1. Introduction

Contemporary methods of designing and testing mechanical vehicles are based on simulation techniques that require the use of precise models of vertical and horizontal dynamics and sequences of random events occurring in road traffic conditions. This problem is relevant to the optimization of vehicles with internal combustion engines, electric vehicles (EV) and hybrid electric vehicles (HEV). To select the most appropriate drive system architecture for a particular vehicle class and driving cycle, it is necessary to optimize the size of components according to their cost functions, such as the lowest $CO_2$ emissions, the lowest weight, fuel savings or any combination of these attributes in the architecture [1, 7, 9 and 18].

Regardless of the simulation technique used: quasi-static using a “Backward-facing” vehicle model or a dynamic simulation with a “Forward-facing” model, understanding of the representative driving cycle is essential. In the first case, for an open-loop system, the time series of speed is imposed on the input of the vehicle model in order to calculate rpm and torque on the wheels. In a closed-loop vehicle model, on the other hand, the driving cycle is a setpoint for the driver block, which generates a suitable engine torque. The time and cost constraints associated with the design and testing of various possible vehicle architectures require methods of driving cycle synthesis that can meet the modelling and simulation requirements of automotive engineers throughout the R&D process. It is not possible to optimize the parameters and gradually increase the autonomy of the vehicles based on standard driving cycles, and such optimization cannot prevent “cycle beating”. To ensure that the synthesized time series based on the collected databases are representative, it is necessary to use algorithms adopting techniques based on stochastic and statistical models [6, 19].

A synthesis technique combining the MCMC method and multifractal analysis has been proposed and verified. The method allows simple determination of the speed profile compared to classic frequency analysis. A combination of multiple criteria is frequently used [2, 4].

The methods of driving cycles construction require quantization of traffic parameters. Depending on their function (emissions estimation, fuel consumption estimation or traffic engineering, etc.), the defined states can be synthesized for micro-trips, segments, heterogeneous classes or modal cycles [17]. Micro-trips are driving models between stops including periods of inactivity. Traffic signals and overloads contribute to “stop-go” driving patterns, and result in increased fuel

---

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie www.ein.org.pl
consumption. Micro-trips are a good representation of fuel consumption and emissions. Segments model situations for various types of roads and traffic conditions classified, for instance, by the Level of Service (LOS). They may start and end with different driving parameters, which is why speeds and accelerations have to be accordingly adjusted when combining segments in the course of cycle synthesis. Driving cycles based on heterogeneous traffic classes determined through a statistical breakdown of data are constructed as kinematic sequences using probabilistic methods and analysis of probability distributions. This method is not aimed at testing emissions and fuel consumption. Modal cycles represent recorded parameters of vehicle traffic for specific acceleration intervals, constant speed or idle periods. In procedures using the theory of stochastic processes to analyse the equations of vehicle dynamics, represented by speed and acceleration, the major trend in recent research focuses on methods based on the Markov chains theory [8, 11]. There have also been attempts at using 3D Markov models in the synthesis of driving cycles, which incorporate the roadway slope [20]. Methods based on multi-dimensional Markov chains enable a realistic assessment of fuel consumption and CO₂ emissions, even after time compression of the synthesized time series [5]. However, such simulations involve a high time cost.

This paper proposes a method for synthesizing naturalistic driving cycles in which the information about instantaneous values of acceleration can be replaced by the degree of multifractality assessed using formalism based on wavelet leaders. This helped reduce the number of Markov chain dimensions in the simulation process. The process was illustrated on the example of the Markov chain Monte Carlo (MCMC) algorithm for the vehicle speed signal. The input for the algorithm was recorded during a series of experiments in real conditions. Statistical factors and mean tractive force (MTF) were used to select and classify road traffic models equivalent to the real conditions.

2. Wavelet leaders multifractal formalism in MCMC technique

Each real driving cycle can be regarded as a sequence of random transitions between defined m-states of the vehicle occurring in road traffic. The frequency of specific states depends on the technical parameters of the car, the intensity of road traffic and the driver’s behaviour. By determining the probability of remaining in the current state or transition into a different state, we can describe the examined phenomenon in the form of a transition probability matrix (TPM) (1):

\[ P = \begin{bmatrix} P_{11} & \cdots & P_{1m} \\ \vdots & \ddots & \vdots \\ P_{nm} & \cdots & P_{mm} \end{bmatrix} \in \mathbb{R}^{mn}, \]

where the entry \( P_{ij} \) is the probability of transition from and to state \( j \) when \( j \neq i \) or remaining in state \( i \) when \( j = i \). The probability \( P_{ij} \) can be calculated using the following equation:

\[ P_{ij} = \frac{N_{ij}}{N_j}, \]

where \( N_{ij} \) is the number of transitions from and to state \( j \). The sum of the values of entries in each row is equal to one. The random process \( \{X_n\}_{n \in \mathbb{N}} \) is referred to as the Markov chain if for any \( n \in \mathbb{N} \) the following equation is true: \( P(X_{n+1} = k | X_n = i) = P(X_{n+1} = k | X_0 = i, X_1, \ldots, X_n) \). It is assumed that the TPM is stationary, which implies that the Markov chain is homogeneous. Therefore, for Markov chains, the conditional distribution of the random variable \( X_{n+1} \) depends only on the currently known value of \( X_n \). Thus, considering the current driving state, the future state can be determined using Monte Carlo simulation based on the transition probability matrix. It is possible to generate a driving cycle of any duration, which may be used to identify a cycle with the required duration, for the assumed equivalence criteria.

The synthesis of a driving cycle using the MCMC method, where in addition to the speed signal – also consider other parameters are taken into consideration, requires a multi-dimensional description of the defined vehicle states, which significantly complicates the determination of the transition probability matrix and extends the implementation time of the algorithm. If the second parameter is acceleration, which is not measured directly in most real driving cycles, it becomes necessary to differentiate the speed signal in order to acquire information about motion dynamics. Where this is the case, the standard 1-second sampling period for the time series of speed does not guarantee a sufficient precision of the acceleration signal.

Papers where road traffic was analysed based on recorded vehicle speed signals indicate the multifractal properties of the dynamics of such traffic [3, 16]. Multifractality can also be observed both in real and standard driving cycles [14]. Our research proposes to eliminate the acceleration signal from the multi-dimensional description of vehicle states using information about driving dynamics represented by multifractal parameters of the speed signal. The iteration in the Monte Carlo simulation was performed for a specific time, with a requirement concerning driving dynamics. The multifractal analysis, which is based on estimated scaling exponents of the signal, is a popular statistical tool used to assess empirical data. In the case of time series, mathematical formalism was initially based on increments of their value, measured as Hölder point exponents \( h \) of time function \( x(t) \) at point \( t_0 \), determined by the supremum of all exponents that, for constant \( C > 0 \), meet the following condition:

\[ |x(t) - P_n(t - t_0)| \leq CT - t_0^h, \]

where \( P_n(t - t_0) \) is a polynomial of degree \( n < h \) [13, 15, 16]. The result of the algorithm is the multifractal spectrum \( D(h) \), i.e. a function describing the fractal dimensions of points with the same Hölder exponent.

The multifractal formalism in the time and frequency domain that is used in the research makes it possible to estimate multifractal parameters using wavelet leaders, which are representatives of local Hölder exponents of the signal. The algorithm is characterized by low computing costs, numerical stability and high versatility with respect to real signals. For coefficients (3) of the discrete wavelet transform (DWT) of function \( x(t) \) and basic wavelet with a compact support \( \psi_0(t) \):

\[ d_j(j,k) = \int_{-\infty}^{\infty} x(t) 2^{-j/2} \psi_j(2^{-j/2}t - k) \, dt, \]

wavelet leaders (4) for the collection of largest coefficients \( d_{j}(j,k) = d_{j,l} \), in the neighbourhood of \( 3\lambda \) are defined in any scale by the following equation:

\[ L_{\lambda}(j,k) = \sup_{\lambda \leq 3\lambda} |d_{j,l}|, \]

where \( j,k \) are integers and \( 3\lambda := 3\lambda_{j,k} = \lambda_{j,k-1} \cup \lambda_{j,k} \cup \lambda_{j,k+1} \) and \( \lambda_{j,k} := \left[ 2^{j/2} (k+1) 2^{j/2} \right] \). It can be demonstrated [10] that Hölder exponents are scaling exponents of wavelet leaders: \( l_n(j,k) \sim 2^nh \). Also, the structural function (5) defined for wavelet leaders is described by a power law where the exponent is a multifractal scaling exponent \( \zeta(q): R \rightarrow R \).
where \( q \) is the order of the structural function, and \( n_j \) is the number of intervals of the multi-resolution analysis.

The function generated using the Legendre transformation of the multifractal scaling exponent \( \zeta(q) \), under mild regularity conditions, is the upper limit of the multifractal spectrum (6) of the investigated signal:

\[
D(h) \leq \min_{q=0} \left[ 1 + qh - \zeta(q) \right]
\]

Coefficients of the Taylor expansion of the exponent \( \zeta(q) \) – log-cumulants \( c_p \) – are an alternative description of the parameters of the multifractal spectrum of the analysed signal:

\[
\zeta(q) = \lim_{j \to 0} \log_j \frac{Z_{j+1}(q,j)}{Z_j(q,j)} = \sum_{p=1}^{\infty} c_p \frac{q^p}{p!} = c_1 q + c_2 \frac{q^2}{2} + c_3 \frac{q^3}{6} + \ldots
\]

In particular: the coefficient \( c_1 \) describes the position of the maximum of the spectrum, and coefficients \( c_2 \) and \( c_3 \) describe its degree of multifractality, i.e. the width of the spectrum and its asymmetry, respectively. The dynamic properties of the systems are successfully described based on the parameters of the multifractal spectra of the representative time series [12]. Approximation of \( \zeta(q) \) (7), i.e. also of the multifractal spectrum \( D(h) \), using coefficients \( c_p \) significantly simplifies the algorithms for a comparative analysis of the investigated systems.

3. Simulation tests of driving dynamics and wavelet leaders of speed signal

The relationship between the time series of acceleration and parameters of the multifractal spectrum of speed has been illustrated on the example of the synthetic signal \( v(n) \) of vehicle speed (Fig. 1a). The signal was resampled to achieve signals with acceleration 2, 4 and 8 times higher. Due to the resampling, the histograms are not identical, but they are comparable. The signals that had been shortened were repeated 2, 4 and 8 times, respectively, to obtain signs with the same number of samples.

An analysis of the singularity spectra demonstrates that the position of their maxima and width depends on the accelerations of the simulated signals. The log-cumulants of synthetic signals (Table 1) and relationships of log-cumulants and acceleration (Fig. 4) were determined. The first and second log-cumulant, describing the position of the maximum of the multifractal spectrum and its width, respectively, were proposed as the synthetic parameters for the assessment of driving dynamics using the multifractal spectrum.

<table>
<thead>
<tr>
<th>Log-cumulant</th>
<th>Signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_1 )</td>
<td>0.8127</td>
</tr>
<tr>
<td>( c_2 )</td>
<td>-0.1640</td>
</tr>
<tr>
<td>( c_3 )</td>
<td>0.1220</td>
</tr>
</tbody>
</table>

Fig. 1. Synthetic signal of vehicle speed (a) and its histograms (b) in the specified time; the same run in a two times shorter time (c) with comparable amplitude distribution (d); the same run in four times shorter time (e) with comparable amplitude distribution (f); the same run in eight times shorter time (g) with comparable amplitude distribution (h).

For the speed signals in figure 1, accelerations were determined (using differentiation – Fig. 2), as well as multifractal spectra (Fig. 3).
4. Implementation of the algorithm for the synthesis of driving cycles and analysis of research results

An algorithm was proposed to generate naturalistic driving cycles using first-order Markov chains and multifractal formalism based on an analysis of wavelet leaders (Fig. 5).

Fig. 5. Block diagram of the MCMC algorithm for synthesizing driving cycles

The paper presents results of tests and analysis of car traffic in actual road conditions, represented by urban driving in a large agglomeration (Fig. 6). The analysis was carried out based on the time series of vehicle speed recorded with a sampling period of 1 s. The research has been described in the paper [14]. Due to the large share of “zero speeds” (idle periods) in the test, amounting to approx. 25%, fragments corresponding to stops were removed from the recorded time series (Fig. 7), which enabled the segmentation and determination of the transition probability matrix (TPM) and testing of driving dynamics through an analysis of log-cumulants. The recorded speeds were divided into 20 even intervals corresponding to increasing speeds other than zero as well as a 21st interval corresponding to stops. Speed resolution is approx. 0.9 m/s. The authors attempted to achieve a fairly good car speed resolution while avoiding intervals with a very low (or zero) probability of occurrence.

Statistical analyses were conducted in the R environment, and the multifractal analysis was carried out in Matlab.

The transition probability matrix (TPM) calculated based on the reference signal of the cycle had the size of 21x21 (Fig. 8).

A simulation of 100 cycles was carried out in accordance with the Metropolis-Hastings algorithm. Three sample cycles – candidates No. 1, 2 and 3 (Figs. 9a–c) were selected to illustrate the results of the algorithm. The main statistics of the speed signal (maximum, minimum, mean and standard deviation) of the sample cycles are similar to the statistics of the reference signal. There was also a fourth cycle shown – candidate No. 4 – generated for verification purposes based on the distribution of speed amplitudes (Fig. 9d).

The first two log-cumulants determined for each cycle (Fig. 10) are the best match of dynamics in comparison with the reference signal for candidate No. 1.

The conformity of probability density distributions to the distribution of the reference cycle has also been verified (Fig. 11a–d). The distribution of the reference cycle has been approximated with an empirical function. The chi-squared test or the Kolmogorov–Smirnov test can be used to check the goodness of fit of empirical data to the
approximated probability distribution function, but these tests reject the null hypothesis for the investigated duration of the driving cycle. For the test to confirm the null hypothesis, the duration of the driving cycle would have to be significantly extended, which is not possible. In such a situation, it is best to estimate the goodness of fit of the theoretical distribution to the observed distribution through a visual comparison. This was done using the probability-probability plot (Fig. 11e). Apart from the cycle of candidate No. 4, the best fit is demonstrated by candidate No. 1.

The accelerations of candidate No. 4 (Fig. 12), which has a speed amplitude distribution that perfectly matches the reference distribution, are entirely different – almost constant. The accelerations of the remaining candidates can only be assessed in terms of their minimum and maximum values.

In the method of modal cycles and speed-based segmentation, which was adopted in this paper, the time series produced using the Markov model are stepped, which means that they had to be smoothed in the next step. The method of local quadratic regression smoothing was selected from among the various methods to smooth the series. Once the iteration and filtration process was completed, the stop periods were added to the time series, and a search was started for the most representative cycles out of all of the cycles produced by the synthesis, for the selected equivalence factor.

In the course of the study, the results of the algorithm for the synthesis of equivalent driving cycles were analysed according to selected statistical parameters and the criterion of the mean tractive force (MTF) (8), i.e. the tractive energy of the vehicle (Table 2) transmitted through the wheels:

\[
\bar{F}_{\text{trac}} = \frac{1}{\tau_{\text{total}}} \int_{t_{\text{trac}}}^{t_{\text{total}}} F(t) v(t) \, dt. \tag{8}
\]

where: total tractive force \( F(t) \) is the sum of the forces of aerodynamic resistance \( F_{\text{air}} \), rolling resistance \( F_{\text{roll}} \) and inertia of the vehicle \( F_{\text{iner}} \), \( v(t) \) and \( a(t) \) are speed and acceleration, respectively, for a driving cycle of the duration \( \tau_{\text{total}} \) and \( \tau_{\text{trac}} \) represents the time intervals during which \( F(t) > 0 \).

In the calculations of the MTF coefficient, the most significant element is the force of inertia, which is proportional to acceleration. The best fit to the real cycle according to the MTF criterion is represented by candidate No. 1. The primary parameters considered in the course of cycle verification are listed in Table 3. Minimum and maximum values of speed and acceleration, subject to initial verification, were omitted.

All synthesized driving cycles have the correct mean value and standard deviation. The selection also cannot be performed based on the distributions of probability density. Candidates No. 1 and No. 4 show the best fit of speed amplitudes probability distribution.

If we assume a discrepancy of the MTF coefficient of the generated cycle with the reference cycle of over 10% to be an unacceptable
in terms of equivalence to real driving conditions, the log-cumulants tested in the phase of synthesis of candidate cycles and the MTF used to verify their equivalence suggest candidate No. 1.

5. Conclusions

The presented research results provide a new perspective on statistical-random methods for synthesizing real vehicle driving cycles. It was demonstrated that driving dynamics represented by acceleration-random methods for synthesizing real vehicle driving cycles. It was possible to carry out cycle synthesis, which took into consideration wavelet leaders for driving dynamics testing made it possible to carry out cycle synthesis, which took into consideration speed and acceleration, using Monte Carlo simulation with a single-dimensional Markov chain. The algorithm for the synthesis of equivalent driving cycles was verified using the criterion of mean tractive force (MTF).

The database used so far included data from tests of vehicles with internal combustion engines. The authors' future research will include an analysis of driving cycle prediction and road traffic modelling for the purpose of drive system control and electricity management in electric vehicles. The expected results will be useful in designing the infrastructure of charging stations for electric cars.

References


Table 3. Summary of selected values characteristic to the investigated cycles

<table>
<thead>
<tr>
<th>Driving cycle</th>
<th>Fit of the distribution</th>
<th>Mean value [m/s]</th>
<th>Standard deviation [m/s]</th>
<th>Log-cumulant 1 [-]</th>
<th>Log-cumulant 2 [-]</th>
<th>MTF [N]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference cycle</td>
<td>+</td>
<td>9.7</td>
<td>4.4</td>
<td>0.72</td>
<td>-0.13</td>
<td>689</td>
</tr>
<tr>
<td>Candidate 1</td>
<td>+</td>
<td>10.0</td>
<td>4.6</td>
<td>0.74</td>
<td>-0.12</td>
<td>743</td>
</tr>
<tr>
<td>Candidate 2</td>
<td>+/-</td>
<td>9.7</td>
<td>4.4</td>
<td>0.58</td>
<td>-0.16</td>
<td>820</td>
</tr>
<tr>
<td>Candidate 3</td>
<td>+/-</td>
<td>10.4</td>
<td>4.5</td>
<td>0.62</td>
<td>-0.10</td>
<td>796</td>
</tr>
<tr>
<td>Candidate 4</td>
<td>+</td>
<td>9.7</td>
<td>4.6</td>
<td>1.06</td>
<td>-0.07</td>
<td>450</td>
</tr>
</tbody>
</table>

Andrzej PUCHALSKI
Iwona KOMORSKA
Department of Mechanical Engineering
Kazimierz Pulaski University of Technology in Radom
Malczewskiego 29, 26-600 Radom, Poland

Marcin ŚLĘZAK
Andrzej NIEWCZAS
Motor Transport Institute
Jagiellońska 80, 03-301 Warszawa, Poland

E-mails: andrzej.puchalski@uthrad.pl, iwona.komorska@uthrad.pl,
marcin.slezak@its.waw.pl, andrzej.niewczas@its.waw.pl