

Article

Quantitative Analysis of the Balance Property in Factorial Experimental Designs 2^4 to 2^8

Ricardo Ramírez-Tapia ¹, Armando Javier Ríos-Lira ^{2,*},, Yaquelin Verenice Pantoja-Pacheco ²,
José Antonio Vázquez-López ² and Edgar Augusto Ruelas-Santoyo ²

¹ Doctorado en Ciencias de la Ingeniería, Tecnológico Nacional de México en Celaya, Celaya 38010, Guanajuato, Mexico

² Tecnológico Nacional de México/Instituto Tecnológico de Celaya, Celaya 38010, Guanajuato, Mexico

* Correspondence: armando.rios@itcelaya.edu.mx

† Departamento de Ingeniería Industrial, Tecnológico Nacional de México/Instituto Tecnológico de Celaya, Antonio García Cubas No. 600, Fovissste, Celaya 38010, Guanajuato, Mexico.

Abstract: Experimental designs are built by using orthogonal balanced matrices. Balance is a desirable property that allows for the correct estimation of factorial effects and prevents the identity column from aliasing with factorial effects. Although the balance property is well known by most researchers, the adverse effects caused by the lack of balance have not been extensively studied or quantified. This research proposes to quantify the effect of the lack of balance on model term estimation errors: type I error, type II error, and type I and II error as well as R^2 , R^2_{adj} , and R^2_{pred} statistics under four balance conditions and four noise conditions. The designs considered in this research include 2^4 – 2^8 factorial experiments. An algorithm was developed to unbalance these matrices while maintaining orthogonality for main effects, and the general balance metric was used to determine four balance levels. True models were generated, and a MATLAB program was developed; then a Monte Carlo simulation process was carried out. For each true model, 50,000 replications were performed, and percentages for model estimation errors and average values for statistics of interest were computed.

Keywords: design of experiments; design matrix; balance; orthogonality; general balance metric

MSC: 62K15; 65C05; 62J10



Citation: Ramírez-Tapia, R.; Ríos-Lira, A.J.; Pantoja-Pacheco, Y.V.; Vázquez-López, J.A.; Ruelas-Santoyo, E.A. Quantitative Analysis of the Balance Property in Factorial Experimental Designs 2^4 to 2^8 . *Mathematics* **2022**, *10*, 3812. <https://doi.org/10.3390/math10203812>

Academic Editor: Kai-Tai Fang

Received: 27 August 2022

Accepted: 21 September 2022

Published: 15 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

An experiment can be defined as a test or a series of tests in which deliberate changes are made to the input variables of a process to observe the reasons for the changes in the output variable [1]. Figure 1 shows the representation of a process where inputs are affected by both controllable (x_p) and uncontrollable (z_q) factors, having an impact on the output or response variable (y).

Factorial designs are commonly used because they contain all possible combinations of factors levels and can estimate factorial effects and interactions. In addition, they are commonly used as factor-filtering designs to identify the factors that have the greatest significance or influence on the response variable.

Two commonly used types of factorial designs include the 2^k and 3^k (where k represents the number of factors to evaluate, and the base represents the number of levels). Other designs include two-level fractional factorial designs, three-level fractional factorial designs, and mixed level designs. This research focuses only on two-level factorial designs, also called 2^k designs.

The 2^k factorial design is a design in which each factor has two levels commonly called high and low, represented by 1 and -1 respectively. These designs can be used to analyze

both quantitative and qualitative factors [1]. The design matrix is an orthogonal array with 2^k runs and k columns; this is illustrated in Table 1.

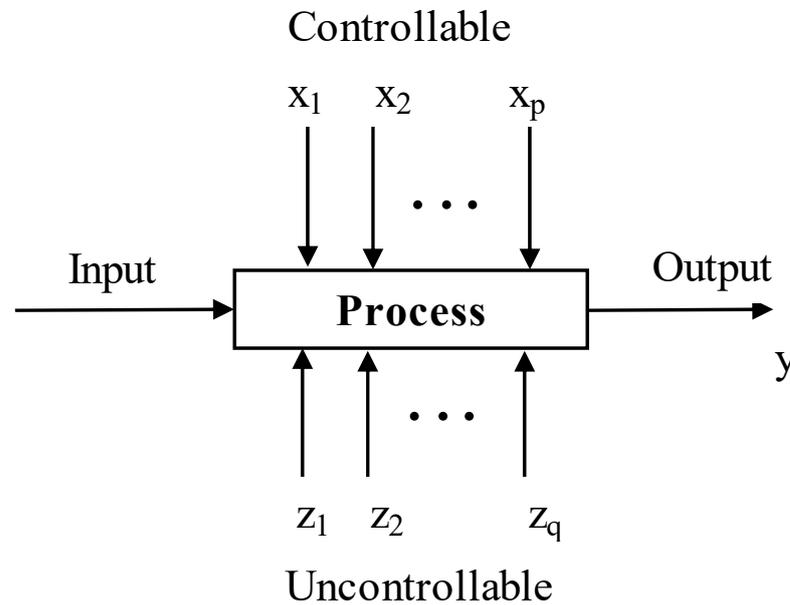


Figure 1. The general model of a process.

Table 1. Example of a 2^3 factorial design matrix.

Run	Factors		
	A	B	C
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1

A design is balanced when each column contains the same number of runs for each factor level [2]. According to Birkes et al. (1999), the balance exists when there is an equal number of observations for all combinations of factor levels [3]. Ríos (2009) mentions that balance is an important property in experimental design, and a design is balanced when in each column each possible factor level appears the same number of times [4].

Balanced experiments have two major advantages when it comes to analysis. First, each experimental combination is estimated with the same precision. Secondly, in a balanced factorial experiment, the effect of each factor can be evaluated independently from other factors [5]. Pantoja et al. (2009), point out that balance is an important property because it prevents the main effects and interactions from aliasing with the interception column β_0 [6]. Balance also plays another important role; it leads to a considerable simplification of the calculations [7].

Balance property has been extensively mentioned in literature. In 1935, the term balance was already included in articles published by Yates; in these articles, it was mentioned that to avoid sacrificing all the information about a possibly important interaction, a balanced arrangement was used wherein the interaction was confused as little as possible.

In 1938, Yates conducted research evaluating the extent to which a balanced experimental design on human nutrition resulted in an efficiency gain, compared to several alternative randomized block designs that could have been used to solve the same problem [8]. In 1965, Nelder introduced the concept of general equilibrium for multi-stratum experiments. He mentions that only in general equilibrium designs can information on any treatment effect be obtained and combined from more than one [9]. In 1972, John and Smith introduced the concept of orthogonal factorial structure for factorial designs; they mentioned that if all the estimation is done in the lowest stratum, only in these designs can the factorial effects be estimated independently [10]. Because of this feature, several later research papers were restricted to this type of design (see, for example, Cotter et al. (1973) [11], John (1981) [12], Mukerjee (1979) [13], Mukerjee (1981) [14] and Gupta (1983) [15]). In 1978, different existing methods for analyzing unbalanced designs were evaluated by using the computer programs of the time [16]. In 1986, David G. Herr published an article in which he talks about the history of unbalanced designs and the methods to deal with them [17].

In 1993, Shaw and Mitchel mention that unfortunately, for various reasons, it is rare for a biological or ecological study to have fully balanced data, and that unbalance requires care in analysis and interpretation. They also try to explain some of the consequences of unbalance and give some guidelines for analyzing unbalanced data for models involving fixed effects [18]. In 1996, Huber and Zwerina mentioned the importance of balance in optimal experimental designs [19]. In 2001, balanced experimental designs were used in conjunction with microarray models to choose an optimal experimental design suitable for gene expression [20]. In 2009, Guo et al. presented the general metric balance, a parameter that explains how to measure the degree of balance in a factorial design [21].

In 2010, Andy Hector et al. performed a study of variance for unbalanced data in the areas of ecology and evolution. They summarized the main developments to deal with unbalanced designs and highlighted the search for the right ANOVA method by which to present one or several models that best fit the objectives of the analysis [22].

Ríos et al. (2011) present a sequential experimentation approach to increase Resolution III fractions. This method was able to overcome the drawbacks of the general methods while maintaining some of their benefits. They mention the importance of maintaining balance and orthogonality [23]. Landsheer and Van Den Wittenboer in 2015 compared balanced and unbalanced 2×2 designs. They used ANOVA correction methods with the sum of squares type II and type III. The purpose was to determine which type of SS provides lower H_0 rejection rates, in their study. They compared 2×2 balanced and 2×2 unbalanced designs with an interaction to determine whether ANOVA correction methods provide satisfactory results in the presence of an interaction. In this study, they compared the methods for calculating ANOVA with balanced data, i.e., with the SS I method, to calculate ANOVA with unbalanced designs they used the ANOVA SS II and SS III methods. They concluded that ANOVA with SS II only works satisfactorily for unbalanced designs when an interaction can be excluded, whereas the application of SS III, when there is an interaction, showed a power about 1% to 5% lower than for the balanced datasets that were simulated. The application of SS II cannot be recommended when there is an interaction in an unbalanced design [24].

Voelkel in 2019 uses systematic methods to create optimal designs by using the balance property and orthogonal arrays [25]. Balanced experiments have two major advantages when it comes to analysis. First, each experimental combination is estimated with the same precision. Secondly, in a balanced factorial experiment, the effect of the intercept or identity column can be evaluated independently from the other factors [5].

Currently, one of the key factors for the success of any organization, whether industrial or service, is to improve and make its products and processes more efficient. This is where experimentation comes in, as it is one of the elements that can contribute the most to the improvement of products and processes. The use of design of experiments (DOE) is an effective tool for understanding and improving processes and products in the industry.

However, in industrial experimentation, it is not uncommon to find that companies resort to unbalanced experimental designs. These designs usually occur due to several circumstances. One such circumstance may arise when the experimenter has made a mistake at the time of experimenting or because some experimental conditions cannot be created once the experiment is running. For example, this often occurs when a machine is unavailable due to a breakdown or failure, or when a batch of raw material is unavailable, either due to a shortage of material or a change of supplier. Another event that happens in industries is the fact that an operator who was being taken as a level of the experiment is missing. It also happens in industries such as the agri-food industry that, due to weather conditions or natural disasters, crops that could be under experimentation are lost or pests arise that are difficult to destroy.

In situations such as this, experimenters may be interested in knowing the consequences of running an unbalanced design. With this research, experimenters will be able to answer whether it is feasible to run an unbalanced design and the consequences of running it. Experimenters will also be able to find out what levels of imbalance are most favorable for the experiment if it is necessary to use an unbalanced design.

In this investigation, we worked with designs 2^4 to 2^8 by using four different balance levels. The levels included: balanced, low unbalance, medium unbalance, and high unbalance. The general balance metric (*GBM*) was used to measure the balance property given that it is simple and effective minimum aberration criterion that measures the balance property of experimental designs. *GBM* is a parameter that determines the degree of balance of a given design matrix. According to Guo et al. (2009), *GBM* can be calculated for a matrix d of size $n \times k$, where n is the number of rows and k is the number of columns, $d^t (t = 1, \dots, k)$ indicates the columns of interactions for the factors from 1 to k , and d^1 represents the main effects matrix [21].

As shown in Equation (1), *GBM* considers the difference between the number of times that a factor level appears in a column and the number of times that it should appear if the design were balanced. Equation (2) shows how these quantities are squared and added for each submatrix of main effects two-factor interactions and three-factor interactions. Finally, an array consisting of k numbers is obtained (Equation (3)) where H^1 indicates the balance for main effects, H^2 for 2-factor interactions, and H^k for k -factor interactions. When a design is balanced for main effects, two-factor interactions, and three-factor interactions, the *GBM* is a vector of zeros (0, 0, 0) [26]. Alternatively, these three quantities can be added to obtain a single number. If this number is zero, the model matrix is completely balanced. The higher this number, the more unbalanced the design will be.

Therefore,

$$H_j^t = \sum_{r=1}^{l_j^t} \left(C_{rj}^t - \frac{n}{l_j^t} \right)^2 \tag{1}$$

$$H^t = \sum_{j=1}^{\binom{t}{k}} H_j^t = \sum_{j=1}^{\binom{t}{k}} \sum_{r=1}^{l_j^t} \left(C_{rj}^t - \frac{n}{l_j^t} \right)^2 \tag{2}$$

$$GBM = (H^1, H^2, \dots, H^k), \tag{3}$$

where

C_{rj}^t is the number of times a level r appears in column j , and

l_j^t is the number of levels that the column j contains.

The design matrices are governed by another property in addition to balance. This property is known as orthogonality. Orthogonality ensures that the effects of the different factors to which the experimental material is subjected. Orthogonality can be estimated separately and without confusion [27]. In other words, orthogonality ensures that effects can be estimated independently so each column provides different information to the design [6].

The orthogonality property can be measured through several highly sophisticated criteria such as minimum generalized aberration [28], minimum aberration moment [29], moment aberration projection [30], J2-optimality [31], J2-modified optimality, variance inflation factors (*VIF*), dot product, simple correlation, and Pearson's correlation index. For practical purposes, we used *VIF* to measure orthogonality. *VIF* is a measure of the amount of multicollinearity in a set of multiple regression variables. Experience indicates that if any of the *VIF* are greater than 5, it is an indication that the associated regression coefficients are poorly estimated due to multicollinearity. *VIF* is calculated as follows,

$$VIF_i = \frac{1}{1 - R_i^2} \quad (4)$$

where R_i^2 is the coefficient of determination of the auxiliary regression of the variable x_i .

To complete this research, we made use of Monte Carlo simulation. The first step consisted of the creation of true models. A true model is an equation that relates a response variable, also known as dependent variable, to a set of regressors or independent variables [32] (see Equation (5)). Independent variables may also include two-factor interactions given that these interactions occur frequently in practice. To establish this relationship, two characteristics of the regression coefficients are determined by the experimenter, magnitude (size of the coefficient), and direction (sign of the coefficient). In addition, a random error from a normal distribution with mean zero and variance σ^2 is introduced to emulate the variability that occurs naturally in the process. The level of significance of the independent variables is directly related to the size of the regression coefficient with respect to the variance of the random error.

It is called a true model because the variables involved, the sizes and signs of the regression coefficients, the level of noise, and the level of significance are determined by the experimenter. True models are useful because they permit the user to simulate data, to manipulate this data, and to observe changes in the regression coefficient estimates with respect to their true values.

The general true model structure depicted in Equation (5) was chosen according to the sparsity of the effects principle, which establishes that most processes are affected by main effects and low order interactions (2FIs and 3FIs):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{123} x_1 x_2 x_3 + \varepsilon. \quad (5)$$

Several unbalanced matrices were created, and their response variables were generated by using the true models previously described; then these arrays were introduced in a MATLAB computer program to perform simulations. The Monte Carlo method is used to solve problems derived from stochastic phenomena, in which physical experimentation is impracticable, and the creation of an exact formula is impossible; it is also adapted to the resolution of complex deterministic problems [33].

The method by which the datasets were analyzed was stepwise regression. The stepwise regression procedure starts by choosing an equation that contains the best variable X and then tries to build with the subsequent additions of X 's one at a time, as long as it is worthwhile. The order of addition is determined by using the partial values of the F test to select which variable to enter next. The highest partial F -value is compared with an F -value to enter (selected or default). After a variable is added, the equation is examined to see if any variables should be removed [34]. In addition to regression coefficients estimates, several performance indicators were analyzed. These include the following.

(i) % No error: indicates the efficiency of the method in picking up the right model:

$$\% \text{ no error} = \frac{\text{number of models without error} \times 100}{\text{number of simulated models}}. \quad (6)$$

(ii) Type I error [30] in statistics means the number of insignificant terms in the model. % Type I error was calculated and measured as the percentage of times that the models showed insignificant terms during the simulations [35]:

$$\% \text{ Type I error} = \frac{\text{number of models with insignificant terms} \times 100}{\text{number of simulated models}}. \tag{7}$$

(iii) Type II error is the number of significant terms missing in the model [36]. % Type II measure the percentage of times that the models showed significant missing terms during the simulations:

$$\% \text{ Type II error} = \frac{\text{number of models with missing terms} \times 100}{\text{number of simulated models}}. \tag{8}$$

(iv) Type I and II error refers to a situation in which both errors are present. % Type I and II error measure the percentage of times that the models presented insignificant and significant missing terms in the model during the simulations:

$$\% \text{ Type I and II error} = \frac{\text{number of models with insignificant and significant missing terms} \times 100}{\text{number of simulated models}}. \tag{9}$$

(v) R^2 is defined as the proportion of the variability in the response that can be explained by the regressors included in the model (see Equation (10)). Clearly, values closer to 1 are more desirable. We have

$$R^2 = \frac{SS_{Model}}{SS_{Total}}, \tag{10}$$

where

SS_{Model} is the sum of squares corresponding to the model, and
 SS_{Total} is the total sum of squares.

The use of *AIC* and *BIC* statistics would be a good complement to determine if the best model has been chosen, but the computational complexity would increase, making the simulations longer; however it would be interesting for future research to see how these indicators behave in the absence of balance.

(vi) The R^2_{adj} is a variation of the ordinary R^2 statistic that reflects the number of factors in the model (Equation (11)). It can be a useful statistic for more complex experiments with several design factors when we wish to evaluate the impact of increasing or decreasing the number of model terms [1]. We have

$$R^2_{adj} = 1 - \frac{SS_E/df_E}{\frac{SS_{Total}}{df_{Total}}}, \tag{11}$$

where df_E and df_{Total} correspond to degrees of freedom (Equation (12)) from error and total degrees of freedom (Equation (13)), respectively. We have

$$df_E = n - k - 1 \tag{12}$$

$$df_{Total} = n - 1, \tag{13}$$

where:

n is the number of treatments in the experiment, and
 k is the number of significant terms in the regression model.

(vii) R^2 can be calculated for prediction based on the *PRESS* (see Equation (14)). Values closer to 1 for this statistic are desirable because this indicates that the model will give reasonable performance in prediction [1]. *PRESS* stands for prediction error sum of squares; it is a measure of how accurate the model is to predict future observations. Small values of *PRESS* are desirable. We have

$$R^2_{pred} = 1 - \frac{PRESS}{SS_{Total}}. \tag{14}$$

2. Method

This research was conducted in four stages, which are represented in Figure 2.

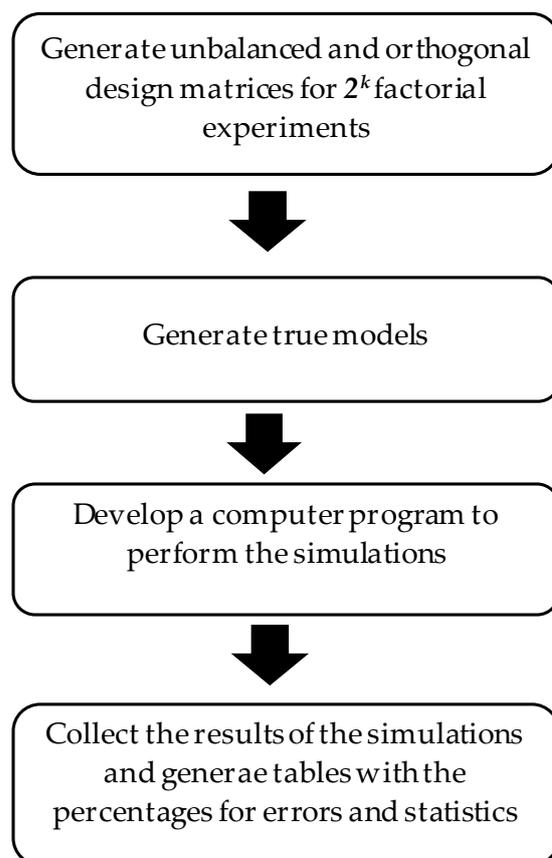


Figure 2. Research method.

2.1. Stage 1 Generate Unbalanced and Orthogonal Design Matrices for 2^k Experiments

The first stage consisted of generating arrays with four balance conditions that preserve the orthogonality property. Given that the objective was to quantify the effect of the lack of balance on the errors and statistics described in the previous section, we did not want orthogonality to be affected. By conducting the experiments in this way, we ensured that the effect on the statistics can be attributed purely to the lack of balance.

An algorithm was developed to alter the signs in a of two-level factorial experimental design to create arrays that are unbalanced but maintain orthogonality for main effects. This algorithm is based on the additive multiplicative properties of equality.

These properties refer to the fact that if we change the sign of the elements of an array, that is, multiplying by minus -1 each element in a row, the array will continue being orthogonal because each element in the row has been multiplied by the same sign, thus respecting the property of multiplicative equality. In addition, by changing the sign of each element of the row, the balance of the matrix will be affected. This procedure would unbalance a design matrix while maintaining the orthogonality property. The metric that we used to measure the level of balance in an array was *GBM*. By using this metric, we were able to determine different degrees for the lack of balance.

Figure 3 shows a design matrix with the respective *VIFs* and *GBM* to measure orthogonality and balance. In this figure, it can be seen that the design matrix is balanced because the *GBM* takes a value of 0; on the contrary, when the *GBM* takes a value greater than zero, the array is unbalanced. *VIFs* have also been observed in 1 so it is an orthogonal design.

		d1				
		A	B	C	Factor	VIF
		-1	-1	-1	A	1
		1	-1	-1	B	1
		-1	1	-1	C	1
		1	1	-1		
		-1	-1	1		
		1	-1	1		
		-1	1	1		
		1	1	1		
Ctj		4	4	4		
		4	4	4		
Htj		0	0	0		
Ht		0				
GBM		0				

Figure 3. Balanced design matrix 2³.

Figure 4 shows how the matrix can be unbalanced by multiplying 2 rows by -1 . This figure shows the design matrix shown in Figure 3 with two modified lines. The lines that were modified are 2 and 4 respectively. After the modifications, the GBM takes a value of 16 which indicates that the matrix is now unbalanced. The table also shows that the VIFs are less than or equal to 5; this means that the design matrix is still orthogonal for main effects.

		Unbalanced				
		d1				
		A	B	C	Factor	VIF
		-1	-1	-1	A	1.13
		-1	1	1	B	1.12
		-1	1	-1	C	1
		-1	-1	1		
		-1	-1	1		
		1	-1	1		
		-1	1	1		
		1	1	1		
Ctj		6	4	2		
		2	4	6		
Htj		8	0	8		
Ht		16				
GBM		16				

Figure 4. Unbalanced design matrix 2³.

An analysis was performed to determine the number of rows that should be multiplied by -1 to achieve a given level of unbalance. The levels of unbalance were defined as follows: low unbalance refers to the lowest unbalance that can be achieved by altering n_1 rows. High unbalance refers to the highest level of unbalance (when a design column becomes a column of 1s or -1 s values) that can be achieved by altering n_3 rows. Medium unbalance is achieved by altering n_2 rows where n_2 is the number between n_1 and n_3 . Table 2 shows these quantities. Note that the number of rows to modify is different depending on the level

of unbalance one wants to achieve and the experimental design in which one is working. In addition, Table 2 shows the expected GBM.

Table 2. Number of rows to be modified to achieve a desired level of unbalance.

Design	Low Unbalance	GBM Value	Medium Unbalance	GBM Value	High Unbalance	GBM Value
2 ³	1	8	2	16	3	24
2 ⁴	2	32	3	42	5	74
2 ⁵	5	140	9	268	13	396
2 ⁶	7	398	14	816	21	1246
2 ⁷	15	1808	30	3640	45	5488
2 ⁸	33	8210	67	16,514	101	24,898

Note that the rows to be modified do not follow a specific order. They can be chosen randomly from the design matrix. For example, if we want to create a medium level of unbalance in a design with five factors, we could invert signs in any subset of nine rows. This could be the first nine rows, the last nine rows, or any other possible combination.

2.2. Stage 2 True Model Generation

The true models generated were first-order models with two-factor and three-factor interactions; interactions between two and three factors were included given that these interactions are present in many real-world situations [37]. The factors and interactions included in the true models were those that presented unbalance after some rows were altered (multiplied by -1); all possible combinations for design size (five designs), level of unbalance (four levels) and level of noise (four levels), were analyzed, and three different models were created for each combination. These models were different regarding the number of factors involved, starting with a model with few terms, then a model with more terms, and finally a model with many terms. These conditions were established in order to observe how the errors behave in the presence of more or fewer factors in the model. In the end, 5 × 3 × 4 × 4 = 240 true models were generated.

The noise level was induced by modifying the size of the regression coefficients in the true model with respect to the error variance. A low noise level is obtained when the regression coefficients are approximately three times the error variance, a medium level is when this relation is 2 to 1, a high level is obtained when the relation becomes 1 to 1, and finally a very high noise level is obtained with a 1/2 to 1 relation; that is, regression coefficients that are half the error variance. If the error variance is fixed to 2, the low, medium, high, and very high noise levels are obtained by assigning values of 6, 4, 2, and 1 to the regression coefficients. The true models used for the different simulations are shown in Tables 3–6. Each model has been assigned an item consisting of a number and the initial of the noise level to which it belongs to identify it in later tables.

Table 3. True models with low noise.

Design	Item	True Models
2 ⁴	1L	Y = 10 - 6A - 6D + N(0,2)
	2L	Y = 10 - 6A - 6D + 6AC + N(0,2)
	3L	Y = 10 - 6A - 6D + 6AC + 6AD + 6ACD + N(0,2)
2 ⁵	4L	Y = 10 + 6A - 6D + 6E + 6AE + 6BD + N(0,2)
	5L	Y = 10 + 6A - 6D + 6E + 6AD + 6AE - 6BC + 6BD + N(0,2)
	6L	Y = 10 + 6A - 6D + 6E + 6AD + 6AE - 6BC + 6BD - 6DE + 6ADE + N(0,2)
2 ⁶	7L	Y = 10 - 6A + 6E - 6F + 6AC - 6AE + 6BE - 6BF + N(0,2)
	8L	Y = 10 - 6A + 6E - 6F + 6AC - 6AE + 6BE - 6BF + 6CF + 6DE + N(0,2)
	9L	Y = 10 - 6A + 6E - 6F + 6AC - 6AE + 6BE - 6BF + 6CF + 6DE - 6DF + 6AEF + N(0,2)

Table 3. Cont.

Design	Item	True Models
2 ⁷	10L	$Y = 10 + 6A - 6F + 6G - 6AE + 6AF + 6AG - 6BC - 6BD + 6BE + N(0,2)$
	11L	$Y = 10 + 6A - 6F + 6G - 6AE + 6AF + 6AG - 6BC - 6BD + 6BE + 6CF + 6DF + N(0,2)$
	12L	$Y = 10 + 6A - 6F + 6G - 6AE + 6AF + 6AG - 6BC - 6BD + 6BE + 6CF + 6DF - 6FG + 6AFG + N(0,2)$
2 ⁸	13L	$Y = 10 + 6A + 6G + 6H - 6AF - 6AG + 6AH + 6BE + 6BF - 6BG + 6CF - 6CG + N(0,2)$
	14L	$Y = 10 + 6A + 6G + 6H - 6AF - 6AG + 6AH + 6BE + 6BF - 6BG + 6CF - 6CG + 6DF - 6DH + N(0,2)$
	15L	$Y = 10 + 6A + 6G + 6H - 6AF - 6AG + 6AH + 6BE + 6BF - 6BG + 6CF - 6CG + 6DF - 6DH + 6GH + 6AGH + N(0,2)$

Table 4. True models with medium noise.

Design	Item	True Models
2 ⁴	1M	$Y = 10 - 4A - 4D + N(0,2)$
	2M	$Y = 10 - 4A - 4D + 4AC + N(0,2)$
	3M	$Y = 10 - 4A - 4D + 4AC + 4AD + 4ACD + N(0,2)$
2 ⁵	4M	$Y = 10 + 4A - 4D + 4E + 4AE + 4BD + N(0,2)$
	5M	$Y = 10 + 4A - 4D + 4E + 4AD + 4AE - 4BC + 4BD + N(0,2)$
	6M	$Y = 10 + 4A - 4D + 4E + 4AD + 4AE - 4BC + 4BD - 4DE + 4ADE + N(0,2)$
2 ⁶	7M	$Y = 10 - 4A + 4E - 4F + 4AC - 4AE + 4BE - 4BF + N(0,2)$
	8M	$Y = 10 - 4A + 4E - 4F + 4AC - 4AE + 4BE - 4BF + 4CF + 4DE + N(0,2)$
	9M	$Y = 10 - 4A + 4E - 4F + 4AC - 4AE + 4BE - 4BF + 4CF + 4DE - 4DF + 4AEF + N(0,2)$
2 ⁷	10M	$Y = 10 + 4A - 4F + 4G - 4AE + 4AF + 4AG - 4BC - 4BD + 4BE + N(0,2)$
	11M	$Y = 10 + 4A - 4F + 4G - 4AE + 4AF + 4AG - 4BC - 4BD + 4BE + 4CF + 4DF + N(0,2)$
	12M	$Y = 10 + 4A - 4F + 4G - 4AE + 4AF + 4AG - 4BC - 4BD + 4BE + 4CF + 4DF - 4FG + 4AFG + N(0,2)$
2 ⁸	13M	$Y = 10 + 4A + 4G + 4H - 4AF - 4AG + 4AH + 4BE + 4BF - 4BG + 4CF - 4CG + N(0,2)$
	14M	$Y = 10 + 4A + 4G + 4H - 4AF - 4AG + 4AH + 4BE + 4BF - 4BG + 4CF - 4CG + 4DF - 4DH + N(0,2)$
	15M	$Y = 10 + 4A + 4G + 4H - 4AF - 4AG + 4AH + 4BE + 4BF - 4BG + 4CF - 4CG + 4DF - 4DH + 4GH + 4AGH + N(0,2)$

Table 5. True models with high noise.

Design	Item	True Models
2 ⁴	1H	$Y = 10 - 2A - 2D + N(0,2)$
	2H	$Y = 10 - 2A - 2D + 2AC + N(0,2)$
	3H	$Y = 10 - 2A - 2D + 2AC + 2AD + 2ACD + N(0,2)$
2 ⁵	4H	$Y = 10 + 2A - 2D + 2E + 2AE + 2BD + N(0,2)$
	5H	$Y = 10 + 2A - 2D + 6E + 2AD + 2AE - 2BC + 2BD + N(0,2)$
	6H	$Y = 10 + 2A - 2D + 2E + 2AD + 2AE - 2BC + 2BD - 2DE + 2ADE + N(0,2)$
2 ⁶	7H	$Y = 10 - 2A + 2E - 2F + 2AC - 2AE + 2BE - 2BF + N(0,2)$
	8H	$Y = 10 - 2A + 2E - 2F + 2AC - 2AE + 2BE - 2BF + 2CF + 2DE + N(0,2)$
	9H	$Y = 10 - 2A + 2E - 2F + 2AC - 2AE + 2BE - 2BF + 2CF + 2DE - 2DF + 2AEF + N(0,2)$
2 ⁷	10H	$Y = 10 + 2A - 2F + 2G - 2AE + 2AF + 2AG - 2BC - 2BD + 2BE + N(0,2)$
	11H	$Y = 10 + 2A - 2F + 2G - 2AE + 2AF + 2AG - 2BC - 2BD + 2BE + 2CF + 2DF + N(0,2)$
	12H	$Y = 10 + 2A - 2F + 2G - 2AE + 2AF + 2AG - 2BC - 2BD + 2BE + 2CF + 2DF - 2FG + 2AFG + N(0,2)$
2 ⁸	13H	$Y = 10 + 2A + 2G + 2H - 2AF - 2AG + 2AH + 2BE + 2BF - 2BG + 2CF - 2CG + N(0,2)$
	14H	$Y = 10 + 2A + 2G + 2H - 2AF - 2AG + 2AH + 2BE + 2BF - 2BG + 2CF - 2CG + 2DF - 2DH + N(0,2)$
	15H	$Y = 10 + 2A + 2G + 2H - 2AF - 2AG + 2AH + 2BE + 2BF - 2BG + 2CF - 2CG + 2DF - 2DH + 2GH + 2AGH + N(0,2)$

Table 6. True models with very high noise.

Design	Item	True Models
2 ⁴	1VH	$Y = 10 - A - D + N(0,2)$
	2VH	$Y = 10 - A - D + AC + N(0,2)$
	3VH	$Y = 10 - A - D + AC + AD + ACD + N(0,2)$
2 ⁵	4VH	$Y = 10 + A - D + E + AE + BD + N(0,2)$
	5VH	$Y = 10 + A - D + E + AD + AE - BC + BD + N(0,2)$
	6VH	$Y = 10 + A - D + E + AD + AE - BC + BD - DE + ADE + N(0,2)$
2 ⁶	7VH	$Y = 10 - A + E - F + AC - AE + BE - BF + N(0,2)$
	8VH	$Y = 10 - A + E - F + AC - AE + BE - BF + CF + DE + N(0,2)$
	9VH	$Y = 10 - A + E - F + AC - AE + BE - BF + CF + DE - DF + AEF + N(0,2)$
2 ⁷	10VH	$Y = 10 + A - F + G - AE + AF + AG - BC - BD + BE + N(0,2)$
	11VH	$Y = 10 + A - F + G - AE + AF + AG - BC - BD + BE + CF + DF + N(0,2)$
	12VH	$Y = 10 + A - F + G - AE + AF + AG - BC - BD + BE + CF + DF - FG + AFG + N(0,2)$
2 ⁸	13VH	$Y = 10 + 2A + 2G + 2H - 2AF - 2AG + 2AH + 2BE + 2BF - 2BG + 2CF - 2CG + N(0,2)$
	14VH	$Y = 10 + A + G + H - AF - AG + AH + BE + BF - BG + CF - CG + DF - DH + N(0,2)$
	15VH	$Y = 10 + A + G + H - AF - AG + AH + BE + BF - BG + CF - CG + DF - DH + GH + AGH + N(0,2)$

2.3. Stage 3 MATLAB Simulations

This stage consisted in performing simulations for each of the true models generated under different levels of balance and noise. To do this, it was necessary to develop a computer program in MATLAB software. Each model was used to generate a response variable by using a Monte Carlo simulation, and 50,000 iterations were performed for each model. In each iteration, the regression coefficients were slightly varied at 10% above and below the original value. This was done to create a diversity of models. Stepwise regression was used to compute the regression coefficient estimates. Then the model errors were computed and averaged to obtain a single data point. This process is shown in Figure 5.

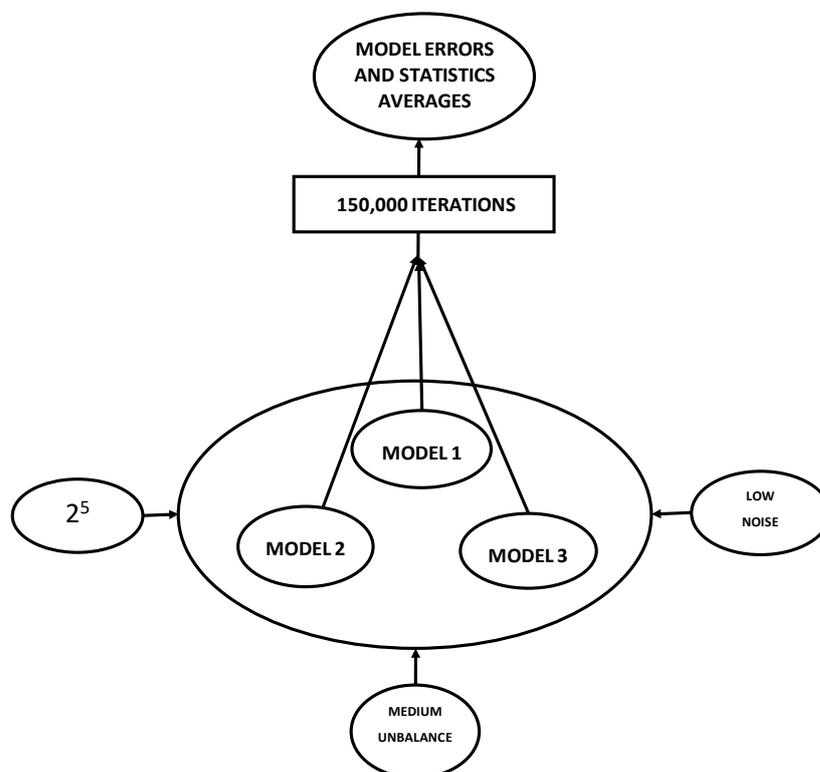


Figure 5. Simulation process.

The simulation strategy was carried out by using design matrices with balance, low unbalance, medium unbalance, and high unbalance. All of them are orthogonal for main effects. The number of iterations was determined by observing that as the number of iterations increased, convergence was achieved for the different errors. By convergence, we mean that the averages stabilized; that is, it did not change significantly as the number of simulations increased, indicating that the averages of the estimated errors were close to their real value. This was accomplished by simulating 50,000 iterations for each model.

3. Results of Simulations

This section presents the results of the simulations in a series of tables and graphs. The objective is to identify how the model errors behave under different conditions of balance and noise. The results are show in Tables 7–10.

Table 7. % Error behavior with low noise.

Design		2 ⁴			2 ⁵			2 ⁶			2 ⁷			2 ⁸		
True Model		1L	2L	3L	4L	5L	6L	7L	8L	9L	10L	11L	12L	13L	14L	15L
Balanced	No error	62	66	74	68	76	85	62	69	77	53	59	66	48	54	59
	Error I	38	34	25	32	24	15	38	31	23	47	41	34	52	46	41
	Error II	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	Error I and II	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Low	No error	63	67	38	70	9	0	63	69	77	54	60	65	1	2	0
	Error I	37	33	12	30	2	0	37	31	23	46	40	34	1	1	0
	Error II	0	0	3	0	0	85	0	0	0	0	0	0	0	0	60
	Error I and II	0	0	47	0	89	15	0	0	0	0	0	0	97	97	40
Unbalance level Medium	No error	64	68	44	0	0	0	62	69	77	49	55	0	0	0	0
	Error I	36	32	14	0	0	0	38	31	23	36	30	0	0	0	0
	Error II	0	0	37	71	81	86	0	0	0	0	0	66	52	58	61
	Error I and II	0	0	5	29	19	14	0	0	0	15	16	34	48	42	39
High	No error	50	46	0	0	0	0	62	69	77	0	0	0	0	0	0
	Error I	21	17	0	0	0	0	38	31	23	0	0	0	0	0	0
	Error II	0	0	78	1	81	86	0	0	0	56	63	66	51	57	60
	Error I and II	29	37	22	99	19	14	0	0	0	44	37	34	49	43	40

Table 8. % Error behavior with medium noise.

Design		2 ⁴			2 ⁵			2 ⁶			2 ⁷			2 ⁸		
True Model		1M	2M	3M	4M	5M	6M	7M	8M	9M	10M	11M	12M	13M	14M	15M
Balanced	No error	66	66	57	68	76	85	61	69	77	60	59	65	49	54	59
	Error I	34	34	19	32	24	14	39	31	23	40	41	35	51	46	41
	Error II	0	0	24	0	0	0	0	0	0	0	0	0	0	0	0
	Error I and II	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Low	No error	67	67	19	70	11	0	62	69	77	54	60	65	2	2	0
	Error I	32	33	6	29	2	0	38	31	23	46	40	34	1	1	0
	Error II	0	0	27	0	0	85	0	0	0	0	0	1	0	0	59
	Error I and II	0	0	48	1	87	15	0	0	0	0	0	0	97	97	41
Unbalance level Medium	No error	68	68	22	0	0	0	63	70	77	45	51	0	0	0	0
	Error I	32	32	7	0	0	0	37	30	23	33	28	0	0	0	0
	Error II	0	0	41	71	81	85	0	0	0	0	0	66	51	59	60
	Error I and II	0	0	30	29	19	15	0	0	0	22	21	34	49	41	40
High	No error	47	47	0	0	0	0	63	69	77	0	0	0	0	0	0
	Error I	17	17	0	0	0	0	37	31	23	0	0	0	0	0	0
	Error II	0	0	77	1	81	85	0	0	0	57	63	65	51	57	60
	Error I and II	36	36	23	99	19	15	0	0	0	43	37	35	49	43	40

Table 9. % Error behavior with high noise.

Design		2 ⁴			2 ⁵			2 ⁶			2 ⁷			2 ⁸		
True Model		1H	2H	3H	4H	5H	6H	7H	8H	9H	10H	11H	12H	13H	14H	15H
Balanced	No error	56	55	35	68	75	84	62	69	77	53	59	66	49	54	60
	Error I	35	30	12	31	24	15	38	31	23	47	41	34	51	46	40
	Error II	8	13	52	0	1	1	0	0	0	0	0	0	0	0	0
	Error I and II	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Low	No error	55	42	4	61	14	0	62	69	77	49	54	45	3	3	0
	Error I	32	21	2	23	3	0	38	31	23	41	35	24	2	2	0
	Error II	9	28	63	3	5	85	0	0	0	0	0	20	0	0	60
	Error I and II	3	8	31	13	78	15	0	0	0	10	10	11	96	95	40
Unbalance level Medium	No error	52	47	3	0	0	0	63	69	77	32	35	35	0	0	0
	Error I	29	22	1	0	0	0	37	31	23	21	18	18	0	0	0
	Error II	11	19	67	70	80	85	0	0	0	10	12	12	52	58	62
	Error I and II	8	12	28	30	20	15	0	0	0	37	34	35	48	42	38
High	No error	34	28	0	0	0	0	62	68	76	0	0	0	0	0	0
	Error I	16	11	0	0	0	0	37	30	22	0	0	0	0	0	0
	Error II	24	28	5	3	80	86	0	1	1	56	63	66	51	58	60
	Error I and II	26	33	25	97	20	14	1	1	0	44	37	34	49	42	40

Table 10. % Error behavior with very high noise.

Design		2 ⁴			2 ⁵			2 ⁶			2 ⁷			2 ⁸		
True Model		1V	2V	3V	4V	5V	6V	7V	8V	9V	10V	11V	12V	13V	14V	15V
Balanced	No error	7	7	2	20	14	9	50	51	54	54	0	65	47	55	59
	Error I	6	6	1	12	6	2	30	24	17	46	0	34	53	46	41
	Error II	69	68	88	52	68	82	14	19	24	0	57	1	0	0	0
	Error I and II	17	19	9	16	12	7	7	6	5	0	43	0	0	0	0
Low	No error	3	3	0	5	14	0	44	46	41	27	31	2	5	6	0
	Error I	3	2	0	2	3	0	26	19	13	23	19	1	4	4	0
	Error II	70	70	79	60	5	86	19	24	37	1	1	63	0	0	59
	Error I and II	24	24	20	33	78	14	10	11	9	49	49	34	91	91	41
Unbalance level Medium	No error	4	4	0	0	0	0	14	38	26	4	4	0	0	0	0
	Error I	2	2	0	0	0	0	7	16	7	2	2	0	0	0	0
	Error II	67	67	77	70	77	86	43	27	49	41	48	63	52	57	60
	Error I and II	27	26	23	30	23	14	36	18	18	53	46	37	48	43	40
High	No error	2	2	0	0	0	0	15	12	12	0	0	0	0	0	0
	Error I	1	1	0	0	0	0	7	4	4	0	0	0	0	0	0
	Error II	65	65	77	21	79	85	43	51	62	52	57	60	51	57	58
	Error I and II	31	32	23	79	21	15	35	32	22	48	43	40	49	43	42

The values of the tables were synthesized by taking the averages for each type of error and balance level. Tables 11–14 show the synthesized values.

Tables 11–13 show similar behaviors for the different types of errors under low, medium, and high noise.

Table 14 shows that when the noise is very high, the percentage of no error decreases dramatically, predominating the type II error and type I and II error.

These tables show a tendency for balanced and low unbalanced designs to have high percentages of no error and type I error. However, it is observed that when the designs have medium and high unbalance, the percentages of no error and error type I decrease, increasing in turn, the percentages of type II error and type I and II error.

The same procedure was used to evaluate the behavior of the R^2 , R^2_{adj} , and R^2_{pred} statistics under different balance and noise levels. The results are shown Tables 15–17.

Table 11. Values for different levels of unbalance with low noise.

Type Error	2^k Design with Low Noise			
	Balanced	Unbalanced Level		
		Low	Medium	High
No error	0.65	0.43	0.33	0.20
Error I	0.35	0.22	0.16	0.09
Error II	0.00	0.10	0.34	0.40
Type I and II error	0.00	0.26	0.17	0.31

Table 12. Values for different levels of unbalance with medium noise.

Type Error	2^k Design with Medium Noise			
	Balanced	Unbalanced Level		
		Low	Medium	High
No error	0.65	0.42	0.31	0.20
Error I	0.34	0.21	0.15	0.08
Error II	0.02	0.11	0.34	0.40
Type I and II error	0.00	0.26	0.20	0.32

Table 13. Values for different levels of unbalance with high noise.

Type Error	2^k Design with High Noise			
	Balanced	Unbalanced Level		
		Low	Medium	High
No error	0.61	0.36	0.28	0.18
Error I	0.33	0.18	0.13	0.08
Error II	0.05	0.18	0.36	0.43
Type I and II error	0.00	0.27	0.23	0.31

Table 14. Values for different levels of unbalance with very high noise.

Type Error	2^k Design with very High Noise			
	Balanced	Unbalanced Level		
		Low	Medium	High
No error	0.33	0.15	0.06	0.03
Error I	0.22	0.08	0.03	0.01
Error II	0.36	0.38	0.59	0.59
Type I and II error	0.09	0.39	0.32	0.37

Tables 15 and 16 show that the R^2 and R^2_{adj} statistics are robust to the different levels of unbalance because their values do not change as the unbalance increases if the noise is maintained in a fixed value. It can also be observed that they resist in a good way the low, medium, and high noise level; however, a significant decrease in their values is noticed in the presence of very high noise.

From Table 17 it can be observed that designs with a high level of unbalance are not good for prediction because the R^2_{pred} reduces drastically. It can also be seen that the designs are not good at predicting when the noise level is very high.

Table 15. Behavior of the R^2 statistic at different noise and unbalanced levels.

		R^2		
Level Noise	Balanced	Unbalanced Level		
		Low	Medium	High
Low	0.99	0.99	0.99	0.99
Medium	0.98	0.98	0.98	0.98
High	0.93	0.93	0.93	0.92
Very high	0.78	0.79	0.79	0.77

Table 16. Behavior of the R^2_{adj} statistic at different noise and unbalanced levels.

		R^2_{adj}		
Level Noise	Balanced	Unbalanced Level		
		Low	Medium	High
Low	0.98	0.99	0.99	0.98
Medium	0.96	0.98	0.98	0.96
High	0.87	0.88	0.88	0.87
Very high	0.65	0.67	0.67	0.66

Table 17. Behavior of the R^2_{pred} statistic at different noise and unbalanced levels.

		R^2_{pred}		
Level Noise	Balanced	Unbalanced Level		
		Low	Medium	High
Low	0.96	0.93	0.88	0.79
Medium	0.92	0.86	0.85	0.77
High	0.74	0.72	0.72	0.70
Very high	0.47	0.49	0.50	0.50

4. Conclusions

This investigation analyzed how the lack of balance affects the experimental design properties, specifically error types and R^2 , R^2_{adj} , and R^2_{pred} statistics. We analyzed how the different types of errors change as the unbalance increases and how the R^2 , R^2_{adj} , and R^2_{pred} statistics are modified in the presence of unbalance and experimental noise.

As a result of this research, it was observed that as the degree of unbalance increases, the type II, as well as type I and II errors, become more probable in the model. It was also observed that, in general, a design with any level of balance resists low, medium, and high noise levels, given that errors become only significant when a very high degree of noise is present.

Balanced designs are more desirable for obvious reasons, but if we have no option and need to run an unbalanced design and it is intended that the errors generated are as minimum as possible, it is advisable to run a design with a low level of unbalance. This is because a design with low unbalance will provide the correct model about 40% of the times as long as the noise level is low to high, and the errors will be distributed in the following way: type I error 20%, type II error 13%, and type I and II error 26%, approximately.

When the design has medium unbalance and the level of noise is low to high, the percentages are distributed in the following way: no error 30%, type I error 15%, type II error 35%, and type I and II error 20%, approximately. If the design has high unbalance

and the level of noise is low to high, the percentages are as follows: no error 19%, type I error 8%, type II error 41%, and type I and II error 31%, approximately.

When the design has a very high noise level, only a balanced design will provide acceptable percentages for error types. In this situation (very high noise), any degree of unbalance will produce results that are not favorable given that no error will decrease and error type II as well as error type I and II will increase significantly.

Regarding the R^2 and R^2_{adj} statistics, they are robust to the different levels of unbalance. Under low, medium, and high noise levels, the R^2 presents values that range from 0.92 to 0.99 and the R^2_{adj} ranges from 0.87 to 0.98. The R^2 and R^2_{adj} statistics reduce significantly in the presence of a very high level of noise, taking values of 0.78 and 0.66, respectively.

Designs with a high level of unbalance are not good at predicting given that values for R^2_{pred} become low, and this situation is even worse if the noise level is very high. The R^2_{pred} begins to have significant changes when a medium level of unbalance is present and reaches its lowest values when the noise level is very high.

As a general conclusion, perfect balance is not necessary in experimentation; a low level of unbalance is tolerable given that it will produce acceptable values for errors and statistics under low to high noise conditions.

Finally, similar behavior for errors and statistics can be expected in other types of designs, such as fractional factorial designs, response surface designs, and mixed-level designs, but an in-depth study is needed to determine for sure how they behave in the presence of unbalance.

Author Contributions: Conceptualization, R.R.-T.; methodology, R.R.-T. and A.J.R.-L.; validation, R.R.-T.; formal analysis, Y.V.P.-P. and J.A.V.-L.; investigation, R.R.-T., A.J.R.-L., Y.V.P.-P. and E.A.R.-S.; writing—original draft preparation, R.R.-T. and A.J.R.-L.; writing—review and editing, R.R.-T., A.J.R.-L. and Y.V.P.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank the editor and reviewers for their helpful comments and suggestion that greatly improved the content and quality of paper.

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

References

1. Montgomery, D.C. *Design and Analysis of Experiments*; John Wiley & Sons: Hoboken, NJ, USA, 2017.
2. Gutierrez, H.; De la Vara, R. *Análisis y Diseño de Experimentos*, 3rd ed.; McGrawHill: New York, NY, USA, 2012.
3. Birkes, D.S.; Seely, J.F.; Vanleeuwen, D.M. Balance and orthogonality in designs for mixed classification models. *Ann. Stat.* **1999**, *27*, 1927–1947. [[CrossRef](#)]
4. Ríos, A.J. Sequential Experimentation, A New Experimentation Approach for Resolution III, Mixed-Level and Robust Designs. VDM Verlag Dr. Müller. 2009. Available online: <http://www.vdm-verlag.de/> (accessed on 6 September 2018).
5. Van Belle, G.; Kerr, K.F. *Design and Analysis of Experiments in the Health Sciences*; John Wiley & Sons: Hoboken, NJ, USA, 2012. [[CrossRef](#)]
6. Pantoja, Y.V.; Ríos, A.J.; Tapia, M. A method for construction of mixed-level fractional designs. *Qual. Reliab. Eng. Int.* **2019**, *35*, 1646–1665. [[CrossRef](#)]
7. Yates, F. Complex experiments. *Suppl. J. R. Stat. Soc.* **1935**, *2*, 181. [[CrossRef](#)]
8. Yates, F. The gain in efficiency resulting from the use of balanced designs. *Suppl. J. R. Stat. Soc.* **1938**, *5*, 70. [[CrossRef](#)]
9. Nelder, J. The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* **1965**, *283*, 147–162. [[CrossRef](#)]
10. John, J.A.; Smith, T.M.F. Two-factor experiments in non-orthogonal designs. *J. R. Stat. Soc. Ser. B Methodol.* **1972**, *34*, 401–409. [[CrossRef](#)]
11. Cotter, S.C.; John, J.A.; Smith, T.M.F. Multi-factor experiments in non-orthogonal designs. *J. R. Stat. Soc. Ser. B Methodol.* **1973**, *35*, 361–367. [[CrossRef](#)]
12. John, J.A. Factorial-balance and the analysis of designs with factorial structure. *J. Stat. Plan. Inference* **1981**, *5*, 99–105. [[CrossRef](#)]

13. Mukerjbe, R. Inter-effect-orthogonality in factorial experiments. *Calcutta Stat. Assoc. Bull.* **1979**, *28*, 83–108. [[CrossRef](#)]
14. Mukerjee, R. Construction of effect-wise orthogonal factorial designs. *J. Stat. Plan. Inference* **1981**, *5*, 221–229. [[CrossRef](#)]
15. Gupta, S.C. Some new methods for constructing block designs having orthogonal factorial structure. *J. R. Stat. Soc. Ser. B Methodol.* **1983**, *45*, 297–307. [[CrossRef](#)]
16. Speed, F.M.; Hocking, R.R.; Hackney, O.P. Methods of analysis of linear models with unbalanced data. *J. Am. Stat. Assoc.* **1978**, *73*, 105–112. [[CrossRef](#)]
17. Herr, D.G. On the history of ANOVA in unbalanced, factorial designs: The first 30 years. *Am. Stat.* **1986**, *40*, 265–270. [[CrossRef](#)]
18. Shaw, R.G.; Mitchell-Olds, T. ANOVA for unbalanced data: An overview. *Ecology* **1993**, *74*, 1638–1645. [[CrossRef](#)]
19. Huber, J.; Zwerina, K. The Importance of Utility Balance in Efficient. *J. Mark. Res.* **1996**, *33*, 307–317. [[CrossRef](#)]
20. Kerr, M.K.; Churchill, G.A. Experimental design for gene expression microarrays. *Biostatistics* **2001**, *2*, 183–201. [[CrossRef](#)]
21. Guo, Y.; Simpson, J.R.; Pignatiello, J.J., Jr. The General Balance Metric for Mixed-level Fractional Factorial Designs. *Qual. Reliab. Eng. Int.* **2009**, *25*, 335–344. [[CrossRef](#)]
22. Hector, A.; Von Felten, S.; Schmid, B. Analysis of variance with unbalanced data: An update for ecology & evolution. *J. Anim. Ecol.* **2010**, *79*, 308–316. [[CrossRef](#)] [[PubMed](#)]
23. Ríos, A.J.; Simpson, J.R.; Vázquez, J.A. Sequential experimentation approach for augmenting of resolution III fractions. *Commun. Stat.-Theory Methods* **2011**, *40*, 2337–2357. [[CrossRef](#)]
24. Landsheer, J.A.; van den Wittenboer, G. Unbalanced 2 X 2 factorial designs and the interaction effect: A troublesome combination. *PLoS ONE* **2015**, *10*, e0121412. [[CrossRef](#)]
25. Voelkel, J.G. The design of order-of-addition experiments. *J. Qual. Technol.* **2019**, *51*, 230–241. [[CrossRef](#)]
26. Naranjo, F.; Ríos, A.J.; Pantoja, Y.V.; Tapia, M. Diseños Factoriales de Taguchi fraccionados. *Ing. Investig. Tecnol.* **2020**, *21*, 1–12. [[CrossRef](#)]
27. Yates, F. *The Principles of Orthogonality and Confounding in Replicated Experiments*; Cambridge University Press: Cambridge, UK, 1933; pp. 108–145. [[CrossRef](#)]
28. Wu, C.F.J.; Xu, H. Generalized Minimum Aberration for Asymmetrical Fractional Factorial Designs. *Ann. Stat.* **2001**, *29*, 1066–1077. [[CrossRef](#)]
29. Xu, H. Minimum Moment aberration for nonregular designs and supersaturated designs. *Stat. Sin.* **2003**, *13*, 691–708.
30. Xu, H.; Deng, L.Y. Moment Aberration Projection for Nonregular Fractional Factorial Designs. *Technometrics* **2005**, *47*, 121–131. [[CrossRef](#)]
31. Xu, H. An Algorithm for Constructing Orthogonal and Nearly-Orthogonal Arrays with Mixed Levels and Small Runs. *Technometrics* **2002**, *44*, 356–368. [[CrossRef](#)]
32. Arias, E.H.; Ríos, A.J.; Vázquez, J.A.; Pérez, R. Estudio comparativo entre los enfoques de diseño experimental robusto de Taguchi y tradicional en presencia de interacciones de control por control. *Ing. Investig. Tecnol.* **2015**, *16*, 131–142. [[CrossRef](#)]
33. Mancilla, A.M. Simulación Herramienta Para el Estudio de Sistemas Reales. *Revista Científica Ingeniería y Desarrollo*. 1999. Available online: <https://rcientificas.uninorte.edu.co/index.php/ingenieria/article/view/2226> (accessed on 19 November 2021).
34. Draper, N.R.; Smith, H. *Applied Regression Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 1998; Volume 326. [[CrossRef](#)]
35. Misra, H.; Ríos, A.J.; Simpson, J.R.; Vázquez, J.A. The Quarterfold, a Sequential Augmentation Procedure for Resolution IV Fractions. *Qual. Eng.* **2013**, *25*, 118–135. [[CrossRef](#)]
36. Ríos, A.J.; Guerrero, G. Sequential experimentation approach for robust design. *Qual. Reliab. Eng. Int.* **2018**, *34*, 1556–1565. [[CrossRef](#)]
37. Jaccard, J.J. *Interaction Effects in Factorial Analysis of Variance*; SAGE Publications: Thousand Oaks, CA, USA, 1998.