Model-based Prognostics for Energy-Constrained Mobile Systems Operating in Stochastic Environments

Applied to the Remaining Driving Range Estimation of Electric Vehicles

DISSERTATION

submitted in partial fulfillment
of the requirements for the degree

Doktor Ingenieur
(Doctor of Engineering)

in the

Faculty of Electrical Engineering and Information Technology
at Technische Universität Dortmund

by

M. Sc. Javier Antonio Oliva Alonso

Guatemala, Guatemala

Date of submission: September 9, 2016

First examiner: Univ.-Prof. Dr.-Ing. Prof. h.c. Dr. h.c. Torsten Bertram
Second examiner: Univ.-Prof. Dr.-Ing. habil. Walter Sextro

Date of approval: 24 May, 2017

"To my lovely Sarah and to my blessings who make each of my days."

Acknowledgements

Even though this work is the result of my own motivation, experience and ambition to achieve my scientific goals, the path I drew on my way to these achievements is the result of the individual effort of many persons that I meet during my stay at RST and, without which, not even a word of this work would have ever been written.

I would like to thank Prof. Dr.-Ing. Prof. h.c. Dr. h.c. Torsten Bertram for letting me be part of the RST team during almost five years and for trusting me to lead the TIE-IN project despite not speaking the German language during my early stages at the chair. But most importantly, I thank him for letting me be creative, which allowed me to find the scientific niche within which I have developed this work.

I also thank apl. Prof. Dr. rer. nat. Frank Hoffmann for being the first person who saw in myself the potential to be the researcher I have become after all this time. Also, his deep insights into every single topic related or unrelated to my research have helped me to contemplate and to approach scientific problems with a total different philosophy. Special thanks go to the ex member of the RST team Felipe Posada, who is not only the first Spanish speaking person I have meet in Dortmund, but he became, among others, the tutor of my master thesis and one of my best friends in Germany. I would also like to thank my office colleges Martin Keller and Malte Oejeklaus, who provided me with very helpful feedback both scientific and technical. All other colleges, namely, Christoph Kripmann, Christian Wissing, Christian Götte, Myrel Alsayeh, Christoph Rösmann, Artemi Makarow and Daniel Schauten, who not only were very valuable for my research but also gave me the sense of belonging and allowed to distract my self, deserve my special thanks. Deploying proper scientific solutions goes hand in hand with the successful development of experimental systems. In this sense, I would like to thank Jürgen Limhoff and Rainer Müller-Burtscheid who always were willing to help me either by providing me software solutions for specific needs or by tightening screws together with me. Also Gabriele Rebbe deserves some words of thanks due to she highly alleviated me from organizational paperwork, what allowed me to focus on my scientific oriented priorities. I would also like to thank the battery competence meet and the competence center TIE-IN for allowing me to carry out experiments used to validate the concepts developed in this work.

I thank my parents and my brother for always supporting me and for encouraging me to come to Germany to pursue my goals.

Finally, and most importantly, I would like to thank my wife Sarah for also supporting me during this phase of my career and for holding me even if my thoughts were often devoted to my research more than due.

Abstract

Due to development of novel and more efficient energy storage systems we bear witness to the dawn of a new era of mobile systems. They have become sophisticated in terms of hardware components and software applications which have made it possible to develop integrated solutions for a large number of imaginable applications ranging from electric vehicles all the way to fully autonomous systems operating in a wide variety of ecosystems, e.g., service, surveillance or bio-inspired robots. Generally it is expected that a mobile system exhibits a sufficient degree of autonomy in the sense of energy availability such that it at least accomplishes the mission objectives for which it is intended. Nevertheless, such autonomy, is influenced to a large extent by the remaining energy that can be retrieved from its energy storage system and by the environment conditions in which the system operates. Assessing the reliability of a mission requires using systems internal and external situational awareness to determine if the available energy at least meets the energy needs demanded by the future operation of the mobile system in order to determine its remaining useful life (RUL). Having this information as soon as possible may allow the decision maker to apply a contingency plan to intervene and reconfigure the mission execution strategy in order to improve the probability of success, in those situations in which the system becomes incapable of achieving the original mission objectives.

Numerous studies have been published for assessing mission reliability and estimating the RUL of mobile systems. However, they deal with structured environment conditions and thus with relatively deterministic loads. Moreover, these approaches neglect the inherent uncertainty which stems from multiple sources such as the lack of knowledge about the true energy available in the mobile system, the noise introduced by sensors or the randomness of the operation environment, just to mention a few. The approach presented in this work is built around the belief that the RUL estimation is formulated as an uncertainty propagation problem. Accordingly, to estimate the RUL multiple sources of uncertainty involved in its estimation are first characterized and then propagated with the aim of computing their combined effect, expressed in terms of a probability density function. The approach developed here achieves this estimation in a Monte-Carlo fashion in which several RUL realizations are simulated in order to accurately estimate its entire probability distribution. The aim of this work is therefore devoted to develop a solution capable of estimating the RUL with application to energy-constrained mobile systems operating in stochastic environments.

Contents

Nomenclature

The following list explains all abbreviations and symbols used throughout this work. In general scalar symbols are represented by normal font letters. Vectors are expressed as bold lower case letters, matrices are indicated by bold upper case letters. Functions are represented either by bold lower or by upper case letters followed by $(\cdot)$ or $[\,\cdot\,]$. For the sake of clarity, the nomenclature presented below is divided in Roman, Greek and calligraphic symbols as well as functions, abbreviations and acronyms. Note, that bold symbols appear at the end of the corresponding sections.

Roman Symbols

Δa_{x}	quantization step size of $\boldsymbol{\mathcal{A}}$
Δt	sampling time
ΔT_{ij}	sojourn time
Δv_{x}	quantization step size of $\boldsymbol{\mathcal{V}}$
$\Delta dist_i$	traveled distance at the i^{th} prediction time
ℓ	number of subsets in the training data set
$\mathbb{B}$	Boolean domain
$\mathbb{E}$	expected value
$\mathbb{R}$	real domain
A	frontal area of the electric vehicle
a_1,a_2, a_3, a_4, b_1, b_2	discrete system parameters of the kinetic battery model
a_5, a_6, b_3, b_4	parameters of the discrete-time transfer function
b_λ	exponent of the number of iterations in the recursion of the particle flow
b_{L_β}, b_{U_β}	lower and upper limits that define the Beta distribution
b_L, b_U	lower and upper bounds of an arbitrary uniform distribution
B_λ	number of iterations in the recursion of the particle flow
B_{x}	stiffness factor of the magic formula
c	cell capacity ratio
c_{w}	aerodynamic drag coefficient
C_n	cell nominal capacity
C_{l}	capacitance of the long-time RC-network
C_{s}	capacitance of the short-time RC-network
C_{x}	shape factor off the magic formula
C	simplification parameter
d	virtual conductance of the cell

d_s	divergence between classes/operating conditions
D_{x}	maximum transmittable longitudinal force
D_{N_T}	set of labelled samples
e_i	normal error distribution of an arbitrary class label
E_{x}	curvature factor of the magic formula
$F_{\mathrm{I,Rot}}$	rotational component of the inertial resistance
$F_{\mathrm{I,Tra}}$	translational component of the inertial resistance
$F_{\mathrm{I,Rot_f}}$	front component of the rotational inertial resistance
$F_{\mathrm{I,Rot_r}}$	rear component of the rotational inertial resistance
F_{A}	longitudinal aerodynamic drag force
F_{G}	climbing resistance
F_{I}	inertial resistance
F_{R}	rolling resistance
F_{x}	longitudinal component of the traction force
$F_{\mathrm{x_W}}$	longitudinal component of the traction force of a tire
F_{Rf}	front component of the rolling resistance
F_{Rr}	rear component of the rolling resistance
F_{xf}	longitudinal traction force on the front axle
F_{xr}	longitudinal traction force on the rear axle
F_{zf}	normal load on the front axle
F_{zr}	normal load force on the rear axle
g	acceleration due to gravity
h	smoothing factor
h_1, h_2	height of the first and second well in the kinetic battery model
H_0	Null-hypothesis
H_{G}	height of vehicle center of gravity measured from the ground
i_{G}	gear ratio of the transaxle (driveline)
I_{batt}	battery cell load
$I_{\mathrm{G,in}}$	rotational inertia of input driveline components
$I_{\mathrm{G,out}}$	rotational inertia of output driveline components
I	current
$J_{\xi,\zeta}$	Jacobian of uncertain parameters transformed from original to standard normal space
k	discrete time variable
k_p	discrete time to prediction
$k_{v_{\mathrm{x}}}$	regenerative braking limiting factor
K_0, K_1,...,K_{10}	parameters of the V_{OC}-SoC model
K_λ	normalization factor
$K_{\mathrm{R_0}}$	static rolling resistance coefficient
$K_{\mathrm{R_1}}$	first speed dependent rolling resistance coefficient
$K_{\mathrm{R_2}}$	second speed dependent resistance coefficient
K_{R}	rolling resistance coefficient
l	arbitrary feature
l_{f}	distance from the center of gravity to the front axle
l_{r}	distance from the center of gravity to the rear axle
m	number of environment states, operating conditions or driving

	scenarios
$m_{\mathrm{V,t}}$	total mass of the electric vehicle
n_λ	number of iterations in the recursion of the particle flow
$n_{\mathbb{X}}$	number of samples describing the present state-parameter uncertainty
$n_{\boldsymbol{\mathcal{A}}}$	dimension of the acceleration state space $\boldsymbol{\mathcal{A}}$
$n_{\boldsymbol{\mathcal{V}}}$	dimension of the speed state space $\boldsymbol{\mathcal{V}}$
n_u	number of realizations characterizing the future load uncertainty
n_x	number of samples describing the present state uncertainty
n_{road}	dimension of the road slope state space $\mathcal{F}_{\mathrm{road}}$
n_{ij}	number of times a driving condition changes from state i to j
N_ψ	dimension of the reduced feature space
N_Ψ	dimension of the full set of features
N_θ	dimension of the parameter vector
N_a	length of a subset in the training data set
N_h	dimension of the high dimensional space
N_n	dimension of the measurement noise vector
N_T	dimension of the training set
N_u	dimension of the input vector
N_v	dimension of the process noise vector
N_x	dimension of the state vector
N_y	dimension of the output vector
$N_{\mathbb{W}}$	dimension of the joint process-parameter noise space
$N_{\mathbb{X}}$	dimension of the joint state-parameter space
N_{o}	number of observation of the open circuit voltage
N_{T+}	number of positive classes
N_{T-}	number of negative classes
N_{RA_i}	number of prediction done up to the i^{th} prediction time
N^{Z}	dimension of the vector of uncertain parameters
p	probability
p_f	probability of failure
$P_{\mathrm{M,ele}}$	electrical power demanded by the electric drive/motor
$P_{\mathrm{M,mec}}$	mechanical power demanded by the electric drive/motor
P_{1ry}	power required by an arbitrary primary auxiliary component
P_{2ry}	power required by an arbitrary secondary auxiliary component
P_{aux}	power required by auxiliary components
P_{value}	probability of results equal to observations when H_0 is true
$P\ (t)$	power required at time t
q	random transition between driving conditions
r	realization of the RDR
$r^{(a)}$	residual of the a^{th} iteration in the leave one out cross validation
r_{W}	effective rolling-radius of the tire
R_{f}	brake-force distribution ratio at the front wheels
R_{l}	resistance of the long-time RC-network
R_{ohm}	ohmic resistance

R_{r}	brake-force distribution ratio at the rear wheels
R_{s}	resistance of the short-time RC-network
s_i	arbitrary environment state
t	continuous time variable
t_e	time to end of life
t_p	time to prediction
T_e	depletion indicator
T_{batt}	temperature
T_{M}	torque of the electric drive/motor
T_{W}	traction/brake torque around the rotational axis of a wheel
T	temperature
u_z	input of the discrete-time transfer function
v_l	transient voltage of the long-time RC-network
v_s	transient voltage of the short-time RC-network
v_∞	wind speed in longitudinal direction
v_{W}	translational speed of the wheel
$v_{\mathrm{x,max}}^{\mathrm{reg}}$	maximum activation speed for regenerative braking
$v_{\mathrm{x,min}}^{\mathrm{reg}}$	minimum activation speed for regenerative braking
V_{batt}	cell terminal voltage
V_{OC}	cell open circuit voltage
$V_{\mathrm{o,C}}$	instantaneous voltage rise
$V_{\mathrm{o,D}}$	instantaneous voltage drop
$V_{\mathrm{re,C}}$	exponential voltage rise
$V_{\mathrm{re,D}}$	exponential voltage decay
V	voltage
w_1,w_2	capacity of the first well in the kinetic battery model
w_k^i	associated importance weight of at sample i at discrete time k
y_z	output of the discrete-time transfer function
z	discrete-time variable in the z-domain
Z_i	arbitrary uncertain parameter
$\mathbf{A}_{\gamma_{\mathrm{re}}}$	system matrix
$\mathbf{F}$	vector of cumulative density functions
$\mathbf{I}$	identity matrix
$\mathbf{K}_G$	Gaussian kernel
$\mathbf{K}_{\mathrm{oc}}$	parameter vector of the V_{OC}-SoC model
$\mathbf{M}^S$	transition probability matrix of driving conditions
$\mathbf{M}^{\mathcal{F}_{\mathrm{road},s}}$	transition probability matrix of road slope in an arbitrary driving condition s
$\mathbf{M}^{\mathcal{F}_s}$	transition probability matrix of the driving behaviour in an arbitrary driving condition s
$\mathbf{n}\,(t)$	measurement noise vector at time t
$\mathbf{n}_k$	measurement noise vector at discrete time k
$\mathbf{O}$	set of observation of the open circuit voltage
$\mathbf{r}$	vector of residuals
$\mathbf{R}_{\theta,k}$	parameter noise covariance matrix at discrete-time k
$\mathbf{r}_{\theta,k}$	parameter noise vector at discrete-time k

$\mathbf{R}_{n,k}$	measurement noise covariance matrix at discrete-time k
$\mathbf{R}_{v,k}$	process noise covariance matrix at discrete-time k
$\mathbf{S}$	set of environment states, operating conditions or driving scenarios
$\mathbf{u}(t)$	input vector at time t
$\mathbf{u}_k$	input vector at discrete time k
$\mathbf{v}(t)$	process noise vector at time t
$\mathbf{v}_k$	process noise vector at discrete time k
$\mathbf{w}$	normal vector to the hyperplane
$\mathbf{X}$	stochastic process
$\mathbf{x}(t)$	state vector at time t
$\mathbf{x}_k$	state vector at discrete time k
$\mathbf{y}(t)$	output vector at time t
$\mathbf{y}_k$	output vector at discrete time k
$\mathbf{Y}_k$	set of observations up to discrete time k
$\mathbf{Z}$	vector of uncertain parameters (in original space)
$\mathbf{z}_k$	model error at discrete time k
$\hat{\mathbf{a}}_{\mathrm{x}}$	vector of predicted longitudinal acceleration
$\hat{\mathbf{H}}$	linearised output matrix around the joint state-parameter vector
$\hat{\mathbf{s}}$	realization of predicted driving conditions
$\hat{\mathbf{S}}_{n_{\mathbb{X}}}$	sample covariance matrix of the prior distribution
$\hat{\mathbf{u}}^{\mathrm{env}}$	output of the stochastic environment model
$\hat{\mathbf{v}}_{\mathrm{x}}$	vector of predicted longitudinal speed
$\hat{a}_{\mathrm{x}}$	predicted longitudinal acceleration
$\hat{d}_s$	enhanced divergence between classes/operating conditions
$\hat{s}_i$	arbitrary predicted driving condition
$\hat{v}_{\mathrm{x}}$	predicted longitudinal speed
$\mathbb{W}_k$	joint process-parameter noise vector at discrete-time k
$\mathbb{X}_k$	joint state-parameter vector at discrete time k
$\hat{\mathbb{X}}_k$	best estimation of the augmented joint state-parameter vector

Greek Symbols

α_{s}	significance level
α	shape parameter of the Beta distribution
$\alpha_{\mathcal{L}}$	Lagrange multiplier
$\alpha_{i,j}$	hyper-parameter (virtual count of the transition $i \rightarrow j$)
$\alpha_{t,i}$	importance factor of an arbitrary uncertain parameter
β	shape parameter of the Beta distribution
β_t	first order reliability index
$\beta_{i,j}$	number of times a transition $i \rightarrow j$ occurs
$\Delta\alpha_{\mathrm{road}}$	quantization step size of $\mathcal{F}_{\mathrm{road},s}$
$\Delta\lambda$	step size in the recursion of the particle flow
η_{G}	efficiency of the driveline
η_{M}	efficiency map of the electric drive/motor
γ_{re}	regularization term
κ_{x}	longitudinal tire slip

λ	progression parameter (pseudo-time)
λ_s	variance scale factor in the mean-variance estimator
λ_t	step length
μ_b	expected value of the acceleration during the next time step
μ_s	mean of an arbitrary class
μ_{x}	longitudinal coefficient of friction
∇	gradient
ω_{M}	rotational speed of the electric drive/motor
ω_{W}	rotational speed of the wheel
$\pi_{i,j}$	transition probability from state i to state j
$\pi_{i,j}^{v_{\mathrm{x}}^o a_{\mathrm{x}}}$	transition probability between $a_{\mathrm{x}}^i \rightarrow a_{\mathrm{x}}^j$ given v_{x}^o and a_{x}^i
ρ_{ESS}	arbitrary parameter of the energy storage system model
ρ_{PC}	arbitrary parameter of the power consumption model
ρ_{air}	density of air
σ_b	expected variance of the acceleration during the next time step
σ_s	variance of an arbitrary class
σ_{RBF}	parameter of the radial basis kernel function
$\sigma_{\mathbf{G}_L}$	variance of the linearised limit-state unction
υ_i	binary class label of the i^{th} instance
ε_1, ε_2, ε_3, ε_4, ε_5	parameters of the terminal voltage difference equation
ϖ_i	a-priori probability of occurrence of an arbitrary class
ξ_i	realization of an arbitrary uncertain in standard normal space
ζ_i	realization of an arbitrary uncertain in original space
$\boldsymbol{\alpha}$	vector of hyper-parameters
$\boldsymbol{\alpha}_{\mathcal{L}}$	vector of Lagrange multipliers
$\boldsymbol{\alpha}_t$	vector of importance factors
$\boldsymbol{\Omega}$	Gram matrix
$\boldsymbol{\pi}_i$	i^{th} row vector of the transition probability matrix
$\boldsymbol{\Psi}$	full set of features
$\boldsymbol{\psi}$	subset of extracted features
$\boldsymbol{\theta}(t)$	parameter vector at time t
$\boldsymbol{\theta}_k$	parameter vector at discrete time k
$\boldsymbol{\Xi}$	standard normal space
$\boldsymbol{\xi}$	arbitrary realization of the vector of uncertain parameters in the standard normal space
$\boldsymbol{\xi}^*$	realization of the vector of uncertain parameters at the most probable point
$\boldsymbol{\zeta}$	arbitrary realization of the vector of uncertain parameters in the original space
$\boldsymbol{\upsilon}$	vector of binary class labels
$\hat{\boldsymbol{\alpha}}_{\mathrm{road}}$	vector of predicted road slope
$\hat{\alpha}_{\mathrm{road}}$	predicted road slope
$\hat{\upsilon}^{(a)}$	classification obtained by leaving out the a^{th} subset of the training data set

Calligraphic Symbols

$\mathcal{B}$	bias term
$\mathcal{E}_{a,\min}$	minimum allowable available energy
$\mathcal{E}_a$	available energy
$\mathcal{S}$	set of weighted samples
$\mathcal{U}_{\rho_{\text{PC}}}, \mathcal{U}_{\rho_{\text{ESS}}}$	uniform distribution of an arbitrary parameter of the power consumption and the energy storage system model.
$\mathcal{X}$	joint state-parameter space
$\mathcal{A}_s$	acceleration state space of an arbitrary driving condition s
$\mathcal{F}_s$	two-dimensional state space of an arbitrary driving condition s
$\mathcal{F}_{\text{road},s}$	road slope state space of an arbitrary driving condition s
$\mathcal{M}_{ESS}$	model of the energy storage system
$\mathcal{M}_{PC}$	model of the power required by the mobile system
$\mathcal{Q}$	sequence of transitions between driving conditions
$\mathcal{V}_s$	speed state space of an arbitrary driving condition s

Functions

$\boldsymbol{\mathcal{L}}(\cdot)$	likelihood function
$\Phi[\,\cdot\,]$	standard normal cumulative density function
$\tilde{\mathbf{f}}(\cdot)$	joint state-parameter dynamics function
$\tilde{\mathbf{h}}(\cdot)$	joint output function
$\mathbf{C}_l(\cdot)$	criterion function
$\mathbf{D}(\cdot)$	log-density function describing the prior
$\mathbf{d}(\cdot)$	density function describing the prior
$\mathbf{f}(\cdot)$	state dynamics function
$\mathbf{f}_b(\cdot)$	Beta probability density function
$\mathbf{F}_i(\cdot)$	cumulative density function of the i^{th} feature
$\mathbf{f}_{Z_i}(\cdot)$	probability density function of an uncertain parameter
$\mathbf{F}_{Z_i}(\cdot)$	cumulative density function of an uncertain parameter
$\mathbf{G}(\cdot)$	Remaining Useful Estimation function
$\mathbf{g}(\cdot)$	operating conditions discrimination function
$\mathbf{g}_b(\cdot)$	mean value of the acceleration during the next time step
$\mathbf{h}(\cdot)$	output function
$\mathbf{h}_b(\cdot)$	variance of the acceleration during the next time step
$\mathbf{K}(\cdot)$	kernel function
$\mathbf{L}(\cdot)$	log-density function describing the incremental likelihood
$\mathbf{l}(\cdot)$	density function describing the incremental likelihood
$\mathbf{t}(\cdot)$	threshold function
$\delta(\cdot)$	Dirac delta function
$\boldsymbol{\Gamma}(\cdot)$	log-density function describing the posterior
$\gamma(\cdot)$	density function describing the posterior
$\boldsymbol{\Gamma}_{\text{g}}(\cdot)$	Gamma function
$\boldsymbol{\phi}(\cdot)$	function that maps training data to high dimensional space
$\varrho(\cdot)$	vector field representing the particle flow
$\boldsymbol{\vartheta}(\cdot)$	function mapping $\mathbf{x}$, $\boldsymbol{\theta}$ and $\mathbf{y}$ to $\mathcal{E}_a$.
$\mathbf{G}_{\text{L}}(\cdot)$	linearisation of the limit-state function

Abbreviations and acronyms

AC	**A**lternate **C**urrent
ACC	**Acc**uracy Measure
ANN	**A**rtificial **N**eural **N**etworks
BAC	**B**alance **A**ccuracy Measure
BMS	**B**attery **M**anagement **S**ystem
CDF	**C**umulative **D**ensity **F**unction
CoG	**C**enter **o**f **G**ravity
CRA	**C**umulative **R**elative **A**ccuracy
DAGSVM	**D**irected **A**cyclic **G**raph **S**upport **V**ector **M**achine
DC	**D**irect **C**urrent
DSSM	**D**ynamic **S**tate-**S**pace **M**odel
EODT	**E**nd-**o**f-**D**ischarge **T**ime
EoL	**E**nd **o**f **L**ife
ESS	**E**nergy **S**torage **S**ystem
EV	**E**lectric **V**ehicle
FWD	**F**ront-**W**heel **D**rive
GPS	**G**lobal **P**ositioning **S**ystem
HL-RF	**H**asofer-**L**ind and **R**ackwitz-**F**iessler
HVAC	**H**eating, **V**entilation and **A**ir **C**onditioning
iFORM	**i**nverse-**F**irst **O**rder **R**eliability **M**ethod
IMU	**I**nertial **M**easuring **U**nit
ind	**ind**ex
KiBaM	**K**inetic **B**attery **M**odel
KKT	**K**arush-**K**uhn-**T**ucker
LBC	**L**i-ion **B**attery **C**ontroller
LOO-CV	**L**eave **O**ne **O**ut **C**ross **V**alidation
LPV	**L**inear **P**arameter-**V**arying
LS-SVM	**L**east-**S**quares **S**upport **V**ector **M**achine
MLE	**M**aximum **L**ikelihood **E**stimation
MP	**M**arkov **P**roperty
MPP	**M**ost **P**robable **P**oint
ODE	**O**rdinary **D**ifferential **E**quation
OFCL	**O**uter **F**eedback **C**orrection **L**oop
OID	**O**ptimal **I**mportance **D**ensity
OvA	**O**ner **v**ersus **A**ll
OvO	**O**ner **v**ersus **O**ne
PDF	**P**robability **D**ensity **F**unction
PF	**P**article **F**ilter
PFPF	**P**article **F**low **P**article **F**ilter
RA	**R**elative **A**ccuracy
RBF	**R**adial **B**asis **F**unction
RC	**R**esistance **C**apacitance
RDR	**R**emaining **D**riving **R**ange
RRT	**R**emaining-**R**untime
RTD	**R**emaining **T**ime to **D**ischarge

RtF	**R**un **t**o **F**ailure
RUL	**R**emaining **U**seful **L**ife
SoC_a	**S**tate **o**f **C**harge (available)
SoC_b	**S**tate **o**f **C**harge (bound)
SCM	**S**ample **C**ovariance **M**atrix
SMC	**S**equential **M**onte **C**arlo
SoH	**S**tate **o**f **H**ealth
SVM	**S**upport **V**ector **M**achine
tn	**t**rue **n**egative
tp	**t**rue **p**ositive
UAV	**U**nmanned **A**erial **V**ehicle
UGV	**U**nmanned **G**round **V**ehicle
UKF	**U**nscented **K**alman **F**ilter
UT	**U**nscented **T**ransform

1 Introduction

1.1. Context and motivation

With the growing development of novel and more cost-effective sensors, lightweight actuators, powerful on-board computers and efficient electrical energy storage systems (ESS) with long lifetime, wide temperature range and high energy density, e.g., lithium-ion batteries, the deployment of energy-constrained systems, e.g., mobile systems, in various fields such as commercial, scientific or space applications, have gained enormous importance. Their range of application varies from human operated systems such as conventional plug-in electric vehicles (EVs) (see figure 1.1 (**a**)) to semi or fully autonomous systems such as unmanned ground (UGVs) and aerial vehicles (UAVs) (see figures 1.1 (**b**) and 1.1 (**c**), respectively).

Despite the differences in complexity of hardware components and software architectures among different applications, we can characterize the core structure of a mobile system analogous to that of a mechatronic system (VDI 2206 2004) (refer to figure 1.1 (**d**)). The basic system comprises both the chassis and the powertrain of the mobile system and is composed mainly of mechanical, electromechanical, hydraulic or pneumatic elements, whose main task is to convert the electrical energy obtained from the energy storage into translational or rotational kinetic energy (locomotion). There also may be cases in which a flow of material exist, e.g., in the case of systems driven by internal combustion engines, however, the scope of this research is limited to those mobile systems driven solely by an electric powertrain, hence with lack of material flow. From the point of view of this work, mobile systems are mechatronic systems in which the energy needed to power the basic system, the actuators, the sensors and the information processing unit is supplied by an energy storage system that is integrated together with the rest of the hardware. This renders thereby remarkable energy-constraints and therefore bounds the autonomy of the mobile system.

Regardless of the pursued application, it is generally expected that: 1) a mobile mechatronic system exhibits a high degree of autonomy in the sense of energy availability, i.e., the system should have sufficient energy resources to at least operate until the mission for which it is intended has been successfully accomplished; 2) a relatively short period of time is required to accrue energy from an external source, such that new missions can be conducted; and 3) the electrical energy storage system can be used for a large number of operating cycles before it irreversibly depletes and consequently needs to be replaced (Olivares et al. 2013). These three aspects are in-

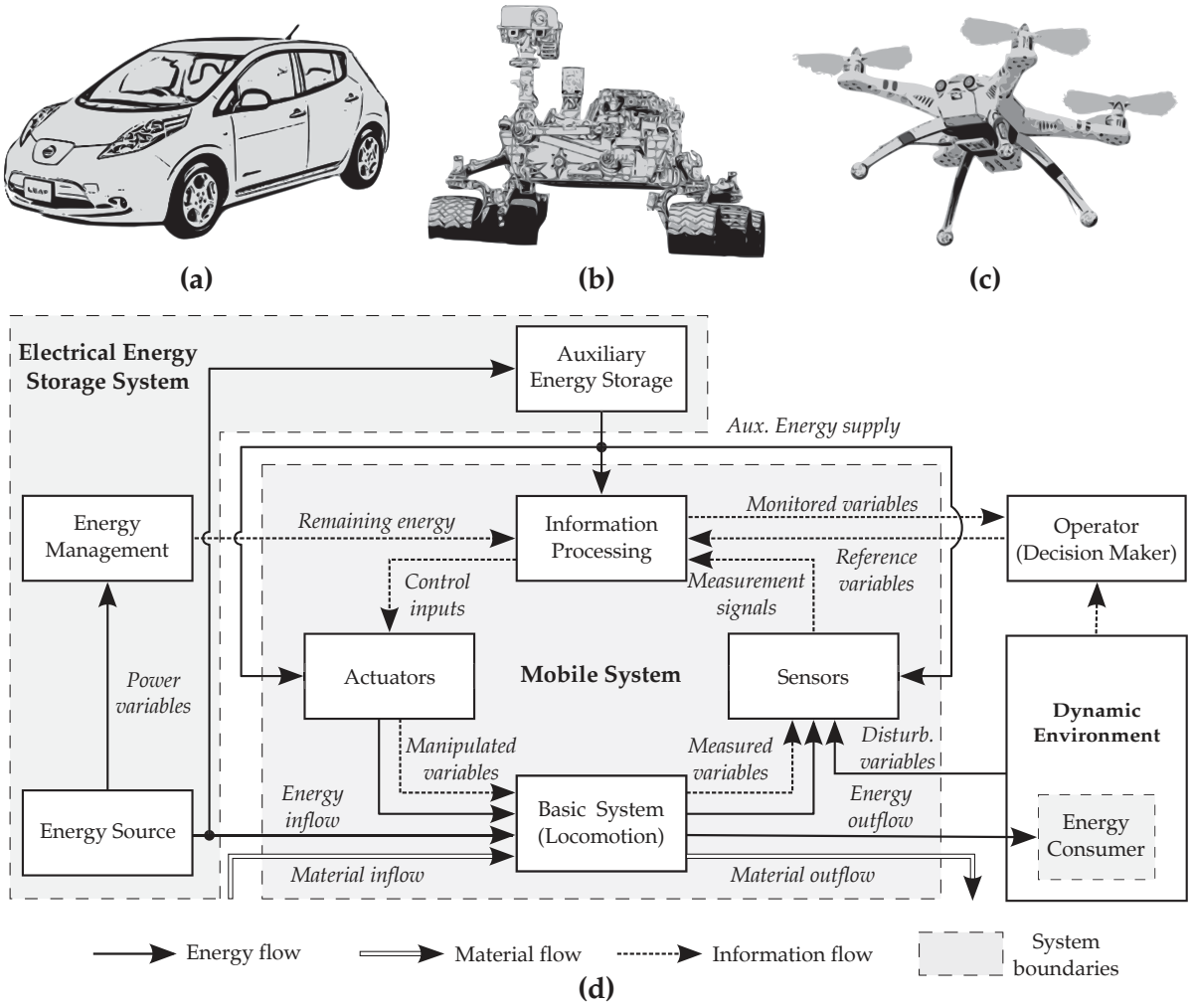

Figure 1.1.: Examples of mobile systems: electric vehicle **(a)**, UGV (**b**) and UAV (**c**). Structure of mobile (mechatronic) system (**d**).

fluenced to a large extent by the degree of depletion of the energy storage system, on the one hand, and by the operating conditions of the mobile system, on the other hand. These two factors added to the fact of on-board energy-constraints constitutes a big challenge for mission planners. For example, mobile systems operating in highly unstructured environments produce stochastic loads, which are difficult to take into account in mission assessment activities.

In most practical applications completing a mission is hardly constrained by the amount of energy that can be retrieved from the energy storage system and is strongly influenced by the dynamic environment in which the system operates. Determining the energy that the energy storage system is able to supply is non a trivial task. In addition, characterizing the influence of the operating environment on the autonomy of the system and on the reliability of a mission can be equally challenging. Because of this it is of paramount importance to have reliable information about the remaining availability of the system, e.g, expressed in terms of time (remaining time to failure) or distance (remaining range). Having unreliable or no information regarding remaining availability of the system may cause undesired consequences. One outstanding example is the *range anxiety* experienced by most owners of electric vehicles. Range

anxiety, i.e., the uncertainty of knowing if the range of the vehicle is insufficient to reach the envisaged destination or at least a charging station, has been pointed out as one of the most important reasons why most drivers are reluctant to acquire an electric vehicle (Rauh, Franke, and Krems 2014). Franke et al. (2012) shows that reliable information about the attainable range may help to alleviate the range anxiety and thereby to increase the acceptance of EVs.

Numerous researchers have developed methods that aim to determine the available energy stored in energy storage systems, e.g., Piller, Perrin, and Jossen (2001). However, relatively little has been expended to date on how to characterize the future energy needs of a mobile system operating in a dynamic environment and specially on how properly incorporate this information into estimating its autonomy.

Prognostics offers solutions to engineering problems with comparable characteristics to the issues exposed before. It aims to predict the time a which a system or component fails to perform a predefined function by comprising current states assessments, previous systems usage and future operating conditions in order to predict failures expressed in terms of the so called remaining useful life (RUL). Interpreting the RUL goes hand in hand with the specific indicator of use of the system under consideration. For example, in prior work prognostics approaches have been developed to forecast remaining operating life in both a macroscopic time scale (cycles) in applications such as batteries (Guo, Li, and Pecht 2015) or fuel cells (Kimotho, Meyer, and Sextro 2014) and in a microscopic time scale (hours or seconds), e.g., bearings (Kimotho and Sextro 2015) or electrolytic capacitors (Celaya et al. 2011). In the context of mobile systems, approaches for fault prognostics have also been developed. LeSage (2013) or Bole, Daigle, and Gorospe (2014) are some of the authors that apply energy-conscious prognostics approaches for estimating the remaining time to discharge (RTD) of energy storage systems powering mobile systems. Oliva, Weihrauch, and Bertram (2013a), for example, extend the definition of the RUL for problems formulated in the space domain and express it in terms of the remaining driving range (RDR). Due to its intrinsic nature, the RUL is a random variable that is influenced by many sources of uncertainty. The lack of knowledge about the system states, the noise presented in measurements or the randomness of the operation environment, are some of the factors that largely contribute to the uncertainty of the RUL. Therefore, properly determining the RUL requires accounting for these sources of uncertainty. The importance of having reliable prognostics approaches that characterize the uncertainty in the RUL has been highlighted by Sankararaman and Goebel (2015). Depending on the application, it might be necessary to employ prognostics algorithms that are capable of adapting the RUL estimation to the most up to date operating conditions. Having this information as soon as possible may allow the decision maker to apply a contingency plan to intervene and reconfigure the mission execution strategy in order to improve the probability of success, in those situations in which the system becomes incapable of achieving the original mission objectives.

This work addresses the still open issue of characterizing and incorporating the uncertainty related to the future operating and environment conditions into the RUL estimation. More precisely, our aim is to develop a solution capable of estimating the RUL of energy-constrained mobile systems operating in highly dynamic environments with application to the remaining driving range estimation of electric vehicles.

1.2. Survey of related work

We first survey previous work describing prognostics approaches developed to estimate the remaining useful life of energy storage systems. Then, we summarize research pertaining the range assessment of energy-constrained mobile systems, paying special attention to the remaining driving range estimation of electric vehicles. Due to the breadth and abundance of literature on topics related to this work, references to these and other topics also appear throughout the dissertation.

Prognostics for energy storage systems

As previously mentioned, the core of prognostics is to estimate the RUL of the system under consideration. Although the literature regarding prognostics has exploded in recent years, there exist no clear consensus among researchers on how to categorize the diverse approaches. This work adopts a categorization scheme which differentiates between approaches according to the sources of information needed for their deployment. Accordingly, we distinct between *data-driven* and *model-based* prognostics approaches. The survey presented here is necessarily incomplete since it covers only those approaches that, to the best of our knowledge, are related to energy storage systems. Moreover, most approaches found in the literature are focused on prognostics solutions under a macroscopic time scale, that is, the pursued RUL estimate reflects the actual remaining lifetime of energy storage systems usually measured in cycles, as it can be seen in the comprehensive survey published by Liao and Kottig (2014). Nevertheless, from the perspective of this work this line of literature is inconvenient since our aim is to assess the remaining operating time of a mobile system within one single cycle given in minutes or maybe hours, i.e., under a microscopic time scale. Hence, the following survey addresses a second line of prognostics approaches focused on estimating the remaining time to discharge (RTD) also referred in the literature to as end-of-discharge time (EoDT) or remaining run-time (RRT).

Data-driven prognostics, also referred to as black-box approach, aims to learn the discharge behavior of the energy storage system on the basis of measurements gathered on the field during operation by converting raw data into relevant pieces of information, from which the energy availability of the system can be inferred. The first line of literature pertaining the prediction of the RTD relies on the basis of the Peukert's law (Peukert 1897). In this sense there exist an empirical static map storing data that explains the relationship between the capacity of the energy storage system and the discharge rate. Pesco et al. (1989) and Matsushima, Ishizuka, and Hashiwaki (1990) introduce relatively simple prediction approaches which take advantage of the linearity presented in some regions of the discharging process. In this case, the time-to-discharge is computed by dividing the available capacity by the load. These approaches, although accurate, have the drawback that the operating conditions are assumed to remain constant during the entire discharging event. To overcome these issues, some researchers have augmented the static maps with the aim of considering time-varying loads (Benini et al. 2003) as well as temperature (Ross and Budney 1995) and aging effects (Pop et al. 2009). Doerffel and Sharkh (2006) demonstrate that, even though the aforementioned approaches show good performance for certain condi-

tions, predictions based on the Peukert's law underperform when the energy storage system is given by Lithium-ion cells and, more important, when the system is subject to loads of stochastic nature, such as in the case of mobile systems.

To circumvent this problematic, purely data-driven methodologies have been published. In this context the literature is divided into two lines of research. The former forecasts the state of charge of the system and the later learns its discharge process. Lu et al. (2010) investigate the applicability of regression analysis, moving average and exponential smoothing to forecast the energy availability of real-time embedded systems and thereby to estimate their RTD. Although this study reports acceptable results, a physical model is still necessary for estimating the energy availability of the system at the time of prediction.

Artificial intelligence approaches overcome this issue by ascertaining the unknown relationship between measurements gathered from the energy storage system and the RTD. Saxena et al. (2012) report from the use of an artificial neural network (ANN) and a least-squared regression approach for estimating the RTD of Li-ion batteries operating under constant operating conditions. These approaches are employed to learn the relationship between different voltage regions of the discharge curve and the RTD. Saxena et al. (2012) show that ANNs are suitable for RTD estimation under constant loads, but that they poorly perform under varying loading conditions. With the aim of accounting for varying loads in the RTD estimation Wen, Wolski, and Krintz (2003) introduce a statistical approach, that takes into consideration varying work loads and battery charge rates. This approach uses a signature extracted from a reference discharge curve. Then, during operation, a statistical model is used to predict deviations from the reference signature and thereby to estimate the RTD.

Personalized or context-aware approaches have also acknowledge the importance of adapting the RTD estimation to varying operating conditions. Kang, Seo, and Hong (2011) propose an approach that assigns possible operating states of a mobile device based on usage patterns from different participants and develop a method that estimate its RTD based on the average load consumed at each state and on the time the mobile device remains on it.

Model-based prognostics, also referred to as white-box approach, has gained in importance during the last decade due to its versatility and ease of implementation in practical engineering applications. In the case of energy storage systems, it involves describing the discharge process by means of a mathematical model, which is generally derived from first principle laws, e.g., chemical, electrical or thermal. The model is employed to predict the evolution of the system capacity and to determine the time at which the system runs out of energy. Approaches falling into this category also benefit from experimental field data since this can be used to identify model parameters and to infer the available capacity.

From the methodologies available in the literature, a model-based approach using particle filters (PF), first developed by Orchard and Vachtsevanos (2009) and later adopted and further developed in the context of energy storage systems by Saha and Goebel (2009), Orchard et al. (2012), Olivares et al. (2013), Zhou et al. (2014) and Tampier et al. (2015), among many other researchers, has emerged as a solid solution. Particle-filtering based approaches for prognostics employ a physics-based model in order to estimate the end-of-discharge time of the system. To this aim a

set of discrete weighted samples, known as particles, is used to first estimate, either directly or indirectly, the available capacity of the system and then to characterize a probability distribution of the RTD by propagating the set of particles forward in time through simulation until the minimum allowable capacity is reached. The computational complexity of this approach is a function of the number of particles used in the capacity estimation and of the time each particle needs to simulate the RTD. It is clear that enhancing the computational performance of this approach requires reducing the number of particles used without sacrificing the accuracy of the RTD estimation.

Daigle, Saxena, and Goebel (2012) aim to solve this issue by adopting a model-based approach (Daigle and Goebel 2010) working on the basis of the Unscented Transform (UT) (Julier and Uhlmann 2004) in which the samples are chosen deterministically instead of using a random sampling method, as in the case of the PF. Although this method is more computationally efficient than the standard PF, the UT may only be applied to nonlinear systems in which all sources of noise are Gaussian, otherwise, this approach should not be used. Oliva and Bertram (2014b) introduce an efficient alternative to the aforementioned model-based approaches and propose the use of a variation of the particle filter based on particle flow and optimal transport methods (Daum and Huang 2008). The idea behind this approach is to reduce the number of particles needed by the PF by introducing a particle flow, in which the particles are progressively transported without needing to randomly sample from any distribution, hence, reducing the number of particles needed and thereby the computational effort in both the capacity estimation and the RTD prediction.

Sankararaman et al. (2013) propose a third line of model-based approaches based on reliability theory with the aim of further reducing the computational complexity. The main contribution of this study is that the RTD is estimated in an analytical manner, eliminating in this manner the randomness introduced by the sampling-based approaches discussed so far. One benefit of this approach is that the significance of each uncertain factor related to the RTD is directed assessed during its estimation, allowing in this manner to perform, for example, uncertainty analysis with no extra computational requirements (Oliva and Bertram 2015).

Range assessment of energy-constrained mobile systems

At first sight it might seem that the task of range assessment can be directly performed with help of the RTD estimation approaches presented above, however, assessing the remaining range of a mobile system requires modeling not only the discharge process of its energy storage system but it also needs information on how the load is generated through the interaction of the mobile system and its surroundings.

Saha et al. (2011) introduce a model-based approach for estimating the end-of-discharge time of an electric UAV using the particle filter framework previously developed by Saha and Goebel (2009). Here, the particle filter employs a cell model to estimate the state of charge (SoC) of a Lithium-ion battery used to power the system. The power requirements of the UAV are characterized from historical data and are clustered based on the flying maneuver they represent, i.e., takeoff, landing, turns or cruise flight. The RTD is estimated by propagating all particles of the filter forward in time until they reach the cut-off voltage of the cell. Due to the flexibility and significance of the experimental system deployed within this research, many new prog-

nostics solutions have been developed on top of it. Quach et al. (2013), for example, create an enhanced prognostics algorithm which is implemented with an improved cell model and replace the particle filter with an unscented Kalman filter (UKF) for state estimation. Furthermore, in contrast to Saha et al. (2011), they characterize the power requirements by means of uniform distributions ranging around mean loads inferred from historical maneuvers. In this manner the prediction becomes more robust against unexpected flying events since it accounts for the uncertainty introduced by the future loading of the UAV. Note, that the accuracy of the approaches discussed so far strongly depends on the characterization of the future flying load. The drawback with the previous characterization schemes is that an RTD estimation can be performed only if the UAV follows a predefined flying plan. These approaches might be not suitable for situations in which the UAV needs to explore unstructured environments, in which the number of characterized maneuvers might be remarkable increased. Instead of augmenting the *pool* of characterized power requirements for every imaginable flying maneuver, Bole et al. (2013) introduce a detailed UAV model which maps the relationship between the aircraft angle of climb, speed, and acceleration and the power requirements of the powertrain. In this manner the RTD estimation reduces to predicting the trajectory of the UAV instead of estimating its future power requirements. Although this approach is more computational expensive, it allows capturing the relationship between the mobile system and its operating environment, which positively impacts the accuracy of the RTD estimation.

Planetary rovers and mobile robots (LeSage 2013; Sadrpour, Jin, and Ulsoy 2013) are also found in the literature as case study for the deployment of RTD estimation solutions. A widespread experimental system is the mobile robot testbed for prognostics-enabled autonomous decision making constructed by Lachat et al. (2006), to carry out experiments in the Antarctic, and further developed by Balaban et al. (2011). Daigle and Sankararaman (2013) present a simulation study in which two model-based prognostics approaches (Daigle, Saxena, and Goebel 2012; Saha and Goebel 2009) are compared for RTD estimation in both constant- and variable-amplitude loading scenarios. The main contribution of this study lies on the uncertainty characterization of the future power requirements of the rover. Such characterization is achieved by using random surrogate variables to parametrize future rover trajectories. Analogous to the work of Bole et al. (2013), the power requirements are computed through a hypothesized future interaction between the rover and the terrain.

Daigle, Sankararaman, and Kulkarni (2015) and later Daigle and Sankararaman (2016) apply the concepts developed within this research in an experimental study in which they evaluate the suitability of model-based approaches working on the basis of Monte Carlo sampling (Saha and Goebel 2009), the Unscented transform (Daigle, Saxena, and Goebel 2012) and reliability theory (Sankararaman et al. 2013) for estimating the end-of-discharge time and the remaining driving range of a planetary rover operating under real conditions with the aim of assisting operational decision-making (Balaban et al. 2013). The evaluations take places under a structured driving scenario, in which the rover navigates following a set of predefined waypoints, and under an unstructured scenario, where the rover executes a sequence of random maneuvers. Both works show the benefits of having reliable RTD and RDR information and highlight their importance for increasing mission feasibility.

Remaining driving range estimation of electric vehicles

Analogous to the prognostics approaches previously surveyed, the categorization of the research regarding the RDR estimation of EVs is based on the sources of information employed, that is, telematics, crowd-sourcing and on-board sensors.

With the growing development of web services and mobile applications, modern electric vehicles are supplied with more precise information about their driving environment. This has allowed researchers to take advantage of telematics to incorporate environmental information such as topology, traffic or even the weather along the road ahead into the RDR estimation. Conradi, Bouteiller, and Hanßen (2011) introduce a method that combines the use of a web server, a digital map and a mobile application. A mobile device sends the position of the vehicle and the current state of charge of the battery to a web server, which first computes the energy consumption along all possible routes and then, based on the current SoC, it calculates the maximum driving range. The achievable RDR is presented to the driver by means of a polygon drawn around the current position of the EV. This idea doesn't remain unheeded by the scientific community and is further developed with minor variants by many other researchers, either by using regression approaches (Ferreira, Monteiro, and Afonso 2011) or by employing detailed models of the EV (Bedogni et al. 2014; Tannahill, Muttaqi, and Sutanto 2016; Zhang et al. 2012), applied to forecast power consumption. Ondruska and Posner (2014a, 2014b) extend this idea and develop a regression approach which generate attainability maps that show the driver the achievable RDR for a given probability of success.

A second line of literature benefits from modern cloud technologies to extend the approach from Conradi, Bouteiller, and Hanßen (2011). Grubwinkler, Brunner, and Lienkamp (2014) use crowd-sourcing to gather 200 000 km of real world driving data and develop a RDR estimation approach based on data mining used to extract statistical features from speed profiles and to forecast power consumption. Ferreira, Monteiro, and Afonso (2012, 2013) also mine data in order to estimate the RDR. This solution extracts information not only related to the driving profile but also regarding the climatic conditions, type of EV and driver behaviors. Other researchers take this approach a step further by augmenting data repositories with information pertaining routes and terrain, vehicle and battery parameters as well as charging history in order to create Big Data frameworks in which the RDR is estimated by means of machine learning (Lee and Wu 2015; Rahimi-Eichi et al. 2015; Rahimi-Eichi and Chow 2014).

Although RDR estimation solutions based on telematics and crowd-sourcing have achieved acceptable results, their deployment is usually linked with high implementation costs. This problematic has served as a motivation for finding more cost effective solutions, which are able to estimate the RDR based solely on sensors installed on-board of the EV. Ceraolo and Pede (2001) introduce a RDR estimation approach which works in two phases. In the first phase the state of charge of the battery is recursively estimated with help of sensor data. Then, in a second phase, the behavior of the electric vehicle as response to a given driving cycle is forecast forward in time through simulation. The RDR is computed by determining the distance traveled, along the driving cycle, at which the minimum allowable SoC is reached. This work has served as basis for many researchers which aim to estimate the RDR under predetermined

standard driving cycles, either using physics-based (Denis et al. 2012; Grewal and Darnell 2013; Liu et al. 2015; Vaz et al. 2015) or data-driven (Bolovinou et al. 2014; Hong, Park, and Chang 2016) models, for computing future power requirements.

The drawback of the approaches discussed so far, with exception of Ondruska and Posner (2014a, 2014b), is that the RDR is treated as a deterministic quantity and that the various sources of uncertainty are not taken into account. Jung et al. (2015) conclude that displaying the uncertainty related to the RDR estimation improves the driving experience towards the EV and reduces the range anxiety. Oliva, Weihrauch, and Bertram (2013a, 2013b) propose approaches based on the prognostics concepts developed by Daigle, Saxena, and Goebel (2012) and Saha and Goebel (2009), respectively, with the aim of including uncertainty information into the RDR estimation. Similarly to Ceraolo and Pede (2001), these approaches proceed in a state estimation and a RDR prediction phase. In the first phase, the SoC of the battery is recursively estimated. In the second phase the RDR is predicted in two steps. First, multiple realizations of the future driving profile given by the speed, acceleration and by the slope of the road are predicted by means of Markov chains. In the second step, the uncertainty represented by the SoC estimate is propagated through the predicted driving profiles until the minimum SoC is reached. The main contribution of this work is that the uncertainty of the estimated RDR is computed at no extra computational cost.

1.3. Research objectives and thesis outline

The survey of related work suggests that prognostics is still a non mature domain that continuously evolves to attain established concepts that can be applied for energy-constrained mobile systems. Indeed, the development of prognostics solutions relies on the type and quality of data and information regarding the past, present and future of the system and its operating environment, the assumptions made about the system and the validity of the models used to described the system behavior.

To select a proper prognostics solution for this work the characteristics of the different approaches are juxtaposed with the requirements imposed by energy-constrained mobile systems and accordingly assessed upon their data/information requirements as well as their characteristics. For the sake of clearness, the approaches are evaluated in the interval $[1, 5]$. In figure 1.2 (**a**) high values mean high priority of having the corresponding data or information, that is, the more area is covered by an approach, the more difficult and costly is its implementation. On the contrary, in figure 1.2 (**b**) high values mean better characteristics of a given approach. This means that the more area is covered by an approach the more suitable is it for our purpose.

A model-based approach seems to be the most suitable solution for assessing the range of energy-constrained mobile systems. Although it requires more information than its data-driven counterpart, it offers more benefits that might result in a better range assessment. For example, having a model that accurately describes the depletion of the energy storage system is mandatory. Since the literature related to its modeling is relatively mature, the drawback regarding this aspect losses its relevance in the context of this work. Other requirements like the failure threshold or the system intrinsic characteristics are easily obtainable by means of experiments or data sheets.

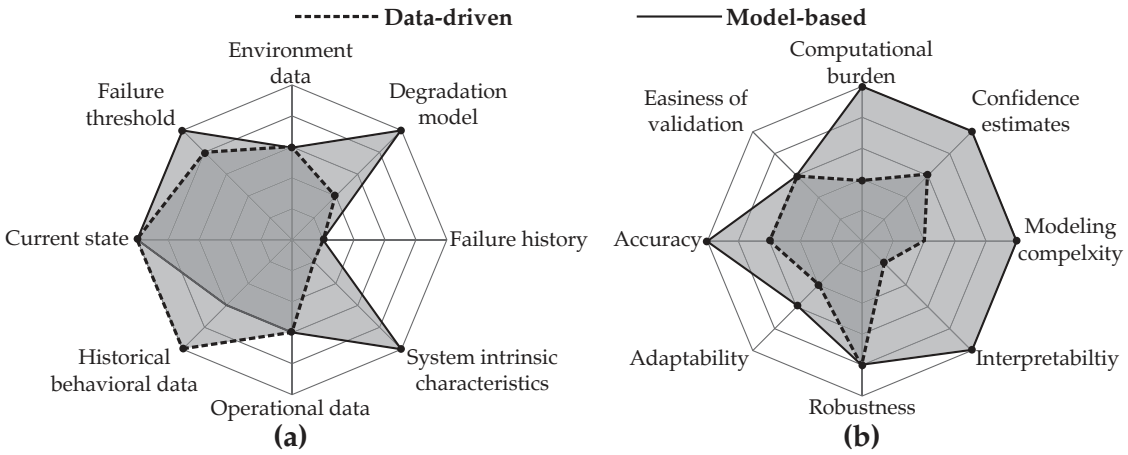

Figure 1.2.: Evaluation of data-driven and model-based prognostics approaches according to the information they require (**a**) and to their characteristics (**b**).

The current state is another piece of information whose availability strongly influences the true assessment of the range. In this work the current state of the system refers to the actual energy that can be used by the mobile system in order to traverse its environment. Having wrong information about this clearly affects the range assessment in a negative manner. A strong benefit of using a model-based approach is that a higher accuracy in the range estimation can be achieved in contrast to the data-driven counterpart, however, this is linked to an increment in the computational burden and in the model complexity. Another advantage of using a model-based solution is that most aspects of the prognostic procedure are interpretable. This allows to gain a better understanding on how the range is actually estimated and, if needed, to carry out changes more quickly. Finally, one of the most important benefits of using models is that it permits not only estimating the range in a deterministic manner, but it allows assessing the uncertainty related to it expressed in terms of confidence estimates. It is clear that a model-based prognostics approach is a reasonable choice, however, as it can be inferred from the survey presented above, state of the art approaches fail to take into account the information related to the stochasticity of the environment and of the operating conditions. This might impose severe limitations on the adaptability and consequently on the robustness of the estimated range.

Given these limitations, this work sets as main research objective to: *Conceive and deploy a novel model-based prognostic solution for assessing the range of mobile systems which operate subject to energy constraints and under randomly-varying environments*. To achieve this objective a comprehensive research plan has been elaborated. Figure 1.3 depicts the pathway that structures both our research plan and this work.

The core of our approach is found in the center of the figure (**chapter 2**), which describes the RUL estimation in the context of model-based prognostics and presents the research methodology and the experimental system employed in this work. The steps depicted in the upper left hand side of the figure, namely **chapter 3** and **chapter 4**, comprise the development, parametrization and validation of physics-based models for computing the power requirements of the mobile system as result of its interaction with its operating surroundings and for accurately capturing the nonlinear capacity behavior exhibited by the energy storage system used to power the mobile system.

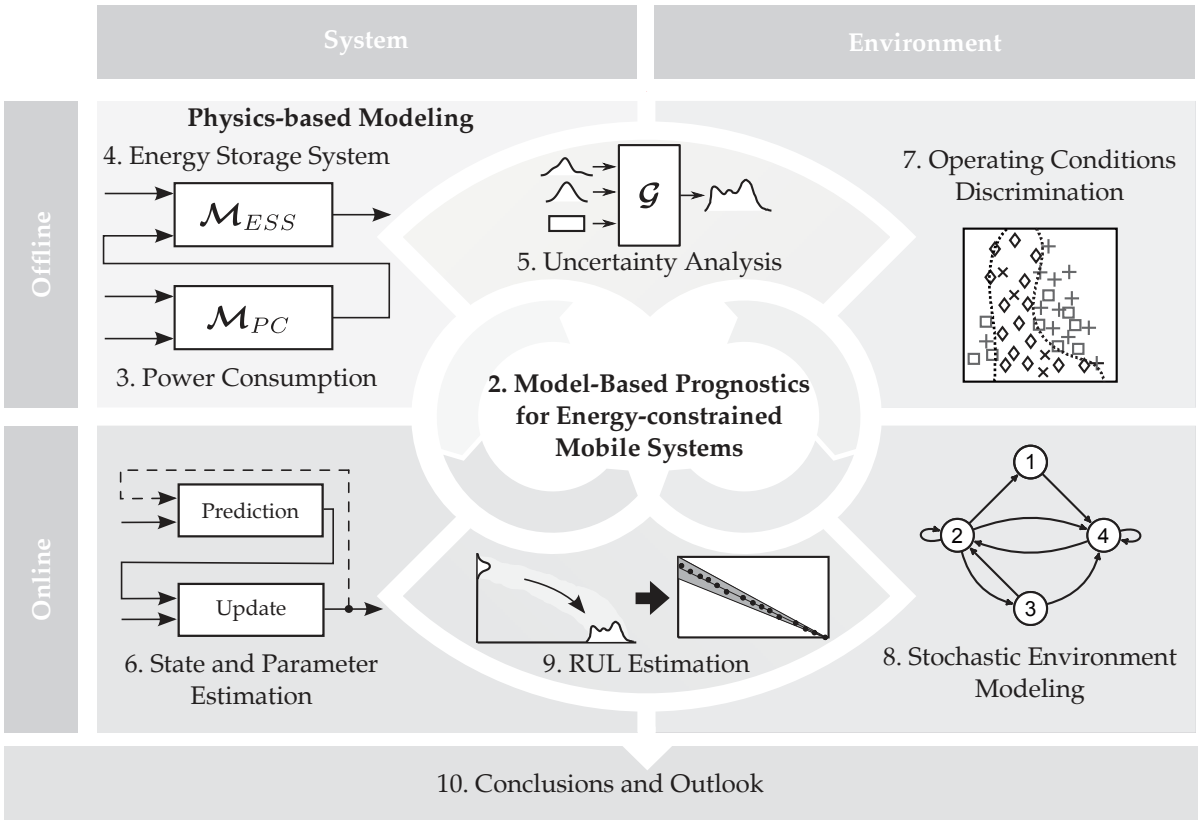

Figure 1.3.: Research plan and structure of the thesis.

With the purpose of reducing the number of uncertain variables, and therefore the computational complexity of the RUL estimation, **chapter 5** focuses on carrying out an uncertainty analysis with the aim of prioritizing the parameters which mostly contribute to the total uncertainty of the estimated RUL.

The bottom left hand side of the figure, i.e., **chapter 6** aims to characterize the uncertainty reflecting the lack of knowledge about the energy available in the energy storage system and introduces the implementation of a nonlinear observer for monitoring it. The right hand side of the figure basically deals with the operating environment. First, **chapter 7** introduces an approach developed to extract, select and compute features from measured variables of the environment, with the aim of classifying, as good as possible, different operating conditions, with the aim of adapting the RUL estimation to changing operating conditions. On this basis, **chapter 8** focuses on modeling the environment from a Markovian point of view, more concretely, on modeling both the evolution of the different operating conditions and the interaction between the mobile system and its surroundings, which conforms the basis for predicting the power requirements dictated by the future operation of the system.

The bottom side of the image (**chapter 9**) integrates the knowledge obtained in previous chapters into an unified solution for estimating the RUL of mobile systems subject to energy constraints with the aim of achieving the research objective by means of a comprehensive series of experiments carried out under controlled and real conditions. Finally, conclusions on the findings gained throughout this work are drawn and an outlook on future work is presented in **chapter 10** .

2

Model-based Prognostics for Energy-Constrained Mobile Systems

We start by formulating the RUL estimation for energy-constrained mobile systems in the context of model-based prognostics. Section 2.2 introduces the methodology employed for addressing the research objective stated in section 1.3 and describes, with help of a case study, the phases incurred for its achievement. Finally, section 2.3 describes the steps for evaluating the proposed methodology.

2.1. Problem formulation

Let us assume that the energy storage of the mobile system incurs an environment-dependent depletion. At any time $t \geq 0$, the system environment is found at one of m states in the finite set $\mathbf{S} = \{s_1, s_2, ..., s_m\}$, with $2 \leq m < \infty$.

Without loss of generality, let the mobile system operate, at any given time $t \geq 0$, in an arbitrary environment-state $s_i \in \mathbf{S}$ and assume that it remains operating at s_i for a random period of time ΔT_{ij}, herein also referred to as sojourn time, before the environment changes to another state $s_j \in \mathbf{S}$. Also assume that the environment randomly evolves over time as a discrete-state stochastic process following a multinomial distribution with parameters $\boldsymbol{\pi}_i = \left(\pi_{i1}, \pi_{i2}, ..., \pi_{ij}, ..., \pi_{im}\right)$, i.e.,:

$$S_i \sim \text{Multinomial}\left(\pi_{i1}, \pi_{i2}, ..., \pi_{ij}, ..., \pi_{im}\right), \tag{2.1.1}$$

where $\pi_{ij} = p\left(s_i \rightarrow s_j\right) = p\left(s_{ij}\right)$ is the probability of the environment incurring a transition from state s_i to state s_j.

In addition, let us describe the energy storage system powering the mobile system with the following continuous-time nonlinear time-variant state space representation:

$$\dot{\mathbf{x}}\left(t\right) = \mathbf{f}\left(\mathbf{x}\left(t\right), \mathbf{u}\left(t\right), \boldsymbol{\theta}\left(t\right), \mathbf{v}\left(t\right), t\right), \tag{2.1.2}$$

$$\mathbf{y}\left(t\right) = \mathbf{h}\left(\mathbf{x}\left(t\right), \mathbf{u}\left(t\right), \boldsymbol{\theta}\left(t\right), \mathbf{n}\left(t\right), t\right), \tag{2.1.3}$$

where t is the continuous time variable, $\mathbf{x}\left(t\right) \in \mathbb{R}^{N_x}$ is the state vector, $\mathbf{u}\left(t\right) \in \mathbb{R}^{N_u}$ represents the input vector, $\boldsymbol{\theta}\left(t\right) \in \mathbb{R}^{N_\theta}$ is the parameter vector, $\mathbf{v}\left(t\right) \in \mathbb{R}^{N_v}$ is the process noise vector, $\mathbf{n}\left(t\right) \in \mathbb{R}^{N_n}$ is the measurement noise vector and $\mathbf{y}\left(t\right) \in \mathbb{R}^{N_y}$ is the output vector. The function $\mathbf{f} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \times \mathbb{R}^{N_\theta} \times \mathbb{R}^{N_v} \times \rightarrow \mathbb{R}^{N_x}$ and the function

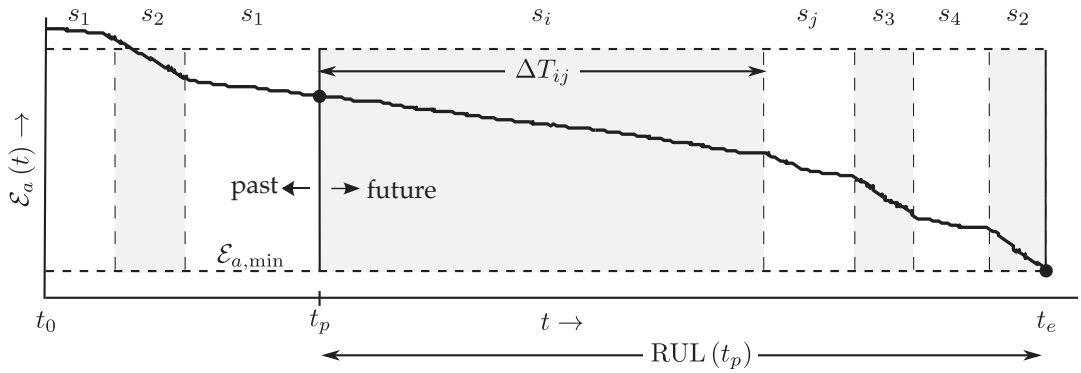

Figure 2.1.: Exemplification of the temporal evolution of $\mathcal{E}_a(t)$ in dependence on the operating conditions.

$\mathbf{h} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_u} \times \mathbb{R}^{N_\theta} \times \mathbb{R}^{N_n} \rightarrow \mathbb{R}^{N_y}$ describe the input-to-state dynamics and the state-to-output relations of the system, respectively. In addition, let:

$$\mathcal{E}_a(t) = \boldsymbol{\vartheta}(\mathbf{x}(t), \boldsymbol{\theta}(t), \mathbf{y}(t), t) \tag{2.1.4}$$

be the energy available in the system at any time $t \geq 0$ and define the function $\boldsymbol{\vartheta} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_\theta} \times \mathbb{R}^{N_y} \rightarrow \mathbb{R}$ as the map between system states, parameters and outputs and the energy available of the system. Furthermore, define:

$$T_e(t) = \mathbf{t}(\mathcal{E}_a(t), t) \tag{2.1.5}$$

as an indicator of the degree of depletion of the energy storage system where the threshold function $\mathbf{t} : \mathbb{R}^{N_x} \times \mathbb{R}^{N_\theta} \times \mathbb{R}^{N_y} \rightarrow \mathbb{R}$ maps the energy available $\mathcal{E}_a(t)$ to the Boolean domain $\mathbb{B} \triangleq [0,1]$, such that:

$$T_e(t) = \begin{cases} 1 & \text{if } \mathcal{E}_a \leq \mathcal{E}_{a,\min} \\ 0 & \text{if } \mathcal{E}_a > \mathcal{E}_{a,\min} \end{cases}, \tag{2.1.6}$$

where $\mathcal{E}_{a,\min}$ is the minimum amount of energy that has to be available, such that the mobile system remains operating according to its specifications.

As it can be seen in the figure 2.1 determining the energy available $\mathcal{E}_a(t)$ anywhere from the time t_p, also referred to as time of prediction, until the time t_e, at which the system runs out of energy, depends not only on $\mathcal{E}_a(t_p)$ but also on the energy consumed by the mobile system during this period of time.

The energy available at any time instant $t \geq t_0$, with $t_0 = 0$, is expressed by:

$$\mathcal{E}_a(t) = \mathcal{E}_a(t_0) + \int_{t_0}^{t} P(t)\, dt. \tag{2.1.7}$$

where $P(t)$ represents a continuous-time continuous-valued random process which describes the power required to run the mobile system at any time $t \geq 0$ as result of its interaction with the operating environment. The RUL is defined at given $t_p \geq 0$ as:

$$\text{RUL}(t_p) = t_e - t_p = \inf\left(t \in \mathbb{R} : t \geq t_p \wedge T_e(t) = 1\right) - t_p, \tag{2.1.8}$$

where t_e represents the end-of-life (EOL) of the system.

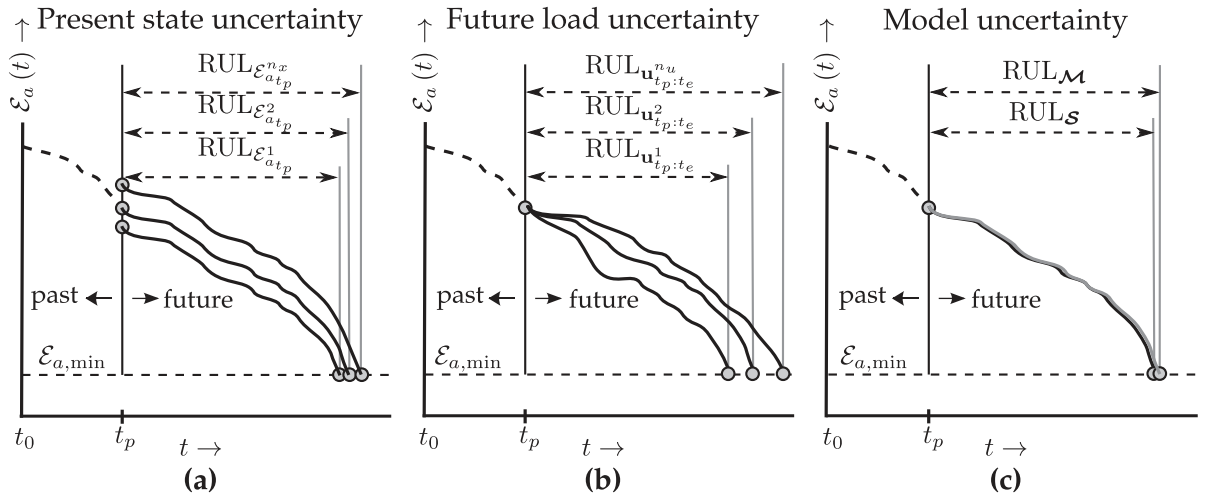

Figure 2.2.: Classification of the sources of uncertainty in present state **(a)**, future load **(b)** and model uncertainty **(c)**.

The definition of the RUL given by equation 2.1.8 is basically built around two premises. The former considers that both $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ exactly describe the temporal behavior of states, parameters and outputs of the system and that at any given time $t > 0$, $\mathcal{E}_a(t)$ can be computed by means of $\boldsymbol{\vartheta}(\cdot)$ without error. The latter assumes that the load of the mobile system, herein represented by $\mathbf{u}(t)$ in equations (2.1.2) and (2.1.3), in the interval $t_p \leq t \leq t_e$ is known precisely. However, in practical applications, due to the presence of uncertainty, none of these premises is totally true.

Identifying and quantifying the sources of uncertainty is the key for properly determining the RUL. Researchers have classified the sources of uncertainty in aleatory (physical variability), and epistemic (lack of knowledge) in order ease the tasks of uncertainty quantification, propagation and management. This classification, however, is not best suited in the context of this thesis. The reason for this is that the RUL is aimed to be estimated for a particular mobile system, thus, the variability among *identical mobile systems* plays an irrelevant role. Sankararaman and Goebel (2015) proposes a more proper uncertainty classification scheme, which considers the core elements involved in estimating the RUL. The *present state uncertainty* reflects the lack of knowledge about the energy available in the system at time t_p. Generally, this quantity is expressed indirectly in terms of system states and parameters, which however, in most cases are not observable and consequently have to be estimated. Figure 2.2 (**a**) shows how the uncertainty related to the initial available energy can affect the RUL of the system. The term n_x represents the number of possibles values of the available energy at time t_p. The *future load uncertainty* mirrors the unpredictability of the future operating and environment conditions of the mobile system and therefore the future power requirements. The lack of knowledge about the future usage of the mobile system represents the largest source of uncertainty in the RUL estimation (see figure 2.2 (**b**)). The term n_{u} represents the number of all possible future load profiles. As it was previously discussed, the core of any model-based prognostic approach is the use of models for describing the behavior of the system states and outputs in time, given by the functions $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$, as a response to the load of the system. Furthermore, the function $\boldsymbol{\vartheta}(\cdot)$ models the temporal evolution of $\mathcal{E}_a(t)$ and the func-

tion $\mathbf{t}(\cdot)$ maps it to the performance of the system. As shown in figure 2.2 (**c**), even if $\mathcal{E}_a(t_p)$ and $\mathbf{u}_{t_p:t_e}$ are assumed to be completely known at the time of prediction t_p, the RUL given by equation (2.1.8) might differ from the true RUL of the system due to the *model uncertainty*. This source of uncertainty comprises uncertainties related to parameters, the structure of the system and the inherent process noise present in $\mathbf{f}(\cdot)$, $\mathbf{h}(\cdot)$, $\boldsymbol{\vartheta}(\cdot)$ and $\mathbf{t}(\cdot)$.

In light of the above discussion, it seems fair to conclude that the RUL is an uncertain variable whose randomness is governed by the uncertainties previously mentioned. It becomes clear that properly estimating the RUL requires quantifying the sources of uncertainty involved in its computation first, and then propagating them in order to compute their combined effect, expressed by a probability density function. Thus, estimating the RUL becomes an uncertainty propagation problem, as stated by Sankararaman and Goebel (2013).

For a better understanding of the RUL estimation problem described from an uncertainty propagation point of view, let us redefine the RUL of the system as:

$$\mathrm{RUL}(t_p) = \mathbf{G}\left(\boldsymbol{\vartheta}(t_p), t_p\right), \tag{2.1.9}$$

where $\mathbf{G}(\cdot)$ is a function that maps the joint probability density function (PDF) of the uncertain variables contained in $\boldsymbol{\vartheta}$, that is, the states estimate $\mathbf{x}(t_p)$, the future parameters $\boldsymbol{\theta}(t_p : t_e)$ and the future inputs $\mathbf{u}(t_p : t_e)$ to the $\mathrm{RUL}(t_p)$, i.e., the probability density function of the RUL of the system at given time $t_p \geq 0$, as depicted in figure 2.3 (**a**). The task is to identify the function $\mathbf{G}(\cdot)$ that allows mapping the uncertain variables to the PDF of the uncertain output variable, i.e., the RUL of the system.

In the effort of finding a solution for the general uncertainty propagation problem, researchers have proposed approaches, which can be classified in analytical and numerical methods (Sankararaman and Goebel 2013), (see figure 2.3 (**b**)). Analytical approaches, e.g., response variability methods (Isukapalli, Roy, and Georgopoulos 1998), structural reliability methods (Zhao and Ono 1999) and stochastic finite element methods (Stefanou 2009) make some assumptions about the distribution of the uncertain input variables and about the function $\mathbf{G}(\cdot)$ evaluated during the uncertainty propagation, in this case the computation of the RUL, such that the statistical information of the output variable, here the RUL, can be conveniently computed in an analytical manner. For those cases in which the function $\mathbf{G}(\cdot)$ is linear and the uncertain input variables are Gaussian, a closed form solution for the propagation of uncertainty might be obtained. However, due to the nonlinearities presented in the capacity behavior of most energy storage systems, the computation of the RUL cannot be achieved using a linear $\mathbf{G}(\cdot)$. The paramount challenge in applying analytical methods in model-based prognostics is the integration of the future load uncertainty into the RUL estimation. In most cases the future the power requirements are governed by a non-Gaussian stochastic process, which cannot be characterized by stationary probability distributions. In addition to this, estimating the RUL, as presented in equation (2.1.9), requires knowledge about $\mathbf{u}(t_p : t_e)$. It becomes obvious that this information is not available in advance, since t_e is the actual outcome of the RUL estimation. The implicitness of t_e makes even more difficult applying the aforementioned uncertainty propagation methods in estimating the RUL.

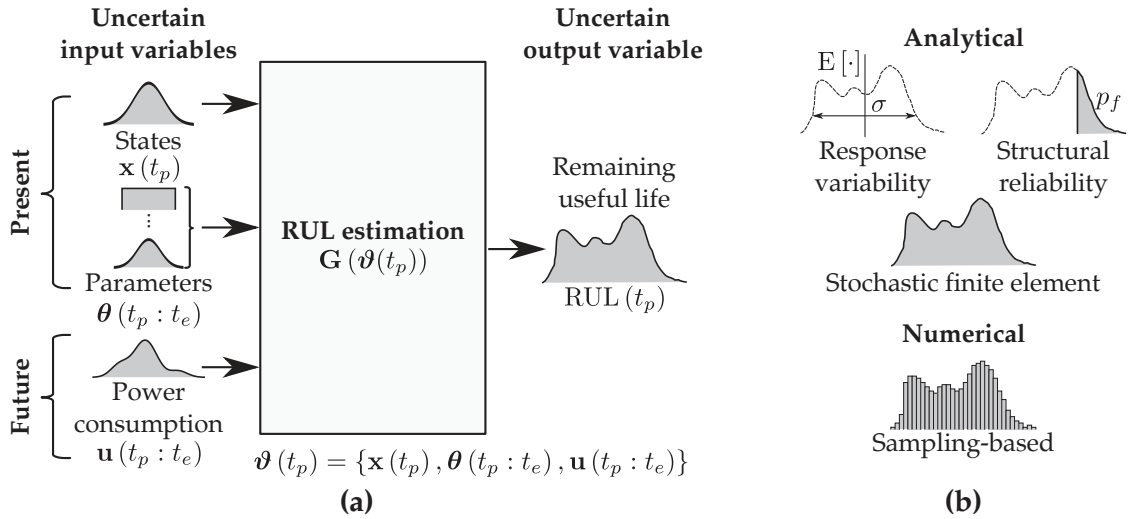

Figure 2.3.: Representation of the RUL estimation as an uncertainty propagation problem (**a**) and classification of uncertainty propagation approaches (**b**).

2.2. Research methodology

One solution to the aforementioned issues is to apply numerical methods such as sampling-based approaches, which rely on random numbers and probability theory in order to approximate the probability density function of the output variable in those uncertainty propagation problems that lack of a closed form solution. Sampling-based approaches simulate multiple realizations of $\mathbf{G}(\cdot)$ by repeatedly changing the values of the input variables based on a pseudo-random sampling, e.g., latin hypercube (McKay, Beckman, and Conover 2000) or importance sampling (Tokdar and Kass 2010). From a sufficient large number of simulations (samples) the aimed solution is approximated. Although these approaches are high accurate, their computational efficiency is strongly affected by the large number of samples needed to converge to the actual probability distribution of the RUL.

Deploying a prognostics solution for a mobile system requires not just executing a *prediction* algorithm. It also processing signals gathered by sensors and analysing reference variables dictated by the operator in order to properly incorporate information about the system itself or from its surroundings into the actual the RUL estimation. As it is shown in figure 1.1 (**d**), all these processes take place in one or in multiple information processing units, e.g., digital computers. Modern mobile systems are equipped with computers usually constrained in both storage and computational resources, causing in this manner remarkable restrictions to their real-time capabilities. As it is shown in equations (2.1.2) and (2.1.3), until know the energy storage system powering the mobile system has been described by a continuous-time nonlinear time-variant state space representation. Nevertheless, such representation is not best suited for algorithms executed in digital computers since they expect time continuous inputs. In order to solve this problem we discretize equations (2.1.2) and (2.1.3) by introducing a sampling time Δt, within which new sensor information must be processed and the prediction algorithm must be completely executed. Such a discretization requires a stepped input $\mathbf{u}(t) \equiv \mathbf{u}_k \forall t \in [k(\Delta t), (k+1)(\Delta t)[$.

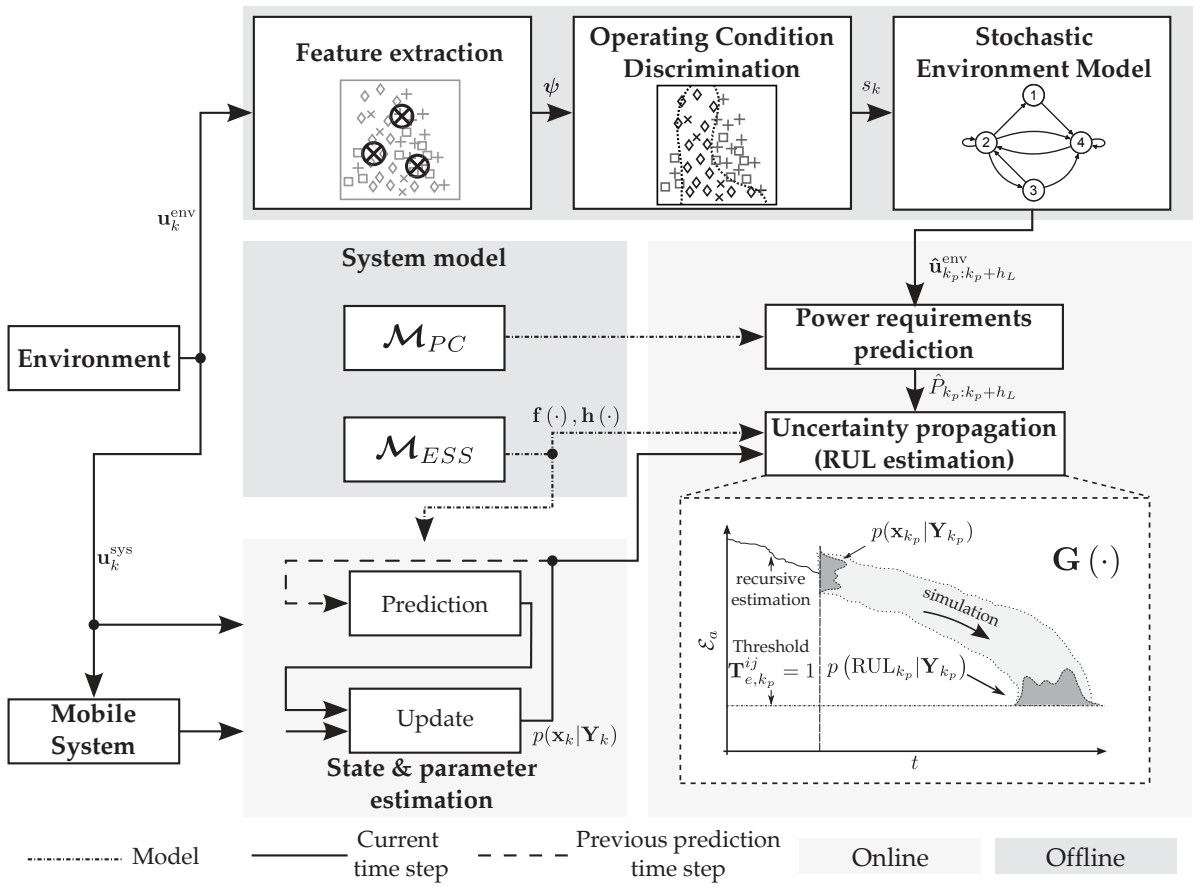

Figure 2.4.: Proposed research methodology for model-based prognostics of energy-constrained mobile systems.

Thus, discretizing equations (2.1.2) and (2.1.3) yields:

$$\mathbf{x}_{k+1} = \mathbf{f}\left(\mathbf{x}_k, \mathbf{u}_k, \boldsymbol{\theta}_k, \mathbf{v}_k, k\right), \tag{2.2.1}$$

$$\mathbf{y}_k = \mathbf{h}\left(\mathbf{x}_k, \mathbf{u}_k, \boldsymbol{\theta}_k, \mathbf{n}_k, k\right). \tag{2.2.2}$$

Van Der Merwe (2004) suggests that describing (2.2.1) and (2.2.2) as a dynamic state-space model (DSSM) eases their application in the context of probabilistic inference. In this sense, the states $\mathbf{x} \in \mathbb{R}^{N_x}$ temporally evolve following a non fully observable first order Markov process according to the conditional probability function $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right)$, which is fully specified by the function $\mathbf{f}\left(\cdot\right)$ together with the process noise distribution $p\left(\mathbf{v}_k\right)$. The observations $\mathbf{y}_k \in \mathbb{R}^{N_y}$ are conditionally independent given $\mathbf{x}_k$ and are generated from the observation likelihood $p\left(\mathbf{y}_k|\mathbf{x}_k\right)$, which in turn is specified by the output function $\mathbf{h}\left(\cdot\right)$ and by the noise distribution $p\left(\mathbf{n}_k\right)$. This description together with the proper characterization of the sources of uncertainty outlined in section 2.1 sets the basis of the sampled-based methodology investigated throughout this work and that is schematically depicted in figure 2.4. In section 2.1 we have mentioned that the RUL estimation problem is concerned with first estimating the energy available $\mathcal{E}_a$ at any given time and then to predict the future operation conditions of the mobile system in order to determine the distribution of the time at which $T_e = 1$. In the sequel we briefly introduce the uncertainty characterization and propagation schemes investigated throughout this work.

Model uncertainty characterization

The first step in deploying a model-based prognostics solution is the development, parametrization and validation of a *system model* that allows describing the temporal behavior of the states and outputs, given by the functions $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$, as a response to the load of the system. As it can be seen in figure 2.4 the system model is composed of $\mathcal{M}_{ESS}$ and $\mathcal{M}_{PC}$. On the one side, $\mathcal{M}_{ESS}$ represents a model that captures the nonlinear capacity behavior exhibited by the energy storage system used to power the system. This model manages to deal with constant or highly dynamic loads, negative or positive loads as well as with changes in the operating temperature. The model is used by the function $\boldsymbol{\vartheta}(\cdot)$ to describe the evolution of the available energy $\mathcal{E}_a$, and also by $\mathbf{t}(\cdot)$ to map $\mathcal{E}_a$ to the availability of the mobile system. The purpose of $\mathcal{M}_{PC}$, on the other side, is to describe the interaction between the mobile system and the environment, more precisely, to compute the power requirements dictated by the operating conditions.

The characterization of the model uncertainty introduced by $\mathcal{M}_{ESS}$ and $\mathcal{M}_{PC}$ is carried out during the stage of parameter identification by quantifying the modeling error, by determining the validity of the model (see chapters 3 and 4) and by a systematic uncertainty analysis (see chapter 5) used to prioritize important variables and to neglect the uncertainty related to those variables whose contribution to the total uncertainty of the estimated RUL is meaningless.

Present state uncertainty characterization

This task is performed in the *state and parameter estimation* phase by a nonlinear observer which monitors the remaining energy available in the energy storage system. The intended outcome of this phase is to represent the most up-to-date knowledge of the system states and parameters at any given time based on the current and on all past observations. To this aim the observer recursively approximates the posterior probability $p(\mathbb{X}_k|\mathbf{Y}_k)$ of the state variables and parameters by a set of $n_{\mathbb{X}}$ weighted samples $\mathcal{S}_k = \{\mathbb{X}_k^i, w_k^i\}_{i=1}^{n_{\mathbb{X}}}$. Here $\mathbb{X}_k^i$ is the set of samples representing the joint state-parameter space, w_k^i are the associated importance weights and $\mathbf{Y}_k = \mathbf{y}_{0:k}$ is the set of all observations generated up to time k. Each sample is drawn from an *a priori* estimation of the joint state space and is propagated through the function $\mathbf{f}(\cdot)$ in the *prediction step*. Then, the value of each sample is updated from new observations through the output function $\mathbf{h}(\cdot)$ in the *measurement update step*.

In this step the weight of each sample is updated according to the likelihood of a new measurement given the sample. Finally, the probability distribution of the joint state variables at discrete time k is approximated by:

$$p(\mathbb{X}_k|\mathbf{Y}_k) \approx \frac{1}{n_{\mathbb{X}}} \sum_{i=1}^{n_{\mathbb{X}}} w_k^i \delta\left(\mathbb{X}_k - \mathbb{X}_k^i\right) \tag{2.2.3}$$

where $\delta(\cdot)$ describes the Dirac delta function located at $\mathbb{X}_k^i$. The *posterior* state estimate establishes the starting point for the second phase, in which the set of samples approximating $p(\mathbb{X}_k|\mathbf{Y}_k)$ are employed for estimating the RUL at given time to prediction k_p. To this aim the posterior estimate $p(\mathbb{X}_{k_p}|\mathbf{Y}_{k_p})$ is set as initial condition.

Future load uncertainty characterization

Characterizing and integrating the uncertainty related to the future power consumption is a paramount challenge during the RUL estimation. This is due to the fact that the future energy needs are mostly governed by a stochastic process, and therefore can not be characterized by merely stationary probability distributions. The prediction of the power consumption of the mobile system, and thereby the future load uncertainty characterization, is achieved by means of a *stochastic environment model* based on a first order Markov process, which captures the behavior dictated by both the temporal evolution of the operating conditions and the actual interaction between the mobile system and its surroundings. The outcome of the stochastic environment model is $\hat{\mathbf{u}}^{\text{env}}_{k_p:k_p+h_L} = [\hat{\mathbf{v}}_{\text{x},k_p:k_p+h_L}\ \hat{\mathbf{a}}_{\text{x},k_p:k_p+h_L}\ \hat{\boldsymbol{\alpha}}_{\text{road},k_p:k_p+h_L}]^T$ which represents a realization of the underlying stochastic process describing the temporal behavior of the longitudinal speed $\hat{\mathbf{v}}_{\text{x},k_p:k_p+h_L} = [\hat{v}_{\text{x},k_p}, \ldots, \hat{v}_{\text{x},k}, \ldots, \hat{v}_{\text{x},k_p+h_L}]$ and the acceleration $\hat{\mathbf{a}}_{\text{x},k_p:k_p+h_L} = [\hat{a}_{\text{x},k_p}, \ldots, \hat{a}_{\text{x},k}, \ldots, \hat{a}_{\text{x},k_p+h_L}]$ of the mobile system as well as the slope profile of the road $\hat{\boldsymbol{\alpha}}_{\text{road},k_p:k_p+h_L} = [\hat{\alpha}_{\text{road},k_p}, \ldots, \hat{\alpha}_{\text{road},k}, \ldots, \hat{\alpha}_{\text{road},k_p+h_L}]$ for a given horizon length h_L measured from the prediction time k_p.

Adapting the stochastic environment model to changes in the environment conditions allows accounting for the most up to date operating conditions, and therefore improves the accuracy of the predicted power requirements. Such an adaptation, however, requires updating the transition probabilities of the Markov model as new information becomes available. The first step towards achieving such an update is to discriminate each operating condition as it occurs, so that transitions between conditions can be counted online. The *operating condition discrimination* is the function $\mathbf{q} : \boldsymbol{\Psi} \subseteq \mathbb{R}^{N_\Psi} \rightarrow s \in \mathbf{S}$, that allows mapping instances of *extracted features* $\boldsymbol{\psi} \in \boldsymbol{\Psi}$ from measurements of speed v_{x} and acceleration a_{x}, to the corresponding class (operating condition) $s \in \mathbf{S}$ where $\mathbf{S}$ is the set of operating conditions. Built on these features, in a second step the function $\mathbf{q}(\cdot)$ is employed for discriminating operating conditions with the aim of updating the parameters of the stochastic environment model. The convergence speed of the updated parameters is specially important at this point. A slow adaptation of the Markov model is inconvenient in this context since they would cause the characterization of the most up to date operating conditions to fail. Therefore, it is desired to have a recursion that progressively fades the influence of older parameters during the updating process.

Uncertainty propagation (RUL estimation)

As it can be appreciated in figure 2.5, a sample-based approach in terms of a Monte Carlo simulation is employed to estimate the RUL of the mobile system. To this aim, the posterior estimate $p(\mathbb{X}_{k_p}|\mathbf{Y}_{k_p})$, is approximated by the finite set of samples $\mathcal{S}_{kp} = \{\mathbb{X}^i_{kp}, w^i_{kp}\}^{n_\mathbb{X}}_{i=1}$. Here, $\mathbb{X}^i_{kp}$, during the *state and parameter estimation* phase, serves as the starting point at given time k_p. In this sense, each sample is independently propagated forward in time through a realization of $\hat{\mathbf{u}}^{\text{env}}_{k_p:k_p+h_L}$, generated by the stochastic environment model, until the $\mathcal{E}_{a,\min}$ is reached. Once this happens, the RUL of all samples is computed and used to approximate the posterior $p(\text{RUL}_{k_p}|\mathbf{Y}_{k_p})$. To account for the fact, that during the sample propagation step the shape of $p(\mathbb{X}_{k_p}|\mathbf{Y}_{k_p})$

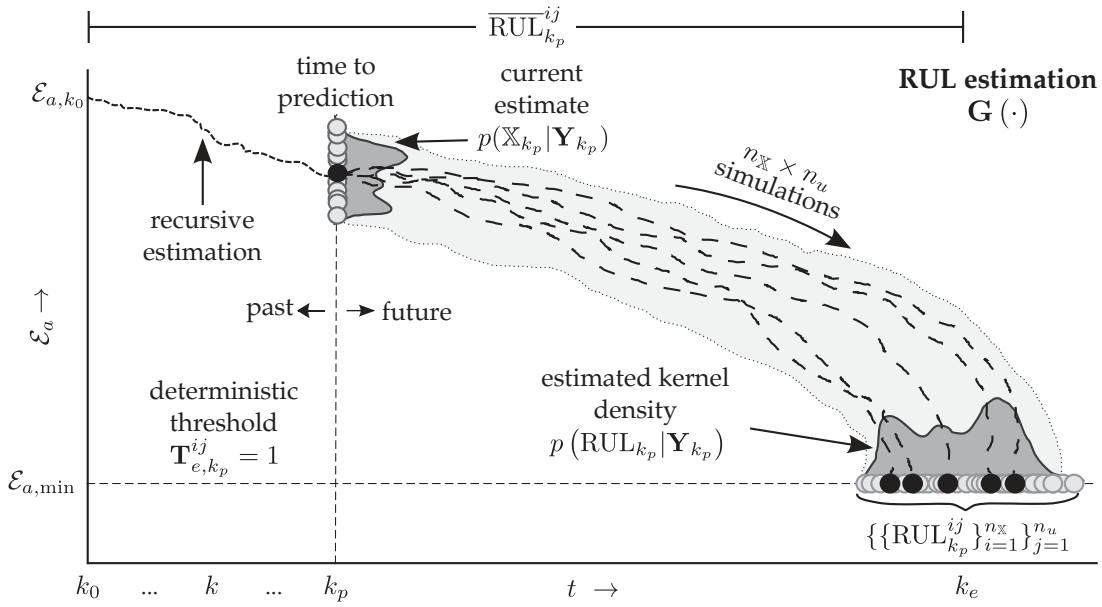

Figure 2.5.: RUL estimation based on a sample-based Monte-Carlo simulation. For the sake of clarity the figure depicts the RUL estimation of just one sample/particle.

may change, due to noise and process nonlinearities, it is required to update the set of weights at each future iteration. However, during this step no new observation, which could serve for updating the weights, can be generated. This implies that an update procedure for the sample weights, as it would happen in a typical filtering problem, cannot be carried out. This issue is addressed by assuming the weights as invariant from the time k_p to $k_p + h_L$. This assumption is justified by considering the uncertainty added by model inaccuracies or by the ignorance about future operation conditions to be large in comparison to the uncertainty which comes from considering constant sample weights. In this way, the set of weighted samples $\mathcal{S}_{k_p}$ is simply propagated forward in time by simulating the behavior of the system as reaction to a hypothesized future operation condition, until $\mathcal{E}_{a,\min}$ is reached. Once all samples have reached this point, i.e., when $T_{e,k_p}^{i} = 1$, the $\mathrm{RUL}_{k_p}^{i}$ of each sample is determined and combined with its weight $w_{k_p}^{i}$ to approximate $p(\mathrm{RUL}_{k_p}|\mathbf{Y}_{k_p})$ as follows:

$$p\left(\mathrm{RUL}_{k_p}|\mathbf{Y}_{k_p}\right) \approx \sum_{i=1}^{n_{\mathbb{X}}} w_{k_p}^{i} \mathrm{RUL}_{k_p}^{i}. \tag{2.2.4}$$

The RUL estimation, as formulated by equation (2.2.4), requires propagating the set of samples through a single hypothesized predicted profile of the future operation conditions of the system. However, such a propagation accounts just for the present state uncertainty but it does not consider the future load uncertainty. Taking this uncertainty into account requires propagating the set of samples through multiple realizations of $\hat{\mathbf{u}}_{k_p:k_p+h_L}^{\mathrm{env}}$, and not just through a single one. Thus, the computational complexity of such a propagation becomes a function of $n_{\mathbb{X}} \times n_u$ (Daigle, Saxena, and Goebel 2012), where n_u is the number of realizations needed to characterize the uncertainty dictated by the future operation conditions. The set of weighted samples is then propagated through multiple profiles until all of them, along all predicted pro-

files, have reached the threshold, i.e., when $T_{e,k_p}^{ij} = 1$. Here j represents each realization of the operation profile. Accordingly, the probability distribution $p(\mathrm{RUL}_{k_p}|\mathbf{Y}_{k_p})$ is approximated by means of kernel density estimation:

$$p\left(\mathrm{RUL}_{k_p}|\mathbf{Y}_{k_p}\right) \approx \frac{1}{n_{\mathbb{X}} n_u h} \sum_{j=1}^{n_{\mathbb{X}} n_u} \mathbf{K}_G \left(\frac{\mathrm{RUL}_{k_p} - \overline{\mathrm{RUL}}_{k_p}^{ij}}{h} \right), \tag{2.2.5}$$

where $\mathbf{K}_G(\cdot)$ is a Gaussian kernel function, h is a smoothing factor known as bandwidth and $\overline{\mathrm{RUL}}_{k_p}^{ij}$ is the mean RUL of the set $\{\{\mathrm{RUL}_{k_p}^{ij}\}_{i=1}^{n_{\mathbb{X}}}\}_{j=1}^{n_u}$ computed as:

$$\overline{\mathrm{RUL}}_{k_p}^{ij} \approx \frac{1}{n_u} \sum_{j=1}^{n_u} \sum_{i=1}^{n_{\mathbb{X}}} w_{k_p}^i \mathrm{RUL}_{k_p}^{ij}. \tag{2.2.6}$$

It must be noted that the importance of each realization of $\hat{\mathbf{u}}_{k_p:k_p+h_L}^{\mathrm{env}}$ are equally weighted by means of $\frac{1}{n_u}$.

To measure the performance of the RUL estimation we employ the relative accuracy (RA) and the alpha-lambda ($\alpha - \lambda$) metric (Saxena et al. 2009). In the context of the RUL estimation the RA at an arbitrary time to prediction k_p is given by:

$$\mathrm{RA}_{k_p} = 100 \left(1 - \frac{\left| \mathrm{RUL}_{k_p}^* - \mathrm{RUL}_{k_p} \right|}{\mathrm{RUL}_{k_p}^*} \right), \tag{2.2.7}$$

where $\mathrm{RUL}_{k_p}^*$ is the ground truth RUL at time k_p and RUL_{k_p} is the estimated RDR at that time. The $\alpha - \lambda$ metric serves to evaluate whether the estimated RUL lies within specified bounds expressed as a percentage above and below of the actual RUL.

2.3. Methodology evaluation

The estimation of the remaining driving range (RDR) of an electric vehicle (Oliva, Weihrauch, and Bertram 2013a) has been chosen as the case study to evaluate and to validate the research methodology introduced in section 2.2. In this context, the RDR estimation is concerned with predicting the power demand of the electric vehicle and with identifying the distance that it can drive with the energy stored in its battery before recharging is required. We consider the state of charge of the battery pack to be the indicator that determines the threshold condition defined in section 2.1 (recall equation 2.1.6), i.e., $T_e(\mathrm{SoC}) = 1$ if $\mathrm{SoC}_{\min}$ is reached and $T_e(\mathrm{SoC}) = 0$, otherwise. The $\mathrm{SoC}_{\min}$ is imposed by the battery management system (BMS) of the electric vehicle to protect the battery cells from a possible total charge depletion. In the following an overview of the system and setup used to carried out the series of experiments performed in this work is presented. Nevertheless, the details of each experiment are presented in the corresponding chapters of this work.

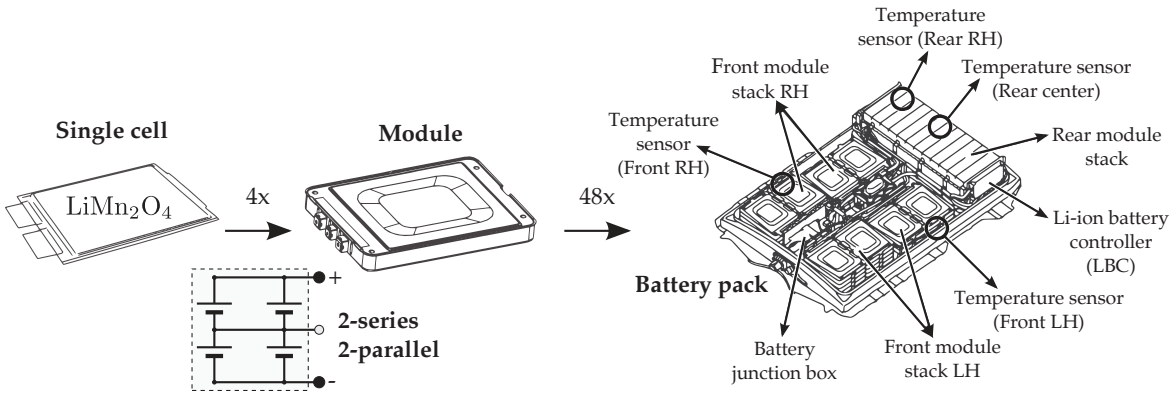

Figure 2.6.: Schematic representation of the energy storage system of the Nissan Leaf.

Experimental system

The electric vehicle used as experimental platform for gathering data and for testing the proposed approach is a Nissan Leaf (2011), which has been provided by the Institute of Control Theory and Systems Engineering of the TU Dortmund. This vehicle is propelled by a 80 kW and 280 Nm synchronous electric motor mounted in the front axle. The vehicle is powered by a 24 KWh Li-ion battery pack with a total capacity of 66.2Ah, according to the disclosure of Nissan. The battery pack is made of 48 modules distributed along two front and one rear stack under the floor of the vehicle, as shown in figure 2.6. Four sheet shaped $LiMn_2O_4$ cells are integrated into a single module with a 2-series, 2-parallel formation (2S2P). Each cell has a 32.5Ah capacity (at 0.3C rate) with a nominal voltage of 3.75 V. The large flat surface of the cells endows them with great cooling characteristics what improves their useful life. In order to keep the cells in a healthy state, it is recommended to maintain their operating voltage between 2.75 and 4.2 V. The pack is equipped with a Li-ion battery controller (LBC) installed next to the rear module stack. It is in charge of measuring the voltage and current of the battery pack, the temperature of each stack and the voltage of each cell, in order to monitor their state of charge (SoC) or state of health (SoH).

Experimental setup

For the experimental evaluation and validation of the sampled-based methodology shown in figure 2.4 three different experimental scenarios have been setup.

Scenario 1: Real (environment, mobile system and ESS): the first scenario, as illustrated in figure 2.7 (**a**), consists of the electric vehicle equipped with a data acquisition system, an inertial measuring unit (IMU) integrated with a differential global positioning system and a real-time system, as shown in figure 2.7 (**b**).

The task of data acquisition system (the DEWE-510 by Dewetron) is to record and synchronize the measurements gathered from both the IMU and the real-time system. The DEWE-510 system is operated with a DC voltage between 18 and 24 V. For experiments carried out on-road, the DEWE-510 is powered by three additional Li-ion batteries, which allow an operation between 3-4 h. The entire architecture of measuring system has been designed so that the energy supply of all devices is independent

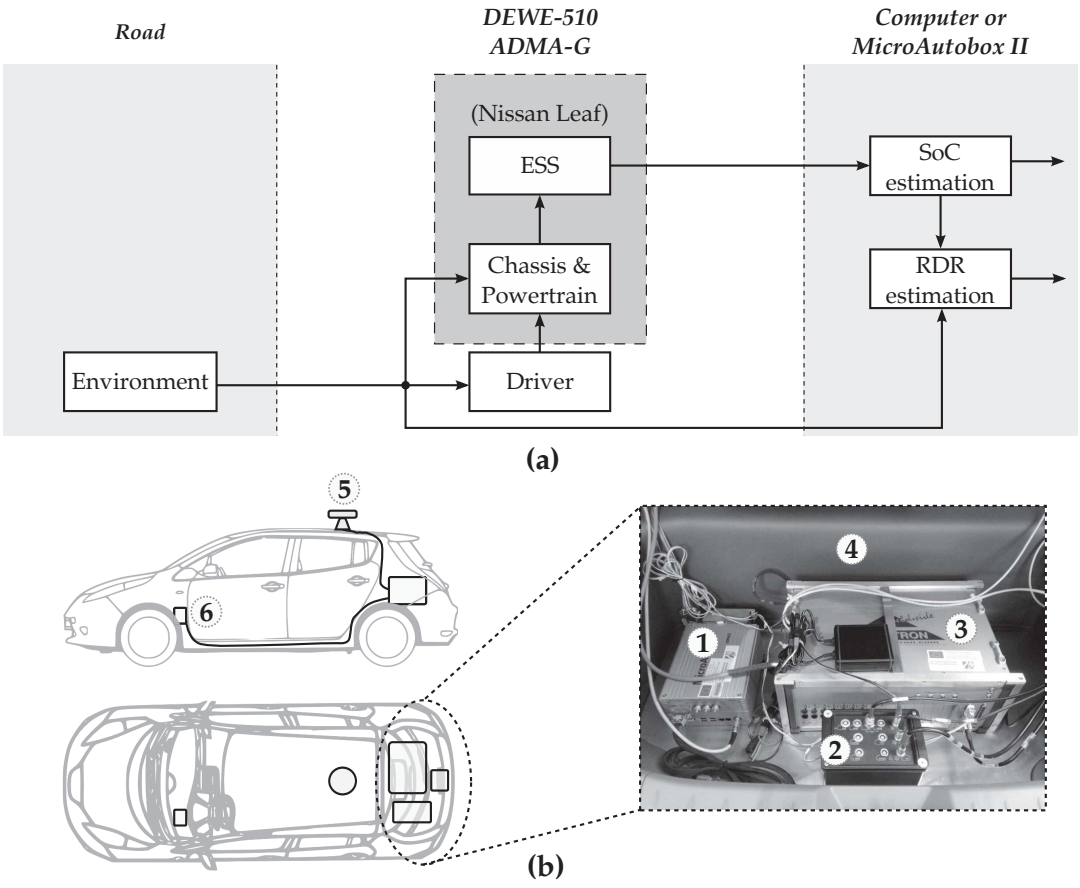

Figure 2.7.: Schematic representation (**a**) and real setup (**b**) of experimental scenario 1 composed of 1-MicroAutoBox II, 2-ADMA-G, 3-DEWE-510, 4-Nissan Leaf, 5-OEM6 and 6-OBD-II.

of the vehicle electrical system. In this manner no distortion of the measurements regarding power consumption is ensured.

The IMU is the Genesis ADMA-G. It allows measuring the position, speed and acceleration of the vehicle in x-, y- and z-direction. The system is a combination of a GPS receiver, the OEM6 of NovAtel, an acceleration sensor and a gyroscope. Its operating voltage is 12 V and is supplied by the DEWE-510. This systems uses a CAN protocol with 1000kBaud to communicate with the DEWE-510 system.

The real-time system, namely the dSPACE MicroAutoBox II, is powered by a DC source of 12 V and communicates with the DEWE-510 system via CAN. It basically serves to decode CAN messages, acquired through the OBD-II interface of the electric vehicle with a 500kBaud. There exist a total of 3 CAN busses that can be accessed via OBD-II, however, not all CAN messages carry relevant information. The messages that are of interest for this work are the vehicle speed in longitudinal direction, the SOC, terminal voltage and load current of the battery pack as well as the voltage of each cell, the temperature at four different locations of the pack, the ambient temperature and the rotational speed of the electric motor.

This setup is used to evaluate most theoretical concepts developed in this work. Firstly, as discussed in chapter 3 and 4, it allows gathering the data necessary to parametrize and validate the models $\mathcal{M}_{PC}$ and $\mathcal{M}_{ESS}$, respectively, as well as to analyze their uncertainty in chapter 5. Secondly, the availability of measurements, not

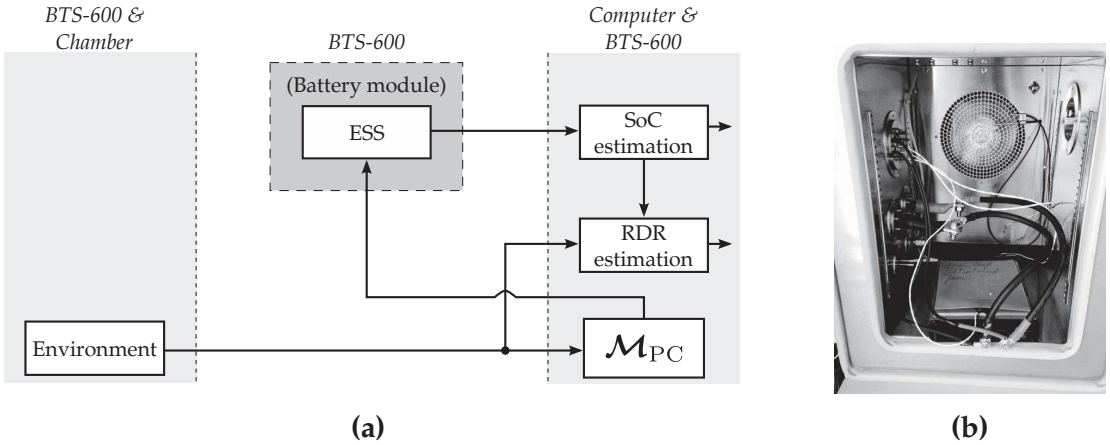

Figure 2.8.: Schematic representation (**a**) and real setup (**b**) of experimental scenario 2.

just of the battery pack but of each cell, is exploited in chapter 6 to evaluate the SOC estimation under real operating conditions. Thirdly and lastly, both the operating condition classifier from chapters 7, the stochastic environment model described in chapter 8 and the RDR estimation presented in chapter 9 also benefit from this setup, since they are primarily built around speed, acceleration and slope measurements.

Scenario 2: Real (environment and ESS) - Simulation (mobile system): the setup of the second scenario (see figure 2.8 (**a**)), allows carrying out tests directly on the modules of the battery pack. To this aim a series of experiments has been performed in cooperation with the battery competence center *Münster Electrochemical Energy Technology* (meet). In this scenario, the environment is simulated by the test bench shown in figure 2.8 (**b**), which consists of a Digatron battery test system BTS-600, a thermostatic chamber 60T by Sunrise and a host computer. The BTS-600 is responsible of charging or discharging the battery cells according to predefined profiles of power demand, which are generated off-line through simulation with help of real speed, acceleration and slope data and the model $\mathcal{M}_{PC}$. The BTS-600 has four channels, which are able to apply a maximum charge/discharge current of 200 A with a recorded accuracy of ± 10 mA. The measured voltage ranges between 0-5 V with an accuracy of ± 10 mV. As it will be further detailed in chapter 4 the capacity behavior and the transient response of the battery cell is highly dependent on the operating temperature. For this reason the ability to carry out tests on the cells at different temperatures has gained in importance during this research, otherwise, temperature dependencies could not be taken into account. To this aim the thermostatic chamber 60T, which is coupled with the BTS-600, has been employed to keep constant temperatures, ranging from 0 °C to 40 °C, with an accuracy of ± 0.25 °C.

During a charge/discharge routine the BTS-600 records the incoming and outgoing capacity of the cell by means of Coloumb Counting together with the terminal voltage, the load current and the temperature of the chamber. These measurements are used in various phases and for different purposes throughout this work, e.g., to parametrize and validate models of the energy storage system in chapter 4, to evaluate the accuracy of the state estimation in chapter 6 or to perform RDR estimations in chapter 9.

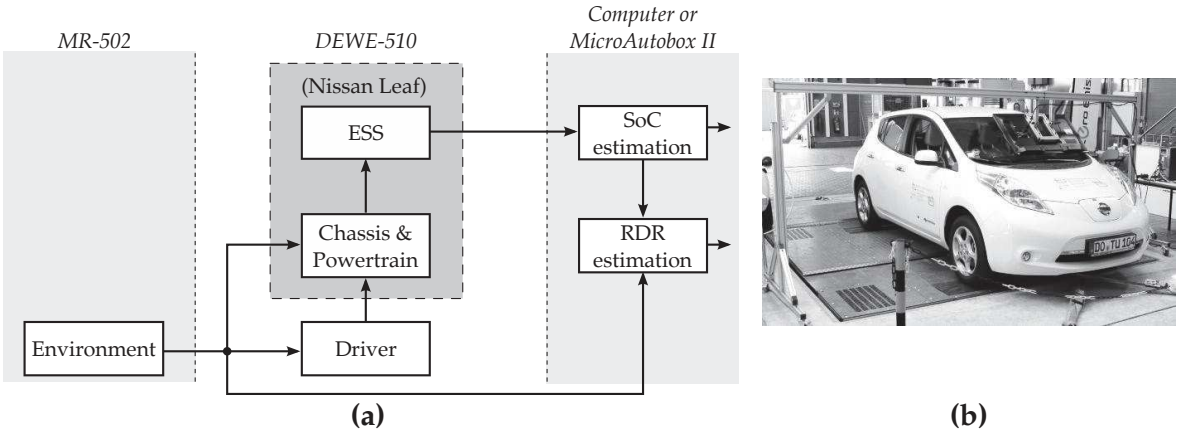

Figure 2.9.: Schematic representation (**a**) and real setup (**b**) of experimental scenario 3.

Scenario 3: Real (mobile system and ESS) - Simulation (environment): the scenario described above allows investigating the behavior of the energy storage system and to draw conclusions about the RDR under controlled conditions. However, to obtain more meaningful insights about the autonomy of the EV it is needed to consider uncertainties related to entire battery pack as well as the ones related to the power consumption of the vehicle. To achieve this, various experiments using the setup shown in figure 2.9 (**a**) have been carried out in cooperation with the *competence center for interoperable electromobility, infrastructure and networks of the TU Dortmund*. The test bench shown in figure 2.9 (**b**) consists of a MAHA MSR-500/2 4WD double roller chassis dynamometer equipped with one Eddy current brake and one electric motor per axle, the Nissan Leaf with the measurement equipment described above, a 42-inch monitor and a communication desk. The environment is simulated by the chassis dynamometer. At each experiment, the MSR-500 adjusts the braking torque at the rollers, such that the driving resistances can be accordingly simulated. The driver accelerates or brakes the vehicle, which is fixed with the wheels sitting on the rollers, with the aim of reaching and keeping the speed shown in the monitor placed in front of the vehicle. The required vehicle-specific coefficients needed to match the roll- to the street driving resistances has been determined using coast down tests.

The driving cycles are represented here by the speed in the form of time series, which as in the previous case of the scenario 2, have been gathered through real test drives on the road using the setup 1. During each experiment the DEWE-510 is used to acquire and synchronize measurements regarding the terminal voltage, both from the entire pack as well as each cell, the load current and the speed of the vehicle, which, instead of being measured with the IMU, is read from the OBD interface of the electric vehicle. The measurements obtained here also serve multiple purposes of this research, e.g., to validate the SoC estimation of the battery pack in chapter 6. Within this scenario it is possible to investigate the RDR estimation, (see chapter 9) under some controlled conditions, but taking into account the uncertainties introduced by the EV itself.

3

Power Consumption Modeling

In this chapter a detailed model of the interaction between the mobile system and its surroundings is developed and validated based on the physical properties of the electric vehicle. First, the approach employed to model the EV as well as its dynamics model are discussed in sections 3.1 and 3.2, respectively. Later, sections 3.3 and 3.4 deal with models for computing the power requirements. Finally, both the identification of parameters and the validity of the model are discussed in section 3.5.

3.1. Physics-based modeling of electric vehicles

From a physical point of view, the power demanded by an electric vehicle can be either modeled by a forward-facing or by a backward-facing approach (Guzzella and Sciarretta 2005). Figure 3.1 shows the direction of action of both approaches. In the forward-facing approach the electric vehicle is controlled by a driver model, which serves as a speed controller, to follow a desired speed. The driver model determines whether the vehicle accelerates or brakes in order to reduce the difference between the reference speed and the actual speed of the vehicle. This approach considers the physical properties of each component of the powertrain and the dynamic interaction between them. The drawback with this modeling approach is the high computational burden required to solve the set of differential equations presented in the model. Reversing the direction of action of the simulation offers the opportunity to evade the complex solution of differential equations. For this purpose, it is assumed that the electric vehicle follows exactly the reference speed. The direction of action of the backward-facing approach is highlighted in black in figure 3.1. This approach is computationally efficient since it computes the forces acting on the wheels and processes them backwards through the powertrain. The computation of the power demand depends only on algebraic equations, decreasing in this manner the computational effort of the model. This approach has the disadvantage that the accuracy of the simulation is low in contrast to the forward-facing counterpart. Despite this, the accuracy achieved by the backward-facing approach while determining the power consumption of the experimental vehicle is sufficient. In the forthcoming explanation we omit expressing all variables of the model as time dependent, since the model presented is described by a set of algebraic equations and thus can be evaluated at any arbitrary time instant.

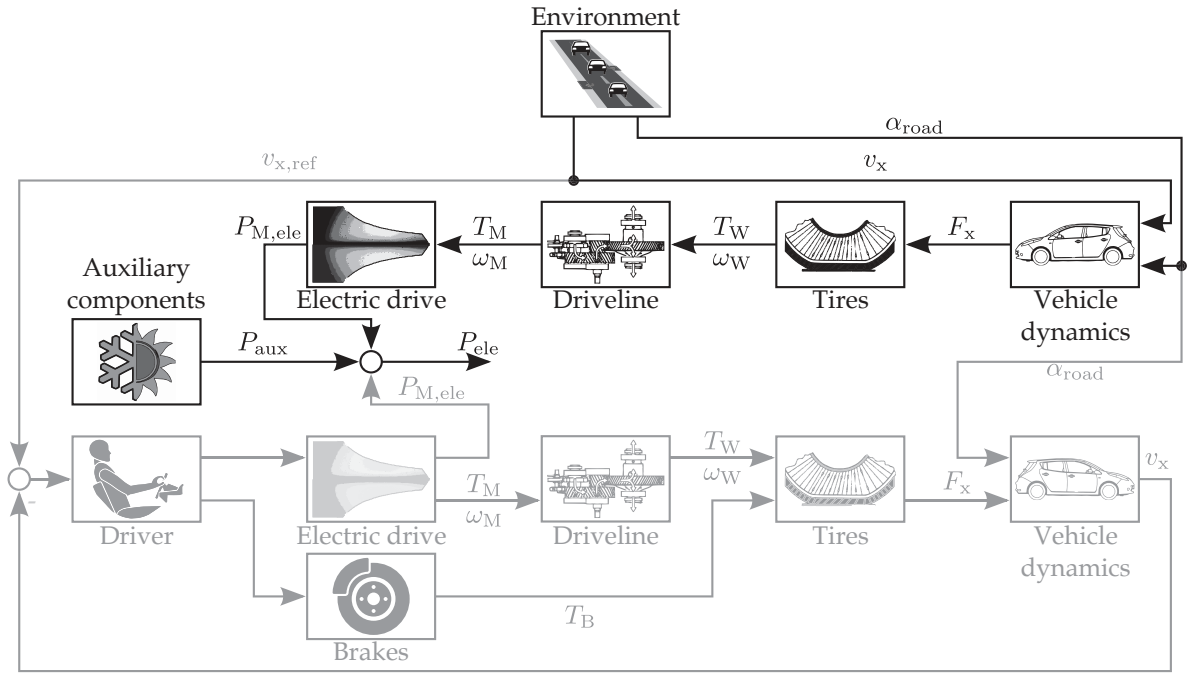

Figure 3.1.: Direction of action of the forward-facing (gray) and the backward-facing (black) modeling approaches.

3.2. Vehicle dynamics model

Both longitudinal, lateral and vertical dynamics are directly coupled. However, if it is assumed that the vehicle drives under the influence of small steering angles, these three components can be accordingly decoupled (El Majdoub et al. 2012). This assumption allows us to allocate all forces acting on the vehicle along its longitudinal axis. The dynamics model derived in this work accounts only for the longitudinal and vertical components. Cornering effects as well as stability issues are not considered as relevant for computing the power consumption and therefore the lateral dynamics of the electric vehicle is ignored in the following.

Longitudinal dynamics

To derive the model let us apply Newton's Second Law along the longitudinal axis of the vehicle, as shown in figure 3.2,:

$$\sum F_{\text{x}} = F_{\text{x}} - F_{\text{A}} - F_{\text{R}} - F_{\text{G}} - F_{\text{I,Rot}} = m_{\text{V,t}} a_{\text{x}} \tag{3.2.1}$$

where $m_{\text{V,t}}$ is the total mass of the electric vehicle (including both curb and cargo weights), a_{x} and F_{x} are the longitudinal component of the acceleration and of the traction force, respectively. The longitudinal aerodynamic drag force F_{A} results from the relative motion of the airflow of around the vehicle while it travels and is expressed as:

$$F_{\text{A}} = \frac{1}{2} \rho_{\text{air}} c_{\text{w}} A v_{\infty}^2, \tag{3.2.2}$$

where v_{∞} is the wind speed, ρ_{air} is the density of air, c_{w} is an experimentally determined aerodynamic drag coefficient and A is the frontal area of the vehicle.

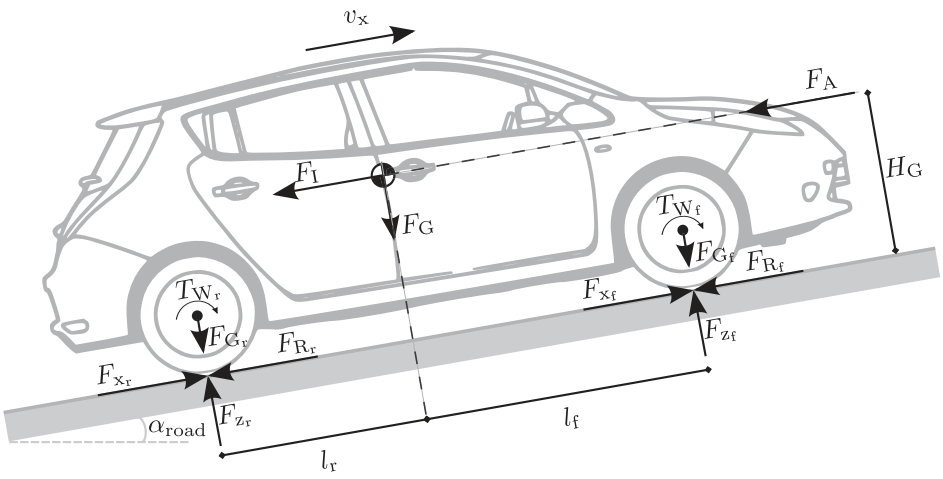

Figure 3.2.: Forces acting on the electric vehicle during its longitudinal motion on a sloped road.

For a vehicle driving in a straight line and on a dry pavement 90%-95% of the rolling resistance F_R is the result from internal hysteresis losses on the tire materials, 2%-10% is due to the friction between the tires and the road and 1%-3.5% is caused by air resistances (Wong 2001, p. 9). Under the assumption that the vehicle drives on a road with small gradient the total rolling resistance can be computed as:

$$F_R = m_{V,t} g K_R, \tag{3.2.3}$$

where g is the acceleration due to gravity and K_R is the rolling resistance coefficient. Heißing and Ersoy (2011, p. 39) shows a dependence of K_R on v_x and suggests that it can be approximated by:

$$K_R = K_{R_0} + K_{R_1}\left(\frac{v_x}{100\text{km/h}}\right) + K_{R_4}\left(\frac{v_x}{100\text{km/h}}\right)^4. \tag{3.2.4}$$

While driving on a sloped road, the sine component of the weight of the vehicle acts on its center of gravity (CoG) along the longitudinal axis (see figure 3.2), either braking it or accelerating it, depending on if the vehicle drives uphill or downhill. The climbing resistance F_G is expressed as:

$$F_G = m_{V,t} g \sin(\alpha_{road}), \tag{3.2.5}$$

where α_{road} is the angle of inclination of the road. In addition to the *steady-state* resistances, which occur at $a_x = 0$, the vehicle experiences a dynamic driving resistance which is attributed to its inertia. The inertial resistance F_I arises during transient motion when the vehicle accelerates or decelerates and is expressed as:

$$F_I = F_{I,Tra} + F_{I,Rot}, \tag{3.2.6}$$

where $F_{I,Tra} = m_{V,t} a_x$ is the translational component, which arises from the transient motion of the mass $m_{V,t}$ and $F_{I,Rot}$ is the rotational component which result from the acceleration or deceleration of the rotating parts of the powertrain, e.g., wheels, transmission and shafts.

Force distribution by acceleration and braking maneuvers

As it was introduced in section 3.1, the power consumption model presented in this work is based on the backward-facing approach. Consequently, the output of the dynamic model is the traction force required to propel the vehicle forward according to the predefined longitudinal speed. Solving (3.2.1) for F_X gives:

$$F_x = m_{V,t} a_x + F_A + F_G + F_{I,Rot}. \tag{3.2.7}$$

In our model F_x is composed of F_{x_f} and F_{x_r}, which represent the traction force on the front and rear wheels, respectively. To better account for the mechanical power demanded to move the vehicle , it becomes necessary to analyse the distribution of F_x between F_{x_f} and F_{x_r} under both acceleration and braking maneuvers.

As it was already mentioned in section 2.3, the vehicle studied in this work is configured with a front-wheel drive (FWD). This configuration causes the rear wheels to be carried along during acceleration maneuvers and when the vehicle travels at constant speed. The force distribution can be accordingly expressed as:

$$F_{x_r} = -F_{R_r} - F_{I,Rot_r}, \tag{3.2.8}$$

$$F_{x_f} = m_{V,t} a_x + F_A + F_G - F_{x_r}, \tag{3.2.9}$$

where F_{R_r} and F_{I,Rot_r} account for the rolling resistance and the rotational inertial resistance caused by the rear wheels.

During braking maneuvers we assume that the brake-force is optimally distributed. According to Mutoh et al. (2007) a proper force distribution can be obtained if we consider the normal load on the front and rear axles, i.e., F_{z_f} and F_{z_r}. By solving the set of equations obtained from the moment balance around the front and rear tire contact point (see figure 3.2), the brake-force distribution ratios of the front R_f and rear R_r wheels are computed by:

$$R_f = \frac{l_r + \frac{-a_x}{g} H_G}{l_r + l_f}, \tag{3.2.10}$$

$$R_r = 1 - R_f, \tag{3.2.11}$$

where l_f and l_r represent the distance from the CoG to the front and rear axles and H_G is the height of vehicle CoG measured from the ground. The ratio $\frac{-a_x}{g}$ is an estimate of the load movement, i.e., the amount of load that is shifted from the rear axle to the front axle while the vehicle brakes.

Analogous to equations (3.2.8) and (3.2.9), the braking forces for the rear and front wheels are expressed as:

$$F_{x_r} = R_r(m_{V,t} a_x + F_A + F_G), \tag{3.2.12}$$

$$F_{x_f} = R_f(m_{V,t} a_x + F_A + F_G). \tag{3.2.13}$$

It is noteworthy that the influence of the rolling resistance F_{R_f} and the rotational inertial resistance F_{I,Rot_f} is not taken into account in the computation of F_{x_f} with equation (3.2.9). Analogously, equations (3.2.12) and (3.2.13) also neglect these resistances while computing F_{x_r} and F_{x_f}. This is due to the fact that in these three cases the wheels are actively actuated, either for driving or for braking. We therefore consider these driving resistances as part of the tire model presented in an upcoming section.

Vertical dynamics

The vertical forces acting on the rear and the front axle consist of a static, a dynamic, an aerodynamic and a gradient component. The static components are determined by the configuration of the vehicle's body and depend solely on l_f and l_r. Acceleration and deceleration maneuvers cause a dynamic load shift which influences the vertical forces. The aerodynamic component also modifies the static components since it causes a force, which can either tend to *lift up* the vehicle from the ground or to *push it down* against it. Analogous to the previous discussion on driving resistances, the gradient of a sloped road also modifies the static component of the vertical force given by the weight of the electric vehicle.

In this work the vertical forces are determined by means of two moment balances around the tire contact point of both rear and front tires, respectively. Accordingly:

$$F_{z_\mathrm{r}} = \frac{-F_\mathrm{A} H_\mathrm{G} - m_{\mathrm{V,t}} a_\mathrm{x} H_\mathrm{G} - m_{\mathrm{V,t}} \mathrm{g} H_\mathrm{G} \sin \alpha_\mathrm{road} + m_{\mathrm{V,t}} \mathrm{g} l_\mathrm{r} \cos \alpha_\mathrm{road}}{l_\mathrm{f} + l_\mathrm{r}} \tag{3.2.14}$$

$$F_{z_\mathrm{f}} = \frac{F_\mathrm{A} H_\mathrm{G} + m_{\mathrm{V,t}} a_\mathrm{x} H_\mathrm{G} + m_{\mathrm{V,t}} \mathrm{g} H_\mathrm{G} \sin \alpha_\mathrm{road} + m_{\mathrm{V,t}} \mathrm{g} l_\mathrm{f} \cos \alpha_\mathrm{road}}{l_\mathrm{f} + l_\mathrm{r}}. \tag{3.2.15}$$

Since the model presented in this work concentrates all forces acting on the vehicle along the longitudinal axis, equations (3.2.14) and (3.2.14) are valid under the assumption that the vehicle's roll moment is neglected.

3.3. Powertrain modeling

Tire model

When a pneumatic tire is subject to a driving torque, either positive or negative, the traction force $F_{\mathrm{x_W}}$ arises at the contact point between the tire and the road. If the driving torque is increased such that the maximum transmission force is exceeded, the tire loses traction and begins to slip. The longitudinal traction force $F_{\mathrm{x_W}}$ transmitted by the tire is described with help of the longitudinal tire slip κ_x, which is defined as:

$$\kappa_\mathrm{x} = \begin{cases} \frac{r_\mathrm{W}\omega_\mathrm{W} - v_\mathrm{x}}{r_\mathrm{W}\omega_\mathrm{W}}, & \text{driving} \\ \frac{v_\mathrm{x} - r_\mathrm{W}\omega_\mathrm{W}}{v_\mathrm{x}}, & \text{braking} \end{cases} \tag{3.3.1}$$

where v_W and ω_W are the translational and rotational rotational speed of the wheel, respectively, and r_W represents the effective rolling-radius of the tire. The traction force of a tire $F_{\mathrm{x_W}}$ depends on the vertical load applied to the tire. This relationship is usually expressed as:

$$\mu_\mathrm{x}(\kappa_\mathrm{x}) = \frac{F_{\mathrm{x_W}}}{F_{z_\mathrm{W}}}, \tag{3.3.2}$$

where $\mu_\mathrm{x}(\kappa_\mathrm{x})$ is the longitudinal coefficient of friction. In the spirit of explaining the nonlinear behavior of $\mu_\mathrm{x}(\kappa_\mathrm{x})$ this work employs the well known semi-empirical model first introduced by Bakker, Nyborg, and Pacejka (1987). The so called *Magic Formula* allows describing the steady-state behaviour of the traction force of a tire subject to pure longitudinal slip by means of:

$$F_{\mathrm{x_W}} = D_\mathrm{x} \sin\left[C_\mathrm{x} \arctan\left\{\mathrm{B_x}\kappa_\mathrm{x} - E_\mathrm{x}\left(B_\mathrm{x}\kappa_\mathrm{x} - \arctan B_\mathrm{x}\kappa_\mathrm{x}\right)\right\}\right]. \tag{3.3.3}$$

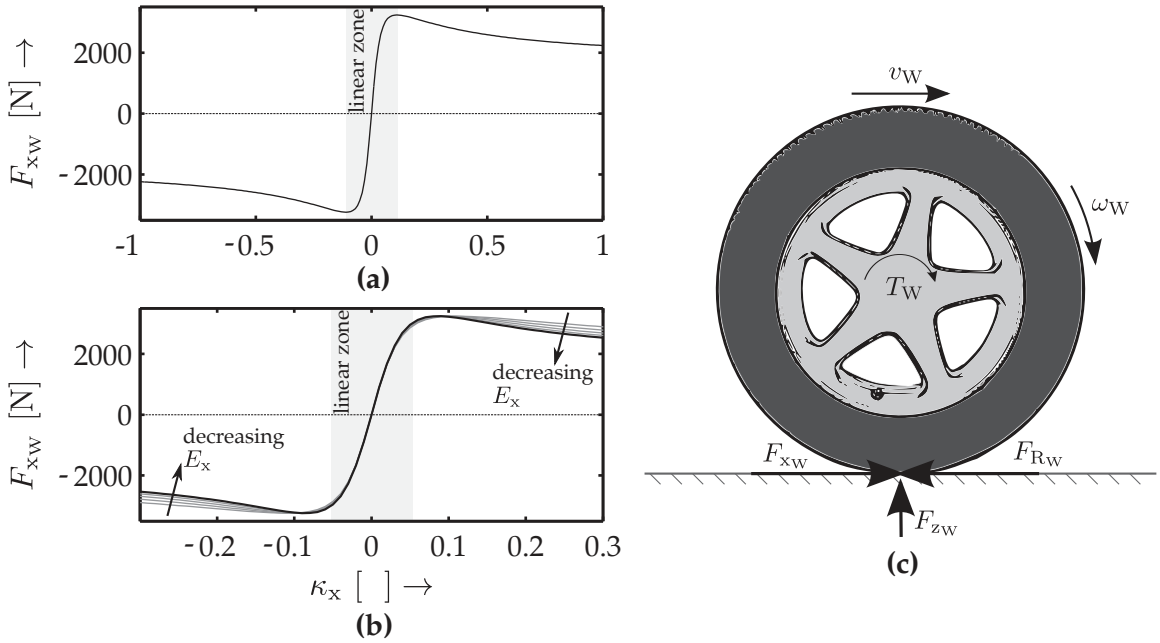

Figure 3.3.: Tire longitudinal force as function of the slip (**a**), by variation of E_x (**b**). Schematic representation of the torque balance around the rotational axis of the wheel (**c**).

Figure 3.3 (**a**) exemplifies the behaviour of F_{x_W} as a function of κ_x. The coefficient B_x determines the slope of the curve at the origin (stiffness factor), C_x is a shape factor, D_x represents the maximum transmittable longitudinal force and E_x is a curvature factor (Pacejka 2005). The tire model under the backward-facing approach pursues to determine the rotational speed of the wheels ω_W and the torque T_W as the result of the longitudinal force computed by equations (3.2.8), (3.2.9), (3.2.12) and (3.2.13). Employing equation (3.3.3) to this aim, however, might be prohibitive since it would require to numerically solve it for κ_x and would therefore diminish the computational benefits presented by the backward-facing approach. One workaround is to neglect the curvature factor, i.e., to assume $E_x = 0$. This assumption is justified by analyzing the behavior of F_{x_W} at different values of E_x (see figure 3.3 (**b**)). As it can be seen, the effect of E_x on the *linear zone* ($-0.05 \leq \kappa_x \leq 0.05$) is relatively small. For the aim of this work this assumption is acceptable, since the driving maneuvers studied here don't incur into the physical limits of the tires, that is, into the nonlinear zone. Thus, solving equation (3.3.3) w.r.t. κ_x yields:

$$\kappa_x = \frac{\tan\left(\frac{\arcsin\left(\frac{F_{x_W}}{D_x}\right)}{C_x}\right)}{B_x}. \tag{3.3.4}$$

After having computed κ_x, ω_W is easily determined by solving equation 3.3.1. As shown in figure 3.3 (**c**)), T_W is computed by performing a torque balance around the rotational axis of the wheel:

$$T_W = I_W \dot{\omega}_W + F_{R_W} r_W + F_{x_W} r_W, \tag{3.3.5}$$

where I_W represents the rotational inertia of the wheel.

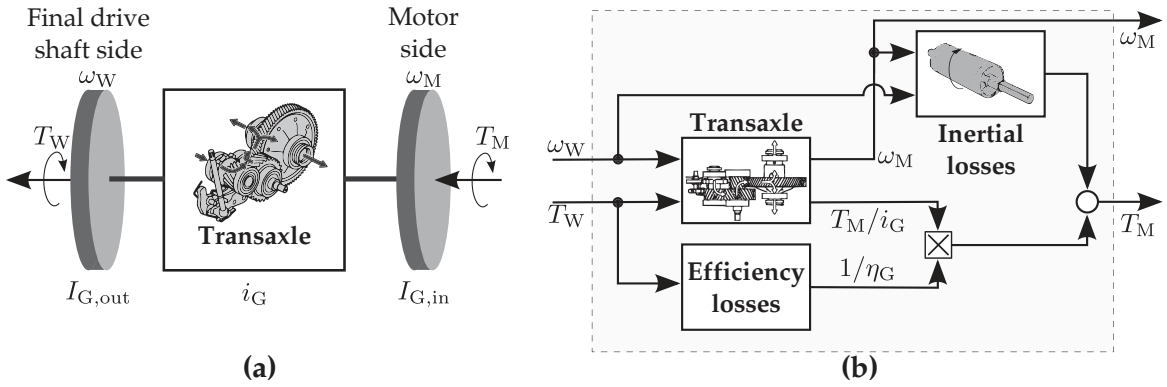

Figure 3.4.: Overview of the transaxle (**a**) and block diagram of the driveline model (**b**).

Driveline model

The experimental vehicle is equipped with a transaxle, i.e., a driveline which combines the gear box and the differential into an integrated assembly and that is direct coupled to the rotor shaft of the electric drive, on the one side, and to the final drive shaft of the front wheels, on the other side (see figure 3.4 (**a**)). The driveline model transforms the rotational speed and the torque demanded by the wheels, that is, ω_W and T_W, into the rotational speed ω_M and torque T_M demanded by the electric drive. ω_M is computed by:

$$\omega_M = \omega_W i_G, \tag{3.3.6}$$

where i_G represents the overall gear ratio of the transaxle.

As it is shown in figure 3.4 (**b**), computing T_M requires knowledge about the efficiency of the transaxle and about the inertia of its rotating components. The efficiency of the driveline η_G with respect to the transmission of power is assumed in this work as constant and is conveniently expressed as:

$$\eta_G(T_M) = \begin{cases} \leq 1, & T_M \geq 0 \\ > 1, & T_M < 0 \end{cases} \tag{3.3.7}$$

such that it is positive if the vehicle is accelerating and negative if a braking maneuver is taking place. The total rotational inertia of the driveline is composed by the inertia of all rotating components on the output side of the transaxle ($I_{G,out}$) and by the inertia of the rotating components on the input side ($I_{G,in}$) including the rotor shaft of the electric drive. Accordingly, T_M is computed by:

$$T_M = \frac{T_W}{i_G \eta_G(T_W)} + (I_{G,out})\,\dot{\omega}_W + (I_{G,in})\,\dot{\omega}_M. \tag{3.3.8}$$

Electric drive model

Figure 3.5 (**a**) depicts a block diagram of the model employed to describe the electric drive. Under the backward-facing philosophy both the torque required from the motor T_M and the rotational speed of the rotor ω_M are used to compute the electrical power $P_{M,ele}$ of electric drive. It can take both positive or negative values, depending on whether the electric drive is operated in motor or in generator mode. The

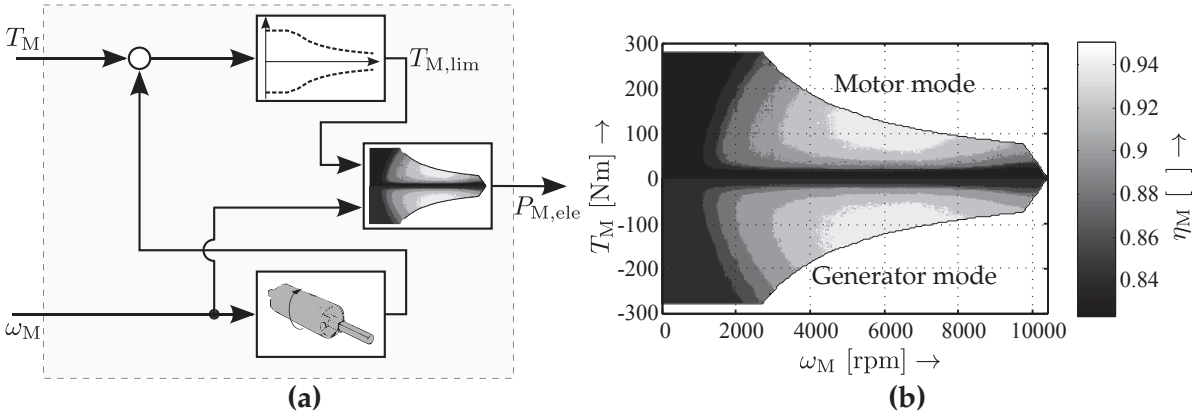

Figure 3.5.: Block diagram of the electric drive model (**a**) and efficiency map used to compute power losses (**b**).

mechanical power $P_{M,mec}$, computed at the rotor of the electric drive, is given by:

$$P_{M,mec} = T_M \omega_M. \tag{3.3.9}$$

As suggested by Guzzella and Sciarretta (2005), the relationship between the mechanical and the electrical power demand of an electric motor can be computed, with a certain degree of accuracy, by employing a stationary map η_M of the electric motor's efficiency as a function of the rotor's rotational speed and the torque demand:

$$P_{M,ele} = \frac{P_{M,mec}}{\eta_M(\omega_M, T_M)}, P_{M,mec} > 0. \tag{3.3.10}$$

Determining $\eta_M(\omega_M, T_M)$ requires computing the power losses of the electric motor for every operating point on an engine dynamometer. The efficiency map comprises the power losses due to the electric motor, the inverter and the AC cables connecting the system.

The efficiency map of the electric vehicle investigated in this work is well defined only for the region operating in motor mode (upper quadrant of figure 3.5 (**b**)). In order to extend the map for the operation in generator mode, Guzzella and Sciarretta (2005) suggest mirroring the power losses as follows:

$$\eta_M(\omega_M, -T_M) = 2 - \frac{1}{\eta_M(\omega_M, T_M)}. \tag{3.3.11}$$

Even though the computed efficiency map obtained by applying equation (3.3.11) slightly differs from the data that might be obtained by measuring the efficiency of the electric motor working as generator, it offers a practical and accurate solution for modeling the electric motor also in generator mode.

One important feature of the experimental vehicle is that certain amount of the kinetic and the potential energy is recovered by means of the **regenerative braking** system. During braking maneuvers the electric motor is operated as a generator, providing in this manner an extra braking torque to the wheels. The recovered energy can then be used to supply power either to the powertrain or to the auxiliary accessories. The amount of braking torque depends on the operation strategy of the

braking system. The operation strategy optimizes the distribution of braking torque between the mechanical and the regenerative brakes in such a way, that the maximum electrical power is generated. The electrical power generated is computed by:

$$P_{\mathrm{M,ele}} = P_{\mathrm{M,mec}} \eta_{\mathrm{M}} \left(\omega_{\mathrm{M}}, -T_{\mathrm{M}}\right) k_{v_{\mathrm{x}}}, P_{\mathrm{M,mec}} < 0. \tag{3.3.12}$$

Since the generated power depends on ω_{M}, it would be a difficult task to supply power to the power bus at low speeds. Because of this, the parameter $k_{v_{\mathrm{x}}}$ is used to limit the usage of the electric motor in generator mode according to equation (3.3.13), so that the mechanical brakes are applied at very low speeds and at high speeds the vehicle is braked mostly by the electric motor. Accordingly, $k_{v_{\mathrm{x}}}$ is given by:

$$k_{v_{\mathrm{x}}} = \begin{cases} 0 & v_{\mathrm{x}} \leq v_{\mathrm{x,min}}^{\mathrm{reg}} \ \mathrm{m/s} \\ \frac{v_{\mathrm{x}} - v_{\mathrm{x,min}}^{\mathrm{reg}}}{v_{\mathrm{x,max}}^{\mathrm{reg}} - v_{\mathrm{x,min}}^{\mathrm{reg}}} & v_{\mathrm{x,min}}^{\mathrm{reg}} < v_{\mathrm{x}} < v_{\mathrm{x,max}}^{\mathrm{reg}} \ \mathrm{m/s} \\ 1 & v_{\mathrm{x}} \geq v_{\mathrm{x,max}}^{\mathrm{reg}} \ \mathrm{m/s} \end{cases} \tag{3.3.13}$$

where $v_{\mathrm{x,min}}^{\mathrm{reg}}$ and $v_{\mathrm{x,min}}^{\mathrm{reg}}$ parametrize the regenerative braking strategy.

3.4. Auxiliary components

Contrary to conventional vehicles, the power needed to run all auxiliary components in an electric vehicle is supplied by the main energy storage system, e.g., the battery pack. Depending on their functionality, auxiliary consumers are grouped in primary and secondary components. To the primary group belong those auxiliaries, which are necessary for the proper operation of the vehicle, e.g., low voltage consumers such as sensors or control devices and high voltage consumers such as the electric power steering or the power-assisted brakes. In contrast to most electric vehicles, the experimental vehicle relies on an air-cooled system for cooling the energy storage system and the power electronics. Thus, no extra power is required to this aim.

The secondary group includes those auxiliary components responsible of lighting, safety and comfort functions. Windshield wipers, headlamps or the telematics system are some of the auxiliaries belonging to this group. Special attention is paid in this work to the heating, ventilation and air conditioning system (HVAC), since it enormously influences the total energy consumed by the vehicle and therefore imposes large constraints to its autonomy.

Even though detailed models for computing the energy consumption of HVAC systems have been published (Valentina et al. 2014), they requires knowledge about the thermodynamic properties of the elements composing the HVAC system, e.g., the heater, the compressor or the evaporator, which in the case of the experimental vehicle, are not available. The identification of such parameters would be out the scope of this work. We instead determine the energy consumed by the HVAC system by means of observations made during experiments carried out under controlled temperature conditions. The total power required by the auxiliary components is given by:

$$P_{\mathrm{aux}} = \sum P_{\mathrm{1ry}} + \sum P_{\mathrm{2ry}}, \tag{3.4.1}$$

where $\sum P_{\mathrm{1ry}}$ and $\sum P_{\mathrm{2ry}}$ stand for the sum of the power required to run both primary and secondary auxiliary components, respectively.

3.5. Model parametrization and evaluation

This section discusses the procedure for setting the nominal values of $\mathcal{M}_{PC}$ and the validation results. For the sake of brevity, we limit the forthcoming discussion to those parameters that have been identified by experimental means, i.e., c_w, K_{R_0}, K_{R_1}, K_{R_4}, r_W, $I_{G,out}$, $I_{G,in}$, $v^{reg}_{x,min}$, $v^{reg}_{x,max}$ and P_{aux}. A discussion on all (*known*) parameters is found in section A.1. The reader is referred to table A.1, which shows the nominal values of all parameters of $\mathcal{M}_{PC}$ used in this work.

The nominal value of the air drag coefficient (c_w) is publicly disclosed. However, the Nissan Leaf has been equipped with a GPS antenna installed on the roof, which causes c_w to increase compared to the nominal value. The nominal tire effective rolling radius (r_W) of the tires is likewise available from public data. Nevertheless, factors such as the inflation pressure or the speed might cause the value of r_W to be found around 92% to 94% of its nominal value. For this reason, we decide to identify it instead of using the standard nominal value. As shown in section 3.2, the coefficient of rolling resistance K_R depends on v_x and is approximated by a 4^{th}-order polynomial (recall equation (3.2.4)) parametrized by K_{R_0}, K_{R_1}, K_{R_4}. Since these values are dictated by the tire-road interaction, which depend among others on the road conditions, they need to be identified. Similarly, the inertia of all rotational components, i.e, $I_{G,out}$ and $I_{G,in}$, is unknown and therefore, has to be identified. Also, as shown if equation (3.3.13), the Nissan Leaf is assumed to apply a regenerative braking strategy. Therefore, to validate this assumption we need to identify values for both parameters $v^{reg}_{x,min}$ and $v^{reg}_{x,max}$. Finally, it has been observed that the Nissan Leaf consumes certain power when all auxiliaries are turned-off. Thus, this *residual power consumption* (P_{aux}) is also considered as part of the identified parameters.

The identification of parameters have been performed by means of a cross-validation based on experimental measurements collected from trips done with the Nissan Leaf through different driving scenarios and under various operating conditions. From these measurements, a total of 13 datasets have been utilized in the context of this chapter. Each dataset consists of measurements of speed (v_x), road slope (α_{road}) and power requirements (P_{ele}) gathered with a sampling rate of 1 Hz. The time series of the v_x and α_{road} can be appreciated in figures B.1, B.2 and B.3. Table B.1 summarizes useful information about the datasets such as the traveled distance, height difference between the start point and the end point of the trip or the number of passengers that occupied the vehicle during each trip. The cross-validation starts by identifying 13 sets of unknown parameters, one for each dataset independently, by minimizing the error (in the square sense) between the real power requirements and the power computed by $\mathcal{M}_{PC}$. Afterwards, each set is cross-validated against all others datasets, thus avoiding a possible over-fitting of the model. The selection criteria of the *best* parameter set is based on the mean of the root mean square error (RMSE) computed by applying each set of parameters among all datasets. We favor this procedure, instead of applying each parameter set against a joined dataset, because, as can be seen in table B.1, not all measurements are gathered under the same road conditions and with the same number of passengers. Figures 3.6 (**a**) and (**b**) qualitatively and quantitatively show that the power required by the Nissan Leaf is properly modeled by applying the 13^{th} parameter set to its corresponding dataset.

Table 3.1.: Mean RMSE from the evaluation of each set of parameters among all data sets.

Dataset	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
RMSE [kW]	0.84	0.58	1.31	0.78	0.99	0.63	0.53	0.53	0.53	0.69	0.70	0.64	0.52

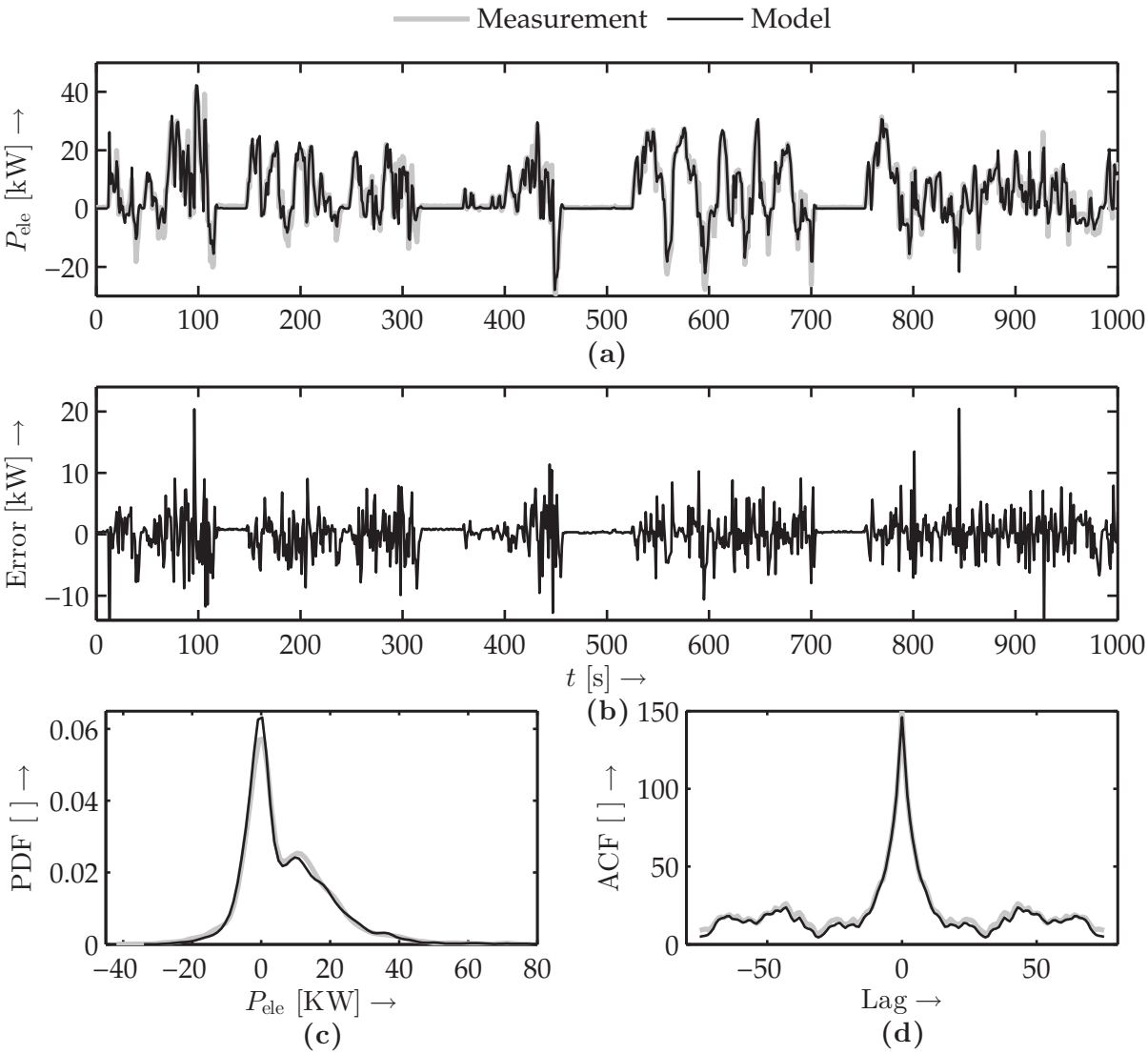

Figure 3.6.: Qualitative (**a**) and quantitative (error) (**b**) validation result obtained by applying the best set of identified parameters. Probability density function (PDF) (**c**) and auto-covariance function (ACF) (**d**) of the measured and simulated power demand.

Furthermore, figure 3.6 (**c**) depicts the probability distribution of the power requirements. As it can be seen, $\mathcal{M}_{PC}$ accurately models the power demand, specially important being the region with negative values, i.e., the distribution of power recovered through the regenerative braking system. The auto-covariance function (see figure 3.6 (**d**)) confirms that the power requirements of the EV can be modeled using a $\mathcal{M}_{PC}$ by using the set of identified parameters and the measured v_x and α_{road} as the input of the model. For the sake of completeness, figures C.1, C.2 and C.3 show validation results by applying the chosen set of parameters on all other datasets.

4

Energy Storage System Modeling

As it was detailed in chapter 3, time-varying operating conditions may cause abrupt changes in the power required by the electric vehicle both in magnitude and in direction. For this reason it is desired to have a model that accounts not only for constant or shock-like loads but also for highly dynamic loads. This requirement becomes even more important since the electrical energy storage system introduced in section 2.3 is recharged thanks to the regenerative braking system of the electric vehicle (see section 3.3). The operating temperature is also an essential factor that is taken into account. As it is demonstrated in the forthcoming sections, the capacity and the transient response of the electrical energy storage system is drastically affected at low temperatures. This chapter introduces the derivation, parametrization and validation of a physics-based model that accurately captures the non-linear capacity behavior exhibited by a single cell of the electrical energy storage system used to power the electric vehicle.

4.1. Single-cell modeling

The first part of this chapter deals with the single cell model (see figure 4.1) introduced by Weihrauch, Oliva, and Bertram (2013) which combines the so called *Kinetic Battery Model* (KiBaM), first introduced by Manwell and McGowan (1994), with a second order equivalent circuit-based model.

The KiBaM describes the reaction of the state of charge (SoC) of a cell to an applied load I_{batt}. It is an improvement of the conventional coulomb counting procedure, since it takes into account well known non-linear capacity effects (see section 4.2). Even though the KiBaM was initially developed for lead acid batteries, it has been shown to be suitable for modeling the capacity behavior of Li-ion cells (Jongerden and Haverkort 2009).

The transient response of the Li-ion cell is described by the ohmic resistance R_{ohm} which captures the instantaneous voltage drop due to a step load current event and by a two-RC network, i.e., $R_{\text{s}}C_{\text{s}}$ and $R_{\text{l}}C_{\text{l}}$, which capture the voltage drops due to the activation and the concentration polarization, respectively (Linden and Reddy 2002, p. 2.1). As it is demonstrated in section 4.4, the parameters of the equivalent circuit-based model exhibit dependencies on the cell temperature (T_{batt}) and on the SoC. For the sake of clarity, these dependencies are expressed in figure 4.1 by the term $(\cdot)$.

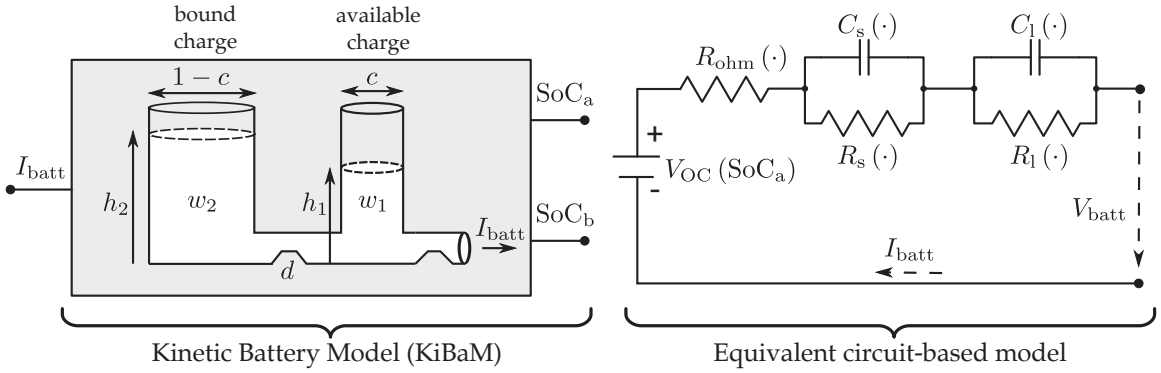

Figure 4.1.: Combined model of single Lithium-ion cell.

Despite the apparent simplicity of the model, it has many advantages in contrast to its electrochemical-based counterpart. Electrochemical cell models stands out in the literature due to their high accuracy, which comes from the detailed description of the chemical processes occurring inside the battery cell. The most important criteria from the point of view of this work in regard of using electrochemical models is that they are characterized by large sets of coupled partial differential equations and therefore require large periods of time, e.g., in the range of hours or even days, in order to simulate the behavior of the cell as response to a charge-discharge cycle (Rao, Vrudhula, and Rakhmatov 2003).

The combined model shown in figure 4.1 has the advantage that it is described through few parameters which can be conveniently identified from simple experiments. Hence, the model possesses a simple structure, exhibits high accuracy and demands low computational effort, which makes it suitable for real-time applications. In this context, the main advantage of the model is given by the flexibility of the second order equivalent circuit. Due to its expandability properties, it can be successively extended in order to increase its accuracy either by using empirical equations to describe the parameters or by adding more electrical components to it such that more effects can be accounted for. Both models are coupled by means of V_{OC} which represents the open circuit voltage of the cell. The reader is referred to the work of Seaman, Dao, and McPhee (2014) for a comprehensive survey on equivalent circuit-based and electrochemical cell models.

4.2. Cell capacity

As it was previously formulated in section 2.1, properly estimating the RUL of the mobile system, i.e., the RDR of the electric vehicle, requires precise knowledge of the energy available in the energy storage system, in this case the Lithium-ion cell. As it was presented in the previous section, the Kinetic Battery Model is used in this work with the aim of describing the load-dependent non-linear behavior of the cell capacity. Under ideal conditions, a discharge event would cause the cell capacity to remain constant independently of the discharge rate. However, in reality the actual capacity of a battery cell decreases proportionally to the rate at which it is discharged. This phenomenon is referred in the literature to as *rate capacity effect*. In addition, after

liberating the battery from a demanding discharge event, certain chemical reactions take place, which lead to an increment of the cell voltage. This process is termed as *recovery effect* (Jongerden and Haverkort 2008).

The Kinetic Battery Model abstracts the chemical processes of the cell discharge to its kinetic properties. The model assumes that the total charge of the cell is distributed with a capacity ratio $0 < c < 1$ between two charge wells (see the left side of figure 4.1). The first well contains the available charge, and delivers it directly to the load I_{batt}, whereas the second well supplies charge only to the first well. The rate of charge that flows from the second to the first well depends on both the parameter d, also referred to as conductance, and on the height difference $h_2 - h_1$, with $h_1 = \frac{w_1}{c}$ and $h_2 = \frac{w_2}{1-c}$. If the first well is empty, then the cell is considered to be fully discharged. By applying load to the cell, the charge in the first well is reduced, which leads to an increment in the height difference between both wells. After removing the load, certain amount of charge flows from the second well to the first well until the height of both wells is the same. In this way the recovery effect is taken into account by the model. The rate capacity effect is also considered. For high discharge currents, the charge in the first well is delivered faster to the load in comparison to the charge that flows from the second well. In this scenario the effective cell capacity is less than the nominal capacity since there is an amount of charge that remains unused (bounded). The consideration of this effect is especially important for applications in electric vehicles, since the unused charge might eventually increase the driving range.

The discrete-time KiBaM yields two difference equations which describe the change of capacity in both wells in dependence of the load $I_{\text{batt},k}$, the conductance d and the capacity ratio c:

$$w_{1,k+1} = a_1 w_{1,k} + a_2 w_{2,k} + b_1 I_{\text{batt},k}, \tag{4.2.1}$$

$$w_{2,k+1} = a_3 w_{1,k} + a_4 w_{2,k} + b_2 I_{\text{batt},k}, \tag{4.2.2}$$

with discrete system parameters:

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} = \mathrm{e}^{\begin{pmatrix} -\frac{d}{c} & \frac{d}{1-c} \\ \frac{d}{c} & -\frac{d}{1-c} \end{pmatrix} \Delta t}$$

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \int_{\tau=0}^{\Delta t} \mathrm{e}^{\begin{pmatrix} -\frac{d}{c} & \frac{d}{1-c} \\ \frac{d}{c} & -\frac{d}{1-c} \end{pmatrix} \tau} \mathrm{d}\tau \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

where Δt is the sampling time used in the discretization (see section 2.2).

Both states w_1 and w_2 do not directly represent the available charge SoC_a and the bound charge SoC_b. Thus, two additional algebraic equalities are introduced with the aim of computing the outputs of the KiBaM as shown in figure 4.1, namely:

$$\text{SoC}_\text{a} = \frac{w_1}{3600 c C_n}, \tag{4.2.3}$$

$$\text{SoC}_\text{b} = (1-c)\left(\frac{w_2}{1-c} - \frac{w_1}{c}\right) \tag{4.2.4}$$

where C_n is the nominal capacity of the cell. From now on, unless otherwise specified, the variable SoC is employed instead of SoC_a.

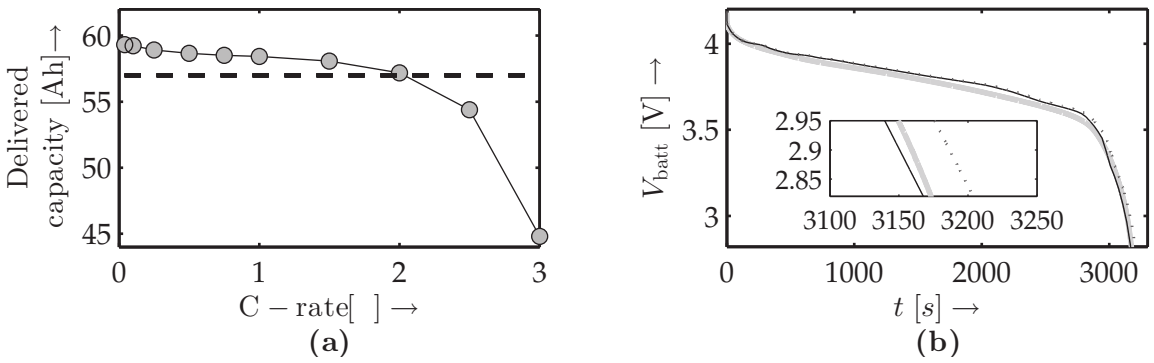

Figure 4.2.: Cell delivered capacity in dependence of the C-rate (**a**). Real discharge curve (*solid gray*) compared to Coulomb counting (*dotted black*) and the KiBa model (*solid black*) (**b**).

Parameter identification: to identify the parameters c and d of the KiBaM multiple experiments consisting in the discharging a battery cell under different C-rates have been performed with help of the second setup of section 2.3. To determine c it is first necessary to identify both $w_{1,0}$ and $w_{2,0}$, i.e., the available and the bound charge when the cell is fully charged. Therefore we assume that the overall charge w_0 is entirely drawn from both wells by applying an infinitesimal small current. Consequently we completely discharge the cell at a rate of $\frac{1}{25}$C, giving a total delivered capacity of 59.5Ah (see the first measurement in figure 4.2 (**a**)). This procedure has been repeated at $\frac{1}{10}$C, $\frac{1}{4}$C, $\frac{1}{2}$C, 1C, 1.5C, 2C, 2.5 and 3C at a temperature of 25 °C.

Figure 4.2 (**a**) shows the delivered capacity by discharging the cell at various C-rates. It can be seen, that the values under a discharge rate of 2C converge to a delivered capacity of 57 Ah, which corresponds to the maximum available capacity that can be delivered at large discharge rates, i.e., $w_{1,0} = 57\text{Ah}$. Thus, the capacity ratio can be computed by $c = \frac{w_{1,0}}{w_0} = \frac{57\text{Ah}}{59.5\text{Ah}} = 0.9580$. Note that all capacity values measured above 2C are not taken into account in determining c. The reason is that the reduced capacities at these C-rates are not due to the diffusion process but instead they is caused by transitions between steep and flat regions of the V_{OC}-SoC curve, as it is further explained in section 4.3. Previous work on the parametrization of the KiBaM, e.g., (Taesic and Wei 2011), has experimented mainly with embedded systems, such as, cell phones or personal computers. Due to the low currents drawn by such devices, even at high C-rates, the low capacity values are attributed to diffusion processes and therefore are taken into account in the computation of c. The parameter d is the last one to be identified from the entire model shown in figure 4.1. Once all parts of the model introduced in the next sections have been parametrized, the conductance d is identified by means of curve fitting. In this sense, d is set as the only optimization parameter and the terminal voltage of the model is fitted against measurements taken by discharging the cell at the aforementioned C-rates. This has lead to different d values for each identification. For this reason, from now on, the mean of all identifications, i.e., $d = 0.8 \times 10^{-5}$ is used and assumed to be constant.

The benefits of the KiBaM with aims of estimating the RDR are shown in figure 4.2 (**b**). As it can be seen Coulomb counting incorrectly quantifies the true capacity that is drawn from the cell and consequently overestimates its remaining time to discharge, which in turn directly relates to the RDR of the electric vehicle.

4.3. Open circuit voltage

The open circuit voltage (V_{OC}) represents the difference of electrical potential between the anode and the cathode of the cell, under idle loading conditions, i.e., at $I_{\mathrm{batt}} = 0$. This difference depends on the concentration of lithium in the electrodes and on the state of charge of the cell, this means: the higher the concentration of lithium in the electrodes, the higher both V_{OC} and the SoC are (Linden and Reddy 2002, p. 2.1). Therefore, having precise knowledge about V_{OC} would allow us to draw conclusions about the SoC of the cell. In an ideal case, finding the relationship between V_{OC} and SoC would require measuring the terminal voltage of the cell (V_{batt}) while totally depleting it under a infinitesimally small load, i.e., with $I_{\mathrm{batt}} \to 0$. Since in practice this procedure is not feasible, many researchers have developed methods for describing the V_{OC}-SoC relationship. With the purpose of characterizing this relationship, we employ in this work a *current-voltage based* approach (Pattipati et al. 2014), due to its practicability and rapid deployment.

To this aim we first gather the set $\mathbf{O} = \{V_{\mathrm{OC},i}, \mathrm{SoC}_i\}_{i=1}^{N_{\mathrm{o}}}$ containing N_{o} observations of the pair $\{V_{\mathrm{OC}}, \mathrm{SoC}\}$ spanned over the range $0 \leq \mathrm{SoC} \leq 1$. Contrary to conventional methods, which gather $\mathbf{O}$ by measuring the terminal voltage of the cell under steady-state conditions, that is, after large resting periods for different SoC values, we collected this data by averaging the charge and discharge curves of battery cell at low C-rates, namely $\frac{1}{25}$, $\frac{1}{10}$ and $\frac{1}{4}$ at 25 °C, (see figure 4.3 (**a**)). According to Plett (2005), averaging low C-rate charge/discharge curves allows mitigating hysteresis and polarisation effects of the terminal voltage. It has been shown that the variation of the $V_{\mathrm{OC}}-\mathrm{SoC}$ curve with respect to temperature is negligible (Lam, Bauer, and Kelder 2011), thus, we assume that the estimated curve accurately resembles the V_{OC} under steady-state conditions.

Having collected $\mathbf{O}$, the next step requires identifying the function $\mathrm{V_{OC}}(\mathrm{SoC})$ that best fits the observations. As it can be appreciated in the figure 4.3 (**b**), the shape of the V_{OC}-SoC curve obtained exhibits characteristics, which are typical of lithium-ion cells. Firstly, towards $\mathrm{SoC} = 0$ and $\mathrm{SoC} = 1$ the voltage abruptly changes due to a minimum or maximum concentration of lithium present in those regions. Secondly, in the range $0.1 \leq \mathrm{SoC} \leq 0.9$ a relative flat region can be observed. This phenomenon, also referred to as plateaus, arises since no significant change of concentration of lithium of the electrode takes place. Thirdly and lastly, multiple regions characterized by a wavy form might appear due to sudden changes in the concentration of lithium.

To accurately describe these characteristics we combine the empirical models introduced by Plett (2004) and by Lam, Bauer, and Kelder (2011) and augmented them by the addition of a sinusoidal and two exponential terms, which allows us to model the low-frequency waves shown in the regions $0.18 \leq \mathrm{SoC} \leq 0.38$ and $0.75 \leq \mathrm{SoC} \leq 0.85$. The function $\mathrm{V_{OC}}(\mathrm{SoC})$ thus results in:

$$\begin{aligned} \mathrm{V_{OC}}(\mathrm{SoC}) &= K_0 - \frac{K_1}{\mathrm{SoC}} - K_2\mathrm{SoC} + K_3 \ln(\mathrm{SoC}) + K_4 \ln(1-\mathrm{SoC}) + \\ &\quad + K_5\mathrm{SOC}^2 + K_6\mathrm{SOC}^4 + K_7 \mathrm{e}^{\left(\frac{-K_8}{1-\mathrm{SoC}}\right)} + K_9 \sin(K_{10}\mathrm{SoC}), \end{aligned} \tag{4.3.1}$$

where the parameter vector $\mathbf{K}_{\mathrm{oc}} = [\ K_0 \quad K_1 \quad \ldots \quad K_{10}\ ]^\top$ is identified by means of the Levenberg-Marquardt algorithm.

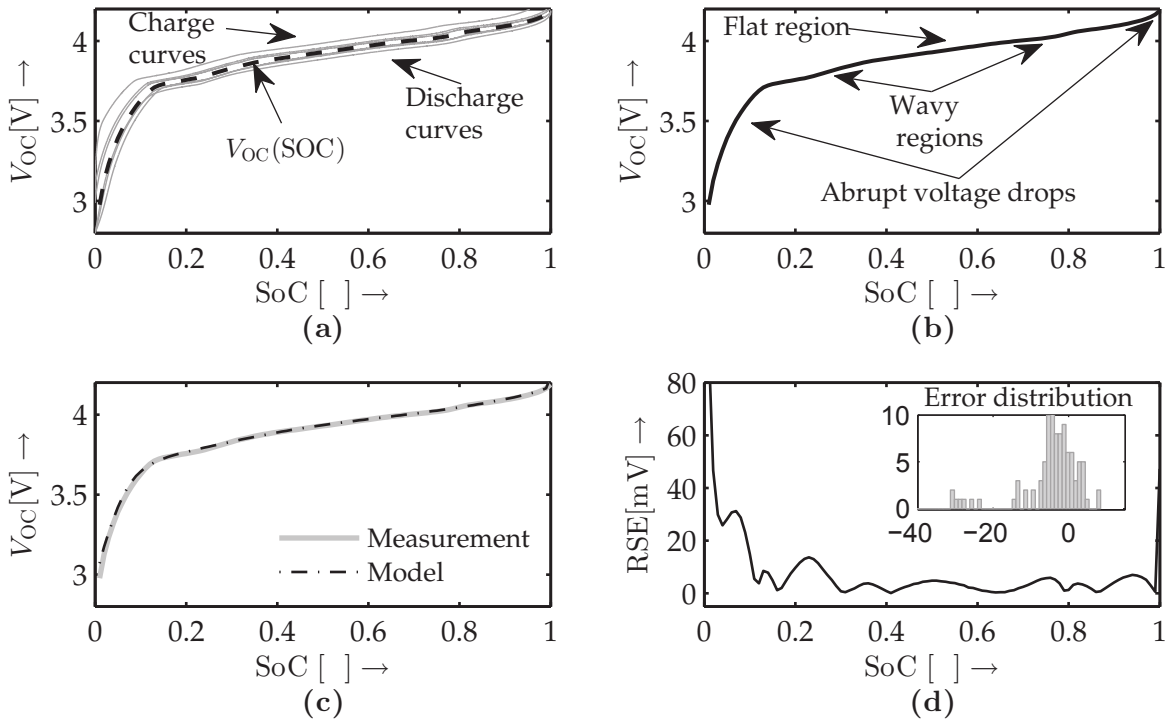

Figure 4.3.: Charge and discharge curves at different C-rates and the averaged V_{OC}-SoC (**a**), characteristics of the lithium-ion cell (**b**) and fit between the measurement and the model of the V_{OC}-SoC curve (**c**) with the corresponding root square error (RSE) (**d**).

Figure 4.3 (**c**) shows that the model relatively good fits the experimental data with root mean squared error (RMSE) of 14.56 mV. Nevertheless, as it can be appreciated in 4.3 (**d**), the model accuracy decreases for SoC ≤ 0.1, where a maximum error of 80 mV can be observed. This error is attributed to the non-smooth shape exhibited by the V_{OC}-SoC curve in this region. In order to properly capture this effect it would be needed to augment equation (4.3.1) with additional terms. To overcome this issue we quantize the V_{OC}-SoC with 100 values within this region and store them in a look-up table so that the V_{OC} can be interpolated between points.

4.4. Transient response

When a load is applied to a cell, the terminal voltage V_{batt} experiences a drop or rise, also known as overpotential, which is attributed to both ohmic and polarization effects. To better understand the overpotential let us analyze the non-linear behavior of V_{batt} as response to stepwise discharge and charge events, as it is shown in 4.4. In the interval $t < t_0$ no load is applied to the cell, thus, V_{batt} is entirely driven by $V_{OC}(\mathrm{SoC}_0)$ and SoC remains constant. By applying a negative load in the interval $t_0 < t < t_1$, V_{batt} first experiences an instantaneous voltage drop $V_{o,D}$, followed by an exponential decay, which is dictated by the V_{OC}-SoC curve. After releasing the load, in the time interval $t_1 < t < t_2$, V_{batt} suddenly rises and then experiences an exponential rise $V_{re,D}$ until it reaches $V_{OC}(\mathrm{SoC}_1)$. In the interval $t_0 < t < t_2$ the charge drawn from the cell is given by $\Delta\mathrm{SoC}_{0\to2} = \mathrm{SoC}_0 - \mathrm{SoC}_1$. Similarly, the response of V_{batt} to a

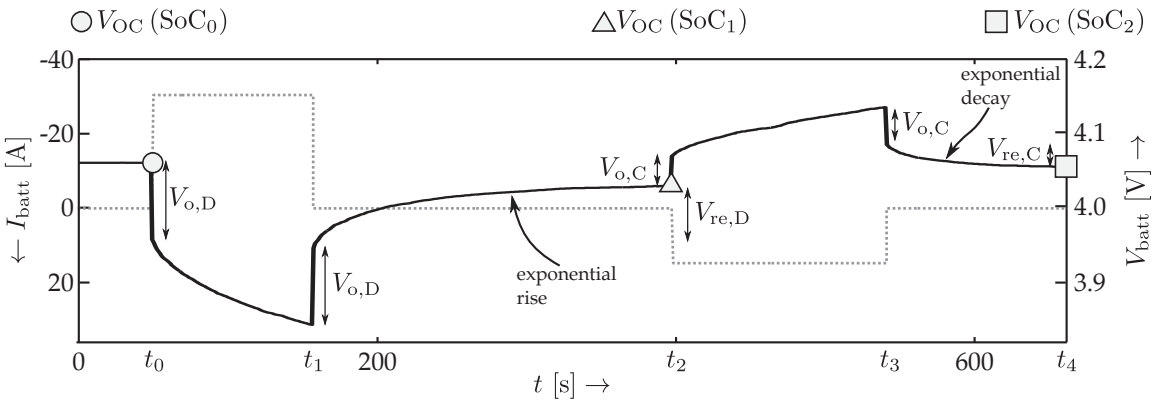

Figure 4.4.: Transient behavior of the cell terminal voltage as response to stepwise discharge and charge events.

charging event, in the time interval $t_2 < t < t_3$, starts with a steep rise $V_{o,C}$ followed by an exponential rise controlled by $V_{OC}-$SoC. Once the load is released, V_{batt} instantaneously decreases, by $V_{o,C}$, and continues decaying exponentially, by $V_{re,C}$, until it reaches $V_{OC}\,(\mathrm{SoC}_2)$. Analogously to the discharge case, the charge supplied to the cell in the interval $t_2 < t < t_4$ is given by $\Delta\mathrm{SoC}_{2\rightarrow 4} = \mathrm{SoC}_2 - \mathrm{SoC}_1$.

The instantaneous voltages drop and rise, i.e., $V_{o,D}$ and $V_{o,C}$, are result of the internal resistance of the cell, which is represented in the equivalent circuit-based model by R_{ohm}. The exponential components of the overpotential, namely $V_{re,D}$ or $V_{re,C}$, occur due to the *relaxation* phenomena of the battery cell. The term relaxation comprises the effect of both the activation and the concentration polarization on V_{batt} (see Saha and Goebel (2009) for a detailed explanation on this phenomena). The voltage drop due to activation polarization is mainly caused by the kinetic of the electrochemical reaction during a charge-transfer phase. Its dynamics is usually found in the range of seconds and is reproduced in our model by means of the R_sC_s network. The voltage drop is caused by the diffusion of ions between electrodes during a charge or discharge event exhibits a *slow* relaxation response in the range of minutes. This effect is represented herein by the R_lC_l network. The equivalent circuit model yields two difference equations which describe the transient response of the cell:

$$v_{s,k+1} = \mathrm{e}^{-\frac{\Delta t}{R_sC_s}} v_{s,k} + \left(-R_s \mathrm{e}^{-\frac{\Delta t}{R_sC_s}} + R_s\right) I_{batt,k}, \tag{4.4.1}$$

$$v_{l,k+1} = \mathrm{e}^{-\frac{\Delta t}{R_lC_l}} v_{l,k} + \left(-R_l \mathrm{e}^{-\frac{\Delta t}{R_lC_l}} + R_l\right) I_{batt,k}. \tag{4.4.2}$$

The transient response, as formulated by equations (4.4.1) and (4.4.2), is just valid for a snapshot of the entire operating range of the cell. In reality, the parameters R_{ohm}, R_s, C_s, R_l and C_l exhibit a strong dependence on the cell operating conditions, which are assumed in this work to be given by the SoC, the temperature and the current direction, i.e., whether the cell incurs a charge or a discharge event. To bear out this assumption, let us analyze the relaxation of the cell under different operating conditions. Firstly, the need of accounting for the parameter dependence on the SoC can be clarified by analyzing figure 4.5 (**a**), which shows the relaxation of the cell after a 1C discharge event at different SoC values by a room temperature of 25 °C. As it

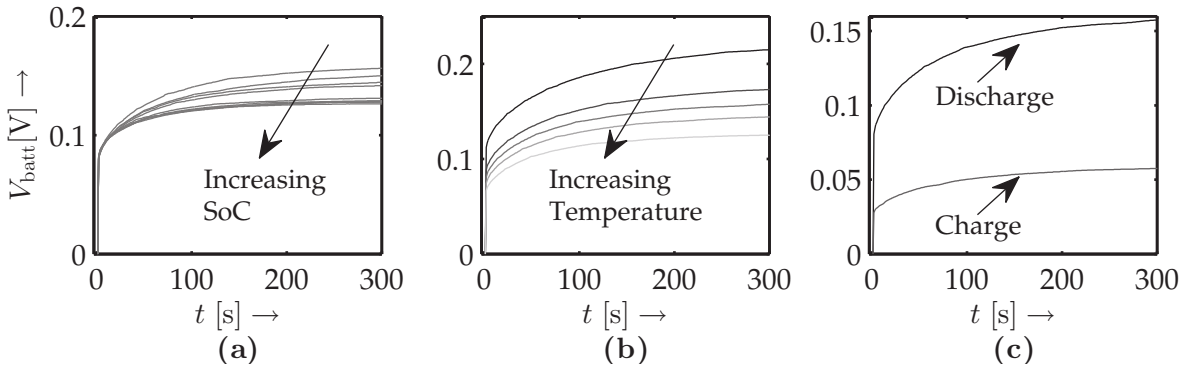

Figure 4.5.: Transient response of the terminal voltage in dependence of the state of charge (**a**), the temperature (**b**) and the current direction (**c**).

can be appreciated, there exist a maximum difference in the end value (after 300 s) of 29.6 mV. This difference suggest that the parameters should be scheduled according to the SoC. Secondly, figure 4.5 (**b**) depicts the cell's relaxation behavior for a SoC of 50% under different room temperatures, namely, 10 °C, 20 °C,25 °C, 30 °C and 40 °C. In general, high temperatures cause the internal resistance to be reduced, which in turn influences the relaxation of the cell. Thus, it becomes also necessary to schedule the parameters on the temperature. Thirdly, figure 4.5 (**c**) shows the relaxation phase after both, a discharge and a charge (flipped upside down) event, for a SoC of 50% and by a room temperature of 25 °C. It is clear that the direction of the current has a strong influence on the relaxation and therefore on the transient response of the cell. In contrast to the parameter scheduling envisaged on the SoC and the temperature, the aim of scheduling on the current direction has to be done following a binary schema. To this aim, as suggested by Hu et al. (2009), we treat the parameter dependencies on SoC and temperature separately according to the current direction. That is, we basically define two sets of parameters, i.e., one that is used for all $I_{\text{batt}} \geq 0$ (charge event or idle condition) and one for $I_{\text{batt}} < 0$ (discharge event).

Parameter identification: in the light of the discussion above, it becomes obvious that to ensure the validity of the single cell model for a wide operating range, the dependence of the parameters of the second order equivalent circuit-based model on the aforementioned operating conditions has to be taken into account. Thus, even though the system given by equations (4.4.1) and (4.4.2) is linear, identifying the model parameters requires treating the system as a linear parameter-varying (LPV) system.

To this aim we discharge and charge the cell in intervals of 0.05 SoC in the range $0 \leq \text{SoC} \leq 1$ ($2.75\,\text{V} \leq V_{\text{batt}} \leq 4.2\,\text{V}$) with pulse currents of 1C and repeat this sequence using the second experimental setup of section 2.3 (see figure 2.8 (**a**)) under different room temperatures, namely, 10 °C, 20 °C, 25 °C, 30 °C and 40 °C. Then, the identification of the model parameters for each SoC and for each temperature is performed by means of curve fitting for both discharge and charge events, respectively. For the sake of clearness, we rely on figure 4.4 to explain the identification procedure.

Let $t_0 < t < t_1$ be the time interval in which the cell is discharged by a total of 5% SoC with a pulse current of 1C and also let $t_1 < t < t_2$ be a sufficiently large resting period, in this case 300 s, which allows the cell to recover the unavailable charge.

In this work we perform the parameter identification by fitting the curve in the

interval $t_1 < t < t_2$ with the real response of the cell. Identifying the parameters in this manner, however, requires an excitation signal, in this case I_{batt}, which is not present during the resting period, i.e., $I_{\text{batt}} = 0$. To overcome this issue we introduce a virtual I_{batt} of 1C and identify the parameters by means of the following system identification procedure. First, let us first define

$$\frac{y_z}{u_z} = \frac{z^2 R_{\text{ohm}} + z\left(b_3 + b_4 - (a_5 + a_6) R_{\text{ohm}}\right) - \left(b_3 a_6 + b_4 a_5 - R_{\text{ohm}} a_5 a_6\right)}{z^2 - z\left(a_5 + a_6\right) + a_5 a_6}, \tag{4.4.3}$$

as the discrete-time transfer function of the transient response of the terminal voltage at certain SoC, i.e., $y_z = V_{\text{batt},z} - V_{\text{OC}}\left(\text{SoC}_z\right)$ due to a load $u_z = I_{\text{batt},z}$. Both b_3, b_4, a_5 and a_6 represent:

$$b_3 = R_{\text{s}}\left(1 - \text{e}^{\left(-\frac{\Delta t}{R_{\text{s}} C_{\text{s}}}\right)}\right) \quad b_4 = R_{\text{l}}\left(1 - \text{e}^{\left(-\frac{\Delta t}{R_{\text{l}} C_{\text{l}}}\right)}\right) \quad a_5 = \text{e}^{\left(-\frac{\Delta t}{R_{\text{s}} C_{\text{s}}}\right)} \quad a_6 = \text{e}^{\left(-\frac{\Delta t}{R_{\text{l}} C_{\text{l}}}\right)}. \tag{4.4.4}$$

By applying the inverse Z-transform to equation (4.4.3) the following second order difference equation describing the terminal voltage of the cell is obtained:

$$\begin{aligned} y_k \;=\; & y_{k-1}\left(a_5 + a_6\right) - y_{k-2}\left(a_5 a_6\right) + u_k R_{\text{ohm}} + u_{k-1}\left(b_3 + b_4 - (a_5 + a_6) R_{\text{ohm}}\right) + \\ & + u_{k-2}\left(-b_3 a_6 - b_4 a_5 + R_{\text{ohm}} a_5 a_6\right). \end{aligned} \tag{4.4.5}$$

For the sake of clearness we rewrite equation (4.4.5) as:

$$y_k = \varepsilon_1 y_{k-1} - \varepsilon_2 y_{k-2} + \varepsilon_3 u_k + \varepsilon_4 u_{k-1} + \varepsilon_5 u_{k-2} \tag{4.4.6}$$

where

$$\begin{aligned} &\varepsilon_1 = a_5 + a_6 && \varepsilon_2 = a_5 a_6 && \varepsilon_3 = R_{\text{ohm}} \\ &\varepsilon_4 = b_3 + b_4 - (a_5 + a_6) R_{\text{ohm}} && \varepsilon_5 = -b_3 a_6 - b_4 a_5 + R_{\text{ohm}} a_5 a_6. \end{aligned} \tag{4.4.7}$$

The desired parameters R_{s}, C_{s}, R_{l}, C_{l} and R_{ohm} are computed by first identifying the parameters ε_1 to ε_5 and then by solving the system of equations (4.4.7). We use the recursive least square (RLS) approach (Isermann and Münchhof 2011) to identify the parameters by iteratively minimizing the least square error between equation (4.4.6) and the real measurements generated during the experiments in the cycler. The identified parameters plotted against the SoC are shown in figures 4.6 (**a**) to 4.6 (**j**). As it can be appreciated, the parameters exhibit a relatively steady behavior with exception of some outliers, which arise when the cell approaches the region of total depletion. In the range SoC ≤ 0.1 the properties of the cell rapidly changes causing, for example, a rapid increment in R_{s}, R_{l} and R_{ohm} together with a decrement of C_{s}, C_{l}. As it is mentioned in section 4.3 the accuracy of the V_{OC}-SoC model has its maximum error also in the range SoC ≤ 0.1, due to the strong non-linearities present in the region towards total depletion. Since the modeled V_{OC}-SoC curve is used during the parameter identification process, as detailed above, an error in the open circuit voltage distorts the identified parameters. Consequently, it is not clearly distinguishable which portion of the parameter changes is attributed to the cell properties and which to the error in the V_{OC}-SoC model. Thus, it is not surprising that the modeling error increases at low SoC values (see section 4.6 for a further analysis on this phenomena).

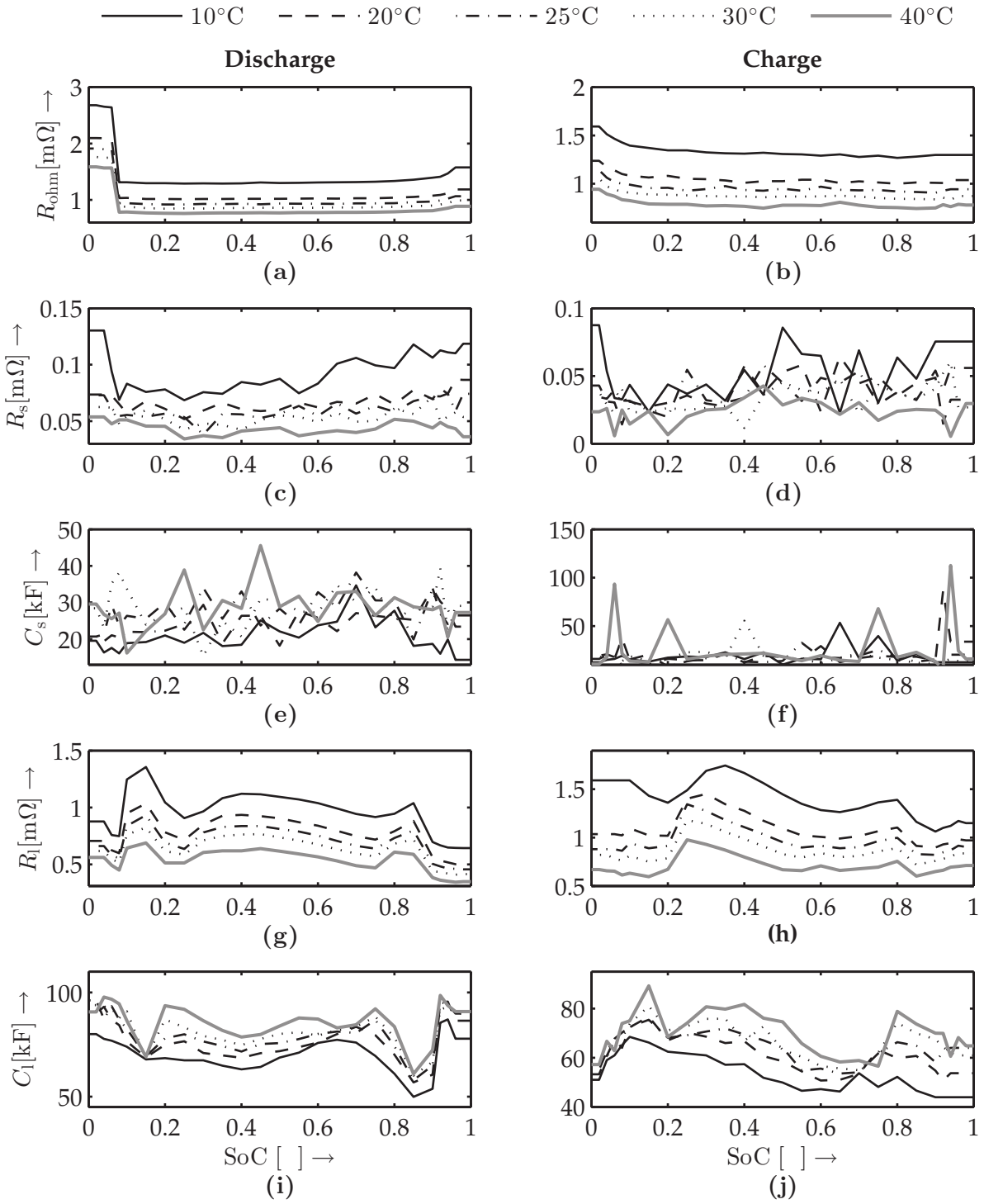

Figure 4.6.: Dependence of the parameter R_{ohm} (**a**)-(**b**), R_{s} (**c**)-(**d**), C_{s} (**e**)-(**f**), R_{l} (**g**)-(**h**) and C_{l} (**i**)-(**j**) on the temperature, the SoC and the current direction.

It can be clearly seen in figures 4.6 (**a**), (**b**), (**c**), (**d**), (**g**) and (**h**) that the value of all resistances increases at low temperatures. This effect becomes more visible in figure 4.7 (**a**), by plotting R_{ohm} against the temperature for different SoC values. A further effect of low temperatures is shown in figure 4.7 (**b**). Here, the modeling error (RMSE) obtained at 10 °C is larger in comparison to that obtained at 40 °C. This statement is based on the validation experiments discussed in section 4.6. Thus, it can be accord-

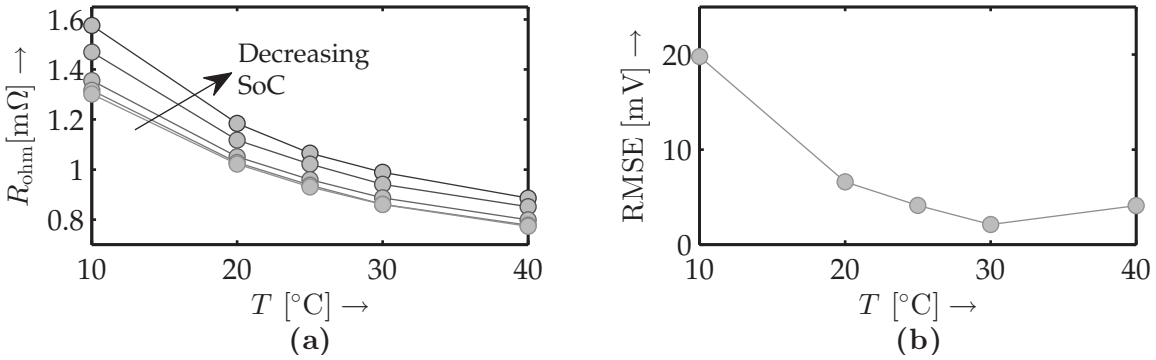

Figure 4.7.: Qualitative representation of the temperature dependence exhibit by both model parameters (**a**) and modeling error (**b**).

ingly concluded that the model accuracy decreases with the temperature. This is due to the SoC, which is used to schedule R_{s}, C_{s}, R_{l}, C_{l} and R_{ohm}, refers to the nominal capacity C_n for its computation (recall equation (4.2.3)). C_n, however, has been assumed to be constant throughout all identification experiments. This assumption is not always valid since C_n also decreases at low temperatures. This issue is overcome by an observer that accounts, among others, for the uncertainty introduced by C_n into the SoC estimation, as it is further detailed in chapter 6.

4.5. Terminal voltage

The output of the system (V_{batt}) is computed at given discrete-time $k \geq 0$ by:

$$\begin{aligned} y_k \;\; = \;\; & V_{\mathrm{OC}}\left(\mathrm{SoC}_k, T_{\mathrm{batt},k}\right) + R_{\mathrm{ohm}}\left(\mathrm{SoC}_k, T_{\mathrm{batt},k}\right) I_{\mathrm{batt},k} + \\ & + v_{\mathrm{s},k}\left(\mathrm{SoC}_k, T_{\mathrm{batt},k}\right) + v_{\mathrm{l},k}\left(\mathrm{SoC}_k, T_{\mathrm{batt},k}\right). \end{aligned} \tag{4.5.1}$$

As presented in figure 3.1, the backward-facing model computes the total electrical power demand P_{ele}. Nevertheless, the battery cell model requires the load current $I_{\mathrm{batt},k}$ as the input variable. Therefore, it is necessary to express $I_{\mathrm{batt},k}$ in terms of P_{ele}. To this aim we first express:

$$V_{\mathrm{batt}} = \frac{P_{\mathrm{ele}}}{I_{\mathrm{batt}}}, \tag{4.5.2}$$

then substitute equation (4.5.2) into equation (4.5.1) and finally solve it with respect to I_{batt}, which yields:

$$I_{\mathrm{batt},k} = -\frac{C - \sqrt{C^2 - 4P_{\mathrm{ele}}(\mathbf{u}_k^{\mathrm{env}})R_{\mathrm{ohm}}}}{2R_{\mathrm{ohm}}}, \tag{4.5.3}$$

with

$$C = V_{\mathrm{OC}}\left(\mathrm{SoC}_k, T_{\mathrm{batt},k}\right) + v_{\mathrm{s},k} + v_{\mathrm{l},k}.$$

$P_{\mathrm{ele}}(\mathbf{u}_k^{\mathrm{env}})$ expresses the dependence of the total electrical power requirements on the input vector $\mathbf{u}_k^{\mathrm{env}} = [v_{\mathrm{x},k}\ a_{\mathrm{x},k}\ \alpha_{\mathrm{road},k}]^T$. The solution with the positive part in the square root term of equation (4.5.3) is neglected, since its consideration would cause some load to be supplied by the battery when $P_{\mathrm{ele}} = 0$, which in practice is not possible.

4.6. Cell model evaluation

The evaluation of the cell model has been performed with aid of the second experimental scenario described in section 2.3. To assess the validity of the model discussed in the previous sections, the load profile shown in figure 4.8 (**a**) is applied to a second (validation) cell. This profile has been designed with the purpose of exiting both the dynamic behavior of the cell capacity as well as of the terminal voltage. The load profile proceed as follows. The cell is first discharged 2 % of its capacity with a C-rate of 1C (approximately 58.5 Ah) and then it idles for a period of 5 min, in which the cell incurs into a relaxation phase. After the idling period, the cell is charged 1 % with a C-rate of 0.25C followed again by an idling period of 2 min, after which, the cell experiences a discharge-charge event, of 1 % at 1C each, which serves to validate the response of the cell to changes in the load direction. This sequence is repeated until the battery is completely discharged. This experiment has been repeated in the cycler at temperatures of 10 °C, 20 °C, 25 °C, 30 °C and 40 °C, such that more meaningfully conclusions about the validity of the model can be drawn.

Figure 4.8 (**b**) shows the experimental results by a temperature of 25 °C. The RMSE is used as the performance metric to asses the accuracy of the model. Noteworthy is that RMSE is computed only on the region $0.1 \leq \text{SoC} \leq 1$, since, as it can be appreciated in the figure, in the range $\text{SoC} \leq 0.1$ the modeling error remarkably increases (see figure 4.8 (**d**)). The reason for this lies on the fact that towards the total depletion of the cell the V_{OC}-SoC curve exhibits an abrupt voltage drop (recall section 4.3). Moreover, there exist some uncertainty about the true value of C_n, which depends on both the C-rate and the temperature, as it has been previously discussed in section 4.4. From multiple experiments it has been noted, that if the nominal value set to C_n slightly differs from its true value, V_{batt} computed by the model at low SoC values drops faster than the real cell. It is observed that the transient behavior is properly captured by the model during both discharge and charge events. This confirms the decision of scheduling the parameters on the SoC and on the load direction. The modeling RMSE of the result shown in figure 4.8 (**b**), in the range $0.1 \leq \text{SoC} \leq 1$, is 4.13 mV with a maximum error of 30 mV (refer to the zoomed region of figure 4.8 (**c**)). The fact that the modeled SoC does not monotonically decreases through the entire operating range, as it is shown in the zoomed area of figure 4.8 (**d**), confirms the suitability of the model in the field of electric vehicles and thereby in the context of this work. This is due to the fact that the model is able to manage charging events caused, for example, by the regenerative braking system. Managing regenerative charging events is specially important, otherwise, their neglect would cause an eventual underestimation of the RDR. For the sake of completeness, figures 4.8 (**e**) and (**f**) as well as table 4.1 show the modeling error achieved at different temperatures.

The validation discussed so far is based on the evaluation of a single cell model. However, as described in section 2.6, the energy storage system of the Nissan Leaf is made up 196 cells. Thus, it might desirable to have a model that describes the behavior of the entire battery pack. Unfortunately, the first experimental setup of section 2.3 allows measuring the load of the entire battery pack and not of each cell. We therefore opt for addressing this issue during the SoC estimation, which is the discussion topic of chapter 6.

Table 4.1.: RMSE of the single cell model at different operating temperatures.

Temperature	10 °C	20 °C	25 °C	30 °C	40 °C
RMSE [mV]	19.78	6.61	4.13	2.12	4.09

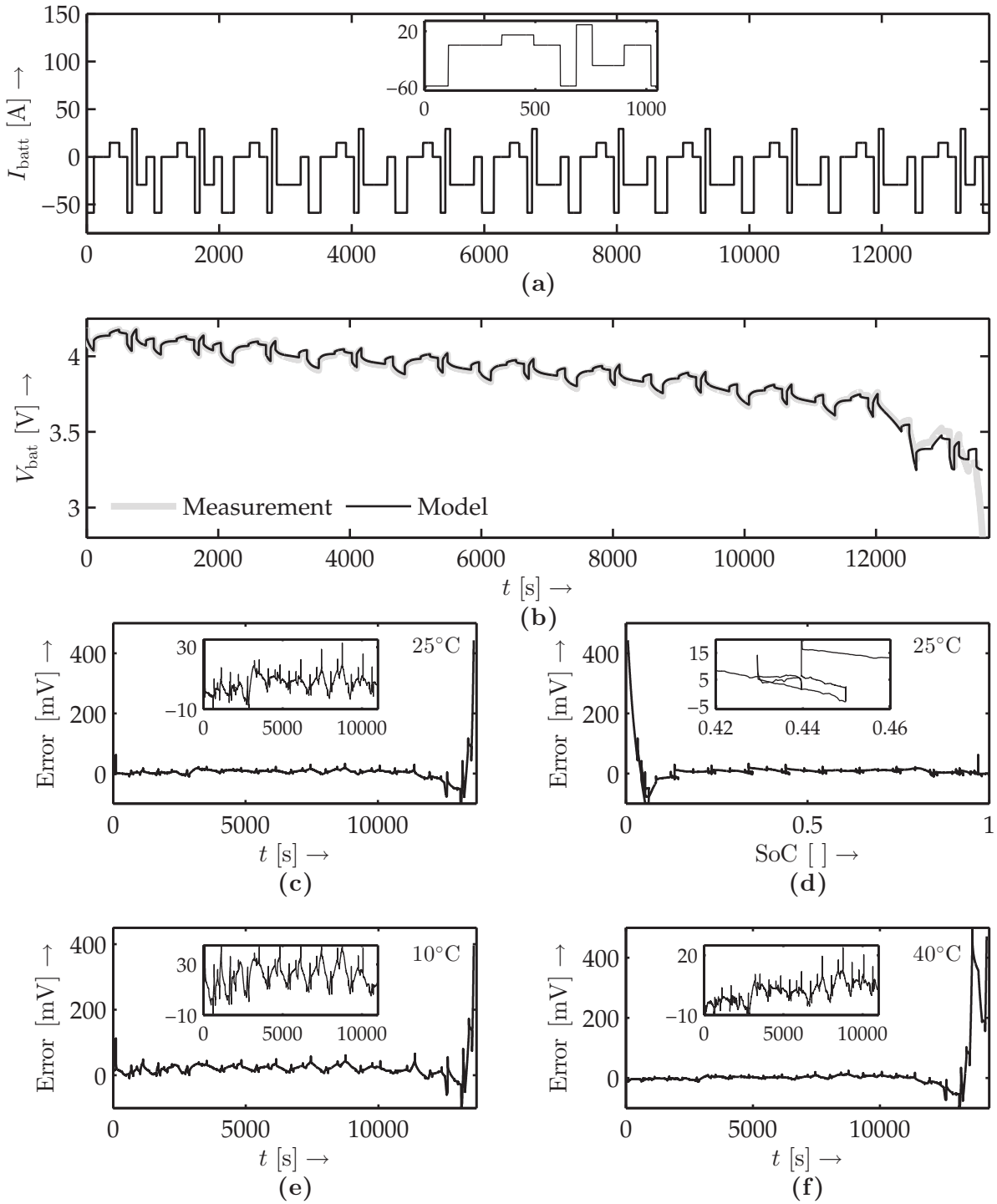

Figure 4.8.: Excitation load profile used for evaluating the single cell model (**a**), qualitative results of modeling the response of the terminal voltage (**b**). Evolution of the modeling error on time (**c**) and on dependence of the SoC (**d**) at 25 °C, 10 °C (**e**) and at 40 °C (**f**).

5
Model Uncertainty Analysis

The estimation of the RDR is affected by many sources of uncertainty. Thus, analyzing the combined effect of those sources is an essential factor in properly determining it. By using of the sample-based methodology introduced in section 2.2 the computational burden of the RDR estimation increases with the number of uncertain variables. Uncertainty analysis offers an efficient way for determining how the uncertainty related to model parameters affects RDR estimation and thereby allows identifying and prioritizing the most relevant uncertain parameters (Campolongo, Saltelli, and Cariboni 2011), such that the uncertainties of those ones with meaningless contribution can be neglected. The methodology presented in this chapter for uncertainty analysis finds its roots in the reliability theory. Reliability-based methods have lately gained importance in prognostics because they are computationally cheaper than, for example, brute-force Monte-Carlo based approaches (Oliva and Bertram 2015). The uncertainties arising from parameters of $\mathcal{M}_{PC}$ and $\mathcal{M}_{ESS}$ are characterized in section 5.1 and propagated through $\mathbf{G}(\cdot)$, i.e., the actual RDR estimation (refer to section 2.1), with the method introduced in section 5.2. Finally, the relative importance quantification and the aimed parameter prioritization is carried out via the uncertainty analysis of section 5.3.

5.1. Characterization of uncertain parameters

The discussion presented in section 3.5 as well as throughout chapter 4 made it clear, that parameterizing $\mathcal{M}_{PC}$ and $\mathcal{M}_{ESS}$ requires having sufficient large data-sets that allows identifying model parameters such that they are valid for the wide variety of conditions, in which the electric vehicle might operate. Moreover, even if enough data is available, the parameter values are largely influenced by the method used for their identification. Parameter uncertainty basically reflects the lack of knowledge about the true distribution of parameter values and arises from measurement errors, limited data-set size and from the decision criteria used to compare model outcomes with experimental data. Identifying $\mathcal{M}_{PC}$ and $\mathcal{M}_{ESS}$ with different data-sets, for example, would result in different parameter values which, in effect, yield different RDR estimations. In the same manner, if $\mathcal{M}_{PC}$ and $\mathcal{M}_{ESS}$ are identified with just one data-set but with many identification procedures, diverse parameter-sets would be also obtained turning out in a variety of RDR estimates. Prior the uncertainty anal-

ysis is required to characterize a distribution that properly represents the uncertainty of each parameter. In the following, the uncertainty characterization is performed based on real measurements and on proper literature references, depending on the parameter under consideration.

$\mathcal{M}_{PC}$ parameters

For the sake of clarity, the characterization of an arbitrary parameter ρ_{PC} of the model $\mathcal{M}_{PC}$ is expressed by an uniform distribution parametrized with an upper bound $\mathrm{b}_{U,\rho_{\mathrm{pc}}}$ and a lower bound $\mathrm{b}_{L,\rho_{\mathrm{PC}}}$, namely, by $\mathcal{U}_{\rho_{\mathrm{PC}}}(\mathrm{b}_{U,\rho_{\mathrm{PC}}}, \mathrm{b}_{L,\rho_{\mathrm{PC}}})$. For convenience of the reader, the following discussion is structured according the description of model parameters of $\mathcal{M}_{PC}$ as presented in chapter 3. Those parameters that are not discussed are considered to be deterministic and their values are set as discussed in section 3.5.

Vehicle dynamics

The **total mass** ($m_{\mathrm{V,t}}$) of the EV comprises the mass of the curb, the passengers (inclusive the driver) as well as the measurement equipment, e.g., DEWE-510, ADMA-G, etc. The mass of both the curb and the measurement equipment are obtained from technical specifications disclosed by the corresponding manufacturers and they are assumed to remain constant during their operating lifetime. The mass of the passengers, however, introduces some uncertainty into this parameter since the vehicle can be occupied by a minimum of one passenger (the driver) and a maximum of four. The Nissan Leaf weights 1520 kg, the mass of the measurement equipment amounts to 50 kg and it is assumed that the average passenger weights 75 kg. Accordingly, the probability distribution describing $m_{\mathrm{V,t}}$ is given by $\mathcal{U}_{m_{\mathrm{V,t}}}(1645, 1870)[\mathrm{kg}]$.

The **air density** (ρ_{air}) plays an important role (recall equation (3.2.2)), since it allows determining the longitudinal air drag. In chapter 3 the value of the air density was set to 1.226 $\mathrm{kg/m^3}$. However, this parameter depends on the ambient temperature, pressure and humidity, which are indirectly dictated by the altitude above the see level. To account for the uncertainty of ρ_{air} we set the ranges to 1.055 $\mathrm{kg/m^3}$ and 1.296 $\mathrm{kg/m^3}$, which correspond to ambient conditions between 100 m and 700 m above the sea level and temperatures between $-5\,^{\circ}\mathrm{C}$ and $-30\,^{\circ}\mathrm{C}$ (Asamer et al. 2016). Thus, the uncertainty is characterized from now on by $\mathcal{U}_{\rho_{\mathrm{air}}}(1.055, 1.296)[\mathrm{kg/m^3}]$.

Just as ρ_{air}, the parameter c_{w} is also used to compute the air drag of the vehicle. The **air drag coefficient** is usually identified by the vehicle manufacturer by means of wind tunnel experiments. The c_{w} of our experimental vehicle is 0.29, as disclosed by Nissan. However, since we have installed the GPS antenna on top of it we assume that c_{w} could increase by a factor of 1.05. For this reason we define $\mathcal{U}_{c_{\mathrm{w}}}(0.29, 0.3045)[-]$ for describing this uncertainty.

According to Hirt et al. (2013) the **gravitational acceleration** (g) varies from 9.76 $\mathrm{m/s^2}$ (measured near the equator), to 9.83 $\mathrm{m/s^2}$ in the arctic ocean. Accordingly, we characterize the uncertainty of g by $\mathcal{U}_g(9.76, 9.83)[\mathrm{m/s^2}]$.

While at high speeds most energy losses are due to air drag, at low speeds rolling resistance is of high importance. As discussed in chapter 3, this parameter depends on many external factors such as tire materials, hysteresis effects or the texture of the asphalt, which is the reason why no standard method for its identification has been established. Ton, Calwell, and Reeder (2003) investigate the influence of the **rolling**

resistance coefficient (K_R) on the fuel consumption of conventional vehicles and assumed K_R-values in the range between 0.0062 and 0.0152 for a total of 17 different tire models. This range is later confirmed by the Thomas and Mezies (2006) which conclude that a range between 0.007 and 0.015 is valid for most conventional vehicles. Although K_R, as formulated in equation (3.2.4), is speed dependent, we assume it as constant during the uncertainty analysis and therefore we define $\mathcal{U}_{K_R}(0.006, 0.015)[-]$.

Powertrain

The tires constitute a fundamental part of the powertrain since they contribute to the stability, handling and to the comfort besides that represent the coupling between the electric vehicle and the operating environment. The **tire effective rolling radius** (r_W) is a parameter that strongly varies depending on the tire properties and on the operating conditions. The Nissan Leaf is equipped with 205/55R16 tires with a static radius of 0.3162 m. Due to many factors such as the inflation pressure, speed, load or the wear of the tear, just to mention a few, r_W can be about 92% to 94% of the static value (Genta and Morello 2009). For this reason we set ranges for r_W between 0.29 m to 0.32 m. Accordingly, the uncertainty is characterized by $\mathcal{U}_{r_W}(0.29, 0.32)[\text{m}]$.

As it was already mentioned, the electric vehicle is propelled by a 80 kW and 280 Nm synchronous electric motor whose power losses are modeled by means of a stationary map. As it can be observed in figure 3.5 (**b**) the **electrical to mechanical efficiency** (η_M), in both motor and generator mode, varies from 0.82 to 0.95 depending on the operating point. Here we assume that η_M also comprises the losses due to power electronics. The **driveline efficiency** (η_G) according to Nissan is above 0.97 for all rotational speeds. Nevertheless, for the sake of reducing the parameter dimensionality during the uncertainty analysis, we integrate η_G into η_M. Thus the uncertainty related to η_M is characterized as $\mathcal{U}_{\eta_M}(0.795, 0.922)[-]$.

Auxiliary components

Finally, as it was highlighted in section 3.4, the **auxiliary power requirements** (P_{aux}) strongly vary depending on operating conditions. On the basis of multiple experiments it was observed that the Nissan Leaf consumes 457 W when all auxiliaries are turned-off and shows a peak of up to 4200 W by switching on the HVAC system together with lightening. For this reason we define $\mathcal{U}_{P_{aux}}(0.457, 4.5)[\text{kW}]$.

$\mathcal{M}_{ESS}$ parameters

Analogously to the previous section, the uncertainty is given by $\mathcal{U}_{\rho_{ESS}}(\text{b}_{U,\rho_{ESS}}, \text{b}_{L,\rho_{ESS}})$ for a given parameter ρ_{ESS} of $\mathcal{M}_{ESS}$.

Cell capacity

The KiBaM discussed in section 4.2 is parametrized by the **conductance** (d) and the **capacity ratio** (c), however, we consider only the uncertainty of the latter since it determines the fraction of the total charge in the cell that is available for use and therefore imposes a direct constraint to the RDR of the vehicle. Here we assume that c takes values between 0.8 and 1. Thus, we define $\mathcal{U}_c(0.8, 1.0)[-]$. Recalling equation (4.2.3), it can be seen that the **nominal capacity** (C_n) of the cell also contributes in determining the SoC. To characterize the uncertainty in C_n discharging events with different C-rates have been performed. It has been found that the withdrawn capacities vary between 57 Ah and 59.5 Ah, see figure 4.2 (**a**). Thus the uniform distribution for C_n is set as $\mathcal{U}_{C_n}(57, 59.5)[\text{Ah}]$.

Although SoC is not a parameter but instead a variable, carrying out a RDR estimation requires knowledge about the SoC at given prediction time k_p (recall section 2.2). Under normal operation the **initial state of charge** (SoC_0) of each cell is found everywhere between 5% to 95%. These bound are set by the battery management system with the aim of protecting the cells from deep discharging and overcharging events. We therefore set $\mathcal{U}_{\mathrm{SOC}_0}(0.05, 0.95)[-]$ to represent the uncertainty of the SOC_0.

Transient response

The second order equivalent circuit-based model used to describe the transient response of the Li-ion cell comprises the **ohmic resistance** (R_{ohm}) and by two two-RC network, i.e., $R_{\mathrm{s}}C_{\mathrm{s}}$ and $R_{\mathrm{l}}C_{\mathrm{l}}$, which describe the fast and the slow relaxation response. As it was previously exposed, these parameters depend on the operating temperature and on the SoC. For this reason, and with the aim of capturing the uncertainty related to the operating temperature, the parameters R_{ohm}, R_{s}, C_{s}, R_{l} and C_{l} are considered to follow an uniform distribution whose upper and lower bounds are obtained from the parameter identification carried out in section 4.4 and that is shown in figures 4.6 (**a**-**j**). Consequently, we define $\mathcal{U}_{R_{\mathrm{ohm}}}(7.43 \times 10^{-4}, 0.0027)[\Omega]$, $\mathcal{U}_{R_{\mathrm{s}}}(1.42 \times 10^{-5}, 1.30 \times 10^{-4})[\Omega]$, $\mathcal{U}_{C_{\mathrm{s}}}(5.91 \times 10^{3}, 1.13 \times 10^{5})[\mathrm{F}]$, $\mathcal{U}_{R_{\mathrm{l}}}(3.43 \times 10^{-4}, 0.0017)[\Omega]$ and $\mathcal{U}_{C_{\mathrm{l}}}(4.40 \times 10^{4}, 9.86 \times 10^{4})[\mathrm{F}]$.

5.2. Analytical uncertainty quantification

The inverse first-order reliability method (iFORM) has been applied in the literature to compute the RUL of various engineering systems, e.g., Xiang and Liu (2010) introduce a method for probabilistic fatigue prognosis based on the inverse-FORM. Similarly, Oliva and Bertram (2015) implement an inverse-FORM based solution for estimating the time to discharge (TTD) of Li-ion battery cells. In the following, we extend the use of the inverse-FORM for uncertainty analysis in prognostics.

To this aim let us first define $\mathbf{Z} = [Z_1, \ldots, Z_i, \ldots Z_{N^{\mathrm{Z}}}]$ as the N^{Z}-dimensional vector containing all uncertain parameters Z_i characterized in section 5.1. In addition, let us recall equation (2.1.9) and redefine $\mathbf{G}(\cdot)$ as a scalar *limit-state function* that maps a given realization $\boldsymbol{\zeta} = [\zeta_1, \ldots, \zeta_i, \ldots \zeta_{N^{\mathrm{Z}}}]$ of $\mathbf{Z}$ to the Boolean domain such that $\mathbf{G}(\zeta_1, \ldots, \zeta_i, \ldots \zeta_{N^{\mathrm{Z}}}) > r$ if $\mathrm{SoC} > \mathrm{SoC}_{\mathrm{min}}$ (safe) and $\mathbf{G}(\zeta_1, \ldots, \zeta_i, \ldots \zeta_{N^{\mathrm{Z}}}) < r$ if $\mathrm{SoC} < \mathrm{SoC}_{\mathrm{min}}$ (failure), wherein r is a given realization the RDR and all points lying on the N^{Z}-dimensional hyperplane $\{\boldsymbol{\zeta} : \mathbf{G}(\boldsymbol{\zeta}) = r\}$ conform the curve termed as *limit-state surface*. The aim of the inverse-FORM is to evaluate a tail of the probability density function of the RDR (see the upper right part of figure 2.3 (**b**)), such that the point along the limit-state surface with the highest probability of occurrence, i.e., the most probable point (MPP), can be found and thereby the RDR of the system that corresponds to a prescribed probability of failure (p_{f}) can be determined.

The first step towards finding this point is to transform all uncertain parameters into the equivalent standard normal space (Ξ) as illustrated figure 5.1. Since in this work we assume that all uncertain parameters contained in $\mathbf{Z}$ are statistically independent with strictly increasing continuous cumulative density functions, we employ the diagonal isoprobabilistic transformation:

$$\xi_i = \Phi^{-1}\left[\mathbf{F}_{Z_i}(\zeta_i)\right], \tag{5.2.1}$$

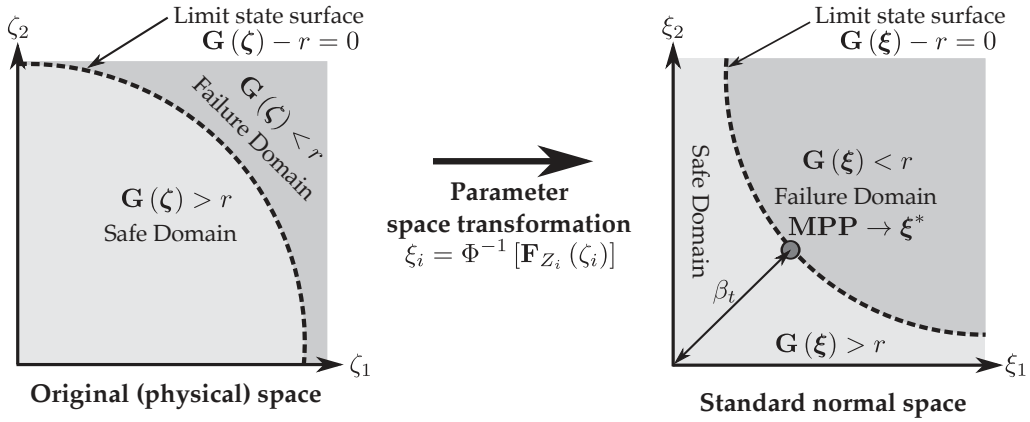

Figure 5.1.: Exemplification of the parameter space transformation from the original (physical) to the standard normal space and the interpretation of the most probable point (MPP).

where ξ_i is an arbitrary uncertain parameter described in the standard normal space, $\mathbf{F}_{Z_i}(\zeta_i)$ is the evaluation of the cumulative density function (CDF) of the uncertain parameter ζ_i expressed in the original (physical) space and $\Phi^{-1}[\,\cdot\,]$ corresponds to the inverse of the standard normal CDF. The term *diagonal* arises from the fact that each uncertain parameter Z_i incurs a transformation independently of all other parameters. Since all uncertain parameters are represented in the standard normal space, the RDR also becomes a normal variable that is be expressed as a linear combination of all uncertain variables and the MPP becomes the point on the limit-state surface that is closest to the origin, as shown in the right hand side of figure 5.1. As it can be seen in figure 5.2 (**a**), in the resulting standard normal space and with help of the MPP concept, multiple realizations r_i of the RDR are evaluated as result of different probabilities p_i defined in the interval $[0, 1]$. The term $p_f = p_i$ represents the probability that $\mathbf{G}(\boldsymbol{\xi_i}) < r_i$, or rather that $\mathrm{SoC} < \mathrm{SoC}_{\min}$ with a given set of parameters $\boldsymbol{\zeta}_i$ (expressed in the physical space) or $\boldsymbol{\xi}_i$ (transformed into the normal space).

In the context of the inverse-FORM the bridge between p_i and the underlying r_i is given by the first-order reliability index $\beta_{t,i}$, which represents the MPP in the standard normal space and is computed by:

$$p_i = \Phi^{-1}(\beta_{t,i})\,. \tag{5.2.2}$$

Hence, computing the CDF of the RDR reduces to find all MPP's along the limit-states surfaces corresponding to each p_i. In this sense, the inverse-FORM is formulated as the following optimization problem:

$$\boldsymbol{\xi}^* = \arg\min\{\,\|\boldsymbol{\xi}\|\,|\,\mathbf{G}(\boldsymbol{\zeta}) = \mathbf{G}(\boldsymbol{\xi}) = r\}\,, \tag{5.2.3}$$

where the vector $\boldsymbol{\xi}^*$ contains the set of values of the uncertain parameters, expressed in the normal space, that correspond to the point along the limit-state surface with the shortest distance to the origin. The aim of the inverse-FORM is to compute $\boldsymbol{\xi}^*$, so that $\beta_t = \|\boldsymbol{\xi}\|$. In this work we solve the optimization problem (5.2.3) by means of the inverse HL-RF algorithm (Rackwitz and Flessler 1978). In this approach the

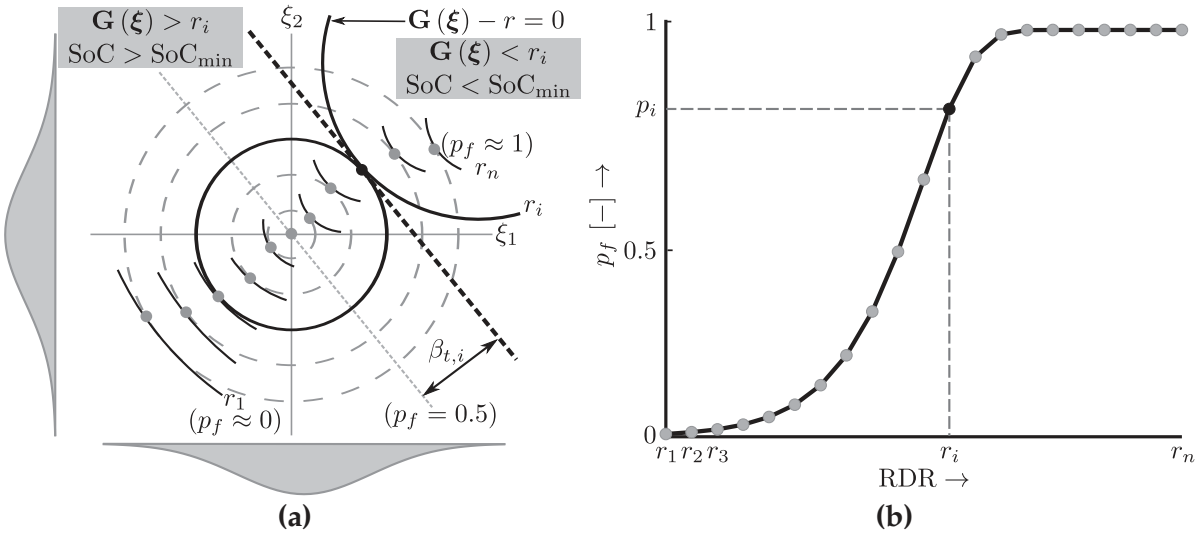

Figure 5.2.: Multiple MPP values corresponding to different probabilities of failure (**a**) and discrete cumulative density function of the RDR computed with the iFORM method (**b**).

optimization problem is iteratively solved by:

$$\boldsymbol{\xi}^{\lambda_t+1} = -\beta_t \frac{\nabla_{\xi}\mathbf{G}\left(\boldsymbol{\xi}^{\lambda_t}\right)}{\left\|\nabla_{\xi}\mathbf{G}\left(\boldsymbol{\xi}^{\lambda_t}\right)\right\|}. \tag{5.2.4}$$

Herein represents $\nabla_{\xi}\mathbf{G}\left(\boldsymbol{\xi}_t^{\lambda}\right)$ a gradient vector computed by means of:

$$\nabla_{\xi}\mathbf{G}\left(\boldsymbol{\xi}^{\lambda_t}\right) = \frac{\partial \mathbf{G}\left(\xi_i\right)}{\partial \zeta_i} \times J_{\xi,\zeta}, \tag{5.2.5}$$

where λ_t is the step-length of each iteration and $J_{\xi,\zeta}$ is the Jacobian of all uncertain parameters transformed from the physical into the standard normal space. Under the assumption that no correlation between the uncertain parameters exist, the Jacobian is computed as follows:

$$J_{\xi,\zeta} = \mathrm{diag}\left(\frac{\Phi\left(\xi_i\right)}{\mathbf{f}_{Z_i}\left(\zeta_i\right)}\right), \tag{5.2.6}$$

where $\mathbf{f}_{Z_i}\left(\zeta_i\right)$ represents the evaluation of the probability density function of Z_i at ζ_i. After having determined $\boldsymbol{\xi}^*$ in the standard normal space a back-transformation into $\boldsymbol{\zeta}^*$ is carried out, so that the RDR can be evaluated with $\mathbf{G}\left(\boldsymbol{\zeta}^*\right)$ in the physical space.

Sensitivity measures

The aim of this chapter is to quantify the sensitivity of the RDR estimation as response to variations of the uncertain parameters on the entire parameter space. Through the uncertainty analysis it is possible to determine the contribution of each uncertain parameter on the overall variance of the RDR estimate. This analysis sets the basis for properly comparing the significance of each parameter, which is the goal of section 5.3. One of the computational advantages of the inverse-FORM is that a set of importance factors and sensitivity measures of the uncertain parameters is provided as by-product while solving equation (5.2.4).

To obtain the sensitivity measures let us first consider $\mathbf{G}_L(\boldsymbol{\xi}^*)$ as the linearization of the limit-state function $\mathbf{G}(\cdot)$ evaluated at $\boldsymbol{\xi}^*$. Kiureghian (2004) suggests that the variance of $\mathbf{G}_L(\boldsymbol{\xi}^*)$ can be approximated by:

$$\sigma^2_{\mathbf{G}_L(\boldsymbol{\xi}^*)} = \left\|\nabla_{\tilde{\xi}}\mathbf{G}_L(\boldsymbol{\xi}^*)\right\|^2 \sum_{i=1}^{N^{\mathbf{Z}}} \alpha^2_{t,i} = \left\|\nabla_{\tilde{\xi}}\mathbf{G}_L(\boldsymbol{\xi}^*)\right\|^2, \tag{5.2.7}$$

where $\alpha^2_{t,i}$ belongs to the unit length vector $\boldsymbol{\alpha}_t = [\alpha^2_{t,1}, \ldots, \alpha^2_{t,i}, \ldots, \alpha^2_{t,N^{\mathbf{Z}}}]$. Thus, it becomes obvious that the contribution of the i^{th} uncertain parameter is proportional to the factor $\alpha^2_{t,i}$, which is computed as by-product in the last iteration of equation (5.2.4) through:

$$\boldsymbol{\alpha}_{t,i} = -\frac{\nabla_{\tilde{\xi}}\mathbf{G}\left(\xi^*_i\right)}{\left\|\nabla_{\tilde{\xi}}\mathbf{G}\left(\xi^*_i\right)\right\|}. \tag{5.2.8}$$

This factor indicates the relative importance of ξ_i to the overall variance of $\mathbf{G}(\cdot)$, i.e., of the RDR estimate. Due to the assumed statistic independence of the elements of $\boldsymbol{\zeta}$ there exist a one to one correspondence between $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$. Hence, the vector of importance factors $\boldsymbol{\alpha}_t$ is also valid in the physical space.

5.3. Remaining driving range uncertainty analysis

The aim of the uncertainty analysis presented in this section is to determine the importance of the uncertain parameters characterized in section 5.1 in regards to their contribution to the total uncertainty of the RDR estimation. The assessment of the importance factors given by equation (5.2.8) is carried out via simulation. To this aim we first generate 3 sufficiently large driving cycles describing urban, rural and highway driving conditions (see section 7.1 for a detailed information on the driving conditions employed throughout this work). In this manner it is possible to investigate how the importance of each uncertain parameter varies according to the operating environment of the electric vehicle.

The basis for the synthetic cycles are the ARTEMIS standard drive cycles (Andre 1996, 2004), which are shown in figures B.5 (**a**), (**b**) and (**c**). Moreover, we also generate a synthetic slope profile (see figure B.5 (**d**)) with the aim of better accounting for the uncertainty introduced by the climbing resistance (recall equation (3.2.5)). As it has been stated before, the uncertainty regarding the operating temperature is not taken into account since it is implied in the parameter ranges characterized in the parameters R_{ohm}, R_{s}, C_{s}, R_{l} and C_{l}.

All importance factors are summarized in table 5.1. The following analysis is based on the first three columns of the table. In this case, the uncertainty of the SoC_0 is characterized as presented in section 5.1. Noteworthy is that the uncertainty related to the mass of the vehicle ($m_{\text{V,t}}$) only slightly contributes to the total variance of the RDR estimation. Similarly, the uncertainty of the environment parameters such as ρ_{air} and g as well as parameters related to the vehicle design, e.g., r_{W}, c_{w} or η_{M} also exhibit a meaningless contribution. Note, that the static rolling resistance coefficient (K_{R}) is certainly important, specially under rural driving conditions. Nevertheless, its maximum contribution ot the total variance of the RDR is around 2.5 %. Thus, it

Table 5.1.: Importance factors of all uncertain parameters for a probability of failure $p_i = 0.5$, under various driving conditions and with various characterizations of the SoC_0 uncertainty.

Uncertain parameter	**Unit**	with SoC_0			without SoC_0		
		$p_i = 0.5$			$p_i = 0.5$		
		Urban	Rural	Highway	Urban	Rural	Highway
		$\mathcal{M}_{PC}$ parameters					
$m_{V,t}$	[kg]	0.001	0.002	0.001	0.027	0.012	0.007
ρ_{air}	[kg/m^3]	0.000	0.001	0.004	0.000	0.010	0.019
c_w	[−]	0.000	0.000	0.000	0.000	0.001	0.001
g	[m/s^2]	0.000	0.000	0.000	0.000	0.000	0.000
K_R	[−]	0.002	0.025	0.007	0.031	0.188	0.171
r_W	[m]	0.000	0.000	0.000	0.000	0.000	0.000
η_M	[−]	0.001	0.008	0.008	0.028	0.010	0.056
P_{aux}	[W]	0.053	0.099	0.006	0.769	0.735	0.463
		$\mathcal{M}_{ESS}$ parameters					
R_s	[Ω]	0.000	0.000	0.000	0.000	0.000	0.000
C_s	[F]	0.000	0.000	0.000	0.000	0.000	0.000
R_l	[Ω]	0.000	0.000	0.000	0.000	0.000	0.011
C_l	[F]	0.000	0.000	0.000	0.000	0.000	0.000
R_{ohm}	[Ω]	0.000	0.000	0.003	0.000	0.004	0.118
c	[−]	0.018	0.019	0.022	0.130	0.036	0.140
SoC_0	[−]	0.923	0.844	0.947	0.000	0.000	0.000
C_n	[Ah]	0.002	0.002	0.002	0.015	0.004	0.014

can also be neglected. On the contrary, the uncertainty of the power consumed by auxiliaries (P_{aux}) has the highest importance of almost 10 %, which suggests that the determination of this parameter might be relevant while estimating the RDR.

The fact that the uncertainty of the parameters R_{ohm}, R_s, C_s, R_l and C_l is meaningless, with a total contribution of zero, is not surprising. This is due to these parameters are in charge of describing the transient response of the battery and not the capacity and thereby the RDR of the vehicle.

The uncertainty analysis discussed above might not be totally unbiased since the uncertainty of the SoC_0 contributes in average to 90.47 % of the total variance of the estimated RDR. This is clear since, as explained in section 2.2, the SoC sets the basis for the load forecasting and thereby for the RDR estimation. In order to draw more meaningfully conclusions about the importance of all uncertain parameters, we characterize the uncertainty of SoC_0 as $\mathcal{N}_{SoC_0}(0.8, 1 \times 10^{-}3)[-]$. A Gaussian distribution for characterizing the uncertainty of SoC_0 is a reasonable choice, since, as it will be further discussed in chapter 6, the SoC is recursively estimated using a Bayes-based approach. The results of the second uncertainty analysis are found in the last three columns of table 5.1. After removing the high uncertainty of the SoC_0 employed in the previous analysis, it can be seen that the uncertainty of P_{aux} remarkably gains in importance. This confirms the suggestion previously stated. It can also be appreciated that K_R and c become more relevant. Thus, we can argue that only the uncertainties related to SoC_0, P_{aux}, K_R and c highly contribute to the variance of the RDR estimation and therefore these parameters have to be carefully determined.

6

System State and Parameter Estimation

This chapter deals with the procedure employed in estimating the state of charge of the energy storage system based on the model presented in chapter 4. To this aim we first formulate the state and parameter estimation in the context of stochastic filtering based on sequential Monte Carlo methods. On this basis, we explain in detail the theoretical foundations of a computationally efficient alternative to state of the art filtering methods, i.e., the particle flow particle filter (PFPF) and afterwards we present the steps needed for its implementation within the prognostics approach presented in section 2.2 for estimating the SoC. Finally, this chapter discusses the experimental results obtained during tests carried out under controlled operating conditions in a cycler, on the one hand, and under real driving conditions with the experimental vehicle, on the other hand.

6.1. Probabilistic state of charge estimation

As it was stated in section 2.1, the states of the energy storage system have to be estimated in order to gain information about the energy available for the mobile system.

Throughout this work it has become clear that the quality of the identified parameters is crucial for accurately modeling the nonlinear behavior of the battery capacity and the terminal voltage. Nevertheless, lithium-ion cells are subject to aging effects which slowly change their electrical behavior over time. Thus, the parameters identified in chapter 4 reflect the true behavior of the cell at that point in time but say few about the negative effects introduced by the aging itself. Furthermore, manufacturing variability between cells causes the quality of an identified model to be not necessarily transferable to other cells. Hence, it is reasonable to think, that properly estimating the state of charge of the energy storage system through its complete lifetime requires accounting for the changes that the parameters may incur due to the aforementioned causes. It is therefore desirable to employ an approach that estimates the model parameters online and thus continuously adapts them to the most up to date behavior of the system, so that both aging and cell-to-cell variability effects can be accounted for and the estimation of the state of charge can be performed more robustly.

To this aim let us first define the augmented state $\mathbb{X}_k = \begin{bmatrix} \mathbf{x}_k^{\mathrm{T}} & \boldsymbol{\theta}_k^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ spanned over the joint state-parameter space $\mathcal{X} \in \mathbb{R}^{N_x+N_\theta}$. Consequently, the task is the *joint estimation* $p\left(\mathbb{X}_k|\mathbf{Y}_k\right)$, i.e., to represent the most up-to-date knowledge about states and

parameters of the model representing the energy storage system at any given time index k based on the current and on all past measurements. While estimating $p\left(\mathbb{X}_k|\mathbf{Y}_k\right)$ it is important to notice, that the behavior of the energy storage system can be analyzed under two time-scales. The former refers to the states of the system, such as the SoC, which vary relatively fast, in the range between minutes and hours. The latter corresponds to the system parameters, e.g., the internal resistance R_{ohm}, which tend to vary slow in the range of months or years, due to, for example, aging effects. It is therefore important to extend the discrete time system equations (2.2.1) and (2.2.2) such that the slow dynamics of the system parameters can be properly taken into account. A commonly accepted approach is to allow the state parameters to evolve obeying a random walk, that is:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{r}_{\theta,k}, \tag{6.1.1}$$

where $\mathbf{r}_{\theta,k} \in \mathbb{R}^{N_\theta}$ is the parameter noise vector. In this manner, the parameters of the energy storage system model become time variant, even though they do not directly depend on time, as discussed in section 6.3. With the aim of formulating the SoC estimation procedure carried out in section 6.2, let us redefine the system equations as:

$$\mathbb{X}_{k+1} = \tilde{\mathbf{f}}\left(\mathbb{X}_k, \mathbf{u}_k, \mathbb{W}_k, k\right), \tag{6.1.2}$$

$$\mathbf{y}_k = \tilde{\mathbf{h}}\left(\mathbb{X}_k, \mathbf{u}_k, \mathbf{n}_k, k\right), \tag{6.1.3}$$

where $\mathbb{W}_k = \left[\begin{array}{cc} \mathbf{v}_k^{\mathrm{T}} & \mathbf{r}_{\theta,k}^{\mathrm{T}} \end{array}\right]^{\mathrm{T}}$, $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_{v,k})$, $\mathbf{r}_{\theta,k} \sim \mathcal{N}(0, \mathbf{R}_{\theta,k})$ and $\mathbf{n}_k \sim \mathcal{N}(0, \mathbf{R}_{n,k})$. The convergence rate and the accuracy of the estimation depend on the choice of the covariance matrices $\mathbf{R}_{v,k}$, $\mathbf{R}_{\theta,k}$ and $\mathbf{R}_{n,k}$. Therefore, they must be properly tuned to ensure the convergence of the SoC estimate. Both, the state function $\tilde{\mathbf{f}} : \mathbb{R}^{N_\mathbb{X}} \times \mathbb{R}^{N_u} \times \mathbb{R}^{N_\mathbb{W}} \rightarrow \mathbb{R}^{N_\mathbb{X}}$ and the output function $\tilde{\mathbf{h}} : \mathbb{R}^{N_\mathbb{X}} \times \mathbb{R}^{N_u} \times \mathbb{R}^{N_n} \times \rightarrow \mathbb{R}^{N_y}$ combine the *fast* and the *slow* dynamics of the system states and parameters, respectively.

Recursive approaches for joint estimation, e.g., sequential Bayesian filtering, have gained in importance in the last decade. Due to their versatility and ease of implementation in practical applications, they conform the basis of the approach implemented in this work.

Let $\mathbb{X}_{0:k} \triangleq \{\mathbb{X}_0, \mathbb{X}_1, ..., \mathbb{X}_k\}$ be a sequence of augmented states generated by $\tilde{\mathbf{f}}(\cdot)$ evolving in time as a first order Markov process and also let $\mathbf{Y}_k \triangleq \{y_0, y_1, ..., y_k\}$ be a sequence of observations generated up to time index k, as previously stated in section 2.2. The goal of the sequential estimation is to compute the optimal estimate of $\mathbb{X}_k$ given $\mathbf{Y}_k$. Computing the optimal solution for the estimation in the mean-squared error sense yields:

$$\hat{\mathbb{X}}_k = \mathbb{E}\left[\mathbb{X}_k|\mathbf{Y}_k\right] = \int_{\mathcal{X}} \mathbb{X}_k p\left(\mathbb{X}_k|\mathbf{Y}_k\right) \mathrm{d}\mathbb{X}_k, \tag{6.1.4}$$

where $\hat{\mathbb{X}}_k$ is the *best* estimate of $\mathbb{X}_k$ and $\mathbb{E}\left[\mathbb{X}_k|\mathbf{Y}_k\right]$ is the conditional mean. As it can be seen, evaluating this expectation means first computing the a-posteriori state estimate $p\left(\mathbb{X}_k|\mathbf{Y}_k\right)$. A recursive procedure for computing this a-posteriori density function can be carried out by applying the Bayes' Theorem under the assumption $\mathbf{y}_k$ is conditional independent of $\mathbf{Y}_{k-1}$ given $\mathbb{X}_k$ (Plett 2006).

In a generic Bayesian framework, the posterior $p\left(\mathbb{X}_k|\mathbf{Y}_k\right)$ is obtained in the prediction step by a single computation of the Bayes' rule given by:

$$\overbrace{p\left(\mathbb{X}_k|\mathbf{Y}_k\right)}^{\text{posterior}} = \frac{\overbrace{p\left(\mathbb{X}_k|\mathbf{Y}_{k-1}\right)}^{\text{prior}}\overbrace{p\left(\mathbf{y}_k|\mathbb{X}_k\right)}^{\text{likelihood}}}{\underbrace{p\left(\mathbf{y}_k|\mathbf{Y}_{k-1}\right)}_{\text{normalization factor}}}. \tag{6.1.5}$$

In principle, evaluating equation (6.1.5) requires solving multi-dimensional integrals. For those systems that exhibit a linear behavior and all noises are assumed to be Gaussian, the solution of equation (6.1.5) is given by the well known filter first introduced by Kalman (1960). Although the model of the energy storage system studied here is linear (see equations (4.2.1), (4.2.2), (4.4.1) and (4.4.2)), the dependence of the parameters on the SoC and on the temperature renders a highly non-linear system. Unfortunately, the integrals involved in equation (6.1.5) for non-linear systems mostly lack of an analytical solution and therefore, have to be approximated, e.g., by applying Monte Carlo integration methods.

Thanks to its robustness and performance in non-linear non-Gaussian applications, sequential Monte Carlo (SMC) filtering have become the state of the art method for joint state-parameter estimation. From all SCM-based methods for filtering available in the literature, the particle filter first introduced by Gordon, Salmond, and Smith (1993) is the most wide spread approach for addressing the joint estimation in the context of model-based prognostics (Orchard and Vachtsevanos 2009), since it is not subject to constraints regarding linearity or Gaussianity of the system.

Recalling equation 2.2.3, the particle filter approximates the posterior probability distribution $p\left(\mathbb{X}_k|\mathbf{Y}_k\right)$ through a set of $n_{\mathbb{X}} \gg 1$ weighted particles $\left\{\mathbb{X}_k^i, w_k^i\right\}_{i=1}^{n_{\mathbb{X}}}$ in a Monte Carlo fashion, satisfying (Doucet, De Freitas, and Gordon 2001):

$$\sum_{i=1}^{n_{\mathbb{X}}} w_k^i \delta\left(\mathbb{X}_k - \mathbb{X}_k^i\right) \overset{n_{\mathbb{X}}\to\infty}{\to} p\left(\mathbb{X}_k|\mathbf{Y}_k\right), \tag{6.1.6}$$

where $\mathbb{X}_k^i$ is the set of particles representing the state space, $w_k^i \geq 0$ are their associated importance weights and $\delta(\cdot)$ describes the Dirac delta function located at $\mathbb{X}_k^i$. Each particle is sampled from an importance density and is propagated through the function $\tilde{\mathbf{f}}\left(\cdot\right)$ in the so called *prediction step*. Then, the value of each particle is updated from measurements through the output function $\tilde{\mathbf{h}}\left(\cdot\right)$ in the *measurement update step*. In this step the weight of each particle is updated according to the likelihood of a new measurement given the particle. In this manner, given new observations the pursued optimal estimate $\hat{\mathbb{X}}_k$ can be accordingly computed as:

$$\hat{\mathbb{X}}_k = \sum_{i=1}^{n_{\mathbb{X}}} w_k^i \hat{\mathbb{X}}_k^i. \tag{6.1.7}$$

As it can be inferred from the discussion above, the performance of the particle filter relies on the selection of the importance density. For non-linear systems, such as the one treated in this work, a proper choice of this density is not an easy task. The most common approach is to sample from the *prior* state estimate $p\left(\mathbb{X}_k|\mathbf{Y}_{k-1}\right)$. Sampling

from the prior, however, usually leads to a high variance of the particle weights, which causes the so called *particle degeneracy* (Daum and Huang 2011), that is, the Monte Carlo approximation given by equation (6.1.7) is mainly dictated by few particles with very high weights. This problem becomes more significant in a joint state/parameter estimation since the dimensionality of the state space can increase considerably. Thus, it becomes obvious that using standard particle filters based on improper importance densities might negatively affect the computational performance of the RUL/RDR estimation approach introduced in section 2.2. Firstly, during the prediction step computational resources are wasted during the propagation of particles with negligible weights through $\tilde{\mathbf{f}}(\cdot)$. Secondly, since these have to be propagated forward in time until they reach the predefined threshold ($\mathrm{SOC}_{\min}$) additional resources are wasted during the prediction step of the prognostics approach.

Doucet, Godsill, and Andrieu (2000) show that the *optimal importance density* (OID) helps to alleviate the degeneracy of particles is the conditional posterior distribution of the states, which however, in most cases has no analytical form. One idea to mitigate the particle degeneracy that has gained in importance over the last years is based on *particle flow* (Daum and Huang 2008) and *optimal transport* (Reich 2011) methods. The main idea behind this approach is to introduce a particle flow, which progressively transforms the prior $p(\mathbb{X}_k|\mathbf{Y}_{k-1})$ into the posterior state-parameter estimate $p(\mathbb{X}_k|\mathbf{Y}_k)$ by gradually moving the particles in an optimal manner as new measurements become available without needing to randomly sample from any distribution.

Oliva and Bertram (2014b) investigate the suitability of particle flow in the context of model-based prognostics and show that a reduction in computational complexity is achieved in contrast to standard bootstrap particle filter since less particles are needed in the state estimation and thereby less number of simulations are performed in order to determine the distribution of the RUL. The results obtained, however, are based on state estimation with identified model parameters. We therefore extend this approach and apply it to the joint state-parameter estimation in order to assess the SoC and the parameters of the single cell model developed in chapter 4. In the sequel, we introduce the basic idea behind the particle flow particle filter. A detailed derivation of this approach is found in the work of Oliva and Bertram (2014b).

6.2. Particle flow for joint state-parameter estimation

The desired flow of particles from $p(\mathbb{X}_k|\mathbf{Y}_{k-1})$ to $p(\mathbb{X}_k|\mathbf{Y}_k)$ is obtained by solving an ordinary differential equation (ODE). By denoting a new set of density functions given by $\boldsymbol{\gamma}(\mathbb{X}_{k,\lambda}|\mathbf{Y}_k) = p(\mathbb{X}_k|\mathbf{Y}_k)$ and $\mathbf{d}(\mathbb{X}_{k,\lambda}|\mathbf{Y}_{k-1}) = p(\mathbb{X}_k|\mathbf{Y}_{k-1})$ it is possible to compute $\boldsymbol{\gamma}(\mathbb{X}_{k,\lambda}|\mathbf{Y}_k)$ in a B_λ-fold recursive manner by progressively introducing the likelihood density, here denoted as $\mathbf{l}(\mathbf{y}_k|\mathbb{X}_k)$, such that the prior $\mathbf{d}(\mathbb{X}_{k,\lambda}|\mathbf{Y}_{k-1})$ gradually deforms into $\mathbf{d}(\mathbb{X}_{k,\lambda}|\mathbf{Y}_{k-1})\,\mathbf{l}(\mathbf{y}_k|\mathbb{X}_k)$. Such transformation is achieved by using a *homotopy* that resembles the Bayes' update step given by equation (6.1.5) and that takes the form:

$$\overbrace{\boldsymbol{\gamma}(\mathbb{X}_{k,\lambda}|\mathbf{Y}_k)}^{\text{posterior}} = \frac{\overbrace{\mathbf{d}(\mathbb{X}_{k,\lambda}|\mathbf{Y}_{k-1})}^{\text{prior}}\overbrace{\mathbf{l}(\mathbf{y}_k|\mathbb{X}_{k,\lambda})^{\lambda}}^{\text{likelihood}}}{\underbrace{K_{k,\lambda}}_{\text{normalization factor}}}. \tag{6.2.1}$$

The term $\lambda \in [0,1]$ is the progression parameter, also referred to as pseudo-time, and $\mathbf{l}\left(\mathbf{y}_k|\mathbb{X}_{k,\lambda}\right)^{\lambda}$ is understood as an incremental likelihood. Thus, equation (6.2.1) represents the prior when $\lambda = 0$ and the posterior when $\lambda = 1$. The number of iterations in the recursion, namely B_λ, depends on the step size $\Delta\lambda$, which determines the rate at which $\lambda_{0\to1}$. For the sake of clarity, from now on we express the augmented state variables as $\mathbb{X}_\lambda$ instead of as $\mathbb{X}_{k,\lambda}$. This is due to the fact that the evolution of the probability distribution as $\lambda_{0\to1}$ always occurs at the discrete time step k. To avoid numerical issues the log-density of equation (6.2.1) is applied yielding to:

$$\mathbf{\Gamma}\left(\mathbb{X}_\lambda\right) = \mathbf{D}\left(\mathbb{X}_\lambda\right) + \lambda\mathbf{L}\left(\mathbb{X}_\lambda\right) - \log K_\lambda, \tag{6.2.2}$$

where the posterior is given by $\mathbf{\Gamma}\left(\mathbb{X}_\lambda\right) = \log\boldsymbol{\gamma}\left(\mathbb{X}_\lambda|\mathbf{Y}_k\right)$, the prior is represented by $\mathbf{D}\left(\mathbb{X}_\lambda\right) = \log\mathbf{d}\left(\mathbb{X}_\lambda|\mathbf{Y}_{k-1}\right)$ and the likelihood is $\mathbf{L}\left(\mathbb{X}_\lambda\right) = \log\mathbf{l}\left(\mathbf{y}_k|\mathbb{X}_\lambda\right)$. The evolution of the probability distribution given by equation (6.2.2) in the pseudo-time is known as *log-homotopy* (Daum and Huang 2008). As it can be seen in figure 6.1 (**a**), the task of this homotopy is to move the particles through a sequence of densities from the prior to the posterior as λ continuously increases from zero to one. As it can be observed in figure 6.1 (**b**), it becomes necessary to find a flow $\frac{\mathrm{d}\mathbb{X}}{\mathrm{d}\lambda}$ that dictates the motion of particles. According to Daum and Huang (2013) this flow can be computed as:

$$\boldsymbol{\varrho}\left(\mathbb{X}_\lambda\right) = -\left[\frac{\partial^2\mathbf{\Gamma}\left(\mathbb{X}_\lambda\right)}{\partial\mathbb{X}_\lambda^2}\right]^{-1}\left[\frac{\partial\mathbf{L}\left(\mathbb{X}_\lambda\right)}{\partial\mathbb{X}_\lambda}\right]^{\mathrm{T}}, \tag{6.2.3}$$

where $\boldsymbol{\varrho}\left(\mathbb{X}_\lambda\right) = \frac{\mathrm{d}\mathbb{X}}{\mathrm{d}\lambda}$ represents a vector field that induces the motion of particles from the prior to the posterior. To solve equation (6.2.3) we first obtain the Hessian $\frac{\partial^2\mathbf{\Gamma}(\mathbb{X}_\lambda)}{\partial\mathbb{X}_\lambda^2}$ in closed form by differentiating twice equation (6.2.2) w.r.t. $\mathbb{X}_\lambda$:

$$\frac{\partial^2\mathbf{\Gamma}\left(\mathbb{X}_\lambda\right)}{\partial\mathbb{X}_\lambda^2} = \frac{\partial^2\mathbf{D}\left(\mathbb{X}_\lambda\right)}{\partial\mathbb{X}_\lambda^2} + \lambda\frac{\partial^2\mathbf{L}\left(\mathbb{X}_\lambda\right)}{\partial\mathbb{X}_\lambda^2}. \tag{6.2.4}$$

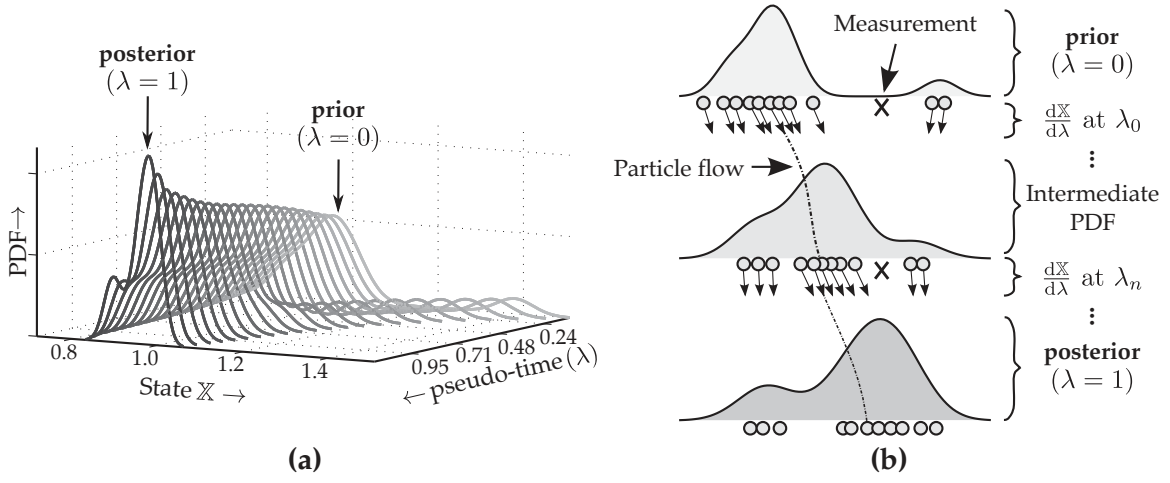

Figure 6.1.: Evolution of the probability distribution from the prior at $\lambda = 0$ to the posterior at $\lambda = 1$ (**a**) and particle flow at different values of λ (**b**).

In this work we use a hybrid approach for computing equation (6.2.4) in which the Hessian $\frac{\partial^2 \mathbf{D}(\mathbb{X}_\lambda)}{\partial \mathbb{X}_\lambda^2}$ is approximated by:

$$\frac{\partial^2 \mathbf{D}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda^2} \approx -\hat{\mathbf{S}}_{N_\mathbb{X}}^{-1}, \tag{6.2.5}$$

where $\hat{\mathbf{S}}_{n_\mathbb{X}}$ is the sample covariance matrix (SCM) of the prior distribution computed from the set of $n_\mathbb{X}$ particles. The SCM offers an unbiased estimate of the true covariance matrix. However, it has to be noted that if the number of particles employed is smaller than the number of states to be estimated the SCM may suffer from high variance. To overcome this issue the Kronecker product expansion can be used to estimate the covariance matrix in high dimensional spaces (Tsiligkaridis and Hero 2013). If it is assumed that the prior $\mathbf{d}\left(\cdot\right)$ is represented by a Gaussian distribution, then the approximation given by equation (6.2.5) is exact. For practical purposes the likelihood function $\mathbf{l}\left(\cdot\right)$ can be assumed to follow an univariate or a multivariate Gaussian distribution. Accordingly, $\mathbf{L}\left(\mathbb{X}_\lambda\right)$ is expressed as:

$$\mathbf{L}\left(\mathbb{X}_\lambda\right) = -\frac{n_\mathbb{X}}{2}\log\left(2\pi\right) - \frac{1}{2}\log\left|\mathbf{R}_n\right| - \frac{1}{2}\mathbf{z}_{k,\lambda}^{\mathrm{T}}\,{\mathbf{R}_n}^{-1}\mathbf{z}_{k,\lambda}, \tag{6.2.6}$$

where $\mathbf{z}_{k,\lambda} = \left(\mathbf{y}_k - \tilde{\mathbf{h}}\left(\mathbb{X}_\lambda\right)\right)$ and $\mathbf{R}_n$ is the covariance matrix of the measurement noise. Computing the gradient of equation (6.2.6) w.r.t. $\mathbb{X}_\lambda$ gives:

$$\frac{\partial \mathbf{L}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda} = \left[\frac{\partial \tilde{\mathbf{h}}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda}\right]^{\mathrm{T}} {\mathbf{R}_n}^{-1}\mathbf{z}_{k,\lambda} = \hat{\mathbf{H}}(\mathbb{X}_\lambda)^{\mathrm{T}}{\mathbf{R}_n}^{-1}\left(\mathbf{y}_k - \tilde{\mathbf{h}}\left(\mathbb{X}_\lambda\right)\right) \tag{6.2.7}$$

where $\hat{\mathbf{H}}\left(\mathbb{X}_\lambda\right)$ is the linearised output matrix around $\mathbb{X}_\lambda$. Computing the Hessian $\frac{\partial^2 \mathbf{L}(\mathbb{X}_\lambda)}{\partial \mathbb{X}_\lambda^2}$ might be computationally expensive. We instead approximate it by computing the expected Hessian by means of the Monte Carlo approximation method as follows:

$$\frac{\partial^2 \mathbf{L}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda^2} \approx \mathbb{E}\left[\frac{\partial^2 \mathbf{L}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda^2}\right] \approx -\frac{1}{n_\mathbb{X}}\sum_{i=1}^{n_\mathbb{X}}\left[\frac{\partial \mathbf{z}_{k,\lambda}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda^{\mathrm{i}}}\right]^{\top} {\mathbf{R}_n}^{-1}\frac{\partial \mathbf{z}_{k,\lambda}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda^{\mathrm{i}}}, \tag{6.2.8}$$

where $\mathbb{E}\left[\,\cdot\,\right]$ is the expected value with respect to the likelihood function. After having computed $\frac{\partial \mathbf{L}(\mathbb{X}_\lambda)}{\partial \mathbb{X}_\lambda}$ and $\frac{\partial^2 \mathbf{L}(\mathbb{X}_\lambda)}{\partial \mathbb{X}_\lambda^2}$ both equations (6.2.3) and (6.2.4) can be evaluated in order to obtain the particle flow. As it can be seen, evaluating equation (6.2.4) requires computing the inverse of $\hat{\mathbf{S}}_{n_\mathbb{X}}$, which can lead to numerical problems if $\hat{\mathbf{S}}_{n_\mathbb{X}}$ is close to be singular. To overcome this issue we apply the matrix inversion lemma known as Woodbury's formula in order to invert equation (6.2.4) as follows:

$$\left[\frac{\partial^2 \mathbf{\Gamma}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda^2}\right]^{-1} = -\hat{\mathbf{S}}_{n_\mathbb{X}} - \hat{\mathbf{S}}_{n_\mathbb{X}}\lambda\frac{\partial^2 \mathbf{L}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda^2}\left(\mathbf{I} - \hat{\mathbf{S}}_{n_\mathbb{X}}\lambda\frac{\partial^2 \mathbf{L}\left(\mathbb{X}_\lambda\right)}{\partial \mathbb{X}_\lambda^2}\right)^{-1}\hat{\mathbf{S}}_{n_\mathbb{X}}. \tag{6.2.9}$$

Note that the rate at which $\lambda_{0\rightarrow 1}$ is determined by the step size $\Delta\lambda$. Numerical experiments presented by Daum and Huang (2013) show that employing a fixed step size, such as in the case of the Euler method, works properly just if the number of particles

is high. Therefore, to reduce the number of particles employed a variable $\Delta\lambda$ has to be used. A proper strategy is to use a very small value of $\Delta\lambda$ at the beginning and to gradually increase it as $\lambda \to 1$, which makes sense since the uncertainty a the beginning of the measurement update step is higher. We therefore use an exponentially increasing step size (George and Powell 2006) given by:

$$\Delta\lambda = 1 - \frac{1}{n_{\lambda}^{b_{\lambda}}}, \tag{6.2.10}$$

where n_{λ} is the number of iteration and $b_{\lambda} \in \left(\frac{1}{2}, 1\right]$.

6.3. Experimental results

To evaluate the performance of the PFPF we first perform a purely SoC estimation, that is, no joint-state space is considered and therefore, the cell parameters previously identified are used. From the discussion presented in chapter 4 the state vector, at discrete time k, is given by:

$$\mathbf{x}_k = \begin{bmatrix} w_{1,k} & w_{2,k} & v_{\mathrm{s},k} & v_{\mathrm{l},k} \end{bmatrix}^{\mathrm{T}}. \tag{6.3.1}$$

Before further detailing the results achieved in this section, the choice of the covariance matrices $\mathbf{R}_{v,k}$ and $\mathbf{R}_{n,k}$ has to be properly discussed, since they strongly influence the convergence rate and the accuracy of the SoC estimation. The measurement noise $\mathbf{n}_k$ is given by an scalar, since V_{batt} is the only output of the system (recall equation (4.5.1)). The variance of the measurement noise is given by the measurement accuracy of the cycler, which in this case is $\pm 10\,\mathrm{mV}$. To determine $\mathbf{R}_{v,k}$ we assume that it is given by a diagonal matrix. Thus, it is only needed to determine four variance parameters, one belonging to the corresponding state in vector in (6.3.1). To this aim we set near to zero all variances belonging to $w_{2,k}$, $v_{\mathrm{s},k}$, $v_{\mathrm{l},k}$, so that the total a-priori knowledge about the transient behavior of the cell can be used while estimating the SoC. Consequently, the variance related to $w_{1,k}$ remains as the only parameter that influences the convergence behavior of the estimated SoC, as in the case of equation (4.2.3). A relative high value about this variance means that the value computed by the KiBaM is not entirely trusted, which gives more importance to the behavior of the terminal voltage. This is beneficial since it allows in this manner to converge to the true SoC even if a wrong initialization has been done.

The reference signal used to asses the accuracy of the estimation is the SoC computed by the cycler through Coulomb counting. Figure 6.2 (**a**) presents the load profile employed during the experiments. The load profile stems from real measurements gathered with the electric vehicle and has been properly down scaled to the range corresponding to a single cell. The experimental results discussed are based on tests performed using the second experimental setup described in section 2.3. The accuracy of the SoC estimation is assessed by means of the RMSE between the reference and the estimated SoC. As it can be appreciated in figure 6.2 (**b**), by an operating temperature of 25 °C the approach introduced throughout this chapter achieves an RMSE of 0.9 %. Figure 6.2 (**c**) details the evolution of the estimation error on time.

Table 6.1.: Results of the SoC estimation of a single-cell under different temperatures.

Temperature	6 °C	12 °C	15 °C	23 °C	25 °C	33 °C	38 °C
RMSE [%]	2.35	0.91	0.52	0.90	0.90	2.02	2.20

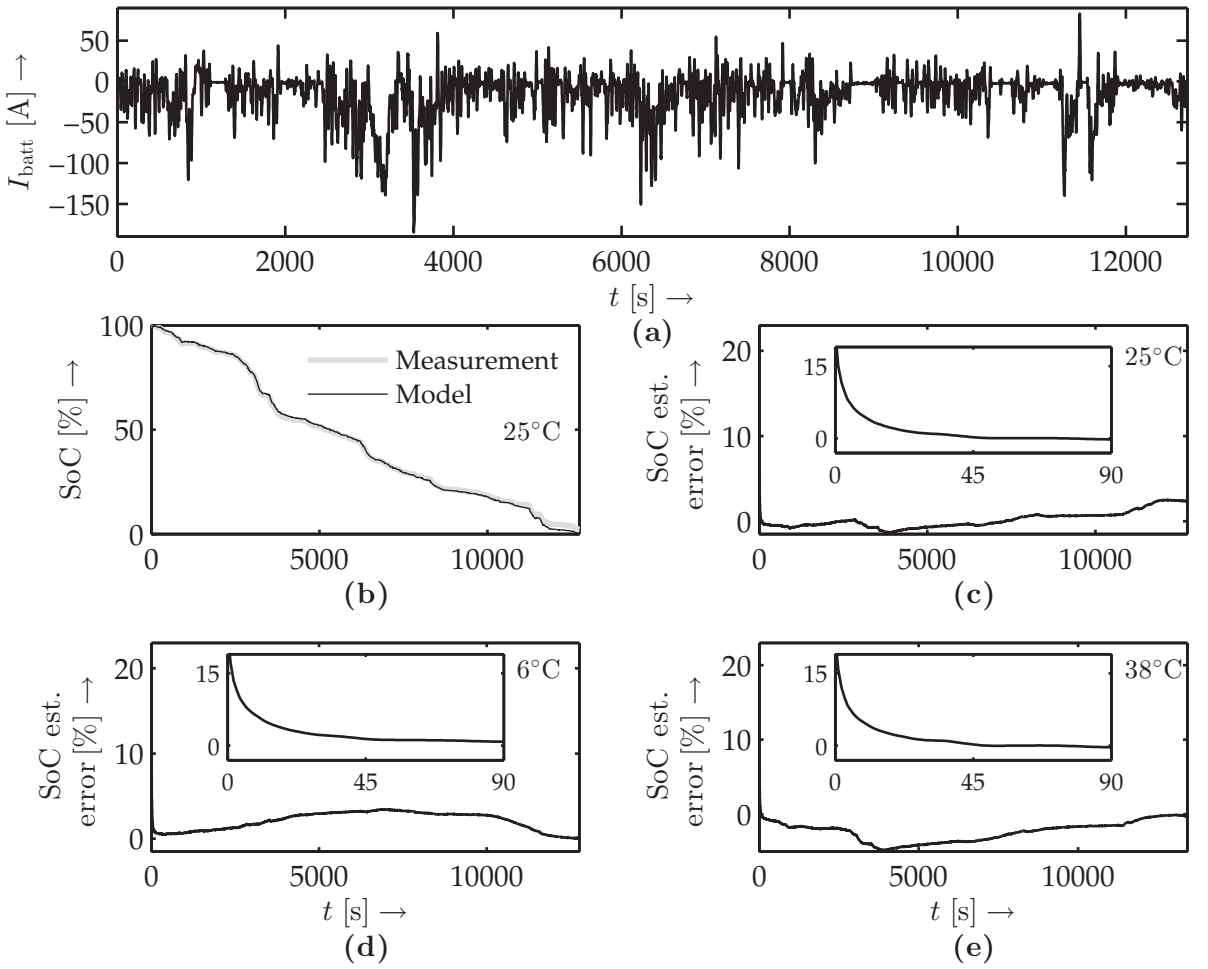

Figure 6.2.: Loading profile used for validating the SoC estimation (**a**). Behavior of the SoC estimation as function of time (**b**), error of the estimation at 25 °C (**c**), 6 °C (**d**) and 38 °C (**e**).

As it can be observed in the zoomed region, the estimation converges to the true SoC within 90 iterations, which is a reasonable convergence rate if we consider that typical applications of electric vehicles lie between 30 min and 3 h. During all experiments presented here the present state uncertainty is captured with 50 particles, that is $n_{\text{x}} = 50$. This number of particles represents a fair trade-off between estimation accuracy and computational complexity. Moreover, it is desired to employ a reduced number of particles during the SoC estimation, because these represent the starting point of the prediction step of the RDR estimation, as explained in section 2.2. For the sake of completeness, table 6.1 summarizes the RMSE achieved while estimating the SoC under 6 °C, 12 °C, 15 °C, 23 °C, 25 °C, 33 °C and 38 °C. Note, that the lowest accuracy is obtained at 6 °C (see figures 6.2 (**d**)), which is reasonable since the accuracy of the model underlying the SoC estimator decreases at low temperatures and therefore, negatively affects the performance of the PFPF.

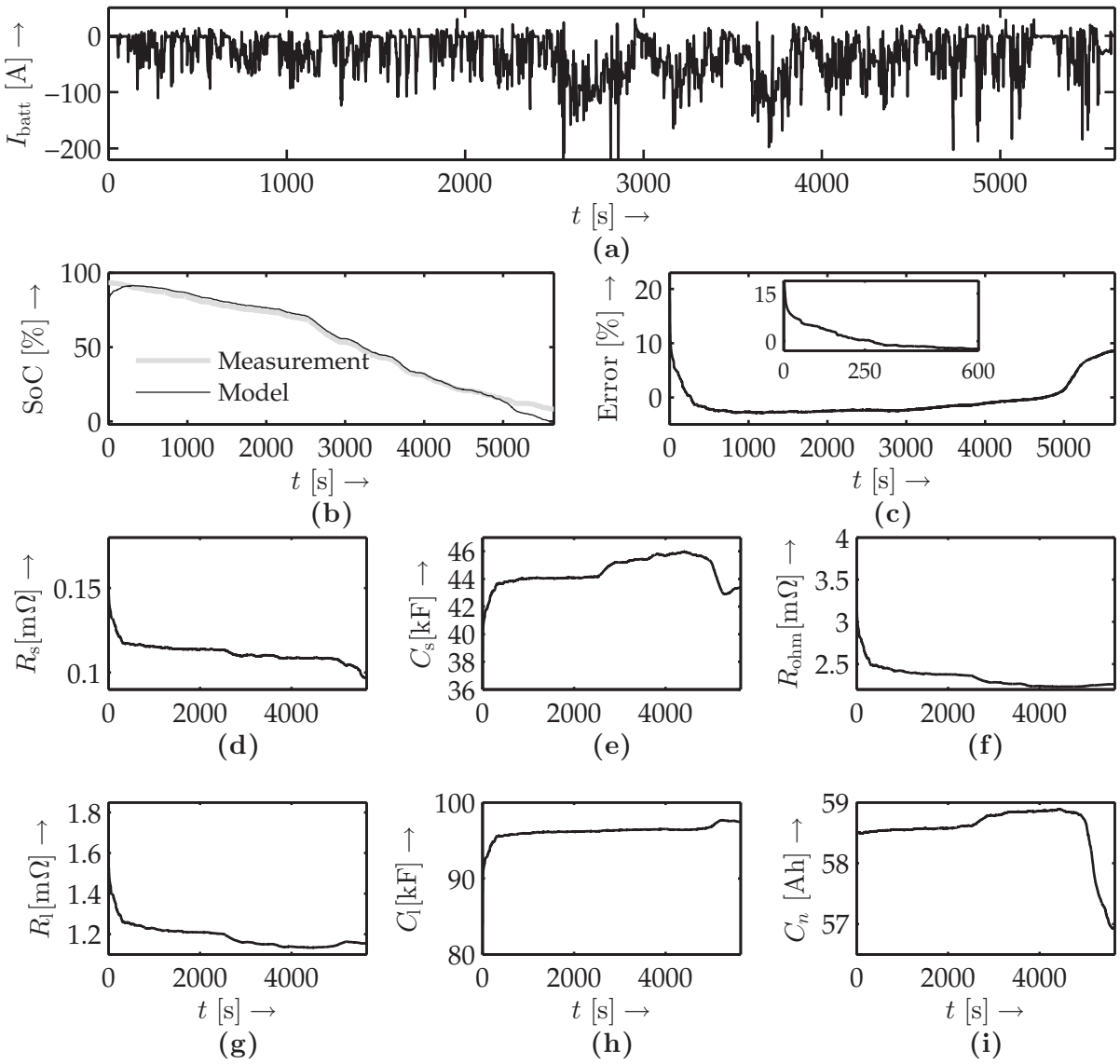

Figure 6.3.: Loading profile measured from the Nissan Leaf (**a**). Behavior of the SoC estimation (**b**) and its error (**c**) as function of time. Temporal evolution of the estimated parameters R_{s} (**d**), C_{s} (**e**), R_{ohm} (**f**), R_{l} (**g**), C_{l} (**h**) and C_n (**i**).

The experiments discussed so far are based on a purely state estimation, that is, the parameters are not estimated. These parameters, however, correspond to the cells identified in the cycler and, even though the cells are of the same type as the ones installed in the battery pack of the Nissan Leaf, they might not be directly applicable to other cells. To this aim we augment the state vector, as described in section 6.1, by $\boldsymbol{\theta}_k = [\ R_{\text{ohm},k}\ \ R_{\text{s},k}\ \ C_{\text{s},k}\ \ R_{\text{l},k}\ \ C_{\text{l},k}\ \ C_{n,k}\ ]^{\text{T}}$. Analogous to $\mathbf{R}_{v,k}$, we assume $\mathbf{R}_{\theta,k}$ to be a diagonal matrix whose diagonal elements are also set near to zero, with exception of the variance related to C_n. The results of the SoC estimation is shown in figures 6.3 (**b**) and (**c**). In this case a RMSE of 2.23 % is achieved over the entire operating range of the battery pack. In this case the reference signals is the SoC estimated by the Nissan Leaf. Figures 6.3 (**d**)-(**i**) show the temporal behavior of the estimated parameters, which resembles the behavior of the identification results shown in figure 4.6. Note, how C_n adapts on time such that the SoC remains properly estimated.

7

Operating Conditions Discrimination

Adapting the RDR estimation to changes in the operating conditions, from now on referred to as driving conditions, requires discriminating them as they occur. The discrimination is a function $\mathbf{q} : \mathbf{\Psi} \subseteq \mathbb{R}^{N_\Psi} \rightarrow s \in \mathbf{S}$, that allows mapping instances of *features* $\boldsymbol{\psi} \in \mathbf{\Psi}$, extracted from measurements of speed v_x and acceleration a_x, to the corresponding driving condition $s \in \mathbf{S}$, where $\mathbf{S}$ is the set of driving conditions. The first part of this chapter deals with the extraction and selection of the feature vector $\boldsymbol{\psi}$, which later serves as the basis for the classifier $\mathbf{q}(\cdot)$ developed for discriminating driving conditions. Finally, this chapter discusses the obtained experimental results.

7.1. Driving conditions classification

A driving condition is made up of environmental factors such as the traffic density and the road type. Properly classifying driving conditions is the first step towards modeling the environment. The choice of proper driving conditions, herein also interchangeably referred to as class, is an important aspect of this chapter, since they constitute the basis for predicting the power requirements of the electric vehicle. Therefore, they must be determined so that differ from each other, as much as possible, in regards to their influence on the power requirements. Andre (1996, 2004) provide a comprehensive work on classifying driving conditions, from the point of view of emissions and fuel requirements of conventional vehicles, based on realistic driving data gathered in Germany, Austria, Switzerland and Italy. Huang, Tan, and He (2011) demonstrate that similar studies can be carried out for electric vehicles, since such classification primarily depends on the speed, the acceleration and the stop times of the vehicle, and therefore, is independent of the powertrain topology of the vehicle. Similarly to the works previously mentioned, in this work the classification of the driving conditions is proposed as *stop and go*, *urban*, *rural* and *highway*. The first class describes a driving situation influenced by a high traffic density and consequently with frequent stops, as it occurs for example during rush hours on the city or on the highway. The remaining classes basically describe the road category under the assumption that the vehicle travels under a fluent traffic flow.

Figures 7.1 (**a**) to (**h**) present the probability density functions of the speed v_x and the probability density function of the power requirements P_ele for each of the proposed driving conditions, together with their mean and standard deviation, i.e., $\mathbb{E}_{v_\mathrm{x}}$

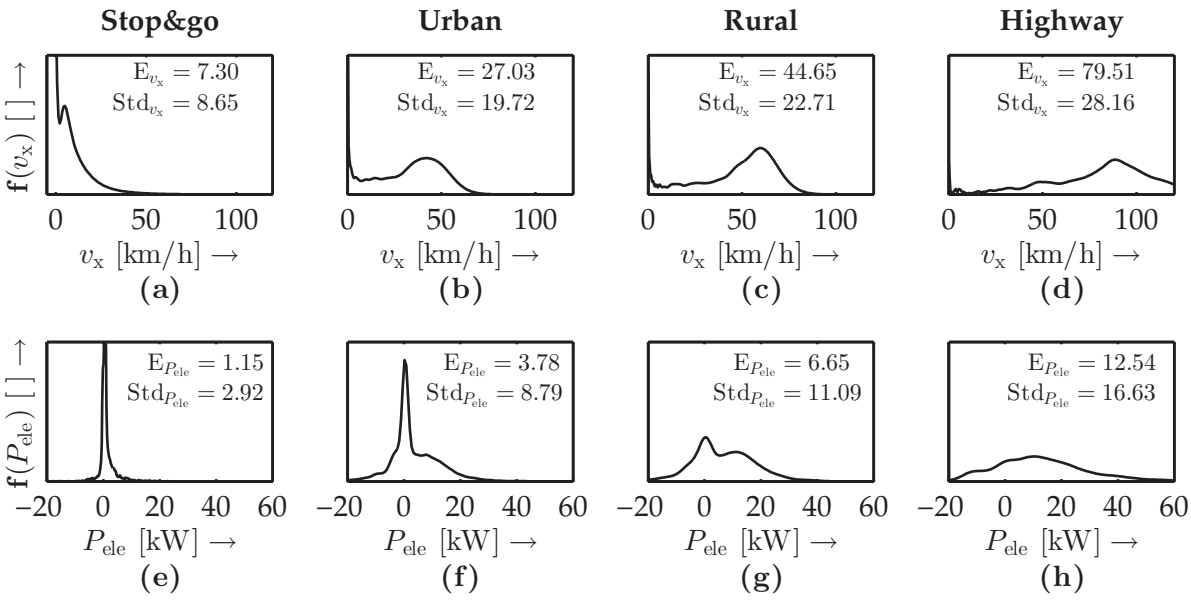

Figure 7.1.: PDFs of the speed (**a**)-(**d**) and road slope (**e**)-(**h**) of the different driving conditions.

Std_{v_x} and $\mathbb{E}_{P_{ele}}$ $\mathrm{Std}_{P_{ele}}$, respectively. As it can be appreciated, the first two statistical moments as well as the shape of all probability densities clearly differ from each other. In addition to this, the flatness of the distributions continuously increases from the class stop and go to the highway driving condition. Noteworthy is the power gained by means of regenerative braking observed among all distributions. As it was mentioned in section 3.3, regenerative braking drops off at low speeds, which becomes clear for stop and go driving conditions (see figure 7.1 (**b**)). For the sake of example, figures 7.2 (**a**) and (**b**) show a segment of measured speed and power profiles, which are superimposed on the corresponding class.

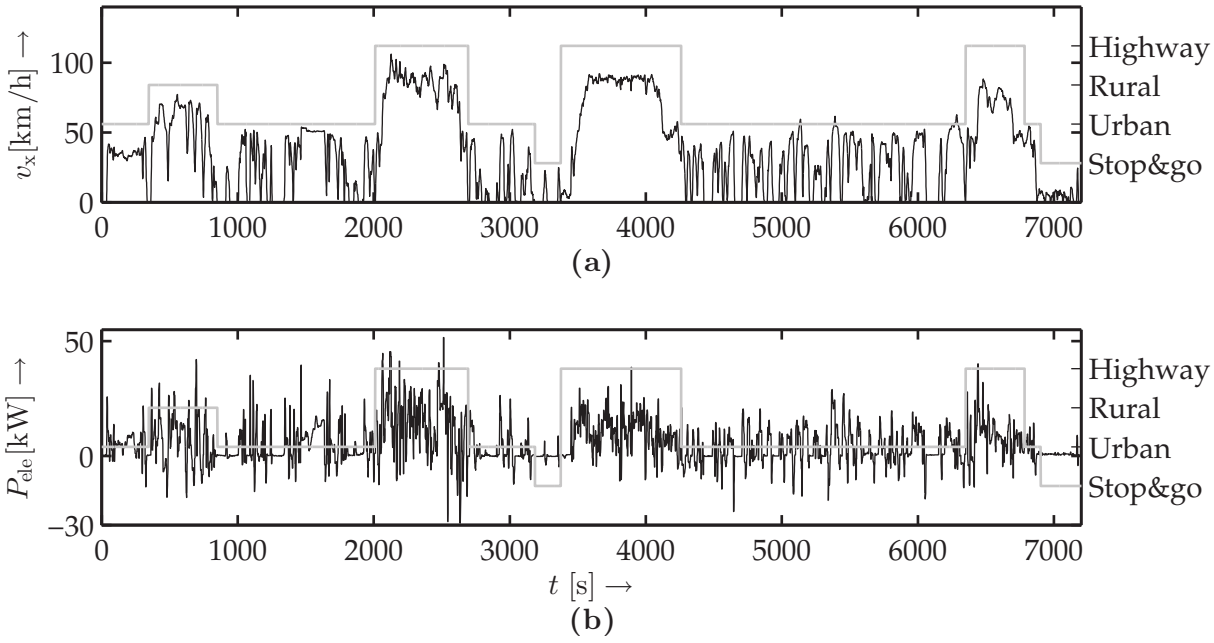

Figure 7.2.: Speed (**a**) and power (**b**) profiles juxtaposed to the underlying driving conditions.

In general terms, the proposed classification allows distinguishing between speed profile and the power requirements. Moreover, these classes are directly interpretable and can be easily understood. For these reasons, the set:

$$\mathbf{S} = \{s_1, ..., s_m\} = \{\text{Stop and go}, \text{Urban}, \text{Rural}, \text{Highway}\}\,, \tag{7.1.1}$$

with $m = 4$, is used from now on to define finite state space of driving conditions.

7.2. Feature extraction and selection

Feature extraction deals with the selection and computation of those characteristics of the original data, e.g., speed and acceleration measurements, with the aim of differentiating, as good as possible, between classes or clusters. Many researchers have developed feature extraction approaches which differ from each other, depending on whether they are performed manually or automatically or if they are formulated on the time or frequency domain (Theodoridis and Koutroumbas 2009, Chapters 6 and 7). This work favors the use of a manual approach, since it allows having a better interpretation and understanding about the selected features. For example, it is intuitively clear that the ratio between the number of stops per kilometer or that the idle time on the highway under normal traffic conditions is significantly lower than in urban traffic. In the literature, there exist a large variety of possible features, which can be extracted in the context of driving conditions. The reader is referred to the studies of Montazeri-Gh, Fotouhi, and Naderpour (2011), Lee, Adornato, and Filipi (2011), Huang, Tan, and He (2011) or Barlow et al. (2009), to obtain an idea of such features.

An overview of the set of features $\mathbf{\Psi}$ considered in this work is given in table C.1. It is noteworthy that many of these features are redundant and, thus, do not really help to improve the classification results. For this reason, a *feature reduction* step is carried out in the next section with the aim of systematically choosing the subset $\boldsymbol{\psi}$ including the most relevant features. As it can be seen, all features have a physical interpretation and can be roughly divided into four groups, namely, based on speed, acceleration, driving characteristics and cumulative parameters, e.g., traveled time or distance. For the sake of example and verification, let us compute all features of table C.1 for the urban, the rural road and the motorway driving cycles developed during the European ARTEMIS project (Andre 2004). Figure 7.3 (**a**) shows as example the scatter diagram of a pair of features plotted against each other. As it can be appreciated, the driving conditions labeled as urban, rural and highway, are properly classified (black centroids) by means of an unsupervised clustering approach, in this case k-means. In this case the experimental data is quantized using measurements (micro-cycles) with a window length $W_k = 600\,\text{s}$. However, if W_k is reduced, for example, to $W_k = 200\,\text{s}$, the arrangement of features changes, which might negatively influence the classification results (see figure 7.3 (**b**)). The use of unsupervised clustering might cause the selected classes to be chosen, such that they do not correspond to any of the driving conditions defined in section 7.1 and eventually with lack of any possible interpretation. Therefore, this work focuses on a supervised classification approach.

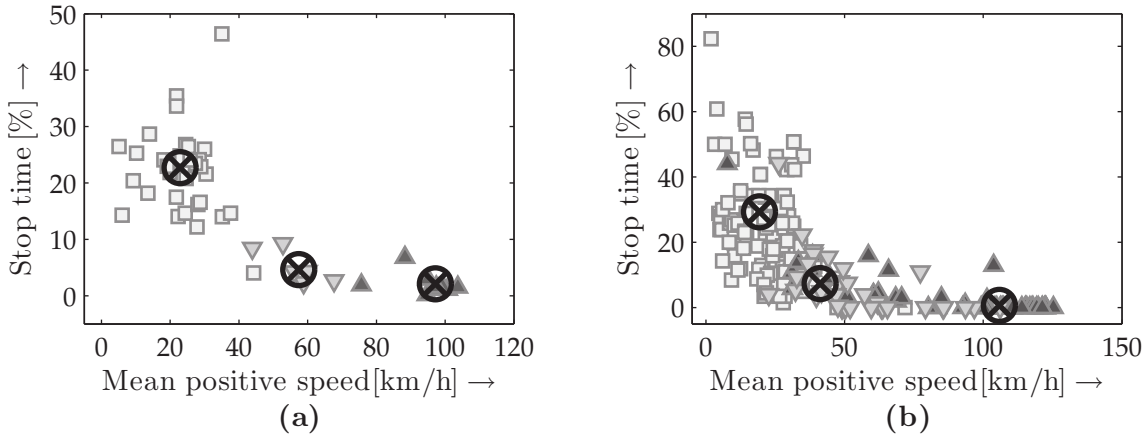

Figure 7.3.: K-means clusters based on two features with $W_k = 600$ s (**a**) and $W_k = 200$ s (**b**) for the classes urban (squares), rural (downward triangles) and highway (upward triangles).

Dimensionality reduction

Properly reducing the dimension of the feature vector $\mathbf{\Psi}$ eases the tasks of data visualization and analysis, reduces training times and improves the classification performance (Guyon and Elisseeff 2003). In the literature dealing with supervised learning for feature selection both filter and wrapper based methods stand out. Whereas filter based approaches rely on rankings obtained with correlation coefficients, e.g., Pearson, to select the most influential features, their wrapper based counterpart evaluates subsets of features, allowing in this manner to asses for possible correlations. Due to its simplicity, easiness of implementation and the quality of the classification results discussed in section 7.4, the *scalar feature selection* (Theodoridis and Koutroumbas 2009, p. 181) is employed in this work. This method introduces the criterion function $\mathbf{C}_l(\cdot), l = 1, \ldots, N_\Psi$, which individually evaluates the classification capacity of each of the features included in $\mathbf{\Psi}$. Basically, any criteria for measuring separability can be used to this aim, e.g., divergence or scatter matrices. In this work the one-dimensional divergence is used. The divergence between two classes i and j is given by:

$$d_{\mathrm{s},ij} = \frac{1}{2}\left(\frac{\sigma^2_{\mathrm{s},j}}{\sigma^2_{\mathrm{s},i}} + \frac{\sigma^2_{\mathrm{s},i}}{\sigma^2_{\mathrm{s},j}} - 2\right) + \frac{1}{2}\left(\mu_{\mathrm{s},i} - \mu_{\mathrm{s},j}\right)^2\left(\frac{1}{\sigma^2_{\mathrm{s},i}} + \frac{1}{\sigma^2_{\mathrm{s},j}}\right), \tag{7.2.1}$$

where $\mu_{\mathrm{s},j}$ is the mean $\sigma^2_{\mathrm{s},j}$ is the variance of the j^{th} class. The divergence separability measure computed with equation (7.2.1) is just valid under the assumption that all features per class are normal distributed. In the case of non-Gaussian feature distributions, a small error in $\left(\mu_{\mathrm{s},i} - \mu_{\mathrm{s},j}\right)^2$ would cause significant errors in $d_{\mathrm{s},ij}$. Theodoridis and Koutroumbas (2009, p. 177) conclude, however, that such error has no remarkable influence in the classification performance and suggest that with the transformation $\hat{d}_{\mathrm{s},ij} = 2\left(1 - \exp\left(-\frac{d_{\mathrm{s},ij}}{8}\right)\right)$ a more robust separability measure is obtained. Thus, the selection criterion is given by:

$$\mathbf{C}_l = \min_{i,j} \hat{d}_{\mathrm{s},ij}, \qquad l = 1, \ldots, N_\Psi. \tag{7.2.2}$$

The result of the feature selection is the vector $\boldsymbol{\psi} = [\psi_{\mathrm{ind}_1}, \ldots, \psi_{\mathrm{ind}_L}]$ where the set $\mathrm{ind} = \{\mathrm{ind}_1, \ldots, \mathrm{ind}_{N_\Psi}\}$, $L \leq N_\Psi$ describes the indexes of the N_Ψ features. The

Table 7.1.: Results of the feature selection with both time- and distance-based measurements.

#	Time-based		Distance-based	
	Feature	Nr.	Feature	Nr.
1	Maximum speed	14	Maximum speed	14
2	Mean speed (without stops)	10	Total duration	2
3	Stops per kilometer	12	Mean acceleration	15
4	Mean speed (including stops)	9	Mean speed (without stops)	10
5	Acceleration per kilometer	11	Mean speed (including stops)	9

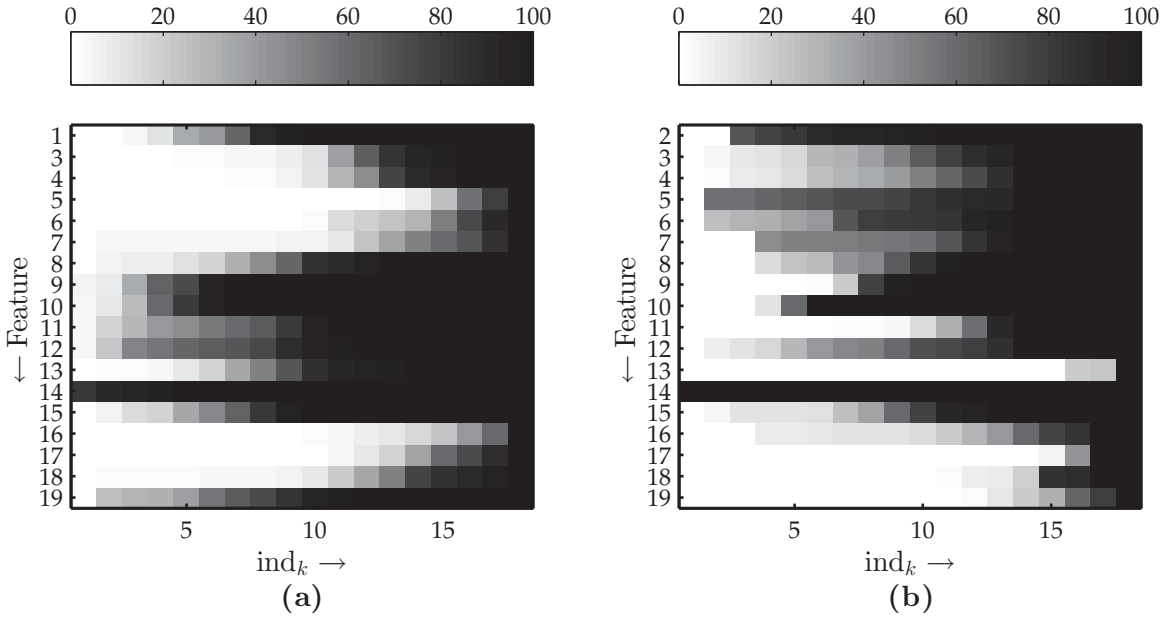

Figure 7.4.: CDF of the feature selection with time- **(a)** and distance-based (**b**) measurements.

figures 7.4 (**a**) and (**b**) show the results obtained with several thousands of evaluations of the selection approach with random mixtures of features and for different lengths W_k of both time and distance-based measurements. The ordinate corresponds to the indexes of all possible features of table C.1, with exception of the total duration of the of the micro-cycle. This feature ranges between 100 s and 500 s and, since it remains constant at each evaluation, it leads to singular matrices. Because of this reason, it is removed from $\mathbf{\Psi}$. In the case of distance-based selection, this is also valid for the feature related to the total traveled distance. The abscissa represents the index ranking k of the most relevant features. The further left a high cumulative probability (dark gray tone) is achieved, the more important is the associated feature. To better rank the features let us analyze the discrete cumulative density function $\mathbf{F}_i(\text{ind}_L)$ for all features. In this context, the $\mathbf{F}_i(\text{ind}_L)$ is defined as

$$\mathbf{F}_i(\text{ind}_L) = \sum_{\text{ind}_k=1}^{\text{ind}_L} (\text{ind}_k), \quad i = 1, \ldots, N_\Psi. \tag{7.2.3}$$

In this work the features with the L highest values are employed. Table 7.1 shows the five best features used in with the classifier introduced in section 7.3. As it can

be seen, the selected features allow high separability, which is why they seem to be a plausible choice. Although a reduction down to five features improves the computational performance, the minimum number of features needed to achieve optimal classification results is investigated and properly defined in section 7.4.

7.3. Online driving condition classification

To allow the RDR estimation to account for changing driving conditions, it is necessary to classify them as they occur. Gathering training data for the classifier represents a challenge, since it is required to operate the mobile system for long periods of time under different operating conditions. This is not feasible in our case due to the constraints in energy availability and the long charging times imposed by the electric vehicle. Because of this, a classifier with low expected generalization error and capable of responding fast to time-varying influences is investigated in this section.

Kotsiantis (2007) juxtaposes different supervised learning classification methods and concludes that Support Vector Machines (SVMs) offer good results for classification problems dealing with multiple dimensions and with continuous features, as in our case. Training the classical SVM, first introduced by Cortes and Vapnik (1995), is relatively slow since it requires solving a convex optimization problem, making it more difficult to implement in an adaptive manner and therefore less suitable for online applications (Kuh 2004). There exist, however, some approaches that avoid using quadratic programming methods (Suykens 2002; Theodoridis and Koutroumbas 2009). Suykens and Vandewalle (1999) formulates a modified least squares SVM, also referred to as LS-SVM, whose training is achieved by solving a set of linear equations and that serves as basis for the classification scheme introduced in the sequel.

LS-SVM based driving condition classification

Let $D_{N_T} = \{\boldsymbol{\psi}_i, v_i\}_{i=1}^{N_T}$ be a set of N_T labeled samples, i.e., training set, where $\boldsymbol{\psi}_i \in \mathbb{R}^{n_\psi}$ is the vector of features in the n_ψ-dimensional feature space and $v_i \in \{-1, 1\}$ is the corresponding binary class label of the i^{th} instance. Furthermore, let us define a discriminant function of the form:

$$\mathbf{q}(\boldsymbol{\psi}) = \mathbf{w}^{\mathrm{T}} \boldsymbol{\phi}(\boldsymbol{\psi}) + \mathcal{B}, \tag{7.3.1}$$

where $\boldsymbol{\phi}\colon \mathbb{R}^{N_T} \to \mathbb{R}^{N_h}, \boldsymbol{\psi} \mapsto \boldsymbol{\phi}(\boldsymbol{\psi})$ with $N_h \leq N_T$ is a nonlinear function that maps the training dataset to a high dimensional feature space $\mathbb{R}^{N_h}$, in which linear separability between samples may exist, $\mathbf{w} \in \mathbb{R}^{N_h}$ is the normal vector to the hyperplane and $\mathcal{B}$ is a bias term, also known as threshold value.

The LS-SVM classification problem is formulated as:

$$\begin{aligned} \min_{\mathbf{w},\mathcal{B},e} \quad & \mathcal{J}(\mathbf{w},\mathcal{B},e) = \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{2}\gamma_{\text{re}} \sum_{i=1}^{N_T} e_i^2 \\ \text{s.t.} \quad & v_i \mathbf{q}(\boldsymbol{\psi}_i) = 1 - e_i, \qquad i = 1 \ldots N_T, \end{aligned} \tag{7.3.2}$$

where $e_i \in \mathbb{R}$ is a normal distributed error of the class label v_i and $\gamma_{\text{re}} \in \mathbb{R}$ is the regularization term that controls the trade-off between fitness error and model complexity,

i.e., the bias-variance dilemma. According to Suykens (2002), solving the optimization problem becomes more tractable by transforming it into its dual space. To this aim the Lagrangian functional of (7.3.2) is obtained and the Karush-Kuhn-Tucker (KKT) optimality conditions are applied, which yields the following linear system of equations:

$$\begin{bmatrix} 0 & \boldsymbol{v}^{\mathrm{T}} \\ \boldsymbol{v} & \boldsymbol{\Omega} + \frac{1}{\gamma_{\mathrm{re}}}\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathcal{B} \\ \boldsymbol{\alpha}_{\mathcal{L}} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}, \tag{7.3.3}$$

where $\boldsymbol{v} = [v_1, \ldots, v_{N_T}]^{\mathrm{T}}$, $\mathbf{1} = [1, \ldots, 1]^{\mathrm{T}}$ and $\boldsymbol{\alpha}_{\mathcal{L}} = \left[\alpha_{\mathcal{L},1}, \ldots, \alpha_{\mathcal{L},N_T}\right]^{\mathrm{T}}$ represents the Lagrange multipliers (support vectors). The Gram matrix $\boldsymbol{\Omega}$, contains the elements $\Omega_{ij} = v_i v_j \boldsymbol{\phi}(\boldsymbol{\psi}_i)^{\mathrm{T}} \boldsymbol{\phi}\left(\boldsymbol{\psi}_j\right) = v_i v_j \mathbf{K}\left(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j\right)$, where $\mathbf{K}\left(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j\right)$ is the kernel function. In this work we investigate two kernel functions (see section 7.4), namely:

$$\text{linear kernel}: \ \mathbf{K}\left(\psi_i, \psi_j\right) = \psi_i^{\mathrm{T}} \psi_j \tag{7.3.4}$$

$$\text{RBF kernel}: \ \mathbf{K}\left(\psi_i, \psi_j\right) = \exp\left(-\frac{\left\|\psi_i - \psi_j\right\|^2}{2\sigma_{\mathrm{RBF}}^2}\right). \tag{7.3.5}$$

The solution of 7.3.3 results in two vectors containing the Lagrange multipliers $\boldsymbol{\alpha}_{\mathcal{L}}$ and the bias term $\mathcal{B}$, respectively, and has a computational complexity of $\mathcal{O}(N_T^3)$ (Ojeda, Suykens, and De Moor 2008), which makes it suitable for online applications. Finally, for a new set of observed features $\boldsymbol{\psi}$, according to the Mercer's theorem, the LS-SVM classifier takes the form:

$$v\left(\boldsymbol{\psi}\right) = \mathbf{q}\left(\boldsymbol{\psi}\right) = \operatorname{sign}\left[\sum_{i=1}^{N_T} \alpha_{\mathcal{L},i} v_i \mathbf{K}\left(\boldsymbol{\psi}, \boldsymbol{\psi}_i\right) + \mathcal{B}\right]. \tag{7.3.6}$$

Multiple classes with unbalanced class distribution

The LS-SVM, as expressed by equation (7.3.6), is a binary-classifier which allows deciding, for example, whether the electric vehicle operates under an urban or a highway driving condition. However, in order to properly classify driving conditions for all $s \in \mathbf{S}$ it becomes necessary to extend equation (7.3.6) to a multi-class classifier. Basically, there exist two philosophies for multi-class classifiers based on SMVs. The former suggests combining multiple binary-classifiers, for example, *one vs. one* (OvO) or *one vs. all* (OvA) (Bishop 2006, p. 339). The latter is based on the idea of a single SVM to discriminate all classes, e.g., the Directed Acyclic Graph SVM (DAGSVM) (Platt, Shawe-Taylor, and Cristianini 1999). Hsu and Lin (2002) and Rifkin and Klautau (2004) evaluate these philosophies and conclude that simple approaches like OvO or OvA are preferable over schemes based on single SVMs, since it is more computationally efficient to solve multiple optimization problems instead of solving a single optimization routine. Due to the good classification results obtained (see section 7.4), in addition to the small computational effort required during the multi-class classification, the OvO is favored in this work.

One vs one: as it was presented in section 7.1, the number of driving conditions considered here is $m = 4$. Under the OvO strategy a total of $m(m-1)/2$ binary LS-SVM classifiers are trained in order to distinguish the samples of one driving

condition from the samples of all other driving conditions. The class is then decided based on the highest number of votes, where each LS-SVM classifier votes for only one driving condition. The main drawback with this approach is that the computational burden grows proportional with the number classes, which makes it computationally inefficient for $m \gg 0$. On the other hand, the OvO approach is relatively fast for small training datasets. The reason for this is that smaller systems of equations have to be solved, since only the samples of two driving conditions have to be considered.

A known problem of the LS-SVM, as formulated in equation (7.3.6), is that if the training data is skewed towards a particular class, then the performance of the classifier might reduced. This issue is important in this work, since the training data set employed (see section B.1) exhibits certain unbalanced class distribution in that urban driving conditions appear more frequently among all driving cycles. For example, during trips on rural areas or on the highway urban scenarios are present at the beginning and at the end of the trip. From the many methods reported in the literature for alleviating the issue of unbalanced data sets (see García et al. (2007) for a detailed survey), the approach implemented in this work is based on the study of Luts et al. (2010). Under this scheme the optimization problem formulated in equation (7.3.2) is weighted as:

$$\min_{\mathbf{w},\mathcal{B},e} \quad \mathcal{J}(\mathbf{w},\mathcal{B},e) = \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{2}\gamma_{\text{re}} \sum_{i=1}^{N_T} \varpi_i e_i^2 \quad , \tag{7.3.7}$$

where ϖ_i is an a-priori probability of occurrence computed by

$$\varpi_i = \begin{cases} \varpi_+ = \frac{N_T}{2N_{T^+}}, & \text{for } v_i = +1, \\ \varpi_- = \frac{N_T}{2N_{T^+}}, & \text{for } v_i = -1, \end{cases} \tag{7.3.8}$$

with N_{T^+} and N_{T^-} being the number of positive and negative classes, respectively. According to Luts et al. (2010), properly weighting is comparable to resampling the training data set such that $N_{T^+} = N_{T^-}$. Figures 7.5 (**a**) and (**b**) exemplify the problem of unbalanced class distribution and the solution achieved through a-priori weighting.

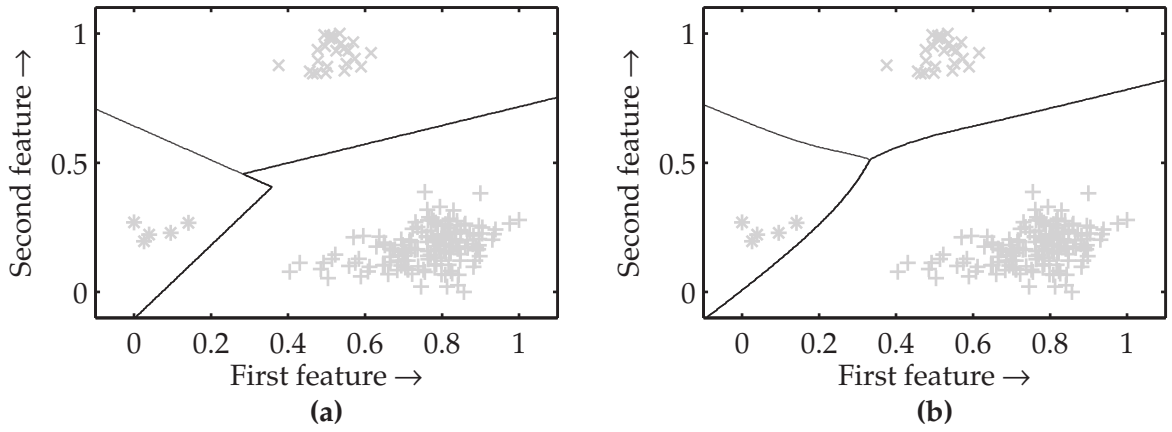

Figure 7.5.: Classification results without (**a**) and with (**b**) a-priori weighting based on a synthetic data of 300 elements separated in 3 classes.

Model selection

As it can be deduced by analyzing equations (7.3.3) and (7.3.6), the model complexity of the LS-SVM depends on at least one parameter, namely, the regularization term γ_{re}. This is only valid, in the case of using a linear kernel. If, however, a RFB kernel is employed, the parameter σ_{RBF} (see equation (7.3.5)) also needs to be properly defined. Choosing the value of σ_{RBF} has to be done carefully, since it has an enormous influence on the classification results. On the one hand, large σ_{RBF}-values prevent an adequate separation between classes, i.e., they cause under-fitting (see figure 7.6 (**a**)). Very small σ_{RBF}-values, on the other hand, cause an over-fitting to the training data (see figure 7.6 (**b**)). There are basically two strategies for choosing the aforementioned parameters, i.e., for *model selection*, namely, based on heuristics or on optimization. Heuristics-based approaches are straightforward to implement. Nevertheless, they are unreliable and highly sensitive to noisy data sets. Even though cross-validation (optimization) based methods are computationally prohibitive, they remain as one of the most widespread model selection techniques among machine learning practitioners.

The *leave-one-out* cross validation (LOO-CV) (Stone 1974) is a technique, which sets the basis of many model selection strategies applied in the context of SVMs. In the case of LS-SVM, the naïve implementation of the LOO-CV has a computational complexity of $\mathcal{O}(N_T^4)$. To overcome this limitation, Cawley (2006) derive an approximation of the LOO-CV of LS-SVMs which reduces the computational effort to $\mathcal{O}(N_T^3)$. An, Liu, and Venkatesh (2007), further improves this approach, namely to $\mathcal{O}(N_T^2)$, and proof that there exist an exact solution for the LOO-CV, which is used herein for selecting the model of the LS-SVM classifier described in this chapter.

For the sake of completeness, a formal description of the LOO-CV used in this work is briefly presented in the following.

Under a LOO-CV the training dataset D_{N_T} is split out into ℓ subsets $\{\psi_j^{(a)}, \upsilon_j^{(a)}\}_{j=1}^{N_a}$ of length N_a, with $a = 1, 2, ..., \ell$ and $\sum_{a=1}^{\ell} N_a = N_T$. Analogously, the Lagrange multipliers of equation (7.3.3) are subdivided into ℓ subsets $\boldsymbol{\alpha}_{\mathcal{L}} = [\boldsymbol{\alpha}_{\mathcal{L}}^{(1)^{\mathrm{T}}}, ..., \boldsymbol{\alpha}_{\mathcal{L}}^{(a)^{\mathrm{T}}}, ..., \boldsymbol{\alpha}_{\mathcal{L}}^{(\ell)^{\mathrm{T}}}]^{\mathrm{T}}$,

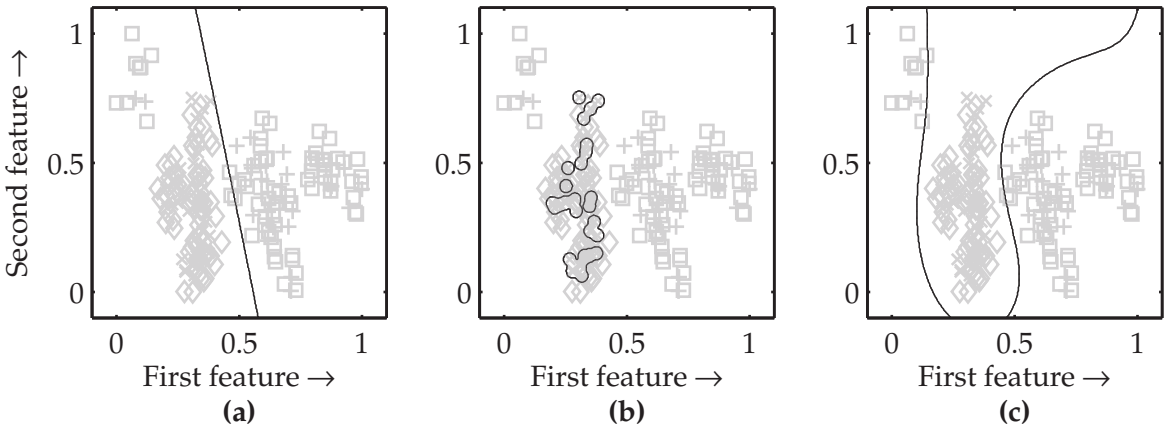

Figure 7.6.: Exemplification of classification results achieved with a RBF Kernel based SVM with an under-fitted ($\gamma_{\text{re}} = 1$ and $\sigma_{\text{RBF}} = 2$) (**a**), an over-fitted ($\gamma_{\text{re}} = 1$ and $\sigma_{\text{RBF}} = 0.01$) (**b**) and an optimized model ($\gamma_{\text{re}} = 1$ and $\sigma_{\text{RBF}} = 2$) (**c**).

where $\boldsymbol{\alpha}_{\mathcal{L}}^{(a)} = [\alpha_{\mathcal{L},1}^{(a)}, ..., \alpha_{\mathcal{L},j}^{(a)}, ..., \alpha_{\mathcal{L},N_a}^{(a)}]^{\mathrm{T}}$. Furthermore, let $\hat{v}^{(a)} = v^{(a)}(\psi)$ be the classifier obtained with equation (7.3.6) by leaving out the a^{th} subset of D_{N_T} and define $r_j^{(a)} = v_j^{(a)} - \hat{v}^{(a)}$ as the residual obtained during the a^{th} iteration. The task of the LOO-CV is to compute the vector of residuals $\mathbf{r}^{(a)} = [r_1^{(a)}, ..., r_j^{(a)}, ..., r_{N_a}^{(a)}]^{\mathrm{T}}$ so that $\hat{\boldsymbol{v}}^{(a)} = [\hat{v}_1^{(a)}, ..., \hat{v}_j^{(a)}, ..., \hat{v}_{N_a}^{(a)}]^{\mathrm{T}}$ for each of the ℓ subsets can be accordingly obtained.

To this aim let us denote $\mathbf{A}_{\gamma_{\text{re}}}$ as the system matrix of equation (7.3.3). Then, with help of the Woodbury inversion lemma the inverse $\mathbf{A}_{\gamma_{\text{re}}}^{-1}$ can be computed as follows:

$$\mathbf{A}_{\gamma_{\text{re}}} = \begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \boldsymbol{\Omega} + \frac{1}{\gamma_{\text{re}}}\mathbf{I} \end{bmatrix} \Rightarrow \mathbf{A}_{\gamma_{\text{re}}}^{-1} = \begin{bmatrix} \frac{1}{d_{\text{L}}} & -\frac{1}{d_{\text{L}}}\mathbf{1}^T\boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1} \\ -\frac{1}{d_{\text{L}}}\boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1}\mathbf{1} & \boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1} + \frac{1}{d_{\text{L}}}\boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1}\mathbf{1}\mathbf{1}^T\boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1} \end{bmatrix}, \tag{7.3.9}$$

which is valid since $\boldsymbol{\Omega}_{\gamma_{\text{re}}} = \boldsymbol{\Omega} + \frac{1}{\gamma_{\text{re}}}\mathbf{I}$ is positive definite and $d_{\text{L}} = -\mathbf{1}^T\boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1}\mathbf{1} \neq 0$. This inversion allows to directly compute $\boldsymbol{\alpha}_L$ from equation (7.3.3). Accordingly:

$$\boldsymbol{\alpha}_L = \left[\boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1} + \frac{1}{d_{\text{L}}}\boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1}\mathbf{1}\mathbf{1}^T\boldsymbol{\Omega}_{\gamma_{\text{re}}}^{-1}\right]\boldsymbol{v}. \tag{7.3.10}$$

Having computed $\boldsymbol{\alpha}_L$, the vector containing all residuals $\mathbf{r} = [\mathbf{r}^{(1)^{\mathrm{T}}}, ..., \mathbf{r}^{(1)^{\mathrm{T}}}, ..., \mathbf{r}^{(\ell)^{\mathrm{T}}}]^{\mathrm{T}}$ is determined by solving the system of equations given by $\text{diag}(\mathbf{A}_{\gamma_{\text{re}}}^{-1}\mathbf{r}) = \boldsymbol{\alpha}_{\mathcal{L}}$.

Finally, $\hat{\boldsymbol{v}}^{(a)} = [\hat{v}_1^{(a)}, ..., \hat{v}_j^{(a)}, ..., \hat{v}_{N_a}^{(a)}]^{\mathrm{T}}$ for all subsets is computed as follows:

$$\hat{\boldsymbol{v}}^{(a)} = \text{sign}\left[\boldsymbol{v}^{(a)} - \mathbf{r}^{(a)}\right], \tag{7.3.11}$$

with $a = 1, 2, ..., \ell$.

7.4. Classifier evaluation

With the help of $\hat{\boldsymbol{v}}$, the expected misclassification can be determined or a performance metric, based on $\mathbf{r}$, can be minimized. In this work, due to the small number of kernel parameters, namely γ_{re} and σ_{RBF}, we employ the grid search method, that is, we first set a proper parameter range and then, by means of an exhaustive search through all parameter combinations, the best parameter setting is defined.

We avoid just quantifying the number of miss classifications, as with the standard accuracy measure (ACC), since unbalanced classes are rejected in favor of an eventual dominant class. For this reason, the balanced accuracy (BAC), which evaluates the misclassification rate per class and averages it across all classes (see for example Sokolova and Lapalme (2009)), is instead employed. The BAC is given by:

$$\text{BAC} = \frac{1}{m}\sum_{i=1}^{m}\frac{(\text{tp}_i + \text{tn}_i)}{N_{T,i}}, \tag{7.4.1}$$

where m is the number of driving situations, N_T is the size of the training set, tp_i is the sum of all classifications of the i^{th} class, that are correctly assigned to that class and tn_i is the sum of all classifications that are correctly classified outside the i^{th} class.

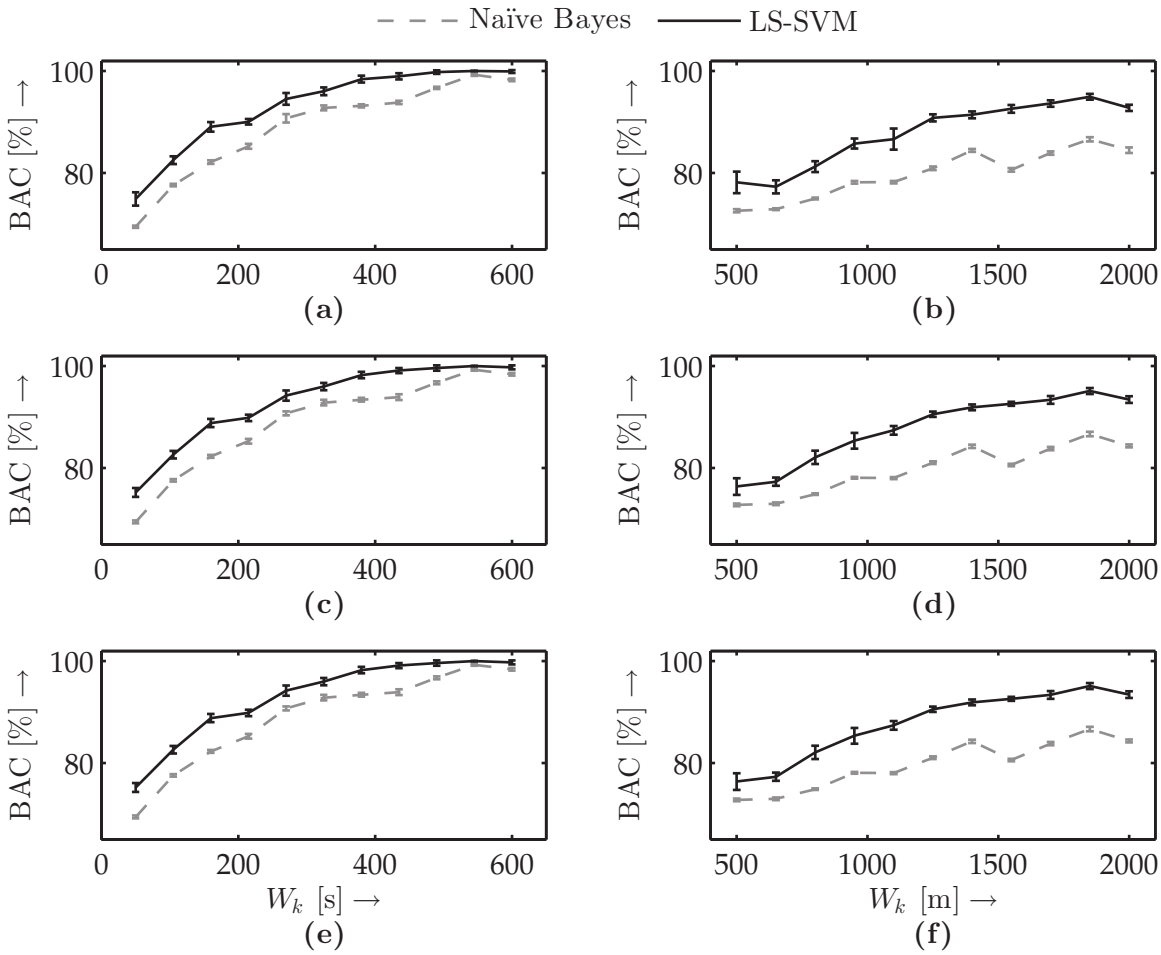

Figure 7.7.: Driving condition classification by using all features: with a RBF kernel and by not removing outliers with time (**a**) and distance-based measurements (**b**), with a RBF kernel and by removing outliers with time (**c**) and distance-based measurements (**d**), with linear kernel and by removing outliers with time (**e**) and distance-based measurements (**f**).

Figure 7.7 shows the classification results achieved using all features (see table C.1) with both the linear and the RBF kernel. The performance metric used to evaluate the classifier is the BAC. For comparison purposes a Bayesian classifier based on a kernel density estimator (Pérez, Larrañaga, and Inza 2009) has been also implemented and evaluated. The evaluation procedure consists of computing the classification error with a 5-fold cross validation in which the window length W_k (both in time and in space domain), used to compute the features in section 7.2, is varied to optimize the classification at each iteration. The training data has been randomly mixed and the entire process has been repeated 20 times, with the aim of computing the first two statistical moments.

As it can be observed in figures 7.7 (**a**) to 7.7 (**f**), all classifiers benefit from a higher window size. The maximum classification rate lies by 99% in the case of the RBF kernel and by using time-based measurements. The classification based on distance-based measurements, on the contrary, generally perform 5%-10% worst that their time-based counterpart. For this reason in this work, specially in chapter 8, the use of time-based measurements is favored for training and validating the stochastic en-

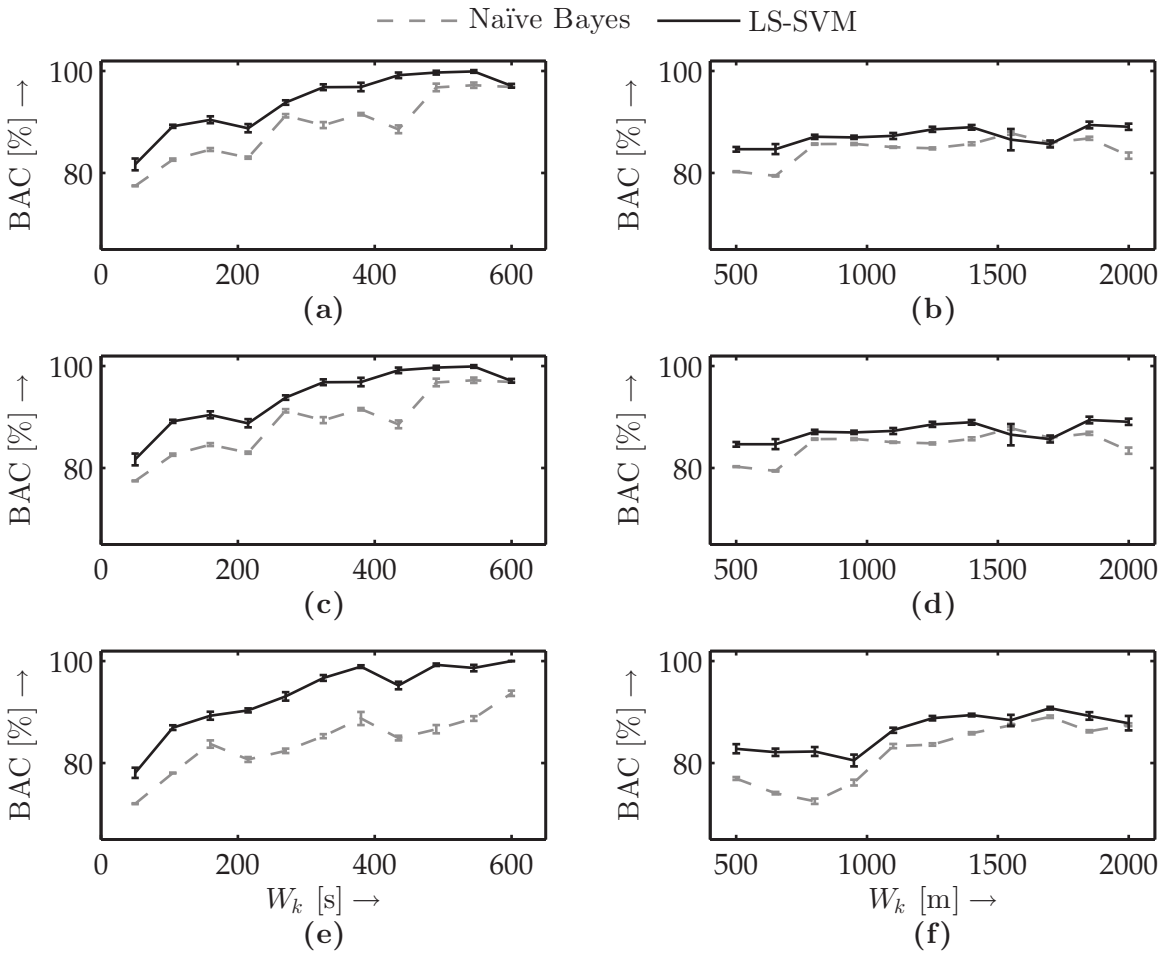

Figure 7.8.: Classification results by using: 5 features with a RBF kernel and by removing outliers with time (**a**) distance-based measurements (**b**), 5 features with a linear kernel and by removing outliers with time (**c**) distance-based measurements (**d**), with 3 features with a linear kernel and by removing outliers with time (**e**) distance-based measurements (**f**).

vironment model. Note, that eliminating outliers in the training data does not really improves the quality of the classification. Nevertheless, in this work this step is carried out to comply with good data treating practices. It is worth mentioning that the naïve Bayes classifier, although its requirements of normally distributed features is not satisfied, also presents acceptable results, however, the LS-SVM outperforms it in most cases.

The effect of the feature reduction, introduced in section 7.2, on the classification results is shown in figures 7.8 (**a**) to 7.8 (**f**). Here, the maximum classification rate lies by 95% in the case of a linear kernel by using just the best three features of table 7.1. The classifier based on a RBF kernel also exhibit superior results, however, the computation effort needed during the classification is higher as in the linear case.

Base on these results it has been decided to employ a classifier with a linear kernel and that uses only three features. Although it has been shown throughout this section that the larger the window size, the better classification results, the choice of this parameter depends on other factors that belong to the stochastic environment model. Accordingly, W_k is investigated and properly determined in chapter 8.

8

Stochastic Environment Modeling

With the aim of predicting the power consumption of the mobile system it is necessary to capture the stochastic behavior dictated by the temporal evolution of the operating environment. The focus of this chapter lies on modeling the environment from a Markovian point of view for both the evolution of the operating conditions and for actual interaction between the system and its surroundings. For the sake of clarity, in the following the operating conditions are given by the driving conditions, as it was already introduced in chapter 7, and the system-surroundings interaction is described by the driving behavior, characterized by the speed and acceleration along the longitudinal axis of the vehicle, and by the slope profile of the road.

8.1. Driving conditions modeling

Stochastic processes with discrete states, both in continuous or discrete time, can be successfully described with of Markov chains (Waldmann and Stocker 2013). Markov chains have been extensively used for generating synthetic drive cycles. For example, Lee, Adornato, and Filipi (2011) employ discrete-time discrete-state Markov chain to model the time evolution of the speed and acceleration of driving cycles. The model developed herein is built around this concept due to the characteristics observed in chapter 7 in regards to the transitions between driving conditions. A stochastic process $\mathbf{X} = \{X_k, k \geq 0\}$ is a sequence of random variables which can be successfully modeled as a first order Markov chain, under the assumption that it satisfies the so called Markov property (MP). This property states that, the future state X_{k+1} depends only on the current state X_k and not on all previous states $X_0, X_2, \ldots, X_{k-1}$. In other words, for all $\{X_k, k \geq 0\}$:

$$p(X_{k+1} = j | X_0 = i_0, \ldots, X_{k-1} = i_{k-1}, X_k = i) \tag{8.1.1}$$

$$= p(X_{k+1} = j | X_k = i), \tag{8.1.2}$$

where $p(X_{k+1} = j | X_k = i)$, with $i_0, ..., i, j \in \mathbf{S} = \{s_1, s_2, s_3, s_4\}$, is the conditional transition probability. All transition probabilities are grouped in the transition probability matrix $\mathbf{M}^{\mathrm{S}}$:

$$\mathbf{M}^{\mathrm{S}} = \begin{array}{c} \text{State} \\ 1_k \\ \vdots \\ m_k \end{array} \begin{array}{c} \begin{array}{ccc} 1_{k+1} & \cdots & m_{k+1} \end{array} \\ \begin{bmatrix} \pi_{1,1} & \cdots & \pi_{1,m} \\ \vdots & \ddots & \vdots \\ \pi_{m,1} & \cdots & \pi_{m,m} \end{bmatrix} \end{array}. \tag{8.1.3}$$

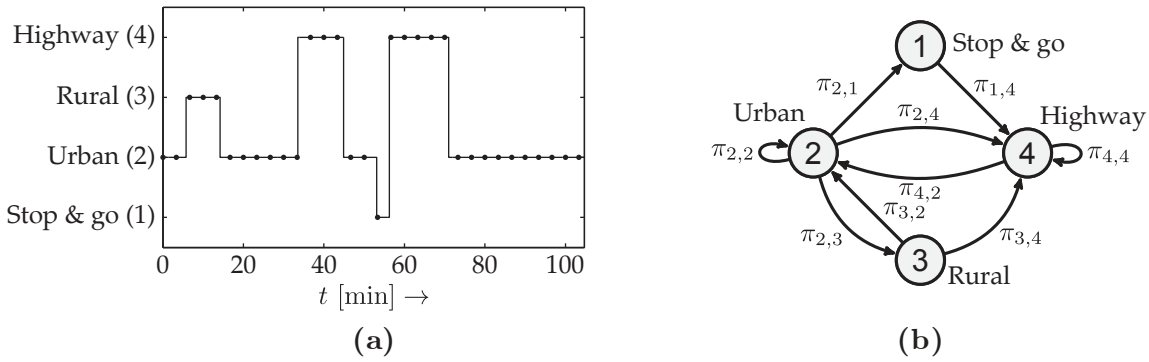

Figure 8.1.: Representation of the evolution of driving conditions in time **(a)** and of the modeled Markov states with their respective transition probabilities **(b)**.

Accordingly, equation (8.1.1) can be expressed as:

$$\pi_{i,j} = \mathbf{M}^{\mathrm{S}}\left(X_{k+1} = j | X_k = i\right), \tag{8.1.4}$$

where $\pi_{i,j}$ is the ij^{th} element of $\mathbf{M}^{\mathrm{S}}$. Since the elements j of $\mathbf{M}^{\mathrm{S}}$ represent the transition probabilities to all other states from i, each row $\boldsymbol{\pi}_i$ satisfies the unit simplex requirement $\sum_{j=1}^{m} \pi_{i,j} = 1$ and $0 \leq \pi_{i,j} \leq 1$. The equation (8.1.4) is said to be *time homogeneous* since $\pi_{i,j}$ is independent of k. The homogeneous Markov chain is completely determined by defining an initial probability $\pi_i(0) := p(X_0 = i)$ with $i \in S$.

Figure 8.1 (**a**) exemplifies the temporal evolution of driving conditions during a test drive over a period of about 100 min. The Markov states as well as the transition probabilities generated therefrom are illustrated in figure 8.1 (**b**). In order to parametrize the Markov model it is necessary to estimate the stationary distributions of $\mathbf{M}^{\mathrm{S}}$. In this work we use the maximum likelihood estimation (MLE) scheme suggested by Lee, Judge, and Zellner (1970) for estimating the time-invariant transition probabilities of $\mathbf{M}^{\mathrm{S}}$. A transition probability $\pi_{i,j}$ is estimated by:

$$\hat{\pi}_{i,j} = \frac{n_{ij}}{\sum_{j=1}^{m} n_{ij}}, \qquad i, j = 1, 2, ..., m, \tag{8.1.5}$$

where m is the number of Markov states, n_{ij} is the number of times a driving condition changes from state i to j and $\sum_{j=1}^{m} n_{ij}$ is the total number of times it changes from state i. This approach is very practical since the estimation can by achieved by simply *counting* the number of times a change in the driving condition occurs. Estimating transition probabilities with MLE properly works as long as the number of observed transitions is large compared to the size of the Markov state space (Teodorescu 2009). In this work $m = 4$ and thus there exist only 16 possible combinations.

Having parametrized $\mathbf{M}^{\mathrm{S}}$, the prediction of future driving conditions for an arbitrary horizon length h_{L}, i.e., a realization $\hat{\mathbf{s}} = [\hat{s}_0, \ldots \hat{s}_k, \ldots, \hat{s}_{h_{\mathrm{L}}}]$ is achieved by generating a Markov chain from a given initial condition $\hat{s}_0$, e.g., the current driving condition. The chain is then generated as:

$$\hat{s}_{k+1} = \mathrm{rand}(\boldsymbol{\pi}_{s_k}), \quad k = 0, \ldots, h_{\mathrm{L}} - 1, \tag{8.1.6}$$

where $\boldsymbol{\pi}_{s_k}$ is the row vector of $\mathbf{M}^{\mathrm{S}}$ indexed by the value of s at time step k. The function $\mathrm{rand}(\cdot)$ randomly draws a sample of the discrete probability distribution given by $\boldsymbol{\pi}_{s_k}$.

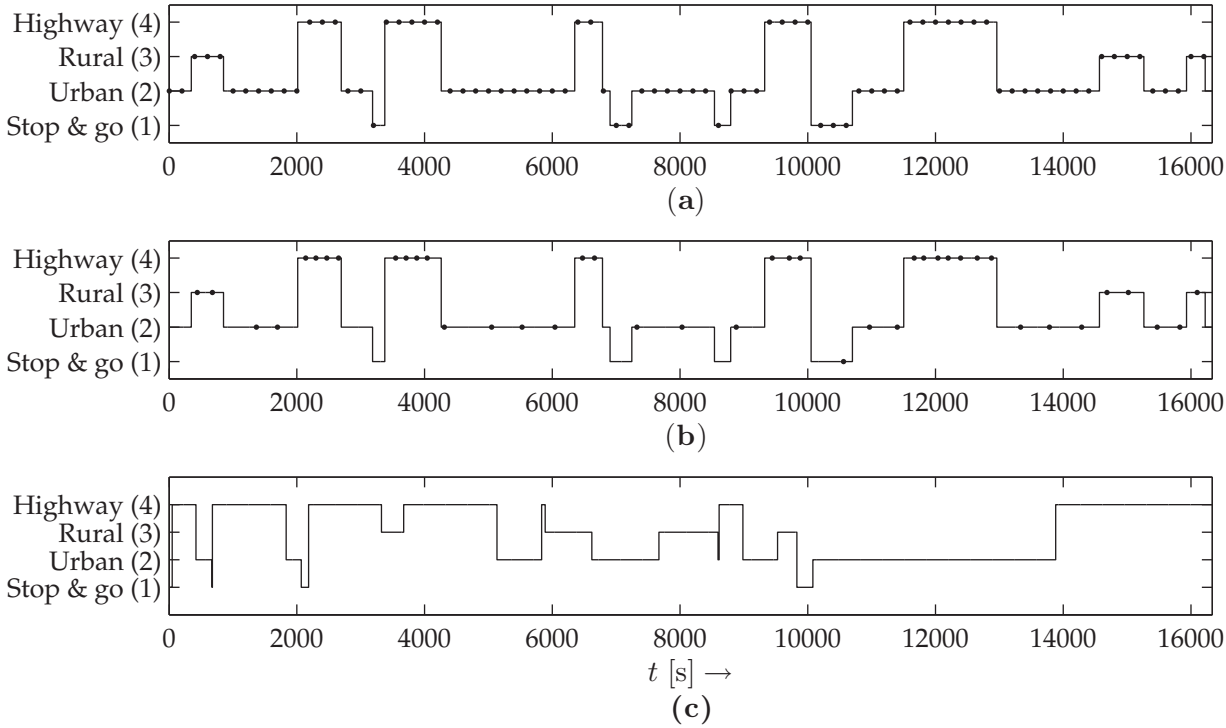

Figure 8.2.: Driving conditions classification results based on time- (**a**) and distance-based (**b**) measurements. Markov chain generated with $W_k = 1$ s (**c**).

As it was previously pointed out, a still pending task is the choice of the type of measurements to be used (time or distance based) as well as its length, i.e., the sampling rate for updating $\mathbf{M}^{\mathrm{S}}$. With help of the classification scheme developed in chapter 7 it is possible to parametrize an arbitrary sequence of driving conditions $\mathbf{s} = [s_1, \ldots, s_M]^\top$ both as time (with a maximum resolution of 1 s) or as spatial series. Since the transition probabilities are updated using the LS-SVM classifier (see section 8.2), the type and length of the measurements chosen for parameterizing both the classifier and the Markov model must be identical. In section 7.4 it has been shown that time-based measurements slightly offer more accurate results than their distance-based counterpart. However, since both of them offer acceptable results for classification, the decision about their choice has to be taken based on the Markov model. Figure 8.2 shows the results of a driving condition classification achieved by using (**a**) time and (**b**) distance based measurements. As it can appreciated, the higher the speed within the classified driving condition, e.g. on the highway, the more samples are needed. This leads to an increment of the values of the transition probabilities for both the rural and the highway situations, which causes an unrealistic transition probability matrix. Based on this criteria, time based measurements are favored in this work. The choice of the sampling rate at which the Markov chain is generated has a lower bound. The reason lies in the fact that the transition between driving conditions exhibits certain dynamics and consequently cannot change arbitrarily fast. If a realization of the Markov chain is generated with a small sampling rate, e.g., 1 s, it is possible that transitions between states occur faster as under real conditions (see figure 8.2 (**c**)). Another implication of employing small sampling rates arises from the fact, that the computational effort required during the classification increases with small

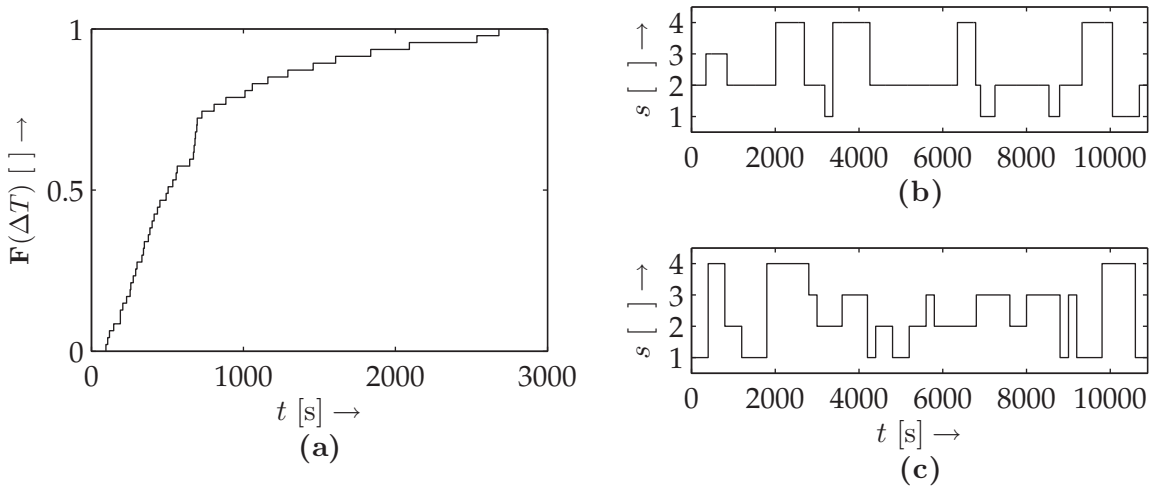

Figure 8.3.: Cumulative distribution of sojourn times (**a**), transition between driving conditions exhibited by the original data (**b**) and by a synthetic profile (**c**).

sampling values of W_k. Furthermore, as it was shown in section 7.4, the accuracy of the classification also decreases with the size of W_k. Choosing W_k imposes a trade-off between computational effort and classification accuracy. Figure 8.3 (**a**) shows the cumulative probability distribution of the sojourn times of all driving situations manually labeled (classified) on the training data. As it can be seen the shortest and the mean sojourn time lies by 145 s and 695 s, respectively. It becomes clear that 50% of all sojourn times lie below 500 s and that 20% under 250 s. A further analysis shows that only 3 sojourn times lie below the quantile 10%, given by 209 s which is considered here as the lower bound. As it was shown in section 7.4, the classification accuracy by $W_k = 200\,\mathrm{s}$ is $(88.50 \pm 1.49)\,\%$ with a linear kernel and $(89.98 \pm 2.13)\,\%$ for the RBF kernel. Accordingly, we set to $W_k = 200\,\mathrm{s}$.

Having defined W_k, $\mathbf{M}^{\mathrm{S}}$ is estimated with equation (8.1.5) resulting in:

$$\mathbf{M}^{\mathrm{S}} = \begin{bmatrix} 0.50 & 0.278 & 0.111 & 0.111 \\ 0.062 & 0.778 & 0.086 & 0.074 \\ 0.115 & 0.231 & 0.615 & 0.038 \\ 0.026 & 0.184 & 0.026 & 0.763 \end{bmatrix}. \tag{8.1.7}$$

A qualitative analysis of an arbitrary realization of the Markov chain generated with W_k, $\mathbf{M}^{\mathrm{S}}$ (see figure 8.3 (**c**)) suggests that a synthetic chain generated with $W_k = 200\,\mathrm{s}$ exhibits characteristics comparable to the original data (figure 8.3 (**b**)) in the sense of the lack of improbable events such as absorbing states or very small sojourn times. However, to make a meaningful statement about the validity of the Markov model it is required to validate the Markov property in order to proof whether the underlying stochastic process $\mathbf{X}$, describing the temporal evolution of driving conditions, can be adequately described by a first-order Markov chain or if higher orders might be necessary. To this aim we test the null-hypothesis:

$$\begin{aligned} H_0 : p(X_{k+1} = j | X_0 = i_0, \ldots, X_{k-1} = i_{k-1}, X_k = i) \\ = p(X_{k+1} = j | X_k = i), \end{aligned} \tag{8.1.8}$$

to proof that $\mathbf{X}$ is memoryless. The null-hypothesis is formulated to test that X_{k+1} is

conditional independent of X_k and of X_{k-1}. To this aim we split the original data in 3 equidistant sets and build contingency tables containing counts of first-order and second-order transitions for each set and test the conditional independence of the data by means of the P_{value} computed with the exact fisher test. The results obtained show an average of $P_{\text{value}}^{\text{1st}} = 0.00049$ and $P_{\text{value}}^{\text{2nd}} = 0.5927$ among all three sets for both contingency tables. This means that H_0 can be rejected just for the first-order model, since $P_{\text{value}}^{\text{1st}} \ll \alpha_{\text{s}}$, where $\alpha_{\text{s}} = 0.05$ is the significance level (Nuzzo 2014). It is therefore validated that $\mathbf{X}$ in our case can be properly described by a first-order Markov model.

8.2. Adaptation of transition probabilities

Parameterizing $\mathbf{M}^{\text{S}}$ relying solely on historical data provides a good estimation of how an electric vehicle travels on the long term, that is, it best describes the driving behavior and conditions on a macroscopic time scale, e.g., weeks or months. However, the way a driver behaves might change depending on the time of the day, the mood or the road condition. This causes the *time homogeneity* assumption of the Markov chain to lose its validity. Therefore, a more proper characterization scheme would require to update the transition probabilities of $\mathbf{M}^{\text{S}}$ as new information about the driving conditions becomes available. This allows to capture the *non-homogeneity* of the underlying Markov process, which might be introduced by changes in the traffic situation or the driving scenario. To this aim we employ in this work a Bayesian posterior probability approach to update the established transition probabilities between Markov states.

Analogous to the discussion of section 6.1, more precisely equation (6.1.5) , applying the Bayes' theorem for updating a transition probability $\pi_{i,j}$ of $\mathbf{M}^{\text{S}}$ requires a *likelihood function* for the new observed driving conditions and an assumption about the *prior distribution* of $\pi_{i,j}$ on each row of $\mathbf{M}^{\text{S}}$. The forthcoming explanation deals with the theoretical foundations for updating any transition probability $\pi_{i,j}$ belonging to the mixture $\boldsymbol{\pi}_i = \left[\pi_{i,1}, \pi_{i,2}, \dots \pi_{i,j}, \dots, \pi_{i,m}\right]$, i.e., to the i^{th} row of $\mathbf{M}^{\text{S}}$ in equation (8.1.3).

Likelihood function: let the random variable q, representing a transition between two Markov states i and j, i.e., two driving conditions, to follow a multinomial distribution (recall equation (2.1.1)). The probability distribution of q can be parametrized by a vector $\boldsymbol{\pi}_i$, where $\pi_{i,j} = p\left(i \rightarrow j\right) = p\left(q_{i,j}\right)$ is the probability of a transition from state i to state j, as it was already stated in equation (8.1.4). Then, the likelihood of a sequence of new transitions $\boldsymbol{\mathcal{Q}} = \{q_1, q_2, ..., q_n\}$ is given by

$$\boldsymbol{\mathcal{L}}\left(\boldsymbol{\pi}_i | \boldsymbol{\mathcal{Q}}\right) = \prod_{j=1}^{n} \pi_{i,j}^{\beta_{i,j}}, \tag{8.2.1}$$

where $\beta_{i,j}$ is the number of times a transition $q_{i,j}$ occurs in $\boldsymbol{\mathcal{Q}}$. Here we express $\beta_{i,j} = \sum \delta_{i,j}$, where $\delta_{i,j} = 1$ if $i \rightarrow j$ occurs and $\delta_{i,j} = 0$, otherwise.

Prior distribution: in the context of Markov chains, the task of the prior is to specify an assumption about the probability distribution of the i^{th} row $\boldsymbol{\pi}_i$ of $\mathbf{M}^{\text{S}}$. Accordingly, it is necessary to find as many prior distributions as the number of Markov states. Updating the transition probabilities under a Bayesian approach works with any kind of prior. However, since we consider the arbitrary set of new transitions $\boldsymbol{\mathcal{Q}}$ to be multinomial distributed, it is mathematically convenient to use a *conjugate prior*. The

use conjugate priors offers the advantage that the posterior distribution has the same functional form of the prior. The conjugate prior of the multinomial distribution is the Dirichlet distribution (Strelioff, Crutchfield, and Hübler 2007). Thus, assuming the transition probabilities of a row from $\mathbf{M}^{\mathrm{S}}$ to be Dirichlet distributed leads to:

$$p\left(\boldsymbol{\pi}_i|\alpha_{i,1},\alpha_{i,2},...,\alpha_{i,m}\right)=\frac{\Gamma_{\mathrm{g}}\left(\sum_{j=1}^{m}\alpha_{i,j}\right)}{\prod_{j=1}^{m}\Gamma_{\mathrm{g}}\left(\alpha_{i,j}\right)}\prod_{j=1}^{m}\pi_{i,j}^{\alpha_{i,j}-1}, \tag{8.2.2}$$

where the hyper-parameter $\alpha_{i,j}$ can be understood as a virtual count of occurrences of $i \rightarrow j$ before considering new observations. Large values of $\alpha_{i,j}$ reflect strong prior knowledge about the distributions of the transition probabilities and small values of correspond to ignorance. The parameter m stands for the number of hyper-parameters that parametrize equation (8.2.2), in this case given by the number of driving conditions. The choice of the Dirichlet distribution as the prior is a fairly intuitive way to explain the meaning of the transition probabilities in $\mathbf{M}^{\mathrm{S}}$. A transition probability $\pi_{i,j}$ as defined by equation (8.1.4), can be understood as the first moment of the Dirichlet distribution evaluated for $\pi_{i,j}$. That is:

$$\mathbb{E}\left[\pi_{i,j}\right]=\pi_{i,j}=\frac{\alpha_{i,j}}{\alpha_0}, \tag{8.2.3}$$

where $\alpha_0=\sum_i \alpha_i$ is the total number of times a transition starting from state i occurs. The Dirichlet distribution satisfies the unit simplex requirement $\sum\mathbb{E}\left[\pi_{i,j}\right]=1$ and $0 \leq \mathbb{E}\left[\pi_{i,j}\right] \leq 1$ complying in this way with the properties of a row $\boldsymbol{\pi}_i$ in $\mathbf{M}^{\mathrm{S}}$. Furthermore, the uncertainty of a transition probability can be computed by:

$$\mathrm{Var}\left[\pi_{i,j}\right]=\frac{\alpha_{i,j}\left(\alpha_0-\alpha_{i,j}\right)}{{\alpha_0}^2\left(\alpha_0+1\right)}. \tag{8.2.4}$$

The parameters of the Dirichlet prior distribution are obtained through MLE of section 8.1. In the absence of prior knowledge about the hyper-parameters of equation (8.2.2), a common approach is to assume all probabilities to be equal, that is, we set all $\alpha_{i,j}=1$, which results in an uniform prior distribution with an expectation value given by $\mathbb{E}\left[\pi_{i,j}\right]={}^1\!/_m$, where m likewise represents the size of the state space.

Posterior distribution: having a multinomial likelihood and a Dirichlet prior, the posterior distribution of $\boldsymbol{\pi}_i$ after observing a new sequence of transitions $\boldsymbol{\mathcal{Q}}$ can be found in a closed form by exploiting the conjugate property of the Dirichlet distribution and the multinomial distribution. Accordingly, the posterior is computed by:

$$p\left(\boldsymbol{\pi}_i|\boldsymbol{\mathcal{Q}},\boldsymbol{\alpha}\right)\propto\boldsymbol{\mathcal{L}}\left(\boldsymbol{\pi}_i|\boldsymbol{\mathcal{Q}}\right)p\left(\boldsymbol{\pi}_i|\boldsymbol{\alpha}\right)=\prod_{j=1}^{m}\pi_{i,j}^{\alpha_{i,j}+\beta_{i,j}-1}. \tag{8.2.5}$$

The posterior is computed on occurrence of new transitions between driving conditions. Considering the fact that in our system just one transition can occur per time step, we can set $\beta_{i,j}=\delta_{i,j}$. Accordingly, the set of hyper-parameter α_i can be recursively updated by setting $\alpha_{i,k+1}=\alpha_{i,k}+1$ if $i \rightarrow i$ or $\alpha_{i,k+1}=\alpha_{i,k}$, otherwise. By employing this Bayesian scheme the updated mean $\mathbb{E}\left[\pi_{i,j}\right]_{k+1}$ and variance $\mathrm{Var}\left[\pi_{i,j}\right]_{k+1}$ of each element in $\boldsymbol{\pi}_i$ can be computed with the help of equations (8.2.3) and (8.2.4).

The posterior computed by equation (8.2.5) keeps the information regarding all transitions occurred up to time step k. Thus, depending on the values of the hyper-parameters, many new observations might be needed in order to converge with the new Markov process. This is inconvenient, since a slow adaptation of transition probabilities would cause the characterization of the most up to date driving condition to fail. Because of this, it would be desirable to find a recursion for both $\mathbb{E}\left[\pi_{i,j}\right]_k$ and $\mathrm{Var}\left[\pi_{i,j}\right]_k$ without needing to deal with any prior knowledge about the hyper-parameters and that can be carried such that the influence of older transitions in the computation of the posterior is progressively faded while keeping the underlying idea of an a Bayesian update. This recursion is achieved by means of the discounted mean-variance estimator (Bertuccelli and How 2008):

$$\mathbb{E}\left[\pi_{i,j}\right]_{k+1} = \mathbb{E}\left[\pi_{i,j}\right]_k + \frac{\mathrm{Var}\left[\pi_{i,j}\right]_k \left(\delta_{i,j} - \mathbb{E}\left[\pi_{i,j}\right]_k\right)}{\lambda_{\mathrm{s},k}\mathbb{E}\left[\pi_{i,j}\right]_k \left(1 - \mathbb{E}\left[\pi_{i,j}\right]_k\right)}, \tag{8.2.6}$$

$$\mathrm{Var}\left[\pi_{i,j}\right]_{k+1} = \frac{\mathrm{Var}\left[\pi_{i,j}\right]_k \mathbb{E}\left[\pi_{i,j}\right]_{k+1} \left(1 - \mathbb{E}\left[\pi_{i,j}\right]_{k+1}\right)}{\lambda_{\mathrm{s},k}\mathbb{E}\left[\pi_{i,j}\right]_k \left(1 - \mathbb{E}\left[\pi_{i,j}\right]_k\right) + \mathrm{Var}\left[\pi_{i,j}\right]_k}, \tag{8.2.7}$$

where $\lambda_{\mathrm{s},k} < 1$ is a factor used to scale the variance at each iteration, which allows making the estimation to be more responsive to new observations. Bertuccelli and How 2008 show that convergence to the true moments is achieved if $\lim_{k\to\infty} \lambda_{\mathrm{s},k} = 1$. We thus consider in this work a decaying factor $\lambda_{\mathrm{s},k} = 1 - \frac{\lambda_\mathrm{s}}{k}$, where $0 < \lambda_\mathrm{s} < 1$ and k denotes the discrete time step. If the initial prior distribution is computed based on historical driving data, then a slow adaptation rate might suffice, as illustrated in figure 8.4 (**a**), however, if a prior with no previous information regarding the driving situation is used, the convergence speed has to be increased, otherwise the parameters describing the most up to date driving situation can not be determined, as seen in figure 8.4 (**b**). Based on the historical driving data sets it can be argued, that the adaptation rate should be found withing one battery discharging cycle, that is, for trips between 120 km and 160 km with durations ranging from 90 min to 120 min approximately. Therefore, changes in the driving situation, such moving from the city to the highway, can be accounted for and meaningless changes, e.g., small traffic-jams, can be neglected. To validate the functionality of the mean-variance estimator, it is investigated, whether the same posterior distribution is obtained for different values of λ_s as in the case when $\lambda_\mathrm{s} = 1$. Figure 8.4 (**c**) illustrates the comparison of 1 out of the 16 probabilities in $\mathbf{M}^\mathrm{S}$ of equation (8.1.7). As it can be seen, this probability converges to the values computed off-line within 150 iterations.

Figure 8.4 (**d**) shows the behavior of mean-variance estimator for a total of 2000 iterations. In this case the initial prior distribution is obtained from the results depicted in figure 8.4 (**c**). Later, after 500 iterations the value of the probability abruptly changes to a new value and remains there until the last iteration. As it can be observed, for $\lambda_\mathrm{s} = 0.95$, the estimator (black line) slowly adapts to the change. Nevertheless, it converges towards the true probability value (gray line). It turns out that the choice of the initial prior distribution of is of high relevance in case $\lambda_\mathrm{s} \neq 1$. Hence, it is advisable to compute the prior distribution off-line with help of historical driving data, so that abrupt changes in the covariance estimation can be avoided.

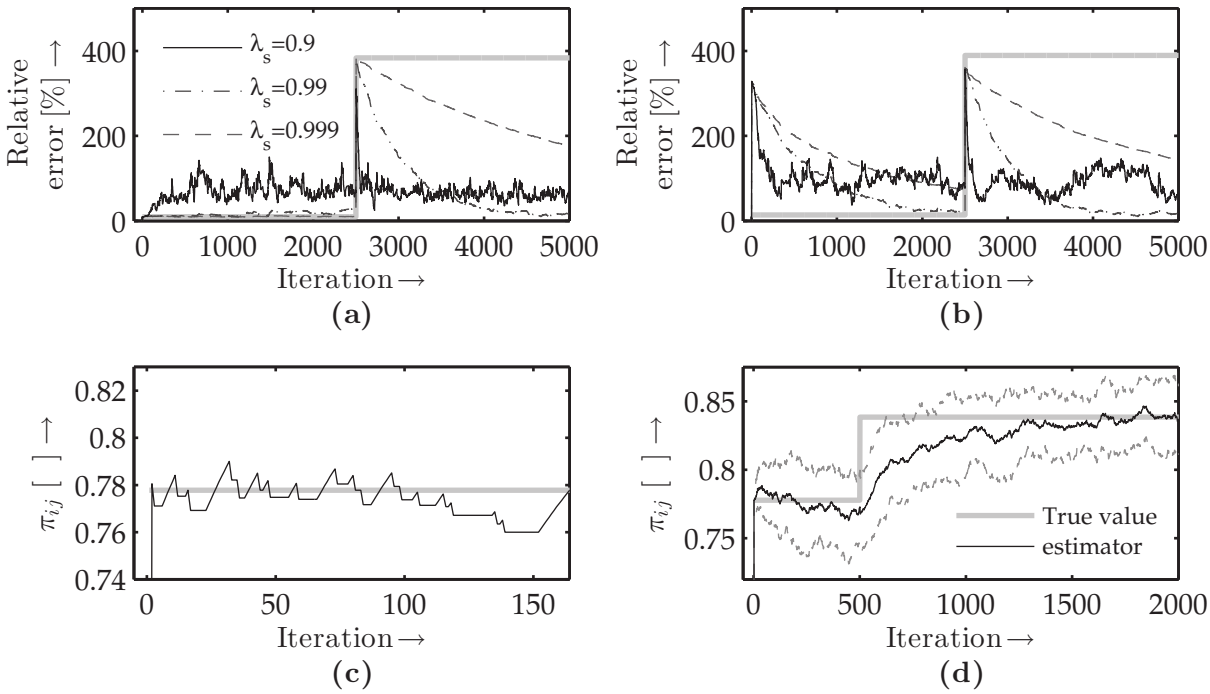

Figure 8.4.: True probability (solid gray) estimation convergence speed with (**a**) and without (**b**) previous information by $\lambda_s = 0.9$, $\lambda_s = 0.99$, $\lambda_s = 0.999$, $\lambda_s = 1$ (**c**), and $\lambda_s = 0.95$ (**d**).

8.3. Driving behavior modeling

As it was already mentioned in the introduction of this chapter, the interaction system-surroundings is partially described by the driving behaviour, characterized by the longitudinal speed v_x and acceleration a_x of the electric vehicle. This section introduces a model for predicting both v_x and a_x in dependence of the condition situation.

For a better understanding of the model let us first define:

$$\mathcal{F}_s = \{v_x \in \mathbb{R}, a_x \in \mathbb{R} : \mathcal{V}_s, \mathcal{A}_s\} \tag{8.3.1}$$

as the two-dimensional state space of v_x and a_x of an arbitrary driving condition $s \in S$, i.e., $\mathcal{V}_s = \{v_{x,\min}, v_{x,\min} + \Delta v_x, \ldots, v_{x,\max}\}$ and $\mathcal{A}_s = \{a_{x,\min}, a_{x,\min} + \Delta a_x, \ldots, a_{x,\max}\}$. Whereas $v_{x,\min}$, $v_{x,\max}$, $a_{x,\min}$ and $a_{x,\max}$ are obtained directly from the training data, the step sizes used for discretizing $\mathcal{V}_s$ and $\mathcal{A}_s$, namely, Δv_x and Δa_x, for each driving condition, are chosen by means of optimization with the aim of minimizing the difference between the mean, the standard deviation and the mean absolute distance between the empirical cumulative distribution (Kolmogorov-Smirnov test) of the measured and the simulated speed. Analogous to the Markov model presented in section 8.1, a conditional transition probability in this model is defined as:

$$\pi_{i,j}^{v_x^o a_x} = p\left(a_{x,k+1} = j | a_{x,k} = i, v_{x,k} = o\right), \tag{8.3.2}$$

where $\pi_{i,j}^{v_x^o a_x}$ is the probability of accelerating at rate a_x^j over the next time step given that the vehicle accelerates with a_x^i at given speed v_x^o in the current time step. Modeling this stochastic process in a discrete state space would require to define a transition probability matrix $\mathbf{M}^{v_x^o a_x}$ for each $v_x^o \in \mathcal{V}_s$ (see figure 8.5 (**a**)). The drawback

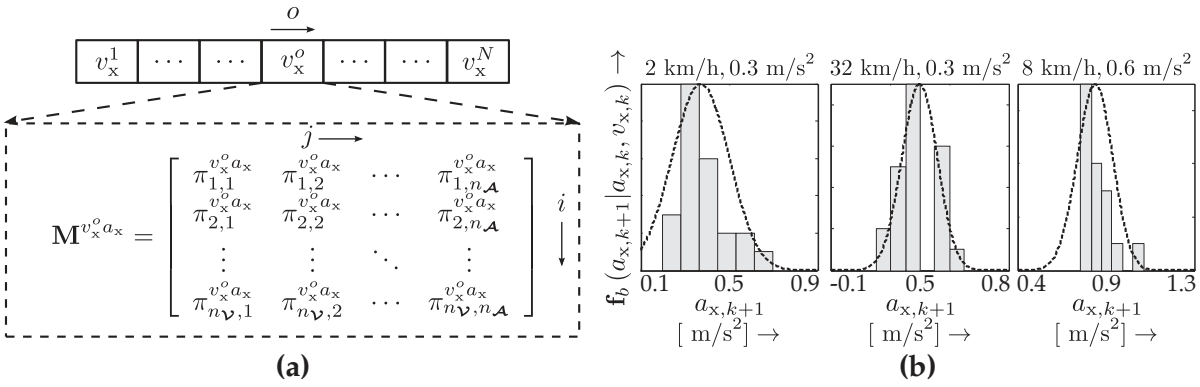

Figure 8.5.: Exemplification of multiple transition probability matrices in dependence of v_x (**a**) and fitted Beta function over the distribution of $a_{x,k+1}$ by different tuple (v_x, a_x) (**b**).

with this approach is that the size of the state space increases proportionally to the size of $\mathcal{A}_s$, which leads to large sparse transition probability matrices, and therefore, more data is needed in order to identify the parameters. Oliva and Bertram (2014a) sort out this issue by finding a suitable probability distribution function of the form $\mathbf{f}_b\left(a_{x,k+1} = j | a_{x,k} = i, v_{x,k} = o\right)$ that can be employed to compute $\pi_{i,j}^{v_x^o a_x}$. The shape of such a function can be better understood by analysing the distribution of $a_{x,k+1}$ at different tuple (v_x, a_x). Billio and Casarin (2011) as well as Ferrari and Cribari-Neto (2004) suggest that a strong candidate for choosing $\mathbf{f}_b$ is the *Beta* distribution. The figure 8.5 (**b**) shows the Beta function fitted over different distributions of $a_{x,k+1}$. The Beta density function is a versatile function which is usually employed for modeling different shapes of probability distributions. The probability density function of the generalized Beta distribution applied to $a_{x,k+1}$ is given by:

$$\mathbf{f}_b\left(a_{x,k+1}|\alpha,\beta,b_{L_\beta},b_{U_\beta}\right) = \frac{\left(a_{x,k+1}-b_{L_\beta}\right)^{\alpha-1}\left(b_{U_\beta}-a_{x,k+1}\right)^{\beta-1}}{\left(b_{U_\beta}-b_{L_\beta}\right)^{\alpha+\beta-1}}, \tag{8.3.3}$$

where α and β are the shape parameters of the Beta distribution and $\left[b_{L_\beta}, b_{U_\beta}\right]$ define the interval for which equation (8.3.3) is defined. The fact that the Beta distribution is defined just over a given interval is exploited in this work in that no accelerations beyond the admissible values, dictated by the performance of the electric vehicle, can be reached. Furthermore, b_{L_β} and b_{U_β} can be conveniently chosen to force any $a_{x,k+1}$, drawn from a Beta distribution given by (8.3.3), to lie within the bounds of the state space of $\mathcal{F}_s$, i.e., $a_{x,k+1} \in \mathcal{A}_s$. The first two moments of $a_{x,k+1}$ are given by

$$\mathbb{E}\left[a_{x,k+1}|\alpha,\beta,b_{L_\beta},b_{U_\beta}\right] = b_{L_\beta} + \left(b_{U_\beta}-b_{L_\beta}\right)\frac{\alpha}{\alpha+\beta}, \tag{8.3.4}$$

$$\mathrm{Var}\left[a_{k+1}|\alpha,\beta,b_{L_\beta},b_{U_\beta}\right] = \frac{\left(b_{U_\beta}-b_{L_\beta}\right)^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \tag{8.3.5}$$

The function $\mathbf{f}_b$ can be reformulated such that the parameters of the Beta distribution depend on the Markov states and that the probability density function is defined only over $\mathcal{A}_s$, i.e., $\mathbf{f}_b\left(a_{x,k+1}|\boldsymbol{\alpha}\left(v_{x,k},a_{x,k}\right),\boldsymbol{\beta}\left(v_{x,k},a_{x,k}\right),a_{x,\min},a_{x,\max}\right)$.

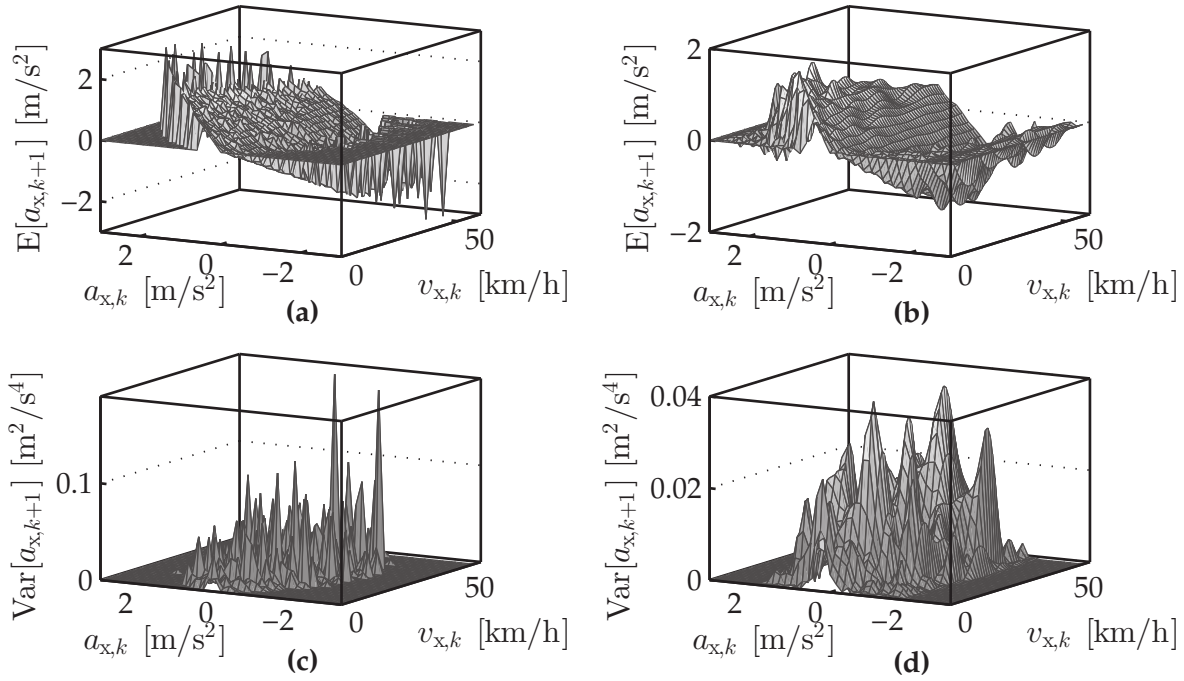

Figure 8.6.: Point-wise definition of $\mathbf{g}_b(v_x, a_x)$ (**a**) and $\mathbf{h}_b(v_x, a_x)$ (**c**) over $\mathcal{F}_s$. Smoothed definition of $\mathbf{g}_b(v_x, a_x)$ (**b**) and $\mathbf{h}_b(v_x, a_x)$ (**d**) over $\mathcal{F}_s$.

The estimation of both $\boldsymbol{\alpha}(\cdot)$ and $\boldsymbol{\beta}(\cdot)$ over the entire state space $\mathcal{F}_s$ is carried out by gathering all samples of $a_{x,k+1}$ for each tuple $(v_{x,k}, a_{x,k})$ observed in the training data. The purpose of the aforementioned step is to sort out the data such that both $\mathbb{E}[a_{x,k+1}]$ and $\mathrm{Var}[a_{x,k+1}]$ can be calculated from the available samples of $a_{x,k+1}$.

As it can be seen in figures 8.6 (**a**) and (**c**) the sparsity of the training data causes both $\mathbb{E}[a_{x,k+1}]$ and $\mathrm{Var}[a_{x,k+1}]$ to be defined point-wise over $\mathcal{F}_s$. Thus, to parametrize $\mathcal{F}_s$ we define two new functions, namely, $\mathbf{g}_b(v_x, a_x)$ and a function $\mathbf{h}_b(v_x, a_x)$ which describe how $\mathbb{E}[a_{x,k+1}]$ and $\mathrm{Var}[a_{x,k+1}]$ continuously vary through $\mathcal{F}_s$.

Johannesson (2005) propose the use of bivariate tensor product B-splines with a predefined sequence, which are accordingly set denser where more information is available in order to better capture the behavior of the most important regions of $\mathcal{F}_s$. The splines describing this variation, namely, $\mathbf{g}_b(v_x, a_x)$ and $\mathbf{h}_b(v_x, a_x)$ over $\mathcal{F}_s$ for the urban driving situation are shown in figures 8.6 (**b**) and (**d**). Having identified $\mathbb{E}[a_{x,k+1}]$ and $\mathrm{Var}[a_{x,k+1}]$ for the entire state space, moment matching (AbouRizk, Halpin, and Wilson 1994) is used for estimating both parameters $\boldsymbol{\alpha}(v_x, a_x)$ and $\boldsymbol{\beta}(v_{x,k}, a_{x,k})$ by evaluating $\mathbf{g}_b(v_x, a_x)$ and $\mathbf{h}_b(v_x, a_x)$ for each state on $\mathcal{F}_s$ and by equating the result to the theoretical moments given by equations (8.3.4) and (8.3.5). Solving the obtained system of equations for $\boldsymbol{\alpha}(v_x, a_x)$ and $\boldsymbol{\beta}(v_{x,k}, a_{x,k})$ leads to:

$$\boldsymbol{\alpha}(v_x, a_x) = \frac{-\left(b_{L_\beta} - \mu_b\right)}{b_{L_\beta} - b_{U_\beta}} - \frac{\left(b_{L_\beta} - \mu_b\right)^2 \left(b_{U_\beta} - \mu_b\right)}{\sigma_b^2 \left(b_{L_\beta} - b_{U_\beta}\right)}, \tag{8.3.6}$$

$$\boldsymbol{\beta}(v_{x,k}, a_{x,k}) = \frac{\left(b_{U_\beta} - \mu_b\right)}{b_{L_\beta} - b_{U_\beta}} + \frac{\left(b_{L_\beta} - \mu_b\right)\left(b_{U_\beta} - \mu_b\right)^2}{\sigma_b^2 \left(b_{L_\beta} - b_{U_\beta}\right)}, \tag{8.3.7}$$

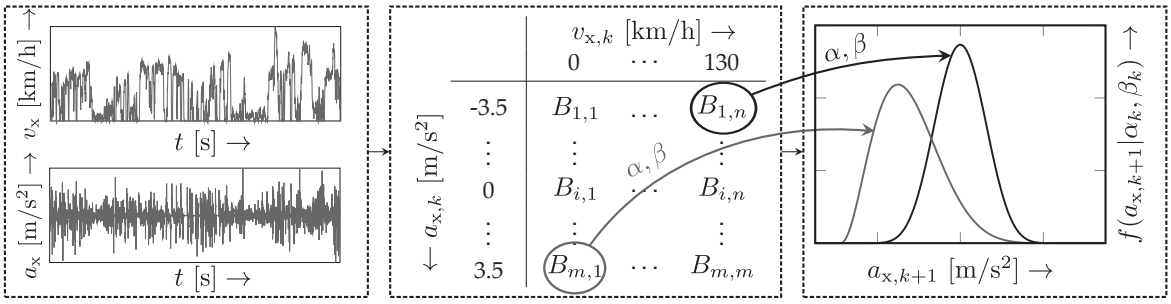

Figure 8.7.: Schematic representation of the procedure for generating $\hat{\mathbf{v}}_{x,0:h_L}$ and $\hat{\mathbf{a}}_{x,0:h_L}$.

where $\mu_b = \mathbb{E}[a_{x,k+1}]$, $\sigma_b^2 = \mathrm{Var}[a_{x,k+1}]$, $b_{L_\beta} = a_{x,\min}$ and $b_{U_\beta} = a_{x,\max}$. The transition probability matrix is expressed analogous to equation 8.1.3, with the difference that both rows and columns correspond to the same time step k and that each cell contains the parameters $B_{ij} = \{\alpha_{ij}, \beta_{ij}\}$ of the Beta distribution that corresponds to the next time step $k+1$ (see figure 8.7). Thus:

$$\mathbf{M}^{\mathcal{F}_s} = \begin{array}{c} \text{State} \\ a_{x,\min} \\ \vdots \\ a_{x,\max} \end{array} \begin{array}{c} \begin{array}{ccc} v_{x,\min} & \cdots & v_{x,\max} \end{array} \\ \begin{bmatrix} B_{1,1} & \cdots & B_{1,n_\mathcal{V}} \\ \vdots & \ddots & \vdots \\ B_{n_\mathcal{A},1} & \cdots & B_{n_\mathcal{A},n_\mathcal{V}} \end{bmatrix} \end{array}. \tag{8.3.8}$$

Simulating a realization of the underlying stochastic process describing the driving behavior, within a given driving condition $s \in \mathbf{S}$ and for a given horizon length h_L, i.e., $\hat{\mathbf{v}}_{x,0:h_L} = [\hat{v}_{x,0}, \ldots, \hat{v}_{x,k}, \ldots, \hat{v}_{x,h_L}]$ and $\hat{\mathbf{a}}_{x,0:h_L} = [\hat{a}_{x,0}, \ldots, \hat{a}_{x,k}, \ldots, \hat{a}_{x,h_L}]$, are performed by randomly drawing a sample $a_{x,j}$ from $\mathbf{M}^{\mathcal{F}_s}$ for the next predicted state $\hat{a}_{x,k+1}$:

$$\hat{a}_{x,k+1} = \mathrm{rand}(\mathbf{f}_b\left(a_{x,k+1}|\boldsymbol{\alpha}\left(v_{x,k}, a_{x,k}\right), \boldsymbol{\beta}\left(v_{x,k}, a_{x,k}\right), a_{x,\min}, a_{x,\max}\right)), \quad k = 1, \ldots, h_L - 1. \tag{8.3.9}$$

Contrary to other methods for generating synthetic driving profiles (Lee, Adornato, and Filipi 2011), our approach compute the value of the speed in the next speed instead of randomly sample it. Here $\hat{v}_{xk+1}$ is given by:

$$\hat{v}_{x,k+1} = \hat{v}_{x,k} + \hat{a}_{x,k}\Delta t, \tag{8.3.10}$$

where Δt denotes the time step size used in the generation of $\hat{\mathbf{v}}_{x,0:h_L}$ and $\hat{\mathbf{a}}_{x,0:h_L}$.

8.4. Road slope modeling

As it is shown in section 3.2, the climbing resistance F_G is proportional to the slope of the road α_{road} (recall equation (3.2.5)), and therefore directly influences both the power that the electric vehicle requires to move forward and the power that can be gained by means of regenerative braking. The slope model serves as basis to simulate multiple time series with identical stochastic characteristics as the training data, i.e., as the slope measurements.

To model the slope a one-dimensional state space has to be first properly defined. Analogous to $\mathcal{F}_s$, the state space is quantized with a step size $\Delta\alpha_{\text{road}}$ in the range between $\alpha_{\text{road, min}}$ and $\alpha_{\text{road, max}}$ for each driving condition defined in section 7.1, i.e., $\mathcal{F}_{\text{road},s} = \{\alpha_{\text{road,min}}, \alpha_{\text{road,min}} + \Delta\alpha_{\text{road}}, \ldots, \alpha_{\text{road,max}}\}$ where $s \in \mathbf{S}$ refers to an arbitrary driving condition. As in the model of the driving behavior, a transition probability matrix $\mathbf{M}^{\mathcal{F}_{\text{road},s}}$ is defined as:

$$\mathbf{M}^{\mathcal{F}_{\text{road},s}} = \begin{array}{c} \text{State} \\ \alpha_{\text{road, min}} \\ \vdots \\ \alpha_{\text{road, max}} \end{array} \begin{array}{c} \alpha_{\text{road},k+1} \\ \left[\begin{array}{c} \alpha_1, \beta_1 \\ \vdots \\ \alpha_{n_{\text{road}}}, \beta_{n_{\text{road}}} \end{array}\right] \end{array}. \tag{8.4.1}$$

A realization of the slope $\hat{\boldsymbol{\alpha}}_{\text{road},0:h_L} = [\hat{\alpha}_{\text{road},0}, \ldots, \hat{\alpha}_{\text{road},k}, \ldots, \hat{\alpha}_{\text{road},h_L}]$ is generated by randomly drawing from $\mathbf{M}^{\mathcal{F}_{\text{road},s}}$ a sample $\alpha_{\text{road},i}$ for the next predicted state $\hat{\alpha}_{\text{road},k+1}$:

$$\hat{\alpha}_{\text{road},k+1} = \text{rand}(\mathbf{f}_b(\boldsymbol{\alpha}(\alpha_{\text{road},k}), \boldsymbol{\beta}(\alpha_{\text{road},k}), \alpha_{\text{road, min}}, \alpha_{\text{road, max}}), \quad k = 1, \ldots, h_L - 1, \tag{8.4.2}$$

where h_L is the length of the realization and $\mathbf{f}_b(\cdot)$ represents a beta function which is parametrized by the tuple $(\alpha_{\text{road},i}, \beta_{\text{road},i})$ indexed in $\mathbf{M}^{\mathcal{F}_{\text{road},s}}$ by the current state $\hat{\alpha}_{\text{road},k}$.

8.5. Evaluation of power requirements prediction

The approach for predicting the power requirements is evaluated through a series of experiments. In order to analyze the performance of the prediction the Kolmogorov-Smirnov test, i.e., the distance between the empirical cumulative distributions of the measured and the simulated speed, slope and power, is employed. In this manner, the complete statistical characteristics of the prediction are taken into account.

The first part of experiments deals with assessing how accurate the driving behavior and the road slope can be predicted for each driving condition. To this aim both $\mathbf{M}^{\mathcal{F}_s}$ and $\mathbf{M}^{\mathcal{F}_{\text{road},s}}$ are parametrized with the datasets labeled as *training* in table B.1 for each driving condition. To this aim we generate multiple realizations of $\hat{\mathbf{u}}^{\text{env}}_{k:k+h_L}$, for each driving condition, and for different horizon lengths (h_L). Figures 8.8 (**a**)-(**c**) and (**d**)-(**f**) show the empirical probability distributions as well as the quantiles 5% and 95 %, of the road slope and the vehicle speed, of both the validation and the simulation data. In general, we can argue that the driving behavior and the road slope are properly simulated within each driving condition. For a deeper analysis, figure 8.9 shows the error between the empirical distributions as a function of h_L. It can be seen that the quality of the prediction increases with h_L. This is clear since the longer h_L the better is described the underlying stochastic process of the driving behavior and the road slope. To evaluate the prediction of the power consumption we employ a reference drive cycle shown in figure 8.10 (**a**) (black line) and assume that the driving conditions (gray line) are known beforehand. In this case the reference cycle is generated from the datasets labeled as *validation* in table B.1. Similarly to the previous case, both $\mathbf{M}^{\mathcal{F}_s}$ and $\mathbf{M}^{\mathcal{F}_{\text{road},s}}$ are parametrized with the datasets labeled as *training* in table B.1 for each driving condition.

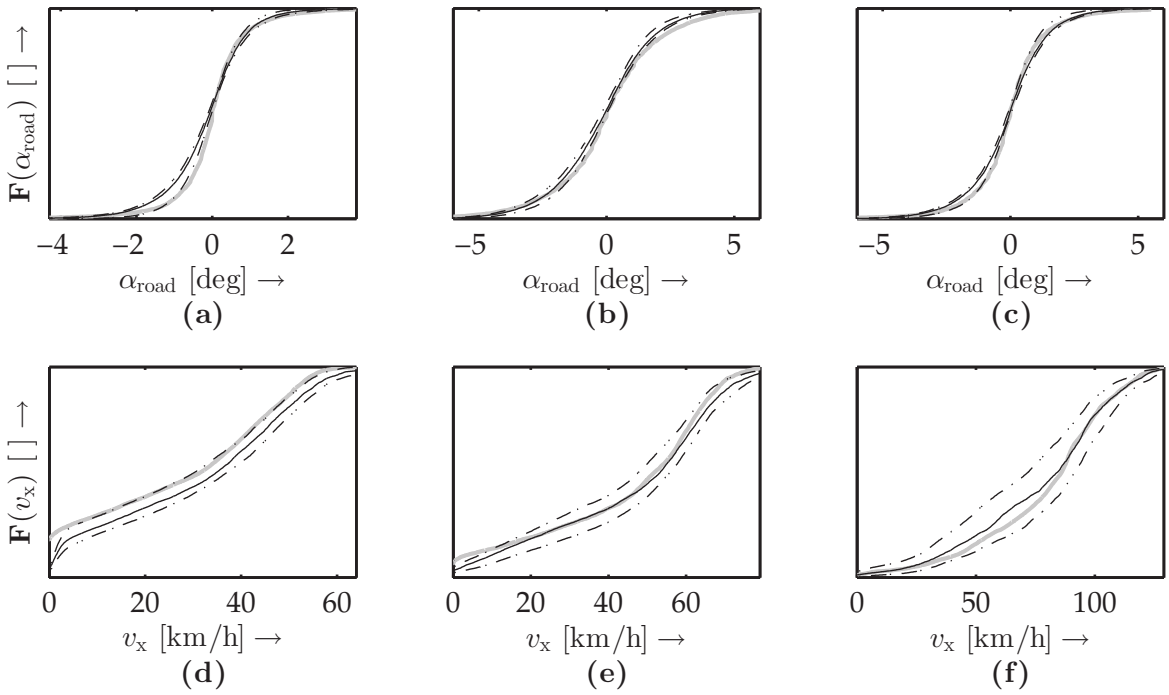

Figure 8.8.: Cumulative density function of speed and road slope from the validation (gray line) and the simulated (black line) data for the urban (**a**) and (**d**), rural (**b**) and (**e**) and highway (**c**) and (**f**) driving scenario.

Figure 8.10 (**b**) shows a simulated driving cycle of the same duration. As already mentioned, the driving conditions are assumed to be known and we set $W_k = 200$ s (recall section 8.1). It can be appreciated, that the driving behavior is properly modeled within each driving condition. Noteworthy is that the simulated speed exhibits larger values in contrast to the validation data. This phenomena is caused because a value of $v_{\text{x,min}} = 130$ km/h has been used while training the $\mathbf{M}^{\mathcal{F}_s}$ corresponding to the highway driving conditions and because the drive cycles of the validation datasets exhibit lower maximum speed on the highway driving condition.

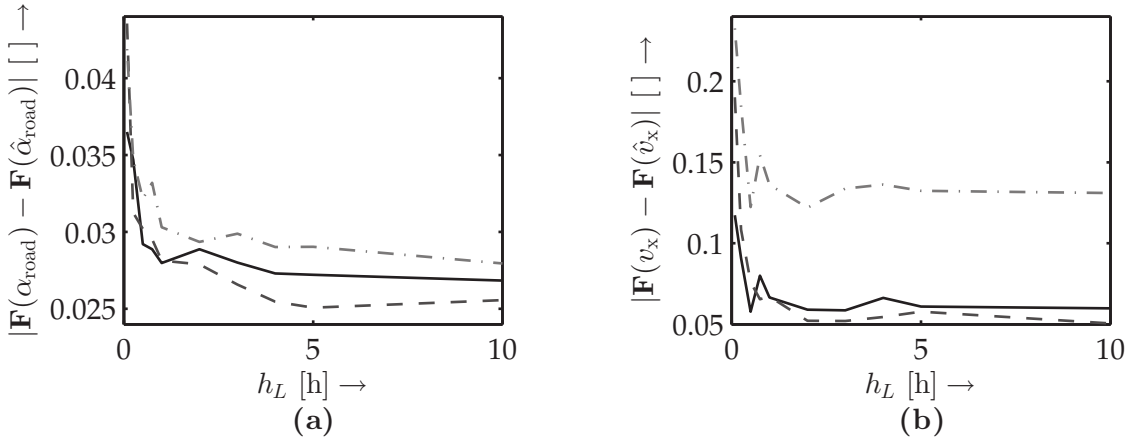

Figure 8.9.: Error between the empirical CDF of the road slope (**a**) and speed (**b**) as function of the prediction horizon length for the urban (black lines) as well as for the rural (black dashed lines) and highway (black dotted lines) driving situations .

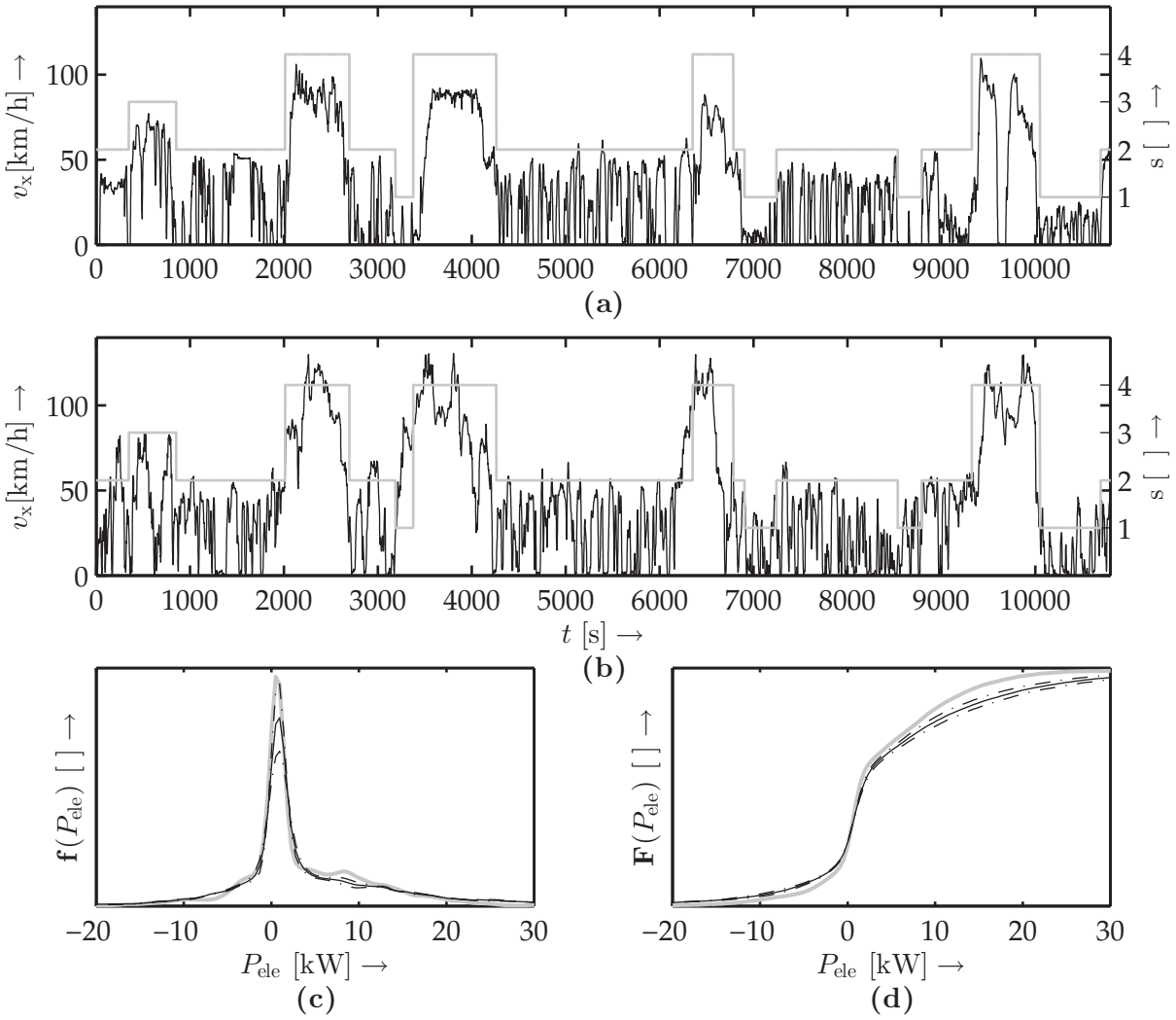

Figure 8.10.: Experimental results of the power prediction for a predefined sequence of driving situations. Speed profile (black line) and driving conditions (gray line) of the validation (**a**) and the simulated (**b**) data. Empirical probability density function (PDF) (**c**) and cumulative density function (CDF) (**d**) of the validation (gray line) and the simulated (gray line) with the corresponding quantiles 5% and 95% (black dotted lines).

The difference between the empirical probability distributions of the real and the simulated power consumption is computed by means evaluations of $\boldsymbol{\mathcal{M}}_{PC}$ as response to multiple realizations of $\hat{\mathbf{u}}^{\text{env}}_{k:k+h_L}$, with h_L set as the duration of the reference drive cycle. The empirical probability distribution and empirical cumulative probability distribution together with the corresponding quantiles 5 % and 95 % are shown in figures 8.10 (**c**) and (**d**). Note that the simulated probability distribution is relatively flat on the right around 10 kW in contrast to the training data. This effect becomes clearer by analyzing the cumulative distribution. It can be appreciated that both distributions diverge towards $\mathbf{F}(P_{\text{ele}}) = 1$. In this case, the error between cumulative distributions is 0.0242 ± 0.0025 and the error of the mean power consumption is (300.2 ± 18) W. The low values of the standard deviation suggest that the power consumption can be properly predicted with the approach developed throughout this chapter.

9

Remaining Driving Range Estimation

This chapter aims to integrate the knowledge gained throughout this work into an unified solution that allows estimating the RDR of the Nissan leaf under different operating conditions. The structure of the discussion presented below allows gradually assessing the impact of model, present state and future load uncertainty, on the RDR estimation. For the sake of comprehension, the reader is referred to sections 2.2 and 2.3 for a detailed explanation about the methodology followed as well the setup used to carry out the experiments of this section.

9.1. Assessing model uncertainty

All results presented in this chapter are based on the driving cycle illustrated in figures 9.1 (**a**) and (**b**), which is the result of joining the datasets shown in figures B.3 and B.4. The benefit of using this cycle, is that the electric vehicle operates under all driving conditions, defined in chapter 7 and modeled in chapter 8, thus, allowing us to draw more meaningfully conclusions about the RDR estimation.

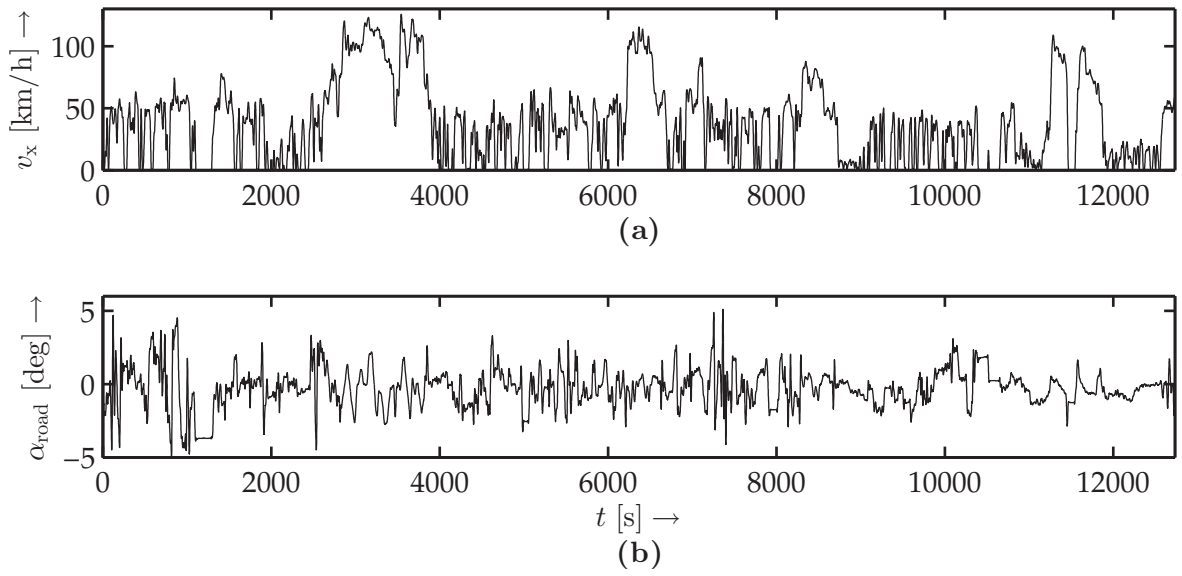

Figure 9.1.: Segment of the time series of speed (**a**) and road slope (**b**) belonging to the driving cycle generated for evaluating the RDR estimation under different operating conditions.

Table 9.1.: Mean and cumulative relative accuracy of the RDR estimation under different temperatures accounting only for model uncertainty.

T_{batt}	0°C	06°C	12°C	15°C	23°C	25°C	33°C	38°C
$\overline{\text{RA}}$[%]	90.27	91.18	88.93	87.8	89.88	90.76	88.96	88.91
CRA[%]	84.60	85.33	84.43	83.91	84.22	85.48	85.03	85.02

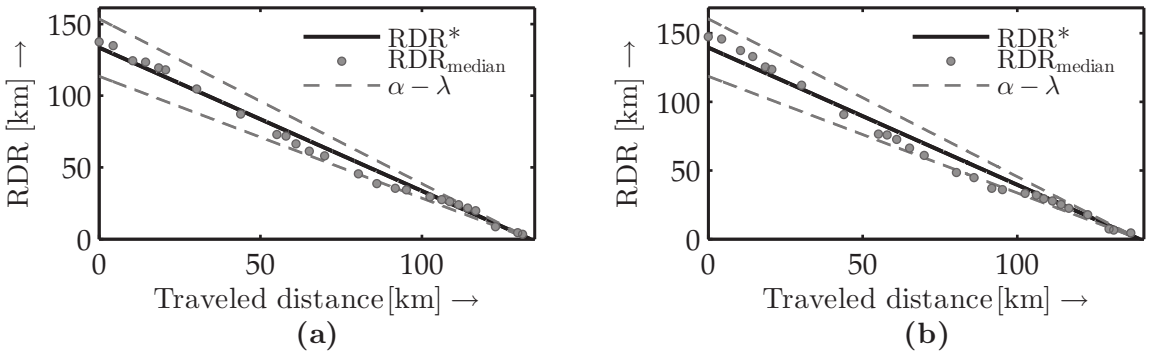

Figure 9.2.: RDR estimation results by accounting only for model uncertainty by a temperature of 25°C (**a**) and 38°C (**b**).

To assess the impact of model uncertainty we estimate the RDR of the EV by simulating $\boldsymbol{\mathcal{M}}_{PC}$ and $\boldsymbol{\mathcal{M}}_{ESS}$ through the driving cycle shown in figure 9.1, under different temperatures, and with a known initial state of charge ($\text{SoC}_0 = 1$). This way, the error between the *simulated* and the *real* RDR is attributed to the uncertainty related to model parameters. Figures 9.2 (**a**) and (**b**) exemplify the RDR estimation results of experiments carried out under 25°C and 38°C. It can be appreciated, the estimated RDR (circles) mostly lies within the predefined $\alpha - \lambda$ bounds. In this case we set $\alpha = 0.15$, that is, the bounds (dashed lines) are defined by a 15% above and below of the actual RDR (black line). Table 9.1 presents the mean relative accuracy ($\overline{\text{RA}}$) computed from predictions performed at different time steps k_p (see table C.2 and figure C.4 for further results on the model uncertainty assessment). Although the RA metric gives information at given time k_p, to better evaluate the performance of the RDR estimation methodology, it might be desirable to heavily weight the RA of RDR estimations that occur nearer the region of total depletion of the battery pack. To this aim we introduce the cumulative relative accuracy (CRA), which is given by:

$$\text{CRA}_i = \frac{1}{N_{\text{RA}_i}} \sum_{i=1}^{N_{\text{RA}_i}} \text{RA}_i \left(1 - \frac{1}{\Delta dist_i}\right) \tag{9.1.1}$$

where $\Delta dist_i$ is the traveled distance and N_{RA_i} is the number of prediction done up to the i^{th} RDR estimation. The RDR estimation achieves a highest $\overline{\text{RA}}$ of 91.8% at 06°C and a CRA of 85.48% at 25°C, which confirms the suitability of our approach.

9.2. Assessing present state uncertainty

Having assessed how model uncertainty affects the RDR estimation, the next step foresees introducing the present state uncertainty.

Table 9.2.: Mean and cumulative relative accuracy of the RDR estimation under different temperatures accounting for model and present state uncertainties.

T_{batt}	0°C	06°C	12°C	15°C	23°C	25°C	33°C	38°C
$\overline{\text{RA}}[\%]$	87.17	86.10	86.95	89.32	90.51	85.99	91.04	91.49
CRA$[\%]$	81.11	80.38	82.59	84.00	86.41	85.51	86.19	86.44

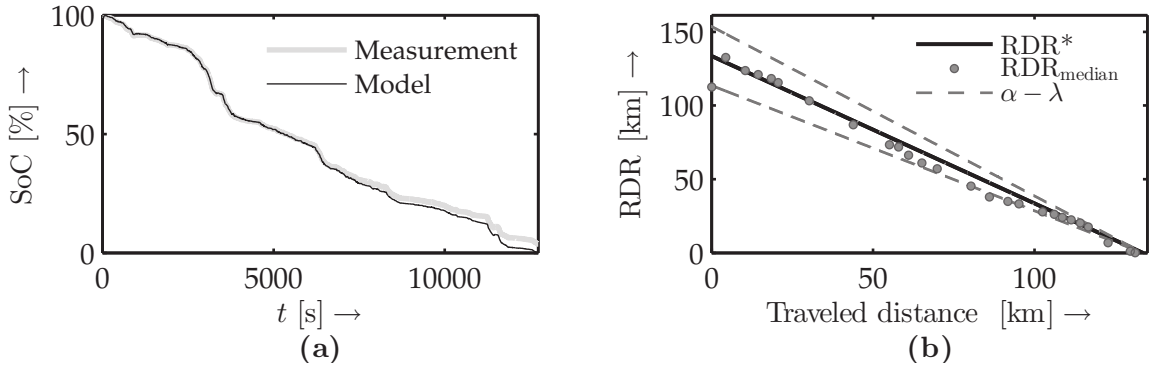

Figure 9.3.: SoC (**a**) and RDR (**b**) estimation results by accounting for model and present state uncertainties at a temperature of 25°C.

As it has been stated multiple times throughout this work, the present state uncertainty represents the lack of knowledge about the state of charge of the energy storage system. Thus, to assess its impact on the RDR estimation the experiments carried out in section 9.1 are repeated with the difference that value of SoC_0 is not set deterministically. We instead, uniformly distribute the set of samples of the particle flow particle filter between the range $0.05 \leq \text{SoC}_0 \leq 1$. As it can be seen in figure 9.3 (**a**), the SoC is successfully tracked for the entire operating range of the cell. The effect of present state uncertainty is reflected during the RDR estimation at $k_p = 1$. As it can be appreciated in 9.3 (**b**), the first RDR estimate fails with a RA of 84.25% (see table C.3 and figure C.5 for a more complete insight into the experimental results). This is due to the fact that the first SoC estimate also fails (recall the experiments presented in section 6.3), since at this point no measurements are available for updating the SoC information. Similarly to the results presented in the previous section, almost each RDR estimation lies withing the $\alpha - \lambda$ bounds, leading to a highest $\overline{\text{RA}}$ of 91.49% and a CRA of 86.44% at 38°C.

9.3. Assessing future load uncertainty

From the discussion above, we can clearly state that the RDR can be successfully estimated with the methodology introduced in section 2.2. However, the fact that the experimental results achieved above are high accurate is not surprising. This is mainly due to two reasons. On the one hand, in chapters 3 and 4 extensively set of experiments have been carried out to identify and to validate $\boldsymbol{\mathcal{M}}_{PC}$ and $\boldsymbol{\mathcal{M}}_{ESS}$, such that the power consumed by the Nissan leaf as well as the behavior of its battery pack can be accordingly simulated for most conventional operating conditions. On the other hand, it has been shown in chapter 6 that the particle flow particle filter

Table 9.3.: Mean and cumulative relative accuracy of the RDR estimation under different temperatures accounting for model, present state and future load uncertainties.

T_{batt}	0°C	06°C	12°C	15°C	23°C	25°C	33°C	38°C
$\overline{\text{RA}}$[%]	87.00	86.35	87.95	89.94	91.99	88.95	91.64	92.09
CRA[%]	82.11	81.41	82.50	84.90	86.75	83.95	87.22	87.40

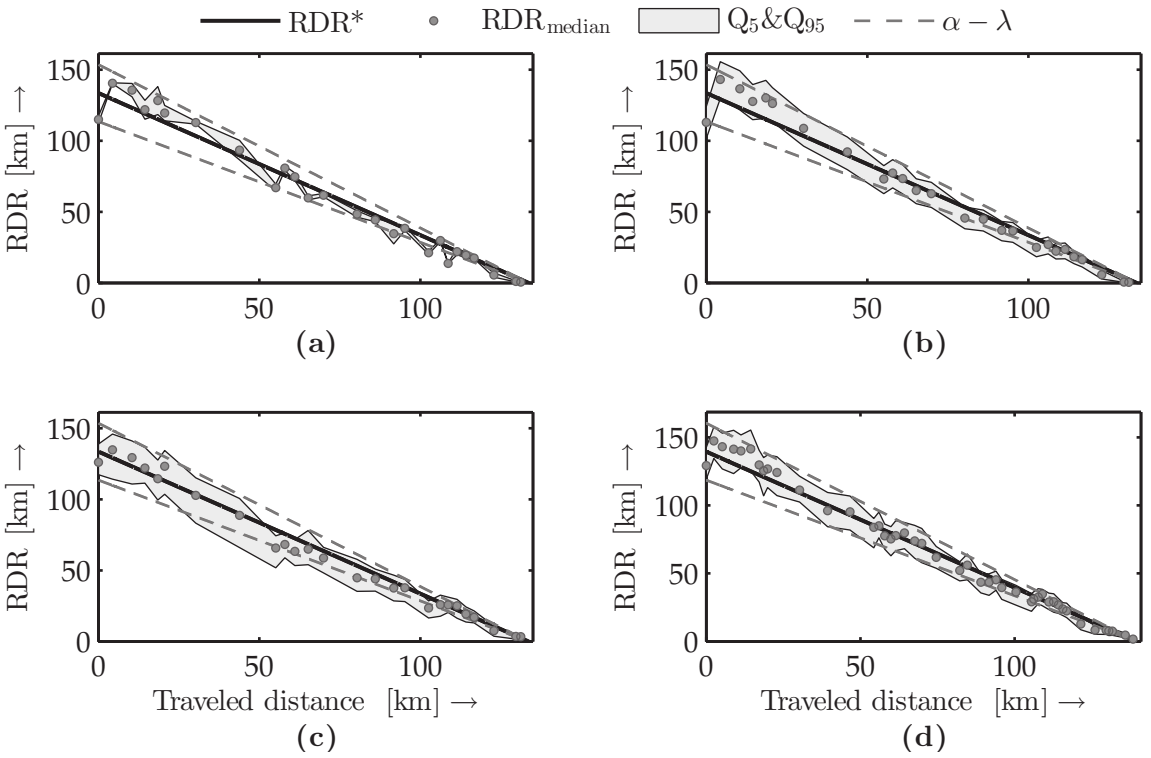

Figure 9.4.: RDR estimation results by accounting for model, present state and future load uncertainties at 25°C with $n_u = 2$ (**a**), $n_u = 1000$ (**b**), $n_u = 50$ (**c**) and with $n_u = 50$ at 38°C (**d**).

is robust against process, parameter and measurement noises and that it succeeds in tracking the true SoC within 100 iterations in the case of the single cell and within 500 iterations in the case of the real battery pack. Thus, it is expected that the RDR estimation, by only considering model and present state uncertainties, to be accurate.

A robust RDR estimation, however, also requires accounting for the uncertainty introduced by the future operating conditions. Thus, it is first necessary to determine how many realizations of the future load are needed in order to properly asses this type of uncertainty. Therefore, the experiments presented in section 9.2 are repeated here with the difference that, instead of following the driving cycle shown in figure 9.1, the models $\boldsymbol{\mathcal{M}}_{PC}$ and $\boldsymbol{\mathcal{M}}_{ESS}$ are evaluated by following a set of n_u driving cycles $\hat{\mathbf{u}}^{\text{env}}_{k_p:k_p+h_L}$ generated with the help of the stochastic environment model introduced in chapter 8. In all experiments both $\mathbf{M}^{\mathcal{F}_s}$ and $\mathbf{M}^{\mathcal{F}_{\text{road},s}}$ are parametrized with the *training* datasets of table B.1 for each driving situation. On the other hand, the transition probabilities of $\mathbf{M}^{\text{S}}$ are initialized with $1/m$, so that no previous information about the driving scenario is available. This allows to investigate the adaptation sywintroduced in section 8.2 in regards to the RDR estimation. Figures 9.4 (**a**) and (**b**) exemplify the

results obtained by estimating the RDR with $n_u = 2$ and $n_u = 1000$. As it can be appreciated, even though estimations performed with $n_u = 2$ exhibit acceptable results, the variance of the RDR estimate (gray region) is not well captured. In addition, the RDR estimation drastically jumps from prediction to prediction, which is results of using very low n_u values. On the other hand, as shown in figure 9.4 (**b**), the accuracy of the RDR estimation increases with $n_u = 1000$. Nevertheless, such a high number of $\hat{\mathbf{u}}^{\text{env}}_{k_p:k_p+h_L}$ realizations negatively affects the computational performance of the RDR estimation. Thus, the choice of n_u sets a trade-off between estimation accuracy and computational burden. It has been found, by repeatedly executing the estimation routine under different n_u values, that the $\overline{\text{RA}}$ and the CRA corresponding to a value of $n_u = 50$ exhibit good results for different temperatures, as shown in table 9.3 and figures 9.4 (**c**) and (**d**). Table C.4 and figure C.6 present further results by setting $n_u = 50$ under different temperatures. It can be appreciated that, even though no prior information about the driving cycle is assumed, the RDR is accurately estimated with $n_u = 50$, achieving a $\overline{\text{RA}}$ of 92.09% and a CRA of 87.40% at 38°C. In section 5.3 it has been concluded that the uncertainty regarding auxiliaries P_{aux} drastically influences the total uncertainty of the RDR. All experiments discussed so far have been performed under the assumption that $P_{\text{aux}} = P_{\text{aux,min}}$, where $P_{\text{aux}} = P_{\text{aux,min}}$ is residual power consumed by all secondary auxiliaries (recall section 3.5 and table A.2). Under real operating conditions, however, this assumption is not completely valid in the case of electric vehicles. The reason for this is that during a trip the driver might constantly turn on and off different auxiliaries, e.g., the HVAC system. This fact adds uncertainty to the future load and therefore it needs to be accounted for. For example, neglecting the component of the future load leads to over-estimations of the RDR (see figure C.7 (**c**)). In this example, a virtual load of 4000 W is turned on after 2000 s, resembling in this manner the sudden use of heating or air conditioning. To overcome this issue and thereby to account for this uncertainty in the RDR estimation, we allow $\boldsymbol{\mathcal{M}}_{PC}$ to compute online P_{ele} as response to v_{x} and α_{road} and measure the actual power consumed. Based on the difference $\Delta P_{\text{ele}} = P_{\text{ele,real}} - P_{\text{ele,sim}}$ and by setting a threshold, we are able to differentiate between two new operating states, that is, $s_{P_{\text{aux}}} = 1$ if important auxiliaries are turned on $s_{P_{\text{aux}}} = 0$, otherwise.

Figure C.7 (**a**) shows the results of the simple classification scheme described above. It can be seen, that this approach successfully recognizes when an external load is turned on with an accuracy of 87.58 % for this example. The results of the RDR estimation are presented in figure C.7 (**c**). In this case a $\overline{\text{RA}}$ improvement of 21.20 % is achieved by including the recognized external load into the RDR estimation.

All experiments presented so far have been carried out with the second scenario described in section 2.3. This means, that even though the time series of v_{x}, α_{road} and P_{ele} correspond to real measurements taken with the Nissan leaf, the inherent uncertainty, imposed by the variability of the single cells conforming the battery pack, is not taken into account. Furthermore, to make more meaningfully statements about the performance of the RDR estimation approach developed throughout this work, it is necessary to test it on the real vehicle and under real operating conditions. This might seem to be a trivial task, however, in practice carrying out such tests is not feasible since it would require to drive the electric vehicle *to empty*, that is, until the battery pack is totally discharged.

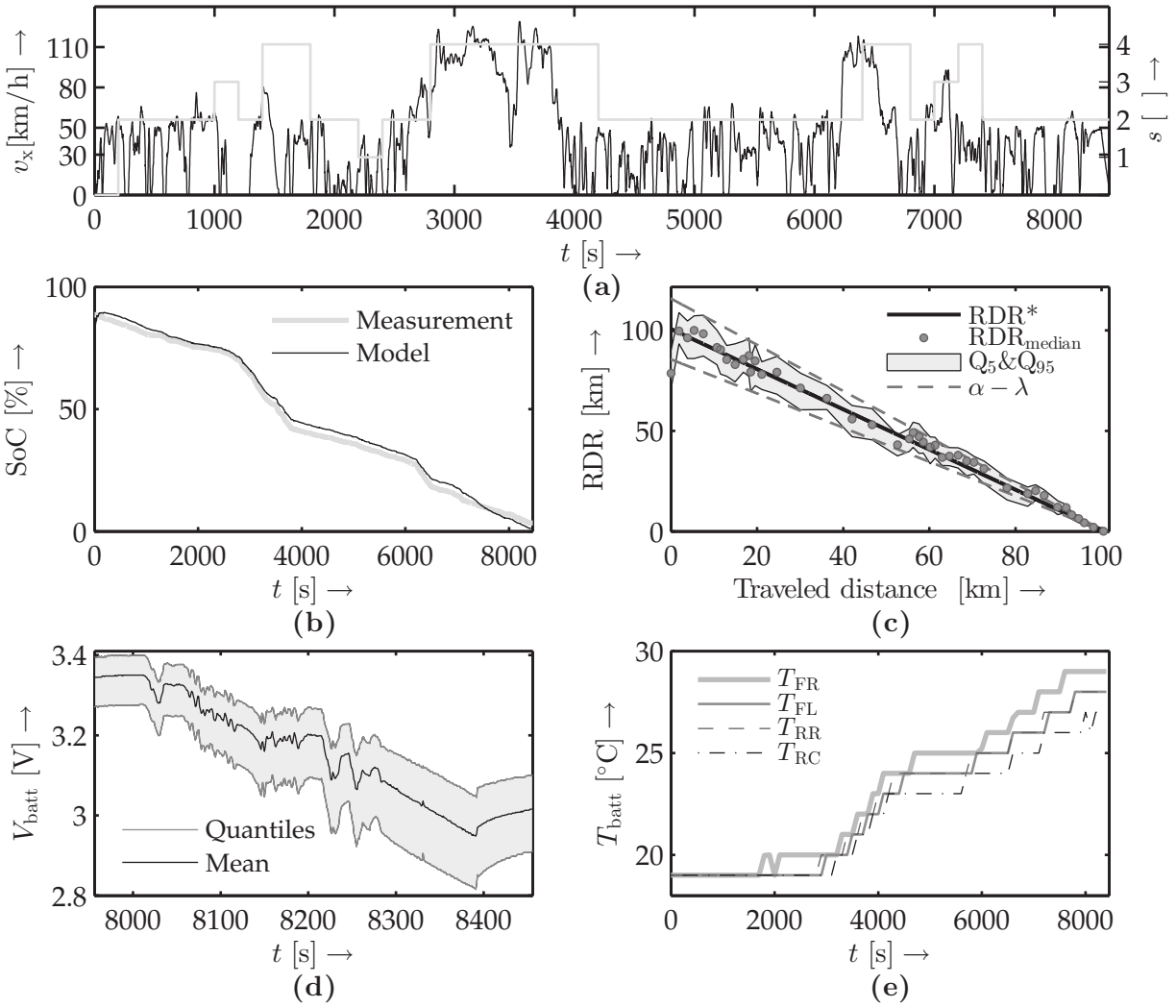

Figure 9.5.: Speed (black line) and driving condition classification (gray line) (**a**). Results of the SoC (**b**) and RDR (**c**) estimation of experiments performed on the Nissan Leaf. Variability in terminal voltage (**d**) and temperature (**e**) of the cells conforming the battery pack.

To overcome this issue a third experimental scenario (recall section 2.3) has been developed. During the experiment the EV follows the predefined drive cycle of figure 9.1. Note, that the speed of the vehicle slightly differs from the predefined one. This is due to the disability of the driver to exactly follow the driving cycle shown on the screen. Figure 9.5 (**a**) shows the actual speed of the vehicle (black line), measured from the tachometer, and the result of the driving condition classification (gray line) applied to this drive cycle. It is observed in figure 9.5 (**b**) that the approach presented in chapter 6 succeeds on tracking the value of the SoC for the entire operating range, with an RMSE 2.45 %. The RDR estimation shown in figure 9.5 (**c**) achieves a $\overline{\text{RA}}$ of 90.87 % and a CRA of 84.15 %. It is observed in figure 9.5 (**d**) that V_{batt} varies from cell to cell. This variability causes the battery management system to turn off the vehicle when an arbitrary cell reaches the predefined SoC_{min}. Moreover, the temperature among all cells vary depending on their location on the battery pack (see figure 9.5 (**e**) and recall figure 2.6). These two factors cause the uncertainty of the battery pack mentioned above, which are successfully accounted in the RDR estimation.

10
Conclusions and Outlook

Mobile systems have become very important in terms of supporting the human being in a broad spectrum of applications both in hostile environments, where the presence of the humans might be prohibitive, or just for increasing the comfort. Nevertheless, the autonomy of mobile systems is strongly affected by multiple internal and external factors and its accurate determination is remarkably distorted by many sources of uncertainty. This work is first and foremost driven by the idea that assessing the autonomy of a mobile system can be formulated under an uncertainty propagation context, in which these uncertainties are characterized and propagated with the aim of computing their combined effect as a probability density function.

The complexity of such estimation increases with the number of uncertain variables taken into consideration. With the purpose of reducing the number of uncertain variables, and therefore the computational complexity of the estimation, this work introduces a novel approach to quantify and analyze uncertainties and therefore prioritize important variables and to neglect the uncertainty of those variables with meaningless contribution. To analyze these uncertainties and to prioritize the set of important uncertain variables, detailed models describing the nonlinear capacity behavior exhibited by the energy storage system used to power the mobile system as well as the power requirements imposed by its intrinsic interaction with the operating environment have been developed based on physical principles. These models have been accordingly parametrized and validated through comprehensive series of experiments carried out in various experimental setups specially designed to address different theoretical concepts and based on the type of data available.

The present state uncertainty reflects the lack of knowledge about the energy available in the system at any given time. This quantity is expressed indirectly in terms of system states and parameters, which however, are not observable and consequently have to be estimated. A novel sample-based nonlinear observer for tracking faulty states of the system, and specially, for estimating the remaining energy available in the energy storage system has been implemented. This observer combines the benefits of two state of the art filters, namely, the UKF and the PF, in that it makes no assumption about the distribution of the states and parameters and that it progressively transports the samples to the correct locations from the prior to the posterior following an homotopy dictated by the Bayes' rule without needing to randomly sample from any distribution, reducing in this manner the number of samples needed and thereby the computational effort of the estimation.

The paramount challenge encountered in this work is given by the uncertainty characterization and integration of the future energy consumption of the system. Due to the unstructured operating environment of the mobile system, the future energy needs are governed by a stochastic process, and therefore can not be characterized by stationary probability distributions. In addition to this, estimating the autonomy requires knowledge about the future system usage from the time at which the estimation is done to the time at which the systems runs out of energy. Unfortunately, this time interval is not available in advance, since it represents the actual outcome of the estimation approach. With the aim of capturing the stochastic behavior dictated by changing operating conditions, a stochastic environment model from a Markovian point of view has been developed in this work. In order to allow the RUL estimation to adapt to changing operating conditions, it is necessary to update the parameters of the Markov model as new information becomes available. The use of a recursive Bayesian posterior probability approach has been proposed to this aim. After having developed the stochastic model, an approach to extract, select and compute features from measured variables of the environment has been developed, with the aim of distinguishing, as good as possible, between the different operating conditions. Gathering training data for the classifier represents a challenge, since it is required to operate the system for long periods of time under different operating conditions. This is not feasible in our case due to the constraints in energy availability imposed by the type of systems considered within this work. Because of this, a classifier with low expected generalization error and capable of responding to time-varying influences is investigated in this work.

On the basis of all experimental results achieved throughout this work, we can argue that the model-based prognostics methodology proposed in section 2.2 can be successfully implemented for assessing the autonomy of mobile systems operating in highly dynamic environments, more precisely, to estimate the RDR of electric vehicles.

As it became clear, the RUL estimation problem exhibits acausal properties, that is, it requires information about the future operating conditions in order accurately forecast the temporal behavior of the states and parameters of system. Furthermore, it is needed to have precise knowledge about the true RUL of the system in order to obtain some feedback that allows evaluating the performance and accuracy of the RUL estimation algorithm. Most evaluation methods available in the literature follow a run-to-failure (RtF) philosophy, in which controlled experiments are carried out in order to gather failure data, and therefore, to experimentally build a probability density function of the RUL. This family of approaches is best suited for evaluating prognostics performance once all data has been gathered and its applicability is limited to offline analysis. Nevertheless, an online evaluation the performance of the RUL estimation is one of the most important factors for prognostics-enabled decision making. Online validation of RUL estimates, however, requires incorporating a more precise characterization of the future operating conditions. For this reason, future work foresees to investigate the use of outer feedback correction loops (OFCLs) to predict, under different prediction horizons, the temporal evolution of states and parameters of the system and to derive metrics that allow mapping the accuracy of short term predictions to the performance of the RUL estimation.

A

Model Parameters

A.1. Discussion on all known power consumption model parameters

The total mass ($m_{V,t}$) is composed of the mass of the Nissan Leaf, the passengers and the measurement equipment. The mass of both the curb and the measurement equipment are obtained from technical specifications publicly disclosed. In this case, the vehicle can be occupied by a minimum of one passenger (the driver) and a maximum of four.

The ambient related parameters, i.e., air density (ρ_{air}) and gravitational acceleration g, correspond to average spring/summer conditions of Dortmund city ant its surroundings. Similarly to $m_{V,t}$, the nominal values of the chassis parameters, that is, A, l_f, l_r and H_G are drawn from technical sheets published by Nissan.

While the magic formula (see equation (3.3.3)) succeeds in describing the steady-state behavior of the traction force of a tire, experimentally determining its parameters requires relatively complex procedures, which are out of the scope of this work. Since the Nissan Leaf is equipped with standard tires, namely 205/55R16, we opt for computing B_x, C_x and D_x based on information published by Pacejka (2005). As discussed in section 3.3, the curvature factor (E_x) is set to zero such that equation (3.3.3) can be conveniently solved w.r.t. κ_x.

The Nissan Leaf is propelled by a 80 kW and 280 Nm synchronous electric motor whose efficiency modeled by means of a stationary map (η_M). The data employed to describe the power losses for the region operating in motor mode and for computing the losses in generator mode has been published by Sato et al. (2011). In the same manner, nominal values of the gear ratio i_G and the transmission efficiency η_G of the transaxle are drawn from technical sheets.

For the sake of completeness, table A.1 summarizes the nominal values of all parameters belonging to the power consumption model.

A.2. Power consumption model parameters

This section presents the parameters of the model describing the power requirements of the electric vehicle grouped according to whether they belong to sub-models of the to the vehicle dynamics, tires, driveline or electric components.

Table A.1.: Nominal values of the $\mathcal{M}_{PC}$ parameters

	parameter	symbol	unit	value
Vehicle dynamics	total mass of the electric vehicle	$m_{\mathrm{V,t}}$	kg	1645
	density of air	ρ_{air}	$\mathrm{kg/m^3}$	1.226
	aerodynamic drag coefficient	c_{w}		0.2931
	frontal area of the electric vehicle	A	$\mathrm{m^2}$	2.27
	acceleration due to gravity	g	$\mathrm{m/s^2}$	9.81
	static rolling resistance coefficient	K_{R_0}		0.014
	first speed dependent rolling resistance coefficient	K_{R_1}		9.22×10^{-4}
	second speed dependent rolling resistance coefficient	K_{R_4}		2.97×10^{-5}
	distance from the center of gravity to the front axle	l_{f}	m	1.188
	distance from the center of gravity to the rear axle	l_{r}	m	1.512
	height of vehicle center of gravity measured from the ground	H_{G}	m	0.5
Tire	effective rolling-radius of the tire	r_{W}	m	0.3053
	stiffness factor of the magic formula	B_{x}		17.09
	shape factor off the magic formula	C_{x}		1.65
	maximum transmittable longitudinal force	D_{x}		2.97×10^{-3}
	curvature factor of the magic formula	E_{x}		0
Driveline	gear ratio of the transaxle (driveline)	i_{G}		7.9377
	efficiency of the driveline	η_{G}		0.97
	rotational inertia of output driveline components	$I_{\mathrm{G,out}}$	$\mathrm{kgm^2}$	0.533
	rotational inertia of input driveline components	$I_{\mathrm{G,in}}$	$\mathrm{kgm^2}$	0.9
Electric components	efficiency map of the electric drive/motor	η_{M}		[0.795,0.922]
	minimum activation speed for regenerative braking	$v_{\mathrm{x,min}}^{\mathrm{reg}}$	km/h	55
	maximum activation speed for regenerative braking	$v_{\mathrm{x,max}}^{\mathrm{reg}}$	km/h	110
	power required by auxiliary components	P_{aux}	W	457

A.3. Energy storage system model parameters

This section presents the parameters of the model describing the energy storage system grouped according to whether they belong to sub-models of the transient response, cell capacity or the open circuit voltage.

Table A.2.: Nominal values of the $\mathcal{M}_{ESS}$ parameters.

	parameter	symbol	unit	value
Transient response	ohmic resistance	R_{ohm}	Ω	$[7.43 \times 10^{-4}, 0.0027]$
	resistance of the short-time RC-network	R_{s}	Ω	$[1.42 \times 10^{-5}, 1.3 \times 10^{-4}]$
	capacitance of the short-time RC-network	C_{s}	F	$[5.91 \times 10^{3}, 1.3 \times 10^{5}]$
	resistance of the long-time RC-network	R_{l}	Ω	$[3.43 \times 10^{-4}, 0.0017]$
	capacitance of the long-time RC-network	C_{l}	F	$[4.4 \times 10^{4}, 9.86 \times 10^{4}]$
Cell capacity	cell capacity ratio	c		0.958
	virtual conductance of the cell	d		0.8×10^{-5}
	discrete system parameter of the KiBaM	a_1		1
	discrete system parameter of the KiBaM	a_2		0.0002
	discrete system parameter of the KiBaM	a_3		0
	discrete system parameter of the KiBaM	a_4		0.9998
	discrete system parameter of the KiBaM	b_1		0.5
	discrete system parameter of the KiBaM	b_2		0
	cell nominal capacity	C_n	Ah	59.5
Open circuit voltage	parameter of the V_{OC}-SoC model	K_0	V	2.095
	parameter of the V_{OC}-SoC model	K_1	V	0.0041
	parameter of the V_{OC}-SoC model	K_2	V	-20.39
	parameter of the V_{OC}-SoC model	K_3	V	-0.2531
	parameter of the V_{OC}-SoC model	K_4	V	-0.004
	parameter of the V_{OC}-SoC model	K_5	V	-17.1769
	parameter of the V_{OC}-SoC model	K_6	V	2.2476
	parameter of the V_{OC}-SoC model	K_7	V	-3.4802
	parameter of the V_{OC}-SoC model	K_8		0.2962
	parameter of the V_{OC}-SoC model	K_9	V	-2.4952
	parameter of the V_{OC}-SoC model	K_10		-2.7897

B
Databases

B.1. On-road measurements

The following figures present the times series of the speed and the road slope of the datasets employed throughout this work either for training or for validation.

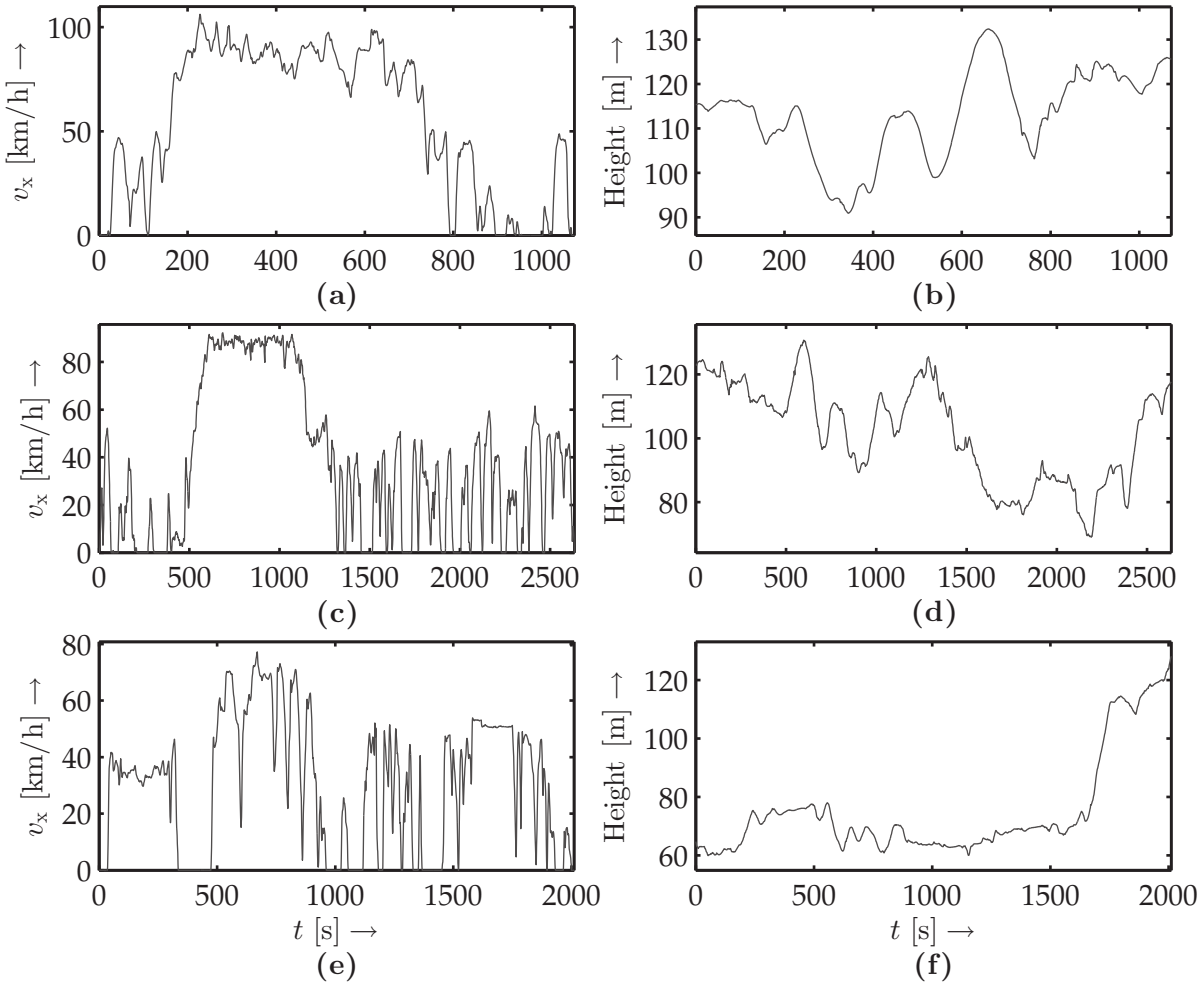

Figure B.1.: Speed (**a**)-(**c**)-(**e**) and road slope (**b**)-(**d**)-(**f**) time series corresponding to 3 out of 8 training datasets.

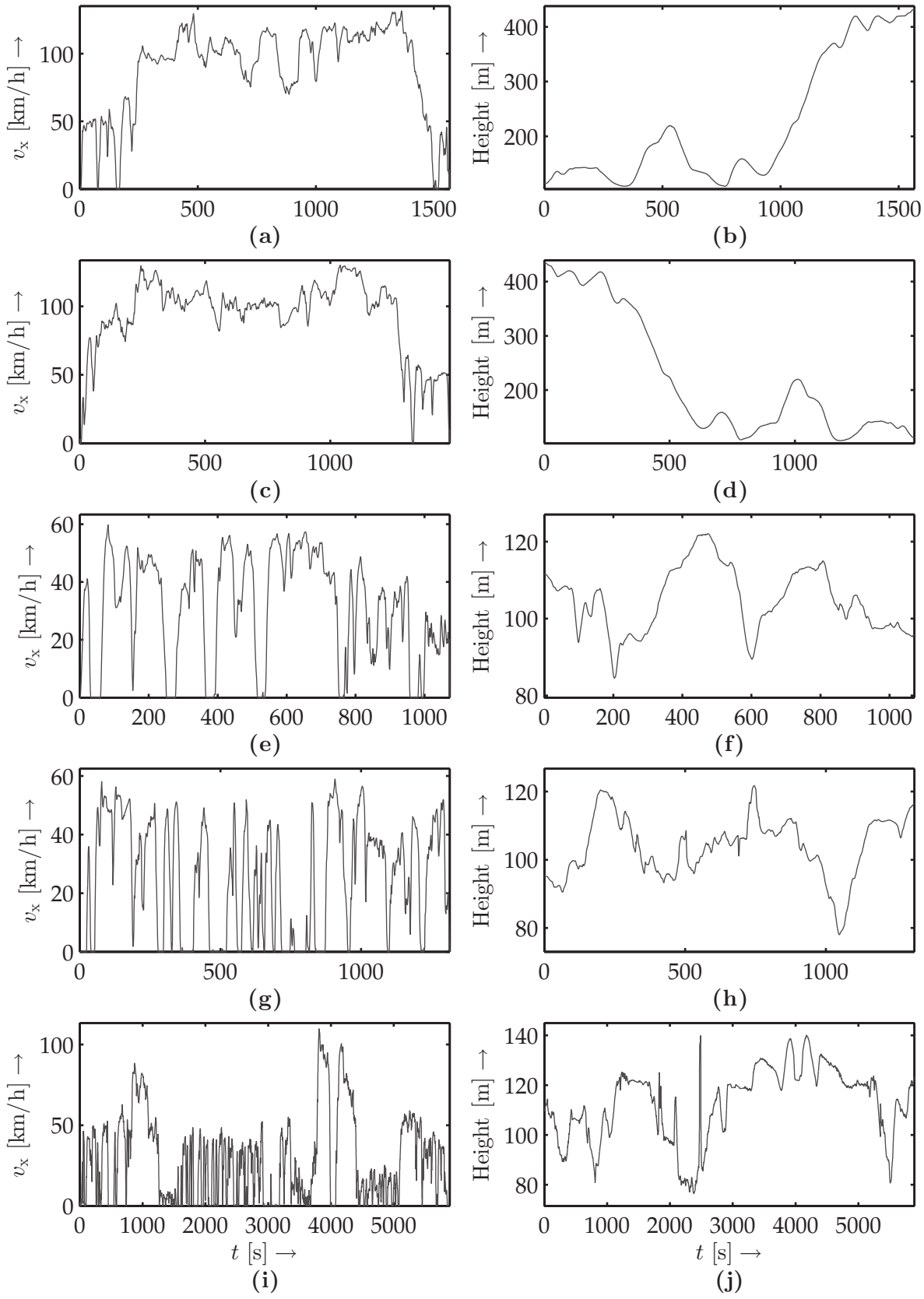

Figure B.2.: Speed (**a**)-(**c**)-(**e**)-(**g**)-(**i**) and road slope (**b**)-(**d**)-(**f**)-(**h**)-(**j**) time series corresponding to 5 out of 8 training datasets.

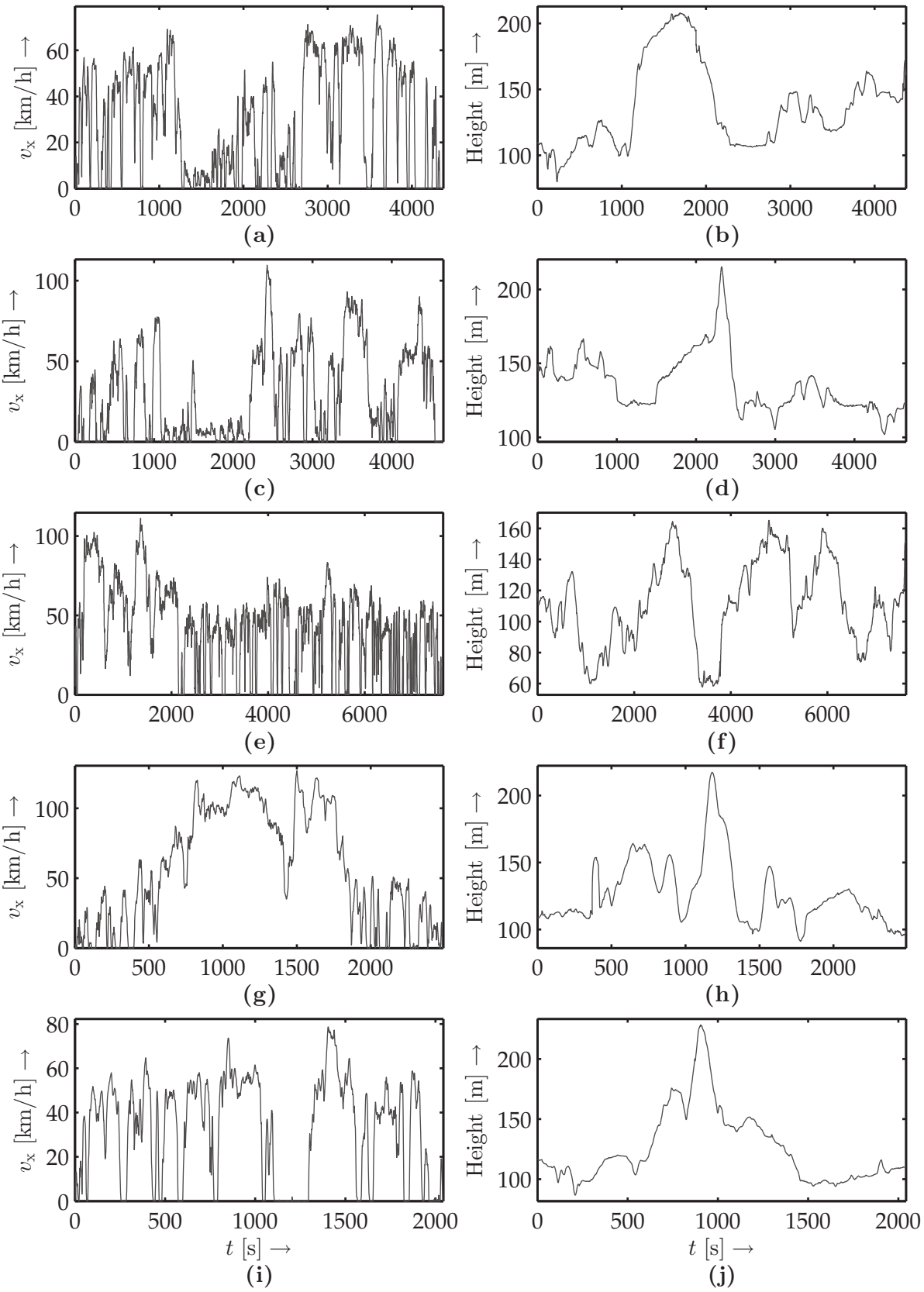

Figure B.3.: Speed (**a**)-(**c**)-(**e**)-(**g**)-(**i**) and road slope (**b**)-(**d**)-(**f**)-(**h**)-(**j**) time series corresponding to 5 out of 7 validation datasets.

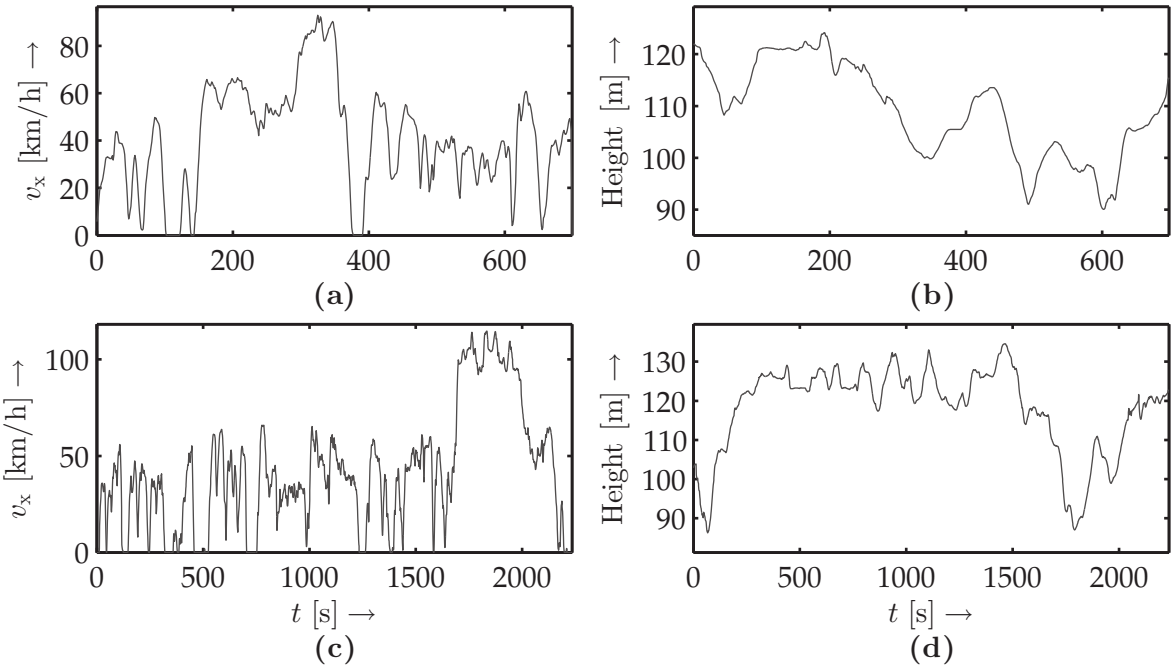

Figure B.4.: Speed (**a**)-(**c**) and road slope (**b**)-(**d**) profiles of 2 out of 7 training datasets.

Table B.1 summarizes all relevant information pertaining both training and validation datasets. For the sake of clarity, a dataset with a road condition labeled as *dry* is assumed to have a rolling resistance coefficient in the range 0.012-0.016. On the contrary, road conditions labeled as *wet* are assumed to have a coefficient ranging between 0.006 and 0.009.

Table B.1.: Summary of driving characteristics of measurements (drive cycles) taken on-road.

	Drive cycle	Traveled distance [km]	Mean speed [km/h]	Height difference [m]	Nr. of drivers [ ]	Road conditions [ ]
Training	Figure B.1 (**a**) and (**b**)	16.47	55.20	10.18	1	dry
	Figure B.1 (**c**) and (**d**)	27.27	37.26	-5.41	2	wet
	Figure B.1 (**e**) and (**f**)	17.03	30.45	63.33	1	dry
	Figure B.2 (**a**) and (**b**)	38.81	89.10	318.51	2	dry
	Figure B.2 (**c**) and (**d**)	38.23	93.11	-322.75	2	dry
	Figure B.2 (**e**) and (**f**)	9.34	31.32	-15.75	2	dry
	Figure B.2 (**g**) and (**h**)	9.17	25.08	21.15	2	dry
	Figure B.2 (**i**) and (**j**)	47.56	29.05	13.44	2	wet
Validation	Figure B.3 (**a**) and (**b**)	35.93	29.57	51.00	2	wet
	Figure B.3 (**c**) and (**d**)	38.11	29.48	-26.25	2	wet
	Figure B.3 (**e**) and (**f**)	90.42	42.66	33.31	2	wet
	Figure B.3 (**g**) and (**h**)	39.41	56.93	-10.42	1	dry
	Figure B.3 (**i**) and (**j**)	18.53	32.64	-5.67	1	dry
	Figure B.4 (**a**) and (**b**)	7.91	40.70	-6.34	1	dry
	Figure B.4 (**c**) and (**d**)	25.87	41.65	19.84	1	dry

Standard driving cycles

This section presents the standard driving cycles (figure B.5 (**a**)-(**c**)) employed in chapter 5 for performing the uncertainty analysis of the RDR estimation. The cycles have been first introduced by Andre (1996, 2004). The road slope profile (figure B.5 (**d**)) has been synthetically generated with aid of the training datasets presented in section B.1.

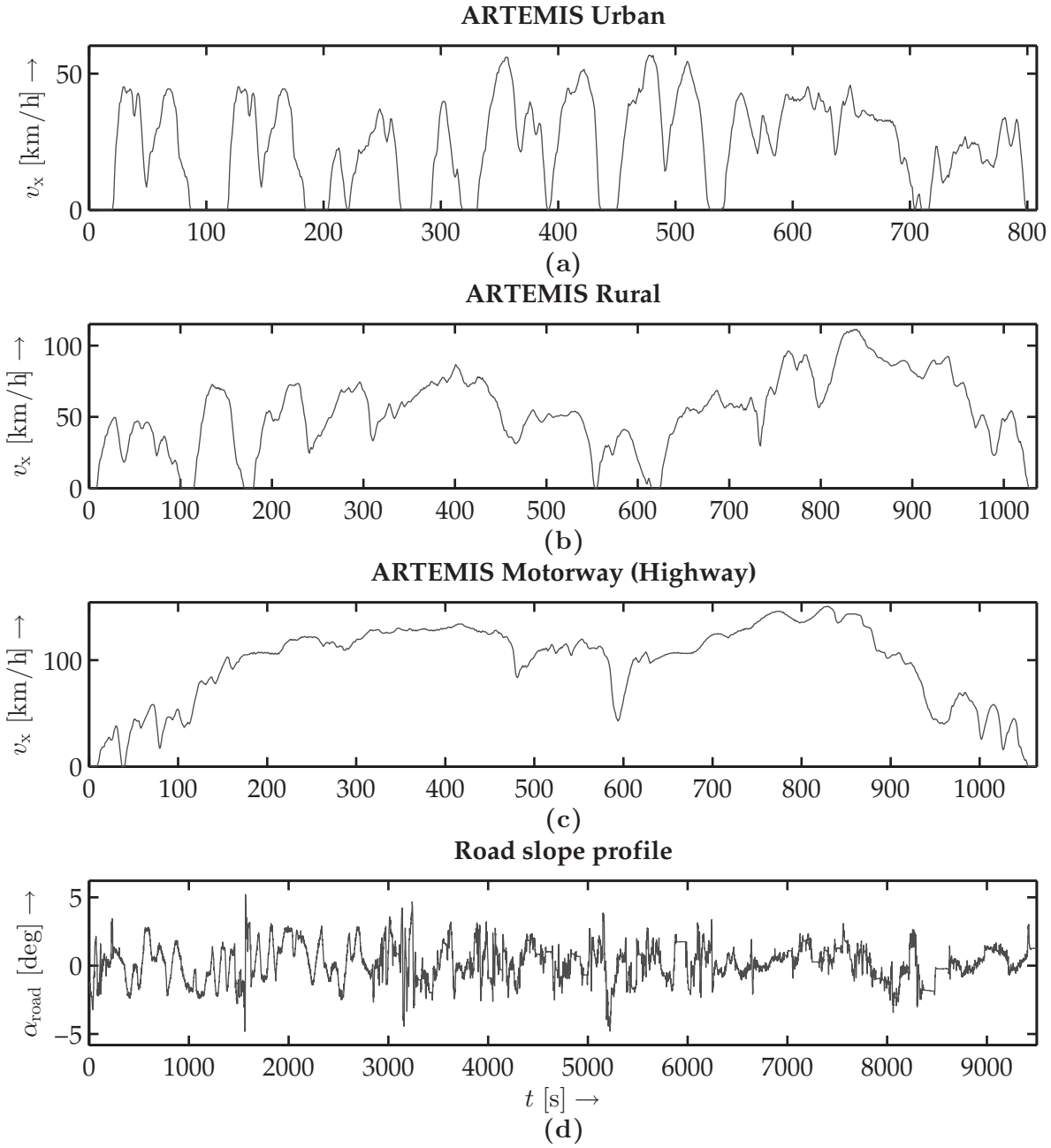

Figure B.5.: Standard driving cycles (**a**)-(**c**)) and synthetic road slope profile (**d**) used in the uncertainty analysis presented in chapter 5.

C
Further Results

C.1. Power consumption model evaluation results

This section presents evaluation results by applying the set of parameters, with the lowest mean RMSE, on the remaining datasets.

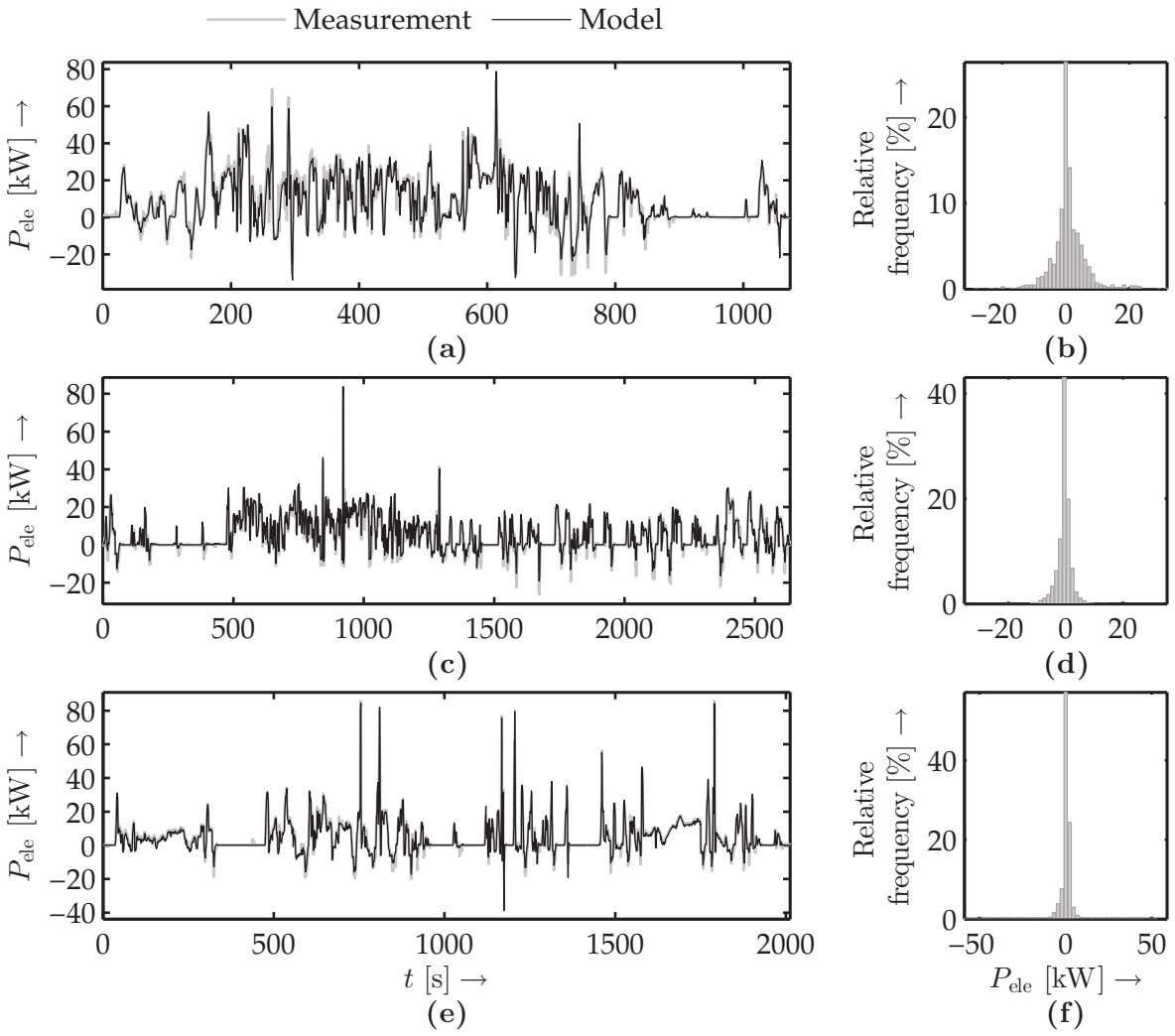

Figure C.1.: Illustration of 3 out of 12 evaluation results of the power consumption model.

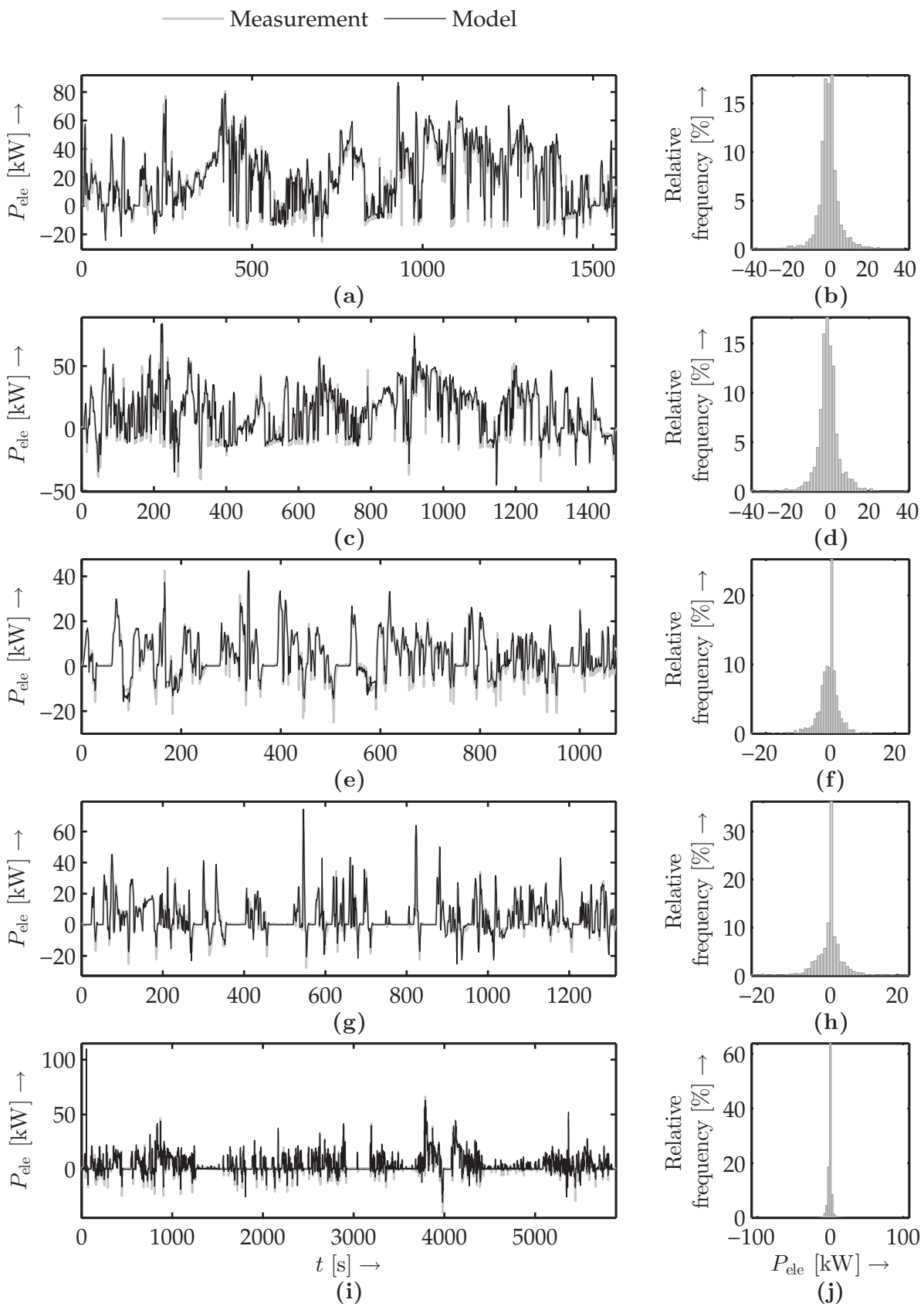

Figure C.2.: Illustration of 5 out of 12 evaluation results of the power consumption model.

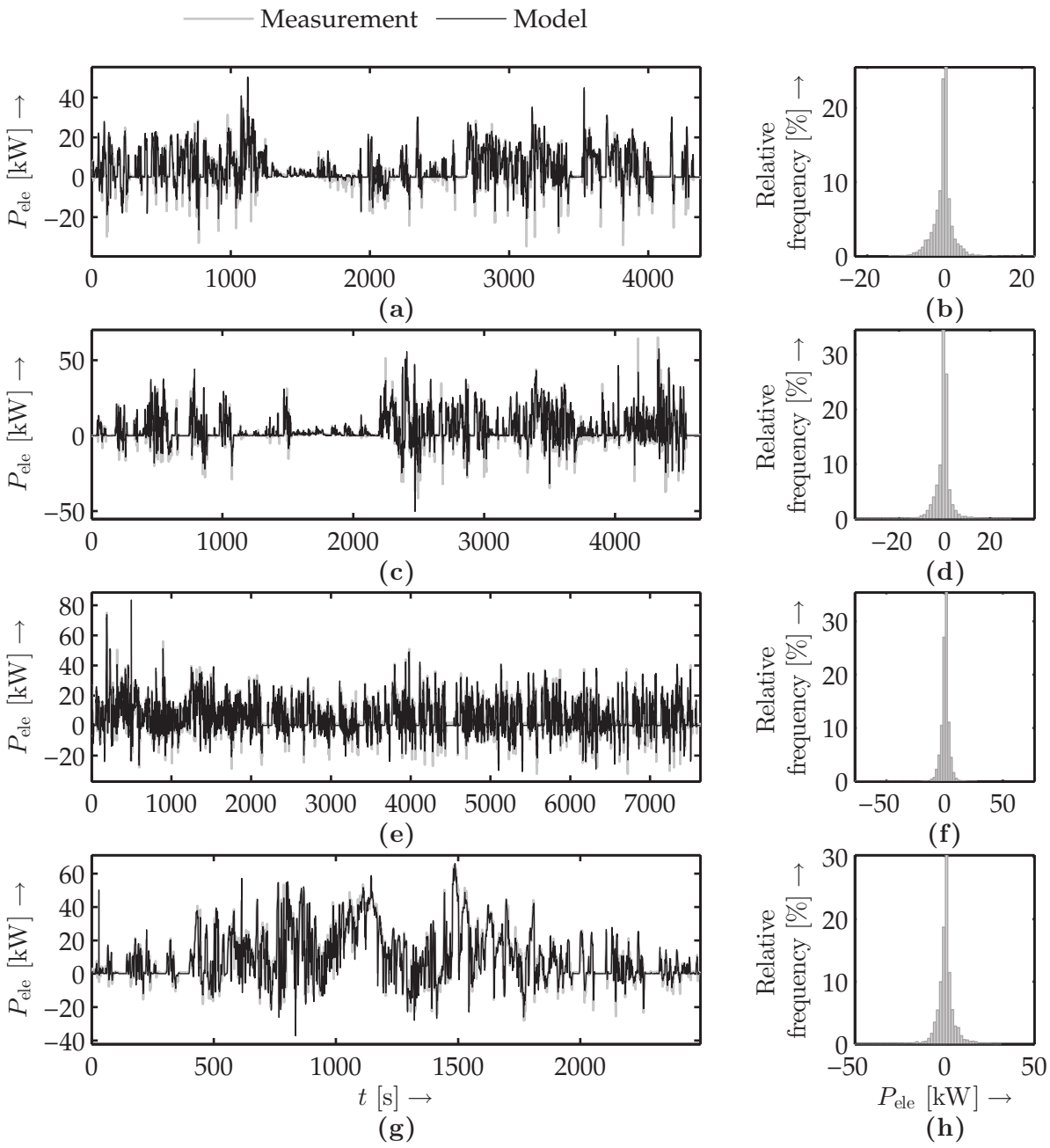

Figure C.3.: Illustration of 4 out of 12 validation results of the power consumption model.

C.2. Features

The features presented in this section are grouped in four categories depending on which information is needed for their computation. Accordingly, we distinct between speed, acceleration, driving characteristic and cumulative values. This features are the result of the feature extraction scheme introduced in section 7.2.

Table C.1.: All features extracted form experimental datasets with aim of driving situation classification.

Cat.	Nr.	Name	Unit
Speed	9	Mean speed	km/h
	10	Mean speed (without stops)	km/h
	13	Speed standard deviation	km/h
	14	Maximum speed	km/h
Acceleration	15	Mean acceleration	m/s^2
	16	Mean positive acceleration	m/s^2
	17	Mean negative acceleration	m/s^2
	18	Acceleration standard deviation	m/s^2
	19	Maximum acceleration	m/s^2
Driving characteristics	3	Relative stop time	[−]
	4	Relative driving time	[−]
	5	Relative acceleration	[−]
	6	Relative deceleration	[−]
	7	Relative braking	[−]
	8	Relative cruise time	[−]
	11	Acceleration per kilometer	1/km
	12	Stops per kilometer	1/km
Cum.	1	Total distance	m
	2	Total time	s

C.3. Complete set of results of the remaining driving range estimation

The discussion presented throughout chapter 9 repeatedly refers to $\overline{\text{RA}}$ and CRA values, which have been computed with more information than the shown in the corresponding sections. For the sake of summarizing, each discussion is based on 2 or 4 figures that schematically depict the experimental results. With the aim of offering deeper insights into the set of experiments performed in chapter 9 we present in the sequel the complete set of RDR estimation results obtained under various temperature conditions and by assessing, analogous to chapter 9, both model, present state and future load uncertainties. The results shown towards the end of this section, more precisely, the experiments shown in figure C.7 are referred in section 9.3 where a discussion, about considering residual power consumption values, takes place. The purpose of the results depicted in figures C.8, C.9 and C.10 is to illustrated the performance of our RDR estimation approach for driving cycles other than the one introduced in figure 9.1 and discussed throughout chapter 9.

Table C.2.: Further results of model uncertainty assessment in the RDR estimation.

	RA [%]							
k_p	0°C	06°C	12°C	15°C	23°C	25°C	33°C	38°C
1	95.09	97.20	93.83	93.91	93.87	97.04	94.38	94.40
4	93.06	96.82	94.38	94.55	94.57	96.45	93.74	93.76
7	94.06	99.09	95.64	96.15	96.53	98.76	97.39	97.42
10	99.60	94.49	94.42	94.14	93.80	94.95	93.01	92.87
13	93.63	88.84	88.89	88.69	88.45	91.03	87.99	87.58
16	92.02	82.48	79.54	79.37	79.60	84.59	77.48	77.47
19	87.11	94.61	90.78	90.76	92.49	99.64	95.87	95.93

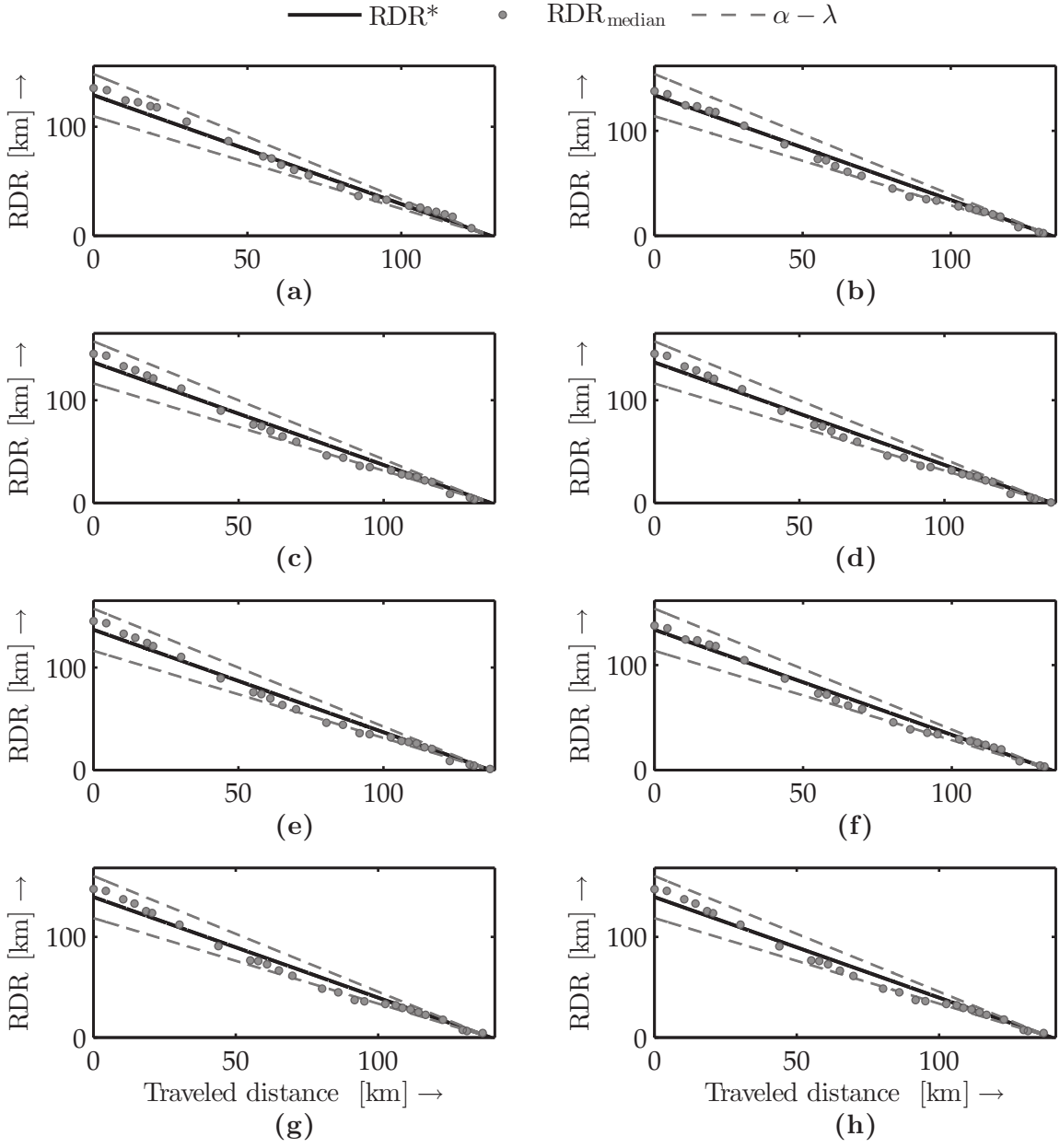

Figure C.4.: Further results of assessing model uncertainty in the RDR estimation under temperatures of 0°C (**a**), 06°C (**b**), 12°C (**c**), 15°C (**d**), 23°C (**e**), 25°C (**f**), 33°C (**g**) and 38°C (**h**).

Table C.3.: Further results of present state uncertainty assessment in the RDR estimation.

	RA [%]							
k_p	0°C	06°C	12°C	15°C	23°C	25°C	33°C	38°C
1	89.99	87.07	86.58	86.55	86.76	84.25	85.63	85.65
4	89.86	94.22	91.39	91.07	90.06	98.46	90.26	90.24
7	93.59	97.03	94.54	94.48	93.00	99.88	94.47	94.28
10	89.48	90.49	94.61	96.34	97.92	94.95	96.59	97.34
13	81.78	80.56	88.45	89.88	92.95	89.51	93.37	94.21
16	76.60	77.80	78.57	79.51	81.40	83.48	82.74	84.28
19	88.45	81.90	89.55	90.66	99.96	95.38	96.67	96.81

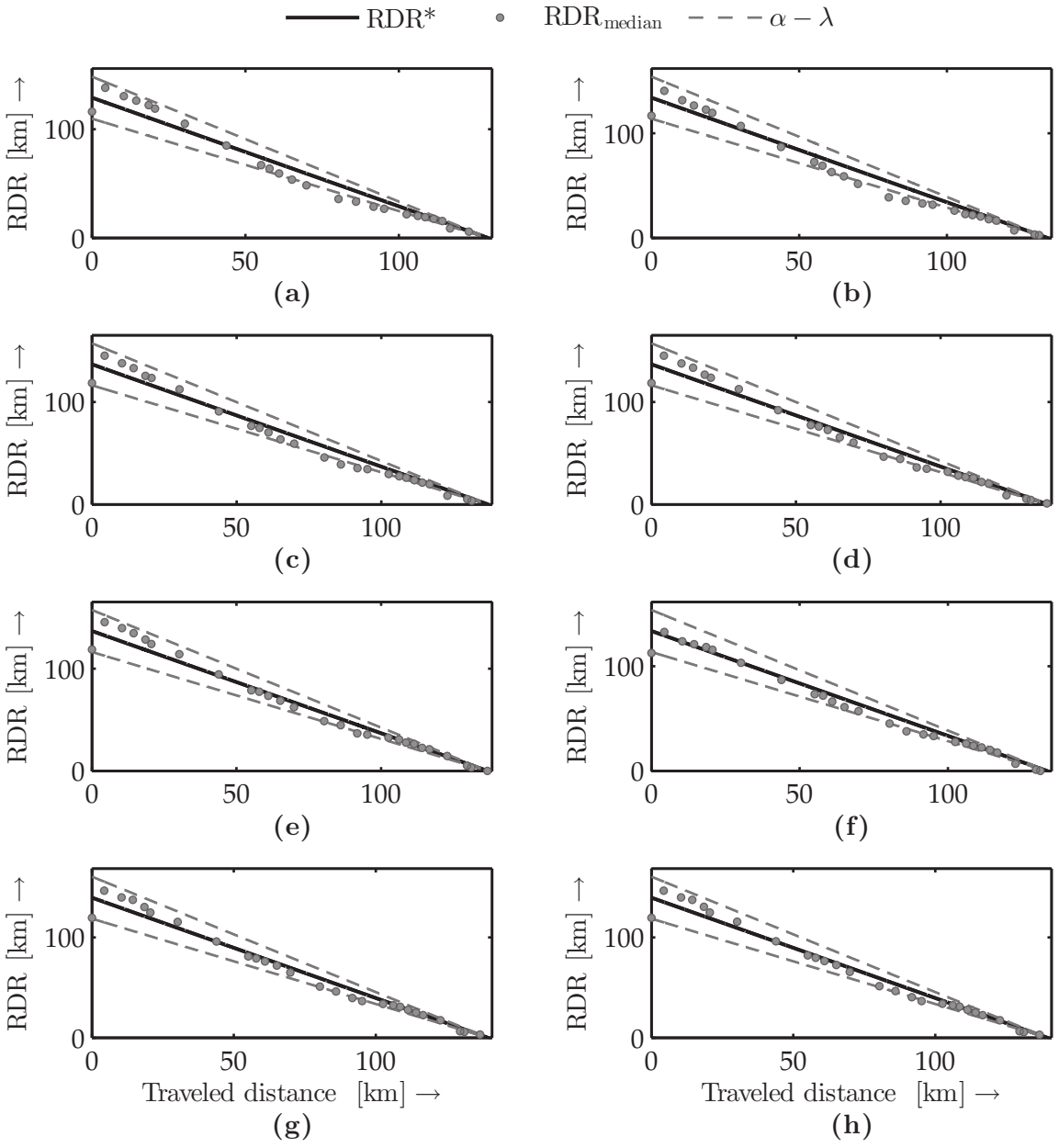

Figure C.5.: Further results of assessing present state uncertainty in the RDR estimation under temperatures of 0°C (**a**), 06°C (**b**), 12°C (**c**), 15°C (**d**), 23°C (**e**), 25°C (**f**), 33°C (**g**) and 38°C (**h**).

Table C.4.: Further results of future load uncertainty assessment in the RDR estimation.

	RA [%]							
k_p	0°C	06°C	12°C	15°C	23°C	25°C	33°C	38°C
1	85.47	82.35	89.46	82.24	90.36	93.88	95.12	92.56
6	88.26	90.20	89.39	96.36	92.34	87.65	91.62	86.91
11	95.69	99.97	92.28	96.25	95.78	94.98	97.97	98.23
16	96.22	94.91	93.30	91.71	96.22	91.13	97.82	95.14
21	84.97	87.85	86.86	97.29	98.91	99.48	99.04	96.61
26	82.44	86.21	81.97	93.96	97.73	96.50	96.13	91.03
31	74.80	81.42	94.00	94.32	94.80	92.41	90.02	93.06
36	94.09	81.75	86.13	96.94	99.13	83.67	97.92	94.51

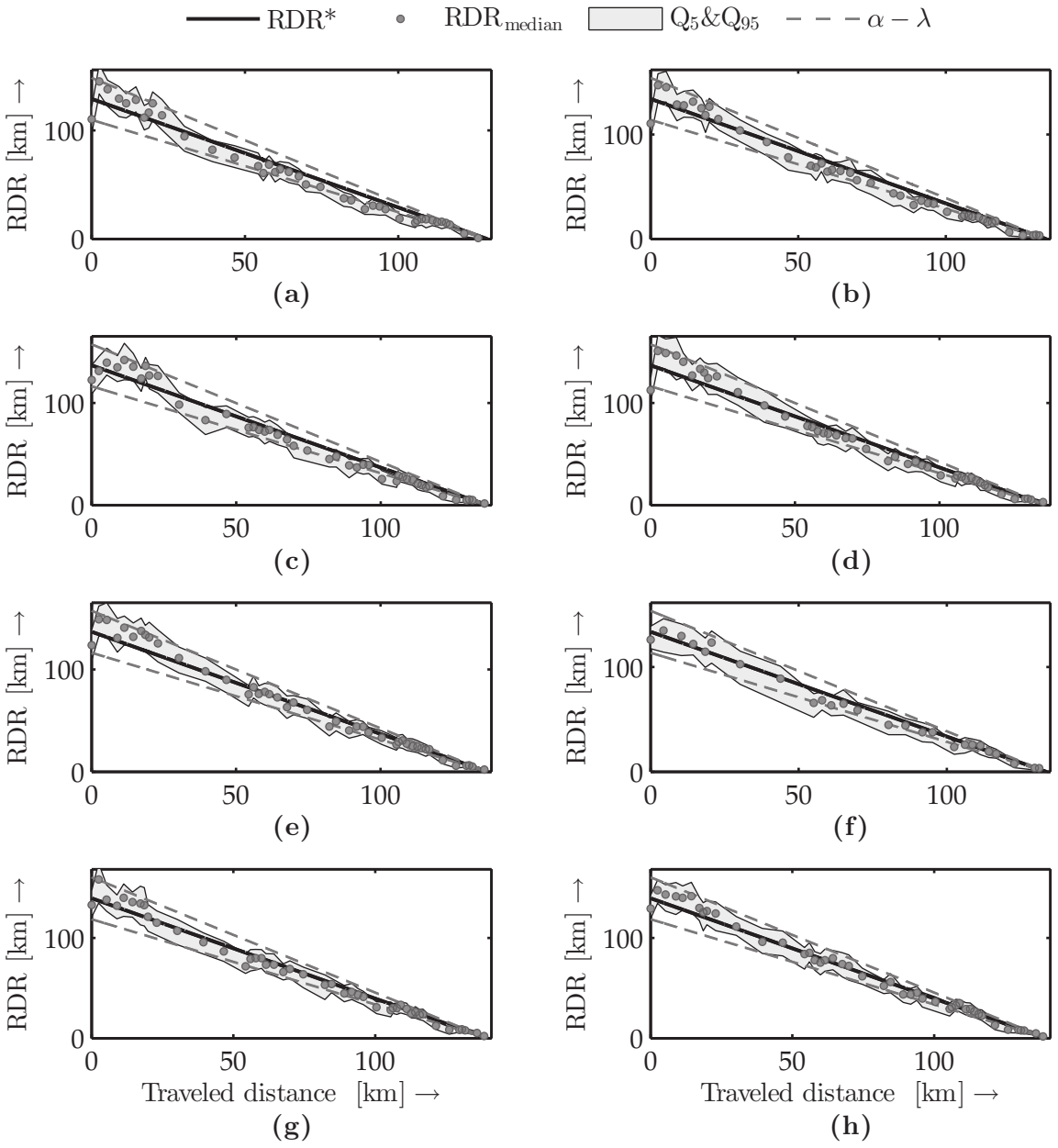

Figure C.6.: Further results of assessing future load uncertainty in the RDR estimation under temperatures of 0°C (**a**), 06°C (**b**), 12°C (**c**), 15°C (**d**), 23°C (**e**), 25°C (**f**), 33°C (**g**) and 38°C (**h**).

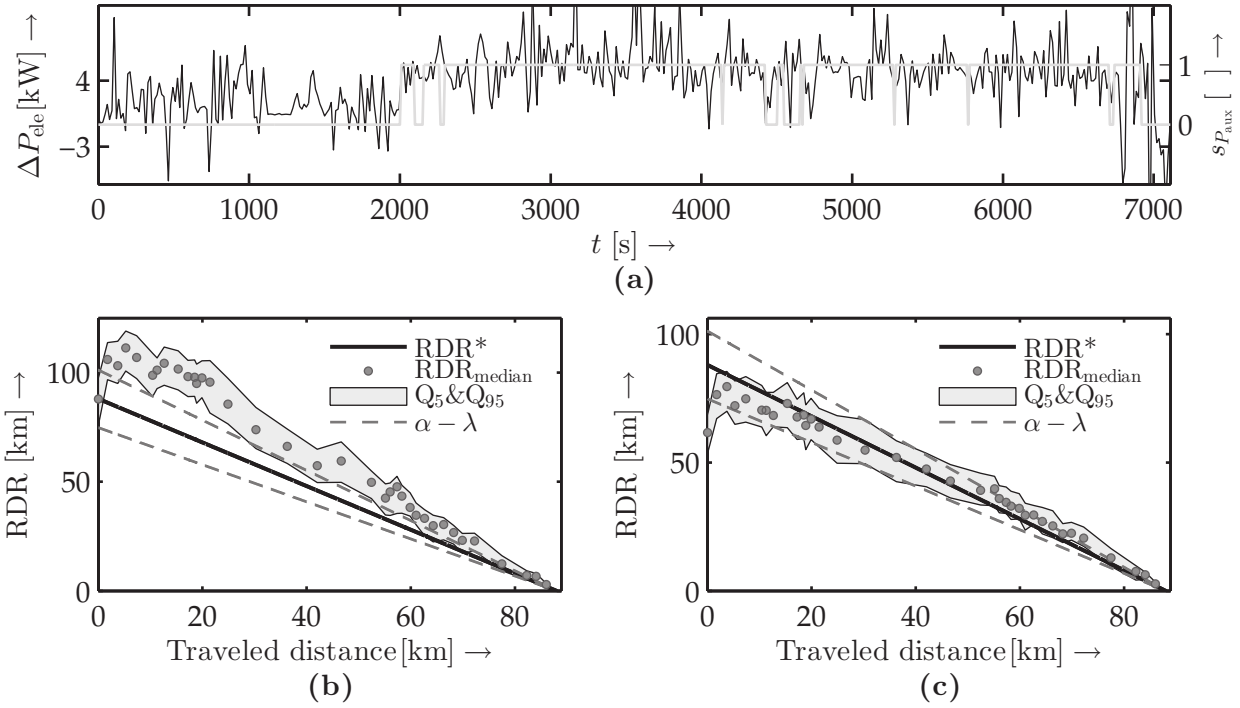

Figure C.7.: Difference between real and simulated power consumption (black line) and state of use of auxiliaries (gray line) (**a**). Experimental results by neglecting (**b**) and by accounting for (**c**) the influence of external loads on the RDR estimation.

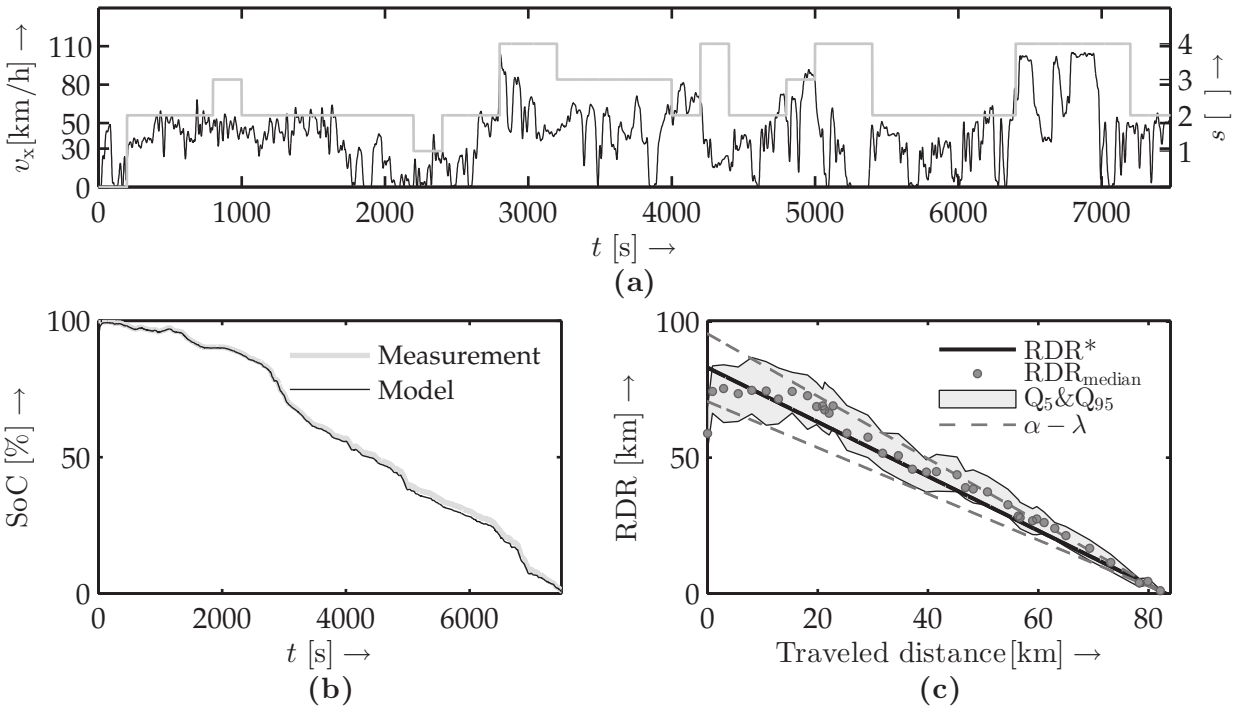

Figure C.8.: Speed (black line) and driving condition classification (gray line) (**a**) of an arbitrary drive profile. Results of the SoC (**b**) and RDR (**c**) estimation of experiments performed in the cycler by a temperature of 12°C with a residual power consumption of 4000 W activated after 2000 s.

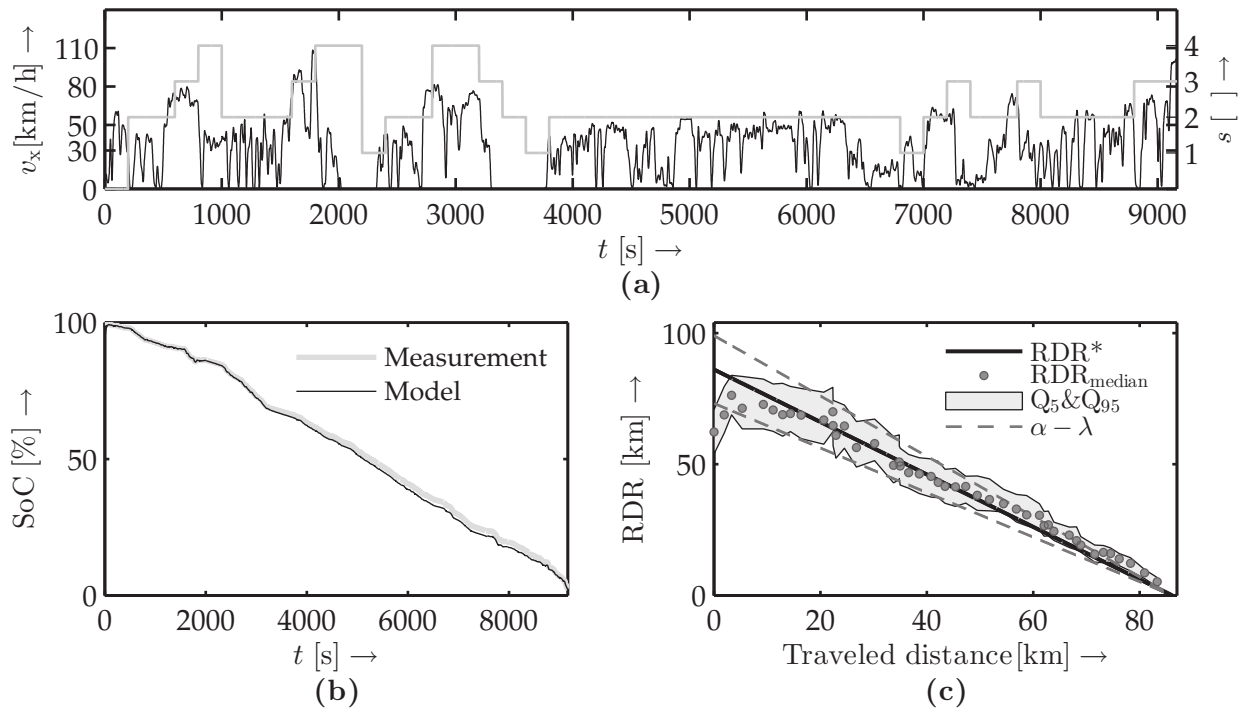

Figure C.9.: Speed (black line) and driving condition classification (gray line) (**a**) of an arbitrary drive profile. Results of the SoC (**b**) and RDR (**c**) estimation of experiments performed in the cycler by a temperature of 25°C with a residual power consumption of 4000 W activated after 2000 s.

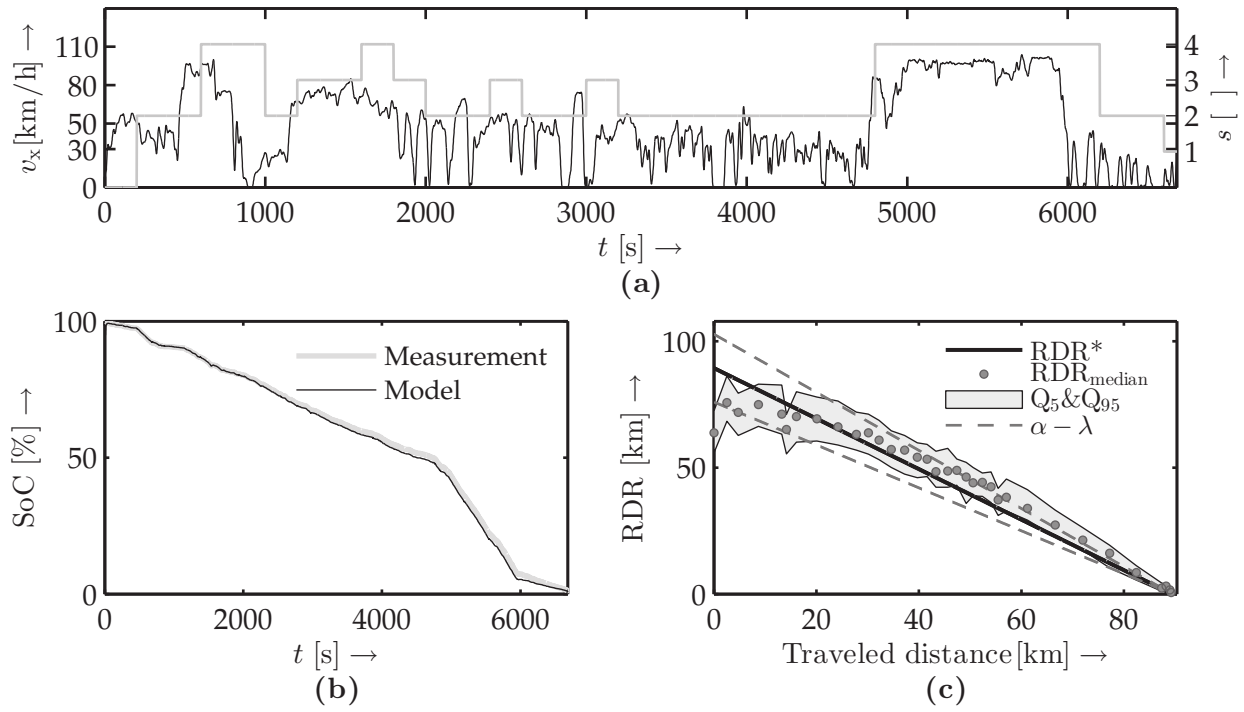

Figure C.10.: Speed (black line) and driving condition classification (gray line) (**a**) of an arbitrary drive profile. Results of the SoC (**b**) and RDR (**c**) estimation of experiments performed in the cycler by a temperature of 38°C with a residual power consumption of 4000 W activated after 2000 s.

Bibliography

References

Parts of the material presented in this work has been originally published in conferences and journals. These publications as well as the resources by other researchers are summarized in the following list:

VDI 2206 (2004): *Design methodology for mechatronic systems*.

AbouRizk, S., D. Halpin, and J. Wilson (1994): Fitting beta distributions based on sample data. In: *Journal of Construction Engineering and Management* Vol. 120, no. 2, pp. 288–305.

An, S., W. Liu, and S. Venkatesh (2007): Fast cross-validation algorithms for least squares support vector machine and kernel ridge regression. In: *Pattern Recognition* Vol. 40, no. 8, pp. 2154–2162.

Andre, M. (1996): Driving cycles development: characterization of the methods. In: *SAE International Technical Report*.

Andre, M. (2004): The ARTEMIS European driving cycles for measuring car pollutant emissions. In: *Science of The Total Environment* Vol. 334, no. 1, pp. 73–84.

Asamer, J., A. Graser, B. Heilmann, and M. Ruthmair (2016): Sensitivity analysis for energy demand estimation of electric vehicles. In: *Transportation Research Part D: Transport and Environment* Vol. 46, no. 1, pp. 182–199.

Bakker, E., L. Nyborg, and H. B. Pacejka (1987): Tyre modelling for use in vehicle dynamics studies. In: *SAE International Technical Report*.

Balaban, E., S. Narasimhan, M. Daigle, J. Celaya, I. Roychoudhury, B. Saha, S. Saha, and K. Goebel (2011): A mobile robot testbed for prognostics-enabled autonomous decision making. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 15–30.

Balaban, E., S. Narasimhan, M. Daigle, I. Roychoudhury, A. Sweet, and C. Bond (2013): Development of a mobile robot test platform and methods for validation of prognostics-enabled decision making algorithms. In: *International Journal of Prognostics and Health Management (IJPHM)* Vol. 4, no. 1.

Barlow, T. J., S. Latham, I. S. McCrae, and P. G. Boulter, eds. (2009): A reference book of driving cycles for use in the measurement of road vehicle emissions. Ed. by T. J. Barlow, S. Latham, I. S. McCrae, and P. G. Boulter. TRL.

Bedogni, L., L. Bononi, A. D'elia, M. Di Felice, M. Di Nicola, and T. S. Cinotti (2014): Driving without anxiety: A route planner service with range prediction for the electric vehicles. In: *International IEEE Conference on Connected Vehicles and Expo (ICCVE)*, pp. 199–206.

Benini, L., D. Bruni, A Mach, E. Macii, and M. Poncino (2003): Discharge current steering for battery lifetime optimization. In: *IEEE Transactions on Computers* Vol. 52, no. 8, pp. 985–995.

Bertuccelli, L. and J. How (2008): Estimation of non-stationary Markov Chain transition models. In: *International IEEE Conference on Decision and Control*, pp. 55–60.

Billio, M. and R. Casarin (2011): Beta autoregressive transition Markov-switching models for business cycle analysis. In: *Studies in Nonlinear Dynamics & Econometrics* Vol. 15, no. 4.

Bishop, C. M., ed. (2006): Pattern recognition and machine learning. Ed. by C. M. Bishop. Springer.

Bole, B., M. Daigle, and G. Gorospe (2014): Online prediction of battery discharge and estimation of parasitic loads for an electric aircraft. In: *European Conference of the Prognostics and Health Management Society*, pp. 23–32.

Bole, B., C. A. Teubert, Q. Cuong Chi, E. Hogge, S. Vazquez, K. Goebel, and V. George (2013): SIL/HIL replication of electric aircraft powertrain dynamics and inner-loop control for V&V of system health management routines. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 325–336.

Bolovinou, A., I. Bakas, A. Amditis, F. Mastrandrea, and W. Vinciotti (2014): Online prediction of an electric vehicle remaining range based on regression analysis. In: *International IEEE Electric Vehicle Conference (IEVC)*, pp. 1–8.

Campolongo, F., A. Saltelli, and J. Cariboni (2011): From screening to quantitative sensitivity analysis. An unified approach. In: *Computer Physics Communications* Vol. 182, no. 4, pp. 978–988.

Cawley, G. (2006): Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs. In: *International Joint Conference on Neural Networks (IJCNN)*, pp. 1661–1668.

Celaya, J. R., C. Kulkarni, G. Biswas, S. Saha, and K. Goebel (2011): A model-based prognostics methodology for electrolytic capacitors based on electrical overstress accelerated aging. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 25–29.

Ceraolo, M. and G. Pede (2001): Techniques for estimating the residual range of an electric vehicle. In: *IEEE Transactions on Vehicular Technology* Vol. 50, no. 1, pp. 109–115.

Conradi, P., P. Bouteiller, and S. Hanßen (2011): Dynamic cruising range prediction for electric vehicles. In: *Advanced Microsystems for Automotive Applications*, pp. 269–277.

Cortes, C. and V. Vapnik (1995): Support-vector networks. In: *Machine Learning* Vol. 20, no. 3, pp. 273–297.

Daigle, M. and K. Goebel (2010): Improving computational efficiency of prediciton in model-based prognostics using the Unscented transform. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 1–12.

Daigle, M., S. Sankararaman, and C. S. Kulkarni (2015): Stochastic prediction of remaining driving time and distance for a planetary rover. In: *International IEEE Aerospace Conference*, pp. 1–9.

Daigle, M., A. Saxena, and K. Goebel (2012): An efficient deterministic approach to model-based prediction uncertainty estimation. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 326–335.

Daigle, M. J. and S. Sankararaman (2013): Advanced methods for determining prediction uncertainty in model-based prognostics with application to planetary rovers. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 262–274.

Daigle, M. J. and S. Sankararaman (2016): Predicting remaining driving time and distance of a planetary rover under uncertainty. In: *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering* Vol. 2, no. 4.

Daum, F. and J. Huang (2008): Particle flow for nonlinear filters with log-homotopy. In: *Proceedings of SPIE - The International Society for Optical Engineering*.

Daum, F. and J. Huang (2011): Particle degeneracy: root cause and solution. In: *Proceedings of SPIE - The International Society for Optical Engineering*.

Daum, F. and J. Huang (2013): Particle flow with non-zero diffusion for nonlinear filters. In: *Proceedings of SPIE - The International Society for Optical Engineering*.

Denis, N., M. R. Dubois, K. A. Gil, T. Driant, and A. Desrochers (2012): Range prediction for a three-wheel plug-in hybrid electric vehicle. In: *International IEEE Transportation Electrification Conference and Expo (ITEC)*, pp. 1–6.

Doerffel, D. and S. A. Sharkh (2006): A critical review of using the Peukert equation for determining the remaining capacity of lead-acid and Lithium-ion batteries. In: *Journal of Power Sources* Vol. 155, no. 2, pp. 395–400.

Doucet, A., N. De Freitas, and N. Gordon (2001): An introduction to sequential Monte Carlo methods. In: *Sequential Monte Carlo Methods in Practice*. Springer, pp. 3–14.

Doucet, A., S. Godsill, and C. Andrieu (2000): On sequential Monte Carlo sampling methods for Bayesian filtering. In: *Statistics and Computing* Vol. 10, no. 3, pp. 197–208.

El Majdoub, K., F. Giri, H. Ouadi, L. Dugard, and F. Chaoui (2012): Vehicle longitudinal motion modeling for nonlinear control. In: *Control Engineering Practice* Vol. 20, no. 1, pp. 69–81.

Ferrari, S. and F. Cribari-Neto (2004): Beta regression for modelling rates and proportions. In: *Journal of Applied Statistics* Vol. 31, no. 7, pp. 799–815.

Ferreira, J. C., V. Monteiro, and J. L. Afonso (2011): Mobile geographic range prediction for electric vehicles. In: *Proceedings of the Conference on Electronics, Telecommunications And Computers (CETC)*, pp. 1–8.

Ferreira, J. C., V. Monteiro, and J. L. Afonso (2012): Data mining approach for range prediction of electric vehicle. In: *Conference on Future Automotive Technology-Focus Electromobility*, pp. 1–15.

Ferreira, J. C., V. Monteiro, and J. L. Afonso (2013): Dynamic range prediction for an electric vehicle. In: *IEEE Electric Vehicle Symposium and Exhibition (EVS27)*, pp. 1–11.

Franke, T., I. Neumann, F. Bühler, P. Cocron, and J. F. Krems (2012): Experiencing range in an electric vehicle: Understanding psychological barriers. In: *Applied Psychology* Vol. 61, no. 3, pp. 368–391.

García, V., J. Sánchez, R. Mollineda, and R. A. J. Sotoca (2007): The class imbalance problem in pattern classification and learning. In: *II Congreso Español de Informática (CEDI)*, pp. 283–291.

Genta, G. and L. Morello, eds. (2009): The automotive chassis - Volume 1: Components design. Ed. by G. Genta and L. Morello. Springer.

George, A. and W. Powell (2006): Adaptive stepsizes for recursive estimation with applications in approximate dynamic programming. In: *Journal of Machine Learning* Vol. 65, no. 1, pp. 167–198.

Gordon, N. J., D. J. Salmond, and A. F. Smith (1993): Novel approach to nonlinear and non-Gaussian Bayesian state estimation. In: *IEE Proceedings F - Radar and Signal Processing*. Vol. 140. 2, pp. 107–113.

Grewal, K. and P. Darnell (2013): Model-based EV range prediction for electric hybrid vehicles. In: *IET Hybrid and Electric Vehicles Conference 2013 (HEVC)*, pp. 1–6.

Grubwinkler, S., T. Brunner, and M. Lienkamp (2014): Range prediction for EVs via crowd-sourcing. In: *International IEEE Vehicle Power and Propulsion Conference (VPPC)*, pp. 1–6.

Guo, J., Z. Li, and M. Pecht (2015): A Bayesian approach for Li-Ion battery capacity fade modeling and cycles to failure prognostics. In: *Journal of Power Sources* Vol. 281, pp. 173–184.

Guyon, I. and A. Elisseeff (2003): An introduction to variable and feature selection. In: *Journal of Machine Learning Research* Vol. 3, pp. 1157–1182.

Guzzella, L. and A. Sciarretta, eds. (2005): Vehicle Propulsion Systems: Introduction to Modeling and Optimization. Ed. by L. Guzzella and A. Sciarretta. Springer Verlag, Heildelberg.

Heißing, B. and M. Ersoy, eds. (2011): Chassis Handbook: Fundamentals, Driving Dynamics, Components, Mechatronics, Perspectives. Ed. by B. Heißing and M. Ersoy. Vieweg+Teubner Verlag, 1st edition.

Hirt, C., S. Claessens, T. Fecher, M. Kuhn, R. Pail, and M. Rexer (2013): New ultrahigh-resolution picture of Earth's gravity field. In: *Geophysical Research Letters* Vol. 40, no. 16, pp. 4279–4283.

Hong, J., S. Park, and N. Chang (2016): Accurate remaining range estimation for electric vehicles. In: 21st *Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp. 781–786.

Hsu, C.-W. and C.-J. Lin (2002): A comparison of methods for multi-class support vector machines. In: *IEEE Transactions on Neural Networks* Vol. 13, no. 2, pp. 415–425.

Hu, Y, S Yurkovich, Y Guezennec, and B. Yurkovich (2009): A technique for dynamic battery model identification in automotive applications using linear parameter varying structures. In: *Control Engineering Practice* Vol. 17, no. 10, pp. 1190–1201.

Huang, X., Y. Tan, and X. He (2011): An intelligent multifeature statistical approach for the discrimination of driving conditions of a hybrid electric vehicle. In: *IEEE Transactions on Intelligent Transportation Systems* Vol. 12, no. 2, pp. 453–465.

Isermann, R. and M. Münchhof (2011): Identification of dynamic systems: an introduction with applications. Springer.

Isukapalli, S. S., A Roy, and P. G. Georgopoulos (1998): Stochastic response surface methods (SRSMs) for uncertainty propagation: application to environmental and biological systems. In: *Risk analysis* Vol. 18, no. 3, pp. 351–363.

Johannesson, L. (2005): Development of a time invariant stochastic model of a transport mission. In: *Technical Report of the Chalmers University of Technology*.

Jongerden, M. R. and B. R. Haverkort (2008): Battery modeling. In: *Technical Report TR-CTIT-08-01, Centre for Telematics and Information Technology, University of Twente*.

Jongerden, M. and B. Haverkort (2009): Which battery model to use? In: *IET Software* Vol. 3, no. 6, pp. 445–457.

Julier, S. and J. Uhlmann (2004): Unscented filtering and nonlinear estimation. In: *Proceedings of the IEEE* Vol. 92, no. 3, pp. 401–422.

Jung, M. F., D. Sirkin, T. M. Gür, and M. Steinert (2015): Displayed uncertainty improves driving experience and behavior: The case of range anxiety in an electric car. In: *Proceedings of the* 33rd *Annual ACM Conference on Human Factors in Computing Systems*, pp. 2201–2210.

Kalman, R. E. (1960): A new approach to linear filtering and prediction problems. In: *Journal of Fluids Engineering* Vol. 82, no. 1, pp. 35–45.

Kang, J. M., S. S. Seo, and J. W. K. Hong (2011): Personalized battery lifetime prediction for mobile devices based on usage patterns. In: *Journal of Computing Science and Engineering* Vol. 5, no. 4, pp. 338–345.

Kimotho, J. K., T. Meyer, and W. Sextro (2014): PEM fuel cell prognostics using particle filter with model parameter adaptation. In: *International IEEE Conference on Prognostics and Health Management*, pp. 1–6.

Kimotho, J. K. and W. Sextro (2015): Comparison and ensemble of temperature-based and vibration-based methods for machinery prognostics. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 1–11.

Kiureghian, A. D. (2004): Engineering Design Reliability Handbook. Ed. by S. Singhal, D. M. Ghiocel, and E. Nikolaidis. CRC Press. Chap. 14: First- and second-order reliability methods.

Kotsiantis, S. B. (2007): Supervised machine learning: a review of classification techniques. In: *Proceedings of the Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies* Vol. 31, no. 3.

Kuh, A. (2004): Least squares kernel methods and applications. In: *Soft Computing in Communications* Vol. 136, no. 1, pp. 365–387.

Lachat, D., A. Krebs, T. Thueer, and R. Siegwart (2006): Antarctica rover design and optimization for limited power consumption. In: *IFAC Proceedings Volumes* Vol. 39, no. 16, pp. 788–793.

Lam, L., P. Bauer, and E. Kelder (2011): A practical circuit-based model for Li-ion battery cells in electric vehicle applications. In: *International IEEE Telecommunications Energy Conference (INTELEC)*, pp. 1–9.

Lee, C. H. and C. H. Wu (2015): A novel big data modeling method for improving driving range estimation of EVs. In: *IEEE Access* Vol. 3, no. 1, pp. 1980–1993.

Lee, T. C., G. G. Judge, and A. Zellner, eds. (1970): Estimating the parameters of the Markov probability model from aggregate time series data. Ed. by T. C. Lee, G. G. Judge, and A. Zellner. North-Holland, 2nd edition.

Lee, T.-K., B. Adornato, and Z. S. Filipi (2011): Synthesis of real-world driving cycles and their use for estimating PHEV energy consumption and charging opportunities:

Case study for Midwest/U.S. In: *IEEE Transactions on Vehicular Technology* Vol. 60, no. 9, pp. 4153–4163.

LeSage, J. R. (2013): "Energy storage-aware prediction/control for mobile systems with unstructured loads". PhD thesis. The University of Texas at Austin.

Liao, L. and F. Kottig (2014): Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. In: *IEEE Transactions on Reliability* Vol. 63, no. 1, pp. 191–207.

Linden, D. and T. Reddy, eds. (2002): Handbook of batteries, 3rd edition. Ed. by D. Linden and T. Reddy. McGraw-Hill.

Liu, G., M. Ouyang, L. Lu, and J. Li (2015): Remaining driving range estimation for electric vehicles based on an advanced battery residual energy model. In: *IEEE Electric Vehicle Symposium and Exhibition (EVS28)*, pp. 1–5.

Lu, J., S. Liu, Q. Wu, and Q. Qiu (2010): Accurate modeling and prediction of energy availability in energy harvesting real-time embedded systems. In: *International IEEE Green Computing Conference*, pp. 469–476.

Luts, J., F. Ojeda, R. Van de Plas, B. De Moor, S. Van Huffel, and J. A. Suykens (2010): A tutorial on support vector machine-based methods for classification problems in chemometrics. In: *Analytica Chimica Acta* Vol. 665, no. 2, pp. 129–145.

Manwell, J. F. and J. G. McGowan (1994): Extension of the kinetic battery model for wind-hybrid power systems. In: *Proceedings of the* 5th *European Wind Energy Association Conference*, pp. 284–289.

Matsushima, T., S. Ishizuka, and M. Hashiwaki (1990): Prediction of remaining battery discharge time in telecommunications systems. In: *International IEEE Telecommunications Energy Conference (INTELEC)*, pp. 543–550.

McKay, M. D., R. J. Beckman, and W. J. Conover (2000): A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. In: *Technometrics* Vol. 42, no. 1, pp. 55–61.

Montazeri-Gh, M., A. Fotouhi, and A. Naderpour (2011): Driving patterns clustering based on driving features analysis. In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* Vol. 225, no. 6, pp. 1301–1317.

Mutoh, N., Y. Hayano, H. Yahagi, and K. Takita (2007): Electric braking control methods for electric vehicles with independently driven front and rear wheels. In: *IEEE Transactions on Industrial Electronics* Vol. 54, no. 2, pp. 1168–1176.

Nuzzo, R. (2014): Statistical errors. In: *Nature* Vol. 506, no. 7487, pp. 150–152.

Ojeda, F., J. A. Suykens, and B. De Moor (2008): Low rank updated LS-SVM classifiers for fast variable selection. In: *Neural Networks* Vol. 21, no. 2, pp. 437–449.

Oliva, J. A. and T. Bertram (2014a): Adaptive driving situation characterization for predicting the driving load of electric vehicles in uncertain environments. In: *European Conference of the Prognostics and Health Management Society*, pp. 2–14.

Oliva, J. A. and T. Bertram (2014b): On the use of particle flow to enhance the computational performance of particle-filtering-based prognostics. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 309–319.

Oliva, J. A. and T. Bertram (2015): Analytische Unsicherheitsanalyse von der Schätzung der verbleibenden Entladezeit elektrischer Energiespeichersysteme. In: *VDI/VDE Fachtagung Mechatronik*, pp. 43–48.

Oliva, J. A., C. Weihrauch, and T. Bertram (2013a): A model-based approach for predicting the remaining driving range in electric vehicles. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 438–448.

Oliva, J. A., C. Weihrauch, and T. Bertram (2013b): Model-based remaining driving range prediction in electric vehicles by using particle filtering and Markov chains. In: *IEEE Electric Vehicle Symposium and Exhibition (EVS27)*, pp. 1–10.

Olivares, B. E., M. A. Cerda Munoz, M. E. Orchard, and J. F. Silva (2013): Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena. In: *IEEE Transactions on Instrumentation and Measurement* Vol. 62, no. 2, pp. 364–376.

Ondruska, P. and I. Posner (2014a): Probabilistic attainability maps: Efficiently predicting driver-specific electric vehicle range. In: *IEEE Proceedings of the Intelligent Vehicles Symposium*, pp. 1169–1174.

Ondruska, P. and I. Posner (2014b): The route not taken: Driver-centric estimation of electric vehicle range. In: 24th *International Conference on Automated Planning and Scheduling*, pp. 413–420.

Orchard, M. E., M. A. Cerda, B. E. Olivares, and J. F. Silva (2012): Sequential Monte Carlo methods for discharge time prognosis in lithium-ion batteries. In: *International Journal of Prognostics and Health Management (IJPHM)* Vol. 3, no. 2, pp. 1–12.

Orchard, M. and G. Vachtsevanos (2009): A particle-filtering approach for on-line fault diagnosis and failure prognosis. In: *Transactions of the Institute of Measurement and Control*. Vol. 31, pp. 221–246.

Pacejka, H. B., ed. (2005): Tire and Vehicle Dynamics, 2nd edition. Ed. by H. B. Pacejka. SAE International.

Pattipati, B, B Balasingam, G. Avvari, K. Pattipati, and Y Bar-Shalom (2014): Open circuit voltage characterization of lithium-ion batteries. In: *Journal of Power Sources* Vol. 269, pp. 317–333.

Pérez, A., P. Larrañaga, and I. Inza (2009): Bayesian classifiers based on kernel density estimation: Flexible classifiers. In: *International Journal of Approximate Reasoning* Vol. 50, no. 2, pp. 341–362.

Pesco, A. M., R. V. Biagetti, R. S. Chidamber, and C. R. Venkatram (1989): An adaptive battery reserve time prediction algorithm. In: *International IEEE Telecommunications Energy Conference (INTELEC)*, pp. 6.1/1–6.1/7.

Peukert, W (1897): Über die Abhänigkeit der Kapazität von der Entladestromstärke bei Bleiakkumulatoren. In: *Elektrotechnische Zeitschrift (ETZ)* Vol. 18, pp. 287–288.

Piller, S., M. Perrin, and A. Jossen (2001): Methods for state-of-charge determination and their applications. In: *Journal of Power Sources* Vol. 96, no. 1, pp. 113–120.

Platt, J. C., J. Shawe-Taylor, and N. Cristianini (1999): Large margin DAG's for multiclass classification. In: *Advances in Neural Information Processing Systems* Vol. 12, no. 3, pp. 547–553.

Plett, G (2005): Results of temperature-dependent LiPB cell modeling for HEV SoC estimation. In: *Proceedings of the* 21st *Electric Vehicle Symposium (EVS21)*, pp. 1–9.

Plett, G. L. (2004): Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 2. Modeling and identification. In: *Journal of Power Sources* Vol. 134, no. 2, pp. 262–276.

Plett, G. L. (2006): Sigma-point Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 2: Simultaneous state and parameter estimation. In: *Journal of Power Sources* Vol. 161, no. 2, pp. 1369–1384.

Pop, V, H. J. Bergveld, P. H. L. Notten, J. H. G. Veld, and P. P. L. Regtien (2009): Accuracy analysis of the State-of-Charge and remaining run-time determination for Lithium-ion batteries. In: *Measurement* Vol. 42, no. 8, pp. 1131–1138.

Quach, C. C., B Bole, E Hogge, S Vazquez, M Daigle, J Celaya, A Weber, and K Goebel (2013): Battery charge depletion prediction on an electric aircraft. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 503–512.

Rackwitz, R. and B. Flessler (1978): Structural reliability under combined random load sequences. In: *Computers & Structures* Vol. 9, no. 5, pp. 489–494.

Rahimi-Eichi, H., P. B. Jeon, M. Y. Chow, and T. J. Yeo (2015): Incorporating big data analysis in speed profile classification for range estimation. In: *International IEEE Conference on Industrial Informatics (INDIN)*, pp. 1290–1295.

Rahimi-Eichi, H. and M.-Y. Chow (2014): Big-data framework for electric vehicle range estimation. In: *Annual IEEE Conference of the Industrial Electronics Society (IECON)*, pp. 5628–5634.

Rao, R., S. Vrudhula, and D. N. Rakhmatov (2003): Battery modeling for energy aware system design. In: *Computer* Vol. 36, no. 12, pp. 77–87.

Rauh, N., T. Franke, and J. F. Krems (2014): Understanding the impact of electric vehicle driving experience on range anxiety. In: *Journal of the Human Factors and Ergonomics Society* Vol. 57, no. 1, pp. 177–187.

Reich, S. (2011): A dynamical systems framework for intermittent data assimilation. In: *BIT Numerical Mathematics* Vol. 51, no. 1, pp. 235–249.

Rifkin, R. and A. Klautau (2004): In defense of one-vs-all classification. In: *Journal of Machine Learning Research* Vol. 5, pp. 101–141.

Ross, W. and P. Budney (1995): Development of a battery runtime prediction algorithm and a method for determining its accuracy. In: *International IEEE Telecommunications Energy Conference (INTELEC)*, pp. 277–283.

Sadrpour, A., J. J. Jin, and A. G. Ulsoy (2013): Mission energy prediction for unmanned ground vehicles using real-time measurements and prior knowledge. In: *Journal of Field Robotics* Vol. 30, no. 3, pp. 399–414.

Saha, B. and K. Goebel (2009): Modeling Li-ion battery capacity depletion in a particle filtering framework. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 2909–2924.

Saha, B., E. Koshimoto, C. C. Quach, E. F. Hogge, T. H. Strom, B. L. Hill, S. L. Vazquez, and K. Goebel (2011): Battery health management system for electric UAVs. In: *International IEEE Aerospace Conference*, pp. 1–9.

Sankararaman, S. and K. Goebel (2013): Why is the remaining useful life prediction uncertain? In: *Annual Conference of the Prognostics and Health Management Society*, pp. 337–349.

Sankararaman, S. and K. Goebel (2015): Uncertainty in prognostics and systems health management. In: *International Journal of Prognostics and Health Management (IJPHM)* Vol. 6, no. 10, pp. 1–14.

Sankararaman, S., M. Daigle, A. Saxena, and K. Goebel (2013): Analytical algorithms to quantify the uncertainty in remaining useful life prediction. In: *International IEEE Aerospace Conference*, pp. 1–11.

Sato, Y., S. Ishikawa, T. Okubo, M. Abe, and K. Tamai (2011): Development of high response motor and inverter system for the Nissan LEAF electric vehicle. In: *SAE International Technical Report*.

Saxena, A., J. Celaya, B. Saha, S. Saha, and K. Goebel (2009): On applying the prognostics performance metrics. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 1–16.

Saxena, A., J. R. Celaya, I. Roychoudhury, S. Saha, B. Saha, and K. Goebel (2012): Designing data-driven battery prognostic approaches for variable loading profiles: Some lessons learned. In: *European Conference of Prognostics and Health Management Society*, pp. 72–732.

Seaman, A., T.-S. Dao, and J. McPhee (2014): A survey of mathematics based equivalent circuit and electrochemical battery models for hybrid and electric vehicle simulation. In: *Journal of Power Sources* Vol. 256, no. 1, pp. 410–423.

Sokolova, M. and G. Lapalme (2009): A systematic analysis of performance measures for classification tasks. In: *Information Processing & Management* Vol. 45, no. 4, pp. 427–437.

Stefanou, G. (2009): The stochastic finite element method: Past, present and future. In: *Computer Methods in Applied Mechanics and Engineering* Vol. 198, no. 9, pp. 1031–1051.

Stone, M. (1974): Cross-validatory choice and assessment of statistical predictions. In: *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 111–147.

Strelioff, C. C., J. P. Crutchfield, and A. W. Hübler (2007): Inferring Markov chains: Bayesian estimation, model comparison, entropy rate, and out-of-class modeling. In: *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* Vol. 76, no. 1, pp. 1–14.

Suykens, J. a. K. and J. Vandewalle (1999): Least squares support vector machine classifiers. In: *Neural Processing Letters* Vol. 9, no. 3, pp. 293–300.

Suykens, J. A. K. (2002): Least squares support vector machines. World Scientific.

Taesic, K. and Q. Wei (2011): A hybrid battery model capable of capturing dynamic circuit characteristics and nonlinear capacity effects. In: *IEEE Transactions on Energy Conversion* Vol. 26, no. 4, pp. 1172–1180.

Tampier, C., A. Pérez, F. Jaramillo, V. Quintero, M. E. Orchard, and J. F. Silva (2015): Lithium-ion battery end-of-discharge time estimation and prognosis based on Bayesian algorithms and outer feedback correction loops: A comparative analysis. In: *Annual Conference of the Prognostics and Health Management Society*, pp. 1–14.

Tannahill, V. R., K. M. Muttaqi, and D. Sutanto (2016): Driver alerting system using range estimation of electric vehicles in real time under dynamically varying environmental conditions. In: *IET Electrical Systems in Transportation* Vol. 6, no. 2, pp. 107–116.

Teodorescu, I. (2009): Maximum likelihood estimation for Markov chains. In: *arXiv*, pp. 1–14.

Theodoridis, S. and K. Koutroumbas, eds. (2009): Pattern recognition. Ed. by S. Theodoridis and K. Koutroumbas. Academic Press.

Thomas, R. and J. Mezies (2006): Tires and passenger vehicle fuel economy. In: *Transportation Research Board Special Report 286*.

Tokdar, S. T. and R. E. Kass (2010): Importance sampling: a review. In: *Wiley Interdisciplinary Reviews: Computational Statistics* Vol. 2, no. 1, pp. 54–60.

Ton, M., C. Calwell, and T. Reeder (2003): Low rolling resistance tires. In: *Choose Green Report*, pp. 1–6.

Tsiligkaridis, T. and A. Hero (2013): Covariance estimation in high dimensions via Kronecker product expansions. In: *IEEE Transactions on Signal Processing* Vol. 61, no. 21, pp. 5347–5360.

Valentina, R., A. Viehl, O. Bringmann, and W. Rosenstiel (2014): HVAC system modeling for range prediction of electric vehicles. In: *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 1145–1150.

Van Der Merwe, R. (2004): "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models". PhD thesis. OGI School of Science & Engineering, Oregon Health & Science University.

Vaz, W., A. K. Nandi, R. G. Landers, and U. O. Koylu (2015): Electric vehicle range prediction for constant speed trip using multi-objective optimization. In: *Journal of Power Sources* Vol. 275, pp. 435–446.

Waldmann, K.-H. and U. M. Stocker, eds. (2013): Stochastische Modelle: eine anwendungsorientierte Einführung. Ed. by K.-H. Waldmann and U. M. Stocker. Springer.

Weihrauch, C., J. Oliva, and T. Bertram (2013): Modellierung einer Li-Ionen Batterie für die Anwendung in Fahrerassistenzsystemen für Elektrofahrzeuge. In: *VDI-Fachtagung AUTOREG, Steuerung und Regelung von Fahrzeugen und Motoren, VDI-Berichte 2196*, pp. 81–91.

Wen, Y., R. Wolski, and C. Krintz (2003): Online prediction of battery lifetime for embedded and mobile devices. In: *Power-Aware Computer Systems* Vol. 3164, no. 1, pp. 57–72.

Wong, J. Y., ed. (2001): Theory of Ground Vehicles. Ed. by J. Y. Wong. John Wiley & Sons, Inc., 3rd edition.

Xiang, Y. and Y. Liu (2010): Inverse FORM method for probabilistic fatigue prognosis. In: 51st *AIAA/ASME/ASCE/AHS/ASC Conference on Structures, Structural Dynamics and Materials*, pp. 1–10.

Zhang, Y., W. Wang, Y. Kobayashi, and K. Shirai (2012): Remaining driving range estimation of electric vehicle. In: *IEEE International Electric Vehicle Conference (IEVC)*, pp. 1–7.

Zhao, Y.-G. and T. Ono (1999): A general procedure for first/second-order reliability method (FORM/SORM). In: *Structural safety* Vol. 21, no. 2, pp. 95–112.

Zhou, Z., Y. Huang, Y. Lu, Z. Shi, X. Li, J. Wu, and H. Li (2014): Lithium-ion battery end-of-discharge time prediction using particle filtering algorithm. In: *International Conference on Prognostics and System Health Management*, pp. 658–663.