Article

Optimization of the Culture Medium Composition to Improve the Production of Hyoscyamine in Elicited *Datura stramonium* L. Hairy Roots Using the Response Surface Methodology (RSM)

Ryad Amdoun [1], Lakhdar Khelifi [1],*, Majda Khelifi-Slaoui [1], Samia Amroune [1], Mark Asch [2], Corinne Assaf-Ducrocq [3] and Eric Gontier [3]

[1] Laboratoire des Ressources Génétiques et Biotechnologie, Ecole Nationale Supérieure Agronomique, 16200 El-Harrach-Alger, Algeria; E-Mails: r_amdoun@yahoo.fr (A.R.); m.khelifi@ensa.dz (K.-S.M.); samia_amr@hotmail.fr (A.S.)

[2] Université de Picardie Jules Verne, UFR of Sciences, LAMFA, CNRS-UMR 6140, 33 rue Saint Leu, 80039 Amiens cedex 1, France; E-Mail: mark.asch@u-picardie.fr

[3] Research Unit EA3900 BioPI-UPJV Biologie des plantes et contrôle des insectes ravageurs, Université de Picardie Jules Verne, UFR of Sciences, Ilot des poulies, 33 rue Saint Leu, 80039 Amiens cedex 1, France; E-Mails: corinne.assaf@u-picardie.fr (A.-D.C.); eric.gontier@u-picardie.fr (G.E.)

* Author to whom correspondence should be addressed; E-Mail: khelifi.lakhdar@gmail.com; Fax: +213-21-822-729.

Abstract: Traditionally, optimization in biological analyses has been carried out by monitoring the influence of one factor at a time; this technique is called *one-variable-at-a-time*. The disadvantage of this technique is that it does not include any interactive effects among the variables studied and requires a large number of experiments. Therefore, in recent years, the Response Surface Methodology (RSM) has become the most popular optimization method. It is an effective mathematical and statistical technique which has been widely used in optimization studies with minimal experimental trials where interactive factors may be involved. This present study follows on from our previous work, where RSM was used to optimize the B5 medium composition in $[NO^{3-}]$, $[Ca^{2+}]$ and sucrose to attain the best production of hyoscyamine (HS) from the hairy roots (HRs) of *Datura stramonium* elicited by Jasmonic Acid (JA). The present paper focuses on the use of the RSM in biological studies, such as plant material, to establish a predictive model

with the planning of experiments, analysis of the model, diagnostics and adjustment for the accuracy of the model. With the RSM, only 20 experiments were necessary to determine optimal concentrations. The model could be employed to carry out interpolations and predict the response to elicitation. Applying this model, the optimization of the HS level was 212.7% for the elicited HRs of *Datura stramonium*, cultured in B5-OP medium (optimized), in comparison with elicited HRs cultured in B5 medium (control). The optimal concentrations, under experimental conditions, were determined to be: 79.1 mM [NO^{3-}], 11.4 mM [Ca^{2+}] and 42.9 mg/L of sucrose.

Keywords: *Datura stramonium*; hyoscyamine; hairy root; medium components; optimization; Response Surface Methodology

---

1. Introduction

Hairy roots (HRs) are an efficient system for the production of secondary metabolites [1–4]. However, the performance of this production system depends on the composition of the nutrient medium, which not only affects the content of secondary metabolites [5–11], but also the response to elicitation [12]. The main aim of optimization is to improve the performance of the system by increasing production at a low cost [13,14].

In general, the experimental procedure of optimization is achieved by studying a single factor at any one time. While this factor is modified to find the optimal response, others are kept at a constant level. This is known as the *one-variable-at-a-time-technique*. Clearly, its major disadvantage is that interactions among the factors are not considered so it does not reflect all the potential effects on the process [15]. Another drawback is the large number of experiments needed, requiring additional time and expense. More efficient analytical techniques are based on Response Surface Methodology (RSM) [14]. This was first proposed by Box and his collaborators in 1951 [16] as a method to determine the optimal conditions which maximize or minimize a response. It enables a large amount of data to be obtained from a reduced number of experiments, including the potential interactions between the studied factors [13]. The RSM can be defined as a group of statistical and technical tools used to study the relationship between a response of interest and several input variables. The model has to describe the behavior of a group of data with a view to making statistical predictions. The aim is the simultaneous optimization of several factors to lead to the best performance of a particular system [14].

During recent years, the RSM has been used extensively for optimization in many areas of industrial research and process development in chemistry and biochemistry. Although Calam [17] was the first to suggest the application of RSM in biotechnology, its field was limited [18].

Its use for the optimization and analysis of biotechnological processes with microorganisms and enzymatic engineering has given good results. For example, the work of Maddox and Richert [19] in optimizing bacterial media and Cheynier *et al.* [20] in optimizing enzymatic activity can be cited. However, the application of this analytical technique to systems with plant tissues has been limited.

This study follows the work of Amdoun *et al.* [12] and presents two main aims:

- First, to use the RSM method for the optimization of nutrients (nitrate, calcium and sucrose) in the culture medium B5 [Gamborg,1968] to improve hyoscyamine production in elicited HRs;
- Second, to show the value of the RSM method in the optimization of several responses in plant material cultures.

2. Materials and Methods

2.1. Plant Materials

The HRs, obtained by genetic transformation and selection, were subcultured in the same conditions as described previously [12].

2.2. Elicitation

The use of the desirability function enabled the Jasmonic Acid concentration (JAC) and the exposure time (ET) to be optimized. The optimal JA concentration of 0.06 mM with an exposure time of 24 h was the optimal compromise. The elicitation treatment was initiated 24 h before harvesting the HRs line and was removed for analysis on the 28th day of culture. This length of time was identified as the best stage for elicitation for the highest hyoscyamine content [12].

The solution of Jasmonic Acid (JA) [(−)-jasmonic acid from Sigma-Aldrich] was prepared by dissolving Jasmonic Acid in an adequate volume of ethanol. The solution was filtered through membrane filter (pore size: 0.2 mm Nalgene) and completed with sterile distilled water. The HRs were elicited with this stock solution at 0.06 mM.

The controls were the same HRs lines in the same conditions of culture but without elicitation (ethanol-water solution without Jasmonic Acid). All cultures were conducted in Petri dishes containing 20 mL of B5 medium, in darkness at 26 ± 1 °C.

2.3. Extraction and Hyoscyamine Analysis

Alkaloids were extracted using a method described by Amdoun *et al.* [12]. Hyoscyamine was analyzed using the GC-MS method previously reported by Kartal *et al.* [21].

2.4. Theory of RSM

The response surface methodology (RSM) consists of an adjustment of empirical models to the data obtained experimentally. Linear (1st degree) or quadratic (2nd degree) mathematical models are employed to describe the system to be optimized [22]. When several factors influence a particular system, it is impossible to screen and to control the contribution of each factor so only those factors which have a major influence must be considered. The screening design, such as the factorial designs 2^k, can be used to meet this objective. They are efficient and economical [15].

To evaluate the form of the true response, a second degree model is used. The factorial designs 2^k are used to determine the first-order effects but these are inefficient when additional effects, like the second-order effects, are so important. Experimental central points in the factorial design 2^k can be added to evaluate the shape. The polynomial function must contain other factors, which include the

interaction between the various experimental variables. To determine a critical point (maximum, minimum or saddle), the polynomial function of second degree must contain quadratic factors.

The model of second degree is given by equation (1), where Y is the response, α_0 is the intercept of the y axis, α_j, α_{jj},..., α_{jl} are the various coefficients of the model (linear and quadratic), X_j and X_l are the independent variables (factors) and ε is the error of model with $\forall i$, $V(\varepsilon_i) = \sigma^2$ (homoscedasticity) and $\varepsilon \sim N(0, \sigma^2)$ (normal distribution).

$$Y = \alpha_0 + \sum_{j=1}^{k} \alpha_j X_j + \sum_{j=1}^{k} a_{jj} X_j^2 + \sum_{j<1} \sum_{l=2}^{k} \alpha_{jl} X_j X_l + \varepsilon \tag{1}$$

The equation system (1) can be resolved by the Method of Least Squares (MLS) and may be written in matrix form (2) [23], where y is the $(n, 1)$ vector of measured responses, X is the (n, p) matrix of the mathematical model of factor levels, and α is the $(p, 1)$ vector of unknown coefficients.

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} 1 & x_{11} & . & . & . & x_{1k} \\ 1 & x_{21} & . & . & . & x_{2k} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & x_{n1} & x_{n2} & . & . & x_{nk} \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ . \\ . \\ . \\ \alpha_k \end{bmatrix}}_{\alpha} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix}}_{\varepsilon} \tag{2}$$

With n responses, and y and p coefficients in the model, there are n equations and $n + p$ unknowns. To solve the system, the Method of Least Squares is used. After some calculations, the solution of system (2) is given by Equation (3):

$$\hat{a} = (^tXX)^{-1}\,{}^tXY \tag{3}$$

The low and high levels of variables (factors) are denoted by -1 and 1, respectively. The levels of X_i variable are coded and obtained from Equation (4), where X_i is the independent variable coded value, A_i is the independent variable real value, A_0 is the independent variable real value on the center point and ΔA_i is the step change value.

$$X_i = \left(A_i - A_0\right)/\Delta A_i \tag{4}$$

The most used second-order symmetric designs for the RSM are: The general factorial design, the Box-Behnken design, the Central Composite Design and Doehlert's design. These differ by the location of the experimental points in the studied region, the number of factor levels kept, the number of experiments and the blocks [14].

2.5. Global Predicted Capacity, Analysis and Diagnostic of the Model

When the relationship between the variables and the response has been established by the modeling, predictions can be made. However, the mathematical model obtained after adjustments to experimental results, sometimes cannot describe the studied domain. It is necessary to analyze and examine the diagnostic of the obtained model to evaluate its pertinence to describe the studied phenomena. If the

analysis and the diagnostic are satisfactory, the defined model can be used in predictions, but only if the conditions are identical and the standard error is present.

The global predicted capacity of a mathematical model is generally explained by the coefficient of determination R^2 A large value of this coefficient does not necessarily suggest that the model is a good one. The R^2 value improves if the variables are high even if they are not statistically significant [24]. Because R^2 alone is not a measure of the model's accuracy, it is also necessary to take into account the value of the Absolute Average Deviation (AAD) [13]. The AAD is a direct method which describes deviations, so its value must be as small as possible between the measured and the predicted data [13].

During the analysis, the statistical global significance of the model is determined by variance analysis (ANOVA, Fisher's Test: *F*-value). The lack of fit test is also applied to estimate the model (fitting). The statistical significance of the model terms is determined by the Student test (*t*-value). The model is adjusted to the experimental data when it is significant and when the lack of fit value of the model is not significant. The truth of the hypothesis related to the homoscedasticity and the normal distribution of variations is very important during the diagnostic. The visual analysis of graphical plots of deviations gives some information about the robustness of the model. Concerning the influential observations on the predictions of the model, the main statistics are the leverage, Cook's distance and the DFFITS.

2.6. Determination of the Optimum

Depending on the object, the optimal point can be characterized as the maximum, the minimum or the saddle. The value of the maximal point can be calculated from the first derivative of the model's equation. The positive values of the explanatory variables, where the first derivative is equal to zero, correspond to the optimal values. The accuracy of the values can be verified by comparing the predicted values obtained with the mathematical model, and the measured values obtained after the experiments with the same conditions.

2.7. Application of the RSM for Optimization of the Culture Medium B5

The aim of this paper is to optimize the composition of the medium, which influences the response to elicitation [12,25,26]. This work succeeds a precedent paper [12], where we applied the screening analysis to the influence of minerals on the hyoscyamine content of elicited HRs.

As has been reported in [12], nitrate $[NO_3^-]$, calcium $[Ca^{2+}]$, the combination $[NO_3^- \times Ca^{2+}]$ and phosphorus $[H_2PO_4^-]$, are the most significant factors in the intensity of the response to elicitation for hyoscyamine production by HRs. Although phosphorus $[H_2PO_4^-]$ is statistically efficient, only $[NO_3^-]$ and $[Ca^{2+}]$ are used for optimization by RSM. Sucrose is also included for optimization due to its effect on biomass and alkaloid production [6,8]. The minimal and maximal values of nitrate and calcium have been selected from the precedent results of Amdoun *et al.* [12]. The real values of the concentrations have been codified by the Equation (4) (see Section 2.4).
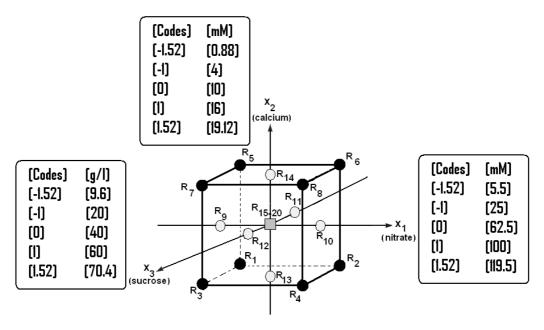
The Central Composite Design (CCD) is used to obtain the measured responses which will be useful for the mathematical model. The CCD was presented by Box and Wilson in the 1950s. It consists of the following two parts:

- A factorial design with at least one experimental point located in the center of the experimental area;
- A star design whose axial points $-\alpha$ and $+\alpha$ are located on the axis of each factor. This design is particularly adapted to the progressive acquisition of results with a factorial design 2^k.

It is necessary to carry out the experiments corresponding to the star design's points and to calculate whether the results are explained by the linear model. Figure 1 shows the studied experimental region, the range and the levels of the studied variables. Five levels have been considered for each of the three studied factors (NO_3^-, Ca^{2+} and sucrose). So, the application of CCD consists of carrying out 20 experiments combining three factors, six of which are located in the center of the studied region. The optimal criterion is orthogonality, in this case α calculated is equal to ± 1.52 [27]. Each point ($R_1...R_{20}$) corresponds to one experiment. The points R_1 to R_8 represent the factorial design 2^k. The points R_9 to R_{14} are the star design. The points R_{15} to R_{20} represent the experiment done in the central experimental design.

Figure 1. Experimental region and levels of each of the three factors of the CCD (X_1: [NO_3^-], X_2: [Ca^{2+}] and X_3: sucrose).

For all experiments, the HRs cultures were performed in Petri dishes containing 20 mL of B5 medium [28]. In order to avoid other modifications of the medium due to the addition of other counterions, each element was chosen and then HNO_3 for [NO_3^-] and $Ca(OH)_2$ for [Ca^{2+}] were added to the medium before adjustment of the pH to 5.8 [29]. The quantity of inoculum for each dish was 0.3 g fresh weight from the root tips of HRs of *Datura stramonium*. The measured response was the hyoscyamine level (mg/L). The averages were calculated from the three replicates. The control corresponds to the same HRs line; in the same B5 medium containing 25 mM [NO_3^-], 1.0 mM [Ca^{2+}] and 3% sucrose; and the same conditions of temperature and darkness. Dry weight was obtained by oven drying the HRs for 48 h at 40 °C. The final dry biomass is expressed by the average of the three values with a precision balance.

3. Results and Discussion

3.1. Global Predictivity of the Model

The coefficient of determination R^2 measures the variability explained by the factors and their interactions in the observed responses [30]. It is 0.97 for the model, from which it can be concluded that 97% of the HS level of elicited HRs is attributed to independent variables and only 3% of the total variability is not explained by the model. The AAD value is low (1.2%) and the value of deviation is $\sigma = 0.9$. Thus, the model explains the global variability; it is globally predictive.

3.2. Analysis of the Quadratic Model

The *F*-value and the value of probability show that the model is statistically significant. The lack of fit test is not significant (Table 1), which indicates that the model fits well with the experimental data.

Table 1. ANOVA table for the quadratic model (results in bold are significant).

Source	Sum of Squares	df	Mean Square	*F*-value	*p*-value
Model	9537.1	10	953.7	414.1	**0.0**
Residual	20.7	9	2.3		
Lack of fit	16.1	4	4.0	4.3	0.1
Pure Error	4.7	5	0.9		
Total	9557.8	19			

All the linear terms related to $[NO_3^-]$, $[Ca^{2+}]$, $[NO_3^-/Ca^{2+}]$ interaction and the sucrose effect are significantly positive (Table 2). The linear effect of the latter is almost identical to that of $[Ca^{2+}]$. As the concentrations of these three elements increase in the B5 medium, the response to elicitation of HRs becomes more significant up to a certain limit where the response is reduced.

Table 2. Analysis terms for the quadratic model (results in bold are significant).

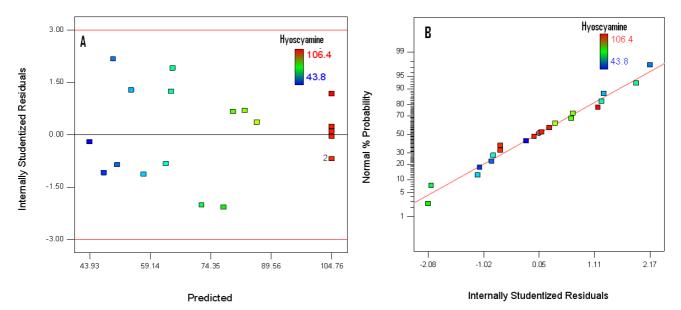| Model Terms	Coefficient Estimate	t-statistic	*p*-value
Intercept	104.8	172.2	-
x_1: nitrate	10.5	24.6	**0.0**
x_2: calcium	5.5	12.9	**0.0**
x_3: sucrose	4.2	9.8	**0.0**
$x_1 x_2$	3.5	6.5	**0.0**
$x_1 x_3$	1.0	1.9	0.1
$x_2 x_3$	1.0	1.9	0.1
x_1^2	−16.4	−35.4	**0.0**
x_2^2	−14.4	−31.1	**0.0**
x_3^2	−14.5	−31.3	**0.0**
$x_1 x_2 x_3$	0.7	1.3	0.2

The beneficial effect of $[NO_3^-]$ is twofold: it improves the biomass [7,8,12] and the HS production by HRs simultaneously [8,12]. It is one of the crucial elements of the medium to improve the level of HS in elicited HRs. Effectively, its deficiency during the first week of culture, negatively affects the

response to elicitation of HRs [12]. Moreover, its excess in the medium (from 95 mM) shows a negative effect on the biomass [8,12,31]. The $[Ca^{2+}]$ is in second position for increasing the HS level after elicitation [12]. The $[Ca^{2+}]$ ion activates Putrescine Methyl Transferase (PMT), which is one of the enzymes involved in the biosynthesis of tropane alkaloids [5,9]. It is a secondary messenger and participates in signal transduction and cellular regulation. Its intracellular flux is activated by a stimulus (such as elicitation) and also induces the defense reactions [26,32]. For sucrose, Saenz-Carbonell and Loyola-Vargas [6] showed that it improves the specific production of hyoscyamine. In our case, sucrose particularly increases the biomass.

3.3. Diagnostic of the Model

3.3.1. Diagnostic Plot

The validity of a model can be evaluated by the residual analysis. We defined as residual, the difference between the observed values and the calculated values obtained by the model. It is the part which is not explained by the equation of the model. This analysis can also detect some outliers among the total data. We use principally graphical methods for the residual analysis, such as the graphical presentation of the residuals as a function of the estimated values [33]. This is essential to check the homoscedasticity hypothesis of the errors. If the selected model is adequate, the residuals are distributed uniformly on a horizontal band of the graph. The analysis of Figure 2A shows a random distribution of the residuals as a function of the predicted values, so the homoscedasticity hypothesis is verified. Another crucial hypothesis on the residual is the normal distribution. The graphical representation of the residuals is an important tool for diagnostics [34,35]. Figure 2B shows a normal distribution of the Studentized residuals, which are independent of each other.

Figure 2. Diagnostic plot: (**A**) Studentized residuals *versus* predicted values; (**B**) normal probability plot of the studentized residuals.

3.3.2. Influential Observations and Accommodation

If the value of an observation diverges from the supposed form of the distribution of the total observations, it is called an outlier. If this form is modified, the observation can converge with the new model [33]. After the detection of outlier observations, the accommodation can be done in different ways, for example by complementary data, re-specifying the model or removing observations [36]. An outlier value is called an influential observation if its presence in the data affects the estimated coefficients of the model [36]. The classical measurement of the influential observations on the predictions of a model are the leverage h_i, Cook's distance and DFFITS (Table 3).

Table 3. Diagnostics for influential observations (Results in bold are outliers).

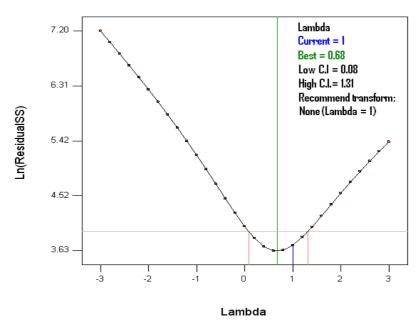| Run Order (*n*)	Variable Code Levels			Measured	Predicted	Residual	h_i Leverage	Cook's Distance	DFFITS
	x_1 [NO_3^-]	x_2 [Ca^{2+}]	x_3 [sucrose]						
Model with Full Terms									
R_1	−1	−1	−1	43.8	43.9	−0.1	0.8	0.0	−0.4
R_2	1	−1	−1	56.8	57.5	−0.7	0.8	0.5	**−2.5**
R_3	−1	1	−1	46.8	47.5	−0.7	0.8	0.5	**−2.4**
R_4	1	1	−1	70.8	72.0	−1.3	0.8	**1.7**	**−5.6**
R_5	−1	−1	1	51.2	49.8	1.4	0.8	**2.0**	**6.5**
R_6	1	−1	1	65.2	64.4	0.8	0.8	0.7	**2.8**
R_7	−1	1	1	55.2	54.4	0.8	0.8	0.7	**2.9**
R_8	1	1	1	86.2	86.0	0.2	0.8	0.1	0.7
R_9	−1.52	0	0	50.0	50.9	−0.9	0.6	0.1	**−1.0**
R_{10}	1.52	0	0	83.6	83.0	0.7	0.6	0.1	0.8
R_{11}	0	−1.52	0	62.3	63.1	−0.8	0.6	0.1	−0.9
R_{12}	0	1.52	0	80.6	79.9	0.7	0.6	0.1	0.7
R_{13}	0	0	−1.52	66.7	64.8	1.8	0.6	0.4	**2.7**
R_{14}	0	0	1.52	75.5	77.5	−2.0	0.6	0.5	**−3.1**
R_{15}	0	0	0	103.8	104. 8	−1.0	0.2	0.0	−0.3
R_{16}	0	0	0	104.7	104.8	−0.0	0.2	0.0	0.0
R_{17}	0	0	0	105.1	104.8	0.3	0.2	0.0	0.1
R_{18}	0	0	0	106.4	104.8	1.6	0.2	0.0	0.5
R_{19}	0	0	0	103.8	104.8	−1.0	0.2	0.0	−0.3
R_{20}	0	0	0	104.9	104.8	0.1	0.2	0.0	0.0
Model with Only Significant Terms (Equation 5)									
R_4	1	1	−1	70.8	74.8	−4.0	0.6	0.9	**−4.7**
R_8	1	1	1	86.2	83.2	2.3	0.5	0.5	**2.4**
R_{14}	0	0	1.52	75.5	77.5	−2.0	0.6	0.5	**−2.2**

We called the leverage h_i, for an observation i as the estimation of the ith variable response which is influenced by the value of the corresponding independent variable X_i; h_i is between 0 to 1 [33]. Table 3 does not show extreme values concerning h_i for the global observations. Nevertheless, the values of the observations R_1 to R_8 are close to 1. The h_i does not take into consideration the residual of the

observation so, to complete the diagnostic, Cook's distance and the DFFITS are examined. The latter are close to leverage and residuals. Cook's distance of an observation is a measurement of the observation's influence on the total predictions of the model. The value of Cook's distance must be <1 [37]. The observations R_4 and R_5 show a main influence on the global predictions of the model (Table 3). The DFFITS measurement of an observation is the measurement of this observation's influence on its value's predictivity by the model. The DFFITS value could be less than 1 [38]. Table 3 shows that the experiments R_2 to R_7, R_9, R_{13} and R_{14} have crucial influences. From this diagnostic, we expect deviations if we are looking at the predictions in the experimental domain near these points. To improve the model's accuracy, the first operation is to select the terms which are most significant. In this case, the model can be written as:

$$\hat{Y}_{HS} = 104.7 + 10.5X_1 + 5.5X_2 + 4.2X_3 + 3.5X_1X_2 - 16.4X_1^2 - 14.4X_2^2 - 14.5X_3^2 \tag{5}$$

After this operation, only the observations R_4, R_8 and R_{14} presented main DFFITS values (Table 3). The data needed no transformation (Y^λ), as confirmed by the Box-Cox Plot exam (Figure 3). We decided to adjust the outliers, which originate from measurement errors, by their removal. After the removal of the R_4 observation, which has the highest DFFITS value, only the R_{14} observation shows an extreme value. After this elimination, the diagnostic of influential observations reveals no outliers. We can observe that the normal distribution of residuals and their independence, the homoscedasticity and the statistical significance of the model are still verified.

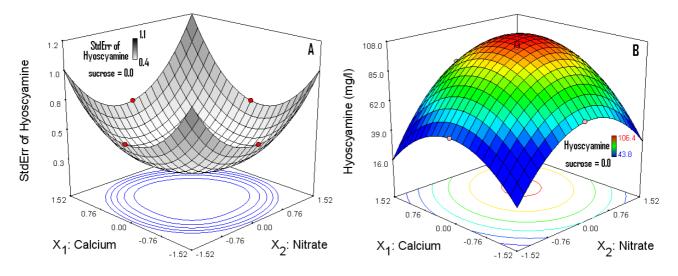**Figure 3.** Box-Cox plot for power transformations (Y^λ).

The removal of the observations R_4 and R_{14} improves the accuracy of the estimated coefficients. The CV% goes from 2.0 to 0.9, in this case the AAD value was calculated as 0.4%. More precisely, it is the model $\hat{Y}_{HS}$ (Equation 6) which has been selected for the response surface methodology analysis and the determination of optimal concentrations.

$$\hat{Y}_{HS} = 104.7 + 11.0X_1 + 6.0X_2 + 4.1X_3 + 4.3X_1X_2 - 16.4X_1^2 - 14.4X_2^2 - 13.7X_3^2 \tag{6}$$

3.5. Response Surface (RS) Analysis and Determination of Optimal Concentrations

The graphical representation of the standard error (StdErr) function (Figure 4A) is symmetric and its highest value (>1) is obtained in the four corners of the experimental area. Nevertheless, its value is 0.3 inside the central area. This low value is related to a good model quality, so the model Equation 6 can be used to evaluate the values of the $\hat{Y}_{HS}$ response in the global experimental domain.

Figure 4. Response Surface plots showing effects of $[NO_3^-]$ and $[Ca^{2+}]$ as a function of standard error (StdErr) (**A**) and the HS level for elicited HRs (**B**).

The RS is the graphical representation of the mathematical model in the experimental domain. Figure 4B shows the response surface for the HS level from elicited HRs as a function of $[NO_3^-]$ (X_1) and $[Ca^{2+}]$ (X_2). The sucrose value has been fixed to its coded value 0.0 (equivalent to 40 mg/L). From this figure, for different combinations of $[NO_3^-]$ and $[Ca^{2+}]$, the responses can be predicted. This is an effective tool to see the simultaneous effect of $[NO_3^-]$ and $[Ca^{2+}]$ on the HS level from elicited HRs. The improvement in the HS level is correlated with the increasing $[NO_3^-, Ca^{2+}]$ concentrations until limits where the response decreases, beyond the central zone of the experimental domain (Figure 4B). The beneficial effect of the combination has been described by Amdoun *et al.* [12]. In the case of a low response with low $[NO_3^-, Ca^{2+}]$ values, the cells are in a critical condition for plant defense reactions [12]. The high $[NO_3^-, Ca^{2+}]$ values have a negative impact on the biomass and also on the HS level. Our results are identical with those obtained by Sikuli and Demeyer [7] for the biomass. These authors found low values of biomass and HS for HRs of *Datura stramonium* cultivated in B5 medium with high $[NO_3^-, Ca^{2+}]$ concentrations. This effect is greater when the $[Ca^{2+}]$ value is high. We also note that the nitrate reductase activity measured in the HRs is high in this medium [7].

All the combinations $[NO_3^-, Ca^{2+}]$ located in the red area (Figure 4B) illustrate the optimal conditions. Nevertheless, the local optimum concentrations correspond to positive values of variables where the first derivatives ($\delta Y_{HS}/x_1$; $\delta Y_{HS}/x_2$; $\delta Y_{HS}/x_3$) of the model Equation 6 cancel each other out. The optimal concentrations in coded values are 0.37 for $[NO_3^-]$, 0.26 for $[Ca^{2+}]$ and 0.14 for sucrose. This is equivalent to 79.1 mM, 11.4 mM and 42.9 g/L respectively in real values and corresponds to the optimized B5 medium (B5-OP). These optimal calculated concentrations are located in the area
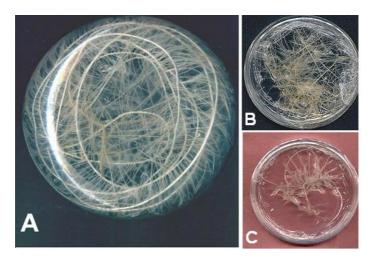
where the standard error function is lower and where the predicted accuracy of the model Equation 6 is confirmed.

According to the simulation of the HS level from elicited HRs, cultured in B5-OP medium, the predicted value is 107.90 mg/L. The calculated prediction interval, at 95%, is *PI* = [106.17; 109.72]. The measured level for the same concentrations and in the same conditions is 110.3 ± 1.4 mg/L. This value is in the 95% *PI* range and validates the optimal concentrations of $[NO_3^-]$, $[Ca^{2+}]$ and sucrose.

3.6. Optimization Level

Figures 5A and 5B show a culture of HRs in B5-OP medium in a Petri dish or Erlenmeyer. It displays a high biomass in comparison with HRs cultivated in B5 control (Figure 5C).

Figure 5. Appearance of HRs on the 28th day of culture in B5-OP (**A**) in 250 mL Erlenmeyer; (**B**) in Petri dishes and in B5 control (**C**) in Petri dishes.

There is a significant improvement in biomass (51.2%) with the B5-OP medium in comparison with the HRs control (Table 4). A significant difference was observed for the specific production of HS (mg/g DW) in HRs with or without elicitation. The optimization was determined as 81% and 101.2% for elicited or non-elicited HRs respectively in B5-OP medium.

Table 4. Biomass and HS production of HRs cultivated in B5 control medium or B5-OP medium after the 28th day of culture [*****: B5 control (25 mM NO_3^-, 1.0 mM Ca^{2+} and 3% sucrose); ******: B5-OP optimized (79.1 mM NO_3^-, 11.4 mM Ca^{2+} and 42.9% sucrose)].

			Production of HS			
		Biomass (g DW/L)	(mg/g DW)		(mg/L)	
			Without Elicitation	With Elicitation	Without Elicitation	With Elicitation
B5 *		8.4 ± 0.6	2.1 ± 0.1	4.2 ± 0.6	17.6 ± 1.6	35.3 ± 2.0
B5-OP **		12.7 ± 0.2	3.8 ± 0.1	8.5 ± 0.3	48.3 ± 2.3	110.3 ± 1.4
Optimization		51.2%	81%	101.2%	173.6%	212.7%
LSD test	differences	−4.3	−1.7	−4.3	−30.6	−75.0
	±limits	0.9	0.2	0.8	4.4	13.8
	significance	significant	significant	significant	significant	significant

Finally, the improvement in the level of HS (mg/L) was 173.6% for non-elicited HRs, and 212% for elicited HRs cultivated in B5-OP, in comparison with the control medium.

4. Conclusion

The selection of significant statistical variables enables the accuracy of the model to be improved. After the diagnostics, some outliers related to the model with the global terms are eliminated after the screening. This selection of outliers must be carried out carefully; their removal is implemented when they are due to pure error. In our case, only two observations were removed after the study of the screening of variables.

Optimization of the B5 medium improves the response to elicitation. Under optimal conditions, the optimization of the HS level is 212.7% for elicited *Datura stramonium* HRs cultivated in B5-OP medium, in comparison with HRs in the B5 control.

The optimal concentrations for the selected line and in our conditions are: 79.1 mM, 11.4 mM and 42.9 g/L for $[NO_3^-]$, $[Ca^{2+}]$ and sucrose, respectively. These values correspond to the solutions obtained with the first derivatives of the model. Nevertheless, all the concentrations included in the optimal area of the surface response can be used.

This new approach to the optimization of B5 medium could be employed to maximize the response to elicitation. The B5-OP medium significantly influenced the biomass and alkaloid production. This was possible by the simulation and the predictability of the quadratic model used. Indeed, unlike classical methods, this mathematical model is not time consuming and a large number of experiments are not needed to optimize the parameters. Mathematical modeling is a powerful tool for biological studies. In this work, by applying the RSM, only 20 experiments were required to optimize the B5 medium composition in terms of $[NO_3^-]$, $[Ca^{2+}]$ and sucrose. This quadratic model led to a demonstration of the effects of these nutrients on the HS level of HRs elicited by Jasmonic Acid (JA).

However, the HS level of the elicited HRs, cultured in B5-OP medium, can still be improved by an irregular deficiency in $[NO_3^-]$ [12]. The feasibility needs to be confirmed and the time and the frequency need to be determined. For production in a bioreactor, it is clear that the optimal concentrations must be added according to the B5 medium volume and the biomass.

References

1. Bourgaud, F.; Bouque, V.; Gontier, E.; Guckert, A. Hairy root cultures for the production of secondary metabolites. *AgBiotech News Inf.* **1997**, *9*, 205–208.
2. Shanks, J.V.; Morgan, J. Plant 'hairy root' culture. *Curr. Opin. Biotechnol.* **1999**, *10*, 151–155.
3. Bourgaud, F.; Gravot, A.; Milesi, S.; Gontier, E. Production of plant secondary metabolites: a historical perspective. *Plant Sci.* **2001**, *161*, 839–851.
4. Raoa, S.R.; Ravishankarb, G.A. Plant cell cultures: Chemical factories of secondary metabolites. *Biotechnol. Adv.* **2002**, *20*, 101–153.
5. Gontier, E.; Sangwan, B.S.; Barbotin, J.N. Effects of calcium, alginate, and calcium-alginate immobilization on growth and tropane alkaloid levels of stable suspension cell line of *Datura innoxia* Mill. *Plant Cell Rep.* **1994**, *9*, 533–536.

6. Saenz-Carbonell, L.; Loyola-Vargas, V.M. *Datura stramonium* hairy roots tropane alkaloid content as a response to changes in Gamborg's B_5 medium. *Appl. Biochem. Biotechnol.* **1996**, *3*, 321–337.

7. Sikuli, N.N.; Demeyer, K. Influence of the ion-composition of the medium on alkaloid production by hairy roots of *Datura stramonium*. *Plant Cell Tissue Organ Culture* **1997**, *47*, 261–267.

8. Nussbaumer, P.; Kapétanidis, I.; Christen, P. Hairy root of *Datura candida × D. aurea*: effect of culture medium composition on growth and biosynthesis. *Plant Cell Rep.* **1998**, *17*, 405–409.

9. Piñol, M.T.; Palazón, J.; Cusidó, R.M.; Ribó, M. Influence of calcium ion-concentration in the medium on tropane alkaloid accumulation in *Datura stramonium* hairy roots. *Plant Sci.* **1999**, *141*, 41–49.

10. Boitel-Conti, M.; Laberche, J.C.; Lanoue, A.; Ducrocq, C.; Sangwan-Norreel, B.S. Influence of feeding precursors on tropane alkaloid production during an abiotic stress in *Datura innoxia* transformed roots. *Plant Cell Tissue Organ Cultrue* **2000**, *60*, 131–137.

11. Lanoue, A.; Boitel-Conti, M.; Dechaux, C.; Laberche, J.C.; Christen, P.; Sangwan-Norreel, B.S. Comparison of growth properties, alkaloid production and water uptake of two selected *Datura* hairy root lines. *Acta Biol. Cracoviensia Serie Bot.* **2004**, *46*, 185–192.

12. Amdoun, R.; Khelifi, L.; Khelifi-Slaoui, M.; Amroune, S.; Benyoussef, E.-H.; Thi, D.V.; Assaf-Ducrocq, C.; Gontier, E. Influence of minerals and elicitation on *Datura stramonium* L. tropane alkaloid production: modelization of the *in vitro* biochemical response. *Plant Sci.* **2009**, *177*, 81–87.

13. Baş, D.; Boyacı, I.H. Modeling and optimization I: Usability of response surface methodology. *J. Food Eng.* **2007**, *78*, 836–845.

14. Bezerra, M.A.; Santelli, R.E.; Oliveira, E.P.; Villar, L.S.; Escaleira, L.A. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta* **2008**, *76*, 965–977.

15. Lundstedt, T.; Seifert, E.; Abramo, L.; Thelin, B.; Nyström, Å.; Pettersen, J.; Bergman, R. Experimental design and optimization. *Chemom. Intell. Lab. Syst.* **1998**, *42*, 3–40.

16. Box, G.E.P.; Wilson, K.B. On the experimental attainment of optimum conditions. *J. R. Stat. Soc. B* **1951**, *13*, 1–45.

17. Calam, C.T. Media for industrial fermentations. *Process Biochem.* **1967**, *2*, 19–22.

18. Saval, S.; Pablos, L.; Sanchez, S. Optimization of culture medium for streptomycin production using response surface methodology. *Bioresour. Technol.* **1993**, *43*, 19–25.

19. Maddox, I.S.; Richert, S.H. Use of response surface methodology for the rapid optimization of microbiological media. *J. Appl. Bacteriol.* **1977**, *43*, 197–204.

20. Cheynier, V.; Feinberg, M.; Chararas, C.; Ducauze, C. Application of Response Surface Methodology to Evaluation of Bioconversion Experimental Conditions. *Appl. Environ. Microbiol.* **1983**, *45*, 634–639.

21. Kartal, M.; Kurucu, S.; Altun, L.; Ceyhan, T.; Sayar, E.; Cevheroğlu, S.; Yetkin, Y. Quantitative Analysis of l-Hyoscyamine in *Hyoscyamus reticulatus* L. by GC-MS. *Turk. J. Chem.* **2003**, *27*, 565–569.

22. Teófilo, R.F.; Ferreira, M.M.C. Chemometrics II: spreadsheets for experimental design calculations, a tutorial. *Quim. Nova* **2006**, *29*, 338–350.

23. Goupy, J.; Creighton, L. *Introduction aux Plans D'éxpériences*; DUNOD: Paris, France, 2006.

24. Myers, R.H.; Montgomery, D.C. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*; John Wiley & Sons, Inc.: New York, NY, USA, 2009.

25. Dunlop, D.S.; Curtis, W.R. Synergistic response of plant hairy-root cultures to phosphate limitation and fungal elicitation. *Biotechnol. Prog.* **1991**, *7*, 434–438.

26. Vasconsuelo, A.; Boland, R. Molecular aspects of the early stages of elicitation of secondary metabolites in plants. *Plant Sci.* **2007**, *172*, 861–875.

27. Myers, R.H. *Response Surface Methodology*; Allyn and Bacon, Inc.: Boston, MA, USA, 1971.

28. Gamborg, O.L.; Miller, R.A.; Ojima, K. Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell Res.* **1968**, *50*, 151–158.

29. Dechaux, C.; Boitel-Conti, M. A strategy for overaccumulation of scopolamine in *Datura innoxia* hairy root cultures. *Acta Biol. Cracoviensia Serie Bot.* **2005**, *47*, 101–107.

30. Haaland, P.D. *Experimental Design in Biotechnology*; Marcel Dekker: New York, NY, USA, 1989; pp. 1–18.

31. Chashmi, N.A.; Sharifi, M.; Karimi, F.; Rahnama, H. Enhanced production of tropane alkaloids by nitrate treatment in hairy root cultures of *Atropa belladonna*. *J. Biotechnol.* **2008**, *136*, 22–71.

32. Zhao, J.; Davis, L.C.; Verpoorte, R. Elicitor signal transduction leading to production of plant secondary metabolites. *Biotechnol. Adv.* **2005**, *23*, 283–333.

33. Dodge, Y. *Statistique: Dictionnaire Encyclopédique*; DUNOD: Paris, France; 1993.

34. Draper N.R.; Smith, H. *Applied Regression Analysis*, 2nd ed.; John Wiley & Sons Publication: New York, NY, USA, 1981.

35. Montgomery, D.C. *Design and Analysis of Experiments*; John Wiley & Sons, Inc.: New York, NY, USA, 2008.

36. Mason, R.L.; Gunst, R.F.; Hess, J.L. *Statistical Design and Analysis of Experiments with Applications to Engineering and Science*, 2nd ed.; John Wiley & Sons, Inc.: New York, NY, USA, 2003.

37. Cook, R.D. Detection of influential observation in linear regression. *Technometrics* **1977**, *19*, 15–18. Available online: http://www.ime.usp.br/~abe/lista/pdfWiH1zqnMHo.pdf (accessed on 17 November 2010).

38. Belsley, D.A.; Kuh, E.; Welsch, R.E. *Regression Diagnostics—Identifying Influential Data and Sources of Collinearity*; John Wiley & Sons, Inc.: New York, NY, USA, 2004.