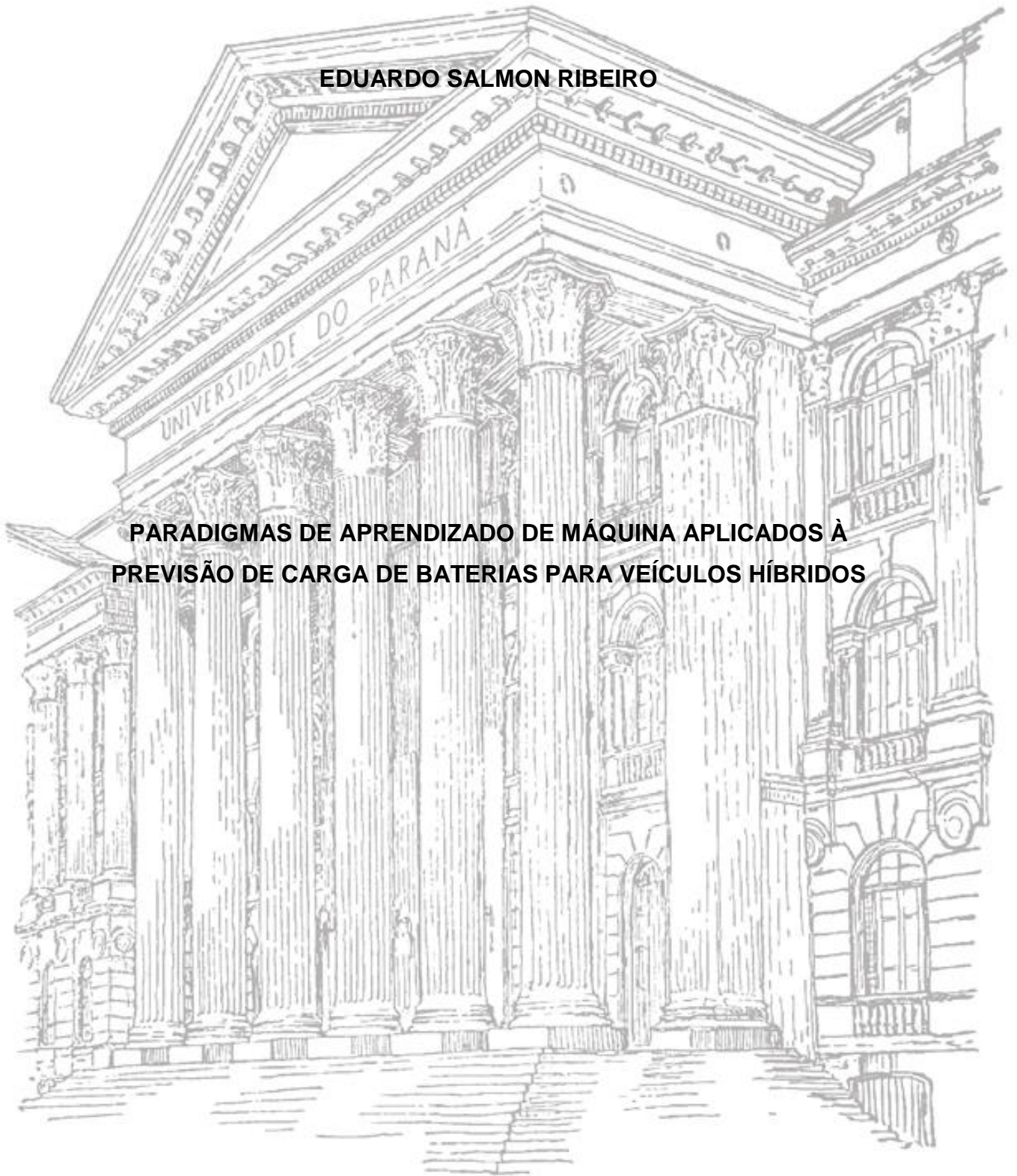


**UNIVERSIDADE FEDERAL DO PARANÁ  
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**EDUARDO SALMON RIBEIRO**

**PARADIGMAS DE APRENDIZADO DE MÁQUINA APLICADOS À  
PREVISÃO DE CARGA DE BATERIAS PARA VEÍCULOS HÍBRIDOS**



**CURITIBA**

**2017**

---

R484p

Ribeiro, Eduardo Salmon

Paradigmas de aprendizado de máquina aplicados à previsão de carga de baterias para veículos híbridos / Eduardo Salmon Ribeiro. – Curitiba, 2017.

90 f. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, 2017.

Orientador: Leandro dos Santos Coelho.

1. Aprendizado de máquina. 2. Comitê de máquinas. 3. Veículos híbridos. I. Universidade Federal do Paraná. II. Coelho, Leandro dos Santos. III. Dagostin, João Luiz Andreotti. IV. Título.

CDD: 621.8

---

**EDUARDO SALMON RIBEIRO**

**PARADIGMAS DE APRENDIZADO DE MÁQUINA APLICADOS À  
PREVISÃO DE CARGA DE BATERIAS PARA VEÍCULOS HÍBRIDOS**

Projeto de Dissertação apresentado como requisito para a obtenção do grau de Mestre, no Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Paraná.

Orientador: Prof. Leandro dos Santos Coelho

**CURITIBA**

**2017**



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
Setor TECNOLOGIA  
Programa de Pós Graduação em ENGENHARIA ELÉTRICA  
Código CAPES: 40001016043P4

### TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA ELÉTRICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **EDUARDO SALMON RIBEIRO**, intitulada: "**Paradigmas de aprendizado de máquina aplicados à previsão de carga de baterias para veículos híbridos**", após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 30 de Agosto de 2017.

LEANDRO DOS SANTOS COELHO  
Presidente da Banca Examinadora (UFPR)

ELIZETE MARIA LOURENÇO  
Avaliador Interno (UFPR)

GIDEON VILLAR LEANDRO  
Avaliador Interno (UFPR)

EMERSON DONAISKY  
Avaliador Externo (PUC/PR)



## **AGRADECIMENTOS**

Agradeço ao professor Leandro dos Santos Coelho, que além de despertar meu interesse para tal tema sempre esteve disposto e ajudar no desenvolvimento mesmo nos momentos mais difíceis, e ao meu filho Henrique que nasceu ao longo do desenvolvimento da dissertação de mestrado.

“Aprenda com o ontem, viva o hoje  
e acredite no amanhã”

Albert Einstein

## RESUMO

Uma das alternativas para a diminuição da poluição, seja esta sonora ou por gases e particulados, são os veículos híbridos, que combinam a autonomia do motor à combustão com os benefícios dos motores elétricos. Entretanto, por estes veículos contarem com baterias, o gerenciamento de energia destes veículos se torna determinante para uma adequada autonomia. Por outro lado, aprendizado de máquina é uma abordagem que trata do projeto e desenvolvimento de algoritmos que melhoram automaticamente com a experiência inspirado no comportamento de aprendizado de humanos. Tal comportamento pode ser obtido por meio do ajuste de parâmetros com base nos dados de entrada que são apresentados e, em alguns casos, nas informações de saída desejadas. Um dos focos da pesquisa em Aprendizado de Máquina é aprender automaticamente a reconhecer padrões complexos e tomar decisões com base em dados. Esta dissertação tem como objetivo principal a modelagem do estado de carga das baterias, usando diferentes técnicas de Aprendizado de Máquina para tarefa de identificação de sistemas, tais como máquinas de vetores de suporte, redes neuro-nebulosas (*neuro-fuzzy*), *bagging*, *boosting*, *subspace* e redes com estado de eco. Para avaliar o desempenho dos modelos matemáticos de Aprendizado de Máquina foram adotados dois índices de desempenho: (i) o erro quadrático médio (*Mean Squared Error*, MSE) e (ii) o coeficiente de determinação ( $R^2$ ). Pelos resultados obtidos, observou-se que alguns métodos de Aprendizado de Máquina apresentaram uma aproximação de boa qualidade quando comparada a saída real das cargas das baterias, isto evidenciado pelos valores de MSE e  $R^2$ .

**Palavras-chave:** Aprendizado de máquina; Identificação de sistemas; Máquinas de vetores de suporte; Comitê de máquinas; Previsão de carga de baterias; Veículos híbridos.

## ABSTRACT

One of the alternatives for the pollution reduction, the hybrid vehicles combines the autonomy of the engine to the combustion with the benefits of the electric motors. However, because these vehicles have batteries, the energy management of these vehicles becomes decisive for an adequate autonomy. On the other hand, machine learning is an approach that deals with the design and development of algorithms that automatically improve with the experience inspired by human learning behavior. Such behavior can be achieved by adjusting parameters based on the input data that is presented and in some cases on the desired output information. One of the focuses of Machine Learning research is to automatically learn to recognize complex patterns and make decisions based on data. This work has as main objective the modeling of the batteries state of charge using different Machine Learning techniques applied to system identification task, such as support vector machines, neuro-fuzzy networks, bagging, boosting, subspace and echo state networks. In order to evaluate the performance of Machine Learning mathematical models, two performance indices were adopted: (i) the mean squared error (MSE) and (ii) the coefficient of determination ( $R^2$ ). From the results obtained, it was observed that some methods of Machine Learning presented a good quality approximation when compared to the actual output of the loads of the batteries, evidenced by the MSE and  $R^2$  values.

**Keywords:** Machine learning; System identification; Support vector machine; Ensemble learning; Battery charge forecasting; Hybrid vehicles.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Primeiro veículo híbrido .....	21
Figura 2 – Motor de partida/alternador.....	23
Figura 3 – As diferentes topologias.....	24
Figura 4 – Estação de carregamento – tipo pantográfica .....	26
Figura 5 – Estação de carregamento – tipo conector .....	27
Figura 6 – Autoestrada eletrificada .....	28
Figura 7 - Esquemático do programa.....	36
Figura 8 – Modelo de McCulloch & Pitts .....	40
Figura 9 – Estrutura de uma ESN .....	44
Figura 10 - Conjunto de treinamento e diferentes hipóteses .....	45
Figura 11 - Hiperplano ótimo de separação de classes .....	47
Figura 12 – Dados de entrada e saída adotados no procedimento de identificação de sistemas.....	53
Figura 13 – Resultados de previsão obtidos por BAGGING01 .....	56
Figura 14 – Resultados de previsão obtidos por BAGGING02. ....	57
Figura 15 – Resultados de previsão obtidos por BAGGING03. ....	57
Figura 16 – Resultados de previsão obtidos por BAGGING04. ....	58
Figura 17 – Resultados de previsão obtidos por BOOSTING01.....	59
Figura 18 – Resultados de previsão obtidos por BOOSTING02.....	59
Figura 19 – Resultados de previsão obtidos por BOOSTING03.....	60
Figura 20 – Resultados de previsão obtidos por BOOSTING04.....	60
Figura 21 – Resultados de previsão obtidos por LS-SVM01. ....	61
Figura 22 – Resultados de previsão obtidos por LS-SVM02. ....	62
Figura 23 – Resultados de previsão obtidos por LS-SVM03. ....	62
Figura 24 – Resultados de previsão obtidos por LS-SVM04. ....	63

Figura 25 – Resultados de previsão obtidos por ANFIS01. ....	64
Figura 26 – Resultados de previsão obtidos por ANFIS02. ....	64
Figura 27 – Resultados de previsão obtidos por ANFIS03. ....	65
Figura 28 – Resultados de previsão obtidos por ANFIS04. ....	65
Figura 29 – Resultados de previsão obtidos por ESN01. ....	66
Figura 30 – Resultados de previsão obtidos por ESN02. ....	67
Figura 31 – Resultados de previsão obtidos por ESN03. ....	67
Figura 32 – Resultados de previsão obtidos por ESN04. ....	68
Figura 33 – Resultados de previsão obtidos por MQ01. ....	69
Figura 34 – Resultados de previsão obtidos por MQ02. ....	69
Figura 35 – Resultados de previsão obtidos por MQ03. ....	70
Figura 36 – Resultados de previsão obtidos por MQ04. ....	70
Figura 37 – Resultados obtidos pelo modelo <i>BAGGING</i> . ....	71
Figura 38 – Resultados obtidos pelo modelo <i>BAGGING</i> . ....	72
Figura 39 – Resultados obtidos pelo modelo LS-SVM. ....	73
Figura 40 – Resultados obtidos pelo modelo ANFIS. ....	74
Figura 41 – Resultados obtidos pelo modelo ESN. ....	75
Figura 42 – Resultados obtidos pelos modelo linear com MQ. ....	76
Figura 43 – Distribuição dos valores de $R^2$ obtidos na fase de treinamento. .	77
Figura 44 – Distribuição dos valores de $R^2$ obtidos na fase de testes. ....	77
Figura 45 – Distribuição dos valores de $R^2$ obtidos na fase de treinamento e testes para o método dos MQ. ....	78
Figura 46 – Resultados de previsão PARA O CENÁRIO 2 obtidos por BAGGING04.	
80	
Figura 47 – Resultados de previsão PARA O CENÁRIO 2 obtidos por ESN02.	80



## LISTA DE TABELAS

Tabela 1 – Resultado em termos dos critérios de desempenho. ....	55
Tabela 2 – Resultado em termos dos critérios de desempenho sem sinal $x_6(t-1)$ como sinal de entrada do modelo. ....	79

## LISTA DE ABREVIATURAS E SIGLAS

BMS	<i>Battery management system</i>
CAN	<i>Controller area network</i>
CM	Comitê de máquinas
CO <sub>2</sub>	Dióxido de carbono
ECU	Unidade eletrônica de controle
EM	Motor elétrico
Detran	Departamento de Trânsito
IBGE	Instituto Brasileiro de Geografia e Estatística
ICE	Motor a combustão interna
ISG	<i>Integrated starter generator</i>
LEZ	Zonas de baixa emissão
LIB	Bateria de Lítio-íon
ML	Aprendizado de máquina
MSE	<i>Mean Squared Error</i>
$R^2$	Coeficiente de determinação
SOC	<i>State of Charge</i>
SVM	Máquinas de vetores de suporte

## SUMÁRIO

1	Introdução .....	14
1.1	Motivação.....	17
1.2	Objetivos .....	18
1.2.1	Objetivo geral .....	18
1.2.2	Objetivos específicos .....	19
1.3	Organização do trabalho.....	19
2	Veículos híbridos.....	20
2.1	Breve histórico .....	21
2.2	Classificação dos veículos híbridos .....	23
2.2.1	Micro-híbridos.....	23
2.2.2	Médio-híbridos ( <i>mild</i> ).....	23
2.2.3	Híbridos completos ( <i>full hybrid</i> ).....	24
2.3	Classificação por topologia .....	24
2.3.1	Veículo convencional .....	25
2.3.2	Veículo elétrico.....	25
2.3.3	Veículo híbrido – série.....	25
2.3.4	Veículo híbrido – paralelo.....	26
2.3.5	Veículo híbrido – <i>plug-in</i> .....	26
3	Fundamentos de identificação de sistemas e modelos de previsão.....	29
4	Abordagens de aprendizado adotadas.....	33
5	Modelagem e parametrização .....	35
5.1	<i>Ensembles</i> .....	37
5.2	<i>Subspace</i> .....	38
5.1	<i>Bagging</i> .....	38
5.2	<i>Boosting</i> .....	39
5.3	Redes neurais artificiais .....	40
5.4	Sistemas nebulosos ( <i>fuzzy systems</i> ) .....	42
5.5	ANFIS .....	43
5.6	ESN.....	43
5.7	Máquina de vetor de suporte (SVM) .....	45
5.8	Parametrização dos métodos.....	49

5.8.1	<i>Bagging</i> .....	49
5.8.2	<i>Boosting</i> .....	49
5.8.3	SVM .....	49
5.8.4	ANFIS.....	50
5.8.5	ESN.....	50
6	Análise dos resultados da modelagem.....	52
7	Conclusão .....	83
	Referências.....	85

# 1 INTRODUÇÃO

Em seu princípio os projetos automotivos tinham como principal objetivo proporcionar a locomoção. Com o passar dos anos a evolução dos meios de transporte e seus projetos agregados ao desejo dos consumidores, levaram à combinação de elementos tais como motores potentes, carrocerias mais seguras e itens que favorecem o conforto e o entretenimento.

A evolução teve suas consequências, e uma delas é a questão ambiental, pois motores convencionais utilizam a queima de combustíveis como fonte de energia, e eliminam partículas nocivas a saúde. Por isso, os órgãos regulamentadores estabelecem limites na quantidade de particulados e poluentes emitidos através de normas de emissões.

De acordo com ROCHA, ALBERTON e OLIVEIRA (2014) a eficiência dos veículos convencionais chega no máximo a 30%, sendo que o restante é dissipado na forma de calor devido ao atrito dos componentes do motor. A organização mundial da saúde, afirma que  $\frac{1}{4}$  das doenças da humanidade está relacionada a exposição prolongada à poluição (KIMANI, 2012) Além da asma e outros problemas respiratórios já conhecidos, também relaciona-se a poluição como o caso apresentado em NISHIOKA et al. (2000) o qual associa a mortalidade neonatal da cidade de São Paulo com o mesmo problema.

Em algumas cidades do mundo, tais como Londres e Tóquio, foram estabelecidas zonas de baixa emissão (*Low Emission Zones*, LEZ) com o intuito de melhorar a qualidade do ar e reduzir a poluição nessas regiões. Geralmente localizadas nos centros das cidades as LEZs não apenas visam a diminuição de dióxido de carbono (CO<sub>2</sub>), mas qualquer poluente tóxico. As cidades suecas Göthenburg, Malmo e Estocolmo, por exemplo, foram ainda mais longe, e em 1996 estabeleceram seu próprio limite. Em 2002 essa mudança representava 80% da regulamentação europeia vigente na época (AUTOMOTIVE WORLD, 2011).

Por outro lado, de acordo com BERGGREN; MAGNUSSON (2012), a indústria automotiva tem tido dificuldades em acompanhar as legislações de emissões, pois nos últimos anos pouco foi desenvolvido quando trata-se apenas da tecnologia dos motores a combustão interna (ICE). Em um trecho

apresentado na fonte anterior, é comentado que a 10 anos atrás seria impossível atingir uma boa eficiência e ter a redução de particulado e emissões proposta pela legislação Euro 6. Mesmo assim, BERGGREN e MAGNUSSON (2012) comentam que mesmo em casos apresentados na Suécia, na eminência entre a regulamentação e a competitividade do mercado, tem-se buscado a expansão do repertório de tecnologias, combinando-as de forma a atingir os patamares estabelecidos pelas normas regulamentadoras.

No Brasil, por termos um potencial agrícola, uma das formas empregadas para diminuir a poluição foi através da utilização de combustíveis alternativos, e neste caso pode-se citar o Biodiesel o qual é obtido através da reação química de óleos vegetais ou gordura animal junto com o álcool comum ou metanol. O Biodiesel é obrigatório desde 2008 conforme o Ministério de Minas e Energia, o produto pode ser extraído de vegetais tais como a mamona, o dendê, a soja e outros, substituir parcialmente ou totalmente o óleo diesel de petróleo (MDA, 2013). De acordo com União Brasileira do Biodiesel e Bioquerosene, o incremento na mistura obrigatória reduz em 5% a emissão de gases de efeito estufa (BIOENERGIA, 2015).

Além da utilização de combustíveis alternativos, o Brasil ainda conta com iniciativas em certas cidades, como na capital paranaense que desde 2012 já tem em circulação diversos ônibus híbridos, e recentemente a cidade testou um veículo elétrico híbrido do tipo *plug-in* que promete uma redução de 75% de diesel, pois parte do trajeto opera apenas em modo elétrico (AUTOMOTIVE WORLD, 2016).

Por combinarem as melhores características do ICE junto com um motor Elétrico (EM), os veículos híbridos tem se destacado na busca de fontes alternativas na mobilidade, tendo um impacto direto na redução de emissões e por obterem índices até 50% inferiores em relação aos limites estabelecidos na Europa (VOLVO, 2014).

De forma a guardar a energia necessária para utilizar com o EM, esses veículos utilizam um conjunto de baterias de Lítio-íon (LIB), mais preferido pelo setor automotivo devido a características como a alta densidade energética (OPITZ et al., 2017). Para operar de forma segura sem danificar os componentes e com um desempenho ótimo, um sistema de controle de



baterias (*Battery management system*, BMS) é necessário, e segundo (DICKINSON,1995) esta é uma ferramenta importante para a comercialização dos veículos desse tipo. Isso ao fato de que o BMS estima parâmetros tal como o SOC (*State of Charge*), ou seja, o indicador da quantidade de energia disponível na bateria, o qual tem papel importante para a melhor utilização da energia e conseqüentemente uma melhor autonomia do veículo.

## 1.1 MOTIVAÇÃO

Uma motivação dessa dissertação é investigar a aplicação de técnicas de identificação de sistemas dinâmicos baseadas em algoritmos de aprendizado de máquina para estimação do estado de carga de baterias de um veículo híbrido.

Em veículos elétricos e híbridos as baterias tem um papel fundamental para a propulsão elétrica, pois são a fonte do sistema de armazenamento da energia. As baterias de Lítio-íon (LIB) em especial, são consideradas a tecnologia mais promissora para tal aplicação devido às suas características e seu potencial de redução de custo (THIELMANN, SAUER, ISENMANN, 2012), quando comparadas com as baterias de chumbo ácido e níquel hidreto metálico. São comercializadas desde 1991, inicialmente em produtos como celulares e computadores portáteis, e desde então diversas pesquisas e patentes são desenvolvidas para melhorar seu desempenho (BRODD, 2002). De acordo com SCROSATI et al. (2015) esse tipo de bateria tende ao longo prazo substituir combustíveis fósseis e tradicionais fontes de energia para aplicações móveis ou estacionárias.

Ao serem utilizadas em veículos, as LIBs são equipadas com um sistema de controle de bateria (BMS), o qual combina *software* e *hardware* para assegurar o melhor uso da energia armazenada, e prevenir o risco de danificá-las (POP et al., 2008), monitorando diversos estados durante a sua operação. Esse monitoramento constante e durante a operação é denominado monitoramento *on-line* da bateria.

Quando trata-se de *hardware*, apenas grandezas como a corrente, tensão e temperatura externa da célula são itens medidos, já o *software*, deve levar em consideração limites seguros para a operação das LIBs como a máxima corrente de carga e descarga das baterias, a tensão de cada célula e os limites de temperatura. Entretanto, casos como a máxima corrente de carregamento, variam quanto a fatores externos, e suas características são alteradas em temperaturas mais baixas para prevenir que durante o carregamento os íons de lítio se depositem no ânodo da bateria, e comprometam o componente (WAAG et al., 2014).

O *software* da BMS estima diversos parâmetros, entre eles o SOC. Utilizado como o majoritário parâmetro da quantidade de energia remanescente dentro do conjunto de baterias, o SOC é comunicado entre diversos outros subsistemas, aplicações e unidades eletrônica de controle (ECUs) do veículo. Definido por TEAM (2008) como a percentagem da máxima carga possível dentro de uma bateria recarregável. O SOC é expresso pela seguinte expressão:

$$SoC = \frac{C_r}{C_{Bat}} \cdot 100\% \quad (1.1)$$

onde  $C_r$  representa a capacidade da bateria naquele momento e  $C_{Bat}$  após a carga completa.

Em outras palavras, pode-se associar o SOC ao medidor do tanque de combustível de um veículo, o qual serve de informação para o sistema utilizar da melhor forma a energia elétrica disponível nas baterias.

Como cada célula do conjunto de bateria é constituída por complexos processos químicos não lineares que podem variar de acordo com a sua construção ou condições do ambiente, e as únicas grandezas que podem ser medidas externamente são a tensão, corrente da bateria e a temperatura externa de cada célula, a determinação do SOC se torna difícil, dessa forma é necessário estimar este parâmetro.

## 1.2 OBJETIVOS

Nesta seção são apresentados os objetivos do trabalho, primeiramente os objetivos gerais e em seguida os objetivos específicos.

### 1.2.1 Objetivo geral

Com a utilização de técnicas de Aprendizado de máquina, esta dissertação visa a modelagem do estado de carga de baterias, com dados obtidos por uma rede CAN (*Controller area network*) de um veículo do tipo híbrido *plug-in*.

### 1.2.2 Objetivos específicos

Os objetivos específicos são os seguintes:

- a) Avaliar diferentes técnicas de modelagem de aprendizado de máquina, e buscar o melhor desempenho de previsão do SOC aliado a um menor custo computacional;
- b) Validar a aproximação das técnicas de Aprendizado de Máquina por meio da avaliação de desempenho do erro médio quadrático (*Mean Squared Error*, MSE) e do coeficiente de determinação ( $R^2$ ) para avaliar a qualidade dos resultados de modelagem.

### 1.3 Organização do trabalho

Esta dissertação está organizada da seguinte maneira. Neste capítulo apresentaram-se a introdução, justificativa e objetivos da dissertação. No capítulo 2 mostram-se os fundamentos dos veículos híbridos. O capítulo 3 traz os fundamentos de identificação de sistemas e modelos de previsão.

O capítulo 4 apresenta as abordagens de aprendizado adotadas. Após, no capítulo 5 os métodos de Aprendizado de Máquina a serem avaliados em procedimentos de modelagem do SOC são descritos. No capítulo 6 mostram-se os resultados obtidos com a aplicação de métodos de identificação de sistemas e aprendizado de Máquina para modelagem do SOC. Finalizando, no capítulo 7 onde são apresentados comentários gerais quanto aos objetivos de dissertação, os resultados obtidos e uma proposta de continuação para a pesquisa.

## 2 VEÍCULOS HÍBRIDOS

Os veículos elétricos, incluindo elétricos puros, e híbridos, usualmente são reconhecidos na mobilidade como uma das soluções para problemas relacionados a crise energética ou a problemas ambientais (WAAG et al., 2014). Isso se dá pela eficiência do seu sistema de tração. No caso dos veículos híbridos, é utilizado em sua tração um sistema de propulsão adicional ao convencional, que consiste em um EM e um sistema de armazenamento de energia. Desta forma os veículos híbridos estão intermediários ao veículo elétrico puro e o veículo convencional, pois combinam as duas formas de tração, a elétrica e a combustão, enquanto o EM é mais adequado para baixas rotações, o ICE tem vantagem para tarefas as quais demandam percorrer longas distâncias (STREHLER et al., 2016).

Diferentemente dos ICE os EM tem um torque melhor, o qual está disponível mesmo em baixas rotações, sendo assim uma boa opção para a propulsão de veículos em baixas velocidades. Ainda é caracterizado pela sua alta eficiência, robustez, vida útil e baixo ruído (SCROSATI et al., 2015). Além disso, também é utilizado no sistema *start/stop* o qual desliga o motor a combustão enquanto o veículo está parado em um sinalizador por exemplo.

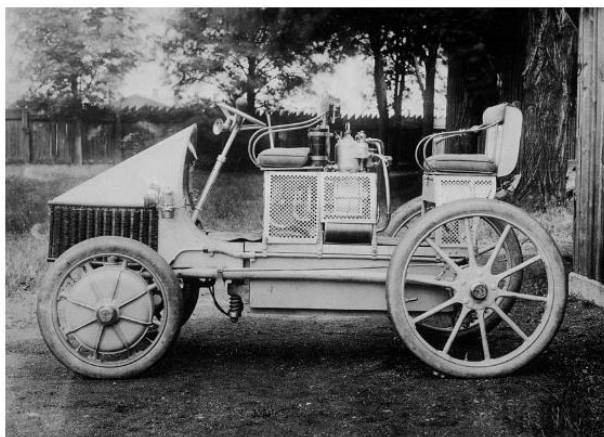
Em comparação com os veículos convencionais, os híbridos oferecem a possibilidade de um ICE menor, e a energia das rodas pode ser fornecida tanto pelo ICE quanto pelo EM (ENANG; BANNISTER, 2017). Outra vantagem, é a energia dissipada durante as frenagens que pode ser reaproveitada, convertendo a energia cinética em energia elétrica, para mais tarde ser armazenada nos conjuntos de baterias. Esse processo é denominado frenagem regenerativa, melhorando a eficiência do veículo em ambientes urbanos (DEES, 1999).

A combinação da mecânica, elétrica, e eletrônica unem as melhores práticas de cada tecnologia, e o sistema de tração adicional proporciona ainda uma maior flexibilidade no uso do motor quanto à força requerida nas rodas.

## 2.1 Breve histórico

Mesmo que muitas das tecnologias presentes nos veículos híbridos tenham sido melhoradas nos últimos anos, o princípio do desenvolvimento não ocorreu neste século. De acordo com HOYER (2008), existem registros da produção de um veículo híbrido já em 1903, o qual provavelmente foi apresentado por Ferdinand Porsche no salão do automóvel em Paris.

FIGURA 1 – PRIMEIRO VEÍCULO HÍBRIDO



FONTE: (CALMON, 2011)

Segundo a mesma fonte, a história dos veículos híbridos e elétricos começou por volta do século XIX, pois em 1859, Glaston Planté demonstrou a primeira bateria chumbo ácido. Alguns anos mais tarde, Thomas Edison desenvolveu a bateria de níquel ferro com uma capacidade de energia 40% melhor que a versão de Planté. Neste mesmo período outras baterias como níquel zinco e zinco ar também foram apresentadas.

Em 1903, Nova York tinha registro de 4 mil automóveis contando com diferentes tecnologias: elétrico, a vapor e a gasolina. Quando a cidade atingiu seu ápice com 30 mil unidades, fatores como a produção em série de veículos, e a gasolina, barateavam o produto. O aumento da demanda em percorrer longas distâncias e a descoberta de poços de petróleo no Texas contribuíram para que a quantidade dos veículos movidos a gasolina fosse pelo menos 30 vezes maior (STRUBEN, 2006).

Quanto a questão ambiental, somente na conferência Rio 92, foi que foram discutidas questões sobre o aquecimento global e suas consequências



pela primeira vez. A conferência resultou na Agenda 21, a qual enfatizava os problemas da utilização de combustíveis fósseis, almejava a redução de tal energia nos países desenvolvidos e buscava fontes alternativas.

Ainda por volta dos anos 90 foi realizada uma nova tentativa para introduzir os automóveis híbridos, desta vez com parcerias público privadas, mas nenhum veículo chegou ao fim da produção. Somente após o lançamento do Toyota Prius em 1997 foi que outras montadoras como a Honda e a Audi tomaram o mesmo rumo. O Prius daquele ano se tratava de sedã híbrido com um motor de 1,5 litros e quatro cilindros acoplado a um motor elétrico alimentado por uma bateria de 274 volts (BARAN, LEGEY, 2010). O modelo fazia 19 Km/litro na estrada, e em uma matéria apresentada na Forbes, Bob Carter, atual vice-presidente da Toyota, lembra que embora os Estados Unidos da América não tenham dado tanto valor ao Prius, na mesma época a empresa vendeu 8,5 milhões de veículos híbridos, e dos quais 5,4 milhões eram Prius (FORBES, 2016).

No Brasil, os veículos elétricos e Híbridos ainda são raros, embora já no mercado há algum tempo a grande maioria disponível são carros topo de linha, como é o caso do Ford Fusion Hybrid, Toyota Prius, Mercedes S400, Lexus CT 200 e Nissan Leaf. Fora esses modelos ainda existem os utilizados em por algumas empresas privadas ou alguns órgãos públicos como é o caso do Renault Zoe e o Renault Twizy. Este último utilizado como transporte interno na usina Itaipu (FELIX, 2015; SILVIO, 2016).

## 2.2 Classificação dos veículos híbridos

Existem basicamente duas formas de classificar os veículos híbridos. Quanto a topologia e a quantidade de tecnologia agregada ao veículo.

Os principais tipos são: micro-híbridos, médio-híbrido (*Mild*), Híbrido completo (*Full Hybrid*). Para esta seção foi utilizada como base o trabalho de (ENANG; BANNISTER, 2017).

### 2.2.1 Micro-híbridos

Os micro HEV (*Hybrid Electric Vehicle*) possuem como fonte de tração apenas um motor a combustão, embora um motor de partida/alternador ISG (*Integrated Starter Generator*) é utilizado para desligar o ICE quando o veículo está completamente parado. O ICE volta imediatamente a operar apenas quando o pedal de freio é solto e o acelerador é pressionado. Tal sistema é conhecido como *Start/Stop*. Exemplos desse tipo podem ser encontrado em veículos como BMW 1 e série 3, Fiat 500, SMART e outros.

FIGURA 2 – MOTOR DE PARTIDA/ALTERNADOR



FONTE: INDIAMART

### 2.2.2 Médio-híbridos (*mild*)

Apesar de ainda contar como principal fonte de tração um motor a combustão, o médio híbrido pode-se dizer que é uma melhoria do micro híbrido, pois neste caso o ISG é maior, podendo ajudar durante a propulsão do veículo. Dessa forma é possível uma economia de 20 a 25% quando comparado aos não híbridos. Exemplos de *Mild Hybrid* são: Chevrolet Malibu com eAssist e o Mercedes-Benz s400 Blue Hybrid.

### 2.2.3 Híbridos completos (*full hybrid*)

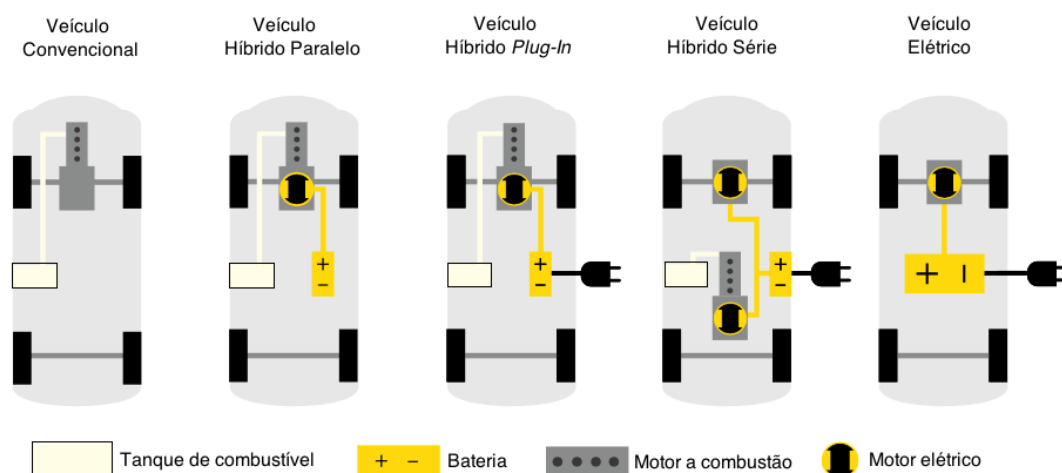
Nos híbridos completos o motor elétrico é significativamente maior do que dos tipos anteriores e pode ser utilizado como a única fonte de energia de propulsão.

Esses veículos possuem um ICE menor quando comparado as anteriores embora necessitem de um sistema de controle de energia mais desenvolvido, e oferecem uma redução de 40 a 45% de combustível quanto aos não híbridos. Exemplos desses são o Toyota Prius, e Camry *Hybrid*.

### 2.3 Classificação por topologia

Para TEIXEIRA DIAS (2013) existe três tipos básicos de configurações em veículos híbridos: série, paralelo e *plug-in*. Para complementar e verificar as diferenças de forma clara, também são apresentados a topologia do veículo convencional e do veículo elétrico. A Figura 3 apresenta as topologias que são descritas.

FIGURA 3 – AS DIFERENTES TOPOLOGIAS



FONTE: SPATH ET AL. (2011)

### 2.3.1 Veículo convencional

O sistema encontrado normalmente nos veículos atuais, seja este a gasolina, álcool ou a diesel, tem como única fonte de energia o motor a combustão, sendo este conectado diretamente às rodas, além da transmissão e embreagem.

Apesar desse sistema de veículo convencional apresentar autonomia, motivo pelo qual foi primeiramente desenvolvido antes aos elétricos ou híbridos, tem uma baixa eficiência. De acordo com ROCHA, ALBERTON e OLIVEIRA (2014) a eficiência dos veículos convencionais chega, no máximo, a 30%, sendo que restante é dissipado na forma de calor devido ao atrito dos componentes do motor. Também são presentes nesse sistema a poluição, devido a queima de combustíveis, e o barulho.

### 2.3.2 Veículo elétrico

No caso dos veículos elétricos, algumas das suas principais vantagens estão em sua alta eficiência energética, sua manutenção menos frequente e a não emissão de gases poluentes, embora como apresentado em BARRETO (1986) e OLIVEIRA, FRANCISCO e SILVA (2007) ainda tenha um alto custo, e apresenta também autonomia limitada a energia armazenada em suas baterias. Além do sistema de armazenamento de energia, o sistema basicamente conta com o motor elétrico para a geração da força motriz.

### 2.3.3 Veículo híbrido – série

Os veículos com configuração em série apresentam apenas o motor elétrico como fonte de tração, enquanto o motor a combustão gera a energia necessária para carregar as baterias e energizar o motor elétrico. Os motores a combustão nestes veículos tendem a ser menores, pois atuam apenas como geradores.

### 2.3.4 Veículo híbrido – paralelo

Nesta configuração os motores a combustão e o motor elétrico podem tracionar. Em pequenas velocidades o motor elétrico propulsiona o veículo e quando é necessário mais força e torque o motor a combustão entra em operação. Da mesma forma como os veículos elétricos e em série, o motor elétrico disponível na topologia paralela pode também atuar como gerador, recarregando as baterias em momentos de desaceleração. Entretanto mesmo presente nessa configuração, o freio regenerativo é limitado, pelo fato do motor elétrico ser menor que o do veículo em série ou elétrico.

### 2.3.5 Veículo híbrido – *plug-in*

A versão *plug-in* em veículos híbridos refere-se à possibilidade destes serem carregados por uma fonte de energia externa, ou seja, não apenas através do motor a combustão do próprio veículo ou pelo aproveitamento da energia dos freios. Para isso, deve haver investimento nas instalações urbanas, disponibilizando assim estações de recarga que irão adaptar a corrente externa para a tensão necessária ao carregamento das baterias do veículo (SCROSATI et al., 2015).

FIGURA 4 – ESTAÇÃO DE CARREGAMENTO – TIPO PANTOGRÁFICA



FONTE: Opp charge, 2017

As estações de recarga já são realidade em países como a Suécia ou os Estados Unidos da América, embora ainda seja possível encontrar casos de estações em algumas cidades brasileiras.

O sistema de recarga depende das opções disponíveis no veículo. Atualmente existem duas possibilidades diferentes, em uma delas o carregador está instalado no veículo e o usuário apenas faz a conexão em uma tomada, outro caso é quando o carregador está externo ao veículo, seja em estruturas como totens ou em pontos de recarga, conforme apresentado nas Figuras 5 e 6.

FIGURA 5 – ESTAÇÃO DE CARREGAMENTO – TIPO CONECTOR



FONTE: ABB GROUP LTD. (2016)

Além das formas apresentadas anteriormente, uma parceria da Scania, Siemens e diversos outros setores públicos e privados da Suécia implementou, recentemente, uma autoestrada elétrica. O sistema é similar a trens elétricos, em que uma linha de alta tensão percorre o traçado e o veículo é conectado por uma pantográfica, entretanto este caso é dedicado a caminhões do tipo *plug-in* e está em uma estrada que conectam duas cidades suecas, conforme mencionado por ATHERTON (2016).

FIGURA 6 – AUTOESTRADA ELETRIFICADA



FONTE: ATHERTON (2016)

### **3 FUNDAMENTOS DE IDENTIFICAÇÃO DE SISTEMAS E MODELOS DE PREVISÃO**

A identificação de sistemas é a área do conhecimento que estuda maneiras de modelar e analisar sistemas a partir de observações, isto é, de dados (AGUIRRE, 2007).

Para melhor compreensão do que se trata a identificação de sistemas, os conceitos de sistema e modelo são apresentados. Segundo HAYKIN e VAN VEEN (2001) um sistema é formalmente definido como uma entidade que manipula um ou mais sinais para realizar uma função, produzindo assim, novos sinais. O termo modelo significa o conjunto de hipóteses sobre a estrutura ou o comportamento de um sistema físico pelo qual se procuram explicar ou prever, dentro de uma teoria científica, as propriedades do sistema (HOLANDA, 1988).

Um modelo de sistema segundo BALAKRISHNAN (2002) pode ser definido como a maneira com que as variáveis deste sistema interagem entre si ou ainda como os valores observados em suas entradas estão relacionados. Isto pode envolver certo grau de formalismo matemático e certa sofisticação para que possa ter uma utilização. Do ponto de vista matemático, um modelo nada mais é do que uma abstração, comumente expressa por meio de equações, de um sistema real (SEBORG et al., 1989).

Pode-se também dizer que a identificação de sistemas lida com problemas de construção de modelos matemáticos para sistemas dinâmicos baseados na observação de dados medidos de um determinado sistema. Um sistema é um objeto em que diferentes tipos de variáveis interagem e produzem sinais observados em suas saídas, e é também afetado por estímulos externos através de suas entradas os quais podem ser manipuladas por seu observador. Em geral, a identificação de sistemas pode ser dividida em cinco etapas principais (LJUNG, 1987):



- i. Obtenção de dados de experimentação do sistema;
- ii. Aplicação de testes aos dados obtidos para detecção de não-linearidades;
- iii. Escolha da estrutura que será utilizada para representar o modelo matemático;
- iv. Estimação dos parâmetros do modelo;
- v. Validação do modelo obtido.

Os procedimentos anteriores são utilizados tanto para a identificação de sistemas lineares quanto não-lineares. A diferença está no modo como cada etapa é projetada e validada.

A coleta de dados é um passo importante para a modelagem de sistemas, pois estes dados são utilizados na detecção de não-linearidades e na estimação dos parâmetros do modelo selecionado. Para isto na experimentação do sistema o sinal de entrada necessita de um espectro de frequência adequado para excitar às dinâmicas desejadas do sistema (LEONTARITIS e BILLINGS, 1987). No caso de sistemas não-lineares, isto requer que os efeitos não-lineares sejam excitados e assim estejam presentes nos dados, ou seja, se os dados não representarem bem o sistema, a modelagem também não será bem representada (PÊSSOA, 2010).

A presença de certos parâmetros, tais como atrito, inércia e tempo morto, fazem com que os sistemas sejam não-lineares. Apesar de alguns sistemas apresentarem uma não linearidade suave, é possível a utilização de modelos lineares, que em torno de um ponto de operação, gerem resultados satisfatórios. Quando a modelagem linear não é suficiente para se obter um bom modelo, é necessário empregar representações não-lineares que obtém uma melhor descrição do comportamento dinâmico do processo, porém estes são mais complexos (PÊSSOA, 2010).

Uma importante etapa na identificação de sistemas é a escolha do modelo a ser utilizado para representar a dinâmica. Esta escolha se torna mais importante no caso de sistemas não-lineares, onde há, muitas vezes, uma acentuada quantidade de elementos não-lineares, e com isso é necessário escolher um modelo complexo que consiga representar a diversas não-linearidades (PÊSSOA, 2010).

Após ter escolhido o modelo que representará o sistema, deve-se estimar os parâmetros, para que o modelo identificado consiga gerar uma dinâmica semelhante à do sistema. Depois desta etapa, a validação do modelo é a última etapa do procedimento de identificação. Nela se verifica se o modelo obtido consegue representar o processo original, com um erro aceitável (PÊSSOA, 2010).

Deve-se ressaltar que os sistemas dinâmicos em sua maioria são não-lineares, mas há casos em que aproximações lineares são suficientes para propósito de utilização prática. Entretanto, em muitos casos, os sistemas lineares não conseguem reproduzir satisfatoriamente certas dinâmicas, com isso é necessária à utilização de representações não-lineares, as quais possuem algoritmos mais aprimorados.

Para a identificação do sistema foram abordadas diferentes técnicas de modelagem, as quais envolvem técnicas de Aprendizado de Máquina (*Machine Learning*, ML) buscando o melhor resultado de previsão aliado ao menor custo computacional.

Aprendizado de máquina pode ser considerada uma sub-área de pesquisa da Inteligência artificial que estuda como desenvolver sistemas capazes de adquirir conhecimento de maneira automática, ou seja, sistemas que melhoram automaticamente seu desempenho com a experiência (MITCHELL, 1997). Os algoritmos de ML constroem sistemas capazes de inferir um modelo a partir de um conjunto de dados de treinamento. Esse modelo pode então ser aplicado a dados não vistos na etapa de treinamento, para realizar previsões a respeito desses dados.

Muitas abordagens de ML permitem melhorar o desempenho na resolução de problemas com o aprendizado ou experiência adquirida “a partir dos dados utilizados, além de descobrir regras, relações ocultas, até chegar numa resposta” (VILLANUEVA, 2006). Os avanços recentes em ML contemplam ampla variedade de tarefas de inteligência artificial (BUCZAK e GUVEN, 2016) e identificação de sistemas (PILONETTO et al., 2014), que têm sido aplicadas em vários problemas de engenharia, economia, biomedicina, computação, física, dentre outros. Tarefas tais como reconhecimento de padrões, previsão de séries temporais, controle adaptativo e detecção de falhas podem ser formuladas em termos da busca de dependências

escondidas em observações empíricas. Tal busca desempenha um papel central no contexto de ML e corresponde ao problema de aprendizado supervisionado. Dentre as várias técnicas de ML estão as Redes Neurais Artificiais (RNAs), Árvores de Decisão (*Decision Trees*) e Máquinas de Vetores de Suporte (*Support Vector Machines*, SVM).

Entretanto, recentemente tem sido propostas abordagens de ML denominadas comitê de máquinas (do inglês *Committee Machines*, CMs) que são guiados pelo princípio dividir-para-conquistar onde cada máquina de aprendizado atua isoladamente. Os algoritmos de CMs são abordagens que permitem combinar as respostas de várias máquinas de aprendizado com o intuito de obter a melhor resposta, buscando, através de aprendizes fracos a determinação de “preditores fortes” (VILLANUEVA, 2006).

Em outras palavras, trata-se da obtenção de uma saída, para um problema de classificação ou de regressão, a partir das saídas individuais propostas por múltiplas soluções alternativas para o mesmo problema, denominadas componentes do *ensemble*. Apesar da falta de uma teoria de *ensemble* unificada, há muitas razões teóricas para combinar múltiplos componentes – classificadores ou regressores que irão compor o *ensemble* – e alguma evidência empírica da eficiência desta abordagem.

## 4 ABORDAGENS DE APRENDIZADO ADOTADAS

O SOC tem uma representação relevante nas baterias, por representar a quantidade de energia disponível, além da utilização para prevenir uma descarga acentuada, melhorar a vida útil da bateria e operação segura da mesma. Embora as baterias sejam fontes de armazenamento de energia baseado na química, variações externas dificultam a verificação do valor, entretanto uma boa precisão de seu valor é necessária para uma boa estratégia energética no veículo.

Na literatura são encontrar técnicas diretas de medição do SOC como o *Open Circuit Voltage* (OCV), apresentadas em CHAGAS (2007) e XING (2014), a qual verifica a diferença de potencial nos terminais, e o *Coulumb Counting* (CC), o qual integra a corrente de entrada e saída da bateria ao longo do tempo. Segundo TRINH (2012) esses não são precisos o suficiente para aplicações em veículos híbridos e elétricos, pois resultam em um alto erro de SOC.

Por sua vez, WANG et al. (2013) mencionam que fatores como descarga de corrente, temperatura interna da bateria, auto descarga, e vida útil dificultam o cálculo. Sendo assim algoritmos adaptativos têm sido utilizados para estimar e analisar com o intuito de melhorar o desempenho e a vida útil das baterias. Dessa forma métodos baseados em modelos, os quais fazem parte do estudo deste trabalho são empregados para estimar o valor do SOC. Para essa forma de aplicação os mais utilizados são filtro de Kalman, redes neurais artificiais, sistemas nebulosos e máquinas de vetor de suporte.

Os modelos baseados em filtro de Kalman estimam o valor do SOC modelando a bateria com parâmetros não específicos requeridos para a medição do SOC (JEON et al., 2015). Em GARCHE (2000) a estimativa é realizada com base na corrente e temperatura da bateria e comparado com o valor verdadeiro de tensão. Entretanto, JEON et al. (2015) afirmam que o filtro de Kalman estendido é mais usado para casos com não linearidade. YAN et al. (2010) mencionam que para ser um estimador “ótimo” o filtro de Kalman estendido necessita de um modelo preciso com conhecimento dos erros

envolvidos, e completa ao dizer que não podem ser utilizados para processamento *on-line* no veículo.

A utilização de filtros de Kalman aplicados para o SOC pode ser encontrado em referencias como em YU et al. (2015) o qual conseguiu uma boa aproximação atualizando a linearização dos parâmetros medidos, LI et al. (2016), PAN et al. (2017) e ZHI et al. (2017).

Já para a aplicação de redes neurais e máquinas de vetores de suporte, este pode ser observado como em QIU et al. (2014) o qual utiliza uma técnica baseada em redes neurais para realizar uma previsão na área energética e ainda a compara com outras técnicas, dentre elas a máquina de vetores de suporte.

A seguir nas próximas seções serão apresentados os fundamentos das abordagens de ML adotadas neste trabalho para tarefa de previsão de carga de baterias para veículos híbridos.

## 5 MODELAGEM E PARAMETRIZAÇÃO

Em parceria com uma empresa do ramo automotivo, foram adquiridos para este estudo dados disponíveis na rede CAN de um veículo *plug-in* do tipo paralelo que encontra-se em funcionamento em um trajeto urbano na Inglaterra.

A base de dados fornecida contém diversas instâncias, 69 ao todo, com 4281 amostras de cada característica. Dentre essas características estão disponíveis desde dados de topografia, emissões, da bateria, do motor a combustão e outros.

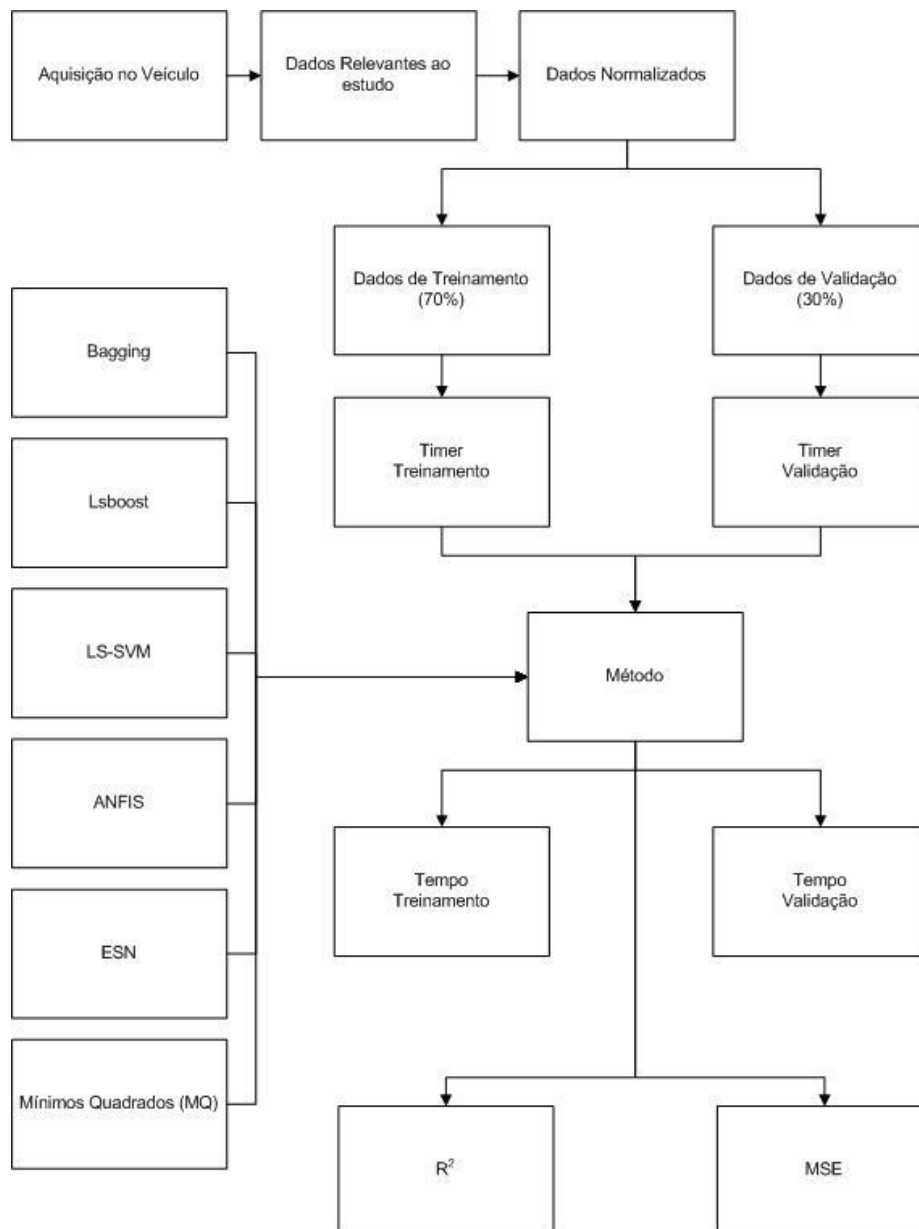
Para este trabalho foram usadas as seguintes instâncias como sinais de entrada: quantidade de combustível injetado ( $x_1$ ) em mg, torque ( $x_2$ ) N.m, velocidade do motor elétrico ( $x_3$ ), posição do pedal do acelerador ( $x_4$ ) em %, velocidade do veículo ( $x_5$ ) em km/h, e SOC ( $x_6$ ) atrasado um instante de tempo ( $t-1$ ) em relação a saída (em %), o que insere a saída  $y(t-1)$  como entrada do sistema. Como sinal de saída  $y(t)$  (a ser identificado) foi adotado o SOC( $t$ ) (*State Of Charge*) no instante atual.

É importante ressaltar que neste caso o SOC já está disponível na rede do veículo e que não pretende-se avaliar a forma como foi adquirido este dado, mas utilizar o dado disponível como parâmetro de entrada em diferentes técnicas de predição a pelo menos um estado de tempo a frente, verificando a confiabilidade do modelo, seu desempenho quanto ao tempo além de comparar com as técnicas aplicadas. Por motivos de confidencialidade, a base de dados foi normalizada e os valores reais foram distribuídos em uma faixa entre 0 a 1.

As técnicas abordadas foram *subspace*, *bagging* e *boosting*, pertencentes ao grupo dos *ensembles*, escolhidas devido ao seu bom desempenho em competições, tais como o promovido pelo Netflix Prize (TORSCHER e JAHNER, 2009); O ANFIS (*Adaptive Neuro- Fuzzy Inference System*) e a rede de estado de eco (*Echo State Network*, ESN) são baseadas em RNAs, e por fim a SVM pelo fato de ser um promissor método de ML.

A seguir, na Figura 7 pode-se observar um esquemático dos passos de projeto adotados para o programa desenvolvido.

FIGURA 7 - ESQUEMÁTICO DO PROGRAMA



FONTE: O AUTOR, 2017

As simulações foram geradas em um computador com processador Intel Core i7-5820K 3,30GHz, 128 Gb de RAM (*Random Access Memory*) com sistema operacional Windows 7 e ambiente computacional Matlab® R2015a. Todos os métodos foram submetidos e modelados com a mesma base de dados, e otimizados com relação aos parâmetros de ajuste, a fim de obter o melhor desempenho para cada modelo.

## 5.1 ENSEMBLES

Um *ensemble* consiste de um conjunto de componentes, cada um representando uma proposta de solução, cujas classificações, no caso de problemas de classificação de padrões, ou estimativas, no caso de problemas de regressão, são combinadas de formas diversas visando um ganho de desempenho. Este paradigma originou-se do trabalho de HANSEN e SALAMON (1990), que mostraram que a habilidade de generalização pode ser significativamente melhorada por meio da composição de várias redes neurais artificiais, ou seja, o treinamento independente de várias redes neurais artificiais e posterior composição das saídas individuais. O bom desempenho dos *ensembles* tem incentivado o seu emprego em todas as linhas de atuação em aprendizado de máquina, de modo que outros componentes vêm sendo adotados para compor o *ensemble*, além de redes neurais artificiais (LIMA, 2004). Um apanhado da literatura de CM é detalhado em REN et al. (2016).

Para casos de regressão e/ou identificação de sistemas, o objetivo é identificar decisões contínuas, agregando diversos modelos com pesos em um único modelo, por meio de:

$$\gamma_r = f_r(x, \theta_r), \gamma_r \in \mathcal{R} \quad (5.1)$$

onde  $x$  é o novo vetor a ser observado,  $\gamma_r$  o resultado,  $f_r(\cdot)$  é a função de regressão e  $\theta_r$  o parâmetro de ajuste da função de regressão.



## 5.2 SUBSPACE

Disponível em Matlab para regressão, o *subspace* é um dos métodos *ensemble* e pode ser utilizado com aprendizes fracos do tipo “*Discriminant*” (*Discriminant Analysis*) ou KNN (k-Nearest Neighbour) (“Statistics and Machine Learning Toolbox: User’s Guide,” 2016).

A análise discriminante é um método de classificação o qual assume diferentes classes de dados gerados baseados em diversas distribuições gaussianas. Para a predição de novas classes de dados o classificador gerado a partir do treinamento tenta encontrar a classes com menor erro de classificação. No Matlab a análise discriminante para predição pode utilizar três classificadores para as observações: *Posterior Probability*, *Prior Probability* e *Cost*, onde a última é uma matriz associada à análise discriminante por observações.

## 5.1 BAGGING

*Bootstrap* é um processo estatístico de reamostragem que consiste em gerar amostras a partir da amostra original, os dados obtidos aleatoriamente são utilizados na formação de cada amostra “*bootstrap*” denominada de pseudovalores e com reposição dos dados da amostra original, assim o resultado depende diretamente da amostra original.

O termo *Bagging* vem da contração de dois termos em inglês *Bootstrap + Agregating*. *Bagging* foi primeiro algoritmo para a construção de CM Ensembles. Como descrito por POLIKAR (2009), “*Bagging* é um das primeiras e mais intuitivas e talvez o mais simples algoritmo baseado em *Ensemble*”, com um desempenho promissor (BREIMAN, 1994).

O *Bagging* treina seus classificadores utilizando diferentes conjuntos de treinamento gerados aleatoriamente a partir de amostragem uniforme do conjunto de dados, com reposição dos exemplos utilizados. Muitos exemplos podem aparecer mais de uma vez no conjunto de treinamento e outros podem não aparecer nenhuma vez. De cada conjunto de treinamento gera-se um preditor ou classificador diferente, sendo combinados no final. Os classificadores *Bagging* são obtidos por réplicas *Bootstrap* dos dados de

treinamento, sendo assim, diferentes subconjuntos de treinamento são criados aleatoriamente, com reposição de todos os dados de treinamento. Cada subconjunto de treinamento é usado para treinar um classificador diferente do mesmo tipo.

Os classificadores individuais são então combinados tomando o voto majoritário simples de suas decisões. Para cada instância dada, a classe escolhida pelo maior número de classificadores é o conjunto de decisão. Para se evitar a sobreposição dos conjuntos de dados o que pode ser feito para aumentar a diversidade é utilizar um subconjunto de treinamento para treinar cada classificador ou utilizar os classificadores mais fracos como “troncos” da árvore de decisão.

## 5.2 BOOSTING

Os modelos baseados em *boosting*, tal como o LSBoost (do inglês *Least Squares Boosting*) (FREIDMAN, 2001) são uma família de poderosas técnicas de ML que mostraram sucesso considerável em uma ampla gama de aplicações práticas. Este tipo de modelo é flexível quanto a sua customização para as necessidades específicas da aplicação, tal como aprender com diferentes funções objetivo (custo). Modelos de *boosting* quando aplicados no contexto de problemas de regressão linear levam a modelos com propriedades estatísticas atraentes.

A principal diferença entre as abordagens *boosting* e *bagging* é que no *boosting* os componentes da agregação são treinados sequencialmente, pois o conjunto de treinamento do próximo componente é gerado com base no desempenho dos demais componentes já obtidos. Com isso, enquanto a abordagem *bagging* atua apenas na redução da variância, a abordagem *boosting* é capaz de reduzir tanto variância quanto a polarização ou *bias* (LIMA, 2004).

Isto se deve à atribuição de maior ênfase às amostras responsáveis pela queda de desempenho durante o treinamento, ou seja, as amostras que mais contribuem para o erro de treinamento dos componentes já treinados têm maior probabilidade de serem escolhidas na composição do conjunto de treinamento do próximo componente. Lembre-se que, na abordagem *bagging*,

a probabilidade de escolha das amostras é sempre uniforme, razão pela qual o treinamento dos componentes pode ser feito em paralelo, ou seja, simultaneamente (LIMA, 2004).

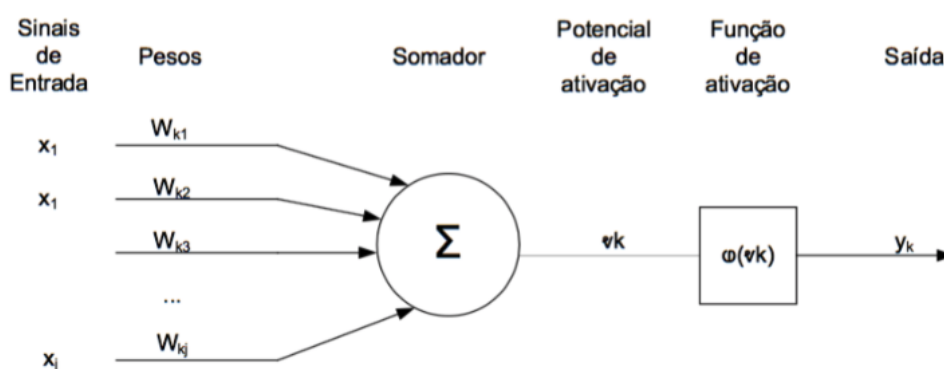
### 5.3 REDES NEURAIS ARTIFICIAIS

Uma rede neural artificial é um sistema de processamento de informações que apresenta certas características em comum com uma rede neural biológica. As redes neurais artificiais são desenvolvidas como generalização de modelos matemáticos de neurônios biológicos com base no seguinte:

- i) o processamento das informações ocorre por intermédio de neurônios;
- ii) os sinais são passados entre os neurônios por meio de conexões;
- iii) cada conexão tem um peso associado;
- iv) para determinar o sinal de saída, cada neurônio aplica na função de ativação a soma dos pesos dos sinais de entrada.

O modelo proposto por McCulloch & Pitts apresenta dois estados: o neurônio disparando  $X_j = 1$  ou em modo de espera com  $X_j = 0$ , este é representado da seguinte forma:

FIGURA 8 – MODELO DE MCCULLOCH & PITTS



Os sinais de entrada, representam as sinapses vindas de outros neurônios, ponderada por seus respectivos pesos sinápticos, ou seja, o

conhecimento adquirido. A soma, o processamento dos estímulos ponderados, gerando assim a entrada para a função de ativação determinando se a informação continuará adiante ou não.

Assim se transcreve-se o modelo neural para fórmula matemática e obtém-se a seguinte formulação para a saída do neurônio:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (5.3)$$

onde:

$x_j$  são as entradas, ou os sinais vindos de outros neurônios;

$w_{kj}$  são os pesos sinápticos do neurônio

$u_k$  é a saída.

Uma rede neural do tipo *feedforward* perceptron multicamadas é composta por vários neurônios distribuídos em camadas. A primeira é chamada de camada de entrada e tem a função de armazenar os dados de entrada na rede. Camadas intermediárias, denominadas de camadas escondidas ou ocultas, ocorrem os processamentos da rede. A última camada é definida como de saída, que recebe os valores já processados pela rede. Os neurônios localizados nas camadas escondidas desempenham um papel importante durante o treinamento, porque agem como detectores de características.

De acordo com JEON et al. (2015), as redes neurais perceptron multicamadas com algoritmo de treinamento da retropropagação do error (*error backpropagation*) possuem uma simples topologia e resolvem problemas não lineares. Estas são compostas por 3 (ou mais) camadas, sendo a primeira com os dados de entrada (tensão, corrente, resistência e temperatura da bateria), uma camada escondida (ou mais) que segundo o autor determina a precisão, e a última camada onde resulta no valor estimado do SOC. Neste caso, YAN et al. (2010) apontam como desvantagem a necessidade de dados de treinamento embora quando devidamente aplicado minimize o erro médio quadrático.

#### 5.4 SISTEMAS NEBULOSOS (*FUZZY SYSTEMS*)

Na teoria clássica de conjuntos, um elemento do universo de discurso (domínio) pertence ou não pertence ao referido conjunto. Na teoria dos conjuntos nebulosos existe um grau de pertinência, no intervalo de  $[0,1]$ , de cada elemento, a um determinado conjunto. A teoria dos conjuntos nebuloso pode ser utilizada para traduzir em termos matemáticos a informação imprecisa expressa por um conjunto de regras linguísticas. O modelo linguístico composto de regras da forma SE-ENTÃO relacionam variáveis de entrada (premissa) e saída (consequente). A estrutura básica de um sistema nebuloso pode ser vista com constituindo-se de:

- i) *base de regras*, que contém um conjunto de regras nebulosas;
- ii) *base de dados*, que define as funções de pertinência das regras nebulosas;
- iii) *mecanismo de inferência*, que executa os procedimentos de inferência sobre as regras e condição para obtenção de uma saída.

Conforme JANG e SUN (1995), a essência dos modelos nebulosos é similar ao princípio de “dividir e conquistar”. Os antecedentes das regras nebulosas dividem o espaço de entrada em um número de regiões nebulosas locais, enquanto os consequentes descrevem o comportamento dentro de uma região através dos seus constituintes. Os componentes dos consequentes resultam em tipos diferentes de modelos matemáticos nebulosos, mas seus antecedentes são essencialmente os mesmos. Em BRUGOS et al. (2015) e SÁNCHEZ et al. (2017) são apresentados algoritmos baseado em modelagem nebulosa para estimação do SOC.

## 5.5 ANFIS

O método ANFIS (*Adaptive Neuro-Fuzzy Inference System*) é uma combinação de RNA com um sistema nebuloso (*fuzzy*) de inferência. ANFIS trata de forma adequada a não linearidade e a incerteza, muitas vezes, presente na modelagem de sistemas complexos (PESSANHA, 2002).

Existem dois tipos de sistemas *fuzzy*, o “*Fuzzy if-then rules*” que baseia-se em expressões condicionais do tipo se então (se A então B) e conseguem capturar as imprecisões do raciocínio humano ao tomar decisões em ambientes incertos ou imprecisos. O “sistema *fuzzy* de inferência” é composto por blocos de funções distintas, contendo blocos de “*fuzzy if-then rules*”, funções associadas às regras *fuzzy*, unidades de tomada de decisão além de interfaces transformação de entradas em saídas pelas regras *fuzzy*.

Segundo JANG (1993) o ANFIS é um sistema *fuzzy* de inferência implementado em uma estrutura de redes neurais adaptativas. O uso do processo de aprendizado híbrido proposto pelo ANFIS permite construir um mapeamento de entradas-e-saídas baseados no conhecimento humano (*fuzzy if- then rules*) e em pares específicos de entradas-e-saídas, sendo assim, é capaz de modelar um sistema não linear, identificar componentes em sistemas de controle e prever séries temporais caóticas.

## 5.6 ESN

As redes com estado de eco (*Eco State Network*, ESN) fornecem uma arquitetura e o princípio de aprendizado supervisionado para redes neurais recorrentes. As ESN fazem uso de uma “grande” rede neural artificial aleatória onde cada unidade da rede recorrente gera sua própria transformação linear dos sinais de entrada. As entradas,  $u(n) = [u_1(n) \dots u_k(n)]^T$ , são linearmente combinadas com os coeficientes na matriz  $W^{in} \in \mathbb{R}^{N \times K}$  e transmitidas para o reservatório de dinâmicas formado por  $N$  neurônios não lineares que representam os estados da rede,  $x(n) = [x_1(n) \dots x_N(n)]^T$ , de acordo com a expressão:

$$x(n + 1) = f\left(W^{in}u(n + 1) + Wx(n)\right) \quad (5.6.1)$$

onde  $W \in \mathbb{R}^{N \times N}$  contém os pesos das conexões recorrentes do reservatório e  $f(\cdot) = [f_1(\cdot) \dots f_n(\cdot)]$  representa as funções de ativação das unidades internas.

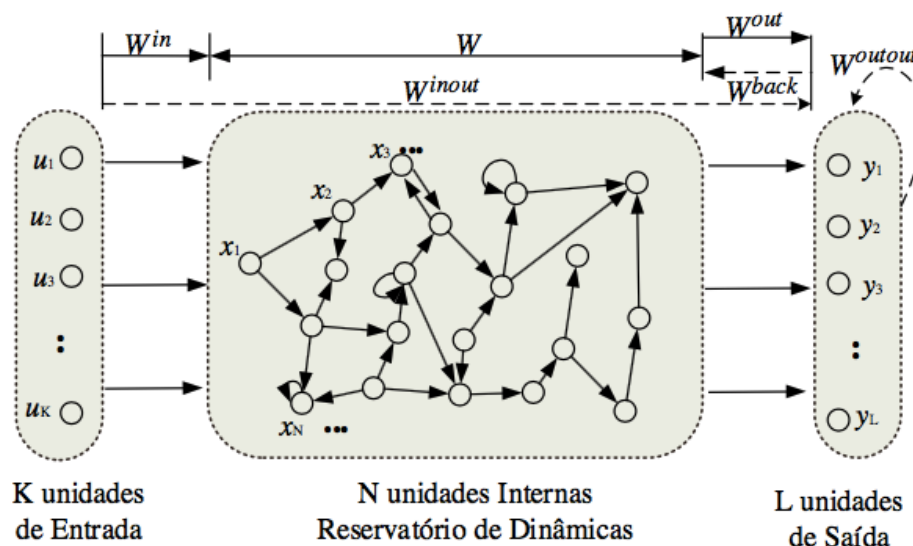
Os sinais de saída são lidos através de algum mecanismo, normalmente uma simples combinação linear dos sinais *Reservoir*.

Representadas pelo vetor  $y(n) = [y_1(n) \dots y_L(n)]^T$  as saídas são determinadas pela expressão:

$$y(n+1) = W^{out}x(n+1) \quad (5.6.2)$$

onde  $W^{out} \in \mathbb{R}^{L \times N}$  é a matriz dos pesos de saída.

FIGURA 9 – ESTRUTURA DE UMA ESN



Os estados da rede ESN são independentes da condição de entrada, sendo assim o estado da rede passa a ser governado pelo histórico recente. Quando esta propriedade é satisfeita, o efeito dos estados iniciais desaparece e a dinâmica do reservatório depende exclusivamente do histórico de entrada, de modo que a rede possui estados de eco.

O *reservoir computing* (RC) foi uma unificação de ideias das linhas de pesquisa de técnicas RNN desenvolvidas independentemente. Dentre as técnicas unificadas estão a ESN, Máquinas de estado líquido (*Liquid State Machine*).

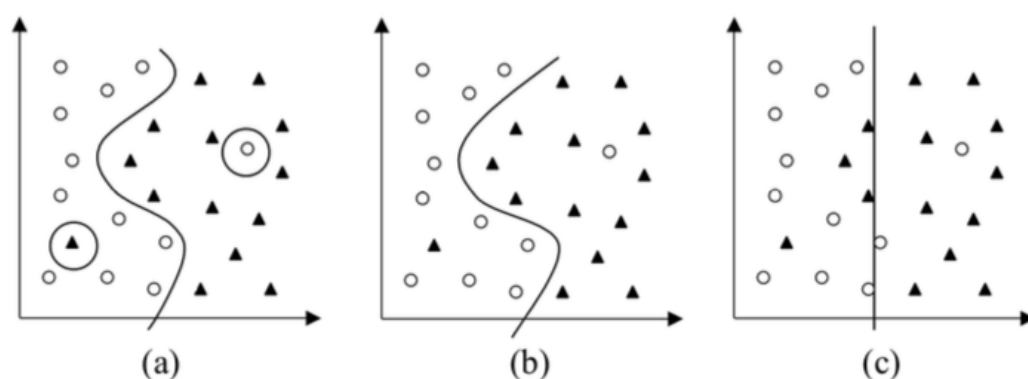
## 5.7 MÁQUINA DE VETOR DE SUPORTE (SVM)

As máquinas de vetores-suporte representam um procedimento de aprendizado não-paramétrica de aprendizado de máquina (VAPNIK, 1995), capazes de manipular eficientemente a flexibilidade dos modelos de classificação e regressão em espaços de elevada dimensão.

O método é baseado na teoria do aprendizado estatístico o qual busca encontrar um bom classificador levando em consideração todo o conjunto de dados, porém sem levar em conta casos particulares.

Considerando um o conjunto de treinamento da Figura 10. O objetivo é separar círculos de triângulos, e a função de classificação representa a linha que separa as classes, ou também conhecidas como fronteiras de decisão, ou hiperplano.

FIGURA 10 - CONJUNTO DE TREINAMENTO E DIFERENTES HIPÓTESES



A Figura 10(a) ilustra a classificação é correta, embora apresente dois ruídos, este caso representa a ocorrência de um super-ajustamento (*overfitting*) do modelo, por ser muito específico. Na Figura 10(c) o classificador desconsidera pontos muito próximos o que resulta em um erro elevado. Já Figura 10(b) seria um caso intermediário, pois não atenta a casos específicos e procura separar a maioria dos casos.

Partindo do princípio que os dados são gerados de forma independente e seguindo a distribuição de probabilidade  $P(x,y)$ , o erro esperado de um classificador  $f$  pode ser quantificado por meio da equação 5.7.1, tal que



$$R(f) = \int c(f(x), y) dP(x, y) \quad 5.7.1$$

Embora a minimização do risco não garanta uma boa generalização, pelo fato da distribuição da probabilidade ser desconhecida, o risco empírico, equação 5.7.1, mede o desempenho do classificador o que limita as funções e previne modelos ruins.

$$R_{emp}(f) = \frac{1}{n} \sum_{i=1}^n c(f(x_i), y_i) \quad 5.7.2$$

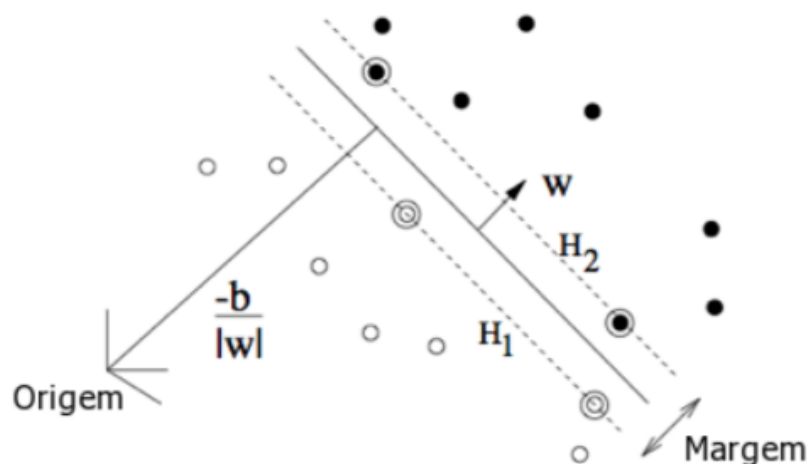
Essas limitações utilizam o conceito de dimensão VC (Vapnik-Chervonenkis), o que segundo (HAYKIN,1999) é o número máximo de exemplos de treinamento que pode ser aprendido pela máquina sem erro.

As SVM lineares separam os dados por meio de um hiperplano apresentado na equação

$$f(x) = w \cdot x + b = 0 \quad 5.7.3$$

Na Figura 11 o hiperplano ótimo maximiza as margens entre os vetores de suporte, e os pontos circulos, estão às margens  $H_1$  e  $H_2$ , ambos com  $w$  e distância perpendicular a origem.

FIGURA 11 - HIPERPLANO ÓTIMO DE SEPARAÇÃO DE CLASSES



Como não existem vetores entre as margens o objetivo passa a procurar o hiperplano ótimo, que tenha máxima distância entre as classes através de:

$$\text{Minimizar}_{w,b} \frac{1}{2} \|w\|^2 \quad 5.7.4$$

A SVM pode também ser utilizada em problemas de regressão e identificação de sistemas gerando então o vetor de suporte de regressão (*Support Vector Regression, SVR*).

Para casos de regressão as margens permitem alguns erros, então utiliza-se a equação de minimização a seguir:

$$\text{Minimizar}_{w,b} \frac{1}{2} \|w\|^2 + C \left( \sum_{i=1}^l \xi_i + \xi_i^* \right) \quad 5.7.5$$

O objetivo do SVM é construir um conjunto de hiperplanos tendo como estratégia a variação da dimensão de Vapnik-Chervonenkis, de modo que o risco empírico e a mencionada dimensão sejam minimizados ao mesmo tempo (CHAGAS, 2009). Consequentemente, a SVM pode produzir um bom desempenho de generalização para problemas de modelagem não-paramétrica e têm sido empregada com sucesso em problemas de classificação e regressão por diversos pesquisadores. Recentemente várias

abordagens de SVM têm sido propostas com sucesso para modelagem de SOC, tais como HANSEN e WANG (2005), ANTÓN et al. (2013), SURENDAR et al. (2015), LIN et al. (2017) e DENG et al. (2017).

O algoritmo de SVM utilizado no trabalho foi o LS-SVM (*Least Squares – Support Vector Machine*) proposto por SUYKENS et al. (2002). Segundo LI e YANG (2010) o LS-SVM é uma ferramenta valiosa para reconhecimento de padrões, em regressão e em problemas de predição, ele fornece funções de aproximação não linear através do mapeamento dos vetores de entrada em espaços característicos de alta dimensão.

## 5.8 PARAMETRIZAÇÃO DOS MÉTODOS

A seguir são apresentados a descrição dos parâmetros utilizados para cada método estudado. Parte dos algoritmos de ML adotados possuem parâmetros cujos valores devem ser especificados pelo usuário. Esses valores para os parâmetros livres, em geral, influenciam diretamente o desempenho de modelos matemáticos induzidos.

### 5.8.1 *Bagging*

Para o método *Bagging* adotou-se um *ensemble* com 10, 20, 50 e 100 árvores de decisão com substituições de ramificações denominadas aqui de Bagging01, Bagging02, Bagging03 e Bagging04.

### 5.8.2 *Boosting*

Para a modelagem do sistema fazendo uso do *Boosting*, fez-se uso nesta dissertação do LSBoost (*Least Squares Boosting*) (FREIDMAN, 2001) composto de um *ensemble* de 10, 20, 50 e 100 árvores de decisão denominados aqui de LSBoost01, LSBoost02, LSBoost03 e LSBoost04, respectivamente. LSBoost é um dos métodos de CM disponíveis em Matlab® para problemas de regressão e áreas afins (Statistics and Machine Learning Toolbox: User's Guide, 2016).

### 5.8.3 SVM

Foram adotadas quatro abordagens de LS-SVM e denotadas por LS-SVM01, LS-SVM02, LS-SVM03 e LS-SVM04. A sintonia do método LS-SVM01 foi realizada por meio da função de sintonia *tuneLssvm* existente na caixa de ferramentas (*toolbox*) disponível para ambiente Matlab e vinculada ao livro SUYKENS et al. (2002). A função realiza a otimização do parâmetro de regularização  $\gamma$  e da largura de banda da função *kernel*, neste caso função de base radial. Os valores obtidos pelo método foram 520 para  $\gamma$  e 7,3861 para a largura da banda.

Quanto aos outras configurações, adotou-se:  $\gamma = 500$  e largura de banda igual a 7 (LS-SVM02),  $\gamma = 550$  e largura de banda igual a 7 (LS-SVM03), e  $\gamma = 600$  e largura de banda igual a 7 (LS-SVM04).

#### 5.8.4 ANFIS

Para a obtenção do modelo ANFIS foram utilizadas da mesma forma que para os outros algoritmos quatro configurações. Em todos projetos de modelos ANFIS adotou-se passo inicial de 0,01 e passos de incremento e decremento iguais a 1,5 e 0,5 para os coeficientes de aprendizado, respectivamente. As quatro configurações testadas foram:

- ANFIS01: adota particionamento em grade (script *genfis1* do Matlab) e 5 épocas de treinamento;
- ANFIS02: da mesma forma que o ANFIS01 utiliza particionamento em grade, mas foi configurado para realizar 100 épocas de treinamento;
- ANFIS03: adota ajuste do modelo nebuloso interpolativo do tipo Sugeno por meio do método de agrupamento (*clustering*) de dados denominado *fuzzy c-médias* com 10 grupos;
- ANFIS04: da mesma forma que o ANFIS03 utiliza o modelo do tipo Sugeno e método *fuzzy c-médias*, no entanto, com 50 grupos.

#### 5.8.5 ESN

A identificação, do modelo estudado, pelo modelo ESN foi realizada fazendo uso de neurônios integradores gotejantes (*leaky*), para a composição do *reservoir*, e dos seguintes parâmetros ajustados por procedimento de tentativa e erro para quatro métodos configurações do ESN nomeadas como ESN01, ESN02, ESN03 e ESN04:

- Número de unidades internas = 10 (ESN01 e ESN02) e 20 (ESN03 e ESN04);
- Raio espectral = 0,5 (ESN01, ESN02, ESN03) e 0,04 (ESN04);
- Escala de entradas = 0,1 (para todas as entradas) para os quatro modelos ESN;
- Deslocamento de entrada = 0;

- Escala de aprendizado = 0,5 (ESN01 e ESN02) e 0,3 (ESN03 e ESN04);
- Deslocamento de aprendizado = 0 (ESN01 e ESN04), -0,001 (ESN02) e -0,005 (ESN03);
- Escala de retorno = 0 para os quatro métodos ESN.

## 6 ANÁLISE DOS RESULTADOS DA MODELAGEM

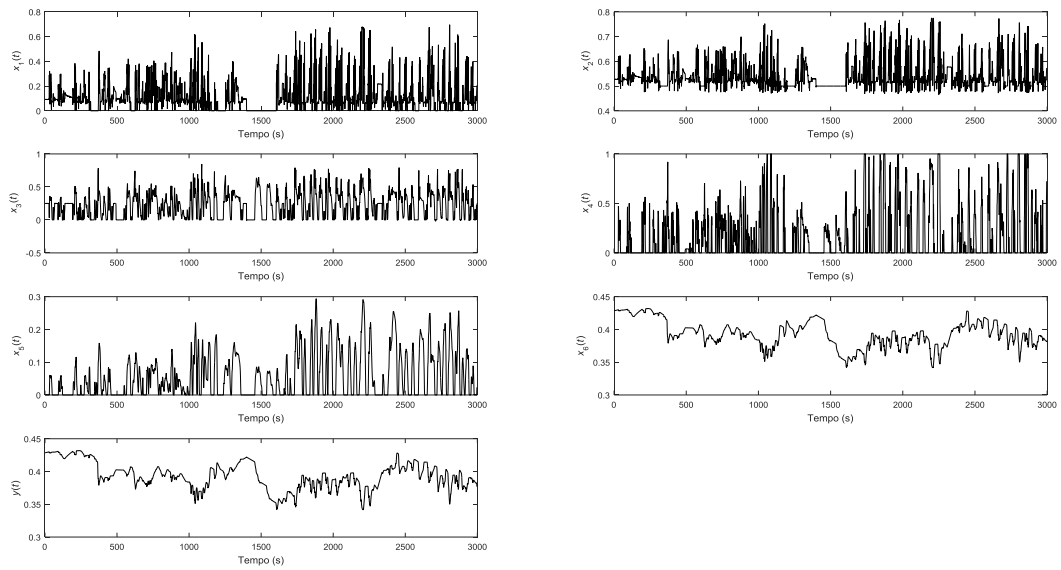
Para fins de comparação com um modelo clássico de identificação linear os métodos de ML foram comparados com um modelo matemático auto regressivo com entrada exógena (do inglês, *Auto-Regressive with eXogeneous inputs*, ARX) com ajuste por meio do métodos de mínimos quadrados (MQ) (LJUNG, 1987). Adotando-se para tal os métodos MQ01, MQ02, MQ03 e MQ04, estes métodos com modelos ARX de 1ª, 2ª, 3ª e 4ª ordens, respectivamente.

Os dados adotados (estes obtidos com período de amostragem de 1s) para identificação do sistema foram divididos em duas parcelas: a de treinamento, abrangendo 70% dos dados (amostras 1 a 2998), e a de validação, contendo as amostras restantes, ou seja, 30% (amostras 2999 a 4282), conforme apresentado na Figura 12.

Tal divisão foi realizada de forma empírica, pois não encontrou-se nenhuma fonte que determinasse qual seria uma divisão “ideal” das amostras na literatura do tema. Baseado no que geralmente é adotado na literatura na maior parte das vezes de 70-90% dos dados para treinamento e de 10-30% para validação, adotou-se tal faixa de divisão, ou seja, 70% para treinamento e 30% para validação nesta dissertação.

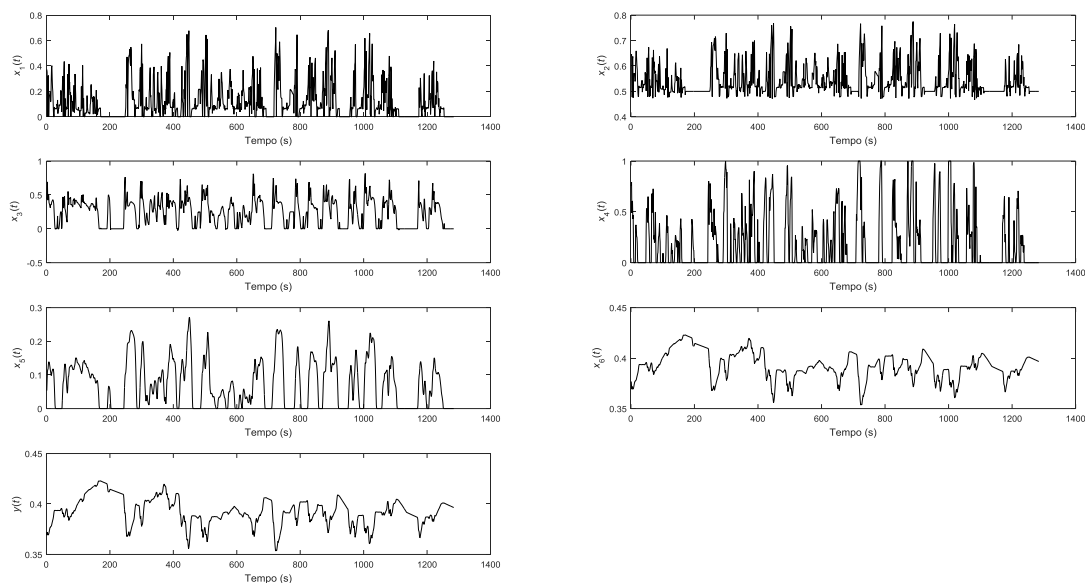
FIGURA 12 – DADOS DE ENTRADA E SAÍDA ADOTADOS NO PROCEDIMENTO DE IDENTIFICAÇÃO DE SISTEMAS.

(a) dados adotados na fase de **treinamento**





## (b) dados adotados na fase de teste



De modo a avaliar o resultado das técnicas com os valores reais, foi empregado erro quadrático médio (do inglês *Mean Squared Error*, MSE), que calcula a diferença entre o estimador e o estimado, no caso a diferença entre o valor real e o calculado. Visando complementar a avaliação, também escolheu-se o critério de coeficiente de determinação ( $R^2$ ) que determina a quantidade de variação na resposta, sendo quanto maior o valor de  $R^2$  melhor a aproximação do modelo obtido. Assim, após rodar a simulação foram obtidos os resultados apresentados na

Tabela 1. O tempo total apresentado na Tabela 1 é o tempo da fase de estimação mais tempo da fase de teste.

TABELA 1 – RESULTADO EM TERMOS DOS CRITÉRIOS DE DESEMPENHO.

Modelo	Estimação (70% dos dados)		Teste (30% dos dados)		Tempo total [s]
	MSE	$R^2$	MSE	$R^2$	
Bagging01	3,8279e-6	0,9957	4,9784e-6	0,9864	0,41
Bagging02	2,8746e-6	0,9970	3,6323e-6	0,9899	0,42
Bagging03	2,6284e-6	0,9974	2,9448e-6	0,9921	0,95
Bagging04	2,1990e-6	0,9978	2,9116e-6	0,9923	1,96
LSBoost01	1,9902e-5	0,9765	2,2875e-5	0,9468	0,72
<b>LSBoost02</b>	9,8716e-6	0,9879	9,9780e-5	0,9721	0,23
LSBoost03	3,8407e-6	0,9953	4,5232e-6	0,9878	0,46
LSBoost04	2,0234e-6	0,9975	2,7324e-6	0,9927	0,93
<b>LS-SVM01</b>	2,2005e-7	0,9997	4,5145e-7	0,9987	0,91
<b>LS-SVM02</b>	2,1680e-7	0,9997	4,5141e-7	0,9987	0,89
<b>LS-SVM03</b>	2,1471e-7	0,9997	4,5121e-7	0,9987	0,90
<b>LS-SVM04</b>	1,8537e-7	0,9998	6,2150e-7	0,9982	44,62
<b>ANFIS01</b>	1,8539e-7	0,9998	6,2151e-7	0,9982	45,19
<b>ANFIS02</b>	4,0481e-7	0,9998	5,7968e-7	0,9983	2,27
<b>ANFIS03</b>	1,9541e-7	0,9998	6,3587e-7	0,9982	38,91
<b>ANFIS04</b>	1,4786e-7	0,9998	9,5231e-7	0,9973	152,33
ESN01	7,1077e-7	0,9991	9,2845e-7	0,9974	0,76
ESN02	7,0178e-7	0,9991	9,4171e-7	0,9973	0,74
ESN03	7,5329e-7	0,9991	1,3031e-6	0,9963	0,89
ESN04	7,4545e-7	0,9991	9,7493e-7	0,9972	1,13
MQ01	-	0	-	0	3,21
MQ02	4,5458e-4	0,6394	2,1929e-4	0,5391	0,46
MQ03	4,8962e-4	0,6188	2,4410e-4	0,5260	0,31
MQ04	-	0	-	0	0,34

Analisando-se os resultados resumidos na

Tabela 1, a abordagem ANFIS produziu o menor erro quadrático seguido de SVM e ESN com os respectivos valores de 0,0006215, 0,0045145 e 0,0092108. Em relação à melhor aproximação ao modelo, por meio dos resultados para  $R^2$ , temos na ordem: SVM, ANFIS e ESN com os respectivos valores de 0,99873 , 0,99823 e 0,99740 (99,87%, 99,82% e 99,74%). Entretanto, conforme citado, busca-se além do melhor resultado, também o menor custo computacional, medido pelo tempo de treinamento e validação. Levando-se em conta esta observação, o melhor método foi SVM para treinamento e *LSBoost* para validação.

Através da relação desempenho versus tempo de processamento, pode-se escolher como sendo o melhor método o SVM tendo como segunda opção ESN. Ambos próximos quanto à aproximação ao modelo (99,8 e 99,7%, respectivamente) e tempos totais de processamento (tempo de treinamento mais tempo de validação) de 1,86773 s e 2,5906 s, respectivamente. ANFIS apresentou também resultados semelhantes à SVM, porém seu tempo total e processamento de 77,71861 segundos é maior do que as duas técnicas escolhidas. Se o desempenho do método em termos de processamento não for requerida então ANFIS seria também uma opção.

A seguir, são apresentados os resultados de previsão dos modelos de previsão testados para os dados de teste visando previsão de um passo à frente (previsão de curtíssimo prazo).

As Figuras 13 a 16 referem-se aos resultados de *Bagging*. Em A são comparados os dados do SOC disponível na rede CAN como o valor estimado. Nota-se que as curvas da saída real e a estimada por *Bagging*, seguem juntas em quase toda a extensão da curva. Ao analisar as Figuras 13(b) a 16(b) e relacionar aos resultados do MSE, apenas nos dois últimos casos valores inferiores a  $2,95e-6$  são encontrados, embora tenham demorado pelo menos o dobro do tempo gasto por *Bagging01* e *Bagging02*.

FIGURA 13 – RESULTADOS DE PREVISÃO OBTIDOS POR BAGGING01

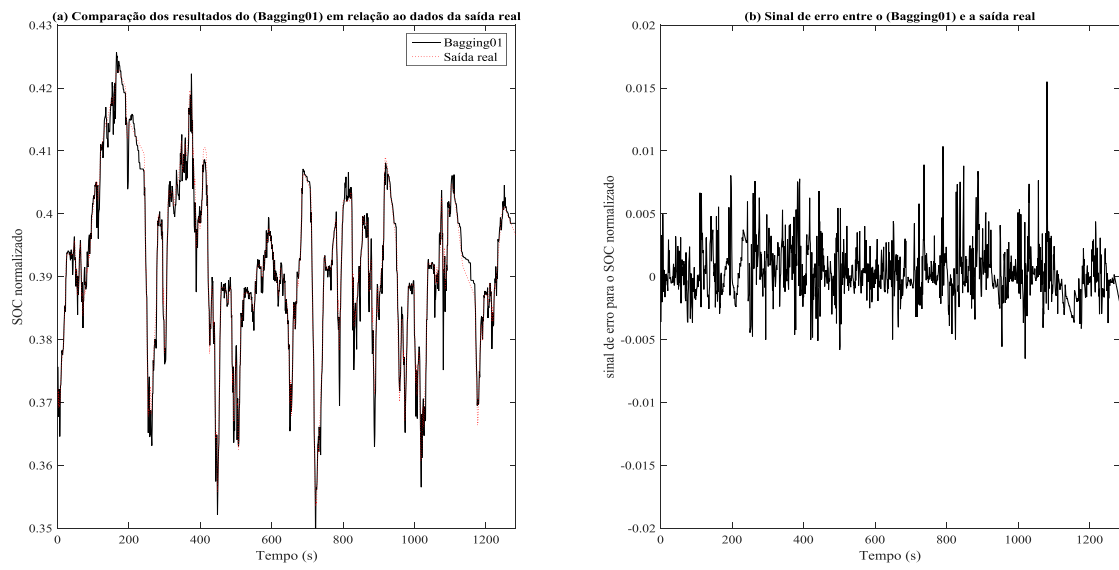


FIGURA 14 – RESULTADOS DE PREVISÃO OBTIDOS POR BAGGING02.

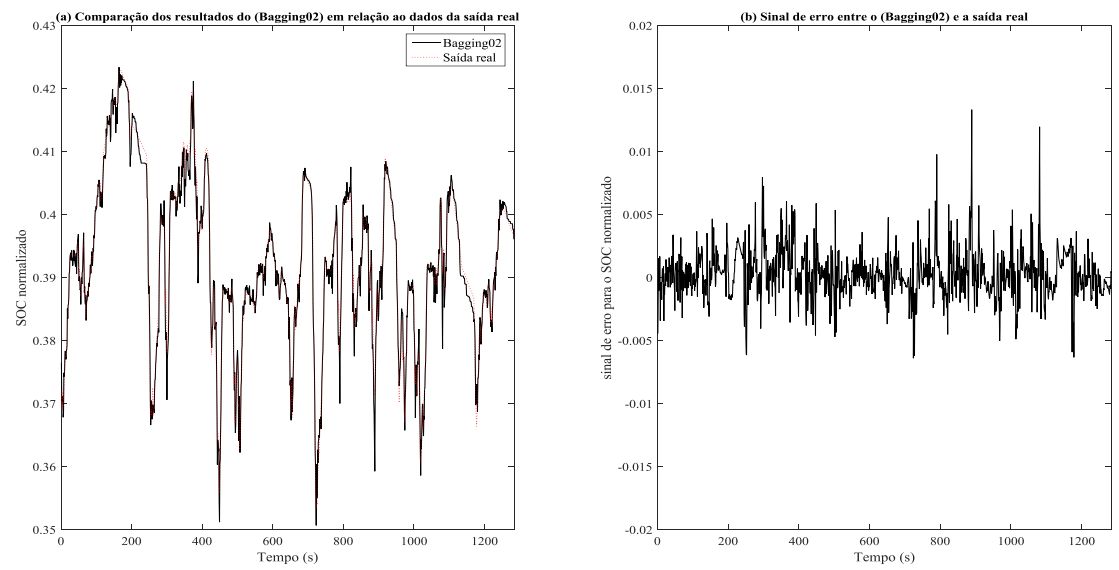


FIGURA 15 – RESULTADOS DE PREVISÃO OBTIDOS POR BAGGING03.

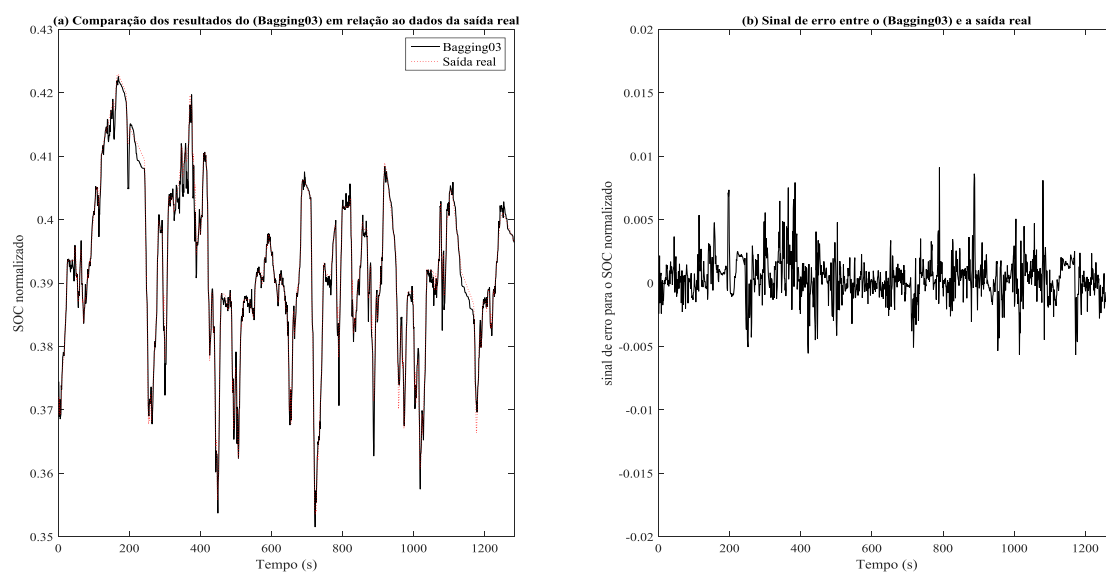
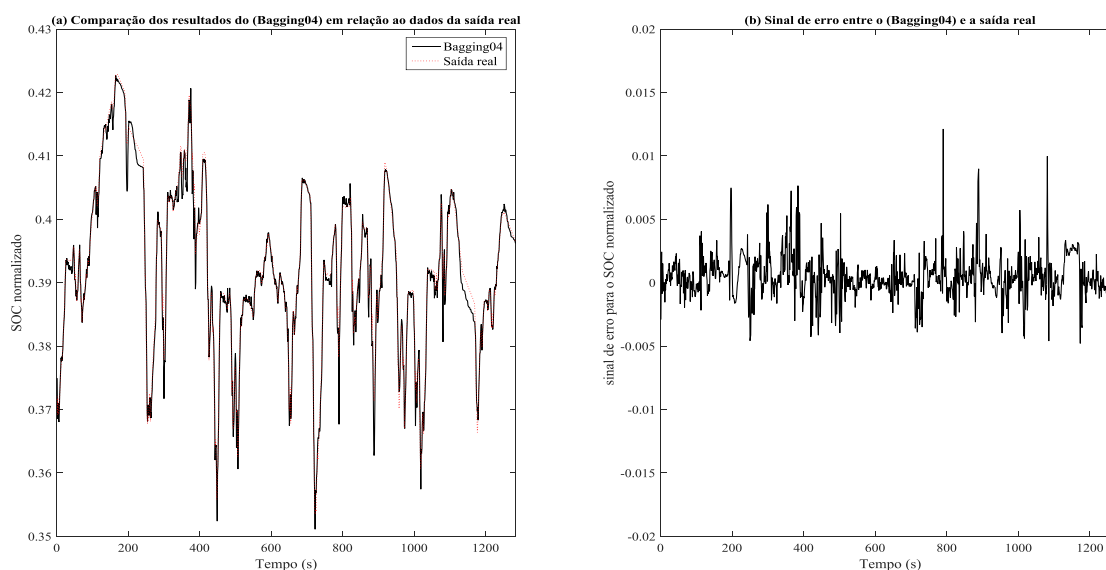


FIGURA 16 – RESULTADOS DE PREVISÃO OBTIDOS POR BAGGING04.



As Figuras 17 a 20 referem-se aos resultados de *Boosting*. Da mesma forma que a o caso anterior, em A são comparados os resultados com os dados do SOC disponível na rede CAN como o valor estimado. Neste caso é mais notável a diferença entre curvas pelo fato da onda quadrada no caso do LSboost, e nota-se a melhora da aproximação ao longo do LSboost01 ao LSboost04. Quanto as Figuras 17(b) a 20(b) o erro possui maior amplitude nos primeiros casos, melhorando gradativamente. Esses fatos são comprovados com os dados observados na Tabela 1 em que LSboost01 apresenta MSE igual a  $2,28e-5$ , e em LSboost04, MSE igual a  $2,73e-6$ . Quanto ao tempo o

LSboost04 apresentou o pior resultado da série, embora uma melhora significativa em termos de  $R^2$  e MSE.

FIGURA 17 – RESULTADOS DE PREVISÃO OBTIDOS POR BOOSTING01.

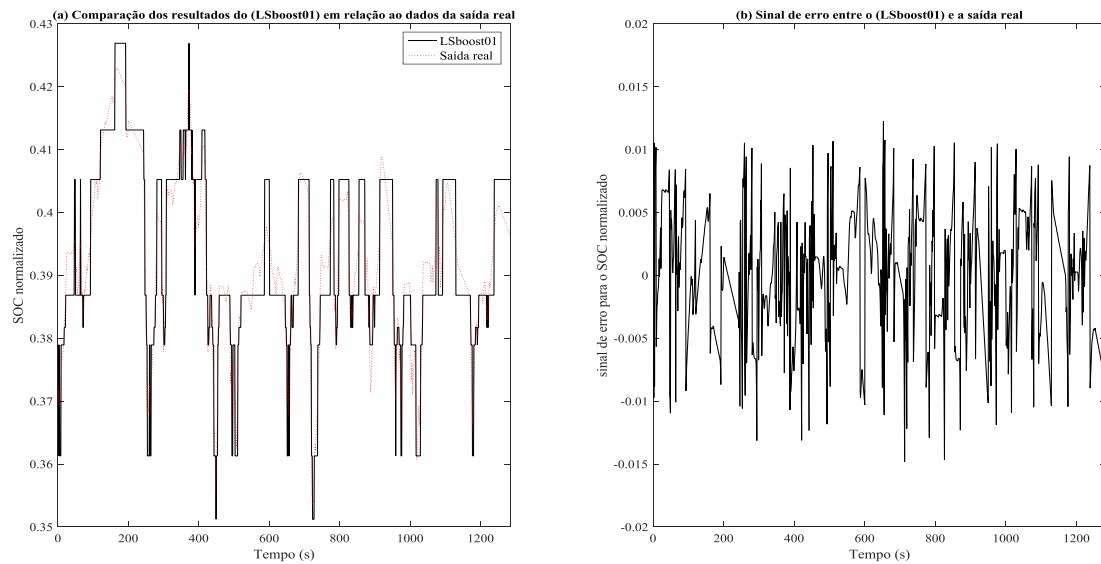


FIGURA 18 – RESULTADOS DE PREVISÃO OBTIDOS POR BOOSTING02.

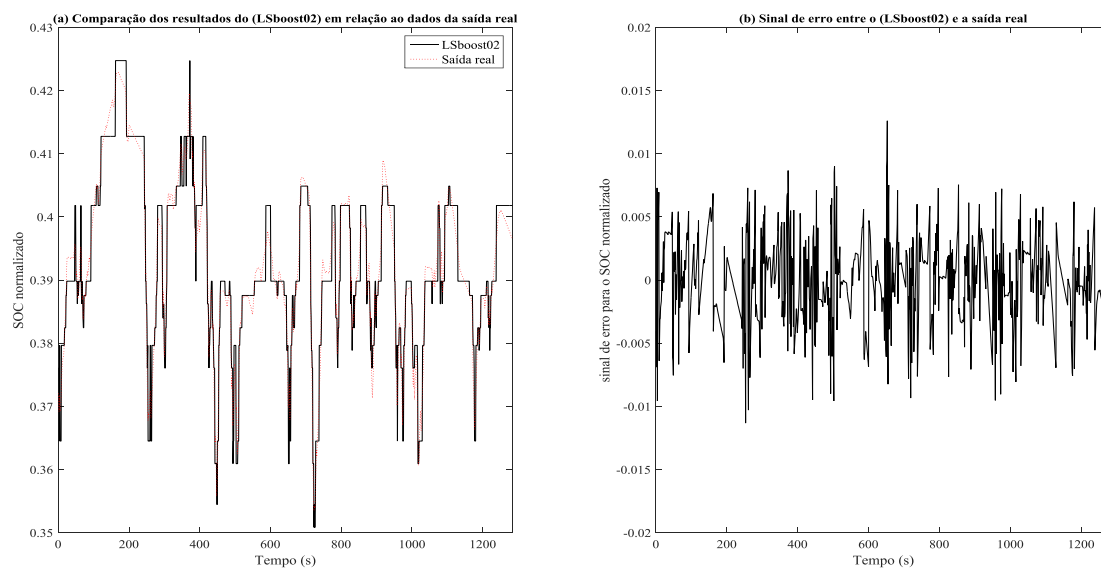


FIGURA 19 – RESULTADOS DE PREVISÃO OBTIDOS POR BOOSTING03.

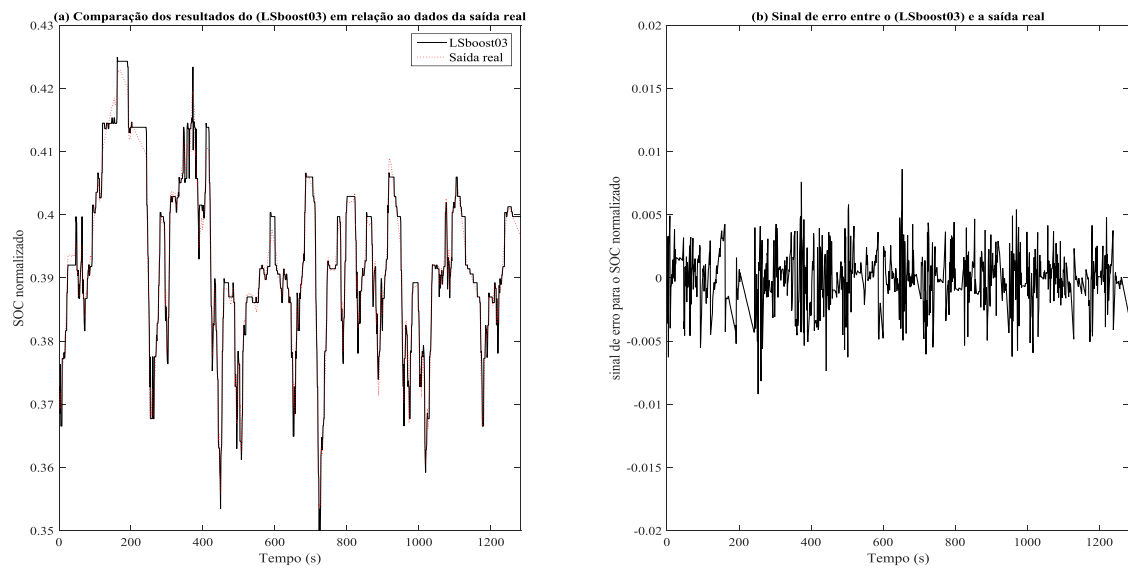
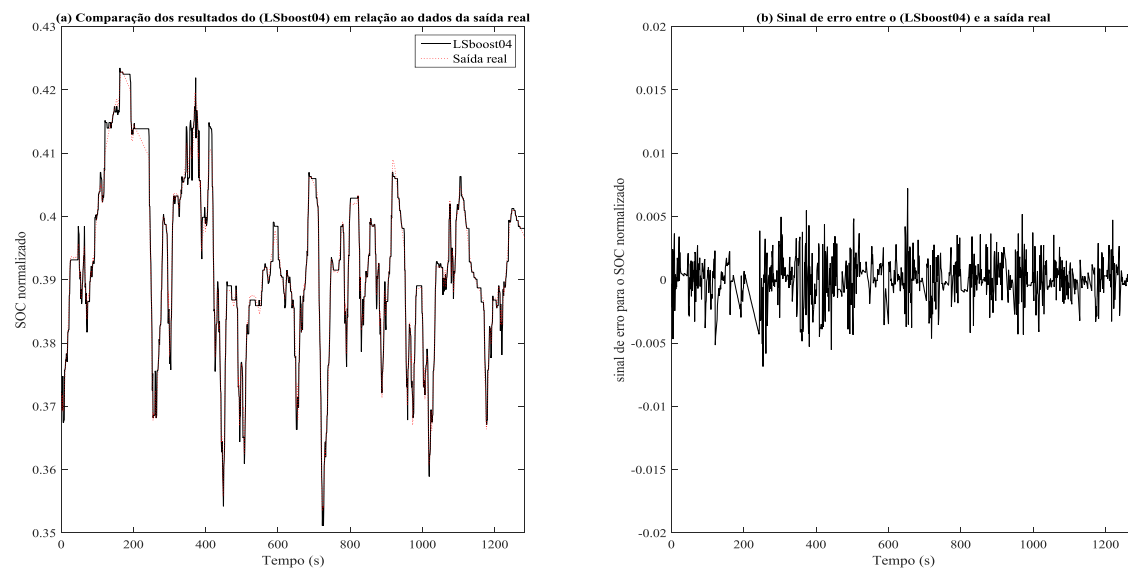


FIGURA 20 – RESULTADOS DE PREVISÃO OBTIDOS POR BOOSTING04.





As Figuras 21 a 24 referem-se aos resultados de *LS-SVM*. Nas Figuras b, ou seja, Figuras 21(b) a 24(b), onde compara-se os resultados com os dados do SOC disponível na rede CAN como o valor estimado, a diferença entre curvas é apenas perceptível nos picos em que a saída é um pouco mais evidente, e Figuras 21(b) a 24(b) o erro possui baixa amplitude em todos os casos testados. Quanto ao MSE com exceção do último caso o valor é igual a  $4,51e-7$  com tempos de processamento em torno dos 0,90 s. Já o *LS-SVM04* embora tenha apresentado um MSE menor, com  $6,21e-7$ , o tempo de processamento significativo quando comparado com os casos anteriores de *LS-SVM*, resultando em 44,62 s.

FIGURA 21 – RESULTADOS DE PREVISÃO OBTIDOS POR *LS-SVM01*.

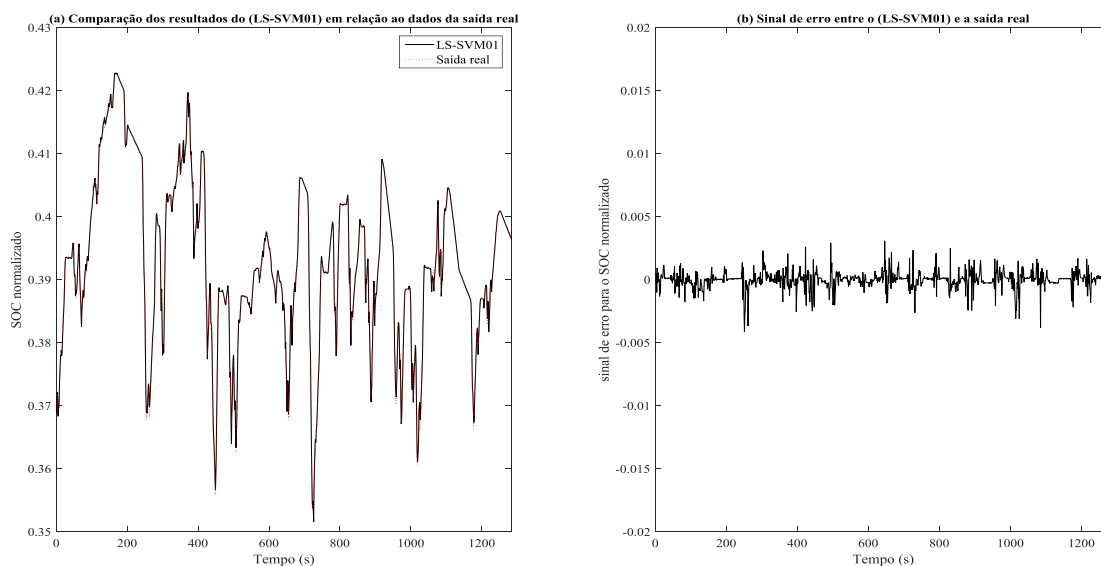


FIGURA 22 – RESULTADOS DE PREVISÃO OBTIDOS POR LS-SVM02.

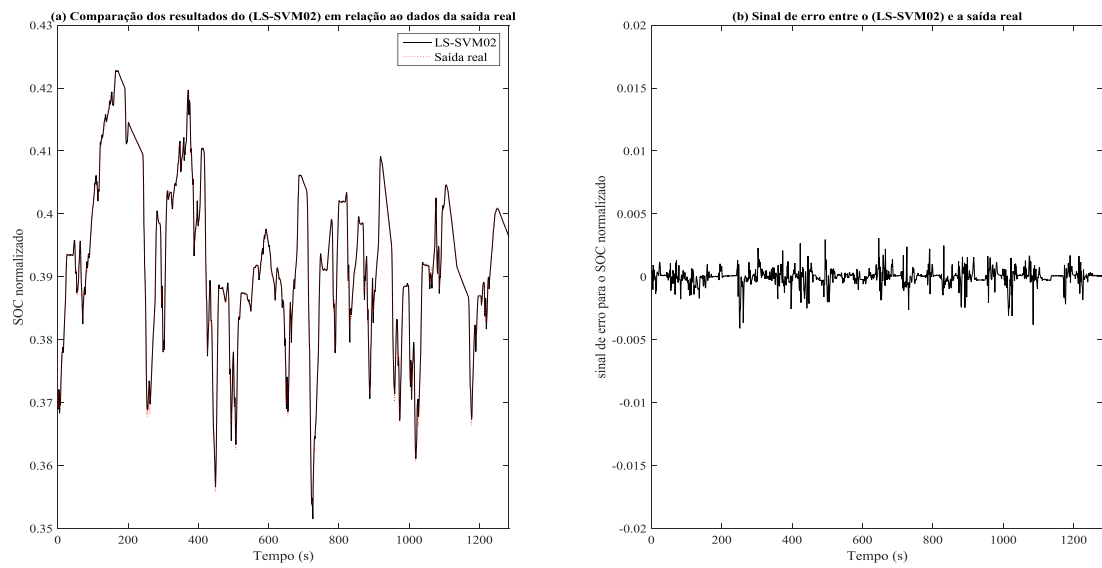


FIGURA 23 – RESULTADOS DE PREVISÃO OBTIDOS POR LS-SVM03.

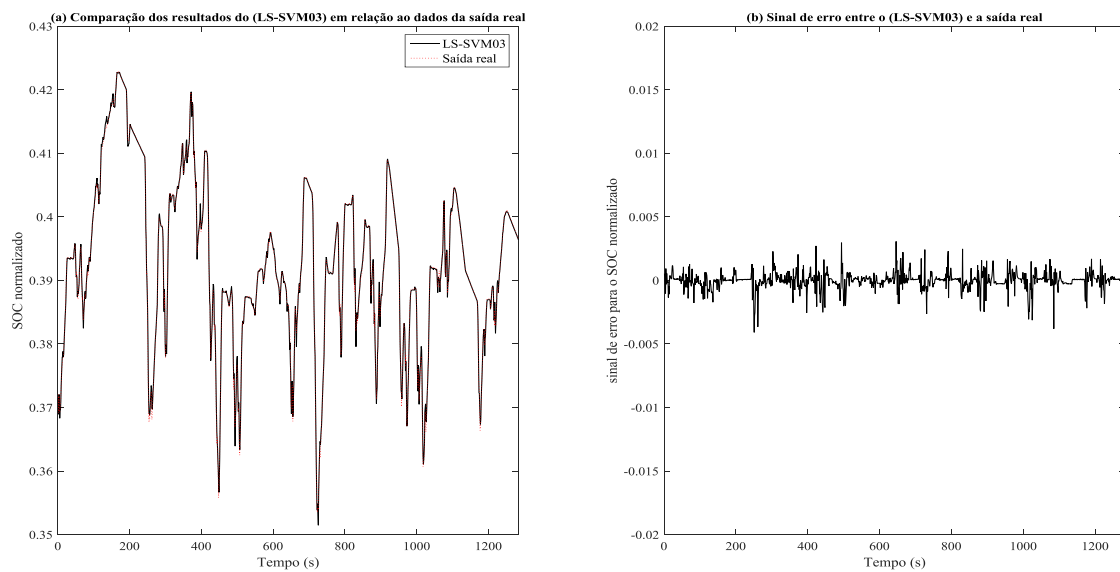
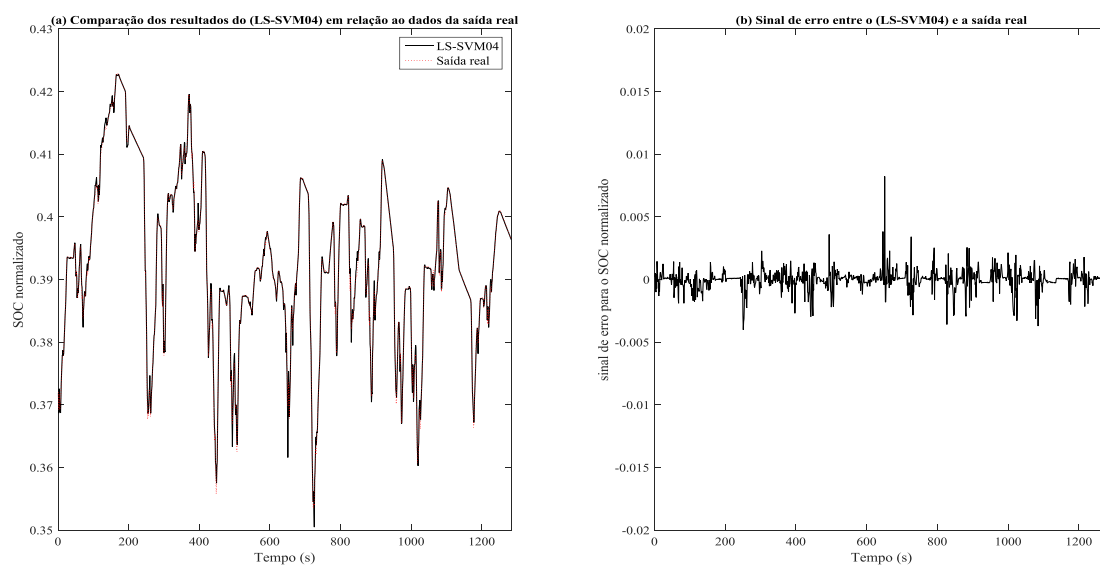


FIGURA 24 – RESULTADOS DE PREVISÃO OBTIDOS POR LS-SVM04.



As Figuras 25 a 28 referem-se aos resultados de *ANFIS*. Da mesma forma que as curvas apresentadas em LS-SVM, os resultados de *ANFIS* apresentam boa aproximação nas Figuras 25(b) a 28(b), onde compara-se os resultados com os dados do SOC disponível na rede CAN como o valor estimado, e em B o erro também possui baixa amplitude em todos os casos testados. Segundo os dados apresentados na Tabela1 os valores de MSE, *ANFIS01* e *ANFIS03* obtiveram  $6,21e-7$  e  $6,35e-7$  com tempo de processamento similares de 45,19 e 38,91 s, respectivamente. Para o caso do *ANFIS02* o MSE se mostrou um pouco pior quando comparado aos anteriores, com  $5,79e-7$  embora seu tempo de processamento ficou bem abaixo dos métodos dessa série em 2,27 s. Por outro lado o *ANFIS04* obteve um MSE igual a  $9,52e-7$ , ou seja um dos menores erros de todo o estudo, entretanto isso impactou de forma significativa no processamento, resultando em 152,33 segundos.

FIGURA 25 – RESULTADOS DE PREVISÃO OBTIDOS POR ANFIS01.

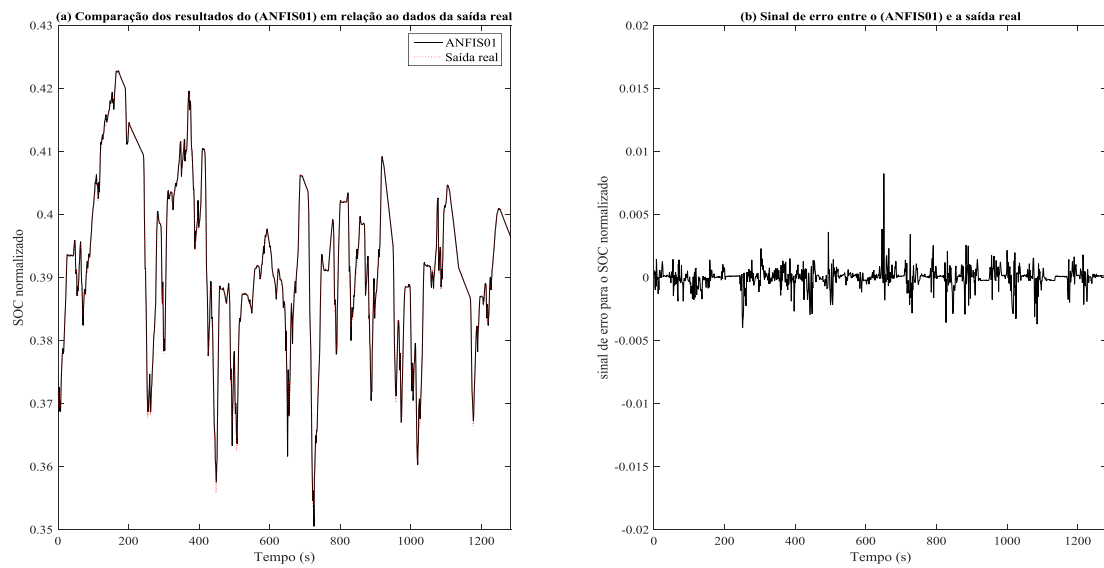


FIGURA 26 – RESULTADOS DE PREVISÃO OBTIDOS POR ANFIS02.

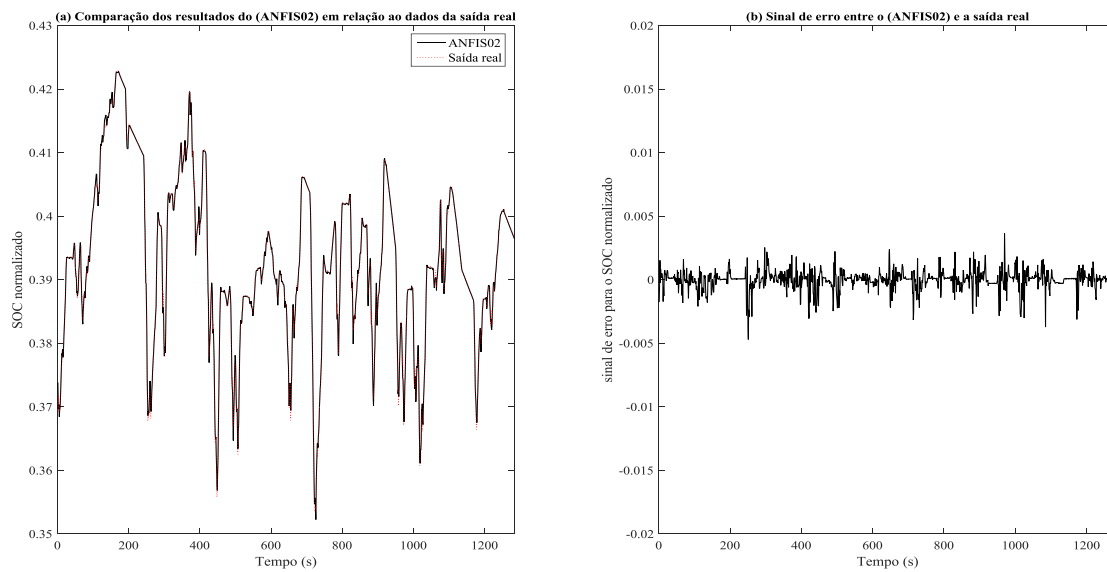


FIGURA 27 – RESULTADOS DE PREVISÃO OBTIDOS POR ANFIS03.

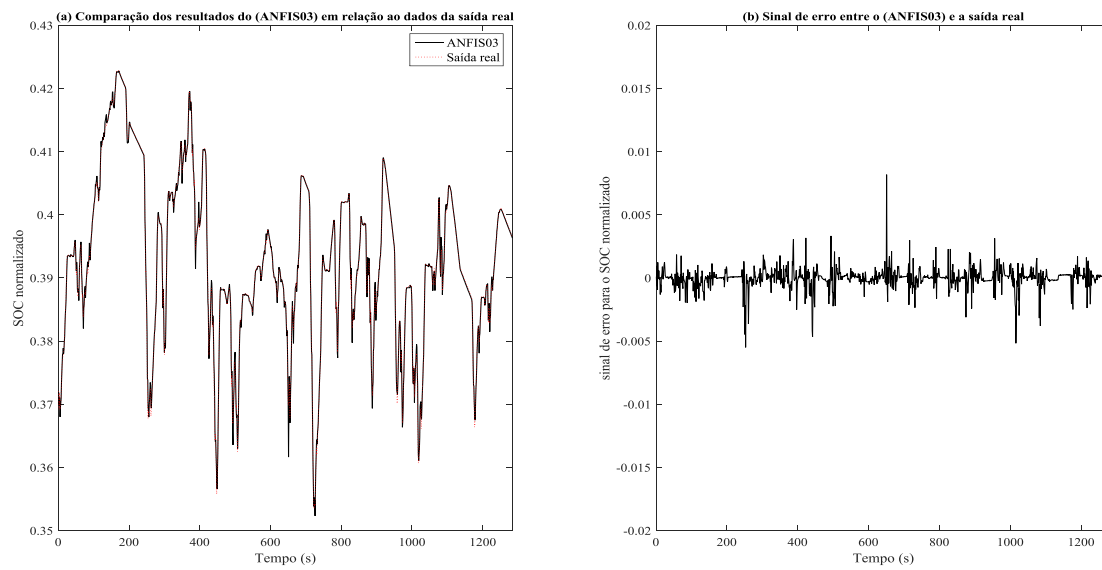
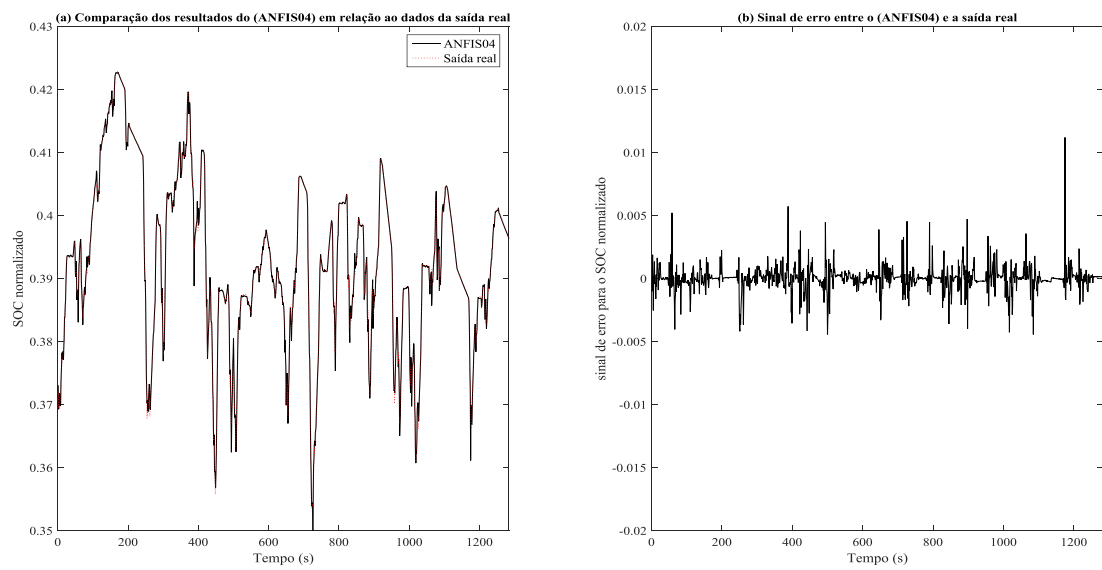


FIGURA 28 – RESULTADOS DE PREVISÃO OBTIDOS POR ANFIS04.



As Figuras 29 a 32 referem-se aos resultados de *ESN*. Nas Figuras 29(a) a 32(a) comparam-se os resultados com os dados do SOC disponível na rede CAN como o valor estimado, e nota-se que a curva do estudo acompanha a saída real. Nas Figuras 29(b) a 32(b), com exceção do ESN03 em que o MSE é igual a  $1,30e-6$ , o erro possui baixa amplitude nos outros casos, com MSE próximos a  $9,5e-7$ . Quanto ao tempo total de processamento o primeiro e segundo caso, se mostraram similares com um tempo médio de 0,75 segundos. Para o ESN03, o qual teve o pior desempenho da série, obteve 0,89 s, enquanto o ESN04 demandou 1,13 s.

FIGURA 29 – RESULTADOS DE PREVISÃO OBTIDOS POR ESN01.

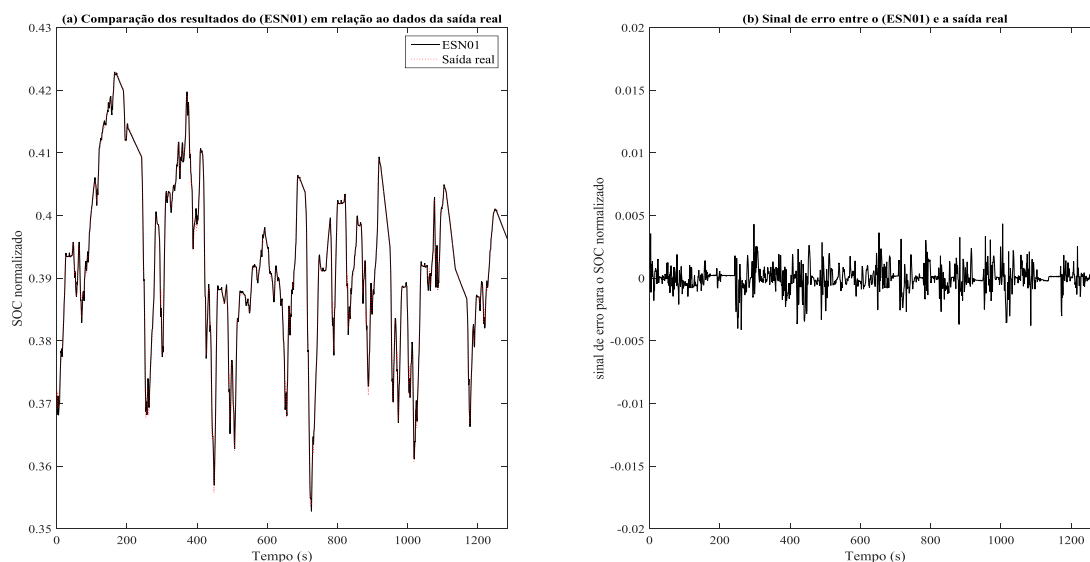


FIGURA 30 – RESULTADOS DE PREVISÃO OBTIDOS POR ESN02.

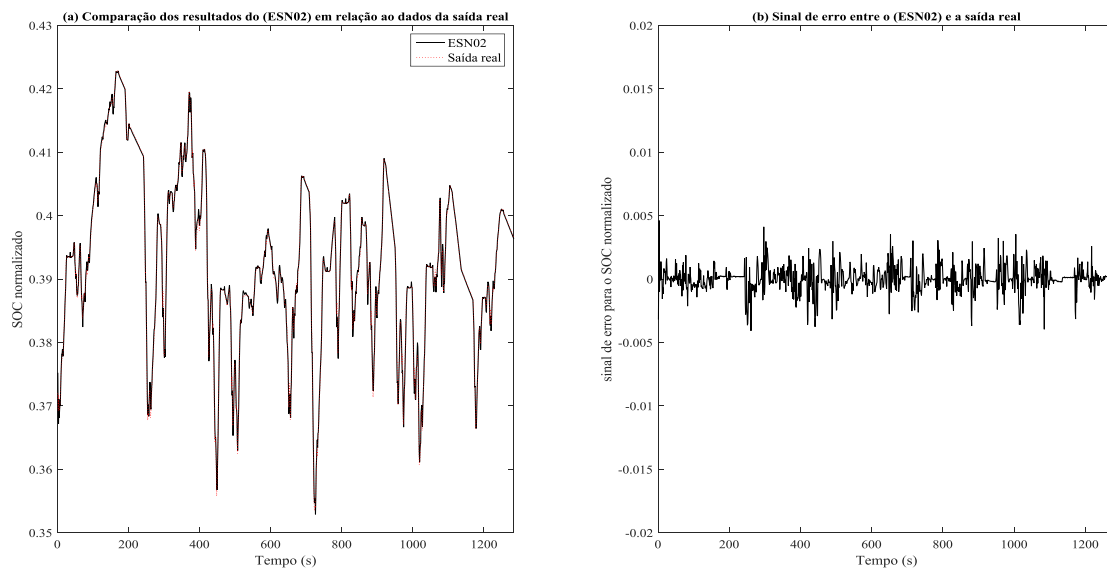


FIGURA 31 – RESULTADOS DE PREVISÃO OBTIDOS POR ESN03.

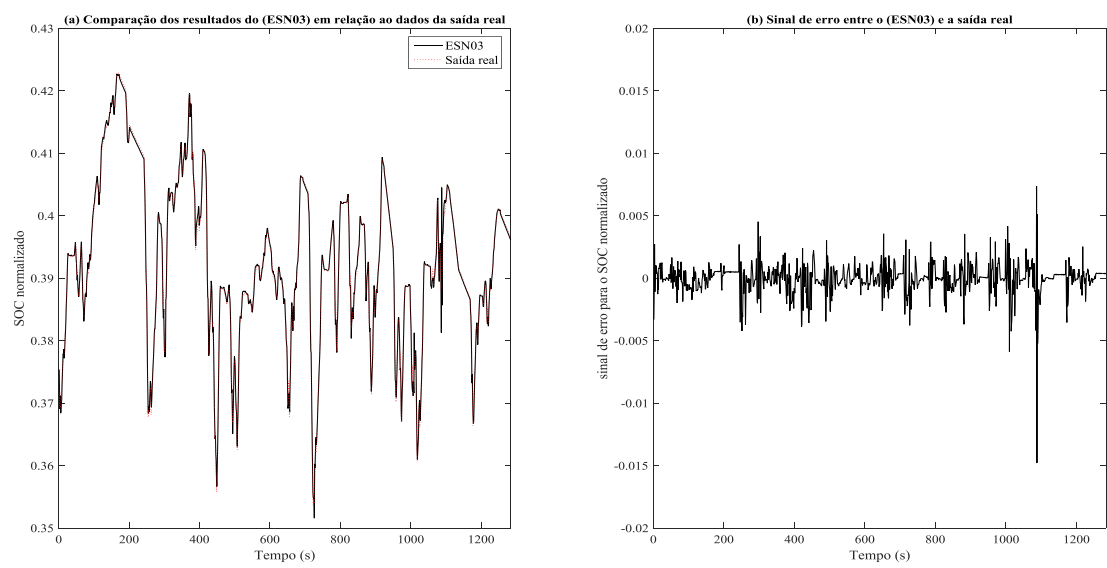
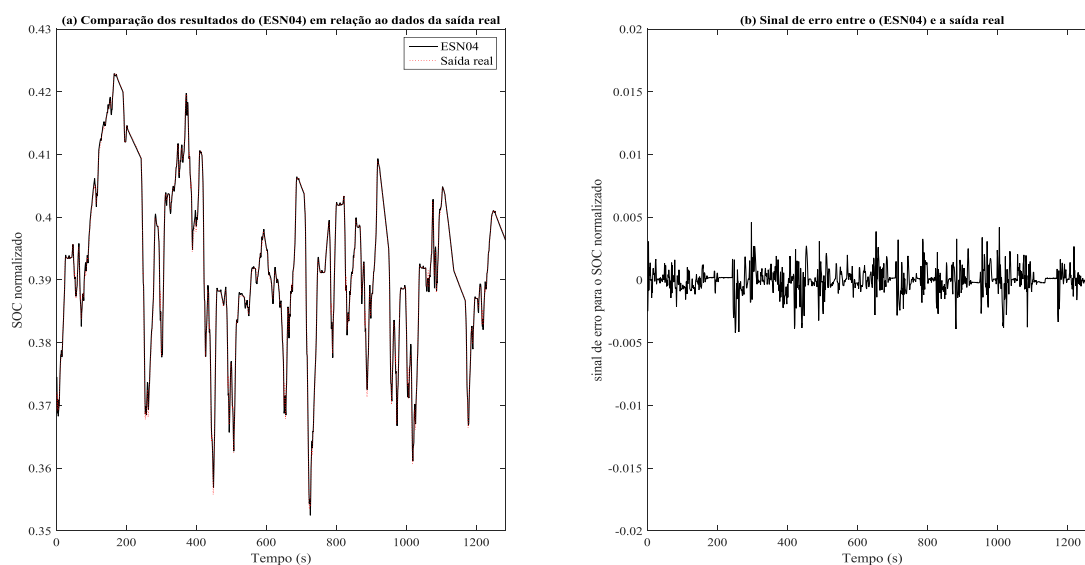


FIGURA 32 – RESULTADOS DE PREVISÃO OBTIDOS POR ESN04.



De forma a complementar o estudo, para as Figuras 32 a 36 comparam para o método dos mínimos quadrados da mesma forma utilizada anteriormente, afim de verificar se tal problema poderia ser solucionado por um método mais simples.

Ao observar as curvas nas Figuras 32(a) a 36(a), nota-se que para MQ01 e MQ04 o método não conseguiu solucionar o problema, pois a curva divergiu do objetivo. Mesmo assim MQ02, e MQ03 obtiveram o MSE em aproximadamente  $2,30e-4$ .

Comparado às técnicas anteriores o MSE apresenta uma baixa eficiência, desta forma para este tipo de problema não linear, este método não é o mais adequado pela abordagem de identificação adotada.



FIGURA 33 – RESULTADOS DE PREVISÃO OBTIDOS POR MQ01.

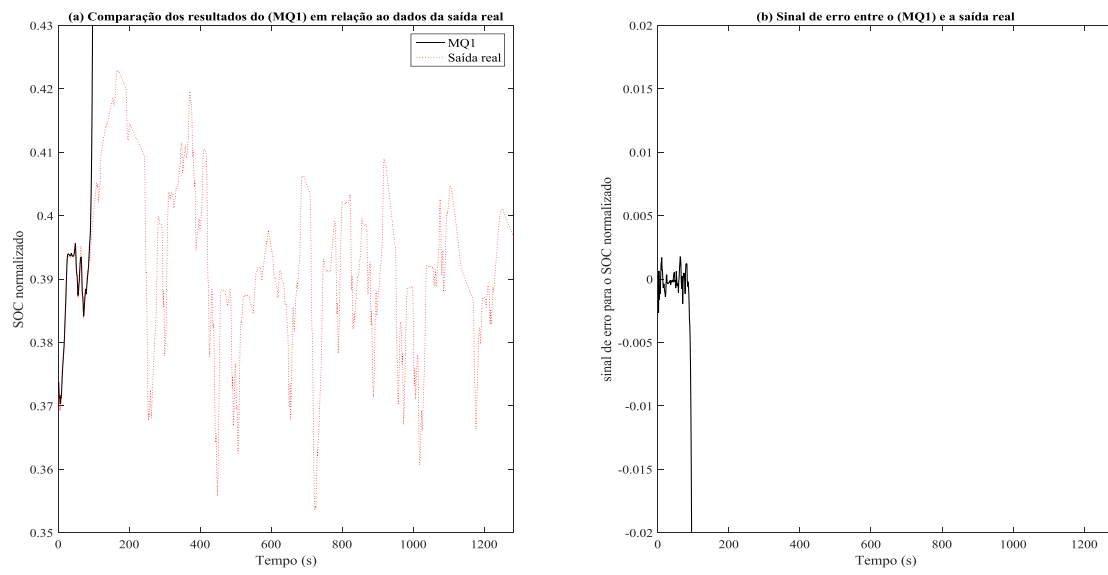


FIGURA 34 – RESULTADOS DE PREVISÃO OBTIDOS POR MQ02.

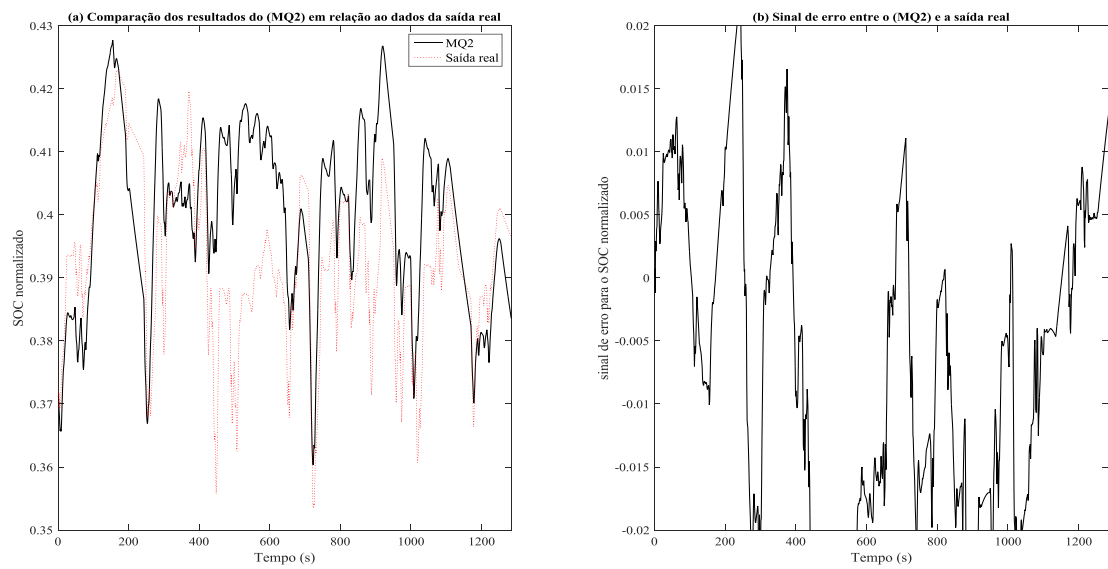


FIGURA 35 – RESULTADOS DE PREVISÃO OBTIDOS POR MQ03.

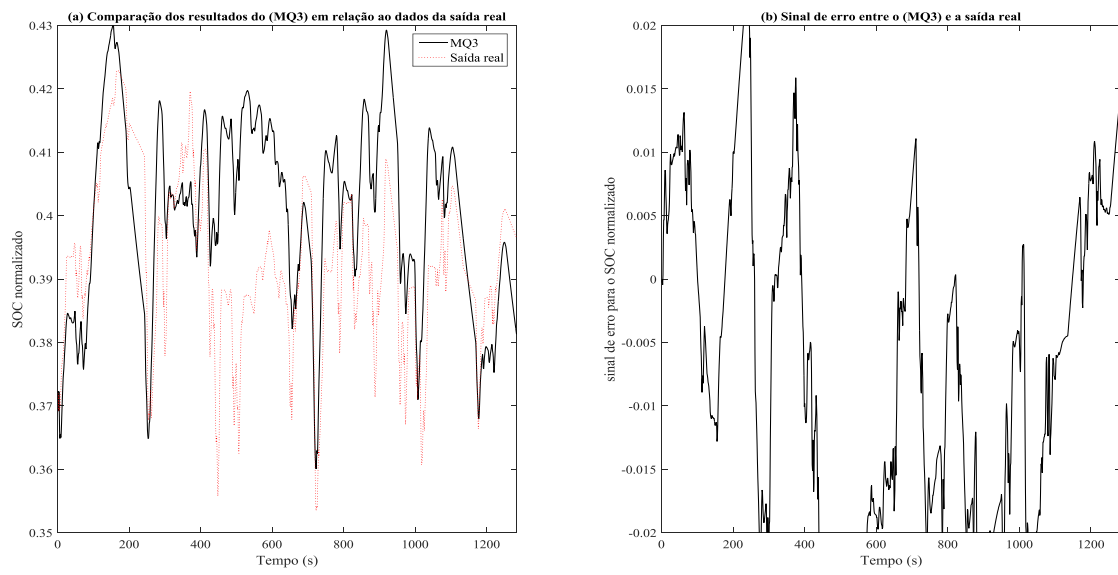
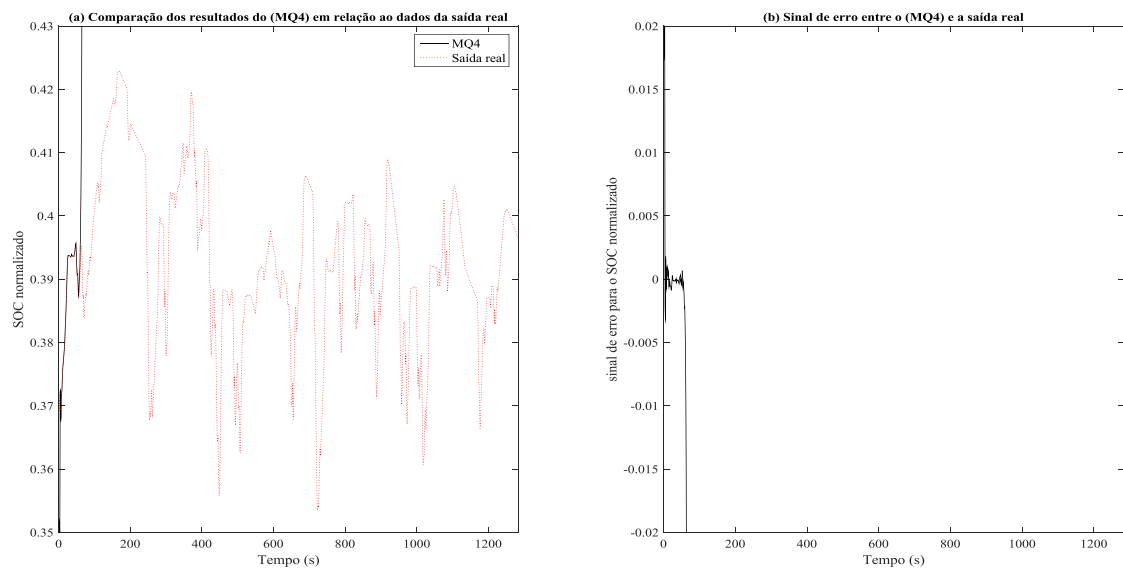


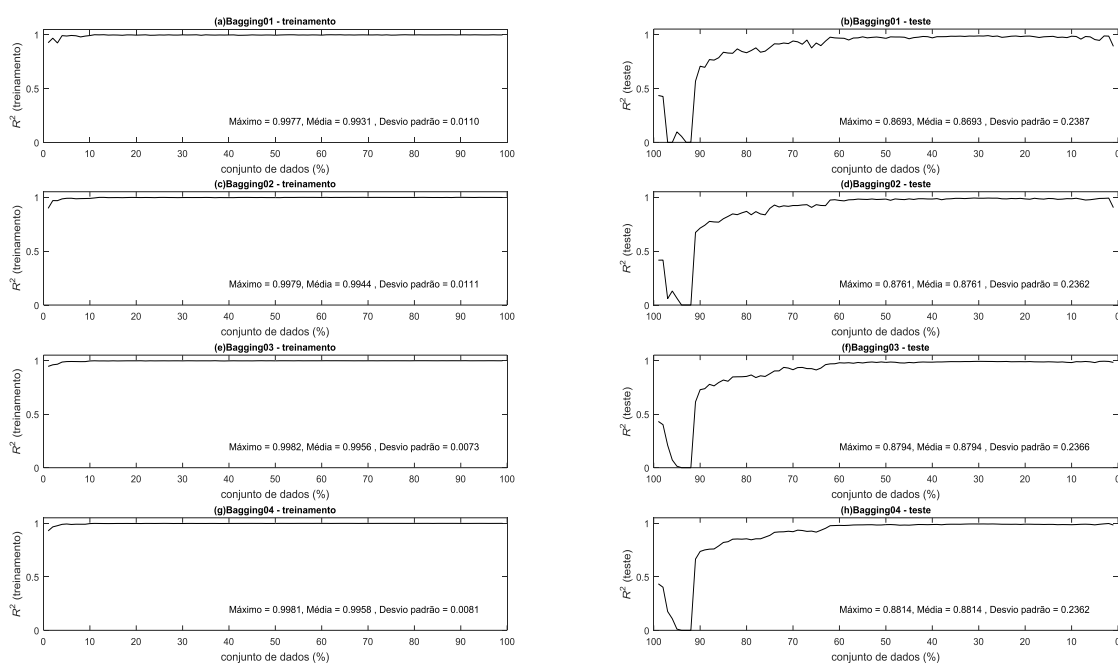
FIGURA 36 – RESULTADOS DE PREVISÃO OBTIDOS POR MQ04.



Com intuito de avaliar a sensibilidade dos métodos de modelagem quanto ao efeito da divisão do conjunto de dados de treinamento e de testes, nas Figuras 37 a 42 são apresentados os resultados para diferentes divisões estas variando de 1% a 99% (tamanho de passo de 1%) para os conjuntos de treinamento e testes em termos do índice  $R^2$ . No caso, lembrando que quanto mais próximo do valor unitário está o  $R^2$  melhor é a modelagem.

Nos resultados obtidos para o modelo *Bagging* de acordo com a Tabela 1 os dois primeiros casos obtiveram valores iguais a 98% enquanto Bagging03 e Bagging04 resultaram em 99%. No treinamento o desvio padrão ficou entre 0,073 e 0,011 enquanto no teste a mesma medida ficou na maioria dos casos em 0,236.

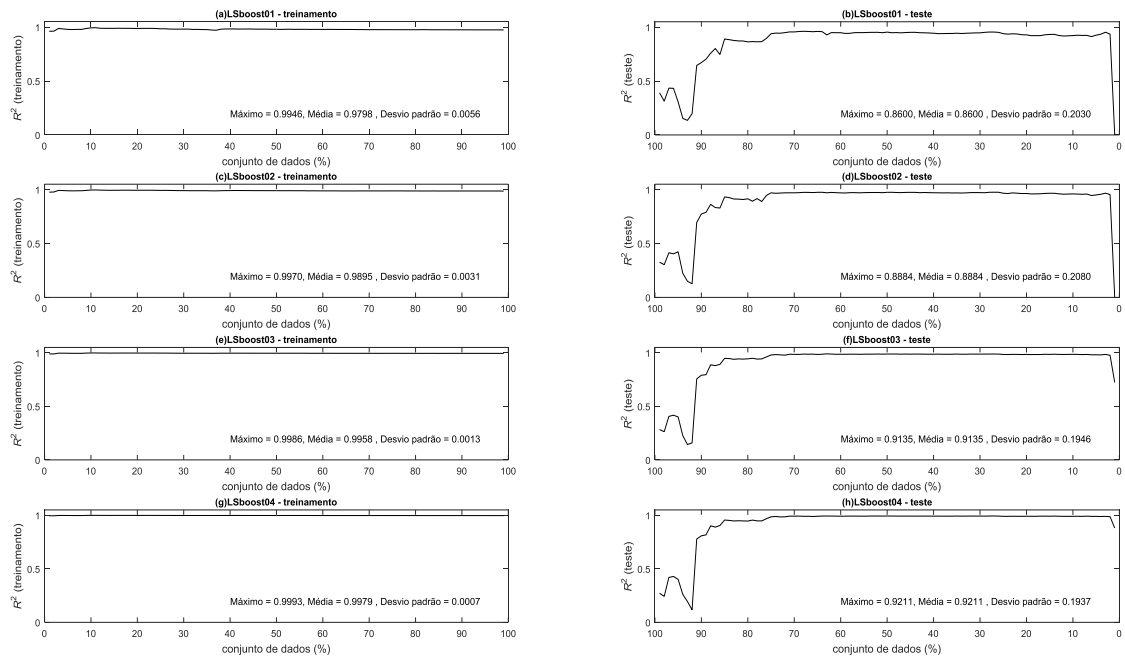
FIGURA 37 – RESULTADOS OBTIDOS PELO MODELO *BAGGING*.



Para as curvas referentes a *Boosting* houve uma melhora gradativa pois no primeiro caso  $R^2$  obtido foi de 94% enquanto no último caso  $R^2$  igual a 99%.

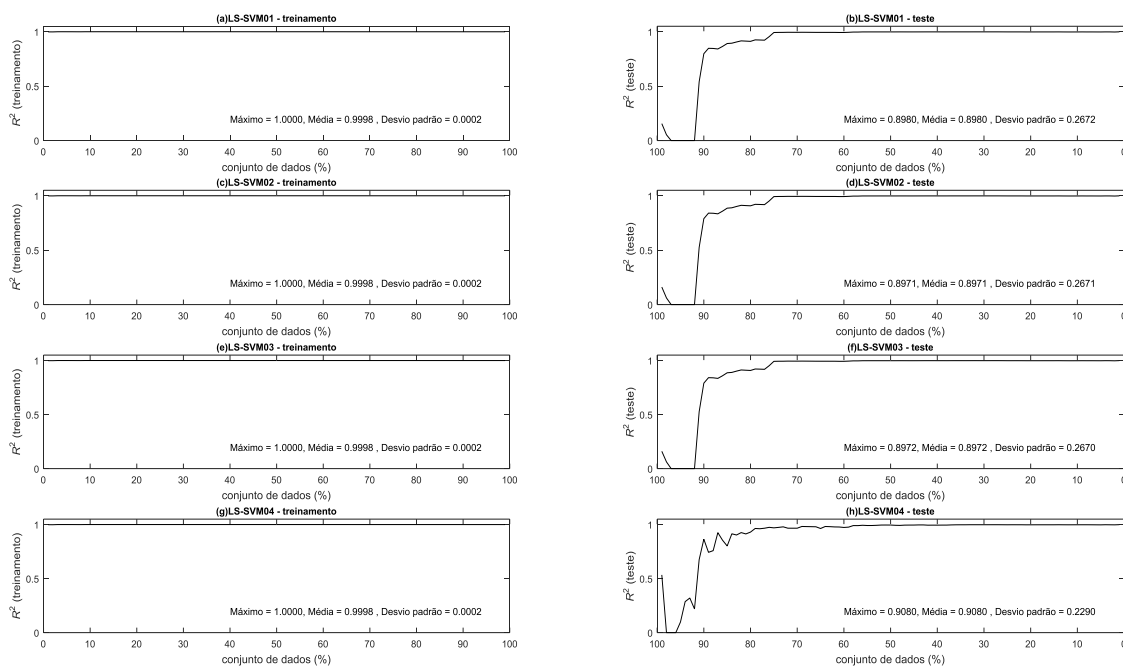
No treinamento o desvio padrão apresentou-se abaixo de 0,006 e com  $R^2$  acentuado, quanto ao teste o desvio padrão ficou em torno de 0,20.

FIGURA 38 – RESULTADOS OBTIDOS PELO MODELO *BAGGING*.



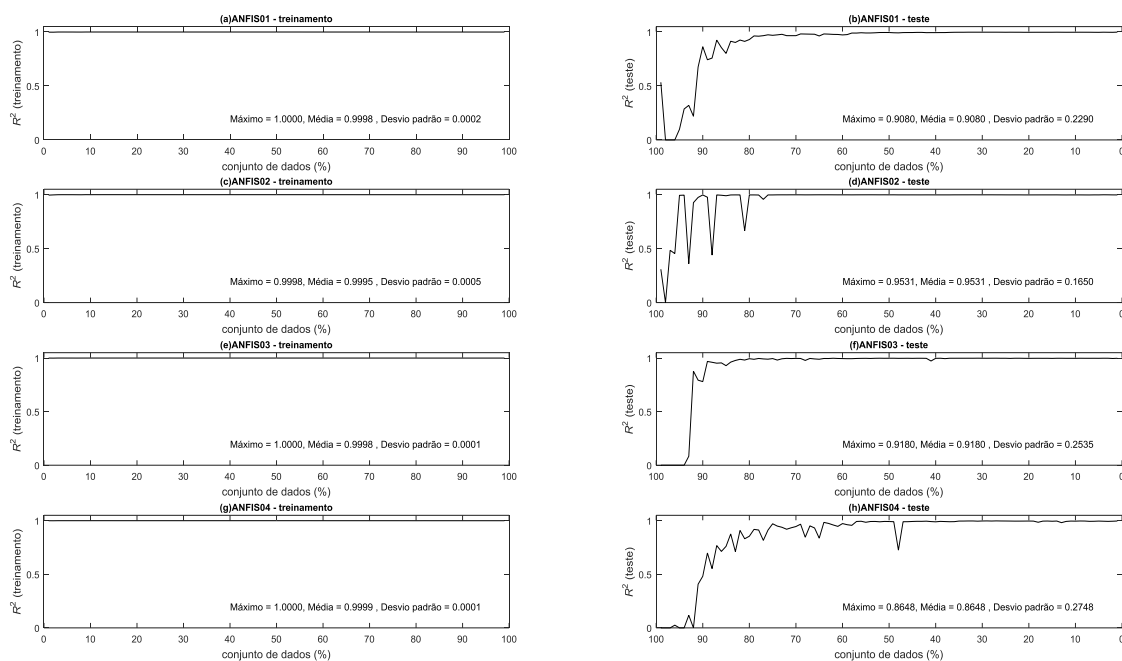
Nos casos de LS-SVM, todos desta classe de algoritmos testada apresentam valores de  $R^2$  superiores a 99,8%. Neste caso o treinamento foram obtidos valores de  $R^2$  próximos ao máximo com desvio padrão em 0,002, por outro lado no teste o desvio padrão foi de 0,267 para os primeiros casos e 0,22 com a LS-SVM04.

FIGURA 39 – RESULTADOS OBTIDOS PELO MODELO LS-SVM.



Em ANFIS os valores de  $R^2$  se mostraram similares em 99,8%, com desvio padrão igual a 0,002 e valor médio próximo a 1. No treinamento e superiores a 0,22 para ANFIS01, ANFIS03, ANFIS04, e valores máximos maiores que 0,90. Embora em ANFIS03 foi obtido 0,16.

FIGURA 40 – RESULTADOS OBTIDOS PELO MODELO ANFIS.



De acordo com os resultados da Tabela 1 os modelos de ESN obtiveram  $R^2$  igual a 99,7%. Para as curvas da Figura 41, no treinamento dos três primeiros casos foram obtidos um baixo desvio padrão e máximos próximos a 1, embora em ESN4 o resultado tenha tomado um pouco mais de tempo para obter resultados similares.

FIGURA 41 – RESULTADOS OBTIDOS PELO MODELO ESN.

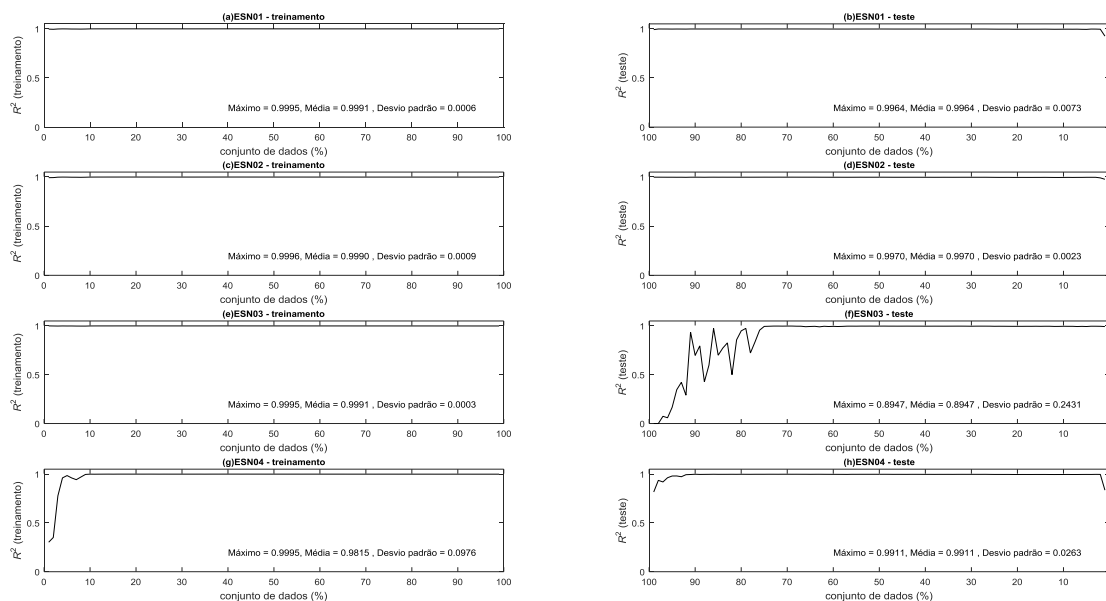
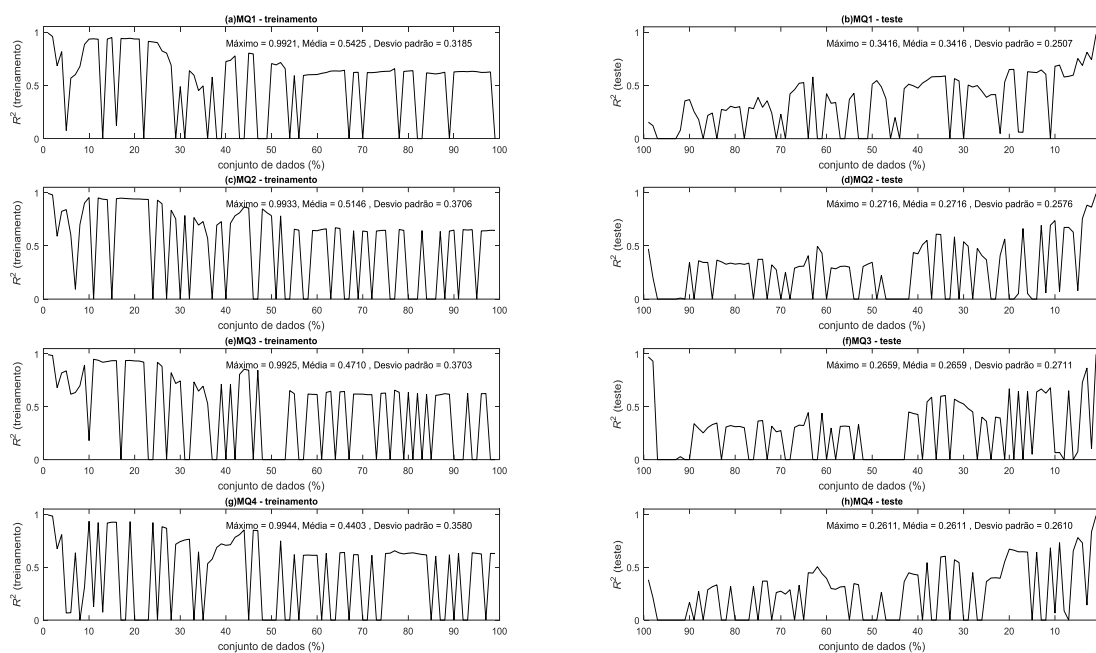


FIGURA 42 – RESULTADOS OBTIDOS PELOS MODELO LINEAR COM MQ.



Nas Figuras 43 e 45 são apresentadas as distribuições dos resultados (incluindo resultados de médias e medianas) das fase de treinamento e teste em termos do  $R^2$ . Não são apresentados os resultados para o método MQ nas Figuras 39 e 40 devido aos valores terem sido acentuadamente diferentes dos obtidos pelos outros métodos. Neste contexto, os resultados para os quatro métodos de MQ são apresentados na Figura 45.



FIGURA 43 – DISTRIBUIÇÃO DOS VALORES DE  $R^2$  OBTIDOS NA FASE DE TREINAMENTO.

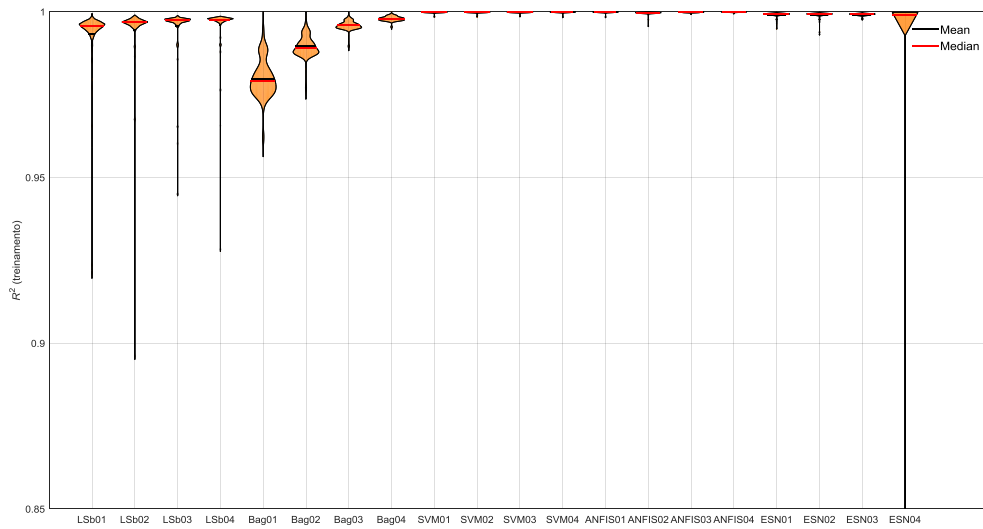


FIGURA 44 – DISTRIBUIÇÃO DOS VALORES DE  $R^2$  OBTIDOS NA FASE DE TESTES.

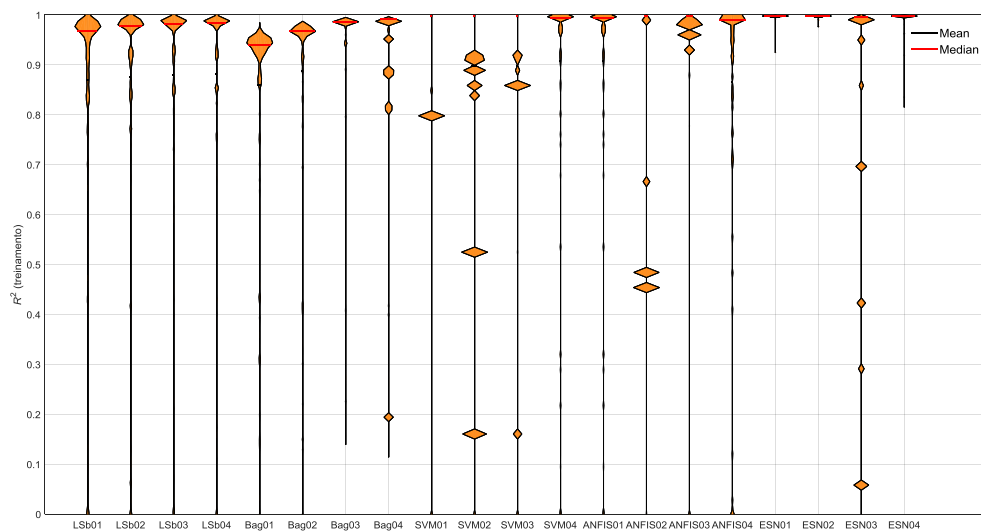
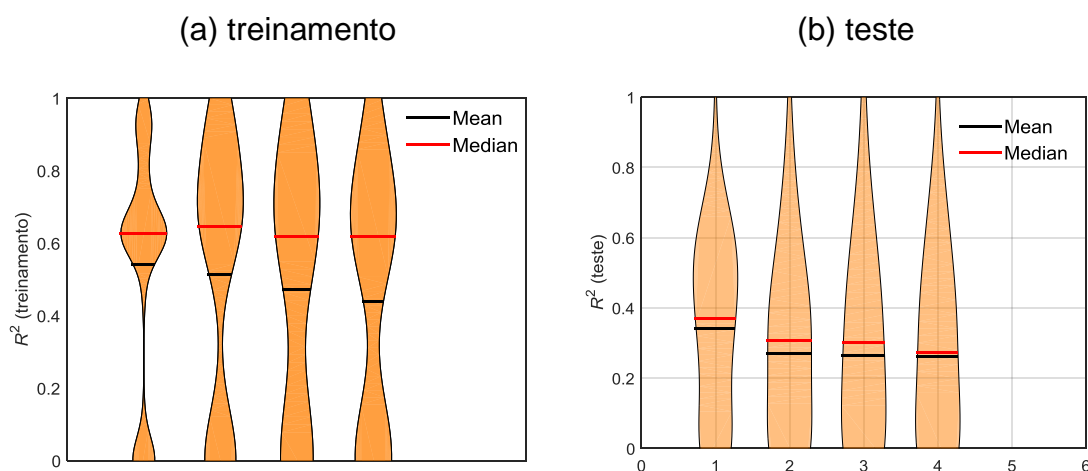


FIGURA 45 – DISTRIBUIÇÃO DOS VALORES DE  $R^2$  OBTIDOS NA FASE DE TREINAMENTO E TESTES PARA O MÉTODO DOS MQ.



Baseado nos resultados da Tabela 1, observa-se que o desempenho de diversos modelos de aprendizado de máquina foi com valores, tanto nas fases de estimação quanto validação com  $R^2$  com valores próximos a 1 (um). No entanto, nota-se que como entradas foi utilizado o SOC ( $x_6$ ) atrasado um instante de tempo ( $t-1$ ) em relação a saída (em %) com intuito de prever o sinal de saída SOC( $t$ ) (*State Of Charge*) no instante atual. O SOC( $t-1$ ) tem acentuada correlação com o sinal SOC( $t$ ). Neste contexto, realizou-se simulações de previsão sem a utilização do sinal SOC( $t-1$ ) ( $x_6$ ) como sinal de entrada, somente  $x_1$  a  $x_6$ . Assim, considerando-se um novo cenário (cenário 2) após realizar as novas simulações sem a utilização do sinal SOC( $t-1$ ) ( $x_6$ ) foram obtidos os resultados apresentados na

Tabela 1.

Em termos dos resultados apresentados na Tabela 2, observa-se que o melhor desempenho nas fases de estimação e validação foram com o Bagging04 (representação na Figura 46) e ESN02 (representação na Figura 47), respectivamente.

**TABELA 2 – RESULTADO EM TERMOS DOS CRITÉRIOS DE DESEMPENHO SEM SINAL  $X_6$  (T-1) COMO SINAL DE ENTRADA DO MODELO.**

Modelo	Estimação (70% dos dados)		Teste (30% dos dados)		Tempo total [s]
	MSE	$R^2$	MSE	$R^2$	
Bagging01	1,4439e-4	0,8171	1,2702e-4	0,5918	0,41
Bagging02	1,4241e-4	0,8216	1,2300e-4	0,5966	0,49
Bagging03	1,4135e-4	0,8253	1,2204e-4	0,5926	1,11
Bagging04	1,3945e-4	0,8295	1,2109e-4	0,5974	2,17
LSBoost01	3,1321e-4	0,4845	1,2687e-4	0,5778	0,78
LSBoost02	3,0162e-4	0,5126	1,2250e-4	0,5720	0,22
LSBoost03	2,8855e-4	0,5426	1,1873e-4	0,5942	0,56
LSBoost04	2,7915e-4	0,5637	1,1235e-4	0,5724	0,74
LS-SVM01	2,5640e-4	0,6107	1,2492e-4	0,5942	0,88
LS-SVM02	2,5570e-4	0,6121	1,2523e-4	0,5940	0,90
LS-SVM03	2,5535e-4	0,6128	1,2551e-4	0,5937	0,92
LS-SVM04	2,5453e-4	0,6142	1,3615e-4	0,5768	10,5
ANFIS01	2,5452e-4	0,6142	1,3615e-4	0,5769	9,83
ANFIS02	2,5758e-4	0,6081	1,2190e-4	0,5965	2,10
ANFIS03	2,1498e-4	0,6885	1,7056e-4	0,5198	27,41
ANFIS04	1,8510e-4	0,7397	2,6199e-4	0,4559	108,90
ESN01	3,5677e-4	0,3860	1,3010e-4	0,5710	0,71
ESN02	3,2272e-4	0,4764	1,1964e-4	0,6298	0,67
ESN03	3,3308e-4	0,4511	3,4139e-4	0,2963	0,81
ESN04	3,6242e-4	0,3728	1,3156e-4	0,5644	0,75
MQ01	4,4184e-4	0,6229	2,3056e-4	0,5296	3,21
MQ02	4,5458e-4	0,6393	2,1929e-4	0,5391	0,46
MQ03	4,8962e-4	0,6188	2,4410e-4	0,5260	0,31
MQ04	4,9222e-4	0,6174	2,4591e-4	0,5250	0,34

FIGURA 46 – RESULTADOS DE PREVISÃO PARA O CENÁRIO 2 OBTIDOS POR BAGGING04.

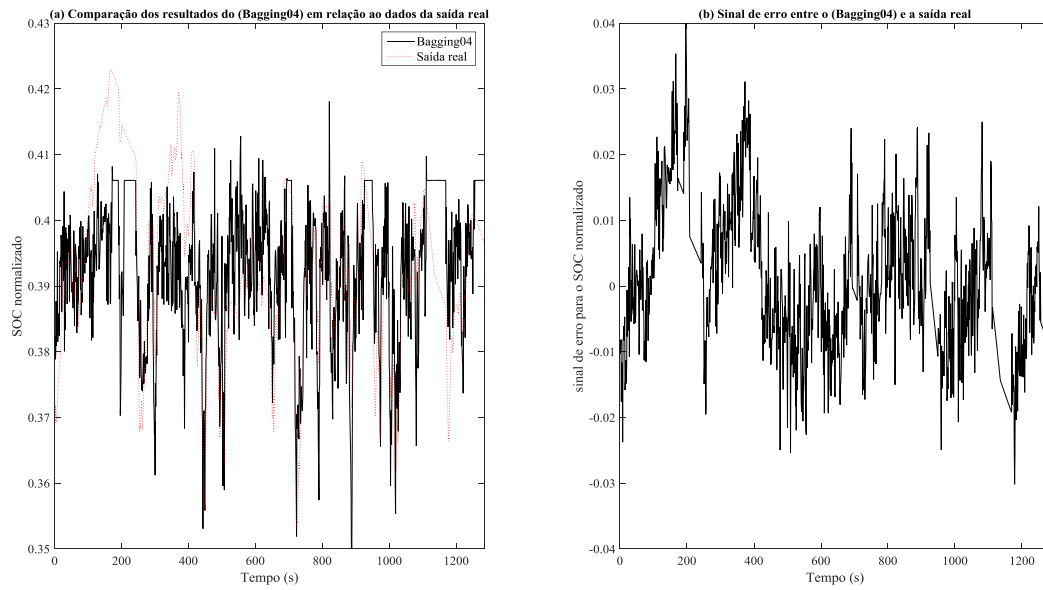
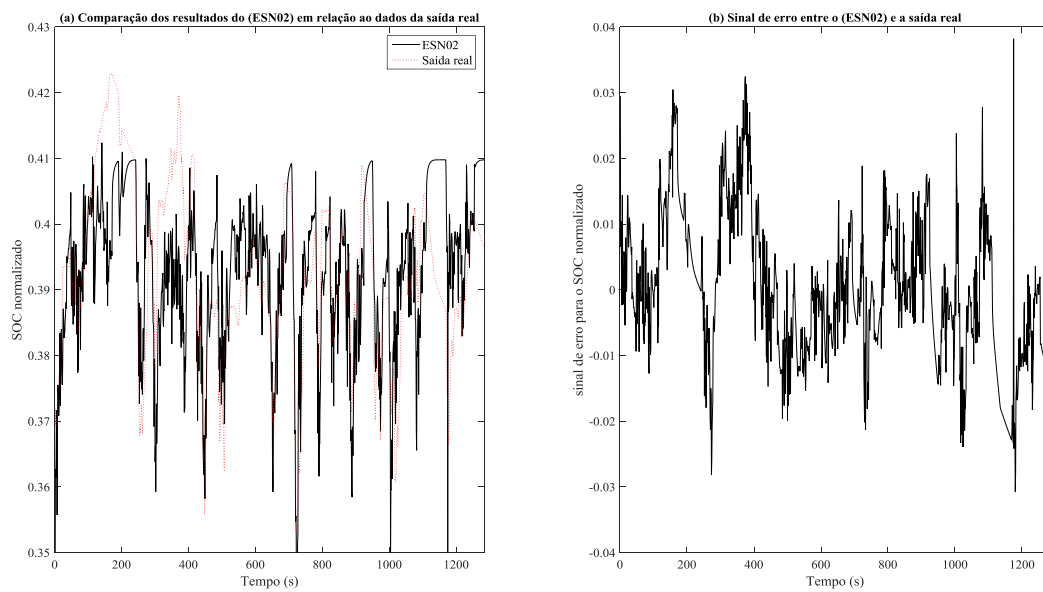


FIGURA 47 – RESULTADOS DE PREVISÃO PARA O CENÁRIO 2 OBTIDOS POR ESN02.



## 7 CONCLUSÃO

Para o estudo foram fornecidos por uma montadora um conjunto de dados que descrevem o comportamento de um veículo híbrido do tipo plug-in em operação. Estes dados foram adquiridos pela rede CAN (*Controller area network*) do veículo durante o funcionamento do mesmo.

Assim este trabalho, avaliou diferentes abordagens de ML no contexto de identificação de sistemas para previsão/modelagem de carga de baterias (SOC) por meio da avaliação de desempenho do erro médio quadrático (MSE) (objetivo minimização) e do coeficiente de determinação ( $R^2$ ) (objetivo maximização), aplicado as técnicas Bagging, Boosting, máquina de vetor de suporte, ANFIS e redes de estado de eco.

Pelo fato que a maioria das aproximações terem ficadas com valores acima de 99% em uma análise preliminar, as técnicas estudadas foram subdivididas em grupos, variando apenas a parametrização dos métodos. Mesmo assim, embora evidenciado as diferentes distribuições de erro e tempo de processamento, os valores em 99% permaneceram pois se tratarmos este caso como identificação a distribuição dos valores já se encontra próximo a 100% conforme mostrado na Figura 44.

O modelo ANFIS projetado gerou o menor erro quadrático seguido de SVM e ESN. Em relação à melhor aproximação ao modelo, através dos resultados para  $R^2$ , temos: SVM, ANFIS e ESN. Entretanto além da busca do melhor resultado buscou-se também o menor custo computacional, este medido pelo tempo. Levando-se em conta esta observação, o melhor método foi SVM para treinamento e *LSBoost* para validação. Se a performance do método em termos de processamento não for requerida então ANFIS seria também uma opção.

Complementando às técnicas citadas, foi ainda aplicado mínimos quadrados, e comprovou-se que tal meio não consegue obter resultados satisfatórios para problemas como o questão que trata de um sistema não linear, pois o resultado divergiu do que era esperado.

Quando analisado o problema como regressão as médias dos métodos não ficaram tão boas quanto quando tratado como identificação. Pela

abordagem de regressão, no treinamento o Bagging representou o melhor método com apenas 80% de aproveitamento, enquanto na abordagem anterior chegou a 99,78%. Desta forma uma falsa impressão pode ser observada caso não escolhido correta a abordagem.

Em termos de futura pesquisa, é sugerido comparar os resultados obtidos com os modelos de ML com outras abordagens de identificação de sistemas não-lineares, tais como modelos Hammerstein, Wiener, Volterra e bilinear. Além disso, adotar algoritmos de otimização, tais como metaheurísticas da computação evolutiva e inteligência de enxames, para a sintonia dos parâmetros de controle dos algoritmos de ML.

## REFERÊNCIAS

- ABB GROUP LTD. EV Charging infrastructure ABB global charging portfolio, ABB library, 12, 2016.
- AGUIRRE, L. A., Introdução à identificação de sistemas: técnicas lineares e não lineares aplicadas a sistemas reais, Belo Horizonte, MG: Editora Universidade Federal de Minas Gerais, 2007.
- ANTÓN, J. C. Á., GARCIA NIETO, P.J., JUEZ, F. J., LASHERAS, F. S., GUTIÉRREZ, M.N. R. Battery state-of-charge estimator using the SVM technique, Applied Mathematical Modelling, v. 37, n. 9, p. 6244-6253, 2013.
- ATHERTON, K. D. A Tiny highway In Sweden Is now electrified. Popular Science. Disponível em: <<http://www.popsoci.com/tiny-highway-in-sweden-is-now-electric>>. Acesso em: 02/07/2017.
- AUTOMOTIVE WORLD. Curitiba, Brazil receives first Volvo electric hybrid bus. Disponível em: <<http://www.automotiveworld.com/news-releases/curitiba-brazil-receives-first-volvo-electric-hybrid-bus/>>. Acesso em: 10/04/2017.
- AUTOMOTIVE WORLD. The electric and hybrid commercial vehicles report. Automotive World Ltd, Cardiff, United Kingdom, 2011. Disponível em <<http://studylib.net/doc/18813428/the-electric-and-hybrid-commercial-vehicles-report>>.
- BALAKRISHNAN, V. System identification: theory for the user (second edition). Automatica, v. 38, n. 2, p. 375-378, 2002.
- BARAN, R.; LEGEY, L. Veículos elétricos: história e perspectivas no Brasil. XIII Congresso Brasileiro de Energia, Brasil, n. Xiii, p. 207–224, 2010. Disponível em <[https://web.bndes.gov.br/bib/jspui/bitstream/1408/1489/1/A%20BS%2033%20Ve%C3%ADculos%20el%C3%A9tricos%20-%20hist%C3%B3ria%20e%20perspectivas%20no%20Brasil\\_P.pdf](https://web.bndes.gov.br/bib/jspui/bitstream/1408/1489/1/A%20BS%2033%20Ve%C3%ADculos%20el%C3%A9tricos%20-%20hist%C3%B3ria%20e%20perspectivas%20no%20Brasil_P.pdf)>. Acesso em: 16/04/2017.
- BARRETO, G. Veículo elétrico à bateria: contribuições à análise de seu desempenho e seu projeto, Dissertação, Mestrado em Ciências, Universidade Estadual de Campinas, Campinas, SP, 1986.
- BERGGREN, C.; MAGNUSSON, T. Reducing automotive emissions-The potentials of combustion engine technologies and the power of policy. Energy Policy, v. 41, p. 636-643, 2012.
- BIOENERGIA, C. -J. DA. O potencial da produção de biodiesel no Brasil. Disponível em: <<http://www.canalbioenergia.com.br/b7-cria-boas-perspectivas-para-a-cadeia-produtiva-2/>>. Acesso em: 16/04/2017.
- BURGOS, C.; SÁEZ, D.; ORCHARD, M. E.; CÁRDENAS, Fuzzy modelling for the state-of-charge estimation of lead-acid batteries, Journal of Power Sources, v. 274, p. 355-366, 2015.
- BREIMAN, L. Bagging predictors. Machine Learning, v. 24, n. 421, p. 1-20, 1994.



- BRODD, R.. Comments on the history of Lithium-ion batteries, *Electrochem*, 2002. Disponível em <<https://www.electrochem.org/dl/ma/201/pdfs/0259.pdf>>. Acesso em: 08/08/2017.
- BUCZAK, A. L.; GUVEN, E. A survey of data mining and machine learning methods for cyber security intrusion detection, *IEEE Communications Surveys & Tutorials*, v. 18, n. 2, p. 1153-1176, 2016.
- CALMON, F. Criado em 1901, Lohner-Porsche Mixte foi o primeiro carro híbrido da história. , 2011. Disponível em: <<https://carros.uol.com.br/noticias/redacao/2011/05/20/criado-em-1901-lohner-porsche-mixte-foi-o-primeiro-carro-hibrido-da-historia.htm>>. Acesso em: 08/08/2017.
- CHAGAS, M. W. P. Novas tecnologias para avaliação de baterias. Dissertação, Instituto de Engenharia do Paraná (LACTEC), Curitiba, PR, p. 97, 2007.
- CHAGAS, S. L. Classificação de sinais de eletroencefalograma usando máquinas de vetores suporte, Dissertação de mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Prebisteriana Mackenzie, São Paulo, SP, 2009.
- CHANG, W. The state of charge estimating methods for battery: a review. *ISRN Applied Mathematics*, vol. 2013, Article ID 953792 (7 pages), Hindawi Publishing Corporation.
- CUBA NISHIOKA, D.; LUIZ, F.; COURA, B.; et al. Estudo dos efeitos da poluição atmosférica na mortalidade neonatal e fetal na cidade de São Paulo, Brasil study of the effects of air pollution in both neonatal mortality and late fetal losses in the city of São Paulo, Brazil, *Revista de Medicina*, v. 79, n. 24, p. 81-89, 2000.
- DEES, D. W. Overview of electrochemical power sources for electric and hybrid/electric vehicles. *Proceedings of the IEEE International Electric Machines and Drives Conference (IEMDC 1999)*, Seattle, Washington, USA, p. 258-259, 1999.
- DENG, Z.; YANG, L.; CAI, Y.; DENG, H. Maximum available capacity and energy estimation based on support vector machine regression for Lithium-ion Battery, *Energy Procedia*, v. 107, p. 68-75, 2017.
- DICKINSON, B.; SWAN, D. EV Battery Pack life: pack degradation and solutions. *Congress of Society of Automotive Engineering, SAE Technical Paper 951949*, p.145-154, 1995.
- ENANG, W.; BANNISTER, C. Modelling and control of hybrid electric vehicles (a comprehensive review). *Renewable and Sustainable Energy Reviews*, v. 74, p. 1210-1239, 2017.
- FELIX, L. Carro “verde” faz fãs no Brasil: quem dirige quer distância do carro comum. Disponível em: <<https://carros.uol.com.br/noticias/redacao/2015/02/13/carro-verde-faz-fas-no-brasil-quem-dirige-quer-distancia-do-carro-comum.htm>>. Acesso em: 09/03/2017.
- FORBES. Como a Toyota está apostando na sustentabilidade para mudar o mercado (e crescer). Disponível em: <<http://www.forbes.com.br/negocios/2016/10/como-a-toyota-esta-apostando-na-sustentabilidade-para-mudar-o-mercado-e-crescer>>. Acesso em: 15/03/2017.
- FREIDMAN, J. H. Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, v. 29, n. 5, p. 1189-1232, 2001.

- GARCHE, J.; JOSSEN, A. Battery management systems (BMS) for increasing battery life time. *Telecommunications Energy Special*, v. 3, p. 81-84, 2000.
- HANSEN, L. K.; SALAMON, P. Neural network ensembles, *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 12, no 10, pp. 993-1001, 1990.
- HANSEN, T.; WANG, C. J. Support vector based battery state of charge estimator, *Journal of Power Sources*, v. 141, n. 2, p. 351-358, 2005.
- HAYKIN, S.; VAN VEEN, B. *Sinais e sistemas*, Porto Alegre, RS, Bookman.
- HOLANDA, A. B. *Dicionário Aurélio básico de língua portuguesa*, Rio de Janeiro, RJ, Nova Fronteira, 1988.
- HOYER, K. G. The history of alternative fuels in transportation: the case of electric and hybrid cars. *Utilities Policy*, v. 16, n. 2, p. 63-71, 2008.
- INDOVAL, R. P. Pronóstico en tiempo real para determinar una alerta contra fenómenos hidrometeorológicos extremos ante la variabilidad climática por medio de: filtro de kalman discreto. *International Journal of General Systems*, v. 17, p. 95-105, 2010.
- JANG, J. S. R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, v. 23, n. 3, p. 665-685, 1993.
- JANG, J. S. R.; SUN, C. -T. Neuro-fuzzy modeling and control, *Proceedings of the IEEE*, v. 83, n. 3, p. 378-406, 1995.
- JEON, S.; YUN, J. -J.; BAE, S. Comparative study on the battery state-of-charge estimation method. *Indian Journal of Science and Technology*, v. 8, p. 26, 2015. Disponível em <<http://www.indjst.org/index.php/indjst/article/view/81677/63028>>. Acesso em: 09/08/2017.
- KALMAN, R. E. A new approach to linear filtering and prediction problems. *Journal of Basics Engineering*, v. 82, p. 35-45, 1960.
- KIMANI, N. G. Environmental pollution and impacts on public health: implications of the Dandora municipal dumping site in Nairobi, Kenya, Urban Environment Unit, United Nations Environment Programme (UNEP), 2012. Disponível em <[http://www.nswai.com/pdf\\_HE/he\\_oct15/Environmental%20Pollution%20and%20Impacts%20on%20Public%20Health.pdf](http://www.nswai.com/pdf_HE/he_oct15/Environmental%20Pollution%20and%20Impacts%20on%20Public%20Health.pdf)>. Acesso em: 05/08/2017.
- LEONTARITIS, I. J.; BILLINGS, S. A. Experimental design and identifiability for nonlinear systems, *International Journal of Systems and Science*, v. 18, n. 1, p. 189-202, 1987.
- LI, Y.; WANG, C.; GONG, J. A combination Kalman filter approach for State of Charge estimation of Lithium-ion battery considering model uncertainty, *Energy*, v. 109, p. 933-946, 2016.
- LI, Z.; YANG, J. PM-25 forecasting use reconstruct phase space LS-SVM. *Proceedings of the 2nd International Conference on Environmental Science and Information Application Technology (ESIAT)*, Wuhan, China, p.143-146, 2010.
- LIMA, C. A. M. *Comitê de máquinas: uma abordagem unificada empregando máquinas de vetores-suporte*, Tese, Faculdade de Engenharia Elétrica e de Computação, UNIVERSIDADE ESTADUAL DE CAMPINAS, 2014.
- LIN, C.; XING, J.; TANG, A. Lithium-ion battery state of charge/state of health estimation

- using SMO for EVs, *Energy Procedia*, v. 105, p. 4383-4388, 2017.
- LJUNG, L. *System identification: theory for the user*. Prentice Hall International, New Jersey, USA, 1987.
- MARIANI, A.; STOCKLEY, T.; THANAPALAN, K. Simple and effective OCV prediction mechanism for VRLA battery systems. *Proceedings of the 3rd International Conference on Mechanical Engineering and Mechatronics*, Paper n. 140, p. 1-10, 2014. Disponível em: <[http://avestia.com/ICMEM2014\\_Proceedings/papers/140.pdf](http://avestia.com/ICMEM2014_Proceedings/papers/140.pdf)>. Acesso em: 05/08/2017.
- MDA. Programa nacional de produção e uso de biodiesel. Disponível em: <[http://www.mda.gov.br/sitemda/sites/sitemda/files/user\\_arquivos\\_64/Biodiesel\\_Book\\_final\\_Low\\_Completo.pdf](http://www.mda.gov.br/sitemda/sites/sitemda/files/user_arquivos_64/Biodiesel_Book_final_Low_Completo.pdf)>. Acesso em: 25/05/2017.
- MITCHELL, T. M. *Machine Learning*. McGraw Hill, New York, USA, 1997.
- PAN, H.; LÜ, Z.; LIN, W.; LI, J.; CHEN, L. State of charge estimation of lithium-ion batteries using a grey extended Kalman filter and a novel open-circuit voltage model, *Energy*, 2017 (in press).
- OLIVEIRA, FRANCISCO; A.M.O.; SILVA, R. P. Uma metodologia para medição de consumo de óleo diesel em ônibus urbanos. *Congresso de Transporte e Trânsito*, Brasília, DF, p. 16, 2007.
- OPITZ, A.; BADAMI, P.; SHEN, L.; VIGNAROOBAN, K.; KANNAN, A. M. Can Li-Ion batteries be the panacea for automotive applications? *Renewable and Sustainable Energy Reviews*, v. 68, p. 685–692, 2017.
- OPP Charge, Fast charging of electric vehicles. Disponível em: <<https://www.oppcharge.org>>. Acesso em: 22/06/2017.
- PESSANHA, J. F. M. Sistema de inferência neuro-fuzzy - ANFIS, Notas de aula, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, RJ, 2002.
- PESSÔA, M. W. Identificação de sistemas não lineares utilizando modelo polinomial NARMAX e nebuloso Takagi-Sugeno-Kang, Programa de Pós-Graduação em Engenharia de Produção e Sistemas, Pontifícia Universidade Católica do Paraná, Curitiba, PR, 2010.
- PILLONETTO, G.; DINUZZO, F.; CHEN, T.; NICOLAO, G.; LJUNG, L. Kernel methods in system identification, machine learning and function estimation: A survey, *Automatica*, v. 50, n. 3, p. 657-682, 2014.
- POLIKAR, R. Ensemble learning. *Scholarpedia*, v. 4, n. 1, p. 2776, 2009. Disponível em: <[http://www.scholarpedia.org/article/Ensemble\\_learning](http://www.scholarpedia.org/article/Ensemble_learning)>. Acesso em: 20/07/2017.
- POP, V.; BERGVELD, H. J.; DANILOV, D.; REGTIEN, P. P. L.; NOTTEN, P. H. L. *Battery management systems accurate state-of-charge indication for battery-powered applications*. 9th edition, Berlin, German: Springer, 2008.
- REN, Y.; ZHANG, L.; SUGANTHAN, P. N. Ensemble classification and regression-recent developments, applications and future directions, *IEEE Computational Intelligence Magazine*, v. 11, n. 1, p. 41-53, 2016.
- ROCHA, B. P. da; ALBERTON, H. B. de S.; OLIVEIRA, L. B. de. Frenagem regenerativa. 2014. Disponível em: <<http://www.ufrgs.br/projenergia3/projetos/trabalhos-2014/>>

- trabalhos-2014-2/GRUPOB.pdf>. Acesso em:10/01/2016.
- SÁNCHEZ, L.; COUSO, I.; BLANCO, C., A class of monotone fuzzy rule-based Wiener systems with an application to Li-ion battery modelling, *Engineering Applications of Artificial Intelligence*, v. 64, p. 367-377, 2017.
- SCROSATI, B.; GARCHE, J.; TILLMETZ, W. (eds.) *Advances in battery technologies for electric vehicles*. A volume in Woodhead Publishing Series in Energy, Elsevier, 2015.
- SEBORG, D. E., EDGAR, T. F., MELLICHAMP, D. A. *A process dynamics and control*, New York, USA, John Wiley.
- SILVA JR, O.; RIBEIRO, E. S.; GRUSKA, E. G.; KLEIN, C. E.; COELHO, L. S. Aprendizado de máquina aplicado à previsão de carga de baterias para veículos híbridos, *Anais do 12th IEEE/IAS International Conference on Industry Applications (Induscon 2016)*, Curitiba, PR, 2016.
- SILVIO. Carros híbridos e elétricos: veja opções que temos no Brasil. Disponível em: <<http://autoblogceara.com.br/carros-hibridos-e-eletricos-veja-opcoes-que-temos-no-brasil/>>. Acesso em: 05/04/2017.
- SPATH, D.; BAUER, W.; ROTHFUSS, F.; VOIGT, S. *Strukturstudie BW e mobil*, 2011. Disponível em: <<http://wiki.iao.fraunhofer.de/images/studien/strukturstudie-bwe-mobil-2011.pdf>>. Acesso em: 15/05/2017.
- Statistics and Machine Learning toolbox: user's guide*, 2016. Natick, MA, USA: The MathWorks Inc. Disponível em: <[http://www.mathworks.com/help/pdf\\_doc/stats/stats.pdf](http://www.mathworks.com/help/pdf_doc/stats/stats.pdf)>. Acesso em: 18/05/2017.
- STREHLER, M.; MERTING, S.; SCHWAN, C. Energy-efficient shortest routes for electric and hybrid vehicles. *Transportation Research Part B: Methodological*, 2017 (in press).
- STRUBEN, J. Identifying challenges for sustained adoption of alternative fuel vehicles and infrastructure. MIT Sloan Working Paper 4625-06, 2006. Disponível em <[https://dspace.mit.edu/bitstream/handle/1721.1/37306/4625-06.pdf?sequence=1&origin=publication\\_detail](https://dspace.mit.edu/bitstream/handle/1721.1/37306/4625-06.pdf?sequence=1&origin=publication_detail)>. Acesso em: 09/08/2017.
- SURENDAR, V.; MOHANKUMAR, V.; ANAND, S.; PRASANNA, V. D. Estimation of state of charge of a lead acid battery using support vector regression, *Procedia Technology*, v. 21, p. 264-270, 2015.
- SUYKENS, J. A. K.; GESTEL, T. VAN; BRABANTER, J. DE; MOOR, B. DE; VANDEWALLE, J. *Least Squares Support Vector Machines*. World Scientific Pub. Co., Singapore, 2002.
- TEAM, M. V. *A guide to understanding battery specifications*, 2008. MIT Electric Vehicle Team. Disponível em: <[http://web.mit.edu/evt/summary\\_battery\\_specifications.pdf](http://web.mit.edu/evt/summary_battery_specifications.pdf)>. Acesso em: 09/08/2017.
- TEIXEIRA DIAS, C. M. Projeto e desenvolvimento de um automóvel híbrido, 2013. Dissertação, Mestrado Integrado em Engenharia Mecânica, Ramo de projeto, desenvolvimento e engenharia automóvel, Faculdade de Engenharia, Universidade do Porto, Porto, Portugal. Disponível em: <[http://sigarra.up.pt/feup/pt/pubs\\_pesquisa.show\\_publ\\_file?pct\\_gdoc\\_id=311150](http://sigarra.up.pt/feup/pt/pubs_pesquisa.show_publ_file?pct_gdoc_id=311150)>.

- Acesso em: 09/08/2017.
- THIELMANN, A.; SAUER, R.; ISENMANN, M. W. Technologie-roadmap energiespeicher für die elektromobilität 2030. Fraunhofer-Institut für Sys- tem- und Innovationsforschung ISI, Karlsruhe, German, 2012.
- TORSCHER, A.; JÄHRER, M. The BigChaos solution to the Netflix grand prize. Netflix prize documentation, August, p. 1–10, 2009. Disponível em: <[http://www.stat.osu.edu/~dmsl/GrandPrize2009\\_BPC\\_BellKor.pdf](http://www.stat.osu.edu/~dmsl/GrandPrize2009_BPC_BellKor.pdf)>. Acesso em: 09/08/2017.
- TRINH, F. A method for evaluating battery state of charge estimation accuracy, Master thesis, Department of Signals and Systems, Division of Automatic Control, Chalmers, University of Technology, Gothenburg, Sweden, 2012.
- VAPNIK, V. N. The natural of statistical learning theory, Berlin, German, Springer Verlag, 1995.
- VILLANUEVA, W. J. P. Comitê de máquinas em predição de séries temporais, Tese, Pós-graduação da Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas (UNICAMP), SP, 2006.
- VOLVO. Ônibus híbrido: tecnologia, economia e respeito ao meio ambiente, 2014. Disponível em <<http://www.volvobuses.com.br/content/dam/volvo/volvo-buses/master/bra/our-offering/buses/hibrido/folheto-onibus-hibrido-2014.pdf>>. Acesso em: 09/08/2017.
- WAAG, W.; FLEISCHER, C.; SAUER, D. U. Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles. Journal of Power Sources, v. 258, p. 321–339, 2014.
- WANG, H.; LIU, S.; LI, S.; LI, G. Study on state of charge estimation of batteries for electric vehicle. Proceedings of the 2nd International Conference on Advanced Signal Processing (ASP), Harbin, China, p. 10-14, 2013. Disponível em: <[http://onlinepresent.org/proceedings/vol18\\_2013/3.pdf](http://onlinepresent.org/proceedings/vol18_2013/3.pdf)>. Acesso em: 09/08/2017.
- XING, Y.; HE, W.; PECHT, M.; TSUI, K. State of charge estimation of lithium-ion batteries using the open-circuit voltage at various ambient temperatures. Applied Energy, v. 113, p. 106-115, 2014.
- YAN, J.; XU, G.; QIAN, H.; XU, Y. Robust state of charge estimation for hybrid electric vehicles: Framework and algorithms. Energies, v. 3, n. 10, p. 1654-1672, 2010.
- YU, Z.; HUAI, R.; XIAO, L. State-of-charge estimation for lithium-ion batteries using a Kalman filter based on local linearization. Energies, v. 8, n. 8, p. 7854-7873, 2015.
- ZHI, L.; PENG, Z.; ZHIFU, W.; QIANG, S.; YINAN, R. State of charge estimation for Li-ion battery based on extended Kalman filter, Energy Procedia, v. 105, pp. 3515-3520, 2017.