

Herberth Birck Fröhlich

**REDES NEURAIS ESPECIALIZADAS PARA INFERÊNCIA DE  
REGIME PERMANENTE EM TESTES DE PERFORMANCE DE  
COMPRESSORES DE REFRIGERAÇÃO**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Engenharia Mecânica.

Orientador: Prof. Carlos Alberto Flesch, Dr. Eng.

Coorientador: Prof. Cesar Alberto Penz, Dr. Eng.

Florianópolis  
2016

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Fröhlich, Herberth Birck

Redes neurais especializadas para inferência de regime permanente em testes de performance de compressores de refrigeração / Herberth Birck Fröhlich ; orientador, Carlos Alberto Flesch ; coorientador, Cesar Alberto Penz. - Florianópolis, SC, 2016.  
130 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Mecânica.

Inclui referências

1. Engenharia Mecânica. 2. Ensaio de desempenho energético. 3. Redes neurais para inferência de regime permanente. 4. Compressores de refrigeração. 5. Inteligência artificial. I. Flesch, Carlos Alberto. II. Penz, Cesar Alberto. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Mecânica. IV. Título.

Herberth Birck Fröhlich

**REDES NEURAIIS ESPECIALIZADAS PARA INFERÊNCIA DE  
REGIME PERMANENTE EM TESTES DE PERFORMANCE DE  
COMPRESSORES DE REFRIGERAÇÃO**

Esta Dissertação foi julgada adequada para obtenção do Título de “Mestre em Engenharia Mecânica”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Mecânica.

Florianópolis, 21 de Dezembro de 2016.

---

Prof. Jonny Carlos da Silva, Dr. Eng.  
Coordenador do Curso

---

Prof. Carlos Alberto Flesch, Dr. Eng. – Orientador  
Universidade Federal de Santa Catarina

---

Prof. Cesar Alberto Penz, Dr. Eng. – Coorientador  
Instituto Federal de Santa Catarina

**Banca Examinadora:**

---

Prof. Carlos Alberto Flesch, Dr. Eng. – Presidente  
Universidade Federal de Santa Catarina

---

Prof. Marco Antônio Martins Cavaco, Dr. Eng. - Examinador  
Universidade Federal de Santa Catarina

---

Prof. Sílvia Modesto Nassar, Dr. Eng. - Examinadora  
Universidade Federal de Santa Catarina

---

Marco Aurélio de Oliveira, Dr. Eng. - Examinador  
UNISOCIESC



Este trabalho é dedicado à minha avó Imelda e à minha mãe, exemplos de mulheres que seguiram sempre em frente.



## **AGRADECIMENTOS**

À minha família pelo apoio incondicional.

À Pamzinha pela paciência e parceria em todos momentos.

Aos meus amigos, próximos ou distantes.

Ao orientador professor Carlos Flesch pela prestatividade e auxílio.

Ao coorientador Cesar Penz por todas sugestões que fizeram o trabalho e o acadêmico melhor.

À Embraco, ao LABMETRO e ao LIAE pelo suporte técnico.

À CAPES pelo financiamento durante a realização do trabalho.

Enfim, a todos que contribuíram de alguma forma para que esse trabalho fosse realizado.





*Oh, the time will come up  
When the winds will stop  
And the breeze will cease to be breathin'  
Like the stillness in the wind  
Before the hurricane begins  
The hour that the ship comes in*

When the Ship Comes in, Bob Dylan



## RESUMO

Ensaio de desempenho energético de compressores de refrigeração têm como objetivo a obtenção de parâmetros de desempenho, dentre eles a capacidade de refrigeração, sob condições preestabelecidas de operação. É uma etapa necessária no desenvolvimento e produção de compressores e que, devido às suas características de execução, torna-se um gargalo no processo de controle de qualidade. Trabalhos anteriores apresentam soluções associadas à inteligência artificial para reduzir o tempo de execução desses ensaios. Para isso, tais ferramentas analisam o comportamento de certas variáveis durante o período transiente do ensaio para inferir o momento no qual ocorre a transição para o regime permanente. Entretanto, a grande variedade de dinâmicas relacionadas a capacidade, temperatura de corpo e pressão de sucção afeta o desempenho das redes neurais. Para contornar esse problema, ensaios provenientes de uma empresa fabricante de compressores foram agrupados manualmente de acordo com regras que levaram em consideração a capacidade inicial e a região da faixa de aceitação pela qual os ensaios adentravam o regime permanente. Isso forneceu 4 grupos, os quais foram utilizados para treinamento de redes especializadas para inferência de regime permanente. Os melhores resultados vieram de redes especializadas treinadas com dois dos quatro grupos, que quando comparadas com aquelas treinadas com ensaios não agrupados dessas mesmas dinâmicas, apresentaram desempenho superior tanto nas taxas de acerto (aumento de até 6%) quanto nas taxas de erro de falso positivo (redução de até 17%). Paralelamente, foram realizadas análises estatísticas entre grupos na busca de comportamentos consideravelmente predominantes. Como destaque, um desses grupos apresentou duração de regime transitório consideravelmente inferior a dos demais. Além disso, como o agrupamento manual é lento e demanda um operador especialista, foi criado um método automático para esse fim que apresentou acerto de 92% em uma primeira análise. De forma geral, os resultados forneceram maior confiabilidade às redes de inferência e motivam a continuação dos estudos sobre as características relacionadas ao processo de treinamento.

**Palavras-chave:** Ensaio de desempenho energético. Compressores de refrigeração. Inteligência artificial. Agrupamento de ensaios.



## ABSTRACT

Energetic performance tests of refrigerating compressors are used to obtain performance parameters, among them refrigerating capacity under pre-established operating conditions. It is a necessary step on development and compressor manufacturing, and due to its execution characteristics, become a bottleneck on quality control process. Previous works presented solutions based on artificial intelligence, which were used to reduce test time. For this, such tools analyse the behaviours of certain variables during the test transient state to infer on which moment occurs the transition to steady state. However, the considerable variety of dynamics related to such used parameters affects the neural network performance. To avoid such issue, tests from a compressor manufacturer were manually grouped accordingly to rules that took into account initial capacity value and acceptance band region by which the test entered steady state. This provided four groups, which were used to train specialized neural networks. The best results came from specialized neural networks trained with two of those groups, which when compared to those neural networks trained with non-grouped tests of these same dynamics, presents superior performance in both of true positive rate (up to 6% increase) and false positive rate (up to 17% decrease). Parallel to it, statistical analysis were conducted between groups in the search for predominant behaviours. As a highlight, one of these groups had both overall duration and transient state duration considerably lower than the other groups. In addition, since manual grouping is slow and requires a specialized operator and automatic method developed for this purpose, which showed a 92% accuracy in a first analysis. In general, results provided greater reliability to the inference neural networks and motivate the continuity of studies about training process related characteristics.

**Keywords:** Energetic performance test. Refrigerating compressor. Artificial intelligence. Test grouping.



## LISTA DE FIGURAS

Figura 1 - Ciclo real de refrigeração por compressão de vapor.....	40
Figura 2 - Circuito para medidor de vazão no estado líquido (método F). .....	43
Figura 3 - Transição do regime transitório para o permanente em ensaio da capacidade de refrigeração.....	45
Figura 4 - Resultados obtidos na redução do tempo de ensaio derivados da melhoria no controle e nova instrumentação (sem aceleração do transitório).....	47
Figura 5 - Esquema geral de ferramenta de inferência de regime permanente.....	48
Figura 6 - Diferentes modelos de aprendizado de máquina. (a) regressão linear (b) classificador (c) clusterizador.....	52
Figura 7 - Neurônio biológico e suas partes. ....	53
Figura 8 - Modelo de neurônio ( <i>perceptron</i> ) de uma rede neural MLP.....	54
Figura 9 - Funções de ativação limiar (a), logística (b) e tangente hiperbólica (c). ....	55
Figura 10 - Exemplo de rede alimentada adiante com múltiplas camadas. .....	57
Figura 11 - Função erro e representação dos mínimos local, global e tamanho do passo via gradiente descendente. ....	58
Figura 12 - Dilema <i>bias</i> variância. (a) Falta de treinamento. (b) Excesso de ajuste. (c) Boa generalização. ....	61
Figura 13 - Aspectos gráficos da validação cruzada e do conceito de generalização. ....	62
Figura 14 - Visualização de um teste de Kolmogorov-Smirnov onde as curvas pertencem à mesma distribuição.....	66
Figura 15- Exemplo de classificadores em um ROC. ....	72
Figura 16 - Curva ROC criada a partir de diferentes limiares. ....	73
Figura 17 - ROC com exemplos de dois classificadores com mesmo AUC. .....	74
Figura 18 - Isolinhas de exatidão. Linhas sólidas para $C=1$ e tracejadas para $C=0,5$ .....	75

Figura 19 - Comparação de classificador por isolinhas de exatidão (cinza) para dois classificadores criados a partir de diferentes matrizes de classificação. $C = 1$ .	76
Figura 20 - ROC para diferentes limiares com isolinhas de exatidão (mais claras) para curva criada a partir de ranqueamento. $C = 1$ .	77
Figura 21 - Exemplo de média de limiar e diferentes intervalos de confiança para os métodos de junção por FPR, TPR e caixa.	78
Figura 22 - Representação das entradas relacionadas a capacidade	83
Figura 23 - Intervalo do ensaio utilizado para treinamento das redes.	84
Figura 24 - Exemplo de aplicação de limiar para discretização de saída contínua de uma rede neural. Limiar = 0,75.	84
Figura 25 - Aspecto geral da dinâmica dos ensaios agrupados. Os limites de 2% correspondem ao regime permanente.	86
Figura 26 - Duração em grupos. (a) Diagrama de caixa (b) Teste de Dunn. $\alpha = 0,05$ .	89
Figura 27 - Momento de entrada em regime permanente para grupos. (a) Diagrama de caixa. (b) Teste de Dunn. $\alpha = 0,05$ .	90
Figura 28 - Variação da temperatura de corpo em grupos. (a) Diagrama de caixa. (b) Teste de Dunn.	91
Figura 29 - Módulo da variação absoluta da pressão de sucção em grupos. (a) Diagrama de caixa. (b) Teste de Dunn.	92
Figura 30 - Módulo da variação da pressão de descarga em grupos. (a) Diagrama de caixa. (b) Teste de Dunn.	93
Figura 31 - Procedimento de treinamento e utilização das redes.	96
Figura 32 - Grupo G1A. Comitê de 9 redes, FPR=0,14 e correlação máxima de 0,53.	97
Figura 33 - Faixas de incerteza em ROC para as redes especializadas G2A e para redes não especializadas G-G2A.	101
Figura 34 - Faixas de incerteza em ROC para as redes especializadas G2B e para redes não especializadas G-G2B.	102
Figura 35 - Diferentes janelas de confirmação para definição do grupo e demonstração de relevância.	104
Figura 36 - Exemplo de áreas de contagem para janela de confirmação de 70%.	105



Figura 37 - Aplicação de dois casos de relevância (pontos em cinza claro): (a) influência na determinação do grupo; (b) permanência em regime permanente. ....	106
Figura 38 - Impacto da suavização com diferentes <i>spans</i> na região que antecede a entrada em regime. ....	107
Figura 39 - Impacto da suavização com diferentes <i>spans</i> na permanência em regime. ....	108
Figura 40 - FDA dos testes de Kolmogoro-Smirnov e histogramas das durações de ensaios. ....	121
Figura 41 - FDA dos testes de Kolmogoro-Smirnov e histogramas dos momentos de entrada em regime permanente dos ensaios. ....	123
Figura 42 - FDA dos testes de Kolmogoro-Smirnov e histogramas da variação da temperatura de corpo. ....	124
Figura 43 - FDA dos testes de Kolmogoro-Smirnov e histogramas da variação absoluta de pressão de sucção. ....	126
Figura 44 - FDA dos testes de Kolmogoro-Smirnov e histogramas da variação absoluta de pressão de descarga. ....	127



## LISTA DE TABELAS

Tabela 1 - Conjunto de regras para o comitê de inferência da entrada em regime permanente da capacidade de refrigeração.....	49
Tabela 2 - Tempos médios de inferência para cada regra.....	50
Tabela 3 - Critérios de parada de treinamento.....	81
Tabela 4 - Incidência dos grupos após agrupamento manual.....	87
Tabela 5 – Ensaios provenientes de apenas uma planta, seus respectivos grupos e aqueles ensaiados na condição de <i>checkpoint</i> .....	88
Tabela 6 - Separação de ensaios por finalidade.....	95
Tabela 7 - Respetivos números de redes, correlações, e valores de FPR para os grupos.....	97
Tabela 8 - Exemplo de aplicação de votação majoritária para obtenção de rede resultante.....	98
Tabela 9 - Resultados obtidos das redes especializadas ( $G^{**}$ ) e não especializadas ( $G-G^{**}$ ) para FPR, TPR, e exatidão.....	99
Tabela 10 - Resultados obtidos das redes especializadas ( $G^{**}$ ) e não especializadas ( $G - G^{**}$ ) para AUC e faixa de variância.....	100
Tabela 11 - Parâmetros de desempenho provenientes da aplicação das redes especializadas (G2A e G2B) e redes não especializadas (G-G2A e G-G2B). Médias e incertezas associadas.....	103
Tabela 12 - Acerto do agrupamento automático para diferentes valores de relevância e janela de confirmação.....	108



## LISTA DE QUADROS

Quadro 1 – Matriz de classificação. ....	69
Quadro 2 - Entradas da rede de inferência de regime permanente. ....	82
Quadro 3 - Variáveis e respectivas diferenciações mais acentuadas provenientes do teste de Kruskal-Wallis e de Dunn. ....	94
Quadro 4 - Regras para agrupamento automático. ....	105
Quadro 5 - Testes de normalidade por Kolmogorov-Smirnov. $H_0$ =ambas amostras pertencem à uma distribuição normal. Rejeita $H_0$ se $K < CV$ . $\alpha=0,05$ . ....	121
Quadro 6 - Resultados da análise de variância. $H_0$ de que não há diferença entre tratamentos. ....	129



## LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
ANNNG	Rede Neural Artificial Não Especializada
ANNS	Rede Neural Artificial Especializada
ANOVA	Analysis of Variance
AUC	Area Under Curve
ASHRAE	American Society of Heating, Refrigerating, and Air-Conditioning Engineers
BC <sub>a</sub>	Bias-Corrected And Accelerated
BIPM	Bureau International Des Poids Et Mesures
CAP	Capacidade de Refrigeração
CON	Potência Elétrica Consumida
COP	Coeficiente de Performance
EBP	Error Back Propagation
FDA	Função de Distribuição Acumulada
FN	Falso negativo
FP	Falso positivo
FPR	False Positive Rate
GL	Graus de Liberdade
GUM	Guide for Uncertainty Measurement
IC	Intervalo de Confiança
INMETRO	Instituto Nacional de Metrologia, Qualidade, e Tecnologia
ISO	International Organization for Standardization
JC	Janela de Confirmação
MQ	Média Quadrática
MLP	Multi Layer Perceptron
MSE	Mean Squared Error
RNA	Rede Neural Artificial
ROC	Receiver Operating Characteristics
RP	Regime Permanente
SQ	Soma Quadrática
TN	True Negative
TP	True Positive
TPR	True Positive Rate





## SUMÁRIO

<b>INTRODUÇÃO</b> .....	<b>35</b>
1.1. OBJETIVOS .....	37
1.2. ESTRUTURA DO DOCUMENTO .....	37
<b>2. ENSAIOS DE DESEMPENHO</b> .....	<b>39</b>
2.1. DESCRIÇÃO GERAL .....	39
<b>2.1.1. Obtenção da capacidade em ensaios de desempenho</b> .....	<b>39</b>
<b>2.1.2. Medição de vazão no estado líquido</b> .....	<b>42</b>
<b>2.1.3. Regimes e aspectos normativos</b> .....	<b>44</b>
2.2. REDUÇÃO DO TEMPO DE ENSAIO .....	46
<b>2.2.1. Melhoramentos do calorímetro</b> .....	<b>46</b>
<b>2.2.2. Soluções utilizando inteligência artificial</b> .....	<b>48</b>
<b>3. INTELIGÊNCIA ARTIFICIAL</b> .....	<b>51</b>
3.1. REDES NEURAIS ARTIFICIAIS .....	52
<b>3.1.1. Modelo de neurônio <i>perceptron</i> de múltiplas camadas</b> .....	<b>53</b>
<b>3.1.2. Arquitetura de redes neurais</b> .....	<b>56</b>
<b>3.1.3. Treinamento de redes neurais</b> .....	<b>57</b>
3.1.3.1. Conjunto de entrada .....	59
3.1.3.2. Generalização e critérios de parada .....	60
<b>3.1.4. Regras para criação de comitês</b> .....	<b>62</b>
<b>4. MÉTODOS ESTATÍSTICOS E DE COMPARAÇÃO DE CLASSIFICADORES</b> .....	<b>65</b>
4.1. MÉTODOS ESTATÍSTICOS .....	65
<b>4.1.1. Teste F para curvas não paramétricas</b> .....	<b>66</b>
<b>4.1.2. Testes de hipótese de Dunn</b> .....	<b>67</b>
<b>4.1.3. Intervalos de confiança para distribuições não normais</b> .....	<b>68</b>
4.2. COMPARAÇÃO DE CLASSIFICADORES .....	69
<b>4.2.1. Curva característica de operação de um receptor (ROC)</b> ....	<b>71</b>
<b>4.2.2. Área sob a curva (AUC) e linhas isométricas</b> .....	<b>73</b>
<b>4.2.3. Avaliação de incertezas em ROC</b> .....	<b>77</b>
<b>5. AGRUPAMENTO DE ENSAIOS E DESENVOLVIMENTO DE CLASSIFICADORES ESPECIALIZADOS</b> .....	<b>81</b>

5.1.	REDE DE INFERÊNCIA DE REGIME PERMANENTE .....	81
5.2.	AGRUPAMENTO DE ENSAIOS .....	85
<b>5.2.1.</b>	<b>Descrição geral do método de agrupamento .....</b>	<b>85</b>
<b>5.2.2.</b>	<b>Análises estatísticas .....</b>	<b>86</b>
5.2.2.1.	Duração do ensaio e momento de entrada em regime permanente em grupos.....	88
5.2.2.2.	Temperatura de corpo em grupos .....	90
5.2.2.3.	Pressão de sucção em grupos .....	91
5.2.2.4.	Pressão de descarga em grupos .....	92
<b>5.2.3.</b>	<b>Resultados referentes ao agrupamento de ensaios .....</b>	<b>93</b>
5.3.	REDES ESPECIALIZADAS .....	94
<b>5.3.1.</b>	<b>Criação dos comitês.....</b>	<b>96</b>
<b>5.3.2.</b>	<b>Comparação entre redes especializadas e não especializadas</b>	<b>98</b>
<b>5.3.3.</b>	<b>Desempenho de redes especializadas .....</b>	<b>102</b>
5.4.	AGRUPAMENTO AUTOMÁTICO .....	103
<b>6.</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>111</b>
6.1.	CONCLUSÕES .....	111
6.2.	SUGESTÃO DE TRABALHOS FUTUROS .....	113
	<b>REFERÊNCIAS.....</b>	<b>115</b>
	<b>APÊNDICE A - TESTES DE NORMALIDADE.....</b>	<b>121</b>
	<b>APÊNDICE B – RESULTADOS DAS ANÁLISES DE VARIÂNCIA POR KRUSKAL-WALLIS.....</b>	<b>129</b>

## INTRODUÇÃO

O objetivo do ensaio de desempenho energético de compressores é a obtenção, sob condições de operação preestabelecidas, dos parâmetros capacidade de refrigeração, potência elétrica ativa consumida e coeficiente de performance. Os métodos de ensaio propostos em norma provêm resultados com incertezas em níveis suficientes para permitir a análise da adequação de um compressor a qualquer condição de operação requerida em um sistema de refrigeração (ISO, 1989).

A automatização e controle das variáveis que influenciam um ensaio de desempenho, mesmo que parcial, minimiza a intervenção do operador, melhora a repetibilidade e reprodutibilidade do ensaio e reduz o tempo de execução, ocasionando aumento da produtividade (FLESCH, 2008). Em bancadas tradicionais de ensaio, o tempo para atingir o regime permanente e viabilizar a obtenção de dados adequados é, em média, de 3 horas (PENZ, 2011). Esse tempo é chamado de regime transitório, que quando somado a uma hora que o compressor deve permanecer em regime permanente acaba resultando em um tempo total de ensaio de, em média, quatro horas (ASHRAE STANDARD, 2005; ISO, 1989). Esse tempo total é um problema que já foi abordado em trabalhos anteriores, e que acaba sendo um gargalo em atividades de pesquisa, desenvolvimento e controle da qualidade da produção.

Como não foram encontrados trabalhos semelhantes nas principais plataformas de divulgação de pesquisa científica, as referências ficaram limitadas aos trabalhos desenvolvidos no Labmetro (Laboratório de Metrologia e Automatização). Isso faz deste laboratório pioneiro nesse campo de pesquisa.

Tais trabalhos abordaram a redução do tempo de ensaio de duas maneiras: controle de hardware/aprimoramento da bancada e por meio da inteligência artificial. A primeira teve foco sobre o melhoramento do hardware e controle da bancada de desempenho de compressores. Nessa linha de pesquisa alguns trabalhos visaram o aumento da confiabilidade metrológica das grandezas medidas (BARBOSA, 2006; FLESCH, 2006; PETROVIC, 2007; POLETTI, 2006). Ainda nessa mesma linha de pesquisa, outros trabalhos buscaram acelerar o regime transitório (LIMA, 2010a; SCUSSEL, 2006; SILVEIRA, 2010). Os resultados dessa primeira abordagem possibilitaram uma redução média de 50% no tempo de ensaio.

Simultaneamente, a segunda abordagem fez uso de técnicas de inteligência artificial para reduzir o tempo de realização de ensaios. Isso

foi possível através de uma ferramenta para inferir o instante de entrada em regime permanente e outra que realizava o prognóstico do valor final da capacidade. Os estudos conduzidos nessa área foram:

HENKLEIN (2006) mesclou redes neurais e lógica *fuzzy* para prognóstico de desempenho. Sua maior contribuição foi evidenciar características importantes de ensaios de compressores que seriam utilizadas posteriormente.

STEINBACH (2008) aplicou redes neurais artificiais para avaliar a estabilidade do ensaio. Isso foi feito a partir da inferência do instante de entrada em regime permanente e no prognóstico dos valores de capacidade de refrigeração, potência elétrica ativa consumida e coeficiente de performance.

CASELLA (2008) avaliou e aplicou redes *fuzzy-bayesianas* para analisar a estabilidade do ensaio a partir da inferência do instante de entrada em regime permanente da capacidade de refrigeração.

LIMA (2010b) desenvolveu e implementou diretamente em bancadas a primeira versão de um conjunto de ferramentas constituída de redes de inferência e prognóstico para potência elétrica ativa consumida, capacidade de refrigeração e coeficiente de performance. Seus resultados apontaram uma redução de 50% de tempo de ensaio em relação aos ensaios padronizados.

PENZ (2011) conseguiu aumentar a confiabilidade dos resultados por meio da implementação de um conjunto de redes aliada ao uso de informações *a priori*. Os resultados obtidos mostraram uma melhora expressiva em comparação às primeiras ferramentas obtidas por LIMA (2010b). No entanto, os estudos realizados por PENZ (2011) ficaram limitados à capacidade de refrigeração, não contemplando o uso de redes para prognósticos da potência elétrica consumida e do coeficiente de performance.

NASCIMENTO (2015) teve acesso a um banco de dados com cerca de 1100 ensaios, os quais foram utilizados para desenvolvimento de ferramentas baseadas em comitês de redes para as outras grandezas de interesse (consumo e coeficiente de performance). Seu trabalho também apontou que as ferramentas de inferência têm desempenho comprometido devido a diferentes comportamentos não previsíveis das variáveis utilizadas para treinamento.

Os comportamentos identificados em NASCIMENTO (2015) estão relacionados à dinâmica dos ensaios, isto é, à maneira como diferentes grandezas se comportam durante o transitório. Dessa forma, torna-se necessário um estudo mais aprofundado sobre o desempenho das ferramentas baseadas em redes neurais frente as diferentes dinâmicas.

## 1.1. OBJETIVOS

O objetivo principal desta dissertação é desenvolver ferramenta baseada em redes neurais especializadas para a inferência de regime permanente em ensaios de desempenho de compressores de refrigeração.

Como objetivos específicos têm-se:

- realizar análise exploratória de dados visando caracterizar o comportamento dinâmico das variáveis envolvidas;
- agrupar ensaios de desempenho pela característica dinâmica de entrada em regime permanente da capacidade de refrigeração;
- desenvolver e comparar o desempenho de redes treinadas com conjunto de dados agrupados (redes especializadas) e não agrupados (redes não especializadas);
- buscar correlações entre comportamentos de diferentes variáveis nos ensaios agrupados;
- desenvolver método para agrupamento de ensaios.

## 1.2. ESTRUTURA DO DOCUMENTO

O capítulo 2 trata sobre ensaios de desempenho, suas principais características e os limites de aceitação estabelecidos por norma. Esse capítulo também discorre sobre as técnicas já utilizadas para redução da duração dos ensaios.

No capítulo 3 está a revisão bibliográfica sobre as redes neurais utilizadas como classificadores neste trabalho, os parâmetros de entrada (*inputs*) e saídas (*outputs*) e construção de comitês de classificadores.

O capítulo 4 trata sobre os métodos estatísticos utilizados e sobre a comparação de classificadores.

O capítulo 5 apresenta o que foi desenvolvido neste trabalho: método de agrupamento de ensaios, criação de redes especializadas e agrupador automático. Esse capítulo ainda apresenta a discussão dos resultados.

No capítulo 6 estão a conclusão acerca dos resultados obtidos e propostas de trabalhos futuros.



## 2. ENSAIOS DE DESEMPENHO

Nesta seção estão os aspectos gerais dos ensaios de desempenho de compressores e as estratégias de redução de tempo de realização desenvolvidas anteriormente.

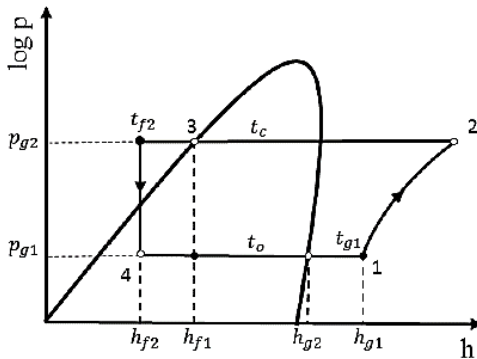
### 2.1. DESCRIÇÃO GERAL

Testes de desempenho são utilizados por fabricantes para garantir que compressores atendam as especificações de catálogo, para criar essas especificações, no controle de qualidade e em pesquisa e desenvolvimento (P&D). O objetivo final do teste de desempenho é a obtenção de vários parâmetros do compressor, entre eles a capacidade de refrigeração ( $\Phi_0$ ).

#### 2.1.1. Obtenção da capacidade em ensaios de desempenho

O ensaio de desempenho energético de compressores utiliza um ciclo de refrigeração por compressão de vapor para obtenção desses parâmetros. A figura 1 apresenta o diagrama pressão por entalpia de um ciclo real desse tipo. Diferente do ciclo ideal, onde não ocorrem perdas de calor para as vizinhanças, no ciclo real são adicionadas etapas para garantir o funcionamento correto. Por exemplo, é acrescentado superaquecimento entre os estágios 4 e 1 para garantir que apenas vapor entre no compressor, e um subresfriador entre os estágios 2 e 3 para garantir que apenas líquido entre no dispositivo de expansão.

Figura 1 - Ciclo real de refrigeração por compressão de vapor.



Legenda:

$h_{f1}$	entalpia no ponto de líquido saturado.	$p_{g2}$	pressão de descarga do compressor.
$h_{f2}$	entalpia do líquido sub-resfriado entrando no processo de expansão.	$t_c$	temperatura de condensação.
$h_{g1}$	entalpia do vapor superaquecido entrando no processo de compressão.	$t_o$	temperatura de evaporação.
$h_{g2}$	entalpia no ponto de vapor saturado.	$t_{f2}$	temperatura do líquido sub-resfriado prestes a expandir.
$p_{g1}$	pressão de sucção do compressor.	$t_{g1}$	temperatura do vapor superaquecido prestes a entrar no compressor.

Fonte: Adaptado de ASHRAE STANDARD (2005)

A partir do monitoramento das pressões e temperaturas em diversos pontos do circuito é possível obter as variáveis de interesse por meio de métodos normalizados. A equação 1 representa a relação entre a capacidade de refrigeração e essas variáveis. Cabe ressaltar que as entalpias podem ser associadas a valores de pressão e temperatura pelo gráfico de ciclo real relativo ao fluido refrigerante utilizado durante o ensaio.

$$\Phi_0 = q_{mf} \frac{V_{ga}}{V_{gl}} (h_{g1} - h_{f1}) \quad (1)$$

Onde:

$\Phi_0$	Capacidade de refrigeração do compressor	[W]
$q_{mf}$	Vazão mássica do fluido refrigerante	[kg/s]



$V_{ga}$	Volume específico do fluido refrigerante (gás) que entra no compressor (durante o ensaio)	$[m^3/kg]$
$V_{gl}$	Volume específico do refrigerante que entra no compressor (estabelecido anteriormente ao ensaio)	$[m^3/kg]$
$h_{gl}$	Entalpia específica do refrigerante que entra no compressor nas condições básicas específicas de ensaio	$[kJ/kg]$
$h_{fl}$	Entalpia específica do refrigerante líquido na pressão correspondente à descarga do compressor	$[kJ/kg]$

Pode-se dizer que a capacidade de refrigeração (dada em watt) é o produto da vazão mássica de fluido refrigerante que passa pelo compressor pela diferença entre as entalpias específicas do refrigerante na entrada do compressor e a entalpia específica do líquido saturado na temperatura correspondente à pressão de descarga na saída do compressor (ISO, 1989).

Além da capacidade, a potência elétrica e o coeficiente de performance também são parâmetros de desempenho energético. A potência elétrica consumida ( $P$ , dada em watt) é definida pela norma (ISO, 1989) como sendo a potência nos terminais do compressor. É a soma do consumo do compressor e de outros dispositivos auxiliares necessários para o funcionamento do mesmo. O coeficiente de performance (COP, adimensional) é definido pela razão de capacidade de refrigeração pela potência elétrica consumida (equação 2) (ISO, 1989).

$$COP = \frac{\Phi_0}{P} \quad (2)$$

Com objetivo de prover confiabilidade e padronização neste tipo de teste, algumas normas internacionais são consideradas. Dentre elas duas se destacam a ISO (1989) e a ASHRAE STANDARD (2005). Ao realizar os testes seguindo as recomendações dessas normas, os fabricantes obtêm a capacidade de refrigeração (CAP), potência elétrica consumida (CON), e o coeficiente de performance de compressores (COP), com resultados suficientemente confiáveis para avaliar se o

compressor está adequado para operar satisfatoriamente em condições padronizadas de operação (condições de *checkpoint*) ou em outras condições de sistema desejadas por clientes. Essas normas citam diversos métodos para medição de vazão mássica. Dois deles são mais frequentemente utilizados pela empresa parceira e referenciados pela norma ISO (1989):

**Calorímetro de fluido secundário (método A):** utiliza um ciclo de compressão de vapor para a medição da vazão mássica do fluido refrigerante por meio da troca de calor entre o fluido refrigerante e um fluido secundário dentro do evaporador.

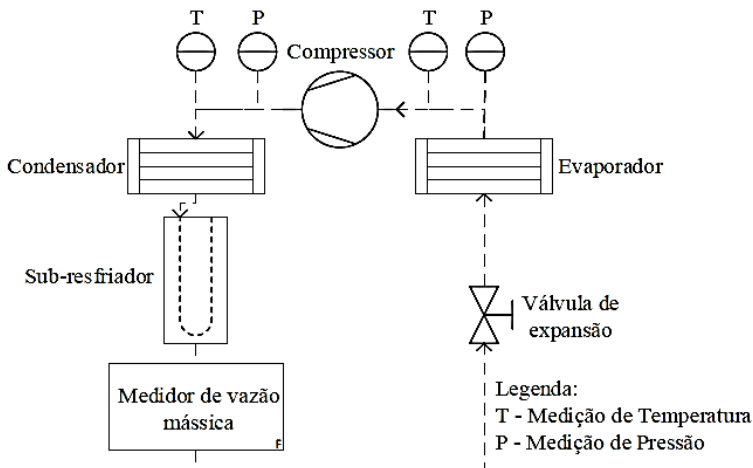
**Medidor de vazão no estado líquido (método F):** faz uso de um medidor de vazão para obter a vazão mássica do fluido refrigerante durante o estado líquido.

Segundo as normas, dois deles devem ser executados simultaneamente e a diferença entre resultados não deve ser superior a 4% para que o ensaio possa ser validado. Neste trabalho todos os ensaios utilizaram o método F para obtenção da variável de desempenho de interesse, a capacidade de refrigeração.

### **2.1.2. Medição de vazão no estado líquido**

Neste método o medidor deve ser posicionado entre o reservatório de fluido e a válvula de expansão (ISO, 1989). Para garantir o estado líquido do fluido ao passar pelo medidor, um subresfriador deve ser adicionado ao sistema conforme apresentado na figura 2. Com o valor obtido no medidor, a capacidade de refrigeração é obtida pela equação 1.

Figura 2 - Circuito para medidor de vazão no estado líquido (método F).



Fonte: Adaptado de ISO (1989).

Após a partida do compressor, a bancada de ensaios inicia o processo de controle para que as variáveis atinjam os valores da condição de operação na qual o compressor será testado. Enquanto esses valores não estabilizarem dentro dos limites da variação em torno do valor de referência, o ensaio estará no transitório. A norma ISO (1989) ainda aponta quais são as variáveis medidas e controladas:

- pressão de descarga;
- pressão de sucção;
- temperatura de sucção;
- tensão e corrente elétrica da alimentação do compressor;
- temperatura e velocidade do fluxo de ar em torno do compressor;
- temperatura interna ao trocador de calor (evaporador);
- temperatura externa ao trocador de calor (evaporador);
- vazão e temperatura da água de condensação;
- subresfriamento do fluido refrigerante na entrada do trocador;

Tal norma também menciona as variáveis que são apenas medidas:

- temperatura de descarga do compressor;
- temperatura do corpo do compressor;
- potência elétrica entregue ao evaporador;
- potência elétrica entregue ao compressor;
- resistência do enrolamento do compressor;
- medidas diversas, de acordo com o tipo de ensaio que está sendo utilizado e interesse do fabricante;

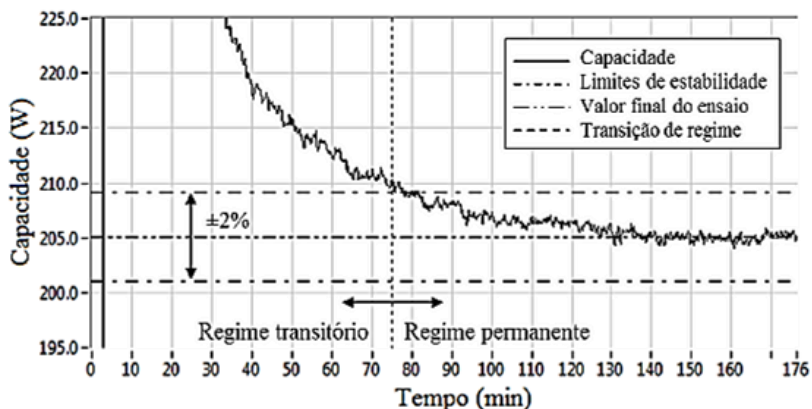
Todas essas diferentes variáveis influenciam diretamente o comportamento dinâmico da capacidade de refrigeração e o tempo necessário para alcançar o regime permanente. Além dessas variáveis medidas, outras também exercem influência nesse comportamento, como modelo do compressor, as condições iniciais da bancada (principalmente temperatura), posição das válvulas, influência do operador e diferentes *setpoints* (condições finais de ensaios).

Segundo NASCIMENTO (2015), a rede para inferência da entrada em regime permanente desenvolvida em trabalhos anteriores utiliza basicamente as medições de três variáveis obtidas no ensaio de desempenho energético: parâmetro de desempenho (capacidade de refrigeração, consumo ou COP), temperatura de corpo do compressor e pressão de sucção.

### **2.1.3. Regimes e aspectos normativos**

O ensaio padronizado tradicional engloba o regime transitório e o regime permanente de operação do compressor. O regime permanente é alcançado quando as flutuações das variáveis medidas estiverem dentro de limites estabelecidos e sem tendência de saírem desses limites (ISO, 1989). Essa progressão pode ser vista na figura 3, onde também é apresentado o limite de aceitação da capacidade de refrigeração.

Figura 3 - Transição do regime transitório para o permanente em ensaio da capacidade de refrigeração.



Fonte: Adaptado de PENZ (2011).

Para a norma ISO (1989), quando essas variáveis alcançam os limites estabelecidos, estas são medidas ao longo de uma hora com intervalos de tempo não maiores que 20 minutos, resultando em pelo menos 4 medições em regime permanente. Ao fim desse período, o valor final de cada variável é a média aritmética dos valores medidos. Já a norma ASHRAE STANDARD (2005) demanda pelo menos três medições em regime permanente. Os limites estabelecidos que devem ser observados para caracterizar o regime permanente são:

- leituras de pressão com variação menor que  $\pm 1\%$  do valor de referência;
- leituras de temperatura com variação menor que  $\pm 3$  K em relação ao valor de referência;
- frequência rotacional com variação menor que  $\pm 1\%$  rpm do valor de referência;
- tensão de alimentação do compressor é de  $\pm 3\%$  V;
- no método F, o valor lido no medidor de fluxo mássico não pode exceder o limite de  $\pm 1\%$  do valor médio da última hora.

Cabe ressaltar que para capacidade de refrigeração, potência elétrica consumida e temperatura de corpo do compressor não há especificação para seus limites de variação na norma. Entretanto,

especialistas da própria empresa que disponibilizou informação para realização deste trabalho definem limites aceitáveis como sendo:

- capacidade:  $\pm 2\%$  do valor final de ensaio;
- consumo:  $\pm 2\%$  do valor final de ensaio;
- temperatura de corpo:  $\pm 1\text{K}$  do valor final do ensaio.

Apesar dos limites dessas grandezas estarem especificados, não se sabe de antemão os valores finais das mesmas. Logo, não é possível identificar durante o ensaio o momento de entrada em regime permanente a partir de uma simples comparação de valores – não há conhecimento dos valores dos limites em termos absolutos. A solução baseada em redes neurais foi concebida para realizar essa identificação a partir da análise do comportamento de diferentes grandezas durante os ensaios.

## 2.2. REDUÇÃO DO TEMPO DE ENSAIO

Duas estratégias principais foram abordadas nos trabalhos anteriores para redução do tempo de ensaio de desempenho: melhoramentos na bancada e ferramentas de inteligência artificial.

Ambas abordagens foram realizadas simultaneamente e com bons resultados. Em seguida estão sumarizados os principais resultados provenientes dessas abordagens.

### 2.2.1. Melhoramentos do calorímetro

Neste cenário destacam-se o trabalho de SCUSSEL (2006), LIMA (2010a) e SILVEIRA (2010). O primeiro autor analisou diretamente os efeitos do aumento das perdas termodinâmicas, já que foi identificado que a temperatura de corpo é a variável que mais influencia a duração do regime transitório. SCUSSEL (2006) também identificou variações das condições elétricas, e chegou à conclusão que uma alimentação do motor com tensão contínua (15V) sobreposta à tensão alternada (115V), somada ao aumento da pressão de descarga (17 bar) e a redução da convecção forçada do compressores resultava na redução de tempo total do ensaio para duas horas e quinze minutos, frente ao total anterior de quatro horas e trinta minutos.

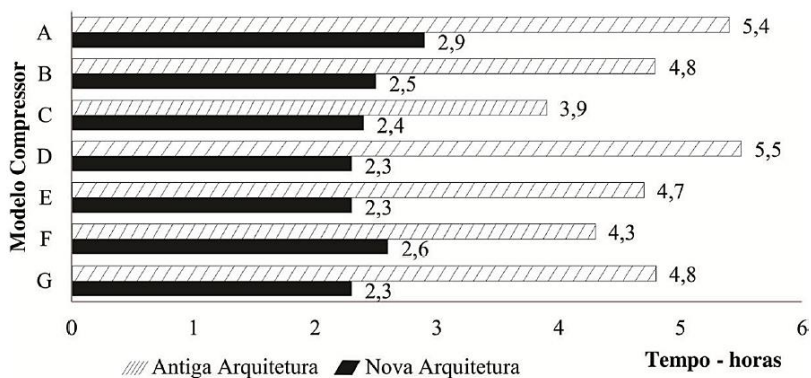
LIMA (2010a), entre outros estudos, propôs análises teóricas e experimentais da estabilização das variáveis do regime transitório utilizando como ferramenta a fonte de alimentação elétrica dos compressores. O autor instrumentou compressores de diferentes tamanhos com termopares em várias partes do compressor para analisar

o comportamento da temperatura principalmente durante o regime transitório. Essa informação foi útil para realizar um aquecimento forçado do compressor de forma a acelerar o transitório. Seus estudos reduziram o tempo de ensaio a uma hora e quarenta e cinco minutos, com a ressalva que todos os experimentos foram realizados em bancadas consideradas estado da arte no ano em questão.

Por fim, SILVEIRA (2010) buscou reduzir a duração do ensaio através da injeção de corrente contínua, método que já havia sido abordado por SCUSSEL (2006). Em seu trabalho, Silveira ressalta que o aquecimento do compressor se dá de dentro para fora, e a temperatura medida externamente é a variável que mais demora para atingir o valor de equilíbrio. Com a metodologia proposta, o autor fez com que o compressor estabilizasse ( $\pm 1$  °C) em menos de trinta minutos, resultando em ensaios com duração total de 1 hora e trinta minutos em média.

Os três autores mencionados anteriormente fizeram uso da aceleração do transitório para reduzir a duração de ensaios. Outros autores (BARBOSA, 2006; FLESCHE, 2006; FLESCHE; J. E. NORMEY-RICO, 2010; PETROVIC, 2007; POLETO, 2006) reduziram o tempo de ensaio por meio da melhoria dos controles e nova instrumentação. A figura 4 apresenta a redução obtida pelo trabalho conjunto desses. Essa redução foi alcançada para diferentes modelos de compressores.

Figura 4 - Resultados obtidos na redução do tempo de ensaio derivados da melhoria no controle e nova instrumentação (sem aceleração do transitório).



Fonte: PENZ (2011).

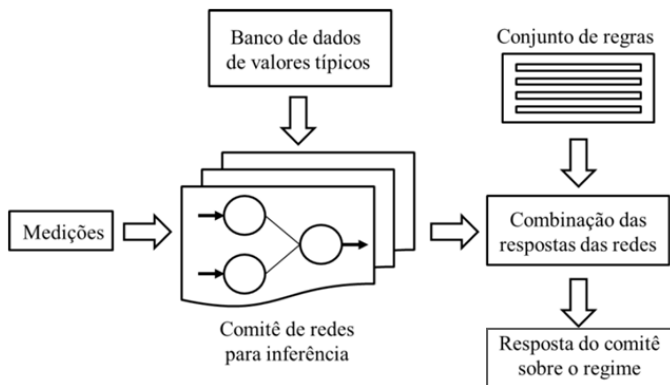
### 2.2.2. Soluções utilizando inteligência artificial

Os avanços obtidos por essa abordagem já foram sintetizados na introdução e terão seus resultados explicados de forma mais detalhada adiante. No entanto, as disposições técnicas sobre as redes de inferência de regime permanente serão apresentadas somente após a introdução dos conceitos de inteligência artificial, que ocorre na seção 3.

A ferramenta de inferência de regime permanente utilizada neste trabalho é resultado do estudo de diversos autores, cada um melhorando o trabalho anterior em algum aspecto e que tem seu estado atual apresentado por NASCIMENTO (2015). A ferramenta é parte de um módulo que também inclui uma ferramenta de prognóstico de desempenho, isto é, uma solução que estima o valor final do ensaio com medições obtidas no momento que a ferramenta de inferência identifica a entrada em regime permanente.

A figura 5 apresenta o esquema geral da ferramenta de inferência de regime permanente. A partir de medições de diferentes variáveis e da utilização de valores típicos obtidos de banco de dados ou de catálogos, treina-se um conjunto de redes (comitê) diferentes entre si. As respostas das redes são combinadas e submetidas a um conjunto de regras que tem como finalidade garantir a confiabilidade metrológica do resultado.

Figura 5 - Esquema geral de ferramenta de inferência de regime permanente.



Fonte: adaptado de PENZ (2011).

O estado da arte da ferramenta tem seus resultados expressos no trabalho de NASCIMENTO (2015), que estendeu e aprimorou os



resultados obtidos por PENZ (2011) e LIMA (2010b). Sua análise foi realizada para as três variáveis de desempenho (consumo, CAP e COP), mas os resultados resumidos nesta seção são apenas para capacidade de refrigeração, a variável analisada neste trabalho.

As entradas da rede advêm das medições de temperatura de corpo, pressão de sucção e capacidade de refrigeração. O banco de dados de valores típicos é a maneira que os autores citados anteriormente definiram para normalizar as variáveis de entrada para tornar a rede insensível ao tamanho do compressor.

Quanto à saída da rede, PENZ (2011) propôs uma estratégia de integração das respostas que foi melhorada posteriormente por NASCIMENTO (2015), e é a que está em uso atualmente. Essas estratégias consistem em utilizar um conjunto de 120 redes, que conforme vão ocorrendo as inferências há um tempo de espera para garantir que a resposta esteja na faixa de  $\pm 2\%$ . A tabela 1 apresenta as regras utilizadas, o número de redes para cada regra e o respectivo tempo de espera.

Tabela 1 - Conjunto de regras para o comitê de inferência da entrada em regime permanente da capacidade de refrigeração.

<b>Regra</b>	<b>Descrição</b>	<b>Tempo de espera (min)</b>
0*	Menos de 40 redes	-
1	Entre 40 e 79	20
2	Entre 80 e 104	15
3	Entre 105 e 114	10
4	Entre 115 e 119	5
5	120 redes	0

\* Se a quantidade de redes que indicam o regime for inferior a 40 a inferência não é confirmada.

Fonte: NASCIMENTO (2015).

Essa estratégia de uso de comitês de redes e de regras resulta em economia de tempo, disposta respectivamente para cada regra na tabela 2. O tempo médio de entrada real é a média dos valores dos ensaios nos quais as redes foram testadas. O tempo médio de inferência é o tempo que as redes pertencentes à respectiva regra levaram para inferir o regime permanente. Nessa questão é possível ressaltar uma característica

desejada para a ferramenta de inferência, que é a inferência após a entrada real. Em outras palavras, é pior que a ferramenta indique que o ensaio entrou em regime quando não entrou, o chamado erro de falso positivo. Por fim, o tempo médio dos ensaios é a duração total daqueles ensaios do conjunto de teste. Na última coluna dessa tabela está o ganho de tempo para cada regra.

Tabela 2 - Tempos médios de inferência para cada regra.

<b>Regra</b>	<b>Tempo médio de entrada em regime real (min)</b>	<b>Tempo médio de inferência (min)</b>	<b>Tempo médio total de ensaio (min)</b>	<b>Ganho de tempo (%)</b>
<b>1</b>	47,3	73,9	141,5	47,7
<b>2</b>	42,0	67,0	125,3	46,6
<b>3</b>	39,8	63,2	125,8	49,8
<b>4</b>	40,3	61,4	128,0	52,0
<b>5</b>	39,1	57,5	124,5	53,8
<b>Média</b>	41,8	64,0	129,6	50,1

Fonte: NASCIMENTO (2015)

Quanto ao desempenho geral, 99,3% de um total de 842 ensaios de teste tiveram inferência do regime. No entanto, NASCIMENTO (2015) aponta que essa inferência é realizada em média 23 minutos após a entrada de regime real, evidenciando que ainda há margem para redução desse tempo.

A correta inferência do regime permanente também é importante para que o prognóstico da capacidade final esteja dentro dos limites de  $\pm 2\%$  definidos pela empresa parceira. NASCIMENTO (2015) obteve como resultado que 98% dos ensaios que tiveram inferência do regime resultaram em prognósticos dentro do limite definido (relativo à capacidade de refrigeração). Outro aspecto mencionado por esse autor é a taxa de inferências com erros de falso positivo (4,31% dos ensaios). Esse panorama abre possibilidade ao estudo da qualidade não só das respostas das regras, mas também do desempenho das redes utilizadas como classificadores, que por sua vez acaba melhorando diretamente o desempenho da ferramenta de prognóstico da capacidade final.

### 3. INTELIGÊNCIA ARTIFICIAL

Apesar de não existir um conceito de inteligência artificial unificado, RUSSEL & NORVIG (2003) a define como um conjunto de ações racionais, sendo idealmente um agente que toma a melhor decisão de acordo com a situação a qual foi apresentado. Ao modelar esses agentes levando em consideração a inteligência natural e biológica, foram criados os chamados sistemas inteligentes. Fazem parte desses algoritmos inteligentes as redes neurais artificiais, computação evolucionária, inteligência de enxame, sistemas *Fuzzy*, entre outros. Dessa maneira, a inteligência artificial pode ser vista de forma multidisciplinar, envolvendo disciplinas exatas e humanas, ao mesclar a lógica, dedução racional, sistemas especializados e aprendizado simbólico de máquina em um único campo de pesquisa (ENGELBRECHT, 2007).

O conceito de aprendizado de máquina se confunde com o de sistemas inteligentes, sendo o campo da inteligência artificial de maior interesse neste trabalho. É definido como método de construção de programas que melhoram sua capacidade de resolver problemas automaticamente (MICHALSKI; CARBONELL; MITCHELL, 1983). Duas das maneiras de aplicação do aprendizado de máquina são a regressão e a identificação de padrões. Regressão busca aproximar entradas de respostas conhecidas, tendo valores contínuos como saída. A identificação de padrões segrega exemplos por classes (classificação) ou agrupa por afinidades de atributos (clusterização) (TAN; STEINBACH; KUMAR, 2009). De forma sucinta, as diferenças entre classificação e clusterização são as seguintes:

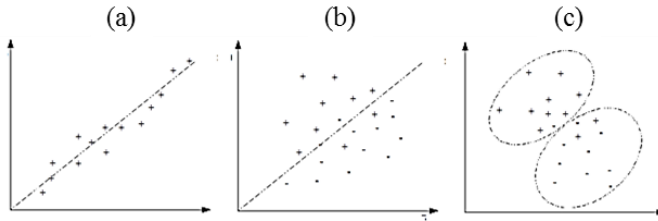
**Classificação:** aprendizado supervisionado, isto é, rotula exemplos a partir de um conjunto de dados com saídas conhecidas. Exemplos de classificadores incluem máquinas de vetor de suporte, árvores de decisão, redes neurais, classificadores baseados em regras e Bayesianos simples. Existe uma etapa de treinamento e teste após a escolha do modelo.

**Clusterização:** realizam a tarefa agrupando um conjunto de entrada por semelhança de atributos e não por rótulos, o que constitui um aprendizado não-supervisionado. São exemplos de algoritmos agrupadores redes auto organizáveis de Kohonen, *K-means* e agrupamento hierárquico.

A diferença entre as abordagens é visualizada na figura 6. Nela, o item (a) apresenta uma reta para regressão linear; o item (b) ilustra como

um classificador binário separa dois conjuntos diferentes; e o item (c) demonstra como a clusterização agrupa dois conjuntos diferentes.

Figura 6 - Diferentes modelos de aprendizado de máquina. (a) regressão linear (b) classificador (c) clusterizador.



Fonte: autor.

Aprendizado de máquina para identificação de padrões via classificação ou agrupamento é uma tendência para as mais diversas áreas e com diferentes níveis de aplicação, como previsão de trocas comerciais por máquinas de vetor de suporte (GERLEIN et al., 2016), previsão de danos em construções com uso de árvores de decisão (TIXIER et al., 2016), classificação de carvão por imagens microscópicas utilizando diversos algoritmos de classificação, como MLP, funções de bases radiais, redes probabilísticas, entre outros (ALPANA; MOHAPATRA, 2016), inferência de performance de sistemas de refrigeração por árvores de decisão, máquinas de vetor de suporte e redes neurais artificiais (CHOU; HSU; LIN, 2014), e detecção de falhar em turbinas a gás por reconhecimento de padrões via clusterização (FONTES; PEREIRA, 2016).

### 3.1. REDES NEURAIS ARTIFICIAIS

Uma rede neural é um processador paralelamente distribuído constituído de unidades de processamentos elementares tal como o cérebro humano, onde cada unidade de processamento é modelada como um neurônio. Assim como no cérebro humano, o conhecimento é adquirido por aprendizagem. O uso de redes neurais oferece as seguintes propriedades e capacidades úteis (HAYKIN, 2001):

- a) aprendizagem por meio de exemplos;
- b) soluções de problemas não-lineares;

- c) adaptabilidade (retreinamento para adequação a novos conjuntos de entrada)
- d) bom desempenho em problemas mal definidos (como em casos onde não existe conhecimento de modelos matemáticos para o problema);
- e) tolerância às falhas e imunidade ao ruído (capacidade de generalização);

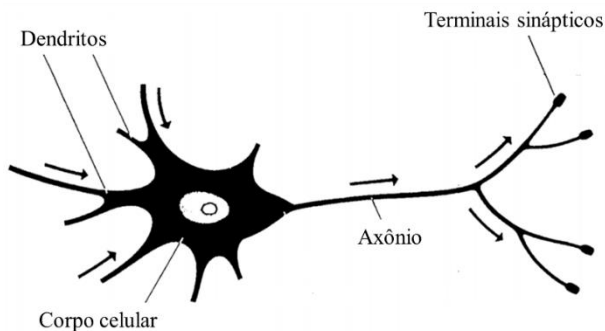
Redes neurais são consideradas modelos complexos com boa flexibilidade mas que exigem muito recursos de sistema, além de possuírem tendência ao excesso de treinamento (*overfitting*). São fatores de complexidade o tempo e características do treinamento, exigência de memória e parâmetros de execução. (KUNCHEVA; WHITAKER, 2003).

Neste trabalho um modelo de rede neural é abordado em detalhes, a rede *perceptron* de múltiplas camadas (ou MLP, do inglês, *Multilayer Perceptron*). Essa é a rede mais popular e chega a ser confundida com o próprio conceito de rede neural. É um modelo com aprendizado supervisionado que desempenha tarefas de regressão ou de classificação (HAYKIN, 2001).

### 3.1.1. Modelo de neurônio *perceptron* de múltiplas camadas

Redes neurais MLP utilizam neurônios artificiais como elemento básico de processamento. Ele busca mimetizar o sistema neural que possui as células nervosas (neurônios biológicos) como componente elementar. A figura 7 apresenta um neurônio biológico e suas partes. Neurônios são conectados entre si pelo axônio de um e dendrito de outro, e essa conexão é denominada sinapse.

Figura 7 - Neurônio biológico e suas partes.

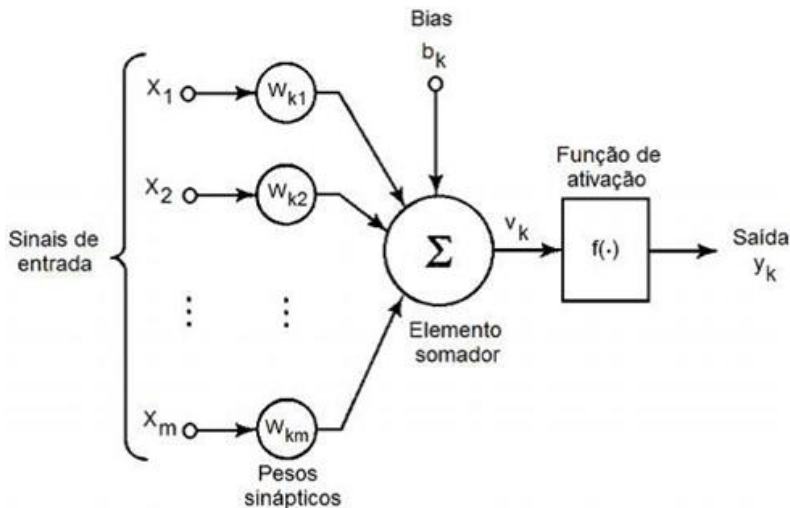


Fonte: adaptado de ARBIB (1989).

O início da busca por um equivalente computacional ao neurônio deu-se nos anos 40, com o trabalho de MCCULLOCH & PITTZ (1943). Mais de uma década depois, ROSENBLATT (1958) propôs o primeiro método de aprendizagem supervisionada para as redes neurais artificiais ao qual deu o nome de *perceptron*. Contudo, um único *perceptron* é capaz de resolver apenas problemas linearmente separáveis (como as funções booleanas ‘E’ e ‘OU’) (ENGELBRECHT, 2007).

O *perceptron* é constituído de três elementos básicos, vistos na figura 8: pesos sinápticos, junção aditiva e função de ativação. Cada uma dessas partes tenta imitar o comportamento do neurônio biológico, sendo os sinais de entrada os dendritos, a junção aditiva o corpo celular e a função de ativação o axônio. Além desses três elementos existe um parâmetro externo, o *bias*, entre a junção aditiva e a função de ativação. Ele é responsável por aumentar ou diminuir, de acordo com seu valor, a entrada líquida da função de ativação. Os pesos sinápticos são caracterizados por um valor próprio, e nestes é armazenado o conhecimento do neurônio. A junção aditiva é um somador de sinais de entrada ponderado pelos pesos sinápticos. Por último, a função de ativação restringe a amplitude de saída do neurônio a um valor finito (HAYKIN, 2001).

Figura 8 - Modelo de neurônio (*perceptron*) de uma rede neural MLP.



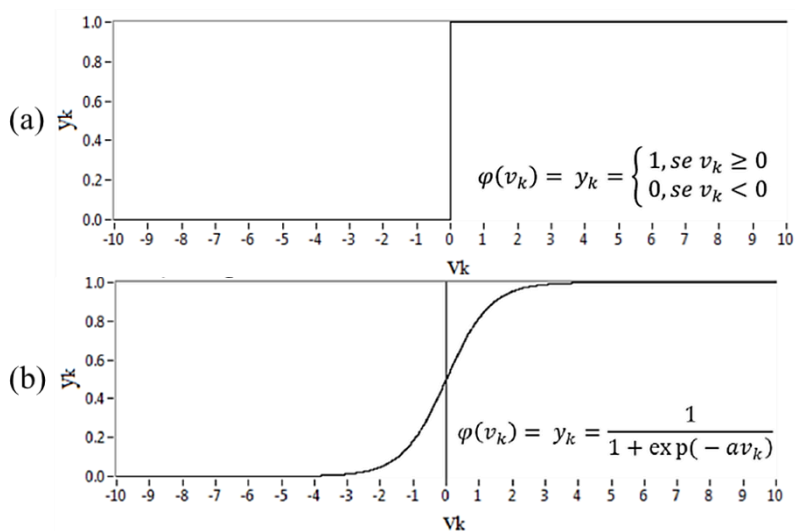
O neurônio pode ser representado matematicamente pelas equações 3 e 4. A equação 3 é o somatório dos produtos entre os pesos sinápticos e suas respectivas entradas. Essa saída,  $u_k$ , é somada ao *bias* quando este existir e inserida na função de ativação (equação 4).

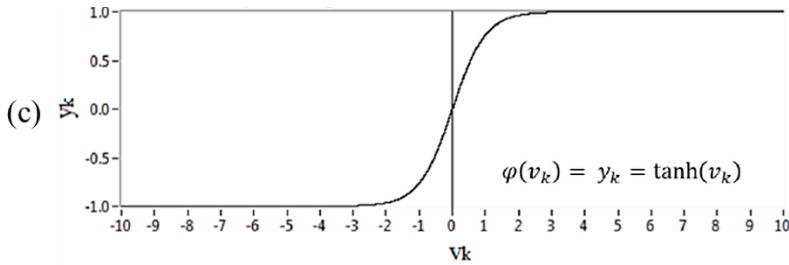
$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (3)$$

$$y_k = f(u_k + b_k) \quad (4)$$

Existem diferentes funções de ativação, escolhidas de forma a se adequarem melhor a cada problema. A mais simples delas é a função limiar (MCCULLOCH; PITTZ, 1943). Outras mais complexas (não-lineares e sigmoidais) têm a característica útil de serem diferenciáveis, o que significa que são contínuas em todos os pontos. A figura 9 ilustra graficamente essas funções em conjunto com suas equações. Nessa figura nota-se que a função limiar apresenta apenas duas saídas, 0 ou 1. Já a função logística possui soluções entre valores de 0 a 1, e a tangente hiperbólica entre -1 a 1.

Figura 9 - Funções de ativação limiar (a), logística (b) e tangente hiperbólica (c).





Fonte: adaptado de HAYKIN (2001).

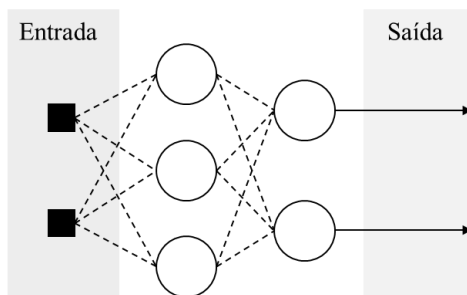
### 3.1.2. Arquitetura de redes neurais

Uma rede neural MLP possui diferentes camadas: a camada de entrada, também chamada de nós de fonte, por onde são inseridos os exemplos; uma ou mais camadas intermediárias, onde encontram-se os pesos nos quais está armazenado o conhecimento da rede; e a de saída. A partir desse conjunto básico o operador pode alterar a quantidade de parâmetros de entrada, o número de camadas e neurônios na ou nas camadas intermediárias e o número de saídas. Ainda é possível modificar o modo como as camadas comunicam-se entre si, isto é, tipo e grau de conexão entre neurônios de acordo com o problema a ser resolvido. É possível identificar três arquiteturas básicas de redes MLP com diferentes características (HAYKIN, 2001):

- **Rede com camada única:** não existe camada intermediária. É dita camada única pois os nós de fonte não são contabilizados. Redes deste tipo são capazes de resolver apenas problemas lineares.
- **Rede alimentada diretamente de múltiplas camadas:** são capazes de extrair estatísticas de ordem elevada pois possuem uma ou mais camadas intermediárias (ditas ocultas). A figura 10 apresenta um exemplo dessa arquitetura.
- **Redes recorrentes:** possuem laços de realimentação que podem estar inseridos nas duas redes mencionadas anteriormente. Esses laços inserem atrasos de tempo e são capazes de fazer processamento dinâmico.



Figura 10 - Exemplo de rede alimentada adiante com múltiplas camadas.



Fonte: adaptado de HAYKIN (2001).

O aumento do número de parâmetros de entrada, número de camadas, neurônios e inserção de novas conexões permite a resolução de problemas mais complexos, mas também implicam aumento do consumo computacional (HAYKIN, 2001).

### 3.1.3. Treinamento de redes neurais

O treinamento de uma rede neural MLP é realizado de forma supervisionada por um algoritmo chamado retropropagação do erro (EBP, *Error Backpropagation*). Antes de iniciar o treinamento, os pesos sinápticos são inicializados. O algoritmo de treinamento consiste então de duas fases (HAYKIN, 2001).

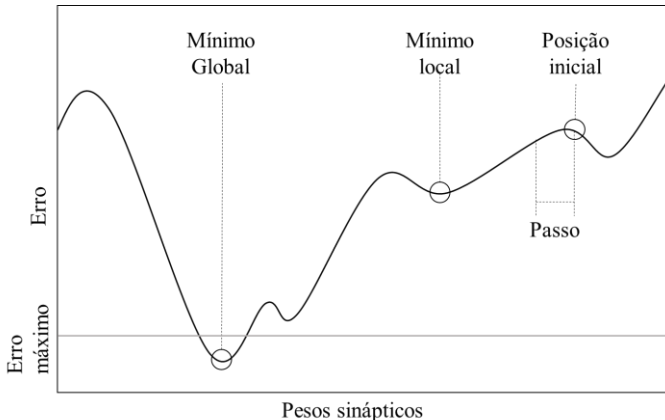
**Propagação:** o vetor de entrada é apresentado aos nós de entrada. O efeito dos valores de entrada é propagado por toda rede, camada a camada, ocasionando um vetor de saída. Durante esta etapa todos os pesos sinápticos são fixos.

**Retropropagação:** os pesos sinápticos são ajustados de acordo com a regra do algoritmo utilizado. Especificamente, a saída da rede é subtraída do valor desejado e um sinal de erro é criado. Esse sinal é retropropagado pela rede na direção dos pesos sinápticos. Esses pesos são atualizados para que a resposta seja mais próxima do valor desejado.

Durante a retropropagação ocorre a busca pelos valores mínimos de erros em um espaço de erros multidimensional. Nesse espaço existem diversas soluções, mas apenas uma é ideal. Espera-se que o algoritmo

ignore mínimos locais e encontre o mínimo global. Esse processo é dificultado por algumas variáveis, como valores da inicialização aleatória dos pesos e o próprio algoritmo de treinamento. A figura 11 ilustra esse espaço de erros.

Figura 11 - Função erro e representação dos mínimos local, global e tamanho do passo via gradiente descendente.



Fonte: autor.

De forma geral o EBP é estável (lida bem com variações nos valores de entrada) e de fácil implementação, mas peca em sua capacidade limitada e convergência lenta. Esses problemas são causados pelo modo como algoritmos simples varrem o espaço de erros (em linha), característica que depende fortemente da correta escolha do passo ( $\alpha$ , ou taxa de aprendizado). Métodos de otimização como gradiente descendente, método de Newton e Gauss Newton visam melhorar esse procedimento (NOCEDAL; WRIGHT, 2006).

Quanto à apresentação do conjunto treinamento à rede, existem duas maneiras principais: em épocas (*offline*) ou incremental (estocástico ou *online*). Quando em épocas, ajusta os pesos somente após a apresentação de todos os exemplos. Logo, cada padrão é avaliado com a mesma configuração de pesos. Tende a ser mais rápido que incremental por necessitar de menos iterações mas necessita de mais memória. Já no método incremental os pesos são atualizados a cada exemplo usando o gradiente do erro de um único padrão. Este último tem chances menores de permanecer em mínimos locais (ENGELBRECHT, 2007).

### 3.1.3.1. Conjunto de entrada

É na relação existente entre os dados de entrada e a saída que está armazenado o conhecimento da rede. Sobre esse conhecimento, ANDERSON e ROSENFELD (1988) apontam quatro regras:

- 1) entradas similares de classes similares normalmente devem produzir representações similares no interior da rede, e portanto devem ser classificados como pertencentes à mesma categoria.
- 2) devem ser atribuídas representações bem diferentes na rede a itens que devem ser categorizados como classes separadas.
- 3) se uma característica particular é importante, então deve haver um grande número de neurônios envolvidos na representação daquele item na rede.
- 4) informação prévia e invariâncias devem ser incorporadas no projeto de uma rede neural, simplificando com isso o projeto da rede por não ter que aprendê-las.

O item 4 é particularmente importante pois ressalta a possibilidade da criação de redes especializadas. Isso se deve ao fato que redes especializadas têm um menor número de parâmetros livres disponíveis para ajuste. Dessa forma requerem um conjunto de entrada menor, aprendem mais rápido e ainda generalizam melhor (RUSSO, 1991).

Tão importante quanto identificar quais são os parâmetros de entrada significativos para o aprendizado é garantir que estes possuam características que assegurem o bom treinamento. A disciplina que trata desse processo é a mineração ou análise exploratória de dados. Inicialmente o conjunto de dados cru pode ser ruidoso, incompleto, inconsistente e não adequado à técnica de aprendizado de máquina desejada. Para aproveitar o conjunto de dados da melhor forma possível são realizados alguns passos pré-processamento (HAN; KAMBER, 2006). As principais tarefas de pré-processamento são:

- limpeza: remoção de ruídos e dados inconsistentes;
- integração: quais conjuntos de dados podem ser agregados para aumentar o valor total;
- seleção: parcela dos dados que pode ser efetivamente utilizado;

- transformação: a partir dos dados selecionados, convertê-los para um formato útil à técnica de aprendizado de máquina em questão. São exemplos de transformações normalização, padronização, e equalização.

GUYON e ELISSEEFF (2003) apontam que a dimensionalidade dos parâmetros de entrada é um problema, e o excesso de variáveis de entrada tende a causar *overfitting*. Esses mesmos autores também mencionam métodos que realizam essa escolha de forma automática, e que nem sempre parâmetros correlacionados são redundantes durante o aprendizado, sendo úteis quando submetidos à transformação matemática.

### 3.1.3.2. Generalização e critérios de parada

Redes neurais possuem uma característica chamada generalização, que se refere à capacidade da rede de estender seus resultados para entradas que não estavam presentes durante o treinamento. Dois termos fazem parte desse conceito: *bias*, a relação entre a função aproximativa e a função de regressão, que representa a incapacidade da rede neural de aproximar com precisão a função de regressão, sendo assim um **erro aproximativo**. Não confundir este *bias* com o termo *bias* presente no *perceptron*; e variância, relacionada ao comportamento da função aproximativa com a amostra de treinamento, representando a inadequação da informação contida na amostra de treinamento acerca da função de regressão. A variância também é chamada de **erro estimativo** (HAYKIN, 2001).

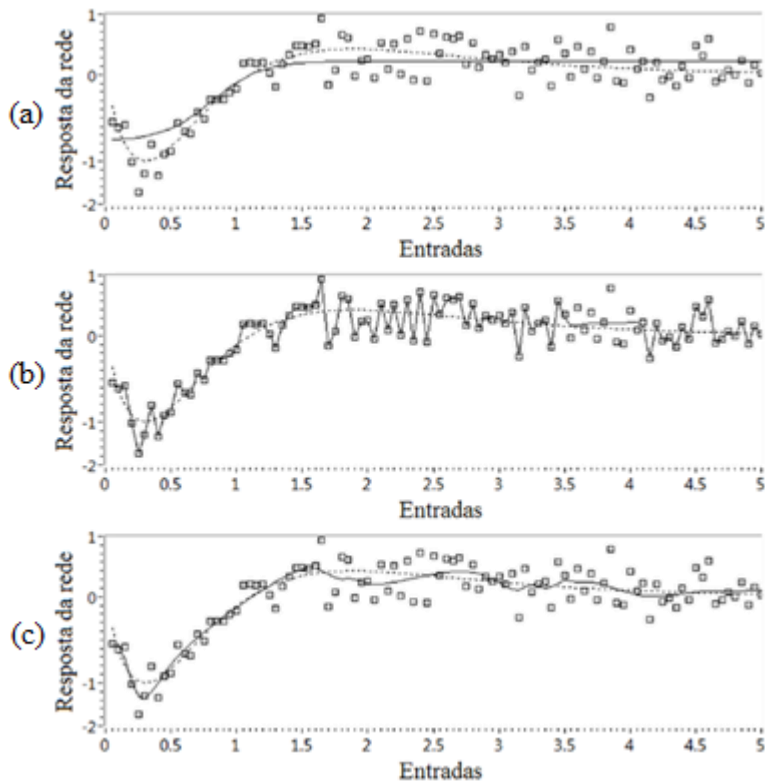
Para que uma rede neural apresente um bom desempenho, é necessário que estas duas parcelas de erros sejam pequenas. Infelizmente, constata-se que para um conjunto de treinamento de tamanho fixo, o preço para um erro aproximativo pequeno (*bias*) é um erro estimativo grande (variância). Por isso o problema de generalização também é conhecido como dilema *bias*-variância ou dilema estabilidade-plasticidade (GEMAN; BIENENSTOCK; DOURSAT, 1992).

No primeiro dos gráficos da figura 12 o modelo apresenta um grande *bias* (também chamado de *underfitting* ou falta de ajuste). No segundo gráfico dessa mesma figura, ocorre grande variância (*overfitting*), onde até os ruídos foram modelados. Por fim, no terceiro gráfico o modelo possui boa generalização.

Para contornar ou amenizar o dilema de *bias*-variância técnicas como a validação cruzada são empregadas. Nesta técnica em específico

são criados dois conjuntos para serem utilizados durante o treinamento: um conjunto de treinamento e um conjunto de validação (menor que o de treinamento).

Figura 12 - Dilema *bias* variância. (a) Falta de treinamento. (b) Excesso de ajuste. (c) Boa generalização.

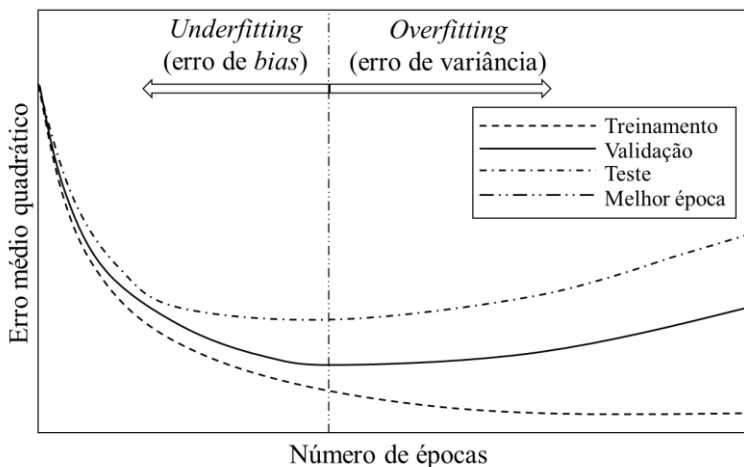


Fonte: Adaptado de GEMAN; BIENENSTOCK & DOUSAT (1992).

É possível notar na figura 13 o erro no conjunto de treinamento caindo conforme os exemplos são apresentados. Contudo é difícil identificar o melhor momento de parada levando em consideração a boa generalização, isto é, parar o treinamento antes que a rede esteja sofrendo excesso de ajuste ou com erro de variância muito alto. Por outro lado, caso o treinamento seja finalizado antes do momento ideal, a rede pode sofrer subtreinamento e o erro de *bias* ser grande. A validação cruzada é

um método para identificar esse momento. Uma parcela do conjunto de treinamento é dedicada à validação, e a cada período o treinamento é interrompido e a rede é testada nesse conjunto. Caso o erro no conjunto de validação ultrapasse um número de épocas definido pelo operador o treinamento é interrompido.

Figura 13 - Aspectos gráficos da validação cruzada e do conceito de generalização.



Fonte: Adaptado de NASCIMENTO (2015).

### 3.1.4. Regras para criação de comitês

Basicamente existem duas possibilidades de garantir que as respostas das redes sejam diferentes (BROWN, 2004):

**Manipulação dos dados de treinamento (amostragem):** uma das maneiras de criação de diferentes conjuntos de treinamento é chamada de *Bootstrap*, onde subconjuntos de treinamento são gerados com reposição a partir de um conjunto original. Um classificador é treinado a partir de cada um desses subconjuntos, gerando vários classificadores (PAPADOPOULOS; EDWARDS; MURRAY, 2001).

**Manipulação da arquitetura das redes:** alterar o número e camadas de neurônios, número de variáveis de entrada ou parâmetros de entrada da rede.

Redes MLP já foram estudadas quanto a essa questão e possuem alta correlação quando o modo de treinamento é por amostragem aleatória. Por isso são necessários métodos de criação de comitês que excluam redes muito correlacionadas, tanto no erro quanto no acerto, para evitar resultados tendenciosos (ULAŞ et al., 2009). A inserção incremental proposta por GOEBEL e YAN (2004) vem sendo utilizada com sucesso na fusão e construção de comitês (KANNATEY-ASIBU; YUM; KIM, 2016; TRABELSI; ELOUEDI; LEFEVRE, 2016), equação 5, assegurando a diversidade do comitê:

$$\rho_n = \frac{nN^f}{N - N^f - N^t + nN^f} \quad (5)$$

onde  $\rho$  é a correlação entre o grupo de redes presente no comitê e a rede candidata,  $n$  o número de classificadores no grupo (incluindo o classificador cuja correlação está sendo analisada),  $N^f$  é o número de índices onde todos classificadores erraram,  $N^t$  o número de índices onde todos classificadores acertaram e  $N$  o número total de índices. A correlação de Goebel é aplicada à relação entre os vetores de rótulo e saída discretizados em sua totalidade. Os passos para criação do comitê foram os seguintes:

1. Definição de conjunto de dados para treinamento das redes neurais
2. Definição de conjunto de dados para criação do comitê
3. Treinamento de um conjunto de  $X$  redes neurais candidatas
4. Aplicação das redes aos ensaios do conjunto de criação do comitê
5. Discretização das saídas contínuas por meio da aplicação do limitar
6. Identificação da rede com maior exatidão
7. Aplicação da correlação de Goebel
8. Criação do comitê definido pelas métricas exatidão simples e correlação de Goebel

Uma vez obtidas as redes menos correlacionadas por meio da aplicação da métrica arbitrária, os resultados podem ser aplicados ao conjunto de teste. Como o conjunto de teste possui mais de um ensaio é necessário obter uma rede resultante deste comitê. Isso foi realizado por votação majoritária. O procedimento ficará mais claro durante a apresentação do método com valores reais, no item 5.3.1.





## 4. MÉTODOS ESTATÍSTICOS E DE COMPARAÇÃO DE CLASSIFICADORES

Nesta seção são abordados métodos estatísticos e de comparação de classificadores. Análises estatísticas são fundamentais para interpretação correta de dados, garantindo que as relações encontradas sejam significativas e não simplesmente causadas aleatoriamente.

Com o objetivo de tornar o texto mais conciso e claro, como as distribuições de dados presentes neste trabalho foram avaliadas predominantemente não-normais via testes de normalidade, os métodos estatísticos utilizados e citados no decorrer desta seção visaram apenas essas distribuições. Essas ferramentas foram utilizadas na comparação de comportamento de variáveis como duração de ensaios, variação de pressão de sucção e também na construção de intervalos de confiança durante a comparação de classificadores. Quanto à seção de comparação de classificadores, em síntese, essa apresenta o ROC (*Receiver Operating Characteristics*), um método gráfico de análise de classificadores, e as maneiras de criar intervalos de confiança durante sua construção.

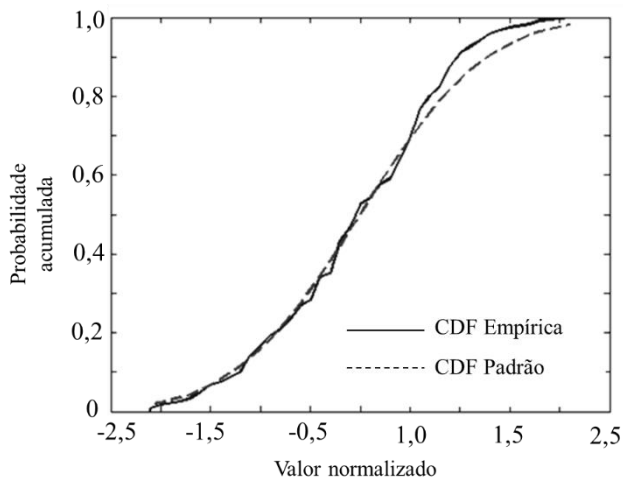
### 4.1. MÉTODOS ESTATÍSTICOS

A estatística apresenta métodos úteis para avaliar e comparar atributos com determinado grau de confiança ou nível de significância. Três ferramentas destacam-se nesse contexto: testes de hipótese (inferência estatística para duas ou mais amostras) e construção do intervalo de confiança (IC).

Esses métodos possuem diferenças quando aplicadas a tipos diferentes de distribuições. Por isso é necessário realizar inicialmente um teste de normalidade para a amostra em estudo. Um dos testes possíveis é o de Kolmogorov-Smirnov, um teste não paramétrico onde testa-se a hipótese nula de que a função distribuição acumulada em questão é igual a outra conhecida com um certo nível de significância (MASSEY, 1951). A figura 14 exemplifica uma comparação entre a curva acumulada hipotética e aquela que está sendo testada.

É possível antecipar que as distribuições presentes neste trabalho não foram consideradas normais após executados os testes de normalidade. Dessa forma, os itens posteriores, sobre análise de variância e testes de hipótese, discorrerão sobre essas aplicações frente às distribuições não normais.

Figura 14 - Visualização de um teste de Kolmogorov-Smirnov onde as curvas pertencem à mesma distribuição.



Fonte: adaptado de MATHWORKS (2016).

#### 4.1.1. Teste F para curvas não paramétricas

Enquanto que para distribuições normais aplica-se ANOVA (teste F), para distribuições não paramétricas a alternativa é a aplicação da análise de Kruskal-Wallis (KRUSKAL, WALLIS, 1952). Este método analisa a variância e compara as medianas dos grupos de dados utilizando chi-quadrado ao invés da estatística-F do ANOVA tradicional. É uma maneira de testar a hipótese nula de que ‘n’ tratamentos são idênticos contra a hipótese alternativa de que alguns tratamentos geram observações maiores que outros. Assim como o ANOVA essa análise não aponta quão grande são essas diferenças, por isso em sequência ao Kruskal-Wallis são realizadas multicomparações para verificar onde ocorrem e quão proeminentes essas diferenças são (MONTGOMERY; 2000). O teste estatístico é definido pela equação 6.

$$H = \frac{1}{S^2} \left[ \sum_{i=1}^a \frac{R_i^2}{n_i} - \frac{N(N+1)^2}{4} \right] \quad (6)$$

onde  $N$  é o número total de observações,  $n_i$  é o número de observações no  $i$ -ésimo tratamento e  $R$  é a soma de *ranks* do  $i$ -ésimo tratamento.  $S^2$  é a variância dos *ranks* e é dada pela equação 7.

$$S^2 = \frac{1}{N-1} \left[ \sum_{i=1}^a \sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4} \right] \quad (7)$$

A rejeição da hipótese nula ocorre quando  $H > \chi_{\alpha, a-1}^2$ . Alternativamente, a abordagem por valor  $p$  também pode ser utilizada, onde esse valor é o menor nível de significância que leva a rejeição da hipótese nula. Em outras palavras, quanto menor o valor  $p$  maiores as probabilidades de rejeição da hipótese nula (MONTGOMERY; RUNGER, 2013).

#### 4.1.2. Testes de hipótese de Dunn

Após a identificação da existência de diferença entre tratamentos é comum a realização de testes de hipóteses pareados para quantificar essa diferença e ordenar os grupos. Uma alternativa é a aplicação de testes tradicionais (chamados t-testes) par a par, desde que as distribuições sejam normais e as amostras independentes.

Todavia, a literatura aponta que se diversos t-testes par a par forem realizados seguidamente em conjuntos com mais de 2 tratamentos ocorre acumulação de erros de tipo 1 (rejeição da hipótese nula quando esta é verdadeira). Por isso são necessários métodos de múltipla comparação, como o método de Dunn (HOCHBERG; TAHMANE, 1987).

O teste de Dunn é um procedimento de múltipla comparação par a par não paramétrico, comumente utilizado após a rejeição da hipótese nula em análises como a de Kruskal-Wallis (DUNN, 1961). A equação 8 apresenta o teste estatístico deste método.

$$|t| = \frac{|\bar{y}_i - \bar{y}_j|}{\sqrt{MSE \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}} \quad (8)$$

onde os valores de  $y$  são as médias dos tratamentos (*ranks* neste caso),  $n$  o número de amostra e MSE a abreviatura para erro médio quadrático.

Rejeita-se a hipótese nula se  $|t| > t_{1 - \eta/2; \nu}$ , (com sendo  $\nu$  o grau de liberdade,  $\alpha$  o nível de significância).  $\eta$  é dada pela equação 9.

$$\eta = 1 - (1 - \alpha)^{1/\binom{k}{2}} \quad (9)$$

com  $k$  sendo o número de grupos em teste.

Este teste utiliza dados provenientes de testes anteriores, logo, os resultados têm a mesma unidade e formato que o teste anterior. Por exemplo, a resposta de um teste de Dunn para uma análise de variância do tipo Kruskal-Wallis seria no formato de *rank*, o que impossibilita uma comparação quantitativa. Isso significa dizer que é possível dizer quais tratamento de grupos são maiores ou menores, mas não o quanto.

#### 4.1.3. Intervalos de confiança para distribuições não normais

Um intervalo de confiança (IC) dita os limites superiores e inferiores criados a partir de um nível de significância, dentro dos quais os valores obtidos estão em uma faixa de aceitação. Significa dizer que se infinitos testes fossem realizados, uma porcentagem estipulada ligada ao nível de significância sempre estaria dentro dessa faixa (MONTGOMERY; RUNGER, 2013).

Quando a amostra pertence à uma distribuição não normal já não é mais possível criar intervalos de confiança da forma tradicional. Ao invés disso, uma alternativa são os métodos de *bootstrap* (já mencionados no item 3.3.1) baseados na obtenção de métricas por meio da seleção aleatória com reposição. Uma das maneiras de se utilizar *bootstrap* é através do  $BC_\alpha$  (*Bias-Corrected and Accelerated*), onde o intervalo de confiança é baseado em percentis de um parâmetro de interesse  $\theta$  proveniente de histogramas criados pela seleção aleatória com reposição. A diferença deste último método para o de percentis é que ele utiliza uma correção de tendência e uma variável de aceleração. De forma objetiva para  $BC_\alpha$ , os limites são aqueles obtidos pela equação 10:

$$\hat{Q}[\alpha] = \Phi \left( z_0 + \frac{z_0(\hat{\theta}) + z^{(\alpha)}}{1 + a(\hat{\theta})(z_0(\hat{\theta}) + z^{(\alpha)})} \right) \quad (10)$$

onde  $\Phi$  é a distribuição resultante,  $z_0$  é a correção de tendência e  $a$  a aceleração, ambos dependentes do comportamento da variável de interesse  $\hat{\theta}$  proveniente de cada reamostragem. Mais detalhes sobre o

método como a obtenção do parâmetro de correção e de aceleração estão disponíveis em EFRON (1981) e EFRON & TIBSHIRAN (1993).

#### 4.2. COMPARAÇÃO DE CLASSIFICADORES

A comparação de classificadores demanda métricas robustas obtidas principalmente a partir de uma matriz de classificação (ou tabela de contingência) (SAMMUT; WEBB, 2010). O quadro 1 apresenta essa matriz para um classificador binário. As colunas representam as saídas da rede, e as linhas os rótulos. As 4 saídas possíveis são:

TP (*True Positive*): verdadeiros positivos. Rótulo = Saída = 1;  
 TN (*True Negative*): verdadeiros negativos. Rótulo = Saída = 0;  
 FP (*False Positive*): falsos positivos. Rótulo = 0, Saída = 1.  
 FN (*False Negative*): falsos negativos. Rótulo = 1, Saída = 0;

Quadro 1 – Matriz de classificação.

		Saídas	
		1	0
Rótulos	1	TP	FN
	0	FP	TN

Fonte: autor.

Parâmetros de interesse provenientes da matriz de classificação são usados para análises mais aprofundadas de desempenho de classificadores. Alguns desses parâmetros estão exemplificados nas equações 11, 12 e 13.

$$TPR = \frac{TP}{(TP + FN)} \quad (11)$$

$$FPR = \frac{FP}{(FP + TN)} \quad (12)$$

$$ACC = \frac{(TP + TN)}{(TP + FN + FP + TN)} \quad (13)$$

onde TPR é a taxa de verdadeiros positivos, FPR é a taxa de falsos positivos e ACC a exatidão<sup>1</sup> ou acurácia. A taxa de verdadeiros positivos exprime a taxa de acertos em relação à saída 1. A taxa de falsos positivos é a taxa de erros que o classificador comete ao indicar classe 1 sobre todos valores 0 de rótulo. A exatidão por sua vez é a taxa de acerto de ambas as classes sobre todas as amostras.

Aparentemente, a exatidão seria suficiente para definir a qualidade de um classificador, mas isso nem sempre ocorre. Quando há desbalanceamento de classes, isto é, existe um maior número de rótulos de uma ou outra classe, a exatidão é mais influenciada pela classe com maior número de amostras, fornecendo uma resposta tendenciosa (FAWCETT, 2004).

Isso fica mais evidente com um exemplo: em um conjunto de dados de mamografia onde 98% das imagens apresentam comportamento negativo e apenas 2% positivo em relação às células cancerígenas, a simples estratégia de classificar o conjunto todo como negativo já traria uma exatidão de 98% ao classificador. Essa característica também é reconhecida como falácia da exatidão. Logo, apenas essa métrica não é suficiente para obter taxas de acerto condizentes com as necessidades do problema, pois não identifica a taxa de acerto de cada classe. A natureza das aplicações que envolvem classes desbalanceadas requer uma taxa de acerto muito grande relacionada à classe com menos exemplos (CHAWLA, 2005).

A literatura apresenta alguns métodos para contornar esse desbalanceamento já durante o treinamento das redes por meio de alterações nos algoritmos ou ainda pela manipulação do conjunto de entrada (GALAR et al., 2012). Uma alternativa mais simples é comparar resultados desbalanceados através de um método de avaliação que desconsidere esse efeito. Para trabalhar com o desbalanceamento no conjunto de testes é necessário saber inicialmente quão desbalanceado ele é e como as métricas derivadas da matriz de classificação são afetadas por esse fenômeno. O termo que sintetiza o desbalanceamento é a razão de custo (FLACH, 2003), equação 14, que tem o mesmo resultado se calculada sobre os rótulos ou sobre as saídas da rede.

---

<sup>1</sup> Neste sentido, o termo exatidão apesar de semelhante ao que é utilizado em metrologia não deve ser encarado como tal. A exatidão da medição como conceito metrológico é o grau de concordância entre o resultado de uma medição e um valor verdadeiro do mensurando, sendo um conceito qualitativo (INMETRO, 2012)

$$C = \frac{(TP + FN)}{(TN + FP)} = \frac{POSITIVOS}{NEGATIVOS} \quad (14)$$

Levando essa razão de custo em consideração a equação 13 toma a forma da equação 15, agora em termos de TPR e FPR.

$$ACC_c = \frac{TPR + C(1 - FPR)}{1 + C} \quad (15)$$

Quando C diminui, ou seja, quando existem mais saídas negativas que positivas, a taxa de falsos positivos (FPR) tem seu peso aumentado. Isso representa a situação onde a rede é penalizada mais por falsos positivos que por falsos negativos, já que positivos são maioria.

Características como exatidão, FPR, TPR e razão de custo são sumarizadas e melhor visualizadas com um recurso gráfico denominado curva característica de um receptor, ou ROC (*Receiver Operating Characteristics*).

#### 4.2.1. Curva característica de operação de um receptor (ROC)

Diferentes razões de custo, TPR e FPR fornecem uma curva denominada ROC, um comparador que permite a análise de classificadores com diferentes desbalanceamentos de classes. Sabendo quais são os rótulos, fixa-se ou encontra-se uma razão de custo única, gerando isolinhas de métrica, o que torna mais fácil a visualização de desempenho de classificadores (FLACH, 2003).

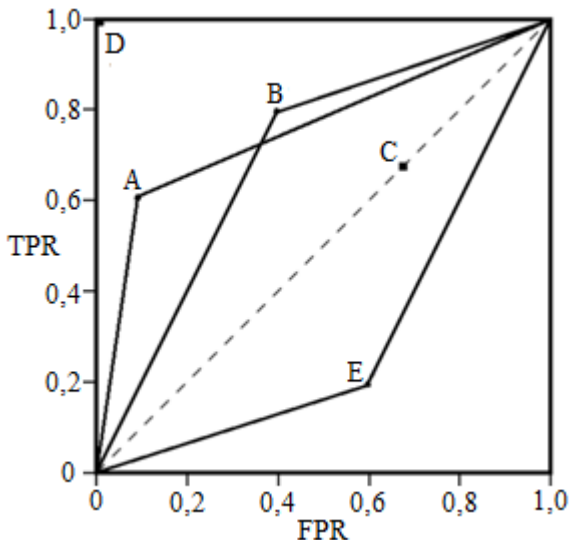
Essa visualização já há muito é utilizada na detecção de sinais (EGAN, 1975; SWETS, 1988), mas foi SPACKMAN (1989) um dos primeiros que a utilizou em aprendizado de máquina. ROC fazem uso de duas informações presentes na matriz de classificação: a taxa de verdadeiros positivos (sensibilidade) e a taxa de falsos positivos (1 – especificidade). Convém ressaltar que a ROC é criada tendo em vista a classe positiva (classe 1), ou regime permanente no caso deste trabalho.

A figura 15 ilustra uma ROC já com classificadores presentes. É um gráfico bidimensional com o TPR no eixo das ordenadas e FPR no eixo das abscissas, onde a linha diagonal que liga os pontos (0,0) e (1,1) representa o desempenho de um classificador dito aleatório. Qualquer ponto (p, p) dessa linha pode ser obtido pela previsão de positivos com probabilidade “p” e negativos com probabilidade “1-p” (FLACH, 2003).

Em uma ROC, classificadores à esquerda do valor de 0,5 de FPR são considerados conservadores, isto é, realizam a classificação com

fortes evidências de acerto. Quando à direita desse valor são ditos liberais, pois classificam com fracas evidências. O desempenho ideal de um classificador está no ponto superior esquerdo do gráfico, coordenadas [0,1]. Quanto mais próximo desse ponto melhor o desempenho do classificador. Por outro lado, quanto mais próximo da linha diagonal mais próximo do desempenho de um classificador aleatório ele se encontra (FAWCETT, 2004).

Figura 15- Exemplo de classificadores em um ROC.



Fonte: FAWCETT (2004).

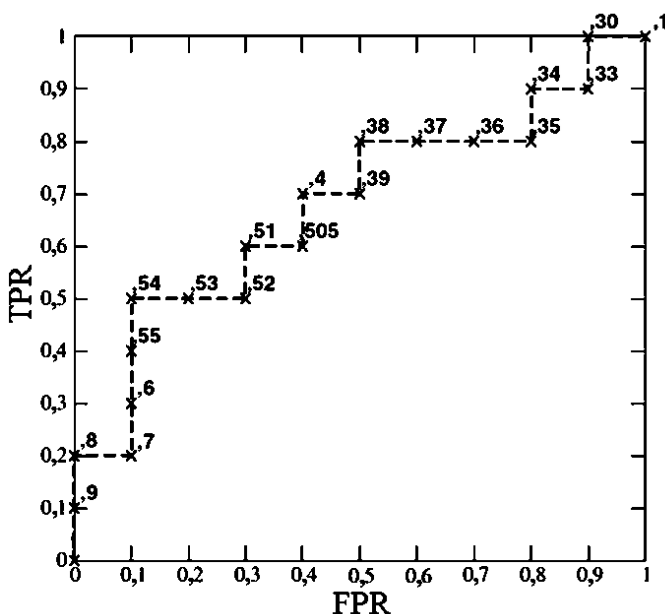
Existem diferentes maneiras para criação de uma ROC. Uma delas é a partir da obtenção de um ponto proveniente da matriz de classificação. Os exemplos de classificadores presentes na figura 15 foram obtidos por esse método. O classificador ideal é representado por “D”, e um classificador aleatório por “C”. “E” é classificador ruim para a classe positiva, mas que pode ser utilizado para a classe negativa, sendo o inverso do classificador “B”. “A” é um bom classificador, mas é necessária uma análise mais detalhada para garantir qual é o melhor entre “A” e “B”.

A outra maneira é obtendo uma matriz de classificação para cada aplicação de diferentes limiares às saídas contínuas de uma rede neural. Limiar é o “ponto de corte” da saída contínua da rede que define a



probabilidade da classe ser negativa ou positiva. Não está diretamente relacionado à função de ativação limiar. A figura 16 ilustra uma curva ROC criada dessa forma. Essa curva é tanto mais suave quanto menor for a variação dos limiares. Para sua construção é necessário inicialmente ordenar de forma decrescente o vetor de rótulos, e com essa mesma ordenação o vetor de saídas. Essa ordenação dos vetores levando em conta a classe positiva que ocasiona o aspecto crescente tanto para TPR quanto para FPR é observada na figura supracitada. O pseudocódigo para criação da ROC por este método encontra-se no trabalho de FAWCETT (2004).

Figura 16 - Curva ROC criada a partir de diferentes limiares.



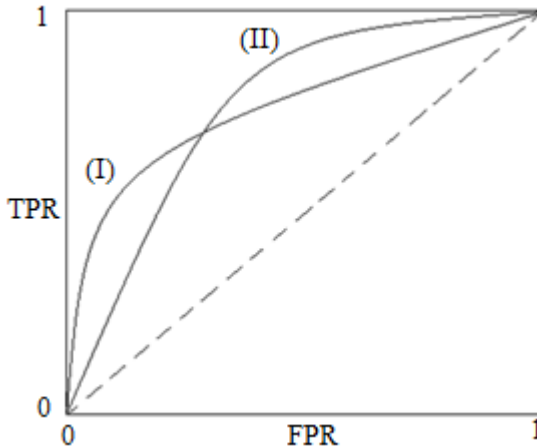
Fonte: FAWCETT (2004).

#### 4.2.2. Área sob a curva (AUC) e linhas isométricas

Para avaliar qual classificador ou limiar produz melhor resposta em relação ao conjunto teste é fundamental que a curva ROC possa ser reduzida a um único valor. Isso é feito por meio da AUC (*Area Under Curve*, ou área sob a curva) (BRADLEY, 1997). O valor de AUC varia entre zero e um, com a ressalva de que um classificador aleatório produz um AUC igual a 0,5 correspondente à área abaixo da linha pontilhada diagonal ascendente da figura 15.

Pode ocorrer de duas curvas possuírem um mesmo AUC, como pode ser visto na figura 17 (curvas I e II). Nesse caso, constata-se apenas qual classificador é mais conservador ou liberal, o que pode ser suficiente dependendo do problema mas que em muitos casos não é conclusivo.

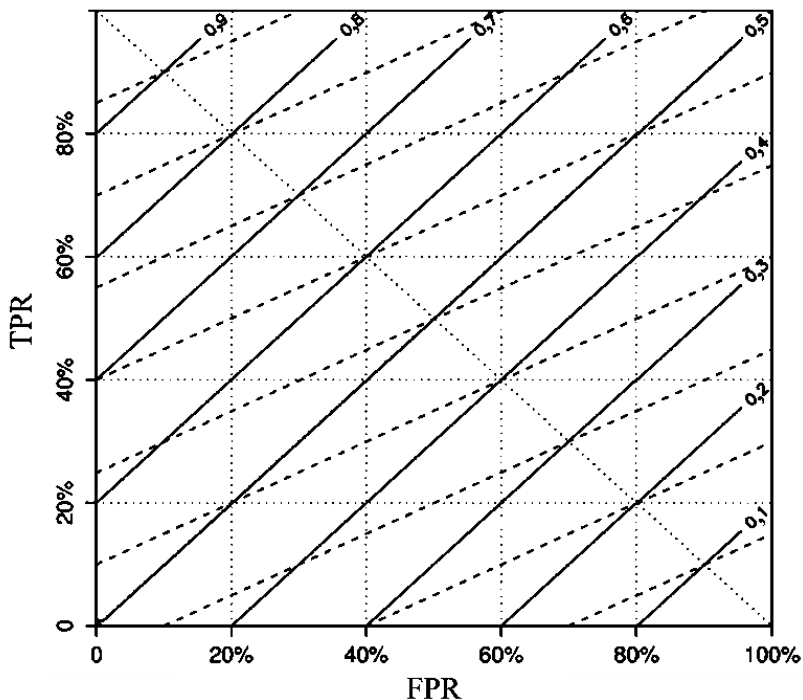
Figura 17 - ROC com exemplos de dois classificadores com mesmo AUC.



Fonte: autor

Quando a conclusão acerca do aspecto geral do classificador não é suficiente há a alternativa de análise da ROC em relação a isolinhas de métricas. Isolinhas são coleções de pontos com o mesmo valor para determinada métrica. Um exemplo de isométrica é a exatidão, apresentada na figura 18, onde cada linha é criada a partir da equação 15 com a fixação de uma razão de custo. Classificadores situados em uma mesma isolinha variam a relação entre TPR e FPR com uma inclinação constante igual a razão de custo (FLACH, 2003).

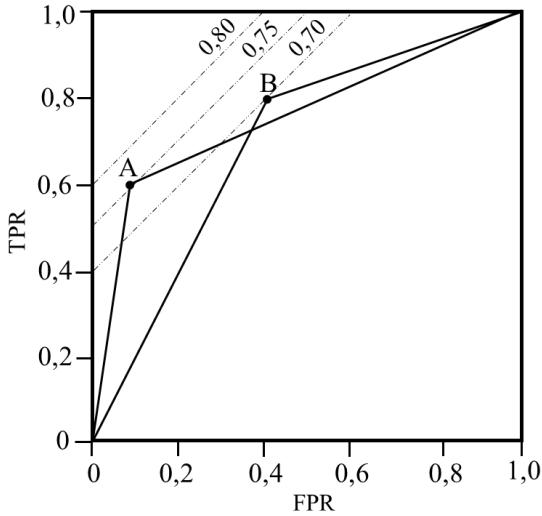
Figura 18 - Isolinhas de exatidão. Linhas sólidas para  $C=1$  e tracejadas para  $C=0,5$ .



Fonte: FLACH (2003).

Com esse conhecimento e fixando  $C=1$ , é possível identificar, por exemplo, qual classificador (A ou B) da figura 15 é melhor. A figura 19 possui isolinhas de exatidão, onde é possível observar que o classificador A é melhor na exatidão (0,76 contra 0,7) e também na AUC (0,75 contra 0,66). Caso os classificadores tivessem razões de custo diferentes, isolinhas próprias para cada classificador deveriam ser traçadas e só assim os resultados seriam comparados de forma adequada.

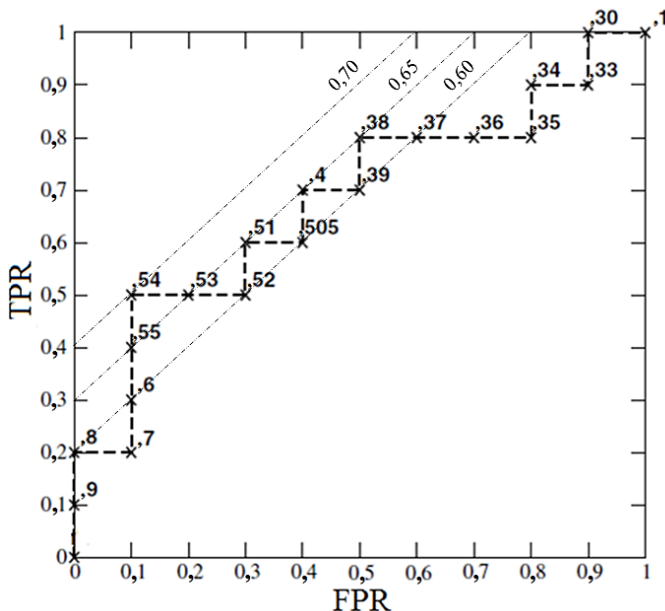
Figura 19 - Comparação de classificador por isolinhas de exatidão (cinza) para dois classificadores criados a partir de diferentes matrizes de classificação.  $C=1$ .



Fonte: adaptado de FAWCETT (2004)

Ainda como exemplo, essa análise também é realizada para o classificador presente na figura 16. Na figura 20, é possível notar que o limiar de 0,54 é melhor que os demais no quesito exatidão (exatidão de 0,7), apresentando TPR igual a 0,5 e FPR igual a 0,1.

Figura 20 - ROC para diferentes limiares com isolinhas de exatidão (mais claras) para curva criada a partir de ranqueamento.  $C=1$ .



Fonte: adaptado de FAWCETT (2004).

#### 4.2.3. Avaliação de incertezas em ROC

Na maioria dos casos, o conjunto de teste possui mais de um ensaio. É possível agrupar todos esses ensaios em um único vetor de saída, resultando uma única curva na ROC. Contudo, esse procedimento inibe a motivação da utilização de diferentes ensaios para teste, que é a obtenção de uma medição da variância dos resultados (FAWCETT, 2004).

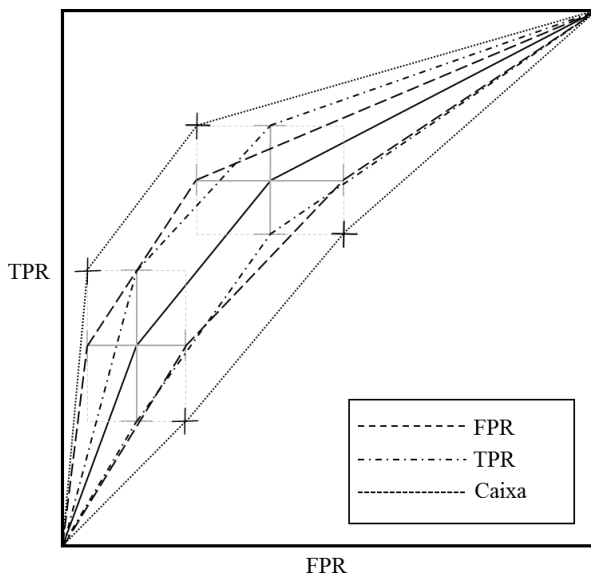
Partindo de várias curvas de saída há mais de uma forma de criar faixas de confiança em ROC. Três possibilidades aplicadas em aprendizado de máquina são (MACSKASSY; PROVOST, 2004):

1. Média vertical: FPR fixo e criação da faixa de confiança a partir de um intervalo de confiança dos valores de TPR.
2. Média horizontal: TPR fixo e criação da faixa de confiança a partir de um IC dos valores de FPR.
3. Média de limiar: limiar fixo e igualmente espaçado, com faixa de confiança criada a partir da junção dos IC tanto de TPR quanto de FPR.

Apesar das médias verticais e horizontais serem de mais fácil obtenção nem sempre essas variáveis são de controle do operador, o que não ocorre quando o limiar é utilizado. A média de limiar percorre as saídas contínuas do classificador do maior para o menor em intervalos igualmente espaçados e em cada um desses valores encontra um IC para TPR e FPR (MACSKASSY; PROVOST, 2004). O pseudocódigo para obtenção dos intervalos de confiança pela média de limiar é encontrada no trabalho de FAWCETT (2004).

Para criação de faixas de confiança os extremos dos intervalos de confiança devem ser conectados. Na figura 21 está um exemplo de conexão partindo da aplicação da média de limiar e dos diferentes métodos de junção, que são por TPR, FPR e por caixa. Os métodos de junção pelos limites de TPR e FPR geram faixa de confiança mais estreitas com diferenças mais evidentes na curva superior. Os pontos da caixa, sinalizados pelos sinais de '+' dessa figura, são constituídos pelos valores extremos dos intervalos de confiança tanto de FPR e TPR, gerando uma faixa mais conservadora que inclusive engloba as faixas criadas pelos outros métodos.

Figura 21 - Exemplo de média de limiar e diferentes intervalos de confiança para os métodos de junção por FPR, TPR e caixa.



A junção por TPR também pode ser considerada conservadora pois considera o IC sobre os acertos do classificador, com a desvantagem de que alguns intervalos de confiança para FPR ultrapassam a faixa resultante principalmente para valores altos de FPR. Entretanto, é possível atribuir o mesmo grau de confiança à faixa obtida por essa junção (MACSKASSY; PROVOST, 2004). Dessa maneira, a faixa tem duas utilidades: criar um IC para AUC e exprimir uma faixa de variância do classificador.





## 5. AGRUPAMENTO DE ENSAIOS E DESENVOLVIMENTO DE CLASSIFICADORES ESPECIALIZADOS

Neste capítulo encontra-se o método de treinamentos de redes para inferência de regime permanente, resultados provenientes do agrupamento dos ensaios, aspectos experimentais utilizados na obtenção desses resultados e da comparação dos desempenhos das redes treinadas com ensaios selecionados e não selecionados.

### 5.1. REDE DE INFERÊNCIA DE REGIME PERMANENTE

A rede neural utilizada neste trabalho possui 2 camadas intermediárias e uma de saída, sendo do tipo MLP alimentada adiante totalmente conectada e sem realimentação. Cada uma das camadas intermediárias possui 10 neurônios, com funções de ativação tangente hiperbólica, e na camada de saída função logística. Essa quantidade de camadas e número de neurônios foi utilizada pois as redes foram treinadas em um computador pessoal sendo o tempo fator limitante.

O treinamento foi do tipo supervisionado com alimentação de dados por épocas e algoritmo de otimização Levenberg-Marquardt. Além do critério de parada de treinamento por validação cruzada, este também era interrompido pelo número máximo de épocas de treinamento, performance ou gradiente de acordo com os limites da tabela 3. A performance remete ao valor mínimo da norma euclidiana do vetor gradiente da superfície de erro. Já a parada pelo gradiente encerra o treinamento caso um determinado número de épocas tenha passado sem que o erro médio quadrático do conjunto de treinamento tenha diminuído.

Tabela 3 - Critérios de parada de treinamento.

<b>Critério de parada</b>	<b>Valor necessário</b>
Número máximo de épocas	500
Performance	0,09
Gradiente	0,0001
Número de épocas de validação	20

Fonte: autor.

Quanto à seleção das variáveis de entrada, os trabalhos anteriores observaram que as grandezas que geraram os melhores resultados no

treinamento das redes foram temperatura de corpo, pressão de sucção, e capacidade de refrigeração. Sobre essas três variáveis realizou-se transformações matemáticas gerando um total de 16 entradas. O quadro 2 apresenta essas entradas.

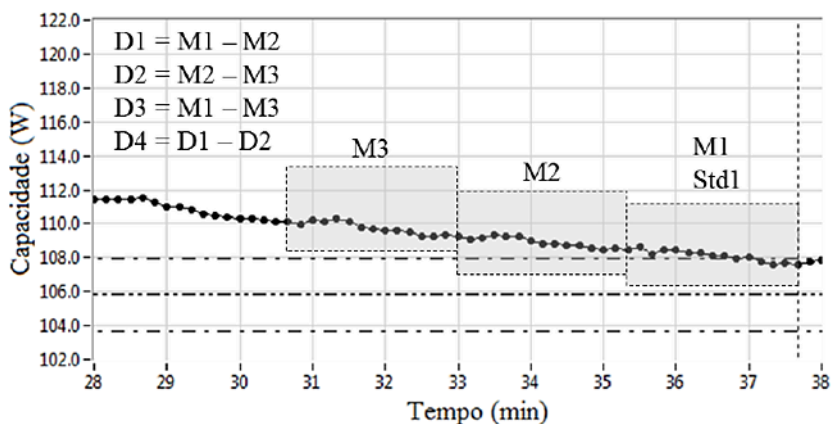
Quadro 2 - Entradas da rede de inferência de regime permanente.

<b>Variável</b>	<b>Entrada</b>	
Capacidade de refrigeração	1	Desvio padrão de 15 amostras da variável precedentes ao instante atual (Std1)
	2	Média de 15 amostras da variável precedentes ao instante atual (M1)
	3	Média de 15 amostras da variável precedentes ao instante atual deslocado de 15 amostras (M2)
	4	Média de 15 amostras da variável precedentes ao instante atual deslocado de 30 amostras (M3)
	5	Diferença entre a 2ª e 3ª entradas (D1)
	6	Diferença entre a 3ª e 4ª entradas (D2)
	7	Diferença entre a 2ª e 4ª entradas (D3)
	8	Diferença entre a 5ª e 6ª entradas (D4)
Temperatura de corpo do compressor	9	Média de 15 amostras da variável precedentes ao instante atual.
	10	Média de 15 amostras da variável precedentes ao instante atual deslocado de 15 amostras.
	11	Diferença entre 9ª e 10ª entrada.
	12	Diferença entre a 9ª entrada e a média de 15 amostras da variável precedentes ao instante atual deslocado de 30 amostras.
Pressão de sucção	13	Média de 15 amostras da variável precedentes ao instante atual.
	14	Média de 15 amostras da variável precedentes ao instante atual deslocado de 15 amostras.
	15	Diferença entre 13ª e 14ª entrada.
	16	Diferença entre a 13ª entrada e a média de 15 amostras da variável precedentes ao instante atual deslocado de 30 amostras.

Fonte: adaptado de NASCIMENTO (2015).

Cabe ressaltar que as médias deslocadas resolvem a limitação das redes utilizadas neste trabalho de lidar com séries dinâmicas (NASCIMENTO, 2015). A figura 22 ilustra como essas médias englobam um espaço da série temporal favorecendo o aprendizado da rede para a capacidade de refrigeração.

Figura 22 - Representação das entradas relacionadas a capacidade



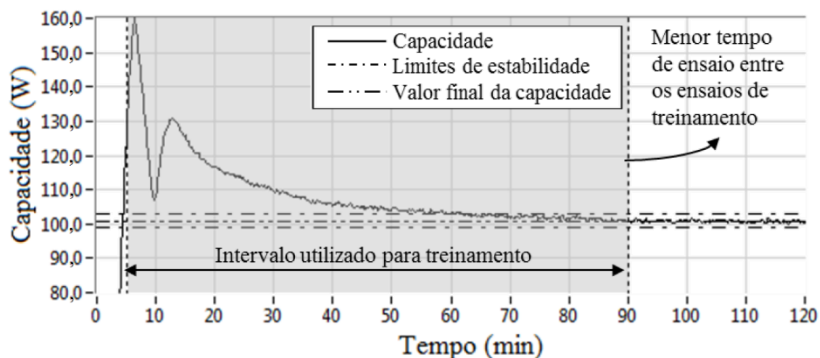
Fonte: NASCIMENTO (2015).

Antes de serem inseridas na rede para aprendizado e após o pré-processamento dos ensaios, as variáveis de entrada sofrem alguns ajustes (equalização). Para evitar que a rede seja sensível à capacidade de refrigeração do compressor, visto que podem ser das mais variadas potências, o valor da capacidade é dividido pela média dos valores da capacidade obtidos no regime permanente. Cabe lembrar que todo o comportamento do ensaio é conhecido durante o treinamento, logo, essa prática é possível. Isso também é realizado para temperatura de corpo e para pressão de sucção. Após a equalização, as variáveis são normalizadas para que possuam média zero e desvio padrão igual a 1. Então são restritas ao intervalo -1 a +1 por um mapeamento mínimo-máximo. Esses procedimentos favorecem o uso de funções de ativação sigmoidais, como visto no item 3.1.1.

Frequentemente as primeiras leituras das variáveis apresentam grande amplitude devido aos ajustes de controle, por isso, os cinco minutos iniciais dos ensaios são descartados (30 aquisições). Além disso,

para padronizar o tamanho dos ensaios, o instante final do intervalo de tempo utilizado corresponde ao tempo do ensaio de menor duração dentre todos presentes no conjunto de treinamento (NASCIMENTO, 2015). A figura 23 apresenta esses detalhes.

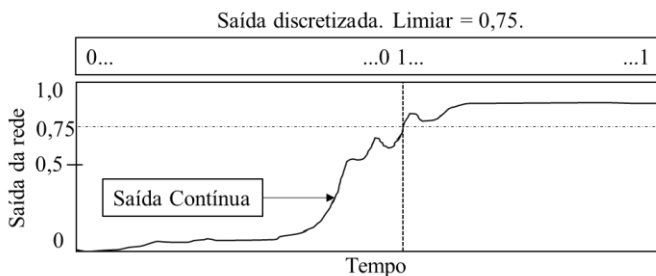
Figura 23 - Intervalo do ensaio utilizado para treinamento das redes.



Fonte: NASCIMENTO (2015).

Por fim, a saída da rede neural é contínua na faixa de zero a um. Para utilização como classificador a saída deve ser discretizada, o que é realizado pela aplicação de um limiar. A figura 24 ilustra um exemplo uma dessas discretizações para um limiar hipotético de 0,75. Assim, á toda resposta contínua da rede maior que esse limiar é atribuído o valor 1 (ou positivo para regime permanente), e valor 0 às respostas de valor inferior.

Figura 24 - Exemplo de aplicação de limiar para discretização de saída contínua de uma rede neural. Limiar = 0,75.



Fonte: autor.

O procedimento de treinamento foi similar ao já utilizado nos trabalhos anteriores, com a exceção das etapas de criação de comitês e fusão das respostas dos mesmos.

## 5.2. AGRUPAMENTO DE ENSAIOS

As respostas de redes neurais dependem diretamente das características das variáveis de entrada utilizadas durante seu treinamento. Quanto mais diversos os exemplos apresentados, mais a rede deve adaptar-se e pior sua resposta para comportamentos que não são observados. Como visto no item 3.1.3.1, ao restringir a diversidade de exemplos que são fornecidos ao treinamento a rede precisa ajustar menos seus pesos, o que resulta em vantagens como aceleração do treinamento e melhora da capacidade de generalização.

Essa redução da diversidade dos ensaios apresentados para treinamento é o escopo do agrupamento de ensaios realizado neste trabalho. Para a realização desse agrupamento foi considerada a maneira como os ensaios progridem em direção ao regime permanente.

Para os métodos de agrupamento que serão apresentados nesta seção é imperativo o conhecimento da faixa de valores de capacidade que define o regime permanente. Para isso, primeiramente foi calculada a capacidade final dos ensaios como a média das últimas 210 aquisições (ou 35 minutos finais), sendo que uma aquisição é realizada a cada 6 segundos nas bancadas. A partir desse valor final de capacidade de refrigeração aplica-se a faixa de aceitação de mais ou menos dois por cento.

### 5.2.1. Descrição geral do método de agrupamento

Duas características da curva de capacidade de refrigeração foram avaliadas para agrupamento manual dos ensaios em grupos: o modo como a curva adentra a faixa de aceitação e a diferença entre o valor inicial da capacidade e o valor desta variável no ponto de entrada em regime permanente. Os ensaios foram então segregados em 4 grupos:

**G1A:** Ensaios que adentram a faixa de aceitação pelo limite superior e que têm valor inicial de capacidade de refrigeração superior ao do ponto de entrada em regime.

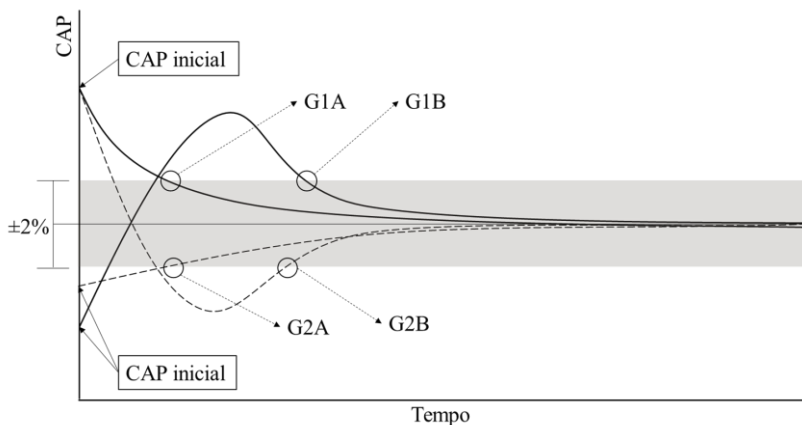
**G1B:** Ensaios que adentram a faixa de aceitação pelo limite superior e que têm o valor inicial de capacidade de refrigeração inferior ao do ponto de entrada em regime.

**G2A:** Ensaios que adentram a faixa de aceitação pelo limite inferior e que têm o valor inicial de capacidade de refrigeração inferior ao do ponto de entrada em regime.

**G2B:** Ensaios que adentram a faixa de aceitação pelo limite inferior e que têm o valor inicial de capacidade de refrigeração superior ao do ponto de entrada em regime.

A figura 25 ilustra o comportamento da variável de desempenho desses grupos. A área em cinza corresponde à faixa de aceitação de entrada em regime (apenas ilustrativa nesse caso). As linhas sólidas correspondem aos grupos que adentram à faixa de aceitação pelo limite superior, e as linhas tracejadas àqueles que adentram pelo limite inferior da faixa.

Figura 25 - Aspecto geral da dinâmica dos ensaios agrupados. Os limites de  $\pm 2\%$  correspondem ao regime permanente.



Fonte: autor.

### 5.2.2. Análises estatísticas

A condição da bancada é um dos fatores que influencia as dinâmicas do ensaio e conseqüentemente sua duração e momento de entrada em regime permanente. Essa condição está associada aos ensaios realizados anteriormente e às configurações impostas pelo operador e/ou sistema de controle, sendo de difícil previsão devido ao grande número

de modelos e compressores que são testados. Algumas das variáveis utilizadas para treinamento das redes são influenciadas pela condição inicial da bancada, o que torna imperativo um estudo sobre seus comportamentos principalmente quando relacionados às condições finais do ensaio para avaliação de sua progressão. O agrupamento dos ensaios além de gerar conjuntos de treinamento com baixa diversidade também permitiu essa análise, que foi feita de forma estatística com o *software* MATLAB - 2014b.

A tabela 4 apresenta a incidência resultante de ensaios em cada grupo sobre um total de 2503 ensaios típicos. Ocorreu a prevalência de ensaios nos grupos G1A e G1B, com cerca de 81% dos ensaios, contra aproximadamente 19% dos grupos G2A e G2B.

Tabela 4 - Incidência dos grupos após agrupamento manual.

<b>Grupo</b>	<b>Incidência</b>	<b>%</b>
G1A	1481	59,17
G1B	549	21,93
G2A	237	9,47
G2B	236	9,43

Fonte: autor.

Esses grupos possuíam ensaios provenientes de duas plantas da empresa parceira. Para evitar qualquer comparação equivocada, foram utilizados apenas ensaios referentes a uma das plantas e ainda segregados aqueles ensaios submetidos à condição de *checkpoint*. Essa planta foi escolhida devido aos resultados mais homogêneos em relação à arquitetura, estratégias de controle, e instrumentação. Nessa mesma planta ainda haviam duas bancadas, mas não houve diferença considerável entre elas no que se refere a duração dos ensaios ou entrada em regime permanente.

Especificamente na condição de *checkpoint*, a pressão de sucção final desejada é de aproximadamente 0,62 bar e a pressão de descarga é de, aproximadamente, 7,6 bar. Esses valores são decorrentes das condições de temperatura de evaporação (-23,3 °C) e condensação (54,4 °C) do fluido de trabalho, derivadas da carta psicrométrica do fluido R600a. A tabela 5 apresenta a quantidade de ensaios restritos à planta escolhida e os respectivos ensaios da condição de *checkpoint*.

Tabela 5 – Ensaios provenientes de apenas uma planta, seus respectivos grupos e aqueles ensaiados na condição de *checkpoint*.

<b>Grupo</b>	<b>Ensaios</b>	<b>Apenas <i>checkpoint</i></b>
G1A	177	150
G1B	134	80
G2A	84	43
G2B	181	162

Fonte: autor.

Nesses grupos de ensaios foram realizadas comparações de duração, momento de entrada em regime permanente, variação de pressão de sucção e variação de pressão de descarga. As distribuições dessas variáveis não são normais (testadas pelo teste de normalidade de Kolmogorov Smirnov com nível de confiança de 95%), sendo esse o motivo da utilização da análise de variância de Kruskal-Wallis seguida do teste de hipóteses de múltipla comparação de Dunn (ambos não paramétricos). A hipótese nula para os testes de variância é de que os grupos apresentavam comportamento igual para a variável em questão, e o nível de confiança para os testes de Dunn foi de 95%.

É necessária uma observação sobre o tamanho das barras de *rank* presentes nos gráficos do teste de Dunn. A análise de Kruskal-Wallis, que fornece os dados para o teste de Dunn, permite a utilização de grupos com número de amostras diferentes. Dessa maneira o tamanho da barra é inversamente proporcional ao número de amostras, e acaba sendo uma forma de tornar mais difícil a rejeição da hipótese nula.

Em cada uma das análises posteriores é apresentado o valor crítico de rejeição da hipótese nula ( $\chi^2$ ) e o respectivo valor p, que quanto mais próximo de zero mais considerável é a rejeição da hipótese nula.

#### 5.2.2.1. Duração do ensaio e momento de entrada em regime permanente em grupos

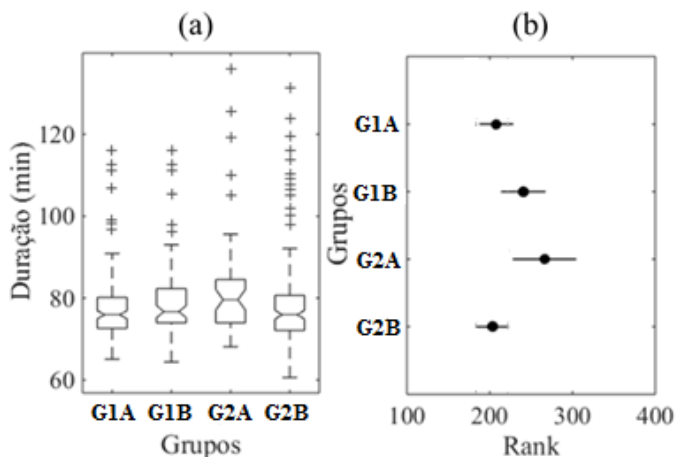
A análise de variância rejeitou a hipótese nula ( $\chi^2 = 12,0$ ; valor  $p > \chi^2 = 0,01$ ) de que os grupos tinham durações de ensaios iguais. Mais detalhes sobre o teste de normalidade e análise de variância podem ser visualizados no Apêndice A – quadro 5 e figura 40, e Apêndice B – quadro 6, respectivamente.

A figura 26 apresenta o diagrama de caixa da dispersão das durações, item (a), e o teste de Dunn com os respectivos *ranks* – item (b).



Em (a) nota-se um maior número de *outliers* em G2B, mas sua mediana e seus quartis são semelhantes aos dos demais grupos. Com o teste de Dunn é possível perceber quais grupos causaram a rejeição da hipótese nula no teste anterior, onde a única diferença relevante ocorreu entre o grupo G2B e G2A.

Figura 26 - Duração em grupos. (a) Diagrama de caixa (b) Teste de Dunn.  $\alpha = 0,05$ .



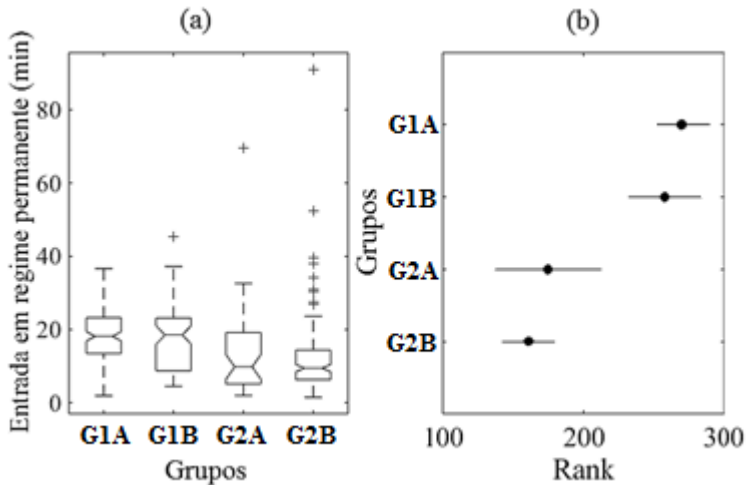
Fonte: autor.

Um ensaio pode continuar sendo executado mesmo após transcorrido o tempo obrigatório ditado pela norma para determinação do regime permanente. Por isso o mesmo procedimento estatístico foi realizado para o momento no qual os ensaios adentravam o regime permanente.

Novamente a hipótese nula foi rejeitada para Kruskal-Wallis ( $\chi^2 = 72,72$ ; valor  $p > \chi^2 \sim 0$ ). Mais detalhes sobre o teste de normalidade e análise de variância podem ser visualizados no Apêndice A – quadro 5 e figura 41, e Apêndice B – quadro 6, respectivamente. O teste de Dunn forneceu a resposta presente na figura 27, item (b). Nessa mesma figura, item (a), está o diagrama de caixa representando a dispersão dessa variável nos grupos. A dispersão do grupo G2B é notavelmente inferior a dos demais grupos no entanto possui mais *outliers*. O teste de Dunn

mostrou que ensaios do grupo G2A e G2B entram em regime consideravelmente mais cedo do que os ensaios dos grupos G1A e G1B.

Figura 27 - Momento de entrada em regime permanente para grupos. (a) Diagrama de caixa. (b) Teste de Dunn.  $\alpha=0,05$ .



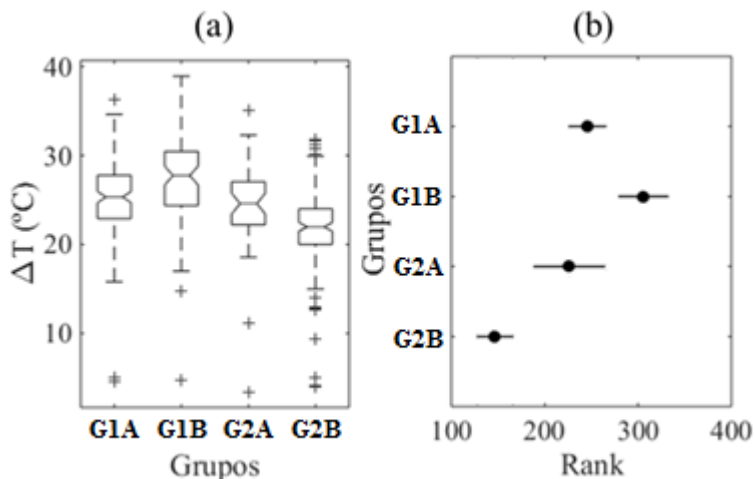
Fonte: autor.

#### 5.2.2.2. Temperatura de corpo em grupos

Foi avaliada a variação da temperatura (entre temperatura inicial e temperatura final) de corpo ao longo do ensaio, com o teste de variância rejeitando a hipótese nula de que os grupos possuem variação de temperatura iguais ( $\chi^2 = 98,45$ ; valor  $p > \chi^2 \sim 0$ ). Mais detalhes sobre o teste de normalidade e análise de variância podem ser visualizados no Apêndice A – quadro 5 e figura 42, e Apêndice B – quadro 6, respectivamente.

A Figura 28 apresenta o diagrama de caixa para variação da temperatura de corpo, item (a), e o teste de Dunn para essa mesma variável, item (b). A dispersão inferior do grupo G2B remete a uma variação menor que a dos demais, o que é comprovado com o teste de Dunn. Este grupo tem variação consideravelmente menor que a dos demais no quesito temperatura de corpo, tendo inclusive *outliers* inferiores a quartis dos demais grupos.

Figura 28 - Variação da temperatura de corpo em grupos. (a) Diagrama de caixa. (b) Teste de Dunn.



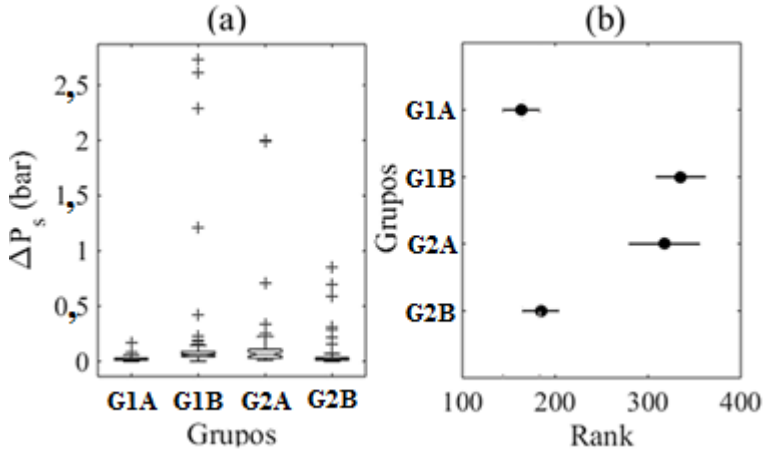
Fonte autor.

### 5.2.2.3. Pressão de sucção em grupos

Assim como a variação da temperatura de corpo, a pressão de sucção também foi avaliada. Porém, como neste último caso existe a possibilidade da pressão de sucção inicial estar acima ou abaixo do valor final (no caso da temperatura o compressor sempre é aquecido), foi utilizado seu módulo para evitar comparações equivocadas. Assim, foi analisada a diferença entre a pressão de sucção inicial após 25 aquisições e a pressão de sucção final obtida, que está relacionada à condição de *checkpoint* (0,63 bar).

A figura 29 (a) apresenta o diagrama de caixa com a dispersão da variação da pressão de sucção em ensaios. A existência de *outliers* dificulta a comparação entre os quartis, mas o teste de Dunn (b) comprova a rejeição da hipótese nula pelo teste de variância ( $\chi^2 = 134,96$ ; valor  $p > \chi^2 \sim 0$ ) e mostra que existe diferença significativa entre os grupos G1A|G2B e os demais grupos. Mais detalhes sobre o teste de normalidade e análise de variância podem ser visualizados no Apêndice A – quadro 5 e figura 43, e Apêndice B – quadro 6, respectivamente.

Figura 29 – Módulo da variação absoluta da pressão de sucção em grupos. (a) Diagrama de caixa. (b) Teste de Dunn.

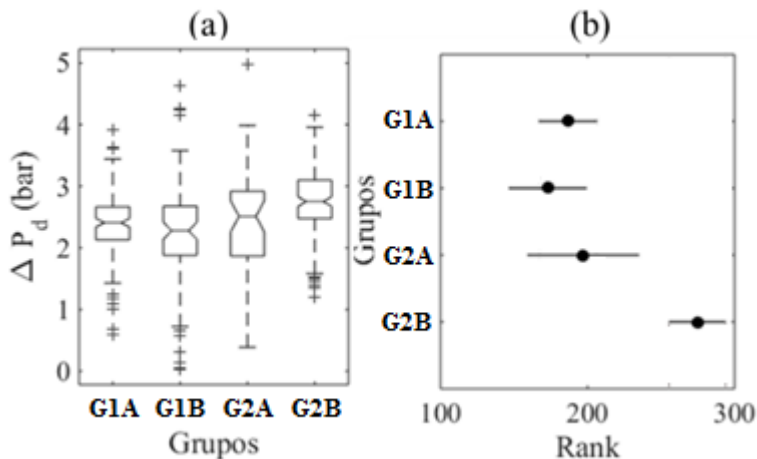


Fonte: autor.

#### 5.2.2.4. Pressão de descarga em grupos

A análise da pressão de descarga seguiu o mesmo procedimento da pressão de descarga. Após a rejeição da hipótese nula ( $\chi^2 = 53,76$ ; valor  $p > \chi^2 \sim 0$ ), o teste de Dunn demonstrou que o grupo G2B tem variação de pressão de descarga consideravelmente maior que a dos demais grupos. Mais detalhes sobre o teste de normalidade e análise de variância podem ser visualizados no Apêndice A - quadro 5 e figura 44, e Apêndice B – quadro 6, respectivamente. A figura 30 apresenta o diagrama de caixa da dispersão, item (a), e o teste de Dunn para a respectiva variável, item (b).

Figura 30 - Módulo da variação da pressão de descarga em grupos. (a) Diagrama de caixa. (b) Teste de Dunn.



Fonte: autor.

### 5.2.3. Resultados referentes ao agrupamento de ensaios

O agrupamento manual segregou os ensaios de desempenho e, para uma análise mais adequada, utilizou para os testes de hipóteses apenas aqueles de *checkpoint* provenientes de uma planta da empresa. Ainda nessa mesma planta, os ensaios são executados em dois painéis diferentes, mas análises estatísticas semelhantes àquelas realizadas nos grupos mostraram que não existe diferença relevante entre esses calorímetros no que se refere a duração e momento de entrada em regime permanente.

Os testes de variância e multicomparação foram realizados em grupos de tratamentos com diferentes números de amostras, mas essa característica é contornada pelos testes de Kruskal-Wallis e de Dunn, cujas análises podem ser aplicadas em distribuições não normais. Os tratamentos (variáveis de interesse) testados foram duração dos ensaios, entrada em regime permanente, variação da temperatura de corpo, módulo da variação da pressão de sucção e módulo da variação de pressão de descarga.

O quadro 3 apresenta as variáveis e as diferenciações mais acentuadas, uma vez que em todos os tratamentos a hipótese nula de igualdade de tratamentos foi rejeitada.

Quadro 3 - Variáveis e respectivas diferenciações mais acentuadas provenientes do teste de Kruskal-Wallis e de Dunn.

Variável (tratamento)	Diferenciação acentuada
Duração	G2B; G1A < G2A
Momento de entrada em regime permanente	G2B; G2A < G1A; G1B
$\Delta T$	G2B < G1A; G1B; G2A
$ \Delta P_s $	G1A; G2B < G1B; G2A
$ \Delta P_d $	G1A; G1B; G2A < G2B

Fonte: autor.

O destaque é o grupo G2B, que possui não só duração consideravelmente menor como também momento em entrada em regime permanente. A primeira característica está associada a baixa variação de temperatura de corpo, que é a variável com progressão mais lenta em um ensaio e que acaba definindo sua duração. Quanto à entrada em regime permanente, o grupo G2B apresenta uma variação do módulo da pressão de sucção mais estreita que a dos grupos G1B e G1A. Em um ensaio ideal é desejado que a variação da pressão de sucção seja a menor possível, pois pequenas variações já influenciam a vazão mássica que por sua vez prejudicam a chegada aos limites de RP desejados. O grupo G2B apresentou também uma variação maior referente ao módulo da pressão de descarga. Esse comportamento não é prejudicial, pois a pressão de descarga não tem tanta influência como a pressão de sucção quando essa última está sob controle (FLESCHE, 2008, 2012; MARTINS et al, 2011; STEINBACH, 2008).

Com esses resultados em mente, conhecer e entender o comportamento das variáveis que influenciam a dinâmica – principalmente aqueles presentes no treinamento das redes – assume grande importância, pois como um dos grupos possui duas características desejáveis (duração e entrada em regime permanente) menores que a dos demais, seria ideal levar o controle da bancada a induzir essa dinâmica, alcançando as vantagens associadas.

### 5.3. REDES ESPECIALIZADAS

Como já foi mencionado no item 3.1.3.1, restringir a diversidade dos exemplos que são apresentados às redes para treinamento favorece o aprendizado e melhora os resultados. Espera-se dessa forma que as

dinâmicas presentes em cada um dos grupos separados manualmente especializem as redes e as tornem aptas a apresentarem melhor desempenho quando apresentadas a ensaios com as características utilizadas durante o agrupamento manual.

Nesta seção são descritos os procedimentos para treinamento das redes ditas especializadas, a criação dos comitês e a comparação entre redes especializadas e não especializadas.

Dezenove redes foram treinadas (*software* MATLAB– 2014b) para cada grupo com diferentes conjuntos de dados criados pelo método *bootstrap*. Este número foi escolhido para facilitar a utilização da votação majoritária se a totalidade das redes fosse escolhida para o comitê.

Diferente das análises estatísticas, cada conjunto de treinamento foi constituído de 236 ensaios (menor quantidade de ensaios, referente ao subgrupo G2B), provenientes das duas plantas e que não estavam restritos apenas às condições de *checkpoint*. Para as redes não especializadas, houve uma seleção aleatória de 236 ensaios das mais diversas dinâmicas e o treinamento e criação do comitê foi feito da mesma forma que o das redes especializadas. A separação das parcelas dos ensaios para cada função deu-se da maneira presente na tabela 6.

Tabela 6 - Separação de ensaios por finalidade.

<b>Função</b>	<b>Contagem</b>	<b>% do total</b>
Treinamento	138	58
Validação	23	10
Criação do comitê	30	13
Teste	45	19

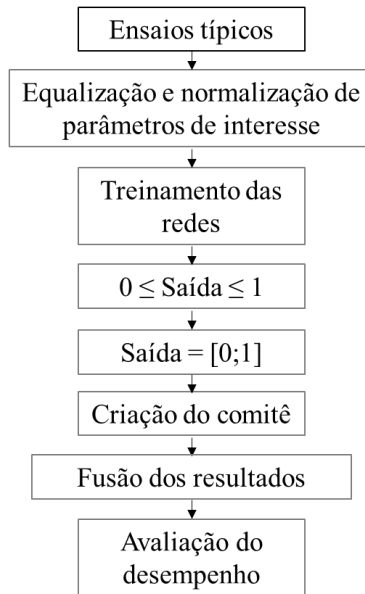
Fonte: autor.

A separação dos ensaios não segue uma regra rígida. É natural dedicar mais ensaios para treinamento pois a grande quantidade favorece a variedade, e nesta etapa é desejado que as redes sejam as mais representativas possíveis. O número de ensaios para validação *online* geralmente é menor que as demais proporções. Foi adicionado uma separação a mais neste caso, a de criação do comitê, já que seria tendencioso utilizar os mesmos ensaios de criação de comitê para testes. Outro detalhe importante é que são os ensaios de treinamento e validação que são escolhidos aleatoriamente durante a utilização do *bootstrap*.

A partir dos ensaios típicos, o treinamento e utilização das redes tanto especializadas quanto não especializadas pode ser resumido pelo

fluxograma da figura 31. Em síntese, a partir de ensaios típicos que podem ser provenientes de diversos modelos, capacidades e fluidos refrigerantes, é realizado uma equalização e normalização. Esses dados transformados são utilizados para treinamento das redes, que fornecem uma saída contínua. Essa saída é discretizada para criação do comitê ou mantida contínua para criação do ROC. Ocorre então a fusão dos resultados (contínuos ou não) e a posterior comparação de desempenho.

Figura 31 - Procedimento de treinamento e utilização das redes.



Fonte: autor.

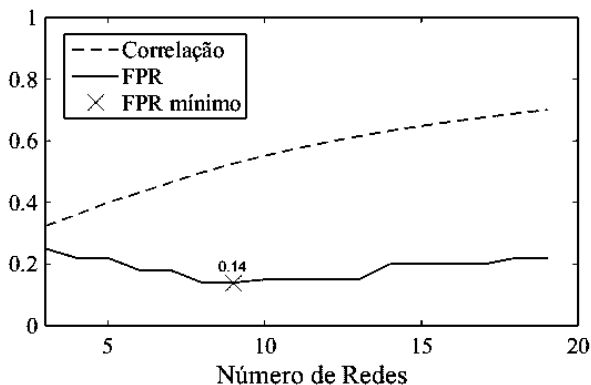
### 5.3.1. Criação dos comitês

As redes para cada grupo foram selecionadas pela correlação de Goebels, e o número ideal foi escolhido pelo FPR mínimo. Ambas métricas foram aplicadas no conjunto de 30 ensaios designado para criação do comitê. A figura 32 apresenta um exemplo da progressão da correlação e da FPR conforme as redes eram incluídas no comitê. Já era esperado que a correlação aumentasse conforme mais redes fossem adicionadas, mas o comportamento de FPR era desconhecido. Como a definição de alta correlação depende dos problemas e é arbitrária, redes



consideradas altamente correlacionadas foram aquelas que apresentaram coeficiente maior que 0,7.

Figura 32 - Grupo G1A. Comitê de 9 redes, FPR=0,14 e correlação máxima de 0,53.



Fonte: autor.

A tabela 7 sintetiza as informações acerca dos comitês para todos os grupos e para as redes treinadas sem agrupamento (grupo “Geral”). Redes presentes no subgrupo G2A foram as que apresentaram menores correlações.

Tabela 7 - Respective números de redes, correlações, e valores de FPR para os grupos.

<b>Grupo</b>	<b>Nº de Redes</b>	<b>Correlação</b>	<b>FPR</b>
G1A	9	0,53	0,14
G1B	13	0,55	0,15
G2A	5	0,11	0,14
G2B	19	0,28	0,08
Não-Especializada	3	0,2	0,29

Fonte: autor.

A menor FPR encontrada está associada às redes do grupo G2B, e que foi obtida sem a remoção de redes. Nota-se a superioridade na criação de comitês desses grupos em relação às redes não especializadas,

onde a correlação foi baixa mas o valor de FPR é alto, quase 30%, o que representa mais que o dobro da média dos valores FPR obtida pelos grupos, que é de 13%.

Do número remanescente de redes em cada grupo foi obtida uma única rede resultante via votação majoritária. Para isso a saída contínua de cada rede foi discretizada com um limiar conservador de 0,5, nas quais foi aplicada a votação. A tTabela 8 apresenta um exemplo desse procedimento para três redes. Da rede resultante é possível constatar o regime permanente a partir da quarta instância.

Tabela 8 - Exemplo de aplicação de votação majoritária para obtenção de rede resultante.

	<b>Instância</b>						
	1	2	3	4	5	6	7
<b>Rede 1</b>	0,1	0,1	0,3	0,7	0,75	1	1
<b>Rede 2</b>	0	0,2	0,6	0,9	1	1	1
<b>Rede 3</b>	0,1	0,5	0,3	0,4	0,7	0,8	1
	<b>Limiar = 0,5</b>						
	1	2	3	4	5	6	7
<b>Rede 1</b>	0	0	0	1	1	1	1
<b>Rede 2</b>	0	0	1	1	1	1	1
<b>Rede 3</b>	0	1	0	0	1	1	1
<b>Rede Resultante</b>	0	0	0	1	1	1	1

**Legenda:** Instância – ponto relacionado ao tempo ou aquisição. Limiar— ponto de corte que discretiza a saída contínua.

Fonte: autor.

### **5.3.2. Comparação entre redes especializadas e não especializadas**

A rede representativa obtida pela votação majoritária foi utilizada para aplicação em um conjunto de aproximadamente 45 ensaios de teste, já que alguns desses ensaios foram removidos pois não cumpriam com os requisitos da criação do ROC, que são ensaios totalmente em regime ou que não entravam em regime na janela utilizada (apenas classes 1 ou apenas classes 0, respectivamente).

Assim, uma rede gerou aproximadamente 45 curvas ROC. Nesse conjunto de ensaios foram aplicadas média de limiar, sendo que para cada limiar pontos com TPR e FPR estavam associados. Para comparação direta, as redes não especializadas foram aplicadas ao conjunto de testes de cada grupo, lembrando que os conjuntos de teste de cada um desses grupos eram constituídos de ensaios com a mesma dinâmica.

O intervalo de confiança para FPR e TPR de cada limiar foi definido com nível de confiança de 95%. A junção por FPR foi preferida pois neste tipo de classificação erros de falso positivo são menos desejados que erros de falso negativos dada a natureza da aplicação das redes. Dessa forma, com a junção por FPR as faixas criadas foram mais conservadoras em relação ao FPR que ao TPR. Dessas linhas retirou-se a área sob a curva (juntamente com seu intervalo de confiança), limiar associado a exatidão máxima e faixa de variância. A tabela 9 sumariza os intervalos obtidos para FPR, TPR e exatidão tanto das redes especializadas quanto das redes não especializadas aplicadas ao conjunto de testes dos grupos selecionados.

Tabela 9 - Resultados obtidos das redes especializadas (G\*\*) e não especializadas (G-G\*\*) para FPR, TPR, e exatidão.

GRUPO	FPR	TPR	EXATIDÃO
G1A	0,14 ± 0,06	0,93 ± 0,03	0,90 ± 0,04
G – G1A	0,14 ± 0,05	0,92 ± 0,03	0,90 ± 0,04
G1B	0,22 ± 0,08	0,96 ± 0,02	0,89 ± 0,04
G – G1B	0,24 ± 0,08	0,95 ± 0,02	0,87 ± 0,05
G2A	0,17 ± 0,06	0,94 ± 0,03	0,90 ± 0,04
G – G2A	0,36 ± 0,11	0,96 ± 0,02	0,86 ± 0,05
G2B	0,18 ± 0,07	0,95 ± 0,02	0,91 ± 0,04
G – G2B	0,19 ± 0,08	0,89 ± 0,04	0,87 ± 0,05

**Legenda:** Gxx – redes especializadas treinadas com dinâmica xx. G-Gxx – redes não especializadas aplicadas a ensaios da dinâmica xx. FPR – Taxa de falsos positivos; TPR: – taxa de verdadeiros positivos; EXATIDÃO – taxa de acerto geral.

Fonte: autor.

Já a tabela 10 apresenta os valores de AUC e de faixa de variância. Dessa tabela e da tabela anterior é possível realizar algumas

afirmações. O desempenho das redes especializadas G1A e G1B é bem similar ao obtido pelas redes não especializadas aplicadas a esses conjuntos de testes, exceto pelo intervalo de exatidão, que é melhor nas redes especializadas. O destaque fica para os grupos G2A e G2B, que apresentaram resultados superiores em praticamente todos os quesitos, incluindo exatidão e área sob a curva.

Tabela 10 - Resultados obtidos das redes especializadas (G\*\*) e não especializadas (G – G\*\*) para AUC e faixa de variância.

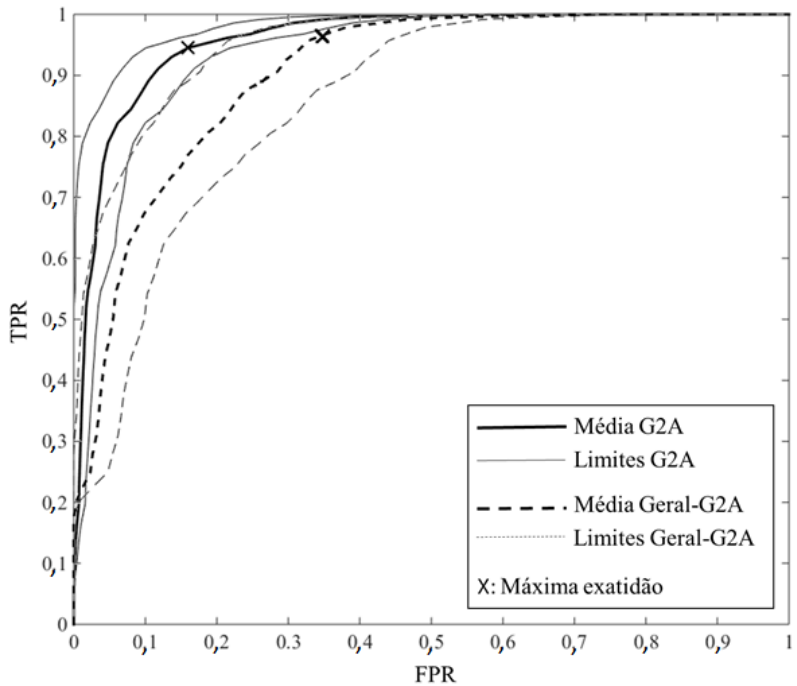
GRUPO	AUC	FAIXA DE VARIÂNCIA
G1A	0,96 ± 0,02	0,03
G – G1A	0,96 ± 0,02	0,04
G1B	0,93 ± 0,03	0,07
G – G1B	0,93 ± 0,03	0,06
G2A	0,95 ± 0,02	0,05
G – G2A	0,89 ± 0,05	0,10
G2B	0,95 ± 0,03	0,05
G – G2B	0,89 ± 0,05	0,10

**Legenda:** Gxx – redes especializadas treinadas com dinâmica xx. G-Gxx – redes não especializadas aplicadas a ensaios da dinâmica xx. AUC – área sob a curva. FAIXA DE VARIÂNCIA – diferença entre os valores superior e inferior das faixas de aceitação.

Fonte: autor.

A figura 33 apresenta os ROCs do grupo G2A e da rede não especializada aplicada aos ensaios de teste desse grupo. Essa figura corrobora a informação das tabelas anteriores. A faixa de variância não especializada é maior, a exatidão é menor e o valor de FPR é maior que os valores da rede especializada.

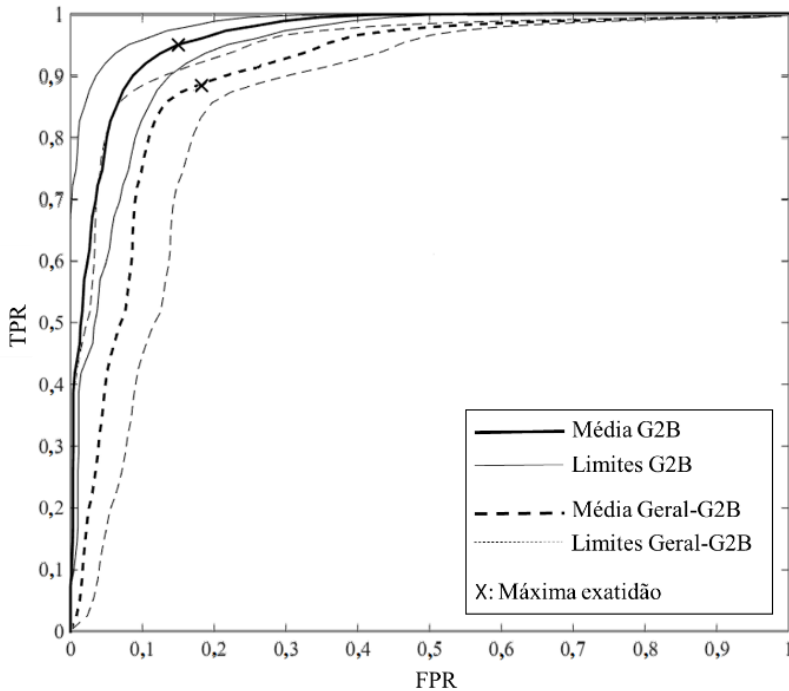
Figura 33 - Faixas de incerteza em ROC para as redes especializadas G2A e para redes não especializadas G-G2A.



Fonte: autor

A figura 34, referente ao grupo G2B, apresenta as ROC da rede especializada e da rede não especializada aplicada aos ensaios de teste desse grupo. Nestes gráficos cabem as mesmas observações realizadas em relação à figura 33.

Figura 34 - Faixas de incerteza em ROC para as redes especializadas G2B e para redes não especializadas G-G2B



Fonte: autor.

### 5.3.3. Desempenho de redes especializadas

As redes especializadas relativas aos grupos G1A e G1B tiveram resultados similares a das redes não especializadas aplicadas aos respectivos conjuntos de teste. Isso já era esperado pois há maior quantidade de ensaios com essa dinâmica no grupo de treinamento das redes não especializadas. Para os grupos G2A e G2B as redes especializadas tiveram desempenho significativamente melhor. A tabela 11 apresenta os parâmetros de desempenho FPR, TPR, exatidão, AUC e faixa de variância. O parâmetro de maior destaque nesta tabela é a taxa de falsos positivos (FPR), particularmente problemático na utilização das redes de inferência de regime permanente. No grupo G2A, FPR chegou a ser maior que o dobro quando as redes não especializadas foram utilizadas. A taxa de verdadeiros positivos (TPR) não apresentou muito

diferença para o grupo G2A, mas para o grupo G2B chegou a ser de 6%. A exatidão também foi melhor para redes especializadas (seus ROC tiveram área sob a curva com diferença de 6% em ambos grupos). Por fim, a faixa de variância é mais estreita, tendo metade da área nos ROC provenientes das redes especializadas. Apesar dos resultados das redes especializadas nos grupos G1A e G1B não serem consideravelmente melhores, sua utilização para os grupos G2A e G2B é comprovadamente melhor que a das redes não especializadas.

Tabela 11 - Parâmetros de desempenho provenientes da aplicação das redes especializadas (G2A e G2B) e redes não especializadas (G-G2A e G-G2B). Médias e incertezas associadas.

	G2A	G-G2A	G2B	G-G2B
FPR	0,17 ± 0,06	0,36 ± 0,11	0,18 ± 0,07	0,19 ± 0,08
TPR	0,94 ± 0,03	0,96 ± 0,02	0,95 ± 0,02	0,89 ± 0,04
EXATIDÃO	0,90 ± 0,04	0,86 ± 0,05	0,91 ± 0,04	0,87 ± 0,05
AUC	0,95 ± 0,02	0,89 ± 0,05	0,95 ± 0,02	0,89 ± 0,05
FAIXA DE VARIÂNCIA	0,05	0,10	0,05	0,10

**Legenda:** Gxx – redes especializadas treinadas com dinâmica xx. G-Gxx – redes não especializadas aplicadas a ensaios da dinâmica xx. FPR – Taxa de falsos positivos; TPR: – taxa de verdadeiros positivos; EXATIDÃO – taxa de acerto geral. AUC – área sob a curva. FAIXA DE VARIÂNCIA – diferença entre os valores superior e inferior das faixas de aceitação.

Fonte: autor.

#### 5.4. AGRUPAMENTO AUTOMÁTICO

O agrupamento manual de ensaios é uma rotina lenta que demanda um especialista dedicado para correta identificação dos grupos. Para aplicações *online* ou para geração de conjunto de treinamento muito extensos isso se torna um problema. Como as redes especializadas forneceram bons resultados, foi desenvolvida uma proposta inicial para automatização do agrupamento manual.

Já existem na literatura algoritmos de clusterização de séries temporais. Contudo, em uma primeira análise eles se mostraram inadequados para o agrupamento automático de dinâmicas pois a grande variedade de comportamentos de curvas presentes nos ensaios vai de

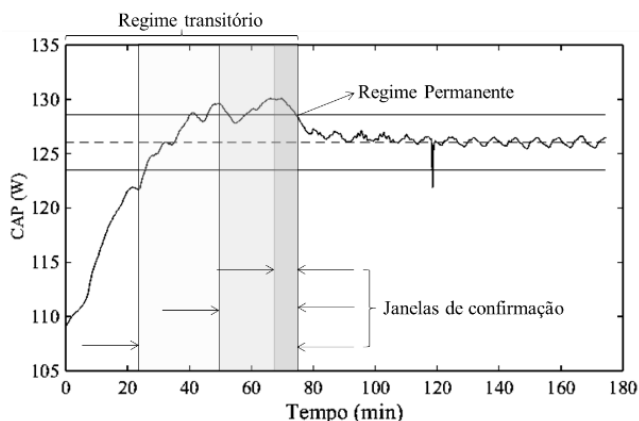
encontro a necessidade que tais algoritmos têm de estipular de antemão o número de grupos (TAN; STEINBACH; KUMAR, 2009).

Contornando esse problema, é apresentada uma metodologia alternativa de agrupamento automático, passível de ser utilizada em futuros trabalhos mas que necessita do ensaio completo para definir a qual grupo ele pertence. Esse método utiliza as seguintes características do ensaio:

1. Momento de entrada em regime permanente.
2. Contagem das aquisições acima ou abaixo da faixa de aceitação durante o regime transitório em um período chamado janela de confirmação.
3. Valor da capacidade inicial.
4. Número de pontos tolerados fora da faixa de aceitação durante o regime permanente.

Dois desses aspectos são variáveis, e quando alterados acarretam em seleções diferentes pois influenciam no momento de entrada em regime obtido pelo algoritmo. O primeiro deles - janela de confirmação (JC) - remete a quantas aquisições o algoritmo observará para confirmar o grupo ao qual o ensaio pertence, sendo uma porcentagem da duração do regime transitório. A figura 35 ilustra algumas opções de JC.

Figura 35 - Diferentes janelas de confirmação para definição do grupo e demonstração de relevância.

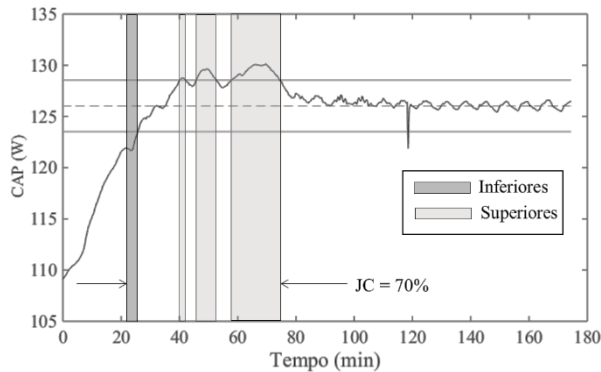


Fonte: autor.



Assim, quando a janela de confirmação é de 100% o algoritmo considerará todas aquisições até o momento de entrada em regime permanente para confirmar o grupo. Uma vez escolhido esse valor, o algoritmo conta quantas aquisições estão acima da faixa de aceitação (ditos valores superiores) e quantas estão abaixo dessa faixa (valores inferiores). Valores dentro da faixa de aceitação são considerados neutros. A figura 36 mostra as áreas de aplicação dessa contagem para uma JC de 70%. Nesse caso o algoritmo confirmaria que o ensaio adentrou a faixa de aceitação por cima, já que a contagem de positivos é maior que a de negativos.

Figura 36 - Exemplo de áreas de contagem para janela de confirmação de 70%.



Fonte: autor.

O quadro 5 resume as regras gerais utilizadas pelo algoritmo para definição dos grupos.  $CAP(0)$  é o valor da capacidade no início do ensaio após desprezar as aquisições com grandes variações (25 aquisições);  $CAP(RP)$  é capacidade no momento de entrada em regime permanente; S, os valores superiores dentro da janela de confirmação; e I os valores inferiores na janela de confirmação.

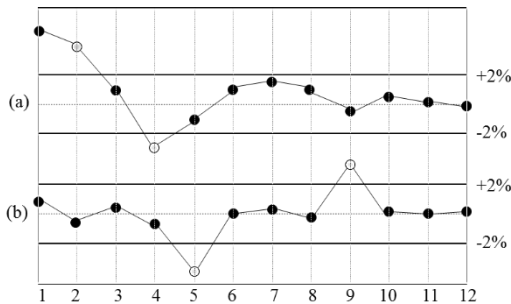
Quadro 4 - Regras para agrupamento automático.

$\Delta = CAP(0) - CAP(RP)$	Contagem	Grupo
$\Delta < 0$	$S > I$	G1A
$\Delta < 0$	$S < I$	G2A
$\Delta > 0$	$S > I$	G1B
$\Delta > 0$	$S < I$	G2B

Fonte: autor.

A segunda característica variável do método é a relevância. Ela define quantos pontos fora da faixa de aceitação serão ignorados mesmo quando o ensaio já se encontra em regime permanente. Com o auxílio da figura 37 é possível visualizar o efeito da relevância em dois casos principais: (a) na determinação do momento de entrada em regime, e (b) ao ignorar *outliers* para assegurar que o ensaio continue em regime permanente. Em ambos os casos, considera-se que a relevância seja de duas aquisições. No primeiro dos gráficos, se não houvesse relevância, o ensaio seria atribuído ao grupo 2, já que a última aquisição está abaixo da faixa de aceitação (ponto 4). Isso não corresponde à tendência geral que indica que o grupo correto seria 1 (entrada pelo ponto 1). No segundo gráfico o ensaio já vem em regime permanente e a relevância serve para garantir que aquisições aleatórias (pontos 5 e 9), independente do motivo, não reiniciem a contagem de entrada em regime.

Figura 37 - Aplicação de dois casos de relevância (pontos em cinza claro): (a) influência na determinação do grupo; (b) permanência em regime permanente.



Fonte: autor.

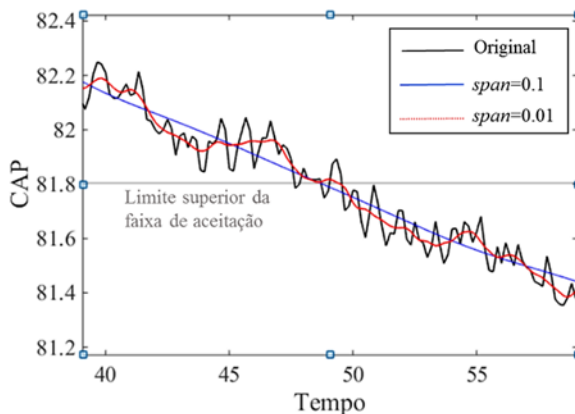
Entretanto, alguns ensaios possuem excessivos *outliers* ou oscilam muito nas aquisições próximas à entrada do regime. Para contornar esses problemas, cada ensaio deveria possuir sua própria relevância, o que é inviável. Isso tornar-se-ia desnecessário se as curvas apresentadas ao algoritmo tivessem esses *outliers* removidos ou as oscilações atenuadas.

Uma maneira de realizar essa tarefa é por meio de suavização (*smoothing*). Alguns dos métodos para isso são médias móveis, filtros ou regressão local. O método que se mostrou útil ao problema foi um de regressão local chamado LOWESS (*Locally Weighted Scatterplot Smoothing*), já implementado no software utilizado no laboratório

(Matlab) e de utilização simples. Este algoritmo aplica uma regressão local por mínimos quadrados ponderados e um modelo polinomial de primeira ordem. A suavização requer a estipulação do parâmetro *span*, que determina a proporção de pontos a serem utilizados na regressão (CLEVELAND, 1979).

A figura 38 apresenta a utilização da suavização com diferentes *spans* e sua influência na região que antecede a entrada na faixa de aceitação. Quanto maior o *span* maior a relevância da regressão frente a variações nos pontos. Buscou-se um *span* que diminuísse a influência de *outliers* e que ainda fornecesse uma correlação de Pearson grande o suficiente (maior que 0,95) entre a curva original e a suavizada

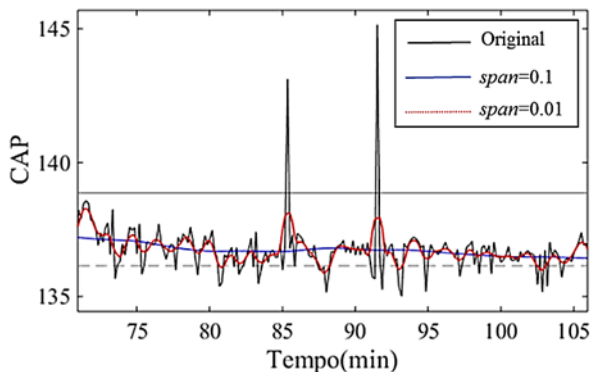
Figura 38 - Impacto da suavização com diferentes *spans* na região que antecede a entrada em regime.



Fonte: autor.

Na figura 39 nota-se o comportamento das curvas suavizadas na ocorrência de um *outlier*. Enquanto com um *span* de 0,1 não só os outliers são ignorados mas também todas as oscilações, com um *span* de 0,01 é possível manter boa parte das oscilações da curva e mesmo assim ignorar os *outliers*. Um *span* de 0,01 mostrou-se suficiente para uma análise que mantivesse a tendência dos ensaios, a alta correlação entre as curvas e minimiza-se o impacto de outliers.

Figura 39 - Impacto da suavização com diferentes *spans* na permanência em regime.



Fonte: autor.

Partindo do pressuposto de que o agrupamento manual é livre de erros, o algoritmo de agrupamento automático foi aplicado aos ensaios agrupados manualmente, logo, sabia-se de antemão o resultado correto dessa comparação. Na tabela 12 estão as diferentes combinações de parâmetros e os acertos correspondentes. A melhor combinação de parâmetros foi aquela com 7 aquisições de relevância e 10% de janela de confirmação, apresentando um acerto de 92%.

Tabela 12 - Acerto do agrupamento automático para diferentes valores de relevância e janela de confirmação.

	Tolerância; Janela de Confirmação (%)						
	3;10	3;30	3;50	7;10	7;30	11;10	11;30
	<b>% de acerto</b>						
G1A	89	90	90	<b>89</b>	89	87	88
G1B	99	97	96	<b>99</b>	97	98	96
G2A	84	86	85	<b>86</b>	85	83	82
G2B	91	81	72	<b>93</b>	81	93	81
Média	91	88	86	<b>92</b>	88	90	86

Fonte: autor.

O método automático é fundamental para adequação do agrupamento de ensaios (logo, a utilização das redes especialistas) à ferramenta de inferência já existente. Contudo, ele é realizado com valores que não são conhecidos no início do ensaio. No momento inicial alguns atributos estão disponíveis, como valores finais de pressão de sucção e de descarga (devido a condição de teste, que é configurada desde o princípio), modelo do compressor e, possivelmente, algumas medições iniciais após os ajustes de controle. Entretanto, como já foi dito o agrupamento automático depende da curva completa de capacidade, que naturalmente só está disponível ao final do ensaio. Dessa maneira é impossível definir a qual grupo pertence o ensaio (ou o compressor em teste) apenas pelo agrupamento automático, sendo necessário algum outro procedimento que o complemente ou que o utilize como complemento. Alguns desses procedimentos são comentados e sugeridos como trabalhos futuros.



## 6. CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo o treinamento de redes de inferência especializadas. Isso foi alcançado através do agrupamento de ensaios de desempenho de compressores de refrigeração. Além disso, a partir desse agrupamento também foi possível observar como diversas variáveis se comportavam em cada grupo viabilizando a identificação de relações, as quais foram embasadas por métodos estatísticos. Como o agrupamento manual é lento e necessita de um operador dedicado ao procedimento, foi proposto também um método automatizado. Nas conclusões a seguir serão abordados os resultados obtidos e mencionados os possíveis trabalhos futuros.

### 6.1. CONCLUSÕES

Neste trabalho foram analisados ensaios de desempenho de compressores de refrigeração provenientes da rotina de uma empresa fabricante de compressores. Esses ensaios foram selecionados manualmente e agrupados de acordo com algumas regras envolvendo valores de capacidade e momento de entrada em regime com dois objetivos: buscar diferenças entre as dinâmicas e treinar redes de inferência de regime permanente ditas especializadas. Como o agrupamento manual é lento e pode ser um problema em aplicações *online* ou para um grande número de ensaios, foi criado um método de agrupamento automático que busca reproduzir as regras utilizadas no agrupamento manual. Esse método utilizou ferramentas de suavização de curvas, relevância de outliers e contagem de pontos em janelas de confirmação para decidir a qual grupo o ensaio completo pertenceria. A melhor combinação desses parâmetros forneceu taxa de acerto média de 92%.

Os ensaios fornecidos pela empresa foram provenientes de duas plantas. Para análises mais criteriosas, esses ensaios foram restritos apenas à uma das plantas (devido à homogeneidade referente ao sistema de controle e instrumentação) e também reduzidos à condição de *checkpoint*. Testes de hipótese foram realizados em distribuições não paramétricas das seguintes variáveis: duração dos ensaios, momento de entrada em regime permanente, variação de temperatura de corpo, variação da pressão de sucção e variação da pressão de descarga. O agrupamento manual evidenciou o comportamento de dois grupos dentre os quatro obtidos, que foram chamados no decorrer do trabalho de G1A e G2B. Ensaios desse grupo são aqueles que possuem capacidade inicial

superior àquela relativa ao momento de entrada em regime permanente. Tais grupo possuem ensaios com duração e momento de entrada em regime permanente consideravelmente inferiores (regime transitório mais rápido). A variação de temperatura de corpo e variação de pressão de sucção também foram consideravelmente inferiores, o que corrobora a última observação realizada. No entanto a variação da pressão de descarga do grupo G2B mostrou-se superior a dos demais. Esse último comportamento, apesar de não estar associado a um ensaio ideal só é problemático caso a pressão de sucção não esteja bem controlada. As características desses grupos fazem deles as dinâmicas ideais que deveriam ser induzidas pelas bancadas para obtenção de vantagens relacionadas ao tempo de execução.

Em outro tópico, as redes especializadas treinadas com ensaios agrupados apresentaram ótimos resultados frente às redes não especializadas. Essas redes foram treinadas com 236 ensaios de cada grupo, sendo que este número foi dividido em treinamento, validação, criação de comitês e testes. A criação de comitês foi realizada pelo método de fusão, sendo o número de redes resultante aquele com menor taxa de falsos positivos. A fusão dos resultados foi realizada por votação majoritária e após aplicação nos conjuntos de testes foram construídas curvas ROC. A comparação foi realizada entre os resultados da aplicação das redes especializadas em conjuntos de testes que possuíam a própria dinâmica presente nos conjuntos de treinamento e as redes não especializadas aplicadas a esses mesmos conjuntos. Nos grupos G1A e G1B, o desempenho das redes especializadas foi tão bom quanto a das redes não especializadas. Esse era um resultado já esperado, pois a proporção de dinâmicas pertencentes a esses grupos é muito mais significativa em conjunto de treinamento de ensaios não agrupados (cerca de 80%). O destaque ocorreu nos grupos G2A e G2B. O desempenho das redes especializadas treinadas com ensaios provenientes dessas dinâmicas superou consideravelmente o desempenho das não especializadas. Estas últimas forneceram não só valores mais baixos de taxa de acerto (TPR) e parâmetros relacionados às curvas de ROC, como também tiveram taxas de falso positivo (FPR) superiores a das redes especializadas em até 19%.

O bom desempenho de redes especializadas em ensaios dos grupos G2A e G2B aliado ao bom comportamento dos ensaios pertencentes ao grupo G2B obtidos durante a análise estatística favorecem a indução desta dinâmica nos controles das bancadas de ensaio de compressores e a continuação dos estudos acerca das características que ocasionam tais dinâmicas.



Por fim, resultados aqui obtidos apontaram um avanço na utilização de redes de inferência e enfatizaram a necessidade do estudo das variáveis que influenciam tais redes. No que diz respeito à utilização do que foi desenvolvido neste trabalho em situações reais de fábrica ainda existe um empecilho. A aplicação das redes especializadas depende do conhecimento a priori do grupo a qual pertence o compressor em teste, o que ainda não é possível. Alguns métodos para contornar esse problema são propostos a seguir.

## 6.2. SUGESTÃO DE TRABALHOS FUTUROS

A análise sobre as dinâmicas que definem os grupos considerou apenas algumas variáveis, que não são os únicos fatores que ocasionam tais comportamentos. Entretanto, como um dos grupos (G2B) se sobressaiu na duração e entrada em regime permanente, seria interessante um trabalho futuro que buscasse nas variáveis iniciais do ensaio características que ocasionam essa dinâmica em específico, ao relacioná-las, inclusive, com as sequências de realização. Essa prática não só complementarria como viabilizaria o agrupamento automático e o uso das redes especializadas.

Quanto ao agrupamento automático ele não é por si só uma solução completa, mas sim uma ferramenta para facilitar um procedimento futuro - o de definir antes do início ou nos momentos iniciais do teste a qual grupo pertence o ensaio em questão. Para isso, uma sugestão é a implementação desse algoritmo juntamente com um classificador de aprendizagem supervisionada, onde os rótulos desse novo classificador seriam provenientes do agrupamento automático.

Outra possibilidade seria a utilização das saídas do agrupador automático como um parâmetro de entrada em um clusterizador. Esse clusterizados agruparia o comportamento inicial dos ensaios, fornecendo o grupo correspondente. Esta é uma sugestão abrangente, pois pode ser utilizada não só em ensaios de desempenho como também em outros testes realizados em compressores que façam uso de séries temporais e que tenham grande variedade de condições iniciais.



## REFERÊNCIAS

- ALPANA; MOHAPATRA, S. Machine learning approach for automated coal characterization using scanned electron microscopic images. **Computers in Industry**, v. 75, p. 35–45, 2016.
- ANDERSON, J. A.; ROSENFELD, E. **Neurocomputing: Foundations of Research**. Cambridge: MIT Press, 1988.
- ARBIB, M. A. **The Handbook of Brain Theory and Neural Networks**. MIT Press, 1989. v. 53
- ASHRAE STANDARD. **ANSI/ASHRAE 23-2005: Methods of testing for rating positive displacement refrigerant compressors and condensing units**. Atlanta, EUA, 2005.
- BARBOSA, F. Q. **Redução das incertezas nas medições de grandezas elétricas em uma bancada de ensaios de desempenho de compressores**. Florianópolis: Universidade Federal de Santa Catarina, 2006.
- BRADLEY, A. P. **The use of the area under the ROC curve in the evaluation of machine learning algorithms**. Pattern Recognition, v. 30, n. 7, p. 1145–1159, 1997.
- BROWN, G. **Diversity in Neural Network Ensembles**. Birmingham: University of Birmingham, 2004.
- CASELLA, A. V. **Ferramentas fuzzy-bayesianas para identificação do instante de entrada no regime permanente em ensaios de desempenho de compressores**. Florianópolis: Universidade Federal de Santa Catarina, 2008.
- CHAWLA, N. V. **Data Mining for Imbalanced Datasets: An Overview**. Data Mining and Knowledge Discovery Handbook. Boston, MA: Springer US, 2005. p. 853–867.
- CHOU, J. S.; HSU, Y. C.; LIN, L. T. **Smart meter monitoring and data mining techniques for predicting refrigeration system performance**. Expert Systems with Applications, v. 41, n. 5, p. 2144–2156, 2014.
- CLEVELAND, W. S. **Robust Locally Weighted Regression and Smoothing Scatterplots**. Journal of the American Statistical Association, v. 74, n. 368, p. 829–836, 1979.
- DUNN, O. J. **Multiple Comparison among means**. Journal of the

American Statistical Association, v. 59, n. 293, p. 52–64, 1961.

EFRON, B. **Nonparametric Estimates of Standard Error: The Jackknife, the Bootstrap and Other Methods**. *Biometrika*, v. 68, n. 3, p. 589–599, 1981.

EFRON, B. **Better Bootstrap Confidence Intervals** Stanford University, 1984.

EFRON, B.; TIBSHIRANI, R. J. **An Introduction to the Bootstrap**. Chapman, 1993.

EGAN, J. P. **Signal Detection Theory and ROC Analysis** New York Academic Press, 1975.

ENGELBRECHT, A. P. **Computational Intelligence: An introduction**. 2. ed. South Africa: University of Pretoria, 2007.

FAWCETT, T. **An introduction to ROC analysis**. *Pattern Recognition Letters*, v. 27, p. 861–874, 2004.

FLACH, P. A. **The Geometry of ROC Space: Understanding Machine Learning Metrics through ROC Isometrics**. *Proceedings of the Twentieth International Conference on Machine Learning*, p. 194–201, 2003.

FLESCHE, R. C. C. **Desenvolvimento e instalação de recursos de automação aplicados a ensaios de compressores de refrigeração da Embraco**. Florianópolis: Universidade Federal de Santa Catarina, 2006.

FLESCHE, R. C. C. **Contribuições ao controle de sistemas monovariáveis e multivariáveis com atraso de transporte**. Universidade Federal de Santa Catarina, 2008.

FLESCHE, R. C. C.; J. E. NORMEY-RICO. **Modeling, identification and control of a calorimeter used for performance evaluation of refrigerant compressors**. *Control Engineering Practice*, v. 18, p. 254–261, 2010.

FLESCHE, S. A. R. C. **Contribuições ao controle de sistemas monovariáveis e multivariáveis com atraso de transporte**. Universidade Federal de Santa Catarina, 2012.

FONTES, C. H.; PEREIRA, O. **Pattern recognition in multivariate time series - A case study applied to fault detection in a gas turbine**. *Engineering Applications of Artificial Intelligence*, v. 49, p. 10–18, 2016.

- GALAR, M. **A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches.** IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, v. 42, n. 4, p. 463–484, 2012.
- GEMAN, S.; BIENENSTOCK, E.; DOURSAT, R. **Neural networks and the bias/variance dilemma.** Neural Computation, v. 4, 1992.
- GERLEIN, E. A. **Evaluating machine learning classification for financial trading: An empirical approach.** Expert Systems with Applications, v. 54, p. 193–207, 2016.
- GOEBEL, K.; YAN, W. **Choosing Classifiers for Decision Fusion.** Proceedings of the Seventh International Conference on Information Fusion, p. 563–568, 2004.
- GUYON, I.; ELISSEEFF, A. **An Introduction to Variable and Feature Selection.** Journal of Machine Learning Research (JMLR), v. 3, n. 3, p. 1157–1182, 2003.
- HAN, J.; KAMBER, M. **Data mining: concepts and techniques.** 2. ed. The Morgan Kaufmann Series in Data Management Systems, 2006.
- HAYKIN, S. **Princípios e práticas.** 2. ed. Porto Alegre: Bookman, 2001.
- HENKLEIN, P. **Análise de ensaios de desempenho de compressores herméticos durante o transitório.** Florianópolis: Universidade Federal de Santa Catarina, 2006.
- HOCHBERG, Y.; TAHMANE, A. C. **Multiple comparison procedures.** John Wiley & Sons, 1987.
- INMETRO. **Vocabulário Internacional de Termos de Metrologia Legal - VIM.**
- ISO. **ISO 917: Testing of refrigerant compressors** Switzerland International Organization for Standardization, 1989.
- KANNATEY-ASIBU, E.; YUM, J.; KIM, T. H. **Monitoring tool wear using classifier fusion.** Mechanical Systems and Signal Processing, v. 85, p. 651–661, 2016.
- KRUSKAL, W H; WALLIS, W. A. **Use of ranks in one-criterion variance analysis.** Journal of the American Statistical Association, v. 47, n. 260, p. 583–621, 1952.
- KUNCHEVA, L. I.; WHITAKER, C. J. **Measures of diversity in**

**classifier ensembles and their relationship with the ensemble accuracy.** Machine Learning, v. 51, n. 2, p. 181–207, 2003.

LIMA, L. C. A. **Controle da alimentação elétrica para a redução de tempo demandado por ensaios de desempenho de compressores e avaliação da influência da qualidade de energia elétrica nos resultados dos ensaios.** Florianópolis: Universidade Federal de Santa Catarina, 2010a.

LIMA, V. R. **Desenvolvimento e avaliação de sistema neural para redução de tempo de ensaio de desempenho de compressores.** Florianópolis: Universidade Federal de Santa Catarina, 2010b.

MACSKASSY, S. A.; PROVOST, F. **Confidence bands for ROC curves: methods and an empirical study.** Workshop on ROC Analysis in AI, 2004.

MARTINS, E. F.; FLESCHE, C. A.; FLESCHE, R. C. C.; BORGES, M. R. **Conceptions of a test bench to generate known and controlled conditions of refrigerant mass flow.** ISA Transactions, v. 50, n. 3, p. 513-520, 2011.

MASSEY, F. J. **The Kolmogoroc-Smirnov Test for Goodness of Fit.** Journal of the American Statistical Association, v. 46, n. 253, p. 68–78, 1951.

MATHWORKS. **Statistics and machine learning toolbox: User's Guide,** 2016.

MCCULLOCH, W. S.; PITZ, W. **A logical calculus of the ideas immanent in nervous activity.** Mathematical Biophysics, v. 5, 1943.

MICHALSKI, R. S.; CARBONELL, J. G.; MITCHELL, T. M. **MACHINE LEARNING: An Artificial Intelligence Approach.** Morgan Kaufmann Publishers, 1983.

MONTGOMERY, D. C. **Design and Analysis of Experiments.** 5. ed. John Wiley & Sons, 2000.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied Statistics and Probability for Engineers.** 6. ed. John Wiley & Sons, 2013.

NASCIMENTO, A. **Desenvolvimento de ferramentas baseadas em redes neurais para prognósticos em ensaios de desempenho de compressores herméticos.** Florianópolis: Universidade Federal de Santa Catarina, 2015.

NOCEDAL, J.; WRIGHT, S, J. **Numerical Optimization**. 2. ed. Springer, 2006.

PAPADOPOULOS, G.; EDWARDS, P. J.; MURRAY, A. F. **Confidence estimation methods for neural networks: a practical comparison**. IEEE Transactions on Neural Networks, v. 12, n. 6, 2001.

PENZ, C. A. **Procedimentos para prover confiabilidade ao uso de inteligência artificial em ensaios de desempenho de compressores herméticos de refrigeração**. Florianópolis: Universidade Federal de Santa Catarina, 2011.

PETROVIC, C. A. **Proposta de arquitetura do sistema de automação aplicado a ensaios de desempenho de compressores herméticos de refrigeração**. Florianópolis: Universidade Federal de Santa Catarina, 2007.

POLETTI, E. L. **Aprimoramento de uma bancada de ensaios de desempenho de compressores herméticos visando reduzir incertezas de medição**. Florianópolis: Universidade Federal de Santa Catarina, 2006.

ROSENBLATT, F. **The perceptron: A probabilistic model for information storage and organization in the brain**. Psychological Review, v. 65, n. 6, p. 386–408, 1958.

RUSSEL, S.; NORVIG, P. **Artificial intelligence: a modern approach**. 2. ed. New Jersey: Prentice hall, 2003.

RUSSO, A. P. **Neural networks for signal processing**. IEEE Conference on Neural Networks for Ocean Engineering. Anais... Washington: 1991

SAMMUT, C.; WEBB, G. **Confusion matrix**, 2010. (Nota técnica).

SCUSSEL, J. N. **Proposta de ações para reduzir o tempo demandado por ensaio de desempenho de compressores herméticos**. Florianópolis: Universidade Federal de Santa Catarina, 2006.

SILVEIRA, A. N. **Diminuição do tempo de transitório em ensaios de desempenho de compressores utilizando injeção de corrente contínua**. Florianópolis: Universidade Federal de Santa Catarina, 2010.

SPACKMAN, K. A. **Signal Detection Theory: Valuable Tools for Evaluating Inductive Learning**. Proceedings of the Sixth International Workshop on Machine Learning. Anais. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1989.

STEINBACH, K. **Identificação do instante de entrada em condições de regime permanente em ensaios de desempenho de compressores.** Florianópolis: Universidade Federal de Santa Catarina, 2008.

SWETS, J. **Measuring the accuracy of diagnostic systems.** *Science*, v. 240, p. 1285–1293, 1988.

TAN, P.; STEINBACH, M.; KUMAR, V. **Introdução ao DATA MINING: mineração de dados.** Rio de Janeiro: Editora Ciência Moderna, 2009.

TIXIER, A. J. P. **Application of machine learning to construction injury prediction.** *Automation in Construction*, v. 69, p. 102–114, 2016.

TRABELSI, A.; ELOUEDI, Z.; LEFEVRE, E. **Comparing dependent combination rules under the belief classifier fusion framework.** *Soft Computing*, 2016.

ULAŞ, A. **Incremental construction of classifier and discriminant ensembles.** *Information Sciences*, v. 179, n. 9, p. 1298–1318, 2009.



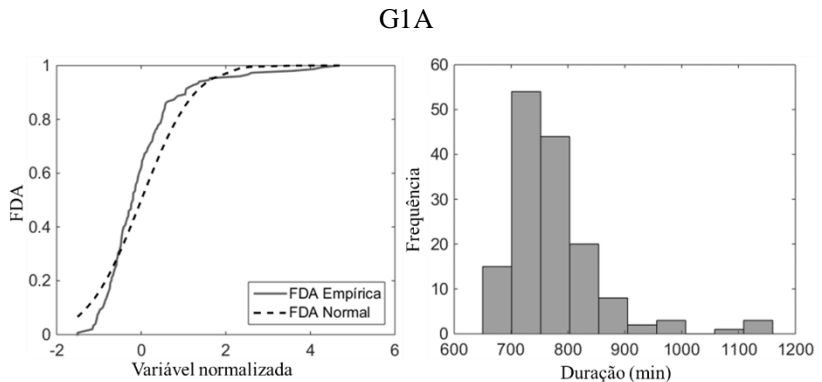
## APÊNDICE A - TESTES DE NORMALIDADE

Quadro 5 - Testes de normalidade por Kolmogorov-Smirnov.  $H_0$ =ambas amostras pertencem à uma distribuição normal. Rejeita  $H_0$  se  $K < CV$ .  $\alpha=0,05$ .

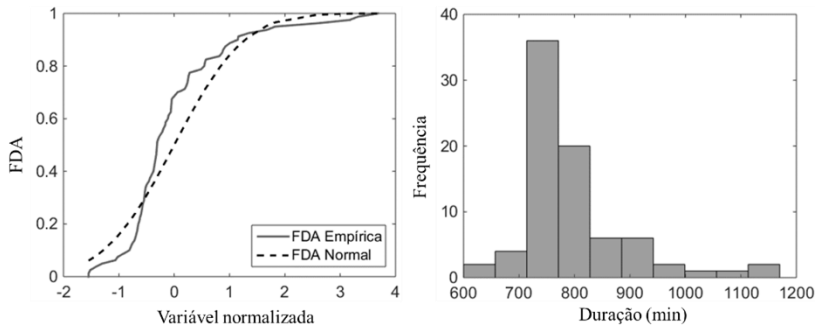
		G1A	G1B	G2A	G2B
Estatística de Teste (K)	Valores Críticos (CV)	0,11	0,15	0,2	0,11
	Duração	1	1	1	1
	Entrada em RP	1	1	1	1
	Delta T	0,98	1	0,98	0,98
	Delta Ps	0,5	0,5	0,5	0,5
	Delta Pd	0,89	0,76	0,77	0,92

Fonte: autor.

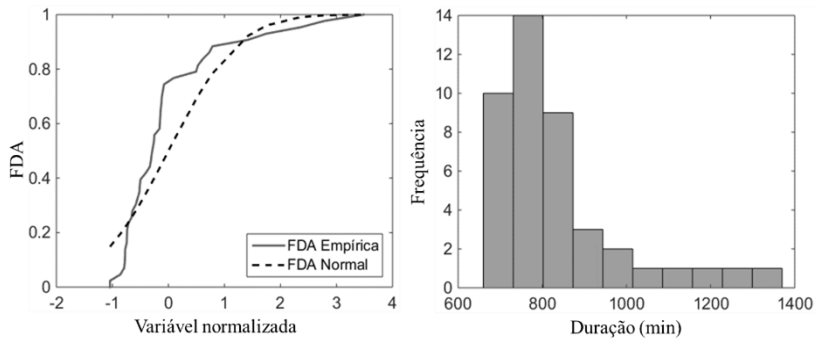
Figura 40 - FDA dos testes de Kolmogoro-Smirnov e histogramas das durações de ensaios.



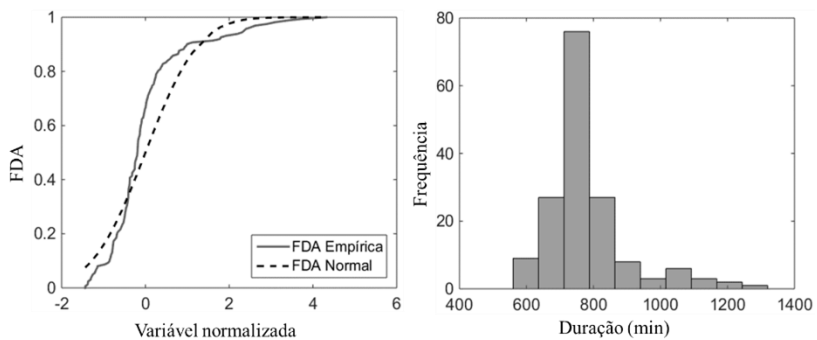
G1B



G2A



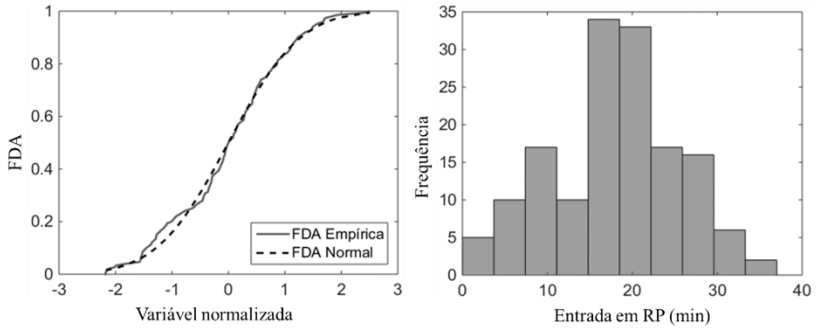
G2B



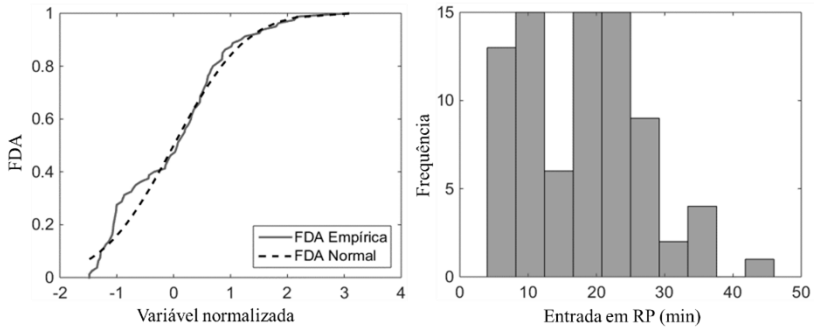
Fonte: autor.

Figura 41 - FDA dos testes de Kolmogoro-Smirnov e histogramas dos momentos de entrada em regime permanente dos ensaios.

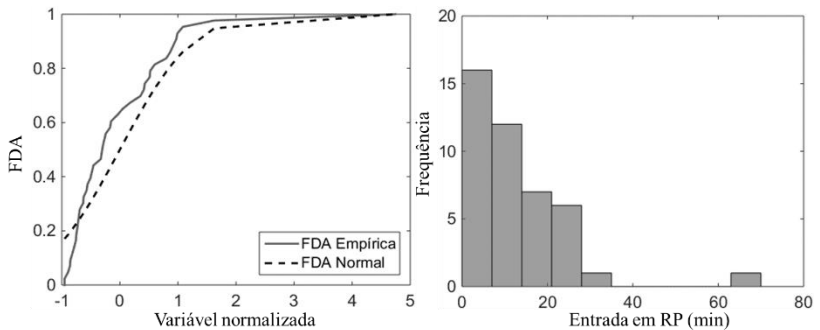
## G1A



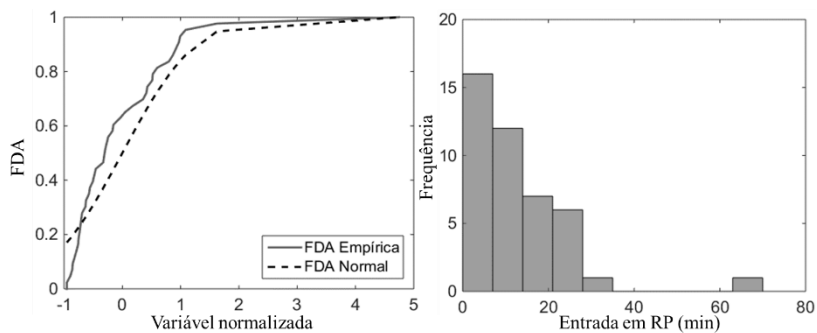
## G1B



## G2A



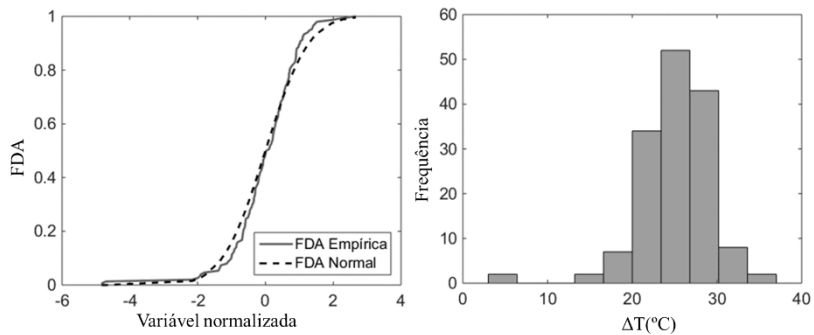
## G2B



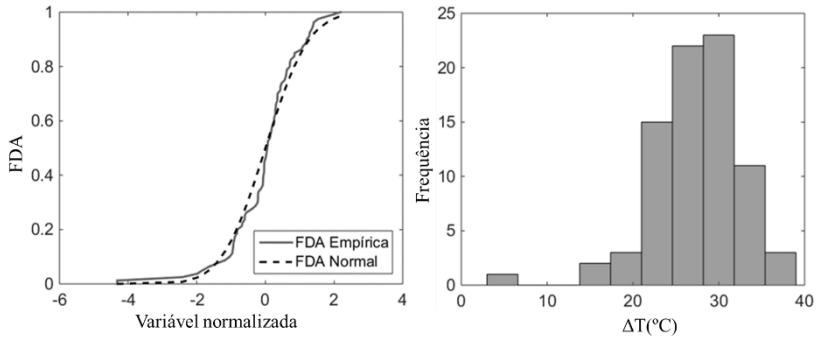
Fonte: autor.

Figura 42 - FDA dos testes de Kolmogoro-Smirnov e histogramas da variação da temperatura de corpo.

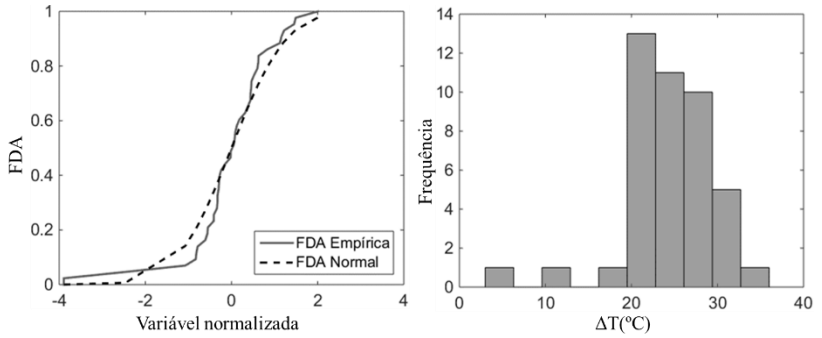
## G1A



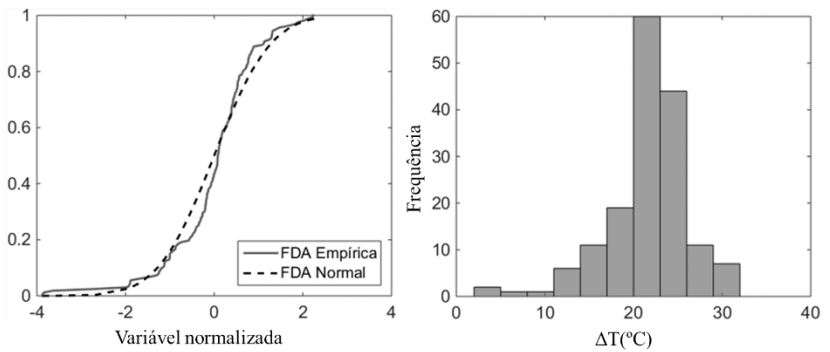
G1B



G2A



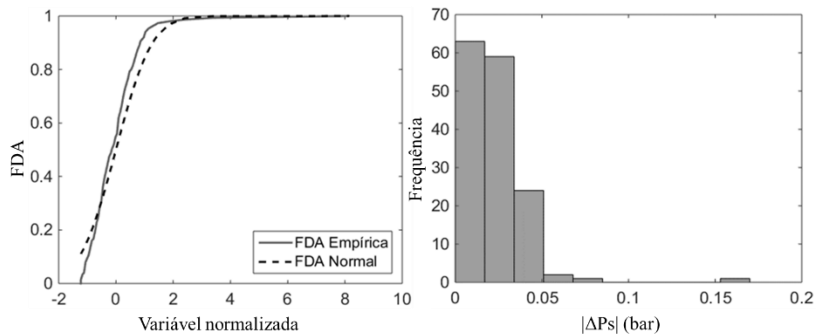
G2B



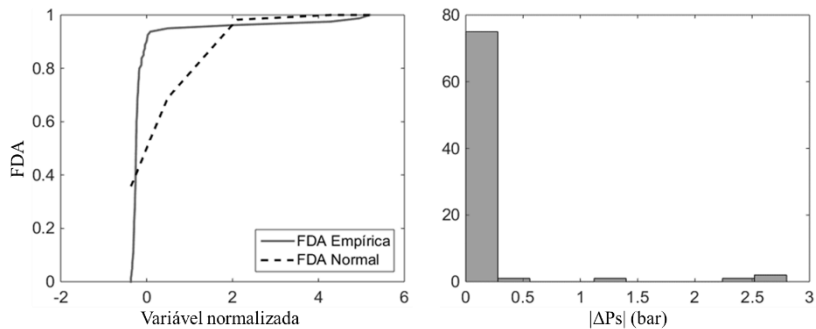
Fonte: autor.

Figura 43 - FDA dos testes de Kolmogoro-Smirnov e histogramas da variação absoluta de pressão de sucção.

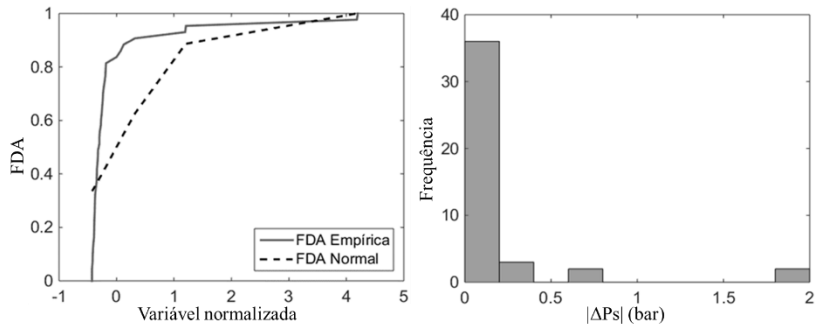
## G1A



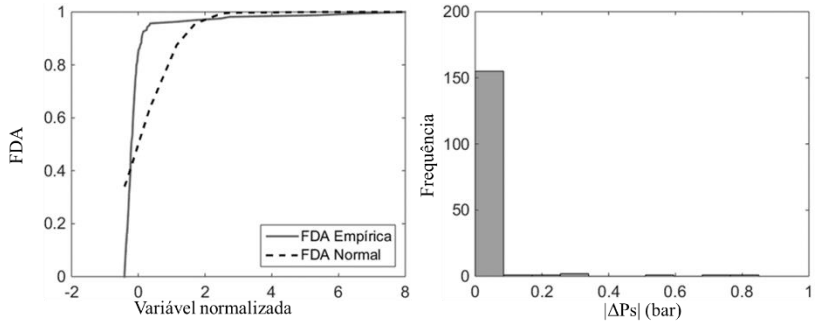
## G1B



## G2A



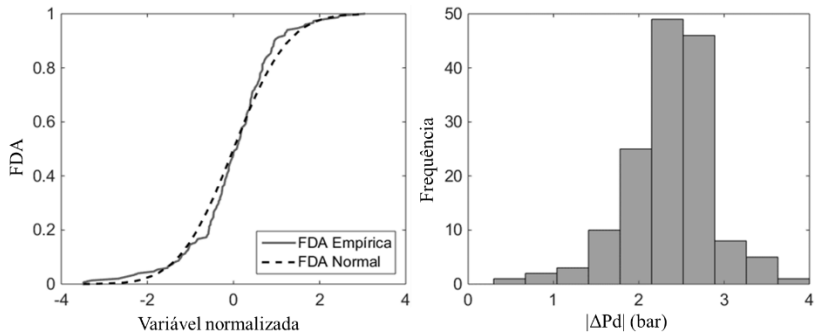
## G2B



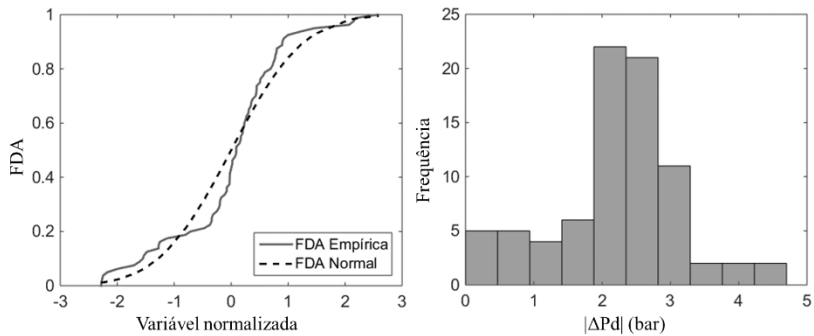
Fonte: autor.

Figura 44 - FDA dos testes de Kolmogoro-Smirnov e histogramas da variação absoluta de pressão de descarga.

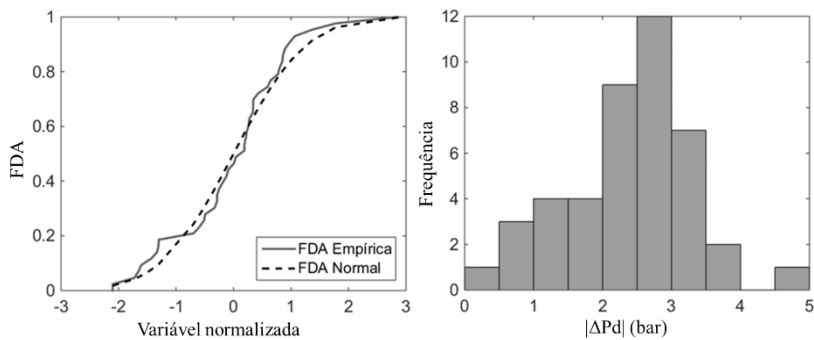
## G1A



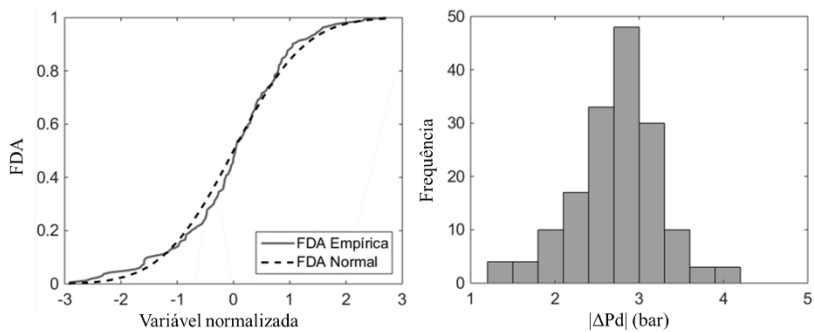
G1B



G2A



G2B



Fonte: autor.



## APÊNDICE B – RESULTADOS DAS ANÁLISES DE VARIÂNCIA POR KRUSKAL-WALLIS

Quadro 6 - Resultados da análise de variância.  $H_0$  de que não há diferença entre tratamentos.

	Fonte	SQ	GL	MQ	$\chi^2$	valor $p > \chi^2$
<b>Duração</b>	Grupo	190822	3	63607	12,07	0,01
	Erro	6668030	431	15471	-	-
	Total	6858853	434	-	-	-
<b>Entrada em RP</b>	Grupo	1149355	3	383118	72,72	0
	Erro	5709755	431	13247	-	-
	Total	6859111	434	-	-	-
<b><math>\Delta T</math></b>	Grupo	1556016	3	518672	98,45	0
	Erro	5303353	431	12304	-	-
	Total	6859370	434	-	-	-
<b><math> \Delta P_s </math></b>	Grupo	2103200	3	701066	134,9 6	0
	Erro	4613210	428	10778	-	-
	Total	6716410	431	-	-	-
<b><math> \Delta P_d </math></b>	Grupo	849715	3	283238	53,76	0
	Erro	6009639	431	13943	-	-
	Total	6859355	434	-	-	-

Fonte: autor.

