CHALMERS | UNIVERSITY OF GOTHENBURG

PREPRINT 2014:18

Estimation of fatigue damage of steering components using vehicle independent load model

ROZA MAGHSOOD
IGOR RYCHLIK

Department of Mathematical Sciences
Division of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg Sweden 2014

Preprint 2014:18

Estimation of fatigue damage of steering components using vehicle independent load model

Roza Maghsood and Igor Rychlik

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg, Sweden
Gothenburg, August 2014

Preprint 2014:18
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2014

Estimation of fatigue damage of steering components using vehicle independent load model

Roza Maghsood and Igor Rychlik
Mathematical Sciences,
Chalmers University of Technology,
SE-412 96 Göteborg, Sweden.

Abstract

Maneuvers at slow speed, curves, and other "steering events" causing large lateral loads, are important for durability assessments of steering components of a vehicle. We focus on modeling of the extreme forces during these steering events with the aim of finding a simple but still accurate estimate of the fatigue damage caused by such loads. The sequence of steering events that forming a vehicle independent part of the load will be modeled by a Markov chain with transition probabilities estimated from CAN (Controller Area Network) bus data. The data is available for all vehicles. The extreme forces during the events, which are driver and vehicle dependent, are assumed to be statistically independent. Distributions of forces are estimated from dedicated field measurements. The main result is an explicit formula for the expected fatigue damage of the proposed model. Usefulness of the formula is validated using measured lateral accelerations and link rod forces.

Keywords: Fatigue damage index; hidden Markov models (HMMs); Markov chain; rainflow cycles; vehicle independent load models; steering events; onboard logging signals; link rod force; lateral acceleration.

1 Introduction

Fatigue is a process of material deterioration caused by variable stresses. For a vehicle stresses depend on external loads, e.g. road roughness, vehicle usage, drivers behavior, and on dynamical properties of the vehicle. Description of service loading independent of the properties of a vehicle is appreciated at design stage of components. Only sections of loads which cause large oscillations of stresses, often described by the range of rainflow cycles, are of interest.

In this paper we will present a model of loads related to the steering events such as driving on curves, slow speed maneuvers etc., which cause large oscillations of the forces acting on steering components. The model consists of two

parts; description of the sequence of steering events and the model for the extreme loads occurring during the events. The sequence of steering events will be modeled by means of a Markov chain. This is a vehicle independent part of the load. The extreme forces during the events are assumed to be statistically independent. Their distributions may depend on the type of steering event, e.g. (left, right) cornering, slow maneuver to the right or to the left etc. The parameters of the distributions are vehicle dependent and need to be estimated using dedicated measurement campaigns or laboratory tests.

For the proposed model an explicit formula for the expected damage is given. The expected damage will depend on frequencies of transitions between events and on the distributions of the extreme forces. In the examples given in Section 4, the Rayleigh distribution will be used to describe the variability of extreme forces. The proposed method is a generalization of the approach presented by Karlsson [15] for lateral loads caused by curves. In that model it was assumed that left and right turns occur independently of the past with probabilities 1/2. However in our studies we have observed that frequencies of left and right turns are usually not equal and that the turns directions (left, right) are not independent, see Maghsood and Johannesson [18, 19]. The proposed model will include the observed dependencies in order of steering events occurrences.

The paper is organized as follows. In Section 2 definition and some properties of rainflow method and fatigue damage calculations are reviewed. The proposed model for loads and means to calculate the expected damage for the model is described in Section 3. Measured data are used to validate models and to illustrate the results in Section 4. The paper closes with conclusions, acknowledgments, references and two appendices. The first appendix contains the proof of the formula for the expected damage while in the second one hidden Markov models (HMMs) based algorithm to detect the steering events is reviewed.

2 Fatigue damage index

Risk for fatigue failure in a component is often measured by means of a damage index. It is computed in the following two steps. First the rainflow ranges h_i, $i = 1, \ldots, N$, in a load x, say, are found, then the damage $D(x)$ is computed using Palmgren-Miner rule [24], [21], viz.

$$D_\beta(x) = \alpha \sum_{i=1}^{N} h_i^\beta, \tag{1}$$

where α and β are material dependent constants determined in constant amplitude tests. The parameter α^{-1} is equal to the predicted number of cycles with range one leading to fatigue failure. Various choices of the damage exponent β can be considered, e.g. $\beta = 3$ which is the standard value for the crack growth process, is often used. For comparison we consider also $\beta = 5$, that is often used when a fatigue process is dominated by the crack initiation phase. Finally, for simplicity only, we let $\alpha = 1$.

2.1 Definition of a rainflow cycle

The rainflow cycle count algorithm is one of the most commonly used methods to count cycles. The method was first proposed by Matsuishi and Endo [20]. Here, we shall use the definition given in [26] which is more suitable for statistical analysis of damage index. The definition is given next for completeness of the presentation.

Assume that a load x has N local maxima. Denote by M_i the height of i^{th} local maximum and by m_i the height of the preceding minimum. For each local maximum M_i, one should find the lowest value in forward direction (m_i^+) until the load exceeds or equal M_i, and the lowest value in backward direction (m_i^-) until the load exceeds M_i. The maximum value of m_i^+ and m_i^- is the rainflow minimum which is denoted by m_i^{rfc}. Then (m_i^{rfc}, M_i) is the i^{th} rainflow pair with the rainflow range $h_i = M_i - m_i^{rfc}$. Figure 1 illustrates the definition of the rainflow cycles.

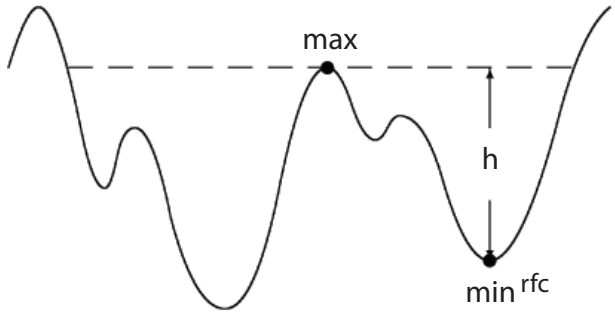

Figure 1: The rainflow cycle.

Note that some local maxima M_i can not be paired with any of local minima in x. (It can happen that the corresponding rainflow minimum m_i^{rfc} lies before or after the period when load was measured.) The sequence of maxima and minima which could not be paired by means of rainflow method is called the residual and has to be handled separately. Here we let maxima in the residual form cycles with the preceding minima (in the residual).

2.2 Rainflow counting distribution and crossings of intervals

In this paper we assume that a measured load x is given in form of time series x_i, $i = 0, \ldots, n$. In reliability applications the variability of the load is often modeled by means of random processes then x is one of many possible realizations of the process. For random loads the rainflow ranges become random

variables and the damage index is a random quantity too. The expected value of the index is an important parameter describing the severity of the load environment. Finding the expected damage is not a simple task and hence there are many approximations proposed in the literature. Maybe the most famous is the so-called narrow-band approximation, proposed by Bendat [2]. It can be proved that the approximation is actually a bound for the damage index defined in Eq. (1), see [29]. The proof is based on the alternative way to compute the damage index which will be presented next.

Suppose that in load x the following rainflow cycles were found (m_i^{rfc}, M_i), $i = 1, \ldots, N$. Variability of the cycles can be described using a cumulative histogram $N^{rfc}(u, v)$, say, called the rainflow counting distribution. The distribution is defined as follows; for any levels $u \leq v$, $N^{rfc}(u, v)$ is equal to the number of rainflow cycles such that $m_i^{rfc} < u \leq v < M_i$. As it was shown in [28] the damage index in Eq. (1) can be evaluated using the distribution, viz.

$$D_\beta(x) = \beta(\beta - 1) \int_{-\infty}^{+\infty} \int_{-\infty}^{v} (v - u)^{\beta - 2} N^{rfc}(u, v)\, du\, dv, \quad \beta > 1. \tag{2}$$

Remark 1. The proof of Eq. (2) follows the following lines. For any $m \leq M$

$$(M - m)^\beta = \beta(\beta - 1) \int_m^M \int_m^v (v - u)^{\beta - 2}\, du\, dv. \tag{3}$$

(For $\beta = 3$ the integral gives the volume of triangular pyramid having area of the base $(M - m)^2/2$ and height $(M - m)$.) Then an alternative formula for damage index (1) is obtained by replacing $h_i^\beta = (M_i - m_i^{rfc})^\beta$ by the double integral (3). Finally Eq. (2) is obtained by changing order of summations and integrations in the formula.

Remark 2. Particularly useful is the property that $N^{rfc}(u, v)$ is equal to the number of times x_i, $i = 1, \ldots, n$, crosses an interval $[u, v]$ in upward direction, see Eq. (A.2) for formal definition. The number of interval upcrossings will be denoted by $N_n^{osc}(u, v)$. The equality between the rainflow counting distribution and the interval crossing was shown independently in [28] and [4]. In the second paper multivalued loads were considered and oscillations between sets were used to define multiaxial rainflow count. For a uniaxial load the oscillations between sets $(-\infty, u)$ and $(v, +\infty)$ is equal to $N_n^{osc}(u, v)$.

Figure 2 illustrates the equality $N^{rfc}(u, v) = N_n^{osc}(u, v)$. It can be seen in the figure that there are three rainflow cycles with top above level 1 and bottom below -1 and there are also three upcrossings of the interval $[-1, 1]$ by the load.

Obviously it is easier to evaluate the damage index using Eq. (1) than by means of Eq. (2). However the formula in (2), with $N^{rfc}(u, v)$ replaced by $N_n^{osc}(u, v)$, viz.

$$D_\beta(x) = \beta(\beta - 1) \int_{-\infty}^{+\infty} \int_{-\infty}^{v} (v - u)^{\beta - 2} N_n^{osc}(u, v)\, du\, dv, \quad \beta > 1, \tag{4}$$

found applications in studies of damage index properties for random loads as will be shown in the following sections.

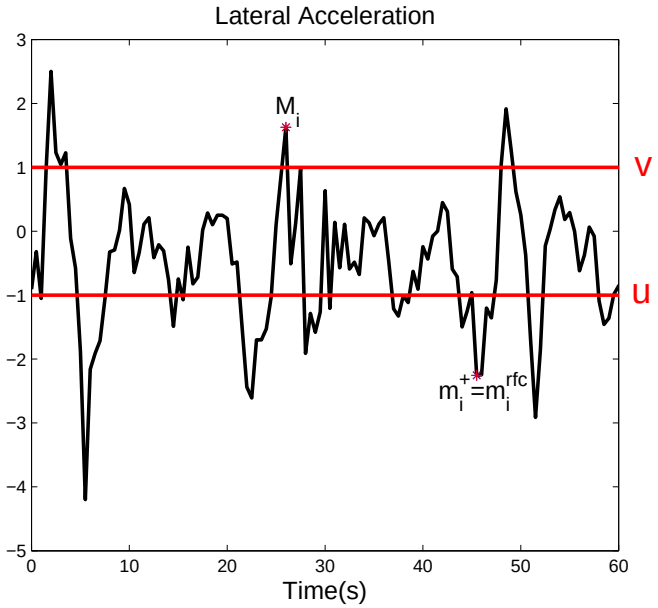

Figure 2: Illustration of crossings of an interval $[u, v]$, $N^{osc}(u, v) = 3$.

2.3 Random loads and damage intensity

Consider a random load $X = (X_0, \ldots, X_n)$. Due to randomness of X the damage index $D_\beta(X)$ will vary between the outcomes of X and becomes random itself. Then the damage intensity, i.e. the average growth of the expected damage index in time unit,

$$d_\beta = \lim_{n\to\infty} \frac{1}{n} E[D_\beta(X)] \tag{5}$$

is often used to measure the severity of the random load.

Here the damage intensity d_β will be computed using the intensity of interval crossings defined by

$$\mu^{osc}(u, v) = \lim_{n\to\infty} \frac{E\left[N_n^{osc}(u, v)\right]}{n}. \tag{6}$$

Now, Eq. (4) with $N_n^{osc}(u, v)$ replaced by $\mu^{osc}(u, v)$, gives

$$d_\beta = \beta(\beta - 1) \int_{-\infty}^{+\infty} \int_{-\infty}^{v} (v - u)^{\beta-2} \mu^{osc}(u, v)\, du\, dv. \tag{7}$$

In the following section we will use Markov chains to model variability of a load. Markov chains are flexible class of random processes and are often used to

model variability of environmental loads. Markov chains are convenient models for several reasons. Maybe the most useful property is that for Markov chains there exist efficient methods to evaluate μ^{osc} and hence the expected damage index, see e.g. [27], [16], [5], [10], [12], [11], [15], [17]. Markov chains can also be used to simulate rainflow filtered loads. More precisely often one wishes to simulate sequence of local maximums and minimums excluding cycles having ranges below some given threshold. This is important for accelerated testing of fatigue strength of components subjected to realistic loads, see [30] and [8] for an alternative approach. A related problem of extrapolating the rainflow counts, using oscillation intensity evaluated for compound Poisson process (a continuous time Markov process) has been considered in [14].

3 Random model of lateral loads based on steering events

Modeling external loads is an important area of transportation engineering as durability studies of vehicle components often require a customer or market specific load description. The most desired properties of the models are robustness and simplicity, so that only a small number of parameters is used to describe loads variability. Approach taken here is to approximate the load by a vehicle independent sequence of steering events, here representing Left and Right turns (LT, RT) or Left and Right steering (LS, RS). In both cases the two events are separated by a section when wheels have approximately zero turning angle, which is called Straight forward (SF). For each steering event only one, the most extreme, value of the load will be modeled by a random variable Y_i, say, and when vehicle is driving straight forward the load will take zero value.

The model is constructed in two steps; first the sequence of steering events is modeled as a Markov chain, then the distributions of extreme loads during events are proposed. It is assumed that the values of extreme loads are statistically independent.

As mentioned above a steering event has two states. The variability of the sequence is modeled by Markov chain Z_i, say, having two states "1" and "2". We denote by P the transition matrix of the chain, viz.

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \tag{8}$$

where p_{ij} denotes the transition probabilities between states. We assume that both p_{11} and p_{22} are smaller than one and hence the chain has the stationary distribution denoted by $\pi = (\pi_1, \pi_2)$. Further it is assumed that the chain is reversible, i.e. $\pi_i\, p_{ij} = \pi_j\, p_{ji}$, for all i, j, and hence one can construct a stationary sequence of steering events Z_i (defined for all integers) such that each Z_i has distribution π and

$$P(Z_i = j_0 | Z_{i-1} = j_1, Z_{i-2} = j_2, \ldots, Z_{i-n} = j_n) = P(Z_i = j_0 | Z_{i-1} = j_1).$$

We turn now to the definition of the extreme loads during steering events Z_i. Let M_i be a sequence of independent and identically distributed (iid) positive random variables while let m_i be iid negative random variables. Assume that the three sequences, Z, M and m are independent. Then the extreme loads Y_i, say, are defined by

$$Y_i = \begin{cases} M_i, & \text{if } Z_i = 1, \\ m_i, & \text{if } Z_i = 2. \end{cases} \tag{9}$$

Finally, the random load X_i is defined by adding zeros between Y_i and Y_{i+1} modeling the load when the vehicle is driving straight forward, viz.

$$X_i = \begin{cases} 0, & \text{if } i \text{ is odd integer}, \\ Y_{i/2}, & \text{otherwise}. \end{cases} \tag{10}$$

Remark 3. The proposed random load X_i in Eq. (10) is defined using the variables Z_i, M_i and m_i. Here Z_i are vehicle independent while M_i and m_i depend on vehicle, driver and other similar factors. Further the sequence Y_i defined in Eq. (9) is stationary.

3.1 Oscillation intensity for the random load

In order to evaluate the interval upcrossing intensity $\mu^{osc}(u, v)$ the following two conditional probabilities are needed;

- $p_1(u, v)$ - the conditional probability that given $Z_0 = 1$, i.e. load has local maximum at time zero, the sequence Y_i, $i > 0$, will visit the set $(v, +\infty)$ before it visits $(-\infty, u)$;
- $p_2(u, v)$ - the conditional probability that given $Z_0 = 2$, i.e. load has local minimum at time zero, the sequence Y_i, $i > 0$, will visit the set $(v, +\infty)$ before it visits $(-\infty, u)$.

It will be shown in Appendix A that for Y_i defined in Eq. (9) the probabilities satisfy the following equation system

$$p_j(u, v) = P(Y_1 > v | Z_0 = j) + \sum_{l=1}^{2} P(u \le Y_1 \le v | Z_1 = l) p_{jl} p_l(u, v), \quad j = 1, 2,$$

and more explicitly

$$\begin{aligned} p_1 &= p_{11} P(M_1 > v) + P(M_1 \le v)\, p_{11}\, p_1 + P(m_1 \ge u)\, p_{12}\, p_2, \\ p_2 &= p_{21} P(M_1 > v) + P(M_1 \le v)\, p_{21}\, p_1 + P(m_1 \ge u)\, p_{22}\, p_2. \end{aligned} \tag{11}$$

In the special case when Z_i are independent, i.e. the probabilities $p_{ij} = 1/2$ in (8), one has that $p_1(u, v) = p_2(u, v)$ and hence

$$p_2(u, v) = \frac{P(M_1 > v)}{P(M_1 > v) + P(m_1 < u)}. \tag{12}$$

We turn now to the main result of the paper an explicit formula for the interval upcrossing intensity by X_i given in the following theorem.

Theorem 4. *For X_i defined in Eq. (10), $\mu^{osc}(u,v)$ is given by:*

$$\mu^{osc}(u,v) = \frac{1}{2}\begin{cases} \pi_2 P(m_1 < u), & u < v < 0, \\ \pi_2\, P(m_1 < u)\, p_2(u,v), & u \le 0 \le v, \\ \pi_1 P(M_1 > v), & 0 < u < v, \end{cases} \tag{13}$$

where $p_2(u,v)$ is a solution to the equation system in (11).

Proof of the thoerem is given in Appendix A.

Example 5. The case of independent Z_i were considered by Karlsson [15]. In this special case $p_2(u,v)$ is given by Eq. (12). Since $\pi_1 = \pi_2 = 1/2$ the intensity of oscillations $\mu^{osc}(u,v)$ is given by

$$\mu^{osc}(u,v) = \frac{1}{4}\begin{cases} P(m_1 < u), & u < v < 0, \\ \frac{P(M_1>v)P(m_1<u)}{P(M_1>v)+P(m_1<u)}, & u \le 0 \le v, \\ P(M_1 > v), & 0 < u < v, \end{cases} \tag{14}$$

If M_i and $-m_i$ are exponentially distributed then Y_i, $i \ge 0$, has Laplace distribution, i.e. having pdf $f(y) = (1/2)\exp(-|y|)$, and

$$\mu^{osc}(u,v) = \frac{1}{4}\begin{cases} e^{u}, & u < v < 0, \\ \frac{e^{u-v}}{e^{u}+e^{-v}}, & u \le 0 \le v, \\ e^{-v}, & 0 < u < v, \end{cases}$$

In the following section $\mu^{osc}(u,v)$ will be used to estimate the expected damage index. More precisely $E[D_\beta(X)] \approx n \cdot d_\beta$ where d_β is computed using Eq. (7) with $\mu^{osc}(u,v)$ given in (13). Basically one is approximating $\mu_n^{osc}(u,v)$ by $n \cdot \mu^{osc}(u,v)$. (Note that $\mu_n^{osc}(u,v) \le n\,\mu^{osc}(u,v)$.) The numerical integration in (7) as well as the rainflow cycle counting have been done using the WAFO (Wave Analysis for Fatigue and Oceanography) toolbox, see [6, 32, 33], which can be downloaded free of charge.

4 Validations

In this section the random load model proposed in Section 3 will be validated. Two data sets will be used. The measured loads will be denoted by x^{obs}. First the steering maneuvers will be detected in x^{obs} using HMM algorithm presented in Appendix B. Then the extreme loads during maneuvers will be found. We assume that each maneuver follows by a driving straight section.

The signal consisting of the extreme loads during maneuvers and zeros for section when vehicle is driving straight will be denoted by $x = (x_0, \ldots, x_n)$

and called the *reduced load*. In Figure 3 part of measured load x^{obs} (lateral acceleration) is shown as the solid line while the reduced load x by dots.

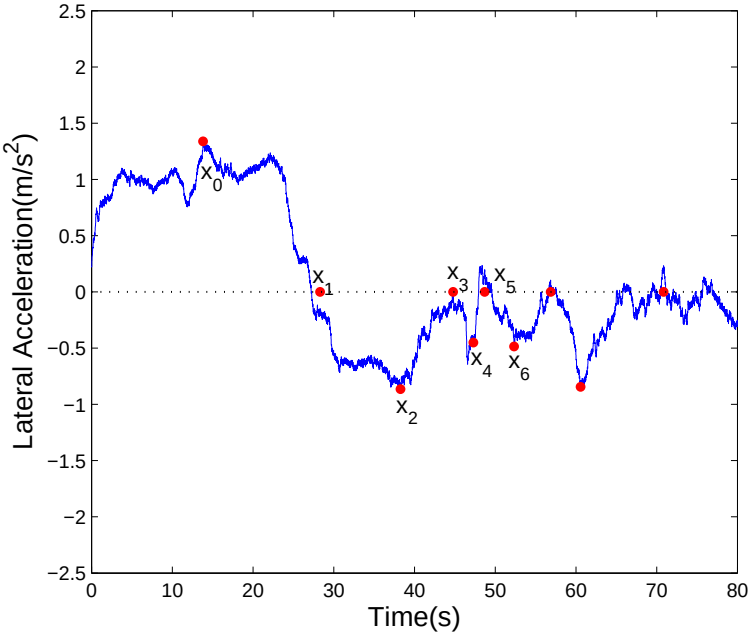

Figure 3: Reduced load x represented by dots compared with observed load x^{obs}, lateral acceleration, represented by the irregular solid line.

In Section 3 a random model for variability of the reduced load x was proposed, viz. $X = (X_0, \ldots, X_n)$, where X_i were defined in Eq. (10). Here the accuracy of the proposed model will be validated. Two issues will be considered:

(I) Firstly, we will investigate whether the reduced load x contains all large rainflow cycles that were found in the load x^{obs}. By this we control whether the assumption that load is zero when the vehicle is driving straight forward is not too crude and whether the HMM algorithm detects correctly the maneuvers.

(II) Secondly, we will investigate whether the random load $X = (X_0, \ldots, X_n)$, defined in Eq. (10), is accurately describing the variability of rainflow ranges counted in the reduced load x.

Investigation (I) consists of comparisons of rainflow cycle counts found in the measured load x^{obs} and in the reduced load x. Furthermore the damage indexes $D_\beta(x^{obs})$, $D_\beta(x)$ and the expected damage index $E[D_\beta(X)]$, for $\beta = 3, 5$, will be evaluated and compared.

The problem (II) will be addressed by studying the variability of the damage index $D_\beta(X)$ and checking whether the index $D_\beta(x)$ does not differ significantly from samples of $D_\beta(X)$. In addition the rainflow range spectrum, see Eq. (17), found in x will be compared with the expected spectrum and with the simulated spectra, i.e. found in samples of X.

4.1 Maneuvering events

In this section we consider the maneuvering events, e.g. driving in or out of a parking lot, standing still but turning steering wheel. (For simplicity of presentation we will not distinguished between driving forward or backward.) A common property of maneuvering is that the speed of vehicle is low. Here the limit is set to 10 km/h. Three maneuvering events are considered; Steering Left (SL), Steering Right (SR) and Straight Forward (SF). The HMMs algorithm, presented in Appendix B, has been applied to detect the maneuvering events from the steering angle speed.

In Figure 4(a), lower plot, the detected time periods when the vehicle was maneuvering are shown. The extreme forces are negative, positive and zero in the three states SR, SL and SF, respectively. There are 42 events detected leading to the following estimate of the transition matrix

$$P = \begin{pmatrix} 0.11 & 0.89 \\ 0.77 & 0.23 \end{pmatrix}. \tag{15}$$

The Rayleigh distributions have been fitted to positive and negative values of the reduced load giving

$$P(M_1 > v) = e^{-\frac{1}{2}\left(\frac{v}{6.2}\right)^2},\ v \geq 0, \qquad P(m_1 < u) = e^{-\frac{1}{2}\left(\frac{u}{6}\right)^2},\ u \leq 0. \tag{16}$$

The transition matrix P given in (15) and the Rayleigh distributions in (16) define the random load X represented in Eq. (10).

Comparison of rainflow counts in measured and reduced loads

The link rod force is used as the load x^{obs}. The load is shown in the top plot of Figure 4(a) as a solid irregular line. The force is not included in the CAN data and it has been separately measured. In the figure, the stars are the extreme link rod forces occurring during maneuvers and constituting the reduced load x. Rainflow cycles have been found both in the load and in the reduced load and are compared in Figure 4(b). The rainflow cycles found in the measured link rod force are marked as dots having coordinates (m^{rfc}, M). One can see that there are few large cycles and many small ones. The rainflow cycles found in the reduced load are presented as circles. As can be seen in Figure 4(b), all large cycles found in the link rod force are also found in the reduced load. Hence, one can expect that the damage index computed for the measured load should be very close to the one computed for the reduced load.

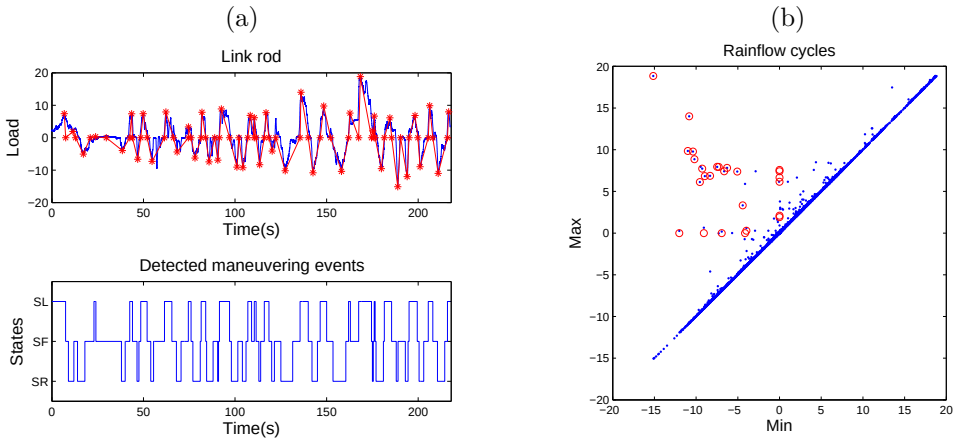

Figure 4: (a) *Top:* solid irregular line is x^{obs} (measured link rod force) while stars represent x (reduced load). *Bottom:* Detected maneuvers. (b) Dots - the rainflow cycles found in x^{obs}. Circles - the rainflow cycles counted in x.

Table 1 shows a comparison of the damage indexes $D_\beta(x^{obs})$ computed for the measured load, $D_\beta(x)$ for the reduced load and the expected damage index $E[D_\beta(X)]$ for the random model of the reduced load with the transition matrix P given in (15) and Rayleigh distributed M and m in (16). Damage indexes $D_\beta(x^{obs})$ and $D_\beta(x)$ are given in columns 2 and 3. As expected, these are almost identical. We conclude that the reduced load models the variability of the measured load well.

Table 1: Comparison of damage indexes $D_\beta(x^{obs})$ computed for the measured load, $D_\beta(x)$ for the reduced load and the expected damage index $E[D_\beta(X)]$.

	$D_\beta(x^{obs})$	$D_\beta(x)$	$E[D_\beta(X)]$
$\beta = 3$	$14.8 \cdot 10^3$	$14.3 \cdot 10^3$	$15.5 \cdot 10^3$
$\beta = 5$	$2.3 \cdot 10^6$	$2.2 \cdot 10^6$	$2.7 \cdot 10^6$

Validation of random load X

In this section we will investigate the problem (II), i.e. we will check if the variability of the reduced load x is well modeled by the random load X. The expected damage $E[D_\beta(X)]$ is given in the fourth column of Table 1. One can see that the observed damage is only 8% smaller than the expected damage evaluated from the model. This is a very small difference for this type of data. However we will further investigate the accuracy of the model by estimating the prediction intervals for $D_\beta(X)$ and comparing the load spectra for reduced load x and random model X.

The load x^{obs} is rather short, only 42 steering events have been detected, and hence statistical uncertainty of the model parameters is large. Neglecting the uncertainty of the model parameters, we have estimated the prediction intervals for the damage index using 10000 simulations of the random load X. The 95% prediction intervals for the damage index $D_\beta(X)$, $\beta = 3, 5$, have been estimated and the results are presented in Table 2. One can see that the variability of the damage index is huge. This demonstrate that the damage index is a rough tool to validate accuracy of the random model X. Consequently, we will use the rainflow range distributions to validate the model instead of the damage indexes.

Table 2: The approximative 95% prediction intervals for $D_\beta(X)$.

	Lower limit	Upper limit
$\beta = 3$	$8.9 \cdot 10^3$	$23.1 \cdot 10^3$
$\beta = 5$	$0.9 \cdot 10^6$	$5.4 \cdot 10^6$

Validation, in statistics, means determining whether a model fits the data well. Often cumulative distribution functions (cdf) and some goodness of fit tests are used. We will not follow this line since it is not taking into consideration the purpose of the model. It is well known that cycles with small ranges do not contribute much to the fatigue damage and hence their distribution not need to be accurately modeled. However these may heavily influence the shape of ranges cdf leading to rejection of practically "good" model. In engineering one often prefers to use the so-called load spectra to compare rainflow cycles distributions. The load spectrum is defined as follows.

Consider a load y having N rainflow cycles and the rainflow ranges with the cdf $F^{rfc}(h)$. Let H be random variable having cdf $F^{rfc}(h)$. Then, the damage index is

$$D_\beta(y) = N \cdot E[H^\beta] = \beta\, N \cdot \int_0^\infty P(h > h)\, h^{\beta-1}\, dh.$$

If for two loads the functions $N \cdot P(h > h) = N(1 - F^{rfc}(h))$ are close for high and moderate h, then for any $\beta > 1$ the damage indexes are close too. Traditionally, one defines the load spectrum $S(h)$, say, to be the inverse of function of $N(1 - F^{rfc}(h))$. Then, the plot of load spectrum $(h, S(h))$ coincides with the graph of the following line

$$\left(N\left(1 - F^{rfc}(h)\right), h\right), \qquad h \geq 0, \tag{17}$$

see [13] for more details.

Load spectrum in (17) was found in measured link rod force x^{obs} and reduced load x. The expected load spectrum was also evaluated by integrating $n \cdot \mu^{osc}(u, v)$ over suitable regions. The interval crossing intensity $\mu^{osc}(u, v)$, defined in (7), is given by (14). The spectra are shown in Figure 5(a). In the

figure one can see that the load spectra are very close to each other except in the region of small ranges where the load spectrum found in x^{obs} is positive while the remaining two spectra take value zero.

Finally in Figure 5(b) the load spectrum of the reduced load x is compared with 10 load spectra computed from simulated samples of the random load. In the figure, the smooth solid line is the expected load spectrum evaluated using $n \cdot \mu^{osc}(u, v)$, $n = 84$, while the thick stairs like line is the load spectrum found in the reduced load. It can be seen that the load spectrum of x does not differ significantly from the simulated load spectra and we can conclude that X is an accurate model for the variability of the reduced load x.

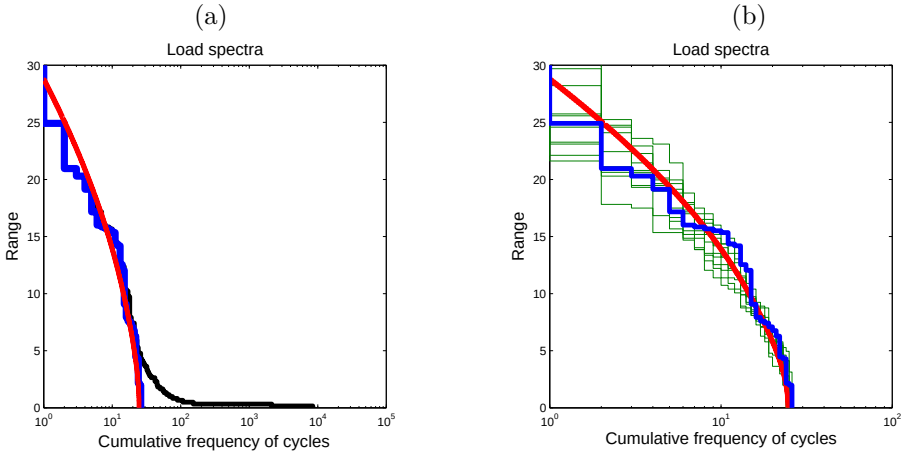

Figure 5: (a) Comparison of the load spectra (rainflow ranges). The regular solid line is the theoretical spectrum computed from $n \cdot \mu^{osc}(u, v)$, $n = 84$. The stairs like functions are the observed load spectra in measured load (link rod force) and the reduced load. (b) Comparison of load spectra found in simulated random load X, see (10), with theoretical load spectrum and the load spectrum of the reduced load (the thick stairs like line).

4.2 Driving through the curves

In this section we consider steering events occurring when vehicle is driving with speed higher than 10 km/h, e.g. when driving in curves. We limit analysis to three cases; Left turn (LT), Right turn (RT) and Straight forward (SF). The lateral acceleration signal and the HMMs algorithm are used to detect the curves. The signal is estimated using CAN supported data; vehicle speed and yaw rate. It has been computed by means of

$$x = \frac{\text{speed} \cdot \text{yaw rate}}{3.6}. \tag{18}$$

Here speed has units [km/h] while yaw rate [s^{-1}]. The signal is shown in Figure 6 (top plot). In the signal, 109 turns have been detected. The detected transitions

between the events are shown in Figure 6 (bottom plot). This signal was used to estimate the transition matrix P, (8), viz.

$$P = \begin{pmatrix} 0.41 & 0.59 \\ 0.19 & 0.81 \end{pmatrix}. \tag{19}$$

The lateral acceleration (18) is selected to be the load acting on a steering component. The reduce load x is shown in Figure 6 (middle plot). The Rayleigh distributions have been fitted to the positive and the negative values of the load giving

$$P(M_1 > v) = e^{-\frac{1}{2}\left(\frac{v}{0.6}\right)^2},\ v \geq 0, \qquad P(m_1 < u) = e^{-\frac{1}{2}\left(\frac{u}{0.4}\right)^2},\ u \leq 0. \tag{20}$$

The matrix P given in (19) and the parameters of Rayleigh distributions in (20) define together the random load X.

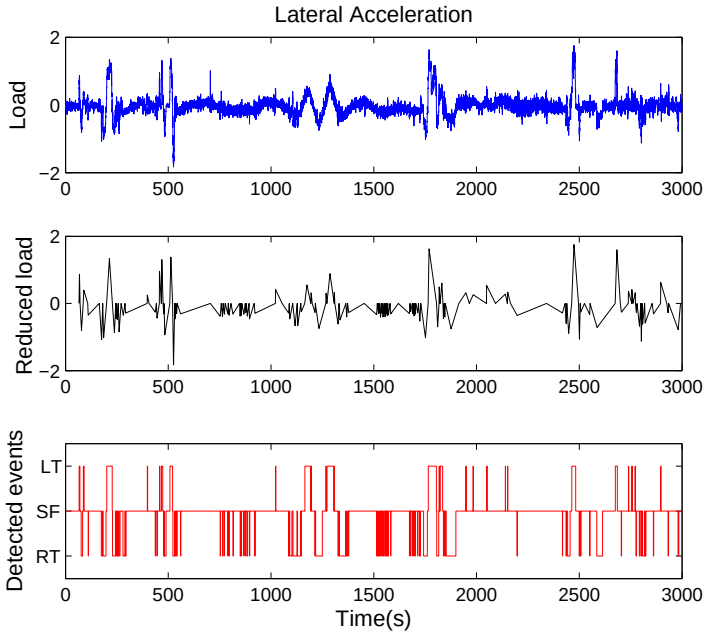

Figure 6: *Top:* Fifty minutes long record of the lateral acceleration (18). *Middle:* The reduced load defined by the extreme accelerations during left and right turns and zero when driving straight. *Bottom:* The corresponding detected curves from the lateral acceleration.

Comparison of rainflow counts in the measured and reduced load

The load (lateral acceleration) is shown in the top plot of Figure 6. In the figure, middle plot shows the extreme accelerations, occurring during curves,

constituting the reduced load. The rainflow cycles have been found both in the load and in the reduced load and compared in Figure 7(a). As before the rainflow cycles found in the load are marked as dots. The rainflow cycles found in the reduced load are presented as circles. One can see that the largest cycles are counted both in lateral acceleration and in the reduced load. However there are also many moderate size rainflow cycles found in the lateral acceleration which are missing in the reduced load rainflow count. We conclude that the largest cycles found in both the lateral acceleration signal and the reduced load are not of higher order of magnitude than cycles occurring during the event SF. And hence one can expect that the damage index computed for the lateral acceleration signal will be higher than the damage index estimated for the reduced load.

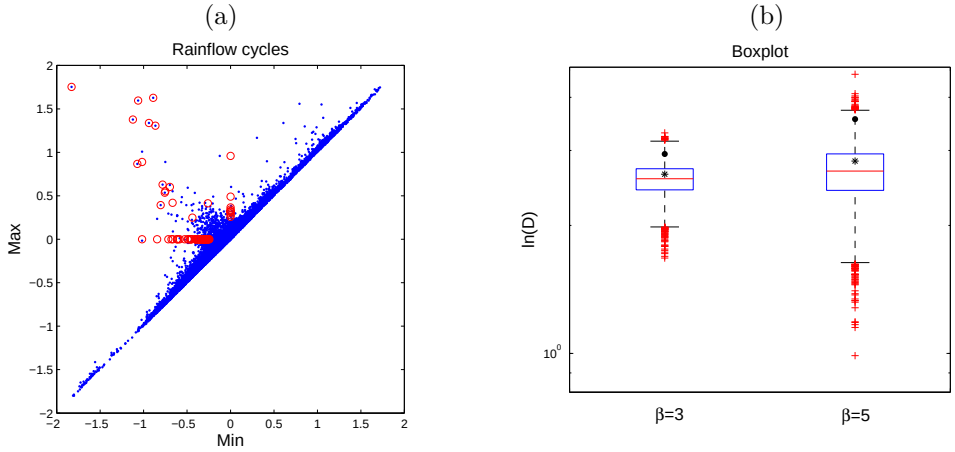

Figure 7: (a) Dots - the rainflow cycles found in the lateral acceleration signal. Circles - the rainflow cycles counted in the reduced load. (b) Boxplots of the simulated damage $D_\beta(X)$ for $\beta = 3, 5$. The star shows the calculated expected damage and the filled circle represents the damage index of the reduced load.

Damage indexes

The damage indexes evaluated for the load (lateral acceleration), reduced load and the expected damage index are compared in Table 3. Since the largest rainflow cycles are found both in lateral acceleration and reduced load, see Figure 7(a), the damage indexes $D_5(x^{obs})$ and $D_5(x)$ are almost identical. The large number of moderate size cycles found when the vehicle was driving straight forward are contributing to $D_3(x^{obs})$ and not to $D_3(x)$ and hence there is larger difference between values of thess two damage indices. The expected damage index $E[D_\beta(X)]$ is clearly smaller than $D_\beta(x)$. Whether this difference is significant will be investigated next.

Table 3: Comparison of damage indexes $D_\beta(x^{obs})$ computed for lateral acceleration signal shown in Figure 6, top plot, $D_\beta(x)$ reduced load, shown in Figure 6 middle plot, and the expected damage index $E[D_\beta(X)]$, for the random model for the reduced load with P given in (19) and Rayleigh distributed M, m (20).

	$D_\beta(x^{obs})$	$D_\beta(x)$	$E[D_\beta(X)]$
$\beta = 3$	25	19	14
$\beta = 5$	36	35	17

We have simulated 10000 sequences $X = (X_0, \ldots, X_n)$, $n = 218$, found rainflow cycles and evaluated the damage indexes. The variability of logarithms of the evaluated damage indexes $D_\beta(X)$ is presented in form of boxplots shown in Figure 7(b). The straight line in the middle of boxplot represents the median of damage. The lines which are extended vertically from the boxes indicating the whiskers. We have used the Tukey boxplot and the ends of the whiskers corresponds to the approximately 99.3% intervals if the logarithm of damage is normally distributed. As can be seen, the damage value calculated by the reduced load (shown by a filled circle) lies within the normal spread of the damage index $D_\beta(X)$, for $\beta = 3, 5$, and hence we conclude that the random load X can be used to model variability of reduced load x. This claim will be further supported by a study of load spectra variability presented in the following subsection.

Validation of random load X - load spectra

In Figure 8(a), the load spectra for the lateral acceleration, reduced load x and the random load X are compared. The observed load spectrum contains much more small and moderately high ranges than the remaining two spectra. Further the expected load spectrum, shown as a smooth line, is close to the load spectra of the reduced load.

In Figure 8(b), the expected load spectrum is compared with 10 load spectra computed from simulated samples of the random load. In the figure, the smooth solid line is the expected load spectrum evaluated using $n \cdot \mu^{osc}(u, v)$ while the thick stairs looking like line is the load spectrum found in the reduced load. Except cycles with very small ranges, one can see that the load spectrum of x does not differ significantly from the simulated load spectra.

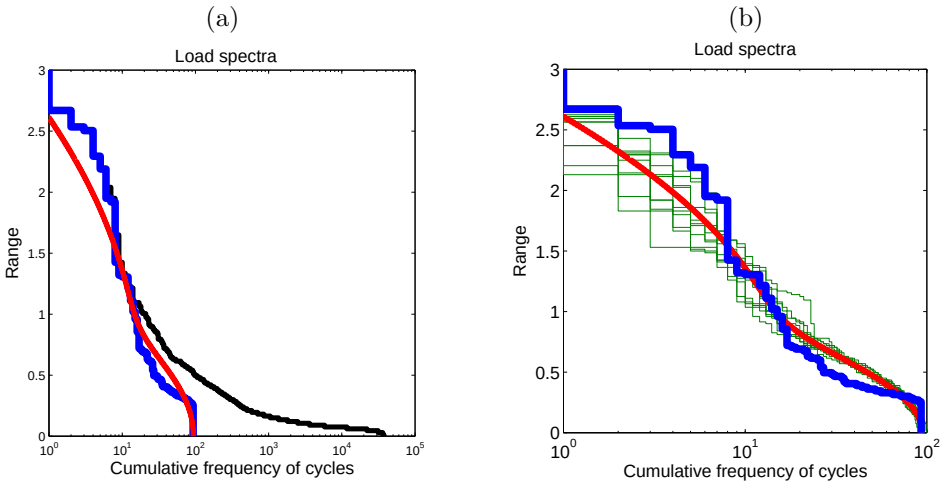

Figure 8: (a) Comparison of load spectra (rainflow ranges). The regular solid line is the theoretical spectrum computed from $n \cdot \mu^{osc}(u, v)$, $n = 109$. The stairs like functions are the observed load spectra in measured load (lateral acceleration (18)) and the reduced load (b). Comparison of load spectra found in simulations of random load X, see (10), with theoretical load spectrum and the load spectrum of the reduced load (the thick stairs like line).

5 Conclusion

A reduced load, i.e. a sequence of the most extreme forces during steering events, was introduced. A random load modeling the variability of the reduced load was proposed. The sequence of steering events, which is vehicle independent information, was modeled using a two states Markov chain. The extreme forces occurring during the steering events were modeled by means of independent Rayleigh distributed variables. For the model, an explicit formula for the expected fatigue damage was presented. The proposed random model depends only on four parameters which could be used to classify and compare the severity of driving environments.

The results were validated using measured data. Two types of steering events were detected; driving through the curve and slow speed maneuvering. All large rainflow cycles found in measured load were also counted in the reduced load. Hence the reduced load can be used to predict fatigue damage of steering components.

Fatigue damage estimated for the reduced loads were in the confidence intervals derived from the random load. The observed load spectrum did not significantly differ from load spectra found in the simulated loads. We conclude that the proposed random load accurately describe the variability of the rainflow ranges for the considered measured loads.

Acknowledgments

We are thankful to Dr. Pär Johannesson for his helpful suggestions and valuable comments which improved the paper. We would like to thank Volvo Trucks for supplying the data in this study and to the members in our research group at Volvo for their valuable advice. Finally, we gratefully acknowledge the financial support from VINNOVA.

References

[1] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.*, 41(1):164–171, 1970.

[2] JS. Bendat. Probability functions for random responses: Prediction of peaks, fatigue damage and catastrophic failures. NASA technical report, 1964.

[3] H. Berndt and K. Dietmayer. Driver intention inference with vehicle onboard sensors. In *IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, pages 102–107, Pune, 11-12 November 2009.

[4] A. Beste, K. Dressler, H. Kötzle, W. Krüger, B. Maier, and J. Petersen. Multiaxial rainflow – a consequent continuation of Professor Tatsuo Endo's work. In Y. Murakami, editor, *The Rainflow Method in Fatigue*, pages 31–40. Butterworth-Heinemann, 1992.

[5] N.W.M. Bishop and A. Sherratt. A theoretical solution for estimation of rainflow ranges from power spectral density data. *Fatigue Frac Eng Mater Struct*, 13:311–326, 1990.

[6] P. A. Brodtkorb, P. Johannesson, G. Lindgren, I. Rychlik, J. Rydén, and E. Sjö. WAFO – a Matlab toolbox for analysis of random waves and loads. In *Proceedings of the 10th International Offshore and Polar Engineering conference, Seattle*, volume III, pages 343–350, 2000.

[7] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via EM algorithm. *J. Roy. Stat. Soc.*, 39(1):1–38, 1977.

[8] K. Dressler, M. Hack, and W. Krüger. Stochastic reconstruction of loading histories from a rainflow matrix. *ZAMM Journal of applied mathematics and mechanics: Zeitschrift fur angewandte Mathematik und Mechanik*, 1997.

[9] G. D. Forney. The Viterbi algorithm. *Proc. IEEE*, 61:268–278, 1973.

[10] M. Frendahl and I. Rychlik. Rainflow analysis - Markov method. *Int. J. Fatigue*, 15:265–272, 1993.

[11] P. Johannesson. Rainflow cycles for switching processes with Markov structure. *Probability in the Engineering and Informational Sciences*, 12:143–175, 1998.

[12] P. Johannesson. *Rainflow Analysis of Switching Markov Loads*. PhD thesis, Lund Institute of Technology, 1999.

[13] P. Johannesson and M. Speckert, editors. *Guide to Load Analysis for Durability in Vehicle Engineering*. Wiley:Chichester, 2013.

[14] P. Johannesson and J-J. Thomas. Extrapolation of rainflow matrices. *Extremes*, 4:241–262, 2001.

[15] M. Karlsson. *Load Modelling for Fatigue Assessment of Vehicles – a Statistical Approach*. PhD thesis, Chalmers University of Technology, Sweden, 2007.

[16] S . Krenk and H. Gluver. A Markov matrix for fatigue load simulation and rainflow range evaluation. *Struct Saf*, 6:247–258, 1989.

[17] G. Lindgren and K. B. Broberg. Cycle range distributions for Gaussian processes - exact and approximate results. *Extremes*, 7:69–89, 2004.

[18] R. Maghsood and P. Johannesson. Detection of the curves based on lateral acceleration using hidden Markov models. *Procedia Engineering*, 66:425–434, 2013.

[19] R. Maghsood and P. Johannesson. Detection of the steering events based on vehicle logging data using hidden Markov models. Submitted to International Journal of Vehicle Design, April 2014.

[20] M. Matsuishi and T. Endo. Fatigue of metals subjected to varying stress. *Japan Society of Mechanical Engineers*, 1968. *In Japanese*.

[21] M. A. Miner. Cumulative damage in fatigue. *Journal of Applied Mechanics*, 12:A159–A164, 1945.

[22] D. Mitrović. *Learning Driving Patterns to Support Navigation*. PhD thesis, University of Canterbury, New Zealand, 2004.

[23] D. Mitrović. Reliable method for driving events recognition. *IEEE Trans. Syst.*, 6(2):198–205, 2005.

[24] A. Palmgren. Die Lebensdauer von Kugellagern. *Zeitschrift des Vereins Deutscher Ingenieure*, 68:339–341, 1924. *In German*.

[25] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

[26] I. Rychlik. A new definition of the rainflow cycle counting method. *International Journal of Fatigue*, 9:119–121, 1987.

[27] I. Rychlik. Rain flow cycle distribution for ergodic load processes. *SIAM J Appl Math*, 48:662–679, 1988.

[28] I. Rychlik. Note on cycle counts in irregular loads. *Fatigue & Fracture of Engineering Materials & Structures*, 16:377–390, 1993.

[29] I. Rychlik. On the "narrow-band" approximation for expected fatigue damage. *Probab. Eng. Mech.*, 8:1–4, 1993.

[30] I. Rychlik. Simulation of load sequences from rainflow matrices: Markov method. *International Journal of Fatigue*, 18:429–438, 1996.

[31] A. J. Viterbi. Error bounds for convolutional codes and an asymptotically optimal decoding algorithm. *IEEE Trans. Informat. Theory*, IT-13:260–269, 1967.

[32] WAFO Group. WAFO – a Matlab toolbox for analysis of random waves and loads, tutorial for WAFO 2.5. Mathematical Statistics, Lund University, 2011.

[33] WAFO Group. WAFO – a Matlab Toolbox for Analysis of Random Waves and Loads, Version 2.5, 07-Feb-2011. Mathematical Statistics, Lund University, 2011.
Web: `http://www.maths.lth.se/matstat/wafo/` (Accessed 24 January 2014).

Appendix

A Proof of Theorem 4

Consider the stationary time series X_i defined in Eq. (10). Let k be the number of maneuvers and random load $X = \{X_i : i = 0, \dots, n\}$, $n = 2k$. We recall that $\mu_n^{osc}(u,v) = E[N_n^{osc}(u,v)]$ is the expected number of upcrossings of the interval $[u,v]$ found in X while the intensity of interval upcrossings

$$\mu^{osc}(u,v) = \lim_{n\to\infty} \frac{1}{n}\mu_n^{osc}(u,v).$$

We begin with a definition of $N_n^{osc}(u,v)$. Consider a load starting at time zero, i.e. infinite sequence $x = (x_0, x_1, \dots)$ of real numbers. For fixed u, v, $u \le v$, $i \ge 0$ and $j > i+1$ define the following sets

$$\begin{aligned} A_i &= \{x_i < u\} \cap \{x_{i+1} > v\}, \\ A_{ij} &= \{x_i < u\} \cap \{x_j > v\} \cap \{u \le x_l \le v \text{ for all } l,\ i < l < j\}. \end{aligned} \tag{A.1}$$

Let $\mathbf{1}_A(x)$ be the indicator function of A, i.e. equal one if $x \in A$ and zero otherwise. Now for fixed $u \le v$, $m \ge 2$ and $x \in R^{m+1}$ we will denote the

number of upcrossings of interval $[u, v]$ found in x, i.e. $N_m^{osc}(u, v)$, by $N_m(x)$. Using the sets A_i and A_{ij}, defined in Eq. (A.1), one obtain that

$$N_m(x) = \sum_{i=0}^{m-1} \mathbf{1}_{A_i}(x) + \sum_{i=0}^{m-2} \sum_{j=i+2}^{m} \mathbf{1}_{A_{ij}}(x). \tag{A.2}$$

Now, we turn to evaluation of $\mu_n^{osc}(u, v) = E[N_n(X)]$ for $n = 2k$. The domain of μ_n^{osc} is divided into three regions; $u \le v < 0$, $0 < u \le v$ and $u \le 0 \le v$.

Region $u \le v < 0$: Since $X_i = 0$ for all odd indexes then $N_n(X)$ is equal to number of $X_{2i} < u < 0$, $0 \le i \le k-1$. Consequently $\mu_n^{osc}(u, v) = k\pi_2\, P(m_1 < u)$ and $\mu^{osc}(u, v) = \frac{1}{2}\pi_2\, P(m_1 < u)$. (Recall that m_i, M_i are independent sequences of iid random variables.)

Region $0 < u \le v$: Similarly $N_n(X)$ is equal to number of $X_{2i} > v > 0$ and hence $\mu_n^{osc}(u, v) = k\pi_1\ P(M_1 > v)$. Consequently $\mu^{osc}(u, v) = \frac{1}{2}\pi_1\, P(M_1 > v)$.

Region $u \le 0 \le v$: Computation of $\mu_n^{osc}(u, v)$ and $\mu^{osc}(u, v)$ when $u \le 0 \le v$ is more complex. First we note that the number of crossings of intervals $[u, v]$, $u \le 0 \le v$ found in sequences X_i, $i = 0, \ldots, n$, and Y_i, $i = 0, \ldots, k$, are equal, i.e.

$$N_n(X) = N_k(Y), \quad Y_i = X_{2i}.$$

Now from the definition of $N_k(Y)$ in Eq. (A.2), it is easy to see that

$$\mu_n^{osc}(u, v) = \sum_{i=0}^{k-1} P(Y \in A_i) + \sum_{i=0}^{k-2} \sum_{j=i+2}^{k} P(Y \in A_{ij}).$$

Since Y_i is a stationary sequence hence $P(Y \in A_i) = P(Y \in A_0)$ and $P(Y \in A_{ij}) = P(Y \in A_{0(j-i)})$ for any $j \ge i + 2$. Consequently, with $P_1 = P(Y \in A_0)$ and $P_l = P(Y \in A_{0l})$, $l = 2, 3, \ldots$,

$$\begin{aligned} \mu_n^{osc}(u, v) &= k\, P_1 + \sum_{i=0}^{k-2} \sum_{j=i+2}^{k} P_{j-i} \\ &= k\, P_1 + \sum_{i=0}^{k-2} \sum_{l=2}^{k-i} P_l = \sum_{i=1}^{k} (k - i + 1) P_i. \end{aligned}$$

Hence the intensity $\mu^{osc}(u, v)$ is given by

$$\mu^{osc}(u, v) = \lim_{n\to\infty} \frac{1}{n} \mu_n^{osc} = \lim_{k\to\infty} \frac{1}{2} \sum_{i=1}^{k} (1 - (i-1)/k) P_i = \frac{1}{2} \sum_{i=1}^{\infty} P_i,$$

by dominated convergence theorem ($\sum_{i=1}^{\infty} P_i \le 1$). Next we will employ Markov property to evaluate $\mu^{osc}(u, v)$.

Let introduce the following sequence of events B_i, $i \geq 1$,

$$\begin{aligned} B_1 &= \{Y_1 > v\} \\ B_i &= \{Y_i > v \text{ and } u \leq Y_l \leq v \text{ for all } 1 \leq l < i\}, \quad i > 1. \end{aligned} \tag{A.3}$$

Using B_i the sum $\sum_{i=1}^{\infty} P_i$ can be written as follows

$$\begin{aligned} \sum_{i=1}^{\infty} P_i &= P(Y \in A_0 \text{ and } Z_0 = 2) + \sum_{i=2}^{\infty} P(Y \in A_{0i} \text{ and } Z_0 = 2) \\ &= \sum_{i=1}^{\infty} P(B_i | Z_0 = 2, Y_0 < v) P(Y_0 < v, Z_0 = 2) \\ &= \pi_2 P(m_0 < u) \sum_{i=1}^{\infty} P(B_i | Z_0 = 2). \end{aligned} \tag{A.4}$$

Since probabilities p_j introduced in Section 3.1 are given by

$$p_j(u,v) = \sum_{i=1}^{\infty} P(B_i | Z_0 = j), \quad j = 1, 2,$$

one has that $\sum_{i=1}^{\infty} P_i = \pi_2 \, P(m < u) \, p_2(u,v)$. This finishes the proof of Eq. (13).

Finally we demonstrate that $p_2(u,v)$ is the solution of Eq. (11). Using Markov property one can evaluate p_j in the following way

$$\begin{aligned} p_j(u,v) &= P(Y_1 > v | Z_0 = j) + \sum_{l=1}^{2} \sum_{i=2}^{\infty} P(B_i | u \leq Y_1 \leq v, Z_1 = l, Z_0 = j) \\ &\qquad \cdot P(u \leq Y_1 \leq v, Z_1 = l | Z_0 = j) \\ &= P(Y_1 > v | Z_0 = j) + \sum_{l=1}^{2} \sum_{i=2}^{\infty} P(B_i | Z_1 = l) P(u \leq Y_1 \leq v | Z_1 = l) p_{jl} \\ &= P(Y_1 > v | Z_0 = j) + \sum_{l=1}^{2} P(u \leq Y_1 \leq v | Z_1 = l) p_{jl} \sum_{i=1}^{\infty} P(B_i | Z_0 = l) \\ &= P(Y_1 > v | Z_0 = j) + \sum_{l=1}^{2} P(u \leq Y_1 \leq v | Z_1 = l) p_{jl} p_l(u,v). \end{aligned}$$

This finishes the proof.

B Hidden Markov models

Hidden Markov models are probabilistic models that can be used for detection of patterns or events in a signal. In some studies, one HMM has been constructed for each type of event, see e.g. Mitrović [22]. They created a training set by

identifying the events manually to build the models and evaluate them. Then for a new observation sequence, they computed the observation likelihoods based on all models and selected the type of driving event with respect to the highest likelihoods. In our suggested method, we have used a single HMM for describing all events instead of constructing several different models where each HMM describes a single event. It should be more simple to estimate the parameters of one model than lots of parameters of different models.

There are two processes in an HMM. The interesting process Z_t describes the events which are not accessible to measure directly. It is thus called hidden and modeled as a Markov chain. However, what can be observed is a process Y_t and its statistical properties depend on the value of Z_t. The problem at hand is to estimate the parameters of the HMM. Based on an observation of Y_t, it is then possible to reconstruct the most probable hidden process and identify events.

We have used HMMs to detect steering events such as curves and maneuvers by using on-board logging signals available on trucks, such as lateral acceleration, vehicle speed and steering wheel angle. Figure B.1 shows a lateral acceleration signal and the corresponding identified hidden states process.

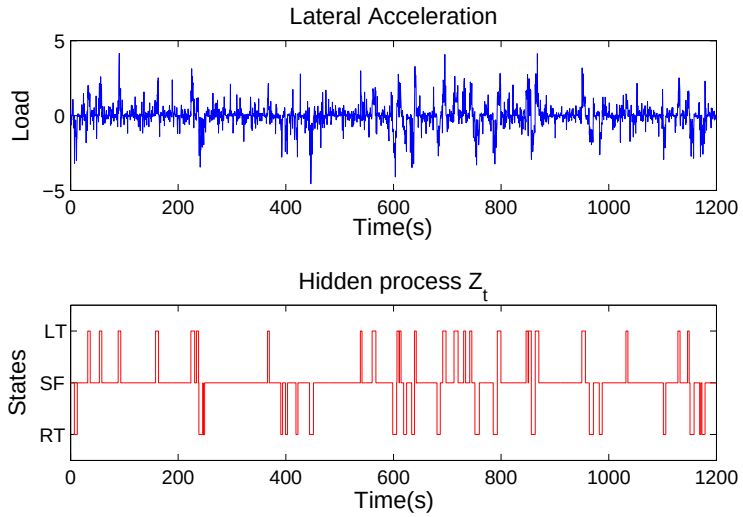

Figure B.1: Lateral acceleration signal and the corresponding hidden states.

Suppose that there are three events Right turn (RT), Left turn (LT) and Straight forward (SF). The idea is to see these three events as three hidden states and construct the HMM based on them. Figure B.2 illustrates three hidden states and the transitions between them and a sequence of observations that can be generated based on the probability distribution of observation symbols.

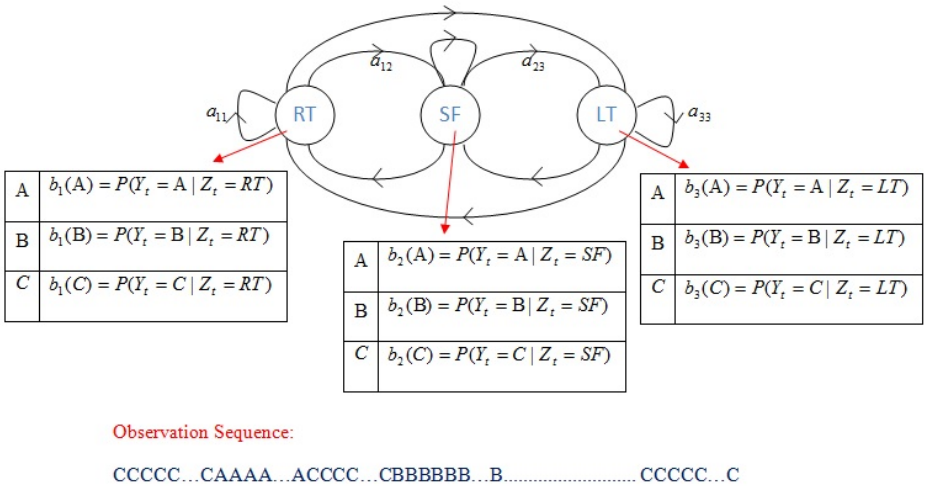

Figure B.2: The hidden state sequence is modeled by a Markov chain and the observation sequence is modeled by the emission probabilities.

Lateral acceleration values have been translated into predefined classes. Here, three classes will be used, $V = \{\text{A, B, C}\}$, that are defined as follows:

- $\text{A} = \left\{\text{"}lateral\ acceleration\text{"} < -0.2\ m/s^2\right\}$,
- $\text{B} = \left\{-0.2\ m/s^2 \leq \text{"}lateral\ acceleration\text{"} \leq 0.2\ m/s^2\right\}$,
- $\text{C} = \left\{\text{"}lateral\ acceleration\text{"} > 0.2\ m/s^2\right\}$.

where the threshold 0.2 m/s^2 has been chosen based on experience. This kind of clustering will create a sequence of observation symbols which has been used to estimate the emission matrix in our model.

It has been demonstrated that a discrete HMM can be good in pattern recognition, see Rabiner [25]. We have also used a discrete HMM $\lambda = (A, B, \pi)$, where λ represents model parameters which contain the transition matrix, the emission matrix and the initial state distribution.

Let $\{Z_t\}_{t=1}^{\infty}$ be a Markov chain where Z_t denotes a hidden state at time t and has possible values $S = \{S_1, S_2, ..., S_N\}$. The transition probabilities between the hidden states are defined by the matrix $A = \{a_{ij}\}$, called transition matrix, where

$$a_{ij} = P(Z_{t+1} = S_j | Z_t = S_i),\ i, j = 1, 2, ..., N$$

and $\sum_{j=1}^{N} a_{ij} = 1$.

Further, there is another process $\{Y_t\}_{t=1}^{\infty}$ where Y_t denoting the observation symbol at time t. The sequence of observation has possible values

$V = \{V_1, V_2, ..., V_M\}$ and it is observable for us. The probability distribution of observation symbols in each state is given by the emission matrix, $B = \{b_j(V_k)\}$, where

$$b_j(V_k) = P(Y_t = V_k | Z_t = S_j),\ k = 1, 2, ..., M$$

and $\sum_{k=1}^{M} b_j(V_k) = 1$.

The state where the hidden process will start is modeled by the initial state probabilities that are denoted by $\pi = \{\pi_1, \pi_2, ..., \pi_N\}$ where

$$\pi_i = P(Z_1 = S_i),\ i = 1, 2, ..., N$$

and $\sum_{i=1}^{N} \pi_i = 1$.
The parameters must be estimated to characterize the model.

In an HMM, a training set is used to estimate the parameters of the model, while a test set is used to validate the model. A training set consists of all necessary information for estimating the model parameters. In our study, the training set contains all history about the curves such as the start and stop points of them.

To estimate model parameters based on an observation sequence, we have used the Baum-Welch algorithm which was introduced by Baum et al. [1]. It is a special case of the EM (expectation-maximization) algorithm, see Dempster et al. [7]. The Baum-Welch algorithm is one of the most well known methods for estimating the model parameters in HMMs on unlabelled sequences. It is an iterative maximum likelihood method and starts with initial parameters that in our case are set based on training data. The algorithm uses a forward-backward procedure to estimate the model parameters for a given sequence of observations. Here, we have used the Baum-Welch algorithm which was described by Rabiner [25].

Another important algorithm in HMMs is the Viterbi algorithm, see Viterbi [31] and Forney [9], which is used to find the most probable sequence of hidden states for a new signal. Suppose that we have an observation sequence $y_1, y_2, ..., y_n$. We would like to find steering events for this new observation. It means that we should find a sequence of hidden states which maximizes the probability of observing this specified observation. The Viterbi algorithm finds the state sequence $z_1, z_2, ..., z_n$ out of the 3^n possible sequences of length n that maximizes:

$$P(Y_1 = y_1, ..., Y_n = y_n | Z_1 = z_1, ..., Z_n = z_n; \lambda).$$

In fact, the Viterbi algorithm gives the most likely sequence of hidden states from which it is possible to identify the steering events.