Engineering Structures 176 (2018) 20–27

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct

A statistical model for lifespan prediction of large steel structures

Jan Kracík[a,b,*], Bohumír Strnadel[a]

[a] VŠB – Technical University of Ostrava, Center of Advanced Innovation Technologies, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic
[b] VŠB – Technical University of Ostrava, FEECS, Department of Applied Mathematics, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic

ARTICLE INFO

Keywords:
Fatigue limit
Steel structure
Fatigue life
Cumulative damage
Mixture model
Log-normal components
Stress amplitudes

ABSTRACT

The fatigue life of steel structures under operating conditions inevitably depends on various random factors. Among the most influential factors are the characteristics of load cycles, such as stress means and amplitudes. A knowledge of their probability distribution is thus crucial for fatigue life analysis and prediction. Finite probabilistic mixture models have previously been used for this purpose. This paper presents a study of the possible benefits of mixture models with log-normal components, using a large experimental data set from the slew bearing substructure of a stacker. The study shows that for this particular situation, the log-normal mixture model performs significantly better than Gaussian mixtures, and thus can be used as a suitable model in similar areas of application.

1. Introduction

Rapid crack propagation in large-scale structures (most frequently initiated by fatigue damage as a consequence of time-variable loading) is highly dangerous [1–3]. In most cases, rapid cracking not only leads to the disintegration of affected structural components; it also causes irreversible accompanying damage to the entire structure [1,2,4]. An essential area of research for engineering practice is the development of methods enabling us to monitor structural states and track the course of fatigue damage up to the end of a structure's operating lifespan [4]. By monitoring damage states, it is possible to manage the ageing of a structure, which brings substantial economic benefits.

However, it is difficult to estimate the degree of structural damage which leads to fracture instability. In most cases, it is not possible reliably to determine the distribution of damage throughout the entire structure or structural component (or to trace the development of this distribution over the course of time); one consequence of time-random loading is thus the random occurrence of failure limit states. Predicting the initiation of limit states in structures is an essential requirement for increasing operational safety.

Because structures are subjected randomly to variable loading during operations, the material fatigue characteristics (usually determined for constant stress amplitude and for fixed load ratio $R = S_{min}/S_{max}$ expressed as a S-N curve) must be transformed, where S_{min} is the minimum stress and S_{max} is the maximum stress. The most frequently used hypothesis for estimating the lifespan of randomly loaded structural components is Miner's rule of cumulative damage [5,6]

$$D = \sum_{i=1}^{m} \frac{N[i]}{N_f[i]} \tag{1}$$

where failure occurs when the damage D equals 1. The real value of the critical damage has been found much lower, $D = 0.5$ [7]. In the case of fluctuating mean stress conditions a further reduction down to a value of $D = 0.2$ may be applied [8]. Leitner et al. [9] revealed that damage sums between $D = 1.0$ and $D = 0.5$ are well applicable for a conservative fatigue assessment on the basis of the IIW-recommended fatigue design curves.

In Eq. (1), $N[i]$ is the real number of cycles at amplitude $S[i]$ and $N_f[i]$ is the number of cycles to failure at constant amplitude $S[i]$. Generally, the number of cycles to failure $N_f[i]$ (for any stress amplitude S) can be formulated as a Wöhler curve [10,11]

$$N_f = \alpha S^{-\beta}, \tag{2}$$

where α and β are parameters characterizing the material fatigue behaviour for a constant load ratio R.

The construction of a new S-N curve for various loading spectra – including the statistical uncertainty of parameters α and β in Eq. (2) – gives a more precise evaluation of the limit state; moreover, it enables us to distinguish between the degree of systematic error in the proposed model and random errors of the loading spectrum. A quick and effective tool for evaluating structural lifespan under variable stress amplitudes

* Corresponding author at: VŠB – Technical University of Ostrava, FEECS, Department of Applied Mathematics, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic.
E-mail address: jan.kracik@vsb.cz (J. Kracík).

https://doi.org/10.1016/j.engstruct.2018.08.065
Received 23 April 2018; Received in revised form 25 July 2018; Accepted 20 August 2018

appears to be the approximation of the S–N curve and its reliability band using a two-parameter Weibull function [12]. This solution enables us to objectively estimate the S–N curve and its variance without any other supplementary information – i.e. purely from the primary data. However, the distribution of stress amplitudes at selected critical structural nodes appears to be a key factor in the objective evaluation of lifespan in randomly loaded structures.

In the majority of large-scale steel structures, random loading always occurs within a certain range of stress amplitudes. These stress amplitudes at each node thus form a time series. Because the amplitudes of stress cycles are influenced by a number of operational and technological factors (often varying widely in character), the options for modelling this series – and particularly the consequences in terms of the reduction of the structure's load capacity – are very restricted. An example of modelling of factors influencing fatigue life can be found in [13] where probabilistic models of passenger volumes in trains on a railroad bridge are employed. For this reason, when dealing with real operating conditions it is only in very exceptional cases that the probability density function (PDF) of loading amplitudes can be derived theoretically; in most cases it must be estimated from experimental data. Because the loading process can, in most cases, be considered ergodic – and thus its statistical characteristics can be derived from records of a sufficiently long sample of historical data – the relations among the individual loading amplitudes can be discounted. However, the PDF of the amplitudes will certainly be a multi-modal one, sometimes with several local extremes. This is due to the different operating modes of the structure during its lifespan, which must be taken into account when modelling the PDF. The statistical model of the loading spectrum must be flexible enough to ensure that these variances are taken into consideration.

Finite mixture models [14] are precisely the type of statistical model that is suitable for the above-described conditions. Klemenc and Fajdiga [15–17] evaluated the statistical distribution of loading amplitudes by means of a mixture of multivariate Gaussian probability density functions. Unknown parameters of a generated distribution were estimated using the maximum likelihood method. This enables the rational extrapolation of the mixture PDF to regions for which experimental data are not available [15]. In order to increase the effectiveness of this method, an optimization of the number of Gaussian probability density functions in the mixture was proposed [16]. Later, it was demonstrated that when modelling loading spectra it is often more suitable to use Student's *t*-PDFs instead of Gaussian PDFs. The application of *t*-PDFs in the mixture is more general, and it produces a much better model of a loading spectrum with high amplitudes [17]. Another application of Gaussian mixtures for modelling of a loading spectrum can be found in [18]. The same idea developing mixture models has been proposed by Ye et al. [4] who formulated loading spectrum of a steel bridge by combining a standard traffic and weather spectrum with a spectrum obtained under typhoon conditions. More complex models of the loading spectrum – e.g. models taking into account the random nature of operating regimes in structural loading in relation to the estimation of the time course of damage and remaining lifespan – require knowledge obtained from larger data sets and sufficiently accurate estimations of parameters.

If the parameters of a complex statistical model are estimated from a small sample of data, even if the resulting fitted model describes the data well, the model can perform poorly during future observations. This effect is referred to as overfitting [19]. The risk of overfitting is even more acute when the amplitudes of the load cycles are not independent. In such cases, the estimated model can be well fitted to a time-local behaviour of the series, but it may poorly describe a long-term behaviour. If a sufficiently long sample of historical data is available, the parameters can be estimated relatively simply, e.g. using the well-known EM algorithm [20] and standard model selection techniques such as Akaike information criterion [21] or Bayesian information criterion [22]. Nevertheless, if no historical data are available, or if the sample is not long enough, the accuracy of the estimates must be taken into account, which makes the problem more complicated. Acquiring the data from operational processes can be a time-consuming and economically demanding procedure. It is therefore highly desirable to acquire as much information as possible from a limited-size sample. The uncertainty about unknown parameters can be systematically treated using the Bayesian approach [18,23].

Despite these undoubtedly positive results, current methods of modelling the PDF of loading amplitudes do not enable us to take into consideration those types of structural loading for which only a restricted set of operational data are available while maintaining a sufficiently accurate estimation of parameters, and for which there is also a lack of precise data on the relative numbers of cycles that occur as random variables in the loading spectrum. An adequate estimate of the PDF is thus an essential prerequisite for a reliable prediction of the fatigue life of the structure. This paper focuses on fatigue damage and lifespan prediction in large steel structures under real operating conditions. PDF of loading amplitudes is estimated using a mixture model with the minimum number of components. The new statistical model is applied in a selected case of a steel structure.

2. Mixture models of load cycle amplitudes

2.1. Damage as a random variable

Consider a situation in which the values of stress cycle amplitudes $S_1, \ldots, S_T$ are known. Miner's hypothesis, given in Eq. (1), for damage caused by T stress cycles, can equivalently be expressed in the form

$$D_T = \sum_{t=1}^{T} \frac{1}{N_f(S_t)}, \tag{3}$$

where $N_f(S)$ is the number of cycles to failure at a constant stress amplitude S. The aim is to predict damage

$$D_M = \sum_{t=T+1}^{T+M} \frac{1}{N_f(S_t)} \tag{4}$$

cumulated during the next M load cycles, or more generally to make decisions related to damage D_M. While damage D_T is (at least theoretically) fully determined by Eq. (3), the future damage D_M is a random variable. In order to predict or make decisions related to D_M, it is therefore necessary to know the probability distribution of D_M. For the sake of simplicity, we assume that the operating conditions vary in such a way that the amplitudes can be considered independent and identically distributed random variables with a probability density function $f(S)$. In view of the expected ergodicity, this condition does not greatly diminish the generality of the solution. The probability distribution of D_M is therefore fully determined by the probability density $f(S)$. In practice, the probability density function $f(S)$ is unknown and must be estimated from the observed amplitudes $S_1, \ldots, S_T$.

The uncertainty about D_M is thus caused both by the randomness of the future amplitudes $S_{T+1}, \ldots, S_{T+M}$ and also by our incomplete knowledge of the amplitude distribution, i.e. the use of an estimate of density $f(S)$ instead of the unknown real probability density function. When making decisions on D_M, the randomness of future amplitudes can be taken into consideration by applying standard tools from statistical decision-making theory [24].

In practice, the number of observed cycles T is usually much lower than M. It can therefore be expected that incomplete knowledge of $f(S)$ will be a substantially greater source of uncertainty than the randomness of future amplitudes. The uncertainty about $f(S)$ may be reduced by using suitable prior information – particularly information on the type of amplitude probability distribution.