

Article

Coagulation: Determination of Key Operating Parameters by Multi-Response Surface Methodology Using Desirability Functions

Marina Corral Bobadilla ^{1,*}, Rubén Lostado Lorza ¹, Rubén Escribano García ²,
Fátima Somovilla Gómez ¹ and Eliseo P. Vergara González ³

¹ Department of Mechanical Engineering, University of La Rioja, Logroño 26004, La Rioja, Spain; ruben.lostado@unirioja.es (R.L.L.); fatima.somovilla@unirioja.es (F.S.G.)

² IK4-LORTEK, 20240 Ordizia, Guipuzcoa, Spain; escribano.engineer@gmail.com

³ Department of Mining Exploitation and Prospecting, University of Oviedo, 33004 Principality of Asturias, Spain; vergaraeliseo@uniovi.es

* Correspondence: marina.corral@unirioja.es; Tel.: +34-941-299-274

Received: 31 October 2018; Accepted: 21 February 2019; Published: 25 February 2019



Abstract: The clarification process removes colloidal particles that are suspended in waste water. The efficiency of this process is influenced by a series of inputs or parameters of the coagulation process, of which the most commonly used are initial turbidity, natural coagulant dosage, temperature, mixing speed and mixing time. The estimation of the natural coagulant dosage that is required to effectively remove these total suspended solids is usually determined by a jar test. This test seeks to achieve the highest efficiency of removal of the total suspended solids while reducing the final turbidity of waste water. This is often configured in iterative fashion, and requires significant experimentation and coagulant. This paper seeks to identify regression models that relate the clarification process parameters to the process outputs (final turbidity and total suspended solid) by the Response Surface Methodology (RSM) based on experiments of Central Composite Design (CCD) of experiments that involve three emerging natural coagulants. Several clarification process scenarios also were proposed and demonstrated using the Multi-Response Surface (MRS) with desirability functions. The experimental results were found to be in close agreement to what are provided by the regression models. This validates the use of the MRS-based methodology to achieve satisfactory predictions after minimal experimentation.

Keywords: clarification process; coagulation; waste water treatment; Multi-Response Surface methodology

1. Introduction

The management of waste water impacts directly aquatic ecosystems' biological diversity and affects significantly our life support systems. The management of waste water is an important component of the ecosystem and functions in all sectors. These include fresh water and marine water. The fresh water is exposed to all types of human activity, the water is captured, diverted, treated and reused to sustain communities and their economies. The water that is used in most human is not returned to the watershed from which it is caned. Instead, 80–90 percent of the waste water of developing countries is discharged into surface water, which contains dissolved suspended matter or dissolved solids that are harmful. Waste water discharges that are not regulated are harmful to the biological diversity and the planet's resilience and its capacity to provide basic ecosystem services. This affects the rural and urban populations. It affects various sectors of human activity, ranging from agriculture, fisheries, industry to health [1]. Coagulation–flocculation is a commonly used

chemical treatment process for water [2] and waste water [3]. It has a wide range of application in water and waste water facilities, because it is efficient and simple to operate [4]. Domestic waste water usually contains pathogens, suspended solids, nutrients and organic materials [5]. Waste water treatment is undertaken to satisfy the requirements for discharging treated water into the environment [6]. The processes of coagulation–flocculation followed by sedimentation and filtration are used globally in the treatment process before discharging the treated water into the river. Solid–liquid separation by a coagulation process is considered to be an important process in primary waste water treatment. Coagulation is a term that is used to describe the process of aggregation of colloidal particles into large aggregates to attain better stability. Aggregation of the particles occurs by two distinct mechanisms—charge neutralization and a sweep-floc mechanism. The charge neutralization may be caused by a specific chemical reaction between positively charged coagulants and negatively charged colloids, and natural organic matter. Alternatively, it may be caused by the shielding of the negatively charged sites, resulting in precipitation [7]. The coagulation process is an effective, efficient and simple method of treating water and waste water and is widely used in treating industrial effluent or water waste. The selection process and the coagulant dose that is added to the waste water are extremely important. Inorganic coagulants are frequently used in the coagulation process. In this regard, many researchers have conducted research of inorganic coagulant (aluminum ferric sulfate and chloride), synthetic organic polymers (polyacrylamide derivatives and polyethylene amine) [8,9]. However, prolonged and excessive usage of them can have detrimental effects on human health and living organisms. In addition, these inorganic coagulants can be ineffective in low-temperature water, result in a high procurement cost and generate significant amounts of toxic sludge. In contrast, natural coagulants, which are favored, are biodegradable, low in toxicity, cheaper and produce less sludge [10,11]. Because natural organic coagulants have proved to be effective in removing colloidal particles and are less expensive than other synthetic and inorganic coagulants [12], their use has been rising in recent years. In addition, organic coagulants do not contain metallic elements or reduce the alkalinity of the medium, because it is not subject to hydrolysis in solutions. Due to these characteristics, this coagulant is an ideal coagulant. The efficiency of the coagulation process is governed by various factors. These include the type, dosage and alkalinity of the coagulant [13–15], effluent pH [16,17], mixing speed and time [18] and temperature [19]. Therefore, optimization of these factors or inputs is crucial in increasing the coagulation treatment efficiency. Consequently, before application of natural coagulants for the clarification process, these factors must be optimized in order to increase the efficiency of the treatment. In this regard, the traditional method of experimentation, which is termed one-variable-at-a-time, involves changing one factor at a time to obtain the optimal conditions. However, it fails to resolve the relationship between multi-variables due to the complex influences of these factors. Further, this approach is time-consuming, laborious, and expensive because it requires a large number of experiments to be carried out [20,21]. To find a solution to this problem, Design of Experiments (DoE) and Response Surface Methodology (RSM) have been employed to study the effect of variables and their responses using a minimum number of experiments. RSM is an effective and efficient mathematical statistical method to build regression models, evaluate the effects of multiple variables and determine the optimal conditions that will provide desirable responses that successfully overcome the limitations of the conventional method [19,22–27]. Estimating the amount of coagulant that is necessary to remove the colloidal particles is normally accomplished by experiment using a jar test. The jar testing, which was developed almost 100 years ago, is still used to determine the optimal coagulation condition. Additionally, in solving waste water treatment engineering problems, soft computing and machine learning-based methods have proved their usefulness [22]. To treat and purify waste water, researchers have modelled and optimized many factors that influence the coagulation–flocculation process by means of regression models that were based on machine learning techniques. Baxter et al. [28], for example, modelled water quality with Artificial Neural Networks (ANNs) considering the water's turbidity and final color as outputs variables. In addition, the optimum dose of chemical flocculants that were necessary for maximum clarification of the waste

water was found. Multiple Linear Regression (MLR) and ANN were used by other researchers [29] to model turbidity and aluminum in treated water. Many factors or inputs were considered in this work including hardness, pH, color, silica, conductivity and turbidity. More recently [30], two types of Support Vector Machine (SVM) that employed different kernel functions were examined for use in predicting with K-Nearest Neighbors (KNN) the necessary coagulant dosage in water treatment plants for various levels of water turbidity. The input parameters in this study were the dosage, pH and temperature. In this case, the differential functions were radial basis functions and polynomial functions.

Several researchers have used also the RSM to obtain the optimal combination of inputs or process variables in the clarification process. However, most of their work is based on the modelling and optimization of relatively few output and input variables. This makes the optimization problem much easier. For example, Trinh and Kang [31] used the RSM technique to optimize the coagulation process in treatment of drinking water. In this case, the input parameters were pH and coagulant dose, whereas the output variables were turbidity and dissolved organic carbon. Other authors, such as Khayet et al. [32], investigated the effects of aluminium sulphate for coagulation/flocculation in removing dye from highly concentrated solutions at different temperatures and initial pH values. Li et al. [33] recently implemented a Multiple Response Surface (MRS) to optimize the coagulation process using polyferric chloride and optimizing three process parameters. They were coagulant dosage, pH and basicity. Subramonian et al. [10] applied RSM based on Central Composite Design (CCD) to investigate the effects of the main process variables, which included the coagulant dosage, pH and mixing time, in optimizing the coagulation process using *Cassia obtusifolia* seed gum and alum.

In the present study, final turbidity (*finTurb*) and total suspended solids (TSS), as the outputs of the clarification process, were considered. Meanwhile, the inputs or parameters of the coagulation process were initial turbidity (*iniTurb*), dosage of natural coagulant (*cuaDose*), process temperature (*T*), mixing speed (*s*) and mixing time (*t*). The main objective of this work was to conduct a comprehensive optimization study on three different and emerging natural coagulants in the treatment of raw domestic waste water. Regression models that relate the clarification process parameters to the outputs of the process were obtained with the use of RSM based on experiments of a CCD. Five clarification process scenarios were proposed and achieved using the MRS with desirability functions. The work is focused on the following ranges of input parameters, in which the regression models and the optimization achieved with MRS were valid: mixing time (*t*) (30–120 sec), mixing speed (*s*) (50–100 rpm), water temperature (*T*) (10–20 °C), initial turbidity (*iniTurb*) (40–140 NTU) and coagulant dosage (*cuaDose*) (1–6 mL). The experimental results were found to be close to the prediction that was derived from the regression models.

2. Materials and Methods

2.1. Materials

The samples of waste water that was used in this study were collected from the Water Waste Treatment Plant (WWTP) in Logroño (Spain). The waste water samples were transported to the laboratory within 15 minutes and then characterized (see Table 1). Three different coagulants were used in this study. They were ECOTAN BIO 90D, ECOTAN BIO 100 and ECOTAN BIO G150 (Servyeco Group, Castellón, Spain). The three coagulants that were studied were cationic organic polymers of essentially vegetal origin and strongly coagulating and flocculating action that act on the systems of colloidal particles. They neutralize charges and create bridges between these particles. This is the process that is responsible for the floc formation and consequently the separation by decantation. None of the coagulants that were used in the research modify the pH of the treated water, because they did not reduce the alkalinity of the medium. However, they were effective in a pH range of 4.5–8.0. The three emerging natural coagulants can be dosed directly to the water to be treated. Their application must be carried out at a point where the agitation guarantees the correct homogenization

of the water with the product, to favor the reaction of the polymer with the finest particles and allow for the formation of flocs. The jar test that was used in the experiments involved a programmable apparatus (FLOCUMATIC, Selecta, Spain [34]) according to the standard American Society of Testing Materials (ASTM) D2035-13 [35]. The apparatus consisted of six paddles on a bench. The paddles were connected to each other by a gear mechanism, and were rotated simultaneously by the same motor at a controlled speed and time. A jar test procedure was established at different temperatures for each experimental run. Waste water samples with a volume of 1000 mL each were transferred to the jars. Then, the required doses of different coagulants were added to each beaker during the flash mixing stage (the speed of the jar test and the time were varied throughout each experiment). Finally, the treated waste water was left to settle for 30 min. After the jar test had been completed, the treated waste water was analyzed to determine the finTurb, TSS and hence the extent of pollutant removal. The finTurb measurement was conducted with a 2100Q Turbidimeter (HACH, Loveland, CO, USA) according to the ASTM D1889-00 [36] standard, and TSS was performed following the standard ASTM D5907-10 method [37].

Table 1. Waste water characterization.

pH	DQO (mg/L)	DBO (mg/L)	Nt (mg/L)	NH ₄ (mg/L)	Pt (mg/L)
7.48	306.71	147.84	34.21	18.54	4.26

Response Surface Method for Optimizing the Clarification Process

The RSM method determines the relationships between input variables and one or more output variables. It was introduced by Box and Wilson in 1951 [38] for experimental data to be used for a model or optimal response. RSM was developed initially to model experimental responses. However, it has found use in optimizing industrial processes and products [39,40]. RSM often appears as a series of statistical techniques that require a regression model that depends on a low-degree polynomial function (Equation (1)):

$$Y = f(X_1, X_2, X_3, \dots, X_k) + e \quad (1)$$

where Y is an experimental response, f consists of cross products of polynomial's terms, $X_1, X_2, X_3, \dots, X_k$ are input vectors, and e is the error. The quadratic (second-order) model is a polynomial function that is frequently used, and is written as Equation (2):

$$Y = b_0 + \sum_{i=1}^n b_i \times X_i + \sum_{i=1}^n b_{ii} \times X_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^n b_{ij} \times X_i \times X_j + e \quad (2)$$

where the linear part constitutes the first summation, the quadratic part forms the second part, and the third part is the product of the pairs of variables. Regression analysis is used to determine the values of the coefficients b_0, b_i, b_{ii} , and b_{ij} . However, good results are not always possible with these functions, if the problem is complex and involves many inputs and nonlinearities. The reason is that the coefficients cannot be adjusted if the data are sparse-like continuous functions that are described by polynomials. The p -value (or Prob. $> F$) is the probability that the result that will equal or exceed the value that was actually observed assuming that the model produces results that are accurate. The p -value can be determined by analysis of the variance (ANOVA). If the model's Prob. $> F$ and no term exceeds the level of significance (e.g., $\alpha = 0.05$), one can consider the model to be acceptable within a $(1 - \alpha)$ confidence interval. Some investigators have used ANOVA to study the effect of inputs or process variables on process outputs [41,42]. If a problem has more than one outputs, the latter is termed MRS. Normally, it causes conflicting solutions, as the optimal configuration may vary significantly from one output to another. A compromise was suggested by Harrington [43]. It involves desirability functions for each output, Equations (3) and (4). It also involves an overall desirability. The latter is the mean value of the desirability (D) of each output (Equation (5)). Equations (3)–(5) can be described as:

$$d_r^{\max} = \begin{cases} 0 & \text{if } f_r(X) < A \\ \left(\frac{f_r(X)-A}{B-A}\right)^S & \text{if } A \leq f_r(X) \leq B \\ 1 & \text{if } f_r(X) > B \end{cases} \quad (3)$$

$$d_r^{\min} = \begin{cases} 1 & \text{if } f_r(X) < A \\ \left(\frac{f_r(X)-B}{A-B}\right)^S & \text{if } A \leq f_r(X) \leq B \\ 0 & \text{if } f_r(X) > B \end{cases} \quad (4)$$

$$D = \left(\prod_{r=1}^R d_r\right)^{1/R} \quad (5)$$

where A and B are limiting values, s indicates the importance of reaching the target value, X is the input vector, and f_r is the model for prediction. It is beneficial to use a second-higher-degree polynomial to optimize responses [44]. In the desirability approach, each estimated response is transformed into a unit-less utility that is bounded by $0 < d_r < 1$. A higher value of d_r signifies a more desirable response value. The optimization function of the R package v.1.6 [45] searches for a combination of importance factors (or weights of 1, 2 or -3) that satisfy the process criteria for each response and input.

2.2. Experiments Design

In order to create accurate models with the minimum amount of data that is necessary to support the initial hypotheses, RSM must establish a DoE [46]. There are several methods available. However, they require the construction of a design matrix (inputs) to measure the responses or outputs. In this case, a CCD [47] that involved three factors and three levels was chosen. The input variables that were selected to develop the DoE were time (t), temperature (T), speed (S), initial turbidity (iniTurb), and coagulant dosage (cuaDose). The factors that were selected by experiment for optimization and their respective ranges were: mixing time (30–120 sec), mixing speed (50–100 rpm), water temperature (10–20 °C), initial turbidity (40–140 NTU) and coagulant dosage (1–6 mL). The model was valid only inside the region of the parameters studied. The variable ranges that are provided in Table 2 were adopted to cover the intervals that are commonly utilized in the literature [48–50].

Table 2. Parameters of the coagulation process and levels.

Input	Notation	Magnitude	Levels		
			−1	0	1
time	t	sec	30	75	120
speed	S	rpm	50	75	100
temperature	T	°C	10	15	20
initial turbidity	iniTurb	NTU	40	90	140
coagulant dosage	cuaDose	mL	1	3.5	6

After the parameters of the coagulation process and levels were set as in Table 1, the design matrix (Table 3) was generated by the R Statistical Software [45]. In order to provide for all possibilities and to select the optimum process of clarification, 54 experiments were required in this case.

After establishing the inputs or parameters of the coagulation process from the DoE, the output variables of the process were determined experimentally (Table 4).

Table 3. Design matrix for clarification of waste water.

Sample	Parameters of the Coagulation Process				
	Time (t) (sec)	Speed (s) (rpm)	Temp (T) (°C)	Initial Turbidity (iniTurb) (NTU)	Coagulant Dosage (cuaDose) (mL)
1	30	100	20	40	1
2*	75	75	15	90	3.5
3	120	50	10	40	6
4	30	50	20	40	6
5	30	50	20	140	1
6	30	100	20	140	6
7	75	75	15	90	1
8	30	100	10	140	6
9	75	75	15	40	3.5
10	120	50	10	140	1
11	30	50	10	140	1
12	75	75	15	90	3.5
13	75	75	15	90	3.5
14	30	50	10	40	6
15	120	100	20	40	1
16	120	100	20	40	6
17	120	100	10	40	1
18	30	75	15	90	3.5
19	120	50	20	40	1
20	75	75	15	90	6
21	120	100	10	140	6
22*	75	75	15	90	3.5
23	120	50	20	140	6
24*	75	75	15	90	3.5
25*	75	75	15	90	3.5
26	75	75	15	140	3.5
27	75	75	20	90	3.5
28	120	50	20	140	1
29	75	50	15	90	3.5
30	30	100	20	40	6
31	75	100	15	90	3.5
32	30	50	20	140	6
33	30	50	10	140	6
34	120	50	10	140	6
35*	75	75	15	90	3.5
36	30	100	20	140	1
37*	75	75	15	90	3.5
38	120	100	20	140	1
39	30	50	10	40	1
40*	75	75	15	90	3.5
41	120	50	20	40	6
42	75	75	10	90	3.5
43	30	100	10	40	1
44	120	100	20	140	6
45	120	100	10	40	6
46	120	50	10	40	1
47*	75	75	15	90	3.5
48	30	50	20	40	1
49	30	100	10	140	1
50*	75	75	15	90	3.5
51	120	75	15	90	3.5
52*	75	75	15	90	3.5
33	30	100	10	40	6
54	120	100	10	140	1

Table 4. Experimental design matrix for the clarification process in waste water treatment.

Sample	Outputs					
	ECOTAN BIO 90D		ECOTAN BIO 100		ECOTAN BIO G150	
	finTurb (NTU)	TSS (mg/L)	finTurb (NTU)	TSS (mg/L)	finTurb (NTU)	TSS (mg/L)
1	20.09	598.3	31.11	696.0	30.20	227
2	12.67	203.7	36.10	689.0	27.30	600
3	11.71	42.0	17.99	318.0	33.30	327
4	19.10	48.0	25.00	631.0	41.30	414
5	34.90	658.0	22.50	134.0	41.70	432
6	9.60	107.8	21.60	394.0	25.90	91
7	12.85	273.3	35.30	528.0	32.70	585
8	13.11	392.0	35.40	110.0	41.30	201
9	9.60	119.5	27.90	429.0	36.70	586
10	6.83	413.0	13.51	555.0	11.50	291
11	15.30	368.0	28.90	507.0	19.10	609
12	9.20	302.0	27.50	694.3	23.90	535
13	11.90	396.0	25.70	629.0	30.40	551
14	28.90	641.9	22.20	5.0	53.90	32
15	10.04	514.7	23.80	423.0	35.10	238
16	6.95	611.3	22.60	29.3	21.70	261
17	6.45	249.0	22.30	414.0	40.70	332
18	18.2	761.9	46.70	285.	63.90	771
19	9.57	354.6	10.70	507.0	35.30	514
20	13.10	596.0	41.20	134.0	39.50	27
21	5.87	563.9	7.63	482.0	48.00	38
22	8.70	707.3	37.20	505.0	32.20	690
23	13.18	362.0	28.70	629.0	46.30	620
24	8.43	387.0	27.40	463.0	30.20	708
25	6.62	261.2	29.90	230.0	33.60	709
26	12.6	388.5	22.70	289.0	14.10	235
27	9.80	576.0	32.10	398.0	27.40	561
28	19.80	57.0	9.83	128.0	14.88	377
29	8.67	423.9	30.20	652.	21.40	390
30	21.30	141.0	26.70	190.0	37.90	215
31	15.20	355.0	19.80	827.9	22.70	211
32	6.30	75.8	20.30	410.0	55.90	166
33	20.10	229.0	53.70	528.0	30.00	546
34	6.82	160.6	66.50	242.0	36.20	571
35	9.63	539.2	39.00	648.0	14.81	260
36	16.10	430.0	13.00	592.0	17.00	255
37	4.97	395.1	39.80	227.0	28.24	473
38	10.70	220.0	31.90	396.0	14.55	670
39	34.30	443.0	28.70	703.0	39.80	534
40	14.60	525.8	49.80	687.0	28.26	198
41	6.63	550.0	28.60	97.0	47.70	4
42	7.68	139.0	49.20	490.0	45.70	357
43	15.60	539.1	17.30	475.0	37.00	383
44	7.65	529.0	5.83	523.0	22.10	276
45	14.34	769.0	20.90	403.0	33.30	275
46	8.45	215.2	25.40	312.0	43.70	307
47	9.10	618.6	60.90	470.0	23.80	486
48	17.55	598.5	35.10	472.0	30.20	420
49	23.10	673.0	24.90	623.0	21.50	463
50	12.80	390.5	47.50	419.2	31.20	467
51	17.82	752.0	38.70	101.0	29.60	520
52	3.87	392.0	19.50	246.0	30.70	604
53	32.60	464.0	13.60	24.0	44.10	74
54	6.83	164.0	12.53	880.0	12.46	432

3. Results and Discussion

This methodology was found to be an efficient and direct way to obtain the optimal dose of coagulant with which finTurb and TSS were reduced, when several inputs or process parameters of the coagulation process were considered. The inputs were: initial turbidity (iniTurb), dosage of natural coagulant (cuaDose), process temperature (T), mixing speed (s) and mixing time (t). The proposed method has been applied to three emerging natural coagulants. The proposed method has been applied to various types of natural coagulants.

3.1. Experimental Results

Table 4 shows the experimental results for the outputs of the process (finTurb and TSS) according to the CCD DoE design matrix for each of the three emerging natural coagulants (Table 3).

3.2. Analysis of Variance

The data that appear in Table 2 were used in fitting Equation (2) to obtain by the R Base package regression equations for the responses [48]. Second-order polynomial models then were constructed for each response. Then, several criteria (*p*-value, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Correlation (Corr.)) were used in the selection of the model that provided the greatest accuracy.

Equations (6)–(11) show the reduced quadratic regression models that were obtained to model finTurb and TSS for each coagulant tested. These equations show how a combination of second-order polynomials is formed by combining the input variables that provide the output. An ANOVA was used to reduce the size of these regression models by removing the insignificant terms of the polynomial [49]. This method, called step-wise regression, automatically removes the features of no significance for each model [50].

$$\begin{aligned} \text{finTurb} = & 36.7506673 - 0.9446782 \cdot t + 0.0039169 \cdot t^2 - 0.2203728 \cdot s + 0.000966 \cdot t \cdot s + 3.5979252 \cdot T + \\ & 0.0096595 \cdot t \cdot T - 0.1570211 \cdot T^2 + 0.009553 \cdot T \cdot \text{iniTurb} - 0.0005356 \cdot \text{iniTurb}^2 + 0.0102462 \cdot t \cdot \text{cuaDose} + \\ & 0.0183188 \cdot s \cdot \text{cuaDose} - 0.1771652 \cdot T \cdot \text{cuaDose} - 0.0170971 \cdot \text{iniTurb} \cdot \text{cuaDose} + 0.2285438 \cdot \text{cuaDose}^2 \end{aligned} \quad (6)$$

$$\begin{aligned} \text{TSS} = & -107.42648 - 36.95502 \cdot t + 0.16738 \cdot t^2 + 0.06127 \cdot t \cdot s - 0.04981 \cdot s^2 + 195.53517 \cdot T + \\ & 0.23366 \cdot t \cdot T - 6.46003 \cdot T^2 + 13.57065 \cdot \text{iniTurb} - 0.07392 \cdot \text{iniTurb}^2 - 101.30295 \cdot \text{cuaDose} + \\ & 1.00312 \cdot t \cdot \text{cuaDose} + 0.90183 \cdot s \cdot \text{cuaDose} - 3.98293 \cdot T \cdot \text{cuaDose} \end{aligned} \quad (7)$$

$$\begin{aligned} \text{finTurb} = & -94.889721 - 0.065567 \cdot t + 3.138794 \cdot s - 0.021709 \cdot s^2 + 0.023985 \cdot s \cdot T + \\ & 0.336475 \cdot \text{iniTurb} - 0.002649 \cdot s \cdot \text{iniTurb} - 0.013975 \cdot T \cdot \text{iniTurb} + 6.61089 \cdot \text{cuaDose} - \\ & 0.049384 \cdot S \cdot \text{cuaDose} - 0.280453 \cdot T \cdot \text{cuaDose} + 0.025552 \cdot \text{iniTurb} \cdot \text{cuaDose} \end{aligned} \quad (8)$$

$$\begin{aligned} \text{TSS} = & 2301.0174 + 19.1258 \cdot t - 0.1315 \cdot t^2 - 67.1235 \cdot s + 0.4737 \cdot s^2 - 0.1496 \cdot T \cdot \text{iniTurb} - \\ & 0.9338 \cdot S \cdot \text{cuaDose} + 4.391 \cdot T \cdot \text{cuaDose} + 0.8546 \cdot \text{iniTurb} \cdot \text{cuaDose} - 17.0285 \cdot \text{cuaDose}^2 \end{aligned} \quad (9)$$

$$\begin{aligned} \text{finTurb} = & -18.290656 - 1.140292 \cdot t + 0.007104 \cdot t^2 + 3.264461 \cdot s - 0.01929 \cdot s^2 - 0.021824 \cdot s \cdot T - \\ & 0.454037 \cdot \text{iniTurb} + 0.012343 \cdot T \cdot \text{iniTurb} - 0.025762 \cdot s \cdot \text{cuaDose} + 0.04546 \cdot \text{iniTurb} \cdot \text{cuaDose} \end{aligned} \quad (10)$$

$$\begin{aligned} \text{TSS} = & 75.98257 - 15.657 \cdot t + 0.08411 \cdot t^2 + 46.44557 \cdot s - 0.30951 \cdot s^2 - 150.0361 \cdot T + \\ & 0.17174 \cdot t \cdot T + 4.19307 \cdot T^2 + 5.53688 \cdot \text{iniTurb} - 0.0266 \cdot \text{iniTurb}^2 + 211.01884 \cdot \text{cuaDose} - \\ & 0.67333 \cdot s \cdot \text{cuaDose} - 28.27164 \cdot \text{cuaDose}^2 \end{aligned} \quad (11)$$

In addition, an ANOVA test was conducted in order to determine whether the variables used in the regression models are statistically significant. It is well known that *p*-value (or Prob. > *F*) is the probability of receiving a result that equals or exceeds what was observed. If the Prob. is greater than the *F* value of the model, and the model has no term with a level of significance that exceeds, for example, $\alpha = 0.05$, then the model is acceptable at a confidence interval of $(1 - \alpha)$. In this case, most

of the variables have a p -value that is less than 0.05. This indicates that the variables that were used on the quadratic models are statistically significant. Tables 5–10 show the results of the ANOVA test for each of the final quadratic models for each coagulant that was tested.

Table 5. ANOVA values for the “finTurb” quadratic model for ECOTAN BIO 90D.

Variable	Df ¹	Sum of Square	Mean Square	F Value	p -value
t	1	916.346	916.346	53.236	8.428×10^{-6}
t ²	1	323.793	323.793	18.811	9.867×10^{-5}
s	1	31.219	31.219	1.814	0.18584
t·s	1	3.638	3.638	0.211	0.64825
T	1	18.473	18.473	1.073	0.30660
t·T	1	102.780	102.780	5.971	0.01917
T ²	1	57.522	57.522	3.342	0.07520
T·iniTurb	1	2.850	2.850	0.166	0.68631
iniTurb ²	1	153.633	153.633	8.925	0.00484
t·cuaDose	1	15.650	15.650	0.909	0.34620
s·cuaDose	1	4.262	4.262	0.248	0.62155
T·cuaDose	1	311.691	311.691	18.108	0.00013
iniTurb·cuaDose	1	136.885	136.885	7.952	0.00751
cuaDose ²	1	30.343	30.343	1.763	0.19200
Residuals	39	671.306	17.213		

¹ Degrees of freedom (Df).

Table 6. ANOVA values for the “TSS” quadratic model for ECOTAN BIO 90D.

Variable	Df ¹	Sum of Square	Mean Square	F Value	p -value
t	1	12123.226	12123.226	0.486	0.48969
t ²	1	142.315	142.315	0.006	0.94016
t·s	1	125559.858	125559.858	5.035	0.03045
s ²	1	5888.921	5888.921	0.236	0.62966
T	1	45.924	45.924	0.002	0.96598
t·T	1	47902.024	47902.024	1.921	0.17345
T ²	1	201652.233	201652.233	8.086	0.00699
iniTurb	1	9905.698	9905.698	0.397	0.53212
iniTurb ²	1	164612.655	164612.655	6.601	0.01403
cuaDose	1	53839.319	53839.319	2.159	0.14957
t·cuaDose	1	371281.918	371281.918	14.888	0.00041
s·cuaDose	1	98334.840	98334.840	3.943	0.05395
T·cuaDose	1	74204.701	74204.701	2.976	0.09225
Residuals	40	997539.737	24938.493		

¹ Degrees of freedom (Df).

Table 7. ANOVA values for the “finTurb” quadratic model for ECOTAN BIO 100.

Variable	Df ¹	Sum of Square	Mean Square	F Value	p -value
t	1	166.585	166.585	1.675	0.203000
s	1	394.790	394.790	3.970	0.053178
s ²	1	2208.677	2208.677	22.209	2.943×10^{-5}
s·T	1	11.438	11.438	0.115	0.736274
iniTurb	1	0.034	0.034	0.000	0.985279
s·iniTurb	1	223.193	223.193	2.244	0.141956
T·tubrIni	1	696.807	696.807	7.007	0.011558
cuaDose	1	118.417	118.417	1.191	0.281708
s·cuaDose	1	414.589	414.589	4.169	0.047805
T·cuaDose	1	311.029	311.029	3.128	0.084605
iniTurb·cuaDose	1	298.428	298.428	3.001	0.090926
Residuals	40	3977.917	99.448		

¹ Degrees of freedom (Df).

Table 8. ANOVA values for the “TSS” quadratic model for ECOTAN BIO 100.

Variable	Df ¹	Sum of Square	Mean Square	F Value	p-value
t	1	3433.556	3433.556	0.134	0.71620
t ²	1	134355.201	134355.201	5.242	0.02714
s	1	11120.751	11120.751	0.434	0.51369
s ²	1	212994.469	212994.469	8.310	0.00619
T·iniTurb	1	641.102	641.102	0.025	0.87509
s·cuaDose	1	407001.590	407001.590	15.879	0.00026
T·cuaDose	1	91379.561	91379.561	3.565	0.06592
iniTurb·cuaDose	1	301130.135	301130.135	11.749	0.00137
cuaDose ²	1	177895.900	177895.900	6.941	0.01175
Residuals	42	1076504.55	25631.061		

¹ Degrees of freedom (Df).**Table 9.** ANOVA values for the “finTurb” quadratic model for ECOTAN BIO G150.

Variable	Df ¹	Sum of Square	Mean Square	F Value	p-value
t	1	270.772	270.772	4.612	0.03742
t ²	1	329.885	329.885	5.619	0.02231
s	1	255.758	255.758	4.357	0.04282
s ²	1	411.999	411.999	7.018	0.01124
s·T	1	332.273	332.273	5.660	0.02186
iniTurb	1	710.517	710.517	12.103	0.00116
T·iniTurb	1	199.801	199.801	3.403	0.07195
s·cuaDose	1	768.059	768.059	13.083	0.00077
iniTurb·cuaDose	1	1120.611	1120.611	19.089	7.74×10^{-5}
Residuals	43	2524.340	58.706		

¹ Degrees of freedom (Df).**Table 10.** ANOVA values for the “TSS” quadratic model for ECOTAN BIO G150.

Variable	Df ¹	Sum of Square	Mean Square	F Value	p-value
t	1	1423.529	1423.529	0.057	0.81278
t ²	1	158135.752	158135.752	6.312	0.01602
s	1	107521.882	107521.882	4.292	0.04463
s ²	1	226926.055	226926.055	9.058	0.00446
T	1	704.326	704.326	0.028	0.86767
t·T	1	63037.289	63037.289	2.516	0.12036
T ²	1	40058.639	40058.639	1.599	0.21319
iniTurb	1	10881.901	10881.901	0.434	0.51354
iniTurb ²	1	25196.398	25196.398	1.006	0.32181
cuaDose	1	293880.040	293880.040	11.731	0.00141
s·cuaDose	1	45925.039	45925.039	1.833	0.18317
cuaDose ²	1	85936.210	85936.210	3.430	0.07122
Residuals	41	1027147.47	25052.377		

¹ Degrees of freedom (Df).

From Table 5, it is observed that for ECOTAN BIO 90D, there is a notable influence of t (p -value = 8.428×10^{-6}) on finTurb, whereas for s (p -value = 0.18584) and for T (p -value = 0.30660), its influence is much smaller. On the other hand, iniTurb has a certain influence on finTurb when it interacts with itself (p -value = 0.00484), as well as when it interacts with T (p -value = 0.00013). Finally, cuaDose has a notable influence on finTurb when it interacts with iniTurb (p -value = 0.00751) and with T (p -value = 0.00013). From Table 6, it is observed that for ECOTAN BIO 90D, there is no significant influence of any input on TSS, at least when the variables do not interact. However, Table 5 also shows that cuaDose has a notable influence when it interacts with t (p -value = 0.00041) and s (p -value = 0.05395).

In addition, *t* has a notable influence when it interacts with *s* (p -value = 0.03045), and with *T* as well when it interacts with itself (p -value = 0.00699). From Table 7, it is observed that for ECOTAN BIO 100, there is a notable influence of *s* (p -value = 0.053178) on *finTurb*, whereas for *t* (p -value = 0.203000) its influence is limited. In addition, this table also shows that *s* has an important influence when it interacts with itself (p -value = 2.943×10^{-5}) and when it interacts with *cuaDose* (p -value = 0.047805). On the other hand, *iniTurb* also has a notable influence when it interacts with *T* (p -value = 0.011558). In the same manner that occurs with ECOTAN BIO 90D, Table 8 shows that for ECOTAN BIO 100, there is no significant influence of any input on TSS, at least when the variables do not interact. However, in Table 8, it is observed that *t* and *s* have a notable influence on TSS when it interacts with itself (p -value = 0.02714 and p -value = 0.00619, respectively). These two variables also have a notable influence on *cuaDose*, which is also significantly influenced when it interacts with *s* (p -value = 0.00026), with *iniTurb* (p -value = 0.00137), and with itself (p -value = 0.01175). Finally, regarding ECOTAN BIO G150 (Table 9), it is observed that *t*, *s* and *iniTurb* have an important influence on *finTurb* (p -value = 0.03742, p -value = 0.04282 and p -value = 0.00116, respectively), and on *cuaDose*, where very low values are obtained for p -value (for *s*·*cuaDose*: p -value = 0.00077; and for *iniTurb*·*cuaDose*: p -value = 7.74×10^{-5}). For this same coagulant, it is also observed that *s* and *cuaDose* have a significant influence on TSS (p -value = 0.04463 and p -value = 0.00141, respectively), while *t* also influences TSS when it is interacting with itself (p -value = 0.01602). *T* has no influence on TSS in any way.

Once the results of the ANOVA were presented, it was concluded that the input *T* has no influence on any of the coagulants for *finTurb* as well as for TSS. Thus, *T* can be eliminated from this study, while *t* and *s* are the most influential variables. By contrast, in ECOTAN BIO G150, *iniTurb* and *cuaDose* have a very important influence on *finTurb* and TSS, in addition to the aforementioned *s* and *t*.

MAE and RMSE were calculated to determine the generalization capacity of the quadratic models using the samples shown in Table 4 according to Equations (12) and (13):

$$\text{MAE} = \frac{1}{m} \cdot \sum_{k=1}^m |Y_{k \text{ Experiment}} - Y_{k \text{ Model}}| \quad (12)$$

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{k=1}^m (Y_{k \text{ Experiment}} - Y_{k \text{ Model}})^2} \quad (13)$$

In this case, $Y_{k \text{ Experiment}}$ are the experimentally obtained responses and $Y_{k \text{ Model}}$ are responses from the quadratic models that RSM and *m* experiments produced. The prediction errors are shown in Table 11. The maximum error corresponds to TSS (the MAE equals to 0.14905 and the RMSE equals to 0.17981). The minimum error corresponds to *finTurb* (the MAE equals to 0.09261 and the RMSE equals to 0.11363).

Table 11. Results of the errors in *FinTurb* and TSS using the quadratic models.

Coagulant	Outputs	Correlation	MAE ¹ Train	RSMSE ² Train
BIO 90D	<i>finTurb</i>	0.87095	0.09261	0.11363
	TSS	0.73405	0.14905	0.18695
BIO 100	<i>finTurb</i>	0.74100	0.11541	0.14416
	TSS	0.74466	0.13621	0.16444
BIO G150	<i>finTurb</i>	0.79713	0.10266	0.13171
	TSS	0.71259	0.14427	0.17981

¹ Mean Absolute Error (MAE), ² Root Mean Squared Error (RMSE).

In order to test the proposed models, 11 new experiments that had not been used during the training process were conducted. They were used to test the proposed second-order regression models with previously unused regression model parameters. These new 11 experiments were generated randomly. Table 12 shows the test matrix for the clarification process in waste water treatment.

Table 12. Test matrix for clarification process in waste water treatment.

Sample	Parameters of the Coagulation Process				
	Time (t) (sec)	Speed (s) (rpm)	Temp (T) (°C)	Initial Turbidity (iniTurb) (NTU)	Coagulant Dosage (cuaDose) (mL)
1	44	84	12.0	78	3.35
2	60	80	17.0	115	4.35
3	88	56	14.5	103	4.80
4	94	57	19.0	133	3.75
5	80	96	15.0	76	2.05
6	69	55	10.0	134	4.20
7	76	81	12.0	97	2.30
8	95	87	16.0	131	1.50
9	83	73	13.0	124	3.30
10	69	81	11.0	59	4.90
11	59	55	15.0	116	3.05

Once the 11 new experiments had been completed, the errors that arose during the testing stage were calculated (see Table 13). This table shows that the maximum error corresponds to TSS (an MAE of 0.1741 and an RMSE of 0.2342), whereas the minimum error corresponds to finTurb (an MAE of 0.1206 and an RMSE of 0.1630). The errors indicate that the adjustment of the regression models and the results of the experiments are relatively accurate.

Table 13. Results of the errors in finTurb and TSS using the second-order regression models.

Errors	ECOTAN BIO 90D		ECOTAN BIO 100		ECOTAN BIO G150	
	finTurb (NTU)	TSS (mg/L)	finTurb (NTU)	TSS (mg/L)	finTurb (NTU)	TSS (mg/L)
MAE	0.14902	0.1600	0.12016	0.17693	0.15815	0.17413
RMSE	0.19515	0.22979	0.16307	0.21898	0.2145	0.23425

After determining the errors in the prediction from the regression models of the training and test data, a scatter diagram of the variables was created. Figure 1 shows the scatter diagrams or relationships between the actual values that were obtained experimentally and the values of finTurb that were predicted by the quadratic models of (Figure 1a,c and e) and TSS (Figure 1b,d and f) for each of the tested coagulants—ECOTAN BIO 90D, ECOTAN BIO 100 and ECOTAN BIO G150. The blue points relate to the 54 data points that appear in Table 2. The red points relate to the 11 additional experiments that the regression models used. They appear in Table 11. If the variables in this case are correlated, the points will follow the diagonal line or curve. The better the correlation, the closer the points will be to the line. Because all of the red dots (test data) are closer to the diagonal line than are most blue dots (training data), their correlation is greater. Because there are fewer test data points than training data points, the MAE and RMSE errors in the testing analysis and training analysis are similar (see Tables 11 and 12).

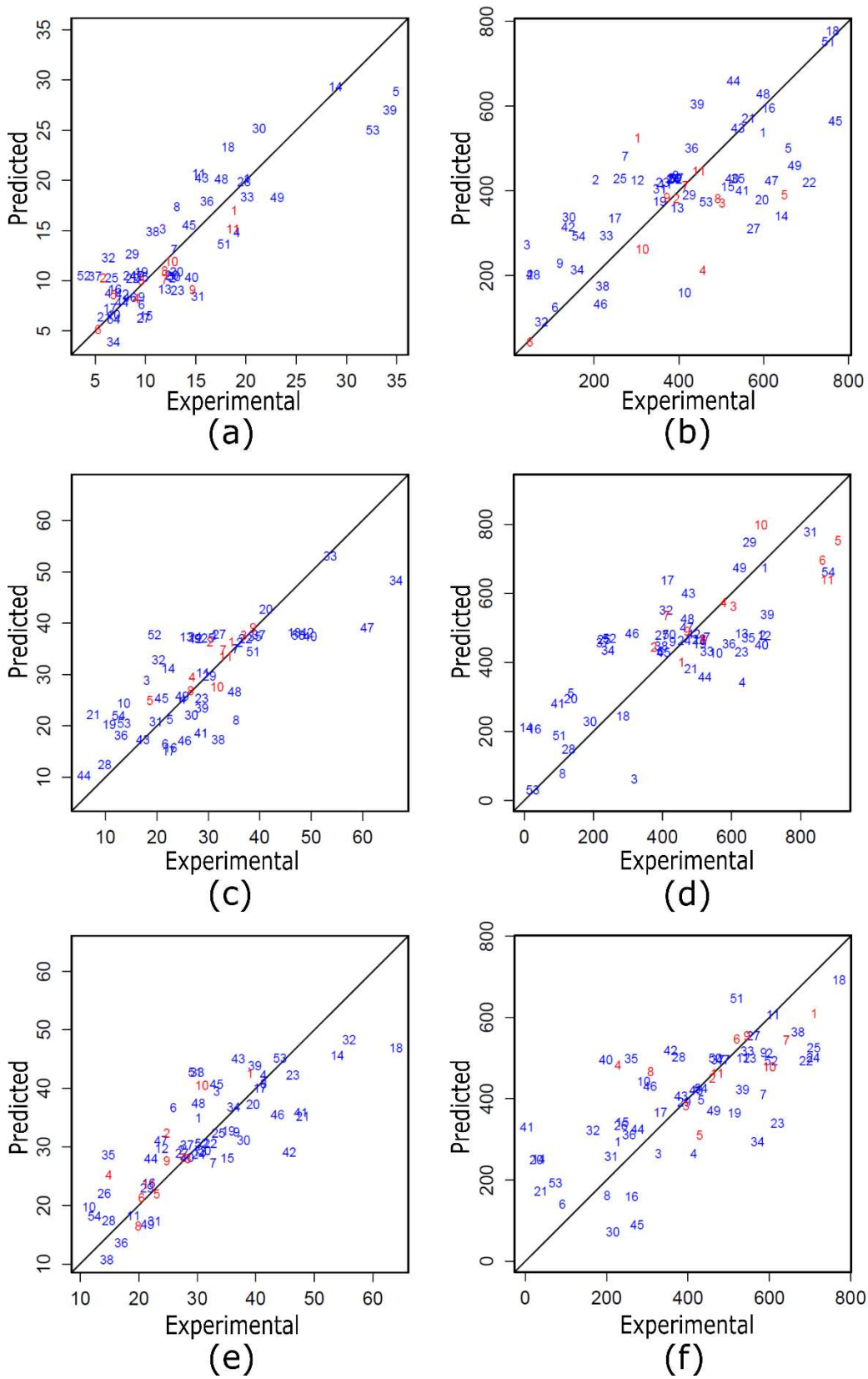


Figure 1. Scatter diagrams of ECOTAN BIO 90D: (a) finTurb and (b) TSS; ECOTAN BIO 100: (c) finTurb and (d) TSS; and ECOTAN BIO G150: (e) finTurb and (f) TSS.

3.3. Multi-Response Optimization

Tables 14–18 show the combination of parameters of the coagulation and outputs of the process that were studied when examining the clarification process in waste water treatment by desirability functions. The examination involved the desirability package [51] for three different emerging natural coagulants and five different criteria. The first column of each table shows the coagulant that was used in the tests. The second and third columns show, respectively, the parameters of the coagulation and the outputs of the process that were studied, as well as the goal that was established for the optimization process. The fourth column indicates the degrees of importance considered in the optimization process. For all cases studied, it was assumed that all objectives had an importance level of 1.0. The fifth column shows the optimized values that were obtained. The desirability values that were obtained appear in the sixth column. Table 14 shows the results when all coagulation parameters and process outputs are equally important and the design requirements are based on minimizing t , s , T , $iniTurb$, $finTurb$ and TSS. In this case, the values of the overall desirability were 0.841 for ECOTAN BIO 90D, 0.797 for ECOTAN BIO 100, and 0.774 for ECOTAN G150. Table 15 shows the results when the design requirements were based on minimizing the energy that the coagulation process consumes (the energy is a function of t , s , and T) and obtaining the highest removal efficiency of turbidity (minimizing $finTurb$) while minimizing TSS. The overall desirability values that were obtained were 0.88 for ECOTAN BIO 90D, 0.791 for ECOTAN BIO 100, and 0.768 for ECOTAN G150. Table 16 provides the results when the design requirements were based on minimizing the energy that was consumed to obtain the highest removal efficiency of turbidity (minimizing $finTurb$) and minimize TSS, when the waste water has a high degree of turbidity ($iniTurb$ is maximum). In this case, the overall desirability values that were obtained for the natural coagulants were 0.857 for ECOTAN BIO 90D, 0.716 for ECOTAN BIO 100, and 0.794 for ECOTAN G150. Table 17 shows the results when the design requirements are based on maximizing the efficiency in removing turbidity (minimizing the $finTurb$) and minimizing TSS, while the waste water has a high degree of turbidity ($iniTurb$ is maximum) and a very low temperature (T is minimum). The overall desirability values that were obtained were 0.986 for ECOTAN BIO 90D, 0.754 for ECOTAN BIO 100, and 0.774 for ECOTAN G150. Finally, Table 18 shows the results when the design requirements are based on maximizing the efficiency in removing turbidity (minimizing the $finTurb$) and minimizing TSS, when the waste water has a high degree of turbidity ($iniTurb$ is maximum) and a high temperature (T is maximum). The overall desirability values obtained were 0.976 for ECOTAN BIO 90D, 0.941 for ECOTAN BIO 100, and 0.774 for ECOTAN G150. Additionally, the results that appear in Tables 14–18 indicate that the outputs of the process that was studied ($cuaDose$ and TSS) are very similar for all different optimization scenarios in a range of inputs or parameters of the coagulation process that was considered. For example, the ranges of values for the $cuaDose$ of the different optimization scenarios that considered for the natural coagulants extend from 3.007 to 5.455 for the ECOTAN BIO 90D, from 1.000 to 5.441 for the ECOTAN BIO 100, and from 1.062 to 5.422 for the ECOTAN G150. In addition, the ranges of values for the TSS of the different optimization scenarios for each natural coagulant extend from 0.000 to 51.901 for the ECOTAN BIO 90D, from 0.000 to 441.478 for the ECOTAN BIO 100, and from 194.109 to 343.679 for the ECOTAN G150. From these results, it appears that the optimal parameters of coagulation for different clarification optimization scenarios are found in a relatively narrow range.

Table 14. The first clarification optimization scenario: all variables are considered to be equally important.

Coagulant	Parameters	Goal	Importance	Value	Desirability
ECOTAN BIO 90D	t	min	1.0	67.197	0.587
	s	min	1.0	51.481	0.97
	T	min	1.0	8.598	0.978
	iniTurb	min	1.0	52.207	0.862
	cuaDose	inRange	1.0	4.704	1.000
	finTurb	min	1.0	12.057	0.736
	TSS	min	1.0	5.010	1.000
	Overall Desirability				
ECOTAN BIO 100	t	min	1.0	32.057	0.977
	s	min	1.0	59.857	0.803
	T	min	1.0	12.027	0.882
	iniTurb	min	1.0	41.924	0.831
	cuaDose	inRange	1.0	5.441	1.000
	finTurb	min	1.0	37.85	0.472
	TSS	min	1.0	51.273	0.947
	Overall Desirability				
ECOTAN BIO G150	t	min	1.0	36.344	0.93
	s	min	1.0	50.476	0.99
	T	min	1.0	13.437	0.738
	iniTurb	min	1.0	42.971	0.936
	cuaDose	inRange	1.0	1.062	1.00
	finTurb	min	1.0	38.788	0.479
	TSS	min	1.0	343.679	0.557
	Overall Desirability				

Table 15. The second clarification optimization scenario: the energy consumption is minimized to obtain the highest efficiency of turbidity removal and TSS.

Coagulant	Parameters	Goal	Importance	Value	Desirability
ECOTAN BIO 90D	t	min	1.0	65.57	0.605
	s	min	1.0	52.019	0.960
	T	min	1.0	8.569	0.980
	iniTurb	inRange	1.0	127.612	1.000
	cuaDose	inRange	1.0	3.162	1.000
	finTurb	min	1.0	5.655	0.942
	TSS	min	1.0	51.901	0.986
	Overall Desirability				
ECOTAN BIO 100	t	min	1.0	31.967	0.978
	s	min	1.0	59.857	0.803
	T	min	1.0	12.027	0.882
	iniTurb	inRange	1.0	40.099	1.000
	cuaDose	inRange	1.0	5.441	1.000
	finTurb	min	1.0	37.882	0.472
	TSS	min	1.0	50.807	0.948
	Overall Desirability				
ECOTAN BIO G150	t	min	1.0	51.271	0.764
	s	min	1.0	50.858	0.983
	T	min	1.0	13.779	0.712
	iniTurb	inRange	1.0	133.808	1.000
	cuaDose	inRange	1.0	3.781	1.000
	finTurb	min	1.0	18.325	0.87
	TSS	min	1.0	328.925	0.576
	Overall Desirability				

Table 16. The third clarification optimization scenario: the energy consumption is minimized to obtain the highest efficiency of turbidity removal and TSS, when the sample water is very turbid.

Coagulant	Parameters	Goal	Importance	Value	Desirability
ECOTAN BIO 90D	t	min	1.0	49.463	0.784
	s	min	1.0	62.784	0.744
	T	min	1.0	10.473	0.838
	iniTurb	max	1.0	164.399	0.995
	cuaDose	inRange	1.0	3.007	1.000
	finTurb	min	1.0	9.222	0.828
	TSS	min	1.0	51.098	0.987
	Overall Desirability				
ECOTAN BIO 100	t	min	1.0	34.399	0.951
	s	min	1.0	55.619	0.888
	T	min	1.0	20.819	0.514
	iniTurb	max	1.0	184.576	0.998
	cuaDose	inRange	1.0	3.490	1.000
	finTurb	min	1.0	28.824	0.621
	TSS	min	1.0	441.478	0.501
	Overall Desirability				
ECOTAN BIO G150	t	min	1.0	51.271	0.764
	s	min	1.0	50.858	0.983
	T	min	1.0	13.779	0.712
	iniTurb	max	1.0	173.808	0.939
	cuaDose	inRange	1.0	3.781	1.000
	finTurb	min	1.0	18.325	0.87
	TSS	min	1.0	328.925	0.576
	Overall Desirability				

Table 17. The fourth clarification optimization scenario: the turbidity and TSS are minimized when sample water is very cold and turbid.

Coagulant	Parameters	Goal	Importance	Value	Desirability
ECOTAN BIO 90D	t	inRange	1.0	81.887	1.000
	s	inRange	1.0	53.278	1.000
	T	min	1.0	8.567	0.980
	iniTurb	max	1.0	160.591	0.966
	cuaDose	inRange	1.0	5.267	1.000
	finTurb	min	1.0	1.213	1.000
	TSS	min	1.0	57.980	1.000
	Overall Desirability				
ECOTAN BIO 100	t	inRange	1.0	110.137	1.000
	s	inRange	1.0	88.552	1.000
	T	min	1.0	9.751	0.977
	iniTurb	max	1.0	173.27	0.933
	cuaDose	inRange	1.0	1.110	1.000
	finTurb	min	1.0	24.719	0.689
	TSS	min	1.0	429.117	0.515
	Overall Desirability				
ECOTAN BIO G150	t	inRange	1.0	81.887	1.000
	s	inRange	1.0	53.278	1.000
	T	min	1.0	10.261	0.980
	iniTurb	max	1.0	177.957	0.966
	cuaDose	inRange	1.0	5.267	1.000
	finTurb	min	1.0	20.331	0.831
	TSS	min	1.0	281.105	0.639
	Overall Desirability				

Table 18. The fifth clarification optimization scenario: the turbidity and TSS are minimized when the sample water is turbid and hot.

Coagulant	Parameters	Goal	Importance	Value	Desirability
ECOTAN BIO 90D	t	inRange	1.0	45.363	1.000
	s	inRange	1.0	79.840	1.000
	T	max	1.0	20.926	0.942
	iniTurb	max	1.0	136.59	0.966
	cuaDose	inRange	1.0	5.455	1.000
	finTurb	min	1.0	3.964	0.997
	TSS	min	1.0	4.730	1.000
	Overall Desirability				
ECOTAN BIO 100	t	inRange	1.0	92.101	1.000
	s	inRange	1.0	78.240	1.000
	T	max	1.0	31.458	0.931
	iniTurb	max	1.0	138.785	0.964
	cuaDose	inRange	1.0	1.197	1.000
	finTurb	min	1.0	13.487	0.874
	TSS	min	1.0	20.112	1.000
	Overall Desirability				
ECOTAN BIO G150	t	inRange	1.0	62.737	1.000
	s	inRange	1.0	99.463	1.000
	T	max	1.0	20.050	0.767
	iniTurb	max	1.0	139.986	0.980
	cuaDose	inRange	1.0	2.5140	1.000
	finTurb	min	1.0	5.4220	1.000
	TSS	min	1.0	194.109	0.752
	Overall Desirability				

After the different clarification optimization scenarios were obtained, five new experiments were conducted using the combination of process variables that appear in Tables 14–18 prepared in order to determine the accuracy of the proposed methodology. Tables 19–21 show the values of different outputs variables (finTurb and TSS) according to the five clarification optimization scenarios that were studied for each of the three emerging coagulants. These tables show that the experimentally obtained values for the five clarification optimization scenarios for the different coagulants that were studied do not differ significantly from those of the MRS methodology produced (see the results of Tables 14–18). To compare the different errors that arise in predicting outputs in this case in accordance with the five design requirements, the normalized data provide MAE and RMSE. Normalizing the data in statistical processes is a common practice. This ensures that the same scale is used by all variables (i.e., 0–1). To normalize the data in this case, the minimum value was subtracted from each original value and the result was divided by the range of each variable, as seen in Equation (14):

$$Y_{k, \text{norm}} = \frac{Y_k - \min(Y)}{\text{range}(Y)} \quad (14)$$

where $Y_{k, \text{norm}}$ are the normalized, experimental outputs of the models that were developed using RSM. The errors in the last columns concern the MAE and RMSE that were normalized for each variable in each clarification optimization scenario to be studied. The normalized MAE and RMSE that appear in the last two rows relate to the errors in the outputs variables that were examined. For example, Tables 19–21 show that the output variable that produce the greatest error in prediction for each of the three natural coagulants studied is the TSS variable (ECOTAN BIO 90D: MAE = 0.08, ECOTAN BIO 100: MAE = 0.08 and ECOTAN BIO G150: MAE = 0.29). These values correspond to the errors that appear in Table 13. In the latter table, the maximum MAE corresponds to the output variable TSS for the three natural coagulants that were studied. However, in Tables 19–21, the third optimization scenario has the highest MAE and RMSE values for the three studied coagulants. The reason for this

may be that the samples of raw water in this optimization scenario required a high turbidity to analyze each of the natural coagulants to obtain the highest efficiency of turbidity removal and TSS. In this case, a greater initial turbidity of the raw water sample, also as a greater requirement in turbidity removal, could produce this maximum MAE for this scenario.

Table 19. Experimental outputs that were obtained according to the five clarification optimization scenarios for ECOTAN BIO 90D.

Opt. Scenario	Experimental Values Obtained			
	finTurb	TSS	MAE	RMSE
1st Scenario	0.86	0.02	0.07	0.10
2nd Scenario	0.48	0.76	0.10	0.10
3rd Scenario	0.65	0.72	0.12	0.13
4th Scenario	0.01	0.90	0.05	0.07
5th Scenario	0.20	0.01	0.03	0.04
MAE	0.07	0.08		
RMSE	0.08	0.10		

Table 20. Experimental outputs that were obtained according to the five clarification optimization scenarios for ECOTAN BIO 100.

Opt. Scenario	Experimental Values Obtained			
	finTurb	TSS	MAE	RMSE
1st Scenario	0.87	0.05	0.08	0.09
2nd Scenario	0.82	0.06	0.10	0.13
3rd Scenario	0.62	0.80	0.11	0.14
4th Scenario	0.48	0.84	0.08	0.10
5th Scenario	0.03	0.01	0.02	0.02
MAE	0.07	0.08		
RMSE	0.10	0.11		

Table 21. Experimental outputs that were obtained according to the five clarification optimization scenarios for ECOTAN G150.

Opt. Scenario	Experimental Values Obtained			
	finTurb	TSS	MAE	RMSE
1st Scenario	0.98	0.62	0.20	0.27
2nd Scenario	0.37	0.55	0.18	0.25
3rd Scenario	0.48	0.57	0.21	0.25
4th Scenario	0.37	0.75	0.12	0.13
5th Scenario	0.03	0.20	0.12	0.15
MAE	0.05	0.29		
RMSE	0.06	0.30		

4. Conclusions

Physical–chemicals methods are usually used for waste water treatment processes. One such physical–chemical method is coagulation, and determining the amount of coagulant that is necessary to clarify a waste water sample is a complicated task that is usually done with a jar test. This article describes a methodology to obtain the optimal combination of the clarification process parameters for efficient removal of turbidity (finTurb), and TSS from a waste water sample collected from the WWTP in Logroño (Spain). In this case, the clarification process parameters considered were: initial turbidity (iniTurb), dosage of natural coagulant (cuaDose), process temperature (T), mixing speed (s) and mixing time (t). The methodology proposed is based on the RSM with desirability functions, and was applied

to three natural coagulants (ECOTAN BIO 90D, ECOTAN BIO 100 and ECOTAN G150). A CCD was used to study the effects of the process parameters and inputs variables on the water clarification process through quadratic regression models. An ANOVA test was performed on the regression models obtained, and it was observed that the process parameter T has no influence on finTurb or on TSS for the studied coagulants. Thus, these parameters could be eliminated in this study. For the natural coagulant ECOTAN BIO G150, iniTurb and cuaDose exercise a strong influence on finTurb and TSS, in addition to those variables mentioned above: s and t. Finally, the optimal combination of process parameters according to five optimization scenarios was determined by applying MRS based on desirability functions. The optimization study reveals that ECOTAN BIO 90D is more efficient than ECOTAN BIO 100 and ECOTAN G150 for removal of turbidity. ECOTAN BIO 90D is recommended for clarification of high-turbidity water. For water that is turbid and hot, both ECOTAN BIO 90D and ECOTAN BIO 100 should be considered. A compromise was made of the removal efficiencies of turbidity and TSS, along with the coagulant dose required. If other parameters are taken into account, such as the low water temperature of samples with high turbidity, the ECOTAN BIO 90D is the coagulant that will demonstrate greatest efficiency in removal of turbidity and TSS. Of the five optimization scenarios that were analyzed, ECOTAN G150 is the natural coagulant that is least efficient in eliminating both turbidity and TSS. This paper has demonstrated that using RMS with desirability functions can be an efficient method for use in selecting and adjusting the parameters of a clarification process using different coagulants, in a short period of time and with the minimal experimentation necessary.

Author Contributions: Experimental work: M.C.B., R.L.L., F.S.G., and E.P.V.G. Contributions to development of predictive models and optimization: R.E.G. Results analysis, revision and improvement of the manuscript: all authors.

Funding: This research received no external funding.

Acknowledgments: The authors wish to thank Servyeco Group for partial support of this study (<https://www.servyeco.com/>).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Corcoran, E.; Nellemann, C.; Baker, E.; Bos, R.; Osborn, D.; Savelli, H. *Sick Water? The Central Role of Wastewater Management in Sustainable Development. A Rapid Response Assessment*; UNEP: Nairobi, Kenya, 2010.
2. Choy, S.Y.; Prasad, K.M.N.; Wu, T.Y.; Ramanan, R.N. A review on common vegetables and legumes as promising plant-based natural coagulants in water clarification. *Int. J. Environ. Sci. Technol.* **2015**, *12*, 367–390. [[CrossRef](#)]
3. Teh, C.Y.; Budiman, P.M.; Shak, K.P.Y.; Wu, T.Y. Recent advancement of coagulation–flocculation and its application in wastewater treatment. *Ind. Eng. Chem. Res.* **2016**, *55*, 4363–4389. [[CrossRef](#)]
4. Lee, K.E.; Morad, N.; Teng, T.T.; Poh, B.T. Development, characterization and the application of hybrid materials in coagulation/flocculation of wastewater: A review. *Chem. Eng. J.* **2012**, *203*, 370–386. [[CrossRef](#)]
5. Boisvert, J.P.; To, T.C.; Berrak, A.; Jolicoeur, C. Phosphate adsorption in flocculation processes of aluminum sulphate and poly-aluminum-silicate-sulphate. *Water Res.* **1997**, *31*, 1939–1946. [[CrossRef](#)]
6. Dawood, A.S.; Li, Y. Modeling and optimization of new flocculant dosage and pH for flocculation: Removal of pollutants from wastewater. *Water* **2013**, *5*, 342–355. [[CrossRef](#)]
7. Ahmad, A.; Ismail, S.; Bhatia, S. Optimization of coagulation-flocculation process for palm oil mill effluent using response surface methodology. *Environ. Sci. Technol.* **2005**, *39*, 2828–2834. [[CrossRef](#)] [[PubMed](#)]
8. Gao, B.Y.; Yue, Q.Y.; Wang, B.J.; Chu, Y.B. Poly-aluminum-silicate-chloride (PASiC) a new type of composite inorganic polymer coagulant. *Colloids Surf. A Physicochem. Eng. Asp.* **2003**, *229*, 121–127. [[CrossRef](#)]
9. Sinha, S.; Yoon, Y.; Amy, G.; Yoon, J. Determining the effectiveness of conventional and alternative coagulants through effective characterization schemes. *Chemosphere* **2004**, *57*, 1115–1122. [[CrossRef](#)] [[PubMed](#)]

10. Subramonian, W.; Wu, T.Y.; Chai, S.P. An application of response surface methodology for optimizing coagulation process of raw industrial effluent using *Cassia obtusifolia* seed gum together with alum. *Ind. Crops Prod.* **2015**, *70*, 107–115. [[CrossRef](#)]
11. Wang, J.P.; Chen, Y.Z.; Wang, Y.; Yuan, S.J.; Yu, H.Q. Optimization of the coagulation-flocculation process for pulp mill wastewater treatment using a combination of uniform design and response surface methodology. *Water Res.* **2014**, *45*, 5633–5640. [[CrossRef](#)] [[PubMed](#)]
12. Ahmad, A.L.; Yasin, N.M.; Derek, C.J.C.; Lim, J.K. Optimization of microalgae coagulation process using chitosan. *Chem. Eng. J.* **2011**, *173*, 879–882. [[CrossRef](#)]
13. Hu, C.Z.; Liu, H.J.; Qu, J.H.; Wang, D.S.; Ru, J. Coagulation Behavior of Aluminum Salts in Eutrophic Water: Significance of Al13 Species and pH Control. *Environ. Sci. Technol.* **2006**, *40*, 325–331. [[CrossRef](#)] [[PubMed](#)]
14. Wang, D.S.; Tang, H.X.; Gregory, J. Relative Importance of Charge Neutralization and Precipitation on Coagulation of Kaolin with PACl: Effect of Sulfate Ion. *Environ. Sci. Technol.* **2002**, *36*, 1815–1820. [[CrossRef](#)] [[PubMed](#)]
15. Teh, C.Y.; Wu, T.Y.; Juan, J.C. Potential use of rice starch in coagulation-flocculation process of agro-industrial wastewater: Treatment performance and flocs characterization. *Ecol. Eng.* **2014**, *71*, 509–519. [[CrossRef](#)]
16. Miller, S.M.; Fugate, E.J.; Craver, V.O.; Smith, J.A.; Zimmerman, J.B. Toward Understanding the Efficacy and Mechanism of *Opuntia* spp. as a Natural Coagulant for Potential Application in Water Treatment. *Environ. Sci. Technol.* **2008**, *42*, 4274–4279. [[CrossRef](#)] [[PubMed](#)]
17. Syu, M.; Chen, B.; Chou, S. A Study on the Sedimentation Model and Neural Network Online Adaptive Control of a Benzoic Acid Imitated Wastewater Oxidation Process. *Ind. Eng. Chem. Res.* **2003**, *42*, 6862–6871. [[CrossRef](#)]
18. Yu, W.Z.; Gregory, J.; Campos, L.; Li, G.B. The role of mixing conditions on floc growth, breakage and re-growth. *Chem. Eng. J.* **2011**, *171*, 425–430. [[CrossRef](#)]
19. Wang, Y.; Chen, K.; Mo, L.; Li, J.; Xu, J. Optimization of coagulation-flocculation process for papermaking-reconstituted tobacco slice wastewater treatment using response surface methodology. *J. Ind. Eng. Chem.* **2014**, *20*, 391–396. [[CrossRef](#)]
20. Wu, T.Y.; Mohammad, A.W.; Jahim, J.M.; Anuar, N. Optimized reused and bioconversion from retentate of pre-filtered palm oil mill effluent (POME) into microbial protease by *Aspergillus terreus* using response surface methodology. *J. Chem. Technol. Biotechnol.* **2009**, *84*, 1390–1396. [[CrossRef](#)]
21. Teh, C.Y.; Wu, T.Y. The potential use of natural coagulants and flocculants in the treatment of urban waters. *Chem. Eng. Trans.* **2014**, *39*, 1603–1608.
22. Granata, F.; Papirio, S.; Esposito, G.; Gargano, R.; de Marinis, G. Machine learning algorithms for the forecasting of wastewater quality indicators. *Water* **2017**, *9*, 105. [[CrossRef](#)]
23. Shak, K.P.Y.; Wu, T.Y. Optimized use of alum together with unmodified *Cassia obtusifolia* seed gum as a coagulant aid in treatment of palm oil mill effluent under natural pH of wastewater. *Ind. Crop. Prod.* **2015**, *76*, 1169–1178. [[CrossRef](#)]
24. Witek-Krowiak, A.; Chojnacka, K.; Podstawczyk, D.; Dawiec, A.; Pokomeda, K. Application of response surface methodology and artificial neural network methods in modelling and optimization of biosorption process. *Bioresour. Technol.* **2014**, *160*, 150–160. [[CrossRef](#)] [[PubMed](#)]
25. Sedano, J.; Curiel, L.; Corchado, E.; de la Cal, E.; Villar, J. A soft computing method for detecting lifetime building thermal insulation failures. *Integr. Comput. Aided Eng.* **2010**, *17*, 103–115. [[CrossRef](#)]
26. Lostado, R.; Fernández, R.; Mac Donald, B.J.; Villanueva, P.M. Combining soft computing techniques and the finite element method to design and optimize complex welded products. *Integr. Comput. Aided Eng.* **2015**, *22*, 153–170. [[CrossRef](#)]
27. Ratnaweera, H.; Fettig, J. State of the art of online monitoring and control of the coagulation process. *Water* **2015**, *7*, 6574–6597. [[CrossRef](#)]
28. Baxter, C.W.; Zhang, Q.; Stanley, S.J.; Shariff, R.; Tupas, R.R.; Stark, H.L. Drinking water quality and treatment: The use of artificial neural networks. *Can. J. Civ. Eng.* **2001**, *28*, 26–35. [[CrossRef](#)]
29. Juntunen, P.; Liukkonen, M.; Pelo, M.; Lehtola, M.J.; Hiltunen, Y. Modelling of water quality: An application to a water treatment process. *Appl. Comput. Intell. Soft Comput.* **2012**, *4*, 14. [[CrossRef](#)]
30. Zhang, K.; Achari, G.; Li, H.; Zargar, A.; Sadiq, R. Machine learning approaches to predict coagulant dosage in water treatment plants. *Int. J. Syst. Assur. Eng. Manag.* **2013**, *4*, 205–214. [[CrossRef](#)]

31. Trinh, T.K.; Kang, L.S. Response surface methodological approach to optimize the coagulation–flocculation process in drinking water treatment. *Chem. Eng. Res. Des.* **2011**, *89*, 1126–1135. [[CrossRef](#)]
32. Khayet, M.; Zahrim, A.Y.; Hilal, N. Modelling and optimization of coagulation of highly concentrated industrial grade leather dye by response surface methodology. *Chem. Eng. J.* **2011**, *167*, 77–83. [[CrossRef](#)]
33. Li, N.; Hu, Y.; Lu, Y.Z.; Zeng, R.J.; Sheng, G.P. Multiple response optimization of the coagulation process for upgrading the quality of effluent from municipal wastewater treatment plant. *Sci. Rep.* **2016**, *6*, 26115. [[CrossRef](#)] [[PubMed](#)]
34. Available online: <http://www.grupo-selecta.com/es/catalogo/subcategorias/5/AGITADORES> (accessed on 10 June 2018).
35. ASTM D2035-13 Standard Practice for Coagulation-Flocculation Jar Test of Water. Available online: <https://www.astm.org/Standards/D2035.htm> (accessed on 10 March 2018).
36. ASTM D1889-00 Standard Test Method for Turbidity of Water (Withdrawn 2007). Available online: <https://www.astm.org/Standards/D1889.htm> (accessed on 10 March 2018).
37. ASTM D5907-10 Standard Test Methods for Filterable Matter (Total Dissolved Solids) and Nonfilterable Matter (Total Suspended Solids) in Water. Available online: <https://www.astm.org/Standards/D5907.htm> (accessed on 15 March 2018).
38. Box, G.E.; Wilson, K.B. On the experimental attainment of optimum conditions. *J. R. Stat. Soc. Ser. B* **1951**, *13*, 1–45. [[CrossRef](#)]
39. Lostado, R.; García, R.E.; Martínez, R.F. Optimization of operating conditions for a double-row tapered roller bearing. *Int. J. Mech. Mater. Des.* **2016**, *12*, 353–373. [[CrossRef](#)]
40. Lostado, R.; Escribano, R.; Martínez, M.A.; Múgica, R. Improvement in the Design of Welded Joints of EN 235JR Low Carbon Steel by Multiple Response Surface Methodology. *Metals* **2016**, *6*, 205. [[CrossRef](#)]
41. Moussaoui, K.; Mousseigne, M.; Senatore, J.; Chieragatti, R.; Lamesle, P. Influence of Milling on the Fatigue Lifetime of a Ti6Al4V Titanium Alloy. *Metals* **2015**, *5*, 1148–1162. [[CrossRef](#)]
42. Gopalsamy, B.M.; Mondal, B.; Ghosh, S. Taguchi method and ANOVA: An approach for process parameters optimization of hard machining while machining hardened steel. *J. Sci. Ind. Res. India* **2009**, *68*, 686–695.
43. Harrington, E.C. The desirability function. *Ind. Qual. Control* **1965**, *21*, 494–498.
44. Kuhn, M. Desirability: Desirability Function Optimization and Ranking. R package v.1.6. Available online: <http://CRAN.R-project.org/package=desirability> (accessed on 25 August 2016).
45. R Core Team R: A Language and Environment for Statistical Computing. R foundation for Statistical Computing. Available online: <https://www.r-project.org/> (accessed on 22 January 2018).
46. Montgomery, D.C. *Design and Analysis of Experiments*; John Wiley & Sons: New York, NY, USA, 2008.
47. Box, G.E.; Behnken, D.W. Some new three level designs for the study of quantitative variables. *Technometrics* **1960**, *2*, 455–475. [[CrossRef](#)]
48. Lenth, R.V. Response-Surface Methods in R, Using RSM. *J. Stat. Softw.* **2009**, *32*, 1–17. [[CrossRef](#)]
49. Sathiya, P.; Aravindan, S.; Haq, A.N. Optimization for friction welding parameters with multiple performance characteristics. *Int. J. Mech. Mater. Des.* **2006**, *3*, 309–318. [[CrossRef](#)]
50. Wilkinson, L.; Dallal, G.E. Tests of significance in forward selection regression with an F-to enter stopping rule. *Technometrics* **1981**, *23*, 377–380.
51. Max, K. Desirability Function Optimization and Ranking. R package Version 1.9. 2015. Available online: <https://CRAN.R-project.org/package=desirability> (accessed on 28 March 2017).

