

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE ENERGIAS ALTERNATIVAS E RENOVÁVEIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA



Modelo para Estimaco do Estado de Carga
e Sade de Baterias de Ltio-on Baseado em
Redes Neurais com Funo de Custo em
Correntropia

Rmulo Navega Vieira

Joo Pessoa
Dezembro - 2021

RÔMULO NAVEGA VIEIRA

**Modelo para Estimação do Estado de Carga e Saúde de
Baterias de Lítio-Íon Baseado em Redes Neurais com
Função de Custo em Correntropia**

Dissertação final apresentado ao Programa de Pós-Graduação em Engenharia Elétrica - PPGEE, da Universidade Federal da Paraíba - UFPB, como requisito final para a obtenção do título de Mestre em Engenharia Elétrica

Orientador: Prof. Dr. Juan Moisés Maurício Villanueva

João Pessoa - PB

2021

Catálogo na publicação
Seção de Catalogação e Classificação

V658m Vieira, Romulo Navega.

Modelo para estimação do estado de carga e saúde de baterias lítio-íon baseado em redes neurais com função de custo em correntropia / Romulo Navega Vieira. - João Pessoa, 2021.

94 f. : il.

Orientação: Juan Moises Maurício Villanueva.
Dissertação (Mestrado) - UFPB/CEAR.

1. Engenharia elétrica. 2. Estado de carga e saúde.
3.
Baterias. 4. Lítio-íon. 5. Correntropia. 6. Redes neurais artificiais. I. Villanueva, Juan Moises Maurício. II. Título.

UFPB/BC

CDU 612.3(043)

UNIVERSIDADE FEDERAL DA PARAÍBA – UFPB
CENTRO DE ENERGIAS ALTERNATIVAS E RENOVÁVEIS – CEAR
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA - PPGE

A Comissão Examinadora, abaixo assinada, aprova a Dissertação

**MODELO PARA ESTIMAÇÃO DO ESTADO DE CARGA E SAÚDE DE BATERIAS
DE LÍTIO-ÍON BASEADO EM REDES NEURAIIS COM FUNÇÃO DE CUSTO EM
CORRENTROPIA**

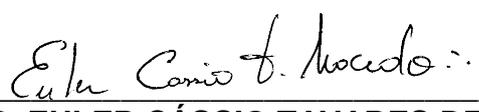
Elaborada por
ROMULO NAVEGA VIEIRA

como requisito parcial para obtenção do grau de
Mestre em Engenharia Elétrica.

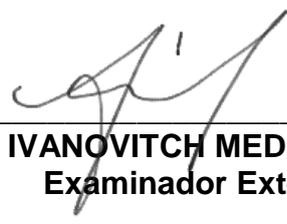
COMISSÃO EXAMINADORA



PROF. DR. JUAN MOISÉS MAURICIO VILLANUEVA
Orientador – UFPB



PROF. DR. EULER CÁSSIO TAVARES DE MACÊDO
Examinador Interno – UFPB



PROF. DR. IVANOVITCH MEDEIROS DANTAS DA SILVA
Examinador Externo – UFRN

Resumo

O grande interesse na identificação de sistemas dinâmicos não-lineares deve-se, principalmente, ao fato de que uma grande quantidade dos sistemas reais são complexos e precisam ter suas linearidades consideradas para que seus modelos possam ser utilizados com sucesso em aplicações, por exemplo, de controle, predição, inferência, entre outros. Um exemplo de um sistema dinâmico não-linear é a bateria. Compreender o envelhecimento da bateria é um processo complexo, pois muitos fatores, desde as condições ambientais às condições operacionais de uso, interagem para gerar diferentes efeitos de envelhecimento e degradação. O presente trabalho analisa a aplicação das Redes Neurais Artificiais na identificação de indicadores de desempenho e saúde em baterias. Normalmente, o aprendizado dessas redes é realizado através de algum método baseado em gradiente, tendo o erro médio quadrático como função de custo. Este trabalho analisa a substituição dessa tradicional função de custo por uma medida da similaridade da Teoria da Informação, a Correntropia. Esta medida de similaridade permite que momentos estatísticos de ordem superior possam ser considerados durante o processo de treinamento. Devido a este fato, ela se torna mais apropriada para distribuições de erro não gaussianas e faz com que o treinamento apresente menor sensibilidade à presença de *outliers*. Para avaliar esta substituição, modelos RNAs são obtidos na identificação de dois estudos de caso: o Estado de Carga (SoC) e Estado de Saúde (SoH) das baterias. Os resultados demonstram que a utilização da correntropia, como função de custo no algoritmo de retropropagação de erro, minimiza os erros de estimação do modelo, bem como permite a rejeição natural de amostras discrepantes. Os resultados finais mostram que o SoC e o SoH foram estimados de forma satisfatória sob todas as diferentes condições de operação consideradas neste trabalho.

Palavras-chave: Identificação de Sistemas Dinâmicos, Estado de Carga e Saúde, Baterias, Função de Custo, Correntropia, *Outliers*.

Abstract

The great interest in identifying non-linear dynamic systems is due, mainly, to the fact that a large number of real systems are complex and need to have their linearities considered so that their models can be successfully used in applications such as control, prediction, inference, among others. An example of a non-linear dynamic system is the battery. Understanding battery aging is a complex process, because many factors, from environmental conditions to operational conditions of use, interact in order to generate different effects of aging and degradation. This work analyzes the application of Artificial Neural Networks in the identification of performance and health indicators in batteries. Normally, the learning of these networks is done through some method based on gradient, having the average quadratic error as a cost function. This work analyzes the replacement of this traditional cost function by a measure of the similarity of Information Theory, the Correntropia. This measure of similarity allows that statistical moments of a higher order can be considered during the training process. Due to this fact, it becomes more appropriate for non-Gaussian error distributions and makes the training less sensitive to the presence of outliers. To evaluate this substitution, ANNs models are obtained in the identification of two case studies: the State of Charge (SoC) and State of Health (SoH) of batteries. The results show that the use of correntropy, as a cost function in the error backpropagation algorithm, makes the identification procedure using Neural Networks more robust to outliers, as well as minimizing model estimation errors. However, this can only be achieved by properly adjusting the width of kernel gaussian, either through techniques based on genetic algorithms or through estimates found in the scientific literature.

Keywords: Identification of Dynamic Systems, State of Charge and Health, Batteries, Cost Function, Correntropy, Outliers.

Lista de ilustrações

Figura 1 – Etapas principais do procedimento de identificação.	27
Figura 2 – Representação de uma RNA recursiva com duas camadas ocultas. . . .	29
Figura 3 – Procedimento de treinamento da rede neural artificial.	30
Figura 4 – Comparação gráfica entre o erro quadrático médio e <i>kernel gaussiano</i> . .	34
Figura 5 – Otimização dos parâmetros η e σ para algoritmo de treinamento sem influência de ruídos e <i>outliers</i>	39
Figura 6 – Diagrama esquemático da arquitetura de identificação utilizada para sistemas sem ruído e <i>outliers</i>	40
Figura 7 – Diagrama esquemático da arquitetura de identificação utilizada para sistemas com ruído e <i>outliers</i>	43
Figura 8 – Arquitetura da Rede Neural Artificial para o processo de identificação - NARX.	45
Figura 9 – Indivíduos da População Inicial - Algoritmo Genético.	50
Figura 10 – Determinação do <i>fitness</i> de cada indivíduo - Algoritmo Genético. . . .	51
Figura 11 – Cruzamento de Indivíduos - Algoritmo Genético.	53
Figura 12 – Curvas de tensão de descarga e estado de carga em único ensaio.	55
Figura 13 – Curvas de corrente de descarga e temperatura de operação	56
Figura 14 – Estado de Carga e Capacidade atual da Bateria de Lítio-íon - Modelo 18650.	60
Figura 15 – Validação dos modelos RNA - Estado de Carga - Resultado de treinamento sem a presença de ruído e <i>outliers</i>	62
Figura 16 – População de Indivíduos - Estado de Carga - Resultado da otimização por algoritmo genético sem a presença de ruído e <i>outliers</i> . (a) População Inicial de indivíduos, (b) população final de indivíduos e (c) evolução do <i>fitness</i> do melhor indivíduo de cada geração.	63
Figura 17 – Curvas de aprendizado dos modelos RNA - Estado de Carga - Resultado de treinamento sem a presença de ruído e <i>outliers</i>	65
Figura 18 – Validação dos modelos RNA - Estado de Carga - Resultado de treinamento com a presença de ruído e <i>outliers</i>	67
Figura 19 – Variação da largura de <i>kernel</i> adaptativo - Estado de Carga - Resultado de treinamento com a presença de ruído e <i>outliers</i>	68
Figura 20 – Curvas de aprendizado dos modelos RNA - Estado de Carga - Resultado de treinamento com ruído e <i>outliers</i>	69
Figura 21 – Medidas Estatísticas do erro de estimação - Estado de Carga - Resultado de treinamento com a presença de ruído e <i>outliers</i>	71

Figura 22 – Validação dos modelos RNA - Estado de Carga - Comparação entre funções de custo MSE e MCC.	72
Figura 23 – População de Indivíduos - Estado de Saúde - Resultado da otimização por algoritmo genético sem a presença de ruído e <i>outliers</i> . (a) População inicial de indivíduos, (b) população final de indivíduos e (c) evolução do <i>fitness</i> do melhor indivíduo de cada geração.	74
Figura 24 – Validação dos modelos RNA - Estado de Saúde - Resultado de treinamento sem a presença de ruído e <i>outliers</i>	75
Figura 25 – Curvas de aprendizado dos modelos RNA - Estado de Saúde - Resultado de treinamento sem a presença de ruído e <i>outliers</i>	76
Figura 26 – Validação dos modelos RNA - Estado de Saúde - Resultado de treinamento com presença de ruído e <i>outliers</i>	78
Figura 27 – Variação da largura de <i>kernel</i> adaptativo - Estado de Saúde - Resultado de treinamento com presença de ruído e <i>outliers</i>	79
Figura 28 – Curvas de aprendizado dos modelos RNA - Estado de Saúde - Resultado de treinamento com presença de ruído e <i>outliers</i>	80
Figura 29 – Medidas Estatísticas do erro de estimação - Estado de Saúde - Resultado de treinamento com a presença de ruído e <i>outliers</i>	82
Figura 30 – Validação dos modelos RNA - Estado de Saúde - Comparação entre funções de custo MSE e MCC.	83

Lista de tabelas

Tabela 1 – Especificações técnicas - Modelo 18650.	54
Tabela 2 – Comparação de desempenho dos modelos RNA - Estado de Carga - Resultado de treinamento sem a presença de ruído ou <i>outliers</i>	66
Tabela 3 – Comparação de desempenho dos modelos RNA - Estado de Carga - Resultado de treinamento com a presença de ruído e <i>outliers</i>	69
Tabela 4 – Medidas Estatísticas do erro de estimação - Estado de Carga - Resultado de treinamento com a presença de ruído e <i>outliers</i>	70
Tabela 5 – Comparação de desempenho com função de custo MCC e MSE - Estado de Carga - Resultado de treinamento com a presença de ruído e <i>outliers</i>	73
Tabela 6 – Comparação de desempenho dos modelos RNA - Estado de Saúde - Resultado de treinamento sem a presença de ruído ou <i>outliers</i>	77
Tabela 7 – Comparação de desempenho dos modelos RNA - Estado de Saúde - Resultado de treinamento com a presença de ruído ou <i>outliers</i>	81
Tabela 8 – Medidas Estatísticas do erro de estimação - Estado de Saúde - Resultado de treinamento com a presença de ruído e <i>outliers</i>	81
Tabela 9 – Comparação de desempenho com função de custo MCC e MSE - Estado de Carga - Resultado de treinamento com a presença de ruído e <i>outliers</i>	83

Lista de abreviaturas e siglas

AG	Algoritmo Genético
BEV	<i>Battery Electric Vehicle</i>
BMS	<i>Battery Management System</i>
CE	Camada de Entrada
CI	Camada Intermediária
CS	Camada de Saída
EV	<i>Electric Vehicle</i>
IBS	<i>Intelligent Battery Sensor</i>
ITL	<i>Information Theoretic Learning</i>
IRPROP	<i>Improve Resilient Backpropagation</i>
MAE	Média dos Erros Absolutos
MAPE	Média Percentual Absoluta do Erro
MCC	<i>Maximum Correntropy Criterion</i>
ME	Erro Médio
MLP	<i>Multilayer Perceptron Network</i>
MSE	<i>Mean Squared Error</i> - Erro Médio Quadrático
NARX	<i>Nonlinear Autoregressive Exogenous Model</i>
NASA	<i>National Aeronautics and Space Administration</i>
NNARX	<i>Neural Network AutoRegressive with Exogenous Input</i>
PDF	<i>Probability Distribution Function</i> - Função de Distribuição e Probabilidade
PHEV	<i>Plug-in Hybrid Electric Vehicle</i>
RMSE	Root Mean Square Error - Raiz do Erro Quadrático Médio
RNA	Rede Neural Artificial

SOC	<i>State Of Charge</i> - Estado de Carga
SOF	<i>State Of Function</i> - Estado de Função
SOH	<i>State Of Health</i> - Estado de Saúde
TDNN	<i>Time Delay Neural Network</i>

Lista de símbolos

a_j	Resposta do j-ésimo neurônio da camada oculta da rede neural
b_j	<i>Bias</i> aplicado ao j-ésimo neurônio da camada oculta da rede neural
e_k	Erro de estimativa para o k-ésimo exemplo de treinamento
ε	Variável aleatória que representa os erros de estimação de modelo
K_σ	Kernel Gaussiano
J_{MCC}	Função custo utilizada pelo algoritmo de retropropagação do erro para maximização da correntropia
N	Número de amostras coletadas do sistema
η	Coefficiente de aprendizado
σ	Largura de <i>kernel</i> da correntropia
P_{ij}	Variável aleatória generalizada auxiliar
u_d	<i>Bias</i> aplicado ao d-ésimo neurônio da camada de saída da rede neural
v_{jd}	Peso sináptico conectando o j-ésimo neurônio da camada oculta ao d-ésimo neurônio da camada de saída
x_i	i-ésima entrada da rede neural
y^*	Saída desejada de referência
y	Saída estimada pela rede neural
w_{ij}	Peso sináptico conectando o i-ésimo nó da camada de entrada ao j-ésimo neurônio da camada oculta da rede neural

Sumário

1	INTRODUÇÃO	12
1.1	Objetivos	14
1.1.1	Objetivo Geral	14
1.1.2	Objetivos Específicos	14
1.2	Contribuições do Trabalho	15
1.3	Organização do Trabalho	16
2	ESTADO DA ARTE	17
2.1	Estado de Carga (SoC)	17
2.1.1	Problemática	17
2.1.2	Estado da Arte	18
2.2	Estado de Saúde (SoH)	20
2.2.1	Problemática	20
2.2.2	Estado da Arte	22
2.2.3	Conclusões	24
3	FUNDAMENTAÇÃO TEÓRICA	25
3.1	Identificação de Sistemas Dinâmicos	26
3.2	Redes Neurais e Estruturas Neurais de Identificação	28
3.3	Medida de Correntropia	32
3.3.1	Propriedades da Correntropia	33
3.3.2	Correntropia e Erro Médio Quadrático	34
3.4	Conclusões	35
4	METODOLOGIA PROPOSTA	36
4.1	Método Proposto - Arquitetura de Identificação	36
4.2	Arquitetura da Rede Neural Artificial	44
4.3	Aplicação do MCC no Treinamento da RNA	45
4.3.1	Algoritmo Genético	48
4.4	Base de Dados - <i>National Aeronautics and Space Administration</i> - NASA	54
4.5	Normalização e Processamento do Conjunto de Dados	56
4.6	Definições de Outliers e Ruídos	57
5	RESULTADOS FINAIS E CONSIDERAÇÕES	58
5.1	Banco de Dados NASA - Conjunto de Treinamento e Validação	59

5.2	Resultados - Estudo de Caso I - Estado de Carga	60
5.2.1	Resultados - Sem presença de ruído e <i>outliers</i>	61
5.2.2	Resultados - Presença de ruído e <i>outliers</i>	66
5.2.3	Resultados - Presença de ruído e outliers - Comparação com MSE clássica .	70
5.3	Resultados - Estudo de Caso II - Estado de Saúde	73
5.3.1	Resultados - Sem presença de ruído e <i>outliers</i>	73
5.3.2	Resultados - Presença de ruído e <i>outliers</i>	77
5.3.3	Resultados - Presença de ruído e outliers - Comparação com MSE clássica .	81
5.4	Conclusões	83
6	CONCLUSÕES	85
7	TRABALHOS FUTUROS E PERSPECTIVAS	87
	REFERÊNCIAS	88

1 Introdução

Nas últimas décadas, a busca pela utilização e aprimoramento dos sistemas de armazenamento de energia elétrica tem se tornado o foco de interesse para pesquisadores em todo o mundo. Levando-se em consideração a restrição ambiental de emissão de CO_2 , o setor de veículos automotivos se encontra em um limiar de mudanças. Com as políticas de redução de efeito estufa, em breve será cada vez mais restrito o uso de gasolina ou óleo diesel, de modo que este setor está iniciando uma migração de veículos a combustão para Veículos Elétricos (EVs), Veículos Elétricos Híbridos *Plug-in* (PHEVs) e Veículos Elétricos a Bateria (BEVs).

Independentemente do veículo utilizado (EVs, HEVs, PHEVs ou BEVs), suas principais fontes energéticas são os bancos de baterias. Estes bancos são compostos de conjuntos de baterias conectadas em arranjos série/paralelo, onde em cada bateria está contida uma grande quantidade de células, a unidade básica de uma bateria (Rosolem, 2012). Neste contexto, existe a necessidade de fomentar o desenvolvimento de baterias mais eficientes, duráveis e seguras. Para tanto, é imprescindível a realização de estudos sobre o comportamento das baterias, especificamente, acerca de métodos para predição de indicadores de saúde e de desempenho.

Como o banco de baterias é um dos elementos mais importantes para aplicações em veículos elétricos e por possuir varias baterias acopladas, existem equipamentos especialmente dedicados em fazer o monitoramento destes componentes no veículo. São eles: O Sistema de Gerenciamento de Baterias (*Battery Management System* - BMS) e o Sensor Inteligente de Bateria (*Intelligent Battery Sensor* - IBS). Ambos são responsáveis por coletar dados como tensão, corrente (carga/descarga) e temperatura e, a partir disto, estimar parâmetros críticos como estado de carga, estado de saúde e estado de função, utilizados para proporcionar o funcionamento adequado e seguro das baterias. Neste trabalho estes parâmetros serão referenciados em termos de suas abreviações em inglês: *State of Charge* (SoC), *State of Health* (SoH) e *State of Fuction* (SoF).

A estimação do estado de carga, que pode ser definido como a quantidade de energia restante em uma bateria, é um ponto chave para que se viabilize a sua utilização de forma segura e eficiente, tanto em EV, quanto em *Smart Grids*. Atualmente, tal estimação é considerada uma tarefa dispendiosa e desafiadora, de modo que foram desenvolvidas, ao longo do tempo, inúmeras técnicas para sua realização (Rahimi-Eichi *et al.*, 2013). O estado de saúde, por outro lado, é uma medida que reflete a condição geral de uma bateria e sua habilidade de entregar uma específica *performance* em comparação a uma bateria nova. O SoC e SoH são índices altamente correlacionados, um compromisso de troca existe

entre estender o ciclo de vida da bateria e estender a faixa de uso do produto. Descarregar a bateria à um nível profundo de descarga não é usualmente recomendado assim como irá reduzir significativamente sua vida útil. Em contraste, carregar a bateria além dos limites estipulados repercutirá em aquecimento excessivo, o que levará novamente a redução do ciclo de vida, inclusive, podendo ocasionar incêndios. (Samadani; Fraser; Fowler, 2012).

Pelo fato dos IBSs e BMSs determinarem vários parâmetros das baterias simultaneamente, existe o desafio de implementar cada um destes bloco de modo que a estimação some em primeiro lugar, uma aproximação correta, uma vez que a determinação destes dados de forma incorreta pode produzir danos aos bancos de baterias reduzindo sua vida útil ou até mesmo causar explosão; em segundo lugar, busca-se sempre um baixo esforço computacional, visto que os sensores devem ser embutidos nos veículos de forma a ocupar o menor espaço possível, além de utilizar um *hardware* de baixo consumo de forma a deixar o equipamento adequado a competitividade do mercado.

Compreender o processo de envelhecimento da bateria é uma tarefa complexa, pois muitos fatores, desde as condições ambientais até a utilização do veículo interage para gerar diferentes efeitos de envelhecimento (Vetter *et al.*, 2005). A degradação da bateria é acelerada com fatores que incluem, mas não são restritos: a frequência de ciclagem, grande mudança do estado de carga, grandes magnitudes de corrente durante carga e descarga, temperaturas elevadas e exposição a temperatura elevada (Smith *et al.*, 2012).

Neste contexto, as Redes Neurais Artificiais (RNAs) vêm se consolidando como uma das alternativas mais utilizadas e eficientes para realizar a identificação de sistemas dinâmicos não lineares, mais especificamente, aqueles relacionados à indicadores de saúde e desempenho de baterias. De acordo com (Haykin, 2001), as RNAs são estruturas paralelas, maciçamente distribuídas, constituídas, por unidades simples de processamento, conhecidas como neurônios, que possuem a propensão natural de armazenar conhecimento experimental e torná-lo disponível para uso.

Levando-se em consideração os dados promissores de crescimento da frota mundial de EVs, associada a busca por melhorias na forma de implementação tanto via *software* quanto via *hardware* de mecanismos para determinação do SoC e SoH, este trabalho visa o desenvolvimento de uma RNA quando utilizada para identificar sistemas dinâmicos não-lineares, mais especialmente àqueles dedicados aos indicadores de saúde e desempenho de baterias, em cenários susceptíveis a incidência de ruídos e *outliers*. Neste cenário, as redes neurais artificiais se apresentam como uma boa opção de estrutura a ser utilizada em problemas de estimação de sistemas não lineares, classificação, predição de séries temporais e controle (Sjöberg *et al.*, 1995). Neste trabalho, as equações de ajuste dos parâmetros da RNA utilizadas pelo algoritmo de retropropagação do erro são modificadas para se adequarem ao novo critério de avaliação, a correntropia.

Enquanto o tradicional critério de avaliação baseado no Erro Médio Quadrático

(MSE) limita-se a analisar apenas a variância da distribuição do erro, o critério de Maximização da Correntropia (MCC) permite que momentos estatísticos de ordem superior possam ser considerados durante o procedimento de aprendizagem, tais como assimetria, curtose, entre outros. Isto é realizado com o intuito de que, estes momentos podem conter informações dinâmicas relevantes que deveriam ser inseridas na estrutura do modelo. Por esta razão, esta medida torna-se mais adequada para distribuições de erro não gaussianas e quando os dados de conjunto de treinamento encontram-se contaminados por ruídos e *outliers*. (Bessa; Miranda; Gama, 2009);(Miranda; Cerqueira; Monteiro, 2006).

No entanto, o sucesso da aplicação da métrica da correntropia nos mais diversos cenários depende do adequado ajuste da largura de seu *kernel* e coeficiente de aprendizado. A largura de *kernel* é um parâmetro ajustável da correntropia que funciona como uma lente de *zoom* capaz de controlar a janela de observação na qual a similaridade entre duas variáveis é avaliada. Na literatura, podemos encontrar alguns métodos que realizam o ajuste desse parâmetro, como a regra de Silverman (Principe, 2010) e algumas formas adaptativas (Liu; Chen, 2013). O coeficiente de aprendizado, por outro lado, está relacionado ao ajuste dos pesos da RNA de forma a convergir em erro nulo. Deste modo, neste trabalho, além de propor uma metodologia de ajuste dos parâmetros da RNA (*bias* e pesos) baseado neste critério de Maximização da Correntropia, também será analisado o desempenho do treinamento da RNA para diferentes valores fixos de largura de *kernel* e coeficiente de aprendizado.

1.1 Objetivos

1.1.1 Objetivo Geral

Este trabalho tem como objetivo desenvolver um estimador do Estado de Carga e Saúde para baterias de lítio-íon utilizando Redes Neurais Artificiais Recursivas com função de custo baseada no critério de Maximização da Correntropia (MCC).

1.1.2 Objetivos Específicos

Como objetivos específicos podem ser citados:

- Desenvolver modelos computacionais baseados em técnicas de inteligência artificial para estimação do SoC e SoH de baterias Li-íon;
- Correlacionar as grandezas medidas e disponíveis em uma bateria para a construção dos modelos, considerando entradas de tensão da bateria, corrente de descarga, temperatura ambiente e valores anteriores de SoC e SoH;

- Propor modificações na estrutura da RNA utilizada para determinação do SoC e SoH considerando a presença de ruídos e *outliers*.

1.2 Contribuições do Trabalho

Neste trabalho, tem-se como objetivo o desenvolvimento de uma metodologia de treinamento de uma rede neural artificial que utiliza como função de custo a Maximização da Correntropia para a estimação do estado de carga e saúde de uma bateria de Li-Ion. No entanto, o sucesso da aplicação da métrica da correntropia nos mais diversos cenários depende do adequado ajuste da largura de seu *kernel* σ , coeficiente de aprendizado η e dos seus demais hiperparâmetros.

Com o propósito de determinar de maneira adequada estes parâmetros nesta metodologia, de acordo com seu cenário de utilização prático, são propostos dois cenários com duas respectivas otimizações destes parâmetros (σ e η).

1. Em cenários na ausência de ruídos e *outliers*, os parâmetros σ e η (largura de *kernel* e coeficiente de aprendizado) são determinados através de estratégias de algoritmos evolutivos. Os algoritmos evolutivos têm se mostrado capazes em determinar soluções ótimas para problemas complexos e não lineares. Como trata-se de um cenário sem ruído ou *outliers*, pretende-se obter os valores mais adequados de σ e η que produzam erro mínimo no sistema de estimação.
2. Em cenários na presença de ruídos e *outliers*, o parâmetro σ é definido através de uma estratégia adaptativa, enquanto o coeficiente de aprendizado é mantido fixo. A escolha por uma solução adaptativa para a definição do valor da largura de kernel, deve-se ao fato de que a escolha de um valor fixo para esse parâmetro que seja eficiente, normalmente, não é uma tarefa simples. Principalmente, quando a correntropia é utilizada em problemas não lineares e na presença de algumas características estocásticas.

Os resultados foram obtidos a partir de uma base de dados de ensaios de baterias advinda de um repositório online mantido pelo centro de pesquisa da NASA (Saha *et al.*, 2011).

Entre as contribuições deste trabalho estão:

1. Ao contrário da função clássica baseada no Erro Quadrático Médio (MSE), que é um critério relacionado apenas com o momento de segunda ordem do erro, o critério de Maximização da Correntropia (MCC) permite que momentos estatísticos de ordens superiores sejam levados em consideração durante o ajuste de parâmetros do

modelo, tais como a assimetria e curtose. Estes momentos podem conter informações dinâmicas relevantes que deveriam ser inseridas na etapa de aprendizagem do modelo.

2. Com a inserção do MCC, existe a possibilidade de que um conjunto de valores muito discrepantes (*outliers*) possa ser desconsiderado ou eliminado, através da seleção adequado da largura de *kernel*.

1.3 Organização do Trabalho

Além deste Capítulo Introdutório, este trabalho é composto por mais quatro capítulos, em que são descritos todo o embasamento teórico para utilização da correntropia como função de custo para estimação do estado de carga e saúde de baterias, bem como as problemáticas associadas aos métodos de estimação atuais.

No Capítulo 2, *a priori* é realizada uma revisão bibliográfica em torno da problemática e estado da arte de estimação do estado de carga e saúde de baterias.

No Capítulo 3 são apresentados fundamentos teóricos importantes para melhor compreensão do trabalho. Inicialmente, são descritos princípios básicos da identificação de sistemas dinâmicos. Em seguida, são apresentados alguns conceitos que constituem a base das redes neurais artificiais. Desta forma, torna-se mais clara a compreensão da arquitetura básica de uma RNA. Dando continuidade ao Capítulo 3, são apresentadas as definições da correntropia como uma medida de similaridade e a sua respectiva representação matemática. Além disso, são listadas as suas propriedades de maior relevância para este trabalho.

O Capítulo 4 é destinado a apresentar a proposta de utilização de correntropia no treinamento das redes RNA. No início do capítulo, é descrita a arquitetura de estimação não-linear utilizada neste trabalho bem como a relação que pode ser estabelecida entre a medida de correntropia e identificação de sistemas dinâmicos (SoC e SoH). Na sequência, a estrutura dos modelos da RNA adotada neste trabalho é devidamente detalhada, bem como são apresentadas as expressões matemáticas aplicadas pelo algoritmo de treinamento de retropropagação do erro para ajustar os parâmetros da rede RNA, considerando a utilização da correntropia em sua função de custo. Neste capítulo também é ressaltada a importância do algoritmo da largura de *kernel* da correntropia, que, na prática, corresponde a um parâmetro adicional do algoritmo de treinamento desenvolvido.

No Capítulo 5, os resultados obtidos nos estudos de caso adotados para avaliar a aplicação de correntropia no treinamento de redes RNA são devidamente detalhados. Finalmente, no Capítulo 6, são apresentadas as conclusões deste trabalho e as perspectivas para o desenvolvimento de futuras pesquisas.

2 Estado da Arte

2.1 Estado de Carga (SoC)

2.1.1 Problemática

O aumento dos preços do petróleo bruto e a conscientização mundial sobre as questões ambientais resultaram em um maior desenvolvimento de sistemas de armazenamento de energia. A bateria é um dos sistemas de armazenamento de energia mais atraentes devido à sua alta eficiência e baixa poluição (Chang, 2012). Existem vários tipos de baterias atualmente em uso na indústria: bateria de Chumbo-ácido, bateria Ni-MH, bateria Ni-Cd e a bateria Lítio-íon. A bateria tem a vantagem de ter alta tensão de trabalho por célula, baixa poluição, baixa taxa de auto descarga e alta densidade de potência. As baterias são comumente usadas para dispositivos portáteis, veículos elétricos híbridos e aplicações industriais (Rao; Wang; Zhang, 2011).

Com o desenvolvimento de veículos elétricos, as tecnologias relacionadas ao sistema de gerenciamento da energia tem-se mostrado de vital importância nos últimos anos. Um dos principais problemas consiste em desenvolver algoritmos para controlar o processo de carga é um parâmetro muito importante para uma estratégia de controle (He; Xiong; Guo, 2012). No processo de carga e descarga, existem muitas relações eletroquímicas contidas na bateria, o que dificulta a determinação de seu estado interno em tempo real. A estimativa exata do estado de carga é útil não somente para a otimização da faixa de uso da bateria (evitando processos de sobrecargas e descargas profundas), mas também para prover informações ao condutor do veículo elétrico por exemplo, sobre a eficiência do veículo, estimar o tempo de carga, julgar o estado de saúde da bateria e permitir que o sistema de gestão faça estratégias de controle racionais para economizar energia (Du; Cheng, 2014),(Li Xue *et al.*, 2014),(Cai; Liu; Luo, 2010).

Uma determinação errônea do estado de carga em sistemas de armazenamento de energia pode resultar em uma acumulação de energia insuficiente, acarretando em uma possível falta de carga em momentos onde o sistema de geração venha a reduzir sua produção. Por outro lado, ao ser gerada uma quantidade de energia acima dos limites de segurança das baterias, é possível que a mesma produza reações químicas que resultem em super aquecimento ou explosão.

Em uma pesquisa de mercado conduzida pelo grupo TNS envolvendo cerca de 7 mil usuários de *Smartphones* em 15 países, mais de 75% das respostas indicaram que a melhoria do ciclo de vida da bateria é a principal ferramenta almejada em um futuro aparelho (Xu; Wang; Yang, 2010). A precisão do SoC permite que a bateria seja utilizada

dentro de seus limites estipulados, e assim, o pacote físico da bateria não precisa ser sobredimensionado. Isso permite em uma bateria mais leve e menor, o que proporciona um custo de produção menor para o fabricante. No entanto, muitos exemplos práticos de baixa precisão e confiabilidade são encontrados na prática, o que tende a ser bastante inconveniente, especialmente quando o dispositivo portátil para de funcionar subitamente mesmo apresentando capacidade suficiente na bateria. Outro exemplo prático é que o usuário é induzido a carregar a bateria todo dia, mesmo quando é indicada capacidade suficiente disponível no aparelho eletrônico. Isto leva a um processo mais frequente de recarga do que o estritamente necessário, gerando um desgaste precoce na bateria.

2.1.2 Estado da Arte

Em muitas aplicações, a acurácia, eficiência e robustez são os principais recursos a serem considerados durante a estimativa do SoC (Hussein, 2015), (Hussein *et al.*, 2012), (Al-Haj Hussein; Batarseh, 2011). Outro fato importante é que, as características da bateria são alteradas de acordo com as condições de operação e envelhecimento. Consequentemente, os erros de modelagem precisam ser minimizados no algoritmo de estimativa. Para determinação do SoC, diferentes métodos de medição direta e indireta, bem como sistemas adaptativos estão atualmente disponíveis (Tertuliano *et al.*, 2020). Quanto mais preciso o SoC é, mais robusto o sistema será e o desempenho será melhorado.

Uma boa estimativa do SoC é possível utilizando lógica *Fuzzy*, redes neurais e filtros de Kalman, porém tais métodos exigem um elevado poder computacional, o que pode ser um problema para sistemas embarcados (Plett, 2004). O método mais utilizado na literatura para estimação do SoC, devido a sua simplicidade, é *Coulomb Counting*. Este método dispõe de um caminho simples para estimar as variações do SoC. No entanto, além da necessidade de se ter um SoC inicial, e a medição de corrente livre de ruídos e *outliers* ser uma tarefa complicada que pode afetar os resultados, a acurácia deste método sob altas temperaturas e correntes de flutuação é baixa.

Na literatura científica, existe uma variedade de alternativas para se determinar o SoC das baterias, de tal modo que obter uma classificação para cada método não é um procedimento simples, uma vez que a maioria dos estimadores apresentam uma combinação de dois ou mais métodos e a inclusão de diferentes ferramentas matemáticas heurísticas e determinísticas em sua construção (Rivera-Barrera; Muñoz-Galeano, 2017). Baseado na classificação realizada em (Waag; Fleischer; Sauer, 2014), a classificação do SoC pode ser feita subdividindo seus métodos em três categorias.

1. Métodos diretos: A estimação do SoC através de métodos diretos é efetuada por meio da medição e uso de parâmetros físicos da bateria como corrente, tensão nos

terminais e temperatura. São exemplos desta metodologia o *Coulomb Counting*, *Look-up Tables*, método de impedância, etc.

2. Sistemas adaptativos: A fundamentação de sistemas adaptativos se encontra na combinação de métodos diretos e a metodologia de estimação baseada em modelos de circuitos equivalentes. Estas técnicas geralmente contém em sua estrutura uma realimentação para ajuste dos parâmetros de saída de acordo com a variação de entrada. Á título de exemplo: filtro de Kalman, redes neurais artificiais, lógica *Fuzzy* e redes neuro-*Fuzzy*.
3. Métodos Híbridos: Os métodos híbridos se beneficiam das vantagens de cada método isolado de estimativa do SoC, o que permite uma estimativa mais robusta a problemas usualmente encontrados no estado da arte. A literatura mostra que os métodos híbridos geralmente produzem uma boa estimativa do SoC, em comparação com os métodos individuais.

Em (Saji; Babu; Ilango, 2019) uma combinação de *Coulomb Counting* com lógica *Fuzzy* foi proposta para aprimorar a acurácia da estimação do SoC. Tal método consiste em conectar as entradas do sistema de identificação as saídas através de lógica *Fuzzy*. Ao combinar a lógica *Fuzzy* com um banco de dados apropriado, as não linearidades no comportamento das baterias podem ser reduzidas em grande extensão. Apesar de ser uma abordagem simples para um problema complexo, este método não leva em consideração as perdas de energia nem o envelhecimento da bateria e a conseqüente perda de capacidade, sendo recomendado efetuar regularmente uma calibração das baterias, e implica que se conheça o estado de carga inicial. Outro ponto importante é que a corrente de operação é o único sinal medido, portanto a medição do sinal deve ser resetada periodicamente caso algum erro na estimação ocorra.

Para superar essa tradicional limitação, em (Haoran *et al.*, 2018) foi proposto um algoritmo de estimação baseado em filtro de Kalman Estendido, o qual supera as deficiências da correção do valor inicial do SoC e pode estimar com precisão o estado de carga da bateria. Neste método, a bateria é vista como um sistema de potência e é descrita como um circuito equivalente. O erro máximo obtido neste algoritmo proposto é cerca de 4%, e o erro médio em torno de 2%. Vale salientar que, o efeito de um aumento de corrente de descarga levando a um aumento na temperatura da bateria não foi levado em consideração neste estudo. Caso considere a influência da temperatura, a precisão do modelo será aumentada, o erro será reduzido e a estimação estará mais próxima do valor de referência. No entanto, seu desempenho é extremamente dependente do modelo elétrico equivalente adotado, o que reflete diretamente na sua complexidade.

Em (Yan; Wang, 2017) é proposto um modelo de previsão do SoC com base em rede neurais artificiais utilizando algoritmos de treinamento *Back-Propagation* ou BP. A

fim de melhorar a precisão do modelo, este artigo propõe um algoritmo genético para otimizar os pesos e bias da rede neural artificial. Como resultado, pôde-se ver que a rede neural otimizada apresenta convergência mais rápida do que as correspondentes e forte adaptabilidade. No entanto, uma desvantagem do modelo é a utilização do MSE como função de custo no algoritmo de treinamento. Ao utilizar este critério, momentos estatísticos de ordem maiores não são levados em consideração durante o ajuste dos parâmetros do modelo, tais como assimetria, curtose, entre outros.

No cenário prático atual, existem duas principais diretrizes de pesquisa para estimação do SoC. A primeira é o método não linear de filtragem baseado no modelo equivalente da bateria, e subsequente o método de inteligência artificial com habilidade de autoaprendizagem (Aslan; Yasa, 2019), que é o método de filtro de Kalman e o método de rede neural artificial, respectivamente. No entanto, conforme mencionado acima, ambas as direções têm suas desvantagens. Se combinadas entre si ou com diretrizes de outros métodos, eles podem ser capazes de superar suas respectivas deficiências.

Desta forma, as informações desta seção podem ser resumidas da seguinte forma.

Vantagem	Desvantagem
Abordagem simples ¹ .	Não considera informações como perda de capacidade, perda de energia e outros problemas de erro polarizados de sensores.
Valores precisos de estimação ² .	A precisão da estimação depende da complexidade do sistema de potência adotado.
Ótimos valores de estimação, não necessita de conhecimentos dos processos intrínsecos da bateria ³ .	A precisão da estimação depende da robustez do banco de dados para treinamento e da função de custo adotada.

Onde ¹ refere-se ao Coulomb Counting, ² ao Filtro de Kalman e ³ Redes Neurais Artificiais.

2.2 Estado de Saúde (SoH)

2.2.1 Problemática

Além do uso de baterias de Lítio-íon em dispositivos portáteis, elas são consideradas como a tecnologia-chave para a próxima geração de veículos (Goodenough; Kim, 2010),(Scrosati; Garche, 2010) e aproximadamente todos os fabricantes de automóveis já introduziram em sua frota um ou mais veículos que utilizam sistemas de acionamentos elétricos. Com o aumento da produção e venda de novos veículos elétricos, o mercado de baterias de energia também inaugurou um crescimento substancial. Apesar das vantagens

majoritárias desta tecnologia, tais como emissão nula de gases poluentes e independência do mercado de petróleo, ainda existem problemas em relação à vida útil das baterias de automóveis. Com o uso generalizado da bateria de Lítio-íon, os problemas em relação a segurança surgiram. Acidentes como a combustão espontânea e explosão causaram sérios danos a segurança pessoal e patrimonial. A fim de melhorar a usabilidade à longo prazo das baterias, bem como sua segurança em aplicações de veículos elétricos, é essencial reunir mais conhecimento sobre os processos químicos que contribuem para seu envelhecimento, assim como monitorar seus o *status* de funcionamento, segurança e confiabilidade em tempo real.

O SoH é um indicador de saúde que determina a "idade" da bateria e é comumente definida como a capacidade de descarga de uma célula envelhecida em comparação com a capacidade de descarga da mesma célula quando era nova. Um $SoH < 80\%$ é comumente considerado como critério de fim de vida para uma bateria (Zhou *et al.*, 2015).

A bateria é um sistema altamente não-linear. A potência que pode ser fornecida por uma bateria eletroquímica é limitada pelas reações químicas que ocorrem dentro da célula. Do ponto de vista elétrico, o principal fator que limita a potência fornecida ou absorvida por uma bateria é a resistência interna. Com o aumento da resistência interna, a capacidade de potência de (des)carga diminui de forma não-linear. Praticamente, a resistência interna é uma função do SoC, temperatura, corrente de operação e SoH. (Hussein, 2017). Em um nível de SoC moderado (entre 30% e 70%), a resistência interna é relativamente baixa e conseqüentemente, a capacidade de potência da célula é alta. Nos extremos do SoC, a capacidade de potência da bateria é relativamente baixa devido ao aumento da resistência interna. Da mesma maneira que, para baixas temperaturas, a resistência interna é alta, enquanto para altas temperaturas, a resistência interna é baixa, isto explica o porque do aumento de potência da célula com a temperatura. Outro ponto é que, com o envelhecimento da célula, a resistência interna decai, resultando na perda de capacidade de potência.

Compreender o envelhecimento da bateria é complicado, pois muitos fatores, desde as condições ambientais até a utilização do veículo interagem para gerar diferentes efeitos de envelhecimento (Ashwin *et al.*, 2018). A degradação da bateria é acelerada com fatores que incluem, mas não são restritos: a frequência de ciclagem, grande mudança no SoC, grandes magnitudes de corrente durante a carga e descarga, temperaturas elevadas e exposição a temperatura elevada (Uddin *et al.*, 2017).

O peso desconhecido e o comportamento não-linear de cada um destes fatores faz com que a estimação precisa do SoH se torna uma tarefa um tão quanto complicada (Barré *et al.*, 2013). É frequentemente difícil estabelecer um razoável e preciso modelo matemático para o processo de carga e descarga. Neste contexto, a rede neural apresenta um bom desempenho quando o quesito é estimativa de problemas com características não-lineares.

2.2.2 Estado da Arte

Uma das tarefas mais importantes do Sistema de Gerenciamento de Bateria (BMS) é estimar os estados da bateria, que incluem principalmente o estado de carga e o estado de saúde. Comparado com a tecnologia de estimativa do SoC, que progride muito atualmente, a metodologia de estimação do SoH não é discutida de maneira ampla nas diversas literaturas atuais. No entanto, alguns pesquisadores têm realizado alguns estudos promissores (Salkind *et al.*, 1999). Em (Topan *et al.*, 2016) é proposto uma metodologia de estimação do SoC e SoH, cujo estudo, o circuito equivalente de Thévenin é usado para descrever a característica de polarização e comportamento dinâmico da bateria, sendo estimada através do filtro de Kalman. No entanto, a acurácia deste método é extremamente dependente do modelo equivalente da bateria adotado.

O método mais convencional de estimativa do SoH é baseado na degradação da capacidade no decorrer do tempo. Este método pode prever o estado de saúde da bateria com relativa precisão, porém é lento e demorado, uma vez que requer muito tempo para completar um ou mais ciclos de carga/descarga. Isso o torna inadequado para estimação de SoH em baterias cuja degradação da capacidade ocorre de maneira rápida. Ser capaz de prever a degradação rápida da bateria é um elemento importante para uma operação segura e confiável. Outro método é relacionar a deterioração da capacidade e o aumento da resistência interna CC à medida que o ciclo de vida da bateria decorre. Ao utilizar um conjunto de dados do processo de envelhecimento, isto é, os resultados de medição da resistência interna CC *offline* e a degradação da capacidade, um modelo estático é construído para prever o indicador de saúde da bateria. No entanto, existem vários pontos fracos ou desvantagens para este método. Um é que o número de ciclos de trabalho pelos quais a bateria passa não estão disponíveis na maioria das aplicações. Outro é que o modelo estático é válido somente para certas condições de operação ou para uma única bateria, e não tem a habilidade de auto ajuste para se adaptar a alterações nas condições de operação, o que limita seu escopo de aplicação. Uma terceira é que a medição da resistência interna CC é feita de maneira *offline*, o que a torna imprecisa para aplicações *online* em tempo real.

A variação do SoH reflete concretamente a variação da resistência interna, capacidade e também outros indicadores. Estudos extensivos foram realizados para a estimativa no SoH nos últimos anos. A resistência interna da bateria foi usada para monitorar o estado de saúde em (Remmlinger *et al.*, 2011),(Galeotti *et al.*, 2015). Esses artigos analisaram a relação entre a resistência e capacidade, e apresentaram um baixo esforço computacional e boa robustez, mas a desvantagem é que poucos fatores influentes na degradação da capacidade são considerados, tais como frequência de ciclagem da bateria, temperaturas elevadas, entre outros. O SoH pode ser estimado por métodos probabilísticos e estatísticos, que são mostrados em (Cheng *et al.*, 2015),(Feng *et al.*, 2013). Este tipo de método

combina conhecimento matemático com conhecimento empírico para construir um modelo empírico ou semi-empírico, o que demandará muito tempo para obter dados experimentais. (Eddahech *et al.*, 2012) apresentou um modelo baseado EIS (Impedância Eletroquímica) e descreve um método de monitoramento de SoH, que usa rede neural recorrente para prever a deterioração no desempenho da bateria. Em (Zhou *et al.*, 2015), é empregado o algoritmo de aprendizado GRNN. O GRNN é um algoritmo de aprendizado bastante avançado e apresenta uma estrutura de rede flexível, propiciando sua aplicação em várias situações de comportamento não-linear. Os resultados mostram um erro médio e variância média entre o SoH estimado e referência abaixo de 4%. Desta forma, a rede neural é uma forma eficaz de estimativa do SoH.

A rede neural artificial em si, é um tipo de modelo estatístico que precisa de um grande número de dados experimentais para treinamento a fim de adquirir resultados com níveis de erro satisfatórios (na ordem de 10^{-3}). A RNA é constituída por simples neurônios interconectados. A rede pode obter as leis inerentes e intrínsecas que regem os dados e apresenta habilidades de mapeamento não-lineares. Assim, a RNA é adequada para problemas imprecisos, de identificação, julgamento, interferência e adaptativa para uma complexa relação causa-efeito. Não é preciso saber de maneira acurada as reações eletroquímicas e é treinada por um algoritmo de aprendizagem repetidamente até que o erro de estimativa seja menor do que o limite predefinido. Para as condições de excitação externas, a rede neural artificial apresenta uma boa habilidade de aprendizado. E pode simular a não-linearidade da bateria para estimação do SoH (Nagahamulla; Ratnayake; Ratnaweera, 2011).

Desta forma, os principais métodos de estimação do estado de saúde podem ser resumidos da seguinte forma.

Vantagem	Desvantagem
Precisão relativamente boa ¹ .	Método lento e demorado de estimação.
Velocidade de estimação rápida ² .	Demonstra robustez somente para certas condições de operação.
Baixo esforço computacional e boa robustez ³ .	Não leva em consideração a degradação da capacidade, influências de temperatura, entre outros..

Onde ¹ refere-se ao Método da degradação da Capacidade, ² refere-se ao Modelo Estático e ³ ao Método da Resistência.

2.2.3 Conclusões

Neste capítulo, foram apresentadas as principais referências literárias de estimação do estado de carga e estado de saúde. Em relação à esta literatura, as Redes Neurais apresentam uma alternativa eficiente tanto para a estimação do SoC quanto para estimação do SoH. O método proposto além de apresentar um conjunto vasto de dados de treinamento presente em um repositório *online* mantido pela NASA, apresenta uma ferramenta que permite uma maior absorção da quantidade de informações presentes no conjunto de treinamento, a Correntropia.

3 Fundamentação Teórica

Realizar a identificação de um sistema dinâmico consiste em determinar um modelo matemático que seja capaz de representar adequadamente as relações dinâmicas existentes entre as variáveis de entrada e saída do sistema. A grande vantagem da identificação em relação, por exemplo, à modelagem fenomenológica, é que ela permite que o modelo seja obtido sem necessidade de um profundo conhecimento das leis físicas que regem o sistema. Desta forma, o procedimento de modelagem pode ser realizado baseado apenas na informação de um conjunto de dados experimentais. Isto é um ponto importante, principalmente, quando o objetivo é determinar modelos de sistemas dinâmicos complexos e não-lineares.

Entre as estruturas não-lineares comumente utilizadas durante o processo de identificação, as redes neurais artificiais merecem destaque. Incentivados pela capacidade de adaptação e aprendizado das redes neurais biológicas, modelos matemáticos de neurônios artificiais foram desenvolvidos de modo a emular a ação de um neurônio biológico e consequentemente reconhecer e prever diversos padrões diferentes. Além do mais, redes neurais NARX, por possuírem uma arquitetura neural dinâmica e recorrente, podem ser utilizadas tanto para previsões à um passo à frente quanto para a previsões a curto, médio e longo prazo (Chatterjee *et al.*, 2012).

Independente da arquitetura neural utilizada no procedimento de identificação, o ajuste de seus parâmetros livres é realizado por algum algoritmo de treinamento que, usualmente, tem como princípio, minimizar uma determinada função de custo, à título de exemplo, o erro médio quadrático. Devido ao objetivo deste trabalho de avaliar a aplicação da rede RNA na identificação de indicadores de desempenho e saúde de baterias, este critério pode ser substituído, por exemplo, pela correntropia. Esta modificação é justificada pelo fato de que, teoricamente, esta medida de similaridade apresenta maior robustez aos efeitos indesejáveis dos grandes valores de erros decorrentes dos *outliers* do que o erro médio quadrático (Munoz; Chen, 2012), bem como permite que momentos estatísticos de ordem superior possam ser considerados durante o processo de treinamento.

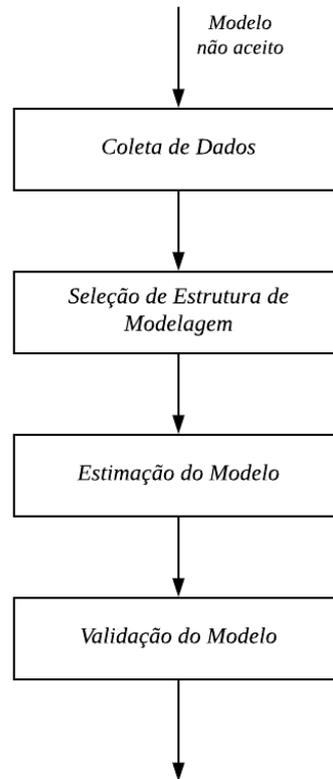
Neste capítulo, inicialmente, são apresentados alguns conceitos básicos na identificação de sistemas dinâmicos, incluindo as etapas comumente utilizadas por este procedimento. Em seguida, são resumidamente apresentados alguns conceitos teóricos das redes neurais artificiais. Por fim, são apresentadas a definição matemática da correntropia e algumas de suas importantes propriedades.

3.1 Identificação de Sistemas Dinâmicos

Segundo (Norgaard *et al.*, 2001), identificar um sistema dinâmico corresponde a inferir um modelo matemático que represente a sua dinâmica a partir de uma série de medições de suas variáveis de entrada e saída. (Aguirre, 2007) afirma que estes modelos têm sido utilizados para os mais diversos fins, por exemplo: projeto de sistemas de supervisão e controle, predição, estimação de estados, simulação e treinamento, detecção de falhas, desenvolvimento de sistemas de inferência, entre outros (Nelles, 2001). Para casos em que a identificação seja baseada exclusivamente em dados experimentais, sem nenhum ou pouco conhecimento sobre os fenômenos físicos envolvidos, o processo de identificação é conhecido como identificação caixa-preta, ou modelagem empírica. Sabe-se que, de forma generalizada, as redes neurais correspondem a estruturas matemáticas não-lineares cujos parâmetros podem ser ajustados a partir de um processo supervisionado, comumente referido como processo de aprendizado ou treinamento. Esse processo, normalmente, é realizado a partir de um conjunto de dados formados por pares entrada-saída de um determinado problema. Desta forma, pode-se notar por essas similaridades que a tarefa de obter um modelo dinâmico de um sistema não-linear, utilizando a identificação caixa-preta, pode ser realizada por algum tipo de rede neural artificial (Linhares; de Araújo; Araújo, 2013).

A Figura 1 apresenta as principais etapas realizadas durante o procedimento de identificação do tipo caixa preta (Norgaard *et al.*, 2001);(Aguirre, 2007);(Linhares *et al.*, 2014). A primeira etapa consiste em intervenções no sistema com o objetivo de se obter um determinado conjunto de amostras que descrevam o comportamento dinâmico do sistema dentro da sua faixa de operação. A ideia, de maneira geral, consiste em variar os sinais de entrada e observar a influência dessa variação em suas saídas. Durante esta etapa, usualmente, o sistema é excitado por sinais de entrada ricos em diversidades de frequência, tais como PRS (*Pseudo Random Signal*) e PRBS (*Pseudo Random Binary Signal*).

Figura 1 – Etapas principais do procedimento de identificação.



Fonte: Autor (2020).

Na etapa seguinte, é necessário selecionar aquela estrutura de modelagem que será utilizada durante o processo de identificação. Essa escolha deve ser realizada com base nas características e propriedades do sistema dinâmico em questão, bem como análise das vantagens e desvantagens de cada uma das estruturas avaliadas. Após a seleção da estrutura de modelagem, o próximo passo consiste em obter um modelo específico desta estrutura que melhor represente o sistema em estudo, levando em consideração a otimização de parâmetros internos ao modelo ou ainda a susceptibilidade do sistema a ruídos. Neste momento, os conjuntos de dados coletados na etapa inicial do processo de identificação são utilizados para estimar as saídas do modelo de acordo com algum critério ou função de custo, como o erro médio quadrático.

O avanço da inteligência artificial e o crescente aumento de desempenho dos computadores contribuíram para que as RNAs se tornassem uma das mais populares técnicas utilizadas em aplicações de identificação não-linear. Entretanto, na realidade, os principais motivos para que isto ocorresse encontram-se nas capacidades de generalização, aprendizado e robustez que as redes neurais possuem. Estas características permitem que as RNAs possam ser utilizadas com sucesso para representar uma ampla classe de sistemas não-lineares. Além disso, outro atrativo é que o procedimento básico para realização da

identificação caixa-preta assemelha-se bastante aos passos seguidos no desenvolvimento de projetos utilizando redes neurais.

3.2 Redes Neurais e Estruturas Neurais de Identificação

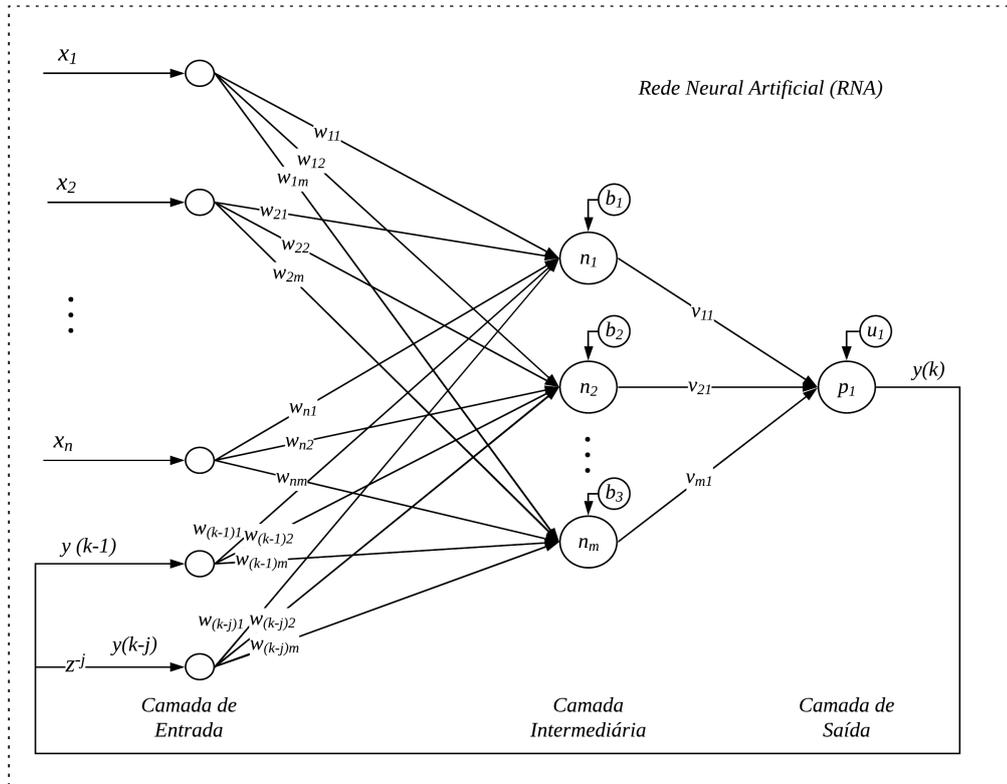
Como explica (Sriram, 1997), as RNAs são técnicas computacionais que geram modelos inspirados na estrutura neural de organismos inteligentes, e que adquirem conhecimento através da análise de experiências prévias. Em sua forma mais geral, uma RNA é uma estrutura projetada para funcionar de maneira similar à forma como o cérebro realiza uma determinada tarefa de interesse. Objetivando atingir um bom desempenho, as RNAs adotam uma estratégia de interligação maciça de células computacionais simples, denominadas neurônios ou unidades de processamento (Haykin, 2001).

As RNAs constituem um paradigma computacional diferente do convencional, que tem por base um elemento processador central controlando o sistema. No paradigma neural o processamento é feito de forma distribuída, através dos neurônios artificiais. Os elementos processadores das RNAs operam de maneira paralela, interagindo-se entre si. A rede aprende a resolver uma tarefa que lhe é atribuída através de um algoritmo de treinamento. O funcionamento do modelo depende da dinâmica dos neurônios e da forma como eles são conectados, o que irá determinar o tipo de tarefa que será realizado pela rede (Corrêa, 2004).

O funcionamento das RNAs se baseia no estabelecimento de relações entre estímulos de entrada e saída de um sistema, e apresentam a vantagem de que, para este fim, não é necessário estabelecer um modelo matemático prévio que defina a forma destas relações. As redes aprendem estas relações dos próprios dados, e a partir de um processo de treinamento, de forma semelhante ao aprendizado do cérebro humano (Lorenzi *et al.*, 2017).

A arquitetura das redes se assemelha à do cérebro em dois aspectos: O conhecimento é adquirido pela rede a partir de registros das condições de entrada e saída de um processo qualquer, por meio de um processo de aprendizagem, que pode ser supervisionado ou não-supervisionado. Pesos sinápticos, que representam as forças de conexão entre os neurônios que formam a rede, são utilizadas para armazenar o conhecimento adquirido.

Figura 2 – Representação de uma RNA recursiva com duas camadas ocultas.



Fonte: Autor (2021).

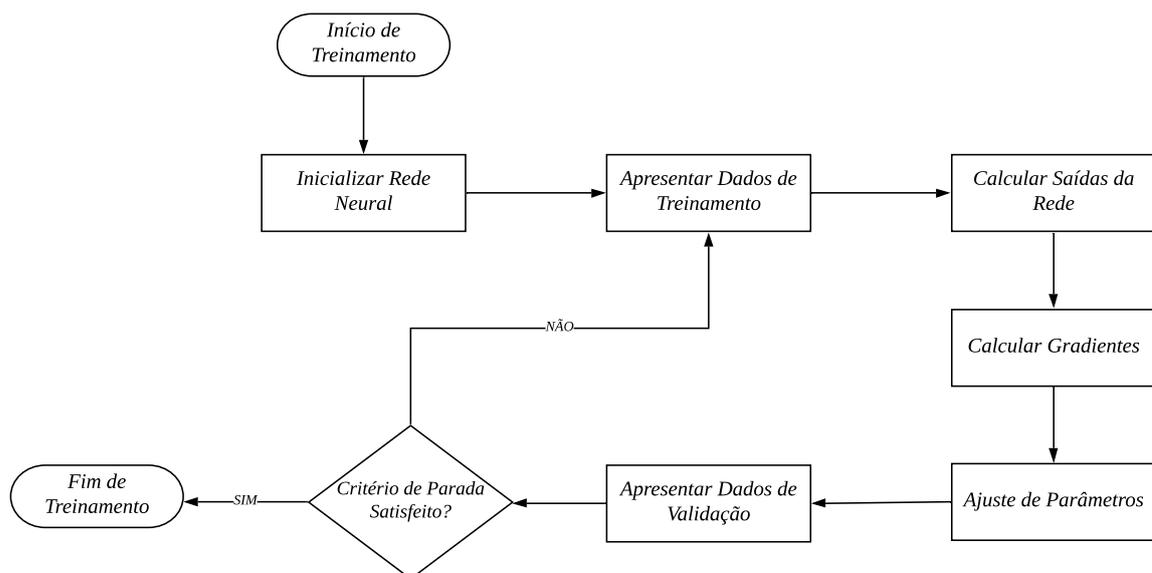
Uma rede típica é constituída por uma camada de entrada (*input layer*), composta de tantos neurônios quanto sejam necessários para codificar as informações conhecidas; de uma ou mais camadas intermediárias (*hidden layers*), que permitem a transformação da informação de acordo com uma estrutura de pesos que foi estabelecida quando a rede foi treinada; e de uma camada de saída (*output layer*), que registra o resultado do processamento realizado. Cada unidade da rede, ou neurônio, guarda apenas um valor de estado, que se modifica dependendo dos estímulos recebidos de todos os neurônios que o antecedem e estão ligados ao mesmo. Ilustra-se na Figura 4 o funcionamento básico de uma rede neural artificial de três camadas. Como representado na figura, cada neurônio da camada de entrada recebe uma série de sinais de entrada, ou estímulos, cada qual ativado com um determinado peso. Os estímulos são combinados através de uma função aditiva, que pode ser influenciada por uma tendência (*bias*) introduzida no sistema. O resultado é filtrado por uma função de ativação. Esta gera um sinal de saída, com uma certa intensidade, que servirá de estímulo para o próximo neurônio. Este tipo de rede neural, permite a recursividade da saída, isto é, a saída da rede neural pode ser realimentada a neurônios da mesma camada ou de camadas anteriores, provendo uma memória dinâmica.

Para implementar uma RNA utilizam-se redes de neurônios, estruturas capazes de

representar superfícies de decisão mais complexas. Conforme (Haykin, 2001), o treinamento é normalmente realizado de forma supervisionada, com o emprego do algoritmo denominado como Retropropagação do Erro. Este algoritmo se baseia na realimentação do erro. Esse procedimento permite que os pesos sinápticos vão sendo ajustados, de trás para frente, buscando reduzir o erro no final de seu treinamento. O mesmo envolve dois passos: a propagação e a retro-propagação. Na propagação, um vetor de entrada é aplicado aos nós sensoriais da rede e seu efeito é propagado através da rede, de forma a produzir um conjunto de estímulos de saída, que caracterizam a resposta da rede. Uma das propriedades mais importantes das RNAs é a habilidade de generalizar o aprendizado, isto é, respondendo a entrada inéditas com bom desempenho. A aprendizagem é um processo pelo qual os parâmetros livres da RNA são adaptados, através de um processo de estimulação, ao ambiente na qual a rede está inserida. Em função da maneira pela qual a modificação dos parâmetros ocorre, se determina qual a estratégia de será utilizada.

Usualmente, para realizar o ajuste de parâmetros da RNA, é aplicado algum algoritmo de treinamento baseado em gradiente. Desta forma, as características da dimensão da estrutura RNA devem ser definidas (número de entradas, número de camadas intermediárias e de saída) antes de executar o procedimento de aprendizagem, e os conjuntos de dados de treinamento e validação devem ser obtidos. Posteriormente, as etapas representadas na Figura 3 e descritas logo na sequência podem ser efetuadas para realizar o treinamento da rede (Linhares *et al.*, 2014).

Figura 3 – Procedimento de treinamento da rede neural artificial.



Fonte: Autor (2021).

- **Inicialização da rede:** Os pesos sinápticos da estrutura são inicializados. De

acordo com (Thuillard, 2001), a inicialização dos parâmetros adaptativos da rede tem influência significativa na velocidade de convergência e desempenho do algoritmo de treinamento;

- **Apresentação dos dados de treinamento:** O conjunto de dados experimentais que compõe o conjunto de treinamento são apresentados para a rede neural artificial. Estes exemplos de treinamento, são representados por um vetor de entrada e uma saída;

Com o intuito de minimizar os problemas provenientes do uso de unidades e faixas de valores distintos entre os parâmetros, é necessário que os dados estejam normalizados, padronizados em uma única escala.

Outro fator que impõe que o conjunto seja normalizado é que, uso de rede neurais com função de ativação sigmoideal exige que os dados que alimentam a rede sejam normalizados para valores pertencentes ao intervalo de 0 a 1. Desta forma, o conjunto de dados de entrada da rede foi totalmente representado dentro do intervalo de 0 a 1, já que a função de ativação sigmoideal foi utilizada. A normalização é realizada através da Equação 3.1:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (3.1)$$

em que: x representa uma matriz com as respectivas entradas do sistema de identificação: para esta pesquisa a corrente, tensão de descarga e temperatura de operação.

- **Cálculo da saída da rede:** Para cada exemplo, as entradas de treinamento são fornecidas para a rede e suas respectivas saídas são calculadas após o fluxo de dados cruzar todas as camadas internas da rede neural;
- **Cálculo do gradiente:** As saídas da rede obtidas na etapa anterior são comparadas com as saídas desejadas fornecidas pelos dados de treinamento. Com base no erro e da função de custo adotado no algoritmo de treinamento, calcula-se o valor do gradiente para cada um dos parâmetros adaptativos da rede a cada amostra de exemplo apresentada;
- **Ajuste dos parâmetros:** As informações do gradiente são utilizadas para realizar o ajuste dos parâmetros da rede RNA, tendo como objetivo, otimizar a função de custo do algoritmo de treinamento;
- **Apresentação de dados de validação:** Após a atualização dos parâmetros da rede, o seu desempenho é avaliado por meio da apresentação dos exemplos presentes no conjunto de validação. A realização da validação é importante para evitar a ocorrência de *overfitting* e verificar a capacidade de generalização da rede.

- **Teste de critério de parada:** É verificado se a rede neural artificial obtida até o presente momento atende aos critérios de parada. Usualmente, adota-se um valor limite máximo para o número de interações e função de custo. Caso algum destes critérios for satisfeito, o treinamento é concluído.

As etapas apresentadas até o presente momento são usualmente utilizadas em qualquer tipo de treinamento de rede neural artificial. A pesquisa de (Haykin, 2001) traz maiores detalhes e informações sobre as estratégias de treinamento supervisionado e uma análise minuciosa acerca de cada uma das etapas mencionadas.

Na breve descrição das etapas de treinamento, nota-se que este procedimento efetua a otimização de um critério de avaliação chamado função de custo, que estabelece uma forma matemática de quantificar o desempenho da rede durante o seu treinamento. Neste trabalho, é avaliada a eficiência das redes RNA na identificação de indicadores de desempenho e saúde de baterias.

3.3 Medida de Correntropia

A correntropia cruzada, ou simplesmente correntropia, corresponde a uma métrica da Teoria da Informação, originalmente desenvolvida pelo grupo de neuroengenharia da Universidade da Flórida. A correntropia é uma medida de similaridade generalizada entre duas variáveis aleatórias X e Y , sendo definida algebricamente pela seguinte expressão (Principe, 2010);(Santamaria; Pokharel; Principe, 2006):

$$v_{\sigma}(X, Y) = E[\kappa(X, Y)] = \iint \kappa(x, y) f_{XY}(x, y) dx dy \quad (3.2)$$

em que: f_{XY} é a função densidade de probabilidade conjunta de X e Y , $E[.]$ é o operador estatístico e $k[.]$ é um kernel simétrico positivo. O *kernel* mais popular utilizado na correntropia é o *kernel* gaussiano, o qual é definido por

$$\kappa(x, y) = K_{\sigma}(x_i - y_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - y_i)^2}{2\sigma^2}\right] \quad (3.3)$$

em que: σ é o desvio padrão, sendo responsável por definir a largura de *kernel*. A largura de *kernel* é um parâmetro ajustável da correntropia que funciona como uma lente de *zoom* capaz de controlar a janela de observação na qual a similaridade entre duas variáveis aleatórias é avaliada.

Em (Santamaria; Pokharel; Principe, 2006), esse parâmetro pode ser interpretado como um fator de resolução pelo qual a correntropia mensura a similaridade em um espaço de alta dimensionalidade. A possibilidade de ajustar o tamanho de *kernel* fornece um

mecanismo que permite que valores muito discrepantes (*outliers*), em um conjunto de dados possam ser desconsiderados ou eliminados (Fontes, 2012).

De maneira geral, a função de densidade e probabilidade estabelecida pela Equação 3.2 é desconhecida, e usualmente, apenas uma quantidade de dados finita de dados está disponível. Diante disso, a correntropia pode ser calculada pelo estimador da correntropia amostral que é dado pela Equação 3.4 (Principe, 2010).

$$v_{\sigma}(X, Y) = \frac{1}{N} \sum_{i=1}^N \kappa_{\sigma}(x_i, y_i) \quad (3.4)$$

Uma característica importante da correntropia pode ser demonstrada através da manipulação matemática da Equação 3.4. Substituindo o *kernel* gaussiano pela sua expansão em série de Taylor na equação citada e assumindo que todos os momentos da função de densidade de probabilidade são finitos, a medida de correntropia pode ser expressa por (Santamaria; Pokharel; Principe, 2006):

$$v_{\sigma}(X, Y) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k \sigma^{2k} k!} E[(X - Y)^{2k}] \quad (3.5)$$

Analisando a Equação 3.4, a presença do termo $E[(X - Y)^{2k}]$ indica que a correntropia contém informação de todos os momentos estatísticos de ordem par da diferença entre as variáveis aleatórias X e Y (Principe, 2010). De acordo com (Santamaria; Pokharel; Principe, 2006), a correntropia é uma generalização da medida de correlação entre sinais aleatórios, pois tal medida consegue extrair tanto informações estatísticas de segunda ordem, quanto de ordem superiores entre os sinais analisados. Embora, por definição, a correntropia seja semelhante à correlação. (Santamaria; Pokharel; Principe, 2006) afirma que é possível demonstrar que esta medida de similaridade é mais adequada do que a correlação ao trabalhar com sistemas não lineares e ambientes não gaussianos, sem ocasionar qualquer aumento significativo no custo computacional.

3.3.1 Propriedades da Correntropia

Nesta seção serão apresentadas três propriedades essenciais para a compreensão da correntropia e de sua aplicação no treinamento de sistemas adaptativos, como as RNAs.

- Propriedade 1: para *kernels* simétricos, a correntropia é simétrica. Esta propriedade implica que $v_{\sigma}(X, Y) = v_{\sigma}(Y, X)$;
- Propriedade 2: a correntropia é positiva e limitada. Os limites da correntropia quando o *kernel* gaussiano é utilizado é dado por $0 < v_{\sigma}(X, Y) < 1/\sigma\sqrt{2\pi}$, atingindo o seu valor máximo se, e somente se, $X = Y$;

- Propriedade 3: para o *kernel* gaussiano, a correntropia é uma soma ponderada de todos os momentos estatísticos pares da variável aleatória $(X - Y)$.

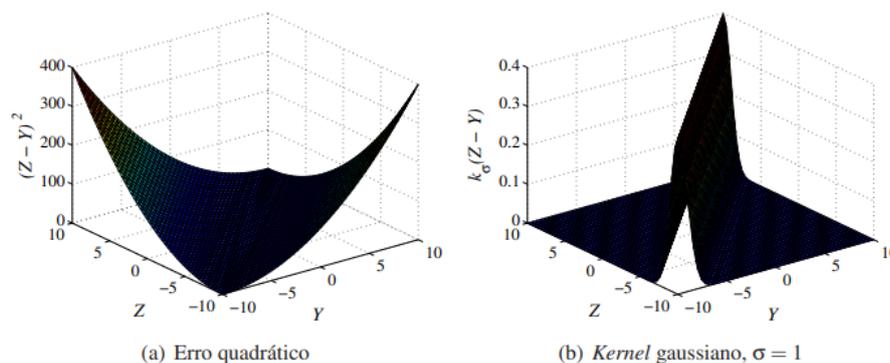
Descrições detalhadas sobre diversas propriedades da correntropia podem ser encontradas nos trabalhos de (Liu; Chen, 2013), (Santamaria; Pokharel; Principe, 2006) e (Principe, 2010). Segundo Príncipe (2010), é importante que, se a largura de *kernel* for suficientemente grande, os momentos estatísticos de segunda ordem dominam a função e a medida se aproxima da correlação enviesada. Adicionalmente, a correntropia é sensível a processos aleatórios variantes no tempo, além de se mostrar robusta a ruídos, uma vez que o produto interno calculado pelo *kernel* gaussiano tende a zero (Santamaria; Pokharel; Principe, 2006).

3.3.2 Correntropia e Erro Médio Quadrático

Os parâmetros de uma rede neural, são ajustados por meio de um algoritmo de treinamento que tem como objetivo minimizar uma determinada função de custo. A escolha desta função é tão relevante que, em (Principe, 2010), diversos tipos de função de custo são avaliadas quando utilizadas em conjunto com o tradicional algoritmo de retropropagação de erro. (Waag; Fleischer; Sauer, 2014) citam algumas características que justificam a popularidade dessa função de custo: primeiramente, ela apresenta uma superfície de erro suave, é continuamente diferenciável, e por último, possui um único mínimo global e o seu cálculo não exige elevado esforço computacional.

A Figura 4 apresenta uma comparação gráfica entre o erro médio quadrático e o *kernel* gaussiano. Essas funções são utilizadas como base pelo MSE e pela correntropia, respectivamente, para efetuar a comparação, por exemplo, entre as variáveis aleatórias Z e Y . Como pode ser visto na Figura 4, o *kernel* gaussiano demonstrado neste exemplo apresenta desvio padrão (largura de *kernel*) unitário.

Figura 4 – Comparação gráfica entre o erro quadrático médio e *kernel* gaussiano.



Analisando o gráfico na Figura 4, pode-se observar que o MSE gera um aumento quadrático para valores afastados da reta $z = y$, amplificando a contribuição de pontos que estão afastados do valor médio da distribuição do erro ϵ . Em razão disso, resíduos com distribuição gaussiana garantem a otimalidade do critério MSE . Em razão disso, resíduos com distribuição gaussiana garantem a otimalidade do critério MSE . Em contrapartida, resíduos com distribuições não-gaussianas tornam o critério MSE não ótimo e, em particular, se as distribuições de erro apresentarem *outliers*, forem assimétricas ou possuírem média não nula (Principe, 2010). Na correntropia, no entanto, a contribuição de cada amostra para o seu valor depende da largura de seu *kernel*. Uma vez que esta medida atua localmente, ela pode-se tornar particularmente útil em casos onde a distribuição do erro não é gaussiana (Bessa; Miranda; Gama, 2009).

3.4 Conclusões

Neste capítulo, foram apresentadas definições importantes na identificação de sistemas dinâmicos, incluindo os passos básicos para a execução deste procedimento. Em relação às estruturas neurais de identificação, demonstrou-se que estas correspondem a generalizações de tradicionais modelos lineares. Uma maior atenção foi dada à estrutura NARX, por se tratar da estrutura adotada como base para obtenção dos modelos RNA deste trabalho.

Por fim, apresentou-se a definição, conceitos básicos e propriedades da correntropia, medida de similaridade da Teoria da Informação que, neste trabalho, substitui o erro médio quadrático como função custo no treinamento das redes RNAs. Deste modo, a partir do conteúdo apresentado neste capítulo, é possível entender como uma rede neural pode ser utilizada para identificar um sistema dinâmico, compreender como as redes RNA são estruturadas e obter uma fundamentação teórica necessária para a devida compreensão da proposta deste trabalho, cujos detalhes serão apresentados no próximo capítulo.

4 Metodologia Proposta

Este capítulo tem como objetivo propor a utilização da Correntropia no treinamento das redes RNAs. Com este propósito, esta rede é utilizada na identificação de sistemas dinâmicos não lineares em presença de ruído e *outliers*. Para realizar a análise de desempenho do algoritmo de treinamento e do modelo RNA obtido pelo procedimento de identificação, são utilizados dois estudos de casos, ambos no âmbito de indicadores de baterias, o estado de carga e o estado de saúde. Primeiramente, é proposto o desenvolvimento de um metodologia de implementação de redes RNA com função de custo baseada na Teoria da Informação, Correntropia. O conteúdo das seções subsequentes apresentam as duas técnicas de treinamento e validação. A primeira técnica de validação consiste na utilização de algoritmos evolutivos com o objetivo de estimar adequadamente os parâmetros (largura de *kernel* σ e coeficiente de aprendizado η) e conseqüentemente, obter a redução do erro. A segunda técnica de validação consiste na utilização de estratégias adaptativas com o objetivo de tornar mais robusto o algoritmo de estimação, quando este, esta sujeito à presença de *outliers* e ruídos. Ambas técnicas são baseadas no conceito de correntropia como função de custo para RNAs. O diferencial de ambas técnicas se encontra na sua aplicação.

4.1 Método Proposto - Arquitetura de Identificação

A utilização de RNAs para a determinação do SoC e SoH de baterias é uma prática amplamente empregada na literatura. A estimação destes indicadores baseada em redes neurais tem exibido um excelente desempenho, em comparação com outros tipos de formas de estimação do SoC, como *Coulomb Counting* e o *Open Circuit Voltage*, para estimação do SoC, ou na utilização de modelos estáticos para estimar o SoH. As principais vantagens da aplicação de redes neurais na identificação dos indicadores de *performance* e saúde das baterias são:

- O modelo considera a bateria como uma caixa preta sem conhecer os complicados mecanismos intrínsecos da bateria e direciona diretamente os dados para previsão do SoC e SoH.
- O modelo pode aproximar qualquer relação não linear, então pode ser utilizado para simular a mudança de comportamento não-linear da bateria de Li-íon.
- Cada vez que a bateria passa por diferentes condições de operação ou ciclos de carga/descarga, os dados de envelhecimento podem ser realimentados para a RNA

para que ela aprenda com os novos dados e autoajuste seus parâmetros do modelo para se adaptar a estas mudanças externas.

Em (Ismail *et al.*, 2017), foi aplicada uma metodologia na qual a arquitetura de identificação foi baseada no conceito da topologia NARX, na qual foram utilizadas como entradas da rede: Corrente e tensão de descarga no instante atual e passado, para estimação do SoC de uma bateria de Lítio-íon genérica, com função de custo fundamentada em MSE. De modo que, este modelo não é capaz de prever com exatidão o SoC para valores de temperatura ambiente diferentes do valor aplicado em seu processo de treinamento. Além do fato que o MSE é um critério relacionado apenas com o momento de segunda ordem da distribuição do erro, portanto, minimizar o erro quadrático é equivalente a minimizar a variância da distribuição. Ao utilizar este critério, momentos estatísticos de ordens maiores não são levados em consideração durante o ajuste dos parâmetros do modelo, tais como a assimetria, curtose, entre outros. Estes momentos podem conter informações dinâmicas relevantes que deveriam ser inseridas na estrutura do modelo, uma vez que as distribuições gaussianas contêm todas as suas informações somente nos dois primeiros momentos estatísticos da distribuição. Logo, se a distribuição do erro não for gaussiana, a obtenção de um modelo a partir de MSE negligência informações presentes em momentos de ordens superiores (Bessa; Miranda; Gama, 2009);(Miranda; Cerqueira; Monteiro, 2006).

O conceito de correntropia pode ser aplicado à identificação de sistemas dinâmicos, cujo objetivo é estimar um modelo capaz de representar, de forma satisfatória, as relações dinâmicas existentes entre as variáveis de entrada e de saída do sistema a ser identificado. Dessa maneira, o principal objetivo é fazer com que, para os mesmos valores de entrada, a saída do modelo obtido durante o processo de identificação, seja o mais próximo possível das saídas do sistema dinâmico.

De acordo com (Liu; Chen, 2013), um modelo de identificação pode ser interpretado como uma função desconhecida $f(X, \theta)$, que para um conjunto de N amostras de entradas $X = x_1, x_2, \dots, x_N$ e um vetor de n parâmetros ajustáveis $\theta_m = [\theta_1, \theta_2, \dots, \theta_n]^T$, fornece um conjunto de P saídas estimadas $Y_m = y_1, y_2, \dots, y_P$. Estes parâmetros θ_m , no contexto deste trabalho, podem ser definidos como os Pesos e Bias de uma RNA típica.

Portanto, esta medida pode ser utilizada, por exemplo, para expressar o quanto as saídas do modelo, $Y_m = f(X, \theta_m)$, assemelham-se às saídas Y do sistema. Neste contexto, a correntropia pode ser aplicada para avaliar o desempenho de modelos obtidos pelo procedimento de identificação.

Uma propriedade essencial para a compreensão da correntropia e de sua aplicação no treinamento de sistemas adaptativos, como RNAs é a de que:

1. A correntropia é positiva e limitada. Os limites da correntropia quando o kernel

gaussiano é utilizado é dado por $0 < v_\sigma(Y_m, Y) \leq 1/\sigma\sqrt{2\pi}$, atingindo o seu valor máximo se, e somente se, $Y_m = Y$.

Sob a definição matemática de correntropia apresentada na Equação (4.1) e desta propriedade de correntropia apresentada, conclui-se que quanto maior for o valor da correntropia, melhor é a qualidade do modelo encontrado, e conseqüentemente, menores são os erros de estimação. Sob este ponto de vista, a identificação do sistema dinâmico pode ser interpretada, portanto, como um procedimento que tem como objetivo ajustar os parâmetros de treinamento do sistemas adaptativos θ_m para maximizar a correntropia, conforme apresentado pela Equação 4.1 (Liu; Chen, 2013).

$$M_{CC} = \underset{\theta_m}{\operatorname{argmax}} \frac{1}{N} \sum_{k=1}^N K_\sigma(Y, f(X, \theta_m)) = \underset{\theta_m}{\operatorname{argmax}} \frac{1}{N} \sum_{k=1}^N K_\sigma(e_k) \quad (4.1)$$

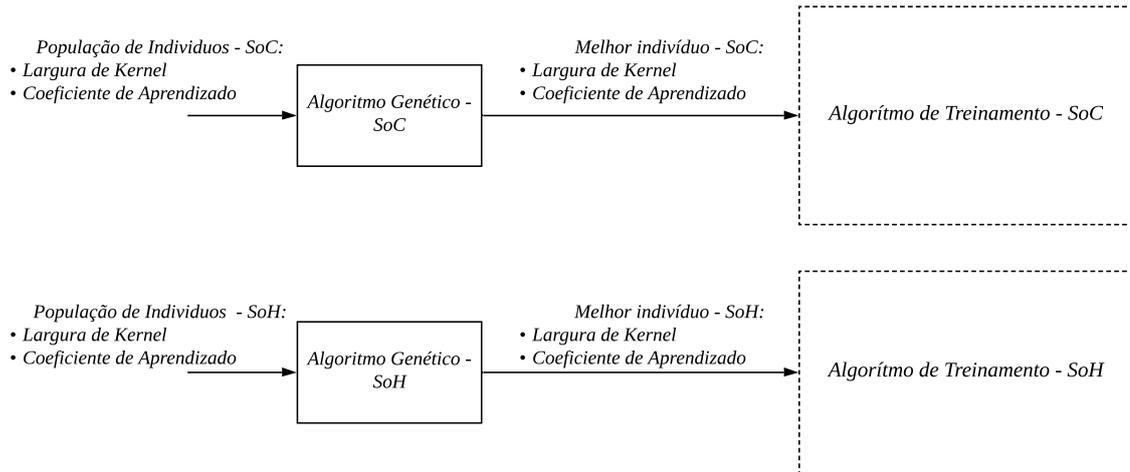
em que: $e_k = y^* - y$ são os erros de estimativa gerados pelo modelo para N amostras experimentais utilizadas na identificação. Essa expressão é conhecida na literatura como Critério de Maximização da Correntropia, ou *Maximum Correntropy Criterion* (MCC) (Liu; Chen, 2013),(Bessa; Miranda; Gama, 2009),(Principe, 2010). Esse conceito é aplicado neste trabalho, com o intuito de substituir o erro médio quadrático no algoritmo de retropropagação do erro. A partir dessa modificação, o algoritmo é aplicado para ajustar os parâmetros do modelo RNA utilizados na identificação de indicadores de performance em baterias, porém, este modelo pode ser generalizado na identificação de qualquer sistema dinâmico sujeito à presença de ruídos e *outliers*.

Com este intuito, o presente trabalho analisa como uma rede RNA com critério de avaliação baseado em Correntropia, se comporta quando utilizada para identificar sistemas dinâmicos não lineares, mais especificamente, àqueles associados ao indicadores de desempenho e saúde das baterias. Enquanto o tradicional MSE limita-se a analisar a variância da distribuição do erro, a correntropia permite que momentos estatísticos de ordem superior possam ser considerados durante o procedimento de aprendizagem. Por esta razão, esta medida torna-se mais adequada para distribuições de erro não-gaussianas e quando os dados do conjunto de treinamento encontram-se contaminados por ruídos e *outliers*. Comparado com a previsão do modelo estático do SoH, o método de previsão proposto tem maiores propabilidades de ter resultados precisos sob condições de operação variadas.

A utilização da correntropia como função de custo no treinamento de redes RNA é avaliada, neste trabalho, a partir da identificação de dois sistemas dinâmicos não-lineares. Como citado anteriormente, ambos estão atrelados ao contexto de desempenho e saúde de baterias. Desta forma, para o treinamento da RNA com conjunto de dados sem a presença de ruídos e *outliers*, é adotada uma estratégia de otimização dos parâmetros de

largura de *kernel* (σ) e coeficiente de aprendizado (η), que é executada previamente ao algoritmo de treinamento baseada em algoritmo genético. Logo, o algoritmo de treinamento é inicializado já com os parâmetros otimizados (largura de *kernel* σ_{best} e coeficiente de aprendizado η_{best}) de forma a minimizar o erro de estimação, conforme a Figura 5.

Figura 5 – Otimização dos parâmetros η e σ para algoritmo de treinamento sem influência de ruídos e *outliers*.



Fonte: Autor (2021).

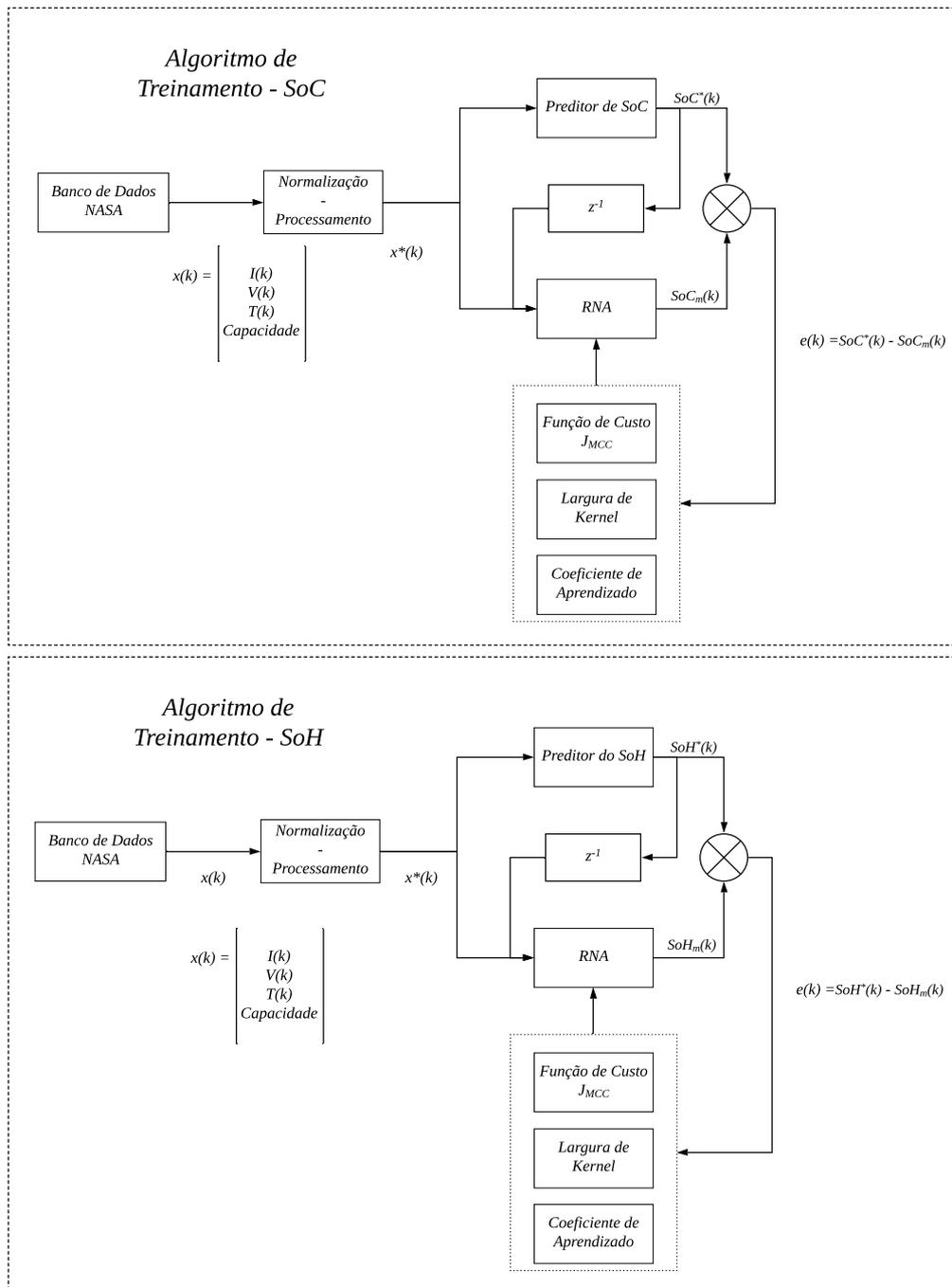
A arquitetura de treinamento apresentada para estimação do SoC e SoH na Figura 6 baseia-se na estrutura de identificação utilizando RNAs descrita em Narendra e Parthasarathy (1990), apresentando algumas adequações devido às características das RNAs, como por exemplo, ao conjunto de dados experimentais utilizados, ao procedimento de aprendizado adotado para ajustar os parâmetros dos modelos RNA e a inserção do algoritmo genético com ferramenta de otimização dos parâmetros intrínsecos à esta metodologia adotada.

Com o intuito de minimizar os problemas provenientes do uso de unidades e faixas de valores distintos entre os parâmetros, é necessário que os dados estejam normalizados, padronizados em uma única escala. Como a função de ativação utilizada na RNA é a tangente sigmoide, os dados que alimentam a rede foram normalizados para valores pertencentes de 0 a 1 de acordo com a Expressão 3.1.

No algoritmo de treinamento para o estado de carga, apresentado na Figura 6, o **Preditor de SoC** corresponde a estimação do estado de carga por *Coulomb Counting*, de acordo com a seguinte expressão:

$$SoC = SoC_0 + \frac{100}{C_n} * \int_{t_0}^{t_0+\tau} Idt \quad (4.2)$$

Figura 6 – Diagrama esquemático da arquitetura de identificação utilizada para sistemas sem ruído e outliers.



Fonte: Autor (2021).

em que: o SoC_0 representa o estado de carga inicial igual a 1, C_n corresponde a capacidade nominal da bateria igual 2150mAh, I representa o fluxo de corrente na bateria, t_0 é o instante de tempo que a carga ou descarga é iniciada na bateria e τ corresponde ao intervalo de tempo entre o começo e fim do ciclo. No preditor do SoC, somente a corrente de operação I e o tempo relativo do ciclo de descarga, proveniente do banco de dados, é utilizada.

O **Preditor de SoH** corresponde a estimação do estado de saúde. O SoH é uma razão da capacidade atual da bateria em relação a capacidade nominal, conforme representado na seguinte expressão:

$$SoH = \frac{C_{atual}}{C_{nominal}} \quad (4.3)$$

em que C_{atual} representa a capacidade da bateria após cada ciclo de carga/descarga e C_n representa a capacidade nominal da bateria. No preditor de SoH, somente a capacidade de cada ciclo de carga/descarga é utilizada.

As entradas do modelo RNA são definidas de acordo com um vetor, que é formado por valores atuais do sinal de entrada (Corrente de operação, Tensão de descarga e Temperatura) e valor passado do sinal de saída do sistema $SoC^*(k-1)$ para estimação do SoC e $SoH^*(k-1)$ para estimação do SoH . Demonstra-se na Figura 6 demonstra que os parâmetros do modelo RNA são ajustados por meio de um algoritmo de treinamento, que é executado a partir do sinal de erro de estimação do modelo $e(k)$. O algoritmo escolhido para realizar esta tarefa é o algoritmo de retropropagação de erro, tendo como função de custo o critério de maximização da correntropia. Ao utilizar este critério, além do coeficiente de aprendizado η , um novo parâmetro é incluído no algoritmo de treinamento: a largura de *kernel*.

O sucesso da aplicação da correntropia depende da escolha adequada da largura de seu *kernel* gaussiano. Portanto, esse parâmetro pode influenciar no desempenho, na taxa de convergência e na robustez do algoritmo de treinamento (Zhao; Chen; Príncipe, 2011),(Liu; Chen, 2013). Segundo (Zhao; Chen; Príncipe, 2011), a largura de *kernel* deve ser definida buscando-se um equilíbrio entre rejeição de *outliers* e eficiência na estimação do modelo.

Na literatura são encontradas algumas abordagens que podem ser aplicadas para determinar a largura de *kernel* da correntropia, por exemplo, a regra de (Silvermann, 1986). Apesar de ser bastante utilizado, esse método não é capaz de fornecer um valor ótimo para esse parâmetro. Normalmente, ele é utilizado para indicar um valor de referência para posterior refinamento (Fontes, 2012). O coeficiente de aprendizado, por outro lado, está relacionado ao ajuste dos pesos da RNA de forma a convergir em erro nulo. Este ajuste é realizado em passos de tamanho definidos. O valor desse passo mínimo é

denominado coeficiente de aprendizagem. Um valor alto da taxa de aprendizado pode significar aproximar do objetivo final mais rápido, mas também pode criar oscilações entre o ponto de aprendizado ideal. De forma similar, um valor muito pequeno pode tornar o treinamento muito lento, contudo pode convergir para o melhor conhecimento possível.

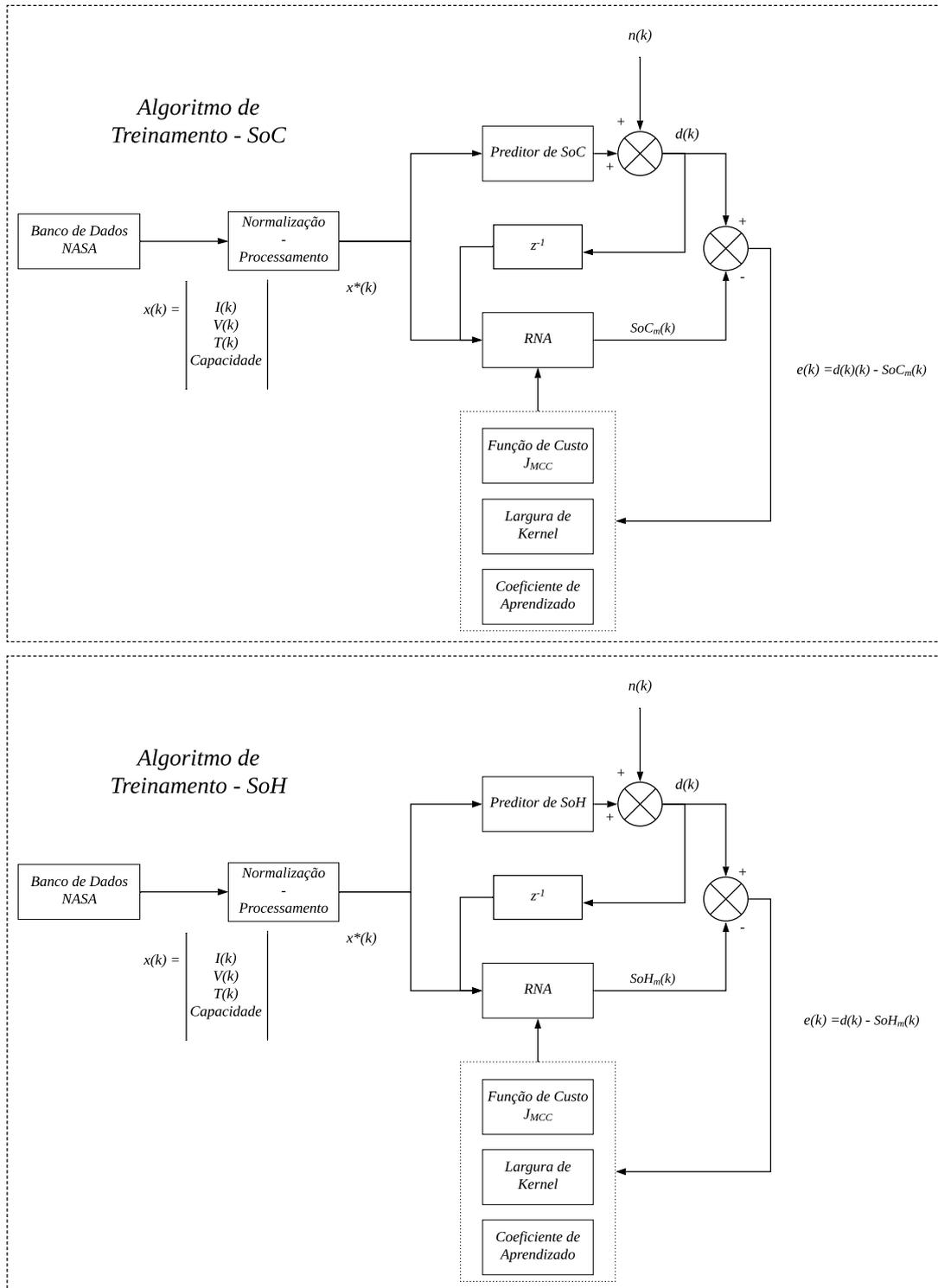
A topologia proposta para identificação de sistemas com a presença de ruído e *outliers* é exposta na Figura 7. cujas entradas do sistema de identificação são definidas de acordo com um vetor, de maneira similar a topologia exposta na Figura 6, com a diferença de que, neste caso, as entradas são formadas por valores atuais do sinal de entrada e valor passado do sinal de saída do sistema $SoC^*(k-1)$ e $SoH^*(k-1)$, sujeito à presença de ruídos e *outliers*. A inserção de ruídos e *outliers* é representado pela entrada $n(k)$. Para esta estrutura proposta de identificação, durante o treinamento da rede, é adotada uma estratégia adaptativa para determinar a largura de *kernel*. A estratégia adotada pela otimização de algoritmo genético não é a mais adequada com a presença de ruídos e *outliers*, visto que, o AG foi proposto com o intuito de determinar os parâmetros (σ e η) de maneira a minimizar o erro. No entanto, minimizar o erro, neste contexto, seria fazer com que o algoritmo aproximasse a saída estimada à saída desejada, ou seja, incluir ruídos e *outliers* em seu processo de treinamento, o que não é adequado. Logo, esta nova estratégia é baseada na Equação 4.4 (Liu; Chen, 2013):

$$\sigma(k+1) = \frac{\max|e_i(k)|}{2\sqrt{2}}, i = 1, 2, \dots, N \quad (4.4)$$

em que: $\sigma(k)$ corresponde ao valor da largura de *kernel* determinada para k -ésima época de treinamento e $e_i(k)$ corresponde ao erro de estimativa gerado pelo i -ésimo exemplo de treinamento apresentado a rede RNA. A escolha por uma solução adaptativa para a definição do valor da largura de *kernel*, deve-se ao fato de que a escolha de um valor fixo para esse parâmetro que seja eficiente, normalmente, não é uma tarefa simples. Principalmente, quando a correntropia é utilizada em problemas não lineares e na presença de algumas características estocásticas, como são a identificação de sistemas não lineares e o procedimento supervisionado de treinamento de rede neural.

Em relação ao treinamento das redes RNA, um ponto importante a ser comentado é a forma como seus parâmetros adaptativos são inicializados. Neste trabalho, os pesos e bias das redes RNA são inicializados com valores aleatórios entre 0 e 1, obtidos a partir de uma distribuição uniforme. (Araújo Junior, 2014) afirma que, embora essa não seja a forma mais eficiente de inicializar os parâmetros da rede, ela é bastante utilizada devido a sua simplicidade. O procedimento consiste em inicializar esses parâmetros de forma que elas sejam igualmente distribuídas e com pelo menos um parâmetro adaptativo ativo em toda a faixa de operação do sistema dinâmico. Uma outra alternativa de inicialização dos pesos e *bias* das redes RNA é conhecida como inicialização de *Xavier* em homenagem ao seu desenvolvedor *Xavier Glorot* (Glorot; Bengio, 2010). Esse inicializador faz com que os

Figura 7 – Diagrama esquemático da arquitetura de identificação utilizada para sistemas com ruído e outliers.



Fonte: Autor (2021).

pesos sejam inicializados em valores que se mantêm em um alcance razoável em todas as camadas da rede neural.

O algoritmo de retropropagação do erro e os modelos RNA avaliados no presente trabalho foram implementados na linguagem de programação MATLAB (Licença do Programa CWS UFPB/2021), instalado em computador pessoal com o sistema operacional *Windows* 10 de 64 bits.

4.2 Arquitetura da Rede Neural Artificial

As RNAs, em sua grande maioria, utilizam arquiteturas projetadas para resolver classes de problemas específicos. Em sua grande parte, existem três camadas em uma rede neural: a Camada de Entrada (CE), a Camada Intermediária (CI) e a Camada de Saída (CS). Estas camadas podem ser arranjadas de tal forma a construírem redes do tipo NARX (do inglês *Nonlinear Autoregressive with Exogenous Inputs*), Auto Regressiva Não-Linear (do inglês *Nonlinear Autoregressive*) ou *Input-Output* Não-Linear.

A rede NARX é aplicada como ferramenta deste trabalho por ser uma rede dinâmica e recorrente com alto desempenho. Segundo (Xie; Tang; Liao, 2009), a rede neural NARX é considerada mais robusta do que as redes recorrentes convencionais. Sua aprendizagem pelo gradiente descendente tende a se tornar mais eficaz devido à sua memória integrada que fornece um caminho menor para propagação do gradiente de informação quando a rede é aberta ao invés de retropropagar o sinal de erro, fornecendo a rede rápida convergência e desempenho robusto.

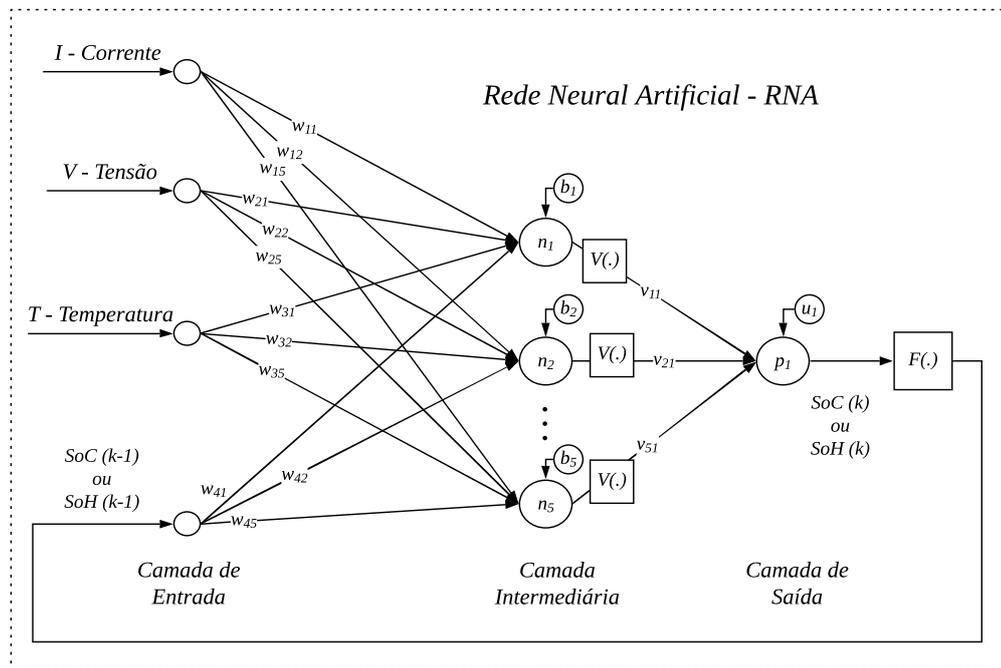
A rede neural artificial aplicada neste trabalho, apresentada na Figura 8, tem como principal característica a realimentação com valores obtidos na saída da rede, que são somados com entradas externas à rede. Desta forma, a saída é uma função das amostras k das entradas atuais em conjunto com as saídas passadas. Desta forma, a estimação do $SoC(k)$ é escrita em função dos parâmetros descritos na Equação 4.5 e o $SoH(k)$ a partir dos parâmetros escritos na Equação 4.6:

$$SoC(k) = f(SoC(k-1), I(k), V(k), T(k)) \quad (4.5)$$

$$SoH(k) = f(SoH(k-1), I(k), V(k), T(k)) \quad (4.6)$$

em que: $I(k)$ representa a corrente de operação, $V(k)$ representa a tensão de descarga, $T(k)$ a temperatura, $SoC(k-1)$ o estado de carga no instante passado e $SoH(k-1)$ o estado de saúde no instante passado.

Figura 8 – Arquitetura da Rede Neural Artificial para o processo de identificação - NARX.



Fonte: Autor (2021).

Para obtenção da rede com maior desempenho, foram realizados vários testes com diferentes parâmetros para testar a desempenho da rede neural artificial recursiva e determinar a escolha dos melhores parâmetros de forma empírica. As características da RNA são: (1) Única camada intermediária; (2) Função de ativação Sigmóide para camada intermediária, representado por $V(\cdot)$; (3) Função de ativação linear para camada de saída, representado por $F(\cdot)$; (4) Entrada com valores atuais e um atraso de saída realimentado e (5) 5 neurônios na camada intermediária.

4.3 Aplicação do MCC no Treinamento da RNA

Os neurônios artificiais são compostos por uma função soma e uma função de ativação. As redes neurais complexas possuem diversas camadas de neurônios. O diagrama da rede neural generalizada utilizada neste trabalho pode ser visualizado na Figura 8. Onde x_i representa as entradas do sistema de identificação dinâmico, w_{ij} e v_{jn} representam os pesos da camada de entrada e intermediária, respectivamente, b_j e u_n representam os *bias* de cada neurônio da camada intermediária e de saída, respectivamente, i o número de entradas, j o número de neurônios da camada intermediária e por fim n representa o número de neurônios da camada de saída. Estes são os parâmetros ajustáveis do modelo da RNA pelo modelo de retropropagação de erro baseado em Maximização da Correntropia

para estimação da saída y . A resposta de cada neurônio, bem como a resposta de saída do sistema, pode ser visualizada nas Expressões (4.7) e (4.8), respectivamente.

$$a_j = \sum x_i \cdot w_{ij} + b_j \quad (4.7)$$

$$y = \sum a_j \cdot v_{jn} + u_n \quad (4.8)$$

Viu-se que ao utilizar o critério MCC, o objetivo é maximizar a correntropia entre duas variáveis aleatórias. No procedimento de treinamento da rede RNA, essas duas variáveis correspondem à saída desejada y^* e à saída estimada y fornecida pela rede. Seja o erro de estimativa do modelo RNA dado por $e = y^* - y$, maximizar a correntropia é equivalente a minimizar a Equação (4.9).

$$J_{MCC} = \frac{1}{N} \sum_{k=1}^N \left(\frac{1}{\sigma\sqrt{2\pi}} - K_{\sigma}(e(k)) \right) \quad (4.9)$$

Em que $e(k)$ é o erro de estimativa para o k -ésimo exemplo de treinamento e N é o número de amostras coletadas do sistema. Como o algoritmo de treinamento será executado em modo sequencial. A Equação (4.10) define a expressão instantânea de J_{MCC} que será utilizada para atualizar os parâmetros pesos e bias após a apresentação de cada par de treinamento.

$$\varepsilon(k) = \frac{1}{\sigma\sqrt{2\pi}} - K_{\sigma}(e(k)) = \frac{1}{\sigma\sqrt{2\pi}} (1 - \exp^{-\frac{e^2(k)}{2\sigma^2}}) \quad (4.10)$$

Para executar o procedimento de aprendizagem da RNA por meio do algoritmo de retropropagação de erro, é necessário calcular as derivadas parciais da função custo escolhida em relação aos parâmetros ajustáveis da rede. A atualização desses parâmetros pode ser realizada a partir do princípio da regra de delta (Haykin, 2001), utilizando a como modelo geral a Equação (4.11).

$$P_{ij}(k+1) = P_{ij}(k) - \eta \frac{\partial \varepsilon(k)}{\partial P_{ij}(k)} \quad (4.11)$$

em que: $P_{ij}(k)$ representa o parâmetros ajustáveis da rede ao longo do procedimento de aprendizagem (pesos e *bias*) e η é o coeficiente de aprendizado.

Desta forma, as derivadas parciais de $\varepsilon(k)$ em relação ao parâmetros ajustáveis pesos (w_{ij} e v_{jn}) e bias (b_j e u_n) são apresentadas nas Equações (4.12) a (4.15).

$$\frac{\partial \varepsilon(k)}{\partial w_{ij}(k)} = \frac{\partial \varepsilon(k)}{\partial e(k)} \cdot \frac{\partial e(k)}{\partial w_{ij}(k)} = -\frac{1}{\sigma^3\sqrt{2\pi}} \exp^{-\frac{e^2(k)}{2\sigma^2}} \cdot e(k) \cdot x(k) \cdot v_{jn}(k) \quad (4.12)$$

$$\frac{\partial \varepsilon(k)}{b_j(k)} = \frac{\partial \varepsilon(k)}{\partial e(k)} \cdot \frac{\partial e(k)}{\partial b_j(k)} = -\frac{1}{\sigma^3 \sqrt{2\pi}} \exp^{-\frac{e^2(k)}{2\sigma^2}} \cdot e(k) \cdot v_{jn}(k) \quad (4.13)$$

$$\frac{\partial \varepsilon(k)}{v_{jn}(k)} = \frac{\partial \varepsilon(k)}{\partial e(k)} \cdot \frac{\partial e(k)}{\partial v_{jn}(k)} = -\frac{1}{\sigma^3 \sqrt{2\pi}} \exp^{-\frac{e^2(k)}{2\sigma^2}} \cdot e(k) \cdot a_j(k) \quad (4.14)$$

$$\frac{\partial \varepsilon(k)}{u_n(k)} = \frac{\partial \varepsilon(k)}{\partial e(k)} \cdot \frac{\partial e(k)}{\partial u_n(k)} = -\frac{1}{\sigma^3 \sqrt{2\pi}} \exp^{-\frac{e^2(k)}{2\sigma^2}} \cdot e(k) \quad (4.15)$$

em que: x representa as entradas do sistema de identificação dinâmico e σ a largura de *kernel*. A partir do cálculo das derivadas parciais de ε em relação aos parâmetros ajustáveis da RNA, pode-se realizar a atualização desses parâmetros a partir do princípio da regra Delta (Haykin, 2001), utilizando as Equações (4.16) a (4.19):

$$w_{ij}(k+1) = w_{ij}(k) - \eta \frac{\partial \varepsilon(k)}{\partial w_{ij}(k)} \quad (4.16)$$

$$b_j(k+1) = b_j(k) - \eta \frac{\partial \varepsilon(k)}{\partial b_j(k)} \quad (4.17)$$

$$v_{jn}(k+1) = v_{jn}(k) - \eta \frac{\partial \varepsilon(k)}{\partial v_{jn}(k)} \quad (4.18)$$

$$u_n(k+1) = u_n(k) - \eta \frac{\partial \varepsilon(k)}{\partial u_n(k)} \quad (4.19)$$

$$1 \leq i \leq 4 \quad (4.20)$$

$$1 \leq j \leq 5 \quad (4.21)$$

$$n = 1 \quad (4.22)$$

em que η é o coeficiente de aprendizado.

Em um processo supervisionado de treinamento, uma época é contada a cada apresentação à rede neural de todos os pares entradas-saídas presentes no conjunto de treinamento. Assim, época após época, a partir das Equações (4.12) a (4.15) em (4.16) a (4.19), respectivamente, o algoritmo realiza o ajuste dos parâmetros da RNA com objetivo de maximizar a correntropia por meio da Equação (4.9). É importante comentar que uma atenção especial deve ser dada à inicialização dos parâmetros Pesos e *Bias*. Uma inicialização inadequada desses parâmetros pode comprometer o treinamento, e

consequentemente, a qualidade do modelo final obtido. Uma escolha adequada incrementa a velocidade de treinamento, resultando em uma rápida convergência.

A substituição do tradicional MSE pelo MCC, utilizando como método de atualização dos parâmetros pelo método de gradiente, adiciona dois novos parâmetros de treinamento ao algoritmo de retropropagação do erro, η e σ . Como explicado anteriormente, o sucesso da utilização da correntropia depende do ajuste apropriado da largura de *kernel*, ou variância, de sua função gaussiana, e seu coeficiente de aprendizado. Desta forma, o novo parâmetro possui papel decisivo no desempenho do algoritmo de treinamento. Se uma escolha inadequada da largura de *kernel* da correntropia e coeficiente de aprendizado for feita, o uso do MCC pode não trazer resultados satisfatórios. (Principe, 2010).

4.3.1 Algoritmo Genético

O algoritmo genético (AG) é um mecanismo de otimização estocástico baseado na "sobrevivência do indivíduo mais apto". No algoritmo genético, a sobrevivência do mais apto na seleção natural é introduzida para otimizar os parâmetros. Desta forma, o algoritmo busca a solução ideal simulando o ambiente de seleção natural e o mecanismo da teoria da evolução biológica de Darwin (Darwin, 1979).

O AG, conforme vislumbrado por (Holland, 1975), consiste de uma população (conjunto de indivíduos, no qual cada indivíduo representa uma solução em potencial). Os indivíduos são compostos por um cromossomo que possui valores que podem representar a solução procurada, e estes cromossomos estão sujeitos a recombinação genética e mutação. Um processo de seleção inspirado na seleção natural se encarrega de eliminar os indivíduos pior adaptados ao problema e permitir a sobrevivência daqueles que se adaptam melhor às condições oferecidas.

Este comportamento é muito importante para aplicação em problemas de otimização, nos quais diversos parâmetros devem ser combinados para gerar a melhor solução. A inicialização aleatória dos indivíduos permite que, possivelmente, exista a melhor solução para o problema, com trechos destas soluções espalhados em diversos indivíduos. Com a seleção e recombinação genética, novos conjuntos de soluções, que combinam partes das soluções anteriores, se formam, levando os indivíduos para uma região do espaço de soluções, que possivelmente representa o espaço de busca onde os melhores resultados podem ser alcançados.

Atualmente, o algoritmo genético tem sido amplamente usado na análise combinatória, aprendizado de máquina, controle adaptativo, processamento de sinal e vida artificial, etc. (Guo; Zhao; Huang, 2017). Devido ao usuário não necessitar de um conhecimento prévio ou outras informações auxiliares, que são substituídas pela função de aptidão para avaliar o indivíduo, optou-se pelo algoritmo genético para a otimização dos parâmetros

da função de custo correntropia: largura de *kernel* e coeficiente de aprendizado, para aplicações em que o conjunto de dados se encontra ausente da presença de ruídos e *outliers*.

O AG usa o método de evolução em grupo. A solução do problema de otimização é codificada de certa forma para produzir indivíduos diferentes. Ao mesmo tempo, de acordo com a "sobrevivência do mais apto", a função *fitness* é usada para orientar a direção da procura da melhor solução. Através das operações no processo de seleção, cruzamento e mutação, uma nova geração de indivíduos são gerados. Estes passos são repetidos até a solução ótima do problema de otimização ser alcançada ou o número de gerações ultrapasse um determinado valor.

A implementação das etapas do algoritmo genético para a otimização dos parâmetros da RNA são descritos como se segue:

1. Geração aleatória de uma população de 250 indivíduos. Cada indivíduo contém informação de uma respectiva largura de *kernel* σ e um coeficiente de aprendizado η . Esta quantidade de indivíduos foi considerada suficiente para cobrir todo o espaço de busca inicial. Os indivíduos da população estão contidos dentro dos limiares:

$$0.2 \leq \sigma \leq 30 \quad (4.23)$$

$$0.2 \leq \eta \leq 15 \quad (4.24)$$

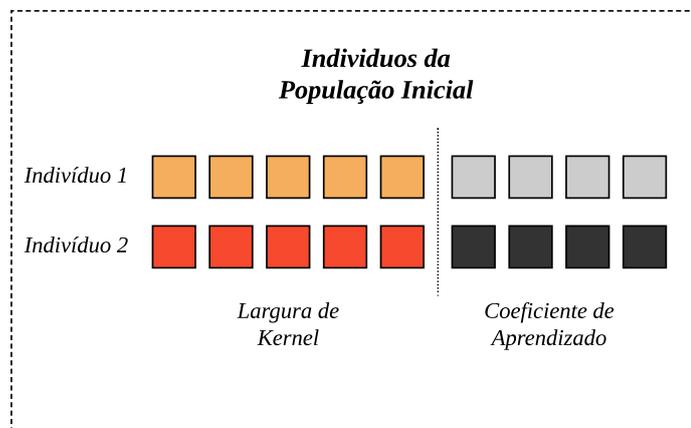
Estes limites foram definidos de forma empírica. Um valor extremamente alto para este estudo de caso ($\eta \gg 15$) na taxa de aprendizado tende a criar oscilações entre o ponto de aprendizado ideal, de forma que o ajuste dos pesos da RNA não convirjam em erro satisfatório. Um valor muito pequeno ($\eta \ll 1$) torna o treinamento do modelo muito lento, contudo o modelo converge para o conhecimento ideal. Uma ponderação entre o tempo de convergência do modelo com o erro em regime foi estabelecida para determinar a faixa de operação da taxa de aprendizado η .

A efetividade da correntropia é controlada pelo valor da largura do *kernel* Gaussiano. Este parâmetro especifica por exemplo, a menor magnitude de erro que será "desconsiderada" pelo estimador durante o ajuste de seus parâmetros. De acordo com (Bessa; Miranda; Gama, 2009) e (Munoz; Chen, 2012), valores elevados de largura de *kernel* tende a fazer com que a correntropia se comporte como o MSE, enquanto valores muito pequenos podem gerar uma região de corte tão restrita que medições corretas no conjunto de dados podem ser tratadas como *outliers* durante o processo de identificação. Assim, para aplicação proposta, definiu-se uma região de operação extensa o suficiente para incluir indivíduos com variadas larguras de *kernel*.

Após definido os limiares da população de indivíduos e a geração aleatória de 250 indivíduos na forma binária, cada indivíduo final, na sua forma binária, contém

informação tanto do coeficiente de aprendizado η quanto da largura de *kernel* σ , conforme mostra se na Figura 9. Nesta figura se ilustra a estrutura de dois indivíduos da população, em que por exemplo, os cinco bits mais significativos contêm a informação da largura de *kernel* e os quatro bits menos significativos contém informação do coeficiente de aprendizado. Este número de bits, conforme na Figura 9, não representa o valor real de bits utilizado neste trabalho. Esta figura tem como objetivo, unicamente de auxiliar a visualização do leitor da forma como um único indivíduo contém informação de duas grandezas diferentes.

Figura 9 – Indivíduos da População Inicial - Algoritmo Genético.



Fonte: Autor (2021).

O valor real do número de bits para largura de *kernel* e coeficiente de aprendizado, considerando quatro casas decimais de precisão para ambas grandezas e os limiares expostos nas Equações (4.20) e (4.21) é dada pela Equação:

$$N_{bits} = \log_2^{Lim_{sup} - Lim_{inf} * 10^4 + 1} \quad (4.25)$$

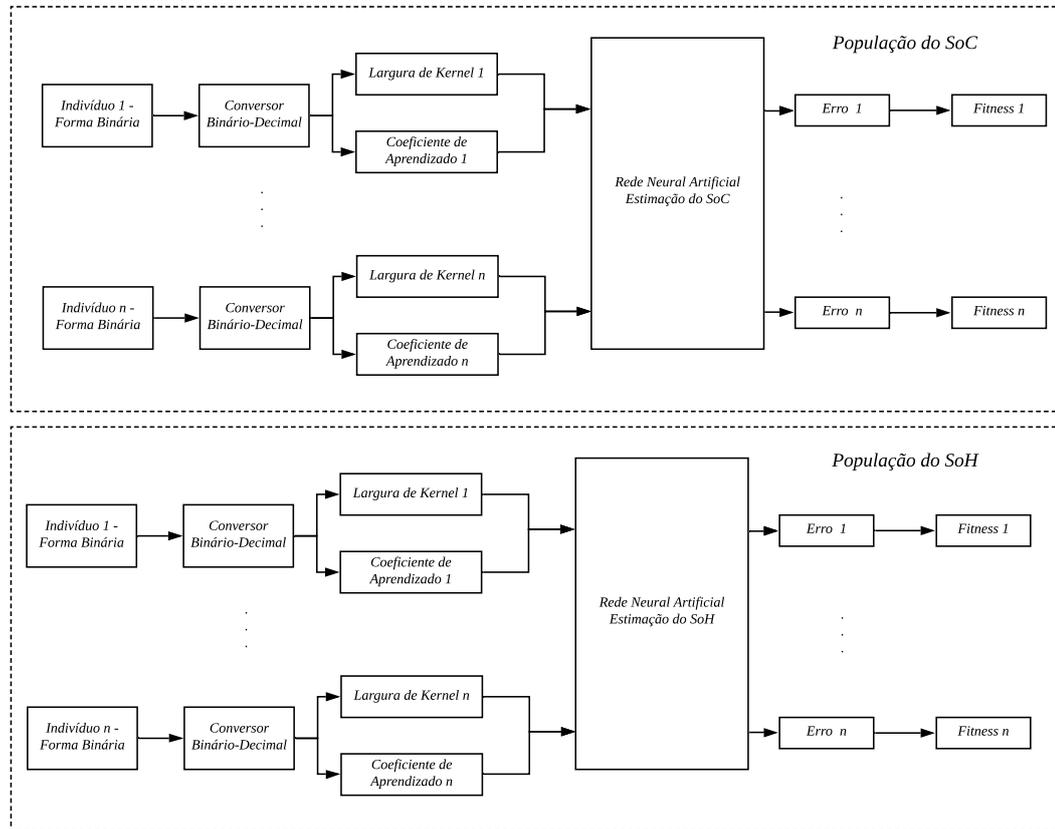
em que: Lim_{sup} representa o limiar superior da grandeza em observação e analogicamente, Lim_{inf} representa o limiar inferior.

Desta forma, o valor real de bits para o largura de *kernel* é 19, considerando que cada indivíduo, quando na forma decimal, apresenta quatro casas decimais e 18 para o coeficiente de aprendizado, apresentando o mesmo número de casas decimais. Portanto, cada indivíduo binário final, terá 37 bits em sua representação, onde os 19 bits mais significativos representam a largura de *kernel* e os 18 bits menos significativos representam o coeficiente de aprendizado.

Após a produção de 250 indivíduos na sua forma binária, é necessário analisar o *fitness* de cada indivíduo, ou de maneira análoga, analisar a resposta do par de grandezas

η e σ tem na rede neural artificial. Para executar esta análise, a informação de largura *kernel* e do coeficiente de aprendizado η são convertidos para forma decimal e aplicados a um treinamento na rede neural conforme observado na Figura 10.

Figura 10 – Determinação do *fitness* de cada indivíduo - Algoritmo Genético.



Fonte: Autor (2021).

O erro quadrático médio é dado através da Expressão (4.23):

$$Erro = \frac{\sum_{k=1}^N e(k)^2}{N} \quad (4.26)$$

em que N representa o número de amostras de treinamento e $e(k)$ o melhor erro obtido durante o processo de treinamento. Desta forma, a função de adaptação ou o *fitness*, de cada indivíduo, é calculada pela Expressão (4.27). Esta expressão procura maximizar a função de *fitness* a medida que o erro quadrático médio diminui, procurando assim, a solução otimizada.

$$Fitness = 100 - Erro \quad (4.27)$$

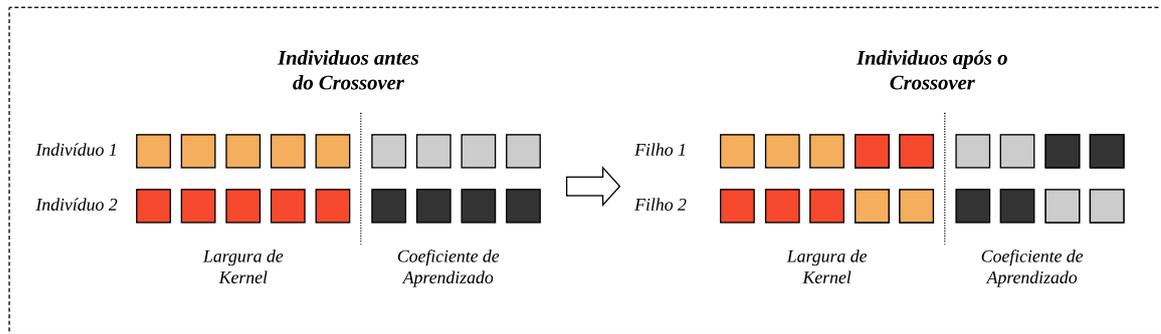
Depois de passar pela função de *Fitness*, os indivíduos são classificados de acordo com os seus valores de *Fitness*. O melhor indivíduo, ou seja, aquele com o menor

valor de erro, é escolhido para integrar uma nova população. Uma forma de tornar o indivíduo diretamente proporcional ao seu valor *Fitness* é produzir uma expressão de acordo com a expressão. (4.27). Desta forma, quanto melhor for o indivíduo, maior será o seu *Fitness*. O valor 100 na expressão acima foi escolhido de modo a que o valor do *Fitness* apresente sempre valores positivos para os estudos de caso e possa assim ser utilizado pelo método da roleta de classificação da população.

É importante salientar que, durante o procedimento de treinamento realizado pela população, todos os indivíduos estão sujeitos as mesmas condições de operações, isto é, o algoritmo de treinamento é inicializado da mesma forma, com os mesmos valores iniciais dos parâmetros adaptativos e com os mesmos critérios de parada. Isto é feito com o objetivo de avaliar unicamente os efeitos que o coeficiente de aprendizado e largura de *kernel* tem sobre a resposta de estimação, excluindo efeitos indesejados como, por exemplo, uma inicialização de parâmetros adaptativos que prejudiquem a resposta final.

- 2. Seleção:** A seleção é realizada através do método da roleta. Nele, cada indivíduo da população é representado na roleta proporcionalmente ao seu valor de aptidão (*fitness*). Assim, para indivíduos com alta aptidão é dada uma porção maior da roleta, enquanto aos indivíduos de aptidão mais baixa, é dada uma porção relativamente menor. Como o tamanho da população é mantido constante ao longo de todo processo evolutivo, na transição de uma geração para a outra, a roleta deve ser girada um número de vezes igual ao tamanho da população. Devido as características da roleta, espera-se que na geração seguinte, a nova população seja formada pelos indivíduos mais aptos da geração anterior e que os menos aptos sejam eliminados. Em outras palavras, espera-se que ao longo do processo de seleção, alcançar uma população de indivíduos que apresentem erros de estimação relativamente baixos.
- 3. Cruzamento:** A função de cruzamento, também chamada de recombinação ou crossover é inspirada nos fundamentos que norteiam a evolução natural. Ela simula, de certa forma, a reprodução sexuada, em que recebe-se dois cromossomos (Indivíduos em sua forma binária) e gera descendentes, resultado da combinação entre eles. (Linden, 2012) afirma que comumente é usado, para a taxa de cruzamento, um percentual alto, entre 60% e 95%. Valores menores foram usados em experimentos nos trabalhos de (Pinho; Montevechi; Marins, 2009) 10%, (Miranda; Pinho; Montevechi, 2011) 50% e (Carvalho, 2017) entre 11% e 90%. O funcionamento do operador de cruzamento para cada indivíduo na população pode ser visto na Figura 11.

Figura 11 – Cruzamento de Indivíduos - Algoritmo Genético.



Fonte: Autor (2021).

Na metodologia aplicada utilizou-se um único ponto de corte escolhido aleatoriamente. Na literatura (Jacinto, 2009) há a possibilidade de mais pontos de corte, caracterizando o cruzamento multiponto. A probabilidade de cruzamento entre dois indivíduos da população na estratégia adotada é de 70%, em que as informações da largura de *kernel* e coeficiente de aprendizado foram iniciadas no algoritmo genético na forma binária.

4. **Mutação:** A taxa de mutação se relaciona diretamente com a diversidade genética da população, essa variedade permite uma maior disseminação dos cromossomos no espaço de busca. No início da execução de um AG essa variedade é alta, considerando a criação da primeira geração de forma aleatória (Barcellos, 2000). No decorrer do processo há uma tendência de convergência genética, diminuição da diversidade, decorrente da seleção e cruzamento de soluções. Portanto, a mutação é um fator que permite a manutenção da diversidade populacional em um AG, aumentando seu poder de exploração. É indicado pela literatura o uso de uma taxa entre 0,5% e 5%, uma taxa de mutação de ultrapasse 5% torna a busca aleatória (Bach, 2003). Apesar disso, outros trabalhos usaram valores para a taxa de mutação acima de 5%, por exemplo, (Pinho; Montevechi; Marins, 2009) até 30%, (Miranda; Pinho; Montevechi, 2011) até 10%, (Carvalho, 2017) até 90%. A probabilidade de mutação adotada no método proposto é de 10%. Este valor, para este trabalho proposto, permite uma manutenção da diversidade de população sem que se torne uma busca aleatória de indivíduos.

4.4 Base de Dados - *National Aeronautics and Space Administration* - NASA

O modelo de identificação proposto foi desenvolvido utilizando informações para treinamento e validação, um conjunto de dados disponibilizados em um repositório *online* mantido pelo centro de pesquisa da NASA. Os dados utilizados para o objetivo geral são gerados na mesa de testes descrita em (Saha *et al.*, 2011). Neste conjunto de dados, existem diversas informações de operação acerca de baterias. As variáveis de interesse neste banco de dados giram em torno da corrente de descarga, tensão nos terminais, temperatura de operação e capacidade da bateria após diversos ciclos de descarga.

Tabela 1 – Especificações técnicas - Modelo 18650.

Parâmetros	LMO
Tensão nominal	3.7 V
Tensão máxima	4.2 V
Tensão de corte	2.6 V
Capacidade nominal	6000 mAh
Impedância interna	48 mΩ
Número de ciclos	300
Geometria	Cilíndrica
Diâmetro	18.35 mm
Comprimento	65.10 mm
Massa	45.5g

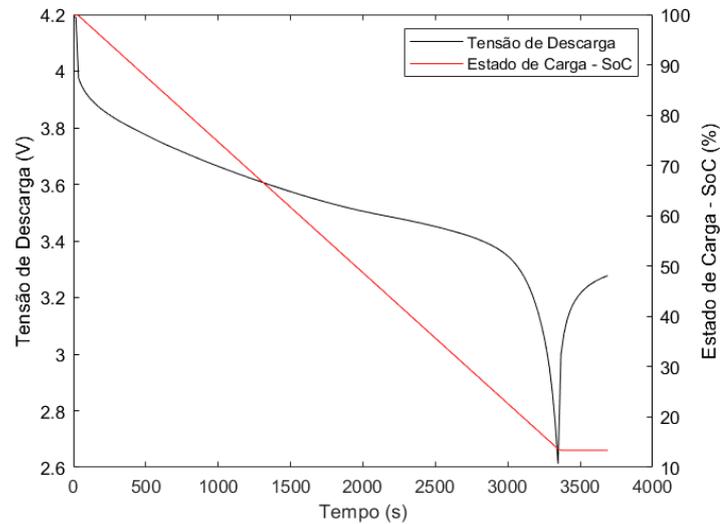
Fonte: NASA (2021)

A base de dados manipulada dispõe de uma ótima variedade de ensaios realizados nas baterias, a título de exemplo, ensaios com a bateria em estado de repouso, ensaios de carga e descarga contínua, testes de estresse dinâmico e entre outros tipos de ensaios. Para o desenvolvimento dos algoritmos neste trabalho, foram utilizados os ensaios de descarga por prover uma boa consistência dos dados, a quantidade de ensaios realizados ser bastante satisfatória para o processo de treinamento e em um estudo mais detalhado a possibilidade de verificar a degradação da bateria ao longo do tempo. Em geral, 93% do conjunto de amostras disponível foram utilizadas para treinamento do modelo de identificação, em que as demais foram subdivididas para a validação. Na Tabela 1 estão dispostas as especificações técnicas da bateria Lítio-íon 18650.

Na seleção de variáveis de entrada para o processo de identificação, considerou-se as principais grandezas que afetam o comportamento da bateria, como corrente, tensão e temperatura.

A princípio, a tensão na bateria foi escolhida como entrada do sistema de identificação por representar o resultado de todo processo intrínseco dinâmico da bateria. A

Figura 12 – Curvas de tensão de descarga e estado de carga em único ensaio.



Fonte: Adaptado de *National Aeronautics and Space Administration* (2021).

corrente foi selecionada por ter uma relação direta com a carga disponível na bateria. Por fim, a última variável que irá compor o modelo é a temperatura, decorrente do fato de que a mesma afeta o comportamento dinâmico interno da bateria, como por exemplo, a elevação da temperatura provoca uma maior movimentação dos elétrons, causando uma perda mais rápida da massa ativa, que é formada pelas placas.

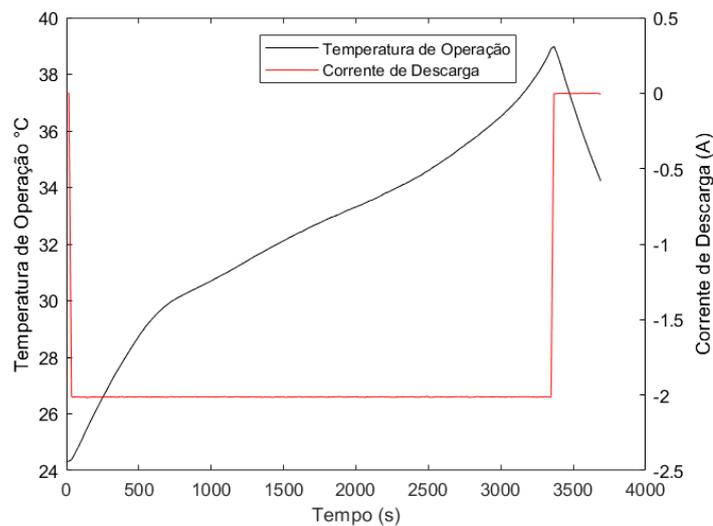
Na Figura 12 e 13 são exibidas curvas de operação das variáveis de interesse para um único ensaio da bateria, à título de exemplo. A tensão de operação permanece dentro de sua faixa especificada, entre o limite máximo (4.2V) e a tensão de corte (2.6V), para uma corrente de descarga contínua de 2A em um ambiente cuja temperatura sofreu uma variação positiva de aproximadamente 15°C. Uma possível limitação do método proposto se encontra ao redor do valor de descarga da corrente de operação. Embora o banco de dados seja extremamente completo no sentido de fornecer ensaios a temperaturas consideradas baixas (5° a 15°C), médias (25° a 35°C) e altas (45° a 55°C), a corrente de operação apresenta apenas 3 (três) valores de descarga, são eles: 1A, 2A e 4A. Essa restrição de valores na corrente de descarga pode ocasionar problemas de estimação do modelo de identificação quando, por exemplo, utiliza-se um possível valor de descarga de 1.5A. Estas limitações podem ser superadas através de uma ajuste fino da rede neural. A ideia básica com o ajuste fino é pegar noutro conjunto de dados muito grande, o que espera-se, algo semelhante ao seu domínio, treinar uma rede neural, e depois ajustar esta rede pré-formada com o seu conjunto de dados mais específicos.

O tempo de amostragem presente no conjunto de dados não é padronizado (10s à 120s), visto que o referido banco de dados foi concebido a partir de vários conjuntos de

	Nº de ensaios	Nº de amostras
Características - Treinamento RNA - SoC/SoH	2626	719785
Características - Validação RNA - SoC/SoH	168	50285

dados doados por várias universidades, agências ou empresas.

Figura 13 – Curvas de corrente de descarga e temperatura de operação



Fonte: Adaptado de *National Aeronautics and Space Administration* (2021).

4.5 Normalização e Processamento do Conjunto de Dados

Em especial, deve-se destacar o processo de normalização presente na entrada dos blocos de previsão de estados. Afim de minimizar os problemas provenientes do uso de unidades e faixas de valores distintos entre os parâmetros, é necessário que os dados estejam normalizados, padronizados em uma única escala.

O uso de rede neurais com função de ativação Sigmóide exige que os dados que alimentam a rede sejam normalizados para valores pertencentes ao intervalo de 0 a 1 e -1 a 1 no caso da função Tangente Hiperbólica. A função de ativação é responsável por passar os sinais das camadas anteriores da rede neural para a próxima camada. Isto é realizado com o objetivo de que valores de entrada numericamente maiores não influenciem erroneamente a saída da rede. Desta forma, o conjunto de dados de entrada da rede foi totalmente representado dentro do intervalo de 0 a 1, visto que a função de ativação Sigmóide foi escolhida, após testes de desempenho realizados na RNA. A normalização é

realizada através da Equação (4.28).

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (4.28)$$

em que: x representa uma matriz com as respectivas entradas do sistema de identificação: a corrente, tensão de descarga e temperatura de operação.

O processamento do banco de dados consistiu em selecionar e adequar às variáveis de interesse (Corrente de descarga, tensão de descarga e temperatura de operação) às características de entrada requeridas pela RNA proposta e sistema de estimação do SoC e SoH.

4.6 Definições de Outliers e Ruídos

Para a implementação do estado da carga e saúde, são necessários vários pesos e análises em relação o modelo de bateria utilizado. No entanto, o modelo continuará a ser sujeito a efeitos de ruído. Estes ruídos podem ser devidos a vários factores tais como: mau contacto; utilização de sensores de má qualidade; A / D com baixa resolução; ruído de rede que pode propagar-se à fonte de energia da bateria; durante o funcionamento. Na maioria dos artigos analisados anteriormente, assume-se que o ruído de medição é ruído branco gaussiano. Assim, em uma das análises do modelo concebido, o estado de carga e saúde são contaminados por amostras com distribuição normal, média zero, e potência igual a 0,01W ou -20dB. A fim de simular medições discrepantes, certos valores são inseridos aleatoriamente no conjunto de validação.

5 Resultados Finais e Considerações

Este capítulo tem como objetivo avaliar a utilização da correntropia no treinamento de redes RNA. Com este propósito, estas redes são utilizadas na identificação de sistemas dinâmicos não-lineares em presença de ruído e *outliers*. Para realizar a análise de desempenho do algoritmo de treinamento e dos modelos obtidos pelo procedimento de identificação, são utilizados dois estudos de caso.

Na avaliação realizada nesses dois estudos de caso, os dados experimentais utilizados pelo procedimento de identificação estão sujeitos à certos cenários de contaminação por ruídos e *outliers*. O algoritmo de retropropagação do erro foi então contaminado por ruído e diferentes percentuais de *outliers*. Após a determinação dos modelos, são analisados alguns aspectos importantes, tais como: velocidade de convergência, desempenho dos modelos encontrados e influência da largura de *kernel* no treinamento da rede. A partir destas informações, é feita uma comparação entre os treinamentos da RNA, realizados a partir da correntropia e do tradicional erro médio quadrático médio. A simulação foi realizada usando Matlab/Simulink num computador Inter Core i7-7500U 2.70GHz com 16GB RAM. O tempo para realizar cada treinamento somado com o tempo para o teste de validação durou uma média de 25 segundos.

Para quantificar o desempenho em relação ao erro de estimação $e_c(k) = y(k)^* - y(k)$ dos modelos foram utilizados três indicadores de desempenho, são definidos por:

Índice MSE - Mean Square Error (Erro Médio Quadrático): Este índice é frequentemente utilizado na verificação da acurácia de modelos numéricos, como descrito, por exemplo, em (Wilks, 2006). Este índice é mais sensível aos grandes erros, por elevar as diferenças individuais ao quadrado. O MSE é sempre positivo, onde $MSE = 0$ indica uma simulação perfeita. Matematicamente expresso por:

$$MSE = \frac{1}{N} \sum_{k=1}^N (y^*(k) - y(k))^2 \quad (5.1)$$

Índice MAPE - Mean Absolute Percentage Error (Erro Percentual Absoluto Médio): Este índice é uma medida de precisão, que mede o tamanho do erro em termos percentuais (Wilks, 2006). Matematicamente expresso por:

$$MAPE = \sum_{k=1}^N \left| \frac{y^*(k) - y(k)}{y^*(k)} \right| \quad (5.2)$$

Índice de Correntropia Percentual - MCP: Este índice, comparado com as métricas clássicas, apresenta uma relação inversamente proporcional, onde a diminuição

dos índices clássicos ocasiona um aumento da similaridade. Estes resultados são esperados, pois, diminuem-se as características temporais durante o regime transitório e o erro em regime permanente.

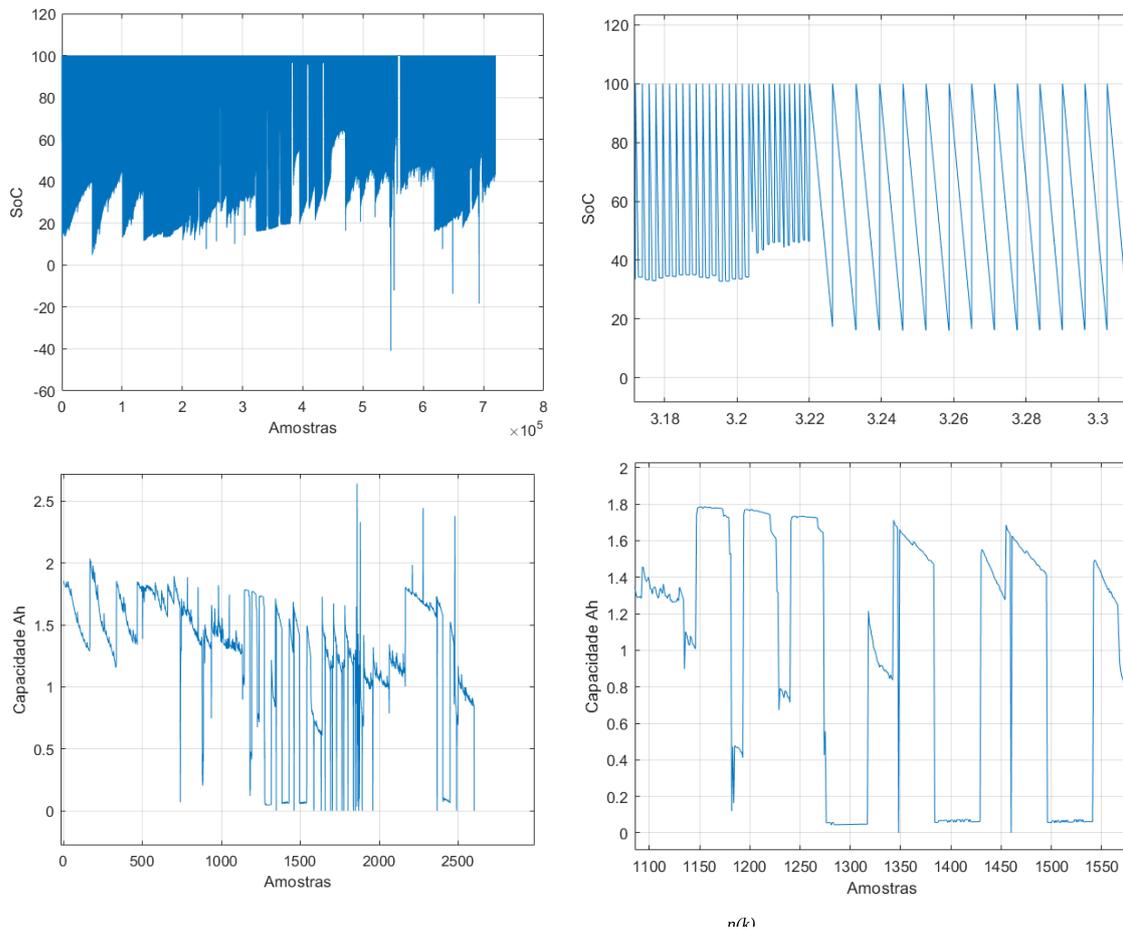
Em sentido oposto, quando a correntropia medida na forma percentual é próxima de zero, a similaridade entre duas variáveis é mínima, isto é, o sinal estimado diverge do sinal referência. Matematicamente pode ser expresso como:

$$MCP = \frac{\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\sum_{k=1}^N (y^*(k) - y(k))^2}{2\sigma^2}\right)}{\frac{1}{\sigma\sqrt{2\pi}}} \quad (5.3)$$

5.1 Banco de Dados NASA - Conjunto de Treinamento e Validação

O modelo de treinamento proposto foi desenvolvido utilizando informações de um conjunto de dados disponibilizados em um repositório *online* mantido pelo centro de pesquisa da NASA. Ao todo, o conjunto de dados utilizado para treinamento abrange 34 baterias, totalizando 2794 ensaios, distribuídos em ensaios de carga e descarga contínua, testes de estresse dinâmico e entre outros tipos de ensaios, conforme apresentado na Figura 13. Na Figura 13(a), é observado o estado de carga e capacidade do conjunto de treinamento das baterias de Lítio-íon. Com o objetivo de melhor visualização desta amostra de dados, a Figura 13(a) é ampliada e observada na Figura 13(b).

Figura 14 – Estado de Carga e Capacidade atual da Bateria de Lítio-íon - Modelo 18650.



Fonte: Autor (2021).

A bateria número 24, com ensaio número 63 foi escolhida como saída desejada para estimação do SoC. Do outro lado, a bateria 9 presente no conjunto de dados da NASA, teve seu ciclo de vida utilizado para estimação do estado de saúde, cuja respostas serão detalhadamente descritas nas próximas seções. Mostra-se na Figura 13 detalhe deste conjunto de dados descrito. Como se trata de um conjunto de grande extensão, uma parcela do estado de carga e do estado de saúde foi ampliada para melhor visualização da resposta da bateria para alguns ensaios de descarga.

5.2 Resultados - Estudo de Caso I - Estado de Carga

Nesta seção são apresentados os resultados obtidos na identificação do estado de carga, quando os dados de treinamento e validação utilizados no procedimento estão com e sem a presença de ruído e *outliers*. Logo, durante os treinamentos das redes RNA, foram utilizados os dados anteriormente apresentados.

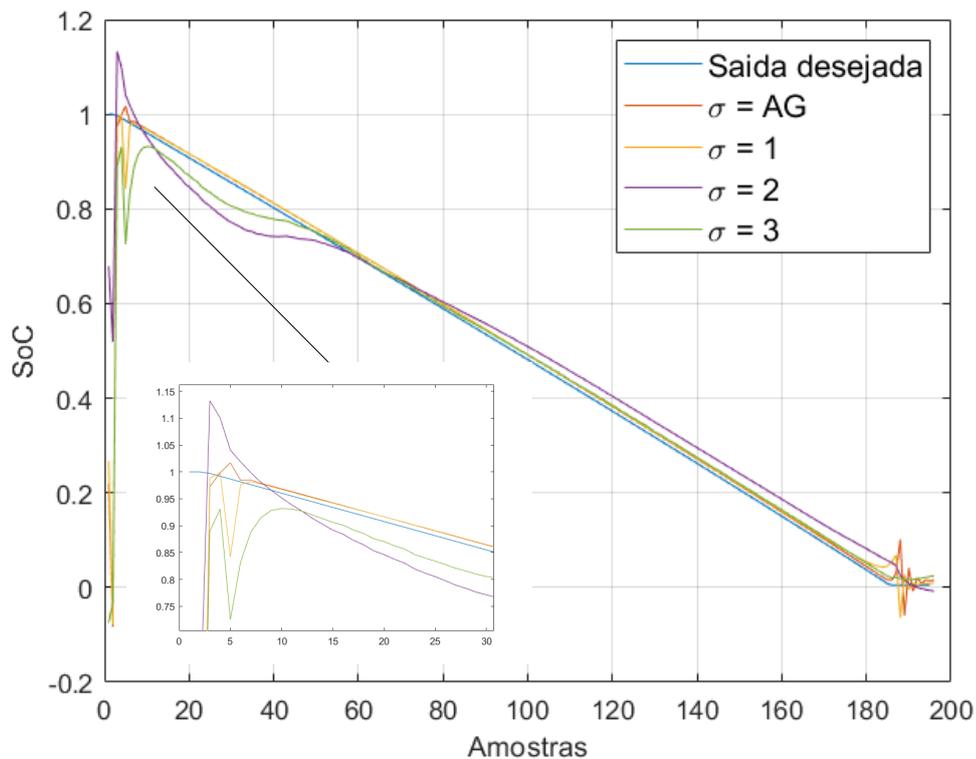
5.2.1 Resultados - Sem presença de ruído e *outliers*

Antes de iniciar a apresentação de resultados, alguns pontos precisam ser estabelecidos. Daqui em diante, será adotada a notação $RNA_{MCC,AG}$ para denominar o modelo obtido a partir da utilização da correntropia como função de custo do algoritmo de retropropagação de erro e algoritmo genético para otimização dos parâmetros η e σ ; De forma análoga, a designação $RNA_{MCC,ADAP}$ será dada ao modelo obtido a partir da utilização da correntropia como função de custo no algoritmo de treinamento adotando uma estratégia adaptativa para determinar a largura de *kernel*.

Para cada combinação de algoritmo, largura de *kernel* e coeficiente de aprendizado foram efetuadas 15 execuções do algoritmo de treinamento. Isto foi realizado com o objetivo de se obter o melhor modelo para cada uma das combinações e permitir calcular algumas métricas de avaliação entre os dados estimados pela RNA e os dados teóricos fornecidos pelo conjunto de dados disponíveis no repositório online da NASA. Foram analisados três diferentes valores de largura de *kernel*: 1, 2 e 3. Esses valores, representam, respectivamente, larguras de *kernel* baixa, média e elevada para este estudo de caso.

Na Figura 15, as respostas do modelo são comparadas com a saída desejada do sistema sem a presença de ruído e *outliers*. É importante ressaltar que durante o treinamento, o conjunto de dados disponibilizados continha dados contaminados e, mesmo sob essas condições, pretende-se que, a partir da utilização da função de custo da correntropia, seja possível obter modelos capazes de representar satisfatoriamente a dinâmica do estado de carga para baterias de Lítio-íon 18650.

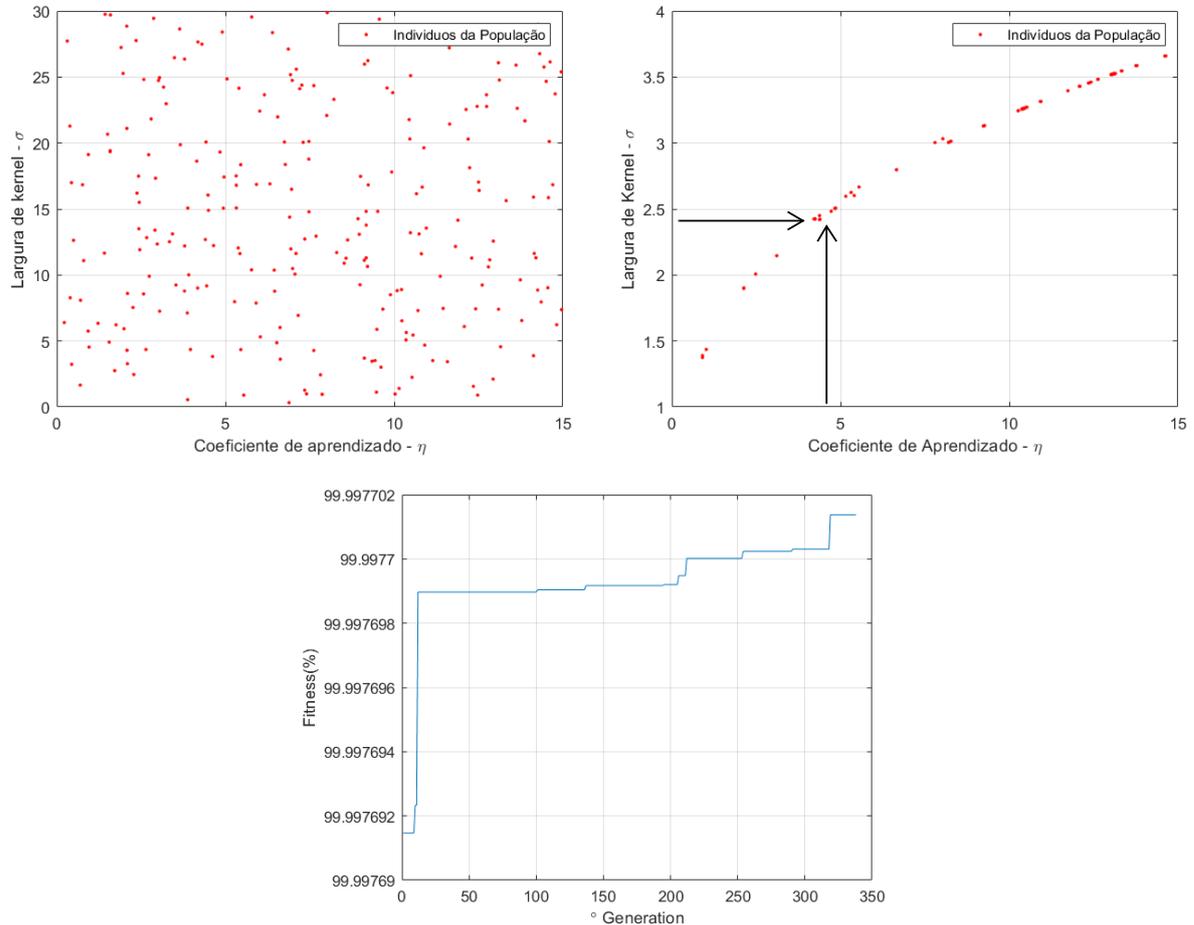
Figura 15 – Validação dos modelos RNA - Estado de Carga - Resultado de treinamento sem a presença de ruído e *outliers*.



Fonte: Autor (2021).

O melhor modelo $RNA_{MCC,AG}$, cuja resposta é apresentada na Figura 15, foi obtido a partir da otimização dos parâmetros de *kernel* e coeficiente de aprendizado pelo algoritmo genético. Através da repetição das operações no processo de seleção, cruzamento e mutação, ao longo das gerações no algoritmo genético, a solução para o problema de otimização apresentou um erro satisfatório. Cada indivíduo da população contém informações sobre uma única largura de *kernel* e coeficiente de aprendizado. Esse processo de seleção inspirado na seleção natural se encarrega de eliminar os indivíduos pior adaptados ao problema de otimização do SoC e permitir a sobrevivência daqueles que se adaptam melhor às condições oferecidas. Ilustra-se na Figura 15 a evolução da população na primeira geração até a última apresentada ao problema de otimização dos parâmetros.

Figura 16 – População de Indivíduos - Estado de Carga - Resultado da otimização por algoritmo genético sem a presença de ruído e *outliers*. (a) População Inicial de indivíduos, (b) população final de indivíduos e (c) evolução do *fitness* do melhor indivíduo de cada geração.



Fonte: Autor (2021).

Os algoritmos genéticos requerem uma forma de mapear a representação do indivíduo (genótipo) em um valor numérico associado ao seu nível de aptidão (*fitness*). Ilustra-se na Figura 16 o nível de aptidão da população inicial apresentada ao algoritmo genético e a geração número 330, respectivamente. Ainda nesta figura, observa-se que o algoritmo genético procura privilegiar indivíduos com melhores aptidões, podendo ser observado pela evolução do *fitness* do melhor indivíduo da população. Desta forma, o AG tenta dirigir a busca para regiões do espaço de busca onde é provável que os pontos ótimos estejam. A região com melhores indivíduos, neste caso, representa um espaço de busca limitado por:

$$1 \leq \sigma \leq 4.5 \tag{5.4}$$

$$1 \leq \eta \leq 15 \quad (5.5)$$

Através da análise da evolução do algoritmo genético, definiu-se como critério de parada a geração número 330. Este número foi definido usando a idéia de *estagnação*, ou seja, quando não se observou melhoria significativa da população depois de várias gerações consecutivas, isto é, quando a aptidão média ou do melhor indivíduo não melhorou mais ou quando as aptidões dos indivíduos de uma população tornaram-se muito parecidas. O melhor indivíduo da população após este processo de otimização contém informações de:

$$\sigma_{best} = 2.4203 \quad (5.6)$$

$$\eta_{best} = 4.3807 \quad (5.7)$$

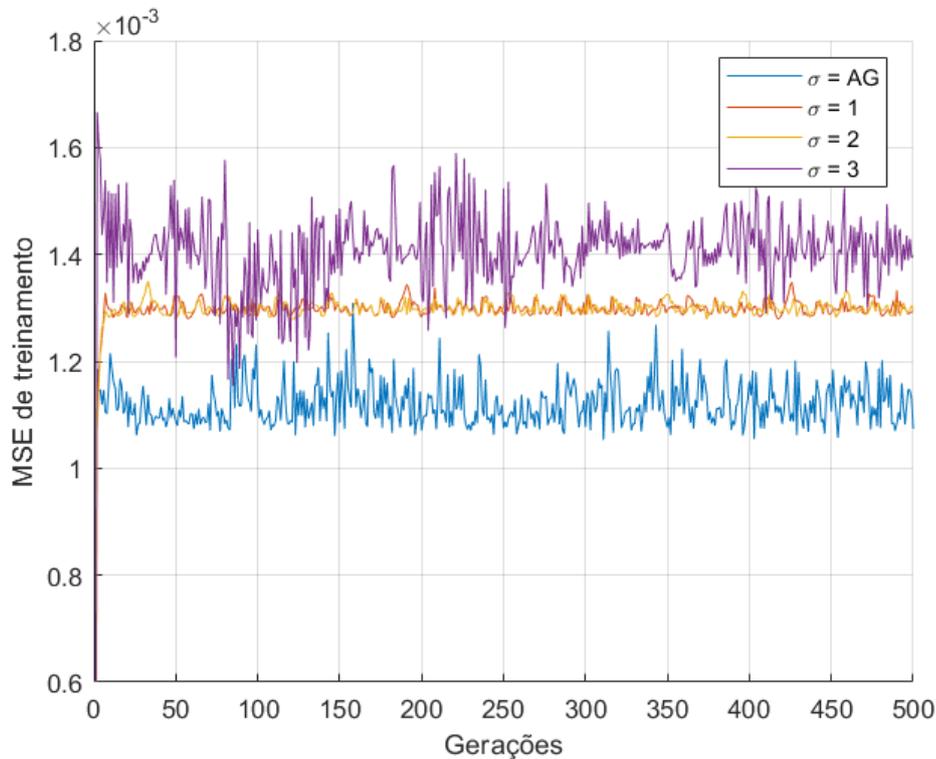
Os níveis de aptidões de todos os indivíduos na última população são bastante similares ao nível ótimo local encontrado pelo algoritmo, e portanto, considerado em um nível adequado.

No ajuste dos parâmetros dos modelos RNA, com exceção do $RNA_{MCC,AG}$, realizado pelo algoritmo de retropropagação do erro, verificou-se, empiricamente, que os melhores resultados foram encontrados ao se definir o coeficiente de aprendizado η com o valor igual a 1. Uma taxa de aprendizado baixa tende a aumentar o tempo de convergência do modelo, visto que o ajuste dos parâmetros da RNA (Pesos e *Bias*) é realizado de maneira mais suave. Em detrimento do aumento de épocas para treinamento, o erro do modelo para tais taxas de aprendizado tende a diminuir em comparação à coeficientes de aprendizado elevados, para as mesmas condições de operação.

Apesar de normalmente, as curvas de aprendizado praticamente se estabilizarem por volta da quinquagésima época de treinamento, alguns bons resultados foram encontrados com a execução de aproximadamente 330 épocas. Por essa razão, definiu-se este último valor como sendo a quantidade máxima de épocas executadas pelo algoritmo de treinamento.

A Figura 17 ilustra as curvas de aprendizado de todas as configurações de modelo avaliadas neste estudo de caso. Nela é apresentada a evolução do erro médio quadrático à medida que as épocas vão sendo executadas pelo algoritmo de treinamento. Os valores de MSE apresentados, foram calculados a partir dos dados de treinamento sem a presença de ruído e *outliers*.

Figura 17 – Curvas de aprendizado dos modelos RNA - Estado de Carga - Resultado de treinamento sem a presença de ruído e *outliers*.



Fonte: Autor (2021).

Demonstra-se na Figura 17 que para todos os valores de largura de *kernel*, as taxas de convergência foram relativamente satisfatórias. Para os modelos RNA_{MCC} obtidos, a Figura 17 indica que o melhor desempenho foi atingido quando a largura de *kernel* e coeficiente de aprendizado foi otimizada pelo algoritmo genético. No entanto, um valor alto da taxa de aprendizado pode significar aproximar do objetivo final rápido demais, criando oscilações entre o ponto de aprendizado ideal, isto pode ser visualmente observado pelas oscilações acentuadas no erro de treinamento para largura de *kernel* $\sigma = 3$ e $\sigma_{best} = 2.4203$. Em relação a largura de *kernel*, por outro lado, verificou-se através de simulação, que para valores $\sigma \leq 1$, a janela de observação do erro tornava-se bastante estreita, prejudicando o desempenho do algoritmo.

Realiza-se na Tabela 2 um comparativo de desempenho entre os modelos RNA encontrados quando os dados experimentais utilizados na identificação do estado de carga quando não estão sujeitos a presença de ruído e *outliers*. São apresentados para cada configuração de modelo, a média do MSE de validação, a variância do MSE de validação, o melhor MSE bem como a média do MAPE e MCP de validação.

Tabela 2 – Comparação de desempenho dos modelos RNA - Estado de Carga - Resultado de treinamento sem a presença de ruído ou *outliers*

Modelo e Treinamento	Largura de kernel	Média do MSE de Validação	Melhor MSE de Validação	Média do MAPE de Validação	Média do MCP de Validação (%)
RNA_{MCC}	$\sigma = 1$	$6,4 \cdot 10^{-3}$	$6,1 \cdot 10^{-3}$	$44,3 \cdot 10^{-3}$	89.86
RNA_{MCC}	$\sigma = 2$	$6,9 \cdot 10^{-3}$	$6,5 \cdot 10^{-3}$	$51,3 \cdot 10^{-3}$	85.40
RNA_{MCC}	$\sigma = 3$	$6,7 \cdot 10^{-3}$	$6,5 \cdot 10^{-3}$	$49,0 \cdot 10^{-3}$	72.45
$RNA_{MCC,AG}$	Otimizado AG	$5,0 \cdot 10^{-3}$	$4,8 \cdot 10^{-3}$	$24,6 \cdot 10^{-3}$	89.85

Fonte: Autor (2021).

De acordo com a Tabela 2, os modelos obtidos a partir da largura fixa de kernel igual a 2 não apresentaram bom satisfatório em comparação aos outros modelos de *kernel*. O modelo de melhor desempenho corresponde ao modelo $RNA_{MCC,AG}$ com otimização dos parâmetros η e σ . Isto pode ser notado tanto pela sua média de MSE de validação quanto pelo média do MCP de validação.

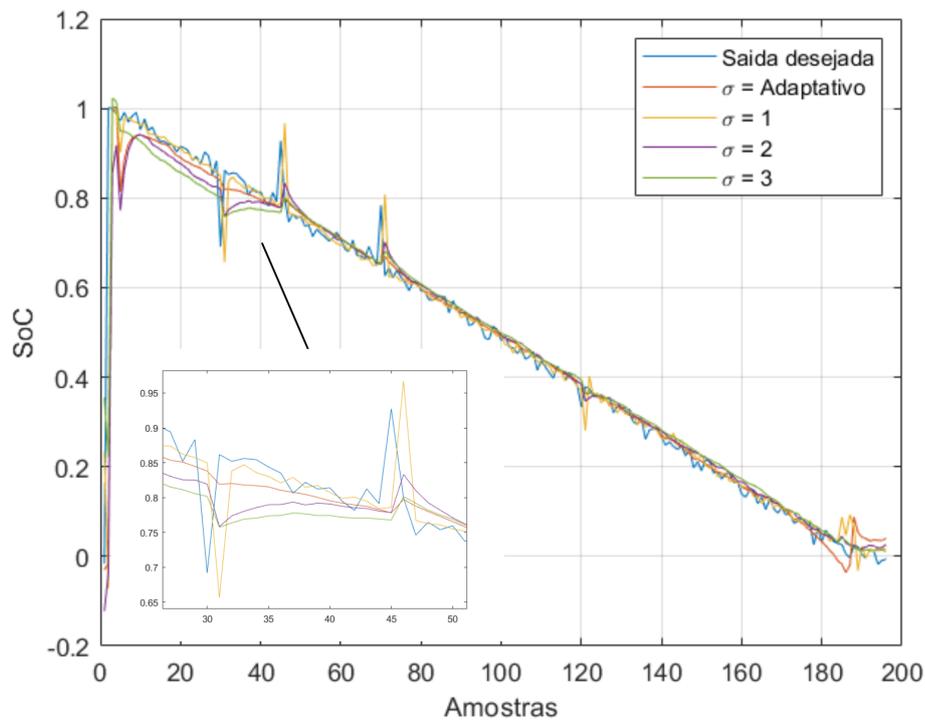
5.2.2 Resultados - Presença de ruído e *outliers*

Nesta seção são apresentados os resultados obtidos na identificação do estado de carga em baterias de Lítio-íon, quando os dados de treinamento e validação utilizados no procedimento estão sujeitos à presença de ruído gaussiano branco e *outliers*.

A metodologia utilizada para obtenção dos resultados aqui apresentada é idêntica a utilizada na seção anterior. Para cada combinação de algoritmo, função de custo e largura de *kernel* foram efetuadas 15 execuções do algoritmo de treinamento. Três diferentes valores de largura de *kernel* (1, 2 e 3) foram avaliados, além do ajuste adaptativo da largura de *kernel*. O valor atribuído ao coeficiente de aprendizado η do algoritmo de retropropagação do erro foi igual a 1 e o número de épocas de treinamento foi estabelecido em 500 épocas.

Apresenta-se na Figura 18 as respostas dos melhores modelos MCC obtidos a partir do conjunto de dados de treinamento contendo ruído gaussiano branco e *outliers*.

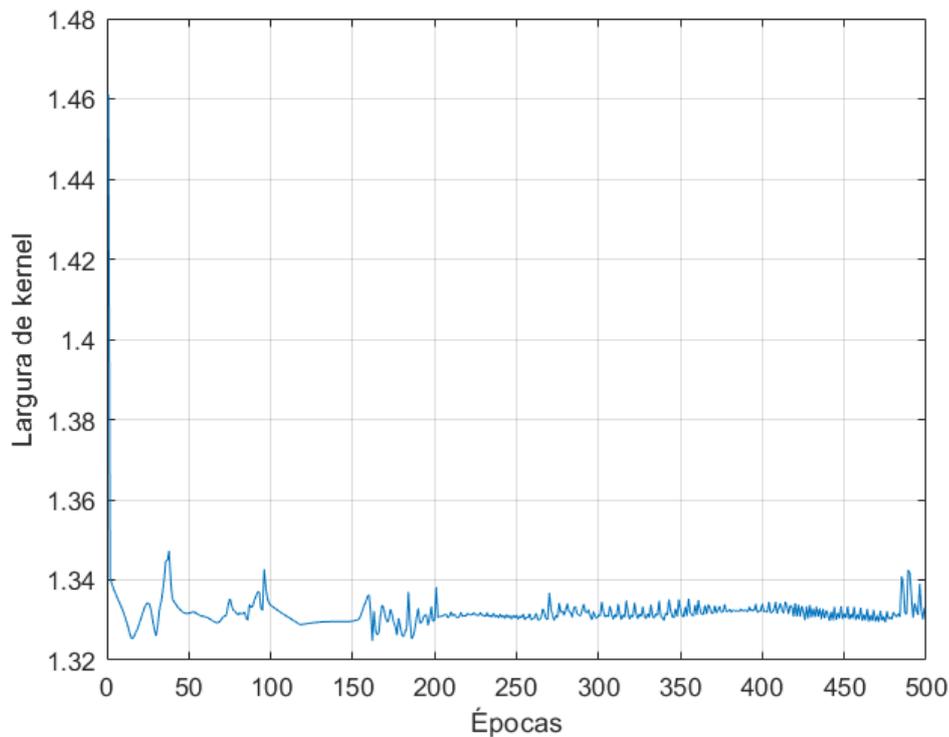
Figura 18 – Validação dos modelos RNA - Estado de Carga - Resultado de treinamento com a presença de ruído e *outliers*.



Fonte: Autor (2021).

A variação da largura de *kernel* adaptativo durante a execução do algoritmo de treinamento é ilustrada na Figura 19. A largura de *kernel* da correntropia é inicializada com o valor unitário, apresentando um acréscimo em seu valor durante as primeiras épocas do algoritmo de treinamento. Em seguida o valor deste parâmetro passa a oscilar, mas com uma tendência de redução, até estabilizar-se em um valor próximo a 1.33. A curva apresentada na Figura 19 corresponde ao ajuste realizado na largura de *kernel* em cada época de treinamento para se obter o melhor modelo RNA_{MCC} , cuja resposta foi apresentada na Figura 19.

Figura 19 – Variação da largura de *kernel* adaptativo - Estado de Carga - Resultado de treinamento com a presença de ruído e *outliers*.

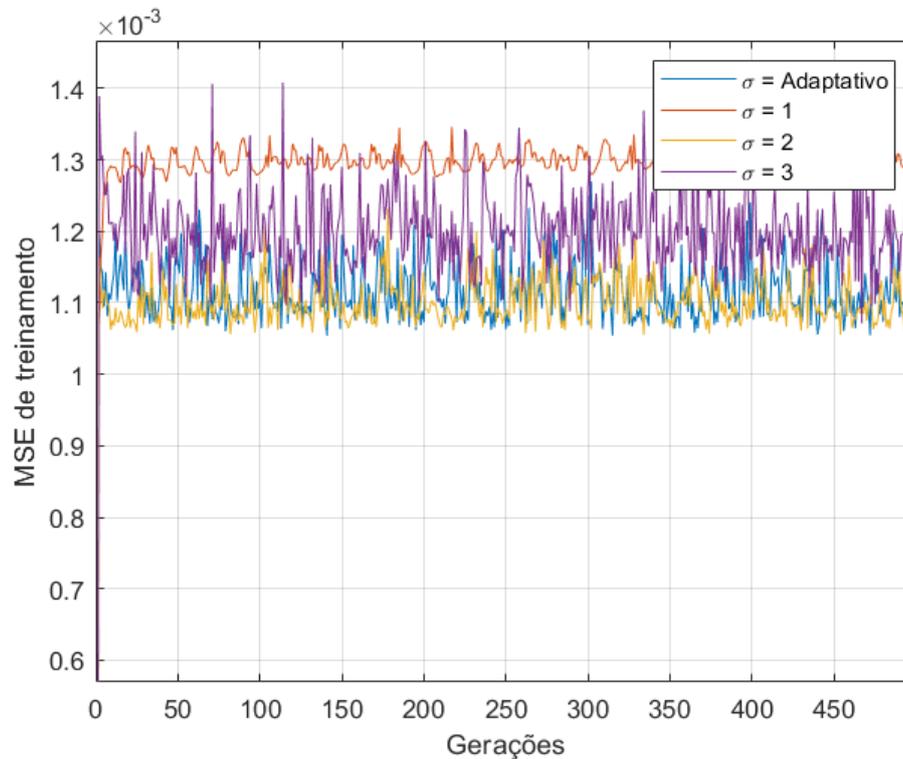


Fonte: Autor (2021).

Na Figura 19 é demonstrada a evolução do erro médio quadrático à medida que as épocas vão sendo executadas pelo algoritmo de treinamento. Os valores de MSE foram calculados à partir dos dados de treinamento em presença de ruído e *outliers*. Devido à escala utilizada nesta figura, algumas curvas estão sobrepostas.

Em relação à curva de treinamento do modelo encontrado a partir da utilização da largura de *kernel* adaptativa, pode-se observar que, a curva de treinamento do modelo $RNA_{MCC,ADAP}$ com largura de *kernel* adaptativa inicia-se com um rápido decaimento e, logo em seguida, começa a oscilar até que, por volta da trigésima época, seu comportamento torna-se praticamente estável.

Figura 20 – Curvas de aprendizado dos modelos RNA - Estado de Carga - Resultado de treinamento com ruído e *outliers*.



Fonte: Autor (2021).

Apresenta-se na Tabela 3 um comparativo de desempenho entre os modelos RNA encontrados quando os dados experimentais aplicados na identificação do estado de carga estão sujeitos a presença de ruído e *outliers*. Novamente, são apresentados para cada configuração do modelo, a média do MSE de validação, a média do MCP de validação e o melhor MSE encontrados em 15 execuções do algoritmo de retropropagação do erro.

Tabela 3 – Comparação de desempenho dos modelos RNA - Estado de Carga - Resultado de treinamento com a presença de ruído e *outliers*.

Modelo e Treinamento	Largura de kernel	Média do MSE de Validação	Melhor MSE de Validação	Média do MAPE de Validação	Média do MCP de Validação (%)
RNA_{MCC}	$\sigma = 1$	$7,4 \cdot 10^{-3}$	$7,1 \cdot 10^{-3}$	$51,5 \cdot 10^{-3}$	$78,54 \cdot 10^{-3}$
RNA_{MCC}	$\sigma = 2$	$6,9 \cdot 10^{-3}$	$6,7 \cdot 10^{-3}$	$49,4 \cdot 10^{-3}$	$84,35 \cdot 10^{-3}$
RNA_{MCC}	$\sigma = 3$	$5,0 \cdot 10^{-3}$	$4,8 \cdot 10^{-3}$	$25,1 \cdot 10^{-3}$	$94,72 \cdot 10^{-3}$
$RNA_{MCC,ADAP}$	Estratégia adaptativa	$2,5 \cdot 10^{-3}$	$2,4 \cdot 10^{-3}$	$6,25 \cdot 10^{-3}$	-

De acordo com a Tabela 3, os modelos obtidos a partir da largura fixa de *kernel* igual a 1 não apresentaram bom desempenho. Por outro lado, de acordo com a média de MSE de validação, o modelo $RNA_{MCC,ADAP}$ com largura de *kernel* adaptativa apresentou

o melhor resultado dentre os modelos com função de custo em correntropia e apresentou maior robustez na presença de *outliers*, ignorando estas medidas discrepantes de seu processo de identificação.

As quantificações dos momentos estatísticos do histograma da Figura 21 são ilustradas na Tabela 4, onde a média está levemente próxima de zero, bem como o desvio padrão. Neste contexto, o parâmetro curtose indica como as caudas de uma distribuição diferem da distribuição normal. Uma distribuição com um valor de curtose positiva indica que a distribuição tem caudas mais "acentuadas" do que a distribuição normal, o que é o caso para todos os modelos descrito acima, o que pode ser descrito como curva de distribuição Leptocúrtica. O parâmetro assimetria indica a assimetria da distribuição. Conforme os dados vão se tornando simétricos, seu valor tende a zero. Para todas as variações de largura de *kernel*, a assimetria é próxima de zero.

Tabela 4 – Medidas Estatísticas do erro de estimação - Estado de Carga - Resultado de treinamento com a presença de ruído e *outliers*.

	Erro de Estimação $\sigma = \text{Adaptativo}$	Erro de Rastreamento $\sigma = 1$	Erro de Rastreamento $\sigma = 2$	Erro de Rastreamento $\sigma = 3$
Média	0.4805	0.4842	0.4793	0.4844
Desvio Padrão	0.2972	0.3000	0.2902	0.2845
Assimetria	-0.0710	0.0113	-0.1272	-0.0863
Curtose	1.7386	1.7667	1.7385	1.7746

5.2.3 Resultados - Presença de ruído e outliers - Comparação com MSE clássica

Nesta seção são apresentados os resultados obtidos na identificação do estado de carga em baterias Lítio-íon, quando os dados de treinamento e validação utilizados no procedimento estão sujeitos à presença de ruído gaussiano branco e *outliers*.

A metodologia utilizada para obtenção dos resultados aqui apresentada é idêntica a utilizada na seção anterior. Para cada combinação de algoritmo, função de custo e largura de *kernel* foram efetuadas 15 execuções do algoritmo de treinamento. O valor atribuído ao coeficiente de aprendizado η do algoritmo de retropropagação do erro foi igual a 1 e o número de épocas de treinamento foi estabelecido em 500 épocas. O objetivo desta seção é comparar os resultados do melhor modelo com função de custo em correntropia e o melhor modelo com função de custo em erro quadrático médio.

Apresenta-se na Figura 22 as respostas do melhor modelo MCC obtido com estratégia adaptativa e o modelo MSE. Comparando-se os modelos MCC e MSE, observa-se que o primeiro apresentou melhor desempenho. Apesar de existir uma porcentagem de *outliers* nos conjuntos de dados de treinamento e validação isto não propocionou um efeito

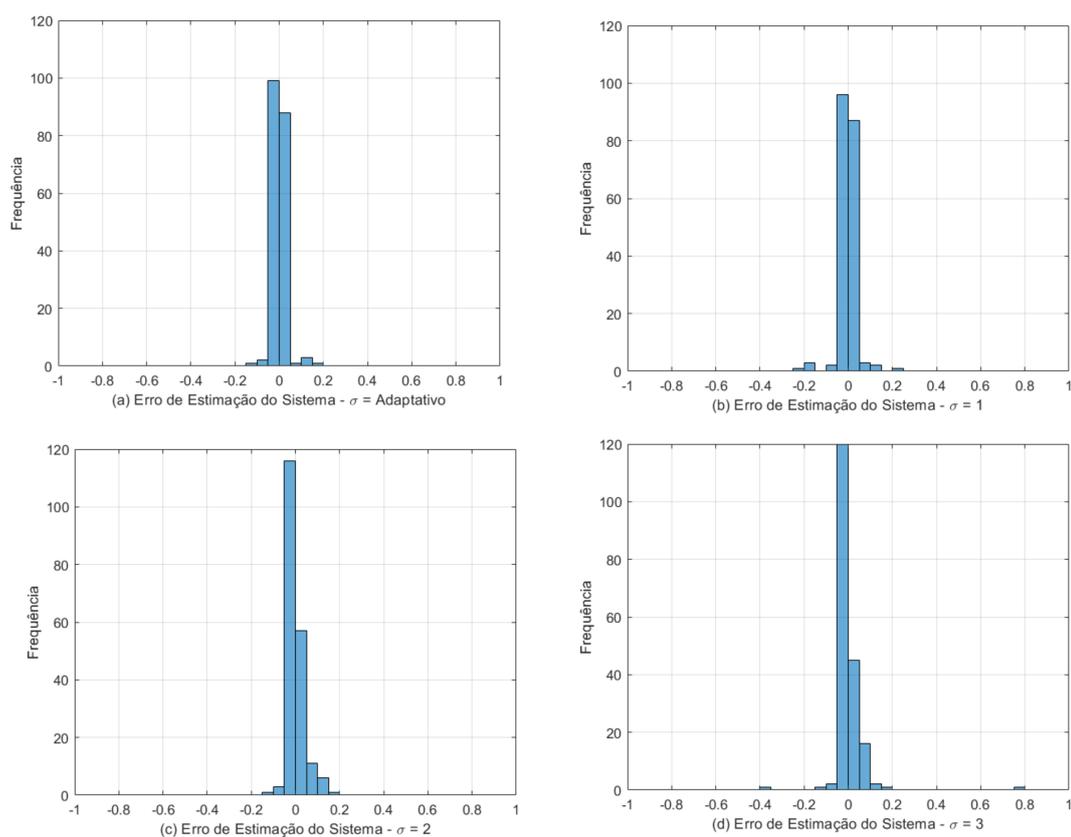
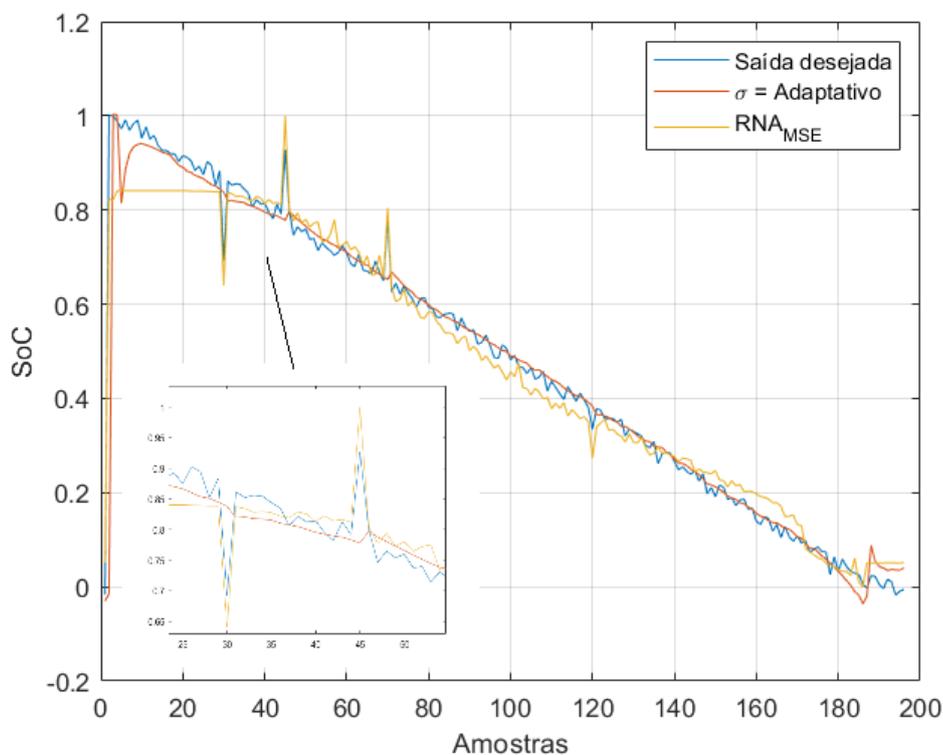


Figura 21 – Medidas Estatísticas do erro de estimação - Estado de Carga - Resultado de treinamento com a presença de ruído e *outliers*.

prejudicial significativo para o modelo MCC. Este modelo foi encontrado pelo algoritmo de retropropagação do erro, utilizando largura de *kernel* adaptativa e com função de custo definida pela maximização da correntropia.

Figura 22 – Validação dos modelos RNA - Estado de Carga - Comparação entre funções de custo MSE e MCC.



Fonte: Autor (2021).

De acordo com a Figura 22, a correntropia é mais robusta que o MSE em relação aos efeitos indesejáveis dos grandes valores de erro decorrentes de *outliers*, isto ocorre devido ao fato de que a função gaussiana da correntropia se aproxima de zero para grandes valores de erro. Nesse contexto, a efetividade da correntropia é controlada pelo valor da largura de seu *kernel* gaussiano. Este parâmetro especifica a menor magnitude de erro que será "desconsiderada" pelo estimador durante o ajuste de seus parâmetros. De acordo com (Bessa; Miranda; Gama, 2009) e (Munoz; Chen, 2012), valores elevados de largura de *kernel* tendem a sobre-suavizar a curva de estimação, enquanto valores muito pequenos podem gerar uma região de corte tão restrita que medições corretas do conjunto podem ser tratadas como *outliers* durante o processo de identificação.

Apresenta-se na Tabela 5 um comparativo de desempenho entre os modelos RNA encontrados quando os dados experimentais presentes na identificação do indicador de carga estão sujeitos a presença de ruído e *outliers*. Novamente, são apresentados para método de estimação, a média do MSE de validação, a variância do MSE de validação e o melhor MSE encontrado em 15 execuções do algoritmo de retropropagação do erro.

Tabela 5 – Comparação de desempenho com função de custo MCC e MSE - Estado de Carga - Resultado de treinamento com a presença de ruído e *outliers*.

Modelo e Treinamento	Média do MSE de Validação	Melhor MSE de Validação	Média do MAPE de Validação	Média do RMSE de Validação
RNA_{MCC}	$2,5 \cdot 10^{-3}$	$2,4 \cdot 10^{-3}$	$6,25 \cdot 10^{-3}$	$20,4 \cdot 10^{-3}$
RNA_{MSE}	$4,2 \cdot 10^{-3}$	$4,1 \cdot 10^{-3}$	$20,2 \cdot 10^{-3}$	$69,5 \cdot 10^{-3}$

5.3 Resultados - Estudo de Caso II - Estado de Saúde

5.3.1 Resultados - Sem presença de ruído e *outliers*

Nesta seção são apresentados os resultados obtidos na identificação do modelo de estado de saúde, quando os dados de treinamento e validação utilizados neste procedimento não estão sujeitos à presença de ruído e *outliers*. Portanto, durante a etapa de ajuste dos parâmetros dos modelos RNA_{MCC} , foram utilizados os conjuntos de dados anteriormente apresentados.

Assim como realizado no primeiro estudo de caso, para cada combinação de algoritmo e largura de *kernel* foram efetuadas 15 execuções do algoritmo de treinamento. Isto justifica-se pela presença de alguns aspectos estocásticos na execução do algoritmo de retropropagação do erro para o treinamento das redes RNA, tais como: inicialização aleatória dos parâmetros da RNA (Pesos e *Bias*) e embaralhamento dos dados experimentais em cada época de treinamento. Portanto, nestas condições, cada execução do algoritmo de treinamento do erro resulta em um modelo RNA diferente.

Para a execução do algoritmo de treinamento dos modelos RNA, foi atribuído ao coeficiente de aprendizado η o valor igual a 1 e o número de épocas de treinamento foi definido como sendo igual a 330. Estes valores foram obtidos a partir de um procedimento empírico semelhante ao realizado no primeiro estudo de caso. Após as execuções do algoritmo de treinamento, com o intuito de comparar o desempenho de cada combinação de algoritmo e largura de *kernel*, foram analisados o desempenho do melhor modelo obtido de cada uma destas combinações e alguns valores estatísticos, tais como a média do desempenho dos modelos RNA encontrados.

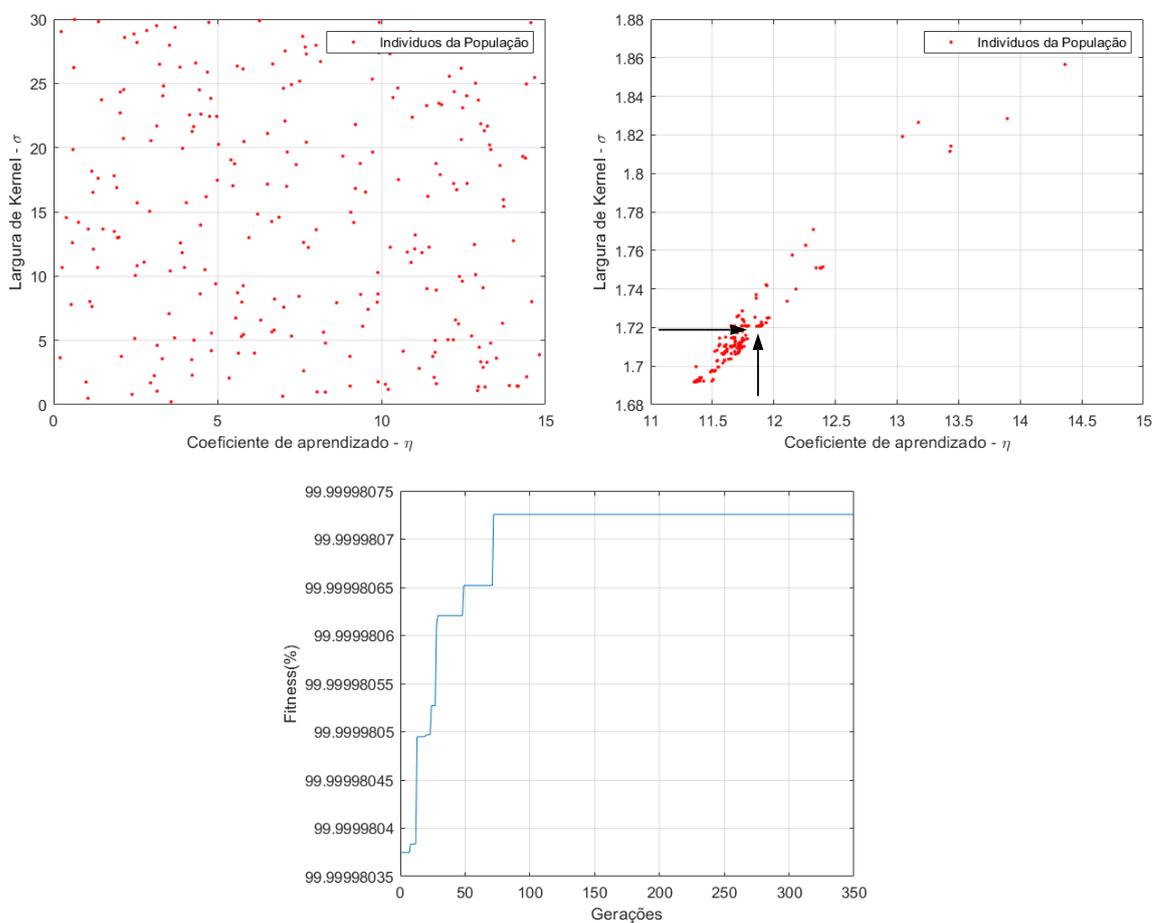
O algoritmo genético, em sua essência, direciona a procura de indivíduos para um espaço de busca onde é provável que os pontos ótimos estejam, isto é, neste contexto, regiões que contenham informações de largura de *kernel* e coeficiente de aprendizado que resultam em uma maximização da função de custo em correntopia, e conseqüentemente redução do erro de estimação. Na Figura 23, observa-se a redução do espaço de busca do algoritmo genético e a centralização de indivíduos próximos a uma região ótima. A região

com melhores indivíduos, neste caso, representa um espaço de busca limitado por:

$$1.68 \leq \sigma \leq 1.78 \quad (5.8)$$

$$11.35 \leq \eta \leq 12.5 \quad (5.9)$$

Figura 23 – População de Indivíduos - Estado de Saúde - Resultado da otimização por algoritmo genético sem a presença de ruído e *outliers*. (a) População inicial de indivíduos, (b) população final de indivíduos e (c) evolução do *fitness* do melhor indivíduo de cada geração.



Fonte: Autor (2021).

Conforme observado na Figura 22(c) e de maneira análoga a estimação do estado de carga, definiu-se como critério de parada a geração número 330, visto que, não foi observado melhora significativa da população depois de várias gerações consecutivas, assim, uma ponderação entre tempo de simulação e resposta proveniente do algoritmo genético foi definida. Desta forma, através da análise da evolução do algoritmo genético, o melhor

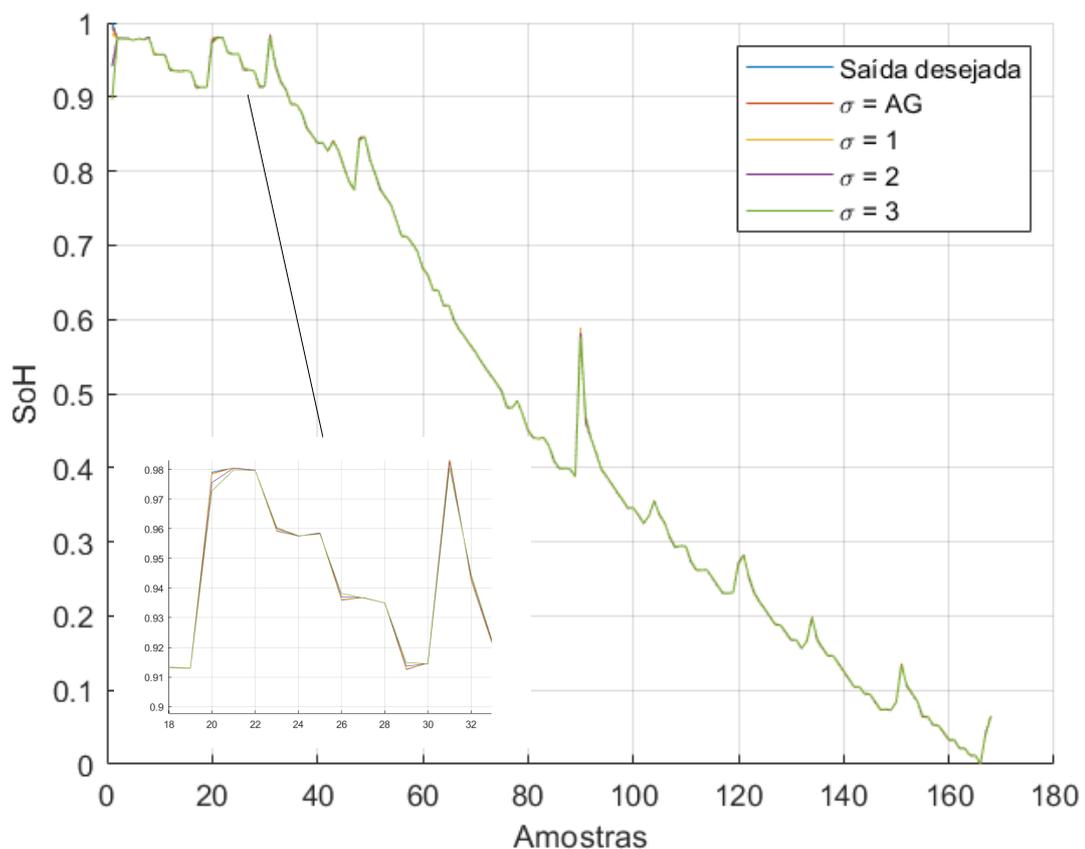
indivíduo da população após este processo de otimização contém informações de largura de *kernel* e coeficiente de aprendizado dado por:

$$\sigma_{best} = 1.7207 \quad (5.10)$$

$$\eta_{best} = 11.8815 \quad (5.11)$$

Apresenta-se na Figura 24 as respostas dos melhores modelos RNA obtidos a partir do conjunto de dados de treinamento sem a presença de ruído e *outliers*. Analisando-se essa figura, pode-se notar que o modelo $RNA_{MCC,AG}$, obtido por meio da otimização por algoritmo genético, apresentou o melhor desempenho. Apesar da resposta do modelo para largura de *kernel* ($\sigma = 1, \sigma = 2$ e $\sigma = 3$) possuírem em geral, resposta bastante próxima da sinal de validação, o modelo proveniente da otimização dos parâmetros possui um comportamento mais suave.

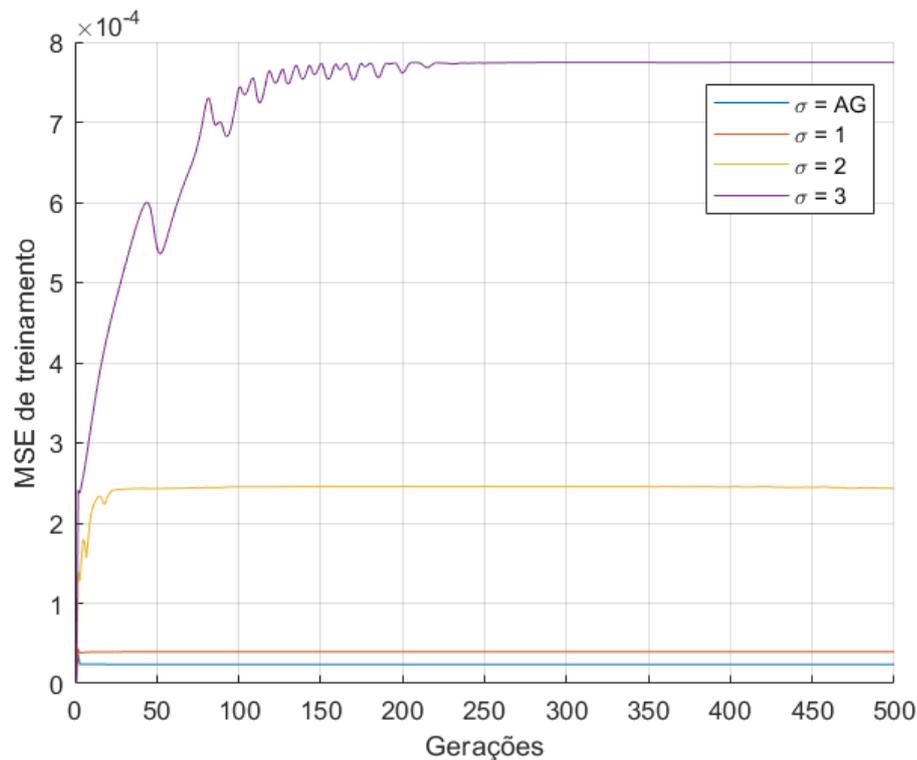
Figura 24 – Validação dos modelos RNA - Estado de Saúde - Resultado de treinamento sem a presença de ruído e *outliers*.



Fonte: Autor (2021).

As curvas de aprendizado de todas as configurações de modelo avaliadas na identificação do estado de saúde são ilustradas na Figura 25. Nesta figura, é demonstrada a evolução do erro médio quadrático à medida que as épocas vão sendo executadas pelo algoritmo de treinamento.

Figura 25 – Curvas de aprendizado dos modelos RNA - Estado de Saúde - Resultado de treinamento sem a presença de ruído e *outliers*.



Fonte: Autor (2021).

Prosseguindo com a análise da Figura 25, pôde-se notar que a velocidade de convergência, quando a largura de *kernel* foi fixada em 3, é maior do que as demais. Todos os modelos de RNA, com exceção do modelo otimizado ($\sigma = AG$), apresentam coeficiente de aprendizado relativamente baixo ($\eta = 1$), o que proporciona uma maior lentidão para alcance de um valor ideal em regime permanente, em detrimento da melhora da resposta do sistema.

Desta forma, o modelo $RNA_{MCC,AG}$ obtido por meio da otimização dos parâmetros η e σ foi a configuração que apresentou melhor desempenho entre todas as avaliadas neste estudo de caso, conforme visto na Figura 25.

Apresenta-se na Tabela 6 um comparativo de desempenho entre os modelos RNA encontrados quando os dados experimentais não estão contaminados por ruídos e *outliers*. São apresentados para cada configuração, a média do MSE de validação, a média do MCP

de validação e o melhor MSE encontrados em 15 execuções do algoritmo de retropropagação de erro.

Tabela 6 – Comparação de desempenho dos modelos RNA - Estado de Saúde - Resultado de treinamento sem a presença de ruído ou *outliers*

Modelo e Treinamento	Largura de kernel	Média do MSE de Validação	Melhor MSE de Validação	Média do MAPE de Validação	Média do MCP de Validação (%)
RNA_{MCC}	$\sigma = 1$	$1.1834 \cdot 10^{-6}$	$1.1832 \cdot 10^{-6}$	3.9070×10^{-4}	99.99
RNA_{MCC}	$\sigma = 2$	$6.6139 \cdot 10^{-5}$	$6.5733 \cdot 10^{-5}$	$1.4333 \cdot 10^{-3}$	99.96
RNA_{MCC}	$\sigma = 3$	$5.6686 \cdot 10^{-4}$	$5.6684 \cdot 10^{-4}$	$3.1000 \cdot 10^{-3}$	99.94
$RNA_{MCC,AG}$	Otimizado AG	$2.0492 \cdot 10^{-7}$	$2.0486 \cdot 10^{-7}$	$1,7460 \cdot 10^{-4}$	99.99

Fonte: Autor (2020).

De acordo com a Tabela 6, todos os modelos apresentaram um desempenho satisfatório para este estudo de caso, apresentando uma média de erro médio quadrático na ordem $\ll 10^{-3}$. O modelo de melhor desempenho corresponde ao modelo $RNA_{MCC,AG}$ com largura de *kernel* otimizada, como pode ser confirmado pela sua média de MSE de validação e pelo MCP de validação. Neste estudo de caso, a diferença de desempenho entre os melhores modelos RNA_{MCC} é menor do que no estudo de caso anterior.

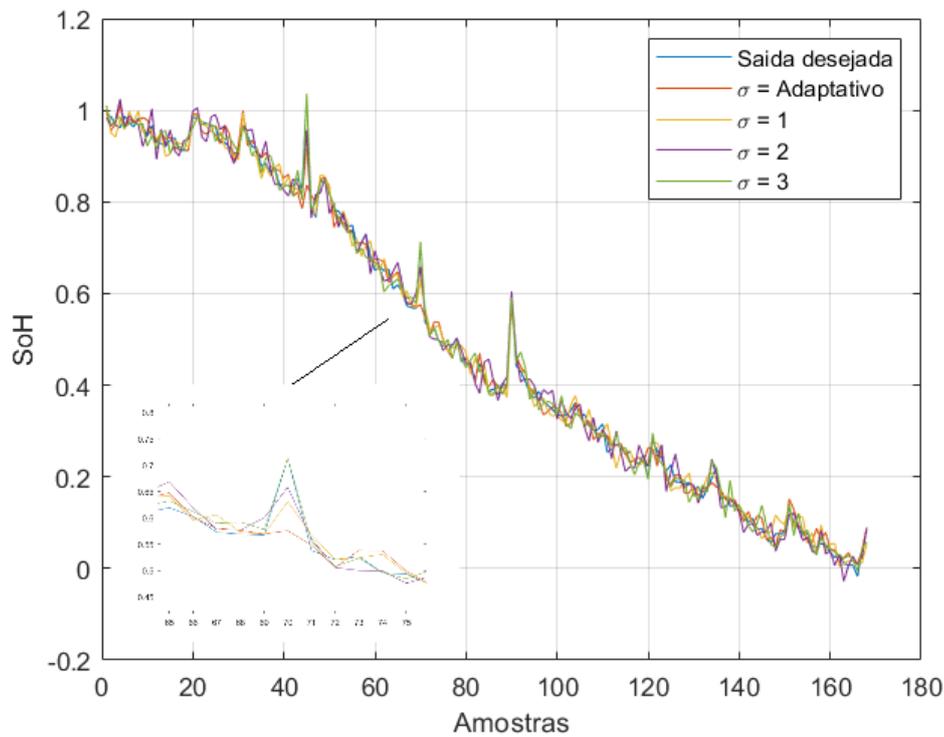
5.3.2 Resultados - Presença de ruído e *outliers*

Nesta seção são apresentados os resultados obtidos na identificação do estado de saúde em baterias de Lítio-íon, quando os dados de treinamento e validação utilizados no procedimento estão sujeitos na presença de ruído gaussiano branco e *outliers*.

A metodologia utilizadas para obtenção dos resultados aqui apresentada é idêntica a utilizada para estimação do estado de carga na seção anterior. Para cada combinação de algoritmo, função de custo e largura de *kernel* foram efetuadas 15 execuções do algoritmo de treinamento. Três diferentes valores de largura de *kernel* (1, 2 e 3) foram avaliados, além do ajuste adaptativo da largura de *kernel*. O valor atribuído ao coeficiente de aprendizado η do algoritmo de retropropagação do erro foi igual a 1 e o número de épocas de treinamento foi estabelecido em 500 épocas.

Apresenta-se na Figura 26 as respostas dos melhores modelos MCC obtidos a partir do conjunto de dados de treinamento contendo ruído gaussiano branco e *outliers*.

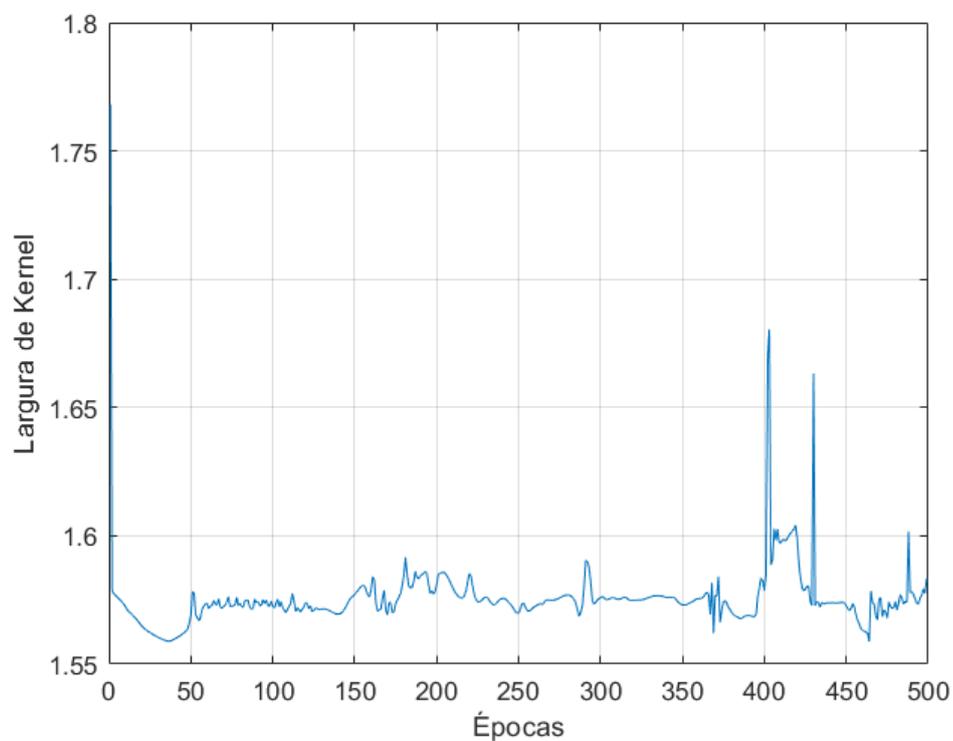
Figura 26 – Validação dos modelos RNA - Estado de Saúde - Resultado de treinamento com presença de ruído e *outliers*.



Fonte: Autor (2020).

A variação da largura de *kernel* adaptativo durante a execução do algoritmo de treinamento é ilustrada na Figura 27. A largura de correntropia é inicializada com o valor unitário, apresentando um acréscimo em seu valor durante as primeiras épocas do algoritmo de treinamento. Em seguida o valor deste parâmetro passa a oscilar, até estabilizar-se em um valor próximo a 1,57. A curva apresentada na Figura 27 corresponde ao ajuste realizado na largura de *kernel* em cada época de treinamento para se obter o melhor modelo RNA_{MCC} , cuja resposta foi apresentada na Figura 27.

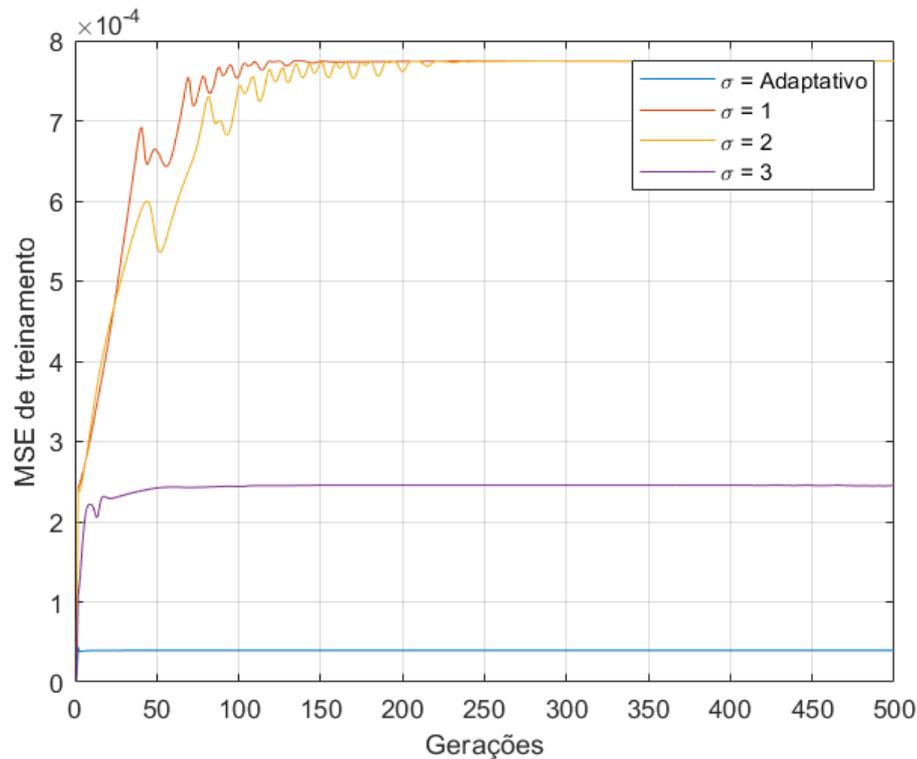
Figura 27 – Variação da largura de *kernel* adaptativo - Estado de Saúde - Resultado de treinamento com presença de ruído e *outliers*.



Fonte: Autor (2020).

As curvas de aprendizado de todas as configurações de modelo avaliadas neste estudo de caso são ilustradas na Figura 28.

Figura 28 – Curvas de aprendizado dos modelos RNA - Estado de Saúde - Resultado de treinamento com presença de ruído e *outliers*.



Fonte: Autor (2020).

Na Figura 28 é demonstrada a evolução do erro médio quadrático à medida que as épocas vão sendo executadas pelo algoritmo de treinamento. Os valores de MSE foram calculados a partir dos dados de treinamento em presença de ruído e *outliers*. Ainda nesta Figura, nota-se que para valores fixos de largura de *kernel* ($\sigma = 1$ e $\sigma = 2$, as taxas de convergência para a obtenção de modelos RNA_{MCC} são bastante próximas. Em relação à curva de treinamento do modelo a partir da utilização da largura de *kernel* adaptativa, pode-se observar que, apesar da inicialização aleatória dos parâmetros da rede, o valor inicial do erro médio quadrático é menor do que para as outras configurações do modelo. A curva de treinamento do modelo RNA_{MCC} com largura de *kernel* adaptativa inicia-se com um rápida oscilação e, logo em seguida, seu comportamento torne-se praticamente estável.

A Tabela 7 apresenta um comparativo de desempenho entre os modelos RNA encontrados quando os dados experimentais aplicados na identificação do estado de saúde estão sujeitos a presença de ruído e *outliers*. Novamente, são apresentados para cada configuração de modelo, a média do MSE de validação, a média do MAPE de validação, a média do RMSE de validação e o melhor MSE encontrado em 15 execuções do algoritmo de retropropagação do erro.

Tabela 7 – Comparação de desempenho dos modelos RNA - Estado de Saúde - Resultado de treinamento com a presença de ruído ou *outliers*

Modelo e Treinamento	Largura de kernel	Média do MSE de Validação	Melhor MSE de Validação	Média do MAPE de Validação	Média do MCP de Validação (%)
RNA_{MCC}	$\sigma = 1$	$6.9817 \cdot 10^{-4}$	$6.8221 \cdot 10^{-4}$	$3.4875 \cdot 10^{-3}$	98.67
RNA_{MCC}	$\sigma = 2$	$6.5813 \cdot 10^{-4}$	$6.5021 \cdot 10^{-4}$	$3.1713 \cdot 10^{-3}$	98.22
RNA_{MCC}	$\sigma = 3$	$4.5012 \cdot 10^{-4}$	$4.4813 \cdot 10^{-4}$	$2.1713 \cdot 10^{-3}$	97.84
$RNA_{MCC,ADAP}$	Estratégia Adaptativa	$1.1832 \cdot 10^{-6}$	$1.1755 \cdot 10^{-6}$	$4.8480 \cdot 10^{-4}$	-

Fonte: Autor (2020).

De acordo com a Tabela 7, os modelos obtidos a partir da largura fixa de *kernel* igual a 1 e 2 não apresentaram bom desempenho. Novamente, o melhor desempenho corresponde ao modelo $RNA_{MCC,ADAP}$ com largura de *kernel* adaptativa. Isto pode ser observado tanto pela sua média de MSE de validação quando pelo seu melhor MSE de validação.

As quantificações dos momentos estatísticos do histograma da Figura 29 são ilustrados na Tabela 8, onde a média está bem próxima de zero, bem como o desvio padrão, para todos os modelos. Ainda neste contexto, o modelo estimado com largura de *kernel* adaptativa apresenta uma curtose positiva mais elevada em relação aos demais modelos, o que implica em concluir que existem erros com maior frequência ocorrendo próximo da média. A assimetria negativa presente, ainda neste modelo, implica que a distribuição está concentrada levemente a esquerda do eixo central.

Tabela 8 – Medidas Estatísticas do erro de estimação - Estado de Saúde - Resultado de treinamento com a presença de ruído e *outliers*.

	Erro de Estimação $\sigma = \text{Adaptativo}$	Erro de Rastreamento $\sigma = 1$	Erro de Rastreamento $\sigma = 2$	Erro de Rastreamento $\sigma = 3$
Média	$1.4356 \cdot 10^{-4}$	$3.1029 \cdot 10^{-4}$	$1.2481 \cdot 10^{-4}$	$-7.4119 \cdot 10^{-4}$
Desvio Padrão	0.0274	0.0230	0.0263	0.0207
Assimetria	-2.9394	-0.9710	0.0375	0.1232
Curtose	21.4687	7.3028	2.8659	3.0520

5.3.3 Resultados - Presença de ruído e outliers - Comparação com MSE clássica

Nesta seção são apresentados os resultados obtidos na identificação do estado de saúde em baterias Lítio-íon, quando os dados de treinamento e validação utilizados no procedimento estão sujeitos à presença de ruído gaussiano branco e *outliers*.

A metodologia utilizada para obtenção dos resultados aqui apresentada é idêntica nas últimas seções. Para cada combinação de algoritmo, função de custo e largura de

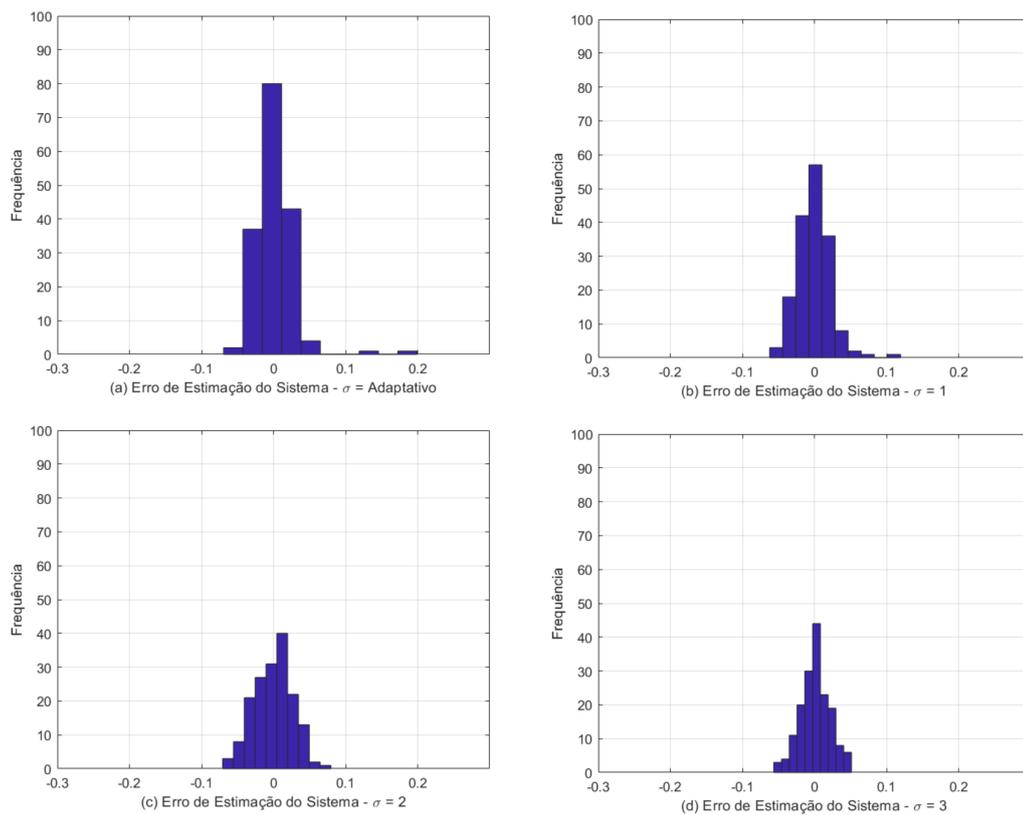
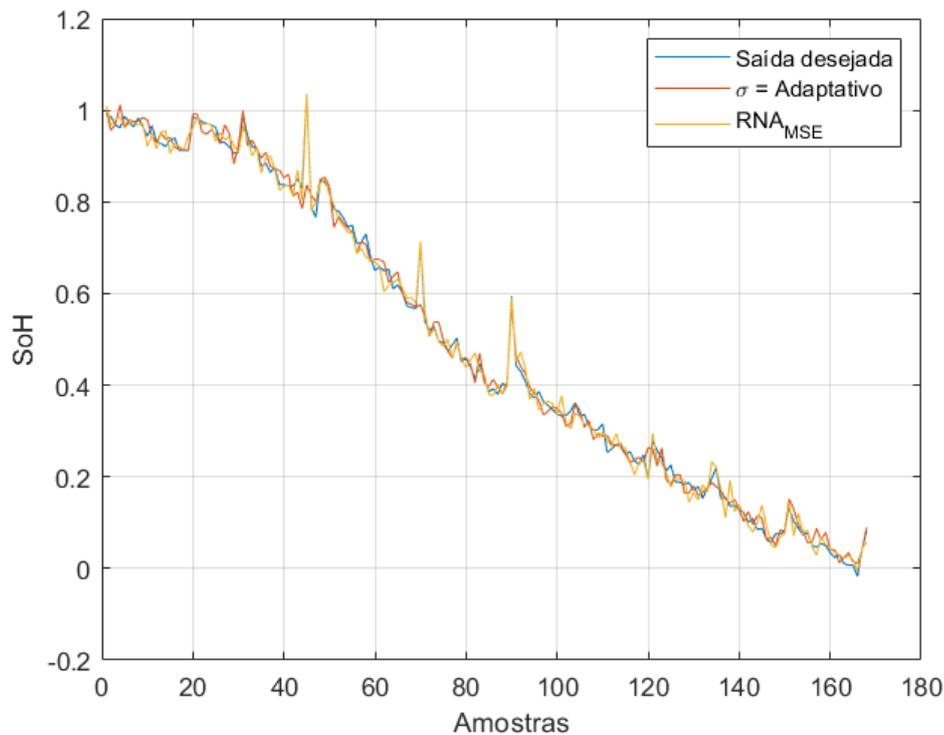


Figura 29 – Medidas Estatísticas do erro de estimação - Estado de Saúde - Resultado de treinamento com a presença de ruído e *outliers*.

kernel foram efetuadas 15 execuções do algoritmo de treinamento. O valor atribuído ao coeficiente de aprendizado η do algoritmo de retropropagação do erro foi igual a 1 e o número de épocas de treinamento foi estabelecido em 500 épocas. O objetivo desta seção é comparar os resultados do melhor modelo com função de custo em entropia e o melhor modelo com função de custo em erro quadrático médio.

Apresenta-se na Figura 30 as respostas do melhor modelo MCC obtido com estratégia adaptativa e o modelo MSE. Comparando-se os modelos MCC e MSE, observa-se que o primeiro apresentou melhor desempenho. Apesar de existir uma porcentagem de *outliers* nos conjuntos de dados de treinamento e validação isto não proporcionou um efeito prejudicial significativo para o modelo MCC. Este modelo foi encontrado pelo algoritmo de retropropagação do erro, utilizando largura de *kernel* adaptativa e com função de custo definida pela maximização da entropia.

Figura 30 – Validação dos modelos RNA - Estado de Saúde - Comparação entre funções de custo MSE e MCC.



Fonte: Autor (2020).

Apresenta-se na Tabela 9 um comparativo de desempenho entre os modelos RNA encontrados quando os dados experimentais presentes na identificação do indicador de carga estão sujeitos a presença de ruído e *outliers*. Novamente, são apresentados para método de estimação, a média do MSE de validação, a variância do MSE de validação e o melhor MSE encontrado em 15 execuções do algoritmo de retropropagação do erro.

Tabela 9 – Comparação de desempenho com função de custo MCC e MSE - Estado de Carga - Resultado de treinamento com a presença de ruído e *outliers*.

Modelo e Treinamento	Média do MSE de Validação	Melhor MSE de Validação	Média do MAPE de Validação	Média do RMSE de Validação
RNA_{MCC}	$2,5 \cdot 10^{-3}$	$2,4 \cdot 10^{-3}$	$6,25 \cdot 10^{-3}$	$20,4 \cdot 10^{-3}$
RNA_{MSE}	$4,2 \cdot 10^{-3}$	$4,1 \cdot 10^{-3}$	$20,2 \cdot 10^{-3}$	$69,5 \cdot 10^{-3}$

5.4 Conclusões

Neste capítulo foram apresentados os dois estudos de caso utilizados para avaliar a aplicação da correntropia como função de custo do algoritmo de retropropagação do erro,

quando este é utilizado para realizar o ajuste dos parâmetros de redes RNA em tarefas de identificação de sistemas dinâmicos não-lineares na presença de ruído e *outliers*.

A adoção de uma estratégia de otimização por algoritmo genético em um contexto ausente de ruído, e a adoção de uma estratégia adaptativa em um cenário na presença de ruídos e *outliers* permitiu a obtenção dos modelos RNA de melhor desempenho para ambos estudos de caso. Com a otimização por algoritmo genético, a largura de *kernel* pode ajustada de acordo com a resposta do sistema dinâmico ao par de entradas η e σ . Na estratégia adaptativa por outro lado, a largura de *kernel* é alterada época a época no procedimento de treinamento.

6 Conclusões

Neste trabalho, foi avaliada a utilização da medida de correntropia no treinamento de Redes Neurais Artificiais para identificação de sistemas não-lineares. Desta forma, o papel da função de custo, tradicionalmente desempenhado pelo erro médio quadrático no algoritmo de retropropagação de erro, é atribuído ao critério de maximização da correntropia. Desta forma, com o objetivo de analisar as possíveis vantagens e desvantagens dessa proposta, foram realizados procedimentos de identificação em que os conjuntos de dados experimentais encontravam-se contaminados por ruído e *outliers*.

Após escrever em detalhes a arquitetura da RNA utilizada neste trabalho, foram apresentadas as equações modificadas do algoritmo de retropropagação de erro resultantes da utilização da correntropia como base de sua função custo, sendo possível notar que mais dois parâmetros ajustáveis foram incluídos ao algoritmo de treinamento, a largura de *kernel* da correntropia e coeficiente de aprendizado. À fim de avaliar esse algoritmo, foram realizados dois estudos de caso que consistiam em aplicar a RNA na identificação do estado de carga e saúde, ambos indicadores de desempenho de baterias.

No treinamento das redes e na consequente obtenção dos modelos RNA, foram atribuídos diferentes valores fixos para a largura de *kernel* da correntropia e inserido ruído e diferentes valores percentuais de *outliers* aos conjuntos de treinamento e validação. Assim, pôde-se notar a importância do adequado ajuste do *kernel* para que o algoritmo modificado possa desempenhar o seu papel de forma satisfatória. Por esta razão, durante a execução do algoritmo de treinamento também foi analisada a estratégia de otimização dos parâmetros por AG e a estratégia adaptativa de ajuste da largura de *kernel* proposta por Liu e Chen (2013).

Em vista dos resultados encontrados dos modelos usados na estimação, alguns comentários podem ser feitos acerca destes:

1. Analisando em um cenário sem a contaminação dos dados de treinamento por ruídos e *outliers*, o modelo com melhor resposta aos dados de validação, em ambos estudos de caso, corresponde ao modelo proposto baseado em uma otimização por algoritmo genético. O algoritmo genético, de maneira simplificada, determinou, a melhor combinação de largura de *kernel* e coeficiente de aprendizado, que resultassem em um erro de estimação reduzido. Isso pode ser constatado nas Tabelas 2 e 6, através dos três indicadores de desempenhos definidos previamente.
2. Analisando em um cenário com a contaminação dos dados de treinamento por ruídos e *outliers*, o modelo com melhor resposta aos dados de validação, em ambos estudos

de caso, foi o modelo proposto baseado em uma estratégia adaptativa. Isso pode ser constatado nas Tabelas 3 e 7, através dos indicadores de desempenho.

3. Analisando o Critério de Maximização de Correntropia, percebe-se que erros demasiadamente grandes elevam o valor da métrica J_{MCC} , o que resulta em uma mudança dos parâmetros (pesos e *bias*) por parte do algoritmo de retropropagação de erro, de modo a reduzir J_{MCC} . A maneira como o algoritmo de retropropagação realiza a mudança destes parâmetros, seja de maneira mais "intensa" ou mais "suave" depende exatamente da largura de *kernel*.
 - Para um cenário sem ruídos e *outliers*, o modelo proposto com largura de *kernel* igual a 1, em ambos estudos de caso, apresentou melhor resultado do que os modelos propostos com largura de *kernel* $\sigma = 2$ e $\sigma = 3$. Demonstrando que o algoritmo atuou de maneira mais incisiva para largura de *kernel* consideradas menores.
 - Para um cenário com ruídos e *outliers*, o modelo proposto com largura de *kernel* igual a 1 não apresentou um resultado satisfatório, isto ocorre devido ao fato de que, pequenos valores de largura de *kernel*, em diferentes tipos de contexto, criam uma janela de observação de erro bastante restrita, o que prejudica o desempenho do sistema. O modelo proposto com largura de *kernel* igual a 2 apresentou melhor resultado do que o modelo proposto com largura de *kernel* igual a 3, mostrando que o algoritmo atuou de maneira incisiva, porém não incisiva o suficiente para prejudicar a janela de observação do erro e conseqüentemente o desempenho do algoritmo.

A comparação dos modelos RNA_{MSE} com os modelos RNA_{MCC} demonstraram que a substituição do erro médio quadrático pela correntropia tornou o algoritmo da retropropagação do erro mais robusto à presença de ruído e *outliers*. Os modelos RNA_{MCC} apresentaram desempenho satisfatório e próximo ao dos modelos RNA_{MSE} obtidos a partir de conjuntos de dados livres destes elementos contaminantes. Logo a aplicação métrica da Correntropia no treinamento das redes *RNA*, quando aplicadas à identificação do *SoC* e *SoH* de baterias Lítio-íon, pode ser considerada uma alternativa interessante para contornar problemas que podem ser causados pela presença de ruídos e *outliers*.

7 Trabalhos Futuros e Perspectivas

Durante os procedimentos de identificação, pôde-se observar a influência significativa da largura de *kernel* no desempenho do algoritmo de treinamento aplicado à rede RNA. A necessidade de ajustar adequadamente este parâmetro pode, inclusive, ser considerada uma desvantagem do modelo apresentado, uma vez que um novo parâmetro é incluído e seu ajuste não é realizado de forma elementar. Mesmo que a estratégia evolutiva e a estratégia adaptativa para o ajuste da largura de *kernel* tenha funcionado satisfatoriamente para os estudos de caso analisados, ela não representa uma solução definitiva. Trabalhos futuros poderão investigar e propor outros mecanismos que tenham como objetivo realizar o ajuste deste parâmetro de forma mais eficiente, permitindo, assim, que as potencialidades da medida de correntropia possam ser melhor aproveitadas. Outra proposta de trabalho futuro seria a de realizar mudanças e testes na Rede Neural de modo aumentar a velocidade de convergência do modelo, como por exemplo, realizar testes com diferentes níveis de camadas, diferentes números de neurônios, diferentes métodos de iniciação dos hiperparâmetros (Pesos e *Bias*).

Neste trabalho foi utilizado o algoritmo básico da retropropagação do erro por se tratar de um algoritmo clássico aplicado ao treinamento de redes neurais. Embora ele tenha apresentado resultados satisfatórios na obtenção dos modelos RNA_{MCC} , outros tipos de algoritmos também podem se ajustar ao uso da correntropia como sua função de custo, ou função de avaliação. Desta forma, pretendem ser realizados trabalhos comparativos entre distintos algoritmos modificados pela utilização da medida de correntropia quando aplicados no treinamento de redes RNA.

Referências

- Aguirre, L. A. *Introdução à Identificação de Sistemas: técnicas lineares e não lineares aplicadas a sistemas reais*. [S.l.]: Editora UFMG, 2007. Citado na página 26.
- Al-Haj Hussein, A.; Batarseh, I. A review of charging algorithms for nickel and lithium battery chargers. *IEEE Transactions on Vehicular Technology*, v. 60, n. 3, p. 830–838, 2011. Citado na página 18.
- Araújo Junior, J. M. *Identificação não linear usando uma rede fuzzy wavelet neural network modificada*. Tese (Doutorado) — Universidade Federal do Rio Grande do Norte, Natal, RN, Brasil, 2014. Citado na página 42.
- Ashwin, T. *et al.* Prediction of battery storage ageing and solid electrolyte interphase property estimation using an electrochemical model. *Journal of Power Sources*, v. 385, p. 141 – 147, 2018. Citado na página 21.
- Aslan, E.; Yasa, Y. A review on the battery state of charge estimation methods for electric vehicle battery management systems. In: . [S.l.: s.n.], 2019. p. 281–285. Citado na página 20.
- Bach, S. Estudo do efeito dos parâmetros genéticos sobre a solução otimizada e sobre o tempo de convergência em algoritmos genéticos com codificações binária e real - doi: 10.4025/actascitechnol.v25i2.2169. *Acta Scientiarum-technology - ACTA SCI-TECHNOL*, v. 25, 04 2003. Citado na página 53.
- Barcellos, J. C. H. *Algoritmos genéticos adaptativos: um estudo comparativo*. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, 2000. Citado na página 53.
- Barré, A. *et al.* A review on lithium-ion battery ageing mechanisms and estimations for automotive applications. *Journal of Power Sources*, v. 241, p. 680 – 689, 2013. Citado na página 21.
- Bessa, R. J.; Miranda, V.; Gama, J. Entropy and correntropy against minimum square error in offline and online three-day ahead wind power forecasting. *IEEE Transactions on Power Systems*, v. 24, n. 4, p. 1657–1666, 2009. Citado 7 vezes nas páginas 14, 34, 35, 37, 38, 49 e 72.
- Cai, Z.; Liu, G.; Luo, J. Research state of charge estimation tactics of nickel-hydrogen battery. In: *2010 International Symposium on Intelligence Information Processing and Trusted Computing*. [S.l.: s.n.], 2010. p. 184–187. Citado na página 17.
- Carvalho, W. L. O. *Estudo de parâmetros ótimos em algoritmos genéticos elitistas*. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Norte, Natal, 2017. Citado 2 vezes nas páginas 52 e 53.
- Chang, W.-Y. State of charge estimation for lifepo4 battery using artificial neural network. *International Review of Electrical Engineering*, v. 7, p. 5874–5880, 10 2012. Citado na página 17.

- Chatterjee, S. *et al.* Software fault prediction using nonlinear autoregressive with exogenous inputs (narx) network. *Applied Intelligence*, v. 37, 07 2012. Citado na página 25.
- Cheng, Y. *et al.* Residual lifetime prediction for lithium-ion battery based on functional principal component analysis and bayesian approach. *Energy*, v. 90, p. 1983 – 1993, 2015. Citado na página 22.
- Corrêa, L. G. *Memória Associativa em Redes Neurais Realimentadas*. Dissertação (Mestrado) — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Paulo, 2004. Citado na página 28.
- Darwin, C. R. *A origem das espécies*. São Paulo, Brasil: Hemus, 1979. Citado na página 48.
- Du, L.; Cheng, X. The research on battery soc estimation within colored noise. In: . [S.l.: s.n.], 2014. p. 1–5. Citado na página 17.
- Eddahech, A. *et al.* Behavior and state-of-health monitoring of li-ion batteries using impedance spectroscopy and recurrent neural networks. *International Journal of Electrical Power Energy Systems*, v. 42, n. 1, p. 487 – 494, 2012. Citado na página 23.
- Feng, X. *et al.* Using probability density function to evaluate the state of health of lithium-ion batteries. *Journal of Power Sources*, v. 232, p. 209 – 218, 2013. Citado na página 22.
- Fontes, A. *Classificação automática de modulação digital com uso de medida teórica de informação para ambientes de rádio cognitivo*. Tese (Doutorado) — Universidade Federal do Rio Grande do Norte, Natal, RN, Brasil, 2012. Citado 2 vezes nas páginas 33 e 41.
- Galeotti, M. *et al.* Synthetic methods for the evaluation of the state of health (soh) of nickel-metal hydride (nimh) batteries. *Energy Conversion and Management*, v. 92, p. 1 – 9, 2015. Citado na página 22.
- Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. *Journal of Machine Learning Research - Proceedings Track*, v. 9, p. 249–256, 01 2010. Citado na página 42.
- Goodenough, J. B.; Kim, Y. Challenges for rechargeable li batteries. *Chemistry of Materials*, v. 22, n. 3, p. 587–603, 2010. Citado na página 20.
- Guo, Y.; Zhao, Z.; Huang, L. Soc estimation of lithium battery based on improved bp neural network. *Energy Procedia*, v. 105, p. 4153 – 4158, 2017. 8th International Conference on Applied Energy, ICAE2016, 8-11 October 2016, Beijing, China. Citado na página 48.
- Haoran, L. *et al.* Lithium battery soc estimation based on extended kalman filtering algorithm. In: *2018 IEEE 4th International Conference on Control Science and Systems Engineering (ICCSSE)*. [S.l.: s.n.], 2018. p. 231–235. Citado na página 19.
- Haykin, S. *Redes neurais: princípios e prática*. second. Porto Alegre, RS, Brasil.: Bookman, 2001. Citado 6 vezes nas páginas 13, 28, 30, 32, 46 e 47.

He, H.; Xiong, R.; Guo, H. Online estimation of model parameters and state-of-charge of lifepo4 batteries in electric vehicles. *Applied Energy*, v. 89, n. 1, p. 413 – 420, 2012. ISSN 0306-2619. Special issue on Thermal Energy Management in the Process Industries. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0306261911005010>>. Citado na página 17.

Holland, J. H. *Adaptation in natural and artificial systems*. Ann Arbor, MI: The University of Michigan Press, 1975. Citado na página 48.

Hussein, A. A. Experimental modeling and analysis of lithium-ion battery temperature dependence. In: *2015 IEEE Applied Power Electronics Conference and Exposition (APEC)*. [S.l.: s.n.], 2015. Citado na página 18.

Hussein, A. A. A neural network based method for instantaneous power estimation in electric vehicles' li-ion batteries. In: *2017 IEEE Applied Power Electronics Conference and Exposition (APEC)*. [S.l.: s.n.], 2017. p. 3122–3126. ISSN 2470-6647. Citado na página 21.

Hussein, A. A. *et al.* Distributed battery micro-storage systems design and operation in a deregulated electricity market. *IEEE Transactions on Sustainable Energy*, v. 3, n. 3, p. 545–556, 2012. Citado na página 18.

Ismail, M. *et al.* Battery state of charge estimation using an artificial neural network. In: *2017 IEEE Transportation Electrification Conference and Expo (ITEC)*. [S.l.: s.n.], 2017. p. 342–349. Citado na página 37.

Jacinto, C. *Acoplamento, simulação e otimização de estratégias de desenvolvimento de campos de petróleo e gás sob incerteza, com aplicações na construção de poços e campos inteligentes*. Tese (Doutorado) — COPPE/UFRJ, Rio de Janeiro, Brasil, 2009. Citado na página 53.

Li Xue *et al.* Effects analysis of model parameters uncertainties on battery soc estimation using h-infinity observer. In: *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*. [S.l.: s.n.], 2014. p. 1647–1653. Citado na página 17.

Linden, R. *Algoritmos genéticos*. third. Rio de Janeiro, Brasil: Editora Ciência Moderna, 2012. Citado na página 52.

Linhares, L. *et al.* A nonlinear system identification approach based on fuzzy wavelet neural network. *Journal of Intelligent and Fuzzy Systems*, v. 28, p. 225–235, 07 2014. Citado 2 vezes nas páginas 26 e 30.

Linhares, L. L. S.; de Araújo, J. M.; Araújo, F. M. U. Nonlinear system identification of a lower limb model by fuzzy wavelet neural networks. In: *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*. [S.l.: s.n.], 2013. p. 3628–3633. Citado na página 26.

Liu, Y.; Chen, J. Correntropy-based kernel learning for nonlinear system identification with unknown noise: an industrial case study. *IFAC Proceedings Volumes*, v. 46, n. 32, p. 361 – 366, 2013. ISSN 1474-6670. 10th IFAC International Symposium on Dynamics and Control of Process Systems. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1474667015382847>>. Citado 6 vezes nas páginas 14, 34, 37, 38, 41 e 42.

- Lorenzi, A. *et al.* Artificial neural networks application to predict bond steel-concrete in pull-out tests. *Revista IBRACON de Estruturas e Materiais (ISSN 1983-4195)*, v. 10, p. 1051–1062, 10 2017. Citado na página 28.
- Miranda, R. C.; Pinho, A. F.; Montevechi, J. A. B. *Análise de significância dos parâmetros de um algoritmo genético através do planejamento de experimentos*. [S.l.]: Cadernos do IME - Série Estatística, 2011. v. 31. Citado 2 vezes nas páginas 52 e 53.
- Miranda, V.; Cerqueira, C.; Monteiro, C. Training a fis with eps0 under an entropy criterion for wind power prediction. In: *Proc. 10th IFAC Int. Symp. on Dynamics and Control of Process Systems*. Stockholm, Sweden: [s.n.], 2006. p. 1–8. Citado 2 vezes nas páginas 14 e 37.
- Munoz, J.; Chen, J. Removal of the effects of outliers in batch process data through maximum correntropy estimator. *Chemometrics and Intelligent Laboratory Systems*, v. 111, p. 53–58, 02 2012. Citado 3 vezes nas páginas 25, 49 e 72.
- Nagahamulla, H. R. K.; Ratnayake, U. R.; Ratnaweera, A. Monsoon rainfall forecasting in sri lanka using artificial neural networks. In: *2011 6th International Conference on Industrial and Information Systems*. [S.l.: s.n.], 2011. p. 305–309. Citado na página 23.
- Nelles, O. *Nonlinear System Identification*. [S.l.: s.n.], 2001. v. 13. Citado na página 26.
- Norgaard, M. *et al.* Neural networks for modelling and control of dynamic systems. *Springer- Verlag London Limited*, 2001. Citado na página 26.
- Pinho, A.; Montevechi, J. A. B.; Marins, F. Análise da aplicação de projeto de experimentos nos parâmetros dos algoritmos genéticos. *Sistemas Gestão*, v. 2, 05 2009. Citado 2 vezes nas páginas 52 e 53.
- Plett, G. L. High-performance battery-pack power estimation using a dynamic cell model. *IEEE Transactions on Vehicular Technology*, v. 53, n. 5, p. 1586–1593, 2004. Citado na página 18.
- Principe, J. *Information theoretic learning: Renyi's entropy and kernel perspectives*. [S.l.]: Springer, 2010. Citado 7 vezes nas páginas 14, 32, 33, 34, 35, 38 e 48.
- Rahimi-Eichi, H. *et al.* Battery management system: An overview of its application in the smart grid and electric vehicles. *IEEE Industrial Electronics Magazine*, v. 7, n. 2, p. 4–16, 2013. Citado na página 12.
- Rao, Z.; Wang, S.; Zhang, G. Simulation and experiment of thermal energy management with phase change material for ageing lifepo4 power battery. *Energy Conversion and Management*, v. 52, n. 12, p. 3408 – 3414, 2011. Citado na página 17.
- Remmlinger, J. *et al.* State-of-health monitoring of lithium-ion batteries in electric vehicles by on-board internal resistance estimation. *Journal of Power Sources*, v. 196, n. 12, p. 5357 – 5363, 2011. Citado na página 22.
- Rivera-Barrera, J. P.; Muñoz-Galeano, N. Soc estimation for lithium-ion batteries: Review and future challenges. *Electronics*, v. 6, p. 102, 11 2017. Citado na página 18.
- Rosolem, M. d. F. N. C. e. a. Bateria de lítio-íon: conceitos básicos e potencialidades. *Caderno CPqD Tecnologia*, 2012. Citado na página 12.

- Saha, B. *et al.* Battery health management system for electric uavs. In: *2011 Aerospace Conference*. [S.l.: s.n.], 2011. p. 1–9. Citado 2 vezes nas páginas 15 e 54.
- Saji, D.; Babu, P. S.; Ilango, K. Soc estimation of lithium ion battery using combined coulomb counting and fuzzy logic method. In: *2019 4th International Conference on Recent Trends on Electronics, Information, Communication Technology (RTEICT)*. [S.l.: s.n.], 2019. p. 948–952. Citado na página 19.
- Salkind, A. J. *et al.* Determination of state-of-charge and state-of-health of batteries by fuzzy logic methodology. *Journal of Power Sources*, v. 80, n. 1, p. 293 – 300, 1999. Citado na página 22.
- Samadani, E.; Fraser, R.; Fowler, M. A review study of methods for lithium-ion battery health monitoring and remaining life estimation in hybrid electric vehicles. In: . [S.l.: s.n.], 2012. Citado na página 13.
- Santamaria, I.; Pokharel, P. P.; Principe, J. C. Generalized correlation function: definition, properties, and application to blind equalization. *IEEE Transactions on Signal Processing*, v. 54, n. 6, p. 2187–2197, 2006. Citado 3 vezes nas páginas 32, 33 e 34.
- Scrosati, B.; Garche, J. Lithium batteries: Status, prospects and future. *Journal of Power Sources*, v. 195, n. 9, p. 2419 – 2430, 2010. Citado na página 20.
- Silvermann, B. W. *Density estimation for statistics and data analysis*. London, England: Chapman and Hall CRC, 1986. v. 26. Citado na página 41.
- Sjöberg, J. *et al.* Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, v. 31, n. 12, p. 1691 – 1724, 1995. Citado na página 13.
- Smith, K. *et al.* Comparison of plug-in hybrid electric vehicle battery life across geographies and drive cycles. *SAE International*, 2012. Citado na página 13.
- Sriram, R. D. *Intelligent Systems for Engineering: A Knowledge-based Approach*. Londres: Springer-Verlag, 1997. Citado na página 28.
- Tertuliano, F. *et al.* Estudo comparativo entre ukf e ekf na estimação de soc em baterias de lítio-íon sob diferentes condições de temperatura. *Sociedade Brasileira de Automática*, 2020. Citado na página 18.
- Thuillard, M. A review of wavelet networks, wavenets, fuzzy wavenets and their applications. *ESIT 2000*, v. 18, 04 2001. Citado na página 31.
- Topan, P. A. *et al.* State of charge (soc) and state of health (soh) estimation on lithium polymer battery via kalman filter. In: *2016 2nd International Conference on Science and Technology-Computer (ICST)*. [S.l.: s.n.], 2016. p. 93–96. Citado na página 22.
- Uddin, K. *et al.* Techno-economic analysis of the viability of residential photovoltaic systems using lithium-ion batteries for energy storage in the united kingdom. *Applied Energy*, v. 206, p. 12 – 21, 2017. Citado na página 21.
- Vetter, J. *et al.* Ageing mechanisms in lithium-ion batteries. *Journal of Power Sources*, v. 147, n. 1, p. 269 – 281, 2005. Citado na página 13.

- 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. ISSN 0378-7753. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378775314002572>>. Citado 2 vezes nas páginas 18 e 34.
- Wilks, D. S. Statistical methods in the atmospheric sciences. *International Geophysics Series*, v. 91, p. 627, 2006. Citado na página 58.
- Xie, H.; Tang, H.; Liao, Y.-H. Time series prediction based on narx neural networks: An advanced approach. In: . [S.l.: s.n.], 2009. v. 3, p. 1275 – 1279. Citado na página 44.
- Xu, D.; Wang, L.; Yang, J. Research on li-ion battery management system. In: *2010 International Conference on Electrical and Control Engineering*. [S.l.: s.n.], 2010. p. 4106–4109. Citado na página 17.
- Yan, Q.; Wang, Y. Predicting for power battery soc based on neural network. In: *2017 36th Chinese Control Conference (CCC)*. [S.l.: s.n.], 2017. p. 4140–4143. Citado na página 19.
- Zhao, S.; Chen, B.; Príncipe, J. C. Kernel adaptive filtering with maximum correntropy criterion. In: *The 2011 International Joint Conference on Neural Networks*. [S.l.: s.n.], 2011. p. 2012–2017. Citado na página 41.
- Zhou, J. *et al.* Battery state of health estimation using the generalized regression neural network. In: *2015 8th International Congress on Image and Signal Processing (CISP)*. [S.l.: s.n.], 2015. p. 1396–1400. Citado 2 vezes nas páginas 21 e 23.