in presence of multiple responses, it is convenient to find an expression that us allows to distinguish the terms of order zero, one and two of the model (10). This model can be reparametrized as

$$\begin{aligned}\widehat{\mathbf{Y}}^{\triangle}_{i_{1\times p}} &= \mathbf{b}_{\,0_{1\times p}} + (\mathbf{x}'_{1\times k}\mathbf{b}_{\,k\times p}) + (\mathbf{x}'\mathbf{B}^{(0)}\mathbf{x}, \mathbf{x}'\mathbf{B}^{(1)}\mathbf{x}, \dots, \mathbf{x}'\mathbf{B}^{(p-1)}\mathbf{x}) \\ &= \mathbf{b}_{\,0_{1\times p}} + (\mathbf{x}'_{1\times k}\mathbf{b}_{\,k\times p}) + (\mathbf{x}'_{1\times k}\mathbf{B}_{k\times kp})(\mathbf{I}_p \otimes \mathbf{x}_{k\times 1})\end{aligned} \tag{17}$$

where $\mathbf{x}_{k\times 1}$ is the vector associated to the k factors of the response surfaces, $\mathbf{b}_{\,0_{1\times p}}$ is the vector whose components are the intercepts of each curve degree, $\mathbf{b}_{\,k\times p}$ is the matrix that contains the coefficients associated to the k linear terms of the response surface for each of the curve degrees, and $\mathbf{B}_{k\times kp}$ is the matrix $(\mathbf{B}^{(0)}, \mathbf{B}^{(1)}, \dots, \mathbf{B}^{(p-1)})$ with $\mathbf{B}^{(m)}$ $(m = 0, 1, \dots, p-1)$ being the $k \times k$ matrix associated to the quadratic form of the response surface for the m-th growth curve degree.

4.2. Optimization

The location of the optimal point is obtained by solving the equation system resulting from the expression

$$\frac{\partial \widehat{\mathbf{Y}}^{\triangle}_{i}}{\partial \mathbf{x}} = \mathbf{b}_{\,k\times p} + 2[\mathbf{B}^{(0)}\mathbf{x} \vdots \mathbf{B}^{(1)}\mathbf{x} \vdots \cdots \vdots \mathbf{B}^{(p-1)}\mathbf{x}] = \mathbf{0} \tag{18}$$

which is demonstrated using properties of differential matrix calculus.

By applying the *vec* operator in system (18), the following system of k variables and kp equations is obtained

$$vec(\mathbf{b}_{\,k\times p}) + 2\begin{pmatrix}\mathbf{B}^{(0)}\mathbf{x}\\ \mathbf{B}^{(1)}\mathbf{x}\\ \vdots \\ \mathbf{B}^{(p-1)}\mathbf{x}\end{pmatrix} = \mathbf{0}$$

$$\mathbf{B}'\mathbf{x} = -\frac{1}{2}vec(\mathbf{b}_{\,k\times p}),$$

which is solved by appending to it a pre-matrix $\mathbf{B}$; hence, the stationary point is

$$\mathbf{x}_0 = -\frac{1}{2}(\mathbf{B}\mathbf{B}')^{-1}\mathbf{B}vec(\mathbf{b}_{\,k\times p}) \tag{19}$$

and the non-singularity of $\mathbf{B}\mathbf{B}'$ is guaranteed by the linear independence of the columns of $\mathbf{X}'\mathbf{X}$.

Let $\gamma_1, \gamma_2, \dots, \gamma_k$ be the characteristic roots of the matrix $\mathbf{B}\mathbf{B}'$, then the nature of the stationary point is determined by

- If $\gamma_v > 0 \; \forall v = 1, 2, \dots, k$, then $\mathbf{x}_0$ is minimum.
- If $\gamma_v < 0 \; \forall v = 1, 2, \dots, k$, then $\mathbf{x}_0$ is maximum.

- In any other case, $\mathbf{x}_0$ is a saddle point.

Using (16) with $\mathbf{C} = \mathbf{I}$, $\mathbf{U} = \mathbf{I}$, $\mathbf{b}' = \mathbf{x}_0$ and $\mathbf{f} = \mathbf{G}^{(m)}$; the confidence bounds for the predicted values of the optimal point in each moment are given by

$$\mathbf{x}_0'\widehat{\boldsymbol{\xi}}\mathbf{G}^{(m)} \pm \left\{ \left(\frac{1}{2n^*+2} F_{(\alpha,1,2n^*+2)} \right) (\mathbf{x}_0'\mathbf{R}_1\mathbf{x}_0) \left[\mathbf{G}^{(m)\prime}\mathbf{E}\mathbf{G}^{(m)} \right] \right\}^{1/2} \tag{20}$$

where $\mathbf{G}^{(m)}$ is a column of the matrix $\mathbf{G}$ and $n^* = (n - s - 2)/2$.

5. Applications

Two applications are analyzed in this Section: the first is an experiment to analyze the plasma silicon concentration and its effect over the dietary ingestion of SZA on the growth of sixty horses (Frey et al. 1992), and the second is an experiment about the waste-water treatment, where the biological oxygen demand (BOD) as a water pollution is studied (Montoya & Gallego 2012).

5.1. Plasma Silicon Concentration

An experiment to analyze the effect of dietary ingestion of SZA on the growth and physiology of sixty horses was reported by Frey et al. (1992). The horses were randomly assigned to four treatments: control (0%) and three levels of dietary SZA (0.66%, 1.32% and 2%). In addition, the plasma silicon concentration was measured in the times: $t = 0, 1, 3, 6$ and 9 hours after ingestion at eighty four days into the diet. This data was previously analyzed by Kshirsagar & Boyce (1995) employing growth curves, but they did not consider the surface responses part. However, Guerrero & Melo (2008) presented an optimization process that combines response surface and growth curves from a univariate approach. The last analysis differs from the work in this paper because we make a parameter estimation which does not depend on the transformation of equation (9). Additionally, the test statistics used in Guerrero & Melo (2008) follow a F distribution approximately, while under the multivariate perspective employed throughout this paper, these tests follow an exact distribution of Wilks's Λ.

Figure 1 shows profiles plot for these data. In this Figure, we see that the silicon concentration in the plasma can be modeled as a cubic polynomial over time. Also, the control group (0%) seems to have a different behavior than other concentrations which suggest a difference among the four treatments.

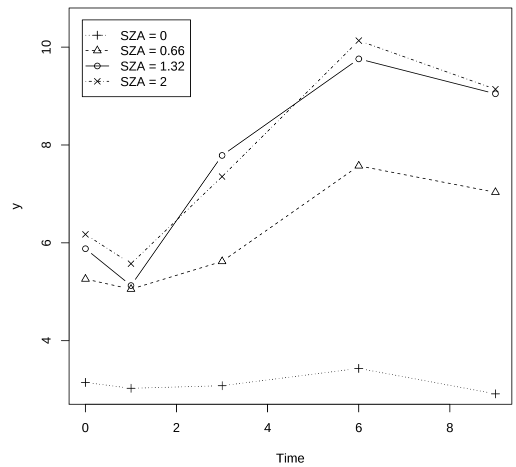

FIGURE 1: Profiles by time for plasma silicon concentration growth.

Fitting the model (6) to this data set, the parameter estimates given by (12) and $\mathbf{P}$ as (11) are

$$\widehat{\boldsymbol{\xi}} = \begin{bmatrix} 3.267 & -0.324 & 0.118 & -0.010 \\ 3.169 & 0.151 & 0.192 & -0.017 \\ -0.921 & -0.127 & -0.024 & 0.002 \end{bmatrix}$$

where the rows are growth curves for the different parameters of the response surface, and the columns correspond to response surfaces for the different growth curve degrees. So, the first row contains the intercepts of the surfaces for the polynomial degrees, the second row contains the linear component of the factor (SZA), and the third row contains the quadratic component of the factor.

Now, the results of the hypotheses testing on the rows (surface parameters) and the columns (curve coefficients) of $\boldsymbol{\xi}$ are shown in Table 2. The hypothesis $H_0 : \boldsymbol{\xi}^{(4)} = \mathbf{0}$ yields a $p - value < 0.001$; therefore, the hypothesis is rejected. This means that the third-order coefficient of the fitted growth curve is significant in the model (see right panel of Figure 2). The hypothesis $H_0 : \boldsymbol{\xi}_{(3)} = \mathbf{0}$ also yields a $p - value < 0.001$, denoting that the quadratic component of the factor is important, too (see left panel of Figure 2). The hypothesis $H_0 : \boldsymbol{\xi}^{(2)} = \mathbf{0}$ is the only one that is not rejected, it corresponds to the linear component of the curve. However, since the degree of the cubic growth curve is significant, the linear component is also included due to the hierarchy of the fitted growth curve.

Table 2: Hypotheses testing on the rows and columns for the effect of dietary ingestion of SZA.

Hypothesis	$\mathbf{C}$	$\mathbf{U}$	Λ	F_c	ndf	ddf	$p-value$
$\boldsymbol{\xi}^{(1)} = \mathbf{0}$	$\mathbf{I}_3$	$(1,0,0,0)'$	0.099	169.45	3	56	< 0.001
$\boldsymbol{\xi}^{(2)} = \mathbf{0}$	$\mathbf{I}_3$	$(0,1,0,0)'$	0.928	1.44	3	56	0.2407
$\boldsymbol{\xi}^{(3)} = \mathbf{0}$	$\mathbf{I}_3$	$(0,0,1,0)'$	0.584	13.29	3	56	< 0.001
$\boldsymbol{\xi}^{(4)} = \mathbf{0}$	$\mathbf{I}_3$	$(0,0,0,1)'$	0.471	20.98	3	56	< 0.001
$\boldsymbol{\xi}_{(1)} = \mathbf{0}$	$(1,0,0)$	$\mathbf{I}_4$	0.350	24.57	4	53	< 0.001
$\boldsymbol{\xi}_{(2)} = \mathbf{0}$	$(0,1,0)$	$\mathbf{I}_4$	0.290	32.38	4	53	< 0.001
$\boldsymbol{\xi}_{(3)} = \mathbf{0}$	$(0,0,1)$	$\mathbf{I}_4$	0.510	12.74	4	53	< 0.001

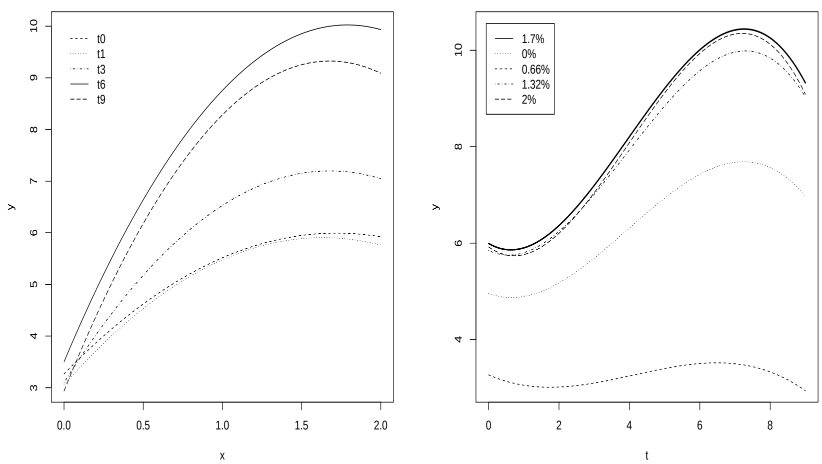

Figure 2: Fitted response surfaces (left panel) and growth curves (right panel).

From the matrix of estimated parameters, it is possible to construct the estimated growth curves for the four treatments of the experimental design. For example, for treatment 0.66%, the growth curve is given by the equation

$$\begin{pmatrix} 1 & 0.66 & 0.66^2 \end{pmatrix} \begin{pmatrix} 3.267 & -0.324 & 0.118 & -0.010 \\ 3.169 & 0.151 & 0.192 & -0.017 \\ -0.921 & -0.127 & -0.024 & 0.002 \end{pmatrix} \begin{pmatrix} 1 \\ t \\ t^2 \\ t^3 \end{pmatrix}$$
$$= 4.957 - 0.279t + 0.233t^2 - 0.019t^3 \tag{21}$$

For the four treatments, the fitted growth curves are summarized in Table 3 and on Figure 2 (right panel). In the same way, we can find the estimation of the response surface parameters at each point in time. From product matrix, $\widehat{\boldsymbol{\xi}}\mathbf{G}$, the equations are derived and summarized in Table 4, and they are plotted on Figure 2 (left panel).

Table 3: Fitted growth curves for the effect of dietary ingestion of SZA.

Treatment (SZA)	Growth curve
0.00	$3.266 - 0.323t + 0.117t^2 - 0.009t^3$
0.66	$4.957 - 0.279t + 0.233t^2 - 0.019t^3$
1.32	$5.845 - 0.345t + 0.328t^2 - 0.027t^3$
2.00	$5.921 - 0.529t + 0.403t^2 - 0.034t^3$

Table 4: Fitted response surfaces for the effect of dietary ingestion of SZA.

Time	Response surfaces
t_0	$3.266 + 3.169(\text{SZA}) - 0.92(\text{SZA})^2$
t_1	$3.051 + 3.495(\text{SZA}) - 1.07(\text{SZA})^2$
t_3	$3.097 + 4.88(\text{SZA}) - 1.456(\text{SZA})^2$
t_6	$3.051 + 7.292(\text{SZA}) - 2.038(\text{SZA})^2$
t_9	$2.936 + 7.616(\text{SZA}) - 2.27(\text{SZA})^2$

On the other hand, the level of SZA that maximizes the plasma silicon concentration regularly well throughout time obtained with (19) is 1.70%, where $\mathbf{b} = (3.169, 0.151, 0.192, -0.017)$ and $\mathbf{B} = (-0.921, -0.127, -0.024, 0.002)$. The confidence bounds in the optimal point constructed using (20) and $\mathbf{x}_0 = (1, 1.7, 1.7^2)$ are shown in Table 5.

Table 5: Parameter estimation and confidence bounds in the optimal point for the effect of dietary ingestion of SZA.

	t_0	t_1	t_3	t_6	t_9
Estimated value	5.99	5.9	7.19	10.01	9.32
Lower limit	5.72	5.66	6.93	9.74	9.06
Upper limit	6.25	6.13	7.45	10.27	9.59

Under the same reasoning used in equation (21), it is possible to construct the growth curve for the optimum point (1.7%), which is given by $5.99 - 0.43t + 0.37t^2 - 0.03t^3$. Figure 2 (right panel) shows the optimum supremacy over all treatments throughout time.

According to the results obtained in this application, we can stand out three facts:

1. in the solution via univariate developed by Guerrero & Melo (2008), in which one time ($t = 1$) was removed to get that the remaining times ($t = 0, 3, 6, 9$)

were equally spaced; the quadratic component SZA factor in the response surface was not significant, and also a linear polynomial for the growth curve was fitted.

2. the hypothesis $H_0 : \boldsymbol{\xi}_{(3)} = \mathbf{0}$ is rejected, justifying the inclusion of the quadratic component in the response surface to fit the plasma silicon concentration. Note that in our proposal the test statistic follows an exact F distribution.

3. Figure 1 clearly suggests that we should fit a cubic model in the growth curve, which is corroborated by the results of the hypothesis $H_0 : \boldsymbol{\xi}^{(4)} = \mathbf{0}$.

5.2. Environmental Pollution

During waste-water treatment it is common inhibitory agents to reduce the negative environmental impact generated by to add substances discharged into the receiving water bodies. Montoya & Gallego (2012) performed a central composite rotatable design adding combinations of detergent (D in ppm) and animal fat (AF in ppm). They studied the residual water BOD and biomass growth and substrate consumption at $t = 12, 24, 36, 48, 60$ hours after the mixture. These components interfere with the biological degradation of organic material during the process of waste-water treatment. In this case, we study the biomass (in mg/l) growth as a water pollution measure. According to Montoya & Gallego (2012), the presence of detergents and animal fat in the affluent waste-water affect the size and shape of the resulting floccules, which produces as a result a decrease in biomass concentration demanding more time for the system retention that translates into a low BOD elimination.

A description of the behavior of the four factorial points (treatments) of the experimental design throughout time is shown in Figure 3 (left panel). This Figure shows a slight increase of biomass between 12 and 24 hours after that the treatments were applied. Furthermore, we see an accelerated growth between 24 and 48 hours and a slight decrease from 48 until 60 hours. This behavior can be approximated by a cubic polynomial throughout time. Moreover, it is noted that the profiles for the four treatments have a very similar behavior, which suggests that there is not a differential effect for factors D and AF.

In order to observe the behavior of biomass growth at each time point, we fitted the univariate response surfaces for each time (see Figure 4). We can see that the fitted surfaces for the first two times ($t = 12, 24$) have a convex shape unlike the three last times ($t = 36, 48, 60$), which have concave shape. The points that optimize each response surface are shown in Table 6; there is a change in the optimal location point between two convex curves and three concave curves.

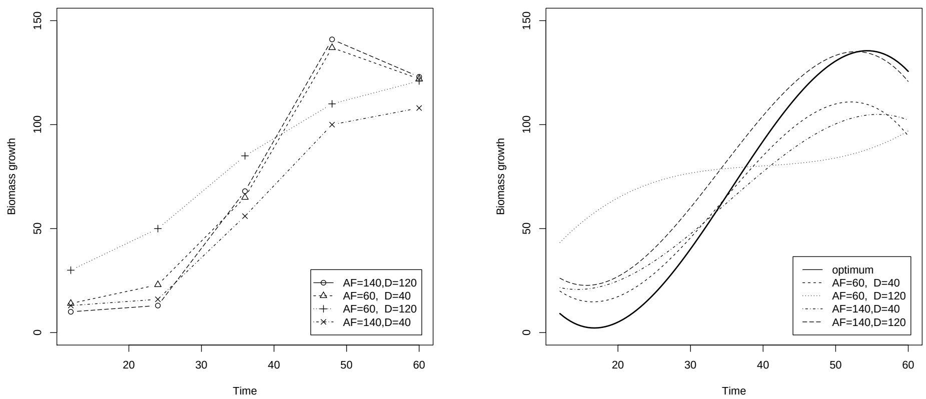

FIGURE 3: Profiles throughout time for the biomass growth (left panel), and fitted growth curves for the four treatments and the optimal point throughout time (right panel).

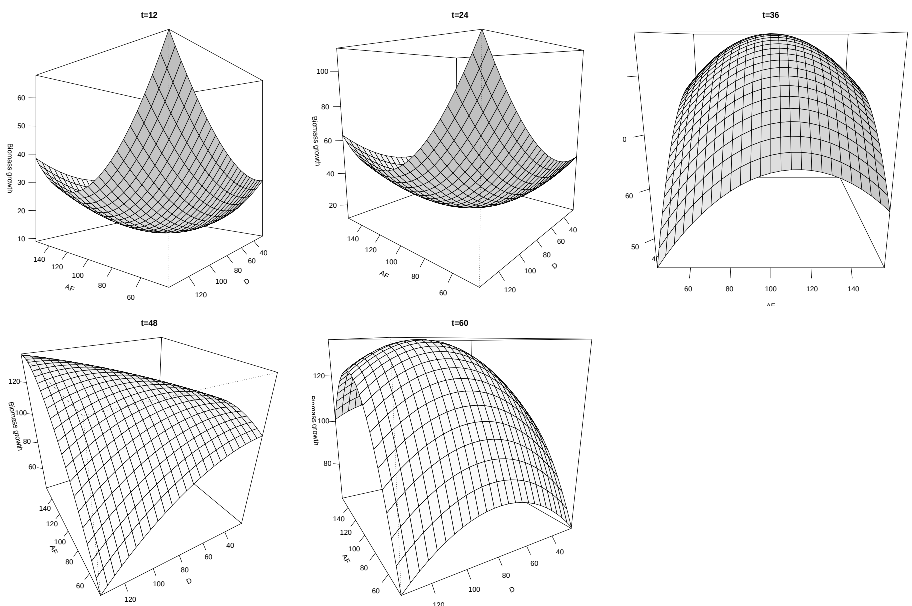

FIGURE 4: Fitted univariate response surfaces.

Table 6: Univariate optimal surfaces.

Time	AF	D	Characterization
t_{12}	100.7	50.8	Minimum
t_{24}	102.7	54.7	Minimum
t_{36}	100.3	104.6	Maximum
t_{48}	183.2	188.7	Maximum
t_{60}	109.2	92.7	Maximum

To fit the model (6), we first analyzed the structure of the matrix $\mathbf{P}$ given in equation (9). Then, we evaluated several possible covariance structures considering the fit to the data and comparing them in terms of Akaike information criterion (AIC). So, it was found that the best covariance structure was an AR(1) with parameter estimates: $\widehat{\phi} = 0.676$ and $\widehat{\sigma} = 23.51$, and the smallest AIC was 492.1. Thus, the estimated parameters matrix using the equation (12) is

$$\widehat{\boldsymbol{\xi}} = \begin{pmatrix} 66.3998 & 0.5581 & 0.0241 & -0.0010 \\ 1.4689 & -0.3724 & 0.0131 & -0.0001 \\ -1.1473 & 0.1479 & -0.0050 & 0.0000 \\ -0.0146 & 0.0029 & -0.0001 & 0.0000 \\ -0.0111 & 0.0018 & 0.0000 & 0.0000 \\ 0.0245 & -0.0036 & 0.0001 & 0.0000 \end{pmatrix}$$

whose first row represents the estimates of the response surface intercepts (ξ_0^m) in the four degrees of growth curve following the expression (7). The second and third rows are associated with the linear effects of factors AF and D, while the third and the fourth rows are associated with the quadratic effects of the factors, and finally; the sixth row estimates the interaction of two factors in all degrees of the growth curve.

Once the above is done, we show in Table 7 the results of the hypotheses testing on the rows (surface parameters) and the columns (curve parameters) of $\boldsymbol{\xi}$. According to the hypothesis $\boldsymbol{\xi}^{(4)} = \mathbf{0}$, a cubic polynomial fit to the growth curve is suitable ($p - value = 0.0039$), while for the hypothesis $\boldsymbol{\xi}_{(2)} = \mathbf{0}, \ldots, \boldsymbol{\xi}_{(6)} = \mathbf{0}$, we do not find evidence of a significant difference between the effects generated by AF and D factors in the experimental design. This is consistent with the behavior seen in Figure 3 (left panel) for the four treatments of central composite rotatable design; however, these factors could be interacting with the time (see Figure 3, left panel) so these components will be kept in the model.

Table 7: Hypotheses testing on the rows and columns for the biomass.

Hypothesis	$\mathbf{C}$	$\mathbf{U}$	Wilks	F_c	ndf	ddf	$p-value$
$\boldsymbol{\xi}^{(1)}=\mathbf{0}$	$\mathbf{I}_6$	$(1,0,0,0)'$	0.057	11.074	6	4	0.018
$\boldsymbol{\xi}^{(2)}=\mathbf{0}$	$\mathbf{I}_6$	$(0,1,0,0)'$	0.046	13.849	6	4	0.012
$\boldsymbol{\xi}^{(3)}=\mathbf{0}$	$\mathbf{I}_6$	$(0,0,1,0)'$	0.028	22.996	6	4	0.005
$\boldsymbol{\xi}^{(4)}=\mathbf{0}$	$\mathbf{I}_6$	$(0,0,0,1)'$	0.026	24.865	6	4	0.004
$\boldsymbol{\xi}_{(1)}=\mathbf{0}$	$(1,0,0,0,0,0)$	$\mathbf{I}_4$	0.768	0.075	4	1	0.978
$\boldsymbol{\xi}_{(2)}=\mathbf{0}$	$(0,1,0,0,0,0)$	$\mathbf{I}_4$	0.438	0.320	4	1	0.848
$\boldsymbol{\xi}_{(3)}=\mathbf{0}$	$(0,0,1,0,0,0)$	$\mathbf{I}_4$	0.831	0.051	4	1	0.988
$\boldsymbol{\xi}_{(4)}=\mathbf{0}$	$(0,0,0,1,0,0)$	$\mathbf{I}_4$	0.271	0.671	4	1	0.710
$\boldsymbol{\xi}_{(5)}=\mathbf{0}$	$(0,0,0,0,1,0)$	$\mathbf{I}_4$	0.492	0.258	4	1	0.879
$\boldsymbol{\xi}_{(6)}=\mathbf{0}$	$(0,0,0,0,0,1)$	$\mathbf{I}_4$	0.327	0.514	4	1	0.764

From the matrix for the estimated parameters the estimated growth curves for the four treatments are constructed. For example, for the treatment AF=140 and D=120, the growth curve is given by the equation

$$\begin{aligned}&\begin{pmatrix} 1 & 140 & 120 & 140^2 & 120^2 & 140(120) \end{pmatrix} \widehat{\boldsymbol{\xi}} \begin{pmatrix} 1 \\ t \\ t^2 \\ t^3 \end{pmatrix} \\ &= 99.30 - 10.86t + 0.45t^2 - 0.0043t^3 \end{aligned} \tag{22}$$

For the four treatments, the estimated growth curves are summarized in Table 8 and Figures 5 and 3 (right panel). Figure 5 compares the estimated curve fitting with the observed profiles where we see that the fitted growth curves provide a good fitting for the data.

Table 8: Fitted growth curves for the biomass.

AF	D	Growth curve
60	40	$96.99 - 11.10t + 0.44t^2 - 0.004t^3$
60	120	$-19.37 + 7.10t - 0.17t^2 + 0.001t^3$
140	40	$58.61 - 5.74t + 0.25t^2 - 0.002t^3$
140	120	$99.30 - 10.86t + 0.45t^2 - 0.004t^3$

In the same way, we can find estimates for the response surfaces at each point in time; these equations are derived from product matrix $\widehat{\boldsymbol{\xi}}\mathbf{G}$ and are summarized in Table 9 and in Figure 6. This Figure shows contour plots constructed for the biomass growth at each point in time.

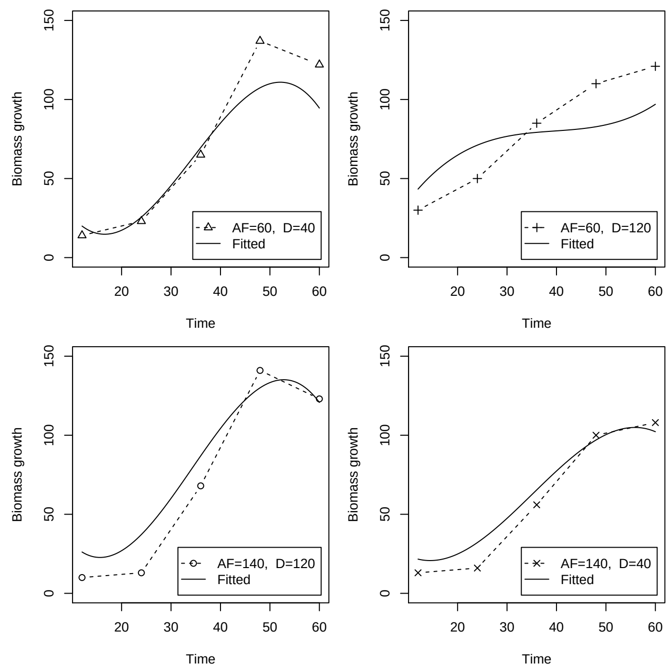

FIGURE 5: Observed and fitted profiles growth curves for the biomass.

It is stressed that the fitted surfaces in the times $t = 12, 24, 48, 60$ capture the behavior concave or convex observed in univariate surface plots (see Figure 4). Moreover, we can conclude from the contour plots that a point located approximately at the coordinate (100, 60) optimizes the process regularly well throughout time, minimizing the fitted surfaces at $t = 12, 24$ and maximizing them at $t = 48, 60$.

TABLE 9: Fitted response surfaces for the biomass

Time	Response surface
t_{12}	$74.75 - 1.30AF - 0.01D + 0.0072AF^2 + 0.0029D^2 - 0.0029AF * D$
t_{24}	$79.20 - 1.44AF + 0.22D + 0.0080AF^2 + 0.0045D^2 - 0.0063AF * D$
t_{36}	$68.88 - 0.10AF + 0.08D - 0.0001AF^2 - 0.0004D^2 + 0.0018AF * D$
t_{48}	$32.99 + 1.59AF + 0.11D - 0.0104AF^2 - 0.0059D^2 + 0.0089AF * D$
t_{60}	$-39.43 + 2.48AF + 0.84D - 0.0124AF^2 - 0.0060D^2 + 0.0025AF * D$

When the model is reparameterized using the expression (17), we obtain the following matrices

$$\mathbf{b} = \begin{pmatrix} 1.4689 & -0.3724 & 0.0131 & -1.104e^{-4} \\ -1.1473 & 0.1479 & -0.0050 & 5.190e^{-5} \end{pmatrix}$$

$$\mathbf{B}^{(0)} = \begin{pmatrix} -0.0146 & 0.0123 \\ 0.0123 & -0.0111 \end{pmatrix} \quad \mathbf{B}^{(1)} = \begin{pmatrix} 0.0029 & -0.0018 \\ -0.0018 & 0.0018 \end{pmatrix}$$

$$\mathbf{B}^{(2)} = \begin{pmatrix} -1.030e^{-4} & 6.371e^{-5} \\ 6.371e^{-5} & -6.445e^{-5} \end{pmatrix} \quad \mathbf{B}^{(3)} = \begin{pmatrix} 9.144e^{-7} & -6.066e^{-7} \\ -6.066^{-7} & 5.798e^{-7} \end{pmatrix}$$

where $\mathbf{b}$ is constructed using the linear effect estimations for the two factors (second and third rows of the estimated parameters matrix, $\widehat{\boldsymbol{\xi}}$). $\mathbf{B}^{(0)}$, $\mathbf{B}^{(1)}$, $\mathbf{B}^{(2)}$ and $\mathbf{B}^{(3)}$ are conformed by the elements of the estimated parameters matrix and kept the reparameterization structure used in the univariate response surface model i.e. the diagonal terms are equivalent to the quadratic effects for each factor, and the off-diagonal elements are equivalent to half of the estimated interaction effects.

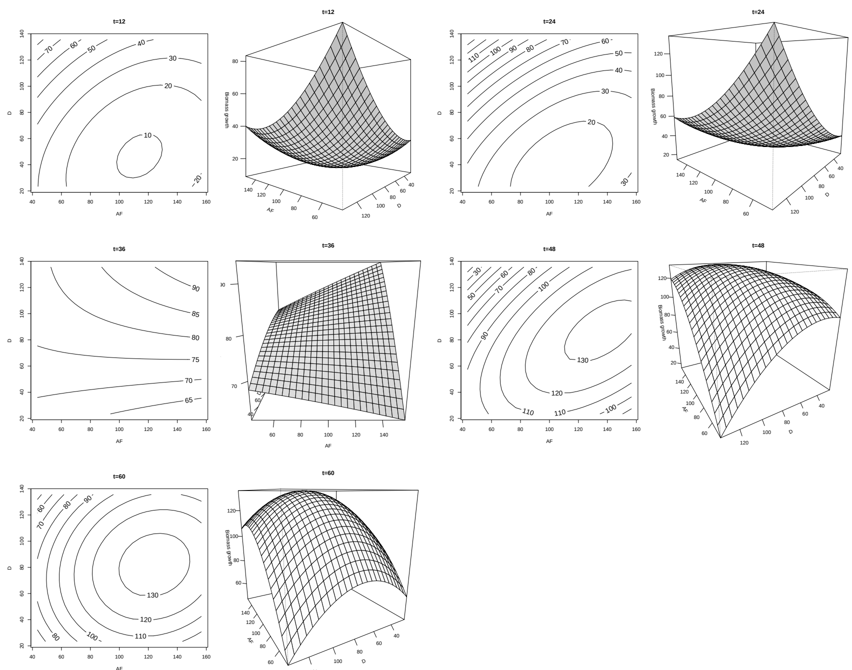

Figure 6: Fitted surface and contour plots for each time in the biomass study.

Thus, following the expression (19), the coordinates for the optimal point that optimizes the process throughout time are found. These are AF= 96.6 and D= 55.3 which are within the observation region of the central composite rotatable design and are in accordance with the behavior seen in the previous contour plots. The confidence bounds for optimum constructed using (20) and $\mathbf{x}_0 = (1, 96.6, 55.3, 96.6^2, 55.3^2, 96.6(55.3))$ are shown in Table 10.

Under the same reasoning used in equation (22), it is possible to construct the growth curve for the optimum found (AF= 96.6 and D= 55.3), which is given by equation $105.2 - 13.7t + 0.53t^2 - 0.005t^3$. This growth curve allows evaluation of the optimization achieved throughout time, minimizing the growth of biomass for times $t = 12, 24$ and maximizing relatively well for time $t = 36, 48, 60$ (see Figure 3, right panel).

Table 10: Parameter estimation and confidence bounds in the optimal point for the biomass.

	t_{12}	t_{24}	t_{36}	t_{48}	t_{60}
Estimated value	9.14	15.16	71.26	125.42	125.62
Lower limit	0.00	0.00	48.84	101.63	101.34
Upper limit	33.42	38.95	93.68	149.21	149.90

6. Conclusions

A joint modelling procedure that gives additional information regarding the interaction of the studied methodologies, as opposed to analyzing them independently, was proposed. In this way, the functional relationship of the response surface parameters with time was modeled by condensing the information of the groups of the usual growth curves analysis. Also, parameter estimation, hypothesis testing, test statistics and confidence bounds were obtained. Finally, under the proposed model, the optimal point that optimizes the response variable regularly well throughout time was found.

In both applications, we studied the optimum combination of factors that optimized our response variable throughout time. Therefore, we fitted a cubic growth curve and a quadratic response surface for the treatments in both situations. In plasma silicon concentration study, it was optimized at a level of dietary ingestion of SZA 1.7% throughout time, so we can say that the plasma silicon concentration has a good growth in horses using this level. In biomass growth, we found that the optimum condition was in the combination of animal fat at a level 96.6 ppm and detergent at a level 55.3 ppm; consequently, using this combination between animal fat and detergent, we optimize this inhibitory behavior during aerobic treatment of waste-water.

Acknowledgments

The authors gratefully acknowledge the comments and suggestions of the anonymous referees that helped to improve the paper immensely. This work was partially supported by Applied Statistics in Experimental Research, Industry and Biotechnology (Universidad Nacional de Colombia).

[Recibido: septiembre de 2011 — Aceptado: mayo de 2013]

References

Box, G. E. P. & Draper, N. R. (1982), 'Measures of lack of fit for response surface designs and predictor variable transformations', *Technometrics* **24**, 1–8.

Box, G. E. P. & Draper, N. R. (2007), *Response Surfaces, Mixtures, and Ridge Analyses*, Wiley Series in Probability and Statistics, New York.

Box, G. E. P. & Wilson, K. B. (1951), 'On the experimental attainment of the optimum conditions', *Journal of the Royal Statistical Society* **13**, 1–45.

Chiou, J., Müller, H. & Wang, J. (2004), 'Functional response models', *Statistica Sinica* **14**, 675–693.

Chiou, J., Müller, H., Wang, J. & Carey, J. (2003), 'A functional multiplicative effects model for longitudinal data, with application to reproductive histories of female medflies', *Statistica Sinica* **13**, 1119–1133.

Davidian, M. (2005), *Applied Longitudinal Data Analysis*, Chapman and hall, North Carolina state university.

Davis, C. S. (2002), *Statistical Methods for the Analysis of Repeated Measurements*, Springer-Verlag, New York.

Draper, N. & Ying, L. H. (1994), 'A note on slope rotatability over all directions', *Journal of Statistical Planning and Inference* **41**, 113–119.

Frey, K. S., Potter, G. D., Odom, T. W., Senor, M. A., Reagan, V. D., Weir, V. H., Elsslander, R. V. T., Webb, M. S., Morris, E. L., Smith, W. B. & Weigand, K. E. (1992), 'Plasma silicon and radiographic bone density on weanling quarter horses fed sodium zelolite A', *Journal of Equine Veterinary Science* **12**, 292–296.

Grizzle, J. E. & Allen, D. M. (1969), 'Analysis of growth and dose response curves', *Biometrics* **25**, 357–381.

Guerrero, S. C. & Melo, O. O. (2008), 'Optimization process of growth curves through univariate analysis', *Revista Colombiana de Estadística* **31**(2), 193–209.

Heck, D. L. (1960), 'Charts of some upper percentage points of the distribution of the largest characteristic root', *The Annals of Mathematical Statistics* **31**(3), 625–642.

Hill, W. J. & Hunter, W. G. (1966), 'A review of response surface methodology: A literature review', *Technometrics* **8**, 571–590.

Kabe, D. G. (1974), 'Generalized Sverdrup's lemma and the treatment of less than full rank regression model', *Canadian Mathematical Bulletin* **17**, 417–419.

Kahm, M., Hasenbrink, G., Lichtenberg-Fraté, H., Ludwig, J. & Kschischo, M. (2010), 'grofit: Fitting biological growth curves with R', *Journal of Statistical Software* **7**, 1–21.

Khatri, C. A. (1966), 'A note on a MANOVA model applied to problems in growth curves', *Annals of the Institute of Statistical Mathematics* **18**, 75–86.

Khatri, C. A. (1973), 'Testing some covariance structures under growth curve model', *Journal Multivariate Analysis* **3**, 102–116.

Khatri, C. A. (1988), 'Robustness study for a linear growth model', *Journal Multivariate Analysis* **24**, 66–87.

Kshirsagar, A. M. & Boyce, S. (1995), *Growth Curves*, Marcel Dekker, New York.

Lucas, J. M. (1976), 'Which response surfaces is best?', *Technometrics* **18**, 411–417.

Magnus, J. R. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley, New York.

Mead, R. & Pike, D. J. (1975), 'A review of responses surface methodology from a biometric viewpoint', *Biometrics* **31**, 830–851.

Molenberghs, G. & Verbeke, G. (2005), *Models for Discrete Longitudinal Data*, Springer, New York.

Montoya, C. & Gallego, D. (2012), Modelo matemático que permita evaluar el cambio de la DBO_5 soluble debido a agentes inhibitorios en un proceso de lodos activados, Master's thesis, Facultad de minas. Universidad Nacional de Colombia.

Pan, J. & Fang, K. (2002), *Growth Curve Models and Statistical Diagnostics*, Springer Series in Statistics, New York.

Potthoff, R. & Roy, S. (1964), 'A generalized multivariate analysis of variance model useful especially for growth curve problems', *Biometrika* **51**, 313–326.

Rao, C. R. (1959), 'Some problems involving linear hypothesis in multivariate analysis', *Biometrika* **46**, 49–58.

Rao, C. R. (1967), Least squares theory using an estimated dispersion matrix and its applications to mesurement of signals, *in* 'Proceeding of the Fifth Berkeley Symposium on Mathematical Statistics and Probability', Vol. I, University of California Press, Berkeley, pp. 355–372.

Singer, J. M. & Andrade, D. F. (1994), 'On the choice of appropiate error terms in profile analysis', *Royal of Statistical Society* **43**(2), 259–266.

Srivastava, M. S. (2002), 'Nested growth curves models', *Sankhyã: The Indian Journal of Statistics, Series A, Selected Articles from San Antonio Conference in Honour of C. R. Rao* **64**(2), 379–408.

Verbyla, A. P. & Venables, W. N. (1988), 'An extension of the growth curve models', *Biometrika* **75**, 129–138.

Wilks, S. S. (1932), 'Certain generalizations in the analysis of variance', *Biometrika* **24**, 471–494.