

Stochastic Real-World Drive Cycle Generation Based on a Two Stage Markov Chain Approach

Andreea Elena Balau and Dennis Kooijman
TNO Automotive

Ignacio Vazquez Rodarte
TUE

Norbert Ligterink
TNO Automotive

ABSTRACT

This paper presents a methodology and tool that stochastically generates drive cycles based on measured data, with the purpose of testing and benchmarking light duty vehicles in a simulation environment or on a test-bench. The WLTP database, containing real world driving measurements, was used as input data. Consequently cycles that contain typical accelerations per velocity and road types are generated, such that these cycles are representative to real driving behavior.

The stochastic drive cycle generator is developed in Matlab and is based on Markov processes. Two separate stochastic generators are used: one for generating the road type and one for generating the vehicle acceleration. First, a random road type profile is generated from the four different road types that are considered in the WLTP database: urban, rural, motorway and high-motorway, each of them with sub-road types based on different velocity bins. For each sub-road type, speed data and acceleration data will be used to partition the data into classes and states. The second random generator function is used to generate the vehicle acceleration corresponding to the already generated road type. Then, the speed profile is derived from the random generated acceleration profile.

CITATION: Balau, A., Kooijman, D., Vazquez Rodarte, I., and Ligterink, N., "Stochastic Real-World Drive Cycle Generation Based on a Two Stage Markov Chain Approach," *SAE Int. J. Mater. Manf.* 8(2):2015, doi:10.4271/2015-01-0488.

INTRODUCTION

Fixed engine or vehicle cycles which serve as input for models and vehicle dyno tests have been the basis for emission legislation and control development so far. However, the focus of OEM's, consumers and emission legislation is shifting towards "real-world" energy consumption and emissions. Therefore effort and resources are put into the development of systems and control strategies that improve the performance of the vehicle for a wide range of vehicle speed profiles. These strategies must all be evaluated, and as test-bench time is expensive, simulation tools are increasingly important. However, in order to have an effective evaluation the system needs to be tested under realistic situations. Emissions [3], component sizing and architecture of hybrid vehicles [4] and Energy Management Strategies (EMS) [1] are all greatly influenced by the drive cycle. Reference cycles that represent the intended vehicle usage are therefore used early in the development phase, and depend on many factors such as geography, type of road and vehicle category. As such a number of driving cycles exist, usually focusing on one specific area, covering average vehicle usage. Efforts as the Artemis European research project, for example, deals particularly with pollutant

emissions [11]. A methodology that generates drive cycles based on measurements of specific vehicle usage and vehicle limitations would improve the real-world performance of the vehicle.

Different methods have been proposed to approach this problem. After analyzing the statistical description of several driving cycles, methodologies have been proposed to build new drive cycles from selected and representative "micro-trips" [2]. The micro-trips consist of sections within the drive cycle which correspond to a Rural, Urban, High-way scenario as determined by characteristics such as variability, average speed and acceleration behavior. This approach considers implicitly the traffic experienced by the vehicle when building the Transition Probability Matrix (TPM). Another proposal to take these characteristics into account is to use a Probability Function to generate a completely new drive cycle [1]. In this case the function is calibrated with data from a data base, but it also allows the explicit modification of parameters in order to control a certain characteristic of the drive cycle such as more variability and average speed among others. The most popular way to synthesize new drive cycles is the use of Markov Chains [5, 6, 7, 8, 9, 10]. This mathematical concept allows representing a driving pattern in a TPM which is,

derived directly from measured data. However, as it is a stochastic process, there is no guaranty that the resulting profile is physically plausible. In order to determine which profiles are to be used a statistical analysis needs to be performed. It is also worth mentioning that the data usually used corresponds to passenger cars [8], [10] or experimental vehicles [7]. Even though the methodologies are generic, the values and some characteristics taken into account are not adequate for commercial vehicles.

This work proposes a two-phase approach to generating new drive cycles. First, the road-type profile is generated which will provide insight on the acceleration of the vehicle. This assumes that the driver will not behave the same when transiting a rural road and a highway; as such the acceleration will have a statistically different behavior. This can then be verified when generating the random cycle, thus assuring that the new profile matches the expectation of actually having the desired sections. The second step consists of creating the acceleration profile which afterwards will be compared against the road-type distribution defined by the user. The relevance of this road-type partitioning is that it has a major impact on simulations results, e.g. emissions [12]. The estimated emissions depend on the mission profile and speed patterns; hence, strong variations among different phases are to be expected in terms of average speed and traffic situation. In this case the traffic is considered as stop-go behavior in the speed profile.

The goal of this work is to develop a tool that stochastically generates random driving cycles based on measured data, with the purpose of testing light duty vehicles in a simulation environment or on a test-bench. The WLTP database was used as input data. This database was created with the help of a number of European countries that collaborated and provided real world driving measurements. Consequently cycles that contain typical accelerations per velocity and road types are generated, such that these cycles are representative to real driving behavior.

The contents of this paper are as follows: first, the Methodology for the development of the Cycle Generator Tool will be presented remarking some of the design decisions and how the general problem was approached. Second, the Graphical User Interface will be presented including an explanation on what parameters need to be defined by the user and how this will affect the resultant drive cycle. Next, the Results section will analyze six drive cycles generated by the tool. Finally a Conclusions section is also included.

DATA DESCRIPTION

For the development of the World-wide harmonized light duty test cycle (WLTC), driving data is collected from several countries. Velocity profiles for a number of vehicles and a number of days of driving is available. In some cases velocities and accelerations were not physically feasible, e.g. velocities above 300 km/h or accelerations above 5 m/s². The data was collected, smoothed with a Hamming filter and grouped into sub-cycles from stop to stop according to the maximal velocity. This data has been made available for generating the input probabilities for the random cycle generator. The data was not further manipulated, except for filling in gaps in between the different sub-cycles.

The design of the WLTC is divided in four phases or sub-cycles (driving between two stops) based on the maximal velocity achieved in that period. These four sub-cycles are *low*, *medium*, *high* and *extra-high*, which are commonly also referred to as urban, rural, motorway and high-motorway.

The amount of high acceleration is limited, which seems to suggest that all countries and participating institutes have selected average, medium-powered vehicles. Likewise, the generated driving behavior may be on the same side for a high-powered vehicle. However, the driving behavior may require occasionally too much power to be achievable by a low power or high mass vehicle. This yields an important restriction on the generated cycle based on vehicle characteristics. In order to circumvent this obstacle the drive cycle generator cross-check the speed profile against the vehicle characteristics specified by the user, more details will be explained later on.

METHODOLOGY

The stochastic drive cycle generator is developed in Matlab and is based on Markov processes. Markov theory for stochastic behavior has been extensively studied. For this paper the main references regarding the theoretical background on Markov chains are [13] and [14].

Two separate stochastic generators are used: one for generating the road type and one for generating the vehicle acceleration. In the first stage of the approach, a random road type profile is generated from the four different road types that are considered in the WLTP database: urban, rural, motorway and high-motorway, each of them with sub-road types based on different velocity bins. In the second stage, a random generator function is used to generate the vehicle acceleration corresponding to the already generated road type.

Data Processing

First, sub-road types are defined based on velocities and then, for each sub-road type, data is clustered in velocity and acceleration bins.

Road-Type Definition from Velocities

In order to distinguish between transient driving and driving at a certain velocity with some fluctuations, the average velocity is determined. If the velocity remains within a band, a fixed "road-type", or velocity band is assumed and the second-by-second driving behavior is generated accordingly. If the average velocity changes, transient driving is assumed with an appropriate acceleration or deceleration incorporated in the driving behavior.

The average velocity and transitions therein is a key aspect of the driving behavior. It separates the typical fluctuations or jitter of constant driving from the deliberate changes in velocity. The associated time-scale must be longer than it takes to accelerate from one road-type to another. A triangular averaging of the velocity (weighing the center point maximal and the edge points minimally) over a window of 31 seconds yields a sufficiently stable result, where

fast fluctuations are removed yet the overall characteristics of the driving behavior is retained. This filter is used to determine the average velocity ($V_i^{average}$) to be used to determine transient driving.

$$V_i^{average} = \frac{(V_{i-15} + 2V_{i-14} + \dots + 15V_{i-1} + \dots + 2V_{i+14} + V_{i+15})}{256} \quad (1)$$

Transient driving is the driving while changing average velocity, for example due to a change in speed limit, speed reduction in congestion, or speeding up after overtaking a slow vehicle. In all cases the average velocity changes from one value to another one.

It is not sufficient to set bands of velocities to distinguish transient behavior as moving from one band to the other. For example, one might drive at a desired velocity in between two velocity bands. Instead, a center of each band-pair is defined. Only if the average velocity crosses the center of the band, a new road-type is selected. If the average velocity crosses another center within a limited time, the intermediate road-type is dropped as this velocity was traversed rather than a target or desired velocity.

Velocity and Acceleration Cluster Algorithm

For each sub-road type, speed data and acceleration data will be used to partition the data into classes and states. The states are related to acceleration and speed profiles, while the classes are solely related to the velocity profile.

The Cluster Algorithm was used for the data classification.

First, an initial guess for the bin centers is defined:

$$C = [c_1 \ c_2 \ \dots \ c_k]^T \quad (2)$$

Second, a new set for the bin's edges is defined such that:

$$X_{speed} = \left[0 \ \frac{c_1 + c_2}{2} \ \frac{c_2 + c_3}{2} \ \dots \ \frac{c_{k-1} + c_k}{2} \ \infty \right]^T \quad (3)$$

Third, based on the new edges, a new set of bins is derived:

$$S(i) = \{v | X_{speed}(i) \leq v \leq X_{speed}(i+1)\} \quad (4)$$

Fourth, the new centers can be defined as:

$$\hat{C} = \bar{S} = \{\bar{s}(1), \bar{s}(2), \dots, \bar{s}(k)\} \quad (5)$$

Fifth, the difference is measured by the least square error (LSE) as:

$$LSE = \|\epsilon\|^2 = \sqrt{\epsilon^T \epsilon} = \sqrt{[\hat{C} - C]^T [\hat{C} - C]} \quad (6)$$

Until the LSE is less than epsilon there is an iteration loop from step two to five.

Speed class definition, determined by using the clustering algorithm, is presented in Table 1.

Table 1. Class definition.

Class	Lower bound	Center	Upper bound
i	$X_{speed}(i)$	$\hat{C}(i) = \bar{S}(i)$	$X_{speed}(i+1)$

Prior to deriving the states, the same clustering algorithm is used as for the classes, except that the bins' edges are now defined as:

$$X_{acc} = \left[-\infty \ \frac{c_1 + c_2}{2} \ \frac{c_2 + c_3}{2} \ \dots \ \frac{c_{k-1} + c_k}{2} \ \infty \right]^T \quad (7)$$

and:

$$S(i) = \{a | X_{acc}(i) \leq a \leq X_{acc}(i+1)\} \quad (8)$$

Acceleration state definition, determined for each class by using the clustering algorithm, is presented in Table 2.

Table 2. States definition.

States	Lower bound	Center	Upper bound	Class
i	$X_{acc}(i)$	$\bar{S}(i)$	$X_{acc}(i+1)$	i

Accordingly, for each speed class "i", a probability matrix is generated, including transitions between acceleration states "j":

$$N\{i\} = \begin{bmatrix} p_{0,0} & \dots & p_{0,j} \\ \vdots & \ddots & \vdots \\ p_{j,0} & \dots & p_{j,j} \end{bmatrix} \quad (9)$$

The total number of state transitions is calculated with

$$M\{i\} = N\{i\} [1 \ 1 \ \dots \ 1]^T \quad (10)$$

In order to guarantee that divisions by zero do not occur it is required that:

$$D\{i\} = \max(M\{i\}, 1) \quad (11)$$

The probability of each state transition to occur is:

$$P\{i\} = D\{i\}^{-1} N\{i\} \quad (12)$$

or with least square estimate:

$$\hat{P} = (D\{i\}^T D\{i\})^{-1} D\{i\}^T N\{i\} = D\{i\}^{-1} N\{i\}$$