

Statistical Fatigue Experiment Design in Medium Density Fiberboard

Mariano Martínez Espinosa^a, Carlito Calil Jr.^b

^a*Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos – SP, Brazil*

^b*Escola de Engenharia de São Carlos, Universidade de São Paulo,
São Carlos – SP, Brazil*

Received: March 21, 2000; Revised: July 31, 2000

Medium Density Fiberboard (MDF) is a wood-based composite widely employed in several industrial applications, in addition to its use in structures subjected to dynamic loads. Its fatigue-related aspects, however, have been consistently ignored. This work proposes to study fatigue in MDF, including the following factors: the basic concepts of MDF and fatigue and the statistical design of fatigue experiments in MDF, with the purpose of obtaining accurate information for analysis by means of statistical methods. The results of our tests revealed that the statistical model is suitable to fit the number of cycles in intermediary S and f levels and to determine the levels of the factors that maximize the total number of cycles to failure. It was also found that the proposed design is of great practical interest for fatigue strength in the tension in wood and wood derivatives.

Keywords: *MDF, fatigue, fatigue properties in wood composites, factorial designs*

1. Introduction

The properties of wood composites, as well as cost and manufacturing considerations, suggest that these are the most suitable materials for several industrial applications such as furniture, civil construction and other uses. These products, called composites, are materials composed of combinations of various elements, mainly wood particles and glue, having different properties of solid wood¹.

One wood-based composite that is widely used in a variety of industrial applications is Medium Density Fiberboard (MDF)⁹. However, the aspects related to fatigue in this product have generally been ignored¹¹, although most service failures associated to mechanical causes of the materials are usually fatigue-related⁷.

Fatigue studies of MDF are possible due to its general properties and characteristics⁸. However, for such purposes, an appropriate statistical experiment design must be made to obtain accurate information for analysis by means of statistical methods in order to reach valid and objective inferences⁶.

The use of statistical experiment design in the study of the fatigue in MDF in particular is justified by several factors such as the reduced number of tests, reduced testing time and, thus, reduced cost, less variability and greater reliability of the results, among others. However, the main

reason is to obtain appropriate data that is analyzable using the appropriate methods and statistical techniques, since the analysis method and technique depend directly on the design used¹⁵.

Therefore, the objective of this work is to carry out an experiment plan for the study of the fatigue strength in the tension in MDF in order to reduce the cost, time and variability of the results. Another objective is to envision a possible appropriate statistical model for MDF fatigue data using the experimental results of the proposed tests.

2. Medium Density Fiberboard (MDF)

Medium Density Fiberboard (MDF) is the generic name used to define a sheet composed mainly of lignin-cellulosic fibers combined with a synthetic resin or some other suitable adhesive system, glued together by means of pressure and heat. The introduction of additives during the manufacturing process can improve the dimensional stability, fire resistance and impermeability of MDF¹⁷.

MDF is a homogeneous, uniform and stable product with a flat surface having good workability, machinability and wide acceptance for coatings with several finishes. It is widely employed by the furniture industry owing to its sturdiness and uniformity, which ensure satisfactory results with conventional techniques (screwing, rabbeting, cutting

^ae-mail: marianom@sc.usp.br

^be-mail: calil@sc.usp.br

to size, gluing, etc.) and its characteristics of mechanical strength allow for its use even in structural panels⁹.

MDF can also be used advantageously in the production of lathe-turned and milled pieces which, until recently, had been made of solid wood, such as curtain rods, parts for chairs etc¹⁹.

3. MDF Fatigue

Fatigue is a structural failure caused by a material being subjected to repeated stress or cyclic loading for a considerable period of service time.

The stresses a material can bear under cyclic loading are much smaller than those bearable under static loading. In other words, a fatigue-related failure may occur at a much lesser stress than the limit of the material's strength.

Three basic factors generally cause fatigue-related failure¹⁰:

- 1) high maximum tensile stress;
- 2) great variation or fluctuation of applied stress;
- 3) a large number of stress application cycles.

A cycle is an interval of time during which a sequence of stresses is cyclically applied to the specimen in the same regular order. Stress is the quotient of intensity of a force by the specimen's surface area.

The stress waves used in fatigue tests are generally triangular, square or sinusoidal and the typical cycles of stress in fatigue are reverse stresses, fluctuating stress and irregular or random stress¹⁰.

The properties of fatigue in wood and wooden products (e.g., laminated wood) are particularly influenced by the following factors¹¹:

- 1) The wood species, place of origin, density, etc.
- 2) The size and shape of the test specimen.
- 3) The moisture content.
- 4) The type of applied load. Experiments normally consist of tension, compression, bending or shear tests or a combination of these. Most fatigue experiments are made with harmonic loads and there are three important factors in this type of load:
 - The R-ratio (the ration between the minimum and maximum tension).
 - The value of the stress.
 - The frequency (loading time), which must be taken in account.

5) Other factors affecting the properties of fatigue in wood are temperature, chemical treatments, adhesives, gluing, etc.

Although MDF has no specific properties of fatigue, its properties and general characteristics facilitate the study of fatigue of this material. Thus, since MDF is a homogeneous and approximately isotropic material, the stages of rupture of an isotropic material subject to fatigue can be considered for its study.

4. Statistical Experiment Design

Statistical experiment design is the process of designing an experiment to obtain appropriate data that can be analyzed using statistical methods in order to reach valid and objective conclusions.

Experiment design is essential in engineering to improve the performance of a manufacturing process. Experiment design is also widely employed in the development of new processes. The application of experiment design techniques during the initial phase of process development can lead to improved process performance, less variability, shorter development time and, thus, reduced overall cost.

The application of experiment design is simple, although certain requisites must be followed. The two main requirements in experiment design are to obtain aleatory and replicable results¹⁵.

4.1. Selection of the variable responses, factors and levels in experiment design

When selecting the dependent response or variable, the experimenter must be certain that the response that will be measured actually provides useful information about the process under study. In experiment design, the average of the measured characteristic is usually the response variable; therefore, replicates should be made⁶.

In experiment design, a factor (independent variable) is a qualitative or quantitative experimental variable, which is investigated to determine its effect on a response. The specific values of factors are called levels and can be determined *a priori*¹³. From a statistical standpoint, it is recommended that two or, at the most, three levels be used per factor since, if the number of levels increases, the number of tests will greatly increase.

4.2. Factorial design

In many experimental engineering situations one knows in advance that some factors produce different responses. One particular case of this is the study of material fatigue. When a response is influenced by certain factors, it can be shown that, in general, factorial design is the most efficient. Factorial design is the type of design in which all the possible combinations of factorial levels are investigated in each test¹⁵.

A complete factorial design with k factors (variables) is obtained by choosing n_1 levels of factor 1, n_2 levels of factor 2, ..., n_k levels of factor k, and then selecting the $n = n_1 \times n_2 \times \dots \times n_k$ experiments obtained, using every possible combination of the selected levels. For example, a 3 x 3 x 3 factorial design requires 27 tests, while a 2 x 2 x 2 factorial design requires 8 tests⁶.

4.2.1 Desirable properties of factorial designs¹⁵:

- They should direct the investigation, since they are particularly useful in the exploratory stage of research.
- They should indicate the size of the sample to be selected.
- They should allow for multiple comparisons, thus facilitating the creation and critique of models.
- They should provide highly efficient parameter estimators.

However, the most significant argument set forth is that they should allow for models to be seen in the data in the specification stage of the model (in the specification of a functional relationship).

4.2.2. Approximate response functions

Engineering, physics, chemistry, biology and other fields of research are often interested in a functional relationship. Such a relationship can be obtained through factorial designing using the following approximations⁶:

$$E(y) = \eta = f(x_1^*, x_2^*, \dots, x_k^*) = f(\bar{X}^*) \quad (1)$$

which relates the expected η value of a y response, such as a number of cycles (or production percentage, purity, viscosity, etc.) to k quantitative or qualitative variables x_1^* , x_2^* , ..., x_k^* , such as stress, frequency, temperature, time, pressure, concentration, etc.

The $f(\bar{X}^*)$ function can assume a great variety of forms: linear, parabolic, exponential, etc. These forms are important because they approximate many real world relationships, in addition to being relatively easy to work with and interpret^{5,6}. However, because reiterated tests are made for the same \bar{X}^* conditions, the measurements of the y response would vary because of measurement and observational errors and the basic variability of the experimental material⁵.

The variations of y , in practice, are the differences between the observed y values and the expected η values. These differences, generally indicated by ε (epsilon, fifth letter of the Greek alphabet), are called residues and are aleatory variables with some degree of distribution of probability around their mean η value⁶. Therefore, our general objective is to investigate certain aspects of a functional relationship affected by an error and expressed as⁵:

$$y = f(\bar{X}^*) + \varepsilon \quad (2)$$

In many practical problems, however, the form of the real relationship between the dependent (y) variable and the independent (\bar{X}^*) variables is unknown. Hence, an appropriate approximation of the real (y) function and the set of independent variables must be determined, normally using first and second order polynomials. The latter is the most frequently used, since it allows a curve of the response to be analyzed, which facilitates modeling the relationship

between the response and each factor using a linear or quadratic function.

Second order models should involve at least three levels of each factor to allow for an estimate of the model's coefficients. There are many factorial designs that could be used to estimate a second order model. In this study, because only the effect of two independent variables was verified, we will use a three-level factorial design (3^k).

4.2.3. Three-level factorial designs (3^k)

Three-level factorial designs are formed by k factors, each one having three levels. In a 3^k design or in any factorial design it is more practical not to have to deal with the actual numerical measurements of the x_i^* variables, working, instead, with coded x_i^c variables, *i.e.*, coding the levels of the factors. Codification is done mainly for the following reasons¹⁴:

- 1) If the independent variables are qualitative, they are not numeric; thus, their values need to be coded to estimate a regression model.
- 2) Although the independent quantitative variables are numeric, they should also be coded to estimate a regression model, for the following reasons:

2.1) To estimate the parameters of the model, a matrix termed (X^*, X^*) should be inverted. Considerable rounding out errors may occur during the inversion process if the absolute values of the numbers in the (X^*, X^*) matrix vary significantly, which usually leads to errors in the estimates of parameters. Coding facilitates the computational inversion of the matrix, reducing calculation errors and allowing for more accurate parameter estimates.

2.2) A second reason for coding quantitative variables is the problem of multicollinearity, which is the existence of an exact or approximately exact linear relationship among the independent variables. The problem of multicollinearity is inevitable in estimates of regression models (*e.g.*, second-order models), particularly when higher-order terms are estimated. In quadratic models, for instance, in which the two x_i^* and x_j^{2*} variables are normally highly correlated, the probability of rounding out errors in the coefficients of regression increases in the presence of multicollinearity. These problems are usually eliminated with coding.

The coding of the levels of the factors of the most frequently used 3^k design is -1, 0 and 1, which denote the factor's low, intermediate and high levels, respectively, as illustrated in Table 1. This code is obtained theoretically by the following equation⁵:

$$x_i^c = \frac{x_i^* - \bar{x}_i^*}{S_i^*} \quad (3)$$

with

Table 1. Notations used in a factorial design 3^2 , with p^* replicates.

Run (i)	Response (Y)	Factors	
		A	B
1	$y_{11} \dots y_{1p^*}$	-1	-1
2	$y_{21} \dots y_{2p^*}$	0	-1
3	$y_{31} \dots y_{3p^*}$	1	-1
4	$y_{41} \dots y_{4p^*}$	-1	0
5	$y_{51} \dots y_{5p^*}$	0	0
6	$y_{61} \dots y_{6p^*}$	1	0
7	$y_{71} \dots y_{7p^*}$	-1	1
8	$y_{81} \dots y_{8p^*}$	0	1
9	$y_{91} \dots y_{9p^*}$	1	1

Note: y_{11}, \dots, y_{9p^*} in the table 1 they indicate the replicates of the response y .

$$\bar{x}_i^* = \frac{x_{-1}^* + x_0^* + x_1^*}{3},$$

$$S_i^* = x_0^* - x_{-1}^*$$

or

$$S_i^* = x_1^* - x_0^*,$$

where x_{-1}^* , x_0^* and x_1^* are the values corresponding to the low, intermediate and high levels, respectively. Thus, (4) is used instead of (1), where (4) is given by:

$$E(y) = \eta = f(x_1^c, x_2^c, \dots, x_k^c) = f(\bar{X}^c) + \varepsilon \quad (4)$$

A particularly important aspect in this study is the design, *i.e.*, two factors each having three levels; our interest in this design lies in determining if the number of cycles to failure depends on stress and frequency factors.

a) Determining the effects of a 3^2 design

The effect of the factors on the response is measured by the change in the average response in two or more combinations of the levels. In this type of factorial design there are 9 combinations of treatments. Thus, there are eight degrees of freedom between these combinations. The main A and B effects each have two degrees of freedom, while the AB interaction has four degrees of freedom. If there are (p^*) replicates, there is a total of $p^* 3^2 - 1$ degrees of freedom, which corresponds to $3^2 (p^* - 1)$ degrees of freedom for the error¹⁵.

The combinations of treatments are written in standard order, which means that one factor at a time is introduced successively, combining each level with the set of the factors' levels. The standard order for a 3^2 design is 00, 10, 20, 01, 11, 21, 02, 12, 22¹⁵.

YATES' algorithm (YATES' modified algorithm generalized for the 3^k series) can be used to calculate the effects and the respective square sum of a 3^2 design. The use of this algorithm is simple and consists of determining some columns of data. A simplified calculation procedure is given below for the specific case of $k = 2$:

1) Calculate the total response of each (i) test, *i.e.*, the total of each replicate. This column will be denoted by Y^* .

2) Using the Y^* column, calculate the Y^{*1} column. The Y^{*1} values are calculated as follows: The first third of the column consists of the sum of each of the three sets of the three values of column Y^* , *i.e.*, $y_1^* + y_2^* + y_3^*$; $y_4^* + y_5^* + y_6^*$ and $y_7^* + y_8^* + y_9^*$. The second third is the third observation minus the first one of each of the three previous sets, *i.e.*, $y_3^* - y_1^*$; $y_6^* - y_4^*$ and $y_9^* - y_7^*$. The last third of column Y^{*1} is the result of the sum of the first and third observations, minus twice the second one, in each set of three observations, *i.e.*, $y_1^* - 2y_2^* + y_3^*$; $y_4^* - 2y_5^* + y_6^*$ and $y_7^* - 2y_8^* + y_9^*$.

3) Calculate the column denoted by Y^{*2} . The values of this column are calculated by the same procedure as that used to determine the Y^{*1} values, using the elements of the Y^{*1} column.

4) Calculate the elements of the divisor column D^* using the expression $2^r 3^t p^*$, where r is the number of factors in the effect considered, t is the number of factors in the experiment minus the number of linear terms (levels 1) in this effect, and p^* is the number of replicates.

1) Calculate the column of effects of the E^* factors, which is obtained by dividing column Y^{*2} by column D^* .

2) Calculate the sums of squares (SQ), which is done by squaring the elements of column Y^{*2} and dividing the result by the corresponding element of the divisor column. The sum of squares column now contains all the information required to construct the table of variance analyses. However, before the table of variance analyses is built, one must check if the data residuals have a constant variance and normal distribution.

One of the tools most commonly used to confirm if the variance is constant consists of plotting the residuals ($\hat{\varepsilon}_{ij}$) vs. the estimated values (\hat{y}_{ij}). This graph should not reveal any obvious pattern. One of the problems that is occasionally observed in such graphs is nonconstant variance, in which the residuals graph compared against the estimated values presents a funnel form; see Fig. 1. Variable variance also generally occurs in cases where the data does not follow a normal distribution. The assumption of normality can be confirmed by plotting a histogram of the residuals, which should have a symmetrical form. For this kind of verification, a normal probability graph (NSCORES) can also be used with the residuals, in which the residuals vs. the normal n-scores are represented. If the graph presents an approximately straight line, the distribution is approximately normal. If the variance is nonconstant and if the

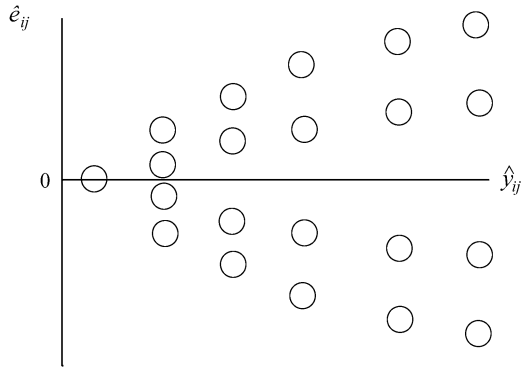


Figure 1. Plot of residuals vs. fitted values (no constant variance).

assumption of normality is not followed, then the original data have to be transformed.

A logarithmic transformation is usually considered in fatigue data to stabilize their variance and lead to their normal distribution³.

The \hat{e}_{ij} residuals for the factorial 3^2 design are calculated by the following equation⁶:

$$\hat{e}_{ij} = y_{ij} - \hat{y}_{ij} \quad (5)$$

With $\hat{y}_{ij} = \bar{y}_{ij}$, where \bar{y}_{ij} are the averages of the observations in the i th test. Therefore, (5) can also be expressed as:

$$\hat{e}_{ij} = y_{ij} - \bar{y}_{ij} \quad (6)$$

5. Materials and Methods

This section discusses the aspects relating to the materials used in the development of the work, directed at obtaining the values of fatigue strength in MDF, considering the statistical factors in mechanical tests¹².

- **Material used in this study:** MDF in sheets, supplied by a sawmill in the city of São Carlos, with dimensions of 1,83 m width, 2,75 m length and 15 mm thickness.

- **Sampling Method:** The specimens were removed from the fiberboard by the simple aleatory method, with 18 specimens selected at random according to the experiment designs indicated in the previous section and based on the ASTM: D1037/96a² code.

- **Dimensioning and preparation of the specimens:** To determine the tensile fatigue strength in MDF, the specimens used followed the dimensions established by the ASTM: D1037/96a² code, as shown in Fig. 2. The T value indicates the thickness of the specimen, which should not exceed 254 mm.

- **Execution of the tests:** The static and dynamic tests were carried out in the DARTEC M1000/RC universal testing machine at LaMEM. The data obtained from the static tests were used to establish the levels of maximum and minimum tension for the fatigue tests. The static tests followed the recommendations of the ASTM: D1037/96a² code and the NBR/7190/97¹⁶ code. Reverse stress cycles

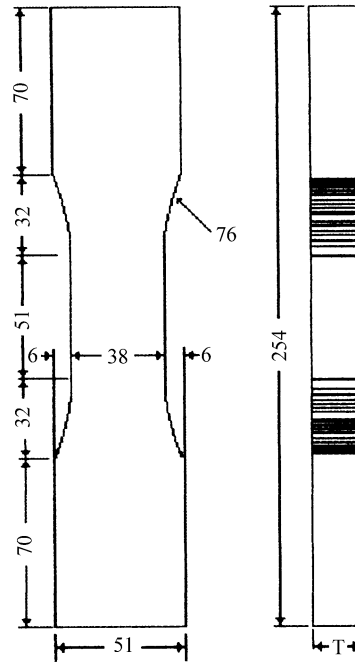


Figure 2. Dimensions of the specimens for tension test parallel to surface (mm).

of sinusoidal form were used in the dynamic tests, considering frequencies of 1; 5 and 9 Hz for the specimens and cyclic loads of 60, 75 and 90 percent of the material's strength, estimated in the static tests (on average $F_r = 9.94$ kN in dry condition). The procedure and results of the static tests employed are part of the work of MARTÍNEZ and CALIL on static tensile strength tests parallel to the surface in MDF¹².

- **Design:** A factorial 3^2 design was used, in which a study is made of the total number of cycles to failure per Y second (response variable), considering two replicates and using combinations of two factors (stress and frequency), each tested with three levels. The levels for each factor, as well as the coding for each level, are given in Table 2, in which the -1, 0 and 1 codes are obtained by Eq. (3), as follows:

$$x_1^c = \frac{x_1^* - 75}{15}, x_2^c = \frac{x_2^* - 5}{4}$$

where,

Table 2. Number of cycles, under three levels of stress and three types of frequencies.

Coded levels	x_i^c	-1	0	1
Stress (S) in %	x_1^*	60	75	90
Frequency (f) in %	x_2^*	1	5	9

$$f_{10}^* = \frac{60 + 75 + 90}{3} = 75, f_{20}^* = \frac{1 + 5 + 9}{3} = 5, \text{ and, } S_1^* = 75 - 60 = 15 \text{ and } S_2^* = 5 - 1 = 4.$$

6. Results

Table 3 presents the experimental results, considering the notation given in Table 2.

7. Data Analyses

In agreement with section 4.2.3 (a), the analysis of the data begins with the graph of the residuals vs. the estimated values (Figs. 3 and 4).

Figure 3 indicates that the variance is nonconstant, *i.e.*, it increases. When this is the case, a logarithmic transformation of the data ($Y^T = \log(y)$) is usually carried out to stabilize the variance.

An analysis of Fig. 4 indicates that the residues are not distributed, since a comparison of the residue graph against the normal scores does not follow a straight line. Proceeding with the transformed data construct, the residuals vs. estimated values are plotted (Fig. 5). Figure 5 indicates a constant variance, while the behavior shown in Fig. 6 is approximately linear, indicating that the residuals follow

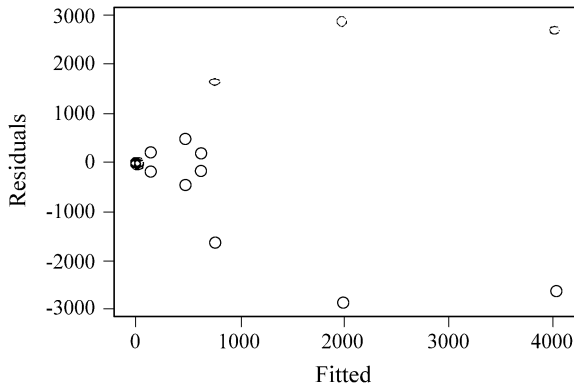


Figure 3. Plot of residuals vs. fitted values.

Table 3. Data of MDF obtained in the factorial design 3^2 .

Run (i)	Response (Y)		Factors	
	y_{i1}	y_{i2}	S	F
1	6201	9423	-1	-1
2	1831	1447	0	-1
3	91	108	1	-1
4	17158	22885	-1	0
5	4405	5362	0	0
6	214	153	1	0
7	43051	37736	-1	1
8	6263	6597	0	1
9	548	482	1	1

an almost normal distribution. Hence, the YATES algorithm is used in the transformed data to determine the effects of the factors.

Table 4 shows that the greatest effects for the number of cycles, in decreasing order, are stress ($S_L = -0.97017$), frequency ($f_L = 0.33887$) and stress to the square ($S_Q = -0.09234$). However, the quadratic effect of S is not as significant as the individual effects of f and S (main effects of f and S), respectively.

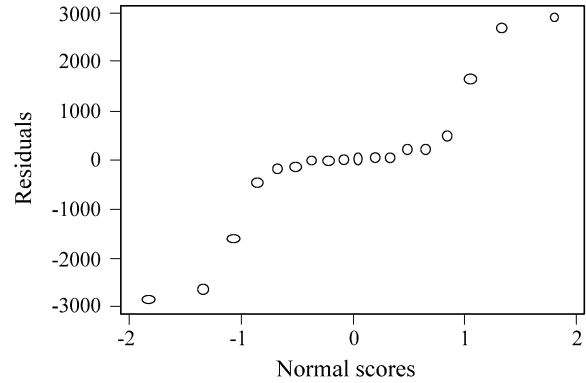


Figure 4. Plot of residuals vs. the normal scores.

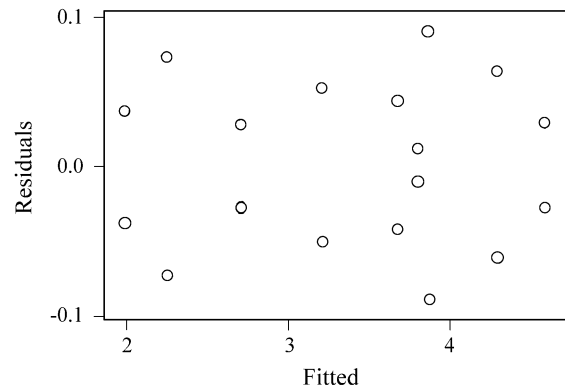


Figure 5. Plot of residuals vs. the fitted values (transformed data).

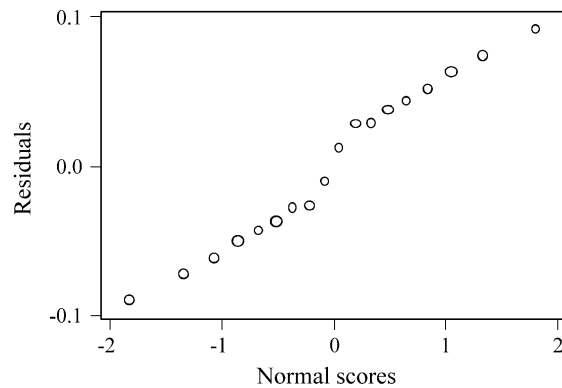


Figure 6. Plot of residuals vs. the normal scores (transformed data).

Another way to detect the significant effects is to calculate the normal scores of the effects and plot the effects vs. the normal scores (Fig. 7). All the effects that are located approximately under the straight line are insignificant, while the significant effects are far removed from it. The effects that are identified in this analysis are the main effects of f and S , and no significant interaction is observed among the factors considered.

One can also use variance analysis to confirm the magnitude of the effects. This analysis is summarized in Table 5, which shows that the linear and quadratic components of stress in the total number of cycles to failure are highly significant. The S_L component is relatively large in comparison to the others, while frequency has a linear effect

Table 4. YATES' algorithm of for the transformed data, of the Table 3.

i	Y^*	Y^{*1}	Y^{*2}	D^*	E^*	Combination	SQ
1	7.76665	25.5714	60.9134	18	3.38408	00	206.135
2	8.59402	21.4126	4.0664	12	0.33887 (**)	10 (f_L)	1.378
3	9.21074	13.9294	-0.5338	36	-0.01483	20 (f_Q)	0.008
4	6.42316	1.4441	-11.6420	12	-0.97017 (**)	01 (S_L)	11.295
5	7.37327	1.1930	-0.0147	8	-0.00184	11 (S_{fLxL})	0.000
6	7.61613	1.4294	0.5947	24	0.02478	21 (S_{fQxL})	0.015
7	3.99247	-0.2106	-3.3243	36	-0.09234 (**)	02 (S_Q)	0.307
8	4.51511	-0.7073	0.4875	24	0.02031	12 (S_{fLxQ})	0.010
9	5.42183	0.3841	1.588(3)-6667				

Table 5. Analysis of the variance for the transformed data.

Source of variation.	SS	DF	MS	F	Significant
Frequency [f]	1.39	2	0.695	115.83	< 0.1%
(f_L)	(1.38)	(1)	1.380	230.00	(< 0.1%)
(f_Q)	(0.01)	(1)	0.010	1.67	
Stress [S]	11.61	2	5.805	967.50	0.1%
(S_L)	(11.30)	(1)	11.300	1883.33	(< 0.1%)
(S_Q)	(0.31)	(1)	0.310	51.67	(< 0.1%)
Interaction [Sf]	0.06	4	0.015	2.50	
S_{fLxL}	(0.00)	(1)	0.000	0.00	
S_{fLxQ}	(0.01)	(1)	0.010	1.67	
S_{fQxL}	(0.01)	(1)	0.010	1.67	
S_{fQxQ}	(0.04)	(1)	0.040	6.67	1%
Residuals	0.05	9	0.006		
Total	69.45	17			

only on the total number of cycles to failure. The $Sf_{Q \times Q}$ interaction is significant at 1%, *i.e.*, there is a slight interaction between the two factors.

Observation 1: Currently there are many commercial computational programs for data analysis. The MINITAB V.10 was used in this study owing to several advantages it offers, such as its easy use, precision, and versatility of the different statistical techniques, among others.

8. Conclusions

1) It has been shown that the use of statistical fatigue experiment design in MDF decreases the number of required tests and can also greatly increase their effective-