Fatigue Reliability Analysis of Wind Turbine Cast Components

Hesam Mirzaei Rafsanjani [1,*], John Dalsgaard Sørensen [1], Søren Fæster [2] and Asger Sturlason [3]

[1] Department of Civil Engineering, Aalborg University, 9220 Aalborg Ø, Denmark; jds@civil.aau.dk
[2] Department of Wind Energy, Technical University Denmark, 4000 Roskilde, Denmark; sfni@dtu.dk
[3] Vestas Technology & Service Solutions, 8200 Aarhus, Denmark; ast@vestas.com
* Correspondence: hmr@civil.aau.dk; Tel.: +45-41-727434

Abstract: The fatigue life of wind turbine cast components, such as the main shaft in a drivetrain, is generally determined by defects from the casting process. These defects may reduce the fatigue life and they are generally distributed randomly in components. The foundries, cutting facilities and test facilities can affect the verification of properties by testing. Hence, it is important to have a tool to identify which foundry, cutting and/or test facility produces components which, based on the relevant uncertainties, have the largest expected fatigue life or, alternatively, have the largest reliability to be used for decision-making if additional cost considerations are added. In this paper, a statistical approach is presented based on statistical hypothesis testing and analysis of covariance (ANCOVA) which can be applied to compare different groups (manufacturers, suppliers, test facilities, etc.) and to quantify the relevant uncertainties using available fatigue tests. Illustrative results are presented as obtained by statistical analysis of a large set of fatigue data for casted test components typically used for wind turbines. Furthermore, the SN curves (fatigue life curves based on applied stress) for fatigue assessment are estimated based on the statistical analyses and by introduction of physical, model and statistical uncertainties used for the illustration of reliability assessment.

Keywords: reliability; casting; fatigue; analysis of covariance (ANCOVA); wind turbines

1. Introduction

Wind energy has become an attractive source of renewable energy, and its installed capacity worldwide has grown significantly in recent years [1]. Offshore wind turbines have been installed many places, especially in the North Sea, and many new, larger wind farms are planned. Due to the harsh environmental conditions for offshore wind turbines, and the access difficulties for maintenance and repairs, it is very important to minimize failures of wind turbine components including fatigue failures of casted components [2] in the wind turbine drivetrain.

Wind turbines are large structures exposed to wave excitations, highly dynamic wind loads and wakes from other wind turbines [3]. Thus, wind turbine components are exposed to stochastic loads that are varying randomly during the design working life. Due to highly variable loads, the components may fail due to fatigue, wear and other deterioration processes [3].

Casting defects are of high importance for the lifetime of structural casting. The fatigue life of cast iron components is often controlled by the growth of cracks initiated from defects such as shrinkage cavities and gas pores [4]. Further, different manufacturers apply different manufacturing processes of cast components, resulting in different fatigue lives of the produced components.

This paper focuses on using analysis of covariance (ANCOVA) for comparing different groups/manufacturing steps of specimens from the casted components. An advantage of ANCOVA is that this method is able to handle different numbers of tests for various groups [5]. The result of

ANCOVA is useful in decision-making processes for companies and manufacturers to choose the better manufacturing process. This will lead to a higher quality of the manufactured components and increase the reliability of the produced components. Further, the results of the ANCOVA analysis are used for the reliability assessment of wind turbine cast components. For this reason, the uncertainties related to model evaluation and the effect of each uncertainty on the reliability level of the components is evaluated. Moreover, a case study is presented to illustrate the application of ANCOVA to compare different manufacturing steps of casted components according to fatigue life results and also the reliability assessment of chosen cast components in wind turbine components based on the ANCOVA results.

2. Reliability Analysis

Wind turbine component's reliability, usually explained as the probability of proper performance of the component throughout its lifetime. Structural reliability methods, such as first order reliability method and second order reliability method, can be used to evaluate the probability of failure/reliability [2].

Based on statistical analyses of statistically homogeneous datasets of fatigue lives, an appropriate stochastic model for the fatigue life can be established as described shortly below. Further, it is shown how a fatigue load model can be formulated based on simulations of the load effects for a specific wind turbine and represented by Markov matrices. Together with stochastic variables representing the uncertainties associated with the load assessment, and following the principles in, e.g., [6,7], the reliability can be estimated by First-Order Reliability Methods (FORM) or by simulation, see, e.g., [8]. Finally, the stochastic model established on the basis of the above statistical analysis in combination with the reliability model for the fatigue failure of wind turbine cast components can be applied for the calibration of partial safety factors to a specified target reliability level, e.g., 5×10^{-4} per year as recommended in [7]. In Section 4, the reliability assessment and methodology to evaluate the probability of failure is briefly explained, see [2] for more details.

3. Statistical Analysis of Fatigue Data Sets

A major challenge in statistical analysis of fatigue data is to establish a statistically homogeneous dataset to be applied as the basis for the statistical analyses. Analysis of variance (ANOVA) is a statistical method that can be used to analyze the differences between mean values of different groups of, e.g., suppliers. ANOVA could be used as a statistical test to investigate equality of the means of several groups, and hence generalizes the classical statistical *t*-test to more than two groups [9]. Based on hypothesis testing, the null hypothesis by the ANOVA method is that the mean values of several groups are equal, and on the other hand, the alternative hypothesis is that the mean values are not equal [10]. The level of significance represents the probability of making a type I error and is denoted by α. The level of significance should not be made too small, because the probability of making a type II error will then be increased. In this paper the value of $\alpha = 0.05$ is chosen.

ANOVAs are useful in comparing (testing) three or more groups. A decision of whether or not to accept the null hypothesis depends on a comparison of the computed values of the test statistics and the critical values. The ANOVA analysis can be performed computationally as follows; for further details see [10].

Step 1: Formulation of hypothesis

If a problem involves k groups, the following hypotheses are appropriate for comparing k group means [10]:

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_k$$
$$H_A : at\ least\ one\ pair\ of\ group\ means\ are\ not\ equal$$

(1)

where μ is the mean value in the above equation.

Step 2: Define the test statistic and its distribution

The hypotheses of Step 1 can be tested using the following test statistic:

$$F = \frac{MS_b}{MS_w} \tag{2}$$

in which MS_b and MS_w are the mean squares quantifying between and within variations, respectively, and F is a random variable following an F distribution with degrees of freedom of $(k - 1, M - k)$; for more details, see [10]. The following table shows the calculation for the ANOVA. In Table 1, k is the number of groups, m_j is the number of data in j-th group, $\overline{X}$ is the mean value of all data, $\overline{X}_j$ is the mean value of data in the j-th group and M is the total number of data [10].

Table 1. Computation table for analysis of variance test. *SS*: Sum of square; *DF*: degrees of freedom and *MS*: mean square.

Source of Variation	Sum of Square	Degrees of Freedom	Mean Square
Between	$SS_b = \sum\limits_{j=1}^{k} m_j (X_j - \overline{X})^2$	$k - 1$	$MS_b = SS_b / (k - 1)$
Within	$SS_w = \sum\limits_{j=1}^{k} \sum\limits_{i=1}^{m_j} (X_{ij} - \overline{X}_j)^2$	$M - k$	$MS_w = SS_w / (M - k)$
Total	$SS_t = \sum\limits_{j=1}^{k} \sum\limits_{i=1}^{m_j} (X_{ij} - \overline{X})^2$	$M - 1$	-

Step 3: The level of significance

The level of significance is chosen to α = 5% in this paper.

Step 4: Collect data and compute test statistic

The data should be collected and used to compute the value of the test statistics (F) in Equation (2). Each data value is categorized according to the different groups to be compared statistically to each other.

Step 5: Determine the critical value of the test statistic

The critical value is a function of the level of significance and the degrees of freedom. If the computed value of Step 4 is greater than the critical value, the null hypothesis of Equation (1) should be rejected and the alternative hypothesis of Equation (1) accepted.

In addition, the ANCOVA always involves at least three variables to be introduced: an independent variable, a dependent variable, and a covariate. The covariate is the variable likely to be correlated with the dependent variable. For application for fatigue data, these variables are chosen as:

- The independent variable is the group types that we consider to compare to each other (for example, test facilities or suppliers);
- The dependent variable is the fatigue life N and it is dependent on the test stress amplitude and the type of group;
- The covariate is the test stress amplitude, σ [5].

The major distinction between the two analyses (ANOVA and ANCOVA) is that in ANOVA, the error term is related to the variation of log$N$ around individual group means, whereas the ANCOVA error term is based on variations of log$N$ scores around regression lines. The effect of the smaller within-group variation associated with ANCOVA is an increase of the power of the analysis. Note that the ANOVA distributions have a larger overlap than the ANCOVA distributions [5].

The analysis of covariance is a combination of the linear models employed in the analysis of variance and regression [11].

If it is assumed that data from *k* groups is available, then the starting point for the ANCOVA is exactly the same as for the ANOVA; the total sum of squares is computed. It is noted that ANCOVA can be used for linear regression methods and therefore the analysis is carried out using the "logarithm of fatigue life (log*N*)" and the "logarithm of test stress amplitude (logσ)". Hence, the covariate variable (*x*) is logσ and the dependent variable (*y*) is log*N*. Assuming that there is a linear relationship between the dependent variable and the covariate, we find that an appropriate statistical model is:

$$logN_{ij} = \mu + \tau_i + \beta\left(log\sigma_{ij} - log\overline{\sigma_{..}}\right) + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, m_j \\ j = 1, 2, \ldots, k \end{cases} \tag{3}$$

where logN_{ij} is the *i*-th observation on the response variable in the *j*-th group, logσ_{ij} is the measurement made on the covariate variable corresponding to logN_{ij}, log$\overline{\sigma_{..}}$ is the mean of the logσ_{ij} values, μ is an overall mean, τ_j is the effect/influence of the *j*-th group, β is a linear regression coefficient indicating the dependency of logN_{ij} on logσ_{ij}, and ε_{ij} is a random error component. It is assumed that the errors ε_{ij} are normally distributed with a mean value = 0 and a standard deviation of σ_ε. The null hypothesis is "the group effect is zero ($\sum_{j=1}^{k} \tau_j = 0$)" and if the null hypothesis is accepted, then logσ_{ij} is not affected by the groups.

3.1. Adjusted Mean

An adjusted mean is the mean dependent variable that would be expected or is predicted for each group, if the covariate variable mean is equal to the grand covariate mean. The grand covariate mean is a unique logσ_u and the adjusted fatigue life in each group is estimated according to this stress amplitude. In this way, for each group there is a unique point (logσ_u, log$N_{j,adj}$). This value would be different in different groups, and these values can be used to compare the state of the regression lines to each other. According to adjusted means and slopes of regression lines, there are four types of regression line configurations, as shown in Figure 1 below.
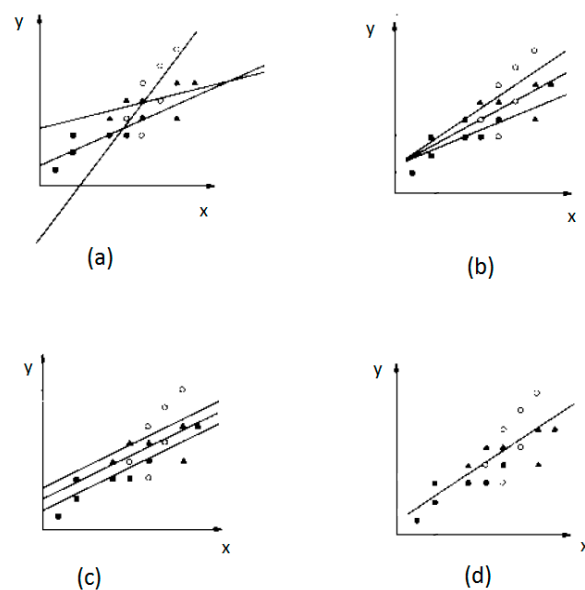
Figure 1. Relation between regression lines. (**a**) Completely different; (**b**) intercept (same intercepts); (**c**) parallel (same slope but without intercept); and (**d**) coincidence (exactly same lines) [12].

By adjustment of the point and slope for each group, comparisons can be performed. As mentioned above, the purpose of the ANCOVA is to test the null hypothesis that two or more adjusted population means are equal. Alternatively, the purpose could be formulated as to test the equality of two or more regression intercepts. Under the assumption of parallel regression lines, the difference between intercepts must be equal to the difference between adjusted means. The formula for the computation of adjusted means is:

$$logN_{j,adj} = log\overline{N_{.j}} - b_w \left(log\overline{\sigma_{.j}} - log\overline{\sigma_{..}} \right) \tag{4}$$

where $log\,N_{j,adj}$ is the adjusted mean for the j-th group; $log\,\overline{N_{.j}}$ is the unadjusted mean (mean of fatigue life) for the j-th group; b_w is the pooled within-group regression coefficient (for details see [5,11]); $log\,\overline{\sigma_{.j}}$ is the mean of the stress amplitude for the j-th group; $log\,\overline{\sigma_{..}}$ is the mean of the $log\sigma_{ij}$ values.

3.2. Testing the Homogeneity of Regression Slopes

An assumption underlying the use of ANCOVA is that the population regression slopes are equal. If the slopes are not equal, the "Type of groups" effects differ at different levels of the stress amplitude; consequently, the adjusted stress amplitudes can be misleading because they do not convey this important information. When the slopes are the same, the adjusted means are adequate descriptive measures because the differences of the fatigue life are the same at different levels of the stress amplitude.

If the slopes for the populations in an experiment are equal, that is, $\beta_1^{(group\ 1)} = \beta_1^{(group\ 2)} = \ldots = \beta_1^{(group\ k)}$, a reasonable way of estimating the value of this common slope from the samples is by computing an average of the sample b_1 values:

$$b_1^{(group\ j)} = \frac{\sum_{i=1}^{m_j} \left(logN_{ij} \right) \left(log\sigma_{ij} \right) - \frac{\left(\sum_{i=1}^{m_j} logN_{ij} \right) \left(\sum_{i=1}^{m_j} log\sigma_{ij} \right)}{m_j}}{\sum_{i=1}^{m_j} \left(log\sigma_{ij} \right)^2 - \frac{\left(\sum_{i=1}^{m_j} log\sigma_{ij} \right)^2}{m_j}} \tag{5}$$

where from Equation (5) we have:

$$b_w = \sum_{j=1}^{k} b_1^{(group\ j)} \tag{6}$$

The slope b_w is the best estimate of the population slope β_1, which is the slope assumed to be common to all groups. As long as $\beta_1^{(group\ 1)} = \beta_1^{(group\ 2)} = \ldots = \beta_1^{(group\ k)}$, the estimate b_w is a useful statistical value to use. Now the problem is to decide whether all groups have the same slope. The homogeneity of the regression F-test is designed to answer the question of the equality of the population slopes. The null hypothesis associated with this test is:

$$H_0 : \beta_1^{(group\ 1)} = \beta_1^{(group\ 2)} = \ldots = \beta_1^{(group\ k)} \tag{7}$$

The steps involved in the computation of the test are described next:

Steps 1 and 2: Computation of within-group sum of squares and within-group residual sum of squares ($SSres_w$)

This parameter can be evaluated by the following equation (for further details, see [5,11]):

$$E_{yy} = \sum_{j=1}^{k} \left[\sum_{i=1}^{m_j} \left(logN_{ij} \right)^2 - \frac{\left(\sum_{i=1}^{m_j} logN_{ij} \right)^2}{m_j} \right] \tag{8}$$

$$E_{xy} = \sum_{j=1}^{k} \left[\sum_{i=1}^{m_j} (logN_{ij})(log\sigma_{ij}) - \frac{\left(\sum_{i=1}^{m_j} logN_{ij}\right)\left(\sum_{i=1}^{m_j} log\sigma_{ij}\right)}{m_j} \right] \tag{9}$$

$$E_{xx} = \sum_{j=1}^{k} \left[\sum_{i=1}^{m_j} (log\sigma_{ij})^2 - \frac{\left(\sum_{i=1}^{m_j} log\sigma_{ij}\right)^2}{m_j} \right] \tag{10}$$

$$SSreg_w = \frac{(E_{xy})^2}{E_{xx}} \tag{11}$$

$$SSres_w = E_{yy} - SSreg_w \tag{12}$$

Step 3: Computation of individual sum of squares residual (*SSres$_i$*)

The third step involves the computation of the sum of squares residual for each group separately, and then adding these residuals to obtain the sum of the individual residual sum of squares (*SSres$_j$*). The difference in the computations of *SSres$_w$* and *SSres$_j$* is that *SSres$_w$* involves computing the residual sum of squares around the single b_w value whereas *SSres$_j$* involves the computation of the residual sum of squares around the b_j values fitted to each group separately (Equation (13)).

$$SSres_j = \left[\sum_{i=1}^{m_j} (logN_{ij})^2 - \frac{\left(\sum_{i=1}^{m_j} logN_{ij}\right)^2}{m_j} \right]$$
$$- \left[\frac{\sum_{i=1}^{m_j} (logN_{ij})(log\sigma_{ij}) - \frac{\left(\sum_{i=1}^{m_j} logN_{ij}\right)\left(\sum_{i=1}^{m_j} log\sigma_{ij}\right)}{m_j}}{\sum_{i=1}^{m_j} (log\sigma_{ij})^2 - \frac{\left(\sum_{i=1}^{m_j} log\sigma_{ij}\right)^2}{m_j}} \right] \tag{13}$$

Step 4: Computation of heterogeneity of slopes sum of squares

The discrepancy between *SSres$_w$* and *SSres$_i$* reflects the extent to which the individual regression slopes are different from the within-group slope b_w; hence, the heterogeneity of slopes *SS* is "$SS_{het} = SSres_w - SSres_i$", see [5].

Step 5. Computation of *F*-ratio

The summary table for the *F*-test is as follows (Table 2). If the obtained *F* is equal to or greater than $F_{[\alpha, j-1, N-2j]}$, then the null hypothesis $H_0 : \beta_1^{(group\ 1)} = \beta_1^{(group\ 2)} = \ldots = \beta_1^{(group\ k)}$ is rejected.

Table 2. Computation table for the heterogeneity of the slope.

Source of Variation	Sum of Square	Degrees of Freedom	Mean Square	*F*
Heterogeneity of slopes	SS_{het}	$k-1$	$MS_{het} = SS_{het}/(k-1)$	MS_{het}/MS_{res_i}
Individual residual (*res$_i$*)	$SSres_i$	$M-2k$	$MS_{res_i} = SS_{res_i}/(M-2k)$	-
Within residual (*res$_w$*)	$SSres_w$	$M-k-1$	-	-

In this method, diagnostic checking of the covariance model is based on residual analysis. Furthermore, the measure of uncertainty is not directly related to the uncertainty of the SN curve. The uncertainty of SN curves used in the next section has been evaluated by the maximum likelihood method (MLM) [2].

4. Reliability Assessment

The fatigue strength of metals is often assumed to follow the Basquin equation (the equation is based on fully reversed fatigue ($R = -1$), i.e., the mean value is zero) and is written [12]:

$$N = K(\Delta\sigma)^{-m} \tag{14}$$

where N is the number of stress cycles to failure with constant stress ranges $\Delta\sigma$; K and m are material parameters dependent on the fatigue critical detail. The fatigue strength $\Delta\sigma_F$ may, e.g., be defined as the value of S for, e.g., $N_D = 2 \times 10^6$ [7]. If one fatigue critical detail is considered, then the annual probability of failure is obtained from:

$$\Delta P_{F,t} = P_{COL|FAT}P(\text{fatigue failure in year t}) \tag{15}$$

where $P(\text{fatigue failure in year t})$ is the probability of failure in year t and $P_{COL|FAT}$ is the probability of collapse of the structure given fatigue failure—modeling the importance of the details/consequences of failure. The probability of failure in year t given survival up to year t is estimated by:

$$\Delta P_{F,t} = P_{COL|FAT}(P(g(t) \leq 0) - P(g(t-1) \leq 0))/P(g(t) > 0) \tag{16}$$

where the limit state equation is based on the application of SN curves and Miner's rule for linear accumulation of fatigue damage, and by introducing stochastic variables accounting for uncertainties in fatigue loading and strength. Note that the probability of failure also depends on the repair and maintenance methods' accuracy and protective methods that have been applied on the component. The design equation can be written as follows, if used in a deterministic code-based verification:

$$D = \sum_i \frac{n_{i,S}T_L}{N(\Delta\sigma_i, \Delta\sigma_F, \gamma_F, \gamma_m, \Delta K, z)} = 1 \tag{17}$$

where $n_{i,S}$ represents the number of cycles per year at a specific stress level and T_L is the design lifetime. It is assumed that for a wind turbine component, for a given fatigue life, the number of cycle according to the stress can be grouped as n_σ, where the number of excitation at the specific stress range i is $n_{i,S}$ per year. In this paper, the Level II reliability method is used to estimate the reliability of the components [7]. The design parameter z is obtained from Equation (17), assuming that the fatigue partial safety factors are given. Consequently, the reliability equation is normalized and it became a function of the partial safety factors and presumed that the component is designed to the limit according to the design parameter z in the design equation.

For a deterministic design, the following partial safety factors are introduced [13]:

γ_f　a fatigue load partial safety factor multiplied by the fatigue stress ranges obtained by, e.g., rainflow counting.

γ_m　a fatigue strength partial safety factor. The design value of the fatigue strength is obtained by dividing the characteristic fatigue strength by γ_m.

The characteristic fatigue strength can be defined in various ways, namely based on:

- the mean minus two standard deviations of log K.
- the 5% quantile of log K, i.e., the mean minus 1.65 times the standard deviation of log K.
- the mean of log K.

The corresponding limit state equation to be used in the reliability analysis is written:

$$g(t) = \Delta - \sum_i \frac{n_{i,S}t}{N(X_W, X_{SCF}, \Delta\sigma_i, \Delta\sigma_F, \Delta K, z)} \tag{18}$$

where X_W is a stochastic variable modeling model uncertainty related to the determination of fatigue loads and X_{SCF} is a stochastic variable modeling model uncertainty related to the determination of stresses given fatigue loads. In addition, Δ models model uncertainty related to Miner's rule for linear damage accumulation [2]. In Equation (18), Δ, X_W and X_{SCF} are assumed to be log-normal distributed with mean values equal to 1 and coefficients of variation COV_Δ, COV_W and COV_{SCF}, respectively, and $N(X_W, X_{SCF}, \Delta\sigma, \Delta\sigma_F, \Delta K, z)$ can be written as following equation:

$$N(X_W, X_{SCF}, \Delta\sigma_i, \Delta\sigma_F, \Delta K, z) = \Delta K \left(\frac{X_W X_{SCF} \Delta\sigma_i}{z\Delta\sigma_F} \right)^{-m} \tag{19}$$

The coefficients of variation are estimated partly subjectively, but generally following the recommendations used as a basis for the material partial safety factors in IEC 61400-1, and also considering information from, e.g., [14]. The importance of the choices of the coefficients of variation is investigated by sensitivity analyses. It is noted that the reliability level obtained is in accordance with the target reliability corresponding to an annual probability failure of the order 5×10^{-4} (annual reliability index 3.3) [15]. In Table 3, the stochastic model is shown. It is noted that m and $\Delta\sigma_f$ are correlated with statistical parameters extracted from [2]. The stochastic model is considered as representative for the fatigue strength represented by SN curves. It is assumed that the design lifetime is $T_L = 25$ year [13].

Table 3. Stochastic model of fatigue strength based on SN curves.

Variable	Definition	Distribution	Expected Value	Coefficient of Variation
Δ	Model uncertainty related to Miner's rule	LN *	1	$COV_\Delta = 0.3$
X_{SCF}	Model uncertainty related to determination of stresses given fatigue load	LN	1	COV_{SCF}
X_W	Model uncertainty related to determination of fatigue loads	LN	1	COV_W
m	Slope SN curve (Statistical uncertainty)	N **	Extracted from test results, [2]	
$\Delta\sigma_f$ (Mpa)	Fatigue Strength (Statistical uncertainty)	N	Extracted from test results, [2]	

* Log-Normal Distribution; ** Normal Distribution.

If the SN curves are obtained by a limited number of tests, then statistical uncertainty has to be accounted for. Table 4 shows indicative values of $\gamma_f \gamma_m$ for the target reliability index equal to 3.3 as a function of the total coefficient of variation of the fatigue load: $COV_{load} = \sqrt{COV_W^2 + COV_{SCF}^2}$. It is noted that more fatigue test data should be investigated to validate the indicative values in Table 4.

Table 4. Indicative partial safety factors $\gamma_f \gamma_m$ as function of *COV* for fatigue load.

$\Delta\beta_{min,FAT}/COV_{load}$	0.00	0.05	0.10	0.15	0.20	0.25	0.30
3,3 (5×10^{-4})	1.62	1.63	1.67	1.74	1.83	1.95	2.11

5. Results and Discussion

This section presents results obtained using representative fatigue test data from test specimens of cast components. The casted specimens are manufactured by two different manufacturing processes. The manufacturing processes are "sand casting" and "chill casting". In this section, the ANCOVA method is used to compare the different test laboratories and different extractions of samples for the tests and obtaining the fatigue test data.

5.1. Comparison of Different Test Laboratories

The fatigue tests were carried out by four different test laboratories. Table 5 shows the number of tests in each group. The comparison between groups (Tests labs) was performed using sand casting results because there was only one group with chill casting. In this configuration, the run-out samples were considered as broken samples with very high cycles in order not to exclude them from the analysis. A summary of the ANCOVA calculations is shown in Table 6.

Table 5. Number of test results by test laboratories.

Testing Laboratories	Sand Casting		Chill Casting		Total
	Broken	**Run-out**	**Broken**	**Run-out**	
Group 1	28	2	-	-	30
Group 2	609	109	302	107	1127
Group 3	57	3	-	-	60
Group 4	19	-	-	-	19
Summation	713	114	302	107	1236

Table 6. Summary of calculations for test laboratories comparison. *SS*: Sum of squares.

Parameter	Value	Parameter	Value
S_{yy}	574	$SSres_w$	196
$SSres_t$	206	SS_{AT}	9.6
$SSreg_w$	377	b_w	−8.9

The value of b_w shows that the stress amplitude and fatigue life were negatively correlated. First it was tested if the choice of test laboratory does not affect the fatigue test results.

The null hypothesis is: H_0, the choice of test laboratory does not affect the fatigue test results.

The *F*-ratio is calculated according to Section 3. The study involves one covariate, four groups and 827 test results. The *F*-ratio calculation is summarized in Table 7.

Table 7. Computation table for the test laboratories comparison.

Source of Variation	Sum of Square	Degrees of Freedom	Mean Square	*F*	*p*-Value
Adjusted (AT)	9.56	3	3.2	13.3	0.00
Error (res_w)	196	822	0.2		
Total residual (res_t)	205	825	-		

The obtained *F* was then compared with the critical value of *F* with three and 822 degrees of freedom and a level of significance equal to 5%; $F_{(0.05, 3,822)}$ equals 2.62, and the null hypothesis of "The choice of test laboratory does not affect the fatigue test results" was rejected. Next, the homogeneity of the slopes of regression lines was considered, i.e., it was tested if the slopes of the different regression lines were equal. The null hypothesis is written: $H_0 : \beta_1^{(group\ 1)} = \beta_1^{(group\ 2)} = \ldots = \beta_1^{(group\ k)}$The summary of the calculation is shown in Table 8.

Table 8. Computation heterogeneity of slopes for test laboratories comparison.

Source of Variation	Sum of Square	Degrees of Freedom	Mean Square	*F*	*p*-Value
Heterogeneity of slopes	1.2	3	0.41	1.71	0.16
Individual residual (res_i)	194	821	0.24		
Within residual (res_w)	196	824	-		

The critical value is $F_{(0.05, 3, 821)} = 2.62$; hence, the null hypothesis was accepted. It was concluded that the slopes of the regression lines are similar. Figure 2 shows the results for the regression lines

(note the results are normalized). It was seen that "Group 2" had the highest fatigue life and "Group 1" had the lowest fatigue life compared to the other testing places and the groups had the same slope.
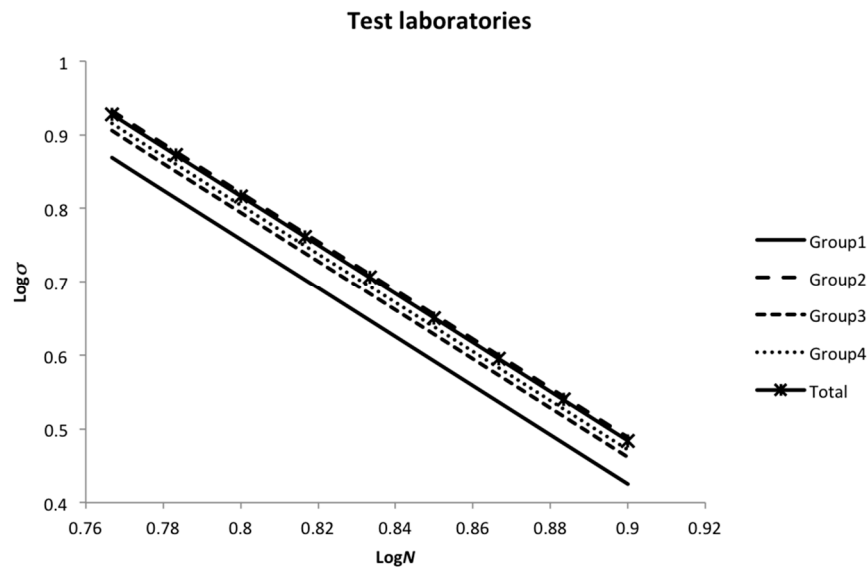

Figure 2. Comparison of "testing laboratories" on logarithmic scale.

5.2. Comparison of Different Cutting Facilities

The components from where samples were taken were cut up at different facilities. The cutting facilities were used to cut the sand casting samples after casting processes, and therefore they could be considered as a part of the manufacturing process for the casting samples. Table 9 shows the number of tests in each group. It is noted that the groups of cutting facilities are not the same as the groups of testing laboratories.

Table 9. Number of test results by cutting facilities.

Cutting Facilities	Sand Casting		Chill Casting		Total
	Broken	**Run-out**	**Broken**	**Run-out**	
Group 1	107	8	96	3	214
Group 2	18	3	-	-	21
Group 3	95	14	-	-	109
Group 4	148	26	107	53	334
Group 5	345	63	99	51	558
Summation	713	114	302	107	1236

In this case, the comparison between groups was performed using sand casting and chill casting. The results of the ANCOVA for sand casting and chill casting are shown in Tables 10 and 11, respectively.

Table 10. Computation table for the cutting facilities based on sand casting method.

Source of Variation	Sum of Square	Degrees of Freedom	Mean Square	F	*p*-Value
Adjusted (AT)	16	4	3.8	16.3	0.00
Error (res_w)	191	821	0.2	-	-
Total residual (res_t)	206	825	-	-	-

Table 11. Computation table for the cutting facilities based on chill casting method.

Source of Variation	Sum of Square	Degrees of Freedom	Mean Square	F	p-Value
Adjusted (AT)	42	2	21.0	111	0.00
Error (res_w)	77	405	0.19	-	-
Total residual (res_t)	119	407	-	-	-

The obtained *F* statistic was then compared with the critical value. In both casting methods, the null hypothesis of "The choice of cutting facilities does not affect the fatigue test results" was rejected. Next, the homogeneity of the regression test (slope of regression lines) was considered. The summaries of the calculation are shown in Tables 12 and 13.

Table 12. Computation heterogeneity of slopes for cutting facilities based on sand casting.

Source of Variation	Sum of Square	Degrees of Freedom	Mean Square	F	p-Value
Heterogeneity of slopes	3.0	4	0.75	3.28	0.011
Individual residual (res_i)	181	817	0.23	-	-
Within residual (res_w)	190	821	-	-	-

Table 13. Computation heterogeneity of slopes for cutting facilities based on chill casting.

Source of Variation	Sum of Square	Degrees of Freedom	Mean Square	F	p-Value
Heterogeneity of slopes	3.4	2	1.68	9.24	0.000
Individual residual (res_i)	73.2	403	0.18	-	-
Within residual (res_w)	76.6	405	-	-	-

Based on Tables 12 and 13, the null hypothesis is rejected and slopes of regression lines are not the same. Figures 3 and 4 show the results for the regression lines of sand casting and chill casting, respectively (note the results are normalized).
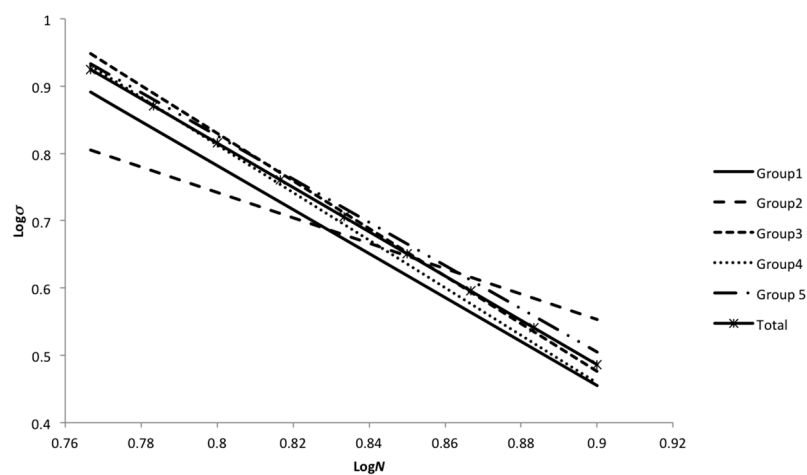
Figure 3. Comparison of "cutting facilities" on logarithmic scale for sand casting.

According to Figures 3 and 4, Group 2 has a different slope compared to the other groups, but the other groups are similar to each other. For sand casting, quite small differences between fatigue lives were obtained for cutting facilities 1, 3, 4 and 5. For chill casting, relatively larger differences in fatigue lives were obtained for cutting facilities 1, 4 and 5.
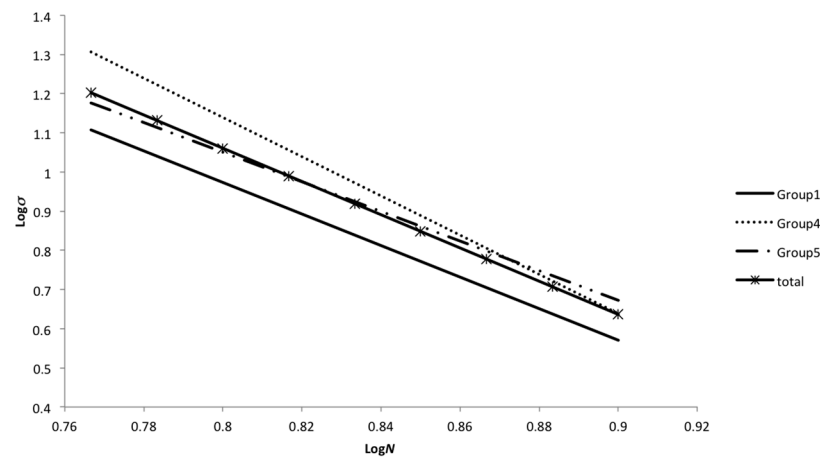
Figure 4. Comparison of "cutting facilities" on logarithmic scale for chill casting.

5.3. Reliability Analysis: Examples

In the following, the SN curves that were derived in the above sections are used for reliability analysis in a case study. The parameters used for analysis are listed in Table 14. By using the design Equation (17), the design values (*z*) are determined for each group, as shown in Tables 15 and 16.

Table 14. Statistical parameters for reliability assessment of samples.

COV_Δ	COV_{SCF}	COV_W	COV_{load}	fl_f	fl_m
0.3	0.1	0.05	0.11	0.95	1.75

Table 15. Design values (*z*) for testing lab groups (obtained from Equation (17)).

| Groups	Group 1	Group 2	Group 3	Group 4
z-value	0.228	0.203	0.214	0.210

Table 16. Design values (*z*) for cutting facilities groups (obtained from Equation (17)).

Groups	Group 1	Group 2	Group 3	Group 4	Group 5
z-value	0.226	0.510	0.178	0.184	0.214

The reliability indices as a function of time for the groups of testing laboratories (Section 5.1) were estimated and are shown in Figure 5. Note that in this figure the annual reliability index is shown. The data from Group 2 result in the largest reliability level compared to the other groups. Moreover, the annual reliability index for cutting samples groups (Section 5.2) is shown in Figure 6. It is seen that the annual reliability indices at the end of the lifetime are of the order of 3.3, corresponding to the target annual reliability index = 3.3.
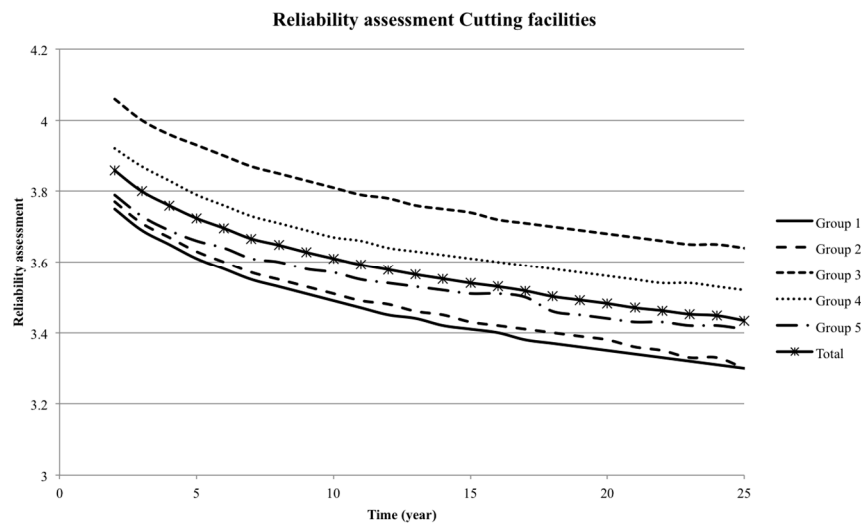
Figure 5. Annual reliability index for different groups based on test laboratories categories (sand casting).
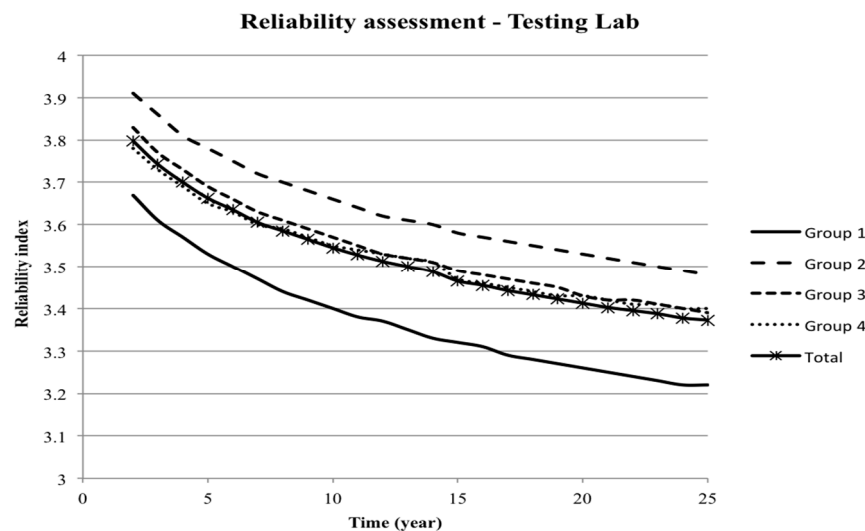
Figure 6. Annual reliability index for different groups based on cutting facilities categories (sand casting).

6. Conclusions

In this paper, the ANCOVA method is used to compare different groups related to the manufacturing process and associated quality control of cast components. The ANCOVA method is applied for fatigue failure data. Test data is used to illustrate the implementation of the ANCOVA method. The different test laboratories and cutting facilities are compared to each other. The results obtained from the ANCOVA analysis can be used for decision-making on which test laboratories and cutting facilities should be included in a statistical analysis in order to make sure that the statistical analyses are performed on data from a statistically homogeneous population. More parameters can be included, if relevant.

Further, it is presented how to apply the statistical results from the ANCOVA analysis to estimate the reliability for generic cases, and to assess the required safety factors for deterministic design such that a given target reliability level is obtained. It is noted that more data and more example

structures are needed to perform a more general reliability-based calibration of safety factors with the proposed approach.

Author Contributions: Hesam Mirzaei Rafsanjani contributed to defining the overall problem and proposed the core scientific idea to solve it. Hesam Mirzaei Rafsanjani established the ANCOVA analysis of SN curve models for fatigue samples that has been tested in the RISØ center (DTU Roskilde campus) and they had been provided by the industrial partner of the project. The test samples provided by "Asger Sturlason" and tests were carried out in the RISØ center by "Søren Fæster". Hesam Mirzaei Rafsanjani conducted all the numerical simulations included in the paper and interpreted all the results with the help of the supervisor (co-author John Dalsgaard Sørenesen). Further, Hesam Mirzaei Rafsanjani wrote the manuscript and revised it according to co-author comments.

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

References

1. Tavner, P.J.; Greenwood, D.M.; Whittle, M.W.G.; Gindele, R.; Faulstich, S.; Hahn, B. Study of weather and location effects on wind turbine failure rates. *Wind Energy* **2012**, *16*, 175–187. [CrossRef]

2. Rafsanjani, H.M.; Sørensen, J.D. Reliability analysis of fatigue failure of cast components for wind turbines. *Energies* **2015**, *8*, 2908–2923. [CrossRef]

3. Rafsanjani, H.M.; Sørensen, J.D. Stochastic models of defects in wind turbine drivetrain components. In *Multiscale Modeling and Uncertainty Quantification of Materials and Structures*; Papadrakakis, M., Stefanou, G., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 287–298.

4. Shirani, M.; Härkegård, G. Damage tolerant design of cast components based on defects detected by 3D X-ray computed tomography. *Int. J. Fatigue* **2012**, *41*, 188–198. [CrossRef]

5. Huitema, B.E. *The Analysis of Covariance and Alternatives Statistical Methods for Experiments, Quasi-Experiments, and Single-Case Studies*; Wiley: Hoboken, NJ, USA, 2011.

6. Sørensen, J.D.; Toft, H.S. Probabilistic design of wind turbines. *Energies* **2010**, *3*, 241–257. [CrossRef]

7. Sørensen, J.D.; Toft, H.S. *Safety Factors—IEC 61400-1 ed.*; 4—Background Document, DTU Wind Energy-E-Report-0066 (EN); DTU: Copenhagen, Denmark, 2014.

8. Madesn, H.O.; Krenk, S.; Lind, N.C. *Methods of Structural Safety*; Dover Publications: Mineola, NY, USA, 1986.

9. Zimmerman, D.W. A note on interpretation of the paired-samples t test. *J. Educ. Behav. Stat.* **1997**, *22*, 349–360. [CrossRef]

10. Ayyub, B.M.; McCuen, R.H. *Probability, Statistics, and Reliability for Engineers and Scientists*, 2nd ed.; CRC/Chapman & Hall: Boca Raton, FL, USA, 2002.

11. Montgomery, D.C. *Design and Analysis of Experiments*, 8th ed.; Wiley: Hoboken, NJ, USA, 2013.

12. Silk, J. *Analysis of Covariance and Comparison of Regression Lines*; University of East Anglia: Norwich, UK, 1979.

13. Sørensen, J.D. Reliability assessment of wind turbines. In Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering ICASP12, Vancouver, BC, Canada, 12–15 July 2015.

14. *Fatigue Design of Offshore Steel Structures*; DNV-RP-C203; Det Norske Veritas: Høvik, Norway, 2010.

15. *Wind Turbine—Part 1: Design Requirements*; IEC 61400; International Electrotechnical Commission (IEC): Geneva, Switzerland, 2015.