

# Improved Estimation of Weibull Parameters Considering Unreliability Uncertainties

Antonio Fernández and Manuel Vázquez

**Abstract**—We propose a linear regression method for estimating Weibull parameters from life tests. The method uses stochastic models of the unreliability at each failure instant. As a result, a heteroscedastic regression problem arises that is solved by weighted least squares minimization. The main feature of our method is an innovative s-normalization of the failure data models, to obtain analytic expressions of centers and weights for the regression. The method has been Monte Carlo contrasted with Benard's approximation, and Maximum Likelihood Estimation; and it has the highest global scores for its robustness, and performance.

**Index Terms**—Heteroscedastic regression, life tests, Weibull distributions, weighted least squares.

## ACRONYMS<sup>1</sup>

<b>HUWE</b>	Heteroscedastic Unreliability Weibull Estimation.
<b>LR</b>	Linear Regression.
<b>MLE</b>	Maximum Likelihood Estimation.
<b>MRSE</b>	Mean Relative Squared Error.
<b>MSE</b>	Mean Squared Error.
<b>WLS</b>	Weighted Least Squares.

## NOTATION

$t$	The time counted from the beginning of a widget's life
$F$	Unreliability, failure probability of a widget
$\eta, \beta$	Scale parameter, and shape parameter of a Weibull distribution
$N_M$	Number of widgets in a life test
$T$	Number of Monte Carlo trials for evaluation of estimators

Manuscript received October 10, 2010; revised February 09, 2011, March 03, 2011, and March 24, 2011; accepted March 26, 2011. Associate Editor: M. Xie.

The authors are with the EUIT Telecomunicación, Universidad Politécnica de Madrid, Spain.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TR.2011.2168652

<sup>1</sup>For the sake of simplicity, the singular and plural of acronyms are spelled the same.

## I. INTRODUCTION

THE Weibull model is the most popular, widely used model in reliability theory [1]. It states the unreliability as a function of time in the form

$$F(t; \eta, \beta) = 1 - \exp \left[ - \left( \frac{t}{\eta} \right)^\beta \right]. \quad (1)$$

The failure profile of a specific type of widgets is described by its  $\eta$  and  $\beta$  values. These parameters are estimated with the help of a *life test*, which is an experiment that outputs the series of failure times  $\{t_k\}$  of a set of widgets of the kind that one wishes to analyse. The ranks  $k \in [1, N_M]$  are natural labels that maintain a coherent order with respect to the failure times, that is  $k_i > k_j \Leftrightarrow t_{k_i} \geq t_{k_j}$ .

Bayesian methods [2] are the most rigorous, complete approaches to parameter estimation, but they are difficult to obtain and interpret. Moreover, the prior conjugate joint distribution for the Weibull parameters is not continuous [1], [3], which makes the approach even harder. Perhaps due to this complexity, Maximum Likelihood Estimation (MLE) [4]–[7], and Linear Regression (LR) [8], [9] approaches are commonly used. Although MLE estimators offer better accuracy when  $N_M$  is large enough, LR is preferred in some environments for the following reasons.

- The MLE shape parameter estimator  $\hat{\beta}$  is s-biased when the sample size is small [10], [11], although resources [5], [12]–[14] have been developed to overcome this drawback.
- MLE leads to equations that are not solvable in explicit form, so numerical procedures are needed, and convergence and numerical problems arise [15], [16].
- Engineers often choose LR methods because of their computational simplicity, and ease of graphical interpretation [17], [18][19].

LR methods proceed transforming the Weibull failure model (1) to obtain an affine model suitable to be fitted by regression techniques. Isolating the exponential in the second term of (1), and taking logarithms leads to a first reliability-related variable

$$\Phi = \ln \left( \frac{1}{1 - F} \right) = -\ln(1 - F) = \left( \frac{t}{\eta} \right)^\beta. \quad (2)$$

Taking logarithms once again leads to a second reliability-related variable

$$Y = \ln \Phi = -\beta \ln \eta + \beta \ln t. \quad (3)$$

From (3), an affine relationship (4) can be established between  $Y$  and the time-related variable  $X = \ln t$ .

$$Y = \beta X - \beta \ln \eta \quad (4)$$

A key problem that arises when regressing on model (4) is how to determine data point coordinates in the  $(X, Y)$  space from the life test. Variable  $X_k$  can be determined from failure times  $t_k$ , and  $Y_k$  can be determined from the unreliabilities at failures  $F_k$  as in

$$\begin{aligned} Y_k &= \ln(\Phi_k) = \ln\left(\ln\left(\frac{1}{1-F_k}\right)\right) = \ln(-\ln(1-F_k)) \\ X_k &= \ln(t_k) \end{aligned} \quad (5)$$

But  $F_k$  are probabilities, so they cannot be directly observed. The most straightforward choice for  $F_k$  are relative frequencies  $F_k = k/N_M$ ; however, this approach is not very accurate, and gets a strong  $s$ -bias under transformations (2) and (3). Indeed, for the last point,  $k = N_M$ , it yields  $F_{N_M} = 1$ , which overflows  $Y_{N_M} = \ln(-\ln(0)) = \infty$ .

A better choice for  $F_k$  is Median Rank [11], [20]. Median Rank cannot be put into explicit form, so its value has to be determined by numerical techniques or approximated methods. A simple analytic approach to Median Rank, which has become a standard way to estimate  $F_k$ , is Benard's approximation [21]:

$$F_k = \frac{k - 0,3}{N_M + 0,4}. \quad (6)$$

Sometimes it is assumed that  $Y_k$  are homoscedastic, i.e. that the  $Y_k$  uncertainties are the same for any  $k$ ; then the regression line should minimize the simple residual

$$\text{residual}(Y) = \sum_{k=1}^{N_M} (Y_k - \hat{Y}_k)^2 = \sum_{k=1}^{N_M} (Y_k - \beta X_k - \beta \ln \eta)^2. \quad (7)$$

In fact,  $Y_k$  are heteroscedastic as its uncertainties are not the same. More elaborate regression algorithms such as Weighted Least Squares (WLS) [22], [23] take into account the heteroscedasticity by weighting terms of the residual with uncertainty-related factors  $w_k$  (8).

$$\text{weighted residual}(Y) = \sum_{k=1}^{N_M} w_k (Y_k - \hat{Y}_k)^2 \quad (8)$$

The weights are commonly chosen as the inverses of the variances of  $Y_k$  (9).

$$w_k = \frac{1}{\sigma_{Y_k}^2} \quad (9)$$

Uncertainties of  $Y_k$  have been considered by some authors. Bergman [24] assumes that the variance  $\sigma_{F_k}^2$  of  $F_k$  is constant in order to find analytically  $\sigma_{Y_k}^2$  in consonance with transformations (2), and (3). Bergman's hypothesis, although it is useful, is not very accurate, as will be shown in Fig. 2, where  $\sigma_{F_k}^2$  is plotted against  $k$ , revealing a considerable variation. Faucher and Tyson [25] evaluate  $F_k$  uncertainty from the binomial model

with prefixed confidence levels. Their method requires carrying numerical tables to evaluate the weights.

In Section II, we develop the Heteroscedastic Unreliability Weibull Estimation (HUWE). HUWE is a new LR method based on the assumption of the  $s$ -normality of  $Y_k$ . This approximation allows us to find, working exclusively with analytical procedures, highly accurate estimates for  $Y_k$  centers and weights in terms of hyper-geometric functions [26]. Analyticity makes HUWE based software straightforward, reliable, and easily portable because it can be designed without numerical tables nor 'ad hoc' iterative solving procedures. It is well known that these two resources compromise the reliability and portability of software: numerical tables because errors while typing can lead to very difficult to detect errors; and 'ad hoc' iterative solving procedures because they involve convergence problems.

To clarify the methodology, an illustrative example of HUWE, applied to a small life test, is given in Section III. In Section IV, a Monte Carlo evaluation procedure is designed to compare HUWE performance with other methods. The results of contrast with Benard and MLE methods are given in Section V. According with these results, HUWE performs the best when we consider a wide range of possible  $N_M$  values. Moreover, HUWE is more computationally efficient and robust than MLE, especially when  $N_M$  is large, and the shape parameter  $\beta$  is far from 1.

## II. HUWE REGRESSION METHODOLOGY

HUWE is a weighted LR method based on a statistical model for  $F_k$  that uses as the centralization, and dispersion of  $Y_k$  in (8) the means  $\mu_{Y_k}$ , and variances  $\sigma_{Y_k}^2$  respectively.

The problem with this approach is that the exact determination of the moments of  $Y_k$  leads to integrals that cannot be put in a closed form. HUWE fixes this problem by assuming  $s$ -normal models for  $Y_k$ . This assumption opens up a way to approximate  $Y_k$  moments as functions of the moments of the intermediate variables  $\Phi_k$ . As the moments of  $\Phi_k$  can be put in closed form, this solves the problem, at the cost of a little information leakage. An outline of the HUWE methodology is depicted in Fig. 1. The steps in Fig. 1 will be described below.

### A. Unreliability Data Points $F_k$ Model

A well established [27] statistical model for  $F_k$  that assumes the principle of indifference on the prior probability yields

$$\rho_{F_k}(f) = N_M \binom{N_M - 1}{k - 1} f^{k-1} (1-f)^{N_M - k}. \quad (10)$$

Eqn. (10) gives the probability density function of the unreliability  $F_k$  just at the precise instant  $t_k$  of the  $k$ -th failure of a total of  $N_M$  widgets. The model (10) can be derived by considering that the number of ways that  $N_M$  failures can consist of  $k-1$  failures before  $t_k - (dt/2)$ ,  $N_M - k$  failures after  $t_k + (dt/2)$ , and just one failure inside  $[t_k - (dt/2), t_k + (dt/2)]$ , is  $N_M! / (k-1)! (N_M - k)!$ .

Eqn. (10) can be also expressed in terms of the Beta Euler function, resulting in

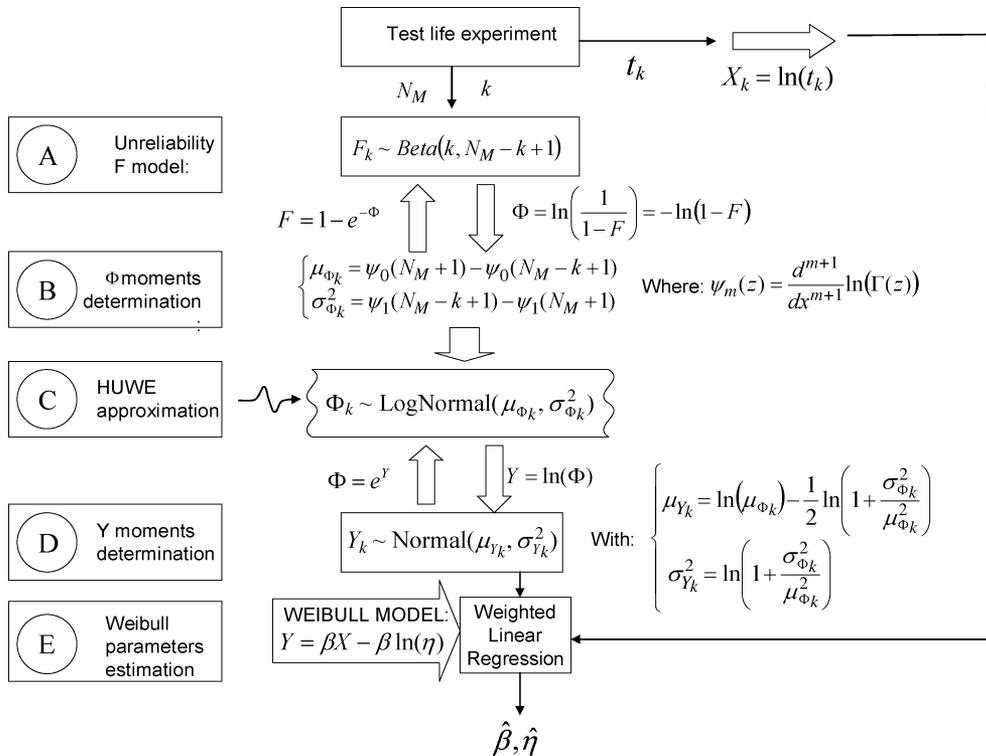
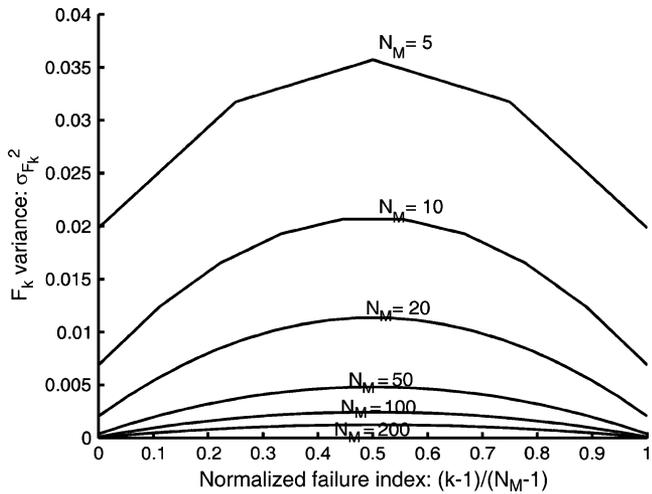


Fig. 1. HUWE methodology explanation.


 Fig. 2.  $\sigma_{F_k}^2$  as a function of the normalized index of failure  $k - 1/N_M - 1$  for several values of  $N_M$ .

$$\rho_{F_k}(f) = \frac{1}{\text{Beta}(k, N_M - k + 1)} f^{k-1} (1-f)^{N_M-k}, \quad (11)$$

or can be summarized in terms of the Beta distribution as

$$F_k \sim \mathbf{Beta}(k, N_M - k + 1). \quad (12)$$

To show the amount of heteroscedasticity of the model (10), in Fig. 2, the variance  $\sigma_{F_k}^2$  has been plotted as a function of the normalized rank,  $k - 1/N_M - 1$ , for several values of  $N_M$ .

From Fig. 2, it is clear that  $\sigma_{F_k}^2$  is not constant, but reach maxima at about  $k - 1/N_M - 1 = 1/2$ , and minima for the

first, and last failures ( $k = 1$ , and  $k = N_M$ ). Moreover, it can be established that the ratio between the maximum and minimum variances (that qualifies the degree of heteroscedasticity) grows with  $N_M$  as  $\sigma_{\max}^2/\sigma_{\min}^2 \approx (N_M + 2)/4$ . This result means that Bergman's approximation [24] gets worse as  $N_M$  increases.

### B. $\Phi_k$ Moment Determination

Now we estimate the moments of  $\Phi_k$  defined as in (13), according to (2).

$$\Phi_k = \ln\left(\frac{1}{1-F_k}\right) = -\ln(1-F_k) \quad (13)$$

By applying the mean and variance definitions to  $\Phi_k$ , we obtain

$$\begin{aligned} \mu_{\Phi_k} &= \int_0^1 -\ln(1-f) \frac{f^{k-1}(1-f)^{N_M-k}}{\text{Beta}(k, N_M - k + 1)} df \\ \sigma_{\Phi_k}^2 &= \int_0^1 (-\ln(1-f) - \mu_{\Phi_k})^2 \frac{f^{k-1}(1-f)^{N_M-k}}{\text{Beta}(k, N_M - k + 1)} df \end{aligned} \quad (14)$$

1) Mean of  $\Phi_k$ : For the mean of  $\Phi_k$ ,  $\mu_{\Phi_k}$ , we have

$$\begin{aligned} \mu_{\Phi_k} &= \int_0^1 \phi_k(f) \rho_{F_k}(f) df \\ &= \int_0^1 -\ln(1-f) \frac{f^{k-1}(1-f)^{N_M-k}}{\text{Beta}(k, N_M - k + 1)} df. \end{aligned} \quad (15)$$

Eqn. (15) can be analytically solved given

$$\mu_{\Phi_k} = H_{N_M} - H_{N_M-k} \quad (16)$$

where  $H_m = \sum_{j=1}^m (1/j)$  is the  $m$ -th harmonic number. Harmonic numbers can be put in terms of the polygamma function [26]  $\psi_m(z) = (d^{m+1}/dz^{m+1}) \ln(\Gamma(z))$  as  $H_m = \psi_0(m+1) + \gamma = \psi_0(m+1) - \psi_0(1)$ , where  $\gamma = -\psi_0(1) = 0.57721\dots$  is the famous Euler-Mascheroni constant, and  $\Gamma(z) \equiv (z-1)! = \int_0^\infty u^{z-1} e^{-u} du$ . Functions  $\psi_m(z)$  are implemented in an accurate, robust way in multiple software packets and programs such as EISPACK, MATLAB, and MATHEMATICA.

We can use the former relationships to rewrite (16) in terms  $\psi_m(z)$ , and arrive eventually at

$$\mu_{\phi_k} = \psi_0(N_M + 1) - \psi_0(N_M - k + 1). \quad (17)$$

2) *Variance of  $\Phi_k$* : To get  $\sigma_{\Phi_k}^2$ , we first calculate the non-centered second moment  $\mu_{\Phi_k^2}$  of  $\Phi_k$  to be

$$\begin{aligned} \mu_{\Phi_k^2} &= \int_0^1 \phi_k^2(f) \rho_{F_k}(f) df \\ &= \int_0^1 (-\ln(1-f))^2 \frac{f^{k-1}(1-f)^{N_M-k}}{\text{Beta}(k, N_M-k+1)} df. \end{aligned} \quad (18)$$

(18) gives  $\mu_{\Phi_k^2} = (H_{N_M} - H_{N_M-k})^2$ . Now, by applying the general identity  $\sigma_{\Phi_k}^2 = \mu_{\Phi_k^2} - \mu_{\Phi_k}^2$ , taking into account (16), and using polygamma functions, we reach

$$\sigma_{\Phi_k}^2 = \psi_1(N_M - k + 1) - \psi_1(N_M + 1). \quad (19)$$

Collecting the above expressions for the mean (17), and the variance (19), a remarkable symmetry becomes evident that widely justifies the transformations of both expressions in terms of polygamma functions (20):

$$\begin{aligned} \mu_{\Phi_k} &= +\psi_0(N_M + 1) - \psi_0(N_M - k + 1) \\ \sigma_{\Phi_k}^2 &= -\psi_1(N_M + 1) + \psi_1(N_M - k + 1) \end{aligned} \quad (20)$$

### C. HUWE Approximation

To determine exactly the relevant moments of  $Y_k$ , the integrals (21) should be performed.

$$\begin{aligned} \mu_{Y_k} &= \int_0^1 \ln(-\ln(1-f)) \frac{f^{k-1}(1-f)^{N_M-k}}{\text{Beta}(k, N_M-k+1)} df \\ \sigma_{Y_k}^2 &= \int_0^1 (\ln(-\ln(1-f)) - \mu_{Y_k})^2 \frac{f^{k-1}(1-f)^{N_M-k}}{\text{Beta}(k, N_M-k+1)} df \end{aligned} \quad (21)$$

Unfortunately, (21) cannot be solved in closed form, so HUWE uses an alternate method to approximate the moments of  $Y_k$  through the moments of  $\Phi_k$ . Basically, the HUWE approximation is the assumption of the  $s$ -normality of  $Y_k$ . This assumption is domain coherent because the domain of a  $s$ -normal distribution is the whole real line  $(-\infty, \infty)$ , which

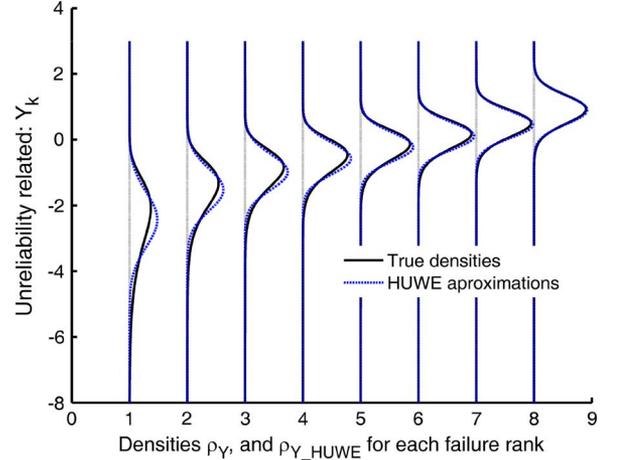


Fig. 3. True and HUWE approximated  $Y_k$  densities evolution along failure rank (for  $N_M = 8$ ).

matches the domain of  $Y_k$  because the successive application of transformations (2) and (3) maps the probability domain  $F_k \in (0, 1)$  to  $Y_k \in (-\infty, \infty)$ . A deductive consequence of the  $s$ -normality of  $Y_k$  is that  $\Phi_k$  should be assumed to be log-normally distributed because both are related by (3). Of course, the domain of  $\Phi_k$  is the transform of  $F_k \in (0, 1)$  by (2); that is,  $\Phi_k \in (0, \infty)$ , as should be expected for a genuine log-normal. The accuracy of HUWE will be discussed in more detail in Section II-D.

### D. $Y_k$ Moments Determination

By applying well known relations (22) between the first moments ( $\mu_N$ , and  $\sigma_N^2$ ) of a  $s$ -normal random variable  $N$ , and the first moments ( $\mu_L$ , and  $\sigma_L^2$ ) of the log-normal distributed variable  $L = e^N$ ,

$$\begin{aligned} \mu_N &= \ln(\mu_L) - \frac{1}{2} \ln \left( 1 + \frac{\sigma_L^2}{\mu_L^2} \right) \\ \sigma_N^2 &= \ln \left( 1 + \frac{\sigma_L^2}{\mu_L^2} \right) \end{aligned} \quad (22)$$

to  $\Phi = e^Y$ , we can express the HUWE approximation as

$$\begin{aligned} \mu_{Y_k} &\approx \ln(\mu_{\Phi_k}) - \frac{1}{2} \ln \left( 1 + \frac{\sigma_{\Phi_k}^2}{\mu_{\Phi_k}^2} \right) \\ \sigma_{Y_k}^2 &\approx \ln \left( 1 + \frac{\sigma_{\Phi_k}^2}{\mu_{\Phi_k}^2} \right) \end{aligned} \quad (23)$$

Hence, (23) is more accurate as  $Y_k$  gets closer to  $s$ -normality.

To qualitatively show the accuracy of the HUWE approximation, Fig. 3 depicts the true, and approximate  $Y_k$  densities. It can be seen that HUWE provides a good approximation, even for low values of  $N_M$  and  $k$ , and that the approximation gets better when  $k$  grows.

A more precise analysis of the error introduced by HUWE approximation can be performed by estimating the Kullbak-Lleiber distance [28] from the true to the approximate density (24). The Kullbak-Lleiber distance is a measure of the information leakage derived by assuming the approximated density  $\rho_{Y\_HUWE}$  instead of the true  $\rho_Y$ .

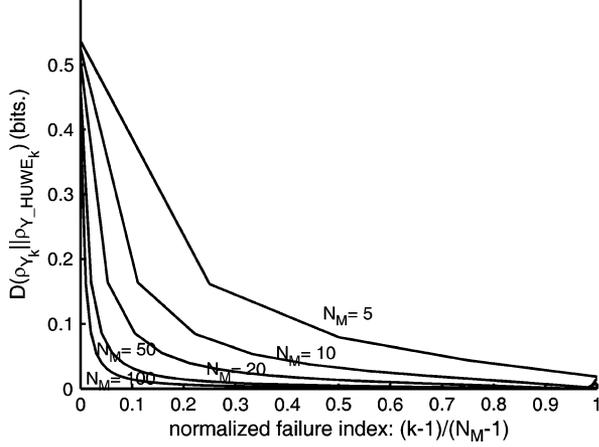


Fig. 4. Kullback-Lleiber information leakage caused by the approximation:  $Y_k \approx Y_{HUWE_k}$ .

$$D(Y||Y_{HUWE}) = \int_{-\infty}^{\infty} \rho_Y(y) \log_2 \left( \frac{\rho_Y(y)}{\rho_{Y_{HUWE}}(y)} \right) dy \quad (24)$$

Eqn. (24) has been evaluated as a function of the normalized failure index  $k-1/N_M-1$  for several values of  $N_M$ , and plotted in Fig. 4. The information leakage remains under 0.6 bits, and gets very small as  $k$  grows, especially for high values of  $N_M$ .

### E. Weibull Parameter Estimation

Once  $\mu_{Y_k}$  and  $w_k = 1/\sigma_{Y_k}^2$  have been determined by (23), they are supplied to a WLS [22] regression algorithm as centralization and weight values for the response variable. The explanatory values  $X_k$  for the regression are derived directly as  $X_k = \ln(t_k)$ .

The moments of  $Y_k$  do not depend on  $\{t_k\}$ , but only on  $N_M$  and  $k$ , so they can be pre-computed before the experiment is done in order to improve time in repeated trials.

Using WLS in step E means that HUWE assumes that the data are uncensored because WLS considers only errors in  $Y_k$ , whereas data censoring also entails errors in  $X_k$  due to the uncertainty of  $t_k$ . As it is clear from Fig. 1, step E is procedurally independent of the previous ones, so HUWE steps A to D can be used also for censored data. However, to address the whole problem, a protocol for modeling time uncertainties has to be defined, and the use of double axis heteroscedastic special regression techniques [29] is required.

## III. EXAMPLE

To clarify the HUWE methodology, we present here a simple example. Suppose we have performed a simple experiment with five widgets  $N_M = 5$ , and we have obtained the failure times shown below.

From test life data in Table I, we put in Table II the relevant quantities for the estimation. We begin by putting failure times in the first column, and associated ranks in the third column. The abscissa coordinates in the second column are set to  $X_k =$

TABLE I  
EXAMPLE TIME SERIES

$t_k$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$
(days)	0.9545	1.0195	1.7962	3.1409	3.8189

TABLE II  
EXAMPLE OF THE HUWE PROCESS

$t_k$	$X_k$	$k$	$\mu_{\phi_k}$	$\sigma_{\phi_k}^2$	$\mu_{Y_k}$	$\sigma_{Y_k}^2$
0.9545	-0.0465	1	0.2	0.04	-1.956	0.6931
1.0195	0.0194	2	0.45	0.1025	-1.0033	0.4096
1.7962	0.5857	3	0.7833	0.2136	-0.3936	0.2987
3.1409	1.1445	4	1.2833	0.4636	0.1254	0.248
3.8189	1.34	5	2.2833	1.4636	0.7019	0.2474

In  $t_k$ . Then, we apply the HUWE methodology in Fig. 1. Step A states the unreliability model (12) for  $N_M = 5$ ; that is,  $F_k \sim \mathbf{Beta}(k, 6 - k)$ ; step B gives (20) to obtain  $\mu_{\phi_k} = \psi_0(6) - \psi_0(6 - k)$ , and  $\sigma_{\phi_k}^2 = -\psi_1(6) + \psi_1(6 - k)$  in the fourth, and fifth columns. Assuming HUWE approximation (step C), use (23) of step D to get  $\mu_{Y_k}$ , and  $\sigma_{Y_k}^2$  for the sixth, and seventh columns.

Finally, in step E,  $X_k$ ,  $\mu_{Y_k}$  and  $\sigma_{Y_k}^2$  are supplied to WLS [22] to achieve the minimization of (8). This result yields the slope  $m$ , and the y-intercept  $b$  of the regression line as

$$m = \frac{-\left(\sum_{k=1}^{N_M} \frac{1}{\sigma_{Y_k}^2}\right) \left(\sum_{k=1}^{N_M} \frac{X_k \mu_{Y_k}}{\sigma_{Y_k}^2}\right) + \left(\sum_{k=1}^{N_M} \frac{X_k}{\sigma_{Y_k}^2}\right) \left(\sum_{k=1}^{N_M} \frac{\mu_{Y_k}}{\sigma_{Y_k}^2}\right)}{\left(\sum_{k=1}^{N_M} \frac{X_k}{\sigma_{Y_k}^2}\right)^2 - \left(\sum_{k=1}^{N_M} \frac{1}{\sigma_{Y_k}^2}\right) \left(\sum_{k=1}^{N_M} \frac{X_k^2}{\sigma_{Y_k}^2}\right)} \quad (25)$$

$$b = \frac{\left(\sum_{k=1}^{N_M} \frac{X_k}{\sigma_{Y_k}^2}\right) \left(\sum_{k=1}^{N_M} \frac{X_k \mu_{Y_k}}{\sigma_{Y_k}^2}\right) + \left(\sum_{k=1}^{N_M} \frac{X_k^2}{\sigma_{Y_k}^2}\right) \left(\sum_{k=1}^{N_M} \frac{\mu_{Y_k}}{\sigma_{Y_k}^2}\right)}{\left(\sum_{k=1}^{N_M} \frac{X_k}{\sigma_{Y_k}^2}\right)^2 - \left(\sum_{k=1}^{N_M} \frac{1}{\sigma_{Y_k}^2}\right) \left(\sum_{k=1}^{N_M} \frac{X_k^2}{\sigma_{Y_k}^2}\right)} \quad (26)$$

Using (25) and (26) for the data in Table II, and  $N_M = 5$  gives  $m = 1.440$ , and  $b = -1.338$ . Thus, according to (4), the estimates  $\hat{\beta}_{HUWE} = m = 1.440$ , and  $\hat{\eta}_{HUWE} = e^{-(b/m)} = 2.53$  days are obtained

## IV. EVALUATION PROCEDURE

In this section, we will describe the procedure, depicted in Fig. 5, that is used to evaluate HUWE. The HUWE performance will be determined by contrasting its ability to guess the (hidden) value of the Weibull parameters  $\{\eta_{true}, \beta_{true}\}$  from a succession of failure times, against the ability of other methods to do the same as references. The successions of failure times that are supplied to the competing methods are synthesized as complete, and  $s$ -independent  $Weibull(\eta_{true}, \beta_{true})$  pseudo-random samples. Each succession defines a trial (that is indexed

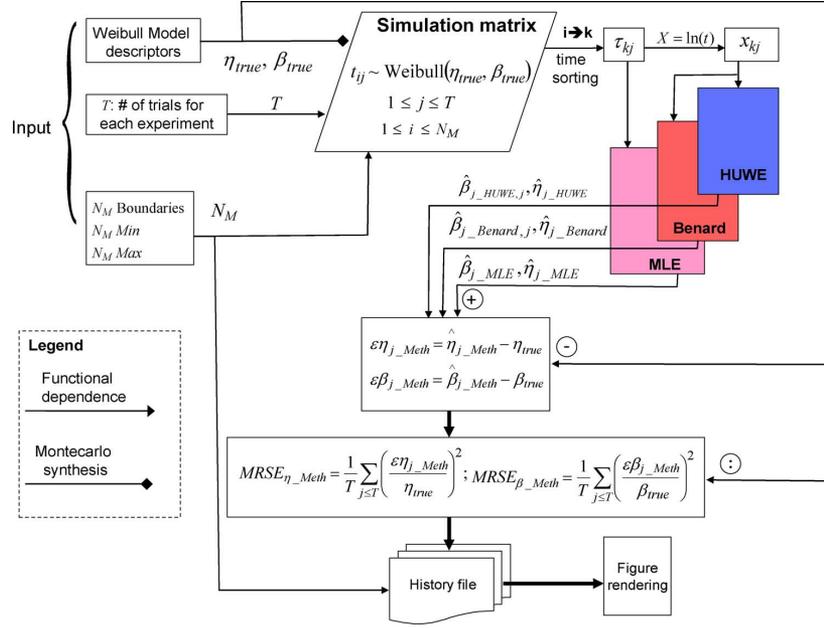


Fig. 5. Monte Carlo test procedure explanation.

as  $j$ ), and a large number of trials  $T$  are generated to accurately evaluate the features of the methods.

The methods used as references are Benard's, and MLE. More specifically, and in the context of the evaluation, we describe the methods.

- Benard's method uses (6) to obtain  $F_k$ , then determines  $Y_k$  and  $X_k$  by (5), and eventually achieves Weibull parameters by ordinary least squares minimization of the residual (7).
- MLE is solved by applying the iterative algorithm described in the classical article [4].

Once the parameter estimations  $\hat{\eta}_{j\_Meth}$ ,  $\hat{\beta}_{j\_Meth}$  offered by each competing method for the  $j$ -th trial have been obtained, the errors ( $\varepsilon\eta_{j\_Meth}$ ,  $\varepsilon\beta_{j\_Meth}$ ) of each are given by

$$\begin{aligned}\varepsilon\eta_{j\_Meth} &= \hat{\eta}_{j\_Meth} - \eta_{true} \\ \varepsilon\beta_{j\_Meth} &= \hat{\beta}_{j\_Meth} - \beta_{true}\end{aligned}\quad (27)$$

Above, the sub-index *Meth* (Method), is a dummy that stands for any particular method name (for instance, to specify the MLE method,  $\varepsilon\eta_{j\_Meth}$  becomes  $\varepsilon\eta_{j\_MLE}$ ).

Examining Fig. 5 in more detail, at the beginning of the evaluation process, the values of the Weibull target parameters ( $\eta_{true}$ ,  $\beta_{true}$ ) are settled, together with the desired range of test life sizes (number of widgets) one wants to investigate [ $N_M Max$ ,  $N_M Min$ ], and the number of trials  $T$ , the number of Monte Carlo independent repetitions of each test life. Then the program expands the  $N_M$  range in logarithm equally-spaced values, and for each one it generates a stochastic  $N_M \times T$  sized matrix  $t_{ij}$  of  $s$ -independent Weibull failure times. Once  $t_{ij}$  has been generated, its columns are ordered giving rise to another matrix  $\tau_{kj}$  such that the  $j$ -th column of  $\tau_{kj}$  will contain the rank-indexed times of failures for the  $j$ -th trial. The data in  $\tau_{kj}$  are then supplied to the three estimation methods under investigation to obtain the estimations  $\hat{\eta}_{j\_Meth}$ ,  $\hat{\beta}_{j\_Meth}$ , and the errors ( $\varepsilon\eta_{j\_Meth}$ ,  $\varepsilon\beta_{j\_Meth}$ ) mentioned above.

To quantify the ability of a method to hit the target values  $\beta_{true}$ ,  $\eta_{true}$ , we define the Mean Relative Square Error (*MRSE*) measures from the error series  $\varepsilon\eta_{j\_Meth}$ ,  $\varepsilon\beta_{j\_Meth}$  as

$$MRSE_{\eta\_Meth} = \frac{1}{T} \sum_{j \leq T} \left( \frac{\varepsilon\eta_{j\_Meth}}{\eta_{true}} \right)^2 \quad (28)$$

$$MRSE_{\beta\_Meth} = \frac{1}{T} \sum_{j \leq T} \left( \frac{\varepsilon\beta_{j\_Meth}}{\beta_{true}} \right)^2 \quad (29)$$

where the sums in (28) and (29) are over all the  $T$  Monte Carlo trials.

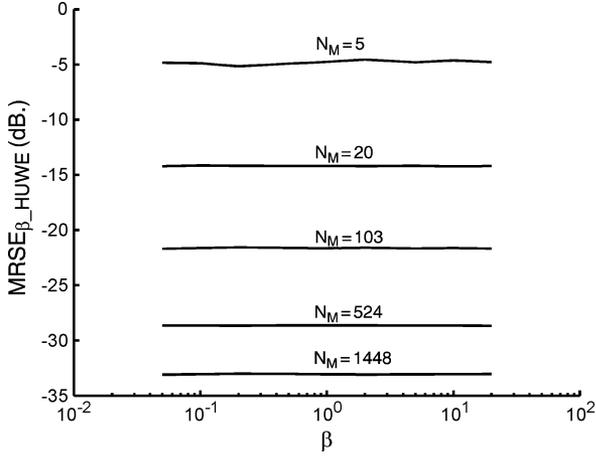
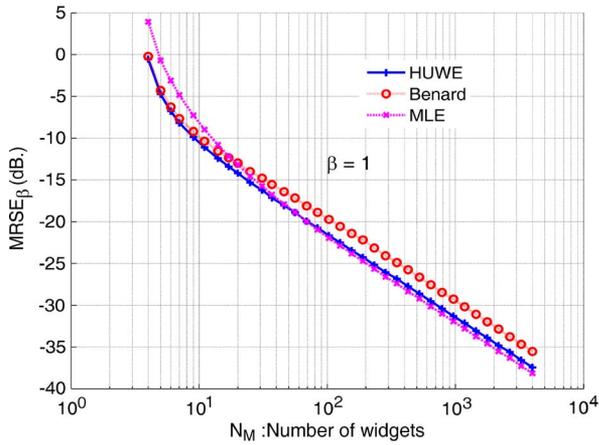
Mean Square Error (*MSE*), often used as a quality measure of estimators, can be obtained in a direct way from *MRSE* by means of

$$\begin{aligned}MSE_{\eta\_Meth} &= \eta_{true}^2 MRSE_{\eta\_Meth} \\ MSE_{\beta\_Meth} &= \beta_{true}^2 MRSE_{\beta\_Meth}\end{aligned}\quad (30)$$

*MRSE* are preferable to *MSE* when we want to compare performances over a wide range of target values of parameters  $\eta_{true}$ ,  $\beta_{true}$ , because *MRSE* are much more stable against changes of such values than *MSE*. In fact, due to scale considerations [11] on the Weibull law (1),  $MRSE_{\eta\_Meth}$  and  $MRSE_{\beta\_Meth}$  do not depend on  $\eta_{true}$  at all. On the other hand,  $MRSE_{\beta\_Meth}$  also appears to be functionally independent of  $\beta_{true}$  as it will be accounted for in the evaluation results of the next section (see Fig. 6). This double independence lets us summarize all the information about the  $\beta$  estimators in a single graph.

Both  $MRSE_{\eta}$ , and  $MRSE_{\beta}$  are dimensionless quadratic measures, so they can be expressed in a natural way in logarithmic units (31).

$$MRSE_{\eta}(dB.) = 10 \log_{10}(MRSE_{\eta})$$


 Fig. 6.  $MRSE_{\beta}$  independence from  $\beta_{true}$  value.

 Fig. 7. Compared studio of  $MRSE_{\beta}$  as a function of  $N_M$ .

$$MRSE_{\beta}(dB.) = 10 \log_{10}(MRSE_{\beta}) \quad (31)$$

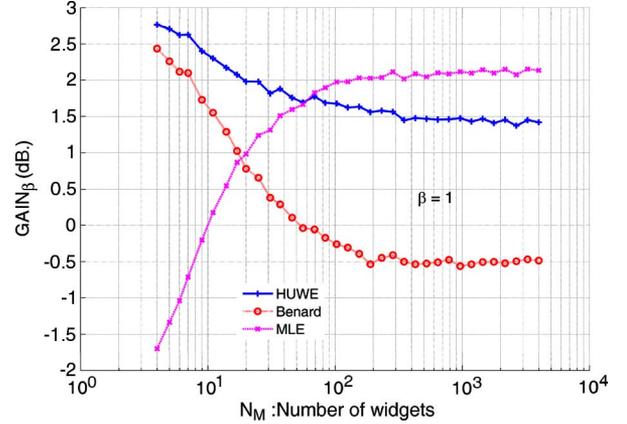
Once  $MRSE$  qualifiers have been obtained for the desired test life sizes, parameters values, and methods of interest, they are stored in files, and post-processed to render figures.

## V. RESULTS

In this section, we will show the results of applying the methodology described above to the evaluation of the three methods that are going to be compared.

To obtain  $MRSE$ , we have established the following environment.

- A number of trials  $T = 40000$  have been synthesized.
- A range of 36 log-spaced values for  $N_M$  have been selected between  $N_{MMin} = 4$ , and  $N_{MMax} = 4000$ .
- As  $\eta$  is a scale parameter, changing its value simply affects the failure times by a multiplicative constant factor [11], so it is enough to perform the evaluations for  $\eta_{true} = 1$ .
- $\beta_{true}$  values have been selected from the range  $0.05 < \beta_{true} < 20$ .


 Fig. 8.  $\beta$  accuracy gains with respect to  $MRSE_{\beta\_Analytic}$  as a function of  $N_M$ .

### A. Shape Parameter ( $\beta$ ) Estimation

We can observe in Fig. 6 that  $MRSE_{\beta}$  remains almost unchanged against changes of the  $\beta_{true}$  target values.<sup>2</sup> This is a good feature of  $MRSE_{\beta}$  as a quality measure of  $\beta$  estimations, and allows us to reduce the analysis to the case where  $\beta_{true} = 1$ . Thus, Fig. 7 shows  $MRSE_{\beta}(dB.)$  for each method, and  $\beta = 1$ .

Fig. 7 shows the quality of  $\beta$  estimators,  $MRSE_{\beta}$  as a function of  $N_M$ . Here,  $MRSE_{\beta}$  are very widespread due to their strong (functional) dependence with  $N_M$ , thus blurring the contrast between the three methods. So to get a graph that highlight better the differences, it is very useful to establish a reference  $MRSE_{\beta\_Analytic} \approx MRSE_{\beta\_Meth}$  competent to model the common dependence of the behavior of the methods with  $N_M$ . Reasonable requirements for  $MRSE_{\beta\_Analytic}$  are that

- it should follow as closely as possible the global behavior of the methods, especially in the asymptotic limits;
- it should be easily computable, without the need of Monte Carlo synthesis resources; and
- it should not contain large or difficult to manage constants.

A heuristic choice for  $MRSE_{\beta\_Analytic}$  that meets well these requirements is

$$MRSE_{\beta\_Analytic} = \frac{1}{(N_M - 3.4)}. \quad (32)$$

By writing (32) in logarithmic units, we have

$$MRSE_{\beta\_Analytic}(dB.) = -10 \log_{10}(N_M - 3.4). \quad (33)$$

By using  $MRSE_{\beta\_Analytic}(dB.)$  as a reference, we can define the accuracy gain figures  $GAIN_{\beta\_Meth}$  for each method (34). Of course, unlike  $MRSE$ , accuracy gains are defined in such a way that are larger as methods perform better.

$$GAIN_{\beta\_Meth}(dB.) = MRSE_{\beta\_Analytic}(dB.) - MRSE_{\beta\_Meth}(dB.) \quad (34)$$

<sup>2</sup>For the sake of clarity, Fig. 6 only shows  $MRSE_{\beta\_HUWE}$ , results for other methods lead to the same conclusion.

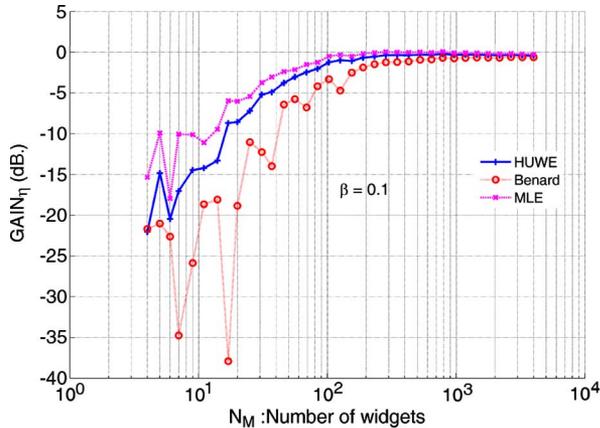


Fig. 9.  $\eta$  accuracy gains with respect to  $MRSE_{\eta-Analytic}$  as a function of  $N_M$  for  $\beta_{true} = 0.1$ .

Under these conventions,  $GAIN_{\beta}$  have been represented in Fig. 8 for  $\beta_{true} = 1$ . Results for other values of  $\beta_{true}$  follow the same behavior.

### B. Scale Parameter ( $\eta$ ) Estimation

To define accuracy gain figures for  $\eta$  estimators, a global reference behavior  $MRSE_{\eta-Analytic}$  has been also defined for them. But unlike  $MRSE_{\beta-Analytic}$ ,  $MRSE_{\eta-Analytic}$  depends on  $\beta_{true}$  in a somewhat complex way. A good heuristic fit for  $MRSE_{\eta}$  has been found to be (35).

$$MRSE_{\eta-Analytic} = \frac{Exp\left(\beta_{true}^{\frac{5}{3}} N_M^{\frac{3}{4}}\right)}{\beta_{true}^2 (N_M - 2)} \quad (35)$$

By writing (35) in logarithmic units, we have

$$MRSE_{\eta-Analytic}(dB.) = -10 \log_{10}(N_M - 2) - 20 \log_{10} \beta_{true} + \frac{10 \log_{10} e}{\beta_{true}^{\frac{5}{3}} N_M^{\frac{3}{4}}}. \quad (36)$$

By using  $MRSE_{\eta-Analytic}(dB.)$  as reference, we can define gain figures for each method.

$$GAIN_{\eta-Method}(dB.) = MRSE_{\eta-Analytic}(dB.) - MRSE_{\eta-Method}(dB.) \quad (37)$$

Under these conventions, in Figs. 9–11, accuracy gains with respect to the reference (36) have been represented for  $\beta_{true} = 0.1$ ,  $\beta_{true} = 1$ , and  $\beta_{true} = 20$ , respectively.

### C. Result Summary

With respect to  $\beta$  estimation, we can conclude from Fig. 8 that

- HUWE performs the best for low values of the number of widgets ( $N_M < 65$ );
- for very low values of the number of widgets  $N_M < 20$ , MLE performs the worst; and
- as  $N_M \rightarrow \infty$ , HUWE tends to perform better (2.0 dB.) than Benard's, and slightly worse (0.6 dB.) than MLE.

With respect to  $\eta$  estimation, we can conclude from Figs. 9–11 that

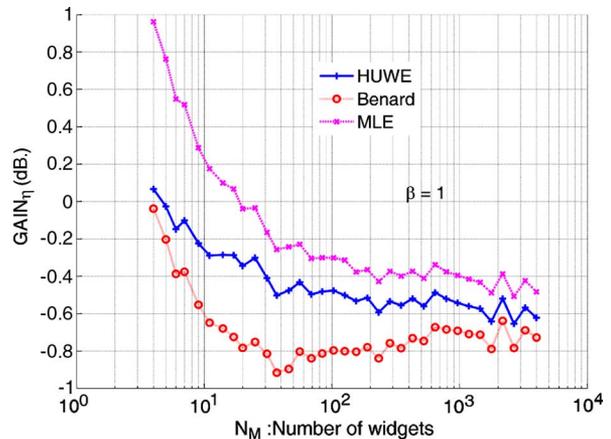


Fig. 10.  $\eta$  accuracy gains with respect to  $MRSE_{\eta-Analytic}$  as a function of  $N_M$  for  $\beta_{true} = 1$ .

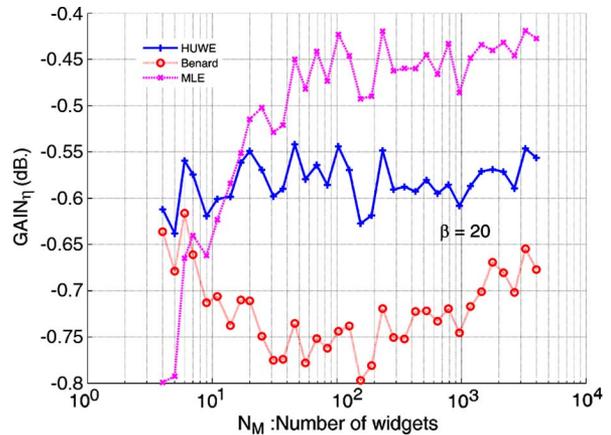


Fig. 11.  $\eta$  accuracy gains with respect to  $MRSE_{\eta-Analytic}$  as a function of  $N_M$  for  $\beta_{true} = 20$ .

- the accuracy levels of the three methods are quite similar, disregarding the fact that when  $\beta_{true}$ , and  $N_M$  are both small, Benard's method does much worse than the others; and
- when  $N_M \rightarrow \infty$ , regardless of the  $\beta_{true}$  value, a small difference of about 0.15dB. remains, where the best is MLE, then HUWE, and then Benard's method.

Note that the advantage of HUWE against Benard's method can have economic significance as it offers an important reduction in the required number of widgets  $-\Delta N_M/N_M$ . For example, regarding  $\beta$  estimation, this reduction can be obtained from Fig. 7, or can be estimated roughly for large values of  $N_M$  (38) by taking into account (32).

$$\lim_{N_M \rightarrow \infty} \frac{-\Delta N_M}{N_M} \approx 1 - 10^{-\frac{2Db}{10}} = 37\% \quad (38)$$

Aside from the accuracy results, MLE has been found to be much more time consuming than other methods, especially for high values of  $N_M$ , even though HUWE pre-computation (see Section II-E) has not been used in the evaluation. Also, when the  $\beta_{true}$  value is far from  $\beta_{true} = 1$ , the MLE computational demand reaches the limits of the numerical representation (64 bit double-precision IEEE 754 [30]) used in the simulation, in

such a way that numerical crashes occurred for  $\beta_{true} < .02$ ; and we experienced a lack of convergence, and infinite loops for  $\beta_{true} > 10$ .

## VI. CONCLUSION

We presented HUWE, an algorithm for estimating Weibull parameters from life tests that offers a good trade-off alternative to the traditional Weibull estimators.

HUWE is based on an approximation, the assumption of  $s$ -normality for the data unreliability related variables  $Y_k$ . This approximation allows us to express analytically the first two moments of  $Y_k$  in a nice symmetrical form. Having determined the means and variances of  $Y_k$ , the Weibull parameters  $\eta$  and  $\beta$  are estimated by fitting the data to a linear model using WLS.

HUWE has been Monte Carlo evaluated against Benard's approximation, and MLE. The evaluation results are as follows.

- HUWE is more robust, and computationally more efficient than MLE. Though in some cases the accuracy of HUWE is slightly lower than that of MLE, the global performance of HUWE is better when one considers the whole range of working cases.
- HUWE significantly improves the performance of Benard's approximation. For large enough numbers of widgets, say  $N_M > \sim 60$ , HUWE is able to get similar confidence levels for the estimators of Weibull parameters than Benard's, but using 37% less widgets.

In short, HUWE is a robust, accurate, flexible algorithm that has global advantages against traditional Weibull estimators, constituting an appealing choice to be embedded in automated test life systems.

## REFERENCES

- [1] D. N. P. Murthy, M. Xie, and R. Jiang, *Weibull Models*. Hoboken, NJ: Ed. John Wiley & Sons, 2004.
- [2] R. Soland, "Bayesian analysis of the Weibull process with unknown scale and shape parameters," *IEEE Trans. Reliability*, vol. R-18, no. 4, pp. 181–184, Nov. 1969.
- [3] M. P. Kaminskiy and V. V. Krivtsov, "A simple procedure for bayesian estimation of the Weibull distribution," *IEEE Trans. Reliability*, vol. 54, no. 4, pp. 612–616, Dec. 2005.
- [4] A. C. Cohen, "Maximum likelihood estimation in the weibull distribution based on complete and on censored samples," *Technometrics*, vol. 7, no. 4, pp. 579–588, Nov. 1965.
- [5] J. Jacquelin, "Generalization of the method of maximum likelihood," *IEEE Trans. Electrical. Insulator*, vol. 28, no. 1, pp. 65–72, Feb. 1993.
- [6] Z. Tan, "A new approach to MLE of Weibull distribution with interval data," *Reliability Engineering and System Safety*, vol. 94, no. 2, pp. 394–403, 2009.
- [7] P. Jiang, J.-H. Lim, M. J. Zuo, and B. Guo, "Reliability estimation in a weibull lifetime distribution with zero-failure field data," *Quality and Reliability Engineering International*, vol. 26, no. 7, pp. 691–701, Nov. 2010.
- [8] J. C. Zhou and D. F. Wu, "Estimation Of Weibull Parameters With Linear Regression Method," *Materials Science and Technology*, vol. 26, no. 8, pp. 1008–1011, Aug. 2010.
- [9] A. Saghafia, A.R. Mirhabibib, and G.H. Yaria, "Improved linear regression method for estimating Weibull parameters," *Theoretical and Applied Fracture Mechanics*, vol. 52, no. 3, pp. 180–182, Dec. 2009.
- [10] S. Amari, "Weibull parameter estimation methods.," *Quaterflash Relx Software Corporation*, vol. 17, pp. 4–5, 2005.
- [11] U. Genschel and W. Q. Meeker, "A comparison of maximum likelihood and median rank regression for weibull estimation," *Quality Engineering*, vol. 22, no. 4, pp. 236–255, 2010.
- [12] R. Ross, "Comparison of methods to reduce the Maximum Likelihood bias on Weibull parameters," in *IEEE 5th International Conference on Conduction and Breakdown in Solid Dielectrics*, 1995, pp. 204–208.
- [13] A. J. Watkins, "On the maximum likelihood estimation for the two parameters Weibull distribution," *Microelectronic Reliability*, vol. 36, no. 5, pp. 595–603, May 1996.
- [14] Z. Yang and M. Xie, "Efficient estimation of the Weibull Shape Parameter based on modified profile likelihood," *Journal of Statistical Computation and Simulation*, vol. 73, no. 2, pp. 115–123, 2003.
- [15] M. M. Bugaighis, "A note on convergence problems in numerical techniques for accelerated life-testing analysis," *IEEE Trans. Reliability*, vol. 37, no. 3, pp. 348–349, Aug. 1988.
- [16] W. Wang and D. B. Kececioglu, "Fitting the weibull log-linear model to accelerated life-test data," *IEEE Trans. Reliability*, vol. 49, no. 2, pp. 217–223, Jun. 2000.
- [17] D. N. Prabhakar Murthy, M. Bulmerc, and J. A. Eccleston, "Weibull model selection for reliability modelling," *Reliability Engineering and System Safety*, vol. 86, no. 3, pp. 257–267, Dec. 2004.
- [18] L. F. Zhang, M. Xie, and L. C. Tang, "A study of two estimation approaches for parameters of Weibull distribution based on WPP," *Reliability Engineering and System Safety*, vol. 92, no. 8, pp. 360–368, Mar. 2007.
- [19] L. F. Zhang, M. Xie, and L. C. Tang, "Bias correction for the least squares estimator of Weibull shape parameter with complete and censored data," *Reliability Engineering and System Safety*, vol. 91, no. 8, pp. 930–939, Aug. 2006.
- [20] Y. Hong, W. Q. Meeker, and L. A. Escobar, "The Relationship Between Confidence Intervals for Failure Probabilities and Life Time Quantiles," *IEEE Trans. Reliability*, vol. 57, no. 2, pp. 260–266, Jun. 2008.
- [21] A. Benard and E. C. Bos-Levenbach, "Het uitzetten van waarnemingen op waarschijnlijkheids-papier (The Plotting of Observations on Probability Paper)," *Statistica Neerlandica*, vol. 7, pp. 163–173, 1953.
- [22] J. O. Rawlings, S. G. Pantula, and D. A. Dickey, *Applied Regression Analysis: A Research Tool*. New York: Springer Verlag, 1998.
- [23] L. F. Zhang, M. Xie, and L. C. Tang, *Recent Advances in Reliability and Quality in Design*. London: Springer Verlag, 2008.
- [24] B. Bergman, "Estimation of Weibull parameters using a weight function," *Journal of Materials Science Letters*, vol. 5, no. 6, pp. 611–614, 1986.
- [25] B. Faucher and W. R. Tyson, "On the determination of the Weibull parameters," *Journal of Materials Science Letters*, vol. 7, no. 11, pp. 1199–1203, 1998.
- [26] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications, 1964.
- [27] C. R. Mischke, *A Distribution-Independent Plotting Rule for Ordered Failures*. New York: American Society of Mechanical Engineers, ASME publication 79-DET-112, 1979.
- [28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, N.J.: John Wiley & Sons, 2006.
- [29] A. Fernández and M. Vázquez, "A generalized regression methodology for bivariate heteroscedastic data," *Communications in Statistics - Theory and Methods*, vol. 40, no. 4, pp. 598–662, 2011.
- [30] *IEEE Standard for Floating-Point Arithmetic IEEE Std 754-2008 (Revision of IEEE Std 754-1985) Microprocessor Standards Committee of the IEEE Computer Society*, 2008.

**Antonio Fernández** was born in Cordoba (Spain) in 1964. He received a Master's in Electronic Engineering in 1989 from UPM (Universidad Politécnica de Madrid), and a Bachelor's in Physics in 1992. From 1990, he has been with the R&D department of Telefónica. He is currently a Professor at UPM. Prof. Fernández is an author of many surveys and recommendations for Telefónica, an international paper, and holds 3 patents.

**Manuel Vázquez** was born in Madrid (Spain) in 1962. He received a Bachelor's in Physics in 1986, and a Ph.D. in 1991 from UCM (Universidad Complutense de Madrid). He was with Alcatel for 10 years. He is currently a Professor at UPM (Universidad Politécnica de Madrid). Prof. Vázquez is the author of 15 papers, and holds 5 international patents.