Proceedings of the ASME 2026 45th International Conference on Ocean, Offshore & Arctic Engineering
OMAE2026
June 7 - 12, 2026, Tokyo, Japan

OMAE2026-181827

DEVELOPMENT OF A PROBABILISTIC SURROGATE MODEL FOR LOADS ON AN EMBEDDED PLATE ANCHOR. PART II: CYCLIC LOADS

E. B. L. Mackay,
University of Exeter, UK

A. C. Pillai,
University of Exeter, UK

K. A. Kwa
University of Southampton, UK

ABSTRACT

The use of probabilistic surrogate models provides a computationally efficient means of generating time series of random loading on anchors over their design lifetimes. In this work we demonstrate how probabilistic surrogates for anchor loads can be developed, using a case study for the IEA 15 MW offshore wind turbine on the VolturnUS-S semi-submersible platform at a site off the east coast of Scotland. Simulations of the dynamic response were conducted for 1297 combinations of significant wave height, peak period, and hub-height wind speed, covering the range of observed conditions at the site. This paper considers the development of a probabilistic surrogate model for cyclic loading, conditional on environmental state, used for assessment of reduction in soil strength. The development of a probabilistic surrogate model for the short-term extreme loads is considered in another paper in the same proceedings. A novel method is presented for modelling the bimodal distribution of load cycle amplitudes, where a model is fitted to the larger amplitude cycles only, and the low amplitude cycles are treated as a nuisance component. The model is shown to accurately capture the accumulated soil damage from cyclic loading. Artificial neural networks are used to model the variation of distribution parameters with environmental conditions. We demonstrate a global fitting scheme, where data from all simulations are fitted simultaneously, after separating into either operational or idling conditions. The resulting surrogate model is shown to have good agreement with the simulation data.

1 INTRODUCTION

This paper is the second part of a study on the development of probabilistic surrogate models for loading on an embedded plate anchor, and continues from another paper in the same proceedings [1]. A general description of the background and requirement for probabilistic surrogate models is presented in the first part of the study. The study considers anchor loads from a floating offshore wind turbine (FOWT), using the IEA 15 MW reference turbine [2] on a semi-submersible platform [3], using a semi-taut mooring system. We refer the reader to Part I for a description of response simulations conducted and the surrogate model developed for the short-term extreme loading. In this second part, we consider the modelling of cyclic loads. As in Part I, we assume that the design of the turbine, platform and mooring are fixed, and we are interested in creating a surrogate for the anchor loading as a function of environmental condition. This surrogate will be applied to a whole-life geotechnical design method in future work. In the present work, the focus is on characterising the cyclic loading conditional on environmental condition.

To recap, an embedded plate anchor is considered to have failed when the applied load exceeds its ultimate capacity Q_{ult}. Under undrained conditions the capacity is linearly proportional to s_u, the volume-averaged strength of the soil [4]. Under cyclic loading conditions, the soil strength s_u decreases as a result of a build up in excess pore water pressure and reduction in soil effective stress. The reduction in soil strength is described in terms of a damage parameter.

The accumulation of damage is similar to fatigue analysis in structures. Rainflow counting is used to convert an irregular time series of loads to an equivalent series of cycles or half-cycles with

associated amplitudes, S, and mean values, R. An R-S-N curve for the soil type (analogous to a Wöhler curve) is then used to calculate the damage to the soil [5, 6] – see Appendix A and [7] for details. The accumulated damage on the anchoring system that results from various environmental (wind and wave) conditions is discussed in [7], where it is shown that high soil damage is linked to extreme environmental conditions. However, cyclic loads in more frequently-occurring operational conditions have the dominant contribution to soil damage when weighted by occurrence. This highlights the need for the surrogate modelling method presented here, which enables efficient modelling of anchor loads during both extreme and operational conditions.

Since the modelling of damage accumulation in soil is analogous to modelling damage accumulation in structures, we begin by giving a brief overview of surrogate modelling for cyclic loading in offshore structures in Section 2. The load cycle model adopted in this work is described in Section 3 and the artificial neural networks used in the surrogate are discussed in 4. Results are presented in Section 5. Finally, conclusions are presented in Section 6.

2 SURROGATE MODELS FOR CYCLIC LOADS

Surrogate modelling strategies can be classified in terms of whether they treat the outputs of the simulator (or a derived quantify such as a damage equivalent load (DEL)) as deterministic or stochastic. Deterministic surrogates reduce the stochastic output to some conditional statistic, such as the expected value $\mathrm{E}[Y|\boldsymbol{X}=\boldsymbol{x}]$. The disadvantage of this approach is that multiple simulations are required for each input condition in order to estimate the mean or variance. For wind turbine applications, averaging over six 10-minute simulations or one 60-minute simulation, as recommended in the IEC 61400-1 design standard [8], has been shown to be insufficient to converge to the expected value in a given environmental condition [9, 10]. To address this, Murcia et al [11] ran 100 simulations with different random inflow conditions for each environmental condition. They then constructed separate surrogates for the mean and standard deviation of fatigue loads on the DTU 10 MW reference turbine [12].

Deterministic surrogates for cyclic loads using artificial neural networks were developed in [13–15]. The use of Gaussian process regression (GPR) was considered in [16–18]. Various comparative studies of ANN, GPR and polynomial chaos expansions have been presented in [19–21]. ANNs generally perform favourably compared with other methods.

Probabilistic surrogate models for fatigue can be categorised in terms of whether they treat the DEL, or the load cycle amplitude as the random variable. Singh et al [22, 23] developed probabilistic surrogates for fixed and floating wind turbines that treat 10-minute DELs as the random variable. In this approach each simulation results in a single data point for modelling. Probabilistic surrogates for the fatigue damage of a mooring line of a FOWT were proposed in [24–26], where an ANN was used to map environmental conditions to the distribution of the tension cycle amplitude distribution. ANN and GPR models for the fatigue damage on a mooring line were compared in [27] and both models were found to perform well.

In this work we develop a surrogate for the cyclic loads on the mooring line tension at the anchor point. Since the surrogate is intended for optimising the anchor sizing, and the soil damage is dependent on the anchor size (see Appendix A), we develop a surrogate model for the the rainflow cycles, rather than for the damage itself.

3 LOAD CYCLE DISTRIBUTION MODEL

The damage accumulation model described in Appendix A and [7] takes a sequence of load cycles as input, identified using a rainflow counting algorithm. In this work, the rainflow counting algorithm used adopts the ASTM E1049 standard [28]. The output of the rainflow counting is a list of values (r_j, s_j, n_j), $j = 1,..,m$, where r_j and s_j are the mean and amplitude of cycle j, and $n_j = 0.5$ or 1, depending on whether it is a half or full cycle. We can think of the sequence of (r_j, s_j) values as being a sample from a random response vector $\boldsymbol{Y} = (R,S)$. To account for the occurrence of half and full cycles when fitting a model, we can double the occurrence of all cycles with $n_j = 1$. When we subsequently sample from the model we draw a total of $M = \sum_{j=1}^{m} n_j$ samples, to remove the effect of double counting.

3.1 Model for Rainflow Matrix Distribution

Various statistical models for the rainflow matrix distribution have been proposed. These are often based on mixture models of Rayleigh and Weibull distributions [29], Weibull and normal distributions [30, 31], Weibull and exponential distributions [32] or the Nakagami distribution [33]. Here we take a slightly different approach.

The purpose of the surrogate model is to be able to replicate the joint distribution of (R,S) *well enough* to allow an accurate calculation of the damage. To this end, we make a number of simplifying assumptions, and verify the impact on the calculated damage, where the damage model has been calibrated to match the damage response measured in element scale cyclic experiments on normally consolidated clays presented in [34, 35]. Firstly, it was found that replacing the cycle mean values $r_1, r_2, ...$ by the mean load over the simulation, made negligible difference to the calculated damage (less than 0.5% in any given condition). The reason for this is that for larger values of S, which contribute the most to the damage, there is a relatively narrow range of values of R (see upper left plot in Figure 1). So approximating R as a constant results in only a small error. This significantly simplifies the problem, from one of modelling the joint distribution of (R,S), to one of modelling just the univariate distribution of S.

Secondly, in many of the simulations, the observed distributions of S were bimodal, with a large number of low amplitude cycles. These low amplitude cycles have a negligible contribution to the damage, since the accumulated damage from an individual cycle is proportional to S^4 (see Appendix A).

Several options for modelling the bimodal distribution of S were investigated. These included (i) filtering the mooring tension signal prior to rainflow cycle counting to remove low amplitude cycles; (ii) using a multimodal distribution model; and (iii) treating the low-amplitude cycles as a nuisance component and just modelling the distribution of large-amplitude cycles. Of these, the last approach proved most effective in our case.

We assume that the observed distribution of S is composed a weighted sum of a small-cycle distribution, F_1 and a large-cycle distribution F_2:

$$F_S(s) = wF_1(s) + (1-w)F_2(s), \tag{1}$$

where $w \in [0,1)$ is a weight parameter. We make two further assumptions. Firstly, we assume F_1 has a finite upper end point s_0, so that $F_1(s) = 1$ for $s > s_0$. And secondly, that load cycles with $s \le s_0$ have negligible contribution to the damage. Under these assumptions, the form of F_1 is irrelevant. For $s > s_0$, the model (1) becomes

$$F_S(s) = w + (1-w)F_2(s), \quad s > s_0, \tag{2}$$

which does not involve F_1. Therefore, if the model is fitted only for values $s > s_0$, then we do not require a model for F_1. In the present work, we found that estimating s_0 as the highest value for which all cycles $s < s_0$ contribute to less than 0.1% of the total damage in a given environmental condition, provided a robust way of filtering out the low amplitude cycles that do not contribute to the damage. It was also found that modelling F_2 using a two-parameter Weibull distribution gave an accurate representation, as discussed further below.

3.2 Inference

For $s > s_0$, the density of (2) is $f_S(s) = (1-w)f_2(s)$. We cannot use this to estimate w by maximum likelihood, since setting $w = 0$ maximises $f_S(s)$ for any given form of $f_2(s)$. We therefore use an alternative method for inference. As the largest cycles contribute the most towards the damage, we use a fitting procedure that is weighted towards the tail of the distribution, by minimising the squared differences between the model and observations, of the log-exceedance probabilities.

Suppose we have a sequence of random cycle amplitudes as $S_1, \ldots, S_m$. The corresponding non-exceedance probabilities $P_1, \ldots, P_m$ are uniformly distributed. Therefore, the log-exceedance probabilities $Z_j := -\log(1-P_j)$ follow an exponential distribution with unit scale. So if we define an ordered sequence $S_{(1)} \le \ldots \le S_{(m)}$, then the corresponding ordered values $Z_{(1)} \le \ldots \le Z_{(m)}$ are order statistics of a standard exponential variable, with expected values [36]

$$z_j := \mathrm{E}\left[Z_{(j)}\right] = \sum_{k=1}^{j} \frac{1}{m-k+1}. \tag{3}$$

Note that $\mathrm{E}\left[P_{(j)}\right] = j/(m+1)$, but due to the nonlinear transformation, $\mathrm{E}\left[Z_{(j)}\right] = -\mathrm{E}\left[\log\left(1-P_{(j)}\right)\right] \neq -\log\left(1-\mathrm{E}\left[P_{(j)}\right]\right)$. The difference is small for lower j but is larger in the upper tail[1].

The Weibull distribution function is $F_{Wbl}(s) = 1 - \exp\left(-(s/\lambda)^k\right)$, where λ and k are the scale and shape parameters respectively. To fit the model we find the values of λ, k and w that minimise the squared differences between z_j and $\hat{z}_j = -\log\left(1-F_S(s_{(j)})\right)$, the negative log exceedance probability from the model. For $s > s_0$ we have

$$\hat{z}_j = (s_{(j)}/\lambda)^k - \log(1-w). \tag{4}$$

3.3 Example

An example of the model fit to observations is shown in Figure 1, for a case with $U_{hub} = 33$ m/s, $H_s = 5.0$ m and $T_p = 8.0$ s. In this case, there are 10 repeat simulations for the same environmental conditions, and the observation data has been pooled over the 10 repeats. The upper left plot shows a scatter plot of R against S. There is a larger range of R for the low cycle amplitudes, but a relatively narrow spread of values for the higher cycle amplitudes, which contribute most to the damage. The upper right plot shows a histogram of the observations together with the estimated threshold value s_0. There are a large number of low amplitude cycles that do not contribute to the damage. The model provides a good fit to upper part of the range. The lower left plot shows the empirical and model exceedance probabilities on a logarithmic scale, together with the corresponding Z values. The good fit of the model in the upper tail is evident. Finally, the lower right plot shows the damage calculated from the 10 individual simulations (without applying any of the simplifying assumptions mentioned above), together with the distribution of damage from the model based on 10^4 simulations from the estimated Weibull distribution. For each simulation we sample $M(1-w)$ points from the Weibull distribution, where M is the mean number of cycles per hour over the ten 1-hour simulations. The range of damage values from the model agrees well with the range from the repeat OrcaFlex simulations, indicating that the model is able to capture both the mean behaviour of the damage and the random variability.

[1] For the largest observation $\mathrm{E}\left[Z_{(m)}\right] = \gamma + \log(m) + O(1/m)$, where $\gamma \approx 0.57721$ is the Euler–Mascheroni constant, whereas $-\log\left(1-\mathrm{E}\left[P_{(m)}\right]\right) = \log(m+1)$. Hence, the difference between the two expressions tends to γ as $m \to \infty$. For example, for a sample size of $m = 100$, $\mathrm{E}\left[Z_{(m)}\right] = 5.19$, whereas $-\log\left(1-\mathrm{E}\left[P_{(m)}\right]\right) = 4.62$. Therefore, ignoring these differences can lead to biases in estimates of the upper tail.

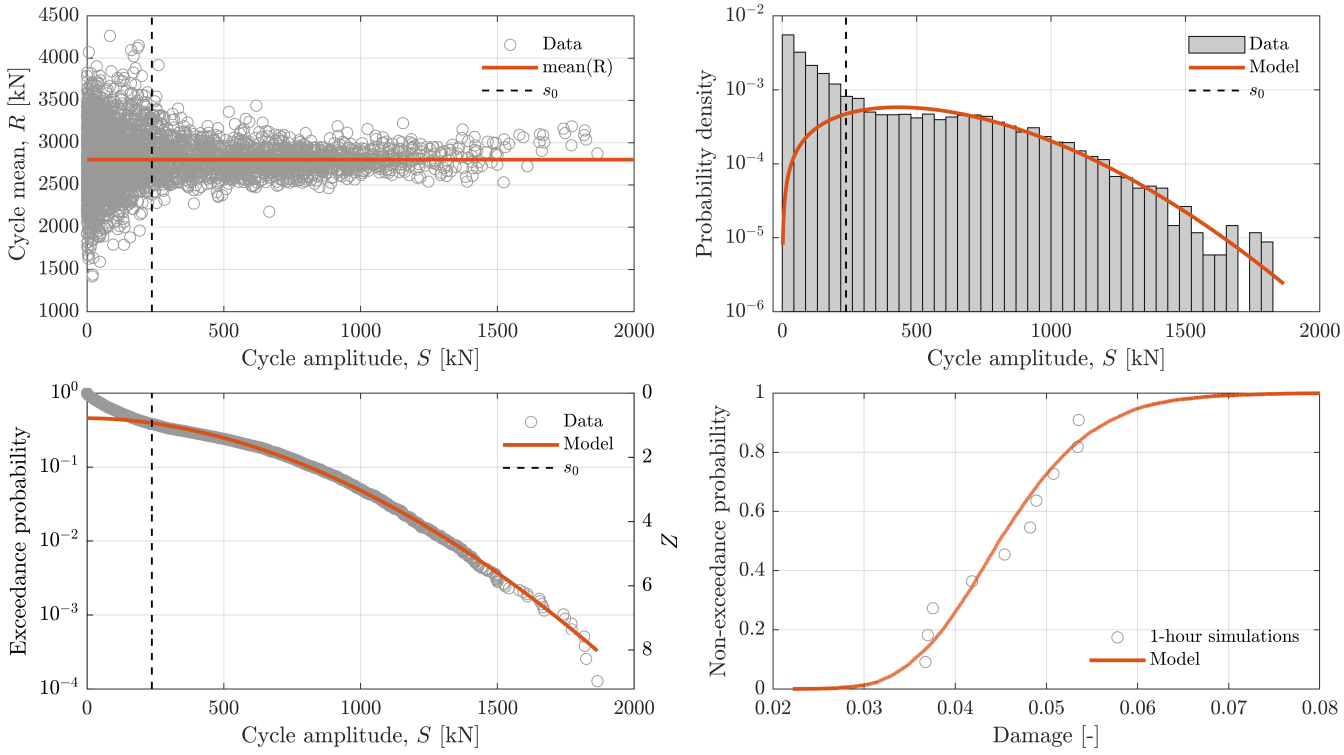

FIGURE 1. EXAMPLE OF MODEL FOR LOAD CYCLE AMPLITUDES FOR SIMULATIONS WITH $U_{hub} = 33$ m/s, $H_s = 5.0$ m, and $T_p = 8.0$ s. TOP LEFT: SCATTER PLOT OF LOAD CYCLE MEAN R AGAINST CYCLE AMPLITUDE S. TOP RIGHT: COMPARISON OF PROBABILITY DENSITY FOR OBSERVATIONS AND MODEL FOR S. LOWER LEFT: COMPARISON OF EXCEEDANCE PROBABILITIES FOR OBSERVATIONS AND MODEL FOR S. LOWER RIGHT: COMPARISON OF EMPIRICAL DISTRIBUTION OF DAMAGE FROM TEN SIMULATIONS OF 1-HOUR DURATION, WITH DISTRIBUTION CALCULATED FROM 10^4 REALISATIONS OF THE FITTED MODEL.

4 NEURAL NETWORK MODELS

As discussed in [1], we fit separate surrogates for operational and idling states due to the sharp transition in the response at the cut-out wind speed. We refer the reader to [1] for details of the general setup and hyperparameter optimisation.

As discussed above, our model for the load cycle distribution, is based on the assumption of a mean value of R per sea state, and modelling only those cycle amplitudes for which $S > s_0$, where the threshold s_0 also varies per sea state. Our model assumes that $M(\boldsymbol{x})$, $\bar{R}(\boldsymbol{x})$ and $s_0(\boldsymbol{x})$ are deterministic functions of environmental condition $\boldsymbol{x}$. We calculate their values for each simulation, then fit ANNs using a least-squares loss function. The ANNs are assumed to smooth any random variability in per-simulation estimates. The appropriate level of smoothing is determined by the hyperparameter optimisation, which finds the hyperparameters that give the best performance (i.e. lowest loss) for the validation data.

Once the ANN for $s_0(\boldsymbol{x})$ has been estimated, we use this model to discard cycles with $S \leq s_0(\boldsymbol{x})$ from each simulation. The ANN for the load cycle amplitude distribution parameters $(w(\boldsymbol{x}), \lambda(\boldsymbol{x}), k(\boldsymbol{x}))$ is then estimated by minimising the loss function

$$L(\mathcal{W}_{(w,\lambda,k)}) = \sum_{(i,j)\in I_{s_0}} [z_{i,j} - \hat{z}_{i,j}]^2, \tag{5}$$

where $I_{s_0} := \{(i,j) : i \in \{1,\ldots,N_{sim}\}, j \in \{1,\ldots,m_i\}, s_{i,j} > s_0(\boldsymbol{x}_i)\}$ as the set of indices of cycles which exceed the low-amplitude threshold $s_0(\boldsymbol{x})$, and $\mathcal{W}_{(w,\lambda,k)}$ is the parameter set of the ANN model for $(w(\boldsymbol{x}), \lambda(\boldsymbol{x}), k(\boldsymbol{x}))$. The values $z_{i,j}$ and $\hat{z}_{i,j}$ are the log-exceedance probabilities for the observations and model, defined in (3) and (4).

In total four ANNs are trained, and tuned using the same hyperparameter tuning framework as described in [1]. Following this process, the s_0 model consisted of 3 hidden layers with 192, 256, and 208 neurons respectively, dropout of 0.2, an initial learning rate of 5×10^{-4}, and weight decay of 3.48×10^{-6}. The model for estimating $\bar{R}$ per sea state consisted of 3 hidden layers with 256, 256, and 16 neurons respectively, dropout of 0.1, an initial learning rate of 10^{-3}, and weight decay of 7.1×10^{-5}. The $M(\boldsymbol{x})$ model consisted of 3 hidden layers with 112, 176, and

160 neurons respectively; dropout of 0.05, an initial learning rate of 7×10^{-4}, and weight decay of 3.49×10^{-6}. Finally, the neural network for estimating the load cycle distribution parameters consisted of a shared trunk for all parameters with 3 hidden layers of 64, 448, and 512 neurons respectively, dropout of 0.1, an initial learning rate of 10^{-4}, and weight decay of 10^{-4}.

All networks employed a learning rate scheduler which monitored the validation loss and reduced the learning rate by a factor of 0.5 after 5 epochs without improvement. The training process lasted for 100 epochs, with early stopping to avoid over-fitting.

5 RESULTS

Scatter plots of the deterministic surrogates for the low-amplitude threshold s_0, mean cycle level $\bar{R}$, and cycles per hour M against the corresponding empirical values are shown in Figures 2 to 4. Values are shown for both the training and validation data. To benchmark the level of scatter observed, the values from the repeat simulations are plotted against the corresponding 10-hour mean value (this is taken as a proxy for a 'perfect' surrogate). For the case of s_0, the level of scatter for the surrogate is very close to that for the repeat data, with an RMS error of 21.1 for the training data and 25.4 for the validation data, compared with 19.6 for the repeat data. The surrogate for $\bar{R}$ has a considerably higher level of scatter than for the repeat cases, but the coefficient of determination R^2 is still good, with a value of $R^2 = 0.99$ for the training and $R^2 = 0.93$ for the validation data. For the total cycle count M, there is considerable scatter between the surrogate and observations, both for the training and validation data. However, as discussed further below, the metocean conditions resulting in large cycle counts tended to be for lower-damage sea states, so the larger scatter here is less consequential.

Figure 5 shows quantile-quantile (QQ) plots for the surrogate cycle amplitude distribution. The left hand plot shows surrogate values of log-exceedance probability Z against empirical values from the validation cases. Grey lines correspond to values in individual validation cases and the red line is the aggregated trend over all validation cases. The right hand plot is the same, but for the model versus empirical cycle amplitude S. Overall the agreement is excellent in both cases, with the aggregated trend very close to the 1:1 line. There is some deviation at the highest values, but these are likely due to sampling effects.

Finally, Figure 6 shows a scatter plot of damage from the surrogate against damage from the OrcaFlex model. In the case of the surrogate, the damage is taken as the median value over 200 simulations from the cycle amplitude distribution. Our assessment is based on an assumed ultimate capacity for the anchor of $Q_{ult} = 14.1$ MN, corresponding to a circular plate anchor with a diameter of 7 m, embedded in a soft seabed at a depth of twice its diameter. However, the value of Q_{ult} only affects the absolute values of the damage, and only has a very small affect on the

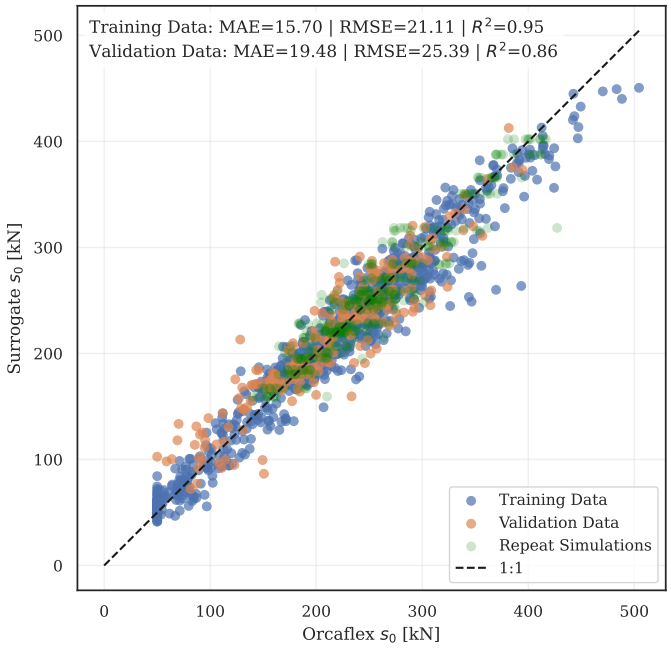

FIGURE 2. COMPARISON BETWEEN THE SURROGATE ESTIMATED LOW CYCLE THRESHOLD VALUES, s_0, AND THE ESTIMATES FROM ORCAFLEX SIMULATIONS. THE DASHED LINE INDICATES PERFECT AGREEMENT (1:1). MODEL PERFORMANCE IS COMPUTED BOTH FOR THE TRAINING DATASET (BLUE POINTS) AND THE INDEPENDENT VALIDATION DATASET (ORANGE POINTS). THE GREEN POINTS REPRESENT THE VARIATION DUE TO RANDOM SAMPLING AND PLOTTED HERE AS A BENCHMARK FOR THE LEVEL OF SCATTER.

agreement between the surrogate model and simulations. Overall, the level of agreement is very good, with an RMS error of 14.6×10^{-3} for the validation data, compared to 9.8×10^{-3} for the repeat cases. Although, the level of scatter between the surrogate and simulations is higher than due to sampling effects alone, the model is able to capture the variation in soil damage with environmental conditions. The higher scatter may in part be due to the uncertainties in the values of $\bar{R}$ and M, which will introduce some additional uncertainty in the damage estimate.

6 CONCLUSIONS

This work has presented a novel method for estimating a probabilistic surrogate model for cyclic loading on an embedded plate anchor. The key novelties lie in (1) the method used to model the load cycle amplitude distribution, where the low-amplitude cycles are treated as noise and ignored, and (2) the global fitting method, where the model is fitted to data from all simulations simultaneously. By modelling the load cycle amplitude distribution as a function of environmental condition, each 1-hour response simulation creates a large sample of data to use

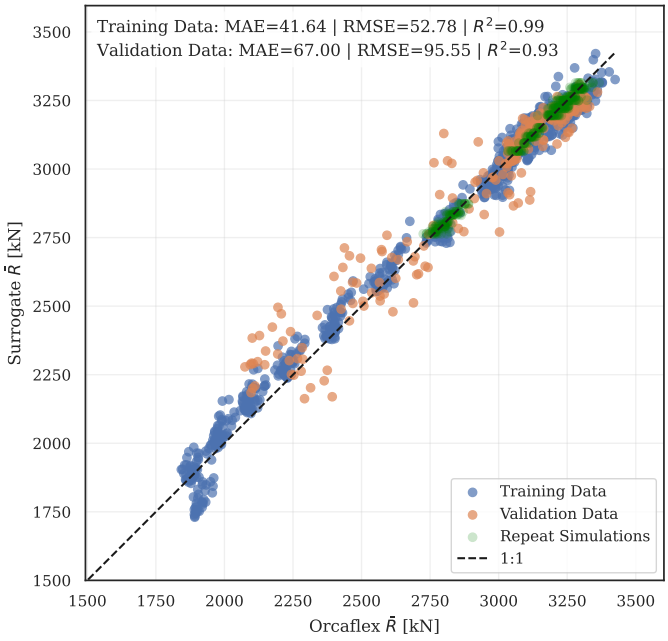

FIGURE 3. COMPARISON BETWEEN THE SURROGATE $\bar{R}$, AND THE ESTIMATES FROM ORCAFLEX SIMULATIONS. THE DASHED LINE INDICATES PERFECT AGREEMENT (1:1). MODEL PERFORMANCE IS COMPUTED BOTH FOR THE TRAINING DATASET (BLUE POINTS) AND THE INDEPENDENT VALIDATION DATASET (ORANGE POINTS). THE GREEN POINTS REPRESENT THE VARIATION DUE TO RANDOM SAMPLING AND PLOTTED HERE AS A BENCHMARK FOR THE LEVEL OF SCATTER.

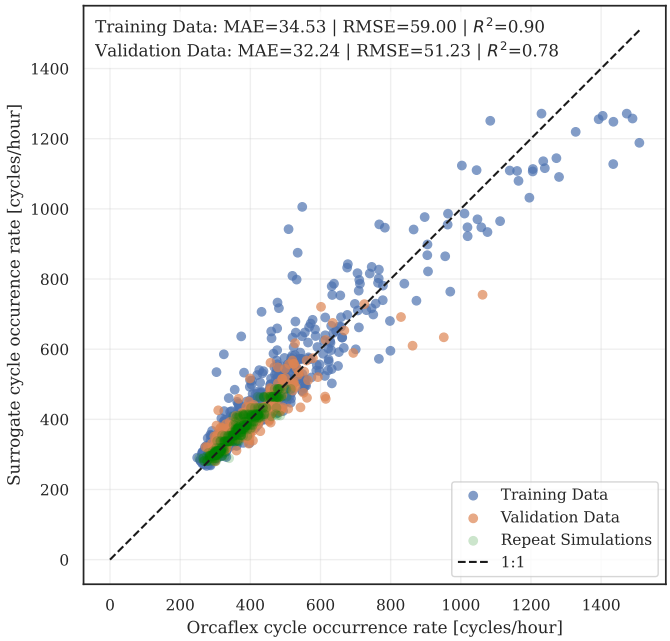

FIGURE 4. COMPARISON BETWEEN THE SURROGATE M, AND THE ESTIMATES FROM ORCAFLEX SIMULATIONS. THE DASHED LINE INDICATES PERFECT AGREEMENT (1:1). MODEL PERFORMANCE IS COMPUTED BOTH FOR THE TRAINING DATASET (BLUE POINTS) AND THE INDEPENDENT VALIDATION DATASET (ORANGE POINTS). THE GREEN POINTS REPRESENT THE VARIATION DUE TO RANDOM SAMPLING AND PLOTTED HERE AS A BENCHMARK FOR THE LEVEL OF SCATTER.

in the inference. In contrast, for surrogates which model derived quantities such as a damage equivalent load, each 1-hour response simulation only produces a single data point.

The modelling approach was shown to give an excellent fit to the load cycle amplitude distribution, and be able to accurately capture the accumulated damage to the soil resulting from the cyclic loading. The good fit to the load cycle amplitude distribution shows that the modelling strategy of treating the low-amplitude cycles as noise is effective, and allows a relatively simple two-parameter Weibull model to be used, reducing the number of parameters that must be estimated.

The present work has considered the accuracy of the surrogate model on a per-sea-state basis. As soil damage is a cumulative process, future work could consider the assessment of the surrogate based on a damage value weighted by the relative occurrence of environmental conditions, as discussed in [7].

As with the model for the short-term load peaks, presented in [1], there is some scope for improving the surrogate model fit, as observed differences between the model and observations were shown to be larger than the expected level of scatter from sampling effects alone. Ways in which the model can be improved will be investigated in future work.

ACKNOWLEDGMENT

This work was funded by the EPSRC Supergen Offshore Renewable Energy Hub, United Kingdom [grant no: EP/Y016297/1] under the NextGen Anchor flexible fund project. A. C. Pillai and K.A Kwa acknowledge support from the Royal Academy of Engineering under the Research Fellowship scheme [Awards RF\202021\20\175 and RF 2122-21-119]. The metocean data was supplied by the UK Met Office.

References

[1] E. Mackay, A. Pillai, and K. Kwa, "Development of a probabilistic surrogate model for loads on an embedded plate anchor. Part I: Short-term extreme loads," in *International Conference on Offshore Mechanics and Arctic Engineering*, 2026, OMAE2026–175 098.

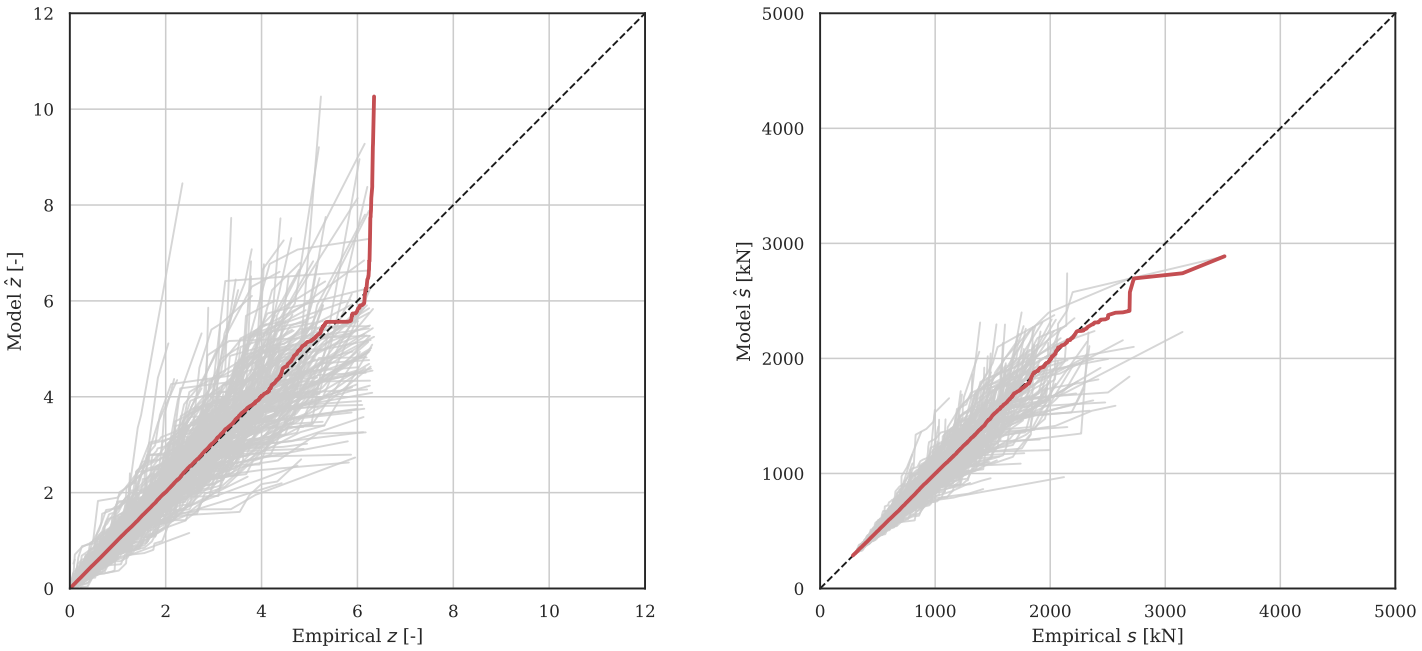

FIGURE 5. QQ PLOTS OF SURROGATE VALUES AGAINST EMPIRICAL VALUES FROM THE ORCAFLEX SIMULATIONS OF LOG-EXCEEDANCE PROBABILITIES Z (LEFT) AND CYCLE AMPLITUDE S (RIGHT). GREY LINES CORRESPOND TO VALUES IN INDIVIDUAL VALIDATION CASES AND RED LINES ARE AGGREGATED TRENDS OVER ALL VALIDATION CASES.

[2] E. Gaertner et al., "Definition of the IEA 15-Megawatt offshore reference wind turbine," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep. NREL/TP-5000-75698, 2020. Available: https://www.nrel.gov/docs/fy20osti/75698.pdf.

[3] C. Allen et al., "Definition of the UMaine VolturnUS-S reference platform developed for the IEA wind 15-Megawatt offshore reference wind turbine," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep. NREL/TP-5000-76773, 2020. Available: https://www.nrel.gov/docs/fy20osti/76773.pdf.

[4] S. Stanier and D. White, "Enhancement of bearing capacity from consolidation: Due to changing strength or failure mechanism?" *Géotechnique*, vol. 69, no. 2, pp. 166–173, 2019. DOI: 10.1680/jgeot.17.T.030.

[5] A. Verruijt, *Computational geomechanics*. Springer Science & Business Media, 1995, vol. 7.

[6] M. Randolph and S. Gourvenec, *Offshore geotechnical engineering*. CRC press, 2017.

[7] K. Kwa, E. Mackay, and A. Pillai, "Linking environmental conditions to cyclic and peak anchor loads for a 15 MW floating offshore wind turbine," in *International Conference on Offshore Mechanics and Arctic Engineering*, 2026, OMAE2026–175 105.

[8] International Electrotechnical Commission, *Wind energy generation systems - Part 1: Design requirements*. IEC 61400-3-1, 2019.

[9] D. Zwick and M. Muskulus, "The simulation error caused by input loading variability in offshore wind turbine structural analysis," *Wind energy*, vol. 18, no. 8, pp. 1421–1432, 2015. DOI: 10.1002/we.1767.

[10] K. Müller and P. W. Cheng, "Application of a monte carlo procedure for probabilistic fatigue design of floating offshore wind turbines," *Wind Energy Science*, vol. 3, no. 1, pp. 149–162, 2018. DOI: 10.5194/wes-3-149-2018.

[11] J. P. Murcia et al., "Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates," *Renewable Energy*, vol. 119, pp. 910–922, 2018. DOI: 10.1016/j.renene.2017.07.070.

[12] C. Bak et al., "The DTU 10-MW reference wind turbine," in *Danish wind power research 2013*, 2013. Accessed: Dec. 16, 2025. Available: https://orbit.dtu.dk/en/publications/the-dtu-10-mw-reference-wind-turbine/.

[13] L. Schröder, N. K. Dimitrov, D. R. Verelst, and J. A. Sørensen, "Wind turbine site-specific load estimation using artificial neural networks calibrated by means of high-fidelity load simulations," in *Journal of Physics: Confer-*

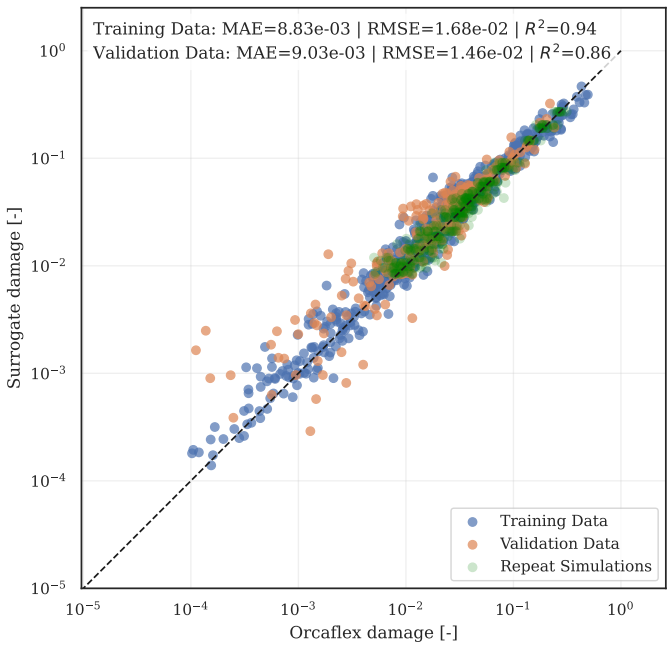

FIGURE 6. COMPARISON BETWEEN THE SURROGATE ESTIMATED DAMAGE AND THE ESTIMATES FROM ORCAFLEX SIMULATIONS. THE DASHED LINE INDICATES PERFECT AGREEMENT (1:1). MODEL PERFORMANCE IS COMPUTED BOTH FOR THE TRAINING DATASET (BLUE POINTS) AND THE INDEPENDENT VALIDATION DATASET (ORANGE POINTS). THE GREEN POINTS REPRESENT THE VARIATION DUE TO RANDOM SAMPLING AND PLOTTED HERE AS A BENCHMARK FOR THE LEVEL OF SCATTER.

ence Series, vol. 1037, 2018, p. 062 027. DOI: 10.1088/1742-6596/1037/6/062027.

[14] N. Dimitrov, "Surrogate models for parameterized representation of wake-induced loads in wind farms," *Wind Energy*, vol. 22, no. 10, pp. 1371–1389, 2019. DOI: 10.1002/we.2362.

[15] K. Shaler, J. Jasa, and G. E. Barter, "Efficient loads surrogates for waked turbines in an array," in *Journal of Physics: Conference Series*, vol. 2265, 2022, p. 032 095. DOI: 10.1088/1742-6596/2265/3/032095.

[16] R. Teixeira, A. O'Connor, M. Nogal, N. Krishnan, and J. Nichols, "Analysis of the design of experiments of offshore wind turbine fatigue reliability design with kriging surfaces," *Procedia Structural Integrity*, vol. 5, pp. 951–958, 2017. DOI: 10.1016/j.prostr.2017.07.132.

[17] L. D. Avendaño-Valencia, I. Abdallah, and E. Chatzi, "Virtual fatigue diagnostics of wake-affected wind turbine via gaussian process regression," *Renewable Energy*, vol. 170, pp. 539–561, 2021. DOI: 10.1016/j.renene.2021.02.003.

[18] X. Li and W. Zhang, "Long-term fatigue damage assessment for a floating offshore wind turbine under realistic environmental conditions," *Renewable Energy*, vol. 159, pp. 570–584, 2020. DOI: 10.1016/j.renene.2020.06.043.

[19] N. Dimitrov, M. C. Kelly, A. Vignaroli, and J. Berg, "From wind to loads: Wind turbine site-specific load estimation with surrogate models trained on high-fidelity load databases," *Wind Energy Science*, vol. 3, no. 2, pp. 767–790, 2018. DOI: 10.5194/wes-3-767-2018.

[20] G. Gasparis, W. H. Lio, and F. Meng, "Surrogate models for wind turbine electrical power and fatigue loads in wind farm," *Energies*, vol. 13, no. 23, p. 6360, 2020. DOI: 10.3390/en13236360.

[21] R. M. Slot, J. D. Sørensen, B. Sudret, L. Svenningsen, and M. L. Thøgersen, "Surrogate model uncertainty in wind turbine reliability assessment," *Renewable Energy*, vol. 151, pp. 1150–1162, 2020. DOI: 10.1016/j.renene.2019.11.101.

[22] D. Singh, R. Dwight, and A. Viré, "Probabilistic surrogate modeling of damage equivalent loads on onshore and offshore wind turbines using mixture density networks," *Wind Energy Science*, vol. 9, no. 10, pp. 1885–1904, 2024. DOI: 10.5194/wes-9-1885-2024.

[23] D. Singh, E. Haugen, K. Laugesen, R. P. Dwight, and A. Viré, "Data-driven probabilistic surrogate model for floating wind turbine lifetime damage equivalent load prediction," *Wind Energy Science Discussions*, vol. 2025, pp. 1–33, 2025. DOI: 10.5194/wes-2025-24.

[24] Y. Kim, H. Kim, and I.-G. Ahn, "A study on the fatigue damage model for gaussian wideband process of two peaks by an artificial neural network," *Ocean Engineering*, vol. 111, pp. 310–322, 2016. DOI: 10.1016/j.oceaneng.2015.11.008.

[25] C. B. Li and J. Choung, "Fatigue damage analysis for a floating offshore wind turbine mooring line using the artificial neural network approach," *Ships and Offshore Structures*, vol. 12, no. sup1, S288–S295, 2017. DOI: 10.1080/17445302.2016.1254522.

[26] C. B. Li, J. Choung, and M.-H. Noh, "Wide-banded fatigue damage evaluation of catenary mooring lines using various artificial neural networks models," *Marine Structures*, vol. 60, pp. 186–200, 2018. DOI: 10.1016/j.marstruc.2018.03.013.

[27] Y. Zhao and S. Dong, "Probabilistic fatigue surrogate model of bimodal tension process for a semi-submersible platform," *Ocean Engineering*, vol. 220, p. 108 501, 2021. DOI: 10.1016/j.oceaneng.2020.108501.

[28] ASTM International, *Standard Practices for Cycle Counting in Fatigue Analysis*. ASTM E1049-85, 2017.

[29] W. Zhao and M. J. Baker, "On the probability density function of rainflow stress range for stationary gaussian

processes," *International Journal of fatigue*, vol. 14, no. 2, pp. 121–135, 1992. DOI: 10.1016/0142-1123(92)90088-T.

[30] M. Nagode, J. Klemenc, and M. Fajdiga, "Parametric modelling and scatter prediction of rainflow matrices," *International Journal of Fatigue*, vol. 23, no. 6, pp. 525–532, 2001. DOI: 10.1016/S0142-1123(01)00007-X.

[31] J. Klemenc and M. Fajdiga, "Improved modelling of the loading spectra using a mixture model approach," *International Journal of Fatigue*, vol. 30, no. 7, pp. 1298–1313, 2008. DOI: 10.1016/j.ijfatigue.2007.08.024.

[32] Y. Zhu, P. Li, and J. Sun, "A probability density model of stress amplitude under bimodal vibration response," *International Journal of Fatigue*, vol. 170, p. 107 540, 2023. DOI: 10.1016/j.ijfatigue.2023.107540.

[33] Q. Han, J. Li, J. Xu, F. Ye, A. Carpinteri, and G. Lacidogna, "A new frequency domain method for random fatigue life estimation in a wide-band stationary gaussian random process," *Fatigue & Fracture of Engineering Materials & Structures*, vol. 42, no. 1, pp. 97–113, 2019. DOI: 10.1111/ffe.12875.

[34] K. Andersen and R. Lauritzen, "Cyclic bearing capacity analysis for gravity platforms; calculation procedure, verification by model tests, and application for the gullfaks c platform," 1988.

[35] K. H. Andersen et al., "Cyclic soil parameters for offshore foundation design," *Frontiers in offshore geotechnics III*, vol. 5, no. 11, 2015.

[36] H. David and H. Nagaraja, *Order Statistics*, Third. Wiley Series in Probability and Statistics, 2003, ISBN: 0-471-38926-9.

[37] Marine Scotland, *National marine plan interactive map*, https://marinescotland.atkinsgeospatial.com/nmpi/default.aspx?layers=745, Accessed: 2025-01-21.

[38] D. Zografou, S. Gourvenec, and C. O'Loughlin, "Response of normally consolidated kaolin clay under irregular cyclic loading and comparison with predictions from the accumulation procedure," *Géotechnique*, vol. 69, no. 2, pp. 106–121, 2019. DOI: 10.1680/jgeot.16.P.340.

[39] K. S. Skau, B. M. Dahl, H. P. Jostad, Y. Suzuki, J. D. Sordi, and O. Havmøller, "Response of lightly overconsolidated clay under irregular cyclic loading and comparison with predictions from the strain accumulation procedure," *Géotechnique*, vol. 73, no. 12, pp. 1087–1099, 2022. DOI: 10.1680/jgeot.21.00139.

APPENDIX A - DAMAGE ACCUMULATION MODEL

Suppose that rainflow counting has been used to identify a sequence of load cycles $\{(r_1,s_1,n_1),...,(r_m,s_m,n_m),\}$, where r_j is the cycle mean, s_j is the cycle amplitude and $n_j = \frac{1}{2}$ or 1, depending on whether it is a cycle or half cycle. For a given anchor ultimate capacity Q_{ult}, we define normalised cycle means and amplitudes as $\tilde{r}_j = r_j/Q_{ult}$ and $\tilde{s}_j = s_j/Q_{ult}$. The damage is calculated sequentially, with the accumulated damage up to cycle j defined as

$$d_j = k_1\,(1+\tilde{r}_j)^{k_5}\left(1-\exp\left(-k_2\,(n_j+b_{j-1})\,(\tilde{s}_j-k_4)_+^{k_3}\right)\right), \quad (6)$$

where $(\cdot)_+ = \max(\cdot,0)$, b_{j-1} is the equivalent number of cycles with mean and amplitude $(\tilde{r}_j,\tilde{s}_j)$ that would give damage d_{j-1}, and $k_1,...,k_5$ are constants describing the soil properties. The soil conditions at the location used in this work are predominantly sediment, i.e. mud and sandy mud [37]. Based on this, throughout this work, the soil parameters are set at $k_1 = 1$, $k_2 = 22.58$, $k_3 = 4$, $k_4 = 0.001$ and $k_5 = 1$, which have been calibrated to result in damage contours presented in [34, 35] for normally consolidated clays.

The value b_{j-1} is found by rearranging (6) to get

$$b_{j-1} = -\frac{1}{k_2(\tilde{s}_j-k_4)^{k_3}}\log\left(1-\frac{d_{j-1}}{k_1(1+\tilde{r}_j)^{k_3}}\right). \quad (7)$$

In the present work, the damage is used only to assess the model for the load cycle distribution. We therefore, assume that the damage starts at zero at the start of each response simulation, and set $d_0 = 0$, so that $b_0 = 0$.

The sequential accumulation of damage defined by equations (6) and (7) is dependent on the ordering of the cycles. Changing the order of the cycles results in a different total damage. The convention is to sort the cycles by increasing $\tilde{r}$ [38, 39]. In our dataset, there was relatively little variation in the cycle mean for a given metocean condition. It was found that replacing $\tilde{r}_1,...,\tilde{r}_m$ by the mean value in the simulation $\bar{r} = \frac{1}{m}\sum_{j=1}^{m}\tilde{r}_j$, resulted in a mean change of 0.2% over all simulations, and less than 1% in any given simulation. When r_j is constant, the damage accumulated from m cycles with variable s_j and n_j is invariant to the ordering of the cycles, and simplifies to

$$d_m = k_1\,(1+\bar{r})^{k_5}\left(1-\exp\left(-k_2\sum_{j=1}^{m}\left[n_j\,(\tilde{s}_j-k_4)_+^{k_3}\right]\right)\right). \quad (8)$$

From (8), the contribution from cycle i to the total damage accumulated over all cycles is $n_i\,(\tilde{s}_i-k_4)_+^{k_3}\,/\sum_{j=1}^{m}\left[n_j\,(\tilde{s}_j-k_4)_+^{k_3}\right]$. In the statistical model for the cycle amplitude distribution, described in Section 3, it is assumed that $n_j = 1$ for $j = 1,...,m$, since full-cycles are counted twice. Moreover, since k_4 is close to zero and $k_3 = 4$ in the present work, the contribution from cycle i to the total damage is (approximately) proportional to $\tilde{s}_i^4$.