

Universidade Federal de Itajubá  
Programa de Pós Graduação em Engenharia Elétrica

**AVALIAÇÃO DA PROBABILIDADE DE  
OCORRÊNCIA DE QUEDA DE ESTRUTURAS  
METÁLICAS DE LINHAS DE TRANSMISSÃO  
USANDO REGRESSÃO LOGÍSTICA E REDES  
NEURAIS ARTIFICIAIS COMBINADAS COM DOE**

**Silvio Antonio Bueno Salgado**

Itajubá, Junho de 2015

Universidade Federal de Itajubá

Programa de Pós Graduação em Engenharia Elétrica

**Silvio Antonio Bueno Salgado**

**AVALIAÇÃO DA PROBABILIDADE DE OCORRÊNCIA DE QUEDA DE  
ESTRUTURAS METÁLICAS DE LINHAS DE TRANSMISSÃO USANDO  
REGRESSÃO LOGÍSTICA E REDES NEURAS ARTIFICIAIS COMBINADAS  
COM DOE**

Tese submetida ao Programa de Pós Graduação em Engenharia  
Elétrica como parte dos requisitos para obtenção do Título de Doutor  
em Ciências em Engenharia Elétrica

Área de concentração: Sistemas Elétricos de Potência

**Orientadores: Prof. Dr. Antonio Carlos Zambroni de Souza  
Prof. Dr. Pedro Paulo Balestrassi**

Itajubá, Junho de 2015

# Dedicatória

*A minha mãe, Maria José Bueno*

# Agradecimentos

- A Deus, sem o qual nada disso seria possível.
- Aos meus orientadores, Professor Doutor Antonio Carlos Zambroni de Souza e Professor Doutor Pedro Paulo Balestrassi, pela paciência e pela amizade que me concederam, e pela excelente orientação neste tempo que trabalhamos juntos.
- Aos Professores e colegas do curso que estiveram comigo nesta caminhada.
- Ao meu colega Alexandre Corrêa Grassi Bissacot, que muito me ajudou.
- A minha mãe Maria José pelo incentivo e apoio incondicional que sempre me deu.

# Resumo

Nesta tese, avalia-se a probabilidade de ocorrência de queda de estruturas metálicas de linhas de transmissão através de duas técnicas: regressão logística e RNAs combinadas com DOE. A primeira, pode ser considerada como técnica padrão e já é consagrada como uma poderosa ferramenta de auxílio na tomada de decisão. A segunda tem sido utilizada em profusão para problemas similares e apresenta-se como excelente alternativa para a regressão logística, pois assume menos restrições de uso, apesar de exigir uma simulação computacional mais complexa. A metodologia estatística de Delineamento de Experimentos (*Design of Experiments* - DOE) é proposta para determinar os melhores parâmetros de uma RNA para o problema de queda de estruturas. Ao invés de se utilizar a técnica mais comum de treinamento de uma RNA, ou seja, tentativa e erro, DOE é proposto como uma metodologia para encontrar a melhor RNA.

**Palavras-Chave:** Redes Neurais Artificiais, Delineamento de Experimentos, Regressão Logística, Curva Roc, Linhas de Transmissão, Quedas de Estruturas.

# Resumo Gráfico

A figura 1 mostra uma representação esquemática do problema estudado nesta tese e também os passos adotados para sua solução.

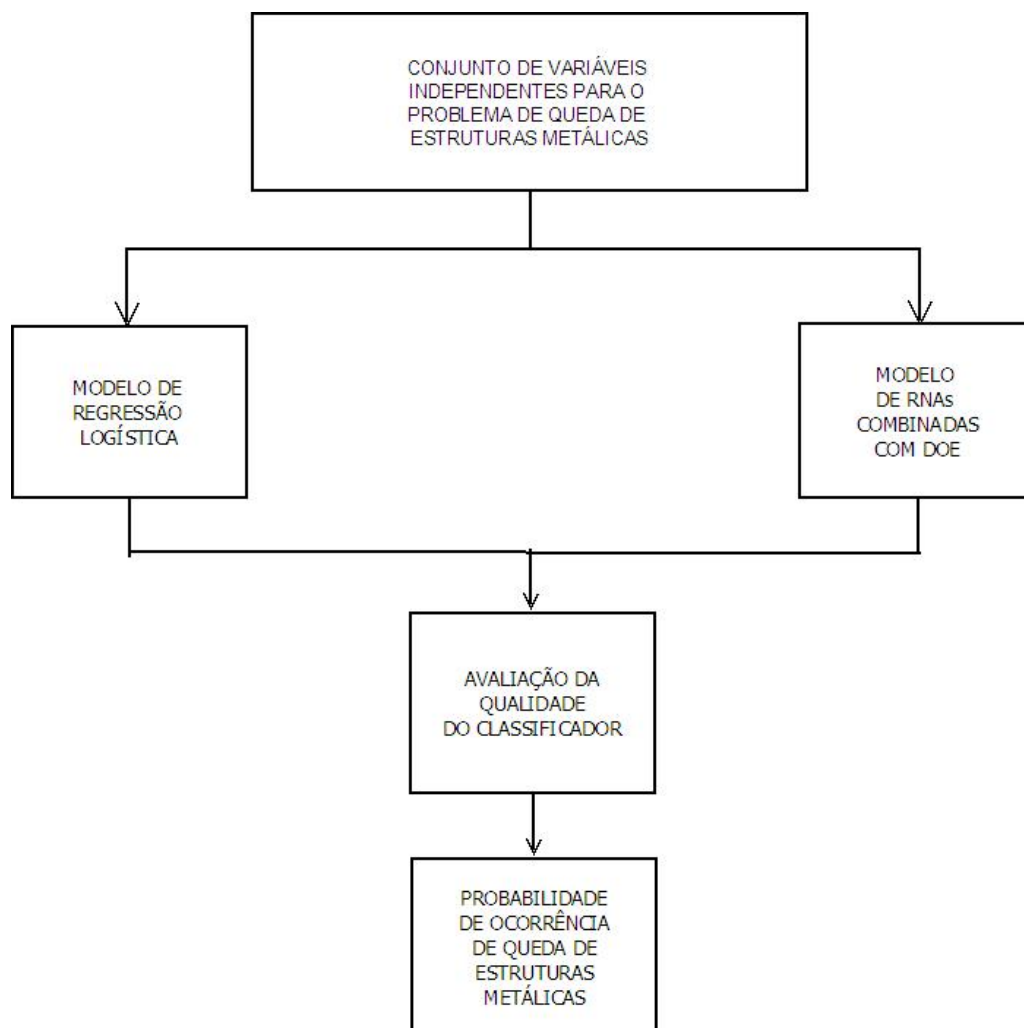


Figura 1: Resumo gráfico da organização dos resultados obtidos

# Abstract

This thesis evaluates the probability of failure of steel structures in transmission lines using two techniques: logistic regression and artificial neural networks. The first may be regarded as a standard technique and has already been established as a powerful helping tool for decision making. The second has been used in profusion for similar problems and presents itself as an excellent alternative to logistic regression, since it assumes fewer restrictions on use, in spite of requiring more complex computer simulation tests. The statistical methodology of Design of Experiments (DOE) is applied to determine the best parameters of an artificial neural network for the structures fall problem. Instead of using the most common technique of training an artificial neural network, namely trial and error, DOE is proposed as a methodology to find the best RNA.

**Keywords:** Artificial Neural Networks, Design of Experiments, Logistic Regression, ROC curves, Transmission lines, Dropped Structures.

# Sumário

<b>Resumo</b>	<b>i</b>
<b>Resumo Gráfico</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Lista de Figuras</b>	<b>vi</b>
<b>Lista de Tabelas</b>	<b>viii</b>
<b>Lista de Abreviaturas e Símbolos</b>	<b>x</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação do Trabalho . . . . .	1
1.2 Revisão Bibliográfica . . . . .	2
1.3 Proposta do Trabalho . . . . .	2
1.4 Contribuição do Trabalho . . . . .	3
1.5 Organização do Trabalho . . . . .	4
<b>2 Conceitos Fundamentais</b>	<b>5</b>
2.1 Regressão Logística . . . . .	5



2.1.1	Modelo de Regressão Logística Simples . . . . .	6
2.1.2	Modelo de Regressão Logística Múltipla . . . . .	10
2.2	Redes Neurais Artificiais e DOE . . . . .	12
2.2.1	DOE para Simulação . . . . .	17
2.2.2	Rede MLP e o DOE de Plackett-Burman . . . . .	21
2.2.3	RNAs e Regressão Logística - Vantagens e Desvantagens . . . . .	26
2.3	Comparação do Comportamento dos Classificadores . . . . .	29
2.3.1	Curvas ROC . . . . .	29
2.4	O Método <i>Desirability</i> . . . . .	32
<b>3</b>	<b>Falhas em Estruturas Metálicas de Linhas de Transmissão</b>	<b>36</b>
3.1	Introdução . . . . .	36
3.2	Linhas de Transmissão . . . . .	37
3.2.1	Estruturas . . . . .	38
3.2.2	Cabos . . . . .	43
3.2.3	Fundações . . . . .	44
3.2.4	Atributos das Estruturas . . . . .	45
3.2.5	Quedas de Estruturas . . . . .	47
<b>4</b>	<b>Avaliação da Probabilidade de Ocorrência de Queda de Estruturas Metálicas de Linhas de Transmissão</b>	<b>49</b>
4.1	Introdução . . . . .	49
4.2	Análise Descritiva do Conjunto de Dados . . . . .	50
4.3	Resultados com Regressão Logística . . . . .	58
4.3.1	Modelo de Probabilidade de Ocorrência de Queda de uma Estrutura . . . . .	58

4.4	Resultados com RNAs Combinadas com DOE . . . . .	66
4.4.1	Planejamento Pré Experimental . . . . .	66
4.4.2	Escolha do Delineamento Experimental . . . . .	73
4.4.3	Realização do Experimento . . . . .	74
4.4.4	Otimização Simultânea das Variáveis de Resposta Média e Diferença . . . .	80
4.4.5	Conclusões e Recomendações . . . . .	85
		<b>87</b>
<b>A</b>	<b>Conjunto de Dados</b>	<b>107</b>
<b>B</b>	<b>Matrizes de Hadamard e o Delineamento de Plackett-Burman</b>	<b>108</b>

# Lista de Figuras

1	Resumo gráfico da organização dos resultados obtidos . . . . .	ii
2.1	Estrutura da RNA <i>feedforward</i> multicamadas . . . . .	12
2.2	Curva ROC com cinco classificadores discretos - Fonte Fawcett (2006) . . . . .	31
2.3	Área sob a curva de dois classificadores probabilísticos -Fonte Fawcett (2006) . . . . .	32
3.1	Queda de torre no Oeste do Paraná - Fonte Portal Medianeira . . . . .	48
4.1	Gráfico ilustrando o número de estruturas com dois tipos de capacidade . . . . .	51
4.2	Gráfico ilustrando o número de estruturas com quatro diferentes tipos de terrenos . . . . .	52
4.3	Gráfico ilustrando o número de estruturas com diferentes massas de ar quente . . . . .	52
4.4	Gráfico ilustrando o número de estruturas com diferentes tipos de massas de ar fria . . . . .	53
4.5	Gráfico ilustrando o número de estruturas com diferentes níveis de tensão . . . . .	53
4.6	<i>Box-plot</i> da variável tensão . . . . .	54
4.7	Gráfico ilustrando o número de estruturas alocadas em diferentes temperaturas . . . . .	54
4.8	<i>Box-plot</i> da variável temperatura . . . . .	55
4.9	Gráfico ilustrando o número de estruturas em diferentes tipos de relevo . . . . .	55
4.10	Gráfico ilustrando o número de estruturas com dois diferentes tipos de estrutura . . . . .	56
4.11	Gráfico ilustrando o número de estruturas com três diferentes tipos de fundação . . . . .	56

4.12	Gráfico ilustrando o número de estruturas com diferentes velocidades do vento . . .	57
4.13	<i>Box plot</i> da variável vento . . . . .	57
4.14	Gráfico de barras ilustrando o número de estruturas que sofreram queda . . . . .	58
4.15	Gráfico da qualidade de ajuste do modelo I . . . . .	65
4.16	Gráfico da qualidade de ajuste do modelo II . . . . .	65
4.17	Curva ROC . . . . .	66
4.18	Diagrama de Pareto para média . . . . .	77
4.19	Diagrama de Pareto para Diferença . . . . .	77
4.20	Plotagem dos efeitos principais da média . . . . .	78
4.21	Plotagem dos efeitos principais da diferença . . . . .	79
4.22	Histograma da variável resposta média . . . . .	80
4.23	Histograma da variável resposta diferença . . . . .	81
4.24	Probabilidade normal para a variável média . . . . .	82
4.25	Probabilidade normal para a variável diferença . . . . .	82
4.26	Simulação da otimização das variáveis resposta . . . . .	83
4.27	Simulação da otimização das variáveis resposta . . . . .	84
4.28	Curva ROC da RNA com parâmetros da tabela 4.12 . . . . .	85

# Lista de Tabelas

2.1	Delineamento Plackett-Burman para doze ensaios . . . . .	25
2.2	Terminologia correspondente entre estatística e RNAs . . . . .	27
2.3	Relação dos artigos que comparam técnicas . . . . .	28
2.4	Matriz de confusão . . . . .	30
2.5	Relação entre a variação $d_i$ com o significado da variável resposta . . . . .	35
3.1	Extensão das linhas de transmissão do SIN . . . . .	38
4.1	Informações dos fatores . . . . .	60
4.2	Tabela de regressão logística . . . . .	61
4.3	Teste de qualidade do ajuste . . . . .	62
4.4	Tabela das frequências observadas e esperadas . . . . .	63
4.5	Medidas de associação . . . . .	64
4.6	Parâmetros de treinamento da RNA para o problema de queda de estruturas metálicas de linhas de transmissão . . . . .	67
4.7	Delineamento Plackett-Burman para 18 fatores . . . . .	74
4.8	Valores obtidos das AUCs . . . . .	75
4.9	Média e amplitude da diferença das AUCs . . . . .	76

4.10	Algumas estatísticas das variáveis média e diferença . . . . .	81
4.11	Configuração para os parâmetros de otimização . . . . .	83
4.12	RNA escolhida para o problema de queda de estruturas . . . . .	84
A.1	Parte do conjunto de dados . . . . .	107
B.1	Estrutura das primeiras linhas de alguns delineamentos de Plackett-Burman . . .	111

# Lista de Abreviaturas e Símbolos

DOE: *Design of Experiments*

ROC: *Receiver Operating Characteristic*

RNA: Rede Neural Artificial

ARIMA: *Autoregressive Integrated Moving Average*

MLP: *Multilayer Perceptron*

RL: Regressão Logística

N-R: Método iterativo de Newton-Raphson

K-S: Teste de normalidade de Kolmogorov-Smirnov

AUC: *Area under Curve*

RBF: *Radial Basis Functions*

GRNN: *General Regression Neural Network*

SVM: *Support Vector Machines*

ARTMAP: *Adaptative Resonance Theory Map*

GARCH: *Generalized Autoregressive Conditional Heteroscedasticity*

SARIMA: *Seasonal Autoregressive Integrated Moving Average*

SARIMABP: *Seasonal Autoregressive Integrated Moving Averages Neural Network*

*Backpropagation*

RSM: *Response Surface Methodology*

IID: Independentes e Identicamente Distribuídos

COPEL: Companhia Paranaense de Energia

ABNT: Associação Brasileira de Normas Técnicas

NBR: Denominação de Norma da ABNT

SIN: Sistema Integrado Nacional

MW: megawatt

kW: quilowatt

*km/h*: quilômetro por hora

$Bin(p, q)$ : Distribuição Binomial com parâmetros  $p$  e  $q$

$Ber(p)$ : Distribuição de Bernoulli com parâmetro  $p$

arg máx: Argumento Máximo

$\mathcal{N}(0, 1)$ : Distribuição normal com média 0 e variância 1

$Unif(a, b)$ : Distribuição uniforme com parâmetros  $a$  e  $b$

$E(Y)$ : Valor esperado da variável aleatória  $Y$

$diag(X)$ : Matriz Diagonal

**I**: Matriz de informação de Fisher

**I<sub>n</sub>**: Matriz identidade de ordem  $n$

$\otimes$ : Produto direto de matrizes

**H<sup>T</sup>**: Matriz transposta

$\varepsilon_i$ : Erro com distribuição normal de média zero e variância constante

$L(\beta)$ : Função de verossimilhança

$l(\beta)$ : Logaritmo natural da função de verossimilhança



# Capítulo 1

## Introdução

### 1.1 Motivação do Trabalho

A energia elétrica é fundamental para a sociedade moderna. Para sua transmissão é necessário o uso de torres que sustentam os cabos transmissores mantendo-os a uma distância segura do solo, interligando as unidades geradoras aos consumidores.

As linhas de transmissão de energia elétrica atravessam lugares nem sempre de fácil acesso, como vales, montanhas e regiões com muitas árvores altas e estão expostas a condições climáticas adversas.

A queda de uma torre de linha de transmissão é um problema comum em sistemas elétricos de grande porte. No Brasil, ventos fortes são responsáveis por uma grande parte de acidentes com torres metálicas usadas em linhas de transmissão. Existem inúmeros registros de acidentes com torres metálicas, que em alguns casos, causaram enormes prejuízos para as empresas do setor elétrico e inúmeros transtornos para os consumidores devido as interrupções do fornecimento de energia. A variável vento pode ser uma variável que impacta positivamente para a queda de uma

estrutura, mas ela não é a única. Existem outras variáveis que combinadas com a ação do vento podem ocasionar a queda da estrutura.

Segundo Rippel (2005) em novembro de 1997 a queda de 10 torres do sistema de transmissão de Itaipu, acarretou um racionamento de energia para as regiões Sul, Sudeste e Centro Oeste do país. O acidente interrompeu a transmissão de 5000 MW ao sistema interligado, obrigando o corte de energia para cerca de um milhão de consumidores no estado de São Paulo.

O problema de queda de uma torre de linha de transmissão é um problema que ocorre em qualquer região do Brasil, podendo ser mais intenso em determinadas regiões. A motivação para elaboração deste trabalho é a importância deste problema no sistema elétrico.

## **1.2 Revisão Bibliográfica**

O único trabalho encontrado, nas bases de conhecimento que tratam de linhas de transmissão, a respeito da avaliação de suscetibilidade das estruturas quanto a quedas, devido a uma série de variáveis independentes, foi o trabalho de Wazen (2011).

## **1.3 Proposta do Trabalho**

Nesta tese, utilizou-se um conjunto de dados contendo 181 estruturas metálicas de linhas de transmissão em operação em um dado sistema elétrico mantido pela COPEL (2015) para avaliar a probabilidade de ocorrência de queda utilizando duas técnicas distintas: regressão logística e RNAs combinadas com DOE. Para tal estudo, considera-se fatores como: velocidade do vento na região da estrutura, relevo de sua base, vegetação de seu entorno, pressão atmosférica e variação de temperatura da região, além da capacidade mecânica máxima e do tipo de fundação da estrutura.

A regressão logística é uma técnica estatística consagrada que permite dizer quais são as variáveis que mais impactam para ocorrência do evento queda. Já a metodologia de RNAs combinadas com DOE permite encontrar um modelo robusto para a avaliação da probabilidade de ocorrência de queda de uma estrutura. O modelo encontrado através das RNAs combinadas com DOE representa a verdadeira contribuição nesse estudo, pois, com o uso das RNAs combinadas com DOE estaremos descartando a tradicional técnica de treinamento de uma RNA, ou seja, tentativa e erro.

Com a solução desse problema será possível avaliar, dentre as estruturas atualmente em operação, quais são mais susceptíveis a danos físicos causados por fatores naturais.

Primeiro, obtém-se um modelo de regressão logística para o conjunto de dados e discute-se as implicações deste modelo. Em seguida, utiliza-se a metodologia de RNAs combinadas com DOE para encontrar um modelo que permita avaliar a probabilidade de ocorrência de queda. Utiliza-se curvas ROC para avaliar o desempenho do modelo de regressão logística e das RNAs na avaliação da probabilidade de ocorrência de falhas de estruturas metálicas.

## 1.4 Contribuição do Trabalho

Wazen (2011) avalia a probabilidade de ocorrência de queda de uma estrutura de linha de transmissão utilizando regressão logística e conjuntos aproximados. Nesta tese, aborda-se o mesmo problema utilizando, além do modelo logístico, o modelo gerado pelas RNAs combinadas com DOE. O uso de DOE, será a maior contribuição do trabalho, pois, a partir dele, é possível estimar os melhores parâmetros da RNA. Com a metodologia de DOE descarta-se a tradicional técnica de treinamento de RNAs, ou seja, tentativa e erro, tendo como consequência um modelo mais robusto para a solução do problema de se avaliar a probabilidade de ocorrência de queda.

## 1.5 Organização do Trabalho

Esta tese está organizada em quatro capítulos, sendo o primeiro capítulo uma introdução ao tema, em que se apresenta a motivação para o problema estudado, a proposta de solução do problema, a contribuição do tema estudado e a forma de organização desta tese.

O capítulo 2, apresenta alguns conceitos fundamentais para o desenvolvimento da tese. Discute-se o modelo de regressão logística simples e multivariado e a estimação dos parâmetros do modelo. Em seguida, apresenta-se diretrizes gerais para o uso de DOE nas RNAs. Discute-se também uma técnica para organizar, visualizar, avaliar e selecionar classificadores baseado em seus desempenhos, conhecida como Análise ROC. Finalmente apresenta-se um método de otimização de múltiplas respostas conhecido por método *desirability*.

O capítulo 3, apresenta um esboço das variáveis de interesse para o estudo do problema de quedas em estruturas metálicas de linhas de transmissão.

O capítulo 4, apresenta os resultados fundamentais desta tese. Avalia-se a probabilidade de ocorrência de queda de estruturas metálicas de linhas de transmissão através de duas técnicas distintas: regressão logística e RNAs combinadas com DOE. Através do modelo logístico, visualiza-se as variáveis que mais impactam para ocorrência de queda e com isso, apresenta-se soluções para as estruturas alocadas em regiões com presença de tais variáveis. Por outro lado, utiliza-se RNAs combinadas com DOE a procura de um modelo robusto que descarte a tradicional técnica de tentativa e erro em busca da melhor rede. De posse do modelo logístico e do modelo de RNAs combinadas com DOE utiliza-se a curva ROC como métrica para avaliar o desempenho de classificação de cada modelo.

Finalmente, apresenta-se as conclusões do trabalho, algumas propostas de trabalhos futuros e as referências que apoiaram todo desenvolvimento deste trabalho.

# Capítulo 2

## Conceitos Fundamentais

Neste capítulo, apresenta-se alguns conceitos fundamentais para o desenvolvimento desta tese. Destacam-se alguns aspectos gerais do modelo logístico univariado e multivariado, das RNAs, do desempenho de um classificador e do método de otimização conhecido por *desirability*.

### 2.1 Regressão Logística

Existem diversos métodos de regressão e que podem ser divididos de diferentes maneiras. Os critérios utilizados na divisão dos métodos são baseados no tipo do modelo gerado. A regressão logística é uma técnica estatística que tem como objetivo produzir, a partir de um conjunto de observações, um modelo que permita a predição de valores tomados por uma variável categórica, frequentemente binária, a partir de uma série de variáveis explicativas contínuas e/ou binárias.

A regressão logística tem se constituído num dos principais métodos de modelagem estatística de dados. Embora seja conhecida desde os anos 50 foi através de Cox que a mesma se tornou popular entre os usuários de Estatística. Atualmente, pesquisadores de diversas áreas do conhecimento utilizam a regressão logística como ferramenta de apoio a pesquisa.

### 2.1.1 Modelo de Regressão Logística Simples

Os métodos de regressão têm como objetivo descrever as relações entre a variável resposta  $Y$  e a variável explicativa  $X$ . Segundo Hosmer e Lemeshow (1989) na regressão logística, a variável resposta  $Y$  é dicotômica, isto é, atribui-se o valor 1 para o acontecimento de interesse (sucesso) e o valor de 0 para o acontecimento complementar (fracasso), com probabilidades  $\pi_i = P(Y = 1|X = x_i)$  e  $1 - \pi_i = P(Y = 0|X = x_i)$ , respectivamente. Para descrever a média condicional de  $Y$  dado  $X$  com distribuição logística, utiliza-se a notação  $\pi_i$ .

Considera-se uma série de eventos binários, em que  $Y = (Y_1, \dots, Y_n)$  são variáveis aleatórias independentes com distribuição de Bernoulli, com probabilidade de sucesso  $\pi_i$ , isto é,  $Y_i \sim Ber(\pi_i)$  e denota-se  $x_i^T = (1, x_i)$  a  $i$ -ésima linha da matriz  $\mathbf{X}$  em que  $i = 1, 2, \dots, n$ .

A probabilidade de sucesso do modelo logístico simples é definida como

$$\pi_i = \pi(x_i) = P(Y = 1|X = x_i) = \frac{e^{(\beta_0 + \beta_1 x_i)}}{1 + e^{(\beta_0 + \beta_1 x_i)}} \quad (2.1)$$

e a probabilidade de fracasso

$$1 - \pi_i = 1 - \pi(x_i) = P(Y = 0|X = x_i) = \frac{1}{1 + e^{(\beta_0 + \beta_1 x_i)}} \quad (2.2)$$

em que  $\beta = (\beta_0, \beta_1)^T$  é o vetor de parâmetros desconhecidos.

Uma diferença importante entre o modelo de regressão logística e o modelo de regressão linear pode ser notada, quando diz respeito a natureza da relação entre a variável resposta e as variáveis independentes. Em qualquer problema de regressão, a quantidade a ser modelada é o valor médio da variável resposta dado os valores das variáveis independentes. Esta quantidade é chamada de média condicional, denotada por  $E(Y|X = x_i)$ , em que  $Y$  é a variável resposta e  $x_i$ , os valores das variáveis independentes. Na regressão linear tem-se  $-\infty < E(Y|X = x_i) < +\infty$  e na regressão

logística, devido a natureza da variável resposta,  $0 \leq E(Y|X = x_i) \leq 1$ .

Na regressão linear,  $E(Y|X = x_i) = \beta_0 + \beta_1 x_i$  e na regressão logística, usando a definição de variáveis aleatórias discretas, tem-se

$$E(Y|X = x_i) = 1P(Y_i = 1|X = x_i) + 0P(Y_i = 0|X = x_i) = \pi_i \quad (2.3)$$

Outra diferença importante entre um modelo de regressão linear e o modelo de regressão logística refere-se a distribuição condicional da variável resposta. No modelo de regressão linear, assume-se que uma observação da variável resposta pode ser expressa por  $Y_i = E(Y|X = x_i) + \varepsilon_i$ , em que  $\varepsilon_i$  é chamado de erro, com distribuição normal com média zero e variância constante. Isto não ocorre quando a resposta é dicotômica.

A transformação de  $\pi_i$ , é interpretada como o logaritmo da razão das chances entre  $\pi_i$  e  $1 - \pi_i$ . Esta transformação é bastante empregada em várias áreas do conhecimento, sendo definida como

$$g(x_i) = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) \quad (2.4)$$

Supondo que  $(x_i, y_i)$  seja uma amostra independente com  $n$  pares de observações,  $y_i$  representa o valor da variável resposta dicotômica e  $x_i$  é o valor da variável independente da  $i$ -ésima observação em que  $i = 1, 2, \dots, n$ . Para o ajuste do modelo de regressão logística simples, segundo a equação 2.1, é necessário estimar os parâmetros desconhecidos:  $\beta_0$  e  $\beta_1$ . O método mais utilizado para estimar esses parâmetros considerando uma regressão linear é o de mínimos quadrados. Neste método a escolha de  $\beta_0$  e  $\beta_1$  é dada pelos valores que minimizam a soma dos quadrados dos desvios para os valores observados ( $y_i$ ) em relação ao valor predito ( $\hat{y}_i$ ) baseado no modelo. No entanto, quando o método dos mínimos quadrados é aplicado para um modelo com variável dicotômica, os estimadores não seguem as mesmas pressuposições do modelo de regressão linear.

O método da máxima verossimilhança é utilizado para estimar os parâmetros de um determinado modelo. Trata-se de um método de estimação pontual para obtenção de estimadores de uso extremamente amplo porque possibilita realizar inferências com propriedades altamente desejáveis. Assim, a partir de um conjunto de dados e dado um modelo estatístico, a estimativa por máxima verossimilhança estima valores para os diferentes parâmetros do modelo.

A função de distribuição da probabilidade de  $Y_i$  para o modelo de regressão logística simples com  $Y_i \sim Ber(\pi_i)$  é dada por

$$f(y_i, \pi_i) = \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \quad (2.5)$$

Como as observações são independentes, a função de distribuição de probabilidade conjunta de  $y_1, \dots, y_n$  será

$$\prod_{i=1}^n f(y_i, \pi_i) = \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i}, \quad y_i \in [0, 1] \quad (2.6)$$

Então a função de verossimilhança é dada por

$$L(\beta) = \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i}, \quad \beta \in \mathbb{R}^2 \quad (2.7)$$

O princípio da máxima verossimilhança consiste em estimar o valor de  $\beta$  que maximiza  $L(\beta)$ .

Aplicando logaritmo natural a equação 2.7 tem-se que

$$l(\beta) = \ln[L(\beta)] = \ln \left[ \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \right] \quad (2.8)$$

Utilizando as propriedades de logaritmo juntamente com as propriedades de produtório, segue que



$$\begin{aligned}
l(\beta) &= \ln[L(\beta)] = \ln \left[ \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \right] \\
&= \sum_{i=1}^n [y_i \ln(\pi_i) + (1 - \pi_i) \ln(1 - \pi_i)] \\
&= \sum_{i=1}^n [y_i \ln(\pi_i) + \ln(1 - \pi_i) - y_i \ln(1 - \pi_i)] \\
&= \sum_{i=1}^n \left[ y_i \ln \left( \frac{\pi_i}{1 - \pi_i} \right) + \ln(1 - \pi_i) \right] \\
&= \sum_{i=1}^n [y_i(\beta_0 + \beta_1 x_i) - \ln(1 + e^{\beta_0 + \beta_1 x_i})]
\end{aligned} \tag{2.9}$$

Para encontrar o valor de  $\beta \in \mathbb{R}^2$  que maximiza  $l(\beta)$ , deriva-se  $l(\beta)$  em relação a cada parâmetro  $(\beta_0, \beta_1)$ , obtendo-se duas equações

$$\frac{\partial l(\beta)}{\partial \beta_0} = \sum_{i=1}^n \left[ y_i - \frac{1}{1 + e^{(\beta_0 + \beta_1 x_i)}} e^{(\beta_0 + \beta_1 x_i)} \right] \tag{2.10}$$

$$\frac{\partial l(\beta)}{\partial \beta_1} = \sum_{i=1}^n \left[ y_i x_i - \frac{1}{1 + e^{(\beta_0 + \beta_1 x_i)}} e^{(\beta_0 + \beta_1 x_i) x_i} \right] \tag{2.11}$$

que igualando-as a zero, geram o seguinte sistema de equações

$$\begin{cases} \sum_{i=1}^n (y_i - \pi_i) = 0 \\ \sum_{i=1}^n x_i (y_i - \pi_i) = 0 \end{cases} \tag{2.12}$$

em que  $i = 1, \dots, n$  e  $\pi_i = \frac{e^{(\beta_0 + \beta_1 x_i)}}{1 + e^{(\beta_0 + \beta_1 x_i)}}$ .

Como o sistema de equações 2.12 é não linear em  $\beta_0$  e  $\beta_1$ , são necessários métodos iterativos para sua resolução.

## 2.1.2 Modelo de Regressão Logística Múltipla

Seja um conjunto com  $p$  variáveis independentes, denotadas por  $\mathbf{x}_i^T = (x_{i0}, x_{i1}, x_{i2}, \dots, x_{ip})$ , o vetor da  $i$ -ésima linha da matriz  $\mathbf{X}$  das variáveis explicativas, em que cada elemento da matriz correspondente ao  $ij$ -ésimo componente ( $x_{ij}$ ), em que  $i = 1, 2, \dots, n$  e  $j = 0, 1, \dots, p$ , com  $x_{i0} = 1$ . Denota-se por  $\beta = (\beta_0, \dots, \beta_p)^T$  o vetor de parâmetros desconhecidos e  $\beta_j$  é o  $j$ -ésimo parâmetro associado a variável explicativa  $x_j$ . No modelo de regressão logística múltipla a probabilidade de sucesso é dada por

$$\pi_i = \pi(\mathbf{x}_i) = P(Y_i = 1 | \mathbf{X} = \mathbf{x}_i) = \frac{e^{\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}}} = \frac{e^{\mathbf{x}_i^T \beta}}{1 + e^{\mathbf{x}_i^T \beta}} \quad (2.13)$$

e a probabilidade de fracasso dada por

$$1 - \pi_i = 1 - \pi(\mathbf{x}_i) = P(Y_i = 0 | \mathbf{X} = \mathbf{x}_i) = \frac{1}{1 + e^{\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}}} = \frac{1}{1 + e^{\mathbf{x}_i^T \beta}} \quad (2.14)$$

No modelo de regressão linear múltipla assume-se que  $Y_i$  tem uma distribuição de Bernoulli com parâmetro de sucesso  $\pi_i$ . O logit para o modelo de regressão logística múltipla é dado pela equação

$$g(\mathbf{x}_i) = \ln\left[\frac{\pi_i}{1 - \pi_i}\right] = \mathbf{x}_i^T \beta = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} \quad (2.15)$$

Assim, o logaritmo da função de verossimilhança pode ser escrito da seguinte forma

$$l(\beta) = \sum_{i=1}^n [y_i \mathbf{x}_i^T \beta - \ln(1 + e^{\mathbf{x}_i^T \beta})] \quad (2.16)$$

A estimação dos parâmetros do modelo de regressão logística múltipla é realizado através do método de máxima verossimilhança que consiste em maximizar a função 2.16.

Segundo Silva (1992), para obter o valor de  $\beta$  que maximiza  $l(\beta)$ , utiliza-se o processo iterativo de N-R, e, para isso, tem-se que

$$\frac{\partial l(\beta)}{\partial \beta_j} = \sum_{i=1}^n \left[ y_i x_{ij} - \frac{e^{\mathbf{x}_i^T \beta}}{1 + e^{\mathbf{x}_i^T \beta}} x_{ij} \right] = \sum_{i=1}^n [y_i - \pi_i] x_{ij} \quad (2.17)$$

dessa maneira, o vetor escore  $\mathbf{U}(\beta)$  é escrito como

$$\mathbf{U}(\beta) = \mathbf{X}^T \mathbf{y} - \mathbf{X}^T \boldsymbol{\pi} = \mathbf{X}^T (\mathbf{y} - \boldsymbol{\pi}) \quad (2.18)$$

e a matriz de informação de Fischer é dada por

$$\mathbf{I}(\beta) = E \left[ -\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta^T} \right] = \mathbf{X}^T \mathbf{Q} \mathbf{X} \quad (2.19)$$

em que  $\mathbf{Q} = \text{diag}[\pi_i(1 - \pi_i)]$  e  $\mathbf{X}$  a matriz de dados.

As soluções para as equações de verossimilhança são obtidas utilizando o método iterativo de N-R.

O conjunto de equações iterativas é dado por

$$\beta^{(t+1)} = \beta^{(t)} + [\mathbf{I}(\beta^{(t)})]^{-1} \mathbf{U}(\beta^{(t)}) \quad (2.20)$$

$t = 0, 1, 2, \dots$ , sendo que  $\beta^{(t+1)}$  e  $\beta^{(t)}$  são vetores de parâmetros estimados nos passos  $t$  e  $t + 1$ , respectivamente.

Para mais detalhes veja Allison (1999).

## 2.2 Redes Neurais Artificiais e DOE

As primeiras aplicações das RNAs ocorreram nos campos da ciência cognitiva e engenharia. Como citado por Tsay (2005), as RNAs são ferramentas gerais e flexíveis para aplicações em previsão.

As RNAs representam um tema popular na moderna análise de dados e podem ser classificadas como um método semiparamétrico. Ao contrário dos métodos não lineares baseados em modelos, as RNA são abordagens baseadas em dados que podem capturar as estruturas não lineares, sem hipótese prévia sobre a relação subjacente em um problema particular.

A figura 2.1 mostra a estrutura de uma RNA amplamente utilizada: uma rede *feedforward* multicamadas formada com o algoritmo *backpropagation*. A RNA possui três tipos de camadas, a saber, a camada de entrada *input layer*, a camada de saída *output layer* e a camada oculta *hidden layers*, a qual é intermediária entre a entrada e as camadas de saída. O número de camadas ocultas é geralmente um ou dois. Cada camada é constituída de neurônios, e os neurônios em duas camadas adjacentes estão completamente ligados com respectivos pesos, enquanto que os neurônios no interior da mesma camada não estão ligados.

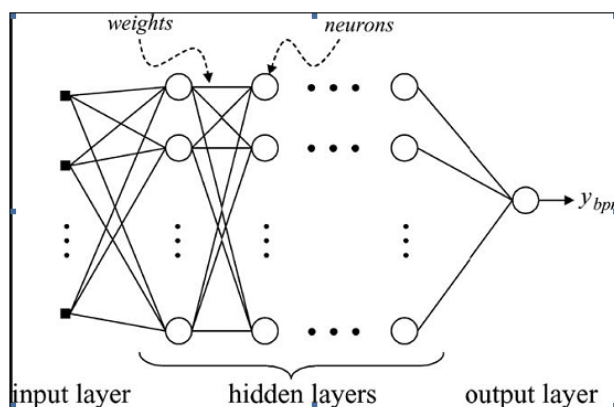


Figura 2.1: Estrutura da RNA *feedforward* multicamadas

Os neurônios na camada de entrada são designados para um atributo nos dados e produz

uma saída que é igual ao valor (escala) do atributo correspondente. Por outro lado, para cada neurônio na camada oculta ou de saída, a seguinte transformação de entrada e saída é geralmente empregada:

$$v = f\left(\sum_{h=1}^H w_h u_h + w_0\right) \quad (2.21)$$

em que  $v$  é a saída,  $H$  é o número total de neurônios na camada anterior,  $u_h$  é a saída e  $h$  o número de neurônios na camada anterior,  $w_h$  é o peso de ligação correspondente,  $w_0$  é a polarização (ou interceptão). A função  $f$  é a função de ativação também usada na camada de saída. A função de ativação, costuma ser denotada por  $\varphi(v)$  e define a saída de um neurônio em termos do campo local induzido  $v$ . Existem algumas funções de ativação que são consideradas clássicas, por exemplo, a função  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ , definida por

$$\varphi(v) = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases}$$

é denominada função de limiar.

Na literatura de engenharia, esta forma de função de limiar é normalmente referida como função de Heaviside. Correspondentemente, a saída do neurônio  $k$  que emprega essa função de limiar é expressa como

$$y_k = \begin{cases} 1 & \text{se } v_k \geq 0 \\ 0 & \text{se } v_k < 0 \end{cases}$$

em que  $v_k$  é o campo local induzido do neurônio, isto é,

$$v_k = \sum_{i=1}^n w_{ki} x_i + b_k$$

Por outro lado, a função  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ , definida por

$$\varphi(v) = \begin{cases} 1 & \text{se } v \geq \frac{1}{2} \\ v & \text{se } -\frac{1}{2} < v < \frac{1}{2} \\ 0 & \text{se } v \leq -\frac{1}{2} \end{cases}$$

é denominada função linear por partes.

Uma função de ativação muito comum é a função  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  tal que

$$\varphi(v) \rightarrow \begin{cases} 1 & \text{se } v \rightarrow +\infty \\ 0 & \text{se } v \rightarrow -\infty \end{cases}$$

que é denominada função sigmóide.

As funções sigmóides são largamente as mais utilizadas em projetos de RNAs e proporcionam um balanceamento entre comportamento linear e não linear. Outra vantagem é que as funções sigmóides são diferenciáveis (diferenciabilidade é uma característica importante da teoria das RNAs). As funções sigmóides mais utilizadas são:  $\varphi(v) = \frac{1}{1 + e^{-v}}$  e  $\varphi(v) = \tanh(v)$ .

A avaliação de RNAs a partir de perspectivas estatísticas e econométricas podem ser encontradas em Cheng (1994). Atualmente, RNAs são usadas em uma grande variedade de modelos e problemas de previsão. Apesar de muitos modelos comumente usados em problemas reais são lineares, a natureza da maioria do conjunto de dados reais sugere que modelos não lineares são mais apropriados para previsão e descrição do problema.

A literatura sobre RNAs é muito vasta e suas aplicações espalharam-se por vários domínios científicos com graus variados de sucesso. Na *M-Competition*, Makridakis et al. (1982), *M2-Competition*, Makridakis et al. (1993) e *M3-Competition*, Makridakis et al. (2000) muitos participantes utilizam RNAs. A principal razão para esse aumento da popularidade das RNAs é

que elas tem se mostrado capazes de aproximar qualquer função não linear arbitrariamente dada (Teorema de Aproximação Universal). Assim, quando aplicada a dados, que são caracterizados por relações dinâmicas de natureza não linear, a RNA irá detectar essas relações e proporcionar uma modelagem superior quando comparado com os modelos lineares.

As RNAs são amplamente utilizadas para resolver o problema de previsão de séries temporais. Diversos artigos tem considerado as RNAs como uma metodologia promissora e que abordam questões importantes. Franses e Homelen (1998) exploraram a habilidade de uma RNA capturar não linearidades como implicação dos modelos SETAR, *Markov-Switching* e GARCH. Kaastra e Boyd (1996) fornecem um guia introdutório, em oito etapas, na concepção de RNAs, para o procedimento de previsão de séries temporais econômicas. Bodyanskiy e Popov (2006) apresentam uma abordagem especial de RNAs para previsão de séries temporais financeiras com base na apresentação da série como uma combinação de componentes quase periódicos. Tseng et al. (2002) propõem um modelo de previsão híbrido, que combinam séries temporais sazonais ARIMA (SARIMA) e RNAs treinadas com o algoritmo *backpropagation*, modelos denominados SARIMABP. Karunasinghe e Liong (2005) investigaram o desempenho de uma RNA como um modelo global sobre os modelos locais utilizados (técnica local das médias e técnica local dos polinômios) em séries temporais caóticas. No artigo de Aitkenhead et al. (2003) óleo, água corrente e as variáveis climáticas foram mensurados durante vários meses em duas situações no nordeste da Escócia, utilizando coletores de dados e outros instrumentos de medida. Para o conjunto de dados, foram utilizados três diferentes métodos no treinamento de RNAs, incluindo um novo, sistema biologicamente plausível. BuHamra et al. (2003) combinam Box Jenkins e a abordagem de RNAs para modelar séries temporais de dados de consumo de água no Kuwait. Shi et al. (1999) investigaram a modelagem de séries temporais não lineares usando o modelo *state dependent autoregressive*.

Niska et al. (2004) estudaram o modelo de qualidade do ar, usando RNAs, uma tarefa difícil devido ao fenômeno caótico e não linear e um espaço amostral de alta dimensão. Zhang (2003) apresenta uma metodologia híbrida que combina o modelo ARIMA e RNAs para retirar vantagem do modelo ARIMA e das RNAs na modelagem linear e não linear. Kim (2004) usa SVM para previsão de séries temporais financeiras. Este estudo se aplica a previsão do índice de preços das ações. Ho et al. (2002) mostram um estudo comparativo de RNAs e modelos ARIMA na previsão de séries temporais. O algoritmo *backpropagation* e RNAs recorrentes apresentam um desempenho satisfatório quando comparados com ARIMA. Kohzadi et al. (1996) comparam ARIMA e RNAs no desempenho de previsão de preços. Terasvirta et al. (2005) examinaram a precisão das previsões do modelo linear autoregressivo, *smooth transition autoregressive* e RNAs para 47 variáveis macroeconômicas mensais das economias do G7. Ghiassi et al. (2005) apresentam um modelo dinâmico de RNA para previsão de eventos em séries temporais que utilizam uma arquitetura diferente dos modelos tradicionais. Balkin e Ord (2000) exploram um método chamado *Automated Artificial Neural Networks*, que é a tentativa de desenvolver um procedimento automático para seleção da arquitetura da RNA para fins de previsão. Cubiles de la Vega et al. (2002) propõem um procedimento para projetar um *perceptron* multicamadas para previsão de séries temporais. Este fato baseia-se na geração de acordo com um conjunto de regras que emergem de um modelo ARIMA, previamente montado de um conjunto de modelos não lineares de previsão. Kalaitzakis et al. (2002) apresentam o desenvolvimento e aplicação de RNAs avançadas para enfrentar com sucesso o problema da previsão da carga elétrica de curto prazo utilizando dados reais de carga a partir do sistema elétrico da ilha de Creta, na Grécia. Zhang e Qi (2001) investigaram o problema de sazonalidade e mostraram que estudos empíricos sobre a previsão de séries temporais sazonais com RNAs produzem resultados mistos. Chiang et al.



(1996) mostraram que as RNAs são superiores aos modelos de regressão.

1. “[...] Em geral, as RNAs apresentam um desempenho satisfatório em problemas de previsão” (BALESTRASSI et al. 2009);

2. Balestrassi et al. (2009) concluíram que:

Existem muitos fatores que podem afetar o desempenho das RNAs. No entanto, não há investigações sistemáticas destas questões. A metodologia de tentativa e erro é adotada em problemas específicos por muitos pesquisadores que é a razão primária para inconsistências na literatura.

3. “[...] Uma quantidade considerável de pesquisa tem sido realizada nesta área, dado o rápido crescimento da literatura” (BALESTRASSI et al. 2009);

### 2.2.1 DOE para Simulação

O processo de treinamento de uma RNA consiste na mudança de entrada dos parâmetros do algoritmo computacional, processar o algoritmo e verificar os resultados. Isto pode ser referido como um estudo de simulação para o problema de RNA.

Apesar da quantidade de sucessos relacionados com aplicações industriais, DOE não é amplamente usado em simulação. Kleijnen et al. (2005) apontam que a falta do uso de DOE para simulação deve-se a:

- Os analistas de simulação não estão convencidos dos benefícios de DOE;
- A investigação em DOE é geralmente encontrada em revistas especializadas raramente lidas por analistas de simulação;
- A maioria dos DOEs foram originalmente desenvolvidos para a experimentação do mundo real, em vez de desenvolvidos especificamente para as configurações de simulação.

Balestrassi et al. (2009) apontam que:

DOE pode detectar informações detalhadas sobre o comportamento do modelo, permite que a equipe envolvida na modelagem discuta em detalhes as implicações de vários pressupostos do modelo, ajuda a formular perguntas, quando os analistas não sabem de antemão que perguntas devem ser feitas, contesta ou confirma as expectativas sobre a direção e a importância relativa dos efeitos de fatores, e até mesmo descobre problemas na lógica do programa. DOE representa uma estratégia adequada para uma determinada aplicação, muito melhor do que tentativa e erro no contexto de RNAs. Conseqüentemente, os profissionais devem estar abertos para a noção de que DOE é uma parte útil e necessária de análise de simulação do complexo.

Por outro lado, Balestrassi et al. (2009) concluíram que:

Traduzindo para a terminologia de simulação pode-se dizer que uma entrada ou um parâmetro na simulação corresponde a um fator no DOE. Normalmente há muitos mais fatores na simulação do que em um experimento no mundo real. Um fator pode ser qualitativo ou quantitativo. Cada fator pode ser definido por dois ou mais valores, chamados de níveis de fator, normalmente codificados numericamente para fins de análise. Um experimento é uma combinação dos níveis de todos os fatores. Ao considerar simulações estocásticas, repetições significam que os diferentes números pseudo aleatórios (*Pseudo Random Number*), são utilizados para simular o mesmo experimento. A natureza da coleta de dados não é aleatória, mas seqüencial. As saídas nas repetições são IID, como a maioria dos métodos estatísticos.

Segundo Kleijnen et al. (2005), desenvolver uma compreensão básica de simulação é equivalente a testar hipóteses sobre os efeitos do fator no DOE.

Outra questão importante quando se utiliza DOE para a simulação é que o objetivo principal aqui não é a otimização. Ao utilizar DOE os esforços são dedicados a encontrar decisões robustas, ao invés de decisões ideais. É claro que encontrar a melhor decisão para um sistema simulado é um tema importante, e foram propostos vários métodos. Fu (2002) e Spall (2003) discutem a atual pesquisa e a prática de otimização para simulação. Estes métodos incluem técnicas heurísticas de busca, tais como algoritmos genéticos, metodologia de RSM, *simulated annealing*, *tabu search*

e métodos que analisam o modelo de simulação para estimar gradientes, tais como análise de perturbação e funções de pontuação. O resultado da otimização está condicionada a pressupostos de distribuições específicas (normalmente assumido como independente) e muitas variáveis de entrada. O termo “melhor” é problemático quando a probabilidade de todas estas suposições, mesmo que por um tempo limitado é efetivamente nulo. Em contraste, uma abordagem de delineamento robusto trata todos estes pressupostos como fatores adicionais durante a execução do experimento. Estes são considerados fatores de ruído (em vez de fatores de decisão), porque eles são desconhecidos ou incontroláveis no ambiente do mundo real. Um sistema robusto funciona bem em uma variedade de condições de ruído que pode ser experimentado. Por isso, a implementação de uma solução robusta é muito menos provável de resultar em resultados inesperados. Esta filosofia de delineamento robusto é inspirada por Taguchi (1987), que usa o delineamento de seleção para identificar configurações de produtos robustos para a Toyota.

Conforme descrito anteriormente, RNAs são candidatas naturais para previsão de modelos não lineares. No entanto, o grande número de parâmetros que precisam ser selecionados para desenvolver um modelo de previsão significa que o processo de concepção ainda envolve muitas tentativas e erros. Os métodos tradicionais de estudar um fator por vez podem conduzir a resultados não confiáveis e enganosos, e às vezes pode produzir conclusões erradas. Estatisticamente, experiências concebidas num desempenho mais eficiente que consideraram simultaneamente múltiplos fatores detectaram interações importantes. Usando DOE como um processo de planejamento de experimentos é possível recolher dados apropriados usando o número mínimo de experimentos ao adquirir as informações técnicas necessárias. Os *insights* sobre os efeitos principais, bem como os efeitos de interação de fatores, incluindo o ruído são úteis na tomada de decisão.

Diversos trabalhos têm estudado o uso de RNAs combinadas com DOE, em um contexto

diferente das propostas desta tese e explorando diferentes fatores e níveis. Khaw et al. (1995) descrevem uma inovadora aplicação do método Taguchi para a determinação dos parâmetros do delineamento, que incluem tanto os aspectos estruturais de micro e macro estrutura de uma RNA. A viabilidade da utilização desta abordagem foi demonstrada por meio da otimização dos parâmetros do projeto de uma RNA treinada com *backpropagation* para determinar as políticas operacionais de um sistema de manufatura. Os resultados extraídos de Taguchi (1987) demonstraram que o método Taguchi proporciona um meio eficaz para melhorar o desempenho da RNA, em termos de velocidade de aprendizagem e precisão. Kim e Yum (2004) apresentam um papel similar. Sukthomya e Tannock (2005a, 2005b) utilizam a mesma idéia do projeto experimental de Taguchi para definir os parâmetros de uma RNA em um processo de formação de complexos. Lin e Tseng (2000) utilizam a mesma abordagem Taguchi para a “Aprendizagem de Quantização Vetorial” de RNAs em um aplicativo para sistemas desviador de bicicleta. Enemuoh e El Gizawy (2003) descrevem um método para a criação de uma RNA robusta para a predição de delaminação, a largura do dano, e a rugosidade da superfície do furo durante a perfuração em epóxi reforçado com fibra de carbono. Os efeitos do número de neurônios, camadas ocultas, a função de ativação e algoritmo de aprendizagem e o erro quadrático médio de previsão do modelo são quantificados. Zhang (2001) e Zhang et al. (2001) apresentam uma avaliação experimental de RNAs para previsão de séries temporais lineares e não lineares. Três principais parâmetros da RNA são examinados através de um experimento simulado através de um computador:

- Nós de entrada;
- Nós ocultos;
- Tamanho da amostra.

## 2.2.2 Rede MLP e o DOE de Plackett-Burman

Coleman e Montgomery (1993) apresentam algumas orientações para a concepção de um experimento que, apesar de ser centrado em experimentos industriais, também pode ser usado para a simulação computacional:

- Definição do problema;
- Escolha dos fatores e níveis;
- Seleção da variável de resposta;
- Escolha do delineamento experimental;
- Realização do experimento;
- Análise estatística dos dados;
- Conclusões e recomendações.

Essas diretrizes são geralmente interativas e a estrutura não é rígida quando aplicado em experimentos reais. Alguns passos são muitas vezes feitos simultaneamente ou também em ordem inversa. Os três primeiros passos são chamados de planejamento pré experimental. Alguns comentários relacionados com o treinamento de uma RNA são apresentadas a seguir:

*Definição do problema.* Isto pode parecer um ponto óbvio, mas em experimentos industriais e também na simulação computacional não é simples de obter uma imagem global de um problema e, geralmente, é necessária uma abordagem de equipe. Para essa etapa, que envolve muitas áreas do conhecimentos, em que as opiniões diferentes são muitas vezes conflitantes, uma abordagem de equipe é apropriada. Além disso, uma declaração clara do problema contribui substancialmente

para a solução do problema. O reconhecimento do problema e sua afirmação correta costumam dar foco para alcançar um objetivo.

*Escolha dos fatores e níveis.* Em experiências do mundo real, apenas um pequeno número de fatores normalmente são variados. Na verdade, é impraticável ou impossível tentar controlar mais do que dez fatores. Muitos experimentos publicados lidam com menos de cinco fatores. Em contraste, existe uma multiplicidade de potenciais fatores para modelos de simulação.

Uma boa programação evita fixar fatores com valores numéricos específicos dentro do código, em vez disso, o código lê valores do fator para que o programa possa ser executado por muitas combinações de valores. O código computacional deve verificar se estes valores são admissíveis, isto é, quais combinações devem estar dentro do domínio experimental. Tal prática pode fornecer automaticamente uma lista de potenciais fatores. Em seguida, os usuários devem confirmar se eles realmente desejam experimentar todos esses fatores ou se querem corrigir alguns fatores em níveis nominais (ou base) *a priori*. Este tipo de codificação ajuda a flexibilizar a mentalidade dos usuários que de outra forma seriam inclinados a se concentrarem em apenas alguns fatores.

Segundo Kleijnen et al. (2005) os analistas devem efetuar aleatorização na ordem de amostragem para se protegerem contra mudanças relacionadas ao tempo no ambiente experimental (como temperatura, umidade, a confiança do consumidor e os efeitos de aprendizagem) e realizar testes estatísticos apropriados para determinar se os resultados foram contaminados. A maioria dos experimentos de simulação são executados sequencialmente, mesmo se eles não são analisados formalmente dessa forma. Se um pequeno número de pontos do experimento são explorados, esta aplicação pode envolver os analistas a alteração dos níveis dos fatores. O aumento da velocidade dos computadores fez com que alguns analistas adicionassem mais detalhes para os seus modelos de simulação. Diferentes analistas podem usar um conjunto diferente de fatores e níveis.

*Seleção da Variável de Resposta.* Na seleção da variável resposta, o experimentador deve ter certeza de que aquela variável realmente fornece informação útil sobre o processo em estudo. Nesta etapa o pesquisador decide qual será a variável resposta adequada para seu problema.

*Escolha do Delineamento Experimental.* A escolha do delineamento experimental depende do objetivo do estudo, natureza do material experimental. A expressão “escolha do delineamento experimental” indica a criação de uma matriz, em que as colunas representam os elementos de entrada e cada linha representa uma combinação de níveis do fator.

De acordo com Crary (2004), alguns pacotes de *softwares* estatísticos como o *Minitab*, *Statistica*, *SPSS*, *JMP*, *Matlab*, entre muitos outros, são bons programas que oferecem uma biblioteca de delineamentos clássicos. Estes delineamentos normalmente são gerados com os níveis codificados e são escolhidos com base no número de fatores e níveis, estrutura de alisamento, confundimento, resolução e quantidade de tempo e recursos para executar os experimentos. Um programa novo e interessante, conhecido por *WebDOE*, ajuda os usuários a projetar sua experiência através de um *easy-to-use* interface de *web*.

Kleijnen et al. (2005), recomendam, ao se usar DOE para simulação, um esquema baseado nos pressupostos da complexidade do sistema e afirmam que é melhor inicialmente se concentrar em um grande número de fatores (respeitando as limitações de tempo, custo, etc.). Dessa forma, os analistas visualizam amplamente através dos fatores no estudo de simulação. Às vezes, a intuição está errada e precisa ser desafiada. Com esta abordagem, a partir de suposições mínimas para uma experiência inicial, o analista terá a mente aberta para novos pressupostos e admite que pouco se sabe sobre a natureza da resposta. O analista tende a reduzir o esforço de coleta de dados iniciais e de fazer hipóteses simplificadoras. Os analistas podem reduzir o esforço computacional, fazendo suposições sobre a natureza do sistema. A melhor idéia é passar de suposições mínimas

e concentrar-se na pequena relação de fatores selecionados após o experimento inicial, mantendo nos demais fatores apenas algumas configurações. Na prática, os resultados de experimentos, muitas vezes precisam ser modificados, ou seja, expandido ou excluídos para obter informações mais detalhadas sobre o desempenho da simulação para uma região menor das combinações de fatores.

*Realização do Experimento.* A execução de um experimento através de simulação é, geralmente, mais fácil do que a realização de experimentos industriais. Se o programa de simulação de código é aberto, é por vezes conveniente para executar a simulação completa como um processo descontínuo. Infelizmente, isso não é o que acontece muitas vezes devido à habilidade de programação necessária e a restrição de *software* sempre imposta. Além disso, essa experimentação iterativa é considerada um erro em experiências industriais. Coleman e Montgomery (1993) salientam que a experimentação deve ser sequencial e, como regra geral, não mais do que cerca de 25% dos recursos disponíveis devem ser investidos no primeiro experimento.

*Análise Estatística dos Dados.* Métodos estatísticos não terão muito sentido, se as orientações anteriores não forem seguidas. Métodos gráficos, análise residual e modelo de adequação desempenham um papel importante nesta fase. A análise estatística adiciona objetividade ao processo de tomada de decisão.

*Conclusões e Recomendações.* Se juntamente com o processo de conhecimento e do senso comum, a análise estatística pode levar a conclusões e recomendações.

Existem diversas metodologias para se planejar um experimento. Balestrassi et al. (2009) utilizaram da metodologia de Taguchi e do fatorial fracionado para estimar os melhores parâmetros de RNAs visando a previsão de séries temporais. Magallanes e Olivieri (2010), atentam para o



fato de que a metodologia Plackett-Burman é extremamente econômica do ponto de vista de ensaios realizados.

A respeito desse delineamento, Montgomery (2009) observa que o mesmo é classificado como delineamento fatorial fracionado de dois níveis desenvolvido por Plackett e Burman (1946). Os arranjos experimentais são criados em tamanhos múltiplos de quatro e são compostos de sinais positivos para os níveis superiores, e sinais negativos para os níveis inferiores. Além disso, os projetos seguem a regra  $k = (N - 1)$ , em que  $k$  denota o número de parâmetros estudados e  $N$  é o total de ensaios realizados. Como ilustração, a tabela 2.1 apresenta uma matriz com doze ensaios,  $N = 12$  e  $k = 11$ , em que a primeira linha é o padrão para delineamentos Plackett-Burman com doze ensaios. A segunda linha é formada a partir da primeira, copiando-se a mesma com um deslocamento da coluna para a direita e assim sucessivamente até a última linha que é formada somente por sinais negativos. No delineamento de Plackett-Burman denota-se por  $+$  o elemento  $+1$  e por  $-$  o elemento  $-1$ . O delineamento de Plackett-Burman está associado as chamadas matrizes de Hadamard. Para mais detalhes a respeito das matrizes de Hadamard e do delineamento de Plackett-Burman veja o Apêndice B.

Tabela 2.1: Delineamento Plackett-Burman para doze ensaios

+	-	+	-	-	-	+	+	+	-	+
+	+	-	+	-	-	-	+	+	+	-
-	+	+	-	+	-	-	-	+	+	+
+	-	+	+	-	+	-	-	-	+	+
+	+	-	+	+	-	+	-	-	-	+
+	+	+	-	+	+	-	+	-	-	-
-	+	+	+	-	+	+	-	+	-	-
-	-	+	+	+	-	+	+	-	+	-
-	-	-	+	+	+	-	+	+	-	+
+	-	-	-	+	+	+	-	+	+	-
-	+	-	-	-	+	+	+	-	+	+
-	-	-	-	-	-	-	-	-	-	-

### 2.2.3 RNAs e Regressão Logística - Vantagens e Desvantagens

A principal vantagem das RNAs em relação às técnicas estatísticas tradicionais é que o modelo não tem a necessidade de ser explicitamente definido antes do início dos experimentos. As principais vantagens e desvantagens das RNAs em relação à regressão logística são apresentadas a seguir.

#### Vantagens

- Requerem de seus desenvolvedores menos conhecimento estatístico formal;
- Podem detectar relações não lineares complexas entre as variáveis independentes e dependentes;
- São capazes de identificar interações entre as variáveis preditoras;
- Podem ser desenvolvidas utilizando diferentes algoritmos de treinamento.

#### Desvantagens

- São como “caixas pretas”, o que limita na explicitação da influência de cada variável independente sobre a resposta. Seus parâmetros podem não ter interpretação estatística;
- Para outros usuários poderem utilizar um modelo treinado de RNAs, há a necessidade da cópia do *software* utilizado para o treinamento ou do conhecimento da matriz de pesos, a qual é difícil de se interpretar, ao contrário dos parâmetros da regressão logística;
- Requerem alto poder computacional;
- Possuem tendência ao sobreajuste, isto é, exigem cuidados para que não se especializem em um determinado conjunto de dados;

- O desenvolvimento de RNAs é realizado de forma empírica. Não existem metodologias definidas para seu projeto.

A tabela 2.2, apresenta a terminologia correspondente entre alguns termos estatísticos e alguns termos utilizados no estudo de RNAs.

Tabela 2.2: Terminologia correspondente entre estatística e RNAs

Estatística	RNAs
Modelo	Rede
Estimação	Aprendizagem
Regressão	Aprendizado supervisionado
Interpolação	Generalização
Observações	Conjunto de treinamento
Parâmetros	Pesos
Variáveis independentes	Entradas
Variáveis dependentes	Saídas
Regressão de Ridge	Decaimento dos pesos

No decorrer dos anos, tem havido um número crescente de artigos que exploram o uso das RNAs como uma metodologia promissora quando comparada à regressão logística. Uma revisão de 40 artigos desde 1996 até 2013 foi realizada e é resumida na tabela 2.3.

A primeira coluna apresenta o artigo revisado. A segunda coluna mostra a natureza do artigo. A terceira coluna apresenta os objetivos do artigo analisado e a quarta expõe qual metodologia obteve um melhor desempenho. Pelo resultado, infere-se que dentre os trabalhos analisados, as RNAs tiveram um desempenho superior à regressão logística em quase metade dos mesmos.

Tabela 2.3: Relação dos artigos que comparam técnicas

Artigo	Natureza	Objetivo	Desempenho (RNA x RL)
Tu (1996)	Conceitual	Prever resultados médicos	Inconclusiva
Schumacher et al. (1996)	Aplicação	Comparação geral de ambos os métodos	Inconclusiva
Vach et al. (1996)	Conceitual	Comparação geral de ambos os métodos	Inconclusiva
Freeman et al. (2000)	Aplicação	Prever morte hospitalar depois de angioplastia	Similar
Leung e Tran (2000)	Aplicação	Prever surto de doenças em camarões	RNA
Borque et al. (2001)	Aplicação	Prever o estágio patológico (pacientes sem próstata)	Similar
Chun et al. (2007)	Aplicação	Prever a probabilidade de câncer de próstata	RL
Kawakami et al. (2007)	Aplicação	Prever a probabilidade de câncer de próstata	RL
Ottenbacher et al. (2001)	Aplicação	Prever rehospitalização de pacientes com enfarto	Similar
Nguyen et al. (2002)	Aplicação	Prever morte ou amputação na doença meningocócica	Similar
DiRusso et al. (2002)	Aplicação	Analisar sobrevivência em traumas pediátricos	RNA
Dreiseitl e Ohno-Machado (2002)	Revisão	Comparação geral de ambos os métodos	RNA
Hajmeer e Basheer (2003)	Aplicação	Classificar crescimento de bactérias	RNA
Ottenbacher et al. (2004)	Aplicação	Prever cenário de vida depois de fratura no quadril	Similar
Lin et al. (2010)	Aplicação	Prever mortalidade em idosos com fratura no quadril	RNA
Ergün et al. (2004)	Aplicação	Classificar a estenose da carótida em pacientes com diabetes	RNA
Yesilnacar e Topal (2005)	Aplicação	Analisar susceptibilidade de deslizamento de terras	RNA
Yilmaz (2009)	Aplicação	Analisar susceptibilidade de deslizamento de terras	RNA
Pradhan e Lee (2010)	Aplicação	Analisar susceptibilidade de deslizamento de terras	RNA
Choi et al. (2012)	Aplicação	Analisar susceptibilidade de deslizamento de terras	RL
Song et al. (2005)	Aplicação	Diferenciar entre tumores benignos e malignos no peito	Similar
McLaren et al. (2009)	Aplicação	Detecção e diagnóstico de lesões no peito	Similar
Green et al. (2006)	Aplicação	Prever a síndrome coronária aguda	RNA
Chiang et al. (2006)	Aplicação	Diferenciar entre lojas online e lojas tradicionais	RNA
Liew et al. (2007)	Aplicação	Prever doença em pacientes após cirurgia bariátrica	RNA
Gutiérrez et al. (2008)	Aplicação	Mapear a infestação de <i>Ridolfia Segetum</i> (praga persistente)	Inconclusiva
Kurt et al. (2008)	Aplicação	Prever a doença cardíaca coronária	RNA
Al Housseini et al. (2009)	Aplicação	Prever o risco de cesariana em nulíparas	RNA
Caocci et al. (2010)	Aplicação	Prever a ocorrência da doença do enxerto contra o hospedeiro	RNA
Pavlekovic et al. (2010)	Aplicação	Reconhecer crianças talentosas em matemática	Similar
Trtica-Majnaric et al. (2010)	Aplicação	Prever a efetividade da vacinação contra gripe	RNA
Chen et al. (2012)	Aplicação	Diferenciar nódulos malignos e benignos no pulmão	RNA
Larasati et al. (2012)	Aplicação	Pesquisa psicológica	RNA
Pourshahriar (2012)	Aplicação	Pesquisa psicológica	Similar
Swiderski et al. (2012)	Aplicação	Avaliar a condição financeira de uma empresa	Inconclusiva
Askin e Gokalp (2013)	Aplicação	Avaliar os resultados de estudantes em matemática	Similar
Morteza et al. (2013)	Aplicação	Prever o nível de albuminúria em diabetes tipo 2	Inconclusiva
Vallejos e McKinnon (2013)	Aplicação	Classificar registros sísmicos	Similar

## 2.3 Comparação do Comportamento dos Classificadores

Nesta seção, aborda-se uma técnica para organizar, visualizar, avaliar e selecionar classificadores baseado em seus desempenhos.

Segundo Dreiseitl e Ohno-Machado (2002), o poder discriminatório é a análise utilizada para comparar o desempenho de classificadores. A discriminação é uma medida de quão bem as duas classes do conjunto de dados são separadas em consequência das características de entrada.

De acordo com Ben-David (2008), a acurácia é a primeira medida para se avaliar o desempenho de um classificador. Entretanto, é uma medida problemática porque não compensa os sucessos que ocorrem devido ao acaso. Por exemplo, em um problema de classificação binária em que as probabilidades de sucesso e fracasso sejam as mesmas, um classificador que acerta metade das classificações em metade dos casos tem o mesmo desempenho que aquele que um classificador aleatório (por exemplo que julgue sempre todas as saídas como sucesso). Ou seja, ambos os classificadores não tem utilidade. O autor atenta para o fato que existem alternativas de medidas que consideram a aleatoriedade e uma das destacadas é a curva da característica de operação do receptor (*receiver operating characteristic* ou ROC). Na maioria dos trabalhos relacionados com regressão logística e RNAs, a área sob a curva ROC foi a métrica utilizada para avaliar o poder discriminatório dos classificadores.

### 2.3.1 Curvas ROC

Segundo Fawcett (2006), a curva ROC foi desenvolvida no contexto de detecção de sinais eletrônicos e problemas com radares durante a segunda guerra mundial. Seu objetivo era quantificar a habilidade dos operadores de radares (chamados originalmente de *receiver operators*) em distinguir um sinal de um ruído. Esta habilidade era chamada de *receiver operating characteristic*.

Quando o radar detectava algo se aproximando, cabia ao operador decidir se o que foi captado era, por exemplo, um avião inimigo ou outro objeto voador sem importância.

A curva ROC é uma demonstração bidimensional do desempenho de um classificador. A comunidade médica possui uma vasta literatura, nos mais diferentes campos, sobre o uso dos gráficos ROC para testes de diagnósticos. Conceitualmente, são simples mas há complexidades que surgem quando aplicados em pesquisa.

Fawcett (2006) afirma que dado um classificador e um exemplo, há quatro possíveis saídas. Se o exemplo é positivo e é classificado como positivo, tem-se um positivo verdadeiro; se é classificado como negativo é conhecido como falso negativo. Agora, se o exemplo é negativo e é classificado como negativo, tem-se um negativo verdadeiro e se é classificado como positivo, tem-se um falso positivo. A tabela 2.4 apresenta uma matriz de confusão.

Tabela 2.4: Matriz de confusão

	Valor Verdadeiro	
	Positivo	Negativo
Positivo	Verdadeiro Positivo	Falso Positivo
Negativo	Falso Negativo	Verdadeiro Negativo

Os gráficos ROC são gráficos bidimensionais nos quais a sensibilidade (ou a taxa de verdadeiros positivos) é plotada no eixo y e a taxa de falsos positivos (ou 1 - especificidade) é plotada no eixo x. A figura 2.2 mostra um gráfico ROC com cinco classificadores discretos nomeados de A até E.

Diversos pontos num gráfico ROC são importantes. O ponto inferior esquerdo (0,0) representa a estratégia de nunca classificar um caso como positivo. Esse classificador não comete erros falsos positivos mas também não faz acertos verdadeiros positivos. A estratégia oposta, de somente conceder classificações positivas é representada pelo ponto superior direito (1,1). O ponto (0,1)

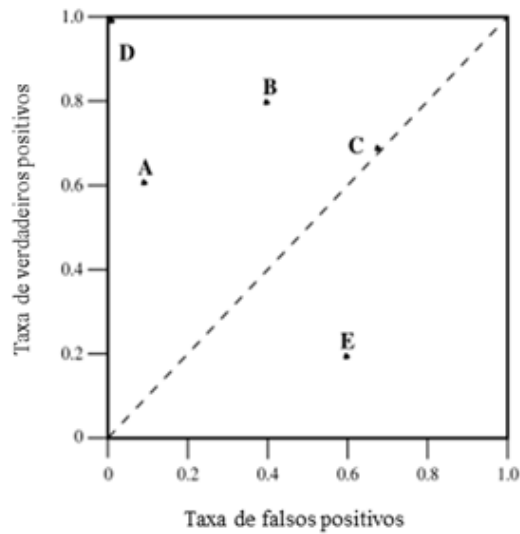


Figura 2.2: Curva ROC com cinco classificadores discretos - Fonte Fawcett (2006)

representa classificação perfeita. O desempenho de D é perfeito como se vê. Um ponto é melhor que outro quanto mais a noroeste ele está no gráfico (taxa de verdadeiros positivos maior e taxa de falsos positivos menor). A medida que o ponto vai descendo no gráfico ele diminui seu desempenho. A e B tem desempenhos similares e satisfatórios, o desempenho de C é exatamente igual a do acaso e o desempenho de E é pior do que aquele encontrado pelo acaso.

Efetivamente, para comparar classificadores, o que se faz é calcular a área abaixo da curva ROC, conhecida como AUC (*area under the curve*). Como a AUC é uma porção de uma área de uma unidade quadrada, seu valor será sempre entre 0 e 1. Entretanto, porque adivinhações aleatórias produzem a linha diagonal entre (0,0) e (1,1), que tem uma área de 0,5, nenhum classificador realístico deve ter uma AUC menor do que 0,5.

A área sob a curva de um classificador tem uma importante propriedade estatística: ela é equivalente à probabilidade do classificador apresentar uma observação positiva escolhida ao acaso com um valor superior à uma observação negativa escolhida ao acaso.

A figura 2.3 apresenta as áreas sob duas curvas ROC, A e B. O classificador B tem uma área maior e, portanto, uma performance média superior, apesar de ser pior do que A em determinadas regiões.

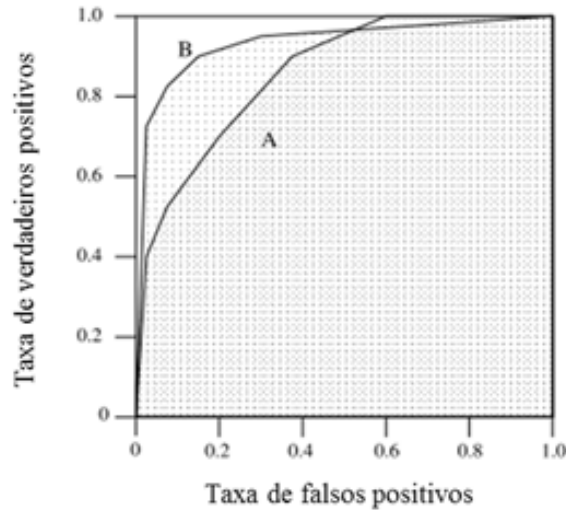


Figura 2.3: Área sob a curva de dois classificadores probabilísticos -Fonte Fawcett (2006)

## 2.4 O Método *Desirability*

O método *Desirability* é um algoritmo criado originalmente por Harrington (1965) e, posteriormente aprimorado por Derringer e Suich (1980), para tratar da otimização simultânea dos modelos de múltiplas respostas.

Suponha-se que cada uma das  $k$  variáveis resposta de um determinado processo estejam relacionadas com  $p$  variáveis independentes por

$$Y_{ij} = f_i(X_1, \dots, X_p) + \varepsilon_{ij} \quad (2.22)$$

$i = 1, \dots, k$  e  $j = 1, \dots, n_i$  em que  $f_i$  denota a relação funcional entre  $Y_i$  e  $X_1, \dots, X_p$ . Se



$E(\varepsilon_{ij}) = 0$  para todo  $i$ , então

$$n_i = f_i(X_1, \dots, X_p), \quad i = 1, \dots, k \quad (2.23)$$

A função *desirability* envolve a transformação de cada variável resposta estimada  $\hat{Y}_i$  para o valor *desirability* denotado por  $d_i$  em que  $0 \leq d_i \leq 1$ . O valor de  $d_i$  aumenta quando a  $i$ -ésima resposta se aproxima dos limites impostos.

A equação 2.24 é utilizada para se encontrar o índice global  $D$ , a partir da combinação de cada uma das respostas transformadas através de uma média geométrica

$$D = (d_1 d_2 \dots d_k)^{\frac{1}{k}} \quad (2.24)$$

A equação 2.24 leva em consideração o fato das respostas terem o mesmo grau de importância. Quando os graus de importância das respostas forem distintos, usa-se a equação

$$D = \left[ \prod_{i=1}^n d_i^{w_i} \right]^{\frac{1}{w}} \quad (2.25)$$

em que  $w_i$  denota o grau de importância de cada resposta e  $w$  a soma dos graus de importância.

O valor  $D$  avalia de maneira geral os níveis do conjunto combinado de respostas. É um índice também pertencente ao intervalo  $[0, 1]$  e será maximizado quando todas as respostas se aproximarem o máximo possível de suas especificações. O ponto ótimo geral do sistema é o ponto ótimo alcançado pela maximização da equação 2.24, calculada a partir das funções *desirability* individuais.

A utilização da média geométrica tem a vantagem de que a solução global seja alcançada de maneira balanceada, permitindo que todas as respostas atinjam os valores esperados e forçando o algoritmo a se aproximar das especificações impostas. Caso isso não seja possível, o algoritmo

retorna uma solução inviável para o problema. Na essência esse método condensa um problema de otimização multivariada em um problema univariado. A desvantagem, segundo Khuri e Conlon (1981) é que nessa transformação, a estrutura de variância-covariância das respostas é totalmente ignorada.

O algoritmo de Derringer e Suich (1980) depende do tipo de otimização desejada para a resposta (maximização, normalização, minimização), dos limites (valores desejados) e das respectivas importâncias de cada resposta.

De acordo com Derringer e Suich (1980), quando se deseja a maximização de uma resposta, a fórmula de transformação empregada é

$$d_i = \begin{cases} 0 & \text{se } \hat{Y}_i \leq Y_{i*} \\ \left( \frac{\hat{Y}_i - Y_{i*}}{Y_i^* - Y_{i*}} \right)^r & \text{se } Y_{i*} < \hat{Y}_i < Y_i^* \\ 1 & \text{se } \hat{Y}_i \geq Y_i^* \end{cases} \quad (2.26)$$

em que  $Y_{i*}$  e  $Y_i^*$  são, respectivamente, o maior e o menor valor aceitável para a  $i$ -ésima resposta. O valor de  $r$  na equação 2.26 denota a preponderância dos limites. Valores maiores do que a unidade devem ser usados quando a resposta cresce rapidamente acima de  $Y_{i*}$ . Portanto,  $d_i$  aumenta vagarosamente, enquanto o valor da resposta vai sendo maximizado. Logo para se maximizar  $D$ , a  $i$ -ésima resposta deve ser bem maior do que  $Y_{i*}$ .

Por outro lado, quando for necessário minimizar a variável resposta, basta considerar o fato que minimizar  $\hat{Y}_i$  equivale a maximizar  $-\hat{Y}_i$ .

As transformações bilaterais ocorrem quando a resposta de interesse apresenta duas restrições: uma de máximo e outra de mínimo. Para esse caso, a fórmula de transformação é dada por

$$d_i = \begin{cases} 0, & \text{se } \hat{Y}_i < Y_{i*} \text{ ou } \hat{Y}_i > Y_i^* \\ \left(\frac{\hat{Y}_i - Y_i^*}{c_i - Y_i^*}\right)^t & \text{se } c_i < \hat{Y}_i < Y_i^* \\ \left(\frac{\hat{Y}_i - Y_{i*}}{c_i - Y_{i*}}\right)^s & \text{se } Y_{i*} < \hat{Y}_i < c_i \end{cases} \quad (2.27)$$

Em outras palavras, o método *desirability* está baseado em construir uma relação entre as variáveis resposta com as variáveis independentes, aliado a transformação unilateral ou bilateral definidas anteriormente. Isso significa dizer que cada variável resposta do conjunto original se transforma em uma função *desirability* individual  $d_i$  que varia de 0 a 1.

No intervalo de 0 a 1, existem faixas de variação que determina o nível de qualidade da variável resposta desejada. A tabela 2.5 apresenta as faixas de variação de  $d_i$  e a qualidade da resposta associada a faixa.

Tabela 2.5: Relação entre a variação  $d_i$  com o significado da variável resposta

Variação de $d_i$	Significado
$0 \leq d_i < 0,30$	A resposta está fora da faixa adequada
$0,30 \leq d_i < 0,40$	A resposta está no limite de aceitação
$0,40 \leq d_i < 0,63$	A resposta está aceitável, porém fraca
$0,63 \leq d_i < 0,80$	A resposta está adequada
$0,80 \leq d_i < 1,0$	A resposta está excelente
$d_i = 1$	A resposta está no seu valor desejável

Com o uso da função *desirability* é possível direcionar as variáveis respostas a seus respectivos objetivos individuais, ou seja, pode-se maximizar algumas e ao mesmo tempo minimizar outras e ainda, delimitar o valor alvo das demais respostas.

## Capítulo 3

# Falhas em Estruturas Metálicas de Linhas de Transmissão

Neste capítulo aborda-se o problema de queda de estruturas metálicas de uma linha de transmissão.

### 3.1 Introdução

Segundo Wazen (2011), linhas de transmissão aéreas estão expostas a diversos riscos associados ao meio, a alterações de características construtivas e variações climáticas. Muitas vezes, esses riscos podem trazer sérios danos, implicando em quedas de estruturas. A queda de uma estrutura pode causar a interrupção do fornecimento de energia de uma localidade por longo período, bem como gerar custos associados na reconstrução de trechos do sistema elétrico, no lucro cessante da concessionária e em indenizações decorrentes da falta de energia.

A intensidade dos ventos incrementa as estatísticas de acidentes com estruturas e sobrepõem tanto as técnicas de projeto quanto as de manutenção. Entretanto, torna-se imprecindível que as

empresas responsáveis pela transmissão de energia elétrica, envolvendo as linhas de transmissão, mais especificamente, tenham planos de contingência montados através de parâmetros padronizados e que funcionem de acordo com a criticidade do sistema. Devido à importância das linhas aéreas, uma análise quantitativa de suas características com o intuito de identificá-las e mitigá-las tem muito a contribuir para o planejamento, operação e manutenção de linhas.

## 3.2 Linhas de Transmissão

Linhas de Transmissão são circuitos que fazem a interligação entre subestações, usinas ou distribuidores de energia. Esses circuitos são compostos por torres autoportantes ou mesmo postes, além de cabos metálicos para permitir o fluxo da energia. Sua principal função é o transporte de grandes volumes de energia elétrica com a menor perda possível.

Os sistemas de transmissão podem ser em corrente alternada ou corrente contínua. Dentre os sistemas, o mais utilizado no Brasil é o de corrente alternada e sua aplicação ocorre em circuitos trifásicos, possuindo atualmente apenas uma linha de transmissão de alta tensão em corrente contínua. Essas são composições para interligação de energia no país, entre diversos centros consumidores, bem como para alimentação de grandes instalações industriais. O sistema brasileiro transmite tensões de 69, 88, 138, 230, 345, 460, 525 e 750kV. A tabela 3.1, extraída do Anuário Estatístico de Energia Elétrica (2014), apresenta a extensão, em quilômetros, das linhas de transmissão do sistema integrado nacional.

Os valores da tabela 3.1 referem-se a rede básica (instalações com tensão maior do que ou igual a 230 kV).

A padronização de tensões faz-se necessária devido à uniformização de equipamentos ( como chaves seccionadoras, disjuntores e transformadores). Assim, o nível de tensão da linha se refere

Tabela 3.1: Extensão das linhas de transmissão do SIN

	2008	2009	2010	2011	2012	2013
230 kV	37.710	41.437	43.185	45.185	47.858	49969
345 kV	9.772	9.784	10.060	10.062	10.224	10272
440 kV	6.6671	6.671	6.671	6.681	6.728	6728
500 kV	31.868	33.196	34.356	35.003	35.726	39123
600 kV	3.224	3.224	3.224	3.224	3.224	7992
750 kV	2.683	2.683	2.683	2.683	2.683	2683
Total	91.928	96.995	100.179	103.362	106.444	116 768

ao valor da tensão entre fases, enquanto os equipamentos são classificados entre o referencial de fase e terra (potencial zero).

A fim de se analisar convenientemente as causas para a descontinuidade de transmissão de energia devido a contingências nas linhas de transmissão provocadas por fatores externos, faz-se necessário uma descrição dos tipos e configurações de estruturas, cabos e fundações, bem como o dimensionamento de uma linha de transmissão para que então se possa analisar e descrever os tipos comuns de quedas e suas causas físicas.

### 3.2.1 Estruturas

As dimensões e formas das estruturas dependem da disposição necessária dos condutores, da distância entre os condutores, das dimensões e forma de isolamento, das flechas projetadas para os condutores, da altura mínima de segurança, além do número de circuitos envolvidos.

O dimensionamento de uma estrutura de linha de transmissão depende tanto da carga que será transportada (associado diretamente à capacidade e comportamento dos cabos) como das

dimensões da própria estrutura a ser utilizada. Esses valores são calculados em função de considerações de segurança e do desempenho em face aos valores de sobretensão e esforços a que as estruturas serão submetidas.

Segundo a ABNT (1985), as distâncias de segurança são os afastamentos mínimos recomendados do condutor e seus acessórios energizados a quaisquer partes. Os valores mínimos devem ser determinados em função de estudos que levam em consideração as várias solicitações elétricas a que a linha será submetida, devidamente coordenada com as condições de vento que ocorrem simultaneamente com cada uma das referidas solicitações.

As estruturas possuem uma isolação em relação aos cabos feita por cadeias de isoladores e as distâncias de segurança também levam em conta a capacidade de isolação do próprio ar. Estas distâncias levam em conta as partes energizadas e as partes das estruturas que realizam os suportes dos cabos, além de considerar a distância mínima ao solo e aos objetos que possam estar próximos aos cabos.

Além da informação de distância de segurança, existe também a preocupação quanto a seu desempenho quando exposta a descargas atmosféricas. É preciso ter em mente que quanto maior a estrutura, mais sujeita a sofrer com descargas ela estará.

Assim, para estruturas de tensões inferiores a  $230[kV]$ , devido às restrições de condutividade de correntes injetadas por descarga, estas deverão ter capacidade de dissipar a energia incidida através de seu sistema de proteção contra descargas (conjunto de cabos, pára raios e aterramento).

Em função do citado problema, uma série de estruturas com dimensões reduzidas foram desenvolvidas nas duas últimas décadas a fim de limitar principalmente a altura e readequando os cabos a esta nova situação. Tais estruturas são denominadas estruturas compactas.

As estruturas compactas também colaboram com a estética das regiões onde são implantadas,

já que seu impacto é inferior ao das estruturas convencionais. Devido aos diversos pontos que devem ser observados, para a determinação dos tipos de estrutura a serem aplicados, pode-se classificar as estruturas em alguns critérios, como:

- *Quanto a sua função.* A função de uma estrutura numa linha de transmissão está associada aos esforços aos quais cada estrutura deverá ser submetida. Esses esforços podem possuir componentes verticais (associados ao peso dos cabos, das cadeias de isoladores e ferragens), componentes horizontais transversais (tração e vento sobre os cabos e equipamentos) e cargas horizontais longitudinais (esforços no sentido da linha). Assim, classificam-se as estruturas como sendo: *Suspensão*: estruturas sujeitas a esforços de componentes verticais e horizontais transversais e *Ancoragem*: estruturas sujeitas a esforços verticais, horizontais transversais e longitudinais. Essas estruturas podem suportar ângulos de altos valores, pois apresentam as cadeias de isoladores ancoradas no corpo da estrutura.
- *Quanto a sua resistência.* Nesta classificação, as estruturas podem ser classificadas quanto ao modo que suportam as componentes incididas sobre elas, ou seja, de que forma ela transfere os esforços sofridos aos materiais aplicados e a região na qual foi realizada sua locação. Segundo Fuchs (1977), as estruturas podem ser classificadas em dois grandes grupos: *Autoportante* : suas ferragens são capazes de suportar todo o esforço aplicado sobre a mesma, transferindo suas componentes verticais e horizontais sobre sua base e distribuindo sobre a área do solo no qual está em contato. *Estaiada* : para a melhor distribuição de esforços, utiliza estais conectados entre pontos de fixação na estrutura e o solo. Desta forma, os pontos mais suscetíveis são neutralizados por cabos metálicos ancorados na terra, ou seja, mais pontos são presos ao solo, gerando mais pontos de distribuição de esforços mecânicos.



- *Quanto a sua composição.* Com relação a composição, as estruturas classificam-se em três grupos. *Metálica:* são geralmente fabricadas com aço-carbono, normais ou de alta resistência, perfiladas ou tubulares. A composição em peça permite a montagem e o relativamente fácil transporte. A galvanização das peças garante a vida útil acima dos 25 anos. No passado, o Alumínio e suas ligas puderam ser utilizados para material estrutural, com sensível redução de peso, porém com grande perda em resistência mecânica. Este tipo de estrutura está amplamente associado a regiões de baixa densidade populacional, devido necessidade de larga faixa de servidão (que varia de acordo com o nível de tensão). *Concreto:* pode ser produzida tanto no processo de centrifugação como de vibração, porém devido ao material empregado, este deve ser fabricado já com o corpo inteiro, dificultando seu transporte e sua montagem. O poste centrifugado apresenta melhor acabamento, necessitando de instalações mais elaboradas para fabricação. O poste vibrado necessita de instalações mais modestas para ser fabricado, sendo que os postes apresentam maior robustez. Os postes feitos em processo de vibração podem ser aplicados nas mesmas situações das torres metálicas com menor custo, porém com aplicação em vãos mais curtos. Já os postes centrifugados, possuem grande aplicação em regiões urbanas, já que são mais leves e requerem uma faixa de servidão bem reduzida, causando menor impacto na área instalada. *Madeira :* apesar de ser feito de material de fácil extração e baixo custo, o poste de madeira não possui grande resistência mecânica. Foram muito utilizados até os anos 90, mas com a elevação dos níveis de tensão aplicados, os esforços requeridos aumentaram, perdendo espaço nas novas construções de linhas.

A escolha do tipo de estrutura é influenciada pelo projeto de uma linha específica, onde o custo dos materiais e mão de obra, os terrenos, a capacidade, a confiabilidade e a manutenção

interferem diretamente na decisão. Por essa razão, a configuração básica é selecionada em função dos seguintes parâmetros:

1. Espaçamento entre fases;
2. Configuração das cadeias de isoladores;
3. Definição dos cabos condutores das fases;
4. Distanciamento necessário entre os condutores e o pára-raios;
5. Ângulo da estrutura em relação ao eixo da linha de transmissão;
6. Folgas necessárias dos componentes energizados, para a realização de manutenções em condições normais de operação;
7. Folga necessária devido à posição do cabo condutor na estrutura, causado por desníveis entre estruturas e ângulos acentuados na estrutura;
8. Folga necessária para a utilização de amortecedores ou pesos adicionais nos cabos, em função de correção de catenárias;
9. Folgas adicionais necessárias para manutenções específicas.

Fuchs (1977) afirma que os tipos e configurações das estruturas em uso são variados. Os projetos de estrutura não estão limitados aos modelos já aplicados. Mas para a definição de um novo modelo, é necessário um grande número de informações para encontrar uma configuração eficiente e que não seja aplicado exclusivamente a uma estrutura de uma série de linhas projetadas. A aplicação de materiais adequados, bem como os níveis de tensão, acaba por tornar ainda mais difícil a definição de estruturas padronizadas.

### 3.2.2 Cabos

De acordo com Fuchs (1977) os cabos podem ser diferenciados pelas diversas funções que possuem numa linha de transmissão, onde podem ser condutores de energia, protetores contra surtos atmosféricos e sobrecorrentes, ou mesmo dissipadores de energia em partes aterradas. O cabo condutor pode ser chamado de parte ativa de uma linha de transmissão, pois serve de guia dos campos elétrico e magnético.

Os materiais dos cabos, geralmente utilizados são de alumínio, cobre e aço galvanizado. Nos cabos condutores, a grande maioria é feita de alumínio possuindo núcleo de aço galvanizado (alma de aço).

Devido ao alto custo para aplicação, cada vez mais os cabos de cobre vem sendo substituídos nas linhas de transmissão. Devido ao valor agregado, muitas linhas de transmissão vêm sofrendo com roubo deste tipo de material para revenda no mercado informal, trazendo grandes prejuízos a empresas do setor elétrico. Mesmo assim, ainda vem sendo utilizados em antigas instalações, mas com pouca perspectiva de continuar em atividade.

No caso dos cabos de alumínio, os materiais utilizados vêm sendo mais aplicados, mas não em sua forma pura, mas em ligas de alumínio aço e ligas de alumínio termo resistente.

Os condutores das linhas aéreas de transmissão podem ser considerados suficientemente flexíveis, quando os pontos onde suas extremidades estão fixadas, estiverem afastados entre si acima de dezenas de metros. Verifica-se sua proximidade com uma catenária. A catenária é uma curva plana que pode ser imaginada como uma corda suspensa pelas suas extremidades e sujeitas à ação da gravidade. Sua equação é dada por

$$y = \cosh\left(\frac{x}{a}\right) \quad (3.1)$$

### 3.2.3 Fundações

Segundo Wazen (2011) as fundações têm por finalidade equilibrar a ação das forças atuantes sobre a parte superior da estrutura e seus equipamentos, e para isto deve levar em conta o tipo de solo onde esta será implantada. Com isto, é necessária realizar uma conciliação destes fatores para que as fundações possam ser mais bem aplicadas, sem gerar excessos de material ou dirimir as forças a qual estará exposta.

- *Fundações de concreto.* Algumas concessionárias utilizam concreto em algumas de suas fundações, sendo que sua utilização depende do preparo do concreto. Para este preparo os concretos devem ser misturados mecanicamente conforme as quantidades estipuladas pelo projeto da estrutura. A quantidade de concreto preparada em cada operação é estritamente necessária para o seu emprego imediato. Os concretos frescos devem sofrer a menor distância possível de transporte e serem lançados logo após a mistura e amassamento. Nos casos de uso de impermeabilizante para concretagens com presença de água, este intervalo para lançamento deve ser extremamente curto, isto para que a massa seja praticamente uniforme. O trabalho com concreto está sempre sujeito a problemas construtivos, desde que o processo de mistura, lançamento e secagem não sofram alterações. Considerando estas situações básicas para a aplicação de concretos, seguem os tipos de fundação aplicados em linhas de transmissão. Dentre os tipos de fundação, destacam-se o *tubulão*, tipo de fundação mais profunda, escavada com pá, picareta, ou trado. Depois de colocada a armadura para o concreto, é realizada a concretagem. Geralmente, usam-se manilhas de concreto armado, para efetuar o escoramento do fuste (camada superficial do concreto) e a *sapata de concreto armado* que são fundações rasas de concreto armado de pequena altura em relação às dimensões da base. Elas são normalmente de base quadrada e de fuste inclinado ou vertical.

No projeto, são utilizados fustes de comprimentos variáveis para facilitar a montagem das estruturas em locais de pés desnivelados. Há a necessidade de armação da face inferior e superior da sapata, em função de os esforços na torre serem ora de compressão, ora de tração.

- *Fundação metálica.* As fundações metálicas são utilizadas em regiões onde o solo possui boa coesão. As fundações puramente metálicas aplicadas em estruturas de linhas de transmissão possuem formato piramidal de forma que a conexão com a estrutura seja feita pelo topo da pirâmide e toda sua área interna seja vazada e preenchida com terra. Este formato faz com que o próprio solo seja uma barreira mecânica que evite que a base da fundação (base da pirâmide) possa vir a subir ou mesmo se deslocar lateralmente. As variações neste tipo de fundação ocorrem quanto a dimensão da base, altura do conjunto, dimensões das longarinas e chapas metálicas a serem aplicadas.
- *Grelha metálica.* A grelha metálica é classificada como fundação rasa, ligada ao pé (montante) da torre, ao nível do terreno. A base da fundação (grelha propriamente dita) é constituída por travessas (perfil U) e por cantoneiras. São fundações rasas com 2 a 4 metros de profundidade, indicadas para terrenos argilosos, arenosos ou siltosos, porém secos e com resistência crescente com a profundidade e com possibilidade de serem escavados a céu aberto.

### **3.2.4 Atributos das Estruturas**

Os atributos selecionados para o problema de queda de estruturas metálicas de linhas de transmissão são: tensão de operação, vento na região, relevo do local, massas de ar, temperatura da região, tipo do terreno, capacidade mecânica da estrutura, função e tipo de fundação da

estrutura.

- Tensão de operação: Os objetos que constituem esse grupo são {69, 138, 230, 550} todos medidos em  $kV$ .
- Vento da região: O atributo vento possui as seguintes faixas {16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26} todos medidos em  $km/h$ .
- Relevo do local: Os atributos dessa variável são planalto, planície, serra e vale.
- Massas de ar (quente ou fria): Para essa variável os atributos são paralela, perpendicular e transversal.
- Temperatura: Os objetos que constituem esse grupo são {16, 17, 18, 19, 20, 21, 22, 23} todos medidos em graus Celsius.
- Tipos de terrenos: De acordo com a NBR 5422, o terreno pode ser diferenciado em quatro categorias, de acordo com os coeficientes de rugosidade definidos para este, e que podem ser definidos por: A) terrenos com vastas extensões de água, áreas planas costeiras e desertos planos; B) terreno aberto com poucos obstáculos; C) terreno com obstáculos numerosos e pequenos; D) áreas urbanizadas e terrenos com muitas árvores altas.
- Capacidade mecânica da estrutura: Com o passar do tempo diversas linhas de transmissão sofreram alterações e estas podem implicar em alterar as características de projeto de aplicação de algumas estruturas. Conforme decisão do projetista algumas são substituídas e outras não, porém podendo sofrer aumento nos esforços a que estará submetida. Os projetos de estruturas até a década de 1980 consideravam ventos com velocidade máxima de  $80 km/h$  ( para estimativa de 50 anos) e o cálculo dos esforços laterais eram sobre a força gerada por

essa velocidade. Após essa década, os novos projetos de estruturas já contemplavam uma atualização nesse valor passando a considerar ventos de  $130 \text{ km/h}$  para a mesma estimativa. Os atributos para essa variável são alta e baixa, considerando as situações descritas anteriormente.

- Função: Para essa variável os atributos são suspensão e ancoragem.
- Fundação: Para essa variável os atributos são *stub*, grelha e estaca.

### 3.2.5 Quedas de Estruturas

Linhas de transmissão transportam grande volume de energia elétrica aos centros de distribuição (consumidores) utilizando cabos metálicos (de cobre ou alumínio) sustentados por estruturas feitas de concreto ou de material metálico. Cada tipo de estrutura possui características próprias de suportabilidade a esforços mecânicos. Nestas, os esforços se dão no sentido longitudinal dos cabos, na componente vertical (associada ao peso) e no sentido transversal aos cabos, ambos laterais às estruturas.

Todo tipo de equipamento está suscetível a sofrer esforços superiores ao que é projetado, podendo gerar uma falha funcional, ou seja, este equipamento deixa de exercer sua função. No caso das linhas de transmissão, existem equipamentos que supervisionam e controlam o transporte de energia, sinalizando a ocorrência de interrupções para atuação das equipes responsáveis por seu restabelecimento.

Além da supervisão feita pela equipe de operação, existe uma equipe de manutenção que realiza inspeções detalhas e de patrulhamento ao longo das linhas. Quando as equipes de manutenção são solicitadas, devido à sinalização de falha, estas realizam um patrulhamento, no qual é possível identificar visualmente qual componente sofreu falha funcional. E com as informações levantadas,

as equipes podem atuar em sua recuperação e disponibilização para retornar à operação normal.

Dentre as possíveis causas de falhas podem ser citadas: curto-circuito fase terra por quebra de isoladores ou por descargas atmosféricas, curto-circuito fase terra por contato de objetos estranhos (árvores, arames, pipas), rompimento de cabos pára raios com ocorrência de curto circuito fase terra, rompimento de condutores com ocorrência de curto-circuito fase-terra ou fase-fase e mesmo quedas de estruturas. Tais situações podem ser causadas por atos voluntários ou involuntários, bem como por questões climáticas (temporais e vendavais). Identificada umas das situações acima, o restabelecimento da carga deverá ser no menor tempo possível e a situação toma características de emergência.

A figura 3.1 apresenta uma imagem de uma estrutura metálica que caiu após um vendaval, localizada na região oeste do Paraná. Nesta imagem, fica claro que, após a queda, os cabos interromperam o fluxo de veículos em uma via.



Figura 3.1: Queda de torre no Oeste do Paraná - Fonte Portal Medianeira



# Capítulo 4

## Avaliação da Probabilidade de Ocorrência de Queda de Estruturas Metálicas de Linhas de Transmissão

Neste capítulo, avalia-se a probabilidade de ocorrência de queda de uma estrutura metálica de linha de transmissão utilizando duas técnicas distintas: regressão logística e RNAs. As estruturas foram selecionadas de forma aleatória.

### 4.1 Introdução

O único trabalho encontrado, nas bases de conhecimento que tratam de linhas de transmissão, a respeito da avaliação de suscetibilidade das estruturas quanto a quedas, foi o trabalho de Wazen (2011).

Utilizando a metodologia de regressão logística e RNAs combinadas com DOE, pretende-se extrair conhecimento sobre quais parâmetros e variáveis influenciam o comportamento mecânico

das linhas em operação e podem ser utilizados para diagnosticar possíveis quedas de torres. Estas informações podem se tornar uma base para direcionar os investimentos de reforço de estruturas, evitando a ocorrência de longos desligamentos e altos custos por decorrência de danos em torres de linhas de transmissão.

Essas análises apresentam como resultado esperado pela combinação das informações de cada torre se há possibilidade de queda, em situações normais de operação, ou praticamente não existe este risco. Obviamente, todas as estruturas em operação tem o risco de sofrerem danos e por isto, a avaliação é feita considerando condições de operação, sem considerar fatores que possam acelerar ou causar a queda de uma estrutura.

Para a classificação das linhas de transmissão quanto a suscetibilidades de falhas propõe-se identificar as linhas mais propensas levando em consideração os seguintes parâmetros: velocidade do vento da região da estrutura, vegetação em seu entorno, pressão atmosférica e variação de temperatura da região, além da capacidade mecânica máxima e do tipo de fundação da estrutura.

## 4.2 Análise Descritiva do Conjunto de Dados

Nesta seção, apresenta-se uma análise descritiva das variáveis explicativas do conjunto de dados do apêndice A. Ao observar uma extensa relação de dados coletados, não é possível obter praticamente nenhuma conclusão, especialmente para grandes conjuntos de dados, com muitas características sendo investigadas.

A análise descritiva é a etapa inicial do processo de estudo de dados coletados. A descrição dos dados tem como objetivo identificar anomalias, até mesmo resultante do registro incorreto de valores, e dados dispersos, aqueles que não seguem a tendência geral do conjunto. As ferramentas descritivas aqui utilizadas são gráficos de barras e *box-plot*.

Utilizou-se um banco de dados do histórico de quedas de estruturas metálicas de linhas de transmissão de um dado sistema elétrico mantido pela COPEL (2015) composto de 181 observações. Esse banco de dados contém dez variáveis independentes (tensão, vento, relevo, massa fria, massa quente, temperatura, terreno, capacidade, função e fundação) e uma variável dependente chamada resultado, que informa se houve ou não queda da estrutura metálica. Do conjunto de dez variáveis independentes as variáveis tensão, vento e temperatura são variáveis quantitativas enquanto que as demais são variáveis qualitativas.

Os gráficos de barras e os *box-plot* do conjunto de dados, foram obtidos através do *software* R. Para os gráficos de barras, o eixo vertical apresenta a quantidade de torres para cada nível da variável especificada. O *box-plot* é um gráfico utilizado para avaliar a distribuição empírica do dados. O *box-plot* é formado pelo primeiro e terceiro quartil e pela mediana.

A figura 4.1 apresenta o número de estruturas metálicas consideradas com capacidade alta e com capacidade baixa. Observa-se que mais da metade das estruturas analisadas possuem capacidade considerada baixa.

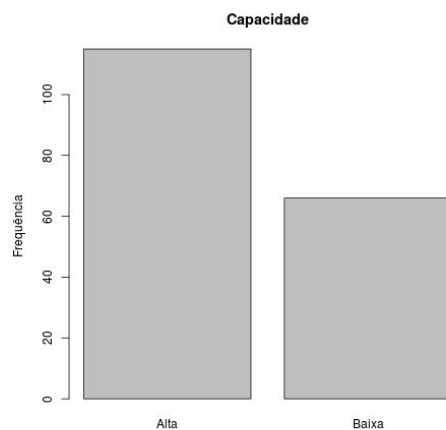


Figura 4.1: Gráfico ilustrando o número de estruturas com dois tipos de capacidade

De acordo com a NBR 5422, um terreno é classificado em quatro diferentes modalidades.

A figura 4.2 apresenta o número de estruturas com quatro diferentes tipos de terrenos.

Observa-se que a maior parte das estruturas foram colocadas em terrenos no tipo B.

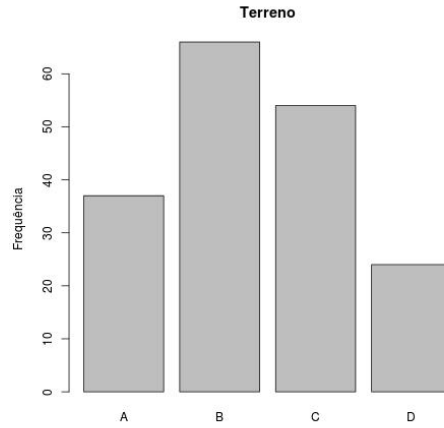


Figura 4.2: Gráfico ilustrando o número de estruturas com quatro diferentes tipos de terrenos

A figura 4.3 apresenta o número de estruturas com diferentes tipos de massas de ar quente.

Convém enfatizar que, se a variável massa de ar quente, assume, por exemplo, o nível paralela, significa que a massa de ar quente age paralelamente ao sentido dos cabos.

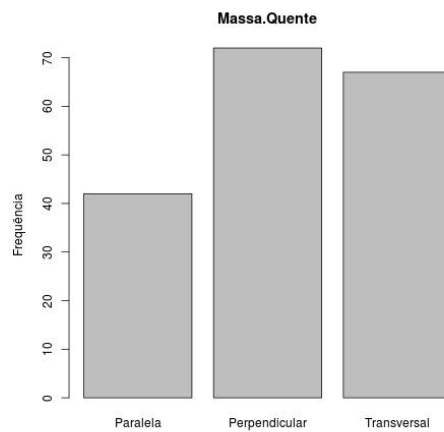


Figura 4.3: Gráfico ilustrando o número de estruturas com diferentes massas de ar quente

A figura 4.4 apresenta o número de estruturas com diferentes tipos de massas de ar fria.

Analogamente a variável massa de ar quente, se a variável massa de ar fria, assume, por exemplo, o nível paralela, significa que a massa de ar fria age paralelamente ao sentido dos cabos.

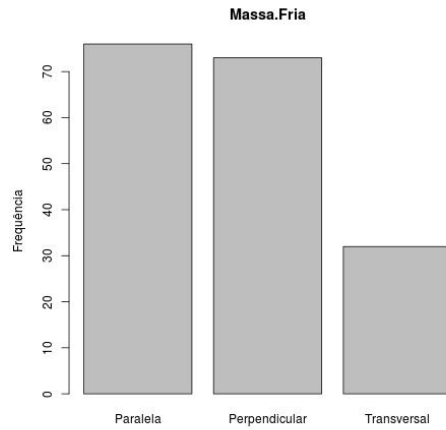


Figura 4.4: Gráfico ilustrando o número de estruturas com diferentes tipos de massas de ar fria

A figura 4.5 apresenta o número de estruturas com diferentes níveis de tensão.

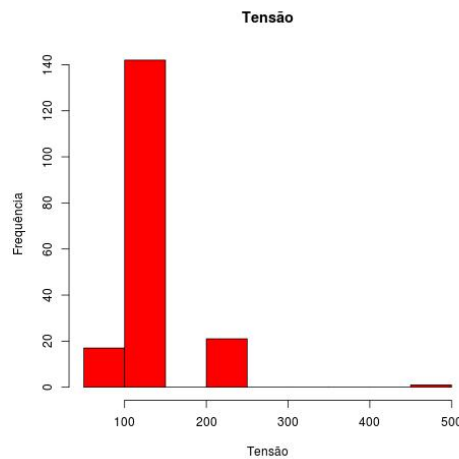


Figura 4.5: Gráfico ilustrando o número de estruturas com diferentes níveis de tensão

A figura 4.6 apresenta o *box-plot* da variável tensão. Através do *box-plot* pode-se observar como a variável tensão está distribuída em relação a homogeneidade dos dados, valores de tendência central, máximos e mínimos e valores atípicos se existirem. Observa-se que os dados da variável

tensão estão concentrados em torno da mediana.

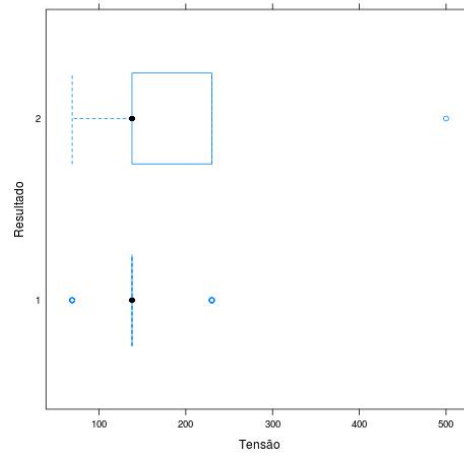


Figura 4.6: *Box-plot* da variável tensão

A figura 4.7 apresenta as estruturas metálicas alocadas em diferentes temperaturas. Observa-se que a maior parte das torres estão alocadas em regiões com temperaturas consideradas inferiores a  $21^{\circ}C$ .

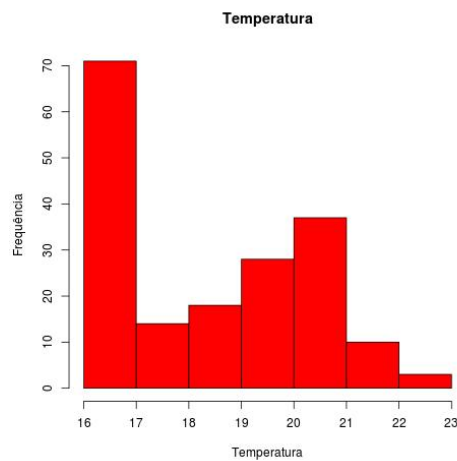


Figura 4.7: Gráfico ilustrando o número de estruturas alocadas em diferentes temperaturas

A figura 4.8 apresenta o *box-plot* da variável temperatura. Observa-se que os dados dessa variável são mais heterogêneos.

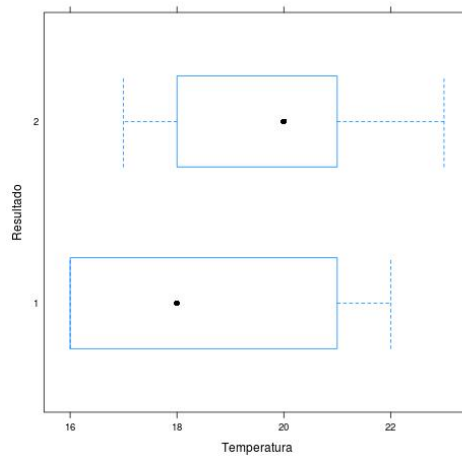


Figura 4.8: *Box-plot* da variável temperatura

O relevo é uma variável que pode impactar positivamente para a queda de uma estrutura metálica. A figura 4.9 apresenta o números de estruturas em diferentes tipos de terrenos. Observa-se que

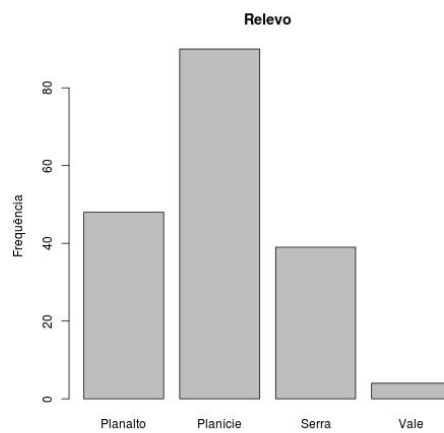


Figura 4.9: Gráfico ilustrando o número de estruturas em diferentes tipos de relevo

A figura 4.10 apresenta o número de estruturas metálicas de linhas de transmissão alocadas com dois diferentes tipos de estrutura. Observa-se que a maior parte das observações apresentam estruturas com a função suspensão.

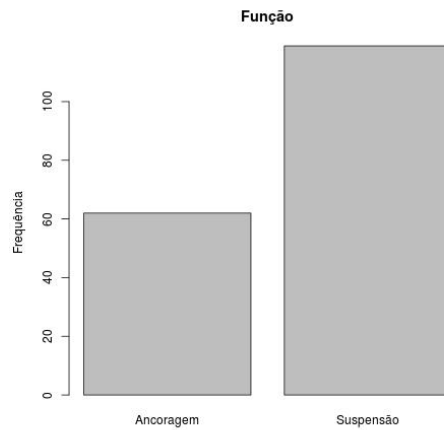


Figura 4.10: Gráfico ilustrando o número de estruturas com dois diferentes tipos de estrutura

A variável fundação assume três níveis. Através da figura 4.11, observa-se que a maior parte das estruturas metálicas de linhas de transmissão foram alocadas com fundação do tipo grelha, ou seja, uma fundação mais rasa.

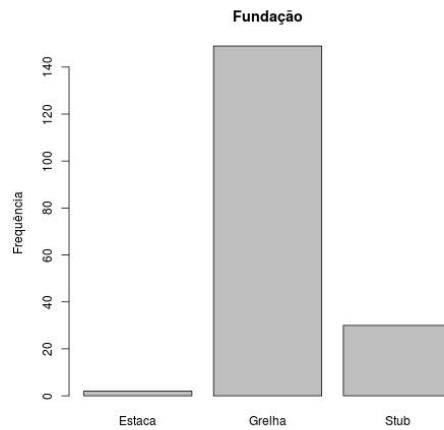


Figura 4.11: Gráfico ilustrando o número de estruturas com três diferentes tipos de fundação

A figura 4.12 apresenta o número de estruturas metálicas de linhas de transmissão alocadas em regiões com diferentes velocidades de vento. No conjunto de dados em estudo, poucas estruturas foram alocadas em regiões com ventos de até  $26\text{km/h}$ .



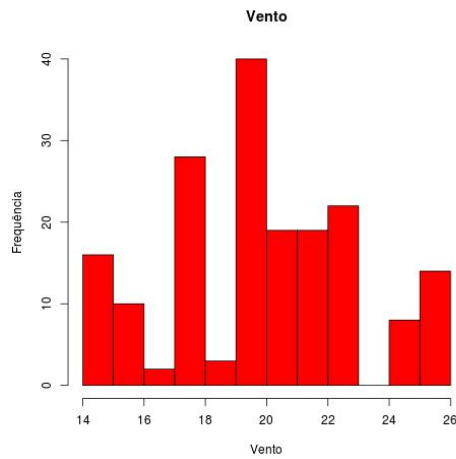


Figura 4.12: Gráfico ilustrando o número de estruturas com diferentes velocidades do vento

Através da figura 4.13, observa-se que a variável vento apresenta os dados mais heterogêneos.

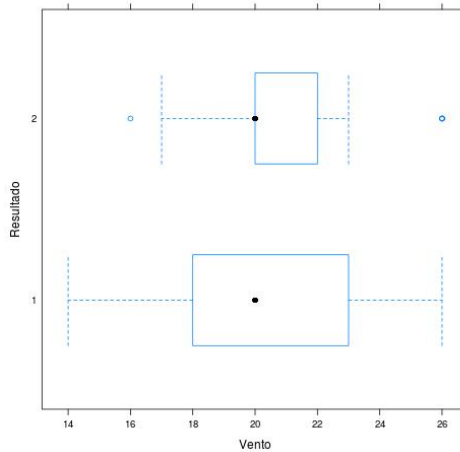


Figura 4.13: *Box plot* da variável vento

A figura 4.14 apresenta o número de estruturas que sofreram queda e as que não sofreram. Do total de estruturas analisadas, 44 caíram e 137 ainda estão em operação.

Neste trabalho, a variável resposta resultado é uma função de dez variáveis independentes ou explicativas. O total de estruturas que caíram representa 24,3% do total de estruturas metálicas analisadas.

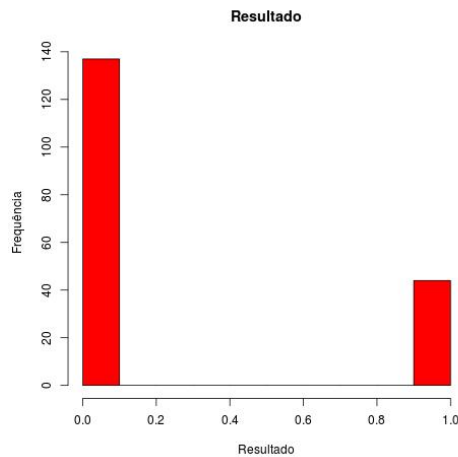


Figura 4.14: Gráfico de barras ilustrando o número de estruturas que sofreram queda

## 4.3 Resultados com Regressão Logística

Nesta seção, apresenta-se o modelo de regressão logística binária para um conjunto de dados históricos de quedas de estruturas de linhas de transmissão de um dado sistema elétrico mantido pela COPEL (2015) e discute-se as implicações deste modelo.

### 4.3.1 Modelo de Probabilidade de Ocorrência de Queda de uma Estrutura

Para o conjunto de dados do apêndice A, pretende-se ajustar um modelo de regressão logística binária. O primeiro passo consiste na definição do modelo de regressão. Para isso, considera-se o vetor aleatório  $\mathbf{Y}$  cujas componentes  $Y_i$ ,  $i \in \{1, 2, \dots, n\}$  são variáveis aleatórias independentes e distribuídas com distribuição binomial com parâmetros 1 e  $p_i$ , ou seja,  $Y_i \sim \text{Bin}(1, p_i)$ . Com a notação utilizada, pode-se escrever que  $E(Y_i) = \pi_i$ .

Por construção do modelo,  $\pi_i$  é uma função das variáveis explicativas apresentadas na matriz  $\mathbf{X}$ , ou seja, no estudo em questão, a matriz  $\mathbf{X}$  contém todas as informações das variáveis inde-

pendentes. O elemento  $x_{ij}$ ,  $i \in \{1, 2, \dots, n\}$  e  $j \in \{1, 2, \dots, m\}$  ( $m$  denota o número de variáveis explicativas) informa o valor da  $j$ -ésima variável explicativa para a  $i$ -ésima observação  $y_i$ .

Para o conjunto de dados em questão, a probabilidade de ocorrência de queda  $\pi_i$  é dada por

$$\pi_i = \frac{e^{\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}}} = \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}}}{1 + e^{\mathbf{x}_i^T \boldsymbol{\beta}}} \quad (4.1)$$

em que  $x_i$  denota a  $i$ -ésima linha da matriz de variáveis explicativas  $\mathbf{X}$  e  $\boldsymbol{\beta}$  é o vetor paramétrico a ser estimado.

Admite-se como pressupostos para o modelo de regressão logística:

- Aditividade. A probabilidade de ocorrência de queda é uma função do preditor linear  $\mathbf{x}_i \boldsymbol{\beta}$ ;
- A variável aleatória  $Y$  deve seguir uma distribuição da família exponencial;
- Não existência de multicolinearidade, ou seja, a não existência de redundância de informações na matriz de variáveis explicativas.

As estimativas pontuais dos parâmetros do modelo são feitas pelo método da máxima verossimilhança. Assumindo que a função de verossimilhança do modelo é dada por

$$L(\boldsymbol{\beta} | Y) = \prod_{i=1}^n f(y_i | \beta x_i \boldsymbol{\beta}) \quad (4.2)$$

segue que

$$\hat{\boldsymbol{\beta}} = \arg \max L(\boldsymbol{\beta} | Y, X) \quad (4.3)$$

O conjunto de dados deste estudo apresenta 181 estruturas metálicas que estavam em operação das quais 137 não sofreram queda e 44 caíram. Com isso, tem-se duas categorias para a resposta

(queda ou nenhum). O valor de resposta que foi designado como o evento de referência é nenhum. Oito fatores com diferentes níveis são considerados no modelo. A tabela 4.1 apresenta esses fatores e seus respectivos níveis.

Tabela 4.1: Informações dos fatores

Fator	Níveis	Valores
Tensão	2	138; outros
Relevo	3	planície; planalto; serra-vale
Massa de ar fria	3	paralela; perpendicular; transversal
Massa de ar quente	3	paralela; perpendicular; transversal
Terreno	4	A; B; C; D
Capacidade	2	alta; baixa
Função	2	ancoragem; suspensão
Fundação	2	grelha; Stub

A estimativa de  $\hat{\beta}$ , obtida no *software* Minitab é apresentada na tabela 4.2. A função de ligação do modelo é a função logit. Uma vantagem da função de ligação do modelo ser a função logit é que ela fornece uma estimativa da razão de chances para cada variável preditora no modelo. Percebe-se que os coeficientes estimados para várias entradas têm p valores inferiores a 0,05, indicando que não há provas suficientes de que os coeficientes não são nulos. Isto significa que vários fatores têm influência sobre o modelo de regressão. Quanto maior o preditor linear, maior a probabilidade de ocorrência do evento queda. A ordem de importância de variáveis significativas para a ocorrência do evento queda é dado pelo valor absoluto dos coeficientes estimados. A interpretação dos coeficientes estimados é relativa ao nível de referência do preditor; coeficientes positivos indicam que um dos preditores impacta mais provavelmente a resposta binária do que

a referência. E vice-versa para os coeficientes negativos. Coeficientes próximos de zero indicam que a associação entre o preditor e a resposta binária pode não ser importante.

Tabela 4.2: Tabela de regressão logística

Preditor	Coefficiente	SE Coef	Z	P	Razão de chances	Intervalo com 95% de confiança
Constante	59,3711	14,6063	4,06	0,000	-	-
Tensão						
Outros	-8,05377	2,29698	-3,51	0,000	0,00	0,00 - 0,03
Vento	-0,282987	0,202494	-1,40	0,162	0,75	0,51 - 1,12
Relevo						
Planalto	-5,14588	1,80962	-2,84	0,004	0,01	0,00 - 0,20
Serra-vale	-7,70186	2,34577	-3,28	0,001	0,00	0,00 - 0,04
Massa de ar fria						
Perpendicular	-0,403916	1,05317	-0,38	0,701	0,67	0,08 - 5,26
Transversal	-5,48150	1,82560	-3,00	0,003	0,00	0,00 - 0,15
Massa de ar quente						
Perpendicular	-3,08682	1,39565	-2,21	0,027	0,05	0,00 - 0,70
Transversal	-3,01681	1,49606	-2,02	0,044	0,05	0,00 - 0,92
Temperatura	-2,21225	0,547453	-4,04	0,000	0,11	0,04 - 0,32
Terreno						
B	1,30160	0,906144	1,44	0,151	3,68	0,62 - 21,71
C	6,32315	2,10111	3,01	0,003	557,32	9,07 - 34243,69
D	4,80950	1,80180	2,67	0,008	122,67	3,59 - 4192,17
Capacidade						
Baixa	-4,16272	1,15683	-3,60	0,000	0,02	0,00 - 0,15
Função						
Suspensão	-2,54554	1,32737	-1,92	0,055	0,08	0,01 - 1,06
Fundação						
Stub	4,39213	2,19840	2,00	0,046	80,81	1,09 - 6008,38

Através dos coeficientes do modelo de regressão da tabela 4.2 é possível avaliar a probabilidade de ocorrência de queda de uma estrutura metálica com determinados atributos. Por exemplo, uma torre colocada numa região com presença de massa de ar fria perpendicular, temperatura de 22°C, fundação do tipo *stub* e capacidade baixa tem probabilidade de queda de

$$P = \frac{e^{59,3711 - 2,21225(22) + 4,39213 - 4,16272}}{1 + e^{59,3711 - 2,21225(22) + 4,39213 - 4,16272}} = \frac{55882,69}{55883,69} = 0,999982$$

ou seja, uma torre com os atributos anteriores certamente cairá.

Em regressão logística, uma estatística importante é a estatística  $G$ , que testa a hipótese nula de que todos os coeficientes associados com preditores nulos contra os coeficientes não todos nulos. Caso o p valor seja nulo, rejeita-se a hipótese nula.

O teste de que todas as inclinações são nulas indica que  $G = 149,335$ ,  $DF = 15$ , p valor = 0,000 e Log-verossimilhança = -25,719.

A qualidade do ajuste de um modelo estatístico descreve o quão bem ele se encaixa em um conjunto de observações. As medidas de qualidade de ajuste normalmente apresentam a discrepância entre os valores observados e os valores esperados sob o modelo em questão.

Os testes de Pearson e Deviance são tipos de resíduos para modelos logísticos. Eles são medidas úteis para avaliar quão bem o modelo selecionado ajusta os dados. Quanto maior o p valor, melhor é a qualidade do ajuste pelo modelo. Ambos os testes, de Pearson e Deviance possuem p valores maiores do que 0,10, indicando que não há evidência suficiente para o modelo não ajustar adequadamente os dados, quando o nível  $\alpha$  é menor do que ou igual a 0,10. O teste de Hosmer-Lemeshow avalia o ajuste do modelo ao comparar as frequências observadas e esperadas. O teste agrupa os dados pelas probabilidades estimadas da menor para a maior, e, então executa um teste qui-quadrado para determinar se as frequências observadas e esperadas são significativamente diferentes. A tabela 4.3 apresenta os resultados dos três testes.

Tabela 4.3: Teste de qualidade do ajuste

Método	qui-quadrado	DF	P
Pearson	122,544	161	0,989
Deviance	51,438	161	1,000
Hosmer-Lemeshow	1,915	8	0,984

A tabela 4.4 permite ver o quão bem o modelo se ajusta aos dados, comparando as frequências

observadas e esperadas. Não há evidência suficiente de que o modelo não se encaixa bem aos dados, como as frequências observadas e esperadas estão bem próximas. Esses resultados apoiam as conclusões feitas pelo testes de adequação do modelo.

Tabela 4.4: Tabela das frequências observadas e esperadas

Valor	1	2	3	4	5	6	7	8	9	10	Total
Nenhum											
Obs	0	2	10	17	17	18	18	18	18	19	137
Esp	0,0	2,1	10,3	16,0	17,6	17,9	18,0	18,0	18,0	19,0	
Queda											
Obs	18	16	8	1	1	0	0	0	0	0	44
Esp	18,0	15,9	7,7	2,0	0,4	0,1	0,0	0,0	0,0	0,0	
Total	18	18	18	18	18	18	18	18	18	19	181

A tabela 4.5 apresenta o número e a porcentagem de pares de observações com diferentes valores-resposta, que são pares concordantes (resposta sucesso com alta probabilidade), pares discordantes (resposta sucesso com baixa probabilidade) e pares ligados (observações iguais às probabilidades previstas).

SomersD apresenta a quantidade de pares mais concordantes do que discordantes, divididos pelo número total de pares.

A estatística Goodman Kruskal Gama mostra quantos pares existem, que sejam mais concordantes do que discordantes, divididos pelo número total de pares, excluindo os pares ligados. Já a estatística Kendall's Tau-a mostra quantos pares existem, que sejam mais concordantes do que discordantes, divididos pelo número total de pares de observações, incluindo os pares com o

mesmo valor resposta.

Para criar os pares usados nessas estatísticas, cada sucesso observado é combinado com todos os fracassos. Anota-se então se a probabilidade de sucesso predita pelo modelo é maior para o sucesso real.

- Se a probabilidade prevista é maior para a observação tipo sucesso, o par é considerado concordante.
- Se a probabilidade prevista é menor para a observação tipo sucesso, o par é considerado discordante.
- Se a probabilidade prevista é igual para a observação tipo sucesso e tipo fracasso, o par é considerado ligado.

Maiores valores para SomersD, Goodman Kruskal Gama e Kendall's Tau-a indicam que o modelo tem melhor habilidade preditiva.

Tabela 4.5: Medidas de associação

Pares	Número	Percentual	Resumo das medidas
Concordantes	5924	98,3	SomersD 0,97
Discordantes	102	1,7	Goodman Kruskal Gamma 0,97
Ligados	2	0,0	Kendall's Tau-a 0,36
Total	6028	100,0	

Houve 98,3% de pares concordantes contra 1,7% de pares discordantes. SomersD e Goodman Kruskal Gamma são iguais, pois há poucos pares ligados.

As figuras 4.15 e 4.16 mostram gráficos usados para identificar fatores que não se ajustaram bem ao modelo. A figura 4.15 mede a variação na estatística da qualidade de Pearson, devido



a remoção de um fator particular. Os fatores que têm valores grandes de delta qui-quadrado (pontos afastados dos demais) correspondem a baixas probabilidades previstas.

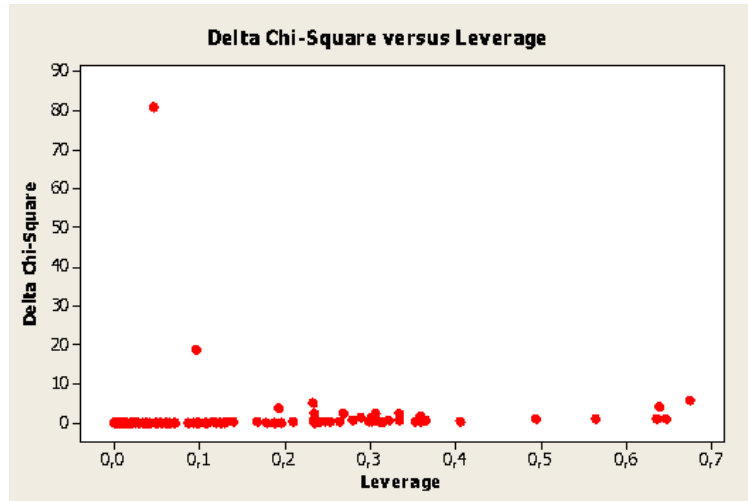


Figura 4.15: Gráfico da qualidade de ajuste do modelo I

Já a figura 4.16 mostra que os fatores que possuem maiores valores de delta qui-quadrado correspondem a menores alavancagens.

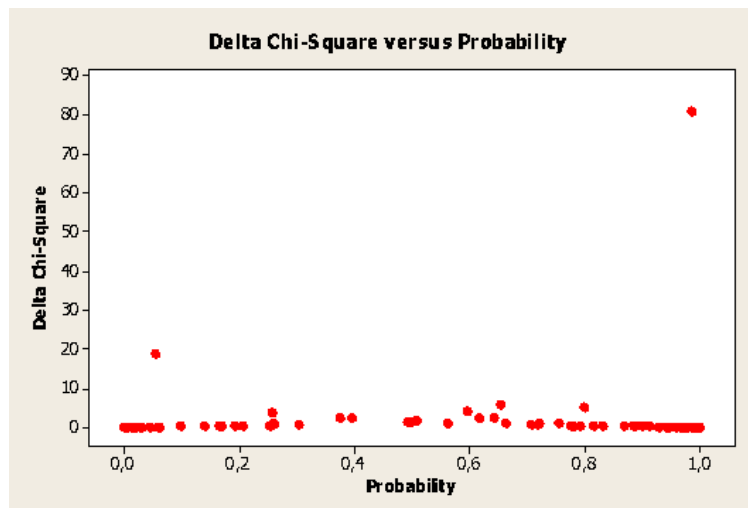


Figura 4.16: Gráfico da qualidade de ajuste do modelo II

Com as probabilidades de ocorrência de queda de estruturas dadas pelo modelo de regressão logística binária, pode-se traçar a curva ROC.

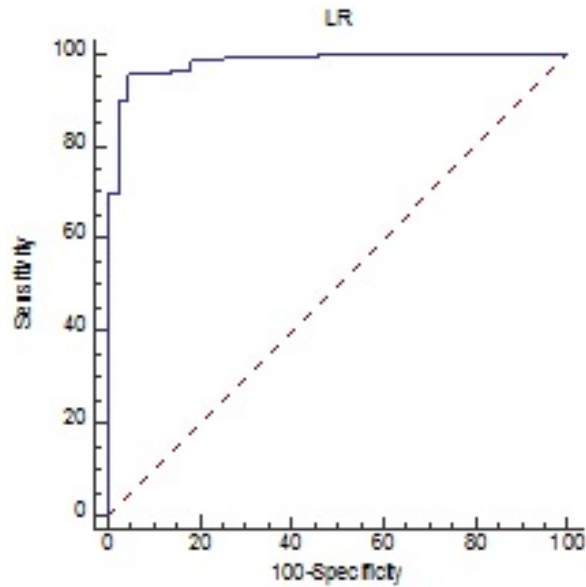


Figura 4.17: Curva ROC

O resultado calculado para a AUC foi de 0,9830, o que indica uma excelente *performance* do classificador.

## 4.4 Resultados com RNAs Combinadas com DOE

Nesta seção apresenta-se os resultados da avaliação de ocorrência de queda de estruturas metálicas utilizando RNAs combinadas com o DOE de Plackett- Burman. A RNA que será encontrada nesta seção deve ser capaz de decidir, para uma dada observação, se a torre caiu ou não, com alto desempenho de classificação.

### 4.4.1 Planejamento Pré Experimental

Neste momento, discute-se a definição do problema, a seleção da variável resposta e a escolha dos fatores e níveis. É importante salientar que para o treinamento de RNAs, o método da tentativa e erro é o mais utilizado e apresenta desvantagens como o consumo elevado de tempo.

A proposta nessa seção é a de apresentar uma metodologia bem estruturada para estimar os parâmetros da RNA.

Para avaliar a eficiência da RNA com relação ao desempenho de sua classificação, a métrica utilizada foi a área sob a curva ROC.

Diversos fatores podem ser considerados no treinamento das RNAs. A tabela 4.6 apresenta os fatores considerados em problemas de classificação utilizando o *software Statistica*. Os detalhes sobre os fatores são descritos na sequência. Para a escolha dos fatores e níveis, além das indicações do próprio *software*, foi utilizado como guia o trabalho de Balestrassi et al. (2009).

Tabela 4.6: Parâmetros de treinamento da RNA para o problema de queda de estruturas metálicas de linhas de transmissão

Fator	Símbolo	Níveis	Número de Níveis
Arquitetura da rede	—	<b>MLP</b> , RBF, GRNN, ARTMAP...	1
Número de camadas ocultas	NC	<b>0,1,2,3</b>	2
Função de erro	FE	Soma de quadrados, entropia cruzada	2
Número de neurônios	NN	$kN$ ou $k(N + 1)$ , N número de entradas	2
Função de ativação da entrada	CE	<b>identidade</b> , tanh, logística, exponencial	1
Função de ativação camada oculta	CO	<b>identidade</b> , tanh, <b>logística</b> , exponencial	2
Função de ativação da saída	CS	Identidade, <b>tanh</b> , <b>logística</b> , exponencial	2
Algoritmo de treinamento - Fase 1	F1	<b>backpropagation</b> , <b>quick propagation</b> , delta bar delta	2
Algoritmo de treinamento - Fase 2	F2	<b>gradiente conjugado</b> , <b>quasi-Newton (BFGS)</b> , Levenberg-Marquardt	2
Épocas- Fase 1	E1	<b>50,100</b>	2
Épocas- Fase 2	E2	<b>50,100</b>	2
Taxa de aprendizagem	TA	<b>0.01, 0.1</b>	2
Método de inicialização	MI	<b>Unif(0,1)</b> , <b>N(0,1)</b>	2
Condições de parada	CP	<b>0; 0.1</b>	2
Melhoria mínima do erro para treinamento	ET	<b>0; -0.1</b>	2
Melhoria mínima do erro para épocas	EE	<b>0;10</b>	2
Poda de neurônio	PN	<b>sem poda</b> , <b>poda de pesos inferiores a 0,05</b>	2
Poda de entrada	PE	<b>sem poda</b> , <b>poda de entradas com baixa influência após treinamento</b>	2
Limiar de classificação	-	limiares especificados manualmente, <b>limiares com menor perda</b>	1
Regularização do decaimento de pesos -Fase 1	D1	<b>sem regularização; 0,01</b>	2
Regularização do decaimento de pesos -Fase 2	D2	<b>sem regularização; 0,01</b>	2
Ajuste backpropagation	-	sem ajustar parâmetros; <b>com ajuste de parâmetros</b>	1
Ajuste Quick propagation	-	sem ajustar parâmetros; <b>com ajuste de parâmetros</b>	1
Método de amostragem	-	<b>aleatório</b> , bootstrapping, validação cruzada	1

- *Arquitetura da RNA.* As RNAs são algoritmos de modelagem não linear. Exemplos de RNAs para a regressão logística são as redes com arquiteturas MLP, RBF, GRNN, SVM, entre muitas outras. As redes MLPs são um dos tipos de rede mais populares e, em muitos domínios de problemas parecem oferecer o melhor desempenho possível.
- *Número de Camadas Ocultas.* Uma camada oculta é um grupo de neurônios que tem uma função específica e são processados como um todo. Normalmente, as redes com arquitetura MLP com uma única camada são capazes de aproximar qualquer função contínua.
- *Função de Erro.* Este fator especifica a função de erro a ser utilizada no treinamento da RNA. Entropia: Se a função de erro de entropia cruzada é escolhida, a rede realiza otimização de máxima verossimilhança, assumindo que os dados são retirados de distribuições exponenciais. Soma dos Quadrados: A abordagem alternativa é a utilização da soma de quadrados.
- *Número de Unidades por Camada.* Este fator especifica o número de neurônios na camada oculta da rede. Quanto maior o número de neurônios, mais complexa e flexível a rede se torna. Nesta tese, optou-se por usar uma quantidade que varia entre o número de entradas e o dobro delas ( $k(N + 1)$ , em que  $N$  é o número de entradas e  $k = 1; 2$ ).
- *Função de Ativação para as unidades da camada oculta.* A escolha da função de ativação, isto é, a função matemática específica, é crucial para construir um modelo de RNA uma vez que está diretamente ligada com o desempenho do modelo. A função de ativação deve ser escolhida tanto para as unidades da camada oculta quanto para as unidades da camada de saída. Via de regra, os neurônios de entrada não têm função de ativação, ou seja, eles usam uma função identidade em que os sinais de entrada não são transformados.

- *Função de Ativação para as unidades da camada de saída.* Com a função identidade o nível de ativação é passado diretamente para a saída. Utilizando a função logística, obtem-se a saída variando de 0 a 1.
- *Algoritmo de treinamento - Fase 1.* Este fator está relacionado com o treinamento do algoritmo para a MLP no primeiro estágio e assume dois níveis:
  - a) *Backpropagation.* Um algoritmo simples com um grande número de parâmetros de ajuste, convergência terminal muitas vezes lenta, porém uma boa convergência inicial.
  - b) *Quickpropagation.* Um algoritmo mais antigo com um desempenho comparável com o *backpropagation* na maioria das circunstâncias, embora pareça desempenhar visivelmente melhor em alguns problemas.
- *Algoritmo de treinamento- Fase 2.* Este fator está relacionado com o treinamento do algoritmo para MLP no segundo estágio e assume dois níveis.
  - a) *Gradiente descendente conjugado.* Trata-se de um bom algoritmo genérico, geralmente, com uma convergência rápida.
  - b) *Quasi Newton.* Um poderoso algoritmo de treinamento de segunda ordem, com convergência muito rápida, mas com altos requisitos de memória. Esse algoritmo trabalha com aproximações numéricas da matriz Hessiana.
- *Épocas.* Uma época é a apresentação de todo o conjunto de treinamento para a RNA em uma dada fase. Ao aumentar esse número, o valor da acurácia do modelo também aumentará (porém com maior tempo). Esse fator está relacionado com o tamanho da amostra. *Épocas - Fase 1:* a) 50 épocas de treinamento. b) 100 épocas de treinamento. *Épocas - Fase 2:* a) 50 épocas de treinamento. b) 100 épocas de treinamento.

- *Taxa de aprendizagem.* A taxa de aprendizagem é um número real entre 0 e 1 que representa a variável de ajuste para o treinamento do algoritmo *backpropagation*, *quickpropagation* e *Delta bar delta*. Taxas mais baixas de aprendizagem exigem mais iterações no treinamento. A taxa de aprendizagem superior permite a rede convergir mais rapidamente, no entanto, as possibilidades de uma solução não ótima são maiores. Os níveis escolhidos aqui foram de 0 e 0,1.
- *Inicialização do método.* Este fator define como os pesos deverão ser inicializados para o treinamento da RNA e assume dois níveis:
  - a) *Variável Uniforme.* Os pesos são inicializados com um valor aleatório uniformemente distribuído dentro de um intervalo cujos valores mínimos e máximos são dados. Neste caso, os valores máximo e mínimo são 0 e 1, respectivamente.
  - b) *Variável Gaussiana.* Os pesos são inicializados com um valor aleatório normalmente distribuído, dentro de um intervalo cuja média e desvio padrão são dados. Neste caso, escolheu-se  $\mathcal{N}(0, 1)$ .
- *Condições de parada (Erro alvo).* Se o erro no treinamento ou teste de seleção ficar abaixo dos valores alvo estimados, considera-se que a rede foi suficientemente bem treinada e o treinamento é finalizado.
- *Melhoria mínima no erro para treinamento.* Este fator representa a melhoria mínima (queda) no erro. O valor padrão zero implica que o treinamento será encerrado se o erro decair. Pode-se também especificar uma taxa de melhoria negativa, o que é equivalente a dar uma taxa máxima de decaimento do erro. A melhoria é medida através de um número de épocas, denominado de “janela”.

- *Melhoria mínima no erro para o número de épocas.* Este fator especifica o número de épocas sobre as quais a melhoria é medida. Alguns algoritmos, incluindo *backpropagation*, possuem ruído nos erros de treinamento e seleção e todos os algoritmos podem demonstrar ruído no erro de seleção. As janelas especificam um número de épocas nas quais as taxas de erro são monitoradas e melhoradas. O treinamento é interrompido se o erro não melhorar em muitas épocas. a) 0, b) 10.
- *Poda de neurônios.* Um neurônio com peso de conexão de saída de baixa magnitude fornece baixa contribuição para as ativações da próxima camada e pode ser podado, levando a uma rede compacta e rápida com *performance* equivalente. a) Sem poda. b) Poda de pesos inferiores a 0,05.
- *Poda de entrada com baixa sensibilidade após o treinamento.* Este fator se aplica à análise de sensibilidade baseada em podas após o treinamento. Uma análise de sensibilidade é efetuada após o treinamento da rede e variáveis de entrada com relações de sensibilidade de treinamento e seleção abaixo de um limiar são podadas. a) Sem poda. b) Poda de entradas com baixa influência após treinamento (relação igual a 1).
- *Limiares de classificação.* No caso de apenas um neurônio de saída, há dois níveis para serem escolhidos. Pode-se escolher manualmente os limiares ou através de uma forma automática em que um coeficiente de perda é escolhido. Um coeficiente de perda igual a 1 indica que as duas classes são igualmente importantes, situação escolhida para o estudo.
- *Regularização no decaimento de pesos - Fase 1.* Este fator refere-se ao uso da regularização do decaimento dos pesos na fase 1 em que foram considerados os algoritmos *backpropagation*, *quickpropagation* e *delta-bar-delta*. Utiliza-se esse fator para encorajar o desenvolvimento

de pesos menores, o que tende a reduzir o problema de sobreajuste ou *overfitting*, assim aumentando o poder de generalização da rede. a) Sem regularização. b) 0,01.

- *Regularização no decaimento de pesos - Fase 2.* Este fator refere-se ao uso da regularização do decaimento dos pesos na fase 2 em que foram considerados os algoritmos Gradiente Conjugado, Quasi-Newton e Levenberg-Marquardt. Serve para encorajar o desenvolvimento de pesos menores, o que tende a reduzir o problema de sobreajuste ou *overfitting*, aumentando assim o poder de generalização da rede. a) Sem regularização. b) 0,01.
- *Ajuste do backpropagation.* Este fator refere-se às opções adicionais para o treinamento do *backpropagation*. Inicialmente, é possível especificar a taxa de aprendizagem e o *momentum* iniciais e finais utilizados. Uma taxa mais alta pode convergir mais rapidamente, mas pode exibir maior instabilidade. O *momentum* é usado para compensar uma convergência lenta. Em segundo lugar, é possível embaralhar a ordem dos casos para cada época de treinamento. E, por último, é possível adicionar ruído gaussiano, para reduzir a tendência de sobreajuste.
- *Ajuste quickpropagation.* Este fator refere-se às opções adicionais para o treinamento do *quickpropagation*. Inicialmente, é possível especificar a taxa de aprendizagem e a aceleração utilizadas. A taxa de aprendizagem é especificada inicialmente, aplicada a primeira época. A aceleração especifica a máxima taxa de crescimento geométrico na mudança dos pesos que é permitida. E, por último, é possível adicionar ruído gaussiano, para reduzir a tendência de sobreajuste.
- *Método de amostragem.* Este fator especifica o método de amostragem (aleatório, validação cruzada ou *bootstrap*). À título de simplificação, adotou-se a amostragem aleatória em que cada variável é colocada em um subconjunto (treinamento, seleção e teste), optou-se ainda



por usar a proporção 8:1:1.

Resumindo, para a etapa do planejamento pré experimental adota-se:

- A definição do problema corresponde a avaliação da ocorrência de queda de estruturas metálicas de linhas de transmissão de um dado sistema elétrico em operação mantido pela COPEL (2015) na qual considera-se um conjunto de dados históricos de quedas contendo dez variáveis explicativas;
- A escolha da variável de resposta foi se a torre caiu (acontecimento de fracasso) denotado por 0 ou se a torre não caiu (acontecimento de sucesso) denotado pelo valor 1.
- A escolha dos fatores e níveis para o problema em questão são os fatores e níveis apresentados na tabela 4.6.

#### **4.4.2 Escolha do Delineamento Experimental**

Nesta etapa, seguindo as orientações de Coleman e Montgomery (1993) para a realização de um experimento, deve-se escolher o delineamento experimental a ser utilizado. Considerando os fatores e níveis apresentados na tabela 4.6, é interessante observar a complexidade de se experimentar todas as possibilidades de combinação, juntamente com a simulação das RNAs. Ao considerar apenas dois níveis para cada fator, seriam necessários  $2^{24}$  experimentos (sem repetições), o que inviabilizaria seu desenvolvimento.

Neste trabalho, o delineamento de Plackett-Burman foi utilizado devido a sua flexibilidade em tratar um elevado número de parâmetros com poucos experimentos. Contudo, devido a esse baixo número de experimentos, a interação entre fatores é negligenciada. O objetivo é justamente verificar o desempenho do DOE mais simples na parametrização da RNA e em seus resultados. Dos

24 fatores apresentados, 6 deles foram definidos *a priori* e excluídos do delineamento experimental (Arquitetura da Rede, Função de Ativação da Camada de Entrada, Limiar de Classificação, Ajuste do *backpropagation*, Ajuste do *quickpropagation* e o Método de Amostragem). Com relação aos parâmetros restantes, existem 18 fatores que serão tratados em dois níveis. O delineamento de Plackett-Burman com 20 experimentos foi escolhido para este trabalho e optou-se por utilizar uma repetição para o cálculo da média e da variação. A tabela 4.7, gerada no *software Minitab*, apresenta o delineamento experimental de Plackett-Burman para o problema em estudo.

Tabela 4.7: Delineamento Plackett-Burman para 18 fatores

NC	FE	NN	CO	CS	F1	F2	E1	E2	TA	MI	CP	ET	EE	PN	PE	D1	D2
2	EC	22	Tanh	Id	quick	QN	50	100	0,10	Unif	0,0	-0,1	0	S	N	S	N
1	EC	11	log	Id	quick	QN	100	100	0,01	Unif	0,1	0,0	0	S	S	N	N
2	SQ	22	Tanh	log	quick	QN	100	50	0,01	Norm	0,1	-0,1	10	S	N	N	N
1	SQ	11	Tanh	Id	back	Grad	50	50	0,01	Unif	0,0	-0,1	0	N	N	N	N
1	SQ	22	Tanh	log	back	QN	100	100	0,10	Unif	0,0	0,0	10	N	S	S	N
2	SQ	22	log	Id	back	Grad	50	100	0,01	Norm	0,0	0,0	10	S	S	N	N
1	EC	22	Tanh	log	quick	Grad	50	50	0,01	Norm	0,0	0,0	0	S	S	S	S
2	SQ	22	log	log	quick	Grad	50	100	0,10	Unif	0,1	0,0	0	N	N	N	S
2	EC	11	Tanh	log	quick	Grad	100	100	0,01	Unif	0,0	-0,1	10	N	S	N	S
2	SQ	11	log	log	back	QN	100	50	0,01	Unif	0,0	0,0	0	S	N	S	S
2	EC	11	log	log	back	Grad	50	50	0,10	Unif	0,1	-0,1	10	S	S	S	N
1	SQ	22	log	Id	quick	QN	50	50	0,01	Unif	0,1	-0,1	10	N	S	S	S
1	EC	22	Tanh	Id	back	Grad	100	50	0,10	Unif	0,1	0,0	10	S	N	N	S
1	EC	22	log	log	back	Grad	100	100	0,01	Norm	0,1	-0,1	0	N	N	S	N
2	EC	22	log	Id	back	QN	100	50	0,10	Norm	0,0	-0,1	0	N	S	N	S
1	SQ	11	Tanh	log	back	QN	50	100	0,10	Norm	0,1	-0,1	0	S	S	N	S
1	EC	11	log	log	quick	QN	50	50	0,10	Norm	0,0	0,0	10	N	N	N	N
2	EC	11	Tanh	Id	back	QN	50	100	0,01	Norm	0,1	0,0	10	N	N	S	S
2	SQ	11	Tanh	Id	quick	Grad	100	50	0,10	Norm	0,1	0,0	0	N	S	S	N
1	SQ	11	log	Id	quick	Grad	100	100	0,10	Norm	0,0	-0,1	10	S	N	S	S

### 4.4.3 Realização do Experimento

Cada linha da tabela 4.7 representa um experimento a ser realizado. Em outras palavras, cada linha da tabela 4.7 representa os parâmetros de uma RNA a ser configurada. Para cada

RNA obtida, calculou-se a AUC correspondente que foi denotada por AUC1. Em seguida, com os mesmos parâmetros, mudando-se apenas o método de inicialização da rede, obteve-se outra AUC, denotada por AUC2. Os valores obtidos das AUC1 e AUC2 estão organizados na tabela 4.8.

Tabela 4.8: Valores obtidos das AUCs

Experimento	AUC1	AUC2	Experimento	AUC1	AUC2
1	0,954214	0,922196	11	0,911828	0,950896
2	0,963006	0,973955	12	0,965163	0,964837
3	0,957200	0,86226	13	0,961513	0,982250
4	0,928500	0,950813	14	0,552513	0,546201
5	0,927505	0,803003	15	0,528202	0,468149
6	0,953052	0,944426	16	0,862807	0,939200
7	0,912989	0,0,960849	17	0,920952	0,954214
8	0,964499	0,978766	18	0,402704	0,662575
9	0,952223	0,965826	19	0,906105	0,920537
10	0,940445	0,950896	20	0,876493	0,657017

A área sob a curva ROC que é denotada por (AUC) é a métrica utilizada para avaliar o desempenho de classificação da rede. Dessa forma, para cada RNA, calculou-se as seguintes quantidades

$$\text{Média} = \frac{AUC1 + AUC2}{2} \quad \text{e} \quad \text{Diferença} = |AUC1 - AUC2| \quad (4.4)$$

ou seja, a média das AUCs e o valor absoluto da diferença entre AUC1 e AUC2, respectivamente. Os resultados desses cálculos são apresentados na tabela 4.9. Esses dois valores foram

calculados porque além de uma posterior maximização nos valores da média das AUCs, espera-se uma maior robustez dos mesmos, que pode ser traduzida como uma minimização da amplitude da diferença encontrada.

Tabela 4.9: Média e amplitude da diferença das AUCs

Experimento	Média	Diferença	Experimento	Média	Diferença
1	0,938205	0,032018	11	0,931362	0,039068
2	0,968481	0,010949	12	0,965000	0,000326
3	0,909713	0,094974	13	0,971882	0,020737
4	0,939657	0,022313	14	0,549357	0,006312
5	0,865254	0,124502	15	0,498176	0,060053
6	0,948739	0,008626	16	0,901004	0,076393
7	0,936919	0,047860	17	0,937583	0,033262
8	0,971633	0,014267	18	0,532640	0,259871
9	0,959025	0,013603	19	0,913321	0,014432
10	0,945671	0,010451	20	0,766755	0,219476

Com o objetivo de reconhecer os parâmetros da RNA que mais influenciam na variável média e na variável diferença, constroeu-se o diagrama de Pareto para essas duas variáveis. O diagrama de Pareto é um método gráfico de apresentação de dados por ordem de tamanho, importância ou prioridade. O diagrama facilita a tomada de decisões, pela facilidade com que se reconhece, quais são as causas vitais do problema. A figura 4.18 apresenta o diagrama de Pareto para a variável resposta média. Algumas conclusões importantes podem ser extraídas da figura 4.18. Observa-se que os fatores PE, PN e MI são os que mais influenciam na variável média. Já os fatores F1, CP e NN possuem pouca influência na variável resposta média.

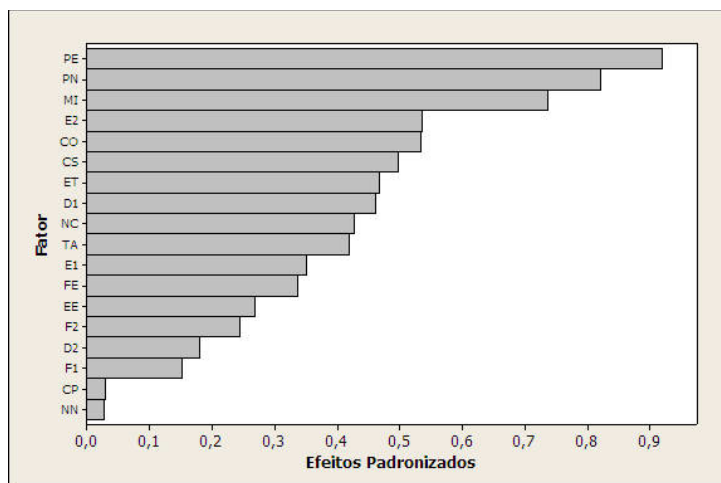


Figura 4.18: Diagrama de Pareto para média

A figura 4.19 apresenta o diagrama de Pareto para a variável resposta diferença. Algumas conclusões importantes podem ser extraídas da figura 4.19. Observa-se que os fatores CS, D1, D2 e TA são os que mais influenciam na variável diferença.

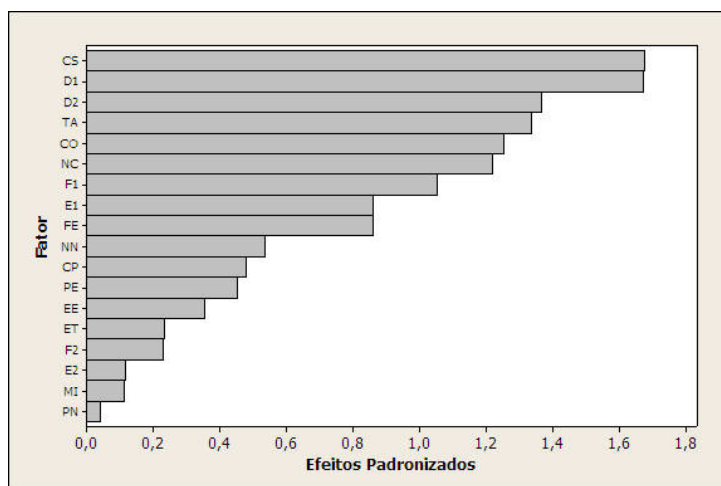


Figura 4.19: Diagrama de Pareto para Diferença

De posse dos resultados da tabela 4.9, pode-se construir os gráficos dos efeitos principais da média e da diferença. Os efeitos principais são definidos como a mudança ocorrida na variável resposta quando se move do nível baixo para o nível alto.

As figuras 4.20 e 4.21 apresentam os gráficos dos efeitos principais para a média e para a diferença (em valor absoluto) das áreas sob a curva ROC. Para o gráfico dos efeitos principais da média, quanto maior a inclinação dos segmentos de reta, maior será a influência desse fator na alteração do valor da área sob a curva ROC. Assim, o fator que mais influencia a média das áreas é o fator PE. Fatores que não influenciam possuem o segmento de reta horizontal. Observa-se que os segmentos de reta associados aos fatores NN e CP são praticamente horizontais, o que demonstra que esses fatores possuem pouca influência para a variável resposta média. Gráficos de efeitos principais permitem extrair conclusões importantes a respeito da variável resposta em questão. Por exemplo, para o gráfico dos efeitos principais da média, ao observar o fator NC, conclui-se que ao passar do nível 1 para o nível 2, ocorre uma redução para a variável resposta média. Por outro lado, ao observar o fator CO, ao passar do nível *tgh* para *log* ocorre um aumento na variável resposta média.

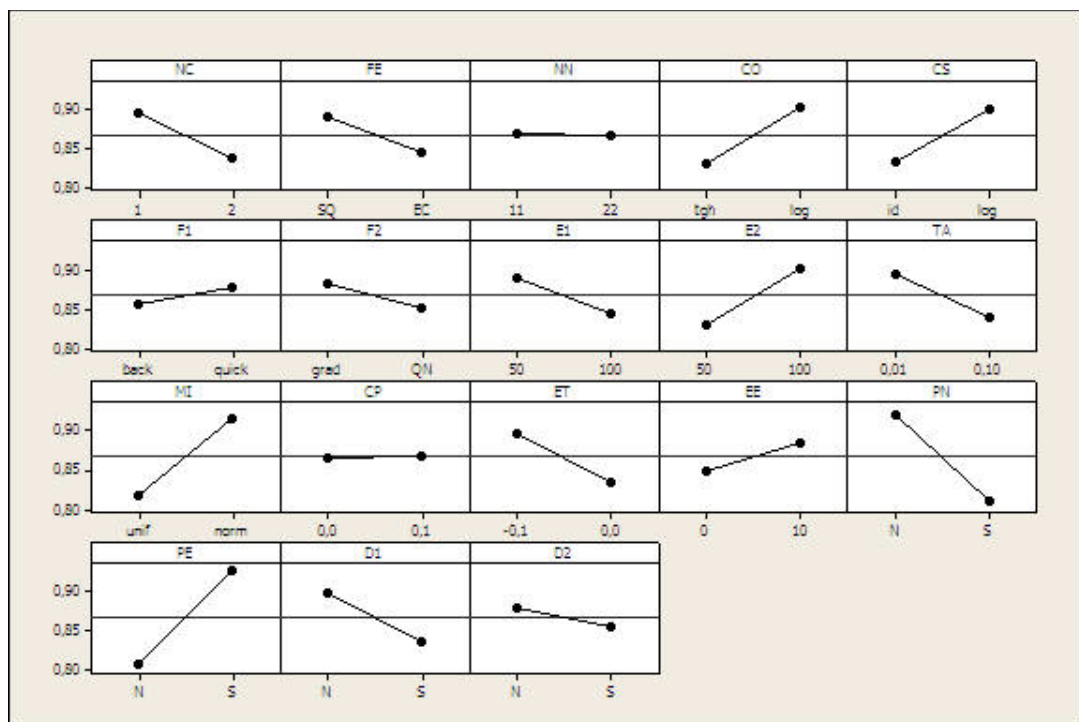


Figura 4.20: Plotagem dos efeitos principais da média

Para o gráfico dos efeitos principais da diferença, quanto maior a inclinação do segmento de reta, maior a influência na amplitude da diferença entre os valores. Os fatores que mais influenciam para a variável resposta diferença são os fatores CS e D1. Fatores que não influenciam possuem o segmento de reta horizontal. Observa-se que os segmentos de reta associados aos fatores E2, MI e PN são praticamente horizontais, o que demonstra que esses fatores possuem pouca influência para a variável resposta diferença.

Além dos gráficos dos efeitos principais permitirem extrair conclusões importantes a respeito da variável em estudo eles constataam as conclusões extraídas dos diagramas de Pareto tanto para a variável média, quanto para a variável diferença. A figura 4.19 apresentou os fatores CS, D1,D2 e TA como os fatores que possuem maior influência para a variável diferença. Essa afirmação é comprovada analisando o gráfico dos efeitos principais da diferença, pois os segmentos de reta associados aos fatores CS, D1,D2 e TA possuem uma maior inclinação.

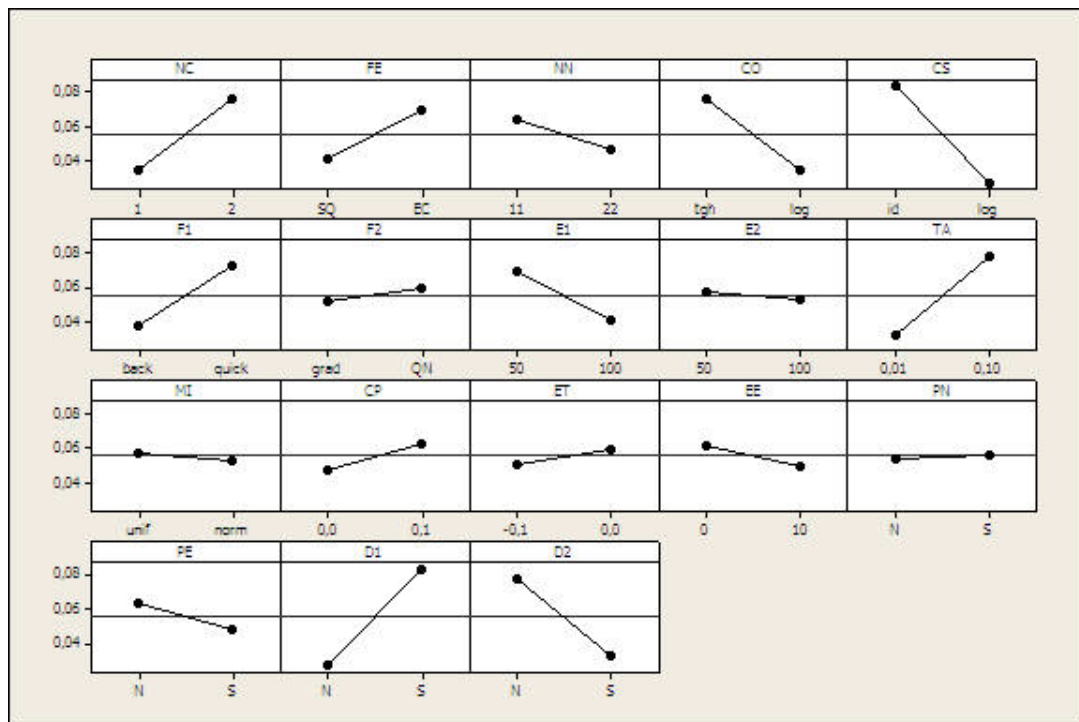


Figura 4.21: Plotagem dos efeitos principais da diferença

#### 4.4.4 Otimização Simultânea das Variáveis de Resposta Média e Diferença

O principal objetivo desta seção é encontrar uma RNA robusta para a avaliação da probabilidade de ocorrência de queda de estruturas metálicas de linhas de transmissão. Com isso, pretende-se maximizar os valores das médias das AUCs e minimizar o valor absoluto da diferença das AUCs. Em outras palavras, deve-se buscar a otimização simultânea das variáveis resposta média e diferença.

Antes de buscar a otimização simultânea pretendida, apresenta-se algumas estatísticas básicas para as variáveis média e diferença.

As figuras 4.22 e 4.23 apresentam os histogramas das variáveis média e diferença, construídos com auxílio do *software Minitab*. A base de cada retângulo representa uma classe (média ou diferença) e a altura de cada retângulo representa a quantidade ou frequência com que o valor dessa classe ocorreu no conjunto de dados. Percebe-se que a maior parte das informações está concentrada em torno da média dessa variável.

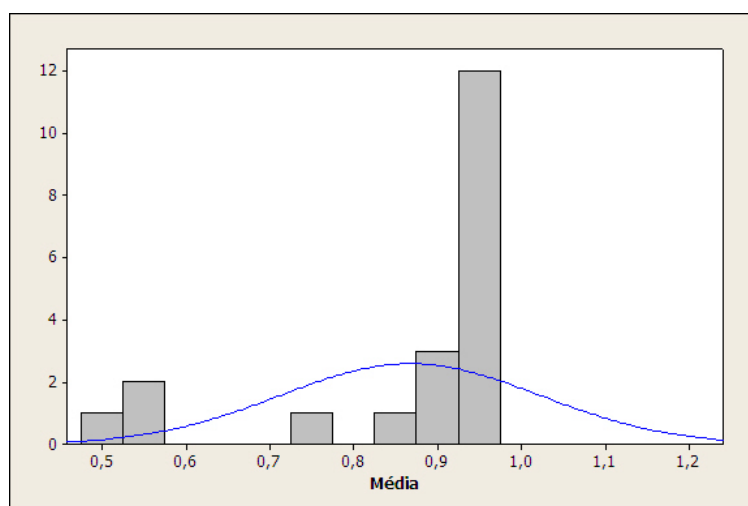


Figura 4.22: Histograma da variável resposta média



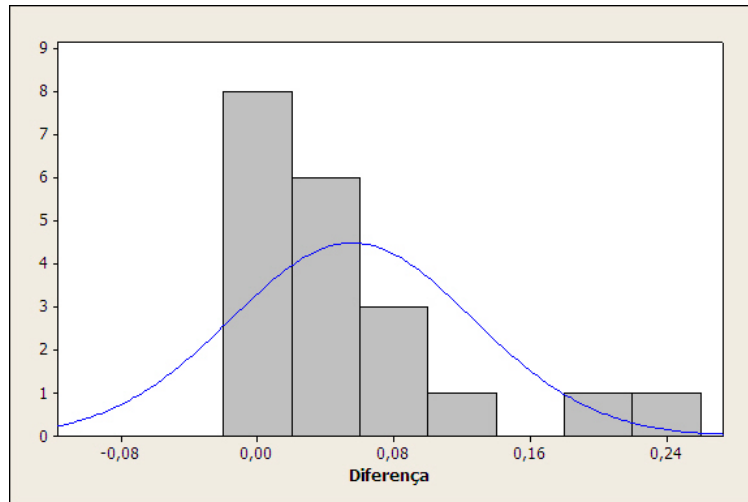


Figura 4.23: Histograma da variável resposta diferença

O *Minitab* também fornece algumas estatísticas básicas de cada uma das variáveis. Essas estatísticas estão resumidas na tabela 4.10.

Tabela 4.10: Algumas estatísticas das variáveis média e diferença

Estatística	Variável Média	Variável Diferença
Média	0,86752	0,055475
Variância	0,02376	0,005049
Valor Mínimo	0,49818	0,000326
Valor Máximo	0,97188	0,259871
1 <sup>o</sup> Quartil	0,87419	0,011612
3 <sup>o</sup> Quartil	0,95645	0,072308

As figuras 4.24 e 4.25, construídas com o auxílio do *software Minitab*, apresentam o teste de normalidade para as variáveis média e diferença. O teste abordado é conhecido por Teste de Kolmogorov-Smirnov. Esse teste observa a máxima diferença absoluta entre a função de distribuição acumulada assumida para os dados, no caso a Normal, e a função de distribuição

empírica dos dados. Como critério, compara-se esta diferença com um valor crítico, para um dado nível de significância.

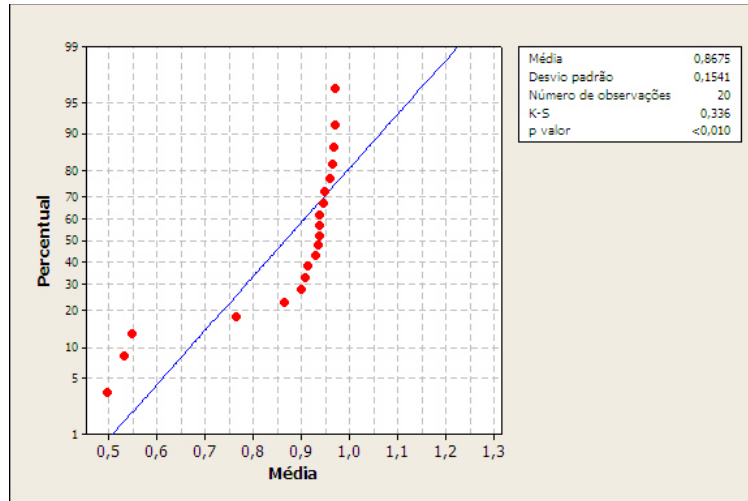


Figura 4.24: Probabilidade normal para a variável média

Como o p valor foi menor do que 0,010, rejeita-se a hipótese de normalidade dos dados da variável média.

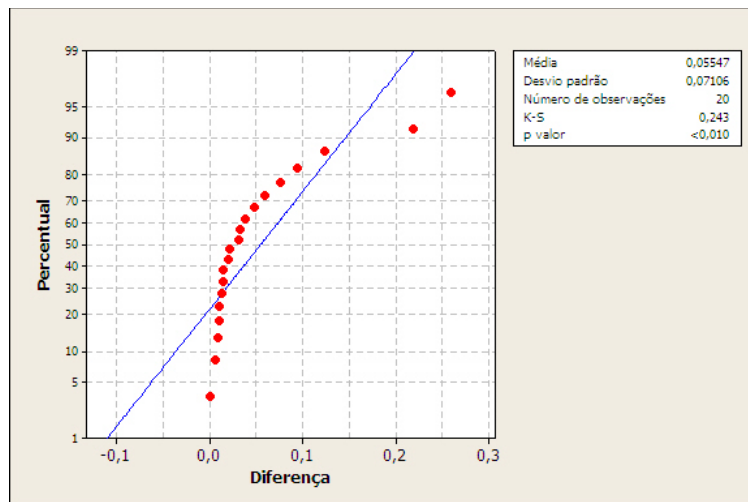


Figura 4.25: Probabilidade normal para a variável diferença

Analogamente ao caso anterior, como o p valor foi menor do que 0,010, rejeita-se a hipótese de normalidade dos dados da variável diferença.

Para maximizar a variável média e minimizar a variável diferença utiliza-se nesta tese o Método *Desirability* com auxílio do *software Minitab*. Para iniciar o algoritmo utiliza-se a configuração apresentada na tabela 4.11. Os valores superior e inferior da tabela 4.11 foram extraídos dos *quartis* da variável média.

Tabela 4.11: Configuração para os parâmetros de otimização

Resposta	Objetivo	Inferior	Alvo	Superior	Peso	Importância
Média	Maximizar	0,95645	1	-	10	1
Diferença	Minimizar	-	0	0,011612	10	0,5

A partir da configuração da tabela 4.11 obtém-se a simulação da otimização das variáveis média e diferença. Assim, a partir dessa etapa iniciam-se as simulações propriamente ditas. As figuras 4.26 e 4.27 apresentam os valores das *desirability* individuais e global para cada variável.

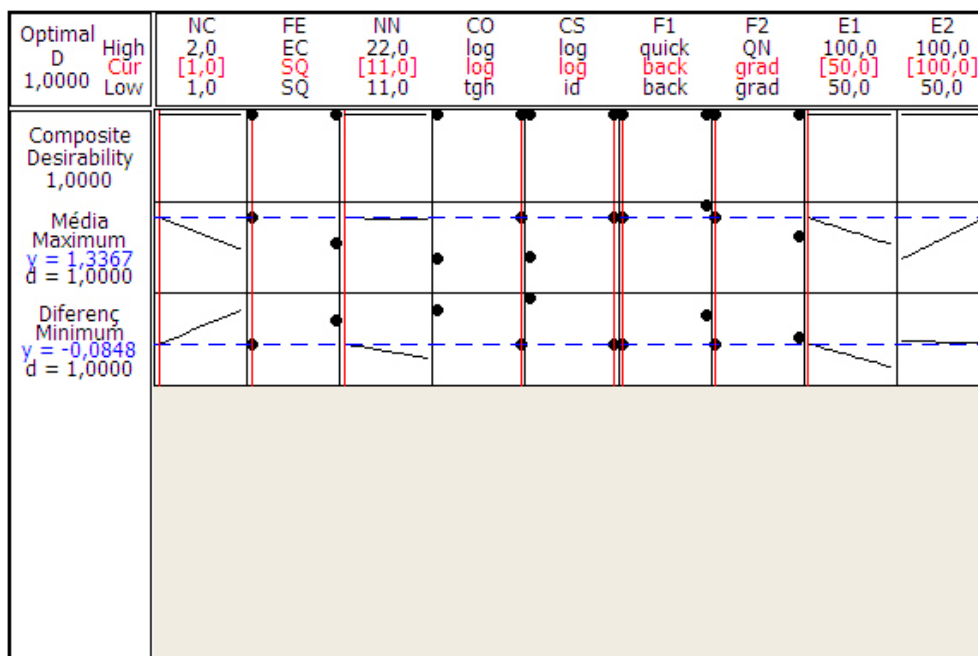


Figura 4.26: Simulação da otimização das variáveis resposta

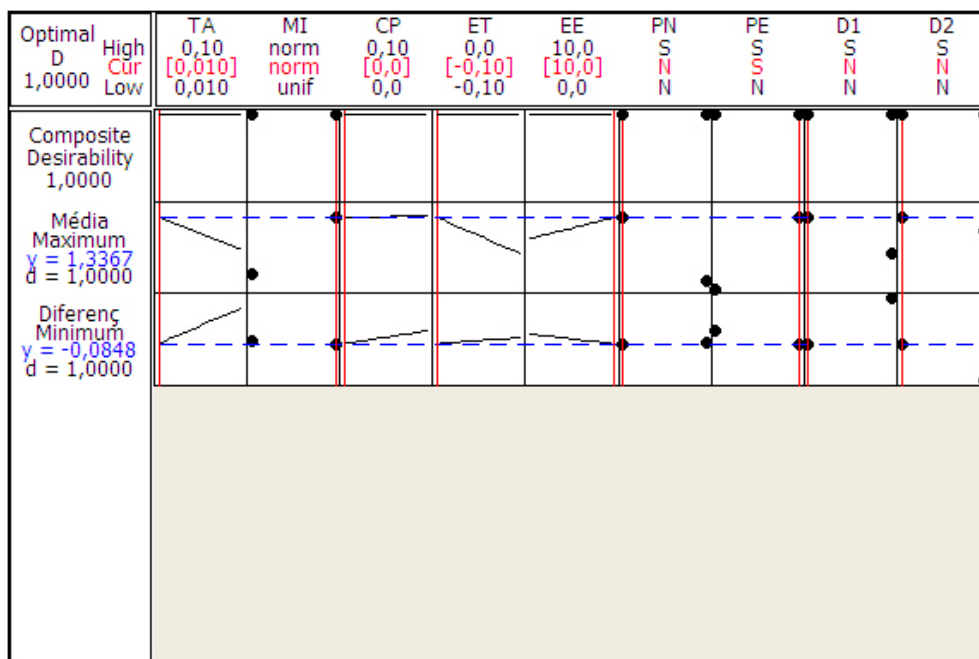


Figura 4.27: Simulação da otimização das variáveis resposta

A solução global obtida pelo algoritmo *desirability* representa a RNA mais robusta que é capaz de avaliar a probabilidade de ocorrência de queda de estruturas metálicas de linhas de transmissão. Para uma dada observação, a rede encontrada é capaz de decidir se a torre caiu ou não. A tabela 4.12 apresenta a rede, retratando todos os níveis dos parâmetros escolhidos.

Tabela 4.12: RNA escolhida para o problema de queda de estruturas

NC	FE	NN	CO	CS	F1	F2	E1	E2	TA	MI	CP	ET	EE	PN	PE	D1	D2
1	SQ	11	log	log	back	grad	50	100	0,01	norm	0	-0,1	10	N	S	N	N

Como a métrica utilizada para avaliar o desempenho de classificação da RNA é a AUC, a partir dos parâmetros da tabela 4.12 e com auxílio do *software Statistica* encontra-se uma AUC de 0,978517. Ao se repetir os mesmos parâmetros para configurar novas redes, os valores encontrados foram muito próximos indicando boa robustez. Ao configurar redes utilizando a técnica tradicional de tentativa e erro, obteve-se, na maioria dos casos, AUCs inferiores a 0,978517.

A figura 4.28 apresenta a curva ROC, construída com o auxílio do *software Statistica*, gerada a partir da RNA cujos parâmetros são apresentados na tabela 4.12.

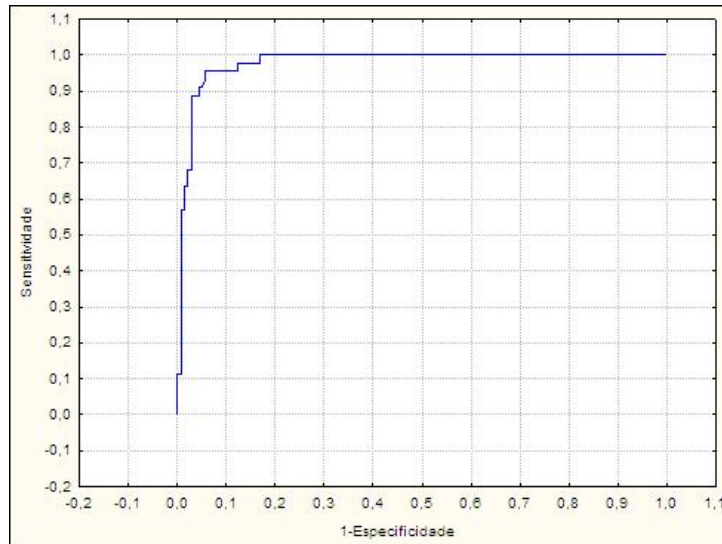


Figura 4.28: Curva ROC da RNA com parâmetros da tabela 4.12

#### 4.4.5 Conclusões e Recomendações

Segundo Coleman e Montgomery (1993), esta é a última etapa da realização do experimento. Para uma dada observação, a RNA da tabela 4.12 é capaz de decidir se a torre caiu ou não com alto desempenho de classificação. Ao afirmar que a AUC gerada pela RNA da tabela 4.12 é de 0,978517, equivale a probabilidade do classificador escolher uma observação positiva escolhida ao acaso com um valor superior a uma observação negativa escolhida ao acaso. Convém destacar que a técnica de RNAs combinadas com DOE apresentou uma RNA com desempenho de classificação superior as redes encontradas pela metodologia tradicional de treinamento, ou seja, tentativa e erro. Suspeita-se que ao combinar RNAs com outros DOEs o desempenho de classificação da rede certamente aumentaria, o que implicaria em um modelo de classificação superior ao modelo de regressão logística.

## Conclusões e Trabalhos Futuros

O presente trabalho visou avaliar a probabilidade de ocorrência de queda de estruturas metálicas de linhas de transmissão utilizando duas técnicas distintas. Os resultados alcançados mostram que, tanto a metodologia de regressão logística, quanto a metodologia de RNAs combinadas com DOE, apresentaram modelos robustos para a solução do problema.

A metodologia de RNAs combinadas com DOE proposta neste trabalho pode ser aplicada a problemas de diversas áreas do conhecimento. Diferentemente da metodologia clássica de treinamento de uma RNA, ou seja, tentativa e erro, as redes combinadas com DOE mostraram um desempenho superior em relação as redes obtidas através da metodologia clássica. Mesmo ao resolver o problema desta tese via regressão logística não houve interesse de comparação de técnicas consagradas na literatura.

A escolha do delineamento de Plackett-Burman para atuar conjuntamente com as RNAs, deve-se, basicamente, a economia de experimentos que seriam realizados. Convém destacar que, o uso de outro DOE que não descarte as iterações, certamente produziria um modelo com maior desempenho de classificação, o que pode ser comprovado no trabalho de Balestrassi et al. (2009).

De acordo com o modelo de regressão encontrado, as variáveis que mais impactam para ocorrência de queda de uma estrutura metálica são fundação do tipo *stub* e os terrenos dos tipos C e D. No trabalho de Wazen (2011) concluiu-se que, no modelo de regressão logística, as variáveis

fundação, terreno e relevo apresentam maior influência para ocorrência do evento queda. Porém, não foi utilizada nenhuma métrica de avaliação do desempenho de classificação.

Com relação aos resultados encontrados no modelo de regressão logística, uma solução para o problema da fundação seria o reforço da mesma. A aplicação de torres em regiões com tipos diferentes de terreno também pode ser avaliada.

A avaliação prévia das estruturas mais frágeis, o estudo destas e a execução de melhorias são ações que podem ser programadas reduzindo ao máximo os custos de desligamento e evitando altos valores de lucro cessante e indenizações a que uma empresa de energia possa estar sujeita, quando se deixa de atuar preventivamente para atuar corretivamente.

O modelo encontrado via RNAs combinadas com DOE representa um modelo com alto desempenho de classificação, visto que a área sob a curva ROC gerada a partir do modelo foi de 0,978517. Por outro lado, ao configurar RNAs utilizando a metodologia clássica de treinamento, ou seja, tentativa e erro, obteve-se, na maioria dos casos, AUCs inferiores a 0,978517, o que comprova a eficácia da metodologia abordada para a solução do problema de queda de torres.

Para a elaboração deste trabalho, fez-se uso de uma série de delimitações necessárias para sua finalização em tempo hábil. Dessa forma, algumas sugestões para trabalhos futuros são:

- Extrapolar a metodologia para novos conjuntos de dados e verificar se os desempenhos obtidos continuam similares;
- A aplicação da metodologia proposta para o estudo de falhas de sistemas elétricos;
- Explorar outros tipos de delineamentos para combiná-los com as RNAs;
- Confrontar resultados obtidos com RNAs combinadas com DOE com RNAs Bayesianas;
- Com relação as RNAs, verificar o desempenho de outras arquiteturas de rede.

# Referências Bibliográficas

- [1] ABDI, H., VALENTIN, D. e EDELMAN, B., Neural Networks, **Sage University Paper Series on Quantitative Applications in the Social Sciences**. Thousand Oaks: Sage Publications, 1999.
- [2] AITKENHEAD, M.J. et al. A novel method for training neural networks for time-series prediction in environmental systems, **Ecological Modeling**, 162, p. 87 – 95, 2003.
- [3] AL HOUSSEINI, A. et al., Prediction of risk for cesarean delivery in term nulliparas: a comparison of neural network and multiple logistic regression models, **American journal of obstetrics and gynecology**, v. 201, n. 1, p. 113.e1-6, 2009.
- [4] ALLISON, P.D., **Logistic regression using the SAS system, theory and application**, SAS intitute, 1999.
- [5] ANDRES, T.H. e HAJAS, W. C., Using iterated fractional factorial design to screen parameter in sensitivity analysis of a probabilistic risk assessment model, In: PROCEEDINGS OF THE JOINT INTERNATIONAL CONFERENCE ON MATHEMATICAL MODELS AND SUPER COMPUTING IN NUCLEAR APPLICATIONS, Germany, p. 19 – 23, 1993.
- [6] ANGELUS A. Electricity price forecasting in deregulated markets, **The Electricity Journal**, 4, p. 32 – 41, 2001.



- [7] ARMSTRONG, J.S. e COLLOPY, F., Error measures for generalizing about forecasting methods: empirical comparisons, **International Journal Forecasting**, 8,p. 99 – 111, 1992.
- [8] ASKIN, O. E. e GOKALP, F., Comparing the Predictive and Classification Performances of Logistic Regression and Neural Networks: A Case Study on Timss 2011, **Procedia - Social and Behavioral Sciences**, v. 106, p. 667-676, 2013.
- [9] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 5422**: Projeto de Linhas de Transmissão de Energia Elétrica, Rio de Janeiro, 1985.
- [10] BAGLEY, S.C., WHITE, H. e GOLOMB, B.A. Logistic regression in the medical literature: Standards for use and reporting, with particular attention to one medical domain, **Journal of Clinical Epidemiology**, 54, 979-985, 2001.
- [11] BALESTRASSI, P. P. et al., Design of experiments on neural networks training for nonlinear time series forecasting, **Neurocomputing**, 72, p. 1160-1178, 2009.
- [12] BALESTRASSI, P.P., SAMOHYL, R. W, e MEURER, R., Métodos de previsão usando a série econométrica de Nelson-Plosser, In: XVIII ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO - ENEGEP, Anais, Niterói: Universidade Federal Fluminense, 1998.
- [13] BALKIN S. D. e J. K. Ord, Automatic neural network modeling for univariate time series, **International Journal of Forecasting**, 16, p. 509 – 515, 2000.

- [14] BASSI, P. et al., Prognostic accuracy of an artificial neural network in patients undergoing radical cystectomy for bladder cancer: a comparison with logistic regression analysis, **British Journal of Urology (BJU) International**, 99, p. 1007-1012, 2007.
- [15] BEN-DAVID, A., About the relationship between ROC curves and Cohen's kappa, **Engineering Applications of Artificial Intelligence**, v. 21, n. 6, p. 874-882, 2008.
- [16] BISHOP, C. **Neural networks for pattern recognition**, Oxford: University Press 1995.
- [17] BODYANSKIY, Y. e POPOV, S., Neural network approach to forecasting of quasi periodic financial time series, **European Journal Oper. Res.** 175, (3), p. 1357 – 1366, 2006.
- [18] BORQUE, A., The use of neural networks and logistic regression analysis for predicting pathological stage in men undergoing radical prostatectomy: a population based study, **The Journal of Urology**, v. 166, n. 5, p. 1672-8, 2001.
- [19] BOX, G.E.P. e JENKINS G.M. **Time series analysis: forecasting and control**, Revised ed. Holden Day, San Francisco, 1976.
- [20] BUHAMRA, S., SMAOUI, N. e GABR, M. The Box Jenkins analysis and neural networks: prediction and time series modeling, **Applied Mathematical Modeling**, 27, p. 805 – 815, 2003.
- [21] CAOCCI, G. et al., Comparison between an neural network and logistic regression in predicting acute graft-vs-host disease after unrelated donor hematopoietic stem cell transplantation in thalassemia patients, **Experimental hematology**, v. 38, n. 5, p. 426-33, 2010.
- [22] CHAN, K.K. e SPEDDING, T.A., On line optimization of quality in a manufacturing system. **Int. J. Prod. Res.** 39, (6), p. 1127 – 1145, 2001.

- [23] CHAVES, N. A. **Bootstrap em Séries Temporais**, Rio de Janeiro, 208f. Tese de doutorado em Engenharia Elétrica, Universidade Federal do Rio de Janeiro, 1991.
- [24] CHEN, H. et al., Performance comparison of artificial neural network and logistic regression model for differentiating lung nodules on CT scans, **Expert Systems with Applications**, v. 39, n. 13, p. 11503-11509, 2012.
- [25] CHEN, T. e CHEN, H. Universal approximation to nonlinear operators by neural networks arbitrary activation functions and its application to dynamical systems, **IEEE Transactions on Neural Networks**, 6, p. 911 – 917, 1995.
- [26] CHEN, R. e TASY, R.S., Functional coefficient autoregressive models, **Journal of the American Statistical Association**, 88, p. 298 – 308, 1993.
- [27] CHENG, B. e TITTERINGTON, D.M., Neural networks: A review from a statistical perspective, **Statistical Science**, 9, p. 2 – 54, 1994.
- [28] CHIANG, W.C., URBAN, T.L., e BALDRIDGE, G.W., A neural network approach to mutual fund net asset value forecasting, **Omega The International Journal of Management Science**, 24, p. 205 – 215, 1996.
- [29] CHIANG, W. K., ZHANG, D. e ZHOU, L., Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression, **Decision Support Systems**, v. 41, n. 2, p. 514-531, 2006.
- [30] CHOI, J. et al., Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS, **Engineering Geology**, v. 124, p. 12-23, 2012.

- [31] CHUN, F. K. H., Initial biopsy outcome prediction-head-to-head comparison of a logistic regression-based nomogram versus artificial neural network, **European Urology**, v. 51, n. 5, p. 1236-40; discussion 1241-3, 2007.
- [32] COLEMAN, D.E. e MONTGOMERY, D.C., **A systematic approach to planning for a designed industrial experiment** **Technometrics**, 35, p.1 – 27, 1993.
- [33] COMPANHIA PARANAENSE DE ENERGIA ELÉTRICA - COPEL, Disponível em < [www.copel.com](http://www.copel.com) >, Acesso em 20 de março de 2015.
- [34] CRARY GROUP, 2004, Disponível em < [www.webdoe.cc/](http://www.webdoe.cc/) >, Acesso em 10 de fevereiro de 2014.
- [35] CUBILES-DE-LA-VEGA, M.D., et al. Building neural network forecasting models from time series ARIMA models: A procedure and a comparative analysis, *Intell data anal.*,v.6, p.53-65, 2002.
- [36] DELONG, E. R., DELONG, D. M. e CLARKE-PEARSON, D. L. Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves?: A Nonparametric Approach, **International Biometric Society**, v. 44, n. 3, p. 837-845, 1988.
- [37] DERRINGER, G. e SUICH, R., Simultaneous optimization of several response variables, **J.Qual.Technol.**, 12, v.4, p. 214 – 219, 1980.
- [38] DIEBOLD, F.X. e NASON, J.A., Nonparametric exchange rate prediction, **Journal of International Economics**, 28, p. 15 – 332, 1990.

- [39] DIRUSSO, S. M. et al., Development of a model for prediction of survival in pediatric trauma patients: Comparison of artificial neural networks and logistic regression, **Journal of Pediatric Surgery**, v. 37, n. 7, p. 1098-1104, 2002.
- [40] DREISEITL, S. e OHNO - MACHADO, L. Logistic regression and artificial neural network classification models: a methodology review, **Journal of Biomedical Informatics**, v.35, p. 352-359, 2002.
- [41] EM: < [www.portalmedianeira.net.br](http://www.portalmedianeira.net.br) >, Acesso em 03 de maio de 2015.
- [42] EMPRESA DE PESQUISAS ENERGÉTICAS (EPE), Anuário Estatístico de Energia Elétrica 2014, Disponível em: < <http://www.epe.gov.br/AnuarioEstatisticodeEnergiaEletrica/Forms/Anurio.aspx> >, Acesso em 17 de abril de 2015.
- [43] ENEMUOH, E.U. e EL-GIZAWY, A.S., Optimal neural network model for characterization of process-induced damage in drilling carbon fiber reinforced epoxy composites, **Machining Science and Technology**, v.3, p. 389 – 400, 2003.
- [44] ENNETT, C.M., FRIZE, M. e CHARETTE, E., Improvement and automation of artificial neural networks to estimate medical outcomes, **Medical Engineering Physics**, 26, p. 321-328, 2004.
- [45] ERGUN, U. U. et al., Classification of carotid artery stenosis of patients with diabetes by neural network and logistic regression, **Computers in biology and medicine**, v. 34, n. 5, p. 389-405, 2004.

- [46] FAWCETT, T., An introduction to ROC analysis, **Pattern Recognition Letters**, v. 27, n. 8, p. 861-874, 2006.
- [47] FRANSES, P.H. e VAN HOMELEN, P., On forecasting exchange rates using neural networks, **Applied Financial Economics**, 8, p. 689 – 696, 1998.
- [48] FREEMAN, R. V. et al., Comparison of artificial neural networks with logistic regression in prediction of in-hospital death after percutaneous transluminal coronary angioplasty, **American heart journal**, v. 140, n. 3, p. 511-20, 2000.
- [49] FU, M.C., Optimization for simulation: Theory vs practice, **Inform Journal of Computing**, 14, p. 192 – 215, 1998.
- [50] FUCHS, R. D., **Transmissão de Energia Elétrica: Linhas Aéreas**, Volume 1, Livros Técnicos e Científicos, Rio de Janeiro, 1977.
- [51] GHIASSI, M., SAIDANE, H. e ZIMBRA, D.K., A dynamic artificial neural network model for forecasting time series events, **International Journal of Forecasting**, 21, p. 341 – 362, 2005.
- [52] GOOIJER, J.G., On threshold moving-average models, **Journal of Time Series Analysis**, 19, p. 1 – 18, 1998.
- [53] GRANGER, C.W.J. e ANDERSON, A.P., **An introduction to bilinear time series models**, Gottingen: Vandenhoeck, Ruprecht, 1978.
- [54] GRANJER, C., e TERASVIRTA, T., **Modelling nonlinear economic relationships**, Oxford University Press, 1993.

- [55] GREEN, M. et al., Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room, **Artificial Intelligence in Medicine**, 38, p.305-318, 2006.
- [56] GUIRELI,, R. C., **Previsão de carga de curto prazo de áreas elétricas através de técnicas de inteligência artificial**, Tese de doutorado apresentada a escola politécnica da USP, São Paulo: 2006.
- [57] GUTIÉRREZ, P. A. et al., Logistic regression product-unit neural networks for mapping *Ridolfia segetum* infestations in sunflower crop using multitemporal remote sensed data, **Computers and Electronics in Agriculture**, v. 64, n. 2, p. 293-306, 2008.
- [58] HAJMEER, M. e BASHEER, I., Comparison of logistic regression and neural network-based classifiers for bacterial growth, **Food Microbiology**, v. 20, n. 1, p. 43-55, 2003.
- [59] HAMILTON, J.D., A new approach to the economic analysis of nonstationary time series and the business cycle, **Econometrica**, 57, p. 357 – 362, 1998.
- [60] HARRINGTON, J., The desirability function, **Industrial Quality Control**, 21(10), p.494-498, 1965.
- [61] HAYKIN, S., **Neural Network: A Comprehensive Foundation**, 2nd edition, New Jersey: Prentice Hall, 1999.
- [62] HEDAYAT, A. e WALLIS, W.D., Hadamard matrices and their applications, *The Annals of Statistics*, v. 6, 6, p. 1184-1238, 1978.

- [63] HO, S.L., XIE, M. e GOH, T.N., A comparative study of neural network and Box Jenkins ARIMA modeling in time series prediction, **Computers Industrial Engineering**, 42, p. 371 – 375, 2002.
- [64] HORNIK, K., Some new results on neural network approximation, **Neural Networks**, 6, p. 1069 – 1072, 1993.
- [65] HORNIK, K., STINCHCOMBE, M. e WHITE, H., Multilayer feedforward networks are universal approximators, **Neural Networks**, 2, p. 359 – 366, 1989.
- [66] HOSMER, D. W e LEMESHOW, S., **Applied Logistic Regression**, 2 ed. New York: John Wiley, 1989.
- [67] HWARNG, H.B., Insights into neural network forecasting of time series corresponding to  $ARMA(p; q)$  structures, **Omega**, 29, p.273 – 289, 2001.
- [68] HWARNG, H.B. e ANG, H.T., A simple neural network for  $ARMA(p, q)$  time series, **Omega**, 29, p. 319 – 333, 2001.
- [69] JACK, V.T., Advantages and Disadvantages of Using Artificial Neural Networks versus Logistic Regression for Predicting Medical Outcomes, **Journal of Clinical Epidemiology**, 49, p. 1225-1231, 1996.
- [70] KABON, M., **TSGP - Time Series Genetic Programming School of Business**, University of Redlands, 2002.
- [71] KAASTRA, I. e BOYD, M., Designing a neural network for forecasting financial and economic time series, **Neurocomputing**, 10, p. 215 – 236, 1996.



- [72] KALAITZAKIS, K., STAVRAKAKIS,G.S. e ANAGNOSTAKIS, E.M., Short term load forecasting based on artificial neural networks parallel implementation, **Electric Power Systems Research**, 63, p. 185 – 196, 2002.
- [73] KARUNASINGHE, D.S.K. e LIONG, S.Y. Chaotic time series prediction with a global model: Artificial neural network, **Journal of Hydrology**, 2, p. 1 – 14, 2005.
- [74] KAWAKAMI, S., Development, validation, and head-to-head comparison of logistic regression-based nomograms and artificial neural network models predicting prostate cancer on initial extended biopsy, **European Urology**,54,p.601-611, 2008.
- [75] KHAW, J.F.C., LIM, B.S. e LIM, L.E.N., Optimal design of neural networks using the Taguchi method, **Neurocomputing**, 7, p. 225 – 245, 1995.
- [76] KHURI,A.I. e CONLON, M., Simultaneous optimization of multiple responses represented by polynomial regression functions, **Technometrics**,v. 23, 4,p.29-30, 1981.
- [77] KIM, K.J., Financial time series forecasting using support vector machines. **Neurocomputing**, 55,p. 307 – 319, 2003.
- [78] KIM, Y.S. e YUM, B.J., Robust design of multilayer feedforward neural networks: an experimental approach, **Engineering Applications of Artificial Intelligence**, 17, p. 249 – 263, 2004.
- [79] KLEIJNEN, J.P.C. et al. State of the art Review: A users guide to the brave new world of designing simulation experiments, **Inform Journal of Computing**, 7 (3), p. 263 – 289, 2005.

- [80] KLEIJNEN, J.P.C., **Statistical Tools for Simulation Practitioners**, Marcel Dekker Inc. Publ.,New York: 1987.
- [81] KOHZADI, N. et al. A comparison of artificial neural network and time series models for forecasting commodity prices, **Neurocomputing**, 10, p. 169 – 181, 1996.
- [82] KRAGER, H. e KUGLE, P., Nonlinearities in foreign exchange markets: a different perspective, **Journal of International Money and Finance**, 12,p. 195 – 208, 1993.
- [83] KURT, I., TURE, M. e KURUM, A. T., Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease, **Expert Systems with Applications**, v. 34, n. 1, p. 366-374, 2008.
- [84] LARASATI, A., DEYONG, C. e SLEVICTH, L., The Application of Neural Network and Logistics Regression Models on Predicting Customer Satisfaction in a Student-Operated Restaurant, **Procedia - Social and Behavioral Sciences**, v. 65, p. 94-99, 2012.
- [85] LEUNG, P. e TRAN, L. T., Predicting shrimp disease occurrence: artificial neural networks vs. logistic regression, **Aquaculture**, v. 187, n. 1-2, p. 35-49, 2000.
- [86] LIEW, P.-L. et al.,Comparison of artificial neural networks with logistic regression in prediction of gallbladder disease among obese patients. Digestive and liver disease?, **Official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver**, v. 39, n. 4, p. 356-62, 2007.
- [87] LIN, T.Y. e TSENG, C.H., Optimum design for artificial neural networks: an example in a bicycle derailleur system, **Engineering Applications of Artificial Intelligence**, 13,p. 3 – 14, 2000.

- [88] LIN, C.C. et al., Comparison of artificial neural network and logistic regression models for predicting mortality in elderly patients with hip fracture, **Injury**, v. 41, n. 8, p. 869-73, 2010.
- [89] MAGALLANES, J. F. e OLIVIERI, A. C., The effect of factor interactions in Plackett-Burman experimental designs, **Chemometrics and Intelligent Laboratory Systems**, v. 102, n. 1, p. 8-14, 2010.
- [90] MAKRIDAKIS, S., M. Hibon, The *M3* Competition: Results, conclusions and implications, **International Journal of Forecasting**, 16, p.451 – 476, 2000.
- [91] MAKRIDAKIS, S. et al. The accuracy of extrapolation (time series) methods: Results of a forecasting competition, **Journal of Forecasting**, 1, p. 111 – 153, 1982.
- [92] MAKRIDAKIS, S. et al. The *M2* Competition: A real time judgmentally based forecasting study, **International Journal of Forecasting**, 9, p. 5 – 23, 1993.
- [93] McLAREN, C. E. et al., Prediction of malignant breast lesions from MRI features: a comparison of artificial neural network and logistic regression techniques, **Academic radiology**, v. 16, n. 7, p. 842-51, 2009.
- [94] MONTGOMERY, D. C., **Design and Analysis of Experiments**, 7<sup>a</sup> edition, John Wiley Sons Inc., 2009.
- [95] MORTEZA, A. et al., Inconsistency in albuminuria predictors in type 2 diabetes: a comparison between neural network and conditional logistic regression. Translational research?, **The journal of laboratory and clinical medicine**, v. 161, n. 5, p. 397-405, 2013.

- [96] MOUNT, T., Market power and price volatility in restructured markets for electricity, **Decision Support Systems**, 30, p. 311 – 325, 2001.
- [97] MYERS, R.H. e MONTGOMERY, D.C., **Response Surface Methodology: Process and Product Optimization Using Designed Experiments**, Wiley, New York: 1995.
- [98] MURATA, N., YOSHIZAWA, S. e AMARI, S., Network information criterion determining the number of hidden units for an artificial neural network model, **IEEE Transactions on Neural Networks**, 5, 865 – 872, 1994.
- [99] NGUYEN, T. et al., Comparison of prediction models for adverse outcome in pediatric meningococcal disease using artificial neural network and logistic regression analyses, **Journal of clinical epidemiology**, v. 55, n. 7, p. 687-95, 2002.
- [100] NISKA, H. et al. Evolving the neural network model for forecasting air pollution time series, **Engineering Applications of Artificial Intelligence**, 17, p. 159 – 167, 2004.
- [101] OI, M. e ZHANG, G.P., An investigation of model selection criteria for neural network time series forecasting, **European Journal of Operational Research**, 132, p. 666 – 680, 2001.
- [102] OTTENBACHER, K.J. et al., Comparison of Logistic Regression and Neural Network Analysis Applied to Predicting Living Setting after Hip Fracture, **Annals of Epidemiology**, 14, p. 551-559, 2004.
- [103] OTTENBACKER, K. J. et al., Comparison of logistic regression and neural networks to predict rehospitalization in patients with stroke, **Journal of clinical epidemiology**, v. 54, n. 11, p. 1159-65, 2001.

- [104] PAI, P.F. e HONG, W.C., Support vector machines with simulated annealing algorithms in electricity load forecasting, **Energy Conversion and Management**, 46, p. 2669 – 2688, 2005.
- [105] PAVLEKOVIC, M., BENSIC, M.e ZEKIC-SUSAC, M., Modeling children’s mathematical gift by neural networks and logistic regression, **Expert Systems with Applications**, v. 37, n. 10, p. 7167-7173, 2010.
- [106] PLACKETT, R. L. e BURMAN, J. P.,The design of optimum multifactorial experiments, **Biometrika**, v. 33, n. 4, p. 305-325, 1946.
- [107] POURSHAHARIAR, H., Correct vs. accurate prediction: A comparison between prediction power of artificial neural networks and logistic regression in psychological researches, **Procedia - Social and Behavioral Sciences**, v. 32, n. 2011, p. 97-103, 2012.
- [108] PRADHAN, B. e LEE, S., Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling, **Environmental Modelling Software**, v. 25, n. 6, p. 747-759, 2010.
- [109] PRIESTLEY, M.B., State dependent models: a general approach to nonlinear time series analysis, **Journal of Time Series Analysis**, 1, p. 47 – 71, 1980.
- [110] REED, R., Pruning algorithms- a survey, **IEEE Transactions on Neural Networks**, 4, ,p. 740 – 747, 1993.

- [111] RIPPEL, L. I., **Estudo em túnel de vento do arrasto aerodinâmico sobre torres treliçadas de linhas de transmissão**, Dissertação de mestrado, UFRGS, Porto Alegre, 2005.
- [112] ROY, A., KIM, L.S. e MUKHOPADHYAY, S., A polynomial time algorithm for the construction and training of a class of multilayer perceptrons, **Neural Networks**, 6, p. 535–545, 1993.
- [113] SANTOS, A.M., Usando Redes Neurais Artificiais e Regressão Logística na Predição da Hepatite A, **Revista Brasileira de Epidemiologia**, 8, p. 117-126, 2005.
- [114] SCHUMACKER, M., ROBNER, R. e VACH, W., Neural networks and logistic regression: Part I, **Computational Statistics & Data Analysis**, 21, p. 661-682, 1996.
- [115] SHI, Z., TAMURA, Y. e OZAKI, T., Nonlinear time series modelling with the radial basis function based state dependent autoregressive model, **International Journal of Systems Science**, 30, (7), p. 717 – 727, 1999.
- [116] SILVA, G.L., **Modelos logísticos para dados binários**, Dissertação de mestrado em Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo, 1992.
- [117] SONG, J. H. et al., Comparative analysis of logistic regression and artificial neural network for computer-aided diagnosis of breast masses, **Academic radiology**, v. 12, n. 4, p. 487-95, 2005.
- [118] SOUZA, E. C. de, **Análise de influência local no modelo de regressão logística**, Dissertação de mestrado. Escola Superior de Agricultura Luiz de Queiroz. Piracicaba, 2006.

- [119] SPACKMAN, K. A., Combining logistic regression and neural networks to create predictive models, In: SIXTEENTH ANNUAL SYMPOSIUM ON COMPUTER APPLICATIONS IN MEDICAL CARE, McGraw-Hill, p. 456-459, New-York: 1992.
- [120] SPALL, J.C., **Introduction to stochastic search and optimization; estimation, simulation, and control**, New York: Wiley, 2003.
- [121] STATSOFT, Inc., **Statistica (data analysis software system)**, version 7.1. [www.statsoft.com](http://www.statsoft.com), 2005.
- [122] SUKTHOMYA, W. e TANNOCK, J., The optimisation of neural network parameters using Taguchis design of experiments approach: an application in manufacturing process modeling, **Neural Comput Applic**, 14, p. 337 – 344, 2005a.
- [123] SUKTHOMYA, W. e TANNOCK, J., The training of neural networks to model manufacturing processes, **Journal of Intelligent Manufacturing**, 16, p. 39 – 51, 2005b.
- [124] SWIDERSKI, B., KUREK, J. e OSOWSKI, S., Multistage classification by using logistic regression and neural networks for assessment of financial condition of company, **Decision Support Systems**, v. 52, n. 2, p. 539-547, 2012.
- [125] TAGUCHI, G., **Systems of Experimental Design** ,vols.1, UNIPUB/Kraus International Publications, NewYork: 1987.
- [126] TASAY, R., **Analysis of financial time series**, 2nd Edition, Wiley Interscience, 2005.
- [127] TERASVIRTA, T., VAN DIJK, D. e MEDEIROS, M.C., Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re examination, **International Journal of Forecasting**, 21, p. 755 – 774, 2005.

- [128] TERRIN, N. et al., External validity of predictive models: A comparison of logistic regression, classification trees, and neural networks, **Journal of Clinical Epidemiology**, 56, p. 721-729, 2003.
- [129] TONG, H., **On a threshold model**. In C. H. Chen (ed.), Pattern recognition and signal processing, Sijhoff Noordhoff, Amsterdam: 1978.
- [130] TROCINE, L. et al., Finding important independent variables through screening designs: a comparison of methods, In: PROCEEDINGS OF THE WINTER SIMULATION CONFERENCE, 2000.
- [131] TROCINE, L., **An efficient, effective, and robust procedure for screening more than 20 independent variables employing a genetic algorithm**, Dissertation, University of Central Florida, Orlando, USA: 2001.
- [132] TRTICA- MAJNARIC , L. et al., Prediction of influenza vaccination outcome by neural networks and logistic regression, **Journal of biomedical informatics**, v. 43, n. 5, p. 774-81, 2010.
- [133] TSENG, F.M., YU, H.C. e TZENG, G.H., Combining neural network model with seasonal time series ARIMA model, **Technological Forecasting Social Change**, 72, (69), p. 71 – 87, 2002.
- [134] TU, J. V., Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes, **Journal of clinical epidemiology**, v. 49, n. 11, p. 1225-31, 1996.



- [135] Vach, W., Robner, R., Schumacker, M., Neural networks and logistic regression: Part II, **Computational Statistics & Data Analysis**, 21, p.683-701,1996.
- [136] VALLEJOS, J. A. e McKINNON, S. D., Logistic regression and neural network classification of seismic records, **International Journal of Rock Mechanics and Mining Sciences**, v. 62, p. 86-95, 2013.
- [137] WANG, Z. et al. A procedure for determining the topology of multilayer feedforward neural networks, **Neural Networks**, 7, p. 291 – 300, 1994.
- [138] WAZEN,, R.N., **Avaliação da suscetibilidade de falhas de estruturas metálicas em linhas de transmissão**, Dissertação de mestrado apresentada no programa de pós graduação em engenharia elétrica da UFPA, 2011.
- [139] YAO, X., Evolving artificial neural networks, **Proceedings of the IEEE**, 87, (9), p. 1423 – 1447, 1999.
- [140] YESILNACAR, E. e TOPAL, T., Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey), **Engineering Geology**, v. 79, n. 3-4, p. 251-266, 2005.
- [141] YILMAZ, I.,Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey) **Computers Geosciences**, v. 35, n. 6, p. 1125-1138, 2009a.
- [142] ZHANG, G.P., An investigation of neural networks for linear time series forecasting, **Computers Operations Research**, 28, p. 1183 – 1202, 2001.

- [143] ZHANG, G.P., Time series forecasting using a hybrid ARIMA and neural network model, **Neurocomputing**, 50, p. 159 – 175, 2003.
- [144] ZHANG, G.P. e QI, M., Neural network forecasting for seasonal and trend time series, **European Journal of Operational Research**, 160, p. 501 – 514, 2005.
- [145] ZHANG, G.P., PATUWO, B.E. e HU, M.Y., Forecasting with artificial neural networks: The state of the art, **International Journal of Forecasting**, 14, p. 35 – 62, 1998.
- [146] ZHANG, G.P., PATUWO, B.E. e HU, M.Y., A simulation study of artificial neural networks for nonlinear time series forecasting, **Computers Operations Research**, 28, p. 381 – 396, 2001.

# Apêndice A

## Conjunto de Dados

A tabela A.1 apresenta parte do conjunto de dados contendo 181 estruturas metálicas de linhas de transmissão em operação em um dado sistema elétrico mantido pela COPEL (2015).

Tabela A.1: Parte do conjunto de dados

Observação	Tensão	Vento	Relevo	Massa fria	Massa quente	Temperatura	Terreno	Capacidade	Função	Fundação	Resultado
1	69	14	planalto	paralela	perpendicular	17	C	alta	ancoragem	grelha	nenhum
2	69	14	planalto	paralela	perpendicular	17	C	alta	suspensão	grelha	nenhum
3	69	14	planalto	paralela	perpendicular	17	C	baixa	ancoragem	grelha	nenhum
4	69	14	planalto	paralela	perpendicular	17	C	baixa	ancoragem	<i>stub</i>	nenhum
5	69	14	planalto	paralela	perpendicular	17	C	baixa	suspensão	grelha	nenhum
7	69	15	planalto	perpendicular	transversal	18	D	alta	suspension	grelha	nenhum
8	69	19	planície	perpendicular	transversal	21	B	baixa	suspensão	grelha	falha
9	69	20	planalto	paralela	perpendicular	17	B	baixa	suspensão	grelha	falha
10	69	20	planalto	paralela	transversal	19	A	baixa	suspensão	grelha	falha
11	69	20	planalto	transversal	perpendicular	17	A	baixa	suspensão	grelha	falha
12	69	23	planície	transversal	paralela	20	A	baixa	suspensão	grelha	falha
13	88	20	planalto	paralela	transversal	22	A	baixa	suspensão	grelha	falha
14	88	20	planície	paralela	transversal	22	B	baixa	suspensão	grelha	falha
15	138	14	serra	paralela	transversal	16	B	baixa	ancoragem	grelha	nenhum
16	138	14	serra	paralela	transversal	16	B	baixa	suspensão	grelha	nenhum
⋮											
179	230	26	planalto	paralela	transversal	17	D	alta	suspensão	<i>stub</i>	nenhum
180	230	26	planalto	paralela	transversal	17	D	baixa	suspensão	<i>stub</i>	nenhum
181	500	17	planície	perpendicular	transversal	20	B	baixa	suspensão	grelha	falha

Fonte: COPEL (2015)

# Apêndice B

## Matrizes de Hadamard e o

## Delineamento de Plackett-Burman

O delineamento de Plackett-Burman está associado as chamadas matrizes de Hadamard. As matrizes de Hadamard possuem diversas aplicações, em particular são usadas em códigos de correção, como Reed-Muller, análise sensorial e experimentos fatoriais.

Uma matriz quadrada cujas entradas são  $-1$  e  $+1$  é chamada matriz de Hadamard de ordem  $n$  se

$$\mathbf{H}\mathbf{H}^T = n\mathbf{I}_n \quad (\text{B.1})$$

A equação B.1 implica que  $\mathbf{H}$  é uma matriz não singular e sua inversa é dada por  $\mathbf{n}^{-1}\mathbf{H}^T$ .

Consequentemente

$$\mathbf{H}^T\mathbf{H} = n\mathbf{I}_n$$

Geometricamente, cada duas linhas distintas numa matriz de Hadamard representam dois

vetores perpendiculares. As matrizes

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$$

são exemplos de matrizes de Hadamard.

Se  $\mathbf{H}$  é uma matriz de Hadamard, então a matriz

$$\begin{pmatrix} \mathbf{H} & \mathbf{H} \\ \mathbf{H} & -\mathbf{H} \end{pmatrix}$$

também é uma matriz de Hadamard.

Existem diversas formas de se construir matrizes de Hadamard. Se  $\mathbf{A}$  é uma matriz do tipo  $n \times p$  com entradas  $a_{ij}$  e  $\mathbf{B}$  é uma matriz do tipo  $r \times s$ , então o produto direto das matrizes  $\mathbf{A}$  e  $\mathbf{B}$ , denotado por  $\mathbf{A} \otimes \mathbf{B}$  é definido por

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1p}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2p}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \dots & a_{np}\mathbf{B} \end{pmatrix} \quad (\text{B.2})$$

Observa-se que a matriz  $\mathbf{A} \otimes \mathbf{B}$  é do tipo  $nr \times ps$ . O teorema B.0.1, cuja demonstração é encontrada em Hedayat e Wallis (1978), apresenta um método para construir matrizes de Hadamard.

**Teorema B.0.1.** *Se  $\mathbf{H}_1$  é uma matriz de Hadamard de ordem  $m$  e  $\mathbf{H}_2$  é uma matriz de Hadamard de ordem  $n$ , então  $\mathbf{H}_1 \otimes \mathbf{H}_2$  é uma matriz de Hadamard de ordem  $mn$ .*

Para se construir uma matriz de Hadamard de ordem 4, basta considerar, por exemplo

$$\begin{pmatrix} + & + \\ + & - \end{pmatrix} \otimes \begin{pmatrix} + & + \\ + & - \end{pmatrix} = \begin{pmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{pmatrix}$$

**Corolário B.0.2.** *Existe uma matriz de Hadamard de ordem  $2^k$  para todo inteiro positivo.*

A pergunta mais importante na teoria das matrizes de Hadamard é o da existência. A conjectura de Hadamard propõe que uma matriz de Hadamard de ordem  $4n$  existe para todo  $n$  inteiro positivo. O teorema B.0.3 apresenta uma condição suficiente para a existência de uma matriz de Hadamard de ordem  $p^\alpha$ . A prova desse teorema é encontrada em Hedayat e Wallis (1978).

**Teorema B.0.3.** *Se  $p^\alpha$  é um número primo e  $p^\alpha + 1 \equiv 0 \pmod{4}$ , então existe uma matriz de Hadamard de ordem  $p^\alpha + 1$ .*

Uma matriz contendo apenas dois símbolos como elementos de entrada com  $N$  linhas e  $k$  colunas é um  $(N, k, p)$  delineamento de seleção se para cada escolha de  $p$  colunas, cada uma das  $2^p$  linhas aparece uma única vez. Delineamentos de seleção são muito úteis em situações em que um grande número de fatores ( $q$ ) é examinado mas apenas ( $k$ ) fatores são importantes.

Delineamentos de seleção surgem a partir das matrizes de Hadamard que tradicionalmente tem sido utilizadas para identificar apenas efeitos principais devido a sua estrutura complexa de alisamento. Sem perda de generalidade, pode-se afirmar que a primeira coluna de alguma matriz de Hadamard contém apenas elementos iguais a um. Então, removendo essa coluna, obtém-se um  $(N, N - 1, p)$  delineamento de seleção com  $p \geq 2$ .

Dentre os delineamentos de seleção, destaca-se o delineamento de Plackett-Burman. Os delineamentos de Plackett-Burman podem ser gerados a partir da primeira linha, que consiste de  $N - 1$  elementos, por um arranjo cíclico. A segunda linha é gerada removendo todas as entradas da primeira linha uma posição para a direita e colocando o último elemento na primeira posição. A terceira linha é gerada a partir da segunda com o procedimento análogo e o processo é interrompido quando  $N - 1$  linhas são geradas. A linha de elementos iguais a -1 é então adicionada como a última linha completando assim o delineamento com  $N$  experimentos e  $N - 1$  colunas. Adicionando uma coluna de elementos todos iguais a um no delineamento de Plackett-Burman com  $N$  experimentos obtém-se uma matriz de Hadamard de ordem  $N$ . Plackett-Burman construíram matrizes de Hadamard de ordem  $N$ , para todo  $N \leq 100$  exceto para  $N = 92$ .

A tabela B.1 apresenta as primeiras linhas de um delineamento de Plackett-Burman para alguns valores de  $N$ .

Tabela B.1: Estrutura das primeiras linhas de alguns delineamentos de Plackett-Burman

---

N=8	+	+	+	-	+	-	-																
N=12	+	+	-	+	+	+	-	-	-	+	-												
N=16	+	+	+	+	-	+	-	+	+	-	-	+	-	-	-								
N=20	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-				
N=24	+	+	+	+	+	-	+	-	+	+	-	-	+	+	-	-	+	-	+	-	-	-	-

---