

Durodola, J., Li, N., Ramachandra, S. and Thite, A. (2017) 'A pattern recognition artificial neural network method for random fatigue loading life prediction', *International Journal of Fatigue*, 99 (Part 1), pp. 55-67.

DOI: <https://doi.org/10.1016/j.ijfatigue.2017.02.003>

This document is the authors' Accepted Manuscript.

License: <https://creativecommons.org/licenses/by-nc-nd/4.0>

Available from RADAR: <https://radar.brookes.ac.uk/radar/items/13792434-2dbd-4559-9f35-5b9baad04d12/1/>

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners unless otherwise waved in a license stated or linked to above. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

# **A pattern recognition artificial neural network method for random fatigue loading life prediction**

**J F Durodola\*, N Li, S Ramachandra and A N Thite**

Oxford Brookes University, Faculty of Technology, Design and Environment, Wheatley  
Campus, Oxford OX33 1HX, UK.

## **Abstract**

Random vibration fatigue loading occurs in automotive, aerospace, offshore and indeed in many structural and machine components. The analysis of these types of problems is often carried out using either time domain or frequency domain methods. Time domain rainflow counting together with Miner's linear damage accumulation assumption is widely accepted as a method of rationalising stress amplitude and mean stress from random fatigue loading and the damage caused to the component. Frequency domain methods provide a faster alternative for the analysis of the same problem but the results are generally conservative compared to those obtained using time domain methods. This paper presents an artificial neural network (ANN) machine learning approach for the prediction of damage caused by random fatigue loading. The results obtained for ergodic Gaussian stationary stochastic loading is very encouraging. The method embodies rapid analysis as well as better agreement with rainflow counting method than existing frequency domain methods.

**Keywords** random fatigue, frequency, time domain, artificial neural networks, Dirlik.

\*Corresponding author, email: [jdurodola@brookes.ac.uk](mailto:jdurodola@brookes.ac.uk); Tel: + 44 (0)1865 483501.

## **1. Introduction**

The paper presents a general machine learning Artificial Neural Network (ANN) method for the analysis of random fatigue loading problem. Virtually all fatigue loading inherently have random loading in addition to deterministic loads. Random fatigue loading is usually analysed in time or frequency domain. The widely accepted time domain approach is based on the rainflow counting method for reducing the loading spectrum into amplitude and mean loading data and the use of Miner's rule for the calculation of cumulative damage. Most wideband frequency domain methods include probability distribution functions with parameters that were determined based on minimising the error of correlation with the predictions based on rainflow counting and Miner's method [1]. The methods used for the determination of the parameters for the probability distribution function included intuitive empirical observations in some cases [1], [2]. While the success of the models developed are respectable, significant and wide discrepancies are still observed in the correlations. This problem may be partly attributed to the fact that the numbers of power spectral density samples used in the development of the models were few. Dirlik [1] used 70 and Petrucci [2] used 45.

Artificial neural network has been known to provide greater scope for non-linear generalisation and ability to deal with a large number of input variables than direct application of optimisation methods [3], [4], [5]. Very little has however been reported in the literature on the use of artificial neural network method on problems related to random fatigue loading problems. A recent publication [6], [7] showed the possibility for ANN to be able to identify a spectral type and use this with various models such as Wirsching-Light, Zhao-Baker, Benasciutti-Tovo and Dirlik to predict damage [1], [8], [9], [10]. This effort did not however progress beyond the use of an idealised spectrum of loading. This paper presents an artificial neural network method for random loading fatigue analysis for predicting fatigue damage

using a broad range of spectral types and material properties. The model is based on a broad coverage of various properties and parameters such as spectral moments, ultimate tensile strength and fatigue strength coefficient and exponent. The results obtained generally demonstrate excellent ability of ANN to make better fatigue damage predictions compared to existing frequency domain methods.

## 2. Theory

This section describes the main theoretical background for the work presented. The main aspects are characterisation of random loading, fatigue life, type of loading, material properties, and artificial neural network method.

### 2.1 Characterisation of random processes and fatigue loading

It is necessary to indicate that this work concerns statistical random processes or loading that are stationary and ergodic. Ergodicity means that the statistical properties of representative samples of the process are the same over the equal time interval. By stationarity, we mean that the statistical properties of the random process at a particular time over many representative samples are the same. These conditions in effect mean that the samples we analyse are representative of the loading over time. The Fourier transform  $X(f)$  of a time domain signal  $x(t)$  is given by equation (1). According to Parseval's theorem [11] the energy under the time and frequency domains are the same, as expressed in equation (2).

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{2\pi i f t} dt \quad (1)$$

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df \quad (2)$$

where  $t$  represents time and  $f$  represents frequency. The input output relationships of random processes are more conveniently described using the frequency domain representation. The power spectral density (PSD) equation (3) describes the variation of the power content of a signal in real terms with frequency.

$$G_x(f) = \lim_{T \rightarrow \infty} E[|X(f)|^2] \quad (3)$$

where  $G_x(f)$  is the PSD and the symbol  $E$  is used to denote the expectation of the value in bracket. The value of  $G_x(f)$  used in this work was normalised with respect to the sampling frequency. In the context of fatigue, the power input during loading into a component is related to the fatigue damage by parameters based on the PSD. The main parameters used are the moments of the PSD which are given by equation (4)

$$m_i = \int_0^{\infty} f^i G_x(f) df \quad (4)$$

where  $m_i$  is the  $i$  th moment of the PSD  $G_x(f)$ ;  $i = 0, 1, 2, 4$  are used for fatigue loading characterisation. Other parameters derived from the moments such as given in equations (5) and (6)

$$E(0) = (m_2 / m_0)^{1/2} \quad (5)$$

$$E(P) = (m_4 / m_2)^{1/2} \quad (6)$$

provide estimates for the number of upward mean crossing and the number of peaks in the signal per second respectively. The irregularity factor,  $\gamma$ , of the signal is given by equation (7)

$$\gamma = E(0) / E(P) = m_2 / (m_0 m_4)^{1/2} \quad (7)$$

reflects the spread of the process,  $\gamma$  tending to a value of 0 or 1 corresponds to a broad band or a narrow band signal respectively. An alternative description of bandwidth characteristics of a signal is given the Vanmarcke's parameter  $q_x = [1 - m_1^2 / (m_0 m_2)]^{1/2}$  [8].

Twelve different forms of spectra including those used by Dirlik, 1985 [1], Tovo [12] and Benasciutti and Tovo [8] as illustrated in Figure 1 were used in this study. The frequency

values  $f_i$ ,  $i = 1,6$ ; spectral amplitudes  $d_i$ ,  $i=1,3$  and shape modification parameters  $d_4$  and  $d_5$  were chosen using the Latin Hypercube Sampling (LHS) [13] experimental design approach. This facilitates maximum coverage of the fatigue loading space. The ranges of the material properties, i.e the ultimate tensile strength,  $\sigma_{uts}$ , and fatigue strength and fatigue strength exponents  $a$  and  $b$ ; and the limits of the spectral moment values  $m_i$ ,  $i = 0,1,2,4$  considered in the work are highlighted in Table 1.

## 2.2 Frequency domain damage prediction models

The Miner's linear cumulative damage rule for different fatigue loading states ( $S_{ai}$ ,  $S_{mi}$  and  $n_i$ ) is given by

$$E(D) = \sum_i n_i / N_i \quad (8)$$

where  $E(D)$  is damage fraction,  $S_{ai}$ ,  $S_{mi}$  and  $n_i$ , in the case of stress fatigue loading analysis, are the amplitude and the mean stress and number of cycles representing the loading corresponding to state  $i$ .  $N_i$  is the number of cycles under the loading state that on its own will cause fatigue failure. The life  $N_i$  is determined in terms of the fatigue properties of the material and the mean stress as highlighted in equations (9) and (10)

$$S_{ai} = \alpha_i a N_i^b \quad (9)$$

$$\alpha_i = (1 - S_{mi} / \sigma_{UTS}) \quad (10)$$

where  $a$  and  $b$  are the fatigue strength and exponents values for the material based on stress amplitude. The mean stress parameter  $\alpha_i$  in equation (10) is the Goodman's correction factor for accounting for the effect of the mean stress and  $\sigma_{uts}$  is the ultimate strength of the material. The frequency domain expression for fatigue damage prediction which includes Miner's rule [14] is generally written as in equation (11)

$$E(D) = E(P) \frac{T}{k} \int_0^\infty S^v p(S) dS \quad (11)$$

where  $T$  is the fatigue loading signal sampling time,  $E(P) = (m_4 / m_2)^{1/2}$  is the number peaks in the signal per second as indicated in equation (6),  $S$  is the stress range variable,  $v = -1/b$  and  $k = (2a)^{-1/b}$  and  $p(S)$  is the probability distribution function expressing the possibility of occurrence of  $S$ . Dirlik [1] in 1985 derived a probability distribution function  $p(S)$  based on the combination of statistical Rayleigh, Normal and additional distributions. The parameters of the distribution were determined using optimisation and heuristic observations to match the form of the distribution with results that obtained using rainflow counting method for 70 different types of spectral. The form of the  $p(S)$  obtained is given in equation (12)

$$p(S) = \left[ \frac{D_1}{Q} e^{-z/Q} + \frac{D_2 Z}{R^2} e^{-z^2/2R^2} + D_3 Z e^{-z^2/2} \right] / 2 \cdot m_0^{1/2} \quad (12)$$

where  $S$  represents stress range,  $Z$  is stress range normalised using the root of the spectral moment  $m_0$ . All other parameters  $R$ ,  $Q$ ,  $D_1$ ,  $D_2$  and  $D_3$  are intrinsically functions of the spectral moments  $m_i$ ,  $i=0,1,2$  and  $4$ . Petrucci and Zuccarello [2] presented a fatigue damage solution in the form shown in equation (13).

$$E(D) = E(P) \exp[\psi(\alpha_x, \beta_x, v, \gamma_p)] / (2\pi k) \quad (13)$$

The function  $\psi$ , and its variables  $\alpha_x$ ,  $\beta_x$ ,  $v$ , and  $\gamma_p$  are also functions of the spectral moments  $m_i$ ,  $i=0,1,2$  and , and  $k$  is a fatigue material property.

### 2.3 Machine learning artificial neural network implementation

An artificial neural network (ANN) is a mathematical model that is designed with the aim for it to function like the neural network in the human brain maps input data to the appropriate output data. ANNs are often used as pattern classifiers or as function approximators. Function approximation is the primary mode for the ANNs used in the work; pattern recognition is intrinsic. Three layers of neuron as illustrated in Figure 3 is generally accepted as sufficient to represent any non-linear function approximation [15]. Each of the hidden and output layer neuron which we here represent by the index  $j$  is connected to the neurons in the preceding

layer. The preceding layer to the output layer is the hidden layer and correspondingly the preceding layer to the hidden layer is the input layer. We have used the index  $i$  here to represent neurons in a preceding layer. The strength of the connection between neurons  $j$  and  $i$  are often represented by the term  $w_{ij}$ . The bias value for each hidden and output neuron is indicated in Figure 2 by the symbol  $\theta_j$ .

Before an ANN can be used, it needs to be trained on data for which the inputs and the corresponding target outputs are known. The training process used in this work is based on the feedforward – backpropagation multilayer perceptron (MLP) method. As highlighted in the foregoing the connection weights  $w_{ij}$  as well as the bias  $\theta_j$  are initially set to some random values, e.g from  $\sim n(0,1)$  which is a standard normal distribution population with 0 mean and variance of 1. The training proceeds by feeding known inputs into the network and obtaining its corresponding predictions for the output. In this process, each internal and output neuron  $j$  receives a weighted sum  $x_j$  of input values  $x_i$  from the preceding neurons  $i$  according to equation (14). The output from the neuron  $j$  denoted by  $y_j = f(x_j)$  is as given in equations (15) and (16) for the internal and output neurons respectively;

$$x_j = \sum_i w_{ij} x_i \quad (14)$$

$$y_j = \frac{1}{1 + \exp(-\sigma x_j + \theta_j)} \quad (15)$$

$$y_j = x_j + \theta_j \quad (16)$$

The sigmoid function in equation (15) is numerically desirable in the perceptron model as it ensures that all values passing to the next neuron lie in the range  $[0,1]$  and more importantly the form of its derivative makes it easy to determine its derivative which is used in the backpropagation process for weight adjustment. The parameter  $\sigma$  in equation (15) modifies the slope of the sigmoidal function. Its value was set to 0.35 which was the magnitude found



to give consistently good prediction in a previous work [3]. The output layer used a linear transfer function, to ensure that erroneous outputs were easy to identify.

The output  $y_j$  from (16) will not in general match the known output corresponding to the inputs used from the data set, at least in the first feedforward through process. The mis-match error, usually the mean square error value, is then used in the backpropagation process to adjust or modify the weights  $w_{ij}$  and bias  $\theta_j$  so that better prediction can be made by the network in the next feedforward iteration process. A number of iterations of feedforward – backpropagation is required before the weights become useful and able to make a right prediction on the current data and subsequently on a previously unseen input data. Various backpropagation algorithms have been devised for the training of networks. The method used in this analysis was based on the +Rprop algorithm which is known to have excellent convergence characteristics [3], [16]. The parameters required for the optimal convergence of the training in this approach has been identified for most problems and are not dependent on trial and error. For research flexibility purposes, the implementation of the ANN in this work was carried out using a set of in house routines developed in a MATLAB [17] environment. The inputs used in this work is discussed in what follows. The output is fatigue damage fraction.

As to be expected the fatigue properties of the material were essential inputs. It is tempting to simplistically expect ANN to be able to deduce the loading characteristics from the random time fatigue loading data that is available directly. Preliminary investigation showed that this approach was not feasible. As highlighted in Section 2.1 of this paper, the moments  $m_0$ ,  $m_1$ ,  $m_2$  and  $m_4$  have been identified as parameters from the frequency domain that characterise the original time domain random fatigue data [1]. These parameters together with the fatigue properties of the material i.e.  $\sigma_{uts}$ ,  $a$  and  $b$  were found useful in this work for the prediction

of fatigue damage using ANN. These seven parameters were fed into the input layer. This is the minimum set of input data required to describe a broadband problem; 3 parameters describe the material properties and the 4 moments describe the spectral characteristics as used in Dirlik [1]. The use of higher moments such as  $m_i, i > 4$  compared to the cases  $i = 0, 1, 2$  and  $4$  does not present additional information that could help with the prediction generalisation, instead they will increase the number of input variables to be considered in the training deduction process. Higher moments are simply products of the PSD and the frequency raised to higher powers. The range  $i = 0, 1, 2$  and  $4$  already contain constant, odd and even powers of the frequency. Any higher odd or even powers of the frequency will not be adding new information. All the input variables used in the study were scaled to fall between 0 and 1. The number of neurons in the output layer was one representing the fatigue damage fraction predicted. Although it is suggested that half the sum of input and output neuron [18], [19] is adequate as the number to be set for the hidden layer neuron, experimentation with this number in this work did not show good prediction. The number of hidden layer found to give good prediction was between 35 and 45; 40 was used for the work presented here. The recommendation to use about 20 times the total number of neurons [19] provided some guidance for the number of patterns constructed for the ANN training and testing in this study. In order to include extreme conditions, the numbers of patterns developed and used ranged from 100 to 20,000.

## **2.4 Overall procedure**

A flowchart describing the overall process followed in the developments for this work is summarised in Figure (3). The analysis starts with the composition of trial space of many different spectral forms as illustrated in Figure 1. The shapes and frequency parameters in the spectra are randomly permuted using design of experiment methods [20]. As highlighted in the foregoing, different sample sizes ranging from 100 to 20000 were analysed in the course

of the study. Fatigue material properties  $\sigma_{uts}$ ,  $a$  and  $b$  were sampled in the range 200 – 2000 MPa, (1.17 – 13.61)  $\sigma_{uts}$  and -0.850 to -0.333 respectively. The range for the strength accommodates most alloys known from copper to maraging steels, both the fatigue slope  $b$  and strength coefficient  $a$  cover all typical values [2] which are dependent on factors such as dimension, surface finish, type of loading and notch factor.

For every combination of spectra parameters, the corresponding time domain signal for the selected spectrum was generated using equation (17) [21],

$$x_n(t) = \sum_{k=1}^N [2G_x(f_k)\Delta f]^{1/2} [\cos(2\pi f t + \phi_{k,n})] \quad (17)$$

where  $n$  is the sample number,  $N$  is the number of discretisation of the spectrum (PSD), with  $f_k = (k - 0.5)\Delta f$  and  $\phi_{k,n}$  are mutually independent random phase angles distributed uniformly over the range 0 to  $2\pi$ . The maximum frequency considered for the fatigue data in the study was 200 Hz. The sampling frequency used varied from the corresponding Nyquist frequency of 400 Hz to 6.40 kHz. The higher sampling frequencies were considered in the light of recent findings [22]. Up to 5000 discretisation of the frequency range and 32000 time steps were considered. The  $x(t)$  obtained was scaled so that the highest peak or deepest valley lied within 5 to 83% of the ultimate tensile strength value. This scaling reduced the possibility of any of the time data leading to extremely low values of oscillation before final fatigue damage occurs. This process provided a pool of input output for the training of the ANN described in section 2.3. The signals were analysed using an in house ANN routine and as well as own rainflow counting routine. The analysis used the material fatigue properties to determine corresponding damage intensity (i.e. damage per second) for each of the signals generated.

### **3. Results**

This section presents a verification of the routines developed against results of fatigue analysis using a spectrum pattern that has been analysed by other researchers. The convergence and robustness characteristics of the pattern recognition artificial neural network method developed are then demonstrated. The predictions made using the method developed was compared against results obtained using existing frequency domain methods such as Dirlik's[1] and Petrucci and Zuccarello's [2] methods. Several spectrum shapes and an extensive range of material properties values were considered.

#### **3.1 Correspondence of frequency and time domains**

The aim of this section is to demonstrate the agreement of energy content in both the frequency and time domain data used because this is fundamental to right predictions using ANN. Although a systematic difference in the values could still lead to good prediction, this type of oversight was avoided in this work. As highlighted in section 2.1 the template spectral in Figure 1 were used to generate time domain data,  $x(t)$ , with the aim to cover a wide variety of spectral shapes as much as possible. It is necessary to scale the time domain data to match the randomly selected alloy ultimate tensile strength property. As highlighted in Section 2.4, this was to limit the possibility of virtual monotonic load leading to instantaneous failure being considered as a fatigue event. After scaling, the spectral shape of the scaled time domain data in the frequency domain was re-determined. Figure 4 shows six of such spectral shapes obtained and the corresponding time domain data that produced them are shown in Figure 5. The shapes of the template spectral can be seen to be preserved although the edges show some oscillation. The oscillation is inconsequential because the power spectral density functions of real fatigue loads is hardly ever smooth. Based on equation 2, Table 2 shows that the energy obtained from the spectral shapes used in the study and the corresponding time domain signals agree to within 0.5% of the spectral value. The power from the spectral shape

is the same as the spectral moment  $m_0$  and the power from the time domain is simply the mean of the square of the peaks and valleys of  $x(t)$ .

### **3.2 Comparison of ANN prediction with rainflow and other frequency domain methods**

The number of patterns used for the ANN training was 5000 for this analysis. The training of the network was carried out using 90% of the pattern data cases. The test of the ANN prediction was carried out using the remaining 10% of the pattern generated. Spectral patterns  $a$  to  $f$  which included the narrow band was first considered and then broad band study excluding it was then considered. The results obtained for the 500 test cases are plotted in Figure 6 and 7. Figure 6 contains the log – log plots of the results of damage predicted using ANN, Dirlik and Petrucci frequency domain methods against the time domain Rainflow counting prediction method. The use of log – log plot allows verification of correlation of data with the rainflow counting methods across the wide range of small damage values, about  $10^{-16}$  to relatively high values of 1. The  $R^2$  correlation value and the slope coefficient of the linear best fit lines are included in the plots to show the relative agreement of the methods with the rainflow counting method. The linear plots of the results are also shown in Figures 8 and 9 to aid the direct comparison of the results obtained eliminating the non-linear effect of the logarithmic scale.

As can be seen in the results in Figures 6 and 7, good correlation was obtained between ANN and the other frequency domain methods against the Rainflow counting method. The  $R^2$  correlation values ranged between 0.905 and 0.999 for the log log plots. The coefficient of linear fit which is supposed to be equal to 1.000 varied between 0.974 and 1.034 for the log log plots. The  $R^2$  correlation values ranged between 0.976 and 0.998 for the linear plots in Figures 8 and 9 and the coefficient of fit varied between 0.972 and 1.299.

The correlation and slope coefficients show the ANN predictions to perform better than Dirlik and Petrucci methods. Both Dirlik and Petrucci methods show significantly higher slope coefficients i.e. 1.293 and 1.17 compared to 0.974 to 0.995 for the ANN. This indicates that the damage predicted using the ANN is averagely closer to the Rainflow counting values. The higher values of the Dirlik and Petrucci methods suggest that they are very conservative prediction methods.

### 3.3 Sensitivity to alternate sampling

This case attempts to detect the sensitivity of accuracy of prediction on alternate sampling compared to the sampling used to train the network. This is a helpful process in order to ensure that the ANN training had not over fitted or ‘memorised’ data. Consequence of overfitting is normally reflected in poor prediction when different unseen data is fed into the network. In order to generate the sample data for the neural network verification, each of the material properties,  $\sigma_{uts}$ ,  $a$  and  $b$  and spectral templates and moments shown in Table 1 were re-sampled randomly but of course within the limits shown in the table. Results from four runs from such sampling are presented in Table 3. Each run consisted of ten batches of data and each batch had 100 independent samplings, i.e 100 of the time domain data  $x(t)$ . The results of the studies were assessed by using the Pearson product-moment correlation coefficient,  $R_{X,Y}$ , equation (18), the coefficient of linear fit,  $m$ , equation (19) and the normalised root mean square deviation,  $E_n$ , equation (20)

$$R_{X,Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (18)$$

$$m = \text{cov}(X, Y) / \text{var}(X) \quad (19)$$

$$E_n = E[(Y - X)^2]^{1/2} / (\hat{X} - \check{X}) \quad (20)$$

where  $X$  and  $Y$  represent rainflow and the predicted  $E(D)$  defined in equation (8) respectively and  $\mu$ ,  $\sigma$ ,  $E$ ,  $var$ , and  $cov$  denote statistical mean, standard deviation, expectation, variance and covariance respectively.  $\hat{X}$  and  $\check{X}$  are the maximum and minimum values of the rainflow damage values.  $R_{XY}$  and  $m$  values are shown Table 3(a) and (b) for two runs and actual damage values are compared in Tables 3(c) and (d) for some randomly selected spectrum. The error  $E_n$  were determined in Table 4 is the average normalised root mean square error for each batch of 100 runs.

### 3.4 Effect of training characteristics and performance on unseen spectral

The effect of training characteristics such as number of signals and fraction of data used to train the ANN on the coefficient of fit with rainflow counting result was studied. All the spectral patterns shown in Figure 1 were used in the study. The number of signals used were 100, 300, 500, 1000, 2000, 5000, 10000 and 20000. The fraction  $n_{tr}$  of data used for training was varied from 0.1 to 0.9 in steps of 0.1. The remaining fractions  $1 - n_{tr}$  were then used to test the ability of the ANN to predict fatigue damage. The results obtained are summarised in Table 5. It can be seen from the table, for number of cases varying from 100 to 500, that the coefficient of fit  $m$  varied widely especially for low fraction of data was used for training. The fit for the higher number of cases from 1000 and over varied less widely compared to the lower number cases. In order to verify the performance of ANN on unseen patterns, training was again conducted using only spectral types  $a$  to  $f$  in Figure 1. The performance of ANN based on this selection of spectral was tested on unseen spectral types  $g$  to  $m$ . The results of the wider study can be exemplified by the data shown in Table 6. The average coefficient of fit  $m$  is 1.0421 with a standard deviation of 0.1731. This figures were obtained by ignoring the contribution of the prediction based on 100 cases for training. This overall performance suggest that many unseen types of spectral may be within the interpolative capabilities of the ANN model.

### 3.5 Verification with published data.

The aim of this section was to use the routines developed in this work to solve a similar problem as has been carried out by other researchers. A two-peaked power spectral spectrum, illustrated as spectrum  $i$  in Figure 1, described in the time domain by the equation (20) was a simplified basis to compare the performance of Dirlik's frequency domain method against the prediction of the time domain rainflow method [14, 23]. The spectrum chosen has two frequency contents  $f_1$  and  $f_2$  which in this case are taken to be 10 and 100 Hz respectively and PSD values  $G_1(x)$  and  $G_2(x)$  were taken as 10000 and 2500  $MPa^2/Hz$  respectively. The stress time history is obtained by adding the contributions from two sine waves corresponding to each of the frequencies in the spectrum.

$$x(t) = A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t) \quad (20)$$

For the sine wave form shown in (18), the amplitude is obtained from  $\sqrt{2}$  times the root mean square value of the corresponding block. In this case for a unit block width of 1 Hz was considered. Therefore  $A_1 = \sqrt{2} \times \sqrt{10,000 \times 1} = 141.42 MPa$ , we obtain similarly  $A_2 = 70.71 MPa$ . By superposition the actual amplitude corresponding to  $f_1$  is  $141.42 + 70.71 = 211.13 MPa$ . The material considered by both [14] and [23] was SAE 1008 which has ultimate tensile strength of 1240 MPa and the corresponding fatigue material strength coefficient and exponents  $a$  and  $b$  in equation (9) are 1145 MPa and -0.18 respectively.

In this simplified analysis, the effect of the mean stress  $S_{mi}$  in equation (10) is not considered hence  $\alpha_i$  in equations 9 and 10 is 1. The longest lives possible under the stress amplitudes 212.30 and 70.71 MPa from equation (9) are  $N_1$  and  $N_2$  which are equal to  $1.17 \times 10^4$  and  $5.23 \times 10^6$  respectively. The number of cycles  $n_1$  and  $n_2$ , for 1 second of loading, corresponding



to frequencies  $f_1$  and  $f_2$  are 10 and 100 cycles respectively. Substituting these values of  $n_i$  and  $N_i$  into equation (8) gives a damage fraction  $E(D) = 8.75 \times 10^{-4}$ . This is an approximate rainflow counting solution. Our computer implementation of Dirlik's method gave a result of  $1.38 \times 10^{-3}$ . Accounting for the higher frequencies used in this work, our results are within 1% of those by [14]. Our computer implementation of rainflow counting and Miner's rule for the time domain signal gave a result of  $8.2 \times 10^{-4}$  and the artificial neural network trained in this work gave a result of  $8.80 \times 10^{-4}$ . These results are within 6% of the simplified calculations. This validation test along with other tests shown in the foregoing demonstrate that the method presented is capable of excellent results and closer to rainflow counting results than Dirlik's method predictions.

In order to broaden the verification of the performance of the ANN model, different combinations of frequency and the amplitude in equation (18) as well as material properties,  $\sigma_{uts}$ ,  $a$  and  $b$  were independently re-sampled to generate 100 set of time domain data. The corresponding *PSD* and spectral moments were obtained and material data were fed into the ANN to make predictions for fatigue damage. The results obtained are shown plotted on log-log and linear scales in Figure 10. Good correlation can be seen in the plots in 10(a) and (b). The plot of Vanmarcke bandwidth parameter for the 1000 data used in training the ANN network and 100 data that were used for testing against the log of the damage fraction is shown in 11 (a). The spread of the bandwidth for the 100 independent cases used for the verification is shown by the plot of the corresponding Vanmarcke bandwidth parameter against the log of damage fraction in Figure 11 (b). It can be seen from that a good breadth of the Vanmarcke's bandwidth parameter was tested and good fit was obtained between ANN and rainflow counting results as can be seen in Figure 10(b).

#### 4.0 Discussion

In general the approach presented produced excellent correlation between predicted and time domain rainflow – Miner’s rule damage target values. The correlation coefficient for predicted and target damage values for Dirlik [1] and ANN developed herein are very similar but the coefficient of fit for the ANN model is much closer to 1.00 than those obtained by Dirlik’s method. The correlation coefficient and coefficient of fit for Petrucci and Zucarello were not as good as those of either the ANN or Dirlik’s method. The ANN approach developed provides a new competitive avenue for random fatigue damage prediction.

The results of the sensitivity study showed that the coefficient of fit between ANN and rainflow counting results was stable for many independent sampling of data. Table 3 for example shows results of 20 batches of independent runs. Each batch contained 100 independently composed spectra. The coefficients did not deviate in any significant way from those obtained during the initial training and testing of the ANN. The table also show raw individual fatigue damage data in 3(c) and 3(d) and how these compare with the rainflow results. Excellent agreement can be seen in most of the cases and better agreement can be seen for the ANN than for the Dirlik and Petrucci and Novo models. When only broadband data is used for the ANN training less error in predictions can be seen in Table 4 compared to the other methods analysed.

The results from the effect of training characteristics, Table 5 can be seen to indicate that about 1000 spectral cases and above are needed to produce ANN models that correlate well with the rainflow counting results. The number of spectral type in this study was 12 as shown in Figure 1. The statistical nature of the ANN approach shows in the weights obtained in Table 5. The table shows that higher number of samples does not necessarily guarantee that perfect ANN weights will be obtained. It can be seen that as low as 0.70 and 0.80 coefficients of fit were obtained for 2000 and 20000 samples respectively. The corresponding  $n_{rr}$  values in

these two cases were 0.8 and 0.6 respectively. The fraction of training  $n_{tr}$  for good coefficient of fit,  $m$ , i.e. close to 1 can be as low as 0.1 as can be seen Table 5. It should be noted however that low number of samples and low  $n_{tr}$  generally represents low coverage of the random fatigue problem space. The observation made in the course of this work was that only ANN models that lead to good coefficient of fit, i.e as close as possible to 1 should be used. As indicated above the higher the value of  $n_{tr}$  the better. For example, for the 5000 case in Table 5,  $n_{tr} = 0.2$  produced  $m = 0.9951$  while  $n_{tr} = 0.8$  produced  $m = 0.9881$ . A test on 100 independent cases of spectral type  $h$  using ANN weights from these two cases produced  $m = 1.0180$  and  $1.0020$  respectively. Clearly, the latter is to be preferred.

It is helpful to highlight the limits of the analysis carried out in this paper. The spectral and material properties considered have been limited to those practically feasible values shown in Table 1. The limits can be however be changed to cover any application that may be found to fall out of these limit. An ANN model is usually only used to solve similar problems as those used for its training. The analysis carried out in Section 3.4 shows that some unseen spectral types are within the interpolative capability of the ANN procedure described herein.

## **5. Conclusions**

Artificial neural network approach has been presented as a new competitive method for the analysis of random fatigue loading problems. Excellent correlation and coefficients of fit were obtained when compared to rainflow – Miner’s rule method. The method yields better agreement with time domain results than notable frequency domain methods such as Dirlik and Petrucci and Zucarello methods. The results from this work shows that ANNs have significant potential that can be tapped in developing more accurate frequency domain fatigue solutions. The interpolative capability of ANN on unseen spectral types demonstrates that the method has potential for wide application as a random fatigue loading analysis method. Future work will consider non Gaussian data.

## **Acknowledgement**

The first author would like to acknowledge the Tan Chin Tuan Fellowship held at Nanyang Technological University (NTU), Singapore in 2015 which provided excellent opportunity to interact and initiate collaborative work with colleagues at the institution. Discussions with colleagues at NTU led to the expediting of activities on return to Oxford Brookes University to complete this work which had been on-going for some years. He would also like to acknowledge his host, Dr Sridhar Idapalapati (Assoc Prof), for making the visit both eventful and exciting.

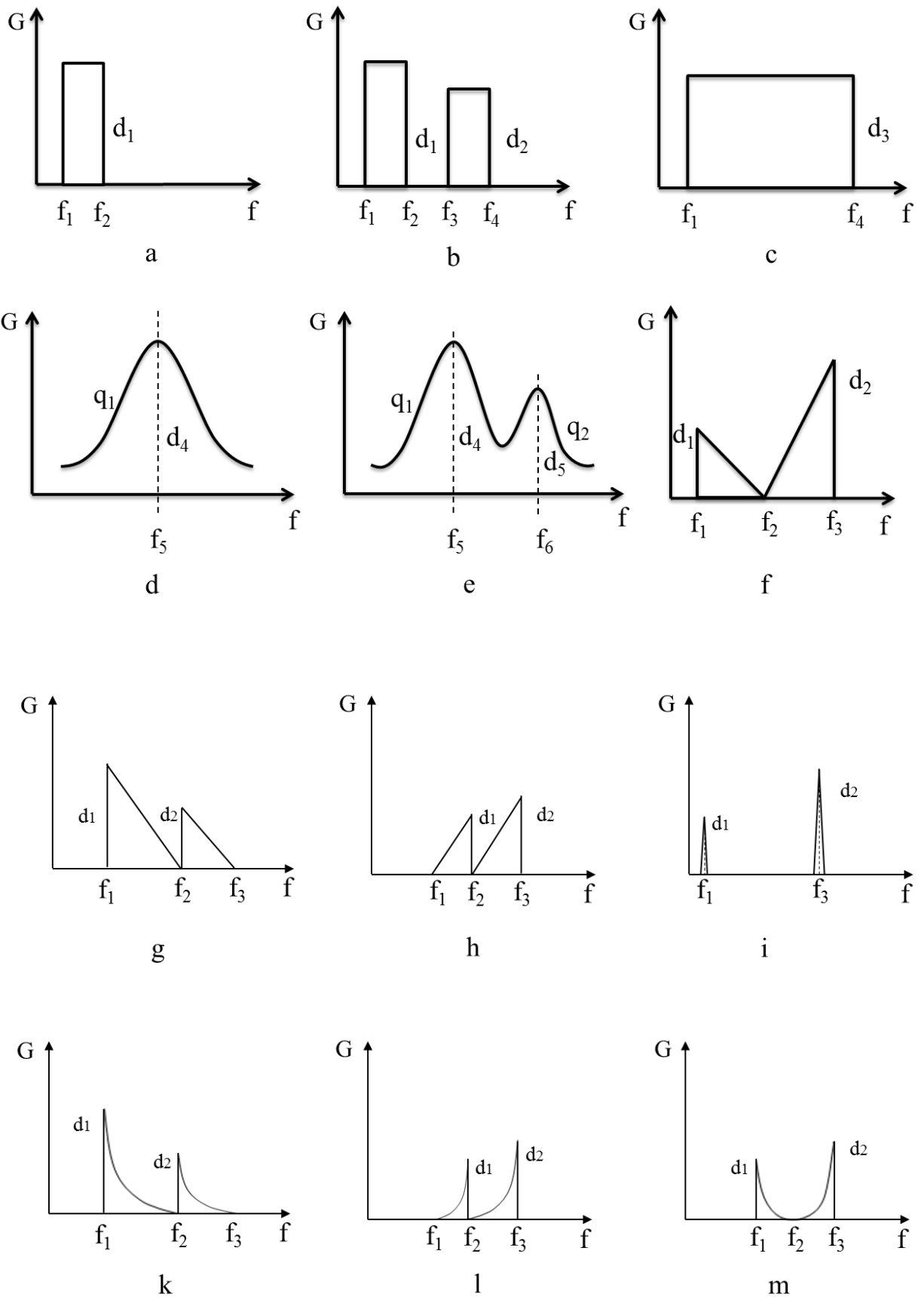


Figure 1

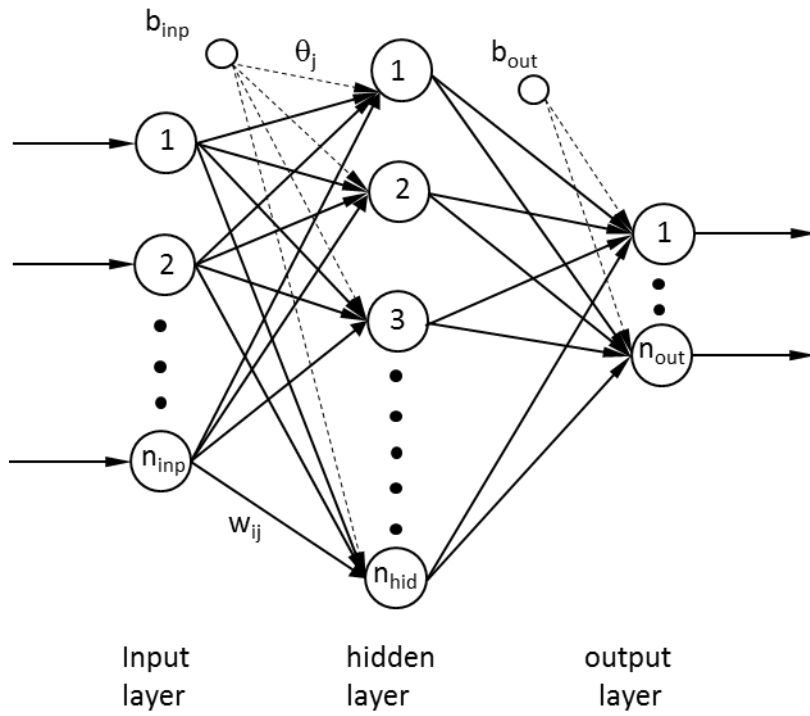


Figure 2

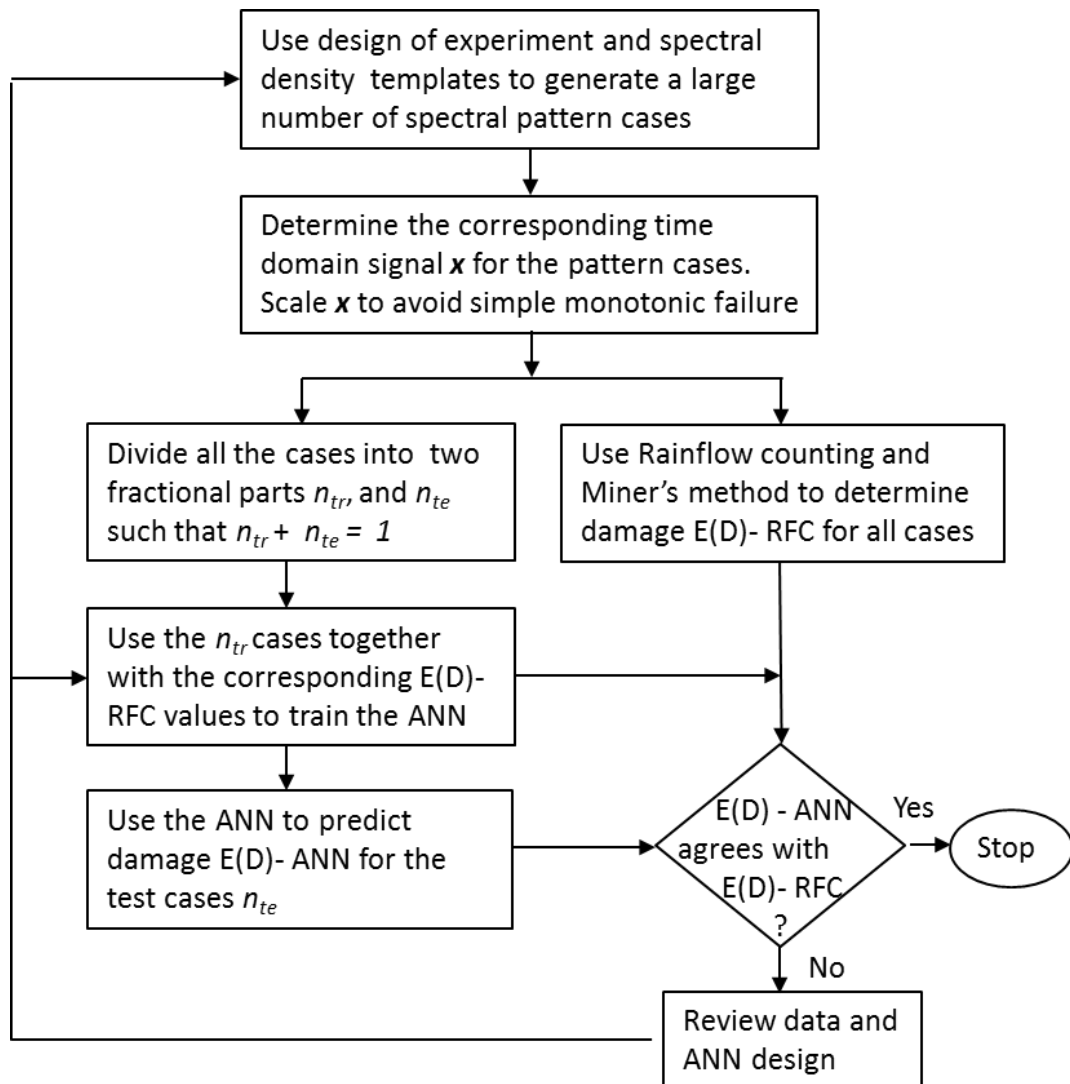


Figure 3

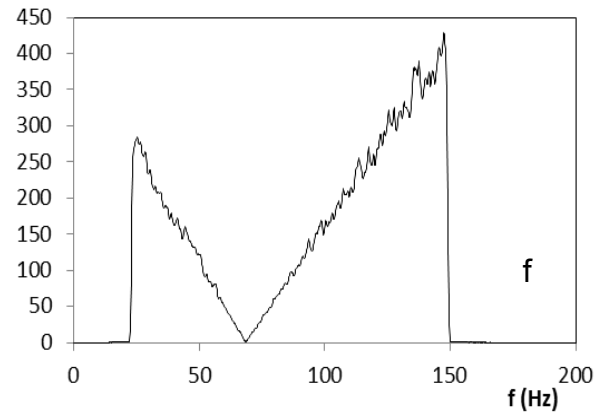
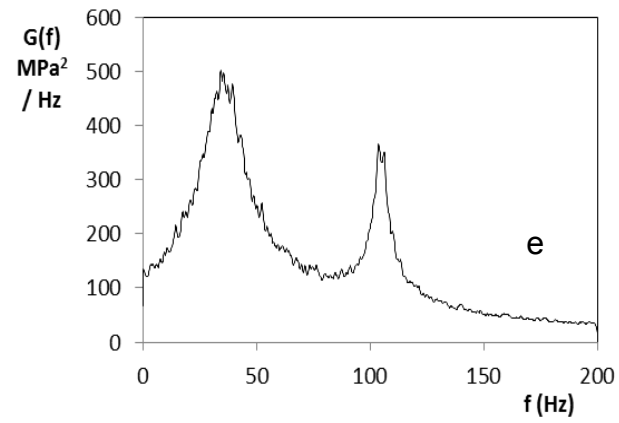
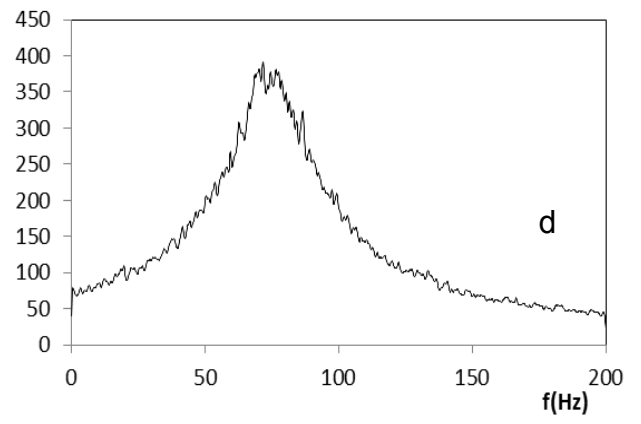
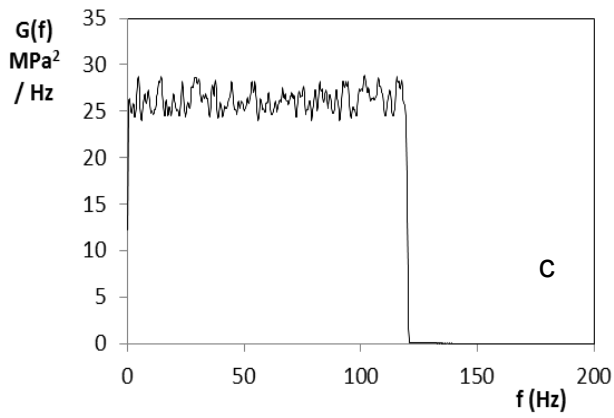
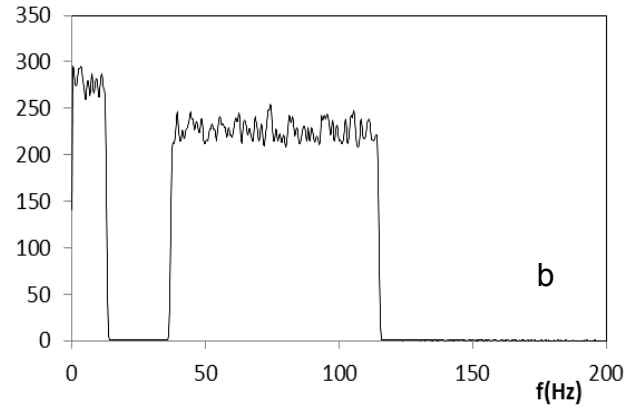
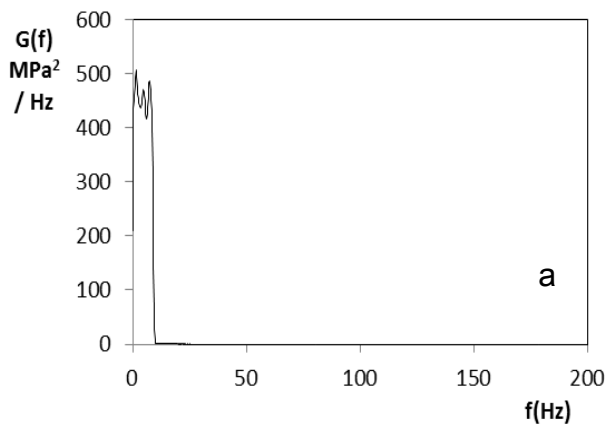


Figure 4

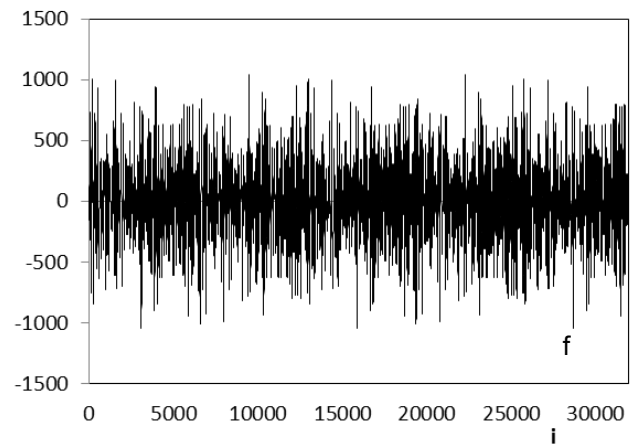
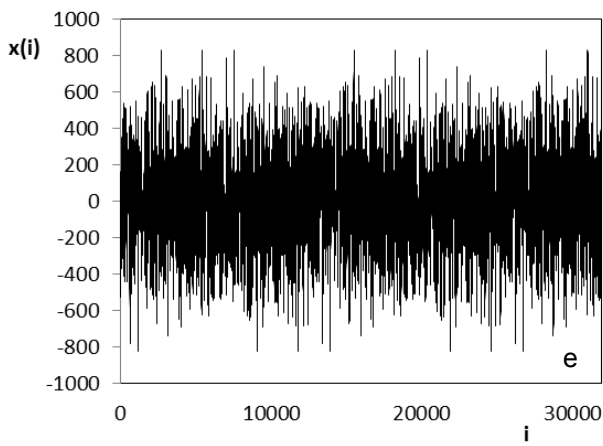
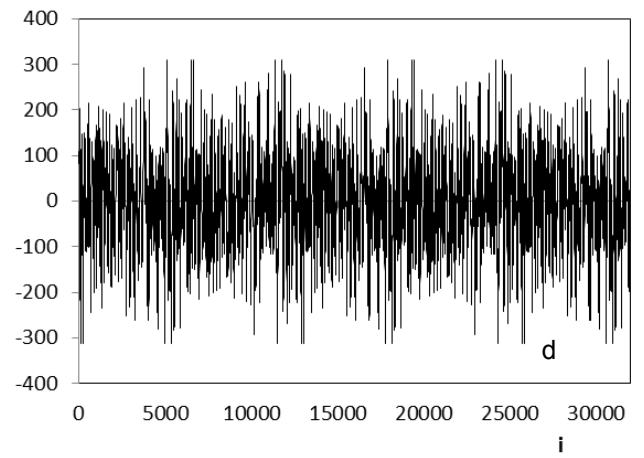
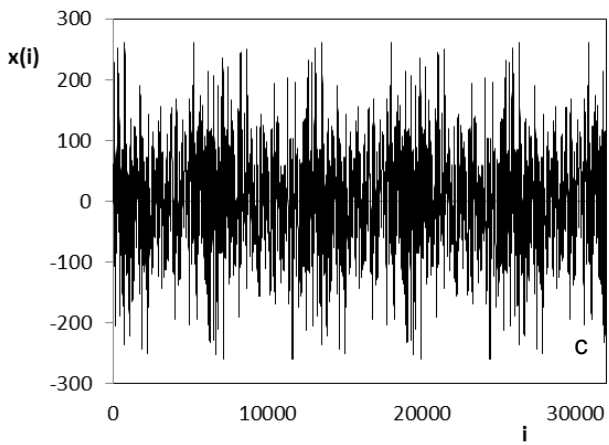
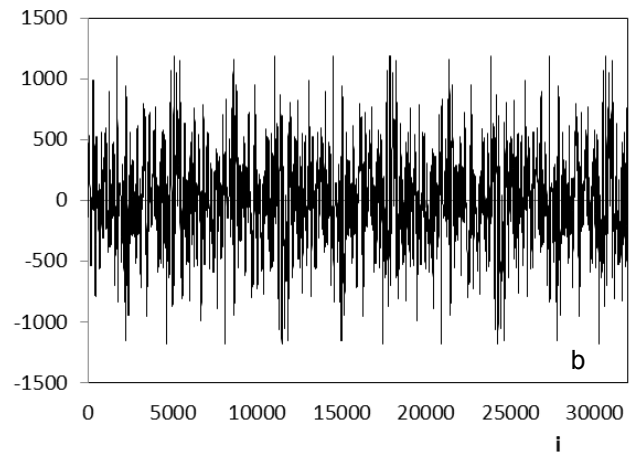
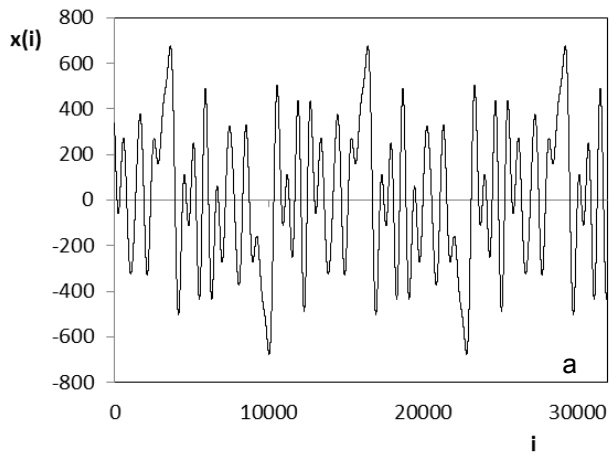


Figure 5



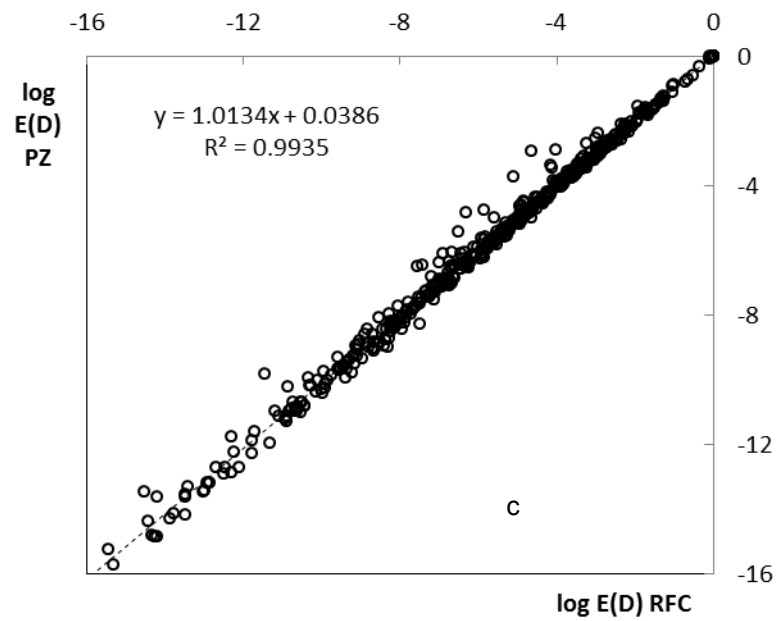
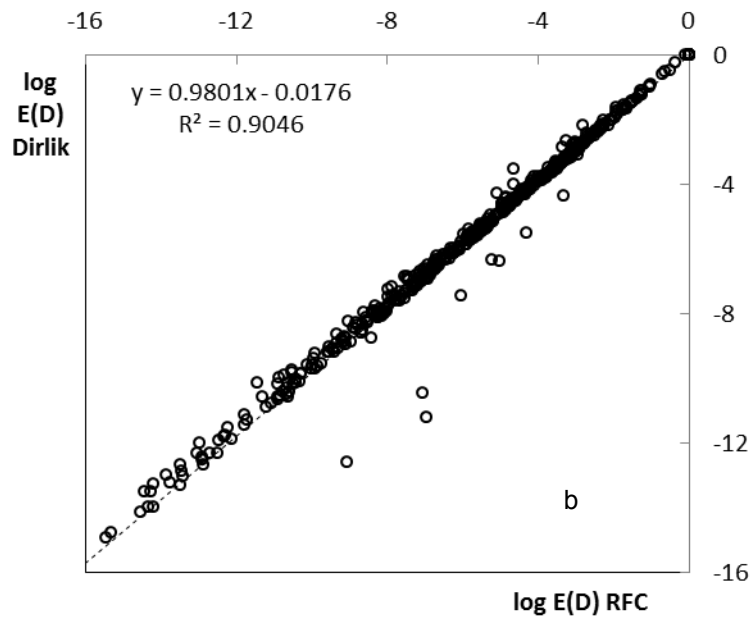
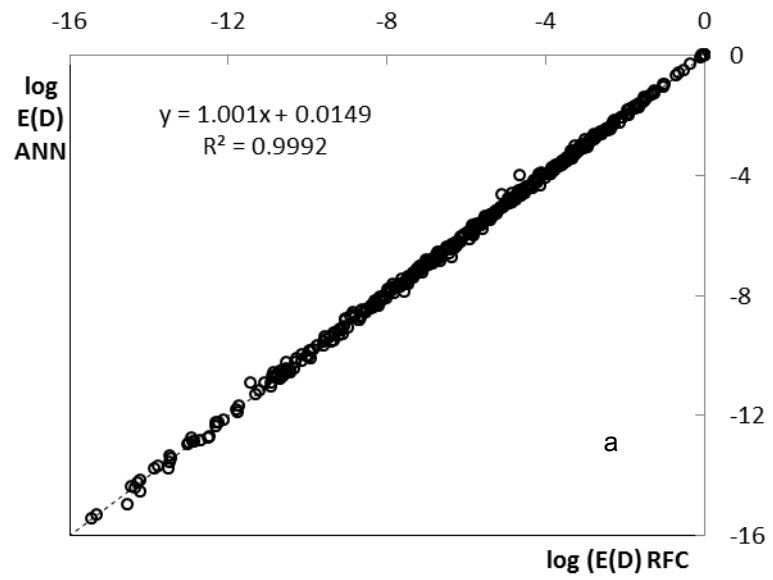


Figure 6

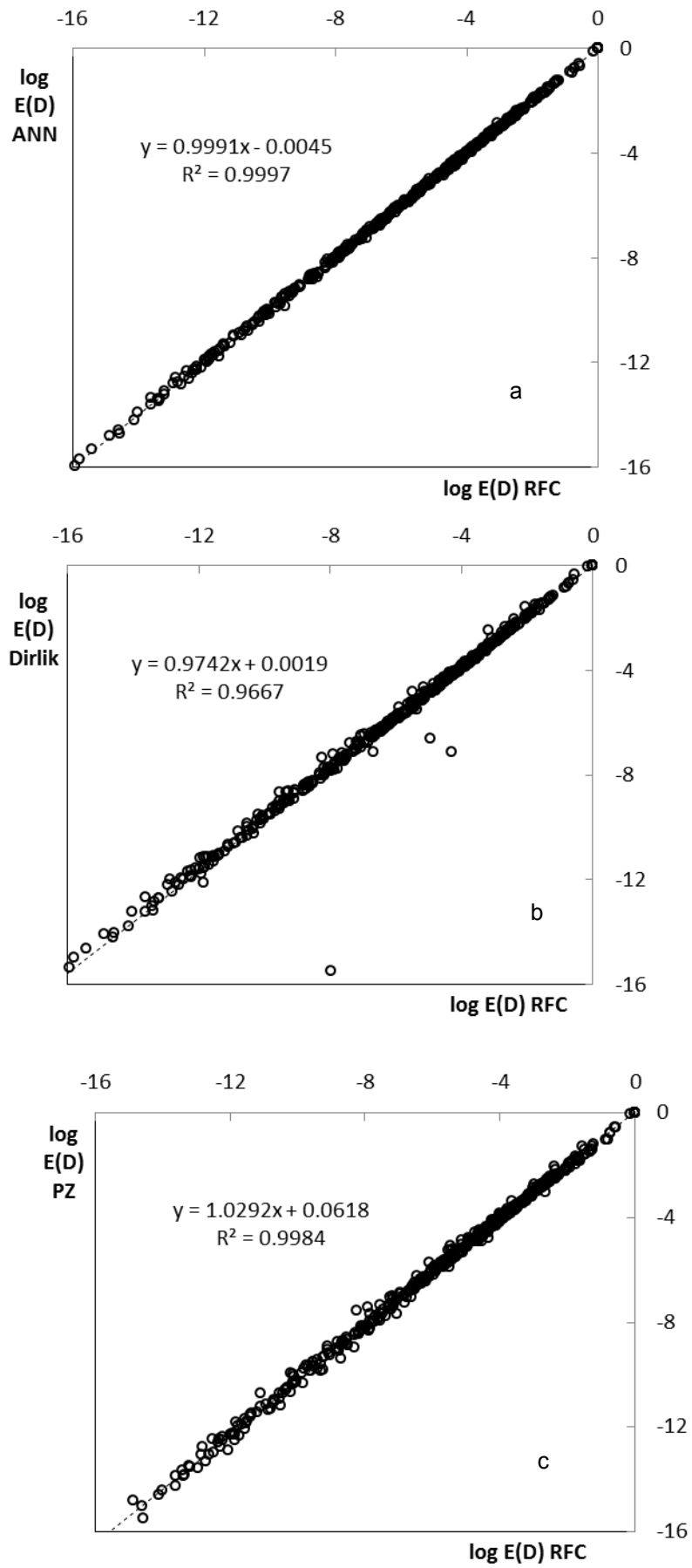


Figure 7

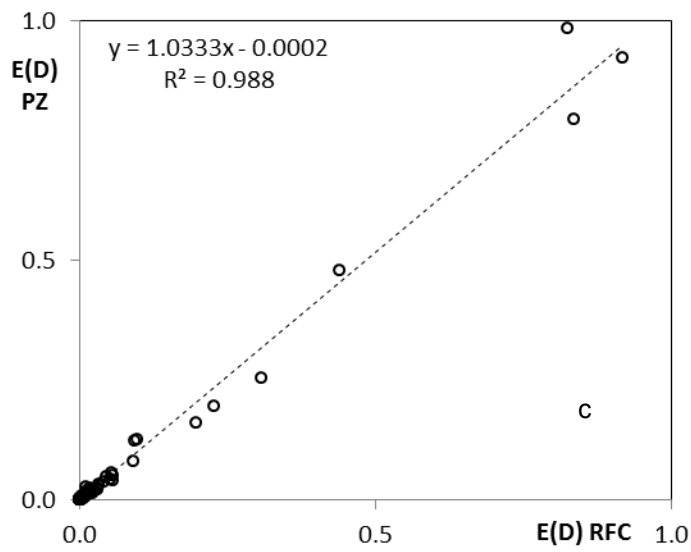
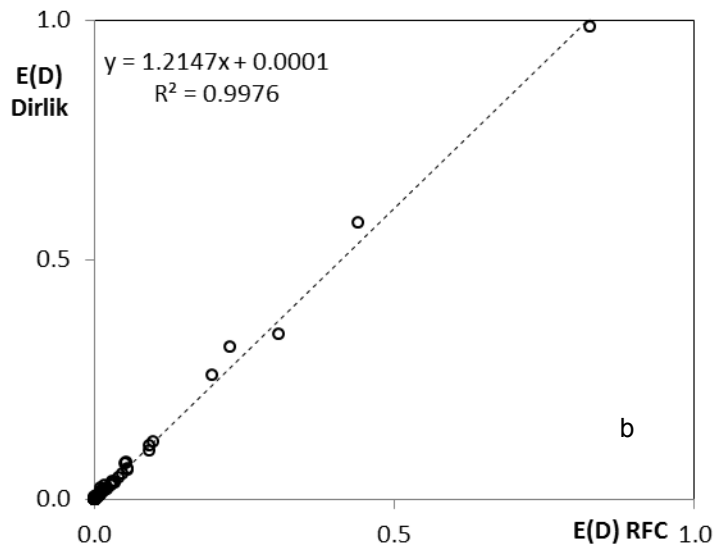
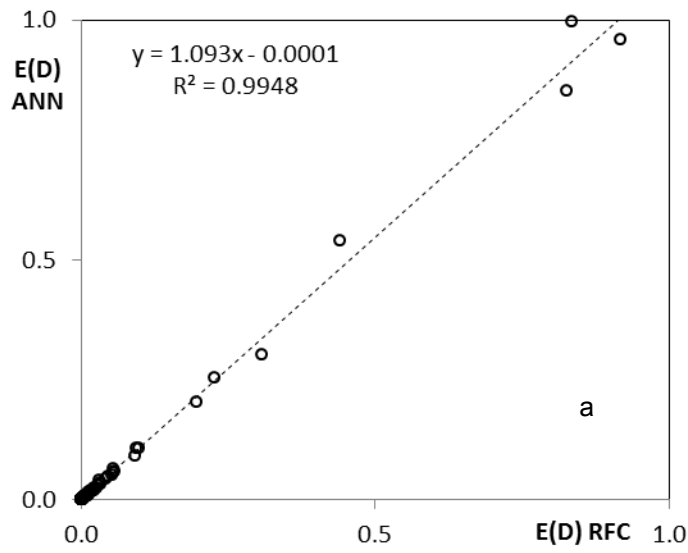


Figure 8

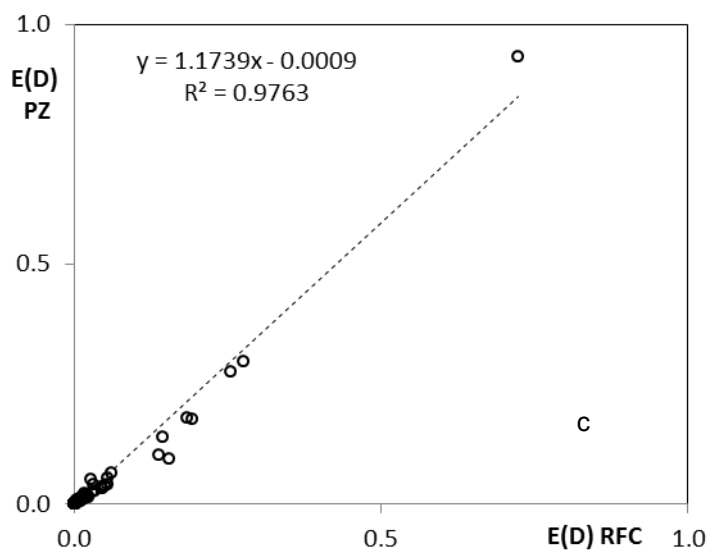
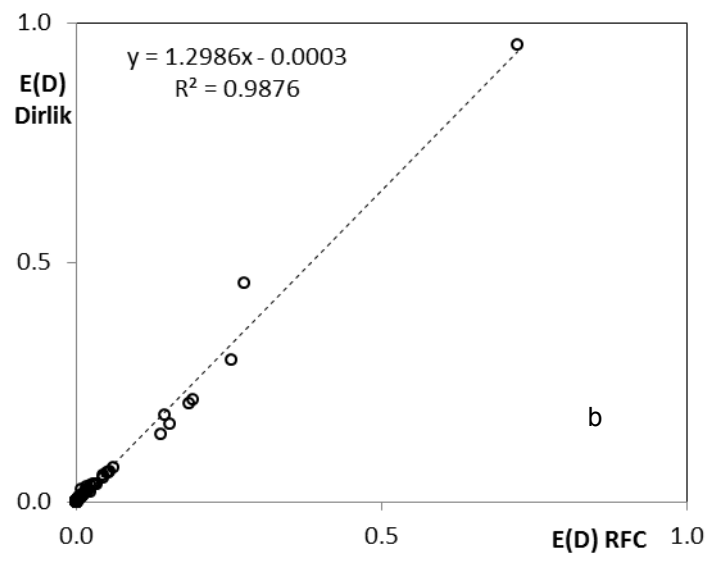
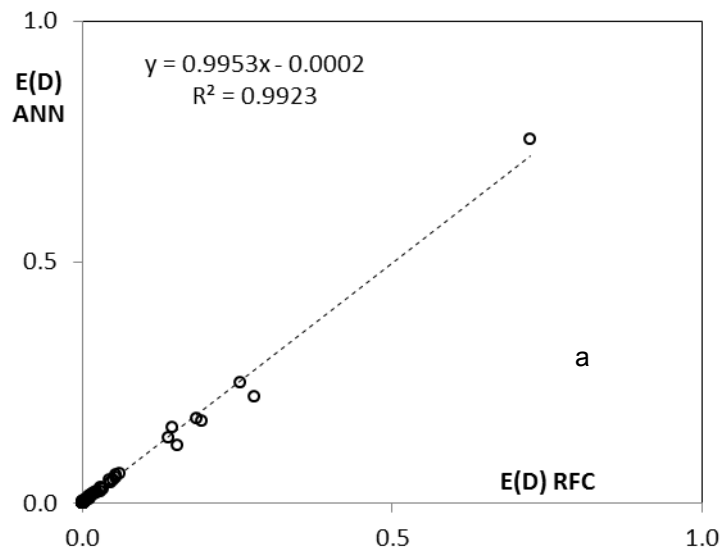


Figure 9

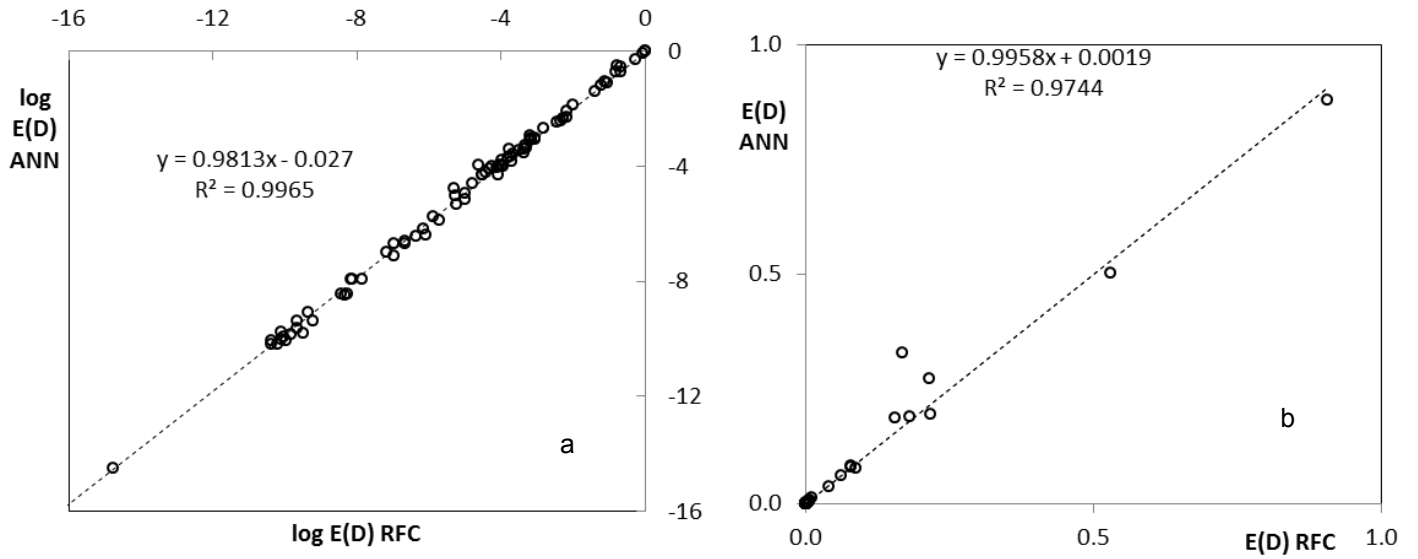


Figure 10

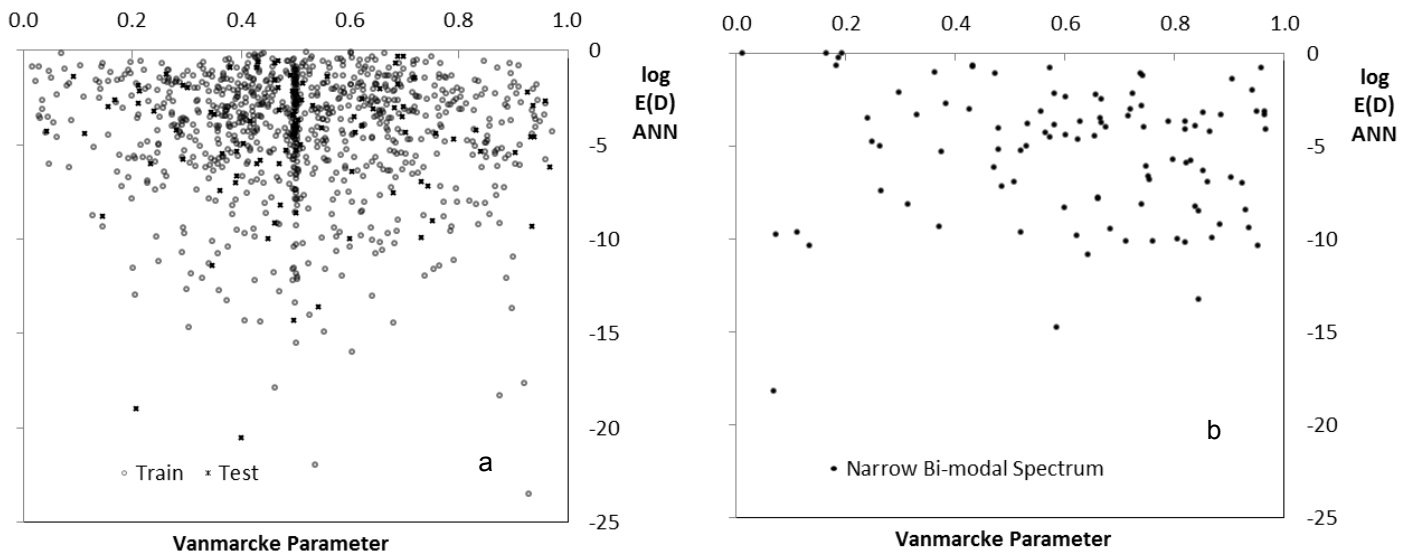


Figure 11

Prop	uts(MPa)	a	b	m <sub>0</sub>	m <sub>1</sub>	m <sub>2</sub>	m <sub>4</sub>
Min	200	261	-0.333	1.86E+01	6.71E+02	1.49E+04	2.85E+06
Max	2000	26543	-0.085	4.13E+05	3.89E+07	5.27E+09	1.07E+14

Table 1

Spec	uts(MPa)	m <sub>0</sub> (MPa <sup>2</sup> )	P <sub>x</sub> (Mpa <sup>2</sup> )	Diff(%)
1	1873	4107	4106	-0.024
2	935	21283	21281	-0.009
3	847	3144	3125	-0.604
4	701	28227	28218	-0.032
5	1054	30685	30825	0.456
6	1552	23022	22949	-0.317

Table 2

Batch	ANN		Dirlik		PZ	
	$R(x,Y)$	$m$	$R(x,Y)$	$m$	$R(x,Y)$	$m$
1	0.9921	0.8715	0.9999	1.0903	0.9916	0.9698
2	0.9925	1.0846	0.9908	1.2159	0.9463	1.1763
3	0.9995	1.1517	0.9993	1.2558	0.9989	1.0128
4	0.9994	1.1664	0.9979	1.1566	0.9985	0.9642
5	0.9989	1.0609	0.9972	1.1677	0.9689	1.2175
6	0.9988	0.9167	0.9967	1.1848	0.9714	1.0431
7	0.9990	0.9773	0.9990	1.1171	0.7918	0.8175
8	0.9918	0.9353	0.9892	1.0704	0.7150	1.3371
9	0.9860	0.9894	0.9925	1.1562	0.9523	0.8420
10	0.9985	0.9151	0.9994	1.0810	0.9982	0.7798
<b>Mean</b>	<b>0.9957</b>	<b>1.0069</b>	<b>0.9962</b>	<b>1.1496</b>	<b>0.9333</b>	<b>1.0160</b>
<b>StDev</b>	<b>0.0047</b>	<b>0.1036</b>	<b>0.0039</b>	<b>0.0604</b>	<b>0.0984</b>	<b>0.1827</b>

Table 3(a) Run 1 Mixture

Batch	ANN		Dirlik		PZ	
	$R(x,Y)$	$M$	$R(x,Y)$	$m$	$R(x,Y)$	$m$
1	0.9993	0.9682	0.9963	1.1196	0.9484	1.0526
2	0.9995	0.9628	0.9994	1.2062	0.9926	0.7952
3	0.9986	1.0551	0.9986	1.2253	0.9958	0.8296
4	0.9999	1.0373	0.9999	1.1702	0.9938	0.9512
5	0.9943	0.8997	0.9954	1.0690	0.9909	0.7308
6	0.9997	0.9100	0.9999	1.1423	0.9997	0.9755
7	0.9955	0.9790	0.9882	1.0971	0.9982	0.9593
8	0.9994	1.0858	0.9969	1.2557	0.9991	1.0085
9	0.9912	1.0400	0.9813	1.4129	0.9854	0.9699
10	0.9993	0.9725	0.9998	1.1670	0.9835	1.0177
<b>Mean</b>	<b>0.9977</b>	<b>0.9910</b>	<b>0.9956</b>	<b>1.1865</b>	<b>0.9887</b>	<b>0.9290</b>
<b>StDev</b>	<b>0.0030</b>	<b>0.0618</b>	<b>0.0061</b>	<b>0.0982</b>	<b>0.0152</b>	<b>0.1063</b>

Table 3 (b) Run 2 BroadBand

Spectrum / x(t)	uts (MPa)	a (MPa)	b	Rainflow	ANN	Dirlik	PZ
1	1569	5731	-0.190	9.57E-07	8.05E-07	1.21E-06	4.56E-07
2	1687	5868	-0.194	1.73E-02	1.34E-02	3.08E-02	3.73E-02
3	1049	6608	-0.199	3.98E-06	5.54E-06	6.89E-06	6.94E-06
4	1525	7004	-0.289	2.30E-03	2.34E-03	2.41E-03	1.80E-03
5	1115	5626	-0.116	1.44E-09	1.43E-09	3.98E-09	7.55E-10
6	1789	11484	-0.313	2.63E-02	3.37E-02	3.86E-02	3.93E-02
7	463	1820	-0.228	1.28E-04	5.31E-05	2.36E-04	2.87E-04
8	1386	16727	-0.179	1.43E-06	1.34E-06	1.65E-06	1.11E-06
9	313	3708	-0.204	1.04E-08	1.17E-08	1.32E-08	6.89E-09
10	1663	22620	-0.106	1.35E-12	8.82E-13	3.77E-12	9.56E-13

Table 3(c) Run 1 Batch 1 Mixture

Spectrum / x(t)	uts (MPa)	a (MPa)	b	Rainflow	ANN	Dirlik	PZ
1	1108	3767	-0.263	3.85E-02	3.96E-02	4.66E-02	6.10E-02
2	1505	4384	-0.201	6.38E-06	6.99E-06	9.66E-06	4.61E-06
3	859	2960	-0.280	2.58E-02	2.58E-02	3.00E-02	2.06E-02
4	317	3802	-0.207	5.72E-11	3.70E-11	6.66E-18	4.19E-11
5	454	1110	-0.098	9.49E-07	1.29E-06	5.63E-06	1.14E-06
6	721	2872	-0.197	2.44E-05	2.63E-05	3.75E-05	2.36E-05
7	1413	5478	-0.142	7.57E-06	4.61E-06	8.79E-06	3.20E-06
8	1591	10674	-0.230	1.52E-04	1.51E-04	1.91E-04	1.35E-04
9	564	6633	-0.329	6.48E-04	7.15E-04	7.79E-04	5.58E-04
10	932	7693	-0.170	8.18E-09	8.86E-09	1.48E-08	4.99E-09

Table 3(d) Run 2 Batch 2 Broadband



Batch	Mixed			Broad Band		
	ANN	Dirlik	PZ	ANN	Dirlik	PZ
1	0.0211	0.0709	0.0205	0.0064	0.0208	0.0473
2	0.0153	0.0711	0.1138	0.0051	0.0225	0.0241
3	0.0157	0.0348	0.0210	0.0092	0.0281	0.0222
4	0.0355	0.0262	0.0281	0.0046	0.0198	0.0133
5	0.0296	0.0458	0.0509	0.0143	0.0129	0.0296
6	0.0306	0.0240	0.0154	0.0094	0.0146	0.0035
7	0.0506	0.0127	0.0168	0.0102	0.0212	0.0076
8	0.0242	0.0790	0.0769	0.0110	0.0328	0.0051
9	0.0310	0.0958	0.0815	0.0245	0.0859	0.0290
10	0.0068	0.0254	0.0282	0.0073	0.0285	0.0308

Table 4

n <sub>tr</sub>	No of cases							
	100	300	500	1000	2000	5000	10000	20000
0.1	9.0836	0.4395	0.5749	1.0279	0.9762	0.9099	0.7485	1.0033
0.2	0.5483	0.7361	1.1393	1.4905	0.9154	0.9951	0.8604	1.0081
0.3	1.5634	1.0516	1.1471	0.8624	0.9302	0.8787	0.9529	0.9764
0.4	0.4277	1.2599	0.9772	0.9413	0.8291	0.9363	1.0249	0.9433
0.5	0.6744	0.7515	0.9762	0.9474	0.9278	0.9309	0.8326	0.9172
0.6	0.9197	0.5562	0.7993	0.8692	0.9249	0.8999	0.9181	0.7972
0.7	1.0172	0.7711	0.7952	0.9654	0.9337	0.9513	0.8565	0.9630
0.8	1.0713	0.6631	0.7058	1.1615	0.7035	0.9881	0.8719	0.8757
0.9	0.9900	0.9454	0.6552	1.0843	1.1001	0.8776	1.0033	0.9583
<b>Mean</b>	<b>1.8106</b>	<b>0.7972</b>	<b>0.8634</b>	<b>1.0389</b>	<b>0.9157</b>	<b>0.9298</b>	<b>0.8966</b>	<b>0.9381</b>
<b>St Dev</b>	<b>2.7477</b>	<b>0.2528</b>	<b>0.2069</b>	<b>0.1951</b>	<b>0.1066</b>	<b>0.0430</b>	<b>0.0873</b>	<b>0.0669</b>

Table 5

$n_{tr}$	No of cases					
	100	300	500	1000	2000	5000
0.1	2.1739	1.0011	0.7600	1.0091	1.3163	1.0889
0.2	0.5248	0.9277	1.1844	1.0218	1.1029	1.0907
0.3	0.4260	0.8821	1.0583	1.2605	1.0427	1.1058
0.4	0.8444	1.1010	0.8365	0.7375	1.0905	1.0727
0.5	0.4756	1.0526	1.7378	0.8055	1.1827	1.1347
0.6	0.4024	1.0251	1.2307	0.8035	1.0027	1.1026
0.7	0.4248	0.9938	1.1014	0.7804	1.0846	1.1100
0.8	0.9581	1.1135	0.9713	0.7785	0.9581	1.0519
0.9	0.3923	1.0218	1.0178	0.8396	1.1563	1.1494
<b>Mean</b>	<b>0.7358</b>	<b>1.0132</b>	<b>1.0998</b>	<b>0.8929</b>	<b>1.1041</b>	<b>1.1007</b>
<b>St Dev</b>	<b>0.5770</b>	<b>0.0746</b>	<b>0.2831</b>	<b>0.1709</b>	<b>0.1060</b>	<b>0.0297</b>

Table 6

### List of Tables

Table 1 Limits of property and spectral moments considered

Table 2 Demonstration of agreement between power values in frequency and time domains

Table 3 Product-moment correlation coefficients between target and predicted values ( $R_{X,Y}$ ) for different runs and batches of data

Table 4 Normalised root mean square error  $E_n$  for different batches of generated sample data

Table 5 Coefficient of fit for tests carried out using different number of cases and different training fractions  $n_{tr}$  for all spectral types  $a$  to  $m$

Table 6 Coefficient of fit for tests carried out using different number of cases and different training fractions  $n_{tr}$  for spectra type  $m$

## List of Figures

Figure 1 Illustration of spectral shapes considered in the study

Figure 2 Illustration of the structure of the artificial neural network used

Figure 3 Flowchart illustrating the process involved in the development of the ANN model

Figure 4 Spectral shapes obtained from the time domain data generated from the templates

Figure 5 Sample time domain data generated from the spectral templates

Figure 6 Comparison of log log plots of damage prediction against target values using a) ANN b) prediction using Dirlik model [1] and c) Petrucci and Zucarello model [2]. Narrow band signal included.

Figure 7 Comparison of log log plots of damage prediction against target values using a) ANN b) prediction using Dirlik model [1] and c) Petrucci and Zucarello [2] model. Broadband signals only were included.

Figure 8 Comparison of linear plots of damage prediction against target values using a) ANN b) prediction using Dirlik model [1] and c) Petrucci and Zucarello model [2]. Narrow band signal included.

Figure 9 Comparison of linear plots of damage prediction against target values a) ANN b) prediction using Dirlik model [1] and c) Petrucci and Zucarello model [2]. Broadband signals only were included.

Figure 10 Verification of ANN for 2 frequency bimodal signal (a) log log (b) linear plot

Figure 11 Scatter plot showing the spread of Vanmarcke's bandwidth parameter a) for training and testing the neural network b) for the 2 frequency bimodal signal used for verification

## References

1. Dirlik, T., *Application of computers in fatigue analysis*. 1985, University of Warwick.
2. Petrucci, G. and B. Zuccarello, *Fatigue life prediction under wide band random loading*. *Fatigue & Fracture of Engineering Materials & Structures*, 2004. **27**(12): p. 1183-1195.
3. Farley, S.J., et al., *High resolution non-destructive evaluation of defects using artificial neural networks and wavelets*. *Ndt & E International*, 2012. **52**: p. 69-75.
4. Farley, S.J., et al., *A Neural Network Approach for Locating Multiple Defects*, in *Advances in Experimental Mechanics Vi*, J.M. DulieuBarton, J.D. Lord, and R.J. Greene, Editors. 2008. p. 125-131.
5. Hernandez-Gomez, L.H., et al., *Locating defects using dynamic strain analysis and artificial neural networks*, in *Advances in Experimental Mechanics IV*, J.M. DulieuBarton and S. Quinn, Editors. 2005. p. 325-330.
6. Pujol, J.C.F. and J.M.A. Pinto, *A neural network approach to fatigue life prediction*. *International Journal of Fatigue*, 2011. **33**(3): p. 313-322.
7. Kim, Y., 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*, 2016. **111**: p. 310-322.
8. Benasciutti, D. and R. Tovo, *Spectral methods for lifetime prediction under wide-band stationary random processes*. *International Journal of Fatigue*, 2005. **27**(8): p. 867-877.
9. Zhao, W. and M.J. Baker, *On the probability density function of rainflow stress range for stationary Gaussian processes*. *International Journal of Fatigue*, 1992. **14**(2): p. 121-135.
10. Wirsching, P.H. and M.C. Light, *Fatigue under wide band random stresses*. *Journal of the Structural Division*, 1980. **106**(7): p. 1593-1607.
11. Newland, D., *An introduction to random vibrations, spectral and wavelet analysis*. 1993. Essex, England: Longman Scientific & Technical.
12. Tovo, R., *Cycle distribution and fatigue damage under broad-band random loading*. *International Journal of Fatigue*, 2002. **24**(11): p. 1137-1147.
13. Mckay, M., R. Beckman, and W. Conover, *A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code*. *Technometrics*, 2000: p. 55-61.
14. Bishop, N.W. and F. Sherratt, *Finite element based fatigue calculations*. 2000: NAFEMS.
15. Bishop, C.M., *Neural networks for pattern recognition*. 1995: Oxford university press.
16. Igel, C. and M. Hüsken. *Improving the Rprop learning algorithm*. in *Proceedings of the second international ICSC symposium on neural computation (NC 2000)*. 2000. Citeseer.
17. The MathWorks Inc(2010), *MATLAB The language of technical computing*. 2010: USA.
18. Heaton, J., *Introduction to neural networks with Java*. 2008: Heaton Research, Inc.
19. [www.statsoft.com](http://www.statsoft.com). *Neural Networks*. July 2016 [cited 2016 July].
20. Montgomery, D.C., *Design and analysis of experiments*. 2008: John Wiley & Sons.
21. Shinozuka, M., *Monte Carlo solution of structural dynamics*. *Computers & Structures*, 1972. **2**(5-6): p. 855-874.
22. Quigley, J.P., Y.-L. Lee, and L. Wang, *Review and Assessment of Frequency-Based Fatigue Damage Models*. *SAE International Journal of Materials and Manufacturing*, 2016. **9**(2016-01-0369).

23. Lee, Y.-L. and H.-T. Kang, *Vibration Fatigue Testing and Analysis*. Metal Fatigue Analysis Handbook: Practical Problem-solving Techniques for Computer-aided Engineering, 2011: p. 333.