5-2017

Battery Aging Studies Based on Real-World Driving

Zifan Liu
Clemson University, zifanl@g.clemson.edu

BATTERY AGING STUDIES BASED ON REAL-WORLD DRIVING

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Automotive Engineering

---

by
Zifan Liu
May 2017

Accepted by:
Dr. Zoran Filipi, Committee Chair
Dr. Andrej Ivanco, Committee Co-Chair, Research Advisor
Dr. Simona Onori
Dr. Srikanth Pilla

Abstract

While being a competitive candidate for energy storage systems in automotive applications, lithium-ion battery still needs to overcome fundamental compromises regarding energy density, power density, lifetime, costs and safety concerns. A significant breakthrough can be expected by understanding the real-world customer usage patterns and leveraging this knowledge to develop an optimized battery design and control. However, the challenges of filtering through massive real-world driving data and identifying the features relevant to the real-world battery operations still remain. This dissertation aims to bridge this gap by linking vehicle drive cycles to battery cell duty cycles, which enables quantifying the impacts of real-world variability on battery performance. In addition to performance and efficiency considerations, the methodology enables battery aging analysis in the context of optimal design and control of hybrid electric vehicles. This will facilitate design decisions that ensure adequate performance over the life span of the vehicle with considerations of the battery health objective.

The novelty of this work lies in a more accurate method of synthetizing representative real-world drive cycles with a new algorithm to classify road and an innovative quantitative metric of driver style. A modified 48V mild hybrid vehicle model was built to relate the real-world drive cycles all the way to the battery cell duty cycles and to validate the impacts from driver aggressiveness on both the fuel efficiency and the battery loads. The cell duty cycles were further analyzed in frequency domain to synthesize characteristic cell test profiles representative of driver styles and road conditions. A battery cell cycle aging experiment was carried out using the synthesized

test profiles. Results validate the positive correlation between driver aggressiveness and cell degradation, and further allow parameter identification of cell electro-chemical model. Modeling effort was extended to generate insights regarding the aging mechanisms, and calibrate a semi-empirical aging model. These tools will enable the inclusion of road conditions and driver styles into the development of battery pack design and propulsion system control hence improving the design assumption fidelity and real-world representativeness of the modeling approach.

Table of Contents

Page

Table of Contents (Continued)

Page

List of Figures

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure Page

List of Figures (Continued)

List of Figures (Continued)

List of Figures (Continued)

Figure Page

List of Tables

List of Tables (Continued)

Table Page

1 Introduction

1.1 Motivation

Growing concerns on global warming and energy security have led to stringent fuel efficiency and CO_2 emission standards for new passenger vehicles. Nine countries and regions (including Japan, the European Union, United States, Canada, China, Australia, South Korea, Mexico, Brazil, and India), which together account for 75% of global fuel consumption by light-duty vehicles, have established mandatory regulations to boost fuel economy and reduce greenhouse gas (GHG) emission in near future [1]. The worldwide standards, as compiled in Table 1-1, mainly differ in the specified metrics (CO_2 emission, GHG emission or fuel economy), the structures (footprint, weight, or weight-class based corporate average), and the test cycles (NEDC, U.S. combined or JC08. However, the stringent trends are unanimous and the enacted targets converge, as shown in Figure 1-1.

Table 1-1. Global comparison of fuel economy/GHG standards for passenger cars [1]

Country or Region	Target Year	Standard Type	Unadjusted Fleet Target/Measure	Structure	Test Cycle
EU	2015 2021	CO_2	130 gCO_2/km 95 gCO_2/km	Weight-based corporate average	NEDC
China	2015 2020 (proposed	Fuel consumption	6.9 L/100km 5 L/100km	Weight-class based per vehicle and corporate average	NEDC
U.S.	2016 2025	Fuel economy/ GHG	36.2 mpg[i] or 225 gCO_2/mi 56.2 mpg[i] or 143 gCO_2/mi	FP-based corporate avg.	U.S. combined
Canada	2016 2025 (proposed)	GHG	217 gCO_2/mi[ii] N/A[iii]	FP-based corporate avg.	U.S. combined
Japan	2015 2020	Fuel economy	16.8 km/L 20.3 km/L	Weight-class based corporate average	JC08
Brazil	2017	Fuel economy	1.82 MJ/km	Weight-based corporate average	U.S. combined
India	2016 2021	CO_2	130 g/km 113 g/km	Weight-based corporate average	NEDC for low-powered vehicle
South Korea	2015	Fuel economy/GHG	17 km/L or 140 gCO_2/km	Weight-based corporate average	U.S. combined
Mexico	2016	Fuel economy/GHG	39.3 mpg or 140 g/km	FP-based corporate avg.	U.S. combined

Notes:

[i] Fuel economy standard by NHTSA assumes manufacturers fully use A/C refrigerant credit.

[ii] In April 2010, Canada announced a target for the light-duty vehicle fleet of 246 g/mi for MY2016. Separate targets for the car and light truck fleets are estimated by the ICCT based on the overall target.

[iii] Canada follows the US standards in the proposal, but the final target value would be based on the projected fleet footprints.

FP: footprint

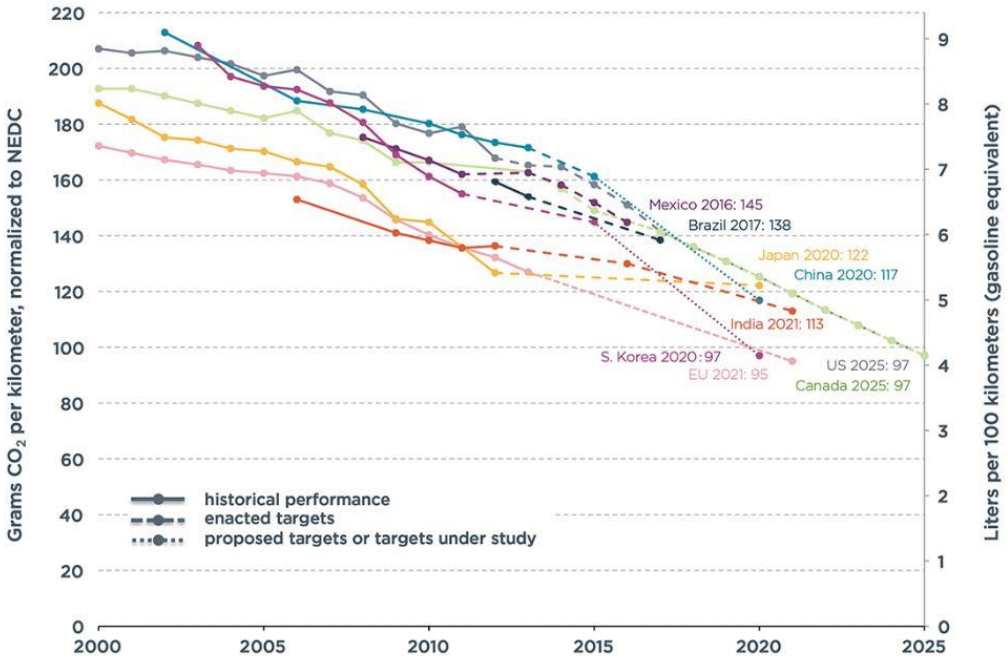


Figure 1-1. Historical fleet CO_2 emissions performance and current or proposed passenger vehicle standards [1]

Manufacturers respond to tightened standards by trying to find the right balance between fuel efficiency and cost of adopting fuel saving technology [2], which often requires complex schemes of fleet analysis and tradeoffs. The leading technological options can be mainly categorized by, but not limited to alternative fuels, advanced engine concepts, advanced transmission designs and electrification. The vehicle propulsion systems with partial or full degree of electrification pave an evolutionary path to sustainable mobility, yet the added costs of alternative energy storage systems are still hindering their wider market penetration. Compared to flywheel, ultra-capacitor and fuel cell, battery is an ideal compromise not only in the balance of energy and power density according to the Ragone plot in Figure 1-2, but also with respect to weight, volume, lifetime and price.

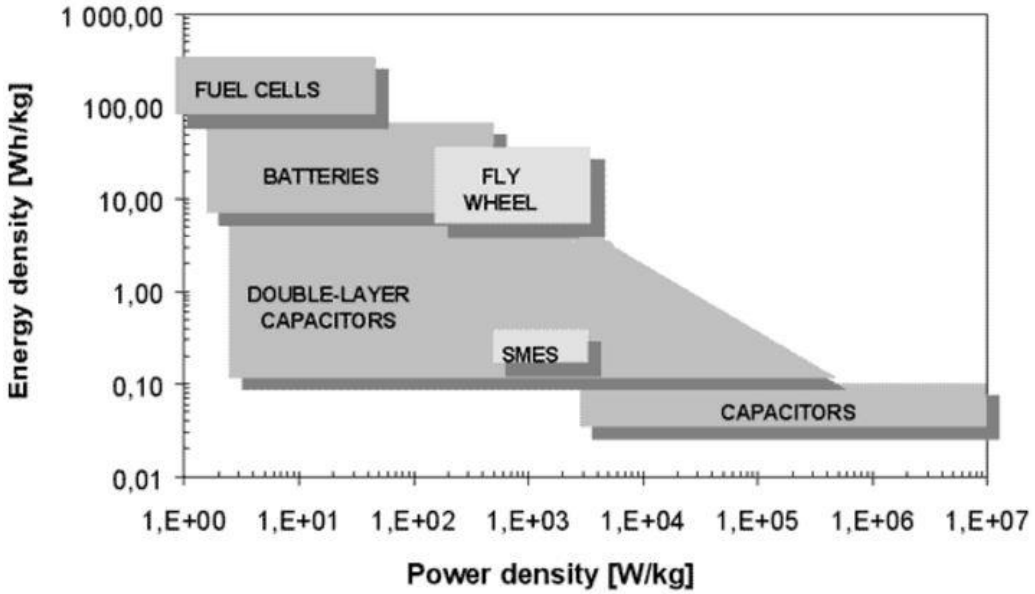


Figure 1-2. The Ragone plot of the energy storage and power handling capacity of alternative storage techniques [2]

The battery chemistries suitable for automotive applications evolve over time. Lead-acid batteries are the cheapest technology despite low energy density, lifetime and dynamic charge acceptance. Therefore, they are still used for Starting Lighting Ignition (SLI) functions and 12V power system, but cannot fully meet the energy storage requirements for electrified propulsion system. Nickel-metal hydride batteries have been used successfully in several commercial hybrid vehicles because of satisfactory power capability. Recently, lithium-ion batteries are taking over the market due to their improving trajectory along the diagonal of the Ragone plot [3] in Figure 1-3, though they are significantly more expensive. Another boost in energy density is expected from potential applications commercially viable Lithium-sulfur and metal-air batteries while challenges lie in power density, safety and lifetime [3].

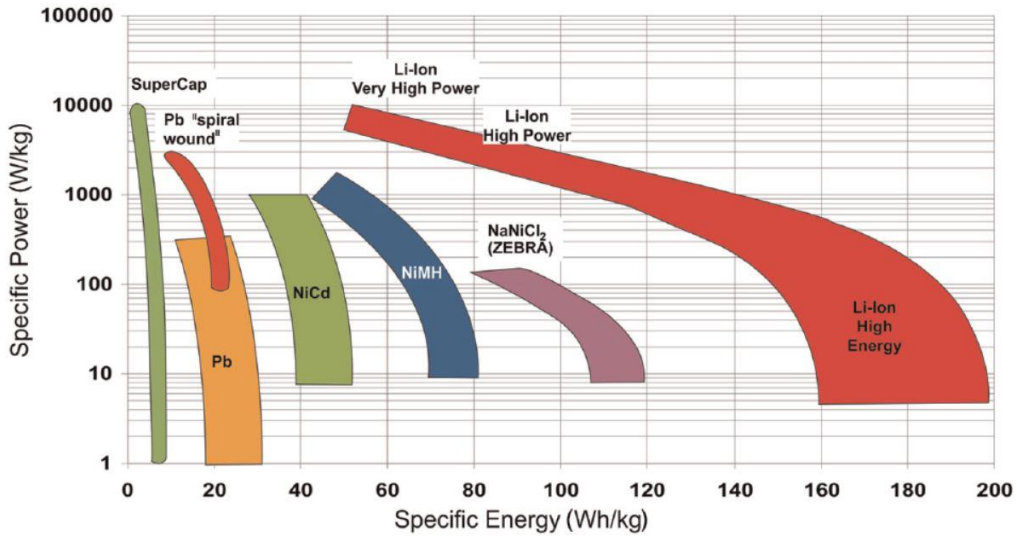


Figure 1-3. The Ragone plot of various battery technologies with specification at cell level for automotive applications [3]

The Lithium-ion batteries comprise a family of chemistries that employ various combinations of anode and cathode materials [4]. The anode materials are mainly carbon-based. The most prominent cathode materials are Lithium nickel cobalt aluminum (NCA), Lithium nickel manganese spinel (LMO), Lithium nickel manganese cobalt (NMC) and Lithium iron phosphate (LFP). The different Li-Ion chemistries are comprehensively compared in six dimensions: specific energy, specific power, performance under impact of temperatures, life span, cost and safety, referred to Figure 1-4. Being highly intertwined, these critical parameters require effective management of their mutual trade-offs, such as battery lifetime and cost.

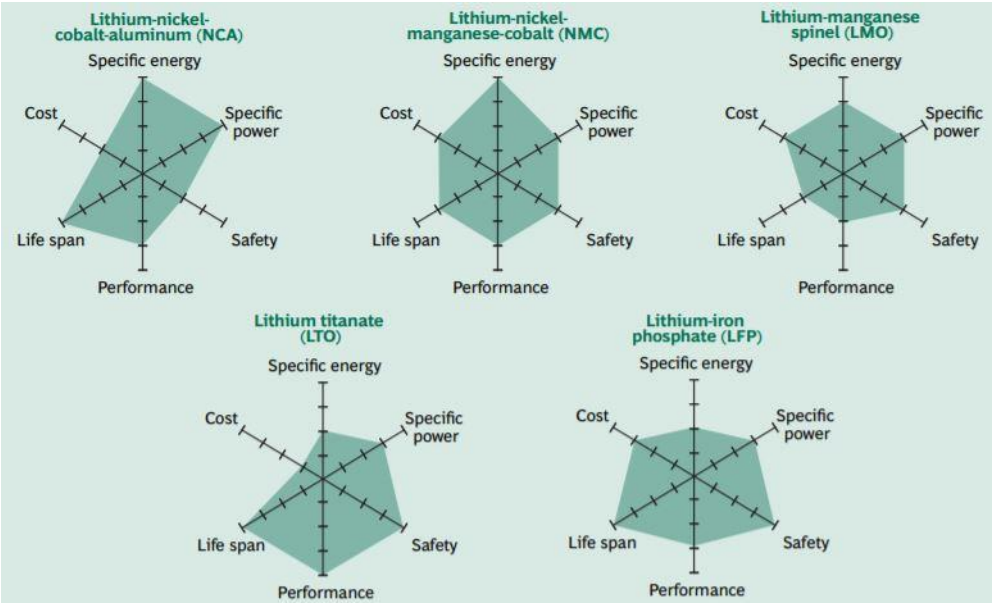


Figure 1-4. Tradeoffs among the five principal lithium-ion battery technologies [4]

As the breakthrough cost of battery packs for the massive commercialization of electrified vehicles is commonly considered at $150/kWh, an exponential reduction due to R&D development and economy-of-scale can be already observed since 2005 and predicted further into 2030 [5], in Figure 1-5. However, the customer acceptance largely depends on driving experience, of which battery lifespan is important besides expected range, acceleration, maintenance and charging issues. To overcome this challenge, manufacturers either oversize the pack to compensate expected degradation, or install less cells but replace them every five to seven years [6]. While such schemes all incur higher costs, current researches focus on achieving required system performance from a significantly downsized battery pack with minimized impacts on battery life under dynamic battery utilization. This requires critical understandings on vehicle operations and the methodologies to relate them to cells.

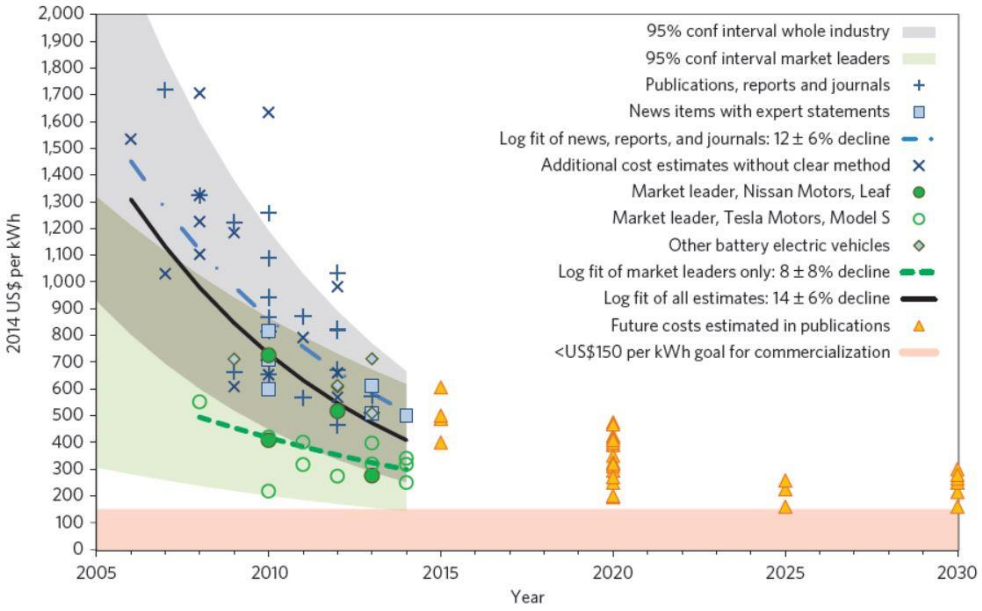


Figure 1-5. Cost of lithium-ion battery packs in battery electric vehicles [5]

The significance of understanding real-world vehicle operations has been translated into several multi-disciplinary collaborative efforts to collect real-world driving data. The National Household Travel Survey (NHTS) [7] summarize people's trip information such as duration, distance, and purpose, with additional information related to demography, geography, economy and vehicle attribute, in the form of questionnaire. The Transportation Secure Data Center (TSDC) by National Renewable Energy Laboratory analysis [8] further incorporates valuable second-by-second speed traces for potential applications including (i) transit planning and travel demand modeling, (ii) congestion mitigation research, (iii) toll and pricing research, (iv) climate change impact studies, (v) homeland security evacuation planning, (vi) validating transportation data from other sources, (vii) alternative fuel station planning, (viii) emission and air pollution modeling, (ix) vehicle energy and power analysis [8]. In certain context, these applications are inter-

dependent. For battery electric vehicles (BEV) and plug-in hybrid electric vehicles (PHEV), the application (i) can derive the cumulative daily distance distribution which can assist to guide the charging infrastructure deployment (vii), quantify the life-cycle cost of electrified mobility (viii), and calibrate the electric range (ix).

1.2 Background and Research Questions

1.2.1 Real-world Drive Cycles

The drive cycle, in the context of this dissertation, is defined as a speed trace time-series vehicle velocity data. The applied drive cycles in the literature can be mainly categorized in two ways: regulatory and real-world speed traces, transient and steady speed traces. The regulatory drive cycles refer to standard speed traces used in either fuel consumption or emissions testing for certification purposes under controlled laboratory settings. Examples include UDDS, HWFET, US06, and LA92 cycles in United States; NEDC, ARTEMIS cycles in Europe; 10-Mode, 10-15 Mode, JC08 cycles in Japan. Some of them are shown in Figure 1-6. In the figure, the NEDC and 10-15 Mode cycles are also the examples of steady speed traces, consisting of different steady speed segments, compared to other transient cycles.

However, beyond the certification testing application, the regulatory drive cycles are in lack of representativeness of real-world driving since they do not reflect the driving aggressiveness in real-world settings and do not have enough flexibility in cycle length [9][10]. Therefore, many efforts have been put into the collection of detailed real-world second-by-second drive cycles. The datasets applied in this dissertation come from the 2001~2002 Southern California Household Travel Survey, the 2010~2012 California

Household Travel Survey [8] and the 2001~2005 Michigan Road Departure Crash Warning System Field Operational Tests [11].

Faced with a large amount of driving data, two main research problems are the data compression and data dimensioning. The compression is to synthesize and condense bulk data into representative candidates for time-consuming and cost-intensive applications. The dimensioning is for categorization and classification based on underlying features of the data for better understanding and applicability. These two problems are both discussed in corresponding chapters to improve compression accuracy and refine two dimensioning features, in which appropriate statistical methodologies will be applied.

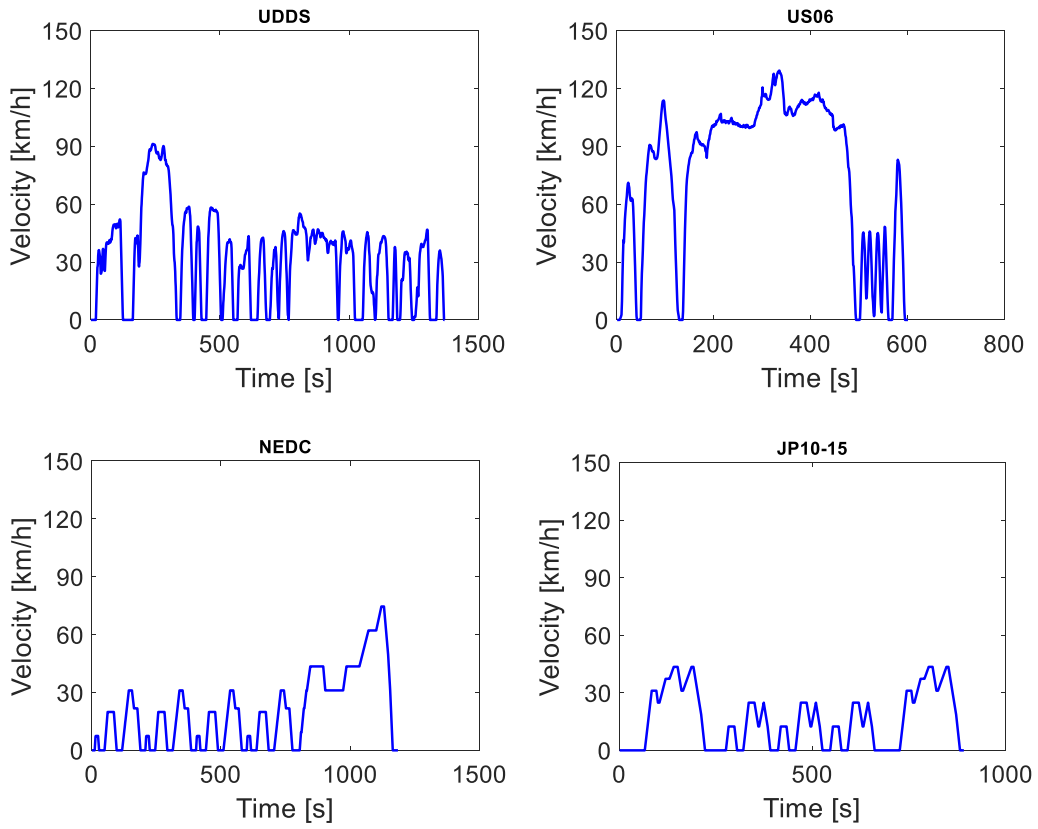


Figure 1-6. Selected regulatory drive cycles from US, Europe and Japan

1.2.2 Electrified Vehicles

To meet the stringent fuel economy and emission regulations, as mentioned above, electrified vehicles have become attractive options since they offer high efficiency by adding alternative energy sources on-board with multiple functions such as start&stop, braking energy regeneration, engine assisting, and electric driving. The electrified vehicles are comprised of Battery Electric Vehicle (BEV), Plug-in Hybrid Electric Vehicle (PHEV) to Hybrid Electric Vehicle (HEV). While BEV and PHEV can be recharged by plugging into the electric grid, they need large battery packs to store excess energy for long electric range with zero tailpipe emissions. But their market penetration is hindered by high prices, long charging time and charging infrastructure availability. This explains the earlier success of HEV, an interim electrification technology, with several widespread commercial vehicles (e.g., Toyota Prius) [12].

The HEV can be classified mainly in two ways [13]. Firstly, they come in series, parallel or series-parallel hybrid topologies as in Figure 1-7. The series HEV is conceptually an ICE (Internal Combustion Engine)-assisted EV to extend the driving range because the ICE is mechanically connected to a generator to convert electricity which charges either the battery or the motor. The parallel HEV combines both the ICE and the motor to mechanically propel the vehicle. The series-parallel HEV adds one more mechanical link between engine and wheel compared to series, and one more generator compared to parallel. Secondly, they can also be differentiated by the hybridization level, the power ratio between the ICE and the electric motor. The micro HEV applies a limited-power motor as a starter-generator to realize start&stop function. The mild HEV

uses a larger motor to achieve engine-assist and regeneration functions in addition to start&stop function. The full HEV integrates the largest motor to ensure occasional electric propulsion. Higher hybridization level usually indicates higher fuel savings.

Practically HEVs come in different topologies with different hybridization levels, which are proposed based on many factors such as cost, complexity and efficiency. And the trade-offs must be evaluated in the initial design phase under the context of real-world driving scenarios. The 48V parallel mild hybridization is a promising near-term technology because it requires minimal powertrain modifications other than coupling a belt-driven or crankshaft-integrated motor to the engine, yet bears great potential for fuel savings. The real-world performance a 48V mild parallel HEV will be discussed in corresponding chapters on both vehicle and cell level.

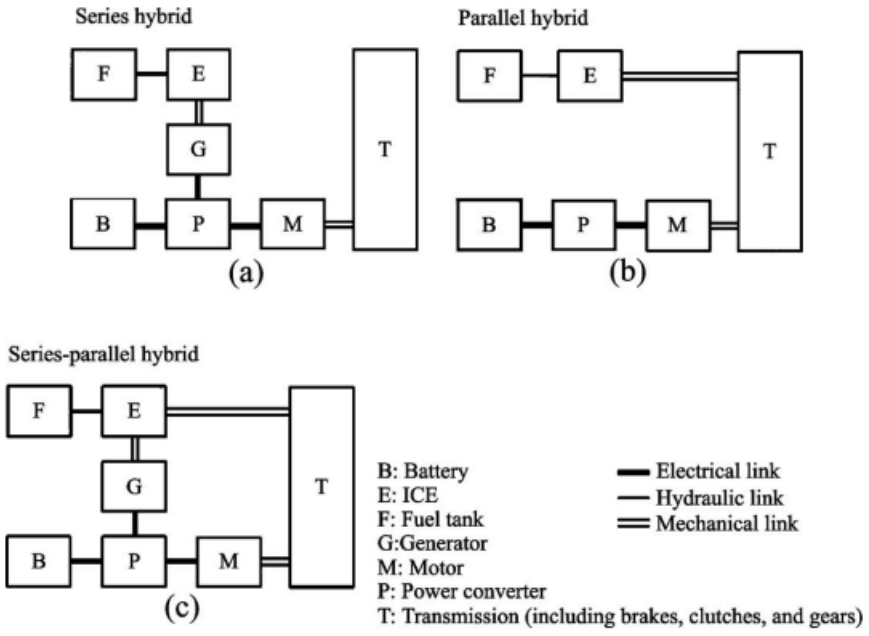


Figure 1-7. Common hybrid topologies of HEV [13]

1.2.3 Lithium-ion Battery Aging

1.2.3.1 Cell Operation

A lithium-ion battery, as in Figure 1-8, is typically composed of positive and negative electrodes immersed with electrolyte solution, a separator which prevents electron flows but allows ions migration, and current collectors at both electrodes to provide pathways for electron flows with external circuits. Common denotations for a battery include: positive electrode, negative electrode, anode, and cathode. They are illustrated in Figure 1-8 [14]. During discharging, the electron flow is from negative electrode to positive electrode, the current flow is in opposite direction; the positive electrode is the cathode accepting positive ions, and negative electrode is the anode accepting negative ions. During charging, current flow is from negative electrode to positive electrode, always contrary to electron flow; negative electrode becomes cathode attracting positive ions, and positive electrode becomes anode attracting negative ions.

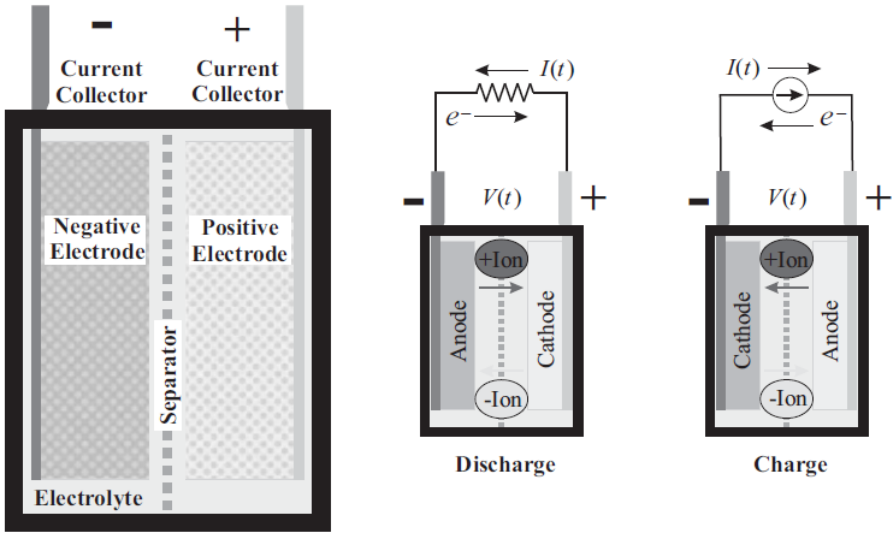


Figure 1-8. Lithium ion cell under charge and discharge [14]

1.2.3.2 Aging Mechanisms

Lithium-ion batteries suffer from degradation of performance over real-life utilizations, explicitly in the loss of available energy and power. As the usable cell capacity losses, available energy decreases affecting the electric driving range; as the cell internal resistance increases, available power reduces impacting the capabilities for acceleration and recuperation. The cell aging can occur under both storage conditions (calendar aging) and operating conditions (cycle aging). While all internal mechanisms are coupled thus complicating the cell aging process [15][16][17], an overview of main mechanisms of lithium-ion cells as in Figure 1-9 is necessary.

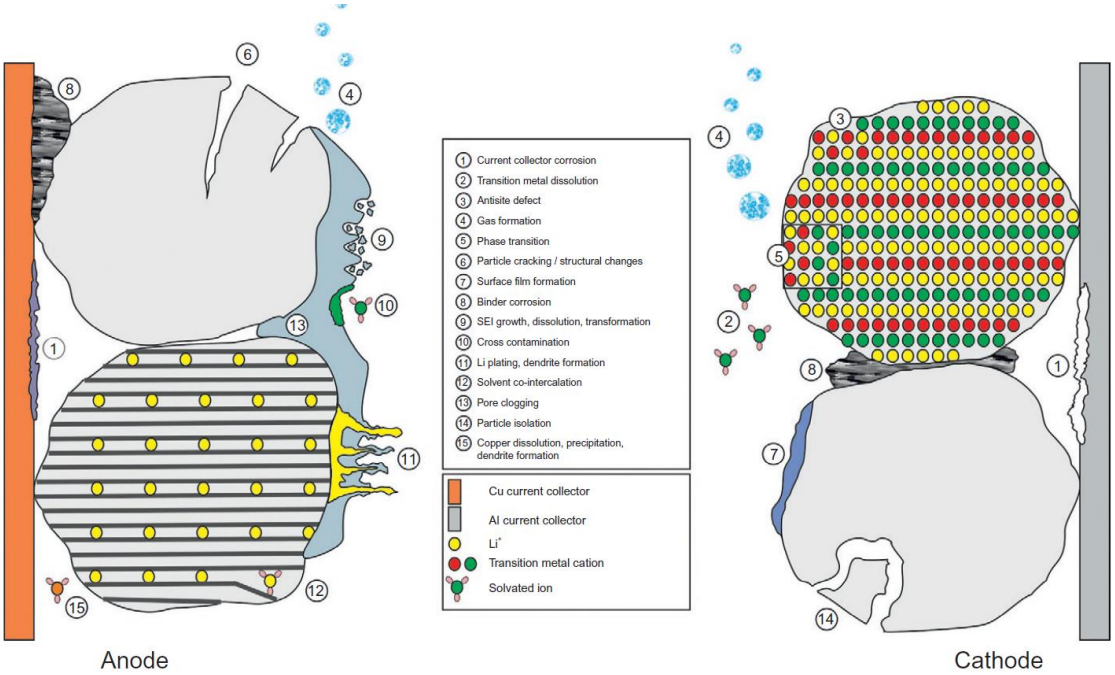


Figure 1-9. Lithium ion cell aging mechanisms [16]

The most popular material for lithium ion cell negative electrode is graphite. The graphite-based negative electrode operates typically at the voltage that is outside the

electrochemical stability window of the electrolyte components. Therefore, reductive electrolyte decomposition during charge accompanied with irreversible consumption of lithium ions occurs at the interface between negative electrode and electrolyte. The change at the interface is believed to be the major source for aging of negative electrode. At the beginning of cycling, the decomposition products build up “protective layers” at the electrode surface. The protective layers act as a so-called solid electrolyte interphase (SEI), which is permeable for lithium ions but rather impermeable for other electrolyte components and electrons. Hence, SEI does not have the perfect unity transference number for unsolved lithium cations, other charged and neutral species still diffuse through the SEI. As a consequence, the Li^+ loss and electrolyte decomposition continue throughout the entire battery life (⑨), though at lower rates compared to the first cycle. On a long time scale, the SEI might even penetrate into pores of the electrode and induce a decrease of accessible electrode active surface area (⑬), leading to impedance rise. The SEI growth process will be accelerated by elevated temperature, which can reversely break down SEI and bring a greater capacity loss (⑨). On the other hand, low temperature also poses significant challenges. It not only slows down Li^+ diffusion, but also results in metallic lithium plating and lithium dendrite growth during fast charging (⑪). Besides, the effects of transition metal dissolution and deposition at negative electrode exist with some positive electrode materials (especially $LiMn_2O_4$ spinel) to accelerate SEI growth (②⑩).

In contrast to negative electrode, positive electrode can be composed of different materials, among which $LiFeO_4$ (LFP), $Li(Ni_xMn_yCo_z)O_2$ (NMC), $Li(Ni_xCo_yAl_z)O_2$

(NCA) and $LiMn_2O_4$ (LMO) are widely used. The impacts of different aging mechanisms vary considerably when different cathode materials are considered. The significant changes on positive electrode include: 1) aging of active material (③⑤⑭) during high cycling current rates; 2) degradation of electrode components like conducting agents, binder, current collector (①⑧) at high temperature and high SOC; 3) oxidation of electrolyte components and surface film formation (⑦) at high temperature and when overcharging; 4) active material dissolution and the interactions with negative electrode (②), especially for LMO chemistry and when overcharging. Specifically, NMC electrode is characterized by a large impedance rise at positive electrode due to the formation of surface films; LFP and LMO electrodes are also affected by considerable transition metal dissolution which contributes to the growth of the negative electrode SEI; NMC and NCA electrodes are prone to phase transitions and anti-site cation defects of the active materials.

1.2.3.3 Aging Modeling

For any battery technology, the life in the long run is critical for its introduction and penetration into the market. Things get more complicated when lithium ion cells are applied in the transportation sector, where complex driving activities and extreme ambient factors are mixed. It has been shown that the real-world battery life depends heavily on the actual driver demands and the operating temperatures. In order to avoid risks and customer expectations, a reliable management of the battery pack with accurate estimations of battery health is required. Then, a battery aging model as a function of different aging stressors (current rate, SOC, temperature and so on) should be

parameterized to quantify the aging effects in real-world usage. Usually constructed from accelerated aging tests in the lab with the concept of design-of-experiments, the aging models can be used to estimate battery SOH (state of health) and predict battery SOL (state of life) even in an extrapolation manner. This dissertation aims to integrate the variability of real-world driving activities into the evaluation of lithium-ion cell aging to achieve improved predictiveness and enable battery health conscientious optimization..

1.3 Objectives and Dissertation Outline

This dissertation aims to integrate the real-world driving variability into the calibration of battery aging models in the context of optimal design and control of electrified vehicles. The inclusion of the real-world drive cycles requires more accurate synthesis and finer categorization. This comprises quantification of a metric for driver aggressiveness and an algorithm to classify road conditions. The drive cycles can then be further related to lithium-ion cell duty cycles in a more representative way via a vehicle simulation for impact characterization. A methodology needs to be proposed to integrate the real-world features extracted from the characterization study and enable their execution in the experimental aging tests in laboratory settings. The aging test results can hint on the aging mechanisms of selected lithium-ion battery chemistry through development of an aging model, but more over can directly identify the most influential features for battery degradation coming from the real-world driving scenarios. The results in this dissertation can support the engineering efforts in the new cell chemistry development but also facilitate the advancement of design decisions and control

strategies to guarantee lifetime performance of electrified vehicles with considerations of battery health objective. This dissertation is organized as follows.

In chapter 2, a higher order Markov Chain model is proposed for improved accuracy of representative drive cycle synthesis [18]. This approach guarantees to condense a large amount of real-world driving cycles into representative ones, useful for time-consuming and cost-intensive simulation studies. Chapter 3 respectively proposes a new algorithm to classify the road conditions [19] and defines an original metric to quantify the driver style [20][21], which allows finer data categorization. Subsequently, the real-world drive cycles are evaluated through a MATLAB/Simulink into a model of 48V mild hybrid vehicle. While identifying relationships between fuel saving potential and battery current load through the impact from the driver styles, the drive cycles are simultaneously related to cell duty cycles, in chapter 4 [21]. A characterization of the cell duty cycles in frequency domain lays the foundation for the synthesis of representative test profiles, repeatable in laboratory settings for aging tests, in chapter 5 [22]. The aging results of commercial 18650 NMC lithium-ion cells are used to identify the parameters of single-particle battery model with a stochastic algorithm, and set up the semi-empirical aging model in chapter 6 [23]. The overall summary and suggestions for future work is discussed in chapter 7.

2 The Synthesis of Real-World Drive Cycles

Previous studies consider the impacts of the drive cycles on electrified vehicle design [24], as well as the large-scale life-cycle energy analysis with electrified fleets [25][26]. They recognized the drawbacks of using just the federal drive cycles are recognized and suggested application of the naturalistic driving patterns instead. Nevertheless, applying the whole naturalistic driving database directly is time-consuming since it involves thousands of executions, while it also makes the results less tractable and difficult to analyze. Vehicle system evaluation needs to be done in a more compact way using only a limited set of highly representative drive cycles. The goal of this chapter is to develop a method for an accurate synthesis of naturalistic drive cycles that provides a realistic representation of real-world vehicle usage patterns.

A procedure for naturalistic cycle synthesis is shown in Figure 2-1. It consists of collection of naturalistic cycles as the first step. Categorization of naturalistic cycles based on variables related to driving pattern and road condition is the second step. The third step synthesizes numerous candidate drive cycles in stochastic manner. Finally, through the regression analysis based on a set of significant cycle metrics, representative cycles are chosen.

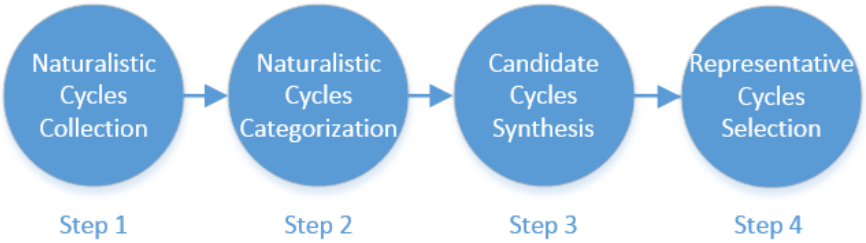


Figure 2-1. Procedure for generating representative cycles

The main objective of drive cycle categorization is to divide cycles into different groups, expecting to aggregate similar drive cycles in each category. A much better chance of success in analysis and synthesis of representative cycles will be achieved if low-speed stop-and-go cycles are separated from high-speed free-flow highway cycles.

The synthesis of candidate drive cycles involves a stochastic recombination of drive cycle segments. Common definitions of drive cycle segments range from micro-trips between vehicle stops [27], to acceleration/cruise/deceleration events related to emissions [28], and to velocity points under maximum recording resolution [26]. The stochastic recombination process relies on the premise that future segment is strongly related to past ones due to vehicle physical limitations and human behavioral patterns. The relationship is captured with probability distribution. Markov chain is a frequently-used mathematical model to capture the probability of driver activities due to its simplicity with "memoryless" character [29]. Analysis assumes that future state only depends on the current state. A data-driven exploration of Markov chain has been done in a previous study [30] but effects from different Markov chain orders have not been addressed, which forms the core part of this comparison study.

After synthesizing numerous candidate cycles for each cycle category, there should be certain criteria upon which some satisfactory cycles can be selected as the representatives for each category. This step is often called representativeness validation. Previous criteria include significant cycle metrics [29] and Speed/Acceleration Frequency Distribution (SAFD) [28]. This study uses 7 significant cycle metrics after a systematic regression analysis [29]. Then the candidate drive cycle whose significant metrics are closest to mean values of these in corresponding cycle category is regarded as the most representative drive cycle.

This chapter starts with naturalistic drive cycle categorization. Impacts of Markov chain of different orders on cycle synthesis performance are assessed next. Solutions and results follow; the chapter ends with the conclusion.

2.1 Naturalistic Drive Cycles Categorization

The data source for this study is a portion of the GPS-enhanced regional household travel survey from June 2001 to March 2002, conducted by the Southern California Association of Government (SCAG). It was made publically accessible by the U.S. Department of Energy's National Renewable Energy Laboratory [8]. This dataset includes filtered drive cycles after removing outlying high and low speeds, data spikes and infeasible accelerations. A total of 612 drive cycles from 292 passenger vehicles are used in this work.

Since our acceptance criteria for a representative drive cycle is the mean-value approximation error, large variance of cycle characteristics will have a detrimental effect. In other words, mixing significantly different driving patterns together, such as low-speed

city driving and highway cruising, will lead to the synthetic cycles might not capture the characteristics from either. This study categorizes the drive cycles based on cycle distance as in a previous study [29].

In Figure 2-2, cycle distances are classified into bins with 2 km increments. Very short cycles less than 2 km are intentionally neglected. The cumulative distribution for the trip distance is also included. We categorize trips by splitting the cumulative distribution equally. Three cycle categories, short, medium and long, are identified. As indicated by the star notation in Fig 2, cycles with the distance less than 4 km cover nearly 33% of people's driving activity, another 33% of trips fall between 4 km and 10 km and the rest 33% are more than 10 km. This categorization is used for subsequent comparison study.

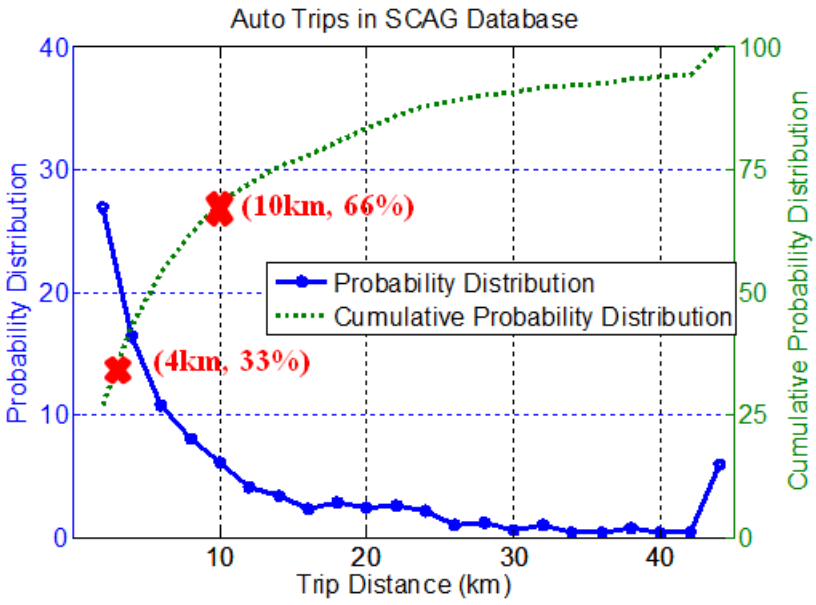


Figure 2-2. Distribution of cycle distances in the SCAG database

Typical driving patterns for three driving distance categories are shown in Figure 2-3. A clear increasing trend of cycle duration and mean velocity is observed when moving from a short cycle to a long cycle, while the idle time percentage and stops per kilometer decrease.

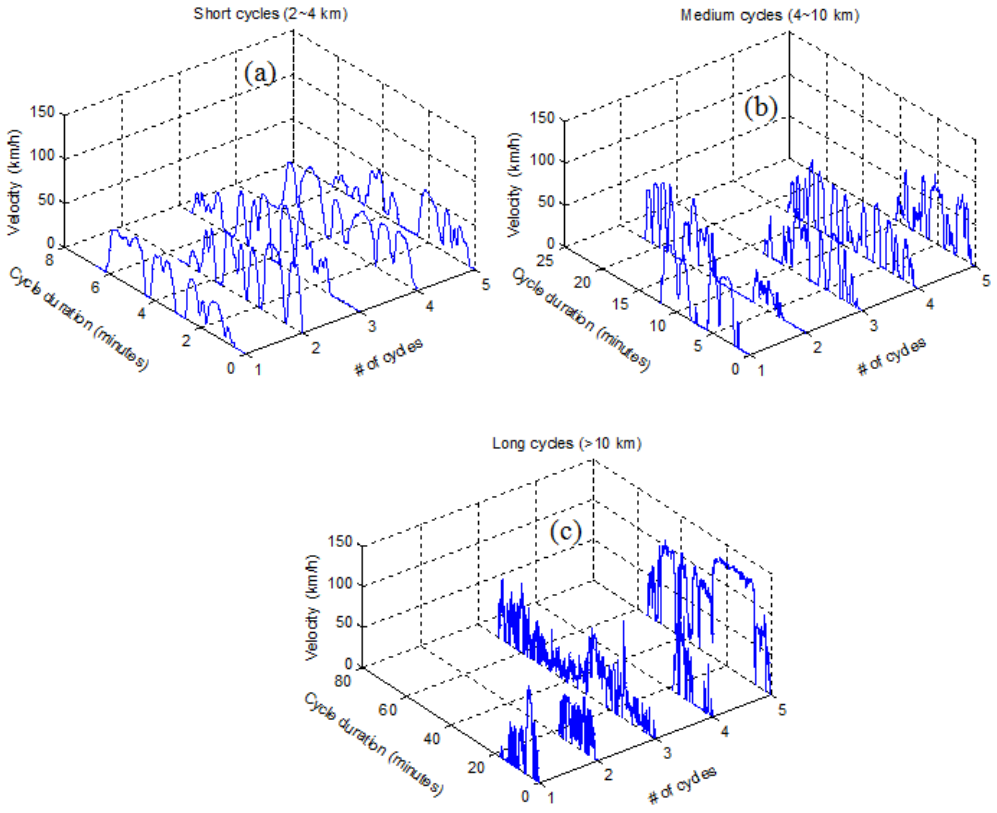


Figure 2-3. Typical naturalistic drive cycles for (a) short, (b) medium and (c) long distance categories

2.2 Problem Formulation

After drive cycle categorization based on cycle distance, cycles in each category are used to generate Transition Probability Matrices (TPM) and subsequently apply Markov chains of different orders. The specific forms of Markov chains are derived next.

The naturalistic driving cycles are considered as a sequence of transitions between different states x_k at time k. Acceleration and speed are reliable representations of vehicle's instantaneous states based on vehicle dynamics. They are combined to represent a state in Markov chain [29]. Each of them is discretized and assumed to take on a finite number of values, $\{a_1, a_2, \ldots, a_m\}$ and $\{v_1, v_2, \ldots, v_n\}$, into which any sequence of observations {a, v} can be mapped. According to Markov Chain property, the future probability of transitions between two successive states depends only on the current state [31][32], shown as:

$$\Pr\{x_{k+1} \mid x_k, x_{k-1}, x_{k-2} \ldots x_1\} = \Pr\{x_{k+1} \mid x_k\} \quad \text{Eq. 2.1}$$

$\Pr\{x_{k+1} \mid x_k, x_{k-1}, x_{k-2} \ldots x_1\}$ denotes the conditional probability distribution of x_{k+1} given the occurrence of $x_k, x_{k-1}, x_{k-2}, \ldots, x_1$. $\Pr\{x_{k+1} \mid x_k\}$ denotes the conditional probability distribution of x_{k+1} given the occurrence of only x_k.

In [26], the state x_k of Markov Chain takes the specific form of (v_k, a_k):

$$\Pr\{v_{k+1}, a_{k+1} \mid v_k, a_k\} \quad \text{Eq. 2.2}$$

Then, v_{k+1} is calculated based on the approximation in Eq. 2.3, Δt is the time resolution, 1 second in this study, the same following.

$$\dot{v}_{k+1} = a_{k+1} \approx (v_{k+1} - v_k) / \Delta t = (v_{k+1} - v_k) \quad \text{Eq. 2.3}$$

Extending Eq. 2.2 by Eq. 2.3 to get Eq. 2.4, we can see that information of previous states is actually taken into account.

$$\Pr\{v_{k+1}, v_{k+1}\text{-}v_k \mid v_k, v_k\text{-}v_{k-1}\} \quad \text{Eq. 2.4}$$

Rewrite Eq. 2.4 in Eq. 2.5:

$$\Pr\{v_{k+1}|v_k, v_{k-1}\} \quad \text{Eq. 2.5}$$

The Eq. 2.5 not only allows feasibility to use only velocity as states instead of Eq. 2.2 for a 2-order Markov chain, but also results in a compact form when trying to consider more past states. A 3-order Markov chain implicitly brings a_{k-1} with the introduction of v_{k-2} in Eq. 2.6.

$$\Pr\{v_{k+1}|v_k, v_{k-1}, v_{k-2}\} \quad \text{Eq. 2.6}$$

Different forms of Markov chain are summarized in Table 2-1. While 2-order Markov chain uses the combination of v_k and v_{k-1} as the vehicle state at time k, 3-order Markov chain adds v_{k-2} to the combination as well.

Table 2-1. Summary of forms of Markov chain in this study

Standard Form	Derivative Form	Notation
$\Pr\{v_{k+1+}, a_{k+1} \mid v_k, a_k\}$	$\Pr\{v_{k+1}\|v_k, v_{k-1}\}$	2-order Markov chain
$\Pr\{v_{k+1}, a_{k+1} \mid v_k, a_k; v_{k-1}, a_{k-1}\}$	$\Pr\{v_{k+1}\|v_k, v_{k-1}, v_{k-2}\}$	3-order Markov chain

Moreover, the exemplary Transition Probability Matrix (TPM) for both 2-order and 3-order Markov chains could be constructed as shown in Figure 2-4 (a) and (b), by sampling the occurrences of velocity states from the naturalistic cycle database. The 0.2 m/s (≈0.8 km/h) for velocity discretization, equaling to nearly 200 values, is used. The real velocity data are mapped onto these state grids. Relying on the derived TPM, the cycle synthesis process begins from standstill, and then moves to each subsequent state in a stochastic manner based on respective cumulative probability distribution, as shown in Figure 2-4 (c), until reaching the destination and ending again in standstill.

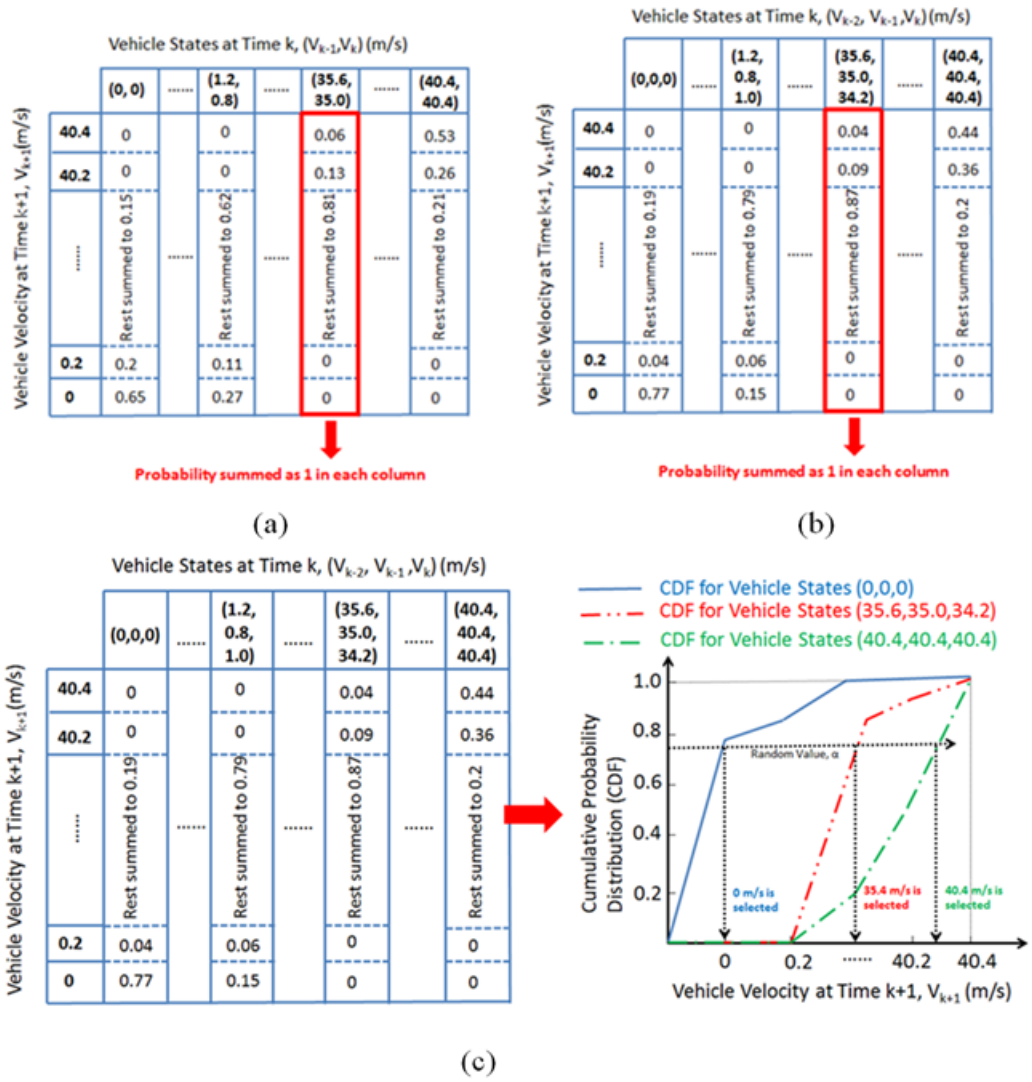


Figure 2-4.a) TPM for 2-order Markov chain; b) TPM for 3-order Markov Chain; c) Different cumulative probability distributions for different vehicle states and one-step vehicle state transition

Next section presents a comprehensive comparison of the cycle synthesis performance by both 2-order and 3-order Markov chain, using the medium-distance (4-10 km) cycle category for illustration.

2.3 Methods

To compare the cycle synthesis performance, a performance index should be defined firstly. The index describes how closely a synthesized candidate drive cycle follows the mean values of significant metrics from the corresponding cycle categories, which are compiled in Table 2-2. Then the distribution of the performance index can be investigated to quantify the cycle synthesis quality achieved by different forms of Markov chain.

2.3.1 Definition of Performance Index

After using Markov chain to generate a candidate cycle of a certain length, say 6.56 km (mean distance of cycles between 4~10km), the ideal outcome would lead to minimal differences between any metric of the synthesized cycle and the mean value of the corresponding metric calculated for the whole category. Therefore in this study, the deviations of the mean values of significant metrics in certain cycle category are summed up and defined as the Metric Errors for any candidate cycle.

$$Metric\ Errors = \sum_{i=1}^{N} \frac{|\bar{X}_i - Y_i|}{|\bar{X}_i|}$$

$\bar{X}_i$: Mean Values of i_{th} interested cycle metrics from certain category of naturalistic cycles.

Y_i: Value of i_{th} interested cycle metrics from a candidate cycle.

N: Number of significant cycle metrics.

Table 2-2. Summary of significant cycle metrics for short, medium and long-distance naturalistic drive cycles and their statistics

Significant Cycle Metrics	Mean Values for Cycles (< 4 km)	Mean Values for Cycles (4~10km)	Mean Values for Cycles (>10 km)
Standard deviation of velocity (km/h)	21.79	26.06	34.50
Mean positive velocity (km/h)	35.30	43.52	62.01
Standard deviation of acceleration (m/s^2)	0.66	0.67	0.59
Minimum acceleration (m/s^2)	-2.47	-2.77	-2.99
Percentage of driving time under negative acceleration (%)	35.99	36.82	44.10
Percentage of driving time under positive acceleration (%)	39.06	40.89	50.01
Number of stops/km (1/km)	1.52	1.03	0.50

2.3.2 Performance Index Comparison

To select representative cycles in each cycle category, numerous candidate cycles need to be synthesized. The distributions of Metric Errors from these candidate cycles are then compared to evaluate cycle synthesis performance with two different forms of Markov chain. Two features of Metric Error distributions are of particular interests, namely, the mean value of Metric Error and the minimum value of Metric Error. They will be used to evaluate the differences between 2-order and 3-order Markov chain cycle synthesis methods. However, regarding the statistical significance of the differences of the mean value of Metric Error and the minimum value of Metric Error, hypothesis tests should be set up to obtain a convincing proof.

2.3.2.1 Hypothesis Test on the Decrease of Mean Metric Error

While numerous candidate cycles can be synthesized, a practical problem arises: how many candidate cycles should be synthesized to estimate the population mean value of

Metric Errors, μ, in certain confidence interval, (1-α)100%, within certain estimation error, ε?

The formula for a (1-α)100% confidence interval for a population mean μ:

$$\mu = \overline{Metric\ Error} \pm \varepsilon = \overline{Metric\ Error} \pm t_{\frac{\alpha}{2},n-1}\left(\frac{s}{\sqrt{n}}\right) \quad \text{Eq. 2.7}$$

$\overline{Metric\ Error}$ is the mean value of the sample, n is the size of the sample, t denotes the t-distribution, s is the standard deviation of the sample.

$$\varepsilon = t_{\frac{\alpha}{2},n-1}\left(\frac{s}{\sqrt{n}}\right) \quad \text{Eq. 2.8}$$

$$n = \frac{\left(t_{\frac{\alpha}{2},n-1}\right)^2 s^2}{\varepsilon^2} \quad \text{Eq. 2.9}$$

As sample size n increases, the t-distribution approaches the standard normal distribution, denoted by Z in Eq. 2.10.

$$n \approx \frac{\left(z_{\frac{\alpha}{2}}\right)^2 s^2}{\varepsilon^2} \quad \text{Eq. 2.10}$$

To solve n, the standard deviation, s, and the estimation error, ε, should be fixed. They can be inferred from a small pilot study on 100 candidate cycles, whose statistics are compiled in Table 2-3.

Table 2-3. A pilot study on mean metric error with a sample of size 100 in medium-distance cycle category

	No. of Candidate Cycles	Mean	Standard Deviation
2-order Markov chain	100	1.2723	0.4220
3-order Markov chain	100	1.1881	0.3776

From the results of the small-scale pilot study, the standard deviations for both forms of Markov chain are used directly, and the estimation error ε is set loosely at 0.02. Besides, the confidence interval is set at 95% so $z_{\frac{\alpha}{2}}$ = 1.96. Then minimal sample sizes for both methods can be determined as:

$$n_{2\text{-order}} = \frac{1.96^2 0.4220^2}{0.02^2} = 1711 \quad \text{Eq. 2.11}$$

$$n_{3\text{-order}} = \frac{1.96^2 0.3776^2}{0.02^2} = 1370 \quad \text{Eq. 2.12}$$

Therefore, the sample size is chosen as 2000 for following application.

Based on the results from the previous subsection, 2000 candidate drive cycles are synthesized by 2-order and 3-order Markov chain respectively. Their statistics are summarized in Table 2-4.

Table 2-4. The study on mean metric error with a sample of size 2000 in medium-distance cycle category

	No. of Candidate Cycles	Mean	Standard Deviation
2-order Markov chain	2000	1.2802	0.4449
3-order Markov chain	2000	1.2430	0.4278

A decrease of mean values of Metric Error from 1.2802 to 1.2430 can be achieved by using a 3-order Markov chain. However, the statistical significance of the decrease should be validated, relying on hypothesis test.

Hypothesis test is oriented towards making decisions and moving ahead with actions. Basic procedures for the test is to (1) formulate a null hypothesis H0 and, if appropriate, an alternative hypothesis H1; (2) design a test procedure by which a decision can be

made; (3) Use statistics to finalize the test procedure, making judgments about the significance level; (4) Apply the test to individual cases; (5) Make decisions.

The two-sample t-test [33] is applied to determine if two population means are equal. The purpose in this study is to test if the 3-order Markov chain is superior to 2-order Markov chain in reducing the mean Metric Error when synthesizing representative drive cycles.

In this study, the two-sample t-test is defined as:

Null hypothesis, $H_0 : \mu_{2\text{-order}} = \mu_{3\text{-order}}$

Alternative hypothesis, $H_a : \mu_{2\text{-order}} > \mu_{3\text{-order}}$

$\mu_{2\text{-order}}$ and $\mu_{3\text{-order}}$ are the population mean of Metric Error by 2-order and 3-order Markov chain respectively.

Test Statistics: T =

$$\frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{{s_1}^2}{N_1} + \frac{{s_2}^2}{N_2}}} \qquad \text{Eq. 2.13}$$

where N_1 and N_2 are the sample sizes, both 2000 in this study; $\overline{X_1}$ and $\overline{X_2}$ are the sample means for 2-order and 3-order Markov chain; ${s_1}^2$ and ${s_2}^2$ are the sample variances for 2-order and 3-order Markov chain.

The significance level α is set at 0.05.

The rejection for the null hypothesis is $T > t_{1-\alpha,v}$, where $t_{1-\alpha,v}$ is the critical value of the t distribution with v degrees of freedom.

$$v = \frac{(\frac{{s_1}^2}{N_1} + \frac{{s_2}^2}{N_2})^2}{\frac{({s_1}^2/N_1)^2}{N_1 - 1} + \frac{({s_2}^2/N_2)^2}{N_2 - 1}} \quad \text{Eq. 2.14}$$

When statistics from Table 2-4 are inserted into above equations, the results are T = 2.6954, and $t_{1-\alpha,v}$ = 1.645. Therefore the criteria of rejection for the null hypothesis is met, $\mu_{2\text{-order}} > \mu_{3\text{-order}}$. In other words, the difference between the mean values of Metric Error population from the candidate cycles by 2-order and 3-order Markov chain is statistically significant; the 3-order Markov chain can indeed improve accuracy.

2.3.2.2 Hypothesis Test on the Decrease of Minimum Metric Error

Another meaningful statistical feature for assessing the performance of a Markov chain is the minimum Metric Error. From a sample of 2000 Metric Errors, only the minimum one will be from the representative cycle. For example, if 10 samples of 2000 candidate cycles, totally 20,000 cycles, are synthesized, only 10 minimum Metric Errors, representative cycles correspondingly, will be achieved. Then two similar questions as above arise again: is there any difference between the minimum Metric Errors? If any, is the difference statistically significant? To answer them, certain amount of minimum Metric Errors should be attained first; in other words, several iterations of synthesizing 2000 candidate cycles should be done. Then the hypothesis test procedures can be set up again by replacing the mean Metric Error with the mean Minimum Metric Error.

The first step is to determine how many minimum Metric Errors should be attained. The statistics of a pilot study of size 10 are in Table 2-5.

Table 2-5. A pilot study on mean minimum metric error with a sample of size 10 in medium-distance cycle category

	No. of 2000-sized Cycle Sample	Mean Minimum Metric Error	Standard Deviation
2-order Markov chain	10	0.4188	0.0455
3-order Markov chain	10	0.3882	0.0309

Setting estimation error at 0.01, refer to Eq. 2.10:

$$n_{\text{2-order, minimum}} = \frac{1.96^2 0.0455^2}{0.01^2} = 79.53$$

$$n_{\text{3-order, minimum}} = \frac{1.96^2 0.0309^2}{0.01^2} = 36.68$$

Therefore, 80 iterations of 2000 candidate cycle synthesis will be conducted for both forms of Markov chain. Thus 80 minimum Metric Errors can be attained for each form of Markov chain.

The statistics of the 80 iterations of 2000 candidate cycle synthesis are summarized in Table 2-6. A two-sample t-test is proposed to test the significance of the decrease of mean Minimum Metric Error.

In this study, the two-sample t-test is defined as:

Null hypothesis, $H_0 : \mu_{\text{2-order, minimum}} = \mu_{\text{3-order, minimum}}$

Alternative hypothesis, $H_a : \mu_{\text{2-order, minimum}} > \mu_{\text{3-order, minimum}}$

Table 2-6. The study on mean minimum metric error with a sample of size 80 in medium-distance cycle category

	No. of 2000-sized Cycle Sample	Mean Minimum Metric Error	Standard Deviation
2-order Markov chain	80	0.4037	0.0477
3-order Markov chain	80	0.3869	0.0442

$\mu_{2\text{-order, minimum}}$ and $\mu_{3\text{-order, minimum}}$ are the population mean of Minimum Metric Error by 2-order and 3-order Markov chain respectively.

According to Eq. 2.13 and Eq. 2.14 at the same significance level (0.05) and the statistics in Table 2-6, the test statistics T = 2.3107 > $t_{1-\alpha,v}$=1.660, indicating rejection to the null hypothesis while accepting the alternative hypothesis. In other words, with 3-order Markov chain, the minimum Metric Error is significantly reduced in a statistical sense.

Summing up, with the new cycle synthesis method based on the 3-order Markov chain, the Metric Error as a performance index decreases in both mean value and minimum value. The significance is statistically validated with hypothesis tests. Though not shown in detail, the decreases by 3-order Markov chain apply to short and long distance cycle categories as well.

2.4 Results

For every cycle category, 80 2000-sized cycle samples have been attained with two forms of Markov chain. From the 160,000 cycles in each cycle category, the 7 significant cycle metrics of the most representative cycles (with least Metric Error) with both forms of Markov chain are compiled in columns 3 and 4 of Table 2-7. The mean values of significant cycle metrics of all naturalistic cycles in each category are also shown in the column 2 of Table 2-7 for better illustration of representativeness.

Table 2-7. The statistics of the most representative drive cycles by 2-order and 3-order Markov chains

Significant Cycle Metrics	Mean Values for Cycles (< 4 km)	2-order Markov Chain (Deviations from mean)	3-order Markov chain (Deviations from mean)
Standard deviation of velocity (km/h)	21.79	21.05 (3.42%)	22.32 (2.42%)
Mean positive velocity (km/h)	35.30	36.11 (2.29%)	35.06 (0.7%)
Standard deviation of acceleration (m/s^2)	0.66	0.74 (11.80%)	0.73 (10.70%)
Minimum acceleration (m/s^2)	-2.47	-2.45 (0.93%)	-2.44 (0.93%)
Percentage of driving time under negative acceleration (%)	35.99	34.36 (4.54%)	35.26 (2.03%)
Percentage of driving time under positive acceleration (%)	39.06	36.81 (5.76%)	36.99 (5.28%)
Number of stops/km (1/km)	1.52	1.41 (7.46%)	1.44 (5.68%)
Metric Errors	N/A	36.19%	27.75%

Significant Cycle Metrics	Mean Values for Cycles (4-10 km)	2-order Markov Chain (Deviations from mean)	3-order Markov chain (Deviations from mean)
Standard deviation of velocity (km/h)	26.06	25.53 (2.04%)	25.53 (2.04%)
Mean positive velocity (km/h)	43.52	41.20 (5.33%)	43.61 (0.20%)
Standard deviation of acceleration (m/s^2)	0.67	0.71 (6.06%)	0.73 (9.43%)
Minimum acceleration (m/s^2)	-2.77	-2.67 (3.56%)	-2.89 (4.48%)
Percentage of driving time under negative acceleration (%)	36.82	34.13 (7.30%)	36.85 (0.08%)
Percentage of driving time under positive acceleration (%)	40.89	39.87 (2.49%)	37.01 (9.48%)
Number of stops/km (1/km)	1.03	1.09 (6.13%)	1.02 (0.26%)
Metric Errors	N/A	32.91%	25.98%

Significant Cycle Metrics	Mean Values for Cycles (>10 km)	2-order Markov Chain (Deviations from mean)	3-order Markov chain (Deviations from mean)
Standard deviation of velocity (km/h)	34.50	33.37 (2.04%)	34.68 (0.51%)
Mean positive velocity (km/h)	62.01	55.42 (10.62%)	57.83 (6.74%)
Standard deviation of acceleration (m/s^2)	0.59	0.57 (2.55%)	0.64 (8.49%)
Minimum acceleration (m/s^2)	-2.99	-3.11 (4.21%)	-2.89 (3.23%)
Percentage of driving time under negative acceleration (%)	44.10	29.45 (26.15%)	30.86 (22.61%)
Percentage of driving time under positive acceleration (%)	50.01	35.08 (20.46%)	35.42 (19.68%)
Number of stops/km (1/km)	0.50	0.50 (0.55%)	0.50 (0.64%)
Metric Errors	N/A	67.83%	61.89%

From above results, 3-order Markov chain is capable of synthesizing more representative drive cycles. Their significant cycles approach mean values better. Worthy of note are the large Metric Errors for long-distance cycle category. One possible reason is that driving patterns in this category have more variability than counterparts in other categories, implying a need for a better cycle categorization in this long-distance range. In fact, using 10km as a transition from medium to long-distance driving activity is imprecise because many city trips would be longer than 10km. However, for a comparison study this inaccuracy is within tolerance. For real applications of representative drive cycle synthesis, this point should be taken into account and finer cycle categorization is recommended. The Figure 2-5 below illustrates the speed traces for the 3 most representative drive cycles.

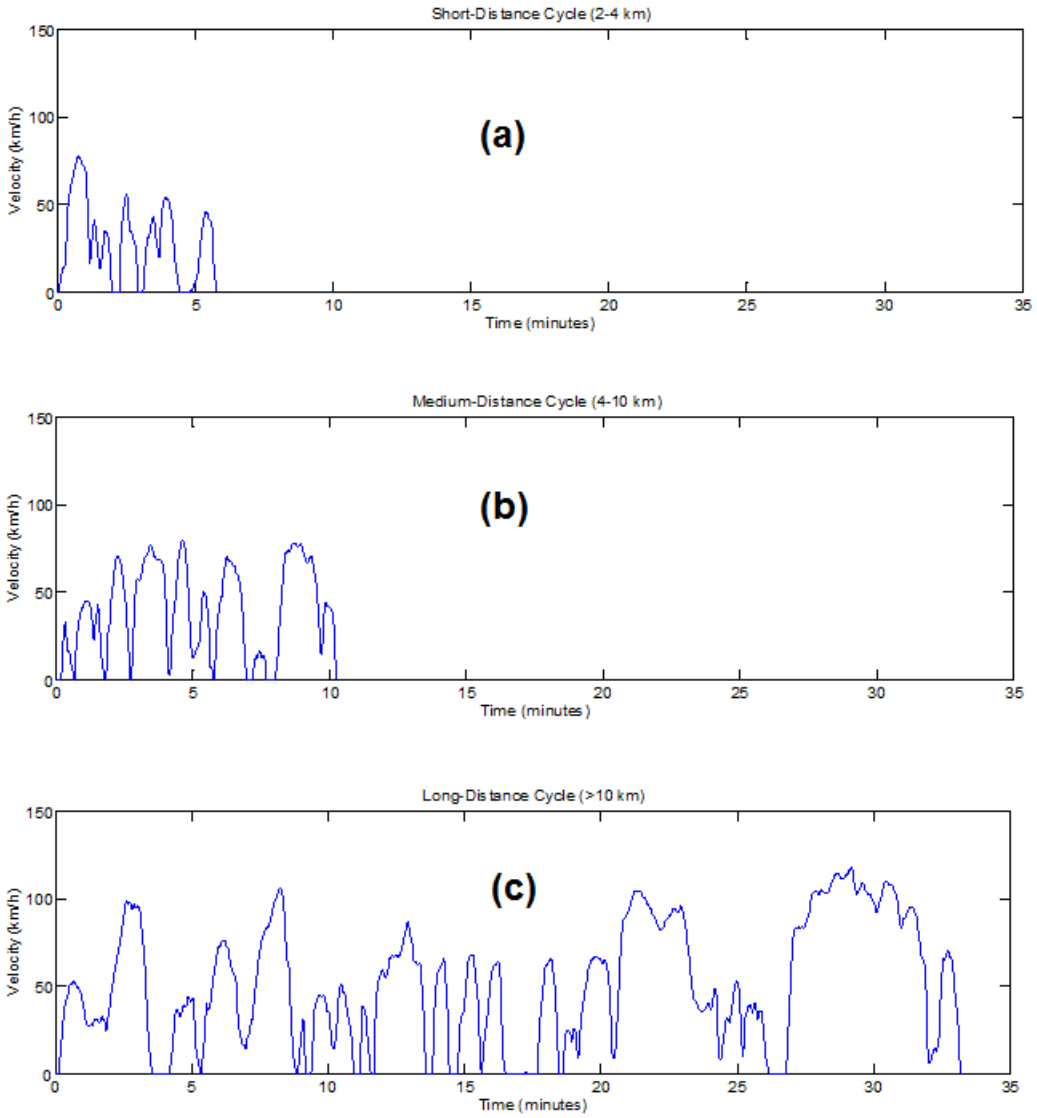


Figure 2-5. Representative drive cycles for (a) short, (b) medium, (c) long-distance real-world trips

2.5 Conclusions

This study investigates synthesis of representative drive cycles by Markov chain of 3-order versus 2-order. Prior to analysis, the naturalistic drive cycles are categorized into short, medium and long, to aggregate similar driving patterns. Then Markov chains of different orders are derived and output is analyzed to determine the potential benefits

from using the higher order. The comparison between 2-order and 3-order Markov chain on cycle synthesis performance is based on the newly defined index, Metric Errors. It sums up the differences on significant cycle metrics between any candidate cycle and the mean values of the naturalistic drive cycles in a corresponding cycle category. The mean Metric Errors and the minimum Metric Errors from numerous synthesized candidate cycles both decrease with a 3-order Markov chain compared to 2-order Markov chain. The significance of the decrease has been validated with statistical hypothesis tests. Hence, the Markov chain of 3-order is preferred when generating representative synthetic cycles from large naturalistic driving database.

3 The Categorization of Real-World Drive Cycles

In chapter 2, the significance of categorizing similar drive cycles has been emphasized for the accuracy of representative cycle synthesis. However, it cannot be emphasized enough that finer categorization is required to validate the product design and control strategy under complex real-world driving scenarios. This chapter proposes an algorithm and a metric, to identify road conditions and quantify driver styles, respectively with only the speed traces available. Demonstration on different datasets indicates the methodologies can be generalized and widely used in different product engineering applications.

3.1 Categorization of Drive Cycles by Road Conditions

Boundary conditions are a critical element of electrified vehicle design process, and they are to a large extent determined by the drive cycles employed in vehicle simulations or testing. Hence, a basic premise of this work is that understanding of customers and their naturalistic driving patterns are essential for the success of the propulsion system. That's the only way to assess the performance and fuel economy that the consumer will experience after acquiring the new vehicle, as well as expected battery degradation over a complete life cycle.

In this study, we rely on the publically accessible portion of a GPS-enhanced regional household travel survey from June 2001 to March 2002 conducted by the Southern California Regional Travel Survey in TSDC [8]. As drive cycle is defined as a series of speed points versus time from "key-on" (zero speed) to "key-off" (zero speed), 2347 drive cycles are collected for 292 "Autos", 513 drive cycles for 66 "Sports Utility

Vehicles", 381 drive cycles for 43 "Vans", 349 drive cycles for 58 "Pickup Trucks", 63 drive cycles for 5 "Other Trucks". Considering the purpose of this study, detailed second-by-second naturalistic drive cycles from the "Pickup Truck" segment are analyzed.

3.1.1 Methods

The main objective of drive cycle categorization is to divide cycles into different groups, based on their dominant features. Since our acceptance criteria for a representative drive cycle will be the error of mean-value approximation, a large variance of cycle characteristics would have adverse effect on determination of representativeness. As an example, it is not reasonable to attempt to synthesize a representative drive cycle from a mixed group of cycles that include low-speed stop-and-go city driving and high-speed free-flow highway driving under the risk of losing features of both driving patterns.

Previously used cycle categorization methods relied on cycle distance and cycle mean velocity and their probability distribution in the database [29][36]. Though the authors found a correlation between the driving distance and some of the statistically significant variables, e.g. mean velocity or mean positive acceleration, one obvious drawback is the sacrifice of other cycle details, e.g. frequency of "hills". This is illustrated by two specific drive cycles from the database in Figure 3-1. Both cycles cover the same distance with the same mean velocity, but Example cycle #1 can be interpreted as a mixed-cycle of highway cruise and congested urban driving, while Example cycle #2 represents an uncongested urban driving interrupted only by traffic lights. With the intuition from Figure 3-1, it is recommended to cut every cycle into several sub-cycles which are self-contained, and correlate them with road conditions to better identify categories. One

approach to determine snapshots of sub-cycles is to look at the driving activities between successive stops, so-called micro-trips, as illustrated in Figure 3-2. In this paper, the defined micro-trip does not include any idling period because the characteristics of running periods can individually describe local road conditions adequately. In Figure 3-2, one can confidently attribute urban stop-and-go traffic to the #1 micro trip, highway traffic to the #2 micro trip, and urban/suburban mixed traffic to the #3 micro trip.

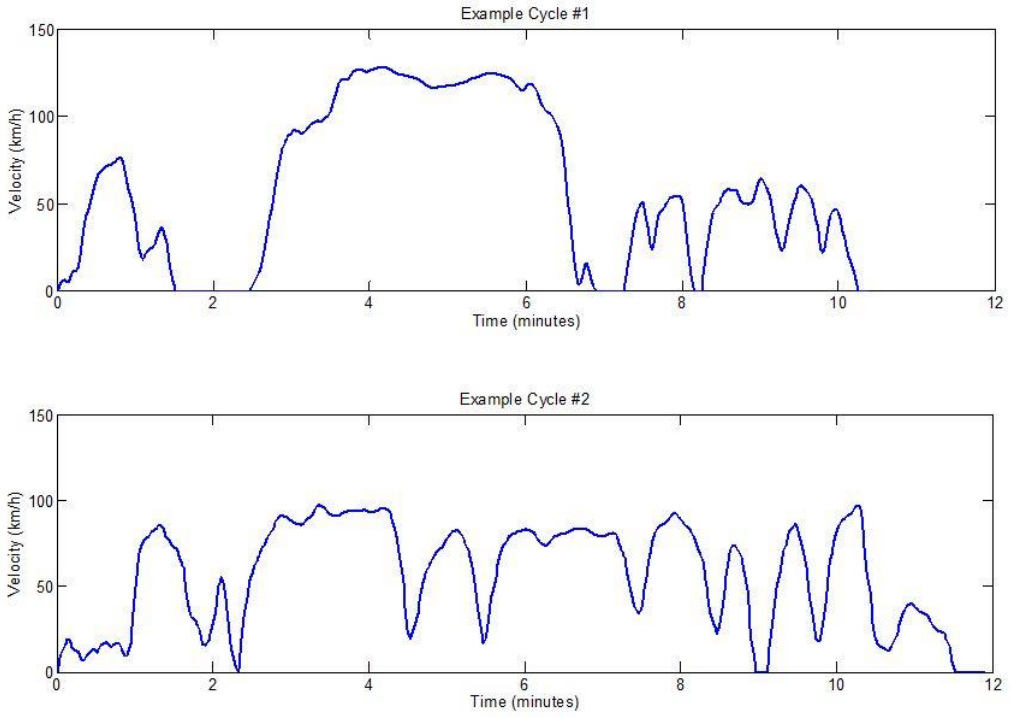


Figure 3-1. Two example drive cycles from the 2001-2002 South California Household Travel Survey

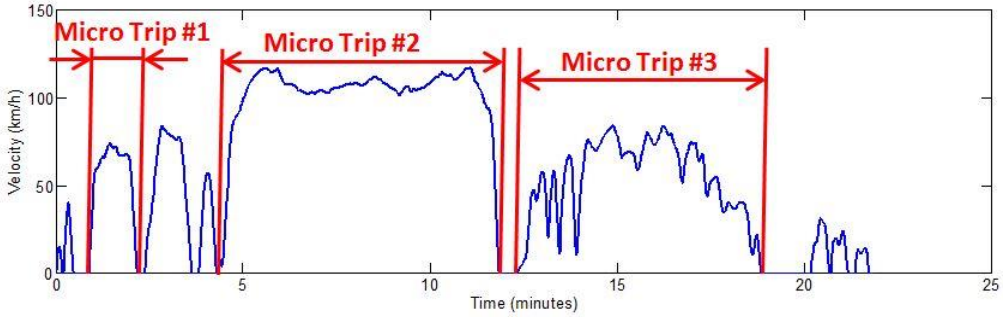


Figure 3-2. Illustration of the micro trip concept

In order to make the technique objective, statistical parameters for characterizing the micro trips ought to be identified. This study uses micro trip duration and micro trip mean velocity for characterization. The two parameters carry explicit driving information. For example, a short-duration micro trip with low mean velocity is associated with city stop-and-go traffic and a medium-duration micro trip with medium mean velocity is indicative of urban-suburban mixed traffic. Figure 3-3(a) shows the micro trip mean velocity vs. micro trip duration for all micro trips extracted from the available pickup truck drive cycles. In Figure 3-3(a), a steep upward trend can be observed for the short micro trips, i.e. part of the range below 10 min duration, followed by leveling-off beyond the 10 minute duration. This is where data clustering concept can be introduced to partition the micro trips. While there are many data clustering techniques, this study uses k-means clustering which utilizes a basic but fast algorithm.

The k-means clustering technique involves an iterative process to partition observations into the requested number of clusters, which equals to k. Necessary steps include the following:

Step 1: Choose k initial cluster centers.

Step 2: For all observations, calculate their distances to each chosen center.

Step 3: Assign each observation to the closest center.

Step 4: Calculate the average distances between all observations and the respective center in each cluster and respectively update the cluster centers.

Step 5: Repeat the steps 2 to 4 until the cluster assignments cease to change.

Reader is referred to Lloyd's algorithm [37] for more details. This study uses the kmeans function in the MATLAB's statistics toolbox.

3.1.2 Results

Since two clusters have been visually identified, all micro trips are firstly clustered based on both normalized micro trip duration and normalized micro trip mean velocity, as seen in Figure 3-3(b). The result clearly separates data that demonstrates a steep increase of mean velocity with trip duration, from the longer trips where this relationship levels off. However, this does not provide sufficient segregation, as micro trips in each cluster still display significant differences. The #1 Cluster alone includes micro trips that reach mean velocity of 5 km/h and 95 km/h but with the same duration. In other words, the two-cluster partitioning is still very rough. Consequently, micro trips in each of the two clusters are further sub-divided into two additional clusters, based solely on micro trip mean velocity, as shown in Figure 3-3(c). All trips with velocity below 30 km/h fall into the #1 Cluster, representing low-speed (congested) urban driving well. While micro trips in cluster #2 represent uncongested urban road conditions with somewhat higher velocities. Clusters #3 and #4 describe urban/suburban mixed, and highway road

conditions, respectively. The duration of these micro-trips is much longer, and mean velocities in cluster #4 are above 80 km/h.

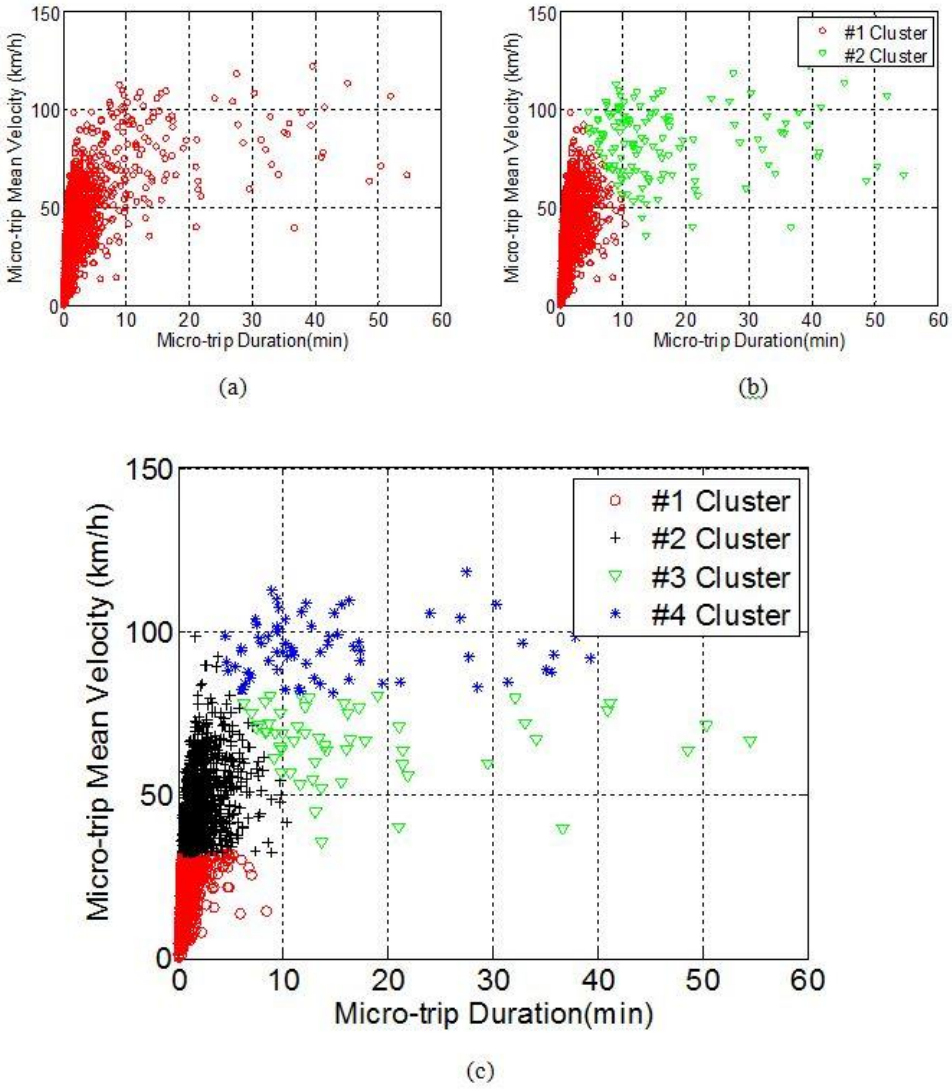


Figure 3-3. (a) Distribution of all micro trips; (b) partitioned by both micro trip duration and micro trip mean velocity into two clusters; (c) further partitioned into four clusters only by micro trip mean velocity.

After clustering all micro trips, the next question is how to categorize every drive cycle based on the variety of its micro trips. In this study, a look-up table is created for manual cycle categorization, see Table 3-1. The specific rules include:

1) If drive cycle contains micro trips from #1 and probably #2 clusters, and the number of micro trips from #1 cluster over total number of micro trips is over ½, the drive cycle is categorized as from Congested Urban road conditions.

2) If drive cycle contains micro trips from #2 and probably #1 clusters, and the number of micro trips from #2 cluster over total number of micro trips is over ½, the drive cycle is categorized as from Uncongested Urban road conditions.

3) If drive cycle contains micro trips from #3 and probably #1 or #2 clusters, and the number of micro trips from #4 cluster is zero, the drive cycle is categorized as from Urban-Suburban road conditions.

4) If drive cycle contains micro trips from #4 and probably #1 or #2 or #3 clusters, and the number of micro trips from #4 cluster over total number of micro trips is below ½, the drive cycle is categorized as from Urban/Suburban-Highway road conditions.

5) If drive cycle contains micro trips from #4 and probably #1 or #2 or #3 clusters, and the number of micro trips from #4 cluster over total number of micro trips is over ½, the drive cycle is categorized as from Highway road conditions.

Figure 3-4 illustrates typical drive cycles for every cycle category.

Table 3-1. Cycle categorization based on micro trip cluster components

Cycle Category	Micro trip Cluster Components	Cycle Characteristics
Congested Urban	#1, (#2)[a]	[#1][b]/[#Total]≥0.5
Uncongested Urban	(#1), #2	[#2]/[#Total][c]>0.5
Urban-Suburban	(#1,#2),#3	[#4]=0
Urban/Suburban –Highway	(#1,#2,#3),#4	[#4]/[#Total]<0.25
Highway	(#1,#2,#3),#4	[#4]/[#Total]≥0.25

a. The micro trip components in () do not necessarily exist
b. The [] represents the number of corresponding cluster of micro trips in a drive cycle
c. The [#Total] represents the total number of micro trips in a drive cycle

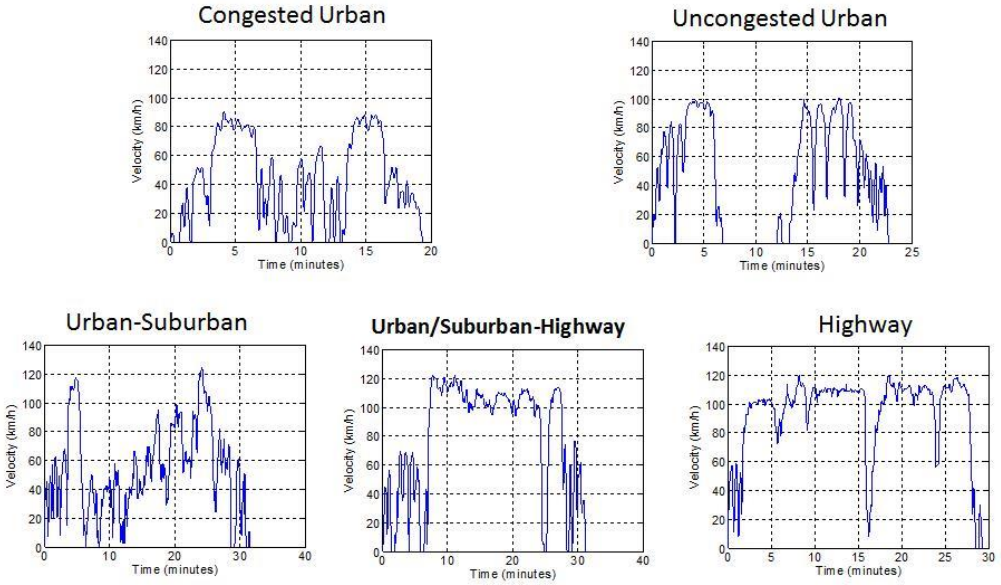


Figure 3-4. Typical drive cycles in each cycle category

Above categorization algorithm allows grouping of cycles with similar features. However, each category contains a huge amount of cycles, and therefore is not practical for powertrain design and/or control studies. Rather, a representative drive cycle should be synthesized for every road condition with the methodology in chapter 2. Following

that logic, five representative drive cycles are selected and their statistics are also inserted in Table 3-2. Their speed traces vs. time are shown in Fig. 6.

Table 3-2. Statistics for naturalistic drive cycles and the representative drive cycle in all cycle categories

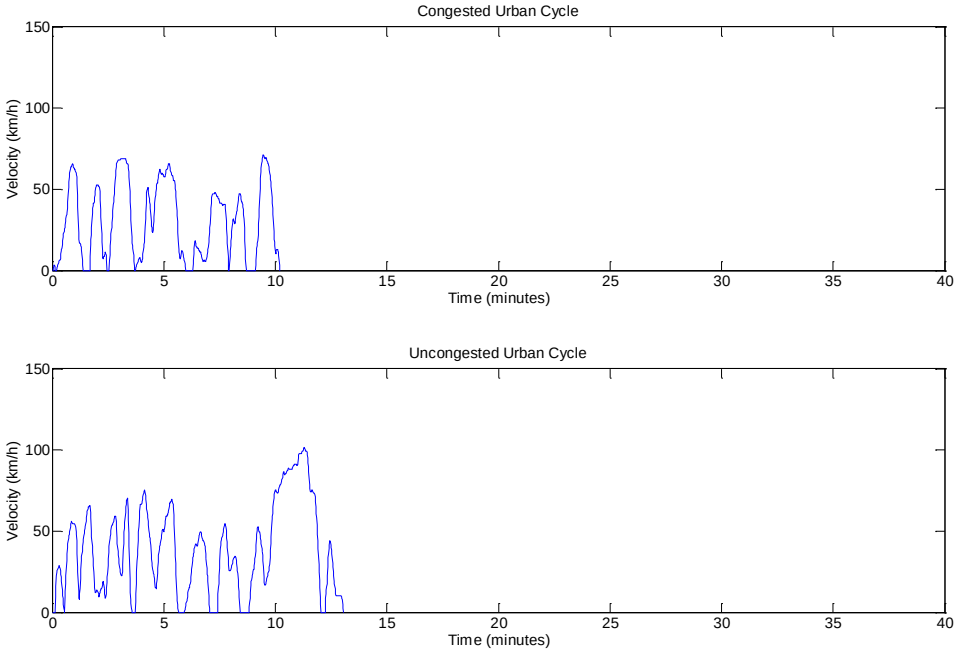
Significant Cycle Metrics	Mean Values Congested Urban Cycles	Representative Congested Urban Cycle
Standard deviation of velocity (km/h)	22.83	23.59
Mean positive velocity (km/h)	34.25	35.18
Standard deviation of acceleration (m/s2)	0.5652	0.6654
Minimum acceleration (m/s2)	-2.4622	-2.4400
Percentage of driving time under negative acceleration (%)	33.50	33.50
Percentage of driving time under positive acceleration (%)	36.12	35.78
Number of stops/km (1/km)	1.61	1.55

Significant Cycle Metrics	Mean Values Uncongested Urban Cycles	Representative Uncongested Urban Cycle
Standard deviation of velocity (km/h)	26.73	27.93
Mean positive velocity (km/h)	44.71	43.63
Standard deviation of acceleration (m/s2)	0.6909	0.7220
Minimum acceleration (m/s2)	-2.9268	-2.8900
Percentage of driving time under negative acceleration (%)	36.94	32.18
Percentage of driving time under positive acceleration (%)	41.40	40.10
Number of stops/km (1/km)	0.90	0.84

Significant Cycle Metrics	Mean Values Urban-Suburban Cycles	Representative Urban-Suburban Cycle
Standard deviation of velocity (km/h)	31.82	31.94
Mean positive velocity (km/h)	56.77	56.17
Standard deviation of acceleration (m/s2)	0.5284	0.5354
Minimum acceleration (m/s2)	-2.6640	-2.6700
Percentage of driving time under negative acceleration (%)	42.08	30.32
Percentage of driving time under positive acceleration (%)	47.60	36.04
Number of stops/km (1/km)	0.33	0.33

Significant Cycle Metrics	Mean Values Urban/Suburban –Highway Cycles	Representative Urban/Suburban –Highway Cycle
Standard deviation of velocity (km/h)	40.41	39.67
Mean positive velocity (km/h)	76.56	74.79
Standard deviation of acceleration (m/s2)	0.5344	0.5463
Minimum acceleration (m/s2)	-3.0461	-3.3300
Percentage of driving time under negative acceleration (%)	41.32	29.19
Percentage of driving time under positive acceleration (%)	45.58	31.22
Number of stops/km (1/km)	0.25	0.27

Significant Cycle Metrics	Mean Values Highway Drive Cycles	Representative Highway Drive Cycle
Standard deviation of velocity (km/h)	36.52	35.64
Mean positive velocity (km/h)	80.40	75.44
Standard deviation of acceleration (m/s2)	0.4711	0.4755
Minimum acceleration (m/s2)	-2.5606	-2.6700
Percentage of driving time under negative acceleration (%)	46.57	30.71
Percentage of driving time under positive acceleration (%)	48.91	32.77
Number of stops/km (1/km)	0.13	0.14



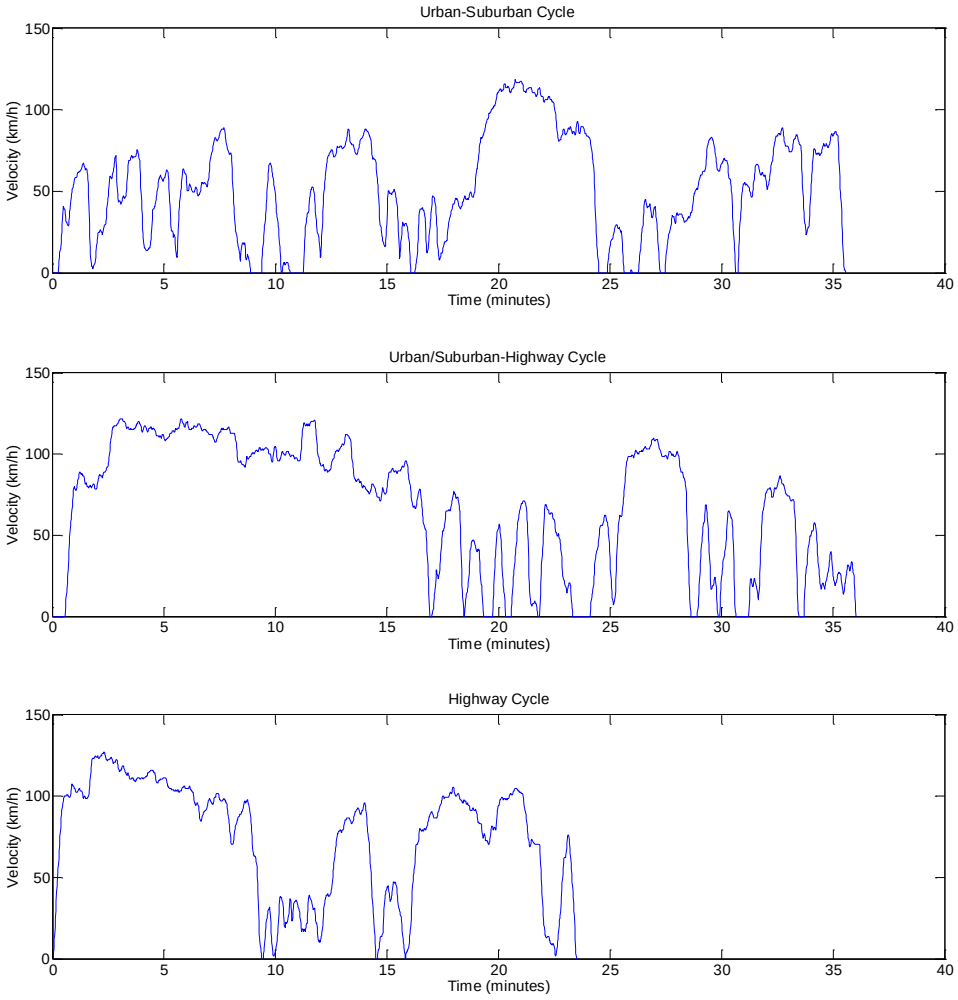


Figure 3-5. Illustration of Representative Drive Cycles for All Cycle Categories

3.1.3 Conclusions – Categorization Based on Road Conditions

An innovative approach based on micro trip clustering is proposed based on road conditions. The naturalistic drive cycles are categorized by analyzing their micro trip components. Micro trips are defined as the driving activity between successive stops. The micro trips are partitioned based on the duration and mean velocity using the k-means data clustering technique to infer the underlying local road conditions. Four typical micro trip clusters were identified. Subsequently, a look-up table was created to categorize drive

cycles based on the features of micro trips. This new method enhances the ability to recognize local underlying road conditions. Next, for each category of road conditions, a representative drive cycle was synthesized from numerous real-world ones. Not surprisingly, closer examination of the results reveals mean-value approximation errors of all representative cycles. The magnitudes are low and therefore very acceptable, but in new large set of data, it will be advisable to assess the appropriate number of drive cycle categories.

3.2 Categorization of Drive Cycles by Driver Style

This part defines a new metric of driver style, derived from the perspective of the cycle speed fluctuations. The speed fluctuations can be divided into two portions: the large-scale low frequency speed traces related to road conditions and the small-scale rapid speed fluctuations normally related to the driver. The latter represent to some extent the driver style and it is well known to affect the vehicle energy consumption and component duty cycles.

Driver style is an abstract term related to vehicle's energy consumption rate over a specific drive cycle. Previous studies had attempted to quantify the driver style with metrics based on velocity, acceleration and jerk [38][39][40][41]. Relationships between these metrics and energy consumption rates had been established to assist in various applications such as component design, eco-drive coaching and hybrid vehicle energy management [41][42][43]. However, the drawbacks of these metrics are that they are mostly average-based, percentile-based or variance-based. In other words, they capture only the general trend of drive cycles, missing some interesting details such as the rapid

fluctuations of the speed traces. These fluctuations cause a significant fuel economy penalty in gasoline/diesel engine powered vehicles, while their impacts on a hybrid vehicle depend on the propulsion system design and control. A metric, denoted as driver aggressiveness, is proposed in this study to define driver style by capturing the rapid speed fluctuations.

Real-world driving data collected with the GPS devices are widely adopted for the evaluation of real-world fuel economy, or the impact on electrified vehicle range and component duty cycles [8][44]. Nevertheless, due to the inherent operating behaviors of these devices, several sources of errors in the collected driving data had been discussed in [45], and summarized as: duplicate records, outlying records, signal drift, signal loss, and signal noises. An important step before downstream applications of the GPS collected data is to separate signal noises. This is visualized in Figure 3-6 with the GPS raw data and smoothed data. Common way of eliminating noise is to implement some filter. However, distinguishing between noises and real speed fluctuations is critical, and a new method is needed to guide proper filtering. While TSDC driving data applies in-house filters to address above errors, the availability of both raw data and processed data offers a chance to examine the potential filtering effects. Besides, the Michigan driving data, which were collected directly from the onboard Controller Area Network (CAN) bus and are assumed to be free of GPS-related high-power noise, are included for comparison as well.

This study aims to distinguish real world driving signals from signal noises and derive a new metric of driver style. The new metric will be validated with vehicle fuel consumption simulation.

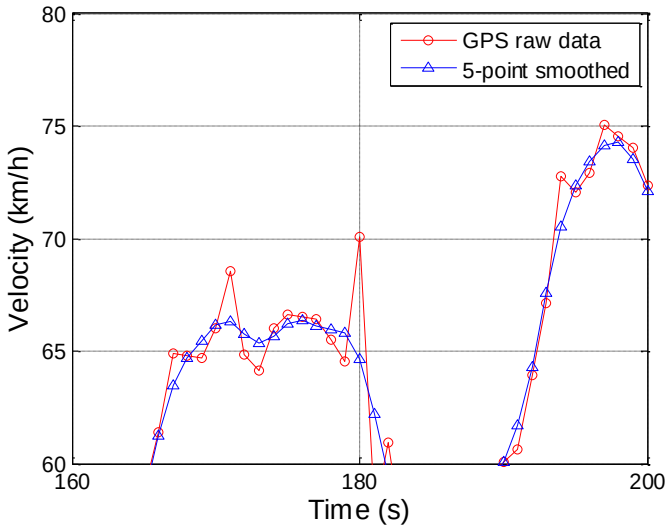


Figure 3-6. An example of GPS raw data and smoothed data

3.2.1 Methods

Every speed trace is a finite-length signal characterizing a driving event (from start to stop). Vehicle speed traces can be regarded as the discretized version of continuous-time driving velocity. In other words, the driving data from the surveys are sampled from real-world driving patterns. The sampling frequency for GPS devices is 1 Hz in this study.

Any discrete-time signal can be studied in both the time domain and frequency domain. The bridge between the two domains is the Fourier theory, a branch of harmonic analysis [46]. Several Fourier operators (direct and inverse) can be applied to discrete-time signals.

- Discrete Fourier Transform (DFT) maps length-N signals into a set of N discrete frequency components;
- Discrete Fourier Series (DFS) maps N-periodic infinite signals into a set of N discrete frequency components;
- Discrete-Time Fourier Transform (DTFT) maps infinite signals into the space of 2π-periodic function of a real-valued argument.

This study deals with finite-length signal; therefore, the DFT operator is chosen as basis for development of new metrics.

3.2.1.1 Discrete Fourier Transform

Given a length-N signal, x(n) in discrete time domain, with the Discrete Fourier Transform (DFT), the Fourier analysis formula of the signal is in Eq. 3.1:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk}, k = 0, \dots, N-1 \quad \text{Eq. 3.1}$$

In Eq. 3.1, $W_N = e^{-j\frac{2\pi}{N}}$

Therefore, a finite length-N signal in time domain can be transferred to frequency domain and decomposed into a set of N sinusoidal oscillatory components, characterized by the X(k) in frequency domain. X(k), which contain the magnitude and initial phase of each oscillator, corresponds to the digital frequency of $\frac{2\pi k}{N}$ $(k = 0, 1, \dots, N-1)$. If denote the abstract dimensionless digital frequency as $\omega_{digital}$, it relates to real frequency f_{real} in Hertz as in Eq. 3.2:

$$f_{real} = \frac{f_s}{2\pi} \omega_{digital} \quad \text{Eq. 3.2}$$

f_s is the sampling frequency of the signal. Due to the phenomenon of *Aliasing* [46], the maximum real frequency for the signal is $f_s/2$ for real-valued signal and f_s for

complex-valued signal. Since for all available speed traces in this study, the speed values are all real and the sampling frequency is 1Hz, the real-frequency range should be (0: 1/N: 1/2).

The transfer from frequency domain to time domain is also feasible with Inverse Discrete Fourier Transfer (IDFT), or say, the Fourier synthesis, as illustrated in Eq. 3.3 for above signal:

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k)\, W_N^{-nk}, n = 0, \dots, N-1 \qquad \text{Eq. 3.3}$$

DFT is a numerical tool which deals with finite signals and defines a finite number of operations, and the Fast Fourier Transform (FFT) is commonly used as the computation-efficient algorithm to accelerate the DFT. There are several techniques to implement the FFT, such as the Cooley-Tukey algorithm [47] and Rader's algorithm [48].

3.2.1.2 The Results of DFT: Periodogram

Based on DFT, one more useful relationship between the X(k) and x(k) of a signal is in the viewpoint of energy conservation, based on the Parseval's theorem [46] as described in Eq. 3.4.

$$\sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N}\sum_{k=0}^{N-1} |X(k)|^2 \qquad \text{Eq. 3.4}$$

As the left side of Eq. 3.4 represents the energy of the finite signal in time domain, it equals to the energy in the frequency domain normalized by 1/N, as shown on the right side of Eq. 3.4. All the terms on the right side of Equations, described as $\frac{1}{N}|X(k)|^2$, make up the so-called *Periodogram*. Specifically for the zero-mean signals, Eq. 3.4 would have

a straightforward physical interpretation: the area under the Periodogram approximates to the variance of the signal in time domain with long signal length and high recording resolution.

With all above techniques, every available speed trace can be observed in both the time domain and the frequency domain, yielding a new insight on the driving data.

3.2.2 Application of the Methods

To apply Fourier analysis on the driving data, some pre-processing is preferred to investigate the relevant details of this study. This section showcases these steps before the final results are described with an example based on the LA92 (Air Resources Board Dynamometer Driving Schedule) drive cycle.

3.2.2.1 Speed Trace

Figure 3-7 shows the speed trace of LA92 drive cycle. Even though the speed trace seems continuous, it is actually discrete-time signals with the sampling time of 1 second.

To better view a speed trace in frequency domain via DFT, several pre-processing steps have to be completed in time domain. The sequential steps include:

- Deleting the Idle-period;

Since the idle-periods do not affect the Ripple Aggressiveness discussed in this study, and would significantly affect the frequency components if included, they should be deleted.

- Mean-Adjusting;

Our investigating Ripple Aggressiveness will not be affected by subtracting the mean value from every speed point. By mean-adjusting, the Direct Current (DC) component

[46] at the 0 Hz in the frequency domain which can be too big and disguise other frequency components will be deleted.

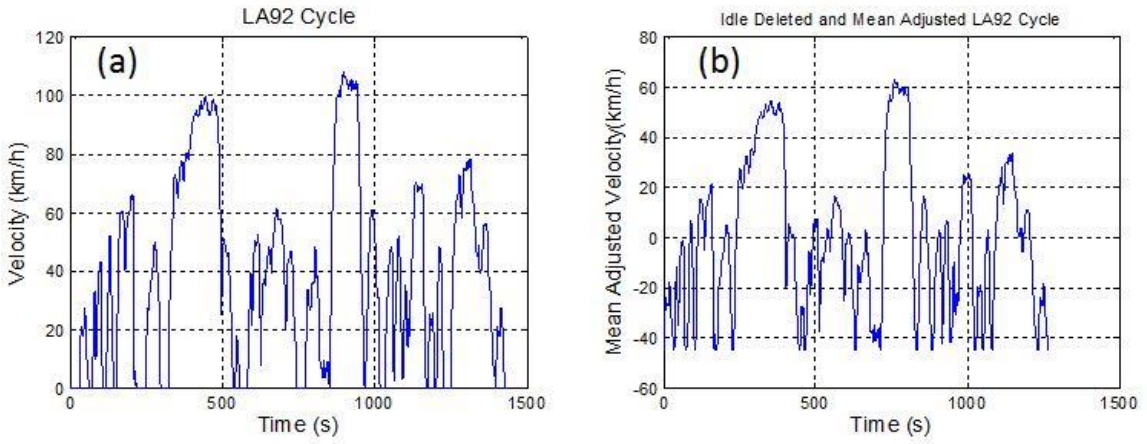


Figure 3-7. (a) Original LA92 drive cycle. (b) Idle-deleted and Mean-adjusted LA92 drive cycle

Converting the processed speed trace in the time domain to the periodogram in the frequency domain, in Figure 3-8 we can find that for LA92's speed trace, nearly all of its significant frequency components are within the low frequency range(<0.1 Hz), implying an inherent harmonic trend. At high frequency range (>0.1 Hz), the frequency components though rather small still exist.

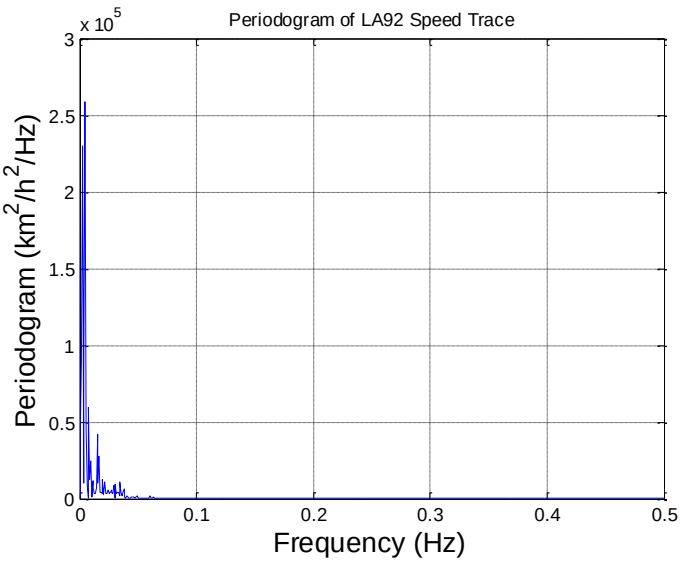


Figure 3-8. Periodogram of LA92 speed trace

One way to quantify the aggregation of low-frequency components for speed traces is based on the Parseval's theorem in above section. Specifically for zero-mean signals, the area under the curve of Periodogram in the frequency domain approximates the variance of the signal in the time domain. Therefore for all the mean-adjusted speed traces, we can move from the origin of the Frequency axis to cover the same proportion of the total variance in time domain, say 99% in this study. The frequency which satisfies this condition will be denoted as the Cutoff Frequency in the following. We apply this notion to all trips with driving distance over 1 km from two survey's processed data, and the results are shown in the Figure 3-9.

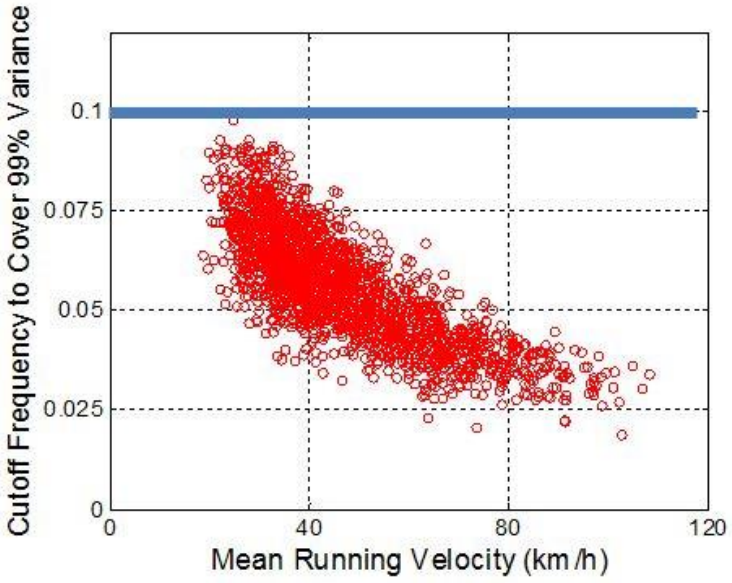


Figure 3-9. Distribution of the cutoff frequency to cover 99% of total variance of the speed traces (idle period excluded)

Figure 3-9 shows a tight relationship between mean running velocity (idle periods excluded) and the cutoff frequency. With increasing mean running velocity, the cutoff frequency inversely decreases from the peak value of 0.1 Hz. In other words, if we focus on the frequency components below 0.1 Hz, all trips can be well transformed with more than 99% of variance covered. This means that 0.1Hz can be enough to cover the basic features of driving.

From above results, we can see that in order to investigate the high-frequency speed fluctuations, it is necessary to firstly eliminate the low-frequency harmonic trend contained in the speed traces. A traditional method that not only de-trends the original trace but also contains explicit physical meaning is to difference the speed traces in viewpoint of time series [49].

3.2.2.2 Acceleration Trace

Differencing of the speed trace yields the acceleration trace. An example of the acceleration trace from LA92 is shown in the Figure 3-10(a). The intention is to eliminate the low-frequency harmonic trend of the speed trace, and the outcome can be evaluated in the Periodogram plot for the LA92's acceleration trace shown in the Figure 3-10(b).

In Figure 3-10(b), the frequency components of the acceleration trace aggregate in the low frequency range (<0.2 Hz) with significant peaks at several frequencies, just as the counterparts from the speed trace. Therefore, an inherent harmonic trend is still contained in the acceleration trace. In order to further de-trend the acceleration trace, a second derivative of speed, or jerk, is examined next.

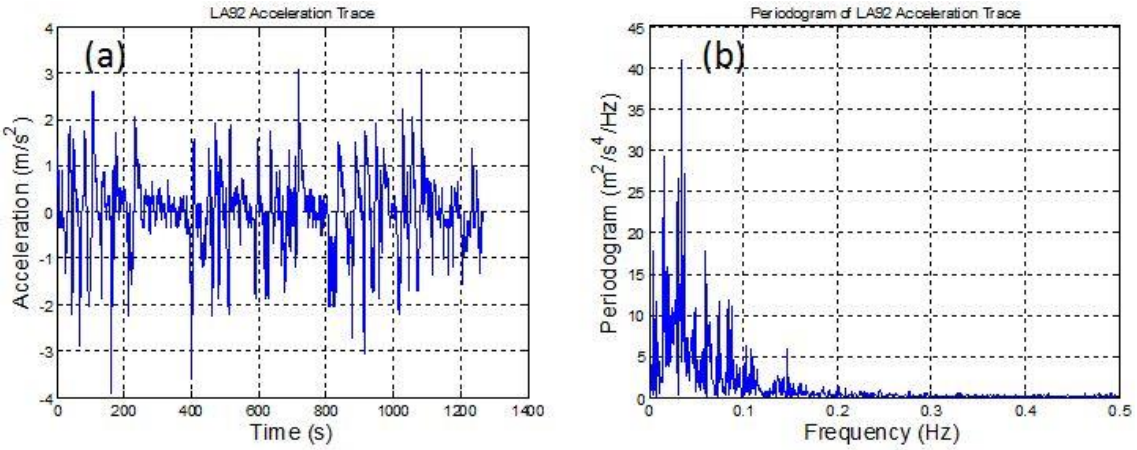


Figure 3-10. (a) LA92's acceleration and (b) its periodogram

3.2.2.3 Jerk Trace

Jerk is defined as the rate of change in acceleration or deceleration. An example calculated for the LA92 is illustrated in Figure 3-11(a). Further converting the jerk trace into Periodogram in the frequency domain is shown in Figure 3-11(b), no obvious peak can be observed and the energy of the trace is almost evenly distributed across the

frequency domain, which is in sharp contrast with counterparts from speed and acceleration traces. In other words, after two-time differencing, the rapid speed fluctuations can be regarded as having been separated from the "traffic trend". Further differentiation is not necessary, and it would not have explicit physical meaning. Therefore, jerk trace is utilized for further analysis.

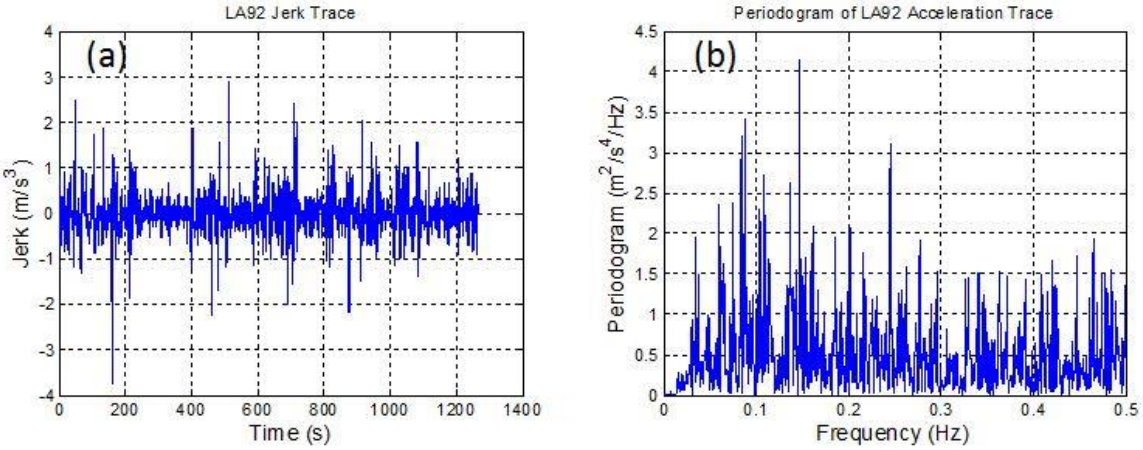


Figure 3-11. (a) LA92's jerk trace and (b) its periodogram

For better illustration, periodograms of several jerk traces from real-world transportation surveys are shown in Figure 3-12. The elimination of high frequency components in Figure 3-12(a) indicates intensive filtering of 2010 California survey's processed data. In stark contrast, the effect of high-power noise is evident in 2001 California survey's raw data, because signal energy aggregates mainly in the high-frequency range (see Figure 3-12(b). Figure 3-12(c) and Figure 3-12(d) show periodograms of the jerk traces from the 2001 California survey's processed data and the 2001 Michigan project's CAN data. Both display a flat distribution similar to the LA92 data.

The greatest difference among the 3 real-world transportation surveys is the proportion of high-frequency energy over the whole frequency domain. While the jerk trace of the 2010 California survey's processed data has the high-frequency components almost eliminated, the jerk trace from the 2001 California survey's raw data has nearly all energy aggregated in the high frequency range.

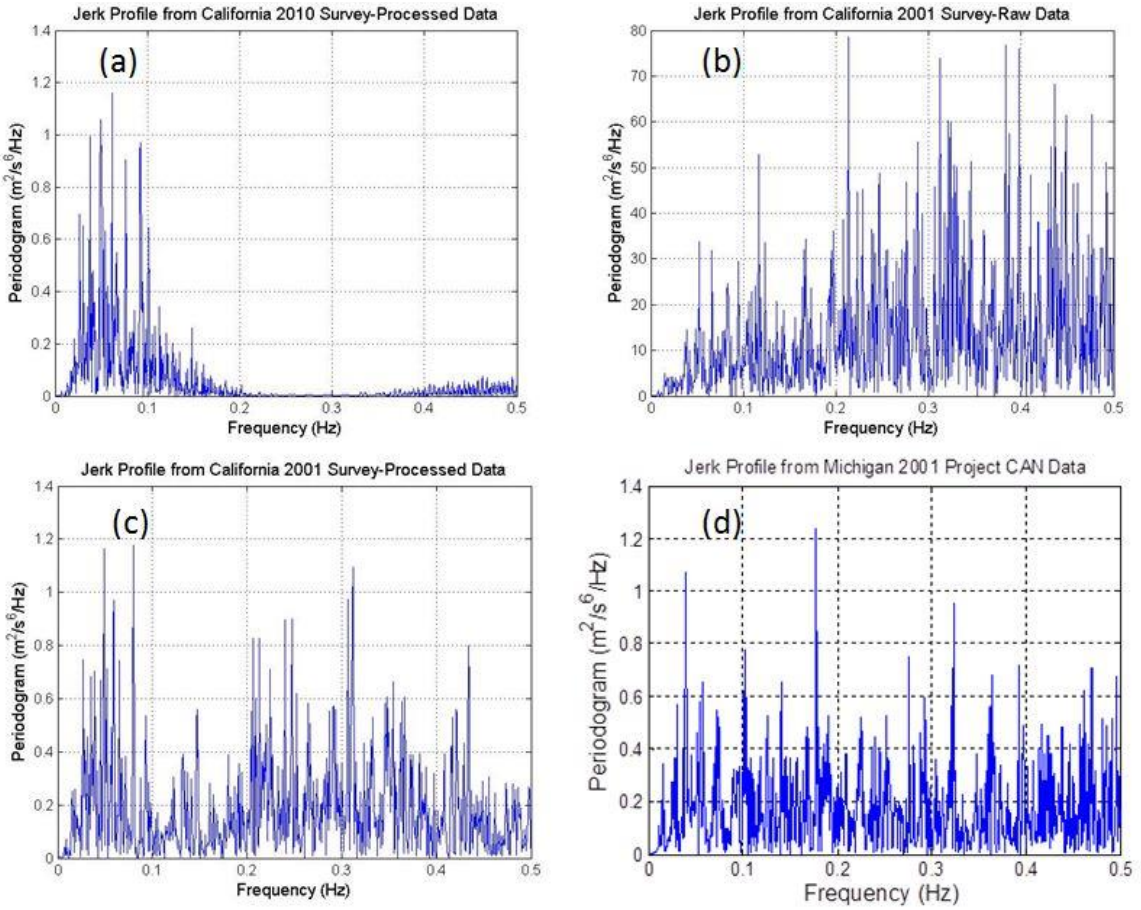


Figure 3-12. Example periodograms of jerk traces from (a) 2010 California survey's processed data; (b) 2001 California travel surveys raw data; (c) 2001 California survey's processed data; (d) 2001 Michigan Project's CAN data

3.2.2.4 The Definition of Ripple Aggressiveness

To formally quantify the variations of the high frequency range observed in Figure 7, the first important step is separation of the low-frequency range and high-frequency range similarly to [50]. This can be referred back to the Cutoff Frequency in Figure 4. As previously mentioned, the frequency components of speed trace below 0.1 Hz can cover at least 99% of energy on time domain for all the speed traces. In other words, for a speed trace, using only the frequency components below 0.1Hz can be enough to cover the basic features of driving determined by traffic

This observation is used to separate frequency regions in the jerk trace. The low-frequency part (<0.1 Hz) in Figure 8 is denoted as the “Traffic” part and attributed to the basic driving situations which cannot be avoided. The high-frequency part (>0.1 Hz) is denoted as the “Ripples”, attributed to the individual driver characteristics and investigated as Ripple Aggressiveness.

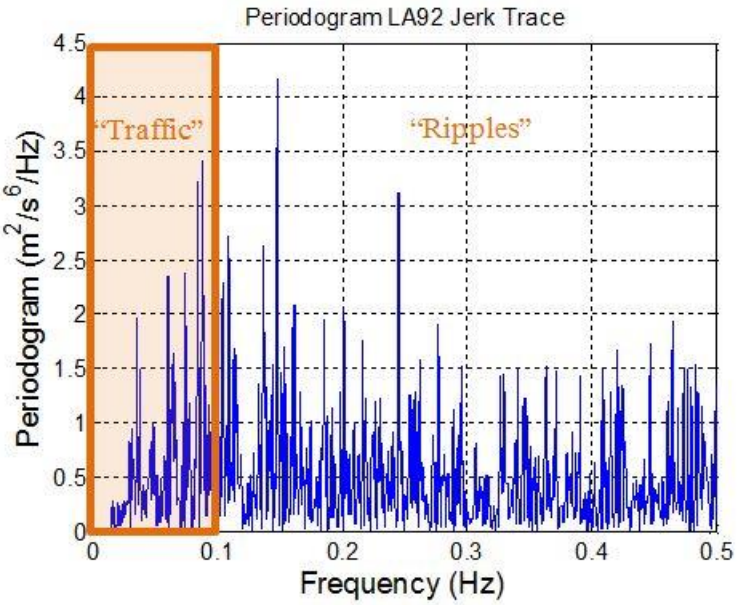


Figure 3-13. Periodogram of LA92's jerk trace

Based on above discussion, the aggressiveness metric can be defined as in Eq. 3.5.

$$\text{Ripple Aggressiveness} = \frac{HF}{LF + HF} \qquad \text{Eq. 3.5}$$

According to Parseval's theorem, LF is the energy aggregated in the low frequency range by summing the low frequency components below 0.1 Hz, while HF is the energy aggregated in the high frequency range by summing the high frequency components above 0.1Hz.

Using this definition, the variation in the proportion of energy between high-frequency range and low-frequency range can be quantified. This metric is automatically normalized between [0, 1], to provide an explicit evaluation for speed traces from any database.

In this work, we apply the Ripple Aggressiveness (RA) defined in Eq. 3.5 to all available speed traces. To better illustrate its distribution, all calculated Ripple

Aggressiveness values are counted and sorted into a 10-bin histogram shown in Figure 3-14. Clearly, the distributions of Ripple Aggressiveness calculated using data from various transportation surveys vary significantly. While the raw data mostly occupy the very upper level of Ripple Aggressiveness (~0.9) due to high-power noise in the signal, the processed data of 2010 California survey show normal distribution around 0.4. We assume the data from onboard CAN bus are subject to much less noise and represent the real-world Ripple Aggressiveness - see the RA distribution of 2001 Michigan project data. Ripple Aggressiveness calculated from 2001 California survey processed data show distributions similar to Michigan CAN; therefore, it can be inferred that 2001 data are more representative for our study, than data from the 2010 survey.

To further test for potential over-filtering effects, a 5-point SG-Binomial filter proposed in [45] is applied to all available speed traces except the processed data of 2010 California survey. As can be seen in Figure 3-14(b), after further filtering, all traces move closer to the traces of the 2010 California survey in Ripple Aggressiveness. Nevertheless, more aggressive filter is needed to create complete overlap with processed data of the 2010 California survey. This leads to a conclusion that the speed traces of California 2010 survey have indeed been over-filtered to reasonably capture the Ripple Aggressiveness.

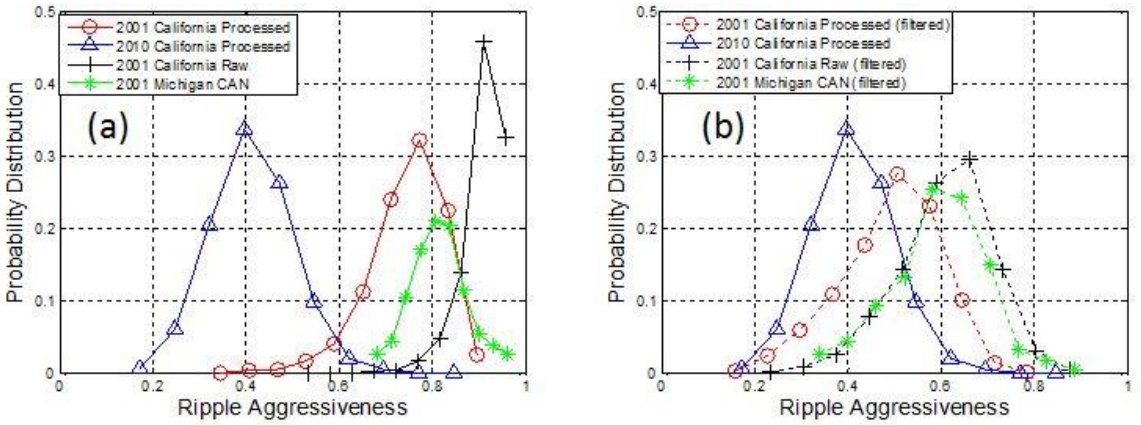


Figure 3-14. (a) Distribution of Ripple Aggressiveness for processed data of California 2010 travel survey, the processed data and raw data of 2001 California survey and CAN data of 2001 Michigan project; (b) Distribution of "Ripple Aggressiveness" for processed data of California 2010 travel survey, the further-filtered processed data and raw data of 2001 California survey and further-filtered CAN data of 2001 Michigan project;

3.2.3 The New Metric of Driver Aggressiveness

This section firstly investigates the relation between Ripple Aggressiveness and fuel consumption. To achieve this goal, a conventional compact vehicle model, adapted from ADVISOR [51], is set up in MATLAB/Simulink as in Figure 3-15. It is a backward/forward-looking model, i.e. the power demands are directly calculated from the drive cycle and fed back upstream to the engine. The specifications of its key components are compiled in Table 3-3. Correctly configured, a reference drive cycle is followed well as shown in Figure 3-16. Then the 1851 trips of 292 passenger vehicles from 2001~2002 Southern California Household Travel Survey [8] with a total driven distance of about 25,000 km are adopted for simulation. Their cumulative distribution of ripple aggressiveness are equivalently split for three drive style categories, namely calm, normal, and aggressiveness in Figure 3-17(a). Their mean fuel consumptions are plot in Figure 3-17(b), yet revealing only a limited positive correlation with ripple aggressiveness.

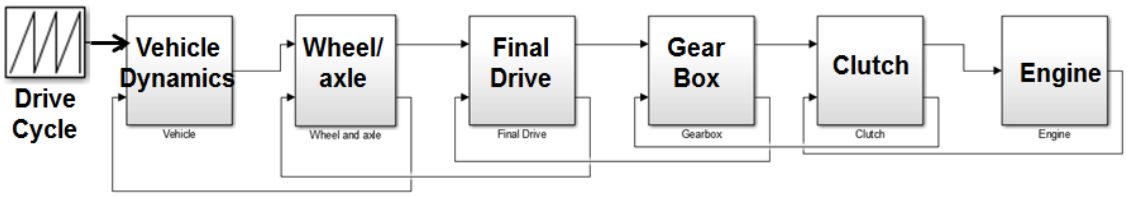


Figure 3-15. Simulink model of a conventional compact vehicle

Table 3-3. Specifications for vehicle/engine/transmission

Specifications	Units	Values
Vehicle Dynamics		
Curb Weights	kg	1300
Passenger and Cargo Weight	kg	136
Drag Area	m^2	0.66
Rolling Resistance Coefficient		0.012
Tire Radius	m	0.287
Engine		
Engine Type		Inline 4 cylinders & Naturally aspirated & Spark Ignition
Engine Displacement	L	1.9
Transmission		
Transmission Type		6-speed manual
Transmission Ratios		[4.148, 2.370, 1.756, 1.155, 0.859, 0.683]
Final Drive Ratio		3.21

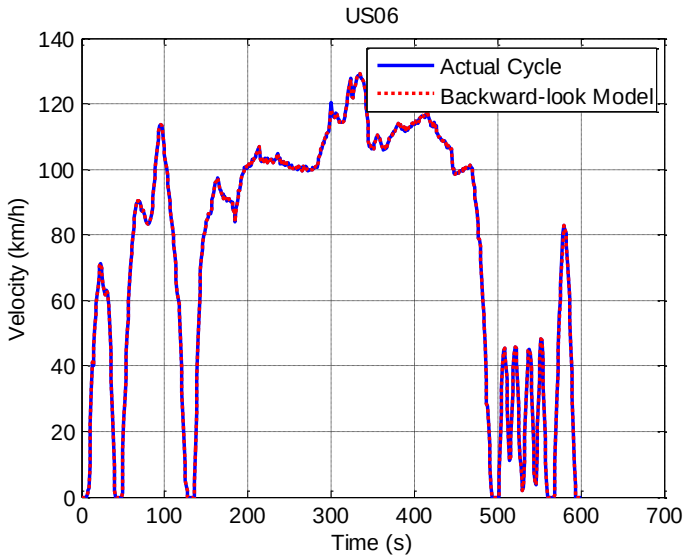


Figure 3-16. Cycle following behaviors by backward/forward-looking vehicle model

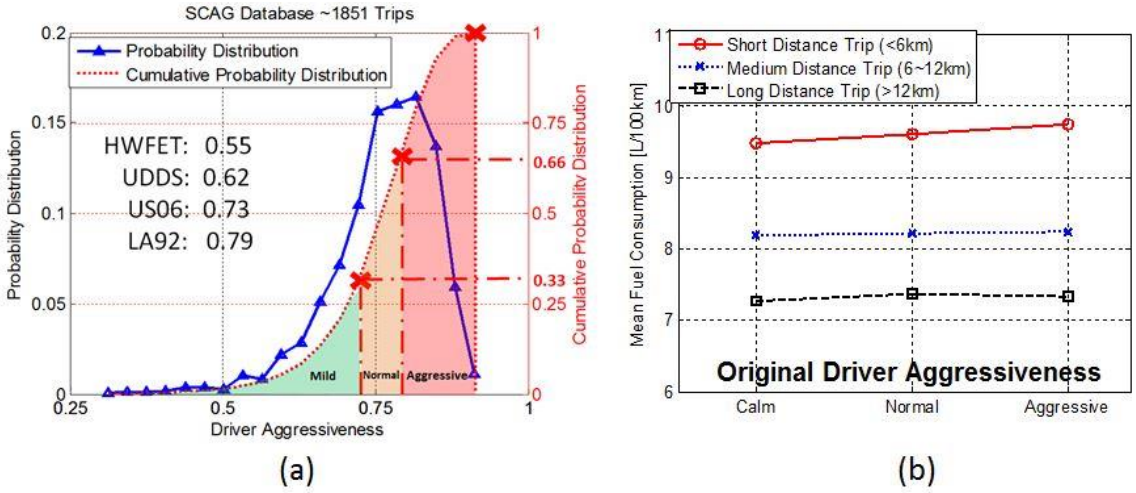


Figure 3-17. (a) Distribution of ripple aggressiveness in SCAG database; (b) Mean fuel consumption predicted for different ripple aggressiveness categories and according to trip distance

Figure 3-18 illustrates an inconsistency that needs to be addressed. Both trips in Figure 3-18 display nearly the same ripple aggressiveness, as defined by Eq. 3.5, even though the acceleration rates of the trip (excluding idling periods) in Figure 3-18(a) are much higher compared to the trip in Figure 3-18(b). A closer examination indicates that

different magnitudes of periodograms might be the cause for the inconsistency, and ought to be included in the metric. Hence, a multiplier is added to the definition of ripple aggressiveness and the new definition is shown in Eq. 3.6. The modified driver aggressiveness is better correlated with the cycle acceleration rates as in Figure 3-18.

$$\text{Driver Aggressiveness} = \text{mean (periodogram)} * \frac{HF}{LF + HF} \qquad \text{Eq. 3.6}$$

The LF is the energy aggregated in the low frequency range by summing the low frequency components below 0.1 Hz; HF is the energy aggregated in the high frequency range by summing the high frequency components above 0.1Hz. The Mean (periodogram) is the mean value of the periodogram across the whole frequency range.

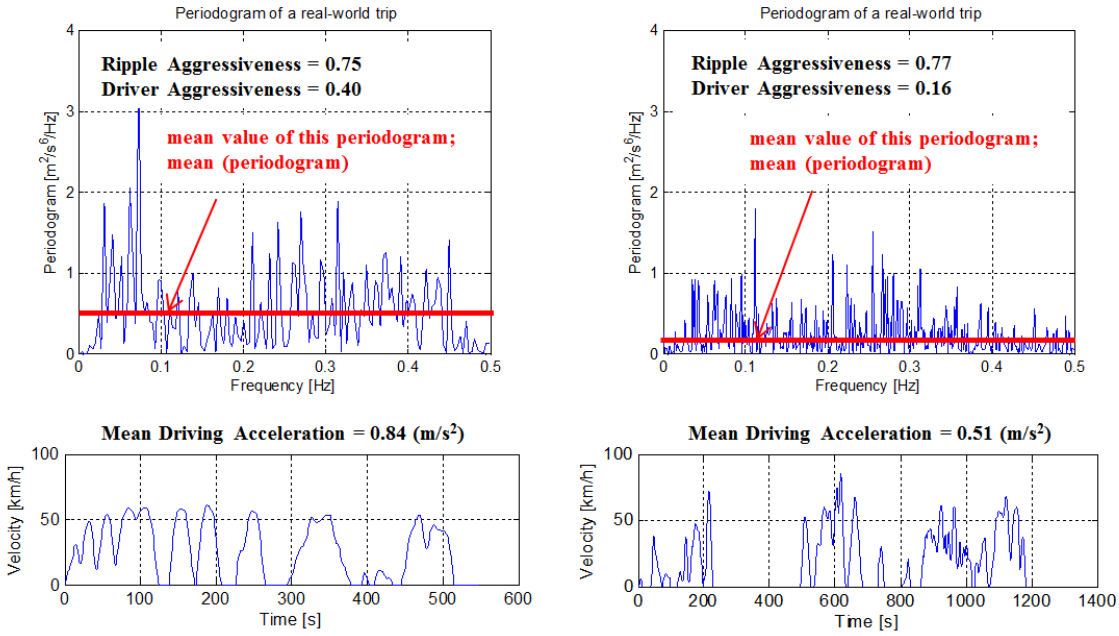


Figure 3-18. Example periodograms of two real-world trips and their corresponding speed traces: two driving activities on secondary road at moderate speeds. Figure 3(a) on the left, 3(b) on the right

Moreover, the driver aggressiveness after categorization into three driver styles are in strong positive correlations with fuel consumptions in Figure 3-19. The key points are summarized as follows:

- The fuel consumption variation between calm and aggressive driver can be as much as 2L/100km for the short distance trips
- The relative fuel consumption values increase with reduction of trip distance as the short trips can be mostly attributed to city driving and frequent stop-and-start situations, with generally higher fuel consumption.
- The relation between fuel consumption and driver aggressiveness is strongest among short-distance trips, seen as the highest positive slope.
- Strength of the correlation degrades with increasing trip distance because operating points move towards the high efficiency region, and the impacts due to driver action is not that pronounced.

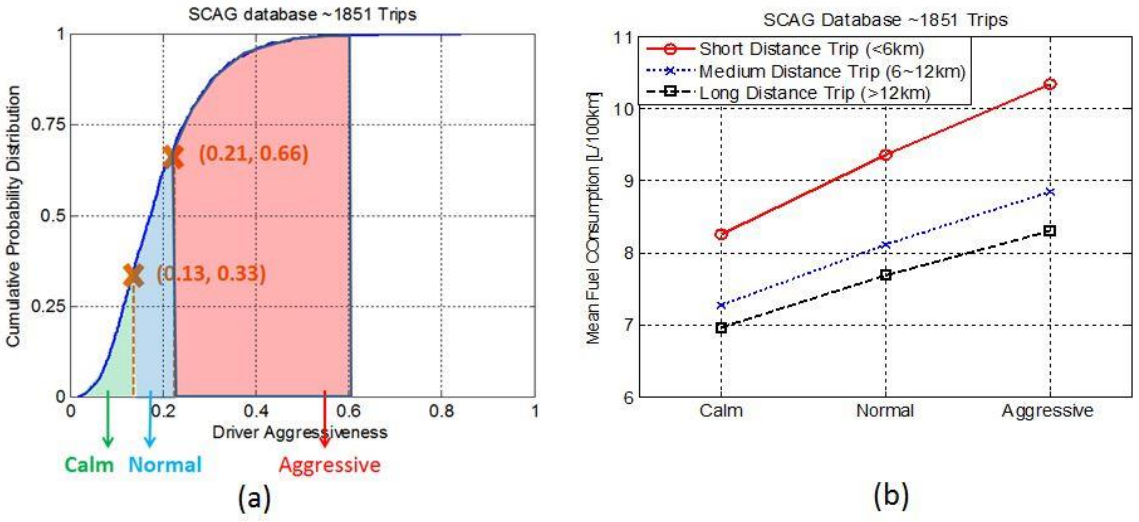


Figure 3-19. (a) Distribution of driver aggressiveness in SCAG database; (b) Mean fuel consumption predicted for different driver aggressiveness categories and according to trip distance

3.2.4 Conclusions – Driver Aggressiveness Metric

This part investigates a tradeoff between speed trace fluctuations and GPS data noises to deduce the proper filtering method for drive cycles in real-world transportation surveys. Analysis of certification cycles led to a conclusion that spectrum of jerk trace characterizes all aspects of driving. Vehicle jerk traces below 0.1 Hz represent driving dictated by traffic, while higher frequencies correspond to driver behaviors. Rapid speed fluctuations in the drive cycle are quantified with metrics called Ripple Aggressiveness (RA). It is calculated based on energy aggregated in the spectrum of vehicle jerk traces.

To assess the relationship between filtering and Ripple Aggressiveness, the analysis methodology is applied to several sets of naturalistic driving data. The probability distribution of Ripple Aggressiveness clearly indicates differences between data acquired with different methodologies. On the left side of the spectrum are over-filtered data from the 2010~2012 California Travel Survey (normal distribution of RA around 0.4), while the raw data from the 2001-2002 Southern California Regional Travel Survey lead to the opposite extreme. The mildly filtered data from the 2001-2002 Southern California Regional Travel Survey produce a Ripple Aggressiveness distribution centered around 0.78, very similarly to the data from Michigan recorded directly from vehicle CAN. Since CAN data are by definition realistic, we conclude that RA distributions can be used to assess the representativeness of data obtained with GPS, and that RA distribution should be centered between 0.6-0.8.

Using a conventional compact vehicle model in the MATLAB/Simulink, the fuel consumption is found to be in loose positive correlation with the newly defined Ripple

Aggressiveness. Modified by the mean value of the periodogram, an innovative definition of driver aggressiveness is found to strongly correlate with fuel consumption, which paves the pathway for following discussions.

4 48V Mild Hybrid Vehicle Modeling and Simulation

The objective of this chapter is to relate real-world drive cycles to battery cell duty cycles. Meanwhile, the impacts of real-world drive cycles on the performance and fuel economy of a generic 48V mild hybrid vehicle are analyzed as well.

The 48V mild hybrid technology is emerging as an attractive option for high-volume vehicle electrification. Compared to high-voltage hybrids, the 48V system has a potential of achieving competitive fuel economy with significantly lower incremental costs. Its typical functions, on top of the Start&Stop, include torque assist and regenerative braking. Furthermore, the electrical shock protection is not required for 48V thus lowering the overall system and maintenance cost [52].

Concerning 48V mild-hybrid technology, many previous studies have been developed. In [53], several mild hybrid system architectures are showcased, while [54][55] discuss engine options and electric component sizing. Benchmark for the fuel saving potential with optimal power management strategy is established in [56][57], and [58][59] propose the implementable rule-based and fuzzy-logic strategy, respectively. Simulation based study of the 48V system efficiency potentials in different vehicle classes is discussed in [60], thus providing useful insights for assessing technology adoption.

Despite previous efforts, one drawback is that they were all based on the certification cycles which lack the representativeness of the real-world driving conditions. Therefore, leveraging a naturalistic driving database is indispensable when analyzing a real-world fuel saving potential and HEV component duty cycles. In this study, 1851 real-world trips from passenger vehicles in the 2001~2002 Southern California Household Survey [8]

are imported and coupled with a vehicle model, which is adapted from ADVISOR (Advanced Vehicle Simulator) [51]. Subsequently, a consistent energy management strategy is implemented in the vehicle simulator and the real-world fuel consumption reductions are quantified for different levels of driver aggressiveness. The strategy is based on the approach proposed in [61], and it separates power demands into the low-frequency range, met by the engine, and high-frequency range fulfilled by the electric motor. Therefore, engine is assumed to carry the low-frequency, high-amplitude base loads for better fuel efficiency, while battery covers only the high-frequency, low-amplitude fluctuations of the power demands. This assures that the power split between the engine and battery is managed consistently regardless of the driver aggressiveness. This also enables fair comparison of the fuel efficiency potential.

4.1 The 48V Mild HEV Simulator

The statistical distributions of travel distances, cycle mean velocity, idle percentage and mean positive acceleration rates of the applied real-world drive cycles are shown in Figure 4-1. A large variety of driving conditions are contained in this database; hence, it well reflects the real-world driving conditions. The drive cycles with distance less than 1 mile (1.6 km) are intentionally excluded.

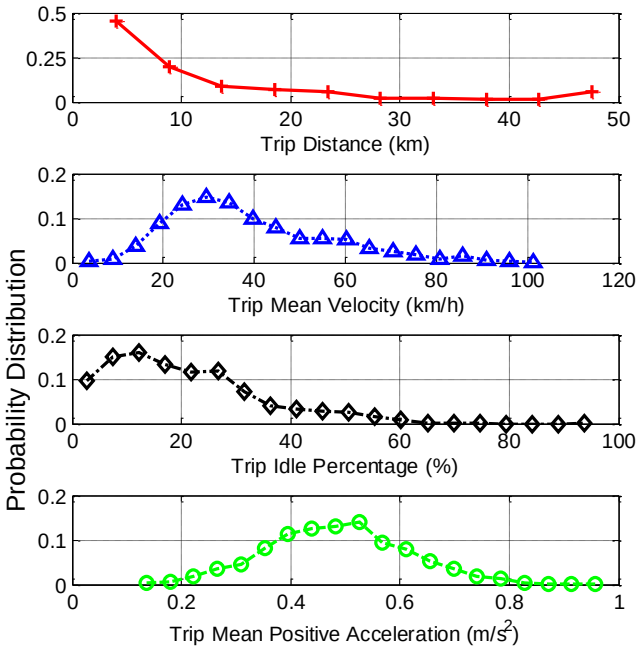


Figure 4-1. Distribution of trip distance, trip mean velocity, trip idle percentage, trip mean positive acceleration in the SCAG naturalistic drive cycle database

For the purpose of this study, the 1851 real-world trips are further categorized according to driver aggressiveness. It abstracts the frequency of changing the vehicle power demand can be referred to Eq. 3.6 and Figure 3-19.

4.1.1 Vehicle Architecture

The conventional compact size car in chapter 3 is selected as a platform for evaluating the impact of the 48V hybrid system. The baseline roughly corresponds to a compact car with a conventional powertrain, which are popular on market with cost-sensitive customers, with vehicle/engine/transmission specifications in Table 3-3.

Typical mild hybrid vehicle combines the separate alternator and starter into a single starter-generator. The electric motor can be connected mechanically to the engine via belt-drive (commonly known as Belt Starter Generator or BSG) or directly coupled to the

engine crank-shaft (commonly known as Integrated Started Generator or ISG). This study assumes the latter, with a pre-transmission location and continuous operation of ISG without an additional clutch.

A vehicle model, adapted from ADVISOR, is set up in MATLAB/Simulink, shown in Figure 4-2. It is a backward/forward-looking model, i.e. the power demands are directly calculated from the drive cycle and fed back upstream to the engine in case of the conventional vehicle, or the engine and ISG in case of a hybrid configuration. The reference drive cycle is again followed closely as shown in Figure 3-16. Hence, there is no noise on top of the measured speed trace that could distort driver aggressiveness calculations.

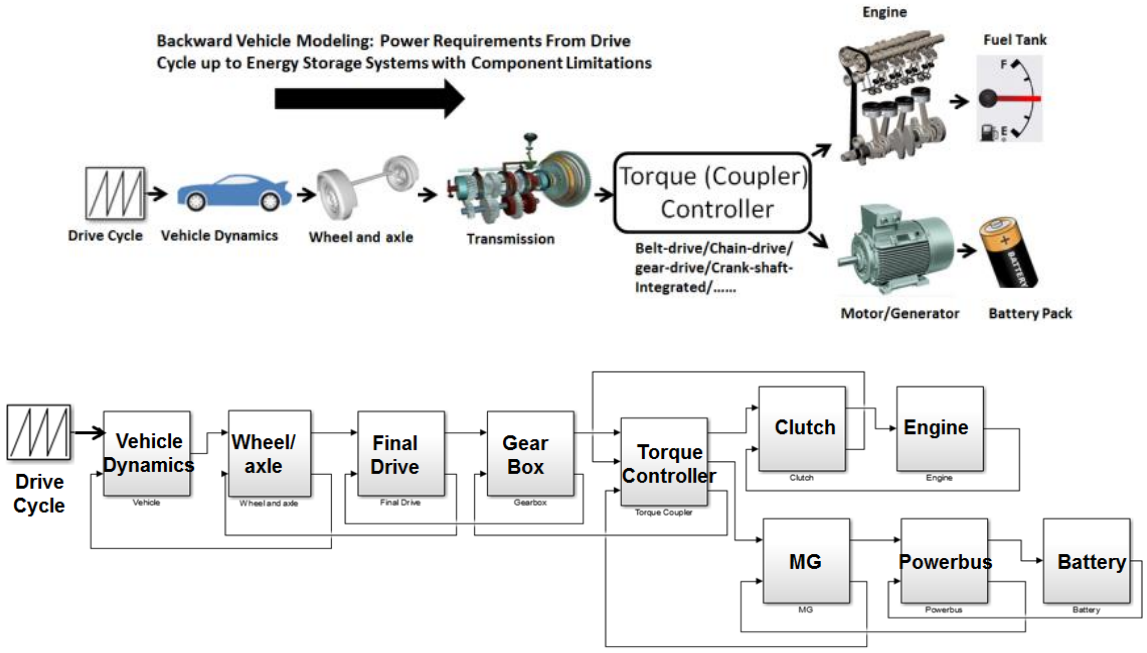


Figure 4-2. A generic of mild hybrid vehicle in the form of backward vehicle model and its Simulink Model

The electric motor is represented by its efficiency map. The specifications of electric motor and battery used in this study are summarized in Table 3. The battery is a zero-order equivalent circuit model, illustrated in Figure 7. The state of charge (SOC) is based on coulomb counting method, as described by Eq. (4).

Table 4-1. Motor and battery specifications

Motor Supply Voltage (V)	42~58
Motor Maximum Power (kW)	25
Motor Continuous Power (kW)	20
Motor Maximum Torque (Nm)	75 @ (0~3000 rpm)
Motor Maximum Speed (rpm)	10000
Battery Cell Type	Lithium-ion (LFP)
Battery Cell Rated Capacity (Ah)	4.2
Battery Cell Max. Continuous Discharge Current (A)	84
Battery Cell Pulse Discharge Current at 30 seconds (A)	110
Battery Cell Max Charge Current (A)	42
Battery Cell Max/Min SOC	0.75/0.35
Battery Pack Configuration	14 Series/4 Parallel

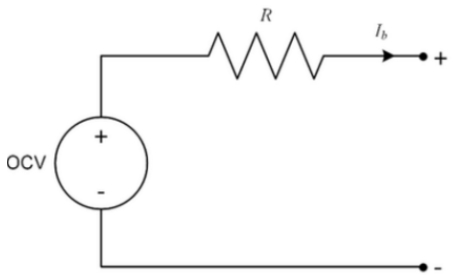


Figure 4-3. Battery zero-order equivalent circuit model

$$\text{SOC} = SOC_{init} - \frac{\int idt}{C_{batt}} \qquad \text{Eq. 4.1}$$

Where SOC_{init} is the initial battery SOC, $\int idt$ is the current integration over time, C_{batt} is the battery capacity.

4.1.2 Power Management Strategy

The power management strategy is derived in frequency domain, so that the impact of different levels of driver aggressiveness can be evaluated in a consistent manner. Similar to the concept proposed in [61], the power demand passes through a low-pass filter, and sends low-frequency and high-frequency power demands to the engine and battery, respectively. In contrast to commonly used approach focused exclusively on efficient operation by moving the operating points towards high-load, lower brake specific fuel consumption (BSFC) region, the strategy developed here aims to separate out the impacts of driver aggressiveness and burden the battery consistently. This fulfills the intent to correlate the effects of driver aggressiveness with fuel consumption. Since the power rating of the ISG is relatively small, consumption penalty associated with the sacrifice of some of the HEV's load leveling ability is not significant.

The benefits of the frequency-based power split strategy include the decrease of engine low-torque/low-BSFC operation, reduced engine transients, and simultaneous decrease of battery power peaks. The constraints and assumptions are as follows:

The filtering effects on power series should ensure that electrical power demands do not exceed component limits. Thus, a simple 2-point weighted moving average filter, as in Eq. 4.2, is adopted.

$$\widehat{P_n} = 0.7P_n + 0.3P_{n-1} \quad \text{Eq. 4.2}$$

Where $\widehat{P_n}$ is the low-pass filtered power demand to engine at current time step n, P_n is the actual power demand in current time step n, P_{n-1} is the actual power demand in previous time step n-1. The coefficient ratios, of which sum equals to 1, are selected considering the frequency of crossing battery power limits with real-world cycles. For example, at high-power discharge peaks, if the ratio before P_n is less than 0.7, the filtered power demand $\widehat{P_n}$ would be smaller, simultaneously the battery power loads $P_n - \widehat{P_n}$ would be larger.

The negative power demands were not filtered, but covered by regenerative braking, while respecting the battery charging limits and considering engine drag torque. When necessary the frictional brakes are blended with the regenerative torque. For the positive power demands, whenever the filtered value is smaller than the actual, the filtered power demand is met by engine while battery makes up the rest of the demand. The application on US06 drive cycle power demand is shown in Figure 4-4.

Prescribed minimum and maximum SOC is set at 0.35 and 0.75, respectively, and the initial SOC at 0.55. The moving average filter is applied between SOC values of 0.45 and 0.65. Outside of that range, the low-pass filter output is further adjusted by a multiplier before being passed on to the engine, as defined in Eq. 4.3. This is done to ensure charge sustaining operation. As an example, if the SOC falls to 0.4, the multiplier according to Eq. 4.3 is (0.35+0.75) / (2*0.4) = 1.375, and the engine power request is multiplied by 1.375. In summary, the engine power can be increased to the level that eliminates power-assist, or even provides battery charging. The same reasoning applies to situations when

SOC is higher than 0.65. In that case, the power assist is enhanced and charging events reduced/eliminated until SOC falls back to the pre-determined target. The aim is to enable a fair comparison of the final fuel economy predictions. Figure 4-5 shows the SOC variation of ISG hybrid model over the US06 drive cycle. Desired SOC levels are maintained over the drive cycle and the final value is close to the initial. After a small correction shown in Eq. 4.3, fuel economy numbers can be fairly compared.

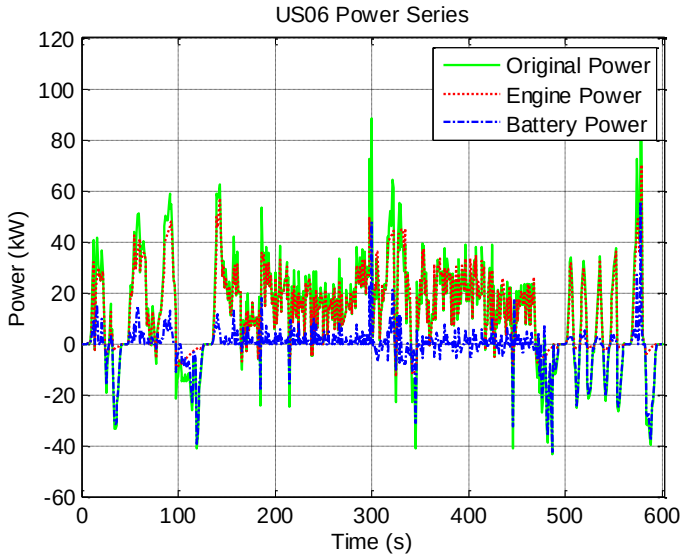


Figure 4-4. Original and filtered power demand of US06 drive cycle

$$(Engine\_Power\_Demand)\ Adjusted\_Multiplier = \frac{SOC_{low\_limit}+SOC_{high\_limit}}{2SOC_{actual}} \qquad \text{Eq. 4.3}$$

The SOC_{low_limit} is the minimum battery SOC, 0.35 in this case; SOC_{high_limit} is the maximum battery SOC, 0.75 in this case; SOC_{actual} is the actual SOC at current time step.

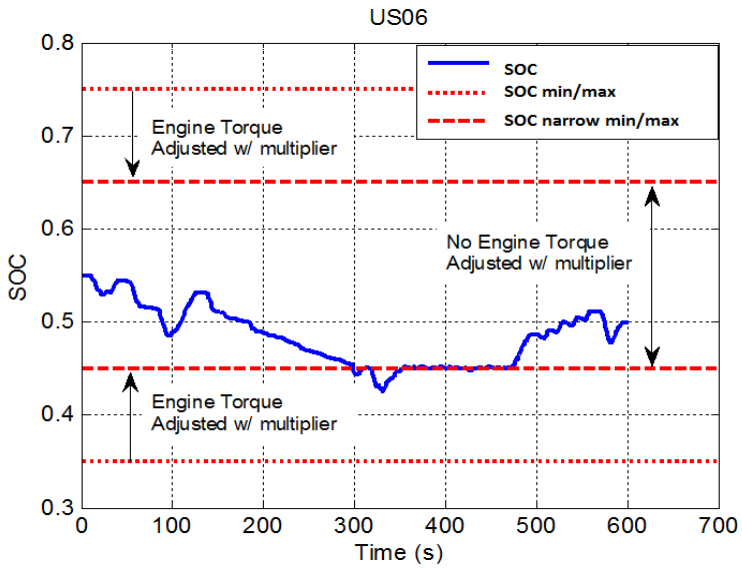


Figure 4-5. Battery SOC for US06 drive cycle

Simulated operation of the 48V mild hybrid powertrain system over the US06 drive cycle is illustrated in Figure 4-6. Figure 4-6(a) shows the extent to which engine operating points can be shifted towards the higher load with power-assist provided by the electric motor. In addition, idling is completely eliminated. Operating points of the electric motor are superimposed on its efficiency map in Figure 4-6(b). Negative torque points correspond to regeneration mode, and one can observe frequent saturation due to a limited power capacity of the 48V system.

Next, the 48V mild hybrid vehicle model is executed 1851 times for both conventional and mild hybrid configuration, using the naturalistic drive cycles from the SCAG database, in order to validate the fuel saving potential of the 48V hybrid under the real-world conditions, and understand the impacts of driver aggressiveness.

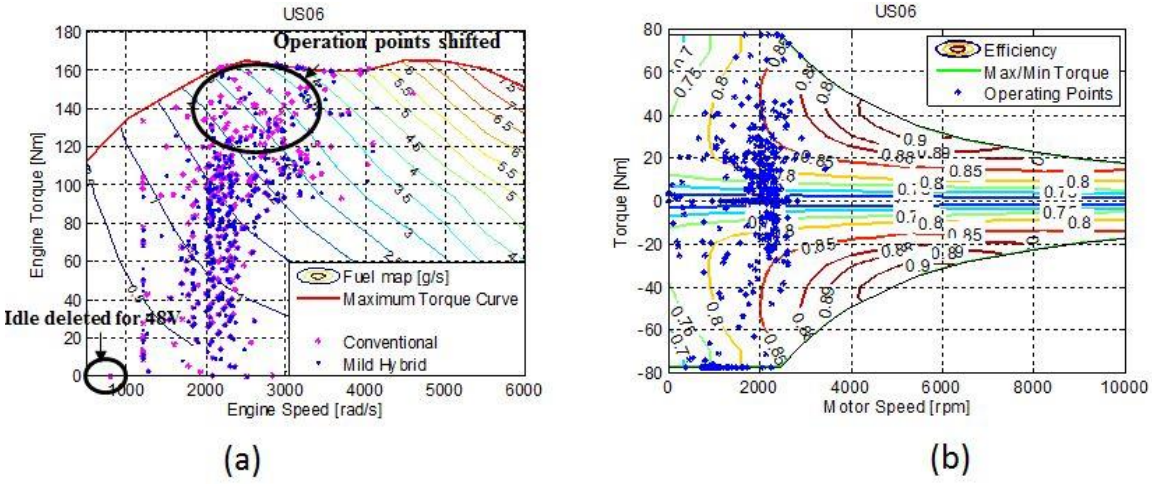


Figure 4-6. Operating points obtained over the US06 drive cycle, super-imposed on efficiency maps of the a) engine, and b) electric motor.

4.2 Results

This section presents the investigation of the potential for reducing fuel consumption with 48V mild hybridization. Specifically, relative contributions of different mechanisms for reducing consumption are quantified, i.e. Start&Stop, power assist and regenerative braking. Furthermore, analysis evaluates the sensitivity of fuel consumption to driver aggressiveness.

The baseline fuel consumtpion in the unit of L/100km is established by simulating the operation of the conventional vehicle, as described in the previous section. In the HEV simulation study, different functions can be enabled or disabled, to assess their respective contributions to the efficiency improvements. Results are presented in a manner that allows assessment of the 48V hybrid's fuel saving potential, as well as its comparison with a lower-cost 12V Start&Stop system. In addition, relative contributions of the Start&Stop function and the power assist/regeneration functions are quantified.

4.2.1 Fuel Saving Potential of 48V vs. 12V Hybridization

Fuel consumption reduction potential of the 48V mild hybrid stems from both Start&Stop function and the power assist/regeneration function. Start&Stop function is particularly effective in city driving. Therefore, Figure 4-7 presents fuel consumption reduction with a 48V system as a function of Idle Percentage (time based). Fuel consumption simulation results were fitted with a third-order polynomial with very tight coefficient of determination of $R^2 = 0.94$ in order to highlight the trends. As the trip idle percentage increases, the fuel savings with a 48V system increases exponentially. This emphasizes the critical role of the Start&Stop function.

To further illustrate this point, the 48V hybrid model was modified to include only the Start&Stop function while excluding power assist and regenerative braking. This effectively approximates a 12V hybrid vehicle. The relationship between fuel consumption reduction percentage and trip idle percentage is shown in Figure 4-7, thus enabling a comparison of 12V with a fully-featured 48V hybrid. Both lines display almost the same shape, and a relatively small downward translation of the 12V hybrid line. This leads to a conclusion that benefits of Start&Stop function dominate under the real-world conditions, even in the case of a 48V system. The specific power management strategy in this study plays a secondary role, since even with a frequency domain power-split e-motor and battery power often experience saturation.

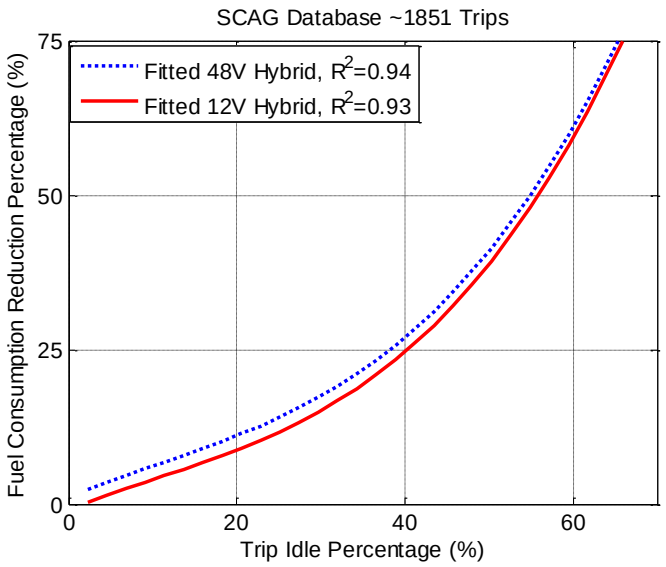


Figure 4-7. Predicted fuel consumption reduction percentage vs. trip idle percentage for both 48V and 12V hybrid vehicle. Results illustrate the critical role of Star&Stop function in the case of mild hybrids

While the dominant role of the Start&Stop function is emphasized whenever driving condition included significant idle time, the marginal benefits from power assist/regeneration functions worth a closer look. The aim is to correlate the fuel consumption improvements with driver impacts. Power management strategy applied in this study facilitates the analysis, since the high-frequency power spikes are mostly handled by the electrical sub-system. To extract the marginal benefits of power assist/regeneration, the Start&Stop function was disabled in the 48V hybrid vehicle model. Distribution of predicted fuel consumption reduction (%) for all naturalistic trips is given in Figure 4-8.

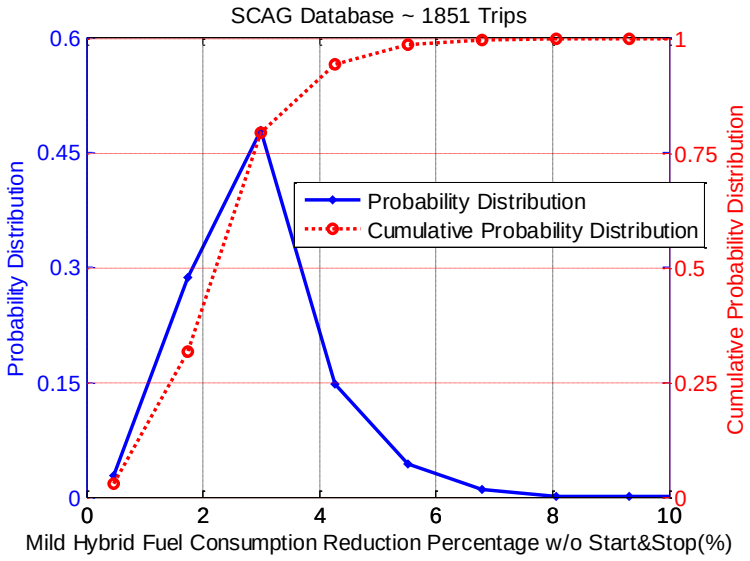


Figure 4-8. Distribution of mild hybridization fuel consumption reduction excluding Start&Stop function calculated for 1851 driving cycles in SCAG.

The most common fuel consumption reduction percentage excluding Start&Stop function is around 3%. According to the cumulative probability distribution, about 1-4% of fuel consumption reduction percentage points can be realized in 80% of trips. The individual gains from power assist/regeneration function are modest but still tangible, as expected considering the limited size of electrical components.

4.2.2 Impacts of Driver Aggressiveness on 48V System

Since a drive cycle's idling period when Start&Stop function engages is not directly related to driver activity, the impacts of driver aggressiveness can be evaluated better by excluding the Start&Stop function. In other words, the fuel saving potential of 48V system employing only the power assist/regeneration function can be correlated to driver aggressiveness. The simulation is setup accordingly, and predicted fuel consumption is assessed for three levels of driver aggressiveness. Statistics of predicted fuel consumption

reduction (%) are compiled in Table 4-2. Clearly, aggressive drivers benefit most from the 48V hybrid function.

Table 4-2. Statistics of fuel consumption reduction percentage excluding Start&Stop function for different driver aggressiveness in SCAG database

	Mean Fuel Consumption Reduction Percentage (%)	Standard Deviation of Mean Fuel Consumption Reduction Percentage (%)
Calm Drivers	2.15	0.87
Normal Drivers	2.82	0.83
Aggressive drivers	3.30	0.93

Results are obtained for 617 trips in each driver category. Before claiming a high-degree of confidence, any differences inferred from the analysis of bulk data should be proven to be statistically significant. In other words, statistical analysis ought to prove that trends hold regardless the variance of data. A hypothesis tests is setup for this purpose.

4.2.2.1 Hypothesis Test on Mean Fuel Consumption Reduction Percentage by Driver Aggressiveness

Hypothesis test is setup to evaluate the statistical significance of v fuel consumption variations due to driver aggressiveness. The basic idea is to compare the data set generated in the actual study against a synthetic data set. Basic procedures comprises several steps, i.e. (1) formulate a null hypothesis H0 and, if appropriate, an alternative hypothesis H1; (2) design a test procedure by which a decision can be made; (3) use

statistics to finalize the test procedure, making judgments about the significance level; (4) apply the test to individual cases; (5) make decisions.

The two-sample t-test [33] is applied to determine if means of both two populations are equal. The purpose is to test whether aggressive drivers achieve more fuel consumption reduction compared to calm drivers.

In this case, the two-sample t-test is defined as:

Null hypothesis, H_0: $\mu_{aggressive} = \mu_{calm}$

Alternative hypothesis, H_a: $\mu_{aggressive} > \mu_{calm}$

$\mu_{aggressive}$ and μ_{calm} are the population mean of fuel consumption reduction percentage with mild hybridization. Test Statistics:

$$T = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{{s_1}^2}{N_1} + \frac{{s_2}^2}{N_2}}} \quad \text{Eq. 4.4}$$

$\overline{X_1}$ and $\overline{X_2}$ are the sample mean fuel consumption reduction percentage of aggressive and calm drivers as in first column of Table 4-2; s_1 and s_2 are the sample standard deviation of fuel consumption reduction percentage of aggressive and calm drivers as in second column of Table 4-2; N_1 and N_2 are the sample size, 617 for each driver category in this study.

The significance level α is set at 0.05, equivalently 95% confidence interval.

The rejection for the null hypothesis is $T > t_{1-\alpha,v}$, which is the critical value of the t-distribution with v degrees of freedom, calculated as:

$$v = \frac{(\frac{{S_1}^2}{N_1}+\frac{{S_2}^2}{N_2})^2}{\frac{({s_1}^2/N_1)^2}{N_1-1}+\frac{({s_2}^2/N_2)^2}{N_2-1}} \quad \text{Eq. 4.5}$$

According to t-distribution table, to compare calm and aggressive driver, $t_{0.95,1226}$ = 1.645.

We further put values from Table 4-2 into Eq. 4.4 and Eq. 4.5, and get T = 22.5452 > $t_{0.95,1226}$ = 1.645. Thus, the criteria of rejection for the null hypothesis is met, the null hypothesis that $\mu_{aggressive} = \mu_{calm}$ should be rejected while accepting the alternative hypothesis $\mu_{aggressive} > \mu_{calm}$. In other words, the difference in mean fuel consumption reduction percentage between aggressive and calm drivers is statistically significant, and our conclusion that aggressive drivers benefit more from mild hybridization stands. The same procedure can be applied to tests between calm and normal drivers, normal and aggressive drivers. Results are $T_{normal-calm}$ = 14.0532> $t_{0.95,1230}$ = 1.645 and $T_{aggressive-normal}$ = 9.4700> $t_{0.95,\ 1216}$ = 1.645, both validating the statistical significance of findings.

4.2.2.2 Hypothesis Test on Battery Loads by Driver Aggressiveness

Simulation of the 48V hybrid generates a plethora of powertrain related variables, therefore allowing a deeper look into impact of driver aggressiveness on powertrain behavior. This presents opportunities to take the study beyond fuel economy predictions. In particular, simulation prediction can provide useful insight into the impact of driver aggressiveness on battery duty-cycle. Since the magnitude and frequency of the energy flows in/out of the battery is known to have a significant impact on its health, battery duty cycle analysis can bring an additional dimension to the 48V HEV study. This will be addressed in chapter 5 and 6.

Battery's current trace representing the electrical loads was generated for every drive cycle. After the zero currents are excluded, the mean values of the average absolute current and the standard deviations of the average absolute current are compiled in Table 4-3.

Table 4-3. Statistics of mean values of average absolute current for different driver aggressiveness in SCAG database

	Mean of Average Absolute Current (A)	Standard Deviation of Average Absolute Current (A)
Calm Drivers	8.8709	2.5227
Normal Drivers	9.6375	2.3028
Aggressive drivers	10.8202	2.5519

Similar to the trend of mean fuel consumption, the mean values of average absolute current increase with driver aggressiveness. This can pose challenges regarding battery life. Hypothesis tests conducted on the variations of mean average absolute current across all driver categories indicates that finding are statistically significant. According to Eq. 4.4 and Eq. 4.5 and values in Table 4-3, the test results are: $T_{aggressive\text{-}calm} = 13.4936 > t_{0.95, 1211} = 1.645$; $T_{normal\text{-}calm} = 5.5749 > t_{0.95, 1231} = 1.645$; $T_{aggressive\text{-}normal} = 8.5467 > t_{0.95, 1228} = 1.645$.

In short, aggressive drivers not only increase the real-world fuel consumption, but also impose highest loads on the battery and increase the chances for its premature deterioration. The proposed methodology in this study can be extended to modify the

supervisory control and explore the trade-off between fuel saving potential and battery life for real-world conditions.

4.3 Conclusions

This chapter translates the real-world drive cycles into cell duty cycles via the simulation of a 48V mild hybrid system. A particular aspect emphasized in the study is the impact of driver aggressiveness on fuel saving potential of the near-term 48V hybridization technology. A naturalistic database based containing 1851 individual trips is processed using a novel definition to determine a distribution of "driver aggressiveness". Methodology splits the signals of vehicle jerk in frequency domain to separate the low-frequency range attributed to traffic, from the high-frequency range influenced by the driving style. The ratio of energy aggregated in the high-frequency range to total, corrected by a mean value of periodogram across the whole range, provides the metric for categorizing trips based on their aggressiveness.

Both the conventional and the 48V mild hybrid configurations of a compact passenger car are modeled using MATLAB/Simulink. Supervisory controller is designed to facilitate consistent power management while assessing the impacts of driver aggressiveness. A low-pass, moving-average filter is used to split the power demand; splitting the low-frequency/large-amplitude signal to the engine, and high-frequency but small-amplitude signal to the battery.

The fuel consumption of the baseline conventional vehicle and the 48V hybrid was simulated for all 1851 driving cycles from the SCAG database. Analysis of results was approached from two angles, i.e. the findings were used to both assess the contribution of

different fuel economy improvement mechanisms in a 48V system, and correlate the driver aggressiveness with the magnitude of improvements. The conclusions are as follows:

- 48V HEV system with Start&Stop, torque assist and regenerative braking functions, provide tangible fuel consumption reduction, strongly correlated with the percentage of idle time
- 48V hybrid can surpass the 12V hybrid by 1-4% of relative fuel consumption improvement in 80% of cases. Therefore, while Start&Stop plays a dominant role, the contribution of electric torque assist/regeneration is still tangible
- Aggressive driving can lead to ~25% increase of fuel consumption with the conventional powertrain. Hybridization can partially mitigate the effects of aggressive driving on fuel consumption, as aggressive drivers benefit the most from the 48V technology. However, insights also indicate that aggressive style leads to highest average battery currents, which can stress the battery and affect its life.

Overall, the results confirm the tangible improvements of fuel economy with a 48V system compared to a 12V HEV, and indicate increased positive impact of hybridization in case of aggressive driving. The quantitative assessments of the driver style aid in optimizing the design and power management under real-world conditions, thus enabling the balance between the regulation requirements and consumer expectations.

This simulation study also translated the real-world vehicle drive cycles to battery cell duty cycles categorized by driver styles and road conditions. The subsequent chapter

serves to characterize the duty cycles and presents a methodology to synthesize them into representative profiles feasible for aging experiments.

5 Synthesis and Experimental Validation of Battery Aging Test Profiles

Electrified powertrains are gaining a larger market share due to regulatory strategies and customer desires. Ranging from micro hybrid to all electric powertrains, however, their success highly depends on satisfactory life-long performance of onboard energy storage systems. In order to guide the battery pack design and management, real-world representative aging tests are essential for a successful product development. Focusing on cycle aging, this study proposes a methodology based on Welch's power spectral density estimation to extract and synthesize characteristic cell aging test profiles from real-world duty cycles. The 48V mild hybrid vehicle model in MATLAB/Simulink relates the real-world vehicle drive cycles to cell-level duty cycles. Compared to other test profiles currently available in the literature, which are mostly composed of constant pulses or real-world transient pulses, the newly designed test profiles combine both clarity and real-world representativeness. Experimental validation is conducted with the proposed aging profiles on nickel manganese cobalt (NMC) lithium-ion batteries. The impacts of real-world driving activities on cell aging are studied and results lay foundation for aging models to identify the most influential aging mechanisms. This chapter focuses on incorporating knowledge of real-world driving activities into battery cycle aging test on cell level.

In contrast to calendar aging tests which typically consider only temperature and SOC (State-of-Charge) [62][63][64], cycle aging tests incorporate additional stress factors, including ΔSOC, C-rate, Ah-throughput and Discharge/Charge events. The design of

cycle aging tests is largely driven by the type of applications, such as hybrid electric vehicles (HEVs), plug-in hybrid electric vehicles (PHEVs) and Electric Vehicles (EVs) [65]. Although battery cycle aging tests are being conducted worldwide, e.g., United States, Japan and Europe, a unified aging test protocol does not exist as different procedures are being used from country to country.

This chapter investigates the battery operations in the 48V mild hybrid vehicle [21] and extracts useful information to synthesize repeatable test profiles to be run in laboratory settings. The cell cycle aging stress factors are closely linked to driver styles as well as road conditions. To account for these uncertain causes, nearly 2000 real-world drive cycles with detailed second-by-second speed traces are adopted in this study from the public travel survey provided by National Renewable Energy Laboratory (NREL) [8].

Cycle aging test profiles from the literature can be summarized in three categories. 1) *Rectangular constant current/power pulses*. These cycles, shown in Figure 5-1(a), are widely used to separate the impacts of different stress factors, usually with no implications from real-world scenarios [66][67][68][69][70][71][72][73][74]. 2) *Transient real-world cell duty cycles.* These traces, shown in Figure 5-1(b), are directly adopted as test profiles to fully consider the intrinsic cycle variability although complex statistical analysis is needed to explore the correlation between aging phenomena and operating conditions [75][76]. 3) *Semi-transient test profiles.* These cycles, as depicted in Figure 5-1(c), which are usually synthesized from statistical analysis of real-world current/power profiles, are the most popular candidates for cycle aging campaign

[77][78][79], as they constitute a well-compromise between clarity and real-world resemblance.

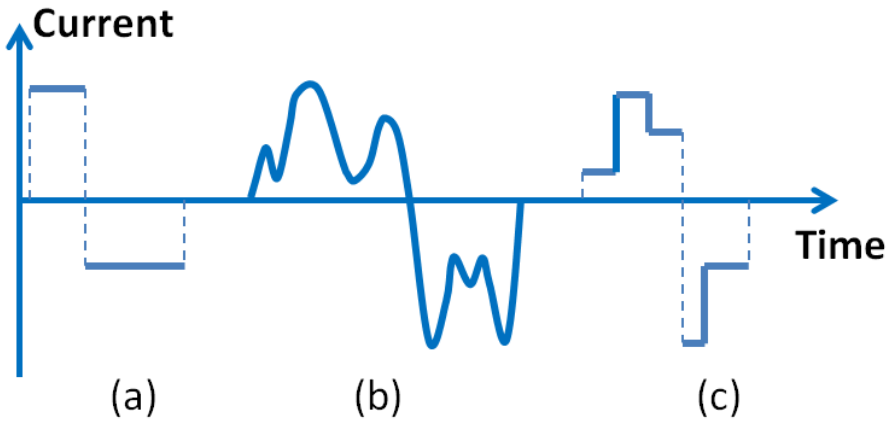


Figure 5-1. Illustration of commonly used cycle aging test profiles

In this study, we propose a novel methodological framework to characterize real-world cell duty cycles in frequency domain with their power spectral density (PSD) estimations, and synthesize them into representative *semi-transient test profiles.* Through our process, the output profiles are dynamic, charge-neutral, within manufacturer operation limits, and most importantly can be well traced back to the real-world driving scenarios. These representative profiles can be coupled with different temperatures to set up the aging test for aging assessment.

Lithium-ion batteries comprise a family of chemistries that employ mostly carbon-based anode and various cathode materials. Each provides disadvantages and advantages in several aspects: safety, performance, cost, specific energy, specific power and lifespan [80][81]. The cycle aging experimental results conducted in this study aim at supplementing the understanding of lifespan of nickel-manganese-cobalt (NMC)-based cathode lithium-ion cells [69][70]. Previous findings on other popular cathode materials,

such as nickel-cobalt-aluminum (NCA) [67][68] and iron-phosphate (LFP) [71][72][73] can be found in the literature. The experimental results obtained in this chapter can be applied to identify cell aging models, useful for the design and control of onboard lithium ion battery pack.

5.1 Approach

This study as above applied the part of GPS (Global Positioning System) enhanced household travel survey from June 2001 to March 2002 conducted by SCAG (Southern California Association of Governments), of which de-identified data are publicly available through the U.S. Department of Energy's National Renewable Energy Laboratory (NREL) [8]. 1851 trips from 292 passenger vehicles with a total driven distance of about 25,000 km are adopted in this study. The statistical distributions of travel distances, cycle mean velocity, idle percentage and mean positive acceleration rates can be referred to Figure 4-1. Our definition of driver aggressiveness abstracts the frequency of changing the vehicle power demand, which is essentially equivalent to the frequency of changing power/brake pedal position, and can be referred to Eq. 3.6 and Figure 3-19.

Table 5-1 compiles the number of trips in each category determined by drive cycle mean velocity and driver aggressiveness, of which the probabilistic distributions are equivalently separated into three parts respectively (calm, normal, aggressive for driver styles; low-, medium- and high-speed for traffic conditions). While aggressive style dominates the low-speed driving, calm style rules the high-speed scenario. This interesting trend implies the real-world variability of driver behaviors.

Table 5-1. The number of real-world trips in each category

Number of Trips	Calm	Normal	Aggressive
Low Speed	**64**	**235**	**318**
Medium Speed	**150**	**234**	**233**
High Speed	**406**	**141**	**70**

The 48V mild hybrid vehicle is described in previous chapter. Since the NMC lithium ion chemistry emerges as a very strong candidate for 48V system, the HEV simulation was modified to align the work in this chapter with relevant industry needs. The NMC cell specifications shown in Table 5-2 were used to develop a model incorporated in the HEV simulation throughout this chapter.

An exemplary drive cycle and its related battery cell current profile are compared in Figure 5-2. The cell discharge and charge periods display significantly different patterns. In contrast to charge current demands, the discharge counterparts with higher limit are more segmented. All drive cycles in the database can then be processed through the 48V hybrid vehicle model to attain related cell current profiles. Next section serves to identify latent features in real-world battery duty cycles and extract characteristic aging test profiles.

Table 5-2. Specifications for battery and starter-generator

Specifications	Units	Values
Battery		
Cell Chemistry		Nickel-Manganese-Cobalt (NMC) Lithium ion
Cell Rated Capacity	Ah	2.0
Cell Max Discharge Pulse Current	A	30
Cell Max Charge Pulse Current	A	12
Battery Pack Configuration		14 series/4 parallel
Starter-Generator		
Supply Voltage	V	42-58
Max/Continuous Power	kW	25/20
Max Torque	Nm	75 @ (0~3000 rpm)
Max Speed	rpm	10000

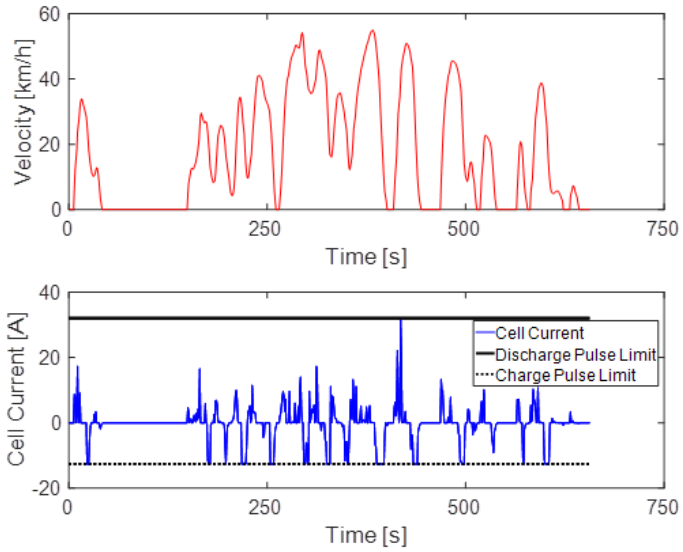


Figure 5-2. A drive cycle speed profile and its battery cell current profile

5.2 Development of Representative Battery Test Profiles

With a snapshot over 460s~540s time window of the current profile in Figure 5-2, discharge and charge micro profiles are defined and displayed in Fig. 6. The zero current snippets during vehicle stops are not considered. This study aims to investigate whether there is any underlying cyclic pattern in discharge/charge micro profiles and how these patterns vary across real-world scenarios. While only sequential data are observed in time domain, cyclic events can be better captured in frequency domain with the Fourier analysis. The introduction of Fourier analysis can be referred back to section 3.2.2.1

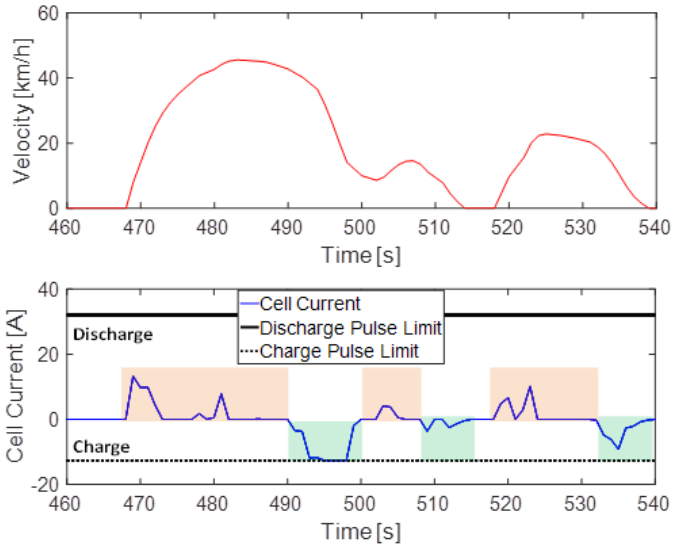


Figure 5-3. The definition of discharge and charge micro profiles

5.2.1 Characterization of Cell Duty Cycles

As the differences between discharge and charge activities had been emphasized, all the discharge micro profiles and charge micro profiles in this study are separately concatenated in each driving data category. Additionally, a pre-processing step, namely mean-adjusting, is preferred to make hidden frequency features more discernible;

otherwise, the component at zero frequency might be too large to disguise other frequency components. For both discharge and charge data, we adjust by flipping every other micro profile, as displayed in Figure 5-4 and Figure 5-5. Considering the length of data and accuracy requirement, we adopt the Welch's method and illustrate the process on the mean-adjusted charge data in Figure 5-5. It applies to the discharge data in Figure 5-4 as well. A hamming window of length 256 is moving along the time series with overlap of length 128; these modified periodograms are averaged to output the estimated PSDs in Figure 5-6 and Figure 5-7 The length 256 and 128 are chosen to achieve satisfactory smoothing level.

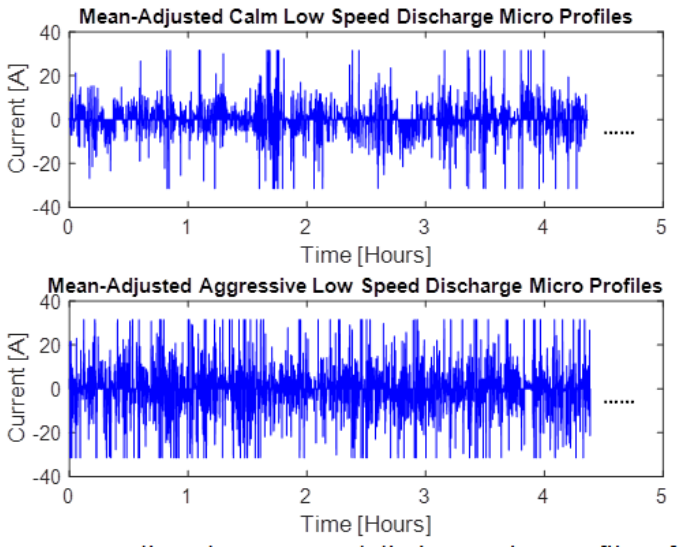


Figure 5-4. Mean-adjusted concatenated discharge micro profiles of calm low speed trips (~4.5 hour data for display)

As the power spectral density indicates the distribution of time series variance across the frequency domain, we emphasize the different peak values of PSDs between calm and aggressive drivers in low-speed driving activities in Figure 5-6 for illustration. In discharge, aggressive drivers have the peak frequency of 0.0215 Hz, translated to a 46-

second sinusoidal period in time domain; calm drivers have the peak frequency of 0.0156 Hz, equal to a 64-second period. In charge, the comparison is 0.0469 Hz of aggressive to 0.0390 Hz of calm; equivalently a period of 22-second to 26-second. The fact that aggressive drivers have shorter discharge/charge events applies for high-speed driving activities in Figure 5-7 as well. Another difference is the area under the curve of PSDs, which is proportional to the variances of concatenated time series summarized in Table 5-4.

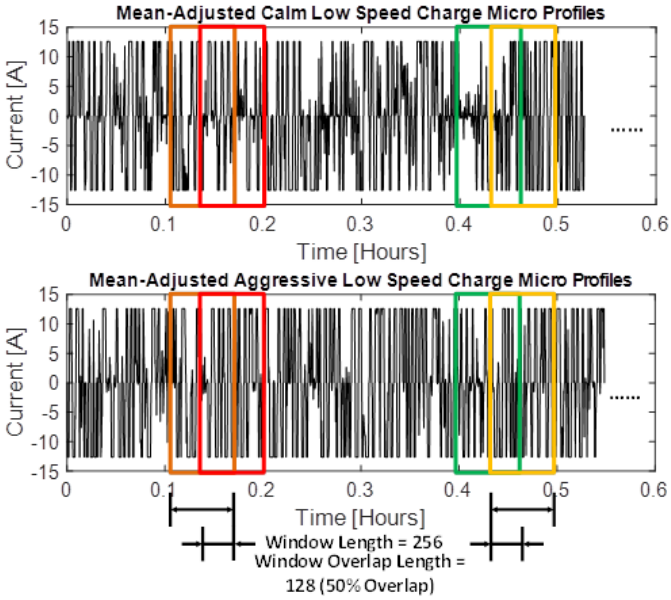


Figure 5-5. Mean-adjusted concatenated charge micro profiles of calm low speed trips (~0.5 hour data for display); also include illustration of Welch's moving window method for power spectral density estimation

The information in Table 5-3 and Table 5-4 can be summarized as: 1) the batteries of aggressive drivers switch faster between discharge and charge; 2) the batteries in low speed driving conditions switch faster between discharge and charge; 3) In each category the charge periods are shorter than discharge counterparts; 4) The differences of current variances in discharge scenarios are significantly enlarged by both driver aggressiveness and driving speed; 5) The differences of current variances in charge scenarios are relatively small because they are mostly saturated by the charge current limits.

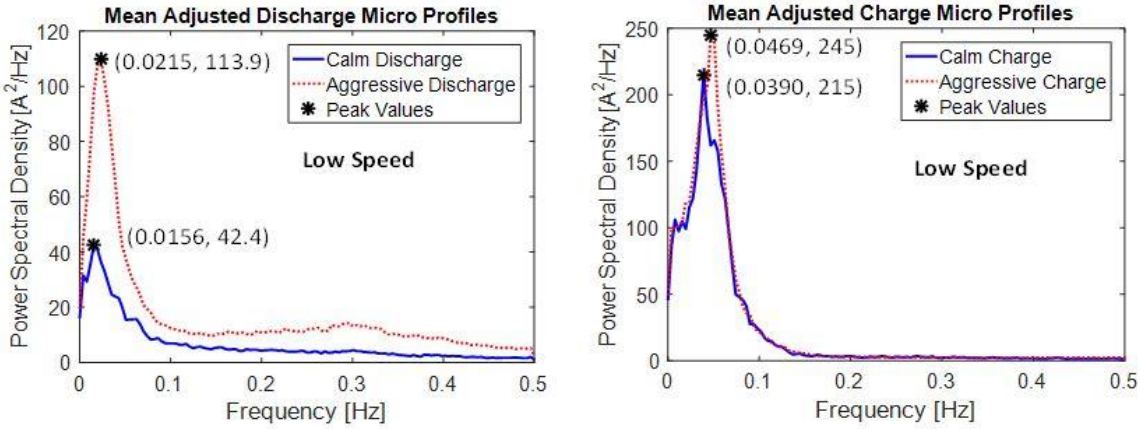


Figure 5-6. Power spectral density estimation for low speed drive cycles

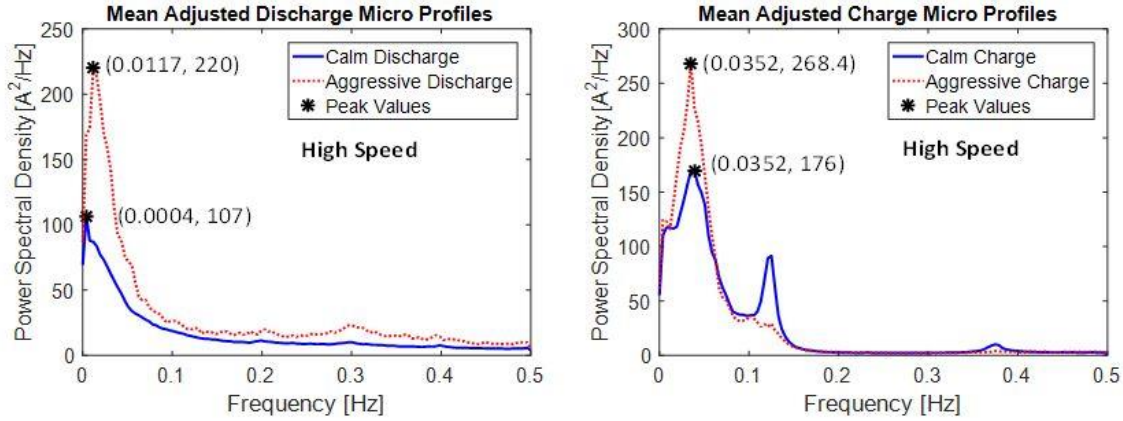


Figure 5-7. Power spectral density estimation for high speed drive cycles

Table 5-3. Significant frequency of concatenated mean-adjusted discharge and charge micro trips for calm and aggressive drivers in low/high-speed drive cycles

Peak Frequency (Hz)	Mean-Adjusted Discharge	Mean-Adjusted Charge
Calm Low Speed	**0.0156**	**0.0390**
Calm High Speed	**0.0004**	**0.0352**
Aggressive Low Speed	**0.0215**	**0.0469**
Aggressive High Speed	**0.0117**	**0.0352**

Table 5-4. Variance of concatenated mean-adjusted discharge and charge micro trips for calm and aggressive drivers in low/high-speed drive cycles

Variance (A^2)	Mean-Adjusted Discharge	Mean-Adjusted Charge
Calm Low Speed	**21.85**	**74.34**
Calm High Speed	**52.96**	**84.00**
Aggressive Low Speed	**54.96**	**82.44**
Aggressive High Speed	**99.03**	**91.10**

5.2.2 Synthesis of Battery Aging Test Profiles

Information on Table 5-3 and Table 5-4 can be used to synthesize characteristic cell aging test profiles based on cycle speeds and driver styles. The synthesized aging profiled are in the form of sinusoidal waves with the peak frequencies from estimated PSDs and the period of mean-adjusted time series. The derivation of such profiles is illustrated, see Figure 5-8, considering the aggressive high speed category. Constructed with half sinusoidal profiles separately from discharge and charge data, the synthesized characteristic profile should comply with current limits and must be Ah-throughput

neutral to guarantee repeatability. Following the same philosophy, the characteristic profiles for calm low speed, calm high speed, aggressive low speed and aggressive high speed driving conditions are compiled in Figure 5-9. For all of them, the charge parts are close to the manufacturer limit, but the discharge parts are significantly different in duration and current magnitude. Aggressive drivers tend to discharge their batteries faster and harder; high-speed driving increases battery discharge duration and current magnitude.

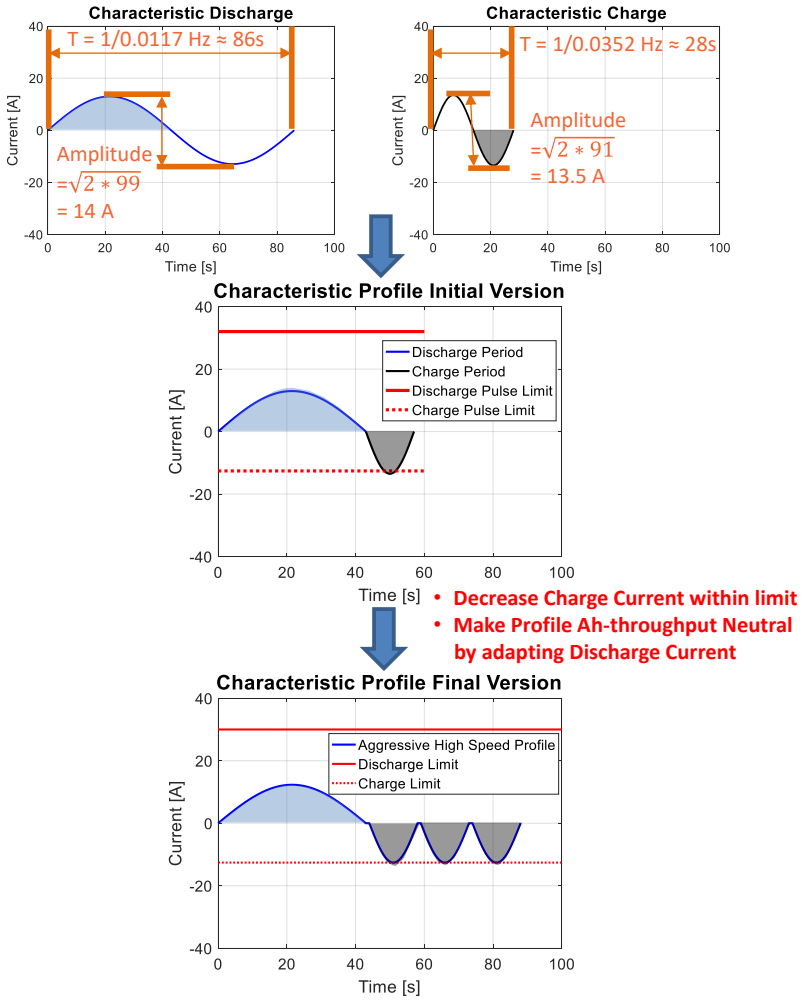


Figure 5-8. The process of synthesizing the characteristic battery test profiles for aggressive high speed category

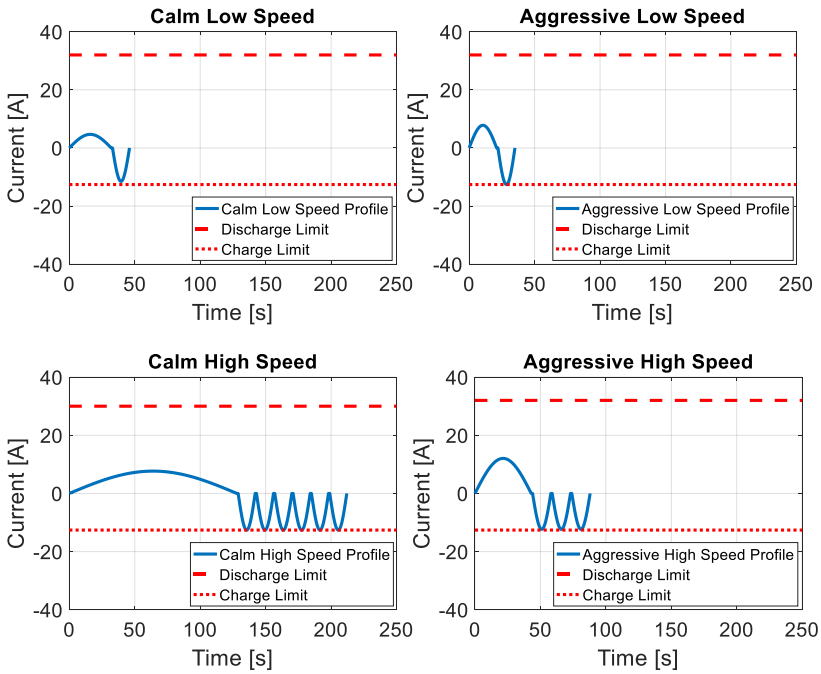


Figure 5-9. Characteristic battery test profiles for four categories of driving data

5.3 The Experimental Setup and Aging Test Results

This section discusses the experimental setup and the aging test results using the synthesized charge-neutral profiles developed in the previoous sections. Experiments were carried out at the Battery Aging and Characterization (BACh) Laboratory at the Automotive Engineering Department at Clemson University. The experimental setup used for the aging campaign, shown in Figure 5-10, is composed of the Arbin BT-2000 battery cycler with a programmable power supply and an electronic load; a MITS Pro data acquisition software for the programming of test profiles and the control of the Arbin cycler; Peltier junctions that are in direct contact with the cell fixtures for thermoelectric temperature control. This study applies cylindrical Nickel-Manganese-Cobalt (NMC)

18650 lithium ion cells with nominal capacity of 2Ah and nominal voltage of 3.6V. Other cell specifications can be referred to Table 5-2.

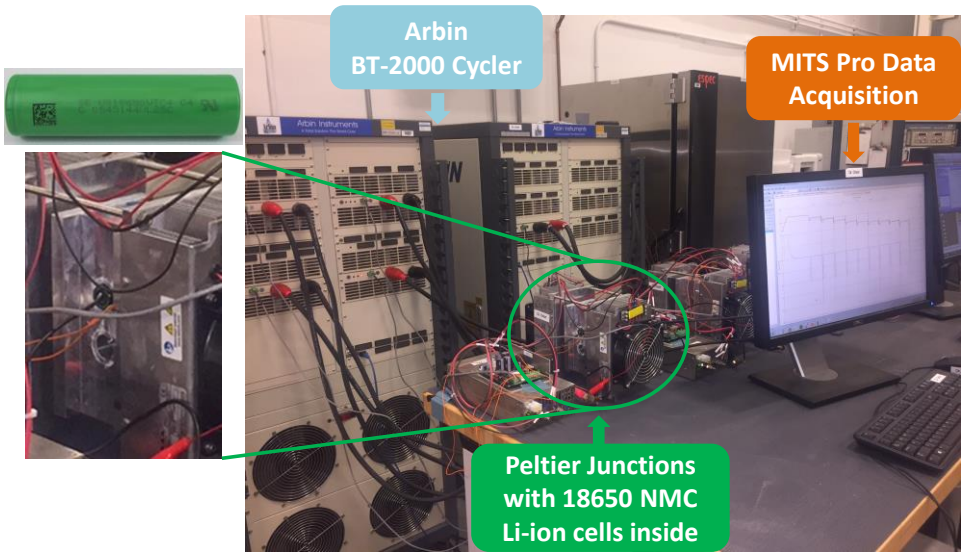


Figure 5-10. Experimental test set-up at the BACh Laboratory, CU-ICAR

In the experimental aging campaign the calm low speed (CLS) and aggressive high speed (AHS) profiles were tested at two temperatures, 23 °C and 45 °C, respectively to form the test matrix shown in Table 5-5. For each test condition, two cells were tested for repeatability concerns. At the beginning and periodically, typically 3 weeks, all cells underwent capacity characterization tests at room temperature of around 25°C. The 1C capacity tests was performed by first charging the cell to 4.2V at constant current of 1C, then holding constant voltage at 4.2V until the current dropped to C/50. The fully charge phase was followed by one hour rest, then finally the 1C discharge until the cut-off voltage of 2.5V was reached. A flowchart of this aging campaign is visualized in Figure 5-11.

Table 5-5. Aging test matrix, numbered cells with aging conditions

Aging Conditions	Cell No. #
Calm Low Speed 23°C (CLS23)	Cell 1 / Cell 2
Calm Low Speed 45°C (CLS45)	Cell 3 / Cell 4
Aggressive High Speed 23°C (AHS23)	Cell 5 / Cell 6
Aggressive High Speed 45°C (AHS45)	Cell 7 / Cell 8

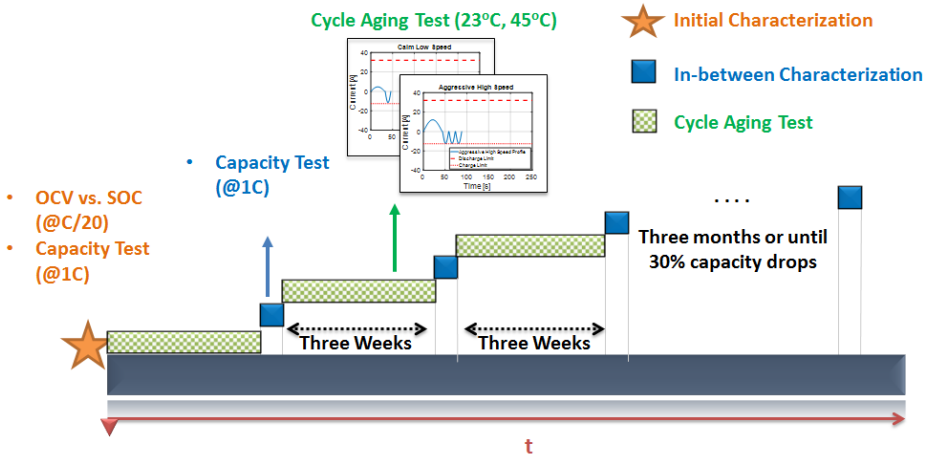


Figure 5-11. The flowchart of the aging test

The normalized capacity loss results from three-month characterization tests are shown in Figure 5-12. The cells undergoing the AHS45 cycle (i.e. cells #7 and #8) only lasted 2 months considering the severe aging rate of the cycle. For better comparison, the x-axis is in Ah-throughput rather than time. Cells #1 and #2, undergoing the CLS23 cycle show no capacity loss after 3-month aging period. A slight capacity recovery is instead observed. Interestingly, for CLS45 Cell#3 and AHS23 Cell#6, some capacity recoveries are also observed during the first few aging characterization stages. But when compared to their repeated counterparts, those recoveries can be regarded as outliers.

The capacity loss severity is in following orders: the aggressive high speed 45°C case as the severest, calm low speed 45°C case as the second, aggressive high speed 23°C as the third, calm low speed 23°C as the mildest. The aging results can be interpreted as following: 1) the impact of temperature on aging is more significant than the discharge current magnitude, or the driving style; 2) the impact of discharge current rate on aging is larger at high temperature; 3) the impact of temperature on aging is larger as the discharge current rate increases. The aging results derived with the test results emphasize the need for an efficient battery thermal management as well as the benefits one might obtain from eco-driving.

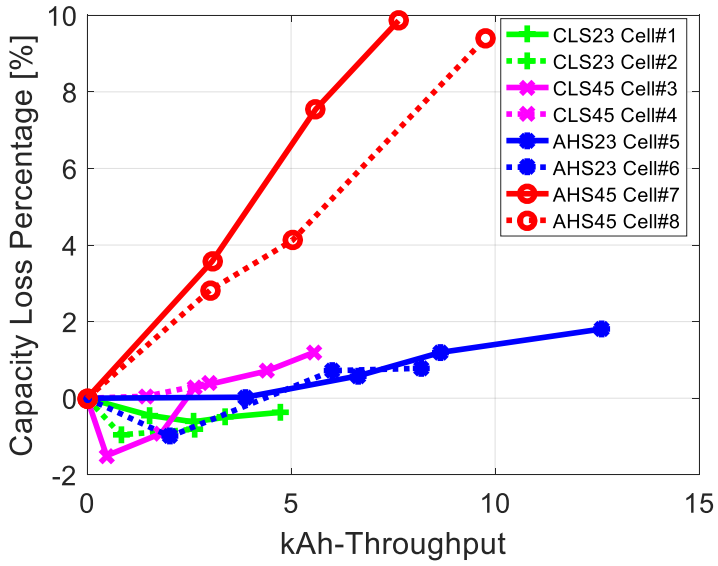


Figure 5-12. Battey aging test results with in-house characteristic profiles

5.4 Conclusions

This chapter analyzes the real-world battery duty cycles and synthesizes characteristic aging test profiles. Using naturalistic vehicle drive cycles, which are categorized by mean

speed and driver aggressiveness, a 48V mild hybrid vehicle model with appropriate power management strategy is constructed in MATLAB/Simulink to generate real-world battery duty cycles. By Fourier analysis on discharge and charge data separately, we conclude that: 1) discharge pulses are generally longer in duration than charge pulses; 2) aggressive driver style shortens the discharge/charge durations while high-speed driving activities lengthen them 3) aggressive driver style and high-speed driving activities both stimulate larger discharge/charge current rates. Recognizing the patterns, characteristic sinusoidal test profiles are synthesized with peak frequencies and amplitudes respectively from power spectral density estimations and time series variances. After adapting those profiles to be within current limits and Ah-throughput neutral, they are qualified for aging tests. Under different temperatures and 'driving styles', the cycle aging results reveal the negative impacts of aggressive driver style at high temperature on cell capacity. The methodologies for the characterization of real-world duty cycles and synthesis of representative profiles can be flexibly extended to new datasets. The aging test results lay the foundation for the cell aging modeling and the investigation of cycle aging mechanisms. This study has also shown that repeatability of tests is an important aspect to account for when aging campaigns are being conducted.

6 NMC Lithium ion Battery Aging Modeling

Lithium-ion chemistry is expanding its dominance from the portable devices to automotive applications. However, to fulfill customer expectations regarding the performance of electrified vehicles in real-world scenarios, lithium-ion batteries are still fundamentally compromised due to effects of aging on regarding electric range, power capability and service life. In response, manufactures have chosen to either oversize the battery pack or provide a smaller pack with a strong warranty program. Both result in added ownership costs. Efficient battery management based on accurate state estimation offers a path to extended life and reduced cost. Among all battery internal states, the state-of-health (SOH) quantifies the capacity loss or impedance rise across usage. Therefore, accurate SOH estimations are critical for pursuing battery aging conscious pack design and including battery aging into supervisory control optimization. This chapter focuses on identification of parameters critical for SOH estimation, and their subsequent calibration of aging models using experimental results generated with battery cycler setup.

The main aging mechanisms of lithium-ion batteries have been discussed in several review papers [15][16][17], and can be briefly summarized based on a phenomena occurring at two electrodes. For carbon-based negative electrode (commonly graphite), there is a passivation protective layer forming at its electrolyte/electrode interface, named solid electrolyte interface (SEI), during first few charging processes. Upon operation tensions, the SEI undergoes thickening and reformation by consuming available lithium ions. The SEI development is accepted as the major reason for lithium ion battery aging.

For positive electrode, despite various material choices, its aging can be attributed to active material loss and a solid permeable interface (SPI) without full passivation. High temperature and state-of-charge (SOC) can aggravate the aging conditions at both electrodes.

Battery aging processes are commonly distinguished by calendar aging under storage and cycle aging under usage, of which an additive degradation relationship is often assumed [16]. While calendar aging is stressed by time, temperature and SOC, cycle aging introduces additional stressors such as Ah-throughput, ΔSOC, C-rate, charge/discharge events. We focus on cycle aging under real-world usage. To separate highly-coupled impacts from the aging stressors in real-world operation, well-controlled test activities are usually set up to age the cells in an accelerated manner. A series of such efforts has been carried out for the family of lithium-ion chemistry which employs different positive electrode materials. Specifically on the cycle aging of $Li(NiMnCo)O_2$ (NMC) 18650 cylindrical cells, [69][82] combined several SOCs with different ΔSOC only at 1 C-rate at 35^{o}C, [83]applied a comprehensive test matrix comprised of 5 ΔSOC , 3 charging rates, 3 discharge rates and 3 temperatures at 50% SOC but aiming at high energy applications (like EVs). A gap clearly exists in the understandings of NMC performance in high power applications. This scarce of knowledge extends to large format NMC pouch cells as well [70][84][85][86]. Taking HEV propulsion as example, with shallow ΔSOC around a fixed SOC for charge sustaining operation, a wide range of discharge/charge current rate and temperature should be the dominating aging factors. Although the NMC chemistry has been generally considered to be limited on high current

rates [64][87], this study tends to evaluate its feasibility in HEV configuration quantitatively with experimental and simulation results.

Our aging campaign builds on a test matrix filled with in-house customized aging profiles, which are extracted from real-world battery duty cycles concerning road conditions and driver styles in chapter 5. They mainly differ in discharge current rate, charge/discharge event duration and Ah-throughput per cycle. A 48V mild hybrid vehicle, which is a near-term affordable electrification solution, is modeled in chapter 4 using MATLAB/Simulink to relate the vehicle drive cycles and battery duty cycles. Further coupled with two temperatures, the NMC cells exhibits capacity loss and impedance rise in varying degrees.

The cycle aging results stimulate the needs to fit appropriate aging models to suggest the internal degradations, imply cell design trade-offs and facilitate online management. A variety of models has been proposed for reasonable inference and can be classified into four categories. 1) Data-driven models, including artificial neural networks [67] and time series analysis [88], are simple yet suffer from inaccuracy especially for unlearned conditions. 2) Empirical functions are directly set up between capacity loss/impedance rise and Ah-throughput, but they require intensive calibration efforts in their prefixed coefficients [69][72][74][85][89] and shed little insights into aging mechanisms. 3）Equivalent circuit based models approximate the batteries with resistors, inductors, capacitors and so on [90][91]; the aging of these electrical components as functions of usage patterns can update the circuit parameters for battery health information [92][93]. Nevertheless, their implications into aging mechanisms are still limited. 4) Physics based

aging models firstly rely on electrochemical battery models, ranging from single-particle model, to porous electrode pseudo-2D, and way up to kinetic Monte Carlo model in the increasing order of complexities [94]. Then either aging sources (mostly the growth of SEI) are explicitly inserted into the governing equations [95][96][97][98][99] or selected model parameters are updated with some curve-fitting procedures [100][101]. To exchange for aging insights and accuracy, they calls for accurate identification of a large parameter set and probably needs model reduction for online applications.

Physics-based aging models typically involve dozens of parameters and are thus under the risk of overfitting. As some parameters (geometry, mass and concentration) can be measured directly, other ones (diffusion, kinetic and stoichiometric) are not easily accessible. The parameter correlations and parameter sensitivities further complicate the problem. They can be summarized as the issue of parameter identifiability. Therefore, the challenge of reliable parameter identification must be tackled before any meaningful conclusions based on model parameters are derived. Several literature have proposed scientific methodologies, such as sensitivity analysis [101], fisher information matrix [102] and Monte Carlo sampling [103]. Among them, Monte Carlo sampling method, a global parameter space searching algorithm, has the potential to evaluate parameter identifiability comprehensively in a simple manner despite its computational burden. On an off-line investigation, this study affords to use the Monte Carlo method to understand the parameter identifiability of the single-particle battery model. Subsequently, without assuming aging mechanisms into the governing equations, an empirical relationship

between the identifiable aging-related parameters and Ah-throughput will be fit for lifetime estimation in other aging conditions.

6.1 Methods and Results

The experimental setup, applied cells, test profiles and test procedures have been reviewed in chapter 5. Besides the capacity loss data reviewed in chapter 5, the internal resistance rise data are also included below for completeness. For calculation of the inner resistance, the hybrid pulse power characterization (HPPC) profile was executed. It applies a 18s 4C discharge pulse followed by a 40s rest period and 10s 3C charge pulse followed by a 40s rest period for every 10% interval between 90% and 20% SOC. An flowchart of the complete aging test is illustrated in Figure 6-1.

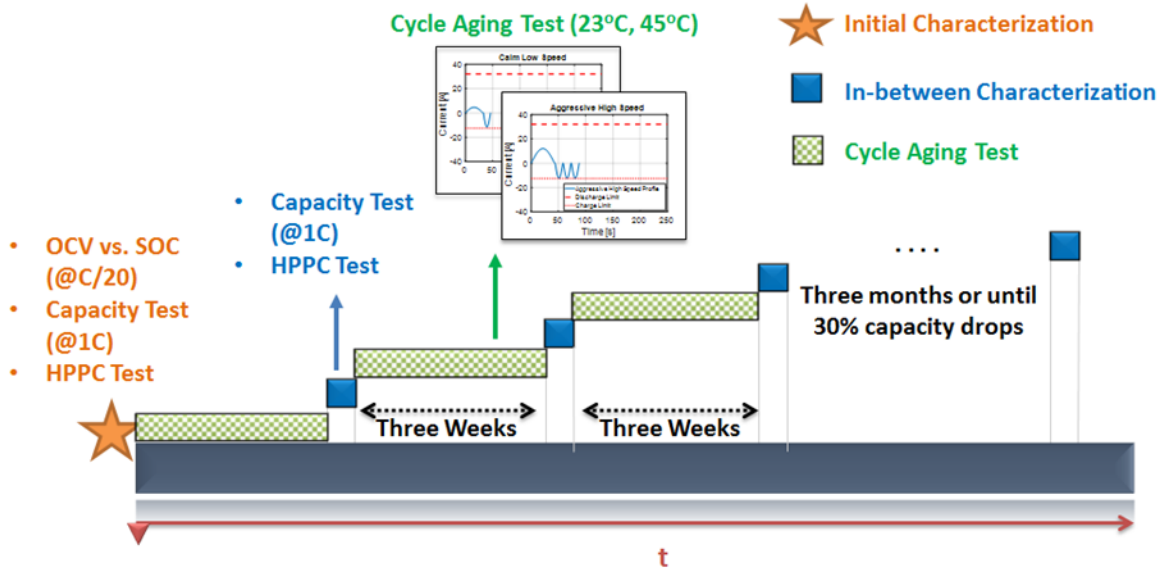


Figure 6-1. The flowchart of the complete aging test

The eight cells for the aging test were numbered and mapped to their corresponding aging conditions in Table 5-5. The initial capacities for the eight cells are plot in Figure

6-2(left), all a bit short of the nominal capacity at 1C. The relationships between open circuit voltage and SOC cells in Figure 6-2(right) are tight among cells.

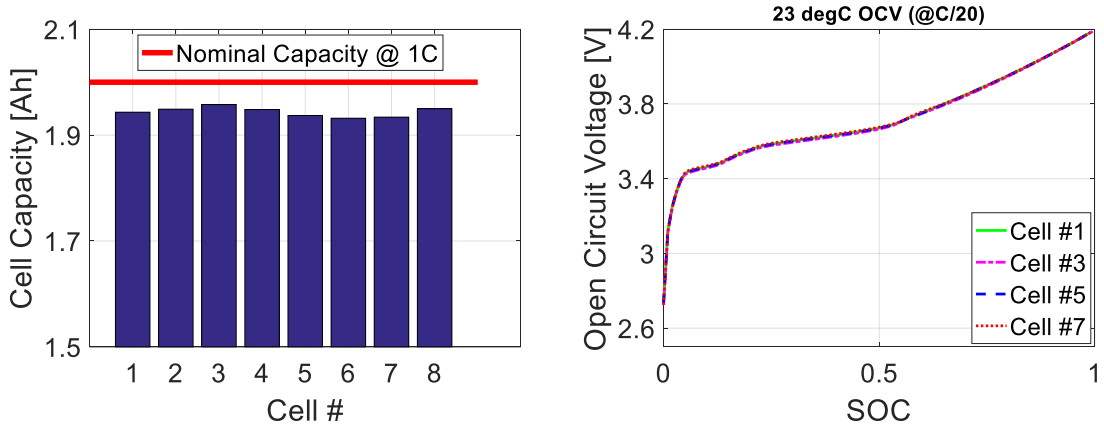


Figure 6-2. Cell initial capacities (left); Cell open circuit voltage vs. SOC (right)

The increase ratios of internal resistance, which is the average 1-second discharge resistance from the HPPC test, can be referred to Figure 6-3. Generally, the degradations of internal resistance are not significant since only the CLS45 conditions reached the 1.1 increase ratio in our aging campaign. In contrast to our intuition, the CLS23 condition overwhelms the AHS23 and AHS45 conditions, and all cells somehow show deviating trends. These may be attributed to the limited resistance degradations so that potential trends are heavily affected by the test uncertainties. In this case, meaningful results on internal resistance increases call for extended aging tests beyond 3 months.

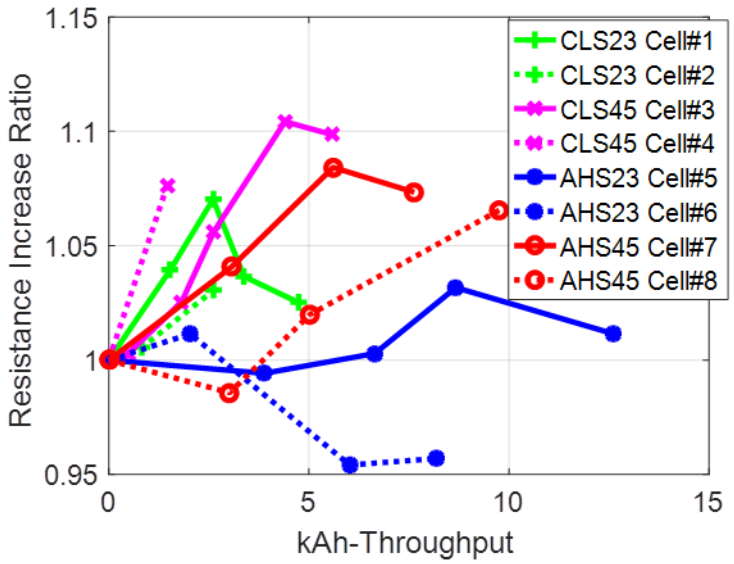


Figure 6-3. The internal resistance increase in the aging test

The normalized capacity losses percentages are compiled in Figure 6-4. The capacity loss severity is in following orders: the AHS45 cases as the severest, the CLS45 case as the second, the AHS23 as the third, the CLS23 as the mildest. Interestingly, except the AHS45, all other conditions witness some capacity recoveries in few beginning characterization stages. The aging results can be interpreted as: 1) the impact from temperature is larger than discharge current magnitude; 2) the impact from discharge current magnitude is larger at high temperature; 3) the impact from temperature is larger at higher discharge current rate. The aging trends derived with the test results emphasize the importance of eco-driving and battery health management.

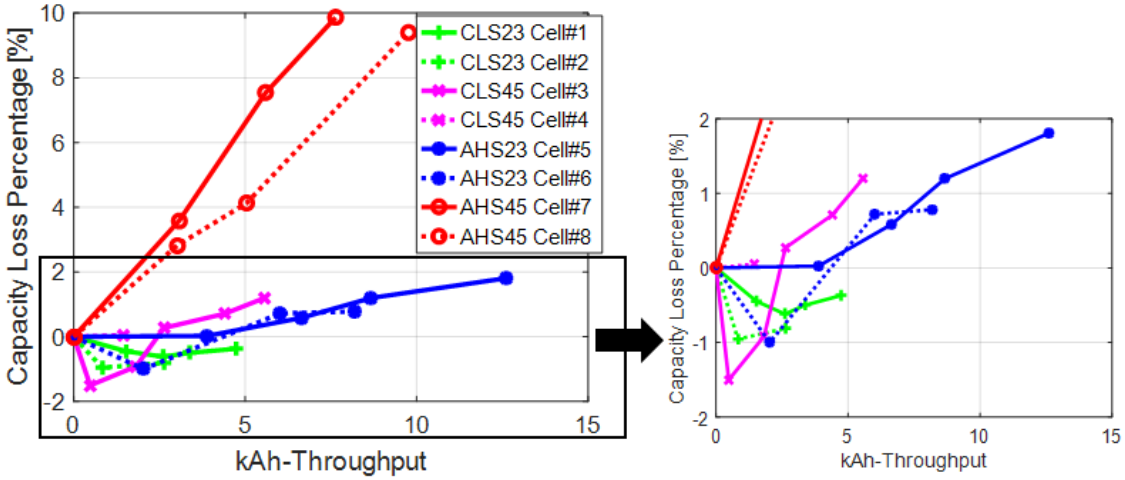


Figure 6-4. The capacity loss in the aging test

6.1.1 Electrochemical Model Parameter Identification

6.1.1.1 Single Particle Model

The single particle model is set up to simulate the cell voltage response. Each electrode is represented by a spherical particle, inside which the lithium ion diffusion follows the Fick's diffusion law as in Eq. 6.1.

$$\frac{\partial c_i}{\partial t} - D_i \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial c}{\partial r}\right) = 0 \qquad \text{Eq. 6.1}$$

Initial and boundary conditions are:

$$D_i \frac{\partial c_i}{\partial r} = 0, \quad at\ r = 0\ and\ for\ t \geq 0 \qquad \text{Eq. 6.2}$$

$$D_i \frac{\partial c_i}{\partial r} = -j_i, \quad at\ r = R_i\ and\ for\ t \geq 0 \qquad \text{Eq. 6.3}$$

Where the subscript $i = n$ for negative electrode, $i = p$ for positive electrode; D_i is the diffusion coefficient of the electrolyte in the solid particle (m^2/s); c_i is the solid-phase

concentration (mol/m^3); R_i is the radius of the solid particle (m); r is the distance from the center of the particle (m); t is the time (s); j_i is the intercalation current density calculated as in Eq. 6.4.

$$j_i = \frac{I}{A_i} = \frac{I}{3/R_i \varepsilon_i V_i} \quad \text{Eq. 6.4}$$

Where I the applied current, positive for discharge is, negative for charge; $A_i$is the equivalent electro-active surface area (m^2); ε_i is the volume fraction of the active material in the electrode; V_i is the volume of the electrode (m^3).

To simplify the Eq. 6.1, a parabolic solid concentration profile in Eq. 6.5 is assumed [104].

$$c_i(r,t) = a(t) + b(t)\frac{r^2}{R_i^2} \quad \text{Eq. 6.5}$$

After insertion, the resultant equations are:

$$\frac{d}{dt}c_{avg,i} + 3\frac{j_i}{R_i} = 0 \quad \text{Eq. 6.6}$$

$$\frac{D_i}{R_i}\left(c_{s,i} - c_{avg,i}\right) = -\frac{j_i}{5} \quad \text{Eq. 6.7}$$

The $c_{s,i}$ is the particle surface lithium ion concentration; the $c_{avg,i}$ is the average lithium ion concentration in the electrode, which can determine the SOC or the stoichiometric number of the electrode in Eq. 6.8 by normalizing on $c_{max,i}$, maximum electrode lithium ion concentration (mol/m^3).

$$SOC_i = \frac{c_{avg,i}}{c_{max,i}} \quad \text{Eq. 6.8}$$

As the SOC_i can be used to infer the electrode open circuit potentials (OCP), $V_{i,ocp}$, it should be further adjusted by the contact resistance voltage drop (R_f), and the over-potentials derived in Eq. 6.9 and Eq. 6.10, which is caused by the resistance of the flow of electrons at the electrode-electrolyte interface.

$$d_{0,i} = F * k_i * \sqrt{c_e * c_{s,i} * (c_{max,i} - c_{s,i})} \quad \text{Eq. 6.9}$$

$$\eta_i = \frac{RT}{0.5F} \text{asinh}(\frac{j_i}{2d_{0,i}}) \quad \text{Eq. 6.10}$$

The cell voltage response output, V_{output}, is expressed in Eq. 6.11.

$$V_{output} = V_{p,ocp} + \eta_p - V_{n,ocp} - \eta_n - R_f * I \quad \text{Eq. 6.11}$$

Where η_i is the electrode surface over-potential; $d_{0,i}$ is the exchange current density (A/m^2); R is the universal gas constant, 8.314 J/kmol; T is the temperature (K); F is the Faraday constant, 96487 (C/mol); k_i is the reaction rate constant (m$^{2.5}$/mol$^{0.5}$/ s); c_e is the solution phase concentration (mol/m^3); 0.5 is the approximate apparent exchange coefficient for lithium ion cell.

In addition to the two electrode OCPs, the single particle model discussed above introduces 14 parameters as in Table 6-1. The initial values are from similar-sized 18650 cells in [95][100] However, the coupling between k_i and c_e in the form of $k_i\sqrt{c_e}$ in Eq.

6.9 indicates that only 13 effective parameters should be identified. In this study c_e is always fixed at 1000 mol/m^3 to allow the estimation of k_i.

Table 6-1. The parameter set with cited values for the discharge capacity test data

Notations	Parameters	Cited Values
R_n	Negative Electrode Equivalent Particle Radius (μm)	5
R_p	Positive Electrode Equivalent Particle Radius (μm)	5
A_n	Negative Electrode Equivalent Active Surface Area (m^2)	2
A_p	Positive Electrode Equivalent Active Surface Area (m^2)	2
D_n	Negative Electrode Solid Phase Diffusion Coefficient, *10^{-14}(m^2/s)	2
D_p	Positive Electrode Solid Phase Diffusion Coefficient, *10^{-14}(m^2/s)	2
k_n	Negative Electrode Surface Reaction Rate Constant, *10^{-11}(m$^{2.5}$/mol$^{0.5}$/ s)	2
k_p	Positive Electrode Surface Reaction Rate Constant, *10^{-11}(m$^{2.5}$/mol$^{0.5}$/ s)	2
$c_{max,n}$	Negative Electrode Maximum Lithium ion Concentration (mol/m^3)	25000
$c_{max,p}$	Positive Electrode Maximum Lithium ion Concentration (mol/m^3)	22000
$SOC_{n,BOD}$	Negative Electrode Beginning-of-Discharge SOC	0.90
$SOC_{p,BOD}$	Positive Electrode Beginning-of-Discharge SOC	0.02
R_f	Cell Contact Resistance (Ω)	0.02
c_e	Solution Phase Concentration (mol/m^3)	1000

Despite the simplicity, the inherent drawback of neglecting solution phase dynamics limits the application of single-particle model in the low current rate scenarios (usually smaller than 1C). In our case, the open circuit voltage measurements (@C/20) and

discharge capacity measurements (@1C) are qualified for the adoption of single-particle model.

The OCPs of the two electrodes are chosen from [105], and validated on the OCV measurements (Cell#7 for example) in Figure 6-5. The three discernible staging phenomena in the OCV measurements are captured well with the selected OCP curves. The electrode OCPs thus are set to be constant and used for 1C discharge capacity measurements as well.

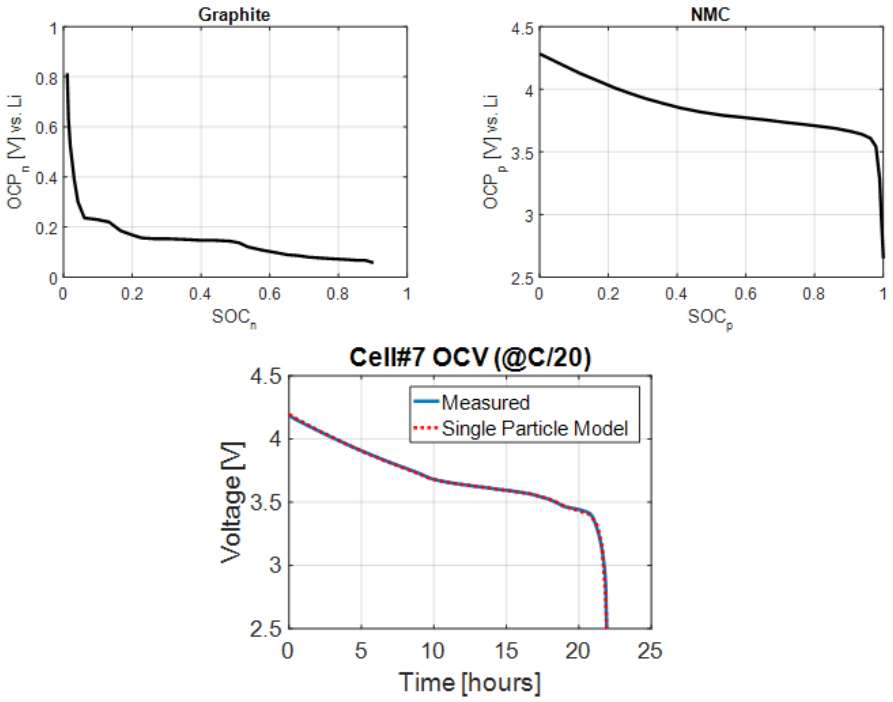


Figure 6-5. Electrode open circuit potential vs. Li and cell open circuit voltage response

6.1.1.2 Aging-Related Parameter Identification

After confirming the electrode OCPs with the fresh cells, we move on to investigate the variations of the parameters of the single-particle model along with the capacity loss.

In other words, the single particle model without including the aging mechanisms in the governing equations is fitted for the discharge capacity test at every characterization stage.

In the context of full discharge capacity test, in previous researches, [100] indicated that $SOC_{n,BOD}$ (negative electrode SOC at the beginning of discharge), $SOC_{p,BOD}$ (positive electrode SOC at the beginning of discharge), A_p (positive electrode equivalent surface area) are the most significant parameters related to aging; [103] evaluated the degradation of D_n (negative electrode solid phase diffusion coefficient), D_p (positive electrode solid phase diffusion coefficient), k_n (negative electrode surface reaction rate constant), k_p (positive electrode surface reaction rate constant) during aging. $SOC_{n,BOD}$, $SOC_{p,BOD}$, D_n, k_n are assumed to account for the SEI form development at the negative electrode; A_p, D_p, k_p are mainly considered to account for active material losses at the positive electrode.

Other non-aging parameters are assumed to be constants and fitted with the OCV data by the *fminsearch* function in MATLAB. The rest of the parameters, aging-related ones, are firstly evaluated on their identifiability with full discharge data in a Bayesian framework.

Table 6-2. The Single-particle model non-aging parameter identification for all cells with OCV measurements

Notations	Cell#1	Cell#2	Cell#3	Cell#4	Cell#5	Cell#6	Cell#7	Cell#8
	Non-aging Parameters							
R_n	5.02	5.03	5.03	5.07	5.02	5.02	5.02	5.04
R_p	5.22	5.22	5.24	5.30	5.24	5.23	5.22	5.29
A_n	2.05	2.05	2.05	1.99	2.05	2.04	2.05	2.06
$c_{max,n}$	23531	23552	23591	23775	23523	23499	23516	23310
$c_{max,p}$	22868	22906	22824	23231	23005	22812	22761	23516
R_f	0.017	0.018	0.018	0.018	0.018	0.018	0.0183	0.018
c_e	1000	1000	1000	1000	1000	1000	1000	1000

Bayesian parameter identification differs from traditional Frequentist counterparts (such as the above *fminsearch* methodology), since it treats any parameter as a distribution. In the domain of Bayesian inference, prior knowledge of parameter values is allowed to be integrated, referred to the Bayes' relation in Eq. 6.12.

$$p(\boldsymbol{\theta}|Data) = \frac{p(Data|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(Data)} = \frac{p(Data|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int p(Data|\boldsymbol{\theta})d\boldsymbol{\theta}} \quad \text{Eq. 6.12}$$

Where $p(\boldsymbol{\theta})$ is the prior distribution, quantifies the prior knowledge of parameter values; $p(Data|\boldsymbol{\theta})$ is the likelihood distribution, quantifies the probability of the observed data under a certain set of parameter values; $p(Data)$ is the normalizing constant, represents the probability of the observed data; $p(\boldsymbol{\theta}|Data)$ is the posterior distribution, quantifies the conditional probability distribution of unknown parameters under the observed data; $\boldsymbol{\theta}$, in the this case, is a vector of parameters (A_p , D_n, D_p, k_n, k_p, $SOC_{n,BOD}$, $SOC_{p,BOD}$).

While the analytical solution of Eq. 6.12 is hard to attain, numerical sampling is desired for approximation of exact distributions. The Metropolis algorithm, which follows a Markov chain Monte Carlo (MCMC) method, is applied in our study according to following steps:

(1) Set the number of sampling iteration N

(2) Determine an initial set of parameters using the $fminsearch$ method

$$\boldsymbol{\theta}^0 = arg\min_{\boldsymbol{\theta}} SE\,; \quad SE = \sum_{j=1}^{n} [\boldsymbol{V}_{model} - \boldsymbol{V}_{exp}(t_j; \boldsymbol{\theta})]^2$$

(3) The error variance of the initial fit:

$$err^2 = \frac{1}{n-p}\sum_{j=1}^{n} SE_{\boldsymbol{\theta}^0}$$

p is the length of the parameter set.

(4) For $i = 1, 2, \dots, N$

Sample $z \sim Normal(0,1)$;

Set $\boldsymbol{R} = 0.01 * \boldsymbol{\theta}^0 * \boldsymbol{I}$;

Generate a new candidate set of parameters in the random-walk manner:

$\boldsymbol{\theta}^* = \boldsymbol{\theta}^{i-1} + \boldsymbol{R}z$;

Compute $SE_{\boldsymbol{\theta}^*} = \sum_{j=1}^{n} [\boldsymbol{V}_{model} - \boldsymbol{V}_{exp}(t_j; \boldsymbol{\theta}^*)]^2$;

Sample $u \sim Uniform[0\ 1]$;

Compute $\alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{i-1}) = min(1, e^{-\frac{[SE_{\boldsymbol{\theta}^*} - SE_{\boldsymbol{\theta}^{i-1}}]}{2err^2}})$;

If $u < \alpha(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{i-1})$

Set $\boldsymbol{\theta}^i = \boldsymbol{\theta}^*, SE_{\boldsymbol{\theta}^i} = SE_{\boldsymbol{\theta}^*}$;

Else

Set $\boldsymbol{\theta}^i = \boldsymbol{\theta}^{i-1}$

End

The basic philosophy of the random-walk metropolis algorithm is to allow a global search in the parameter space towards the best fit on the voltage response. While any set of parameter that generates a better fit must be accepted, the ones with a poorer fit can still be accepted with some probability. After enough iteration, the distribution of each parameter in the set can be numerically approximated.

Starting directly with the cited parameter values in Table 6-3, a MCMC sampling of the six aging-related parameters are executed and visualized in Figure 6-6. The traces of the parameters after 50,000 iterations have different characteristics. While the A_p and $SOC_{n,BOD}$ quickly reach stability, the k_p is the slowest one taking nearly 30,000 iterations. Therefore, the last 20,000 iterations can be used to approximate the parameter distributions, which can be referred to Figure 6-7. From it, both the A_p and $SOC_{n,BOD}$ vary in narrow windows of deviation from their mean values, inside (-0.1%, 0.1%). In contrast, other parameters have rather wide variation windows, thus indicating un-identifiability. The D_p, D_n, k_n, and $SOC_{p,BOD}$ all have uncertainty intervals with width over 50%. In other words, among the six pre-selected aging-related parameters, only A_p and $SOC_{n,BOD}$ can be identified with high confidence under the scenario of full discharge test. Any further aging investigations based on these two parameters should be convincing.

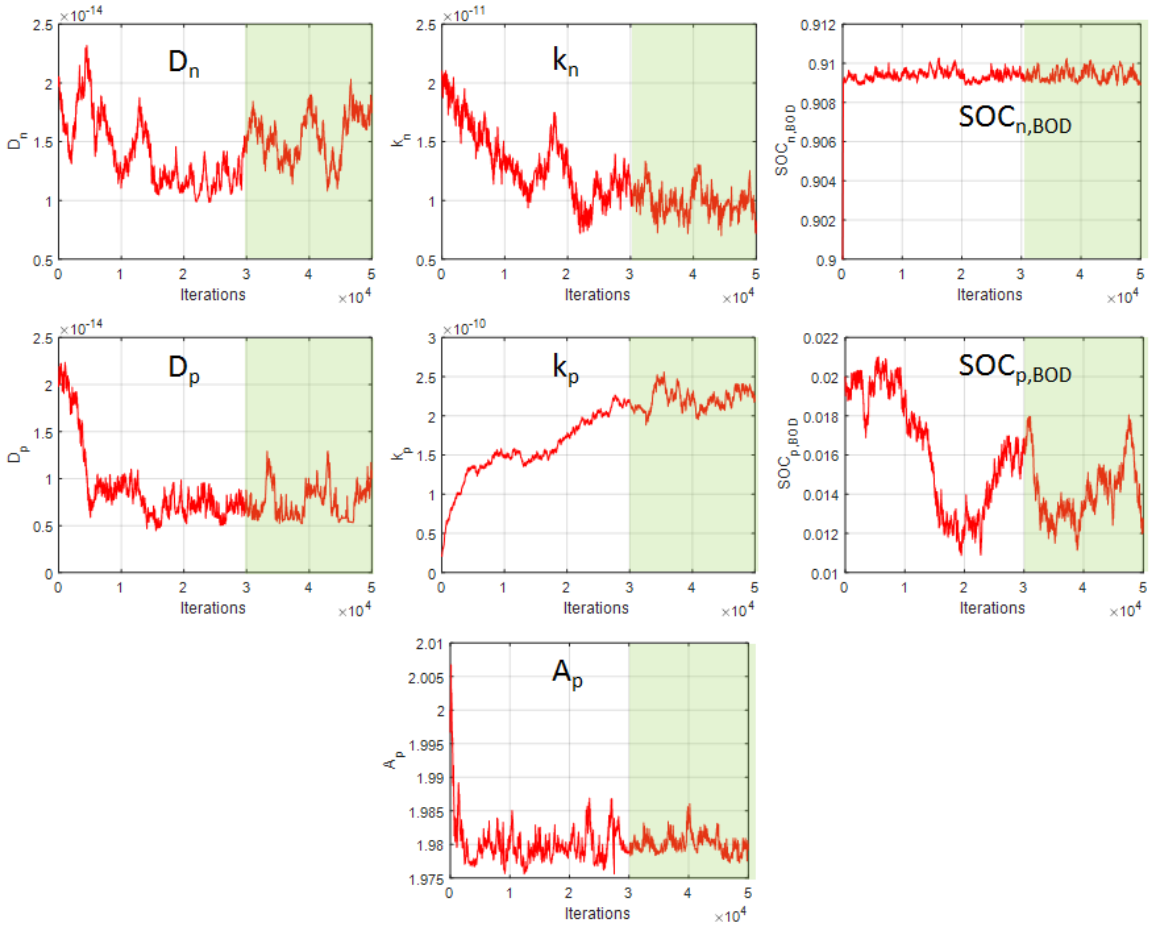


Figure 6-6. The traces of aging-related parameters in the MCMC exploration with 50,000 iterations.

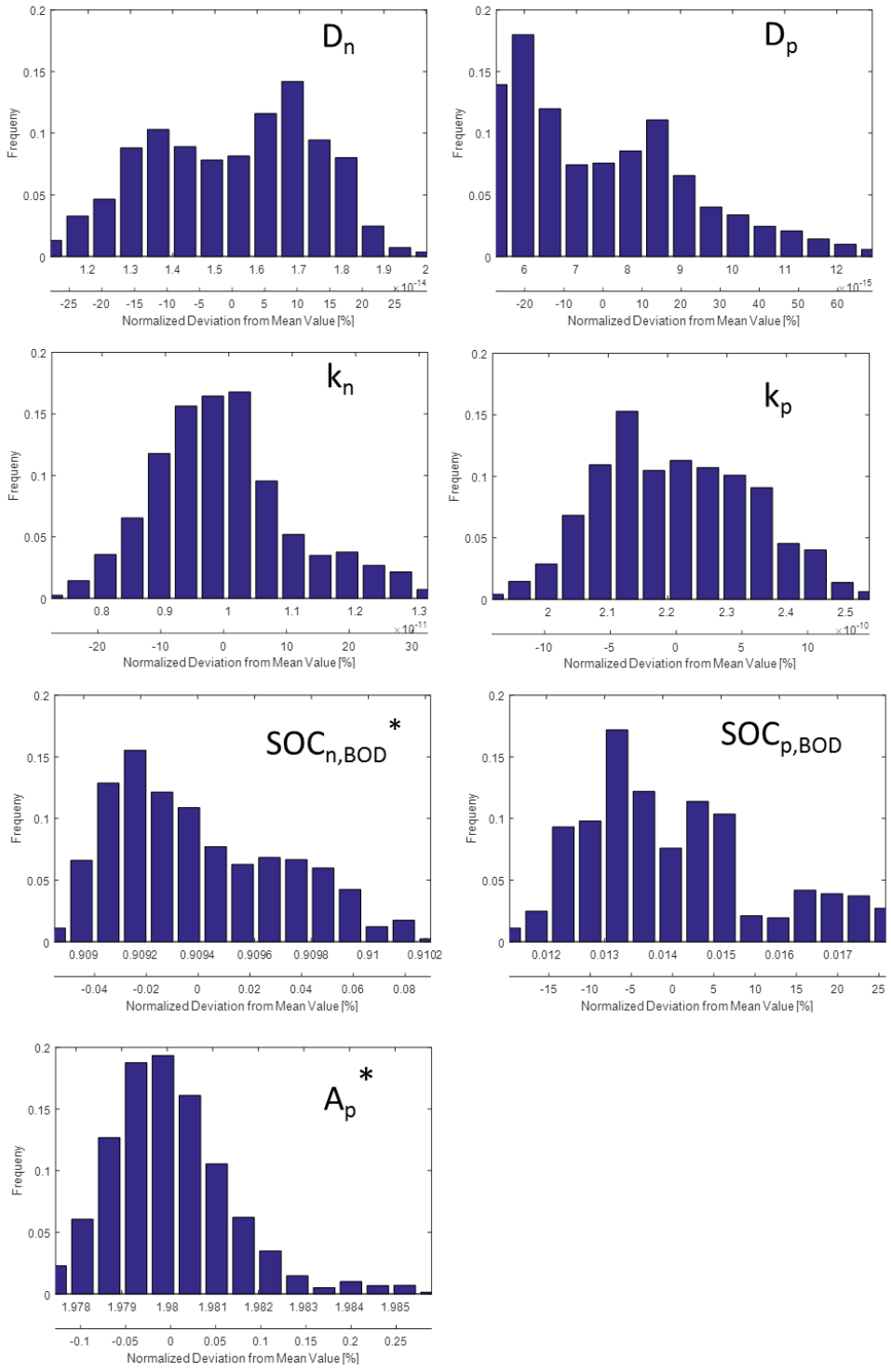


Figure 6-7. The distributions of the aging-related parameters with the last 20,000 iterations of the parameter traces above; (*) highlights the well identified parameters.

To get the values of A_p and $SOC_{n,BOD}$ at different characterization stages, the 1C discharge test are used for fitting, referred to the Cell#8 in Figure 6-8. The discharge curves are well fitted; for four cases, the root mean squared errors between experiment and simulation data are all under 0.02V. The fitted values of these two parameters are compiled in Table 6-3. The A_p, which indicates the positive electrode active material loss and SPI degradation, decreases in all cases. The $SOC_{n,BOD}$, which implies the active lithium ion loss and SEI development, is overall reducing except few characterization stages.

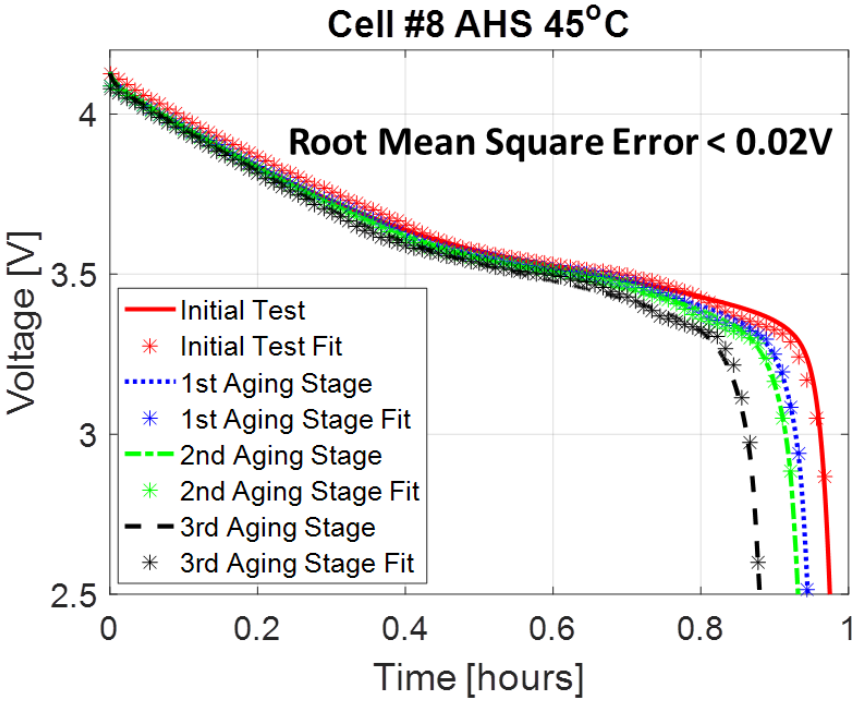


Figure 6-8. The experiment and simulated 1C discharge data with the Cell#8 as example

Table 6-3. The fitted values of A_p and $SOC_{n,BOD}$ for the 1C discharge data at different aging stages.

Aging Parameters	Initial	1st Aging Stage	2nd Aging Stage	3rd Aging Stage	4th Aging Stage	5th Aging Stage
	CLS 23°C Cell#1					
A_p	1.949	1.945	1.948	1.940	1.940	
$SOC_{n,BOD}$	0.906	0.911	0.912	0.908	0.908	
	CLS 23°C Cell#2					
A_p	1.989	1.953	1.948			
$SOC_{n,BOD}$	0.906	0.913	0.912			
	CLS 45°C Cell#3					
A_p	2.001	1.966	1.958	1.935	1.925	1.920
$SOC_{n,BOD}$	0.907	0.918	0.914	0.903	0.899	0.896
	CLS 45°C Cell#4					
A_p	1.986	1.885				
$SOC_{n,BOD}$	0.915	0.915				
	AHS 23°C Cell#5					
A_p	1.978	1.914	1.906	1.887	1.880	
$SOC_{n,BOD}$	0.905	0.905	0.900	0.892	0.887	
	AHS 23°C Cell#6					
A_p	1.967	1.944	1.938	1.956		
$SOC_{n,BOD}$	0.904	0.912	0.896	0.897		
	AHS 45°C Cell#7					
A_p	1.942	1.869	1.798	1.751		
$SOC_{n,BOD}$	0.904	0.873	0.838	0.814		
	AHS 23°C Cell#8					
A_p	1.983	1.847	1.817	1.717		
$SOC_{n,BOD}$	0.910	0.883	0.871	0.823		

Further correlating the loss of the two significant aging parameters with the capacity loss in Figure 6-9, it is found that the loss of $SOC_{n,BOD}$ is strongly positively related to capacity loss (R^2=0.9926), yet the relationship between A_p loss and capacity loss is loose (R^2=0.6757). In the perspective of aging mechanisms, these correlations should mean that

in our aging period (around 3 months), the degradation at the negative electrode (mainly SEI film development) is the most significant cause of capacity loss; the degradation at the positive electrode do exist but affect the capacity loss in a limited manner. Extended aging test beyond 3 months should yield new insights on the shift of aging mechanisms.

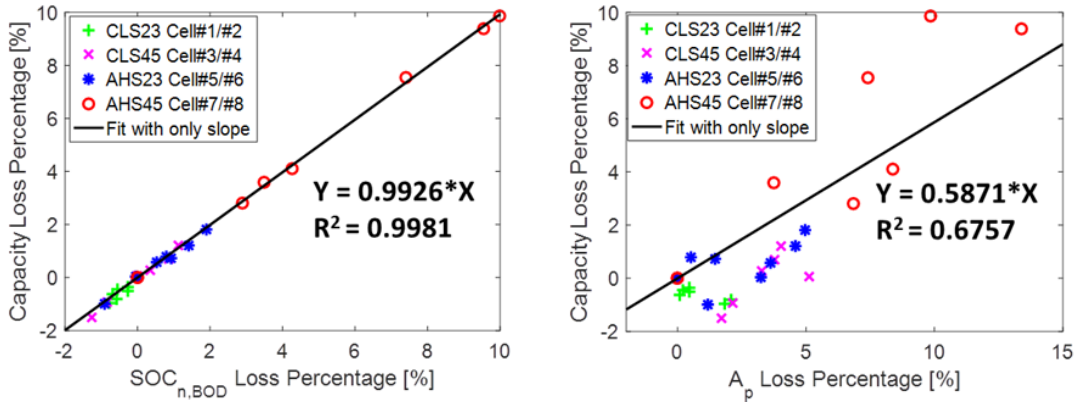


Figure 6-9. The relationships of capacity loss vs. $SOC_{n,BOD}$ loss, capacity loss (left), capacity loss vs. A_p loss (right), for all cells

6.1.2 Battery Aging Estimation in New Test Conditions

After establishing the correlation between significant aging parameters (e.g. $SOC_{n,BOD}$) and capacity loss, it is meaningful to apply it to set up some aging model. New conditions, which might further accelerate the battery aging conditions, can be tested for useful recommendations in real-world vehicle applications.

Two new test profiles, in Figure 6-10, are constructed by adding and deleting one charge micro profile to adapt the current level when the discharge time period is kept unchanged. In this way, new condition #1 and #2 respectively has an increased average current at 16.1A and 5.3A. They are meant to be tested both under the temperature of 45^{o}C for enough aging effects.

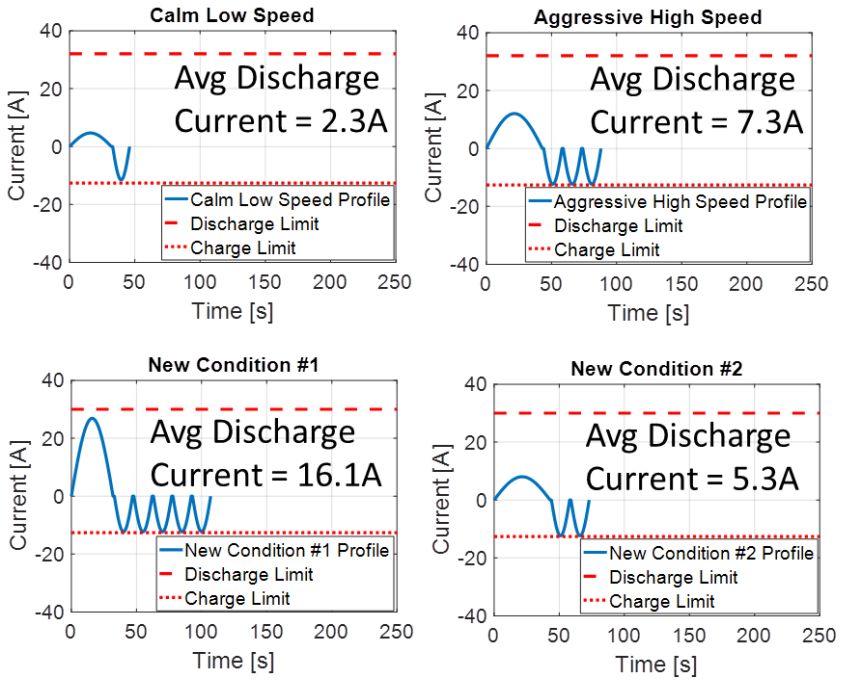


Figure 6-10. Two newly constructed test profiles (bottom)

As the $SOC_{n,BOD}$ is highly correlated with the capacity loss, its evolution with Ah-throughput can be modeled for its updates to predict capacity loss across aging. In Figure 6-11, for both CLS45 and AHS45 cases, a linear relationship between $SOC_{n,SOC}$ loss and Ah-throughput is identified. Since that the $SOC_{n,BOD}$ loss is further linearly correlated to capacity loss, a linear relationship between capacity loss and Ah-throughput is actually assumed in this scenario, which has been reviewed in [16]. Coupling the average discharge current of the profiles with the $SOC_{n,SOC}$ loss evolutions, the predictions for the two new test conditions with uncertainty bounds can be referred to Figure 6-11 as well. Interestingly, an aging characterization result under the new condition #1 falls far below its prediction, on par with the AHS45 case. While the test uncertainty is a potential explanation, the impact of discharge rate on capacity loss is also probably limited beyond

certain level on the NMC chemistry, i.e. the capacity loss might deviate from a linear relationship for high C-rates. Extended and repeated tests are highly recommended.

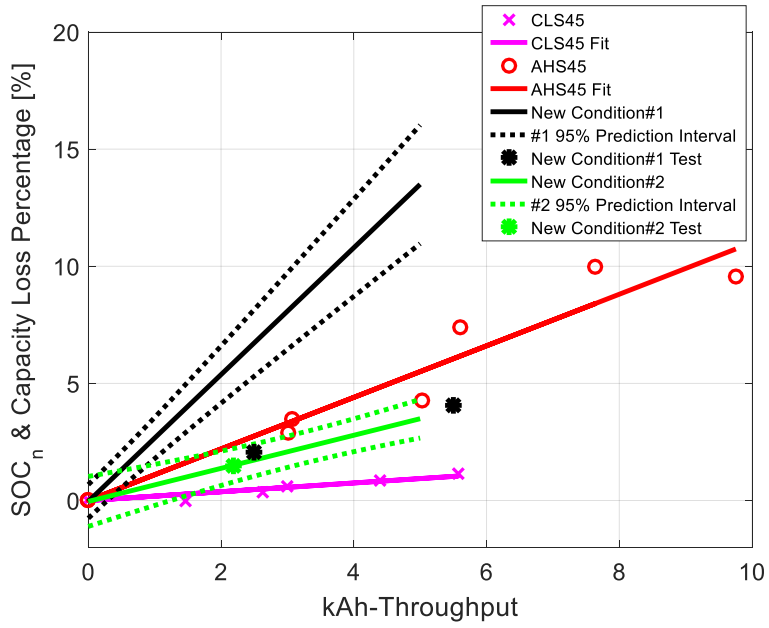


Figure 6-11. SOC_n loss (equivalently capacity loss) vs. Ah-throughput of calm low speed and aggressive high speed 45°C cases; the extrapolated case, the new condition #1 with its 95% prediction bounds; the interpolated case, the new condition #2 with its 95% prediction bounds

6.2 Conclusions

This study tries to supplement the understandings on the cycle aging of 18650 NMC lithium ion batteries. A set of in-house customized test profiles based on real-world driving data analysis were applied under different temperatures. After a 3-month aging campaign, the capacity loss is severer than the internal resistance rise and shows negative impacts from elevated temperature and discharge current rate. As the single-particle is capable to simulate the battery voltage response under low current scenarios, its model parameters are identified with the 1C discharge data at different aging characterization

stages. Out of seven aging-related parameters, two are identified as the significant ones due to their narrow uncertainty interval from a Bayesian analysis. The loss of negative electrode state of charge at the beginning of full discharge ($SOC_{n,BOD}$) is found to strongly linearly correlate with the cell capacity loss, suggesting the SEI surface film development at the negative electrode as the main aging mechanism in the 3-month scope. Built upon above relationship, the $SOC_{n,BOD}$ is further fit against Ah-throughput to facilitate capacity loss estimation with the battery usage history. The derived aging model predicts the battery capacity degradation in two new test conditions with different discharge current rates. While validating the medium-current test condition, the aging model highly overestimates the aging under very high current rate, which might be attributed to either test uncertainty or the inherent feature of NMC chemistry.

7 Summary and Contributions

7.1 Summary

The lithium-ion battery has emerged as an important energy storage system due to high energy and power density for electrified powertrains. However, its aging across usage has always been a problem for automotive applications which require consistent performance throughout life. While manufacturers typically have to oversize the battery pack with increased costs, a promising solution to eliminate this trade-off is a significantly downsized battery pack with minimized impacts on battery life under dynamic battery utilization. This concept requires comprehensive understandings of real-world vehicle operations and the methodologies to incorporate the knowledges in the optimal design and control of battery packs. This dissertation presents a battery aging model derived from aging test results with implications on driver style and road conditions based on real-world driving data.

7.2 Contributions

The main contributions of this dissertation are as follows:

- The methodology for synthesis of representative vehicle drive cycles from a large amount of real-world drive cycles. Previously published work indicates that recorded vehicle speed traces can be can be modeled using a Markov chain. This dissertation extends the approach to consider a higher-order Markov chain, with the goal of improved accuracy. The 3-order Markov chain demonstrated a consistently higher accuracy than the 2-order, and is therefore recommended for future efforts requiring

translation of a large amount of naturalistic driving data into a small number of representative cycles.

- A new algorithm for classifying real-world road conditions based on statistical analysis of a large number of recorded vehicle speed traces. The method relies on micro trips (the speed profiles between consecutive stops) extracted from the real-word drive cycles, and subsequently separated into categories using the k-means clustering. Clustering of micro trips based on duration and mean speed defines four micro trip categories. Finally, a systematic study enabled relating the key features of driving cycles to micro trips, and classification of all vehicle speed profiles in a large database into typical road conditions, e.g. congested urban, uncongested urban, urban-suburban, highway etc.
- An original methodology for developing a metric to quantify driver aggressiveness through analysis of a large amount of real-world speed traces. It is defined in the frequency domain based on the Fast Fourier Transform (FFT) analysis of the jerk traces, which are the second-order derivative of the speed traces. Multiplying the average magnitude of the frequency components by the high-frequency components normalized across the full spectrum leads to the innovative metric. The aggressiveness metric and the ability to synthesize driving cycles representative of selected road conditions and driver aggressiveness are critical for bringing a high degree of reality to battery aging studies.
- A 48V mild hybrid electric vehicle (HEV) was modeled in MATLAB/Simulink to enable relating vehicle drive cycles to battery cell duty cycles. A heuristic supervisory

control strategy includes separation of the power demands between high- and low-frequency , such that the battery meets the high-frequency demands, while the engine responds to low-frequency demand. In addition to generating the battery duty-cycles for subsequent studies of battery aging, the 48V HEV simulation is utilized to assess the sensitivity of the vehicle fuel economy to driver aggressiveness. Results also revealed a strong negative trade-off between battery current loads and driver aggressiveness metric. This emphasizes a need to consider driving style when optimizing the design or control of an HEV.

- An original methodology to synthetize battery current profiles usable for experimental studies of battery aging. The real-world battery current profiles generated by the HEV simulator were characterized in frequency domain to estimate their power spectral density (PSD). Increased driver aggressiveness leads to increased peak frequency and time-series variance of the battery current profiles, implying faster charge/discharge switching and higher C-rates. The information about the peak frequency and time-series variance are used to synthesize representative profiles for experimental testing of cells. The method ensures that the test profiles are charge-neutral and within manufacturer current limits for maximum current and the teat the test profiles are directly relational to anticipated usage in real-world driving scenario.
- An experimental campaign employing the synthesized representative test profiles at two temperature levels, namely 23^{o}C and 45^{o}C. The test was designed to quantify the negative impacts from high temperature and high current rate (driver aggressiveness) combined impact on cell capacity loss.

- Identification of parameters for the electro-chemical, single-particle battery model for the NMC chemistry. Parameters were identified across different aging stages by a stochastic Bayesian analysis with the Monte Carlo Markov Chain algorithm. Results of the simulation study indicate that the loss of one parameter, negative electrode state of charge at the beginning of full discharge ($SOC_{n,BOD}$), strongly correlates with the cell capacity loss. This suggests that the SEI surface film development at the negative electrode is the main aging mechanism.
- A semi-empirical aging model based on the relationship between $SOC_{n,BOD}$ characterized in the identification study, and the cell capacity. Aging model predictiveness was assessed for two validating test conditions corresponding to medium and very high current rates. Excellent agreement was observed for the case with medium current rate, while the model overestimates the capacity loss in the case of very high current rate, thus indicating favorable behavior of the NMC chemistry at high-C rates.

7.3 Future Work

The tools and findings of this dissertation can inspire various future applications:

- The information of the large amount of real-world driving data can also be used for drive cycle prediction for real-time adaptive control studies.
- The quantification of driver style can assist in promoting eco-driving. Real-time feedbacks on driver aggressiveness can be beneficial for adaptive hybrid vehicle power management strategy.

- The methodologies of characterizing and synthesizing real-world cell duty cycles can be unified as a framework and even extended to general analysis of time series data related to vehicle operation.
- A cell aging model with detailed aging functions coupled in the single-particle model can be constructed and compared to the semi-empirical model used in this dissertation in terms of speed and accuracy.
- The battery aging model can be integrated in the battery management system for online estimation of battery health conditions and real-time prediction of battery remaining-useful-life (RUL).

References

[1]. http://www.transportpolicy.net/index.php?title=Global_Comparison:_Light-duty_Fuel_Economy_and_GHG

[2]. Klier, T., & Linn, J. (2016). The effect of vehicle fuel economy standards on technology adoption. Journal of Public Economics, 133, 41-63.

[3]. Budde-Meiwes, H., Drillkens, J., Lunz, B., Muennix, J., Rothgang, S., Kowal, J., & Sauer, D. U. (2013). A review of current automotive battery technology and future prospects. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 227(5), 761-776.

[4]. Dinger, A., Martin, R., Mosquet, X., Rabl, M., Rizoulis, D., Russo, M., & Sticher, G. (2010). Batteries for electric cars: Challenges, opportunities, and the outlook to 2020. Boston Consulting Group.

[5]. Nykvist, B., & Nilsson, M. (2015). Rapidly falling costs of battery packs for electric vehicles. Nature Climate Change, 5(4), 329-332.

[6]. Book, M., Groll, M., Mosquet, X., Rizoulis, D., & Sticher, G. (2009). The comeback of the electric car. How Real, How Soon, and What Must Happen Next.: Studie der Boston Consulting Group Inc.

[7]. Santos, A., McGuckin, N., Nakamoto, H. Y., Gray, D., & Liss, S. (2011). Summary of travel trends: 2009 national household travel survey (No. FHWA-PL-ll-022).

[8]. Transportation Secure Data Center." (2015). National Renewable Energy Laboratory. Accessed January 15, 2015: www.nrel.gov/tsdc.

[9]. Carlson, R., Lohse-Busch, H., Duoba, M., & Shidore, N. (2009). Drive cycle fuel consumption variability of plug-in hybrid electric vehicles due to aggressive driving. SAE Technical Paper, 01-1335.

[10]. Stockar, S., Tulpule, P., Marano, V., & Rizzoni, G. (2009). Energy, economical and environmental analysis of plug-in hybrids electric vehicles based on common driving cycles (No. 2009-24-0062). SAE Technical Paper.

[11]. LeBlanc, David. "Road departure crash warning system field operational test: methodology and results. Volume 1: technical report." (2006).

[12]. Chan, C. C. (2007). The state of the art of electric, hybrid, and fuel cell vehicles. *Proceedings of the IEEE*, *95*(4), 704-718.

[13]. Chan, C. C., Bouscayrol, A., & Chen, K. (2010). Electric, hybrid, and fuel-cell vehicles: Architectures and modeling. *IEEE transactions on vehicular technology*, *59*(2), 589-598.

[14]. Rahn, C. D., & Wang, C. Y. (2013). *Battery systems engineering*. John Wiley & Sons.

[15]. Broussely, M., Biensan, P., Bonhomme, F., Blanchard, P., Herreyre, S., Nechev, K., & Staniewicz, R. J. (2005). Main aging mechanisms in Li ion batteries. Journal of power sources, 146(1), 90-96.

[16]. Vetter, J., Novák, P., Wagner, M. R., Veit, C., Möller, K. C., Besenhard, J. O., ... & Hammouche, A. (2005). Ageing mechanisms in lithium-ion batteries. Journal of power sources, 147(1), 269-281.

[17]. Barré, A., Deguilhem, B., Grolleau, S., Gérard, M., Suard, F., & Riu, D. (2013). A review on lithium-ion battery ageing mechanisms and estimations for automotive applications. Journal of Power Sources, 241, 680-689.

[18]. Liu, Z., Ivanco, A., & Filipi, Z. (In Press). Naturalistic Driving Cycles Synthesis by Markov Chain of Different Orders. *International Journal of Powertrains*

[19]. Liu, Z., Ivanco, A., & Filipi, Z. (2015). Naturalistic drive cycle synthesis for pickup trucks. Journal of safety research, 54, 109-e29.

[20]. Liu, Z., Ivanco, A., & Filipi, Z. (2015). Quantification of drive cycle's rapid speed fluctuations using Fourier analysis. SAE International Journal of Alternative Powertrains, 4(2015-01-1213), 170-177.

[21]. Liu, Z., Ivanco, A., & Filipi, Z. S. (2016). Impacts of Real-World Driving and Driver Aggressiveness on Fuel Consumption of 48V Mild Hybrid Vehicle. SAE International Journal of Alternative Powertrains, 5(2016-01-1166).

[22]. Liu, Z., Onori, S., & Ivanco, A. (2016). Synthesis and Experimental Validation of Battery Aging Test Profiles Based on Real-World Duty Cycles for 48V Mild Hybrid Vehicles.. IEEE Transactions on Vehicular Technology (Conditionally Accepted).

[23]. Liu, Z., Onori, S., & Ivanco, A. (2016). An Aging Model of Nickel-Manganese-Cobalt (NMC) Lithium-ion Batteries Based on Electrochemical Model Parameter Identification. Manuscript submitted for publication

[24]. Patil, R., Adornato, B., & Filipi, Z. (2009). Impact of naturalistic driving patterns on PHEV performance and system design (No. 2009-01-2715). SAE Technical Paper.

[25]. Marshall, B. M., Kelly, J. C., Lee, T. K., Keoleian, G. A., & Filipi, Z. (2013). Environmental assessment of plug-in hybrid electric vehicles using naturalistic drive cycles and vehicle travel patterns: A Michigan case study. Energy Policy,58, 358-370.

[26]. Adornato, B., Patil, R., Filipi, Z., Baraket, Z., & Gordon, T. (2009, September) b). Characterizing naturalistic driving patterns for plug-in hybrid electric vehicle analysis. In Vehicle Power and Propulsion Conference, 2009. VPPC'09. IEEE(pp. 655-660). IEEE.

[27]. Austin, T. C., DeGenova, F. J., Caelson, T. R., Joy, R. W., & Gianolini, K. A. (1993). Characterization of driving patterns and emissions from light-duty vehicles in California (No. ARB-R-949-508).

[28]. Lin, J., & Niemeier, D. A. (2002). An exploratory analysis comparing a stochastic driving cycle to California's regulatory cycle. Atmospheric Environment, 36(38), 5759-5770.

[29]. Lee, T. K., & Filipi, Z. S. (2011). Synthesis of real-world driving cycles using stochastic process and statistical methodology. International Journal of Vehicle Design, 57(1), 17-36.

[30]. Bishop, J. D., Axon, C. J., & McCulloch, M. D. (2012). A robust, data-driven methodology for real-world driving cycle development. Transportation Research Part D: Transport and Environment, 17(5), 389-397.

[31]. Gubner, J. A. (2006). Probability and random processes for electrical and computer engineers. Cambridge University Press.

[32]. Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2013). Time series analysis: forecasting and control. John Wiley & Sons.

[33]. Snedecor, G. W. (1956). Statistical methods, 8th ed., Iowa State University Press.

[34]. Gonder, J., Markel, T., Thornton, M., & Simpson, A. (2007). Using global positioning system travel data to assess real-world energy use of plug-in hybrid electric vehicles. Transportation Research Record: Journal of the Transportation Research Board, (2017), 26-32.

[35]. Smith, K., Earleywine, M., Wood, E., Neubauer, J., & Pesaran, A. (2012). Comparison of plug-in hybrid electric vehicle battery life across geographies and drive cycles (No. 2012-01-0666). SAE technical paper.

[36]. Lee, T., & Filipi, Z. (2011). Representative real-world driving cycles in Midwestern US. Les rencontres scientifiques d'lFP energies nouvelles-RHEVE.

[37]. Lloyd, S. (1982). Least squares quantization in PCM. Information Theory, IEEE Transactions on, 28(2), 129-137.

[38]. Ericsson, Eva. "Independent driving pattern factors and their influence on fuel-use and exhaust emission factors." Transportation Research Part D: Transport and Environment 6, no. 5 (2001): 325-345.

[39]. Rafael, M., M. Sanchez, V. Muciño, and J. Cervantes. "Impact of driving styles on exhaust emissions and fuel economy from a heavy-duty truck: laboratory tests." International Journal of Heavy Vehicle Systems 13, no. 1 (2006): 56-73.

[40]. Murphey, Yi Lu, Robert Milton, and Leonidas Kiliaris. "Driver's style classification using jerk analysis." In Computational Intelligence in Vehicles and Vehicular Systems, 2009. CIVVS'09. IEEE Workshop on, pp. 23-28. IEEE, 2009.

[41]. Langari, Reza, and Jong-Seob Won. "Intelligent energy management agent for a parallel hybrid vehicle-part I: system architecture and design of the driving situation identification process." Vehicular Technology, IEEE Transactions on54, no. 3 (2005): 925-934.

[42]. Kwon, Jay, Jeongmin Kim, Eric Fallas, Sylvain Pagerit, and Aymeric Rousseau.Impact of drive cycles on PHEV component requirements. No. 2008-01-1337. SAE Technical Paper, 2008.

[43]. Barkenbus, Jack N. "Eco-driving: An overlooked climate change initiative."Energy Policy 38, no. 2 (2010): 762-769.

[44]. Transportation Research Board of the National Academies of Science. (2013). “The 2nd Strategic Highway Research Program Naturalistic Driving Study Dataset.” Available from the SHRP 2 NDS InSight Data Dissemination web site: https://insight.shrp2nds.us.

[45]. Duran, Adam, and Matthew Earleywine. “GPS Data Filtration Method for Drive Cycle Analysis Applications.” No. 2012-01-0743. SAE Technical Paper, 2012.

[46]. Prandoni, Paolo, and Martin Vetterli. Signal processing for communications. CRC Press, 2008.

[47]. Cooley, James W., and John W. Tukey. "An algorithm for the machine calculation of complex Fourier series." Mathematics of computation 19, no. 90 (1965): 297-301.

[48]. Rader, Charles M. "Discrete Fourier transforms when the number of data samples is prime." Proceedings of the IEEE 56, no. 6 (1968): 1107-1108.

[49]. Brockwell, Peter J., and Richard A. Davis, eds. Introduction to time series and forecasting. Vol. 1. Taylor & Francis, 2002.

[50]. Tarvainen, Mika P., J-P. Niskanen, J. A. Lipponen, P. O. Ranta-Aho, and P. A. Karjalainen. "Kubios HRV—a software for advanced heart rate variability analysis." In 4th European Conference of the International Federation for Medical and Biological Engineering, pp. 1022-1025. Springer Berlin Heidelberg, 2009.

[51]. Wipke, Keith, Matthew Cuddy, and Steven Burch. "ADVISOR 2.1: A user-friendly advanced powertrain simulation using a combined backward/forward approach." Vehicular Technology, IEEE Transactions on 48, no. 6 (1999): 1751-1761.

[52]. Kuypers, M. (2014). Application of 48 Volt for Mild Hybrid Vehicles and High Power Loads (No. 2014-01-1790). SAE Technical Paper.

[53]. Baldizzone, S. (2012). Performance and Fuel Economy Analysis of a Mild Hybrid Vehicle Equipped with Belt Starter Generator.

[54]. Vallur, A. R., Khairate, Y., & Awate, C. (2015). Prescriptive Modeling, Simulation and Performance Analysis of Mild Hybrid Vehicle and Component Optimization (No. 2015-26-0010). SAE Technical Paper.

[55]. Ivanco, A., & Filipi, Z. (2013). Vehicle modeling and evaluation of the engine options in conventional and mild-hybrid powertrain (No. 2013-01-1449). SAE Technical Paper.

[56]. McGehee, J., & Yoon, H. S. (2013). An optimal powertrain control strategy for a mild hybrid electric vehicle (No. 2013-01-0482). SAE Technical Paper.

[57]. Hsieh, F. C., Chou, T. W., Chen, Y. C., Huang, Y. D., Lin, Y. W., & Peng, Y. W. (2014). Development of Power Management Strategy using Dynamic Programming for BSG Mild HEV (No. 2014-01-1811). SAE Technical Paper.

[58]. Guan, J. C., Chen, B. C., Huang, Y. D., & Chiu, Y. J. (2015, April). Adaptive power management strategy for Hybrid Electric Vehicle with Belt-driven Starter Generator. In Networking, Sensing and Control (ICNSC), 2015 IEEE 12th International Conference on (pp. 503-508). IEEE.

[59]. Shaohua, L., Changqing, D., Fuwu, Y., Jun, W., Zheng, L., & Yuan, L. (2012, November). A rule-based energy management strategy for a new BSG hybrid electric vehicle. In 2012 Third Global Congress on Intelligent Systems (pp. 209-212). IEEE.

[60]. Rick, A., & Sisk, B. (2015). A simulation based analysis of 12V and 48V microhybrid systems across vehicle segments and drive cycles (No. 2015-01-1151). SAE Technical Paper.

[61]. Kim, Y., Lee, T. K., & Filipi, Z. (2012). Frequency domain power distribution strategy for series hybrid electric vehicles. SAE International Journal of Alternative Powertrains, 1(2012-01-1003), 208-218.

[62]. Spotnitz, R., 2003. Simulation of capacity fades in lithium-ion batteries. Journal of Power Sources, 113(1), pp.72-80.

[63]. Kassem, M., Bernard, J., Revel, R., Pelissier, S., Duclaud, F. and Delacourt, C., 2012. Calendar aging of a graphite/LiFePO 4 cell. Journal of Power Sources, 208, pp.296-305.

[64]. Stiaszny, B., Ziegler, J.C., Krauß, E.E., Zhang, M., Schmidt, J.P. and Ivers-Tiffée, E., 2014. Electrochemical characterization and post-mortem analysis of aged LiMn 2 O 4–NMC/graphite lithium ion batteries part II: Calendar aging. Journal of Power Sources, 258, pp.61-75.

[65]. Conte, M., Conte, F.V., Bloom, I.D., Morita, K., Ikeya, T. and Belt, J.R., 2010, November. Ageing testing procedures on lithium batteries in an international collaboration context. In The 25th World Battery, Hybrid and Fuel Cell Electric Vehicle Symposium & Exhibition (pp. 5-9).

[66]. Bloom, I., Cole, B.W., Sohn, J.J., Jones, S.A., Polzin, E.G., Battaglia, V.S., Henriksen, G.L., Motloch, C., Richardson, R., Unkelhaeuser, T. and Ingersoll, D., 2001. An accelerated calendar and cycle life study of Li-ion cells. Journal of Power Sources, 101(2), pp.238-247.

[67]. Jungst, R.G., Nagasubramanian, G., Case, H.L., Liaw, B.Y., Urbina, A., Paez, T.L. and Doughty, D.H., 2003. Accelerated calendar and pulse life analysis of lithium-ion cells. Journal of Power Sources, 119, pp.870-873.

[68]. Thomas, E.V., Case, H.L., Doughty, D.H., Jungst, R.G., Nagasubramanian, G. and Roth, E.P., 2003. Accelerated power degradation of Li-ion cells. Journal of Power Sources, 124(1), pp.254-260.

[69]. Schmalstieg, J., Käbitz, S., Ecker, M. and Sauer, D.U., 2014. A holistic aging model for Li (NiMnCo) O 2 based 18650 lithium-ion batteries. Journal of Power Sources, 257, pp.325-334.

[70]. Käbitz, S., Gerschler, J.B., Ecker, M., Yurdagel, Y., Emmermacher, B., André, D., Mitsch, T. and Sauer, D.U., 2013. Cycle and calendar life study of a graphite| LiNi 1/3 Mn 1/3 Co 1/3 O 2 Li-ion high energy system. Part A: Full cell characterization. Journal of Power Sources, 239, pp.572-583.

[71]. Zhang, Y., Wang, C.Y. and Tang, X., 2011. Cycling degradation of an automotive LiFePO 4 lithium-ion battery. Journal of Power Sources, 196(3), pp.1513-1520.

[72]. Wang, J., Liu, P., Hicks-Garner, J., Sherman, E., Soukiazian, S., Verbrugge, M., Tataria, H., Musser, J. and Finamore, P., 2011. Cycle-life model for graphite-LiFePO 4 cells. Journal of Power Sources, 196(8), pp.3942-3948.

[73]. Stroe, D.I., Świerczyński, M., Stan, A.I., Teodorescu, R. and Andreasen, S.J., 2014. Accelerated lifetime testing methodology for lifetime estimation of Lithium-ion batteries used in augmented wind power plants. IEEE Transactions on Industry Applications, 50(6), pp.4006-4017.

[74]. Safari, M., Morcrette, M., Teyssot, A. and Delacourt, C., 2010. Life prediction methods for lithium-ion batteries derived from a fatigue approach ii. capacity-loss prediction of batteries subjected to complex current profiles. Journal of The Electrochemical Society, 157(7), pp.A892-A898.

[75]. Lohmann, N., Fischnaller, M., Melbert, J., Musch, T., Lamp, P., Scharner, S. and Liebau, V., 2012. Cycle Life Investigations on Different Li-Ion Cell Chemistries for PHEV Applications Based on Real Life Conditions (No. 2012-01-0656). SAE Technical Paper.

[76]. Barré, A., Suard, F., Gérard, M., Montaru, M. and Riu, D., 2014. Statistical analysis for understanding and predicting battery degradations in real-life electric vehicle use. Journal of Power Sources, 245, pp.846-856.

[77]. Spagnol, P., Onori, S., Madella, N., Guezennec, Y. and Neal, J., 2010. Aging and characterization of li-ion batteries in a hev application for lifetime estimation. IFAC Proceedings Volumes, 43(7), pp.186-191.

[78]. Mayer, S.T., 1994. Electric vehicle dynamic-stress-test cycling performance of lithium-ion cells (No. UCRL-ID--116443). Lawrence Livermore National Lab., CA (United States).

[79]. Morita, K., Akai, M. and Hirose, H., 2009. Development of Cycle Life Test Profiles of Lithium-ion Batteries for Plug-in Hybrid Electric Vehicles. Proc. EVS-24, May, pp.13-16.

[80]. Dinger, A., Martin, R., Mosquet, X., Rabl, M., Rizoulis, D., Russo, M. and Sticher, G., 2010. Batteries for electric cars: Challenges, opportunities, and the outlook to 2020. Boston Consulting Group.

[81]. Nitta, N., Wu, F., Lee, J.T. and Yushin, G., 2015. Li-ion battery materials: present and future. Materials today, 18(5), pp.252-264.

[82]. Ecker, M., Nieto, N., Käbitz, S., Schmalstieg, J., Blanke, H., Warnecke, A., & Sauer, D. U. (2014). Calendar and cycle life study of Li (NiMnCo) O 2-based 18650 lithium-ion batteries. Journal of Power Sources, 248, 839-851.

[83]. Schuster, S. F., Bach, T., Fleder, E., Müller, J., Brand, M., Sextl, G., & Jossen, A. (2015). Nonlinear aging characteristics of lithium-ion cells under different operational conditions. Journal of Energy Storage, 1, 44-53.

[84]. Ecker, M., Gerschler, J. B., Vogel, J., Käbitz, S., Hust, F., Dechent, P., & Sauer, D. U. (2012). Development of a lifetime prediction model for lithium-ion batteries based on extended accelerated aging test data. Journal of Power Sources, 215, 248-257.

[85]. Waag, W., Käbitz, S., & Sauer, D. U. (2013). Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application. Applied Energy, 102, 885-897.

[86]. Marongiu, A., Roscher, M., & Sauer, D. U. (2015). Influence of the vehicle-to-grid strategy on the aging behavior of lithium battery electric vehicles. Applied Energy, 137, 899-912.

[87]. Nam, K. W., Yoon, W. S., Shin, H., Chung, K. Y., Choi, S., & Yang, X. Q. (2009). In situ X-ray diffraction studies of mixed LiMn 2 O 4–LiNi 1/3 Co 1/3 Mn 1/3 O 2 composite cathode in Li-ion cells during charge–discharge cycling. Journal of Power Sources, 192(2), 652-659.

[88]. Kozlowski, J. D. (2003, March). Electrochemical cell prognostics using online impedance measurements and model-based data fusion techniques. In Aerospace Conference, 2003. Proceedings. 2003 IEEE (Vol. 7, pp. 3257-3270). IEEE.

[89]. Onori, S., Spagnol, P., Marano, V., Guezennec, Y., & Rizzoni, G. (2012). A new life estimation method for lithium-ion batteries in plug-in hybrid electric vehicles applications. International Journal of Power Electronics, 4(3), 302-319.

[90]. Hu, X., Li, S., & Peng, H. (2012). A comparative study of equivalent circuit models for Li-ion batteries. Journal of Power Sources, 198, 359-367.

[91]. Prasad, G. K., & Rahn, C. D. (2012, October). Development of a first principles equivalent circuit model for a lithium ion battery. In ASME 2012 5th Annual Dynamic Systems and Control Conference joint with the JSME 2012 11th Motion and Vibration Conference (pp. 369-375). American Society of Mechanical Engineers.

[92]. Saha, B., Poll, S., Goebel, K., & Christophersen, J. (2007, September). An integrated approach to battery health monitoring using Bayesian regression and state estimation. In 2007 IEEE Autotestcon (pp. 646-653). IEEE.

[93]. Tröltzsch, U., Kanoun, O., & Tränkler, H. R. (2006). Characterizing aging effects of lithium ion batteries by impedance spectroscopy. Electrochimica Acta, 51(8), 1664-1672.

[94]. Ramadesigan, V., Northrop, P. W., De, S., Santhanagopalan, S., Braatz, R. D., & Subramanian, V. R. (2012). Modeling and simulation of lithium-ion batteries from a systems engineering perspective. Journal of The Electrochemical Society, 159(3), R31-R45.

[95]. Safari, M., Morcrette, M., Teyssot, A., & Delacourt, C. (2009). Multimodal physics-based aging model for life prediction of Li-ion batteries. Journal of The Electrochemical Society, 156(3), A145-A153.

[96]. Lin, X., Park, J., Liu, L., Lee, Y., Sastry, A. M., & Lu, W. (2013). A comprehensive capacity fade model and analysis for Li-ion batteries. Journal of The Electrochemical Society, 160(10), A1701-A1710.

[97]. Ramadass, P., Haran, B., Gomadam, P. M., White, R., & Popov, B. N. (2004). Development of first principles capacity fade model for Li-ion cells. Journal of the Electrochemical Society, 151(2), A196-A203.

[98]. Ning, G., & Popov, B. N. (2004). Cycle life modeling of lithium-ion batteries. Journal of The Electrochemical Society, 151(10), A1584-A1591.

[99]. Safari, M., & Delacourt, C. (2011). Simulation-based analysis of aging phenomena in a commercial graphite/LiFePO4 cell. Journal of The Electrochemical Society, 158(12), A1436-A1447.

[100]. Zhang, Q., & White, R. E. (2008). Capacity fade analysis of a lithium ion cell. Journal of Power Sources, 179(2), 793-798.

[101]. Marcicki, J., Todeschini, F., Onori, S., & Canova, M. (2012, June). Nonlinear parameter estimation for capacity fade in Lithium-ion cells based on a reduced-order electrochemical model. In 2012 American Control Conference (ACC) (pp. 572-577). IEEE.

[102]. Schmidt, A. P., Bitzer, M., Imre, Á. W., & Guzzella, L. (2010). Experiment-driven electrochemical modeling and systematic parameterization for a lithium-ion battery cell. Journal of Power Sources, 195(15), 5071-5080.

[103]. Ramadesigan, V., Chen, K., Burns, N. A., Boovaragavan, V., Braatz, R. D., & Subramanian, V. R. (2011). Parameter estimation and capacity fade analysis of lithium-ion batteries using reformulated models. Journal of The Electrochemical Society, 158(9), A1048-A1054.

[104]. Subramanian, V. R., Diwakar, V. D., & Tapriyal, D. (2005). Efficient macro-micro scale coupled modeling of batteries. Journal of The Electrochemical Society, 152(10), A2002-A2008.

[105]. Chikkannanavar, S. B., Bernardi, D. M., & Liu, L. (2014). A review of blended cathode materials for use in Li-ion batteries. Journal of Power Sources, 248, 91-100.