Bayesian Reliability Modelling of a Tidal Turbine Pitch System

Fraser J. Ewing[#1], Philipp R. Thies[*2], Iraklis Lazakis[#3], Jonathan Shek[#4]

[#]*Industrial Doctoral Centre for Offshore Renewable Energy, University of Edinburgh*
Sanderson Building, The Kings Buildings, Edinburgh, UK

[1]f.ewing@ed.ac.uk

[3]i.lazakis@strath.ac.uk

[4]j.shek@ed.ac.uk

[*]*University of Exeter*
College of Engineering, Mathematics and Physical Sciences, Penryn Campus, Penryn, UK

[2]p.r.thies@exeter.ac.uk

***Abstract*— To date there has been no published reliability data for Tidal Stream Turbines (TST's). Thus, reliability assessments of failure critical components are highly uncertain. To reduce the Levelised cost of energy (LCoE) of TST technology it is critical to be able to perform accurate reliability assessments of devices at the design stage.**

Based on experience from the wind industry, the power take off (PTO) is a failure critical assembly representing a high proportion of total turbine failures and downtimes. Many studies have found that the pitch system (PS) contributes to most turbine failures [1], [2]. Bayesian reliability methods are of interest in industries where data is scarce or commercially sensitive as they allow for the use of surrogate data sources along with domain knowledge; they also allow for this knowledge to be updated as new data becomes available [3].

This research develops a Bayesian reliability model of a Horizontal axis TST PS using state of the art surrogate failure data and domain knowledge. The paper discusses the rationale behind Bayesian modelling and provides a framework for TST developers, researchers, consultants and the like to use their own device failure data when it becomes available to make robust PS reliability assessments with quantified uncertainty levels. The components of the PS focussed on in this research are the Dynamic Seal, Roller Bearing and Electric Motor. These are seen to be failure critical areas. Empirical Physics of Failure (PoF) equations are used to determine individual failure rates for these parts and then Monte Carlo methods are used to calculate the combined failure rates. This combined part failure rate distribution is then updated using representative 'real' PS failure data (from the wind industry) to highlight the Bayesian updating process. This paper represents a step change in current reliability assessment methods in the Tidal industry by presenting a method to attribute specified levels of uncertainty to the underlying parts (e.g. design parameters and correction factors) of the failure rate calculations.

***Keywords*— Bayesian Reliability, Tidal Turbines, Pitch System, Failure Rate, Uncertainty**

I. INTRODUCTION

There exist very few studies focused on the development of probabilistic component failure models for the TST industry because of the severe lack of applicable data.

In the tidal industry, there have been several studies that have applied Bayesian methods to reliability assessment of TST devices. In [5] Val et al propose a notional TST device and analyse the reliability of a seal and rolling bearing using Bayesian methods. Handbook style physics of failure equations [6], with certain parameters as random variables, are used to determine prior failure rates of these components based on the Central Limit Theorem. The likelihood function is used to define the likelihood of observing a pre-determined number of failures. Ultimately the posterior distribution is defined to highlight the effect of a hypothetical failure on a components failure rate.

In [7] Thies et al develop a Bayesian method to assess the uncertainty around the failure rate of a notional wave energy converter (WEC) umbilical cable. OREDA data is used to establish the prior. A hypothetical likelihood is derived for two hypothetical cases (unknown failure mode and fatigue) based on the confidence bounds of the prior distribution. The Bayesian framework is used to determine the effect on the failure rate for each of these hypothetical test cases.

This research seeks to build on the approaches taken in [4], [5] and further them by representing all parameters as random variables, allowing for quantified uncertainty levels to be assigned to the random variables and creating a complete Bayesian updating framework which links different components and also allows for the future application of test data (when it becomes available).

The components of the PS focussed on in this research are the bearing, seal and electric motor. These are some of the most failure critical components. Where possible, design parameters have been used as per the designs of a leading horizontal axis TST developer.

II. BAYESIAN RELIABILITY THEORY

Reliability is a probabilistic measure of the ability of a component/system to perform its stated function for a required amount of time [3]. Traditional frequentist approaches to probability involve taking samples of data and determining point values of the unknown distribution parameters, typically via Maximum Likelihood Estimation (MLE) or Ordinary least squares regression (OLS). If the sample size is small, as is the case in pre-commercial industries such as Tidal Energy, then it is often difficult to obtain accurate inference on the distribution parameters via frequentist methods. The Bayesian approach can help combat this issue via its differing philosophical view of probability. Rather than finding the 'true' value of the distribution parameters based on repeated sampling, the Bayesian approach uses prior knowledge of the parameters and updates this knowledge using the sample. Thus, the parameters of the reliability model are themselves represented as random variables with an inherent measure, or 'degree of belief', in their uncertainty.

Bayes Theorem states that:

$$P(\lambda|t) = \frac{P(t|\lambda)P(\lambda)}{\int P(t|\lambda)P(\lambda)} \quad (1)$$

Where $P(\lambda|\theta)$ is the posterior density of the model parameter, $P(t|\lambda)$ is the likelihood function of the data and $P(\lambda)$ is the prior distribution of the model parameter.

These are the 3 key parts of the Bayesian reliability model. For non-trivial distributions, the integral has no closed form solution so a Markov Chain Monte Carlo (MCMC) algorithm such as the Metropolis-Hastings (MH) must be employed [3]. The model parameter (λ) represents the failure rate of the PS. The Bayesian reliability model goal is to utilise prior information about PS failure rates, gained from the PoF equations found in [6], and then apply representative PS failure data (in the form of surrogate data from the wind industry) to see the effect on the failure rate. This allows for a more robust understanding of the PS failure rate and its uncertainty than if the typical deterministic approach were taken. Also, the framework developed here allows for specific uncertainties to be assigned to the correction factors and design parameters.

A. *The Prior Distribution*

The prior distribution of the PS failure rate is constructed using empirical physics of failure (PoF) equations taken from [6]. These equations were developed to relate the failure rates (λ) of various mechanical components with their physical characteristics, operating behaviours and material properties. An example of the PoF equation for a dynamic seal is:

$$\lambda = \lambda_B C_q C_h C_f C_v C_t C_n C_{pv} \quad (2)$$

where each of the correction factors (Cv, Cf etc.) constitutes a formula that represents a failure critical characteristic of the component (e.g. seal leakage, seal size, fluid viscosity etc.) and λ*se,b* represents the 'base' failure rate taken from relevant literature.

These formulas represent the outcome of thousands of hours of accelerated life tests. The formulas are empirically derived and relate relevant physical and design parameters to failure rates. The design values used in this paper to populate the correction factor formulae are based on a leading tidal developers design and thus are highly confidential and cannot be published.

B. *The Generalised Lognormal Distribution*

For all the correction factors that are represented as random variables (some are deterministic, hence are point values), a generalised lognormal distribution (GLN) with 3 parameters is used. The generalised probability density function (PDF) of a lognormally distributed random variable (λ) is:

$$P(\lambda) = f(\lambda|\theta,\sigma,m) = \frac{1}{(\lambda-\theta)\sigma\sqrt{2\pi}} e^{-\frac{(\ln(\frac{\lambda-\theta}{m})^2)}{2\sigma^2}} \quad (3)$$

where σ is the shape parameter, m is the scale parameter and θ is the location parameter (which is equal to zero).

The lognormal distribution is commonly used in the reliability engineering field to model data arising from accelerated life tests [7]. In its generalised form the distribution is more flexible and can allow for heavier and lighter tails. The distribution occurs only on the positive axis which is important as negative correction factor values are undesirable.

The two non-zero parameters in the GLN are the shape (σ) and scale (m) scale. The scale parameter is used to define the centre point of the distribution and is set equal to the calculated value of a given correction factor. For example, when calculating the *Cv* factor using the deterministic PoF calculations in [6], the output value is used as the m parameter of the GLN. To completely specify the distribution, the shape parameter (σ) is also needed. With location = 0 and m = 1, the relationship between the standard deviation (*std*) and σ is:

$$std = \sqrt{e^{\sigma^2}(e^{\sigma^2}-1)} \quad (4)$$

Therefore, it is possible to specify the standard deviation (or dispersion) of a lognormally distributed random variable in terms of its shape parameter. For different scale parameter values this relationship changes as can be seen in Fig. 1.

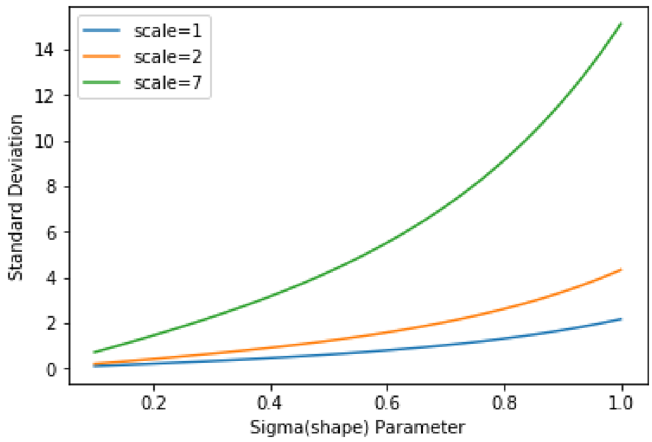

Fig. 1 Relationship between standard deviation and shape and scale parameters (with location = 0)

The shape parameter can then be chosen to reflect the amount of standard deviation required. The uncertainty in the scale parameter is attributed to the size of the standard deviation value. For example, for a scale value of 7, we may wish to have a standard deviation of 3. The corresponding shape parameter can then be selected from the graph.

Three levels of uncertainty are used in this work as shown in Table 1. A high uncertainty in a correction factor is reflected by a standard deviation value that is 100% of the correction factor value. A low uncertainty is reflected by a standard deviation value that is 10% and a medium uncertainty is reflected by a standard deviation value that is 50% of the calculated correction factor value.

TABLE I
UNCERTAINTY LEVELS FOR LOGNORMAL DISTRIBUTIONS OF CORRECTION FACTORS

Standard Deviation	Uncertainty Level
10% of *m*	Low
50% of *m*	Medium
100% of *m*	High

The use of the GLN with its shape and scale parameters means that it is possible to calculate the expected value of a correction factor, set this equal to the mean of the GLN and then define a standard deviation about the mean which reflects the uncertainty in the calculated mean value.

C. The Likelihood Distribution

The likelihood function should represent relevant failure data taken from operational experience or perhaps field tests. For TST there is unfortunately no failure data that is commercially available, hence surrogate data from the wind industry is used in this research to represent tidal turbine data. The data consists of over 1000 real operational PS failures experienced by turbines rated between 1-2MW (which are of a similar rating to early commercial TST).

Several thousand PS failure events are considered in this research. An exponential model is used as the likelihood function, representing exponentially distributed failure occurrences with a constant failure rate:

$$P(t|\lambda) = f(t|\lambda) = e^{-\lambda t} \tag{5}$$

III. METHOD

The approach taken in this paper is summarised in the following stages.

A. Correction Factor Calculations

Based on the empirical Physics of Failure (PoF) equations in [6] (and turbine specific design parameters), correction factor (CF) values are calculated and represented as GLN random variables with defined mean and standard deviation. The mean is equal to the calculated value and the standard deviation is assigned based on engineering judgement about the uncertainty in the empirical PoF equation (and the design parameter data). Shape and scale parameters for the GLN are calculated based on the assigned mean and standard deviation.

B. Monte Carlo Sampling from Correction Factors

To combine the correction factor GLNs for each part (e.g. bearing, seal and motor), a random sample is selected from each distribution and the samples are multiplied per the relevant reliability model equation e.g. Equation (2). 100,000 samples are generated (to allow for model convergence without massive computational cost).

C. Monte Carlo Sampling from Failure Rate Distributions

The failure rate distributions for each part are then combined to make up the PS failure rate distribution. A random sample is selected from each distribution and the samples are summed as per the relationship between independent components with assumed exponential failure times:

$$R_s = \prod_{i=1}^{n} R_i = e^{-(\lambda_1 + \lambda_2 + \lambda_3 \ldots \lambda_n)t} \tag{6}$$

where *Rs* is the combined reliability and *Ri* is the reliability of each part. 100,000 samples are again generated and the resulting histogram is fit with a GLN. This GLN (with defined parameters) forms the prior failure rate distribution of the complete PS.

D. Construct Likelihood distribution

From representative operational PS failure data, the likelihood distribution is then constructed as per Equation (5).

E. Compute prior distribution of failure rate

The posterior distribution of the PS failure rate distribution is then calculated using the representative operational PS failure data via Markov Chain Monte Carlo methods in a Bayesian updating framework.

IV. RESULTS

For each of the parts of the pitch system, the results of the correction factor calculations and their fitted distributions are now presented.

Roller Bearing

The reliability model for the Bearing is:

$$\lambda = \lambda_B C_y C_v C_{cw} C_t \tag{7}$$

The factors listed in Table II are calculated from the design parameters in Table III using the equations in [6]. The lubrication viscosity parameter (*Cv*) is a function of the viscosity ratio *k* (lubrication condition of the bearing). This is calculated from the equations in [6] and is mainly a function of the diameter and motion of the bearing (the maximum rpm of the bearing is assumed to be akin to the maximum blade pitch rate of 60rpm). A low uncertainty level (UL) is assigned to this *Cv* value; thus, the standard deviation value of 10% of the scale parameter (m=0.15) results in a shape parameter of 0.1. This defines the GLN for the lubrication viscosity factor.

TABLE II
BEARING CORRECTION FACTOR AND BASE FAILURE RATE DISTRIBUTION PARAMETERS AND UNCERTAINTY LEVELS (UL)

Parameter	ID	UL	σ	m
Lubrication viscosity	C_v	10%	0.1	0.15
Applied Load	C_y	10%	0.09	0.028
Lubricant contamination	C_{cw}	50%	0.51	1.08
Operating temperature	C_t	-	-	1
Base failure rate (1/million hrs)	λ_B	10%	0.1	0.06

The applied load factor (*Cy*) is based on the dynamic load rating and equivalent dynamic load, taken from the manufacturers specifications based on a bearing of the dimensions listed in Table III [8]. The lubricant contamination factor (*Ccw*) is calculated as 1.08 based on % water content of lubricant value of 0.04 [8]. The uncertainty level (UL) for this figure is medium (50%) as it is unknown how subsea operation will impact the water content of the lubricant. The operating temperature factor is 1 as the temperature inside the nacelle will not reach levels higher than 40 degrees Celsius. The base failure rate is taken from the manufacturers catalogue [8] and the UL is assumed to be low.

TABLE III
BEARING DESIGN PARAMETERS

Name	*Description*	*Value*
d	Outer bearing Bore Diameter (mm)	200
w	Blade Pitch Rate (rpm)	60
D	Outer bearing External diameter (mm)	250
alpha	% water content of lubricant	0.04
P	Equivalent Dynamic Load kN	500
C	Dynamic Load Rating kN	1470

The Monte Carlo simulation is then carried out which randomly samples from each of the correction factor GLNs and combines into the bearing failure rate distribution shown in Fig. 2.

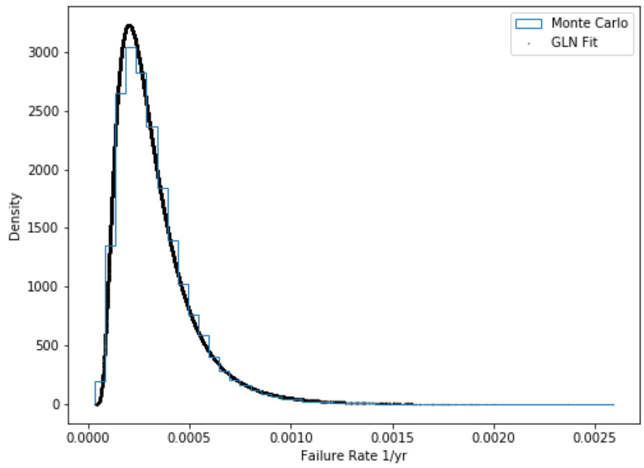

Fig. 2 Bearing failure rate distribution

The fit statistics for the Bearing failure rate GLN are shown in Table IV. The coefficient of variation (COV), a standardised measure of dispersion, highlights the uncertainty in the failure rate distribution.

TABLE IV
BEARING DESIGN PARAMETERS

Mean	***Standard Deviation***	***Coefficient of Variation***
0.000324	0.000194	60%

Dynamic Seal

The reliability model for the Seal is as follows:

$$\lambda = \lambda_B C_q C_h C_f C_v C_t C_n C_{pv} \quad (8)$$

The correction factors and base failure rate for the dynamic seal are listed in Table V (along with the parameters and uncertainty levels for their respective distributions).

TABLE V
DYNAMIC SEAL CORRECTION FACTORS AND BASE FAILURE RATE GLN PARAMETERS AND UNCERTAINTY LEVELS (UL)

Parameter	ID	UL	σ	m
Allowable leakage	C_q	-	-	4.2
Contact Pressure	C_h	10%	0.1	7.28
Surface Roughness	C_f	10%	0.1	1.37
Fluid Viscosity	C_v	-	-	6.31
Seal face temperature	C_t	-	-	0.21
Contaminants	C_n	50%	0.43	4
PV limit	C_{pv}	50%	0.45	0.058
Base failure rate	λ_B	10%	0.1	0.021

No leakage into the PS is allowed therefore the allowable leakage factor is at its maximum value (equating to zero leakage). The contact pressure factor is calculated using a Zero Wear Model (as done in [5]) and uses the RMS of the surface profile which is assumed to be 25 micro inches. The fluid viscosity factor is taken directly from [6]. The seal face temperature factor is also taken directly from [6] because the operating temperature is greater than 4 degrees Celsius. The contaminants factor is taken from [5] and has a medium uncertainty because of a lack of visibility on how the empirical relationship was derived. The PV factor is taken from [5] and has medium uncertainty. The base failure rate is taken from [6].

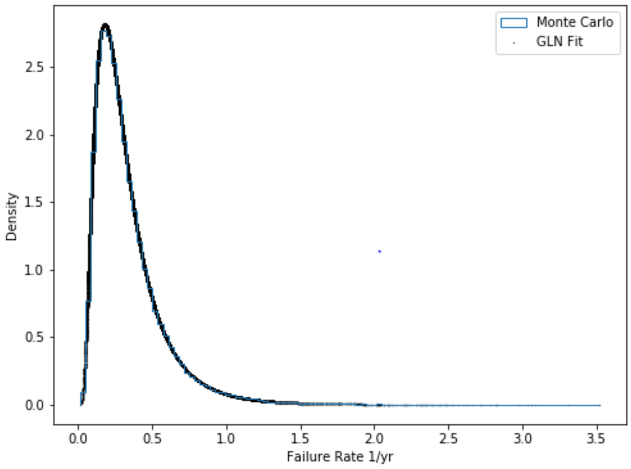

Fig. 3 Dynamic seal failure rate distribution

The fit statistics for the Dynamic Seal GLN can be found in Table VI. The standard deviation of the fit is high which reflects the medium uncertainty that was assigned to some of the underlying correction factors (*Cn* and *Cpv*). The large uncertainty here represents epistemic uncertainty (systematic uncertainty) which can be reduced given a better knowledge of the design conditions. It is likely that a higher fidelity dynamic seal reliability model (using manufacturers test data) would result in lower uncertainty bounds on the correction factors and hence a lower standard deviation of the resulting GLN fit.

TABLE VI
FIT STATISTICS FOR DYNAMIC SEAL GLN

Mean	**Standard Deviation**	**Coefficient of Variation**
0.334	0.242	72%

Electric Motor

The reliability model for the AC single phase Electric Motor is:

$$\lambda_M = (\lambda_{MB} C_{sf}) + \lambda_{WI} + \lambda_{BS} + \lambda_{ST} + \lambda_{AS} \tag{9}$$

This reliability model differs slightly from the other two components. Rather than empirical formulas relating physical parameters to failure rates, this model utilises historical failure rate data of sub components. The values for the failure rates are taken from [6] and [9]. The motor load service factor is based on a medium impact load.

TABLE VII
ELECTRIC MOTOR CORRECTION FACTORS AND BASE FAILURE RATE GLN PARAMETERS AND UNCERTAINTY LEVELS (UL)

Parameter	*ID*	UL	σ	µ
Motor base failure rate	λ_{MB}	10%	0.1	0.06
Motor load service factor	C_{sf}		-	2
Motor windings failure rate	λ_{WI}	10%	0.12	0.175
Brushes failure rate	λ_{BS}	10%	0.1	0.028
Stator housing failure rate	λ_{ST}	10%	0.1	0.001
Armature shaft failure rate	λ_{AS}	10%	0.1	0.001

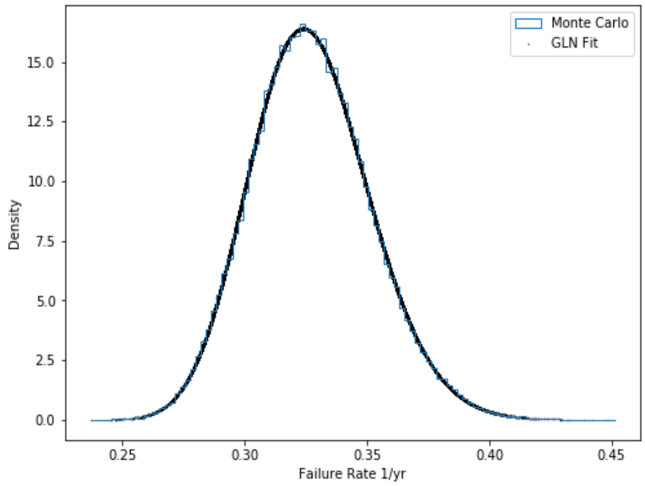

Fig. 4 Electric motor failure rate distribution

The failure distribution for the Electric Motor can be seen in Fig. 4. The fit statistics are in Table VIII. The coefficient of variation for the Electric Motor GLN is much lower than for the bearing and dynamic seal. This highlights a much lower level of uncertainty. The underlying factors in the reliability model of the electric motor are all failure rates based on historical data rather than empirical equations. It is often cited that the failures of electrical components are much better understood than mechanical ones and this is reflected here by the low COV [6].

TABLE VIII
FIT STATISTICS FOR ELECTRIC MOTOR GLN

Mean	**Standard Deviation**	**Coefficient of Variation**
0.327	0.024	7%

V. BAYESIAN UPDATING

Having constructed the prior failure rate distribution based on PoF equations and design parameters, it is necessary to utilise the representative operational failure data and complete the Bayesian Updating framework.

The prior distribution of the PS failure rate is updated with the operational PS failure data to determine the posterior distribution of the PS failure rate. This updating procedure provides insight into the uncertainty surrounding the failure rate parameter.

The PS failure rate prior distribution can be seen in Fig. 5. This is the result of the Monte Carlo sampling of the part failure rate distributions.

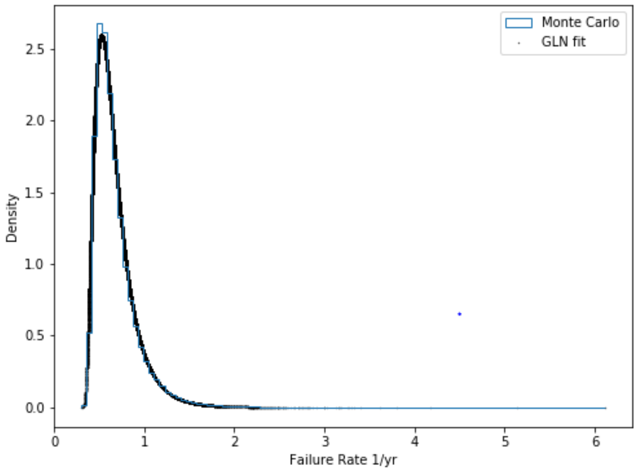

Fig. 5 Prior PS failure rate distribution

The mean value for the PS failure rate distribution is 0.669 failures per year, with a standard deviation of 0.229 failures. The COV is 34% and the tail of the fitted GLN is quite large, indicating that there is large uncertainty in the mean value. It is however important to note that this uncertainty is now quantified rather than being undefined as per the traditional approaches in the literature.

TABLE IX
FIT STATISTICS FOR ELECTRIC MOTOR GLN

Mean	**Standard Deviation**	**Coefficient of Variation**
0.669	0.229	34%

The representative PS failure data (taken from operational wind turbines from [10]) can be seen in Fig. 6. The failures are exponentially distributed with rate parameter = 1.88 failures per year. This represents the likelihood function for the Bayesian update

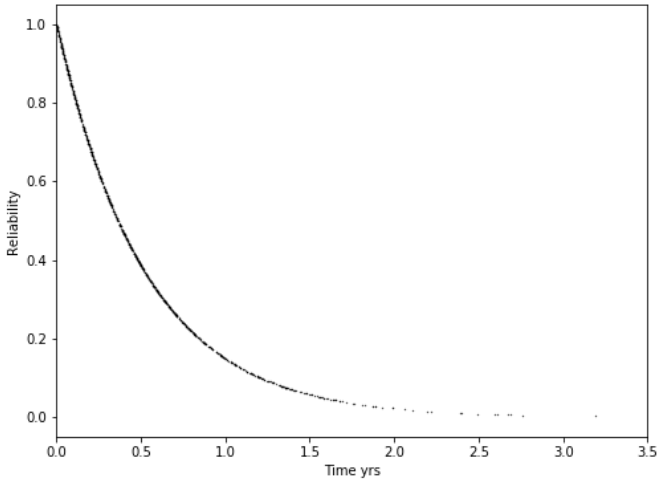

Fig. 6 Reliability function for exponentially distributed wind turbine PS failures

Once the Bayesian update of the prior distribution has taken place, the posterior distribution is produced and can be seen in Fig. 7

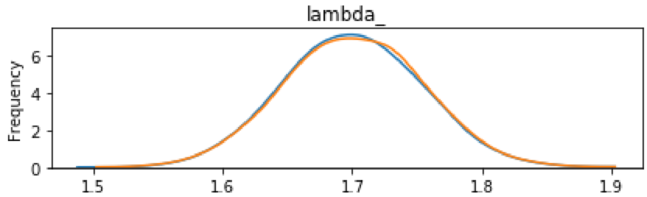

Fig. 7 Posterior distribution of failure rate

The two traces in Fig. 7 represent two MCMC sampling chains that start from slightly different points and try to converge. The *Rhat* statistic is a measure of this convergence. A value close to 1 signifies that the solution has indeed converged upon the ‘true’ value of the parameter. Table X shows that the *Rhat* statistic is very close to 1

TABLE X
MEAN AND RHAT STATISTIC FOR POSTERIOR DISTRIBUTION

Mean	**Rhat**
1.70	1.000075

The Bayesian equivalent of the frequentist confidence interval is the Highest Posterior Density (HPD). The 95% HPD is shown in Fig. 8. It highlights the range within which a 95% certainty can be assigned. The mean value is 1.7 and this has the highest density. 1.596 and 1.803 represent the lower and upper bounds of the HPD.

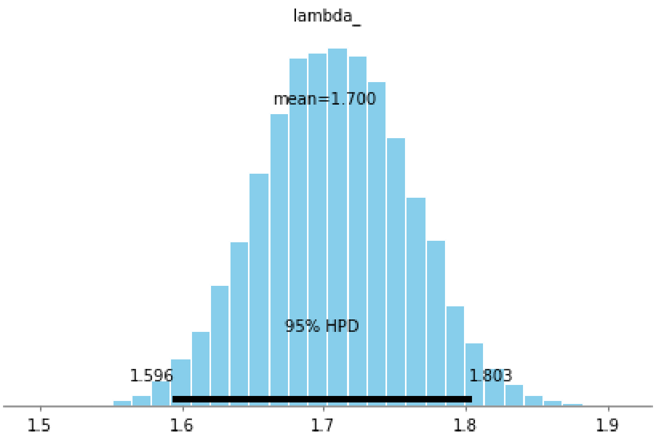

Fig. 8 95% Highest Posterior Density (HPD) of failure rate

The results of the Bayesian update have defined a posterior distribution for the failure rate parameter λ. Thus, the uncertainty surrounding the PS failure rate is completely defined.

The effect of the updating procedure increases the failure rate mean value from a prior 0.669 failures per year to 1.7 failures per year. The representative test data has a mean value of 1.88 failures per year; the prior information constructed from the physics of failure equations has the effect of reducing this value. This is an important observation for tidal developers. Basing reliability projections on limited test data (represented by wind data in this study) does not portray the full picture. Although the numbers are only meant as a guide, it can be seen how incorporating prior knowledge about the failure rate in the form of a prior distribution can result in a parameter value that is very different to one calculated solely on limited test data. Also, the level of confidence in the failure rate parameter can now clearly be seen. A large HPD can indicate that there is a high uncertainty in the underlying correction factors and design parameters. A small HPD shows that there is a high confidence in the calculated failure rate parameter mean value.

VI. CONCLUSIONS

This work has developed a Bayesian Updating framework for Tidal Turbine Pitch Systems based on empirical Physics of Failure equations and operational failure data from wind turbines. The parts of the pitch system that were focused on were the bearing, seal and electric motor as these are seen to be failure critical.

Failure rate distributions for each part were calculated using empirical correction factors and base failure rates. Each correction factor has been represented using the Generalised Lognormal distribution allowing for a mean value and standard deviation to be specified. The uncertainty in each parameter has been quantified and is based on engineering judgement of the robustness of the empirical equations and turbine design parameters. This representation of the uncertainty in the correction factors and design parameters represents a step change from the current state of the art in tidal turbine reliability assessment. Being able to quantify levels of uncertainty in each parameter is important as it allows developers to realise which parts of the system under analysis are potential weak links. The effects of these uncertain parameters are then visible as they propagate through into the final Bayesian failure rate update calculation.

The Bayesian framework developed in this work can be adopted by Tidal Turbine developers to utilise their own component failure data as it becomes available.

Further work to be carried out in this area of tidal reliability assessment should focus on incorporating repairable data into the likelihood distribution as the assumption of constant failure rates used in this work simplifies the reality of turbine component failure behavior.

ACKNOWLEDGMENT

The support of the ETI and RCUK Energy Program funding for IDCORE (EP/J500847/1) is gratefully acknowledged.

REFERENCES

[1] M. Wilkinson, B. Hendriks, F. Spinato, and T. Van Delft, “Measuring wind turbine reliability, results of the reliawind project,” *Eur. Wind Energy Assoc. Conf.*, pp. 1–8, 2011.

[2] F. Spinato, P. J. Tavner, G. J. W. Van Bussel, and E. Koutoulakos, “Reliability of wind turbine subassemblies,” *Renew. Power Gener. IET*, vol. 1, no. 1, pp. 10–16, 2007.

[3] M. S. Hamada, A. G. Wilson, C. S. Reese, and H. F. Martz, *Bayesian Reliability*. 2008.

[4] P. R. Thies, G. H. Smith, and L. Johanning, “Addressing failure rate uncertainties of marine energy converters,” *Renew. Energy*, vol. 44, pp. 359–367, 2012.

[5] C. Iliev and D. Val, “Tidal current turbine reliability : power take - off train models and evaluation,” *3rd Int. Conf. Ocean Energy*, pp. 1–6, 2010.

[6] N. Logistics Technology Support Group and S. W. C. (CDNSWC), *Handbook of reliability prediction procedures for mechanical equipment*, no. Jan. 2010.

[7] J. Martín and C. J. Pérez, “Bayesian analysis of a generalized lognormal distribution,” *Comput. Stat. Data Anal.*, vol. 53, no. 4, pp. 1377–1387, 2009.

[8] “Rolling bearings : SKF Group,” *Skf Bear.*, pp. 97–114, 2012.

[9] M. G. Pecht and F. R. Nash, “Predicting the Reliability of Electronic Equipment,” *Proc. IEEE*, vol. 82, no. 7, pp. 992–1004, 1994.

[10] J. B. Gayo, “Final Publishable Summary of Results of Project ReliaWind,” 2011.