



UFRJ

PREVISÃO A CURTÍSSIMO PRAZO DE EVENTOS METEOROLÓGICOS EXTREMOS
PARA A ÁREA DE INFLUÊNCIA DO AEROPORTO INTERNACIONAL TOM JOBIM
UTILIZANDO TÉCNICAS DE REDES NEURAIAS ARTIFICIAIS

Alessana Carrijo Rosette

Dissertação de Mestrado do curso de Pós-graduação
em Meteorologia, IGEO, da Universidade Federal do
Rio de Janeiro.

Orientador: Gutemberg Borges França

Rio de Janeiro
Agosto de 2011

PREVISÃO A CURTÍSSIMO PRAZO DE EVENTOS METEOROLÓGICOS EXTREMOS
PARA A ÁREA DE INFLUÊNCIA DO AEROPORTO INTERNACIONAL TOM JOBIM
UTILIZANDO TÉCNICAS DE REDES NEURAIAS ARTIFICIAIS

Alessana Carrijo Rosette

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE
GEOCIÊNCIAS DE PÓS-GRADUAÇÃO DA UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO
GRAU DE MESTRE EM METEOROLOGIA.

Aprovada por:

Prof. Gutemberg Borges França, Ph.D.

Prof. Nelson Francisco Favilla Ebecken, D. Sc.

Manoel Valdonel de Almeida, D. Sc.

Prof^a. Célia Maria Paiva, D. Sc.

Prof. Wallace Figueiredo Menezes, D. Sc.

Rio de Janeiro
Agosto de 2011

Rosette, Alessana Carrijo

Previsão a Curtíssimo Prazo de Eventos Meteorológicos Extremos para a área de influência do Aeroporto Internacional Tom Jobim utilizando Técnicas de Redes Neurais Artificiais./Alessana Carrijo Rosette - Rio de Janeiro: UFRJ/IGEO, 2010.

xxii, 102 f. : il. ; 21,8 cm.

Orientador: Gutemberg Borges França

Dissertação (mestrado) – UFRJ/ IGEO/ Programa de Pós-graduação em Meteorologia, 2011.

Referências Bibliográficas: p. 87-94.

1. Redes Neurais Artificiais 2. Previsão de curtíssimo prazo 3. Eventos Convectivos Extremos I. França, Gutemberg Borges *et al.* II. Universidade Federal do Rio de Janeiro, IGEO, Programa de Pós-graduação em Meteorologia. III. Título.

*A Deus, por estar sempre presente em minha
vida, e tornar tudo possível.*

AGRADECIMENTOS

Primeiramente, agradeço a Deus por me abençoar nessa caminhada.

À minha família: minha mãe Ana Cristina Martins Carrijo Rosette minhas irmãs Aleana Carrijo Rosette Bandeira e Annelise Carrijo Rosette, pelo enorme incentivo em momentos de angústia. Em especial ao meu pai Alexandre Carvalhaes Rosette que muito colaborou com conselhos valiosos.

Agradeço o auxílio do meu orientador Prof. Gutemberg Borges França e colaboradores D. Sc. Manoel Valdonel de Almeida, que com críticas construtivas e todo apoio necessário, foram os responsáveis por um melhor desempenho e exatidão na concretização deste trabalho.

A todos que colaboraram direta ou indiretamente com informações de grande relevância para a realização desta pesquisa, destacando-se entre eles os professores Wallace Figueiredo Menezes e Maria Gertrudes Alvarez Justi da Silva e os amigos Victor Hugo Pezzini de Meireles e Aline Gomes Cerqueira.

Também agradeço imensamente, sem esquecer o oportuno e benquisto gesto, o favor e o préstimo dos meus amigos do Laboratório de Meteorologia Aplicada – LMA, em especial, ao Igor Balteiro e Victor Bastos Daher .

Resumo da Dissertação de Mestrado do curso de Pós-graduação em Meteorologia, IGEO, da Universidade Federal do Rio de Janeiro.

PREVISÃO A CURTÍSSIMO PRAZO DE EVENTOS METEOROLÓGICOS EXTREMOS
PARA A ÁREA DE INFLUÊNCIA DO AEROPORTO INTERNACIONAL TOM JOBIM
UTILIZANDO TÉCNICAS DE REDES NEURAIAS ARTIFICIAIS

Alessana Carrijo Rosette

Agosto/2011

Orientador: Gutemberg Borges França.

Programa: Meteorologia

Na tentativa de desenvolver um sistema de previsão a curtíssimo prazo para o Aeroporto Internacional Tom Jobim – RJ, foram utilizados 13 anos de dados meteorológicos da Estação Meteorológica de Superfície e Estação Meteorológica de Altitude do Destacamento de Controle do Espaço Aéreo do Galeão, junto com os dados de acumulados horários de precipitação e registros de raios da Ilha do Governador e proximidades. Esses dados passaram por uma análise rigorosa de correções de erros e preenchimento da série correspondente aos dados ausentes. Desse grupo de dados já pré-processados foram escolhidas minuciosamente duas séries: uma para o treinamento da Rede Neural Artificial; e outra a fim de validar o processamento feito pela rede. Em busca de um melhor desempenho da rede, foram utilizados os registros de raios, obtendo-se como médias de acertos e alarmes falsos do sistema de previsão para as três horas, as respectivas respostas: a) 88% e 15% para primeira hora, b) 87% e 14% para segunda hora, c) 85% e 19% para terceira hora.

Abstract of Master Thesis presented to PPGM/UFRJ as a partial fulfillment of the requirements for the degree of Mester of Meteorology.

NOWCASTING OF EXTREME WEATHER EVENTS FOR INTERNATIONAL AIRPORT
TOM JOBIM IN FLUENCE AREA BY APPLICATION OF TECHNIQUES OF ARTIFICIAL
NEURAL NETWORKS

Alessana Carrijo Rosette

August/2011

Advisors: Gutemberg Borges França.

Department: Meteorology

In an attempt to develop a system of short-term prediction for the International Airport Tom Jobim in the state of Rio de Janeiro, Brazil, were used 13 years of meteorological data from the meteorological station of surface and elevation of the Detachment of Airspace control in Rio, along with the data accumulated hourly precipitation and ray records Governor's Island and vicinity. The data have a undergone a rigorous analysis of bug fixes and completion of the series corresponding to the missing data. Of this group of data has been preprocessed carefully chosen two series: one for training the Artificial Neural Network, and another to validate the processing done by the network. In search of a better network performance, was used ray records, obtained as averages of correct responses and false alarms of the forecasting system for three hours, the answers: a) 88% e 15% for the first hour, b) 87% e 14% for the second hour, c) 85% e 19% for the third hour.

SUMÁRIO

LISTA DE FIGURAS	xi
LISTA DE TABELAS	xiii
LISTA DE SIGLAS E ABREVIATURAS	xv
CAPÍTULO 1 – INTRODUÇÃO.....	1
1.1 MOTIVAÇÃO.....	3
1.2 JUSTIFICATIVAS.....	3
1.3 OBJETIVO.....	5
CAPÍTULO 2 – FUNDAMENTAÇÃO TEÓRICA	7
2.1 CONCEITOS BASILARES	8
2.2 PROCESSOS TERMODINÂMICOS	10
2.2.1 ÍNDICES DE INSTABILIDADE	10
2.3 PROCESSOS DINÂMICOS.....	16
2.3.1 FLUXOS DE BAIXOS NÍVEIS	17
2.4 PREVISÃO METEOROLÓGICA	19
2.4.1 PREVISÕES METEOROLÓGICAS E AVIAÇÃO	20
2.5 REDES NEURAIS ARTIFICIAIS.....	23
2.5.1 INTRODUÇÃO A REDES NEURAIS ARTIFICIAIS.....	23
2.5.2 APRENDIZADO.....	25
2.5.3 FUNÇÃO DE ATIVAÇÃO	31
2.5.4 TIPOS DE REDES NEURAIS.....	34
2.5.5 REDES NEURAIS PROBABILÍSTICAS	36
2.5.6 ALGORÍTMO DE CORRELAÇÃO EM CASCATA	37
CAPÍTULO 3 – DADOS	38
3.1 INTRODUÇÃO.....	38
3.2 PRÉ-PROCESSAMENTO	41
3.3 ANÁLISE PRELIMINAR DOS DADOS	43
3.4 CRIAÇÃO DE VARIÁVEIS DERIVADAS	47
3.5 SELEÇÃO DO EVENTO EXTREMO E CRIAÇÃO DAS VARIÁVEIS DE SAÍDA DA RNA	49
3.6 ESCOLHA DAS VARIÁVEIS DE ENTRADA DA RNA.....	50
CAPÍTULO 4 – MÉTODOS.....	52
CAPÍTULO 5 – RESULTADOS E ANÁLISES	58
5.1 AVALIAÇÃO DA RNA COM DADOS DE RAIOS	70
5.2 ESTUDO DE CASO.....	76
CAPÍTULO 6 – CONCLUSÕES E SUGESTÕES	83
6.1 CONCLUSÕES	83

6.2 SUGESTÕES.....	84
APÊNDICE I	95
ANEXO A – CARTA DE VISIBILIDADE.....	96
ANEXO B – FORMATAÇÃO DOS DADOS DE PRECIPITAÇÃO.....	98
ANEXO C – VARIÁVEL PESO	100
ANEXO D – CORRELAÇÃO.....	101
ANEXO E – IEPV105-78	102

LISTA DE FIGURAS

Figura 1.1 – Aeroporto Internacional do Galeão. A imagem foi obtida na internet pelo site Ambiente Brasil, publicada em 3 de maio de 2010	04
Figura 2.1 – Componentes principais de um neurônio biológico (Fonte: Braga <i>et al.</i> , 1998).....	23
Figura 2.2 – Neurônio de McCulloch e Pitts (Fonte: Braga <i>et al.</i> , 1998).....	24
Figura 2.3 – Aprendizado Supervisionado (Fonte: Braga, 2000).....	27
Figura 2.4 – Aprendizado por Reforço (Fonte: Braga, 2000).....	28
Figura 2.5 – Aprendizado não-supervisionado (Fonte: Braga, 2000).....	28
Figura 2.6 – Gráfico de uma função de ativação limiar.....	32
Figura 2.7 – Exemplo de gráfico de uma função sigmóide logística.....	33
Figura 2.8 – Gráfico genérico de uma função gaussiana.....	33
Figura 2.9 – Exemplos de arquiteturas de RNA (Fonte: Braga, 2000).....	34
Figura 3.1 – Imagem de satélite da região de estudo (Fonte: Google Maps)	38
Figura 3.2 – Distribuição dos sensores da RINDAT no território brasileiro (Fonte: RINDAT).....	40
Figura 3.3 – Localização dos postos pluviométricos da Fundação Geo-Rio (Fonte: Alerta Rio)	40
Figura 3.4 – Pesos correspondentes ao grau de certeza das variáveis dos dados de altitude	42
Figura 3.5 – Histograma de frequências das faixas de tempo das estações do Galeão, Santos Dumont, Campos dos Afonsos e Santa Cruz, no período de janeiro de 1997 a dezembro de 2009	44
Figura 3.6 – Frequência horária de chuva forte e/ou trovada durante os meses dentro do período de 1997 a 2009, na área do Aeroporto Internacional do Galeão e adjacências, compreendidas nas coordenadas geográficas de 23° a 22° 36' S de latitude e 43° 27' a 43° 03' W de longitude	45
Figura 3.7 – Frequências horárias de chuva forte e/ou trovada na área do Aeroporto Internacional do Galeão e adjacências, compreendidos nas coordenadas geográficas de 23° a 22° 36' S de latitude e 43° 27' a 43° 03' W de longitude, para cada hora do dia no período de janeiro de 1997 a dezembro de 2009	46

Figura 3.8 – Peso relativo ao percentual da frequência de chuva forte e/ou trovoadas durante o dia no período de janeiro de 1997 à dezembro de 2009	47
Figura 4.1 – Fluxograma da metodologia aplicada à previsão de curtíssimo prazo de eventos extremos para o Aeroporto do Galeão – RJ.....	52
Figura 4.2 – Ilustração do método de divisão dos dados da rede	53
Figura 4.3 – (a) Histogramas de frequência das faixas de tempo do arquivo geral; (b) do arquivo de validação da rede	53
Figura 5.1 – Desempenho médio da Rede Neural, avaliando sensibilidade e especificidade de cada faixa de tempo para primeira, segunda e terceira hora de previsão.....	60
Figura 5.2 – (a)-(h) Frequência e porcentagem cumulativa das faixas de tempo dos conjuntos de treinamento. Os eventos meteorológicos são relativos ao período de janeiro de 1997 a dezembro de 2007	61
Figura 5.3 – Gráfico de análise da sensibilidade e especificidade para 1ª hora	65
Figura 5.4 – Gráfico de análise da sensibilidade e especificidade para 2ª hora	67
Figura 5.5 – Gráfico de análise da sensibilidade e especificidade para 3ª hora	69
Figura 5.6 – (a) Imagem do satélite GOES-10 para o dia 12/12/2007 às 2345Z; (b) Imagem do satélite GOES-10 para o dia 13/12/2007 às 0030Z	77
Figura 5.7 – Imagens de radar entre os dias 12 e 13 de dezembro de 2007: (a) 12/12 – 22:15 h, (b) 12/12 – 23:00 h, (c) 13/12 – 00:00 h, (d) 13/12 – 01:00 h.....	78
Figura 5.8 – Quantidade de precipitação acumulada com escala em milímetros das 12 Z de 12/12/2007 para as 12 Z de 13/12/2007	79
Figura 5.9 – Análise dos dados de tempo presente registrados por observadores de EMS comparados com registros de descargas atmosféricas obtidos da RINDAT	81

LISTA DE TABELAS

Tabela 2.1 – Classificação de intensidade do chuvisco (fonte: MCA 105-2/2010)	9
Tabela 2.2 – Classificação de intensidade da chuva (fonte: MCA 105-2/2010)	9
Tabela 2.3 – Tabela de classificadores e fenômenos de tempo que foram encontrados nos dados da estação SBGL no período de 01/01/1997 a 18/10/2009 (fonte: Adaptado do MCA 105-10/2010)	10
Tabela 2.4 – Probabilidade de formação tormentas com o Índice K. (fonte: Adaptado de Dias, 2008)	12
Tabela 2.5 – Valores de referência para a probabilidade de formação tormentas com o Índice CAPE. (fonte: Adaptado de Dias, 2008).....	14
Tabela 2.6 – Probabilidade de formação tormentas com o Índice TT. (fonte: Adaptado de Dias, 2008)	15
Tabela 2.7 – Probabilidade de formação de tormentas com o Índice Showalter (fonte: Adaptado de Dias, 2008)	16
Tabela 3.1 – Quantidades e porcentagens de dados de superfície adquiridos para as Estações Meteorológicas de Superfície encontradas no município do Rio de Janeiro	43
Tabela 3.2 – Frequência relativa ao número de ocorrências horárias e porcentagem dos fenômenos pertencentes as faixas de tempo 0, 1, 2 e 3 dos dados obtidos para as estações do Galeão, Santos Dumont, Campo dos Afonsos e Santa Cruz, no período de janeiro de 1997 a dezembro de 2009	45
Tabela 3.3 – Classificação das condições do tempo obtidas dos METARES	48
Tabela 3.4 – Valores de referência da variável MAIOR_PRECIPITAÇÃO.....	49
Tabela 3.5 – Variáveis de altitude, derivadas de sondagens do Galeão	50
Tabela 3.6 – Variáveis de superfície que foram utilizadas	51
Tabela 4.1 – Tabela de contigência 2x2 adaptada (fonte: Wilks, 2006).....	54
Tabela 4.2 – Descrição do desempenho de um modelo de previsão em relação à sensibilidade e à especificidade (Fonte: Almeida, 2009).	56
Tabela 5.1 – Teste de previsão para a primeira hora das quatro faixas de tempo	58
Tabela 5.2 – Teste de previsão para a segunda hora das quatro faixas de tempo.....	59
Tabela 5.3 – Teste de previsão para a terceira hora das quatro faixas de tempo.....	59

Tabela 5.4 – Quadro de tabelas de contingência relativas aos resultados da RNA sobre o conjunto de validação para primeira hora	64
Tabela 5.5 – Quadro de tabelas de contingência relativas aos resultados da RNA sobre o conjunto de validação para segunda hora.....	66
Tabela 5.6 – Quadro de tabelas de contingência relativas aos resultados da RNA sobre o conjunto de validação para terceira hora.....	68
Tabela 5.7 – Análise do desempenho da rede para as três horas de previsão.....	70
Tabela 5.8 – Resumo do melhor resultado que utilizou os registros de raios gerados apenas por operadores de EMS, para a primeira hora de previsão	71
Tabela 5.9 – Resumo do resultado utilizando os registros de raios gerados por operadores de EMS junto com os registros de raios da RINDAT, para a primeira hora de previsão	72
Tabela 5.10 – Quadro de comparação dos melhores resultados obtidos para a primeira hora de previsão	73
Tabela 5.11 – Resumo do melhor resultado que utilizou os registros de raios gerados apenas por operadores de EMS, para a segunda hora de previsão	73
Tabela 5.12 – Resumo do resultado utilizando os registros de raios gerados por operadores de EMS junto com os registros de raios obtidos da RINDAT, para a segunda hora de previsão	74
Tabela 5.13 – Quadro de comparação dos melhores resultados obtidos para a segunda hora de previsão	74
Tabela 5.14 – Resumo do melhor resultado obtido, utilizando os registros de raios gerados apenas por operadores de EMS, para a terceira hora de previsão	75
Tabela 5.15 – Resumo do resultado utilizando diferentes registros de raios, para a terceira hora de previsão	75
Tabela 5.16 – Quadro de comparação dos melhores resultados obtidos para a terceira hora de previsão	76
Tabela 5.17 – Registro da faixa de tempo observada e prevista para as respectivas horas dos dias 12 e 13 de dezembro de 2007.....	80
Tabela 5.18 – Registro da faixa de tempo observada e prevista para as respectivas horas do dia 18 de março de 2009	80
Tabela 5.19 – Índices de instabilidade referentes às horas estudadas	82
Tabela 6.1 – Resultado da previsão de eventos extremos na área de influência do AIRJ para as três horas futuras	84

LISTA DE SIGLAS E ABREVIATURAS

AIRJ – Aeroporto Internacional do Rio de Janeiro.

AIRMET – Airman's Meteorological.

AMS – American Meteorological Society.

APP – Approach (Controle de Aproximação de Aeródromo).

ART – Adaptive Resonance Theory.

AVHRR – Advanced Very High Resolution Radiometer.

AWS – Air Weather Service.

BES – Balanço de energia na superfície.

CAPE – Convective Available Potential Energy.

CB – Cumulonimbus.

CEPTEC - Centro de Previsão de Tempo e Estudos Climáticos.

CINE – Energia de inibição da convecção.

CK – índice CK.

DECEA – Departamento de Controle do Espaço Aéreo.

DIR – Direção.

DME - Distance Measuring Equipment.

EMA – Estação Meteorológica de Altitude.

EMS – Estação Meteorológica de Superfície.

ENIAC – Electrical Numerical Integrator and Computer.

ETA – Modelo Numérico Regional (modelo de previsão numérica) cujo nome é derivado da coordenada vertical η .

F – *Fog* (nevoeiro).

FAA - Federal Aviation Administration (Administração Federal de Aviação).

FN – Falso negativo.

FP – Falso positivo.

G – Fluxo de calor no solo

GAMET – General Aviation Meteorological Information.

GEOPCALC – Geopotencial Calculado.

GEOPOBS – Geopotencial Observado.

GOES – Geostacionary operational environmental satellites.

GRNN – Generalized Regression Neural Networks.

H – *Haze* (névoa)

IAG – Instituto Astronômico e Geofísico

ICEA - Instituto de Controle do Espaço Aéreo.

IFR - Instrument Flight Rule (Regra de Vôo por Instrumento).

INFRAERO - Empresa Brasileira de Infra-Estrutura Aeroportuária.

INMET – Instituto Nacional de Meteorologia.

INPE - Instituto Nacional de Pesquisa Espacial.

IPCC – Intergovernmental Panel on Climate Change.

IS – Índice Showalter.

JBN – Jato de baixo nível.

K – índice K.

L – Chuvisco.

LE – Fluxo de calor latente.

LFM – Limited –Area Fine Mesh.

LR – Índice Lapse Rate.

METAR – Meteorological Aerodrome Report.

MCA – Manual do Comando da Aeronáutica.

MOR - Meteorological Optical Range (Alcance Óptico Meteorológico).

MLP – Multi Layer Perceptron.

NCA – Nível de condensação por ascensão.

NCAR - National Center for Atmospheric Research (Centro Nacional para Pesquisa Atmosférica).

NCEP – National Center for Environmental Prediction.

MCP – McCulloch e Pitts.

NDB - Non Directional Bearing (Radiofarol não-direcional).

NEMEF0 - Neural Meteorological Forecast.

NEN – Nível de empuxo neutro.

NCC – Nível de condensação convectiva.

NWS - National Weather Service (Serviço Meteorológico Nacional).

PNN – Probabilistic Neural Network.

PTEMP – Temperatura potencial.

PTEMPGRAD – Gradiente da temperatura potencial.

QNH – Sigla oriunda do antigo *código Q* que significava a pergunta: Qual é a pressão atmosférica? Essa sigla indica a pressão atmosférica na estação meteorológica ajustada para a pressão atmosférica ao nível médio do mar, para fins aeronáuticos.

RAL - Research Applications Laboratory.

RBF – Radial-Basis Function.

REDEMET – Rede de Meteorologia do Comando da Aeronáutica.

RH – Umidade relativa.

RHGRAD – Gradiente da umidade relativa.

RINDAT – Rede Integrada Nacional de Detecção de Descargas Atmosféricas.

Rn – Radiação líquida.

RNA – Redes Neurais Artificiais.

RVR - Runway Visual Range (Alcance visual na pista).

SATVAP – Pressão de saturação do vapor.

SIGMET – Significant Meteorological Information.

SODAR - Sound Detecting and Ranging (Detecção e Alcance através do Som).

T – Trovoada.

Ta – Temperatura do ambiente.

TAF – Terminal Aerodrome Forecast

Td – Temperatura do ponto de orvalho.

TEMPGRAD – Gradiente de temperatura.

Tp – Temperatura da parcela.

TT – Total totals.

USP – Universidade de São Paulo.

UWND – Gradiente zonal do vento.

UWNDGRAD – Gradiente vertical do vento zonal.

VAPPRESS – Pressão de vapor.

VEL – Velocidade do vento.

VIS – Visibilidade.

VN – Verdadeiro negativo.

VP – Verdadeiro positivo.

VPP – Valor Preditivo Positivo.

VWND – Componente meridional do vento.

VWNDGRAD – Gradiente vertical da componente meridional do vento.

ZCAS - Zona de Convergência do Atlântico Sul.

ZCPS - Zona de Convergência do Pacífico Sul.

VHF - Very High Frequency - Frequência Muito Alta.

UTC - *Universal Time, Coordinated*, Tempo Universal Coordenado.

W – Pancada.

Capítulo 1 – INTRODUÇÃO

As fortes precipitações que atingem o município do Rio de Janeiro e seu entorno trazem à tona uma trágica realidade de desabamentos, deslizamentos e enchentes. Um exemplo claro e recente desses acontecimentos se refletiu na região serrana do Rio de Janeiro em janeiro de 2011, quando um forte temporal atingiu a região entre a noite de 11 de janeiro e a manhã do dia seguinte, provocando deslizamentos de terra e deixando mais de mil mortos e desaparecidos, causando destruição e prejuízos materiais e imateriais.

Mediante esses acontecimentos, a previsão do tempo torna-se cada vez mais importante para os vários setores da sociedade: a agricultura, a construção civil, a economia, a saúde pública, o transporte etc. Destacando-se a aviação: uma atividade bastante sensível às condições do tempo; a ocorrência de atrasos e cancelamentos de voos ou mesmo incidentes/acidentes aéreos em razão das condições adversas do tempo apontam a necessidade do enfrentamento desses possíveis problemas.

Por isso, o conhecimento das condições futuras do tempo é um fator essencial para o planejamento de rotas aéreas seguras. Esse planejamento, que é uma das fases mais importantes do voo, promove a antecipação da adoção de estratégias que minimizem possíveis danos que eventos meteorológicos extremos podem causar ao transporte aéreo. Daí, a importância de estudos sobre as condições meteorológicas favoráveis à formação de sistemas que gerem eventos extremos bem como o estabelecimento de critérios de classificação de eventos extremos segundo os riscos existentes para uma aeronave em voo e nos momentos de pouso e decolagem.

Karl e Easterling (1999); Groisman *et al.* (1999); Hegerl *et al.* (2006); Alexander *et al.* (2006) e Solow (1999), investigaram as tendências históricas de eventos climáticos e reconheceram possíveis ligações com os eventos extremos que ocorrem atualmente. Por definição, as características do que é chamado evento extremo do tempo podem variar de acordo com o lugar. Na verdade, as definições variam, mas em linhas gerais, um evento extremo é um evento pouco frequente dentro de sua distribuição de referência estatística em um lugar específico (MARENGO *et al.* 2004).

Existem diversos estudos no propósito de identificar tais fenômenos meteorológicos através de padrões atmosféricos de cada região. Segundo Easterling *et*

al. (2000), eventos meteorológicos extremos são eventos extraordinários que causam danos econômicos e sociais. Wisner *et al.* (1994) definem eventos extremos como aqueles que aumentam em intensidade e frequência os riscos naturais.

Na tentativa de identificar eventos extremos de precipitação, Konrad (1997) estudou, em sua pesquisa realizada no sudeste dos Estados Unidos, 312 eventos de chuva intensa utilizando dados horários de precipitação e campos gerados pelo modelo “*Limited-Area Fine Mesh*” (LFM) do *National Center for Atmospheric Research* (NCAR) e definiu evento extremo de precipitação como aquele em que o acúmulo de chuva registrado no pluviógrafo supera a marca de 50 mm em uma ou mais estações num período de 6 horas. Junker *et al.* (1999) fizeram, através de uma amostra que abrangeu nove estados do meio oeste dos Estados Unidos, uma climatologia sinótica utilizando as análises do sistema de assimilação de dados regional e imagens de satélite GOES, concluindo que tais eventos caracterizam-se por uma precipitação superior a 50 mm num período de 24 horas.

Liebmann *et al.* (2000) definiram como um parâmetro para aferição desses eventos, um percentual diário que superasse a média sazonal ou anual. Enquanto que Espírito Santo e Satyamurty (2002), no estudo sobre eventos extremos de precipitação na região sudeste do Brasil, utilizaram o banco de dados meteorológicos do Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) no período de 1997 a 2001 e as observações sinóticas do Instituto Nacional de Meteorologia (INMET) para a região do Brasil compreendida entre 15° S e 25° S. Esses pesquisadores examinaram os eventos extremos de precipitação acumulada em 24 horas, sendo adotado como critério de seleção para limiares de precipitação da ordem de 100 mm e 150 mm em 24 horas. Os autores salientaram a importância das previsões por conjunto para produzir previsões de probabilidade de ocorrência de eventos extremos.

A definição de limiares de precipitação para a identificação de eventos de chuvas intensas é bastante variável. Teixeira e Satyamurty (2007), através do estudo de casos de chuvas intensas na região sul do Brasil, consideraram evento extremo de precipitação como os eventos nos quais a isoietas de 50 mm cobrisse uma área de no mínimo 10.000 km² em 24 horas. Eventos desse tipo tornaram-se mais frequentes nos últimos anos (GROISMAN *et al.*, 2005). Groisman *et al.* (2005) identificaram tendências positivas de aumentos sistemáticos de extremos de chuva na região subtropical. Os autores também consideraram que o sudeste, desde 1940, tem mostrado aumentos na frequência de chuvas intensas de até quase 58% em 100 anos.

1.1 Motivação

Segundo o Departamento de Controle do Espaço Aéreo (DECEA), a definição e notificação de eventos extremos são vitais para a segurança das operações aéreas. A notificação quando informada em tempo hábil, contribui no estabelecimento de rotas mais seguras. Para o piloto, o conhecimento das condições meteorológicas nos aeródromos de partida e destino é imprescindível para a realização ou não do seu voo. A direção e a velocidade do vento ao longo da rota, assim como a ocorrência de formação de gelo, turbulência e de tempestades seguidas de trovoadas associadas a nuvens *Cumulunimbus* (CB) podem acarretar desvios de rota e, por conseqüência, maior consumo de combustível. Além disso, o piloto precisa saber a que horas deverão ocorrer essas mudanças e por quanto tempo prevalecerão.

Para o controle de Aproximação de Aeródromo (APP) também é necessária uma previsão da hora de início, duração e extensão da formação de trovoadas associadas a nuvens CB. Assim é possível avaliar o comprometimento operacional em virtude da limitação do espaço aéreo.

Desta forma, a motivação deste estudo está na necessidade de se ter um bom entendimento dos fenômenos causadores de chuvas intensas para que se possa estabelecer uma técnica de previsão de curtíssimo prazo e viabilizar uma potencial aplicação dos resultados obtidos no sentido de proporcionar uma previsão efetiva, visando minimizar os prejuízos causados pelas condições adversas do tempo à aviação.

1.2 Justificativas

O transporte aéreo apresenta grande importância na indústria do turismo, na economia e em diversos setores relevantes para o crescimento do país. A aviação brasileira cresceu muito nos últimos anos com o surgimento de novas companhias aéreas e a modernização da maioria delas. Segundo informações da Empresa Brasileira de Infraestrutura Aeroportuária (INFRAERO), o Aeroporto Internacional do Rio de Janeiro (AIRJ), também chamado de Aeroporto Internacional do Galeão ou Maestro Antonio Carlos Jobim, tem importante contribuição na economia do país, pois além de fazer a ligação de todo o Brasil por vôos domésticos, o AIRJ tornou-se um grande pólo em importações e exportações. Apresentando-se como ponto de ligação de mais de 30 localidades nacionais e 19 destinos internacionais e possuindo dois sistemas de pistas

com operações simultâneas de pousos e decolagens, o aeroporto mostra capacitação e estrutura para ser um local chave tanto no âmbito dos negócios como a serviço da economia. Ainda segundo a INFRAERO, o complexo aeroportuário do AIRJ também conta com a maior pista de pouso e decolagem do Brasil, com 4.000 m x 45 m, assim como um dos maiores, mais modernos e bem equipados terminais de logística de cargas do continente, o que justifica a relevância do aeroporto internacional do Galeão, região alvo desta pesquisa, para o país. A imagem desse aeroporto está exibida na Figura 1.1.



Figura 1.2 – Aeroporto Internacional do Galeão (AMBIENTE BRASIL, 2010).

Portanto, a escolha do tema desta pesquisa foi tomada mediante a constatação da influência dos fatores associados aos eventos extremos de chuva forte com atividades convectivas na segurança em voo, principalmente nos momentos de pouso e decolagem, para o AIRJ, objetivando uma previsão meteorológica qualificada, baseada em um modelo de previsão que seja capaz de identificar os fenômenos meteorológicos que se apresentam como risco potencial a aviação.

Segundo Moura (2007), as previsões são feitas com modelos elaborados em forma de um sistema de equações matemáticas, resolvidas em computadores de alto desempenho, a partir das condições atmosféricas que são fornecidas por operadores de

estações meteorológicas. Entretanto, existe um limite na qualidade das previsões devido à não-linearidade dos processos físicos e limitações nos modelos.

Em face dos tipos de questões a serem enfrentadas, as técnicas de previsão estão sempre a exigir cada vez mais uma tecnologia rápida e com menor número de erro possível – a computacional. Para Moura (1986), embora este avanço tecnológico (computação) tenha sido de grande valia no sentido da evolução das “projeções” meteorológicas de hoje, as previsões de curto prazo (*nowcasting*) e localizadas, envolvendo sistemas de precipitação intensa, apresentam algumas carências. Segundo Holton (1992), essas carências se devem ao fato de que existem movimentos de pequena escala que não podem ser representados pelos modelos atuais.

Ainda que os movimentos de pequeníssima escala fossem representados pelos modelos atuais, segundo Pinto (2002), existe um intervalo de tempo necessário para análise, assimilação, processamento dos dados e execução do processo. Este intervalo forma uma “janela” temporal, denominada *ponto morto*.

Em meio a estes problemas, justifica-se a elaboração desta pesquisa na qual faz-se o uso de uma ferramenta capaz de desenvolver previsões de qualidade observando as variações que ocorrem nos parâmetros não-lineares do tempo. Segundo Haykin (2002), uma Rede Neural Artificial (RNA) produz o mapeamento de qualquer função não-linear com precisão em um curto espaço de tempo. Sendo assim, a RNA se apresenta como uma possível ferramenta para a solução do assunto em questão: a não-linearidade dos processos físicos e a limitação dos modelos atuais para previsões do tipo *nowcasting*.

1.3 Objetivo

O conhecimento prévio das condições atmosféricas e a previsão dos fenômenos do tempo danosos a população é de extrema importância para os mais variados segmentos sociais e econômicos, dentre eles, o transporte aéreo.

Considerando que os modelos de previsão atuais deixam uma “lacuna” a ser preenchida ao não serem capazes de fazer previsões em curtíssimos prazos e de não representar movimentos de pequenas escalas com maior precisão. Nesse contexto, existe o caso do estudo da previsão a curtíssimo prazo de eventos extremos no aeroporto do Galeão, que é ainda um assunto pouco explorado.

Configura-se, mediante as questões apresentadas acima, a pergunta que esta pesquisa pretende responder: existe a potencialidade de prever eventos extremos de precipitação associados a convecções profundas em curtíssimo prazo e com precisão de intensidade, início e duração para o aeroporto internacional do Galeão, preenchendo o vazio deixado pelos modelos atuais?

Em atenção a essas considerações, este estudo tem como proposta investigar, através do uso conjunto de dados provenientes do ar-superior, observações de superfície, dados de precipitação e descargas atmosféricas, associações entre variáveis que identifiquem características importantes para ocorrência de eventos convectivos extremos; e assim formular uma previsão a curtíssimo prazo desses eventos utilizando redes de processamentos capazes de solucionar problemas simulando o comportamento e as funções de um cérebro humano: as Redes Neurais Artificiais (RNA), focando o seu objetivo na previsão imediata (*nowcasting*) dos sistemas de precipitação que possam trazer prejuízos a aviação no âmbito do Aeroporto Internacional do Galeão.

Capítulo 2 – FUNDAMENTAÇÃO TEÓRICA

É visto que avaliar isoladamente os fatores contribuintes na configuração de um evento extremo e identificar os padrões atmosféricos associados com o tempo severo são processos fundamentais no planejamento estratégico aeronáutico. Segundo Wells (2004), a análise correta da condição meteorológica de um aeródromo e o desenvolvimento de práticas seguras no programa de prevenção de acidentes, aumentam significativamente a segurança durante o voo. Mas, para garantir a proteção ao voo, é necessário compreender os fundamentos que regem os eventos relacionados ao tempo severo no local escolhido. Katz e Brown (1992), usando teorias estatísticas, demonstraram que tais eventos dependem das alterações climáticas de cada lugar.

No território brasileiro, a diversidade climática é muito grande devido a diversos fatores naturais, dos quais destacam-se: a fisionomia geográfica, a extensão territorial, o relevo e a dinâmica das massas de ar (LEAL *et al.*, 2010). Este último fator é de suma importância porque atua diretamente tanto na temperatura quanto na pluviosidade, provocando as diferenciações climáticas regionais (MICROSOFT CORPORATION, 1993-1999). De acordo com a classificação climática de Koppen-Geiger, as massas de ar que interferem mais diretamente são a equatorial (continental e atlântica), a tropical (continental e atlântica) e a polar atlântica. Segundo Vitte (2004), no estado do Rio de Janeiro predominam o clima tropical (baixadas) e tropical de altitude (planalto). Entretanto, na região metropolitana do Rio de Janeiro, domina o clima tropical semi-úmido, caracterizado por invernos secos com temperaturas amenas e chuvas abundantes no verão, normalmente, quando costumam ocorrer fenômenos de precipitações intensas.

Para Diniz *et al.* (2006), na avaliação dos mecanismos físicos associados com atividade convectiva, existe a necessidade de analisar o relacionamento entre processos dinâmicos e termodinâmicos na formação e organização dos sistemas convectivos. Outro estudo realizado por Diniz (2008), avaliou a influência relativa da associação entre variáveis atmosféricas na evolução de sistemas convectivos e ocorrência de chuvas extremas constatando que estes eventos estão altamente correlacionados com o teor de umidade do ar, a instabilidade atmosférica e a convergência em baixos níveis. Segundo Reboita (2010), dois ou mais desses fatores atuando em conjunto, podem provocar uma convecção profunda responsável por taxas elevadas de precipitação.

Sendo assim, é importante compreender que para que ocorra a formação de convecção profunda, produtora de chuvas intensas, em um determinado local, é necessário o acontecimento simultâneo de dois fatores determinantes neste local:

1) Condição termodinâmica favorável, normalmente caracterizada pela ocorrência de instabilidade termodinâmica na atmosfera (Menezes, 1998);

2) Efeito forçante ou “gatilho” para o disparo da convecção, representado por algum mecanismo que produza levantamento deste ar instável, de forma que a convecção seja iniciada (Lima, 2005).

2.1 Conceitos Basilares

Neste segmento serão introduzidos alguns conceitos fundamentais para o entendimento dos processos levantados durante esta pesquisa. São eles:

Névoa (H)

Deve ser usada a abreviatura H (*haze*) para indicar névoa seca quando a obstrução à visibilidade consistir predominantemente de litometeoros e a visibilidade predominante tenha se reduzido a 5.000 metros ou menos, em consequência do fenômeno informado. Deve ser usado H para indicar névoa úmida quando a obstrução à visibilidade consistir de gotículas d'água ou cristais de gelo e a visibilidade predominante deve ser de 1.000 metros ou mais, até o limite de 5.000 metros, inclusive (Manual do Comando da Aeronáutica – MCA 105-12/2010).

Nevoeiro (F)

Existe nevoeiro quando a obstrução à visibilidade consistir de gotículas d'água ou cristais de gelo (nevoeiro ou nevoeiro congelante). Para a abreviatura F (*fog*) ser informada, a visibilidade predominante deve ser menor que 1.000 metros (MCA 105-12/2010).

Fumaça (K)

Registra-se a ocorrência de fumaça no aeródromo quando a obstrução à visibilidade consistir predominantemente de litometeoros e a visibilidade predominante tenha se reduzido a 5.000 metros ou menos, em consequência da suspensão no ar de pequenas partículas produzidas por combustão (MCA 105-12/2010; MCA105-2/2010).

Chuvisco (L)

A intensidade do chuvisco é definida pelo acúmulo no pluviógrafo (ou pluviômetro) ou pela visibilidade, conforme exposto na Tabela 2.1.

Tabela 2.1 – Classificação de intensidade do chuvisco (MCA 105-12/2010).

Intensidade	Acúmulo	Visibilidade
Leve	De traços até 0,09 mm/h	Maior que 1.000 m
Moderada	De 0,1 a 0,4 mm/h	Entre 500 e 1.000 m
Forte	Acima de 0,4 mm/h	Menor que 500 m

Chuva (R)

A intensidade da chuva é definida apenas pelo acúmulo no pluviógrafo (ou pluviômetro), conforme Tabela 2.2.

Tabela 2.2 – Classificação de intensidade da chuva (MCA 105-12/2010).

Intensidade	Acúmulo
Leve	De traços até 2,4 mm/h
Moderada	De 2,5 a 9,9 mm/h
Forte	Acima de 9,9 mm/h

Trovoada (T)

O qualificador T (*tunderstorm*) é utilizado para informar ocorrência de trovoada, sempre que forem ouvidos trovões ou detectados raios e relâmpagos no aeródromo, no período de 10 minutos que precede a hora da observação. Quando for o caso, T deverá ser seguido imediatamente, por abreviaturas para indicar qualquer precipitação observada. Quando não for observada qualquer precipitação associada, deverá ser usado somente T (MCA 105-12/2010).

Pancada (W)

A abreviatura W (pancada) foi usada quando houvesse precipitação, com intensidade forte, proveniente de nuvens convectivas. A pancada se caracteriza por

começo e final bruscos, e, em geral, por mudanças de intensidade fortes e rápidas (MCA 105-10/2010).

A Tabela 2.3 apresenta os indicadores de intensidade e descritores do fenômeno de tempo (qualificadores) juntamente com as abreviaturas dos fenômenos de tempo encontrados no período de dados analisados.

Tabela 2.3 – Tabela de classificadores e fenômenos de tempo constantes nos dados no período de 01/01/1997 a 31/12/2009 (MCA 105-10/2010).

QUALIFICADOR		FENÔMENO DE TEMPO	
Intensidade	Descritor	Precipitador	Obscurecedor
- (LEVE)	W (PANCADA)	L (CHUVISCO)	H (NÉVOA)
sem sinal (MODERADA)	T (TROVOADA)	R (CHUVA)	K (FUMAÇA)
+ (FORTE)	—	—	F (NEVOEIRO)

2.2 Processos Termodinâmicos

Grandes centros urbanos, como nas proximidades do AIRJ, implicam em um armazenamento de calor. Segundo Sutton (1977), durante o dia prevalece o aquecimento da superfície devido a radiação. No fim da tarde e durante a noite prevalece a emissão da radiação termal da superfície para a alta atmosfera, permitindo a variação vertical da temperatura. O que favorece o movimento convectivo e aumenta a instabilidade do ar (SOLOMON, 2007).

Dentro desses processos termodinâmicos, existem estimativas convenientes que podem servir como um bom indicativo para a ocorrência de instabilidade na atmosfera. Em geral, são expressões simples e de fácil cálculo, associadas a diferenças entre variáveis meteorológicas em níveis arbitrários, chamadas **índices de instabilidade**.

2.2.1 Índices de Instabilidade

Segundo Costa *et al.* (2006), uma das maiores dificuldades quanto a análise de resultados de modelos numéricos de previsão de tempo está no correto posicionamento

dos núcleos e na quantificação dos totais de precipitação. Não é nada incomum que estes, devido a algumas limitações, também atrasem ou adiantem a chuva ocorrida em uma região. Cabe ao previsor analisar outras características da atmosfera, para então confirmar ou não a possibilidade de chuvas intensas. Algumas dessas características podem ser identificadas pelo uso de índices de instabilidade, que normalmente tomam como base, as diferenças em níveis de pressão (ou altura), temperatura do ponto de orvalho, temperatura potencial, temperatura potencial específica, entre outros. Para Lima (2005), esses índices combinados entre si ou mesmo individualmente com outras variáveis, são de grande utilidade para análise de previsão de tempo severo. Portanto, o uso dos índices de instabilidade, vem se tornando uma ferramenta de grande utilidade no auxílio da previsão de eventos de tempestades (LIMA, 2005; NASCIMENTO, 2005).

Escobar (2007) denomina índice de instabilidade como o valor numérico que permite expressar a instabilidade da atmosfera. Segundo Escobar, a maior parte dos índices de instabilidade, também chamados de “indicadores do tempo severo” foram desenvolvidos para latitudes médias. Porém, existem trabalhos que foram propagados no sentido de ajustar estes índices com testes e avaliações da aplicabilidade real para as regiões Sul e Sudeste do país (BENETI e SILVA DIAS, 1986; FOGACCIA e PEREIRA FILHO, 2002; NASCIMENTO e CALVETTI, 2004), fornecendo parâmetros de estudo meteorológicos consistentes e permitindo previsões climáticas confiáveis, sobretudo para a aviação.

Para Dias (2008), os índices de instabilidade mais utilizados em centros de previsão de tempo, que auxiliam os previsores no tocante a formação de tempestades, são: **índice K**, **índice dos totais** (TT – *total totals*), **índice Showalter** (IS), além da **energia potencial convectiva disponível** (CAPE – *Convective Available Potential Energy*), também chamada de índice CAPE. Mas também existe outro índice utilizado no Brasil - o **índice CK**, sugerido por Lima (2005) para a identificação de áreas propícias a ocorrência de eventos extremos em resultados de modelos numéricos de previsão de tempo. No estudo feito por Lima, foi mostrado que o índice CK tem aplicabilidade para regiões tropicais.

Existe também um índice eficaz na representação adequada de um sistema de instabilidade atmosférica, chamado **índice Lapse Rate** (LR). De acordo com Menezes (1997), o índice LR avaliado junto ao índice TT se mostrou eficiente no ponto de vista

de “marcar” um ciclone em seus campos simulados. A seguir, é apresentada uma breve explicação sobre os índices de instabilidade que foram citados.

Índice K

Costa *et al.* (2006) define índice K pela soma das temperaturas de bulbo seco (T) e ponto de orvalho (T_d) em 850 hPa, subtraída da depressão do ponto de orvalho em 700 hPa e da temperatura de bulbo seco em 500 hPa, vide a Equação 2.1.

$$K = (T_{850} - T_{500}) + Td_{850} - (T_{700} - Td_{700}), \quad (2.1)$$

onde T indica a temperatura em graus Celsius, Td é a temperatura do ponto de orvalho em graus Celsius e os sub-índices indicam o nível de pressão (em hPa) correspondente.

Para Nascimento (2005), o índice K tende a melhor captar condições favoráveis à ocorrência de tempestades em ambientes úmidos. Essa condição pode ser explicada conforme Silva Dias (2006), ao mostrar que para o Índice K ser significativo deve haver calor em baixos níveis e umidade em abundância desde a superfície até 700 hPa. Segundo Silva Dias (2006) uma camada seca em 700 hPa reduz o valor de K. Desta maneira o K não é um bom índice para avaliar atmosferas favoráveis a formação de tempestades severas com ventos fortes, já que estas são associadas com a presença de uma camada de ar seco em níveis médios. Assim, K é um índice mais adequado para determinar apenas condições de chuvas fortes e não ventanias.

Ainda conforme Silva Dias, quanto mais positivo for o índice K, maior será a probabilidade de tempestades. Além disso, os valores do índice K variam conforme a estação do ano e localização. Valores de K acima de 30 são indicadores de alta probabilidade de ocorrência de convecção profunda. A Tabela 2.4 indica os valores de referências do índice K de acordo com a probabilidade de tormentas.

Tabela 2.4 – Probabilidade de formação de tormentas com o Índice K (DIAS, 2008).

VALOR DE K	PROBABILIDADE DE TORMENTAS
$K \leq 20$	Nenhuma
$20 < K \leq 25$	Tormentas isoladas
$25 < K \leq 30$	Tormentas amplamente dispersas
$30 < K \leq 35$	Tormentas dispersas
$35 < K \leq 40$	Muitas tormentas

Índice CK

Lima (2005) sugeriu o uso da análise combinada do índice K e da convergência em baixos níveis, que representam, respectivamente, condição termodinâmica da atmosfera forçante para o levantamento de ar e disparo de convecção. Segundo Lima, essa combinação seria uma ferramenta relevante em termos de caracterizar possíveis regiões de formação dos sistemas convectivos intensos.

Dessa forma, Lima (2005) definiu o índice CK como sendo o produto entre o índice K e a divergência em 925 hPa, conforme mostra a Equação (2.2):

$$CK = K \cdot \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right), \quad (2.2)$$

onde u é a componente zonal do vento, v é a componente meridional e K é o índice K.

O índice CK é válido para valores positivos de K , sendo assim, Lima (2005) menciona que CK denota situações favoráveis ao disparo de células convectivas, quando este apresenta valores negativos, pois K é um número positivo e a “convergência” (que é a forçante) é representada pelos valores negativos da divergência. Quanto menores (mais negativos) forem os valores, mais favorável será a situação para o disparo de células convectivas.

Índice CAPE

Uma tempestade é sustentada pelo empuxo proveniente das correntes ascendentes, o qual é responsável pelo fornecimento da energia necessária para que a parcela de ar continue seu movimento de ascensão, ou seja, a **energia potencial convectiva disponível** - CAPE.

A CAPE é calculada através da integral do empuxo que a corrente sofre à medida que sobe da base da nuvem até uma determinada altura na atmosfera, onde o empuxo é neutro. Vide a Equação 2.3.

$$CAPE = \int_{NCC}^{NEN} g \left(\frac{T_P - T_A}{T_A} \right) dz, \quad (2.3)$$

onde, NEN é o Nível de Empuxo Neutro (próximo à tropopausa), NCC, o Nível de Condensação Convectiva, T_p é a temperatura da parcela, T_a é a temperatura do ambiente e g é a aceleração da gravidade.

No geral, o índice CAPE fornece uma medida da máxima energia cinética possível que uma parcela estaticamente instável pode adquirir, assumindo que a parcela ascende se mistura com o ambiente e se ajusta instantaneamente à pressão do ambiente. Quanto maior é a CAPE, maior é a força das correntes ascendentes da tempestade, portanto maior será o grau de instabilidade da atmosfera. Os valores de referência para a condição de formação de tormentas em relação a CAPE são apresentados na Tabela 2.5.

Tabela 2.5 – Valores de referência para a probabilidade de formação de tormentas com o índice CAPE (DIAS, 2008).

CAPE	POTENCIAL CONVECTIVO
0 a 1000	Marginalmente instável
1000 a 2500	Moderadamente instável
2500 a 4000	Acentuadamente instável
> 4000	Extremamente instável

Índice *Total Totals* (TT)

O índice *total totals* é utilizado para identificar potenciais áreas com desenvolvimento de tormentas convectivas. Este índice foi definido por Miller e Maddox (1975), conforme apresentado na Equação 2.4.

$$TT = T_{850} + Td_{850} - 2 \cdot T_{500}, \quad (2.4)$$

onde T e Td referem-se às temperaturas do bulbo seco e ponto de orvalho em graus Celsius, respectivamente, e os sub-índices indicam o nível de pressão (em hPa) correspondente.

O índice TT denota interpretações semelhantes ao índice K quando a massa de ar é quente e úmida, mas quando o ar é mais frio e seco o *total totals* tem melhores resultados por dar um peso maior à presença de ar frio em níveis médios (BENETI e SILVA DIAS, 1986). Como o índice TT exige calor e umidade em 850 hPa e o índice K depende de uma atmosfera seca em 700 hPa, quando K é baixo e TT é alto, indica propensão à tempestades severas, pois haverá umidade em abundância em baixos níveis

(até 700 hPa), e ventania em superfície, já que K depende do grau de secamento da atmosfera em 700 hPa (NASCIMENTO, 2010). Além disso, o índice TT não restringe a umidade em 700 hPa, conforme o índice K, portanto se houver uma camada seca em 700 hPa não afetará o índice TT (SILVA DIAS, 2000).

Valores de TT acima de 40°C indicam situações favoráveis a tempestades, e acima de 50°C indicam chances de tempestades severas (DIAS, 2008). Esses valores, para convecção, são apresentados na Tabela 2.6.

Tabela 2.6 – Probabilidade de formação de tormentas com o Índice TT. (fonte: Adaptado de Dias, 2008)

TT	ATIVIDADE CONVECTIVA
< 44	Potencial fraco para tormentas
44 a 49	Tormentas moderadas
50 a 55	Tormentas fortes
> 55	Tormentas muito fortes

Índice Lapse Rate (LR)

O índice LR representa a variação da temperatura de acordo com a altura de uma parcela de ar (dT/dz) para uma determinada camada atmosférica (níveis médios, entre 700 hPa e 500 hPa), de modo que quanto maior for a LR para uma dada camada, maior será a instabilidade atmosférica para convecção profunda (MOURÃO *et al.*, 2006). Ou seja, é a taxa de queda da temperatura ambiental com a altura, representada pela Equação 2.5.

$$LR = -\frac{dT}{dz} = -\frac{1000 \cdot (T_{500} - T_{700})}{\text{geo}_{500} - \text{geo}_{700}}, \quad (2.5)$$

onde **T** indica a temperatura em graus Celsius, **geo** indica a *altura geopotencial*¹ e os sub-índices indicam o nível de pressão (em hPa) correspondente.

¹ Representa a altitude acima do nível do mar em que está um determinado nível de pressão, embora não seja exatamente a altitude verdadeira medida em metros. Esta grandeza é útil porque um dado nível de pressão não é paralelo a superfície terrestre, e estas variações de altitude do nível de pressão indicam regiões de alta e baixa pressão no geopotencial escolhido.

Segundo Nascimento (2005), valores de LR menores que $6^{\circ}\text{C km}^{-1}$ indicam condição estável; entre $6^{\circ}\text{C km}^{-1}$ e $9.5^{\circ}\text{C km}^{-1}$ o ambiente é condicionalmente instável, tornando-se absolutamente instável (situações de instabilidade extrema) a partir de $9.5^{\circ}\text{C km}^{-1}$.

O índice showalter (IS)

O índice *showalter* é calculado através de diagramas termodinâmicos e avalia apenas a instabilidade da camada, apresentando valores negativos quando há umidade e calor no nível de 850 hPa e ar frio em 500 hPa (DIAS, 2008). Quanto mais negativo este índice, maior será o empuxo e, portanto, haverá maior chance de ocorrência de tempestades. Os valores de referência para probabilidade de formação de tormentas em relação ao índice *Showalter* (IS) estão apresentados na Tabela 2.7.

Tabela 2.7 – Probabilidade de formação de tormentas com o Índice Showalter (fonte: Adaptado de Dias, 2008).

IS	PROBABILIDADE DE TORMENTAS
3 a 1	Possíveis tormentas (necessita de forte gatilho)
0 a -3	Atmosfera Instável
-4 a -6	Atmosfera Muito instável
< -6	Atmosfera Extremamente instável

2.3 Processos Dinâmicos

Na literatura, alguns autores destacaram a influência dos processos dinâmicos no desenvolvimento de convecção profunda. Araújo *et al.* (1988) estudaram os aspectos dinâmicos associados às chuvas intensas de fevereiro na região sudeste e constataram uma situação típica de bloqueio de frente fria no oceano Atlântico próximo ao litoral sudeste do Brasil, acompanhada de acentuada atividade convectiva tropical no interior do continente.

Para Diniz *et al.* (2006), o aquecimento radiativo tem um papel fundamental no processo de formação das tempestades, pois à proporção que as áreas na superfície terrestre são aquecidas a diferentes taxas, o fluxo de energia cria zonas de convergência, movimentos ascendentes e condições favoráveis à formação e desenvolvimento das nuvens.

Dereczynski, *et al.* (2008), analisaram os eventos de chuvas intensas no município do Rio de Janeiro concluindo que 77% dos 160 casos selecionados, foram provocados por sistemas frontais, que ocorrem durante todo o ano, com menor frequência no inverno.

Segundo Severo *et al.* (1994), chuvas intensas associadas à atividade convectiva é o resultado de mecanismos físicos de diferentes escalas. Os processos de grande escala contribuem para desestabilizar o ambiente local e os de escala menor fornecem as condições para o levantamento inicial, como mecanismos de levantamento do ar potencialmente instável próximo à superfície. Dentre os processos dinâmicos denominados “forçantes” que podem gerar eventos dessa natureza, estão:

- Frente fria, definida como a borda dianteira de uma massa de ar frio, em movimento ou estacionária;
- Linha de instabilidade pré-frontal, que segundo Scolar, *et al.* (1982) são perturbações, frequentemente, mais intensas que a própria frente;
- Frente quente, ou seja, a parte dianteira de uma massa de ar quente em movimento;
- Pseudo frente ou frente de rajada, formada pela ação conjunta das correntes descendentes de um aglomerado de tempestades ou complexo convectivo de mesoescala;
- Intersecção de duas linhas de atividade, como uma linha de instabilidade e uma frente quente.

As forçantes, por sua vez, atuam de modo a empurrar o ar para cima, fornecendo energia para que ele possa superar a região estável criando zonas de convergência através dos fluxos em baixos níveis. É ao longo dessas zonas que as tempestades tendem a se desenvolver criando sistemas de convecção profunda que pode contribuir de várias formas para uma condição de tempo severa.

2.3.1 Fluxos (jatos) de baixos níveis

Dentre as diversas definições existentes para Jatos de Baixos Níveis (JBN), a utilizada neste contexto será aquela proposta por Stull (1995), modificada em função do que foi apresentado por Mahrt (1999) e Andrade (2003). Considerando JBN como uma estreita corrente de ar situada a uma altura inferior a 600 m, que representa um máximo

relativo local no perfil vertical de velocidade de tal forma a apresentar, velocidade de, no mínimo, 2 m/s a mais que o escoamento situado acima e abaixo dele. JBNs podem ter uma extensão horizontal de várias centenas de quilômetros e velocidades de vento que podem exceder 10 m/s (OLIVEIRA e FITZJARRALD, 1993).

Os Jatos de Baixos Níveis já foram discutidos em diversos trabalhos como um processo influente para o disparo de convecção profunda (ocasionalmente severa). Na América do Sul os JBNs ocorrem a leste da Cordilheira dos Andes, desde latitudes equatoriais até latitudes subtropicais (BERRI e INZUNZA, 1993). Marengo, *et al.*, (2004), analisaram a climatologia dos JBNs da América do Sul utilizando-se de dados de reanálise do NCEP/NCAR 1950-2000 e perceberam que o JBN está relacionado ao transporte de umidade desde a região Amazônica até as terras férteis do sul do Brasil e norte da Argentina, provendo enorme quantidade de ar úmido potencialmente instável, podendo ocorrer intensas tempestades nessa região e seus arredores.

Apesar dos JBNs terem sido descritos primeiramente na década de 30 na África, foi a partir da década de 50 que os estudos sobre JBNs se intensificaram. Uma das razões desse interesse na década de 50 até os dias de hoje é proveniente do fato de que o JBN tem se mostrado relacionado com profundas atividades convectivas. O que mostra ser de grande importância considerá-los nos estudos do clima global e local.

Esses jatos são conhecidos por ocorrerem sobre a América do Norte (Bonner 1968; Douglas 1993 *apud* Stensrud (1996)), América do Sul (Virji 1981, 1982 Nogués-Paegle e Mo, 1997; Douglas *et al.*, 1998 e 1999; Dias, 2000; Marengo *et al.*, 2000; Misra *et al.*, 2000; Nicolini e Saulo, 2000; Saulo *et al.*, 2000, Marengo e Soares, 2002, Marengo *et al.*, 2004 *apud* Stensrud (1996)), África (Findlater 1969; Ardanuy 1979; Kelbe 1988; Jury and Spencer-Smith 1988; Jury e Tosen 1989 *apud* Stensrud (1996)), Austrália (Wilson 1975; Brook 1985; Keenan 1989 *apud* Stensrud (1996)), Ásia (Findlater 1969; Tão ad Chen 1987 *apud* Stensrud (1996)), Antártica (Schwerdt-feger 1975; Chiba e Kobayashi 1986 *apud* Stensrud (1996)) e Europa (Sladkovic e Kanter 1977; Kraus 1985 *apud* Stensrud (1996)). Essas regiões de ocorrências frequentes de JBNs estão localizadas quer ao leste de uma grande gama de montanha ou onde existem gradientes de temperatura terra-mar.

Para Stensrud (1996), somente os jatos que possuem apreciável cisalhamento vertical do vento horizontal devem ser chamados de JBN. Segundo Menezes (1997),

simulações numéricas de tempestades mostraram que em situações em que há cisalhamento unidirecional (o vetor cisalhamento não gira com a altura) a tendência é de desenvolvimento de multicélulas, sem região preferencial em relação à célula convectiva inicial.

JBNs são mais freqüentes durante os meses de verão e tipicamente formados durante a noite, atingindo o seu auge horas antes do amanhecer. Mas também existem casos no inverno, como observado por Sias e Silva Dias (2002), ao analisarem um caso de precipitação extrema em julho de 1989 na cidade de São Paulo, com uma precipitação acumulada diária superior a 50 mm acompanhada a esses jatos.

A vantagem de se estudar este fenômeno é devido à possibilidade de se questionar e explicar os mecanismos físicos responsáveis pelo desenvolvimento de eventos extremos de precipitação já que, Berri e Inzunza (1993) destacam o papel do JBN no transporte de umidade da região tropical em direção aos pólos passando por regiões como o município do Rio de Janeiro, durante esse transporte.

2.4 Previsão meteorológica

É claro que a Meteorologia como ciência e atividade operacional faz parte de uma atividade colaborativa em diversos setores, posto que é importante monitorar o desenvolvimento dos fenômenos meteorológicos a cada instante para poder prever a sua trajetória, sua intensidade e impactos que podem trazer.

A idéia de previsão através de processos numéricos de resolução de equações que representam o comportamento da atmosfera foi publicada pela primeira vez por Lewis Richardson, um matemático britânico, em 1922. O pioneirismo de Richardson levou a primeira previsão de tempo bem sucedida, realizada no Instituto de Estudos Avançados da Universidade de Princeton, em 1950, utilizando-se o computador ENIAC (*Electrical Numerical Integrator and Computer*), uma máquina rudimentar que funcionava a válvulas e com pequena memória. Segundo Moura (2007), de 1950 até os dias atuais, a metodologia experimental denominada “subjéitiva” foi superada pelos métodos físicos e numéricos na previsão do tempo denominada “objéitiva”.

Hoje, as previsões são feitas através de modelos codificados em computadores. Esses modelos codificam as leis da física e da dinâmica atmosférica e oceânica, tais como

a lei da conservação da massa, conservação de energia, lei de Newton e a dos movimentos, convertendo-os em dados para análise e previsão. Para Almeida (2009), tais modelos meteorológicos são bastante eficientes para investigar mudanças meteorológicas a médio e longo prazo, porém, os fenômenos de escalas inferiores aos que podem ser mensurados pelos modelos de observações são tratados de forma “parametrizada”, o que induz a erros. Segundo Holton (1992), não importa quão fina seja a resolução de grade, haverá sempre movimentos em escalas ainda menores para serem devidamente representados no modelo numérico.

Outro grande problema na previsão do tempo hoje, como destaca Pinto (2002), é saber com precisão onde ocorrerão os eventos previstos pelos modelos atuais. Para isso é necessário um meio de previsão que leve em consideração detalhes do tempo não somente no local de estudo, mas também nas vizinhanças.

2.4.1 Previsões meteorológicas e aviação

No Brasil, o estudo e as aplicações da meteorologia aeronáutica são realizados por meteorologistas efetivados ao Comando da Aeronáutica e Infraero. O trabalho desses meteorologistas consiste em fazer observações visuais das condições do tempo nos aeroportos, colher dados meteorológicos de estações e radares meteorológicos, divulgar esses dados aos usuários e interpretar os dados produzindo previsões meteorológicas específicas para a região do aeródromo e rotas aéreas. Por se tratar de uma análise local, as informações dadas por esses meteorologistas são mais detalhadas do que outras previsões obtidas dos modelos meteorológicos.

De acordo com o Manual do Comando da Aeronáutica 105-2 (2010), durante a observação são aferidas as condições meteorológicas no aeródromo e seus arredores, através da verificação visual e da leitura de determinados instrumentos como: termômetro, anemômetro, psicômetro entre outros. Essas observações são horárias em Tempo Universal Coordenado (UTC). Elas são inseridas em um sistema que pode ser verificado pelos usuários por meio do código meteorológico METAR (*Meteorological Aerodrome Report*) que proporciona aos pilotos informações como: velocidade e direção do vento, cobertura celeste do aeródromo, visibilidade horizontal, altura das nuvens; tempo presente, temperatura do ar, ponto de orvalho (temperatura em que a umidade do ar se condensa) e pressão atmosférica.

É importante ressaltar que, o código METAR representa uma medida “pontual”, representando as condições meteorológicas em um exato momento. Além disso, existe a avaliação subjetiva em algumas informações contidas no código, já que algumas variáveis são estimadas pelos próprios operadores de estações meteorológicas de superfície (EMS). Dentre as informações que estão mais sujeitas a erros estão a condição do tempo presente e a inspeção visual, que, embora subjetiva, é bastante utilizada pelos previsores. Segundo o MCA 105-2, a visibilidade observada pelo operador da EMS é registrada conforme o posicionamento do observador em relação aos pontos de referência que são detalhados em duas cartas, uma contendo os pontos de referência situados a menos de 2 km e outra contendo os pontos de referência situados a mais de 2 km até um limite de 20 km, conforme pode ser observado no Anexo A.

Apesar da subjetividade, as informações de uma previsão aeronáutica são as mais detalhadas dentre todos os tipos de previsões meteorológicas, pois é fundamental na aviação, tanto civil (relatório anual de segurança operacional, 2009) como militar, que se saiba com a melhor precisão possível quais serão as condições meteorológicas que o piloto irá enfrentar durante o voo e no momento do pouso quando chegar ao destino, sendo fator decisivo na escolha da rota e melhor hora para decolagem.

Para confecção dos códigos de previsões como o *Terminal Aerodrome Forecast* (TAF), *Significant Meteorological Information* (SIGMET), *General Aviation Meteorological Information* (GAMET) e *Airmen’s Meteorological Information* (AIRMET), entre outros, são utilizados, não somente dados meteorológicos de superfície, como o METAR, mas também registrados por meio da radiossondagem através de um pequeno transmissor (radiossonda) acoplado à um balão inflado com gás hidrogênio. A radiossondagem é a principal fonte de dados do ar superior utilizada em pesquisas (servindo como base de dados para a previsão numérica do tempo) e em serviços operacionais (tais como a previsão de vento e temperatura nos níveis de vôo, turbulência, formação de gelo em aeronaves, cálculo da probabilidade de trovoadas, formação de nuvens, trilhas de condensação e, mais recentemente, nas avaliações do movimento e da dispersão de nuvens de cinzas vulcânicas e nuvens radioativas).

De acordo com o Manual do Comando da Aeronáutica 105-9 (2011), os setores responsáveis pela coleta e tratamento dos dados meteorológicos obtidos pela radiossondagem são as estações meteorológicas de altitude (EMA). A radiossondagem é

feita de acordo com a hora oficial de observação do ar superior, que é estabelecida pelo serviço meteorológico regional. Esta hora deve ser a mais próxima possível da hora padrão ou horas sinóticas principais (06:00 UTC, 12:00 UTC, 18:00 UTC e 00:00 UTC). Entretanto, quando somente duas observações forem realizadas, elas deverão ser feitas às 12:00 e 00:00 UTC, que é o caso da estação meteorológica de altitude do Galeão.

O estudo do ar em altitude tem sua aplicação primordial na aviação, ao prover os centros meteorológicos com informações dos diferentes níveis da atmosfera para fins de plotagem e análise das cartas de altitude, para posterior previsão do tempo e fornecer informações meteorológicas em altitude para a navegação aérea, com divulgação internacional, além de obter dados unitários básicos, para fins de estatísticas meteorológicas e climatológicas que poderão fornecer um prognóstico mais detalhado sobre as condições do tempo no aeródromo. As observações do ar superior, conjugadas com as observações de superfície fornecem uma imagem tridimensional da estrutura da atmosfera.

Dentre os códigos de previsão confeccionados pelos previsores subordinados ao Comando da Aeronáutica, o mais utilizado pelos aeronavegantes é o código TAF que contém informações sobre o grupo de identificação de localidade, dia e hora de confecção e período de validade; vento à superfície previsto; visibilidade predominante prevista; tempo significativo previsto; nuvens previstas (ou visibilidade vertical); temperaturas previstas e mudanças significativas esperadas.

A fim de atender o planejamento operacional dos vôos para aeródromos nacionais e internacionais, o TAF é renovado a cada seis horas com períodos de validade que deverão iniciar-se às 0000, 0600, 1200 e 1800 UTC, tendo duração de 12, 24 ou 30 horas, deixando uma “janela” de informações que são confiadas no que foi registrado no último TAF. Com isso, verifica-se a urgência de fornecer previsões detalhadas com um menor intervalo de renovação, possibilitando a eliminação de “janelas” na previsão do tempo. Para tanto, é necessária a utilização de uma ferramenta que seja capaz de produzir resultados de previsão em tempo “quase real” que possa proporcionar uma maior operacionalidade na previsão voltada para a atividade aérea no aeroporto em questão (AIRJ).

2.5 Redes Neurais Artificiais

2.5.1 Introdução a Redes Neurais Artificiais

As primeiras informações sobre neurocomputação surgiram em 1943, em artigos do neurofisiologista Warren McCulloch, do Instituto Tecnológico de Massachusetts, e do matemático Walter Pitts, da Universidade de Illinois. Os autores fizeram uma analogia entre células nervosas vivas e o processo eletrônico, em um trabalho publicado sobre “neurônios formais” que simulavam o comportamento do neurônio natural, por isso, recebeu o nome de **Redes Neurais Artificiais**.

Segundo McCulloch e Pitts (1943), o elemento constitutivo de base das redes neurais artificiais é o **neurônio artificial**, que tem como modelo o neurônio biológico. O neurônio biológico é uma célula que pode ser dividida em três seções: o corpo da célula, os dendritos e o axônio; cada uma com funções específicas, porém complementares. A Figura 2.1 ilustra as componentes principais do neurônio biológico.

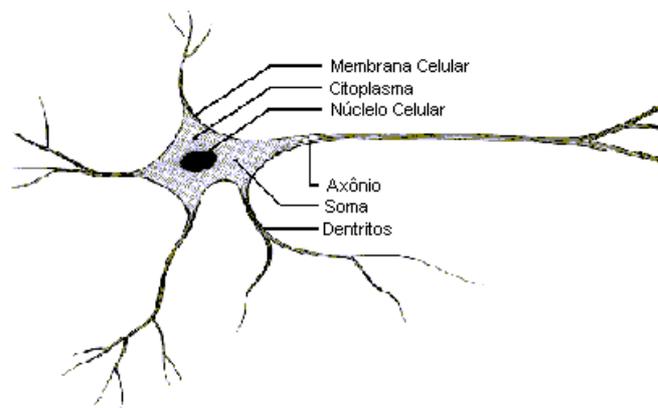


Figura 2.1 – Componentes principais de um neurônio biológico (fonte: Braga, *et al.*, 1998).

Segundo Kovacs (2002), os dendritos têm por função receber as informações, ou impulsos nervosos, oriundos de outros neurônios e conduzi-los até o corpo celular, onde a informação é processada, e novos impulsos são gerados. Esses impulsos são transmitidos a outros neurônios, passando através do axônio até os dendritos dos neurônios seguintes. O ponto de contato entre a terminação axiônica de um neurônio e o dendrito de outro é chamado de sinapse. São pelas sinapses que os nodos se unem funcionalmente, formando redes neurais.

Os sinais oriundos dos nodos pré-sinápticos são passados para o corpo do neurônio, onde são comparados com outros sinais recebidos pelo neurônio. Se o percentual em um intervalo curto de tempo é suficientemente alto, a célula “dispara”, produzindo um impulso que é transmitido para as células seguintes (nodos pós-sinápticos). Este sistema simples é responsável pela maioria das funções realizadas pelo nosso cérebro. Segundo Braga *et al.* (1998), a capacidade de solucionar funções complexas surge com a operação em paralelo de todos os 10^{11} nodos do nosso cérebro.

O funcionamento do primeiro neurônio artificial, proposto por McCulloch e Pitts (MCP) (1943) é uma simulação do que se conhecia a respeito do neurônio biológico naquela época. Assim, o neurônio artificial tem terminais de entrada (x_0, x_1, \dots, x_n), que representam os dendritos do neurônio biológico, e um único terminal de saída y , que representa o axônio do neurônio biológico. Para limitar por emulação o comportamento das sinapses, os terminais de entrada do neurônio têm *pesos* acoplados (w_0, w_1, \dots, w_n), cujo valores podem ser positivos ou negativos. O efeito de uma sinapse particular i no neurônio pós-sináptico é dado por: $x_i w_i$. Os pesos determinam “em que grau” o neurônio deve considerar sinais de disparo que ocorrem naquela conexão. Uma ilustração do neurônio artificial proposto por McCulloch e Pitts é apresentada na Figura 2.2.

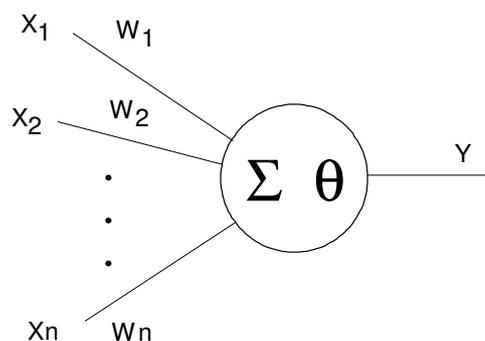


Figura 2.2 – Neurônio de McCulloch e Pitts (fonte: Braga *et al.*, 1998).

O corpo do neurônio, por sua vez, é emulado por um mecanismo simples que faz a soma dos valores $x_i w_i$ recebidos pelo neurônio (soma ponderada). Um neurônio dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação chamado de *threshold*. Portanto, o neurônio de McCulloch e Pitts é ativado por meio de uma **função de ativação**, que ativa a saída ou não, dependendo da soma ponderada das suas

entradas. A Equação 2.6 descreve a função de ativação de um neurônio de McCulloch e Pitts.

$$\sum_{i=1}^n x_i w_i \geq \theta, \quad (2.6)$$

onde n é o número de entradas do neurônio, w_i é o peso associado à entrada x_i , e θ é o limiar (*threshold*) do neurônio. Essa característica do neurônio McCulloch e Pitts é chamada de *propriedade tudo-ou-nada* (HAYKIN, 1998, p.39).

Segundo Braga *et al.* (1998), as redes neurais artificiais são compostas por unidades de processamentos simples (nodos) que computam determinadas funções matemáticas (normalmente não lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente, unidirecionais. Na maioria dos modelos, estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede.

Em RNAs, o procedimento usual na solução de problemas passa inicialmente por uma fase de aprendizagem, onde um conjunto de exemplos é apresentado para a rede, que extrai automaticamente as características necessárias para representar a informação fornecida (KOVACS, 2002). Essas características são utilizadas posteriormente para gerar respostas para o problema. A capacidade de aprender através de exemplos e de generalizar a informação aprendida são, os atrativos principais da solução de problemas através das RNAs.

Sendo assim, uma rede neural pode ser caracterizada por três aspectos principais:

- 1) Capacidade de aprendizado;
- 2) Habilidade funcional;
- 3) Tipo de rede (arquitetura).

2.5.2 Aprendizado

Segundo Braga (1998), as RNAs possuem a capacidade de aprender através de exemplos e fazer interpolações e extrapolações do que aprenderam. Uma definição geral

do que vem a ser aprendizagem pode ser expressa pela forma sugerida por Haykin (2002).

“Aprendizagem é o processo pelo qual os parâmetros de uma rede neural são ajustados através de uma forma continuada de estímulos fornecidos pelo ambiente no qual a rede está operando, sendo que o tipo específico de aprendizagem realizado, é definido pela maneira particular como ocorrem os ajustes realizados nos parâmetros” (Haykin, 2002).

Cada tipo de aprendizagem define de maneira particular como ocorrem os ajustes nos parâmetros. Isto pode ocorrer pela minimização da medida do erro, pela auto-organização da informação através da medida da correlação entre os dados ou pela maximização de recompensas através de um processo de tentativa e erro.

Nesse contexto, foram desenvolvidos diversos métodos para treinamento de redes, podendo ser agrupados em dois paradigmas principais: *aprendizado supervisionado* e *aprendizado não-supervisionado*. Outros dois paradigmas bastante conhecidos são os de *aprendizado por reforço* (que é um caso particular de aprendizado supervisionado) e *aprendizado por competição* (que é um caso particular de aprendizado não-supervisionado). Estes paradigmas serão estudados, juntamente com dois mecanismos de aprendizado encontrados na literatura: a *correção de erros* e o *aprendizado Hebbiano*.

Aprendizado Supervisionado

Este método de aprendizado recebe esse nome porque a entrada e saída desejadas para a rede são fornecidas por um supervisor externo. O supervisor indica explicitamente um comportamento bom ou ruim para a rede, visando a direcionar o processo de treinamento. A rede tem sua saída corrente comparada com a saída desejada, recebendo informações do supervisor sobre o erro da resposta atual. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos (MELO, 2004). Um esquema ilustrado pela Figura 2.3 mostra o funcionamento do aprendizado supervisionado.

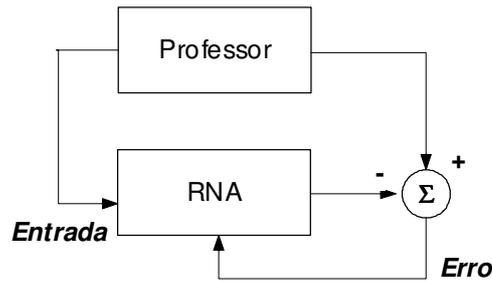


Figura 2.3 – Aprendizado Supervisionado (fonte: Braga, 2000).

Os exemplos mais conhecidos de algoritmos de Aprendizado Supervisionado são a *regra delta* (ou correção de erros) e a *aprendizagem por reforço*.

Correção de erros

O termo $e(t)$ do erro deve ser escrito como: $e(t) = d(t) - y(t)$, onde $d(t)$ é a saída desejada e $y(t)$ é a resposta atual (calculada) no instante de tempo t . A minimização dessa diferença é incremental, já que pequenos ajustes são feitos nos pesos a cada etapa de treinamento, de tal forma que estes caminhem – se houver solução – para uma solução. A forma genérica para alteração dos pesos por correção de erros é mostrada na Equação 2.7.

$$w_i(t+1) = w_i(t) + \eta \cdot e(t) \cdot x_i(t), \quad (2.7)$$

onde η é a taxa de aprendizado e $x_i(t)$ é a entrada para o neurônio i no tempo t .

A soma dos erros quadráticos de todas as saídas é normalmente utilizada como medida de desempenho da rede e também como função de custo a ser minimizada pelo algoritmo de treinamento, de acordo com a Equação 2.8.

$$F(w) = \frac{1}{2} \cdot \sum_{i=0}^k (d_i - y_i(w))^2, \quad (2.8)$$

onde d_i é a saída desejada, y_i é a saída da rede e w o peso.

Aprendizado por reforço

Sutton e Barto (1998), conceitualizaram o aprendizado por reforço como:

“Se uma ação tomada pelo sistema de aprendizagem é seguida de estados satisfatórios, então a tendência do sistema de produzir esta ação particular é reforçada. Se não for seguida de estados satisfatórios, a tendência do sistema de produzir esta ação é enfraquecida”. (Sutton & Baroto, 1998)

O Aprendizado por reforço possui um crítico externo que avalia as respostas fornecidas pela rede e direciona o ajuste dos pesos. O aprendizado é feito por um processo de tentativas e erros, que visa maximizar um dado índice de desempenho, denominado de sinal de reforço. O esquema de funcionamento do aprendizado por reforço é ilustrado pela Figura 2.4.

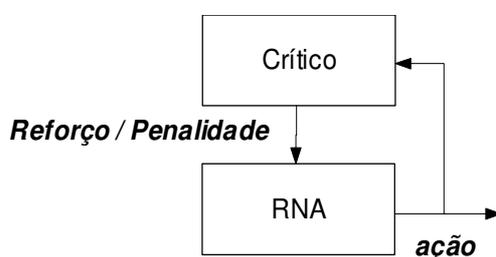


Figura 2.4 – Aprendizado por reforço (fonte: Braga, 2000).

Embora, neste caso, haja apenas dados de entrada, há um agente externo (supervisor) que retorna um sinal de reforço ou de penalidade a cada saída da rede.

Aprendizado Não Supervisionado

No aprendizado não supervisionado, como o próprio nome sugere, não há um professor ou supervisor para acompanhar o processo de aprendizado. Este método é ilustrado na Figura 2.5.



Figura 2.5 – Aprendizado Não Supervisionado (fonte: Braga, 2000).

Para este tipo de aprendizado, somente os padrões de entrada estão disponíveis para a rede, ao contrário do aprendizado supervisionado, cujo conjunto de treinamento possui pares de entrada e saída. A partir do momento em que a rede estabelece uma

harmonia com as regularidades estatísticas da entrada de dados, desenvolve-se nela uma habilidade de formar representações internas para codificar características da entrada e criar novas classes ou grupos automaticamente. Este tipo de aprendizado, só se torna possível, quando existe redundância nos dados de entrada. Sem redundância seria impossível encontrar quaisquer padrões ou características dos dados de entrada.

Uma das vantagens do aprendizado não supervisionado pode ser visualizada na ausência do supervisor ou professor (responsável pela indicação da saída desejada). A desvantagem está intimamente ligada ao conjunto de dados para treinamento, onde este deve ser redundante para que a rede consiga extrair características em seu treinamento.

Aprendizado Hebbiano

Segundo Braga *et al.* (1998, p.2), Hebb “mostrou que a aprendizagem de redes neurais é conhecida através da variação dos pesos de entrada dos nodos”. Os resultados obtidos por Hebb motivaram os primeiros métodos de aprendizado em RNAs. Como originalmente postulado, a regra de Hebb propõe que o peso de uma conexão sináptica deva ser ajustado, caso exista sincronismo entre os “níveis de atividade” das entradas e saídas. Caso dois neurônios, em lados distintos da sinapse, sejam ativados sincronamente, haverá um fortalecimento desta sinapse. Entretanto, se os neurônios forem ativados assincronamente, a sinapse será enfraquecida ou mesmo eliminada. Em outras palavras, se o neurônio pré-sináptico tiver grande influência na ativação do neurônio pós-sináptico, a conexão entre eles deve ser reforçada.

A sinapse Hebbiana possui quatro características principais:

- 1) *Mecanismo interativo*: dentro do aprendizado Hebbiano, não se pode analisar atividades pré-sinápticas de forma isolada; qualquer modificação na sinapse Hebbiana depende (estatisticamente ou deterministicamente) da interação entre os dois tipos de atividades (pré e pós-sinápticos);
- 2) *Mecanismo local*: uma sinapse Hebbiana é uma transmissão com sinais contínuos que produz modificações sinápticas locais que são entradas específicas. É a partir do mecanismo local que sinapses Hebbianas efetuam aprendizado não supervisionado;

- 3) *Mecanismo dependente do tempo*: as modificações em uma sinapse Hebbiana dependem do momento exato de ocorrência das atividades pré e pós-sinápticas.
- 4) *Mecanismo correlacional ou conjuncional*: a sinapse Hebbiana pode ser chamada de sinapse conjuncional pelo fato de que a ocorrência conjunta de atividades pré e pós-sinápticas é suficiente para que haja uma modificação. Além disso, pode também ser chamada de sinapse correlacional, porque uma correlação entre estas mesmas atividades é suficiente para gerar mudanças.

Pode-se expressar o postulado de Hebb em termos matemáticos a partir da Equação 2.9, que consiste em uma regra para a mudança do peso sináptico w_{ij} (regra do produto de atividade).

$$\Delta w_{ji}(t) = \eta \cdot x_j(t) \cdot y_i(t), \quad (1.9)$$

onde η é uma constante positiva que determina a taxa de aprendizado, y_i é a saída desejada no instante de tempo t e x_j é a entrada do neurônio j no instante de tempo t .

Apesar dos vetores de entrada e saída serem fornecidos (como no aprendizado supervisionado), a regra de Hebb é classificada como aprendizado não supervisionado, já que não existe supervisor externo para verificar a qualidade da resposta da rede para prover os ajustes de pesos. Neste caso, o treinamento da rede é feito, independentemente, da resposta atual, através de um mecanismo local à sinapse.

Aprendizado por competição

Aprendizado por competição é um caso particular de aprendizado não supervisionado. A idéia neste caso é, dado um padrão de entrada, fazer com que as unidades de saída disputem entre si para serem ativadas (MELO, 2004). Existe, portanto, uma competição entre as unidades de saída para decidir qual delas será a vencedora e, conseqüentemente, terá a sua saída ativada e seus pesos atualizados no treinamento. As unidades de entrada são diretamente conectadas às unidades de saída, sendo que estas últimas também podem estar ligadas entre si via conexões laterais inibitórias, ou negativas. A unidade de saída com maior ativação inicial terá mais chance de vencer a

disputa com as outras unidades, que perderão o poder de inibição ao longo do tempo sobre as unidades de maior ativação. A unidade mais forte fica ainda mais forte e seu efeito inibidor sobre as outras unidades de saída torna-se dominante. Com o tempo, todas as outras unidades de saída ficarão completamente inativas, exceto a vencedora (MELO, 2001). Este tipo de inibição mútua também é conhecido como *winner tasks all*.

Um algoritmo simples de aprendizado competitivo é:

1. apresentar um vetor de entrada;
2. calcular a ativação inicial de cada unidade de saída;
3. deixar as unidades de saída competirem até que apenas uma permaneça ativa;
4. aumentar os pesos sobre as conexões entre as unidades de saída ativa e as unidades de entrada ativas. Com isso, a unidade de saída terá maior probabilidade de ficar ativa na próxima repetição do padrão.

Um problema neste algoritmo é que uma unidade de saída pode se tornar dominante e ficar ativa todo tempo, podendo captar para si todo o espaço de entrada. A solução para tal problema está em racionar os pesos, de tal forma que a soma dos pesos sobre as linhas de entrada de uma unidade seja limitada a 1. Para aumentar o peso de uma conexão, é preciso diminuir o peso de outra, como mostra a Equação 2.10.

$$\Delta w_j = \frac{\eta \cdot x_j}{M - \eta \cdot w_j}, \quad (2.10)$$

em que para todo $j = 1, 2, \dots, n$, onde w_j é o peso da conexão da unidade de entrada j com a unidade ativa, x_j é o valor da j -ésima entrada, M é o número de unidades de entradas ativas no vetor de entrada e η é a taxa de aprendizagem.

2.5.3 Função de ativação

A partir do modelo proposto por McCulloch e Pitts foram derivados vários outros modelos que permitem a produção de uma saída qualquer, não necessariamente zero ou um, e com diferentes funções de ativação. Estas processam um conjunto de entradas

recebidas e o transformam em estado de ativação. As funções de ativação típicas envolvem adições, comparações e transformações matemáticas. A seguir são descritos alguns tipos mais comuns de funções de ativação (HAYKIN, 2002; KOSKO, 1992):

Função Limiar

Utilizada no modelo de McCulloch e Pitts, a função limiar modela a característica do “tudo ou nada” deste neurônio, e é expressa conforme Equação (2.11).

$$f(v) = \begin{cases} 1, & \text{se } v \geq 0; \\ 0, & \text{se } v < 0; \end{cases} \quad (2.11)$$

onde a saída $f(v)$ será igual a 0, caso o valor de ativação v seja negativo e 1 nos casos em que o valor de ativação seja positivo. A Figura 2.6 correspondente ao gráfico de uma função limiar em que o valor de ativação v é composto pelo combinador linear e pelo bias.

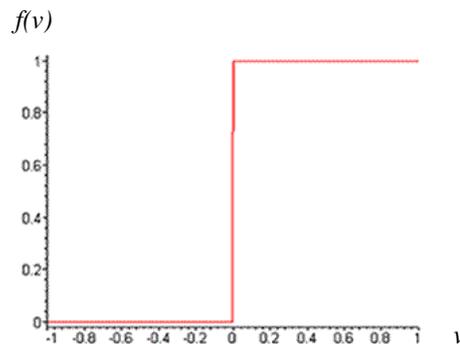


Figura 2.6 – Gráfico de uma função de ativação limiar.

Função Sigmóide Logística

Essa função, ao contrário da função limiar, pode assumir todos os valores entre 0 e 1. Conhecida também como *S-shape*, é uma função semi-linear, sendo possível definir várias funções sigmoidais. Uma das funções sigmoidais mais importantes é a função logística definida pela Equação 2.12.

$$y = f(x) = \frac{1}{1 + e^{-x/T}}, \quad (2.12)$$

onde o parâmetro T determina a suavidade de curva. O gráfico desta função é apresentado na Figura 2.7.

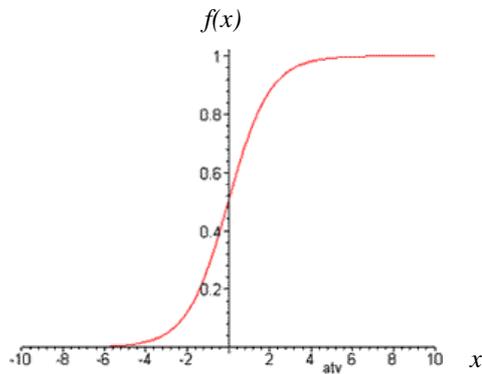


Figura 2.7 – Exemplo de gráfico de uma função sigmóide logística.

Função de Base Radial

As funções de base radial (RBF) são caracterizadas por uma resposta que decresce ou cresce monotonicamente com a distância a um ponto central. Os parâmetros mais importantes a serem definidos em uma função de base radial são o centro e a taxa de crescimento, ou decréscimo, da função. Uma função de base radial típica é a Gaussiana, definida pela Equação 2.13.

$$y = f(x) = e^{-\frac{t^2}{2 \cdot \sigma^2}}, \quad (2.13)$$

onde σ representa o desvio padrão e t o instante de tempo. A função Gaussiana está representada pelo gráfico da Figura 2.8.

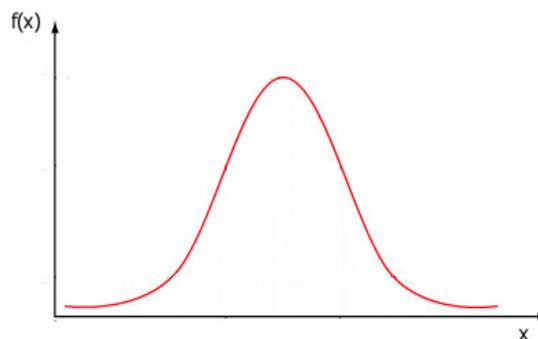


Figura 2.8 – Gráfico genérico de uma função Gaussiana.

2.5.4 Tipos de Redes Neurais

De um modo geral, o tipo da rede neural é tido através da sua arquitetura. A definição da arquitetura de uma RNA é um parâmetro importante na sua concepção, uma vez que ela restringe o tipo de problema que pode ser tratado pela rede. Fazem parte da definição da arquitetura os seguintes parâmetros: número de camadas da rede, número de nodos em cada camada, tipos de conexão entre os nodos e topologia da rede. Alguns exemplos de arquiteturas de RNAs são apresentadas na Figura 2.9.

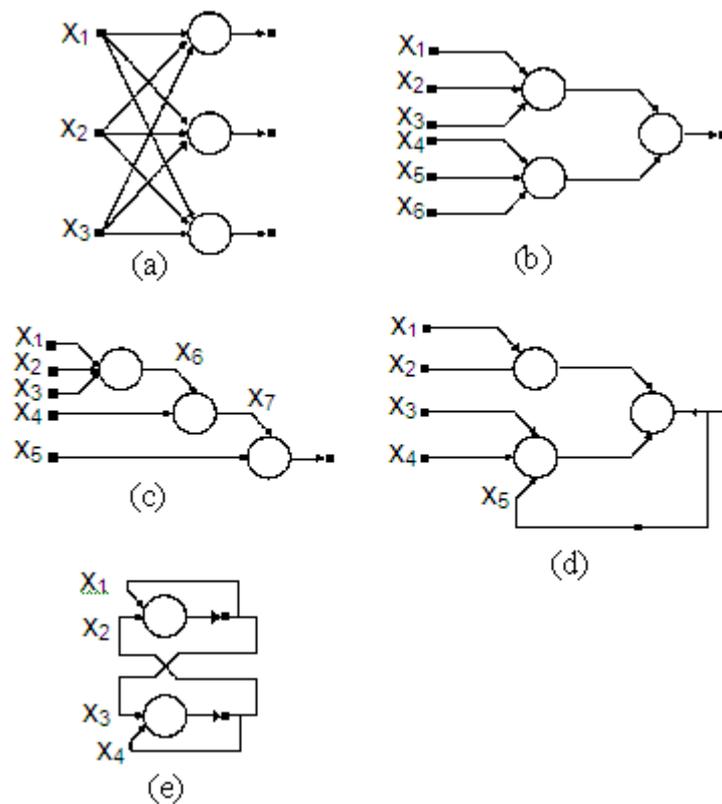


Figura 2.9 – Exemplos de arquiteturas de RNA: (a) única camada; (b, c, d) múltiplas camadas; (a, b, c, d) acíclica ou alimentada adiante (*feedforward*); (e) cíclica ou recorrente (*feedback*) (fonte: Braga, 2000).

Quanto ao número de camadas, pode-se ter:

- *Redes de camada única* – só existe um nó entre qualquer entrada e qualquer saída da rede. Vide Figuras 2.9 (a) e (e);
- *Redes de múltiplas camadas* – existe mais de um neurônio entre alguma entrada e alguma saída da rede. Vide Figuras 2.9 (b), (c) e (d).

Quanto ao tipo de conexões dos nodos, pode-se ter:

- *Feedforward ou acíclica* – a saída de um neurônio na i -ésima camada da rede não pode ser usada como entrada de nodos em camadas de índice menor ou igual a i . Vide Figuras 2.9 (a), (b) e (c);
- *Feedback ou cíclica* – a saída de algum neurônio na i -ésima camada da rede é usada como entrada de nodos em camadas de índice menor ou igual a i . Vide Figuras (d) e (e).

Quanto à conectividade da rede, pode-se ter:

- *Rede fracamente (ou parcialmente) conectada*. Vide Figuras (b), (c) e (d);
- *Rede completamente conectada*. Vide Figuras (a) e (e).

Em termos mais técnicos, o número de camadas define a capacidade de representação das relações entre o espaço de entrada e o de saída. A inexistência da camada intermediária, característica do modelo *perceptron* (classificador binário), condiciona-o a representar bem somente relações linearmente independentes. A existência de camadas intermediárias, característica do modelo *perceptron de múltipla camada* (MLP – *Multi Layer Perceptron*), retira tal limitação. Se houver apenas uma camada intermediária, o MLP pode representar (com qualquer grau de aproximação, por menor que seja) qualquer função contínua. Duas ou mais camadas ampliam o universo de representação a qualquer função, contínua ou não.

As arquiteturas do tipo MLP constituem os modelos neurais artificiais mais utilizados e conhecidos. Tipicamente, esta arquitetura consiste de um conjunto de unidades sensoriais que formam uma camada de entrada, uma ou mais camadas intermediárias (ou escondidas) de unidades computacionais e uma camada de saída. Os sinais de entrada são propagados camada a camada pela rede em uma direção positiva, ou seja, da entrada para a saída.

O algoritmo de treinamento mais usado para este tipo de rede é o *backpropagation*, que baseia-se na aprendizagem por correção de erro (HAYKIN 2002). As redes do tipo MLP tem sido utilizadas com sucesso para a solução de vários problemas envolvendo altos graus de não-linearidade, porém, alguns trabalhos mostraram que a rede RBF (*Radial-Basis Function*), obteve um melhor desempenho perante alguns tipos de problemas (Moura, *et al.*, 2002; Schmitz, 2006; Almeida, 2009).

Assim como as redes MLP, as redes RBF são redes neurais multicamadas. A principal diferença é que as redes RBF utilizam funções de base radial nos nós da camada oculta (BRAGA *et al.*, 2000). Esse tipo de função é definida pela Equação 2.14.

$$f(x) = \sum_{i=1}^N w_i \phi(\|x - x_i\|) \quad (2.14)$$

onde N é o número de neurônios na camada escondida, w_i são os pesos dos i neurônios de saída linear e $\|x - x_i\|$ é a distância euclidiana entre a entrada x e a entrada x_i .

As redes RBF possuem duas variantes, são elas: *Probabilistic Neural Networks* (PNN) e *Generalized Regression Neural Networks* (GRNN). As redes PNN e GRNN, possuem sistemas similares, mas existe uma diferença fundamental entre elas: as redes probabilísticas realizam uma classificação onde a variável alvo é categórica, já as GRNN executam um regressão onde a variável alvo é contínua.

Neste contexto, foi utilizado para esta pesquisa um *software* da *Ward Systems Group*, o *NeuroShell Classifier*, que aplica Redes Neurais Probabilísticas (*Probabilistic Neural Networks*) com algoritmo de correlação em cascata (*Cascade Correlation*) - CC, que é um algoritmo construtivo capaz de sintetizar automaticamente a arquitetura de uma rede como parte do seu processo de treinamento.

2.5.5 Redes Neurais Probabilísticas

As Redes Neurais Probabilísticas ou *Probabilistic Neural Network* (PNN) surgiram com os trabalhos pioneiros de Donald F. Specht (1988, 1990). Basicamente trata-se de uma implementação, via rede neural artificial de um mecanismo *Classificador de Parzen* que estima a função de distribuição de probabilidade.

Redes do tipo PNN's podem ser utilizadas em diversas áreas de estudos, como por exemplo, em problemas relacionados ao diagnóstico e prognóstico de doenças, em particular, de diversos tipos de câncer.

Na meteorologia, as redes do tipo PNN já foram utilizadas na resolução do problema da estabilidade e divisão de fases na simulação de processos de separação de sistemas complexos (SCHMITZ, 2006). Almeida (2009) utilizou a rede PNN como um

sistema de previsão a curtíssimo prazo do teto de nuvens e da visibilidade horizontal no Aeroporto de Guarulhos – SP, obtendo respostas com uma média de 90% de acertos. Bankert (1994) utiliza uma rede neural probabilística na classificação de nuvens das imagens do AVHRR (*Advanced Very High Resolution Radiometer*) em regiões marítimas. Utilizando cinco classes (nuvens baixas, altostratus, nuvens altas, nuvens precipitantes e céu claro), o autor obteve precisão de classificação em 91,2% das amostras.

2.5.6 Algoritmo de correlação em cascata

O algoritmo de correlação em cascata foi desenvolvido por Fahlman *et al.* (1991). Este algoritmo controla o número de neurônios a serem utilizados em uma rede neural durante o treinamento da mesma através de um critério de seleção de neurônio baseado na correlação que este neurônio tem com a saída da rede. O controle do número de neurônios em uma rede neural é importante, pois configurações com número de neurônios maior que o necessário tenderá ao sobre-ajuste (*overfitting*), isto é, a rede treinará bem, mas terá um desempenho ruim no conjunto de testes, entretanto, se o número de neurônios for muito baixo, a rede será incapaz de atingir o desempenho desejado, mesmo durante o treinamento (VONK, 1997).

Queiroz (2004) apresentou um estudo comparativo entre duas técnicas de classificação de imagens de satélite, das quais uma delas partia da inteligência artificial baseada em redes neurais artificiais, usando o algoritmo de Correlação em Cascata. O motivo que levou o autor a optar por este tipo de rede foi o fato de não ser necessária a configuração do número de neurônios da camada intermediária, uma vez que o algoritmo de Correlação em Cascata proporciona uma Rede Construtiva. Isto constitui uma vantagem, visto que em trabalhos que usam outros tipos de redes são necessários vários testes com diferentes números de neurônios na camada intermediária, a fim de se obter a quantidade ideal de neurônios para melhor aprendizagem das redes.

Capítulo 3 – DADOS

3.1 Introdução

Para realização deste estudo foram utilizados os dados de observações meteorológicas de superfície e ar superior, contando também com dados de descargas atmosféricas e totais horários de precipitação na área do aeroporto internacional do Galeão e adjacências, localizada entre as coordenadas geográficas de 23° a 22° 36' S de latitude e 43° 27' a 43° 3' W de longitude e limitada pela faixa litorânea que se estende, aproximadamente, do sul da região ocupada pela Barra da Tijuca ao norte da Ilha do Governador, cobrindo uma área em torno de 1829 km², ilustrada na Figura 3.1.

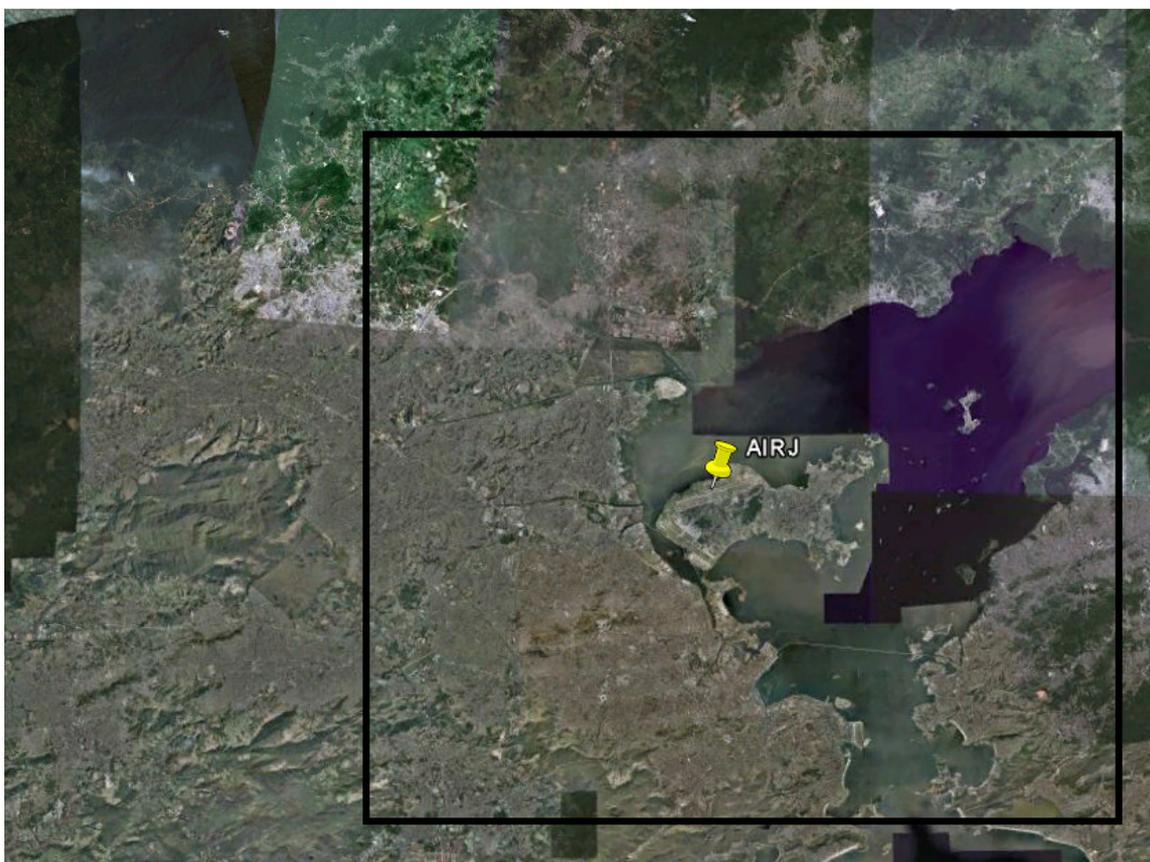


Figura 3.1 – Imagem da região de estudo representada pelo quadrante na cor preta localizada entre as coordenadas geográficas de 23° a 22° 36' S de latitude e 43° 27' a 43° 03' W de longitude (fonte: GOOGLE MAPS).

Os dados de observações meteorológicas de superfície consistem em METARes confeccionados a cada hora do dia por operadores do Comando da Aeronáutica para as seguintes Estações Meteorológicas de Superfície (EMS): EMS-AF (Afonso), EMS-GL (Galeão), EMS-RJ (Santos Dumont), EMS-SC (Santa Cruz) e EMS-JR (Jacarepaguá); todos obtidos por meio da Universidade de *Wyoming*. Esses dados podem ser visualizados na Rede de Meteorologia do Comando da Aeronáutica (REDEMET). Entretanto, por ter sido fundada em 2003, esta rede contém apenas os dados desse período até os dias atuais.

Das informações geradas pelo código METAR, concorreram como dados de entrada para a rede **a velocidade e direção do vento, cobertura celeste do aeródromo, visibilidade horizontal, altura das nuvens, tempo presente, temperatura do ar, ponto de orvalho e pressão atmosférica da EMS-GL**. Das demais estações meteorológicas, extraiu-se apenas os **registros de trovoadas no campo pertinente ao tempo presente na estação**, o que caracteriza uma atmosfera instável aos arredores da EMS-GL, localizada no AIRJ.

Os dados de ar superior e seus derivados foram obtidos da Estação Meteorológica de Ar Superior do Galeão (EMA-GL) pela *National Oceanic and Atmospheric Administration* (NOAA). Esses dados são confeccionados às 00:00 Z e às 12:00 Z, ou seja, às 21:00 hs e às 9:00 hs do horário local, informando as condições gerais da estrutura vertical da atmosfera.

Os dados de descargas atmosféricas não serviram como dados de entrada da rede, foram utilizados para avaliação e análise a fim de aumentar o desempenho dos resultados. Esses dados foram obtidos através das Furnas Centrais Elétricas S. A. (ELETROBRAS FURNAS) e consistem em latitude, longitude, data, hora e minuto de descargas localizadas pelos sensores da Rede Integrada Nacional de Detecção de Descargas Atmosféricas (RINDAT). Na Figura 3.2 é apresentada a distribuição dos sensores da RINDAT e as instituições pertencentes a rede.

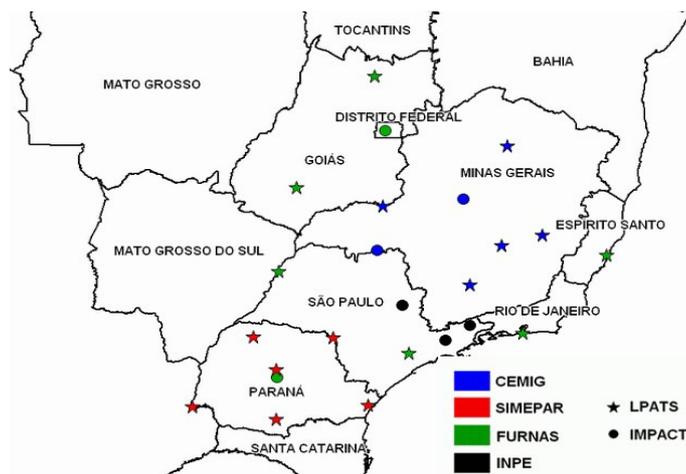


Figura 3.2 – Distribuição dos sensores da RINDAT no território brasileiro (fonte: RINDAT).

Os dados de precipitação que serviram como base para este trabalho referem-se aos observados nos postos da rede pluviométrica instalada pela Fundação Geo-Rio, que estão espalhados em 32 estações no Município do Rio de Janeiro, obtidos do site do Alerta Rio, onde também está disponível a localização de cada estação pluviométrica, conforme Figura 3.3.

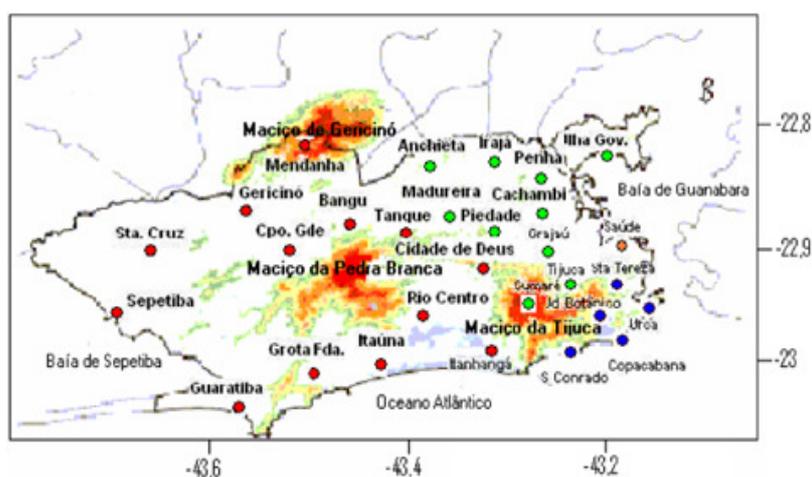


Figura 3.3 – Localização dos postos pluviométricos da Fundação Geo-Rio (fonte: Adaptado de <http://www2.rio.rj.gov.br/georio/site/alerta/estacao.local.htm>).

Criada em 25 de setembro de 1996, a Fundação Geo-Rio visa atender ao sistema de alerta de chuvas intensas, e como pode ser visto na Figura 3.3, as estações não são homoganeamente distribuídas. A data de criação da Fundação Geo-Rio é o motivo pelo qual o período de dados desta pesquisa iniciou-se em 01 janeiro de 1997 e estendeu-se até

o dia 31 de dezembro de 2009, obtendo-se assim, 13 anos para análise. Utilizando-se de 30 estações, em que as informações são registradas em tempo real, foram gerados os acumulados horários de precipitação integrando os acumulados de 15 minutos fornecidos pelo Alerta Rio.

3.2 Pré-processamento

No pré-processamento dos dados, primeiramente, foram inseridos os dias e as horas que faltavam para as séries de dados de superfície, altitude e precipitação. No lugar dos espaços destinados aos dados que não foram registrados, colocou-se a formatação “xxx” significando ausência de dados naquele momento.

Os acumulados horários de precipitação apresentavam-se dispostos de forma que cada arquivo obtido correspondia a uma estação. Dentro do arquivo, as linhas ofereciam informações sobre dia, mês e ano, enquanto que as colunas representavam os horários que foram calculados os acumulados, conforme mostrado no Anexo B. Para que os dados de precipitação ficassem na mesma disposição dos dados de altitude e de superfície, ou seja, onde cada linha corresponde a cada hora do dia e cada coluna um dado diferente (no caso da precipitação, uma estação diferente), foi feito um transporte na matriz acúmulo horário de precipitação, e colocado sucessivamente cada coluna posterior embaixo de sua antecessora, através do *script* no programa MATLAB, descrito no Apêndice I.

Estas operações possibilitaram a disposição dos dados de precipitação da seguinte forma: 30 colunas correspondentes às estações da Geo-Rio utilizadas na pesquisa e 113952 linhas correspondentes ao registro do acúmulo horário da precipitação durante 13 anos, conforme mostra o Anexo B.

Nos dados de ar superior foram selecionados os seguintes níveis de pressão: 1000 hPa, 850 hPa, 700 hPa e 500 hPa. Estes níveis foram escolhidos por estarem presentes em todas as sondagens realizadas e por serem níveis padrões, ou seja, níveis que por convenção são selecionados para fins de análise e pesquisas atmosféricas. Além disso, considerando a operacionalidade relativa ao pouso e decolagem de aeronaves e às formações de sistemas convectivos, decidiu-se pela exclusão dos níveis acima de 500 hPa. Como também pela exclusão do nível padrão de 925 hPa, pois devido a sua proximidade ao nível de 1000 hPa, apresentava alta compatibilidade entre suas

informações (correlação de aproximadamente 96%), que influenciaram na escolha de apenas um deles (o que fosse necessário para confecção dos índices de instabilidade utilizados na pesquisa), portanto, optou-se pelo nível de 1000 hPa.

Neste momento surgiu uma problemática em relação a disposição dos dados de altitude, pois era o único dado de entrada da rede que não estava ordenado de forma horária. Conforme o MCA105-2, as sondagens são feitas às 12Z (09:00 hs) e às 00Z (21:00 hs), o que, em princípio, dificultou a homogeneização desses dados com os de superfície e precipitação. Para solucionar este problema, a sondagem das 12Z foi repetida até a sondagem das 23Z do mesmo dia, enquanto a sondagem das 00Z foi repetida até a sondagem das 11Z do dia seguinte. Este procedimento foi feito para cada dia do ano utilizando uma variável chamada “PESO_1” capaz de indicar o grau de certeza dos dados de ar superior que foram repetidos. Essa variável foi calculada a partir de uma progressão aritmética decrescente onde o primeiro termo admite o valor 1, correspondente a 100% de certeza, enquanto que os demais termos admitem valores correspondentes a uma razão de $-1/12$. Um extrato dos dados de altitude com os pesos correspondentes se encontra no Anexo C. A Figura 3.4 ilustra os valores adquiridos pela variável PESO_1 de acordo com a hora.

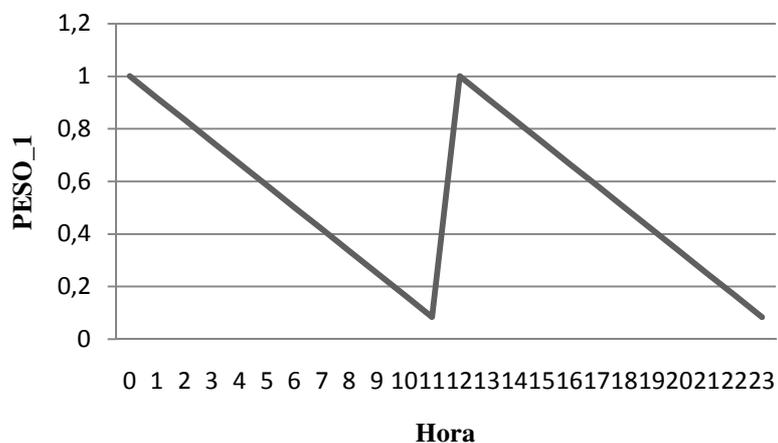


Figura 3.4 – Pesos correspondentes ao grau de certeza das variáveis dos dados de altitude.

Criada a variável PESO_1, pôde-se agrupar os dados de superfície, altitude e precipitação, pois todos passaram a apresentar uma organização horária, disponibilizando, assim, a análise preliminar desse conjunto de dados.

3.3 Análise preliminar dos dados

Dados de superfície – A análise preliminar dos dados de superfície excluiu a estação de Jacarepaguá, pois a quantidade de dados obtidos para esta estação no período considerado não colaborou com o desempenho da pesquisa. A quantidade de dados de superfície (METAR) obtidos para cada estação encontra-se ilustrada na Tabela 3.1.

Tabela 3.1 – Quantidades e porcentagens de dados de superfície adquiridos para as Estações Meteorológicas de Superfície localizadas no município do Rio de Janeiro.

Estações	Quantidade de METAR	Porcentagem de dados
EMS-GL	110533	97
EMS-RJ	66092	58
EMS-AF	47860	42
EMS-SC	39883	35
EMS-JR	31907	28

Dados de altitude – Na análise dos dados do ar superior, foi observada a ausência aleatória de algumas sondagens dentro do período estudado, contando com um total de 28320 dados, que corresponde a 75% do total de dados constantes de 01 de janeiro de 1997 a 31 dezembro de 2009 para a estação meteorológica de altitude do Galeão.

Dados de precipitação – A fim de se obter um melhor resultado na previsão dos eventos extremos de precipitação para o AIRJ, fez-se não somente a análise dos dados obtidos no aeroporto, mas também um estudo do acúmulo horário da precipitação das estações próximas, contando com um total de 30 estações. Incluindo-se, entre elas: Anchieta, Bangú, Cachambi, Campo Grande, Cidade de Deus, Copacabana, Gericinó, Grajaú, Grota Funda, Guaratiba, Ilha do Governador, Irajá, Itanhangá, Itauna, Jardim Botânico, Madureira, Mendanha, Penha, Piedade, Rio Centro, Santa Cruz, Santa Tereza, São Conrado, Saúde, Sepetiba, Sumaré, Tanque, Tijuca, Urca e Vidigal (vide Figura 3.3).

Dados de descargas atmosféricas – Para análise desse tipo de dados na área de influência do AIRJ foi feito um estudo de ocorrência de descargas atmosféricas no quadrante do local de estudo compreendido entre -23° a -22,6° de latitude e -43,45° a -43,05° de longitude, ou seja, o equivalente 1829 km² de área (vide Figura 3.1).

Após o pré-processamento, ao agrupar os dados de superfície, altitude e precipitação, notou-se que em diversas variáveis estavam presentes a expressão “xxx”, que significa “ausência de dados”. Ao excluir as linhas que continham dados ausentes, permaneceram 62.320 de um total de 113.952 linhas de dados, representando cerca de 55% do total. Uma análise preliminar desse conjunto de dados mostrou que eventos de **chuva forte e/ou trovoadas** não ocorrem com muita frequência em relação a outros eventos como por exemplo a névoa úmida ou névoa seca. Conseqüentemente esses eventos (chuva forte e/ou trovoadas) são difíceis de prever. O gráfico da Figura 3.5 ilustra a frequência dos eventos registrados no período analisado, onde as condições de tempo presente foram divididas conforme as seguintes “faixas de tempo”:

- **Faixa 0** – Representa a ocorrência de nevoeiro, fumaça, névoa úmida, névoa seca ou quando nenhum tempo significativo for registrado;
- **Faixa 1** – Chuva leve (Traços a 2,4 mm/h) sem trovoadas;
- **Faixa 2** – Chuva moderada (2,5 a 9,9 mm/h) sem trovoadas;
- **Faixa 3** – Chuva forte (acima de 9,9 mm/h) e/ou eventos com trovoadas.

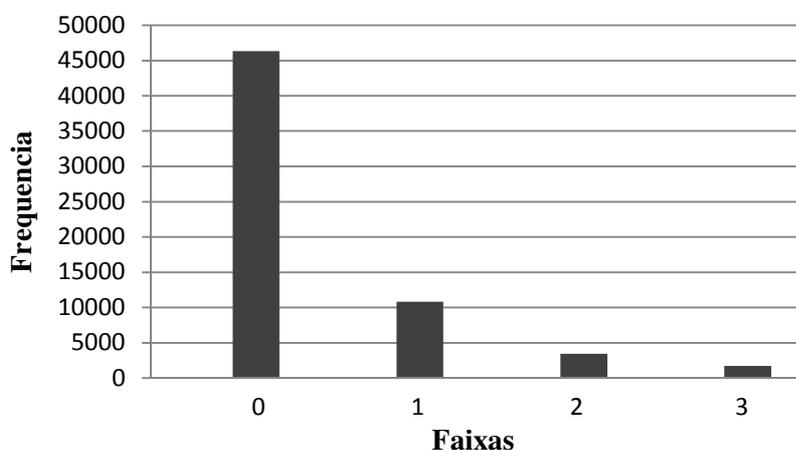


Figura 3.5 – Histograma de frequência relativa ao número de ocorrências horárias dos fenômenos pertencentes às faixas de tempo 0, 1, 2 e 3 dos dados obtidos para as estações do Galeão, Santos Dumont, Campos dos Afonsos e Santa Cruz, no período de janeiro de 1997 a dezembro de 2009.

O gráfico da Figura 3.5 está representado na Tabela 3.2, onde se pode observar a quantidade exata e a porcentagem correspondente de dados para cada faixa de tempo presente das estações meteorológicas analisadas.

Tabela 3.2 – Frequência relativa ao número de ocorrências horárias e porcentagem dos fenômenos pertencentes as faixas de tempo 0, 1, 2 e 3 dos dados obtidos para as estações do Galeão, Santos Dumont, Campo dos Afonsos e Santa Cruz, no período de janeiro de 1997 a dezembro de 2009.

Faixas	Frequência	Porcentagem
0	46326	74,33%
1	10806	17,34%
2	3446	5,53%
3	1742	2,80%
Total	62320	100,00%

Ao avaliar as quantidades de ocorrência de eventos extremos na área de influência do Aeroporto Internacional Tom Jobim, observou-se uma variação na frequência desses eventos de acordo com os meses do ano. O que já era de se esperar, pois eventos de chuva forte e eventos de trovoadas são características que dependem de diversos fatores, entre eles a *sazonalidade* (ou seja, característica própria de uma estação), a qual interfere de modo a proporcionar maior ou menor intensidade das chuvas (LIEBMANN *et al.*, 2000).

As variações na ocorrência de certos eventos meteorológicos conforme os meses do ano puderam ser vistas em todos os anos pertencentes à série de dados obtidos, conforme mostrada na Figura 3.6.

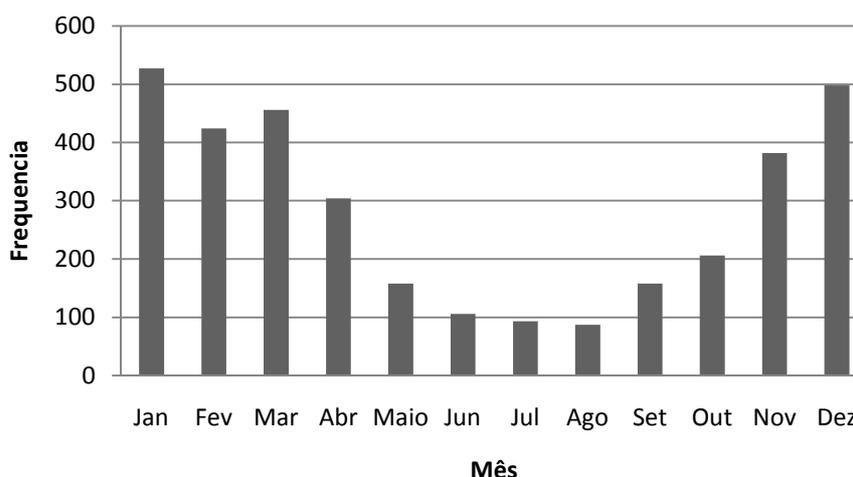


Figura 3.6 – Frequência horária de chuva forte e/ou trovada durante os meses dentro do período de 1997 a 2009, na área do Aeroporto Internacional do Galeão e adjacências, compreendidos nas coordenadas geográficas de 23° a 22° 36' S de latitude e 43° 27' a 43° 03' W de longitude.

A Figura 3.6 mostra que o período com mais eventos de chuvas fortes ou trovoadas na região estudada, ocorre no verão, apresentando normalmente, um máximo pluviométrico em janeiro e um mínimo no mês de julho/agosto, período da estação mais seca no Rio de Janeiro. Essa variação no número de ocorrência dos eventos citados acima é dada também de acordo com a hora do dia. Para todos os meses, observou-se que esses eventos ocorrem com máximos frequentemente no período noite/madrugada, e mínimos durante o dia entre o final da manhã e início da tarde, conforme mostra a Figura 3.7.

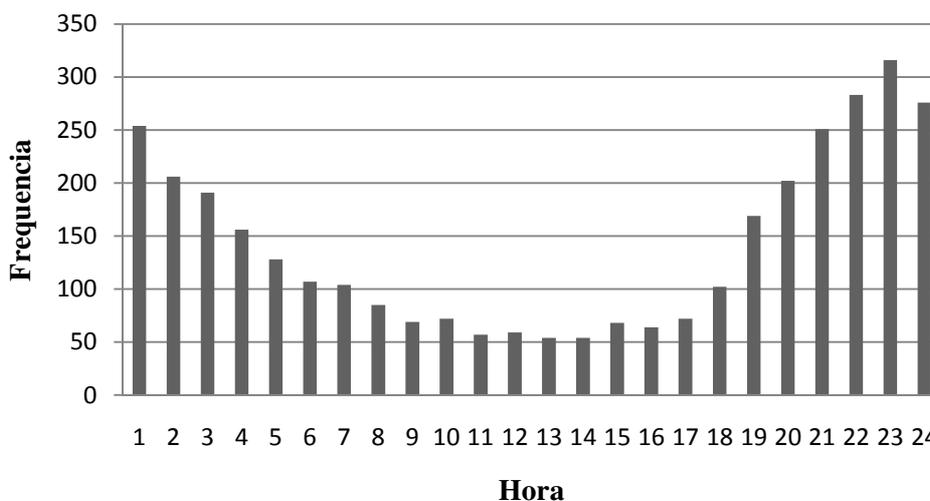


Figura 3.7 – Frequências horárias de chuva forte e/ou trovoada na área do Aeroporto Internacional do Galeão e adjacências, compreendidos nas coordenadas geográficas de 23° a 22° 36' S de latitude e 43° 27' a 43° 03' W de longitude, para cada hora do dia no período de janeiro de 1997 a dezembro de 2009.

Dessa análise, sugeriu-se a criação de uma nova variável chamada PESO_2, que avalia o grau de influência da hora do dia na ocorrência ou não de fenômenos meteorológicos extremos, como a chuva forte ou a trovoada, em função do percentual de frequência dos mesmos em cada hora. A variação desse percentual com a hora está ilustrada na Figura 3.8.

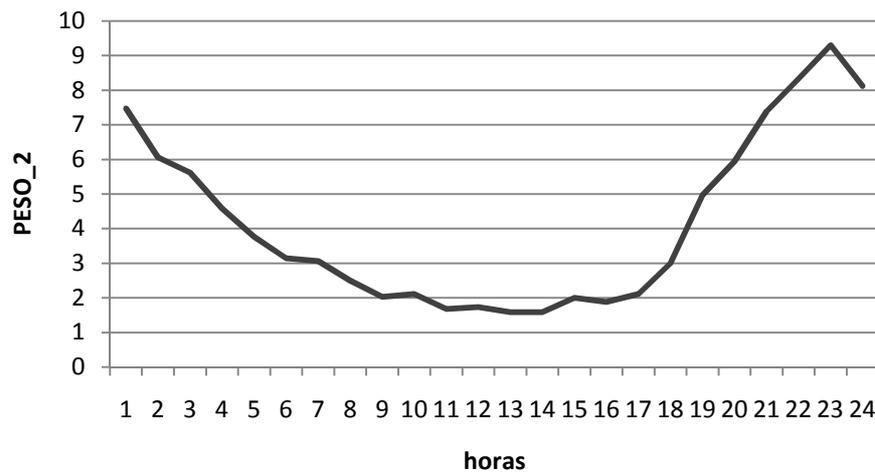


Figura 3.8 – Peso relativo ao percentual da frequência de chuva forte e/ou trovoada durante o dia no período de janeiro de 1997 à dezembro de 2009.

A inclusão da variável PESO_2 mostrou-se importante, pois aumentou em cerca de 2,5% o desempenho dos resultados gerados pela rede neural.

3.4 Criação de variáveis derivadas

Mediante a constatação de que a inclusão de uma nova variável poderia aumentar o desempenho do modelo (vide a inclusão da variável PESO_2), gerou-se a idéia de criar variáveis derivadas das já existentes. Através dos dados da radiossondagem do Galeão calculou-se índices de instabilidade, que auxiliaram como variáveis de entrada para a RNA. Os índices K, TT e LR, foram calculados mediante as fórmulas citadas no item 2.2.1.

Para obtenção desses índices foi necessário o cálculo da *temperatura do ponto de orvalho* que é aquela em que o vapor de água contido no ar começa a condensar, produzindo orvalho ou neblina. Para esse cálculo foi utilizada a Equação 3.1.

$$Td = \frac{B}{L \left[\frac{A\varepsilon}{wp} \right]}, \quad (3.1)$$

onde:

A = 2,53 x 10¹¹ em Pascals

B = 5,42 x 10³ em Kelvin

L = Logarítmo neperiano

p = pressão em Pascais

$\varepsilon = 0,622$ (relação entre as massas moleculares do ar seco e do vapor d'água)

$w = \frac{\varepsilon e}{p - e}$ (razão de mistura, onde e é a pressão de vapor em Pascais)

Além dos índices de instabilidade, outras variáveis foram geradas a fim de melhorar o entendimento da rede neural sobre o comportamento dos dados. Essas novas variáveis constam de informações de 1 hora, 2 horas, 3 horas e 24 horas anteriores à hora atual dos dados de registros da pressão da estação (QNH) e dos fenômenos de tempo encontrados nos METARes, classificados na Tabela 3.3.

Tabela 3.3 – Classificação das condições do tempo obtidas dos METARes.

CÓDIGO	FENÔMENO	CLASSE	CÓDIGO	FENÔMENO	CLASSE
H	Névoa	0	R	Chuva moderada	3
K	Fumaça	0	RF	Chuva moderada com nevoeiro	3
F	Nevoeiro	0	R+	Chuva forte	4
L-	Chuvisco leve	1	R+ F	Chuva forte com nevoeiro	4
L- F	Chuvisco leve com nevoeiro	1	RW	Pancada de chuva	4
L	Chuvisco moderado	1	RW+	Pancada de chuva forte	4
LF	Chuvisco moderado com nevoeiro	1	T	Trovoada	5
L	Chuvisco forte	1	TL	Trovoada com chuvisco moderado	5
R-	Chuva leve	2	TRW-	Trovoada com pancada de chuva leve	6
R- H	Chuva leve com névoa	2	TRW	Trovoada com pancada de chuva moderada	6
R- F	Chuva leve com nevoeiro	2	TRW+	Trovoada com pancada de chuva forte	6

Durante os testes realizados com a rede neural, as informações de horas anteriores e índices de instabilidade renderam um acréscimo de cerca de 6% no aprendizado da mesma, justificando assim, a inclusão de tais variáveis.

3.5 Seleção do evento extremo e criação das variáveis de saída da RNA

Na seleção dos casos de eventos extremos de precipitação consideraram-se aqueles que de acordo com a tabela de classificação da precipitação do MCA105-2/2010 foram classificados como chuva forte, ou seja, aqueles em que o acúmulo do pluviógrafo fosse superior a 9,9 mm/h. Para isso foi criada a variável MAIOR_PRECIPITAÇÃO, onde constam as maiores quantidades horárias de precipitação registrada por, pelo menos, uma das 30 estações através da função MAIOR do programa *Excel*, recebendo como variação uma numeração de 1 a 3, equivalente ao grau de intensidade da variável MAIOR_PRECIPITAÇÃO, de acordo com a Tabela 3.4.

Tabela 3.4 – Valores de referência da variável MAIOR_PRECIPITAÇÃO.

Intensidade	MAIOR_PRECIPITAÇÃO
Leve	1
Moderada	2
Forte	3

As variáveis de saída da rede têm um formato binário que configura não só o grau de intensidade da maior precipitação registrada dentre as 30 estações, mas também a presença ou não de trovoadas nos METARes do Galeão, Santos Dumont, Santa Cruz e Campo dos Afonsos. Portanto quando a variável MAIOR_PRECIPITAÇÃO for igual a 3 e/ou quando a variável CLASSE (Tabela 3.4) for igual a 5 ou 6, estará configurada a condição do evento extremo, o qual pretende-se prever.

A partir daí, foram criadas três variáveis de saída para a RNA (PREV_1H, PREV_2H e PREV_3H), classificadas da seguinte forma:

- 0 – Condições do tempo em que não ocorreu chuva forte ou trovoadas.
- 1 – Condições do tempo em que ocorreu chuva forte ou trovoadas.

Onde:

- PREV_1H – condição 0 ou 1 da primeira hora seguinte;
- PREV_2H – condição 0 ou 1 da segunda hora seguinte;
- PREV_3H – condição 0 ou 1 da terceira hora seguinte.

3.6 Escolha das variáveis de entrada da RNA

Como variáveis independentes (preditoras), candidatas a variáveis de entrada da RNA concorre o conjunto de variáveis obtidas de três grupos de dados:

1. Dados de Altitude da EMA-GL;
2. Dados de Superfície (METAR) das estações: EMS-GL, EMS-RJ, EMS-SC, EMS-AF;
3. Acumulados horários de precipitação de 30 estações da Geo-Rio.

Os dados de altitude são compostos pelas variáveis apresentadas na Tabela 3.5 para cada nível de pressão utilizado nesta pesquisa (100 hPa, 850 hPa, 700 hPa e 500 hPa).

Tabela 3.5 – Variáveis dos dados de altitude, derivadas das sondagens do Galeão.

CÓDIGO	DADO DE ALTITUDE
GEOPOBS	Geopotencial observado
GEOPCALC	Geopotencial calculado
T	Temperatura
Td	Temperatura do ponto de orvalho
TEMPGRAD	Gradiente da temperatura
PTEMP	Temperatura potencial
PTEMPGRAD	Gradiente da temperatura potencial
VAPPRESS	Pressão de vapor
SATVAP	Pressão de saturação do vapor
RH	Umidade relativa
RHGRAD	Gradiente da umidade relativa
UWND	Componente zonal do vento
UWNDGRAD	Gradiente vertical do vento zonal
VWND	Componente meridional do vento
VWNDGRAD	Gradiente vertical da componente meridional do vento

Contudo, dos dados de superfície foram selecionadas as variáveis expostas na Tabela 3.6.

Tabela 3.6 – Variáveis dos dados de superfície que foram utilizadas.

CÓDIGO	DADO DE SUPERFÍCIE
QNHsup	Ajuste de valor da pressão na superfície com relação ao nível do mar
DIR	Direção do vento
VEL	Velocidade do vento
VIS	Visibilidade
TEMP	Temperatura
DEW	Temperatura do ponto de orvalho
RH	Umidade relativa
WW	Tempo presente

Com o objetivo de identificar as variáveis meteorológicas de maior influência no processo de formação e evolução dos sistemas convectivos foi aplicada a técnica de estatística de correlação. Esta correlação foi realizada para todas as variáveis candidatas a preditoras da rede com a variável de saída. Os resultados estão na Tabela D.1 do Anexo D.

Após análise da correlação entre as variáveis, fizeram-se testes na rede neural, abandonando as variáveis candidatas que tiveram um alto índice de correlação (acima de 80%) com outra variável candidata, ou as que tiveram baixa correlação (abaixo de 50%) com as variáveis de saída (variáveis dependentes), o que diminuiu cerca de 3% no índice de acerto da Rede. Assim sendo, a opção recaiu pela utilização de todas as variáveis que foram testadas.

Capítulo 4 – MÉTODOS

Neste capítulo será apresentada passo a passo a metodologia utilizada para a realização deste trabalho. Um resumo ilustrativo do que foi feito durante o processo de manipulação dos dados e utilização da rede neural pode ser observado através da Figura 4.1.

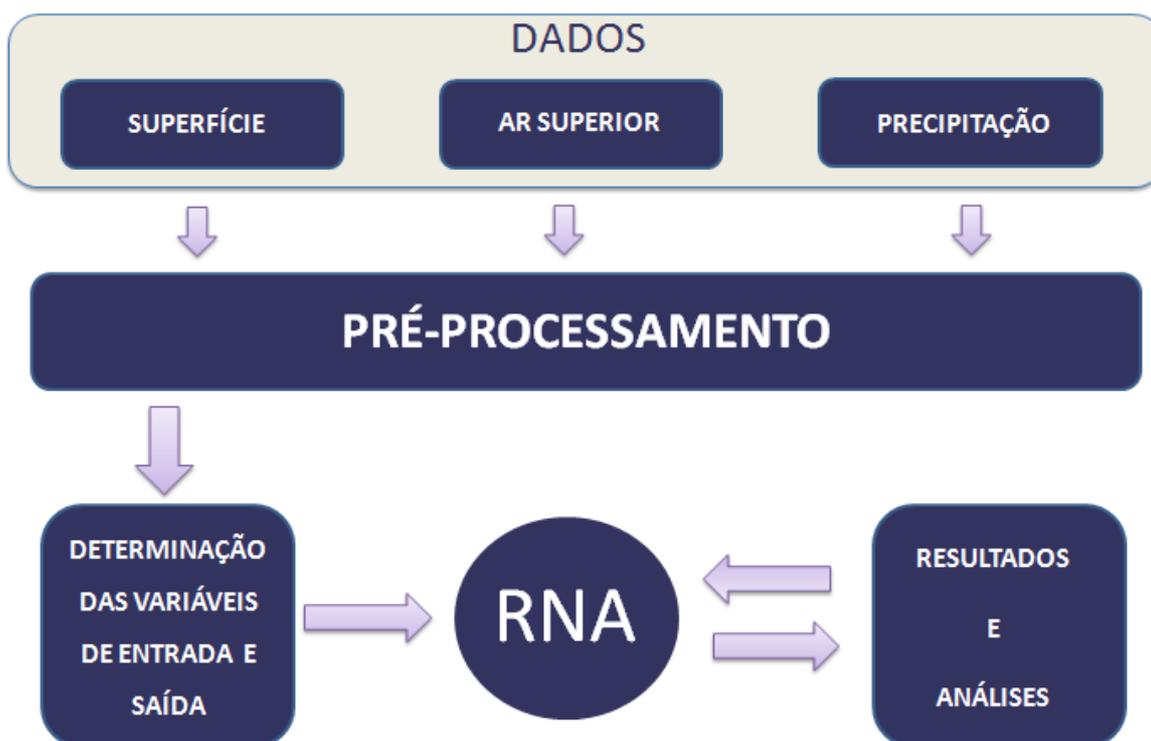


Figura 4.1 – Fluxograma da metodologia aplicada à previsão de curtíssimo prazo de eventos extremos para o aeroporto internacional do Galeão – RJ.

1º PASSO: Coleta e análise preliminar dos dados (conforme itens 3.1 e 3.2);

2º PASSO: Pré-processamento dos dados (apresentado no item 3.3);

3º PASSO: Determinação das variáveis de entrada e saída (itens 3.4 e 3.5);

4º PASSO: Aplicação da RNA.

Na aplicação da RNA foram utilizadas as redes PNN treinadas com o programa *NeuroShell Classifier*, o qual tem a capacidade de resolver problemas de classificação, por meio do algoritmo *Cascade Correlation* (CC).

Para aplicação dos dados selecionados e pré-processados na RNA é necessária a formação dos conjuntos de treinamento e validação. Neste contexto, os dados foram divididos de modo a contar cerca de 70% para o treinamento da rede e 30% para a validação da mesma, conforme ilustra a Figura 4.2.

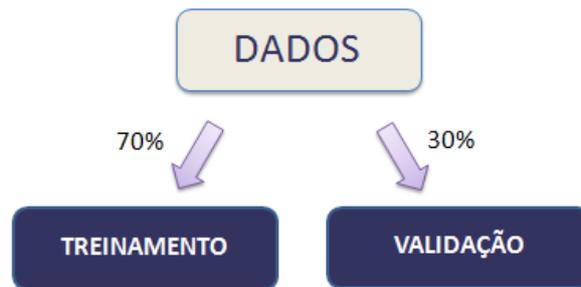


Figura 4.2 – Ilustração do método de divisão dos dados da rede.

Validação – Para seleção de um “bom” arquivo de validação da rede, é necessário que o mesmo retrate um quadro real da atmosfera. Portanto os dados escolhidos para validação da rede deverão refletir a realidade, ou seja, devem estar equivalentes ao que foi registrado durante toda a série de dados do arquivo geral. Para isso, foi feita uma análise comparando a quantidade de cada “faixa de tempo” entre o arquivo contendo 100% dos dados e os 30% retirados para validar a rede, conforme mostram as Figuras 4.3 (a) e (b).

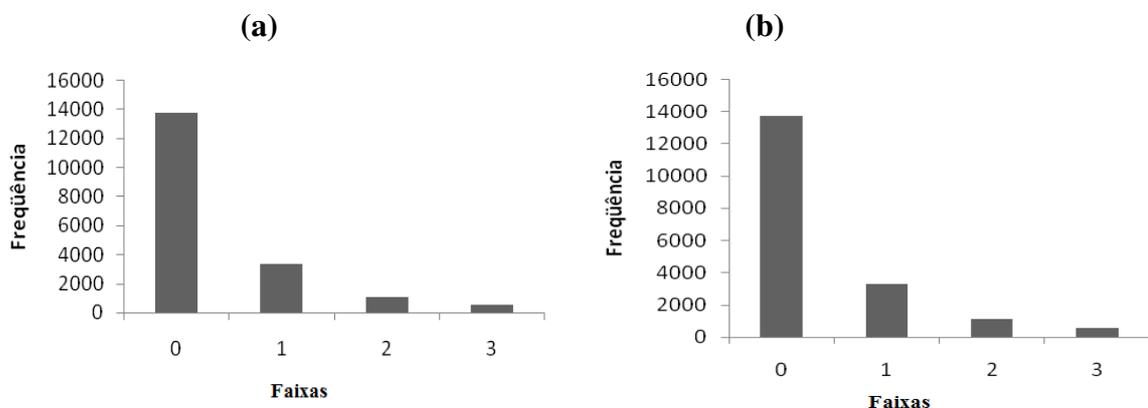


Figura 4.3 – (a) Histograma de frequência das faixas de tempo constantes no arquivo geral (treinamento e validação); (b) frequência das faixas de tempo do arquivo de validação.

Verifica-se, portanto, que as frequências das “faixas de tempo” presentes nos arquivos analisados (arquivo geral e arquivo de validação) apresentam-se de forma harmônica, respeitando as devidas proporções. Isso mostra que o arquivo de validação da rede configura a realidade dos dados encontrados no arquivo geral.

Treinamento – Foram gerados oito conjuntos de treinamento da RNA, conforme o exposto no Capítulo 5, páginas 61 – 63.

5° PASSO: Resultados e análises.

Neste processo foram gerados, a partir do arquivo validação, resultados de previsões para uma, duas e três horas futuras. Esses resultados foram analisados através de medidas de avaliação do desempenho até obter um resultado considerado “ótimo”.

Segundo Wilks (2006), essas medidas de avaliação do desempenho podem ser derivadas de uma tabela de contingência 2x2. Uma adaptação da tabela de contingência idealizada por Wilks é apresentada a seguir na Tabela 4.1.

Tabela 4.1 – Tabela de contingência 2x2 adaptada (WILKS, 2006).

Previsto	Observado		Total de eventos/não-eventos previstos
	NÃO-EVENTO	EVENTO	
NÃO-EVENTO	a	b	a+b
EVENTO	c	d	c+d
Total de eventos/não-eventos observados	a+c	b+d	n = a+b+c+d

Os valores representados por a, b, c e d na Tabela 4.1 são:

- a – representa o número de acertos do não-evento;
- b – representa o número de eventos observados que não foram previstos;
- c – representa o número de eventos previstos que não foram observados;
- d – representa o número de acertos do evento.

Ao analisar um modelo de previsão, percebe-se que podem ocorrer erros de precisão, ou mesmo de interpretação na detecção de um evento meteorológico específico. Quando o resultado for positivo para determinado evento, pode haver a ocorrência do evento (verdadeiro-positivo – VP), como pode não haver (falso-positivo – FP). Da mesma forma, quando o resultado for negativo, pode não ocorrer o evento (verdadeiro-negativo – VN), ou ocorrer (falso-negativo – FN). Tais resultados podem ser expressos em valores absolutos, representados pelas letras **d**, **b**, **a** e **c**, respectivamente, ou através de valores relativos conforme mostram as Equações 4.1 (a) – (d):

$$VP = \frac{d}{b+d} \quad (4.1.a)$$

$$FP = \frac{c}{a+c} \quad (4.1.b)$$

$$VN = \frac{a}{a+c} \quad (4.1.c)$$

$$FN = \frac{b}{b+d} \quad (4.1.d)$$

Os valores de VP e VN indicam os índices de acerto do modelo, portanto, quando esses valores são iguais a um configura-se um ótimo resultado, enquanto que valores iguais a zero indicam resultado péssimo. Já, os valores de FP e FN, indicam os índices de erro do modelo, devendo, portanto, tender a zero.

A partir desses conceitos serão definidas as medidas de verificação do desempenho que foram geradas automaticamente após a aplicação da RNA. São elas:

- *Sensibilidade* – mede a capacidade do modelo em identificar corretamente um determinado evento entre aqueles que possuem, ou seja, quão sensível é o modelo. Equivale a probabilidade de detecção do modelo – POD (*Probability of Detection*). Essa medida proporciona a taxa de previsões corretas do evento.

- *Especificidade* – mede a capacidade do modelo em excluir corretamente aqueles que não são o evento a ser previsto, ou seja, o quão específico o modelo é.

A sensibilidade é a fração das respostas positivas do modelo entre todos os eventos. Esta fração pode ser descrita pela Equação 4.2:

$$Sensibilidade = \frac{VP}{VP + FN} \quad (4.2)$$

A especificidade é a fração das respostas negativas do modelo entre todos os não-eventos. Esta fração pode ser descrita pela Equação 4.3:

$$Especificidade = \frac{VN}{VN + FP} \quad (4.3)$$

Um bom teste possui um alto valor para a sensibilidade e para a especificidade, pois ele identificará corretamente os eventos e os não-eventos. Testes com alta sensibilidade e baixa especificidade emitem alarme falso (falso-positivo), superestimando o resultado positivo. Em contra partida, testes com baixa sensibilidade e alta especificidade, erram muito, causando o efeito do falso-negativo e consequentemente subestimando o resultado positivo. Na Tabela 4.2 é apresentado um resumo sobre o desempenho de um modelo de previsão em relação à sensibilidade e à especificidade.

Tabela 4.2 – Descrição do desempenho de um modelo de previsão em relação à sensibilidade e à especificidade (ALMEIDA, 2009).

Sensibilidade	Especificidade	Desempenho
Alta	Alta	Ótimo
Alta	Baixa	Superestimação
Baixa	Alta	Subestimação
Baixa	Baixa	Péssimo

Além das medidas de desempenho geradas pela rede, existem outras duas medidas que foram propostas por Wilks (2006) e adaptadas de acordo com a Tabela 4.1. São elas:

- *Bias* – É a comparação da previsão com a observação. É representado pela razão entre número de previsões do evento e as observações do evento. Previsões empíricas apresentam *Bias* igual a 1, indicando que o evento foi previsto o mesmo número de vezes que foi observado. Valores maiores que 1 indicam previsões superestimadas e menores que 1, previsões subestimadas. A Equação 4.4 mostra o preceito desta medida.

$$B = \frac{d + c}{d + b} \quad (4.4)$$

- *POFD* – Probabilidade de falsa detecção (*probability of false detection*). É a razão do **alarme falso** pelo total de não-eventos observados. Nesse caso, valores mais próximos de zero tendem a resultados ótimos enquanto que valores próximos de um tendem a resultados péssimos. Essa medida equivale ao valor do falso-positivo, conforme mostra a Equação 4.5.

$$POFD = \frac{c}{c + a} \quad (4.5)$$

Capítulo 5 – RESULTADOS E ANÁLISES

Para a produção dos resultados, formaram-se três variáveis de saída para a RNA:

- VALID_1H – Resultado gerado pelo conjunto de validação da rede para as quatro faixas de tempo na previsão da primeira hora;
- VALID_2H – Resultado gerado pelo conjunto de validação da rede para as quatro faixas de tempo na previsão da segunda hora;
- VALID_3H – Resultado gerado pelo conjunto de validação da rede para as quatro faixas de tempo na previsão da terceira hora.

A partir dos 70% de dados que restaram após a formação do conjunto de validação da rede, foi criado o conjunto de treinamento, a partir do qual a rede foi treinada. Com o conjunto de validação da rede foram gerados resultados no objetivo de fazer previsões das faixas de tempo para a primeira, segunda e terceira horas, a fim de comparar o desempenho obtido pela rede entre eventos variados e o evento que esta pesquisa pretende prever (faixa “3”). Para cada faixa obteve-se resultados diferentes, expostos nas Tabelas 5.1, 5.2 e 5.3.

Tabela 5.1 – Teste de previsão para a primeira hora das quatro faixas de tempo.

VALID_1H	Faixa "0"	Faixa "1"	Faixa "2"	Faixa "3"	Total
Classificado como "0"	13128	880	82	108	14198
Classificado como "1"	569	2055	401	64	3089
Classificado como "2"	10	301	502	118	931
Classificado como "3"	23	67	110	278	478
Total	13730	3303	1095	568	18696
Verdadeiro-positivo	0,9562	0,6222	0,4584	0,4894	
Falso-positivo	0,2155	0,0672	0,0244	0,011	
Verdadeiro-negativo	0,7845	0,9328	0,9756	0,989	
Falso-negativo	0,0438	0,3778	0,5416	0,5106	
Sensibilidade	95,62%	62,22%	45,84%	48,94%	
Especificidade	78,45%	93,28%	97,56%	98,90%	

Tabela 5.2 – Teste de previsão para a segunda hora das quatro faixas de tempo.

VALID_2H	Faixa "0"	Faixa "1"	Faixa "2"	Faixa "3"	Total
Classificado como "0"	13067	1139	194	238	14638
Classificado como "1"	647	1951	633	185	3416
Classificado como "2"	11	194	260	117	582
Classificado como "3"	4	19	8	28	59
Total	13729	3303	1095	568	18695
Verdadeiro-positivo	0,9518	0,5907	0,2374	0,0493	
Falso-positivo	0,3164	0,0952	0,0183	0,0017	
Verdadeiro-negativo	0,6836	0,9048	0,9817	0,9983	
Falso-negativo	0,0482	0,4093	0,7626	0,9507	
Sensibilidade	95,18%	59,07%	23,74%	4,93%	
Especificidade	68,36%	90,48%	98,17%	99,83%	

Tabela 5.3 – Teste de previsão para a terceira hora das quatro faixas de tempo.

TESTE_0_3h	Faixa "0"	Faixa "1"	Faixa "2"	Faixa "3"	Total
Classificado como "0"	13070	1367	261	304	15002
Classificado como "1"	646	1843	724	199	3412
Classificado como "2"	12	93	110	64	279
Classificado como "3"	0	0	0	1	1
Total	13728	3303	1095	568	18694
Verdadeiro-positivo	0,9521	0,558	0,1005	0,0018	
Falso-positivo	0,389	0,1019	0,0096	0	
Verdadeiro-negativo	0,611	0,8981	0,9904	1	
Falso-negativo	0,0479	0,442	0,8995	0,9982	
Sensibilidade	95,21%	55,80%	10,05%	0,18%	
Especificidade	61,10%	89,81%	99,04%	100%	

Os resultados mostram que a ocorrência frequente da faixa de tempo “0” nos dados de entrada da rede fez com que ela “*aprendesse mais*” a existência desse tipo de

evento em relação aos outros, causando um efeito de superestimação da faixa de tempo “0”. Por outro lado, a baixa frequência de eventos com chuva forte e/ou trovoadas (faixa “3”) nos dados de entrada da rede fez com que a rede “*aprendesse menos*” a existência de eventos deste tipo, causando um efeito de subestimação da faixa “3”.

Na Figura 5.1 são apresentados os resultados para a 1ª, 2ª e 3ª horas de previsão de cada faixa de tempo. Esses resultados foram obtidos através de uma média aritmética entre as seguintes medidas de avaliação do desempenho:

- Sensibilidade;
- Especificidade.

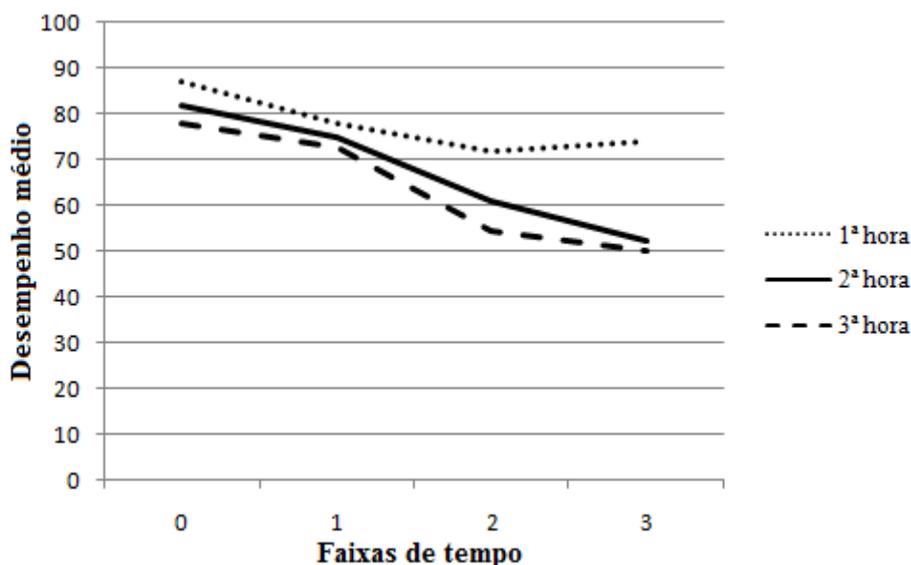


Figura 5.1 – Desempenho médio da Rede Neural, avaliando sensibilidade e especificidade de cada faixa de tempo para primeira, segunda e terceira hora de previsão.

Para resolver a problemática da superestimação da faixa “0” e a subestimação da faixa “3”, foram desenvolvidos oito conjuntos de treinamento da rede, gerados a partir do arquivo original de treinamento (70% dos dados). Nesses novos conjuntos foram acrescentados (duplicados) eventos da faixa “3” e subtraídos eventos da faixa “0”. Com isso, a rede “aprenderá um pouco mais” a faixa “3” em relação a faixa “0”. A composição de cada faixa de tempo nos conjuntos de treinamento da rede pode ser observada nas Figuras 5.2 (a) – (h).

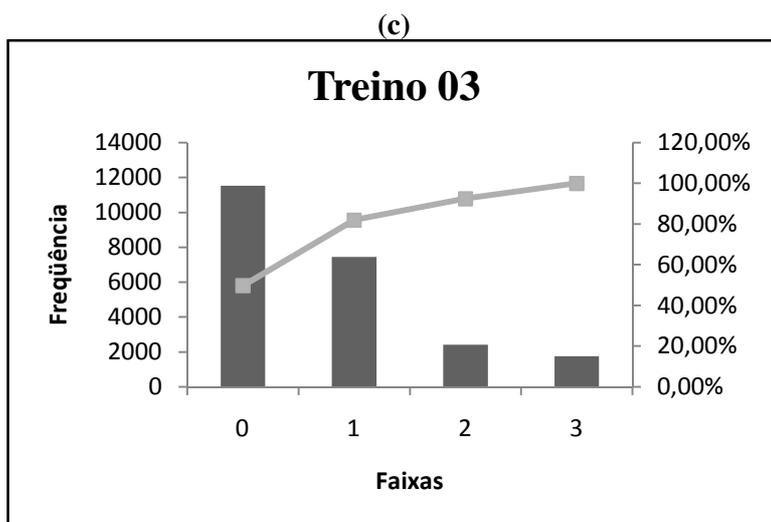
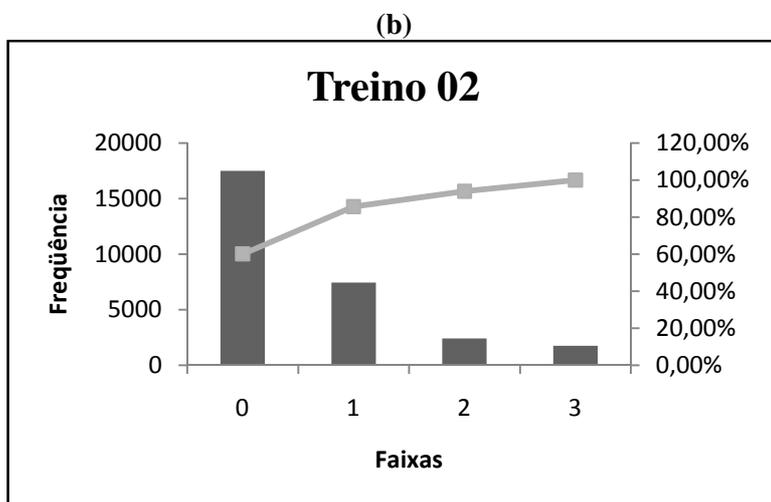
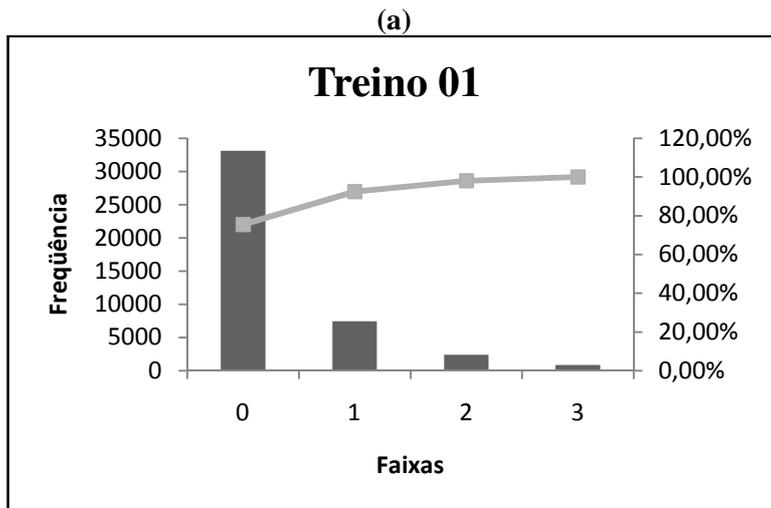


Figura 5.2 – Frequência e porcentagem cumulativa das faixas de tempo dos conjuntos de treinamento: (a) Treino 01; (b) Treino 02; (c) Treino 03. Os eventos meteorológicos são relativos ao período de janeiro de 1997 a dezembro de 2007.

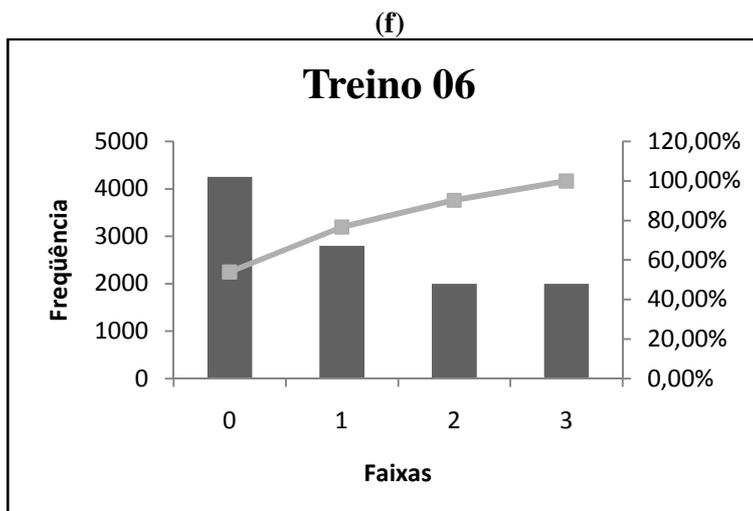
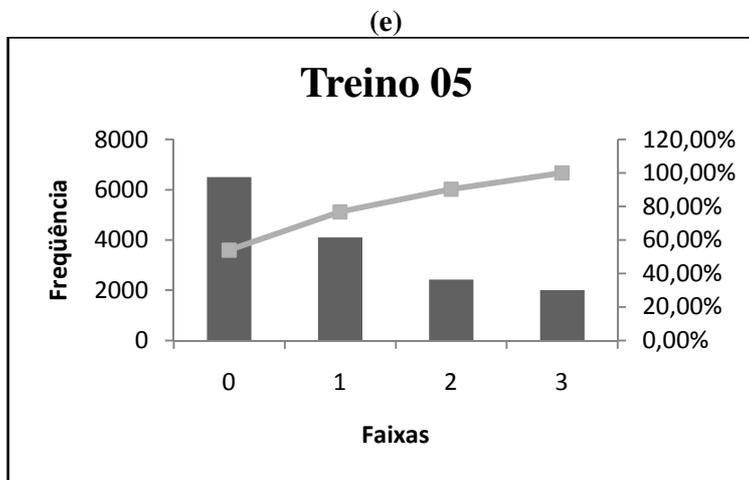
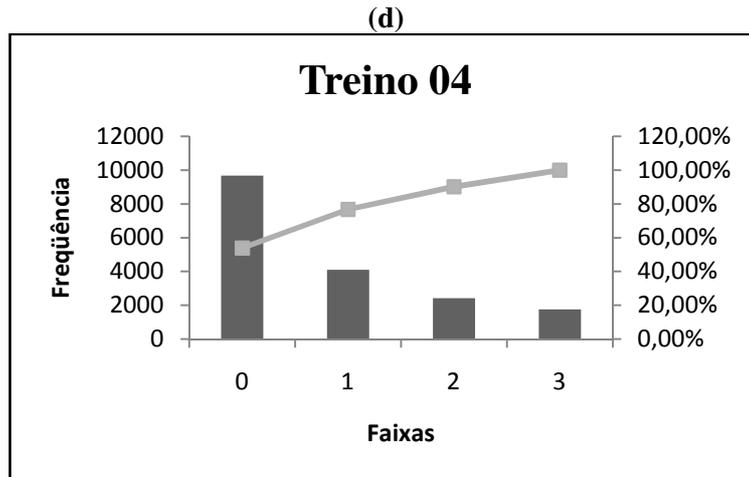


Figura 5.2 – Frequência e porcentagem cumulativa das faixas de tempo dos conjuntos de treinamento: (d) Treino 04; (e) Treino 05 e (f) Treino 06. Os eventos meteorológicos são relativos ao período de janeiro de 1997 a dezembro de 2007.

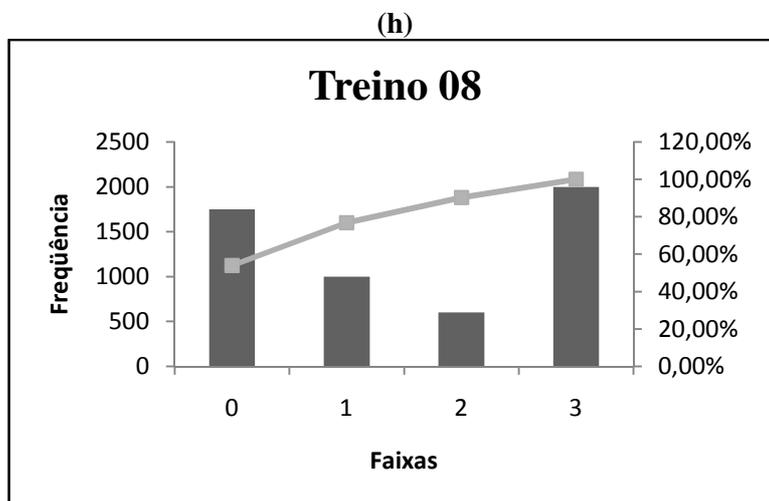
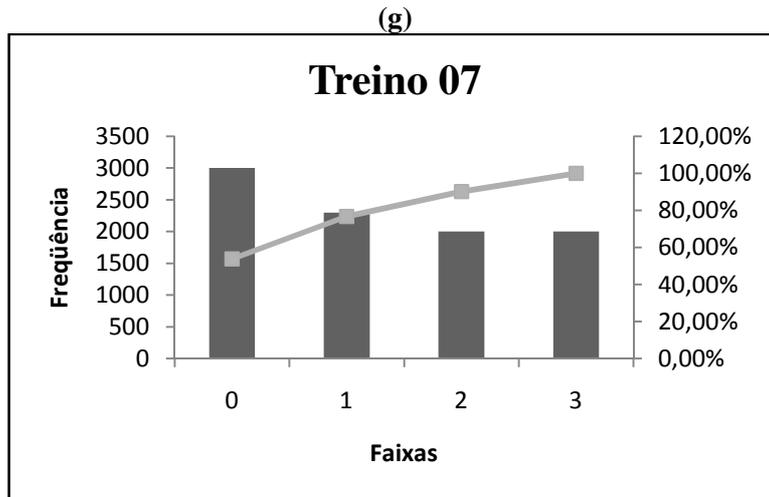


Figura 5.2 – Frequência e porcentagem cumulativa das faixas de tempo dos conjuntos de treinamento: (g) Treino 07; (h) Treino 08. Os eventos meteorológicos são relativos ao período de janeiro de 1997 a dezembro de 2007.

A partir deste momento, serão feitos treinamentos apenas para a condição meteorológica que se pretende prever: **ocorrência de trovoadas e/ou chuva forte**. Portanto, a variável de saída terá duas classes, respeitando o estabelecido no item 3.3.3, as quais são:

- “0” – para quaisquer outros eventos em que não ocorra chuva forte e/ou trovoadas, ou seja, faixas 0, 1 e 2 (NÃO-EVENTO);
- “1” – para ocorrência de chuva forte e/ou trovoadas, ou seja, faixa 3 (EVENTO).

RESULTADOS PARA 1ª HORA

Tabela 5.4 – Quadro de tabelas de contingência relativas aos resultados da RNA sobre o conjunto de validação para primeira hora.

TESTE 01				TESTE 05			
		Observado				Observado	
Previsto	"0"	"1"	Total	Previsto	"0"	"1"	Total
"0"	17561	227	17788	"0"	15572	110	15682
"1"	298	333	631	"1"	2287	450	2737
Total	17859	560	18419	Total	17859	560	18419
Verd-pos.	0,9833	0,5946		Verd-pos.	0,8719	0,8036	
Falso-pos.	0,4054	0,0167		Falso-pos.	0,1964	0,1281	
Verd-neg.	0,5946	0,9833		Verd-neg.	0,8036	0,8719	
Falso-neg.	0,0167	0,4054		Falso-neg.	0,1281	0,1964	
Sensibilidade	98,33%	59,46%		Sensibilidade	87,19%	80,36%	
Especificidade	59,46%	98,33%		Especificidade	80,36%	87,19%	
TESTE 02				TESTE 06			
		Observado				Observado	
Previsto	"0"	"1"	Total	Previsto	"0"	"1"	Total
"0"	17191	181	17372	"0"	15173	100	15273
"1"	668	379	1047	"1"	2686	460	3146
Total	17859	560	18419	Total	17859	560	18419
Verd-pos.	0,9626	0,6768		Verd-pos.	0,8496	0,8214	
Falso-pos.	0,3232	0,0374		Falso-pos.	0,1786	0,1504	
Verd-neg.	0,6768	0,9626		Verd-neg.	0,8214	0,8496	
Falso-neg.	0,0374	0,3232		Falso-neg.	0,1504	0,1786	
Sensibilidade	96,26%	67,68%		Sensibilidade	84,96%	82,14%	
Especificidade	67,68%	96,26%		Especificidade	82,14%	84,96%	
TESTE 03				TESTE 07			
		Observado				Observado	
Previsto	"0"	"1"	Total	Previsto	"0"	"1"	Total
"0"	17561	227	17788	"0"	14833	94	14927
"1"	298	333	631	"1"	3026	466	3492
Total	17859	560	18419	Total	17859	560	18419
Verd-pos.	0,9311	0,7625		Verd-pos.	0,8306	0,8321	
Falso-pos.	0,2375	0,0689		Falso-pos.	0,1679	0,1694	
Verd-neg.	0,7625	0,9311		Verd-neg.	0,8321	0,8306	
Falso-neg.	0,0689	0,2375		Falso-neg.	0,1694	0,1679	
Sensibilidade	93,11%	76,25%		Sensibilidade	83,06%	83,21%	
Especificidade	76,25%	93,11%		Especificidade	83,21%	83,06%	
TESTE 04				TESTE 08			
		Observado				Observado	
Previsto	"0"	"1"	Total	Previsto	"0"	"1"	Total
"0"	16047	124	16171	"0"	14582	93	14675
"1"	1812	436	2248	"1"	3277	467	3744
Total	17859	560	18419	Total	17859	560	18419
Verd-pos.	0,8985	0,7786		Verd-pos.	0,8165	0,8339	
Falso-pos.	0,2214	0,1015		Falso-pos.	0,1661	0,1835	
Verd-neg.	0,7786	0,8985		Verd-neg.	0,8339	0,8165	
Falso-neg.	0,1015	0,2214		Falso-neg.	0,1835	0,1661	
Sensibilidade	89,85%	77,86%		Sensibilidade	81,65%	83,39%	
Especificidade	77,86%	89,85%		Especificidade	83,39%	81,65%	

Ao avaliar a sensibilidade e especificidade, obteve-se o exposto na Figura 5.3 para a primeira hora de previsão.

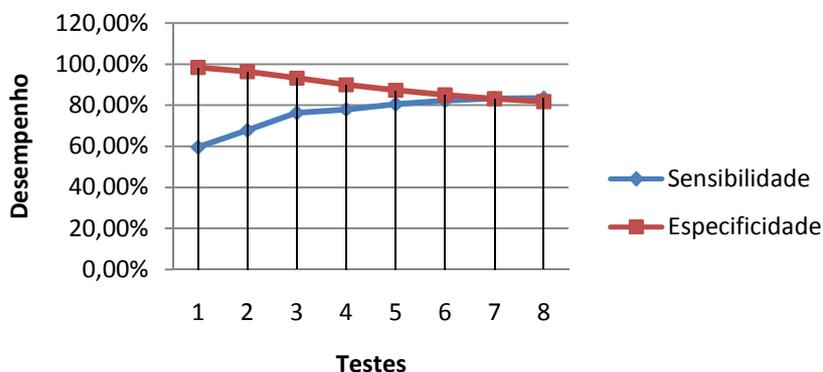


Figura 5.3– Gráfico de análise da sensibilidade e especificidade para 1ª hora.

O resultado gerado pelo Teste 07 mostrou alto desempenho quanto a sensibilidade (83,21%) e a especificidade (83,06%). Ao analisar outros métodos de avaliação do desempenho para o mesmo resultado, obteve-se:

$$B = 6,24 \qquad \text{POFD} = 0,17$$

Onde PODF representa uma medida de erro, devendo, portanto, ser o mais próximo de zero, e B igual a 6,24 denota que a previsão foi superestimada em aproximadamente seis vezes.

Uma análise no tocante a precipitação, mostrou que em 38% dos casos em que foi prevista chuva forte e/ou trovoada, mas não houveram registros dessa natureza no AIRJ, foi verificada a ocorrência de precipitação moderada entre 8 mm/h a 9,9 mm/h em pelo menos uma das estações pluviométricas estudadas (próximas ao local de estudo). A aplicação dessa possível margem de acerto (38%) nos resultados, certamente diminuiria a *bias*, mas, como o objetivo desta pesquisa está exatamente na previsão de eventos **extremos**, essa análise serviu apenas para mostrar que o modelo registrou muito bem o estado de instabilidade da atmosfera, mesmo quando em alguns casos tenha sido considerado o erro de superestimação do modelo.

RESULTADOS PARA 2ª HORA

Tabela 5.5 – Quadro de tabelas de contingência relativas aos resultados da RNA sobre o conjunto de validação para segunda hora.

TESTE 01		Observado		Total
Previsto	“0”	“1”		
“0”	17660	277	17937	
“1”	199	283	482	
Total	17859	560	18419	
Verd-pos.	0,9889	0,5054		
Falso-pos.	0,4946	0,0111		
Verd-neg.	0,5054	0,9889		
Falso-neg.	0,0111	0,4946		
Sensibilidade	98,89%	50,54%		
Especificidade	50,54%	98,89%		

TESTE 02		Observado		Total
Previsto	“0”	“1”		
“0”	17635	252	17887	
“1”	224	308	532	
Total	17859	560	18419	
Verd-pos.	0,9875	0,55		
Falso-pos.	0,45	0,0125		
Verd-neg.	0,55	0,9875		
Falso-neg.	0,0125	0,45		
Sensibilidade	98,75%	55,00%		
Especificidade	55,00%	98,75%		

TESTE 03		Observado		Total
Previsto	“0”	“1”		
“0”	17439	205	17644	
“1”	420	355	775	
Total	17859	560	18419	
Verd-pos.	0,9765	0,6339		
Falso-pos.	0,3661	0,0235		
Verd-neg.	0,6339	0,9765		
Falso-neg.	0,0235	0,3661		
Sensibilidade	97,65%	63,39%		
Especificidade	63,39%	97,65%		

TESTE 04		Observado		Total
Previsto	“0”	“1”		
“0”	16974	181	15806	
“1”	885	379	2613	
Total	17859	560	18419	
Verd-pos.	0,9504	0,6768		
Falso-pos.	0,3232	0,0496		
Verd-neg.	0,6768	0,9504		
Falso-neg.	0,0496	0,3232		
Sensibilidade	95,04%	67,68%		
Especificidade	67,68%	95,04%		

TESTE 05		Observado		Total
Previsto	“0”	“1”		
“0”	15652	154	15806	
“1”	2207	406	2613	
Total	17859	560	18419	
Verd-pos.	0,8764	0,725		
Falso-pos.	0,275	0,1236		
Verd-neg.	0,725	0,8764		
Falso-neg.	0,1236	0,275		
Sensibilidade	87,64%	72,50%		
Especificidade	72,50%	87,64%		

TESTE 06		Observado		Total
Previsto	“0”	“1”		
“0”	15112	101	15213	
“1”	2747	459	3206	
Total	17859	560	18419	
Verd-pos.	0,8462	0,8196		
Falso-pos.	0,1804	0,1538		
Verd-neg.	0,8196	0,8462		
Falso-neg.	0,1538	0,1804		
Sensibilidade	84,62%	81,96%		
Especificidade	81,96%	84,62%		

TESTE 07		Observado		Total
Previsto	“0”	“1”		
“0”	14211	87	14298	
“1”	3648	473	4121	
Total	17859	560	18419	
Verd-pos.	0,7957	0,8446		
Falso-pos.	0,1554	0,2043		
Verd-neg.	0,8446	0,7957		
Falso-neg.	0,2043	0,1554		
Sensibilidade	79,57%	84,46%		
Especificidade	84,46%	79,57%		

TESTE 08		Observado		Total
Previsto	“0”	“1”		
“0”	13344	81	13425	
“1”	4515	479	4994	
Total	17859	560	18419	
Verd-pos.	0,7472	0,8554		
Falso-pos.	0,1446	0,2528		
Verd-neg.	0,8554	0,7472		
Falso-neg.	0,2528	0,1446		
Sensibilidade	74,72%	85,54%		
Especificidade	85,54%	74,72%		

Ao avaliar a sensibilidade e a especificidade para a segunda hora de previsão obteve-se o resultado exposto na Figura 5.4.

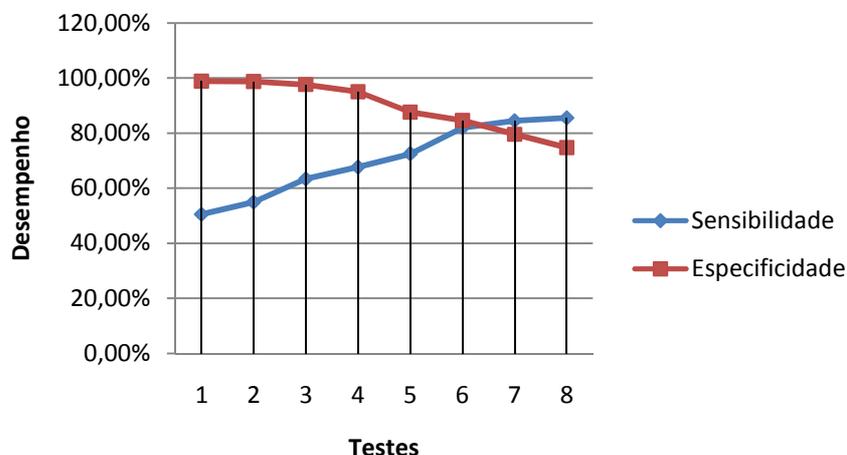


Figura 5.4– Gráfico de análise da sensibilidade e especificidade para 2ª hora.

A análise do resultado gerado pelo Teste 06 para a 2ª hora de previsão, onde o índice de sensibilidade atingiu 81,96% e especificidade 84,62%, manifesta o alto desempenho da RNA também para a segunda hora futura. Os métodos de avaliação do desempenho propostos por Wilks também foram analisados para o Teste 06, obtendo-se:

$$B = 5,73$$

$$POFD = 0,15$$

Bias igual a 5,73 indica que a rede superestimou em aproximadamente seis vezes o resultado. A medida de erro POFD igual a 0,15 indica um índice de erro próximo de zero, o que seria o ideal.

Na análise dos dados de precipitação da Geo-Rio foi confirmado que em 19% dos casos em que o modelo disse ter o evento quando o mesmo não ocorreu, houve precipitação moderada na faixa de 8 mm/h a 9,9 mm/h. Isso mostra que ao estender a faixa de precipitação, o valor do *bias* diminuiria.

RESULTADOS PARA 3ª HORA

Tabela 5.6 – Quadro de tabelas de contingência relativas aos resultados da RNA sobre o conjunto de validação para a terceira hora futura.

TESTE 01		Observado		
Previsto	“0”	“1”	Total	
“0”	17836	513	18349	
“1”	23	47	70	
Total	17859	560	18419	
Verd-pos.	0,9987	0,0839		
Falso-pos.	0,9161	0,0013		
Verd-neg.	0,0839	0,9987		
Falso-neg.	0,0013	0,9161		
Sensibilidade	99,87%	8,39%		
Especificidade	28,60%	97,73%		

TESTE 02		Observado		
Previsto	“0”	“1”	Total	
“0”	17782	478	18260	
“1”	77	82	159	
Total	17859	560	18419	
Verd-pos.	0,9957	0,1464		
Falso-pos.	0,8536	0,0043		
Verd-neg.	0,1464	0,9957		
Falso-neg.	0,0043	0,8536		
Sensibilidade	99,57%	14,64%		
Especificidade	14,64%	99,57%		

TESTE 03		Observado		
Previsto	“0”	“1”	Total	
“0”	17601	375	17976	
“1”	258	185	443	
Total	17859	560	18419	
Verd-pos.	0,9856	0,3304		
Falso-pos.	0,6696	0,0144		
Verd-neg.	0,3304	0,9856		
Falso-neg.	0,0144	0,6696		
Sensibilidade	98,56%	33,04%		
Especificidade	33,04%	98,56%		

TESTE 04		Observado		
Previsto	“0”	“1”	Total	
“0”	17459	347	17806	
“1”	400	213	613	
Total	17859	560	18419	
Verd-pos.	0,9776	0,3804		
Falso-pos.	0,6196	0,0224		
Verd-neg.	0,3804	0,9776		
Falso-neg.	0,0224	0,6196		
Sensibilidade	97,76%	38,04%		
Especificidade	38,04%	97,76%		

TESTE 05		Observado		
Previsto	“0”	“1”	Total	
“0”	16785	237	17022	
“1”	1074	323	1397	
Total	17859	560	18419	
Verd-pos.	0,9399	0,5768		
Falso-pos.	0,4232	0,0601		
Verd-neg.	0,5768	0,9399		
Falso-neg.	0,0601	0,4232		
Sensibilidade	93,99%	57,68%		
Especificidade	57,68%	93,99%		

TESTE 06		Observado		
Previsto	“0”	“1”	Total	
“0”	16242	195	16437	
“1”	1617	365	1982	
Total	17859	560	18419	
Verd-pos.	0,9095	0,6518		
Falso-pos.	0,3482	0,0905		
Verd-neg.	0,6518	0,9095		
Falso-neg.	0,0905	0,3482		
Sensibilidade	90,95%	65,18%		
Especificidade	65,18%	90,95%		

TESTE 07		Observado		
Previsto	“0”	“1”	Total	
“0”	14302	143	14445	
“1”	3557	417	3974	
Total	17859	560	18419	
Verd-pos.	0,8008	0,7446		
Falso-pos.	0,2554	0,1992		
Verd-neg.	0,7446	0,8008		
Falso-neg.	0,1992	0,2554		
Sensibilidade	80,08%	74,46%		
Especificidade	74,46%	80,08%		

TESTE 08		Observado		
Previsto	“0”	“1”	Total	
“0”	14100	107	14207	
“1”	3759	453	4212	
Total	17859	560	18419	
Verd-pos.	0,7895	0,8089		
Falso-pos.	0,1911	0,2105		
Verd-neg.	0,8089	0,7895		
Falso-neg.	0,2105	0,1911		
Sensibilidade	78,95%	80,89%		
Especificidade	80,89%	78,95%		

A Figura 5.5 apresenta uma análise ilustrativa relacionando as duas medidas de avaliação do desempenho geradas pela RNA (sensibilidade e especificidade).

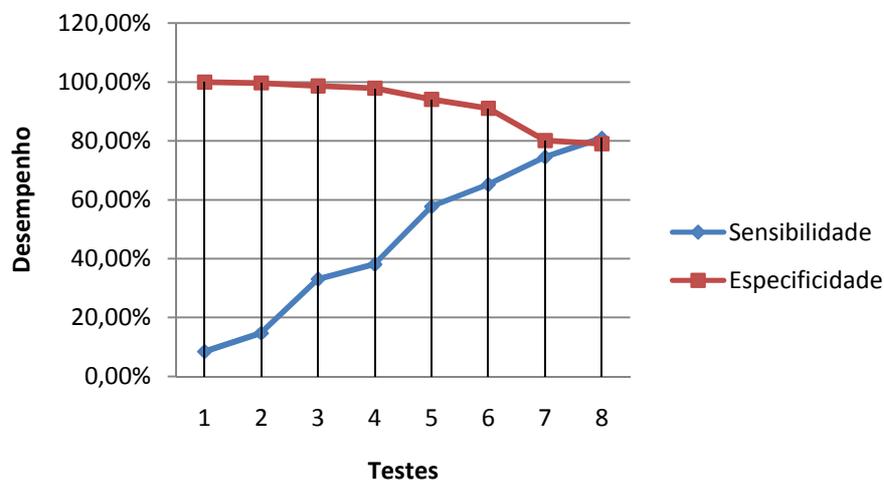


Figura 5.5 – Gráfico análise da sensibilidade e especificidade para 3ª hora.

O melhor resultado obtido para previsão da 3ª hora futura foi gerado pelo conjunto de treinamento 08, onde sensibilidade e especificidade apresentaram valores iguais a 80,89% e 78,95%, respectivamente. A análise das outras medidas de desempenho utilizadas por esta pesquisa no resultado obtido do Teste 08 apresentou:

$$B = 7,72$$

$$POFD = 0,21$$

A análise dos dados de precipitação mostrou que em cerca de 14% dos resultados superestimados do modelo houve a ocorrência de chuva moderada na faixa de 8mm/h a 9,9 mm/h. Denotando, um alto desempenho do modelo quanto a identificação de uma condição instável da atmosfera.

Um resumo dos resultados preliminares obtidos dos testes de melhores resultados desta seção para as três horas de previsão são apresentados na Tabela 5.7. Nela estão colocados os valores obtidos pelas seguintes medidas de desempenho: sensibilidade, especificidade, bias, POD e POFD.

Tabela 5.7 – Análise do desempenho da rede para as três horas de previsão.

Medidas de avaliação do desempenho	1ª Hora	2ª Hora	3ª Hora
Sensibilidade	0,83	0,82	0,81
Especificidade	0,83	0,85	0,80
Bias	6,24	5,73	7,72
POD	0,83	0,82	0,81
POFD	0,17	0,15	0,21

É evidente que para as três horas de previsão as medidas de sensibilidade, especificidade e POFD apresentaram índices satisfatórios relatando um alto desempenho do modelo para a previsão do evento em questão. Mas por outro lado, as “bias” mostraram que o modelo superestimou o evento a ser previsto.

Neste momento, é importante lembrar que, os registros dos dados utilizados para gerar os resultados obtidos foram feitos por operadores de estações meteorológicas, o que pode acarretar em erros. Principalmente a respeito da visibilidade e tempo presente, variáveis que são estimadas pelos próprios operados de EMS. Além disso, os dados utilizados são pontuais, ou seja, denotam a condição atmosférica de um local em um determinado momento. Por isso, o ideal seria que todas as variáveis fossem registradas de forma contínua e sem estar expostas a erros. Para tanto, será feito o uso de dados que represente a instabilidade de forma contínua, através dos registros de descargas atmosféricas obtidos do sistema de detecção de raios da Rede Integrada Nacional de Detecção de Descargas Atmosféricas (RINDAT), o que irá gerar novos resultados.

5.1 Avaliação da RNA com dados de raios

Segundo Valença (2005), como em qualquer modelo de previsão, as variáveis de entrada possuem uma grande importância no desempenho final do modelo. Entretanto, os dados de raio utilizados para gerar os resultados exibidos anteriormente, foram obtidos através da **observação** da ocorrência ou não desse fenômeno, dada por operadores de EMS, o que introduz a erro posto que no momento da observação o fenômeno pode não

ter ocorrido, mas sim minutos/segundos antes ou após, ou também, o observador pode não ter avistado relâmpago ou ouvido o trovão enquanto os mesmos tenham ocorrido.

Portanto, resolver essa problemática tornou-se um novo objetivo a ser alcançado por esta pesquisa. Para tanto, foi considerado que a rede indicou a condição de instabilidade da atmosfera apontada pelos registros de raios obtidos através do sistema de detecção de descargas atmosféricas da RINDAT (e não pelos registros de operadores de EMS) como acerto da rede.

A análise dos dados da RINDAT mostrou que houveram registros de raios na região de estudo (vide Figura 3.1) em 11,66% das previsões superestimadas para a primeira hora, apesar de não ter sido registrado nos METARes. Para segunda hora esse número chegou a 10,13% e 9,26% para a terceira hora.

Uma comparação entre os resultados que utilizaram apenas os registros do METAR e os resultados que utilizaram dados de raios da RINDAT segue da seguinte forma:

RESULTADOS PARA 1ª HORA

O resultado gerado através do conjunto de treinamento 07, apresentou melhor desempenho para a primeira hora futura, configurando-se conforme exibe a Tabela 5.8.

Tabela 5.8 – Resumo do melhor resultado que utilizou os registros de raios gerados apenas por operadores de EMS, para a primeira hora de previsão.

TESTE 08_1H (RESUMO)	Observado		Total
	"0"	"1"	
Previsto			
"0"	14833	94	14927
"1"	3026	466	3492
Total	17859	560	18419

Ao considerar o exposto na Tabela 5.8, foram calculadas as medidas de desempenho obtendo-se:

- Sensibilidade = 0,83;
- Especificidade = 0,83;

- Bias = 6,24;
- POFD = 0,17.

Supondo que a rede indicou a condição de instabilidade da atmosfera apontada pelos raios registrados (e não apenas observados) como acerto da rede. Na 1ª hora a rede acertaria 11,66%, dos 3026 casos superestimados. Obtendo-se, assim, um novo resultado, mostrado na Tabela 5.9.

Tabela 5.9 – Resumo do resultado utilizando os registros de raios gerados por operadores de EMS junto com os registros de raios da RINDAT, para a primeira hora de previsão.

NOVO_TESTE_1H	Observado		Total
	"0"	"1"	
Previsto			
"0"	14833	94	14927
"1"	2673	819	3492
Total	17506	913	18419

Ao considerar o exposto na Tabela 5.9, obtém-se as seguintes medidas de desempenho:

- Sensibilidade = 0,90;
- Especificidade = 0,85;
- Bias = 3,82;
- POFD = 0,15.

Esses novos resultados mostram a otimização do modelo mediante o aumento no desempenho das medidas consideradas, destacando a notável depressão da *bias*, caracterizando um modelo “*menos superestimado*”. Isso confirma a influência da subjetividade dos dados em relação a variável **tempo presente**. Uma comparação do primeiro resultado (utilizando apenas os registros de tempo presente observado) com o segundo (utilizando registros de tempo presente observados e detecção de raios), ambos gerados a partir conjunto de treinamento 08 para a primeira hora futura, é apresentada mediante a Tabela 5.10.

Tabela 5.10 – Quadro de comparação dos melhores resultados obtidos para a primeira hora de previsão.

RESULTADOS	Sensibilidade	Especificidade	Bias	POFD
1º Resultado	0,83	0,83	6,64	0,17
2º Resultado (utilizando os dados da RINDAT)	0,90	0,85	3,82	0,15

RESULTADOS PARA 2ª HORA

O resultado gerado pelo conjunto de treinamento 06 para a segunda hora futura, apresentou resultados conforme exibe a Tabela 5.11.

Tabela 5.11 – Resumo do melhor resultado que utilizou os registros de raios gerados apenas por operadores de EMS, para a segunda hora de previsão.

TESTE 06_2H (RESUMO)	Observado		Total
Previsto	"0"	"1"	
"0"	15112	101	15213
"1"	2747	459	3206
Total	17859	560	18419

Considerando o exposto na Tabela 5.11, obtêm-se as seguintes medidas de desempenho:

- Sensibilidade = 0,82;
- Especificidade = 0,85;
- Bias = 5,73;
- POFD = 0,15.

Supondo que a rede indicou a condição de instabilidade da atmosfera apontada pelos raios registrados (e não apenas observados) como acerto da rede. Para a 2ª hora equivale dizer que a rede acertou 10,13%, dos 2747 casos de superestimação. Sendo assim, um novo resultado configura-se conforme mostra a Tabela 5.12.

Tabela 5.12 – Resumo do resultado utilizando os registros de raios gerados por operadores de EMS junto com os registros de raios obtidos da RINDAT, para a segunda hora de previsão.

NOVO_TESTE_2H	Observado		
Previsto	"0"	"1"	Total
"0"	15112	101	15213
"1"	2469	737	3206
Total	17581	838	18419

Considerando o exposto na Tabela 5.12, tem-se:

- Sensibilidade = 0,88;
- Especificidade = 0,86;
- Bias = 3,83;
- POFD = 0,14.

A comparação dos resultados para a segunda hora de previsão será apresentada pela Tabela 5.13.

Tabela 5.13 – Quadro de comparação dos melhores resultados obtidos para a segunda hora de previsão.

RESULTADOS	Sensibilidade	Especificidade	Bias	POFD
1º Resultado	0,82	0,85	5,73	0,15
2º Resultado (utilizando os dados da RINDAT)	0,88	0,86	3,83	0,14

RESULTADOS PARA 3ª HORA

O resultado do conjunto de treinamento 08 para a terceira hora futura, apresentou resultados conforme exibe a Tabela 5.14.

Tabela 5.14 – Resumo do melhor resultado obtido, utilizando os registros de raios gerados apenas por operadores de EMS, para a terceira hora de previsão.

TESTE 08_3H (RESUMO)	Observado		Total	
	Previsto	"0"		"1"
"0"		14100	107	14207
"1"		3759	453	4212
Total		17859	560	18419

Considerando o exposto na Tabela 5.14, tem-se:

- Sensibilidade = 0,81;
- Especificidade = 0,79;
- Bias = 7,72;
- POFD = 0,21.

Supondo que a rede indicou a condição de instabilidade da atmosfera apontada pelos raios registrados (e não apenas observados) como acerto da rede, para a 3ª hora a rede acertaria 9,26%, dos 3759 casos classificados como “1” e que foram registrados “0” pelos operadores de EMS. Os novos resultados configuram-se conforme a Tabela 5.15.

Tabela 5.15 – Resumo do resultado utilizando os registros de raios gerados por operadores de EMS junto com os registros de raios da RINDAT, para a terceira hora de previsão.

NOVO_TESTE_3H	Observado		Total	
	Previsto	"0"		"1"
"0"		14100	107	14207
"1"		3411	801	4212
Total		17511	908	18419

Considerando o exposto na Tabela 5.15, tem-se as seguintes medidas de desempenho:

- Sensibilidade = 0,88;

- Especificidade = 0,81;
- Bias = 4,64;
- POFD = 0,19.

A Tabela 5.16 apresenta uma comparação dos resultados gerados para a terceira hora futura.

Tabela 5.16 – Quadro de comparação dos melhores resultados obtidos para a terceira hora de previsão.

RESULTADOS	Sensibilidade	Especificidade	Bias	POFD
1º Resultado	0,81	0,79	7,72	0,21
2º Resultado (utilizando os dados da RINDAT)	0,88	0,81	4,64	0,19

Os resultados mostraram que redes do tipo PNN, treinadas com o programa *NeuroShell Classifier* utilizando o algoritmo Cascade Correlation (CC) têm a capacidade e o alto desempenho no seguinte mister – a identificação de padrões que antecedem um evento meteorológico significativo, possibilitando a previsão de eventos de chuva forte e/ou trovoadas na **área de influência** do AIRJ. Ou seja, esse tipo de rede é capaz de caracterizar uma atmosfera instável em um perímetro próximo do aeroporto internacional do Galeão com uma informação de grande relevância para o piloto que executa procedimentos de aproximação nessas áreas.

5.2 Estudo de Caso

1º Caso: Para este momento será investigado um caso de tempestade severa, a partir do relacionamento entre dados reais de descargas elétricas atmosféricas, imagens do satélite geostacionário GOES-10 e do radar da aeronáutica localizado no Pico do Couto (Petrópolis/RJ) e dados de precipitação das estações da Geo-Rio.

Na madrugada do dia 12/12/2007 para o dia 13/12/2007, a região metropolitana do Rio de Janeiro foi atingida por um sistema frontal subtropical que favoreceu a organização de uma zona de convergência de umidade entre o Oceano Atlântico, a região sudeste, o centro-oeste e a região norte do país, ou seja, a formação de um episódio de

ZCAS (Zona de Convergência do Atlântico Sul). Segundo informações da Geo-Rio, os acumulados de precipitação registrados durante a madrugada do dia 12/12/2007 para o dia 13/12/2007 no município do Rio de Janeiro chegaram até 96 mm em alguns pontos da cidade. No aeroporto Tom Jobim foi registrada uma rajada de vento de 34KT, ou seja, próximo de 68 km/h.

O caso estudado implicou em um incidente aéreo com um Airbus A-319 da TAM, que fazia o vôo 3950, de São Paulo para o Rio de Janeiro, com chegada prevista às 23:45 hs no AIRJ. Somente após decolar de Congonhas, na zona sul da capital paulista, o piloto foi informado sobre o mau tempo nas proximidades do AIRJ que deixou o aeroporto fechado por alguns minutos, coincidentes com a chegada da aeronave da TAM para o pouso. A aeronave não possuía combustível suficiente para retornar ao aeroporto de Congonhas, tendo, portanto, que realizar um pouso de emergência no aeroporto internacional de Cumbica, Guarulhos (São Paulo) por volta das 02:30 hs do dia 13/12/2007.

Adiante, nas Figuras 5.6 (a) e (b), pode-se observar a representação das imagens de satélite do canal infravermelho IR (realçada), em que mostra a presença de um sistema convectivo de mesoescala no Rio de Janeiro, caracterizado por nebulosidade convectiva profunda.

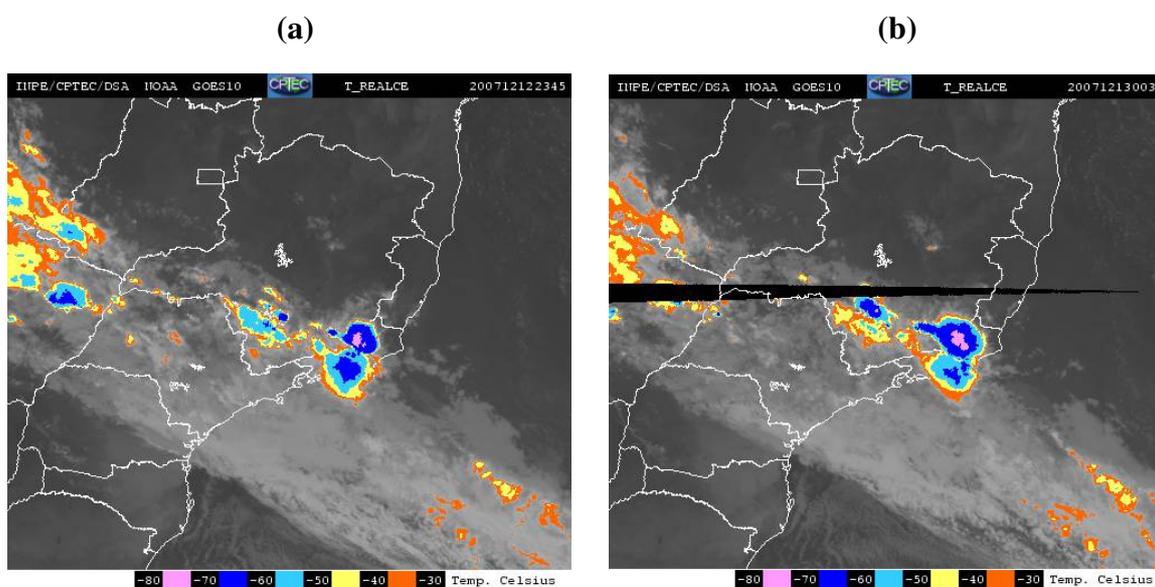


Figura 5.6 – (a) Imagem do satélite GOES-10 para o dia12/12/2007 às 2345Z; (b) Imagem do satélite GOES-10 para o dia13/12/2007 às 0030Z (fonte: REDEMET).

A Figura 5.7 apresenta as imagens do radar meteorológico do Pico do Couto, onde pode ser observada a tempestade de caráter intenso que atingiu o Rio de Janeiro. Através das imagens, pode ser visto que, no momento do incidente com a aeronave da TAM, a região metropolitana do Rio de Janeiro foi uma das mais castigadas pela chuva forte que atingiu grande parte do Estado.

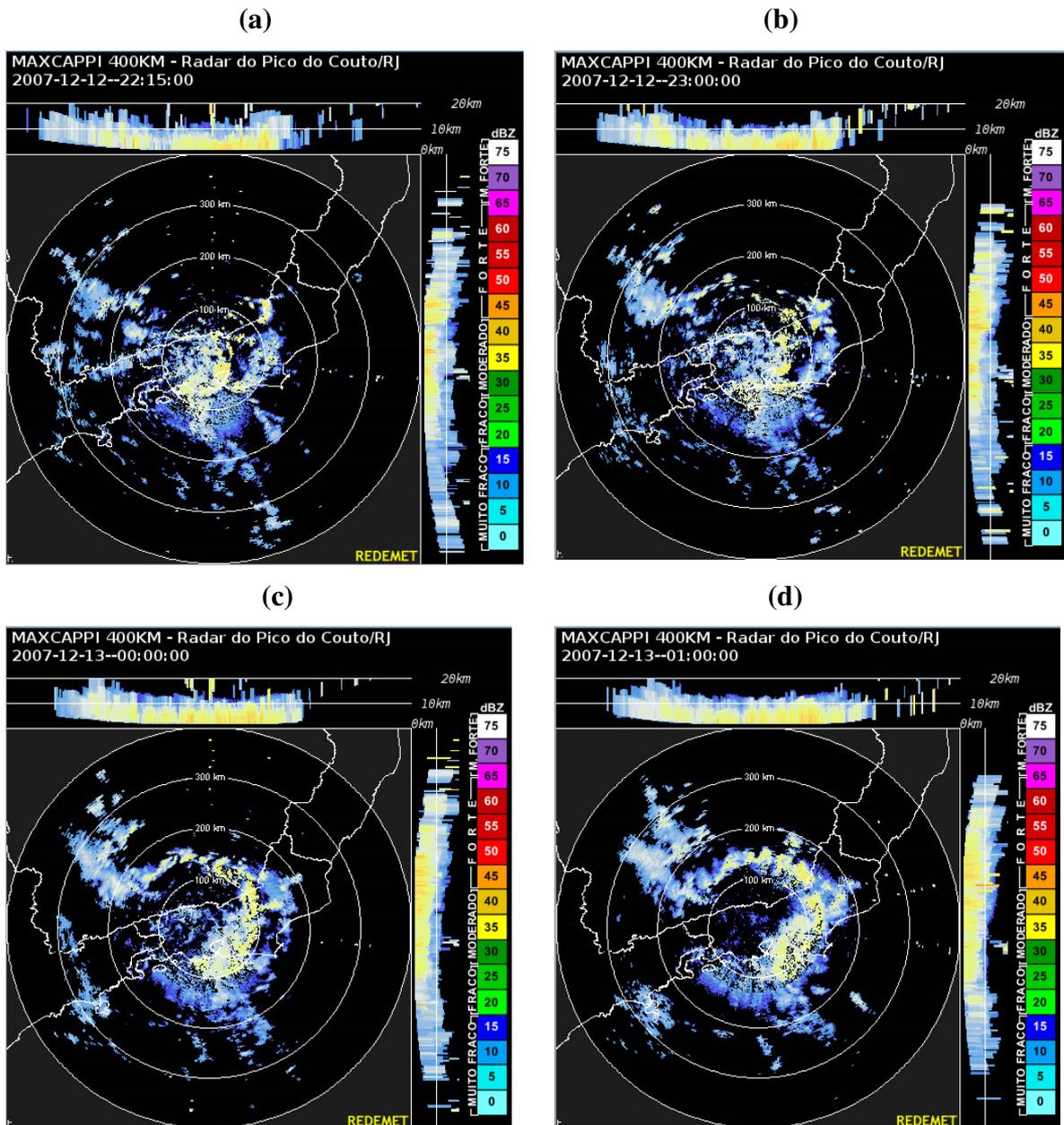


Figura 5.7 – Imagens de radar entre os dias 12 e 13 de dezembro de 2007: (a) 12/12 – 22:15 h, (b) 12/12 – 23:00 h, (c) 13/12 – 00:00 h, (d) 13/12 – 01:00 h.

A Figura 5.8 mostra a quantidade de precipitação acumulada entre 12 Z de 12/12/2007 até 12 Z de 13/12/2007. A chuva ocorreu, principalmente, durante a noite do dia 12/12 e madrugada do dia 13/12, com precipitação máxima registrada na estação de Gericinó de 67 mm/h. Na estação da Ilha do Governador, a precipitação máxima chegou a 28 mm/h.

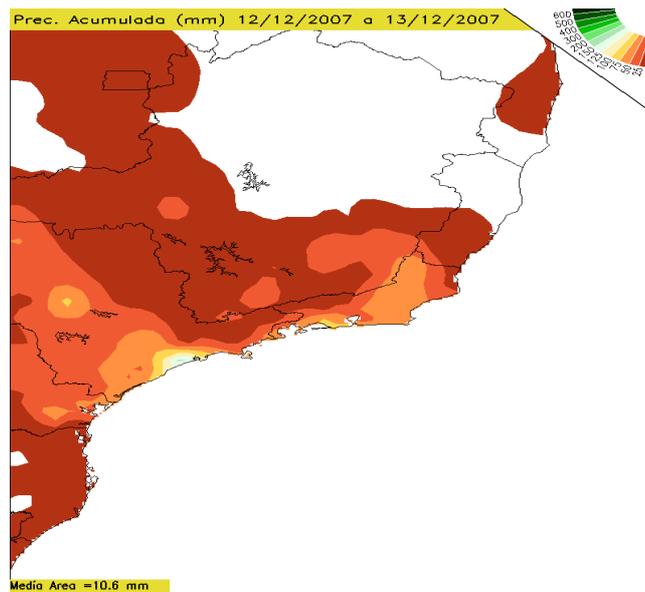


Figura 5.8 – Quantidade de precipitação acumulada com escala em milímetros das 12 Z de 12/12/2007 para as 12 Z de 13/12/2007.

Um aviso prévio desse evento poderia ter evitado o incidente com a aeronave da TAM. Nos testes feitos utilizando o modelo com redes neurais obteve-se resultados positivos para previsão do evento de precipitação extrema no Galeão com até 3 horas de antecedência. Ao ser informado das condições atmosféricas no Galeão, com essa antecedência, o piloto poderia optar por uma rota mais segura, **antes** de sua decolagem, em São Paulo. A Tabela 5.17 exhibe os resultados previstos e observados durante o evento, mostrando a eficiência do modelo no caso estudado.

Tabela 5.17 – Registro da faixa de tempo observada e prevista para as respectivas horas dos dias 12 e 13 de dezembro de 2007.

Ano	Mês	Dia	Hora	Tempo presente observado	Faixa de tempo observada	Faixa de tempo prevista
2007	3	12	22	R	2	3
2007	3	12	23	TRW-	3	3
2007	3	13	00	R+	3	3
2007	3	13	01	R-	1	1

2º Caso: No objetivo de identificar uma razão para os valores relativamente altos das *bias* foram investigados casos em que o modelo indicou a ocorrência de evento extremo que não tenha sido observado, ou seja, superestimou. Neste estudo, foram encontrados alguns casos da superestimação do modelo, conforme o ocorrido no dia 18 de março de 2009 às 17:00, 18:00 e 19:00 hs. Nessas horas o modelo previu condições de instabilidade atmosférica, classificando o tempo presente como a faixa de tempo “3”, ou seja, ocorrência de chuva forte e/ou trovoada enquanto que os registros de tempo observados pertenceram as faixas de tempo “1” e “2”, ou seja, ocorrência de chuva leve e moderada, respectivamente, ambos sem a presença de trovoada . Os registros do grupo data/hora, faixas de tempo observado e faixa de tempo previsto pelo modelo estão exibidos na Tabela 5.18.

Tabela 5.18 – Registro da faixa de tempo observada e prevista para as respectivas horas do dia 18 de março de 2009.

Ano	Mês	Dia	Hora	Tempo presente observado	Faixa de tempo observada	Faixa de tempo prevista
2009	3	18	15	H	0	0
2009	3	18	16	TRW-	3	3
2009	3	18	17	R	2	3
2009	3	18	18	R-	1	3
2009	3	18	19	H	1	3
2009	3	18	20	TRW-	3	3
2009	3	18	21	R+	3	3
2009	3	18	22	T	3	3
2009	3	18	23	TRW+	3	3

É visto que a rede “errou” ao dizer que haveria tempo presente pertencente a faixa “3” às 17:00, 18:00 e 19:00 hs. Porém pode-se observar que a partir das 20:00 hs ocorrem eventos de chuva forte e trovoadas, indicando uma condição instável da atmosfera. Dessa análise pode-se concluir que o modelo captou muito bem a condição de instabilidade na área de influência do AIRJ. Um estudo feito com os dados da RINDAT mostrou a presença de descargas elétricas no perímetro determinado pela área da Figura 3.1. Uma análise desse estudo pode ser observada através do gráfico da Figura 5.9.

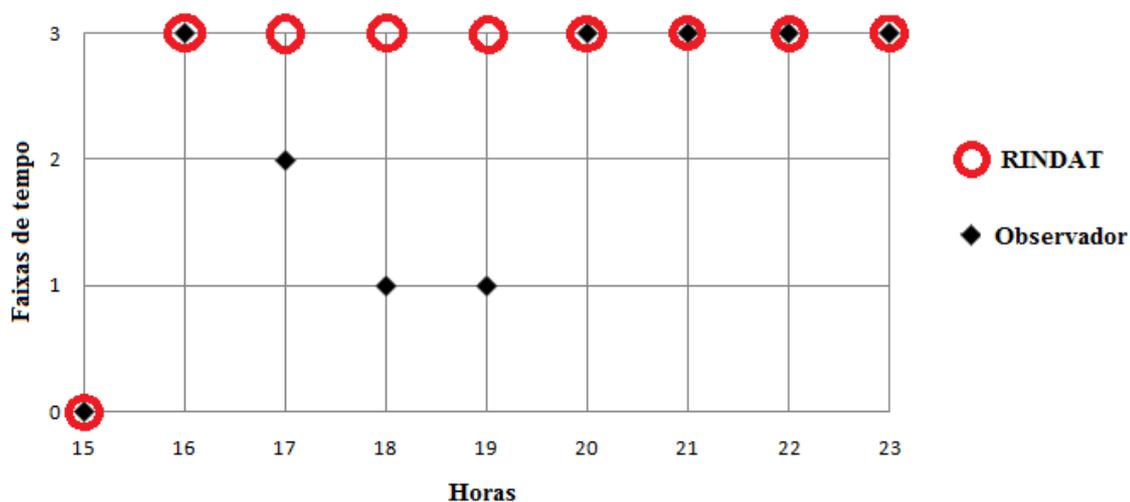


Figura 5.9 – Análise dos dados de tempo presente registrados por observadores de EMS comparados com registros de descargas atmosféricas obtidos da RINDAT.

A análise referente aos índices de instabilidade revelou que nos horários em que o modelo superestimou, a atmosfera permanecia em condição instável. Visto pelo exposto na Tabela 5.19.

Tabela 5.19 – Índices de instabilidade referentes às horas estudadas.

Ano	Mês	Dia	Hora	Índice K	Índice LR	Índice TT
2009	3	18	15	33,64 Tormentas dispersas	5,50 Atmosfera muito instável	44,97 Tormentas moderadas
2009	3	18	16			
2009	3	18	17			
2009	3	18	18			
2009	3	18	19			
2009	3	18	20			
2009	3	18	21			
2009	3	18	22			
2009	3	18	23			

Com isso, fica clara a capacidade do modelo em detectar um quadro de instabilidade presente na região de estudo, tendo em vista que as variáveis registradas (índice K, LR e TT) apontam para tal condição atmosférica.

Capítulo 6 – CONCLUSÕES E SUGESTÕES

Ao comprovar que o conhecimento prévio dos fenômenos meteorológicos que causam danos a população, é um fator necessário para os mais variados segmentos, dentre eles, a aviação, e que os modelos atuais deixam uma “janela” a ser preenchida por modelos de previsão a curtíssimo prazo, este estudo teve como proposta a utilização de redes neurais artificiais do tipo PNN treinadas com o programa *NeuroShell Classifier* através do algoritmo *Cascade Correlation* (CC) na previsão de eventos extremos para a área de influência do Aeroporto Internacional do Galeão.

6.1 Conclusões

A utilização prática da inteligência artificial na previsão de eventos meteorológicos considerados extremos para a região de estudo foi aplicada através de testes gerados a partir de conjuntos de treinamento, fornecendo diversos resultados até que se chegasse a um resultado considerado “ótimo” para o fim que se destina.

Os resultados mostraram que a aplicação de técnicas de redes neurais artificiais para previsão de eventos extremos de precipitação e ocorrência de trovoadas no aeroporto internacional do Galeão e adjacências é bastante eficaz no objetivo de prevenir os efeitos danosos que tais condições meteorológicas trazem à aviação.

Os estudos de casos mostraram que mesmo quando existiram indicações de erro de superestimação do modelo, em alguns casos (em média 10,35% dos casos) o modelo acertou ao prever a condição do evento extremo.

A inclusão de variáveis derivadas, como os índices de instabilidade, gerou um aumento do desempenho do modelo, assim como a utilização de dados de descargas atmosféricas na confirmação ou não do evento previsto.

Portanto, a RNA é capaz de reconhecer o perfil termodinâmico da atmosfera e sua aplicação no sistema de previsão de curtíssimo prazo de chuva forte e/ou trovoadas, seria uma importante contribuição no ramo da aviação no âmbito do aeroporto internacional do Galeão, conforme mostram os resultados apresentados na Tabela 6.1.

Tabela 6.1 – Resultado da previsão de eventos extremos na área de influência do AIRJ para as três horas futuras.

Hora futura de previsão	Sensibilidade	Especificidade	Bias	POFD
1ª hora	0,90	0,85	3,82	0,15
2ª hora	0,88	0,86	3,83	0,14
3ª hora	0,88	0,81	4,64	0,19

Através de um estudo de caso pode-se comprovar que o modelo obteve resultados positivos para previsão do evento de precipitação extrema ocorrido no Galeão com até 3 horas de antecedência, o que poderia ter evitado o incidente com a aeronave da TAM no dia 13 de dezembro de 2007.

6.2 Sugestões

Uma sugestão a partir deste trabalho pese ao fato de que o estudo analisou um período relativamente pequeno de dados (13 anos) o que pode ser um fator limitante no desempenho da RNA.

Também, um estudo aprofundado sobre os fenômenos que ocorrem com frequência associados a chuvas intensas funcionaria como mais um instrumento na análise dos eventos extremos de precipitação. Como por exemplo, a análise na combinação de uma componente dinâmica com padrões termodinâmicos presentes próximos ao local de estudo contribuiriam para a identificação de padrões que antecedem um evento meteorológico significativo.

Outra sugestão seria quanto ao estudo no campo dos vetores de vento e JBNs, pois a umidade que eles carregam pode causar precipitações que, junto a outros fatores colaboram para a ocorrência de fenômenos de precipitação intensa. Além da importância desse vetor (vento) no momento do pouso e decolagem de uma aeronave.

Nesta pesquisa, a utilização da questão referente a sazonalidade foi abordada de forma abrangente ao colocar as datas de forma corrida dentro do ano em vigor. Propõe-se para trabalhos futuros a utilização de redes neurais apropriadas para cada estação, ou mesmo que aborde outras questões referentes a sazonalidade.

Como sugestão para trabalhos futuros, serve o uso de registros do formulário IEPV105-78 (vide Anexo E) ao invés do METAR como dados de entrada da rede. Os dados do METAR são confeccionados a fim de atender a operacionalidade na comunicação entre a torre de controle e o piloto de uma aeronave. Para tanto, essa mensagem (METAR) utiliza-se de códigos que deixam intrínsecas informações de grande valia para formatação de uma previsão. Além disso, tanto os dados de superfície, quanto os de altitude, foram confeccionados de forma a identificar as condições atmosféricas em um exato momento contendo algumas informações sujeitas a erro. Uma solução seria a utilização de uma ferramenta que pudesse captar informações meteorológicas do local de estudo de forma contínua e objetiva.

É importante ressaltar que a inclusão de outros índices de instabilidade além dos que foram utilizados nesta pesquisa e de dados que fossem registrados de forma contínua deprimiria a deficiência no registro dos dados.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ALERTA RIO. **Sistema de alerta de chuvas intensas e de deslizamentos em encostas da cidade do Rio de Janeiro.** Disponível em: <http://www2.rio.rj.gov.br/georio/site/alerta/alerta.htm>. Acesso em: 19 jan. 2010.
- [2] ALEXANDER, L. V., ZHANG, X., PETERSON, T.C. *Global observed changes in daily climate extremes of temperature and precipitation*, *Jn. Geophys Res.*, 111 (D05109), 2006.
- [3] ALMEIDA M. V. **Aplicação de técnicas de Redes Neurais Artificiais na previsão de curtíssimo prazo da visibilidade e teto para o aeroporto de Guarulhos**, Universidade Federal do Rio de Janeiro, 2009.
- [4] AMBIENTE BRASIL, Disponível em: <http://www.ambientebrasil.com.br/>. Acesso em 20 fev. 2010.
- [5] ANAC (Agência Nacional de Aviação Civil), **Relatório anual de segurança operacional de 2009.**
- [6] ANDRADE, E. S. **Estudo sobre a influência de jatos de baixos níveis na estrutura da camada limite estável acima do pantanal mato-grossense**, Dissertação (Mestrado em Meteorologia), fl.110, INPE, São José dos Campos, 2003.
- [7] ARAÚJO M. A. A., SANTOS, A. P., VIOLA P. E., ARAÚJO, C. M. A. **Aspectos Sinóticos das chuvas intensas sobre a região do sudeste do Brasil, Fevereiro de 1988**, Anais do V Congresso Brasileiro de Meteorologia v.2, pp. 39-43, 1988.
- [8] BANKERT, R. L. *Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network*. *Journal of Applied Meteorology*, 33, 909-918, 1994.
- [9] BENETI, C. A., SILVA DIAS, M. A. **Análise da performance de índices de instabilidade como previsores de tempestades na região de São Paulo**, Anais do IV Congresso Brasileiro de Meteorologia, Brasília/DF, Soc. Bras. Meteorologia, v.2, p. 65-70, 1986.

[10] BERRI, G. J., INZUNZA, J. B. *The effect of the Low-Level Jet on the pole ward water vapor transport in the central region of South America*, *Atmospheric Environment*, v. 27A, n. 3, p. 335-341, 1993.

[11] BRAGA, A. P., CARVALHO, A. C. P. L. F., LUDERMIR, T. B. **Fundamentos de Redes Neurais Artificiais**, IX Escola Brasileira de Computação, v.1, p. 245, Rio de Janeiro, 1998.

[12] BRAGA, A.P., CARVALHO, A. C. P. L. F., LUDERMIR, T. B. **Redes Neurais Artificiais: teoria e aplicações**, Livros Técnicos e Científicos – LTC, 1ª ed., v. 1, p. 250, Rio de Janeiro, 2000.

[13] BRASIL DECEA, **Manual do Comando da Aeronáutica – MCA 105-2: Manual de Estações Meteorológicas de Superfície**, Rio de Janeiro, agosto 2010.

[14] BRASIL DECEA, **Manual do Comando da Aeronáutica – MCA 105-9: Manual de Estações Meteorológicas de Altitude**, Rio de Janeiro, fevereiro 2011.

[15] BRASIL DECEA, **Manual do Comando da Aeronáutica – MCA 105-10: Manual de Códigos Meteorológicos**, Rio de Janeiro, janeiro 2010.

[16] **Classificação climática de Koppen Geiger**. In: Wikipédia: a enciclopédia livre. Disponível em: http://pt.wikipedia.org/wiki/Classificacao_climatica_de_Koppen-Geiger. Acesso em: 07 ago. 2010.

[17] COSTA, R. S., LIMA, D. O, SANTOS, I. A., LANDAU, L. **O uso do índice CK como indicador de chuvas intensas – o caso Maceió**, Congresso Brasileiro de Meteorologia, XIV (CBMET), 2006.

[18] DECEA. **Departamento de Controle do Espaço Aéreo – DECEA**, Disponível em: <http://www.decea.gov.br/espaco-aereo/meteorologia-aeronautica>. Acesso em 13 fev. 2010.

[19] DEREZYNSKI, C. P., OLIVEIRA, J. S., MACHADO, C. O., **Climatologia da precipitação no Município do Rio de Janeiro**, Revista Brasileira de Meteorologia, v.24, n.1, 24-38, 2008.

[20] DIAS, V. S. **Análise dos índices de instabilidade para a previsão de formações convectivas severas para a região do aeródromo do Galeão, Aeroporto Tom Jobim**, Relatório de Pesquisa, 2008.

[21] DINIZ, M. C. S., CORREIA, M. F., ARAGÃO, MARIA R. S., MELO, E. C. S. **Importância Relativa de Variáveis Meteorológica na Formação da Convecção Profunda: Uma aplicação da análise fatorial de componentes principais**, XIV Congresso Brasileiro de Meteorologia, Florianópolis, 2006.

[22] DINIZ, M. C. S. **Variabilidade Climática e ocorrência de cheias na zona semi-árida da Bacia Hidrográfica do São Francisco**, 2008.

[23] EASTERLING, D. R., MEEHL, G. A., PARMESAN, C., CHANGNON, S. A., KARL, T. R., MEARNNS, L. O. *Climate extremes: observations, modeling, and impact*, *Science*, v. 289, pp. 2068-2074, 2000.

[24] ESCOBAR, G. **Meteorologia Sinótica Prática Aplicada à Previsão do Tempo**, Rio de Janeiro, 2007.

[25] ESPIRITO SANTO, C., SATYAMURTY, P. **Eventos extremos de precipitação na região sudeste do Brasil e redondezas no período de 1997-2001**, Anais do XII Congresso Brasileiro de Meteorologia, Foz do Iguaçu / PR, 2002.

[26] FAHLMAN, S. E.; LIBIERE, C. *The Cascade-Correlation Learning Architecture*. In CMU-CS-90-100, School of Computer Science Carnegie Mellon University, Pittsburgh, PA, 1991.

[27] FOGACCIA, C. V. C., PEREIRA FILHO, A. J. **Turbulência e cisalhamento do vento na área do Aeroporto Internacional de São Paulo/Guarulhos**, Anais do XII Congresso Brasileiro de Meteorologia, Foz do Iguaçu / PR, 2002.

[28] GOOGLE MAPS. Disponível em: <http://maps.google.com.br/>. Acesso em: 7 abr. 2011.

[29] GROISMAN, P. Y., KARL, T. R., EASTERLING, D. R., *Changes in the Probability of Heavy Precipitation: Important Indicators of Climatic Change*, *Climatic Change* 42, pp. 243-283, 1999.

[30] GROISMAN, P. Y., *et al.* **Trends in intense precipitation in the climate record.** *Journal of Climate*, n.18, p. 1326-50, 2005.

[31] HAYKIN, S. **Neural networks: a comprehensive foundation**, Macmillan College, Publishing Company, 2nd ed. Inc., Englewood Cliffs, NJ, USA, 1998.

[32] HAYKIN, S. **Redes Neurais Princípios e Prática**, 2^a Ed., Editora Bookman, Porto Alegre, 2002.

[33] HEGERL, G. C., KARL T. R., ALLEN, M., BINDOFF, N. L., GILLET, N. **Climate Change Detection and Attribution: Beyond Mean Temperature Signals**, *Jn. Climate* 19, pp. 5058-5077, 2006.

[34] HOLTON, J. R. **Introduction to Dynamic Meteorology**. 3rd ed. San Diego: Academic Press, 1992.

[35] INFRAERO. **Empresa Brasileira de Infraestrutura Aeroportuária**. Disponível em: <http://www.infraero.gov.br/aeroportos/rio-de-janeiro/aeroporto-internacional-do-rio-de-janeiro.html>. Acesso em: 10 mai. 2010.

[36] JUNKER, N. W., SHNEIDER, R. S., FAUVER, S. L. **A study of heavy rainfall events during the great Midwest flood of 1993.** *Weather and Forecasting*, v. 14, pp. 701-712, 1999.

[37] KARL, T.R.; EASTERLING, D.R. **Climate Extremes: Selected Review and Future Research Directions**, *Climatic Change* 42, pp. 309-325, 1999.

[38] KATZ, R.W., BROWN, B. G. **Extreme Events in a Changing Climate: Variability is More Important than Averages**, *Climatic Change*, v. 21, n. 3, pp. 289-302, 1992.

[39] KONRAD, C.E. **Synoptic-scale features associated with warm season heavy rainfall over the interior southern United States.** *Weather and Forecasting*, v. 12, pp. 557-571, 1997.

[40] KOSKO, B. **Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence**. Prentice Hall, Englewood Cliffs, NJ, USA, 1992.

[41] KOVACS, Z. L. **Redes Neurais Artificiais – Fundamentos e Aplicações**, 3^a ed., livro publicado por Livraria da Física Editora, São Paulo, 2002.

- [42] LEAL, N. A., CRUZ, D. S., LEITE, A. F. **Caracterização do Comportamento Sazonal das Chuvas em Bairros das Zonas Norte e Oeste do Município do Rio de Janeiro.** In: Anais XVI Encontro Nacional dos Geógrafos, p. 5, Porto Alegre-RS, 2010.
- [43] LIEBMANN, B., JONES, C., CARVALHO, L. M. V. DE. *Interannual variability of daily extreme precipitation events in the State of São Paulo, Brazil.* *Journal of Climate*, v. 14, pp. 208-218, 2000.
- [44] LIMA, D. R. O. **Diagnóstico de chuvas e previsão meteorológica para a bacia hidrográfica do Rio Manso,** Dissertação (Mestrado em Engenharia Civil), COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2005.
- [45] MARENGO, J. A., *et al.* *Climatology of the Low-Level Jet East of the Andes as Derived from the NCEP–NCAR. Reanalyses: Characteristics and Temporal Variability,* *Jn. Climate* 17, n. 12, pp. 2261-2280, 2004.
- [46] MARENGO, J. A., **Mudanças Climáticas Globais e seus Efeitos sobre a Biodiversidade: Caracterização do Clima Atual e Definição das Alterações Climáticas para o Território Brasileiro ao Longo do Século XXI,** Ministério do Meio Ambiente, Secretaria de Biodiversidade e Florestas, Brasília, DF, 2006.
- [47] MARTH, L., *Stratified Atmospheric Boundary-Layers,* *Boundary-Layer Meteorology*, v. 90, n. 3, pp. 375-396, Mar. 1999.
- [48] MCCULLOCH W.S., PITTS W. *A Logical Calculus of the Ideas Immanent in Nervous Activity,* *Bulletin of Mathematical Biophysics*, v. 5, pp. 115–133, 1943.
- [49] MELO, J. D. **Aspectos de Implementação das Redes Neurais,** DCA-CT-UFRN, julho 2004.
- [50] MOURÃO, C. E. F., *et al.*, **Estudo do Comportamento de indicadores de tempo severo em casos de tempestade sobre o Rio de Janeiro,** Anais do XIV Congresso Brasileiro de Meteorologia, Florianópolis, SC, SBMET, v.1, p. 3, 2006.
- [51] MENEZES, W. F. **Tempestades severas: um modelo para latitudes subtropicais,** Tese de Doutorado, Departamento de Ciências Atmosféricas, Inst. Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, 1997.

- [52] MICROSOFT CORPORATION **Enciclopédia Microsoft Encarta**, 1993-1999.
- [53] MILLER, R. C., MADDOX, R. A. **“The use of sweat and spot index in operational severe storm forecasting”**. In: *9th Conference Severe Local Storms. American Meteorological Society*, Boston, 1975.
- [54] MOURA, A. D. **Entrevista à Revista Mensal do CREA-RS**, v. 37, set. 2007.
- [55] MOURA, A. D. **Meteorologia – é possível resgatar o tempo perdido?**, Revista Brasileira de Tecnologia, Brasília, v. 17, n.1, p. 5-14, jan.-fev., 1986.
- [56] MOURA JR, J. R. V., MARTINS, G, JACOMINE, L. *et al.* **“Redes Neurais Probabilísticas e Aplicação em Otimização de Servidores de Internet”**. In: 47^a Reunião Anual Região Brasileira da Sociedade Internacional de Biometria, Rio Claro-SP, 2002.
- [57] NASCIMENTO, E. L., CALVETTI, L. **Identificação de condições precursoras de tempestades severas no sul do Brasil utilizando-se radiossondagem e parâmetros convectivos**, Anais do XIII Congresso Brasileiro de Meteorologia, Fortaleza / CE, 2004.
- [58] NASCIMENTO, M. X., *et al.* **Aplicação de indicadores de tempo severo na caracterização de um ciclone intenso**. Anais do XVI Congresso Brasileiro de Meteorologia, Belém, PA, 2010.
- [59] NASCIMENTO, E. L., **Previsão de tempestades severas utilizando-se parâmetros convectivos e modelos de mesoescala: uma estratégia operacional adotável no Brasil?**, Revista Brasileira de Meteorologia, v. 20, n. 1, p. 121-140, 2005.
- [60] NOAA. *National Oceanic and Atmospheric Administration*. Disponível em: <http://www.ncdc.noaa.gov/oa/climate/igra/index.ph>. Acesso em: 26 de novembro 2010.
- [61] OLIVEIRA, P. A., FITZJARRALD, D. R. **The Amazon River Breeze and Local Boundary Layer: I. Observations**, *Boundary-Layer Meteorology*, v. 63, n. 1-2, p. 141-162, Feb. 1993.
- [62] PINTO, H. S. Artigo na **ComCiência – Revista Eletrônica de Jornalismo Científico**. 2002.
- [63] QUEIROZ, R. B., RODRIGUES, A. G., GÓMEZ, A. G. **Comparativo entre as**

técnicas máxima verossimilhança gaussiana e redes neurais na classificação de imagens IR-MSS CBERS 1. In: *I Workshop de Computação da Região Sul*, Florianópolis, SC, 2004.

[64] REBOITA, M. S., GAN, M. A., ROCHA, R. P., AMBRIZZI, T. **Regimes de precipitação na América do Sul: uma revisão bibliográfica**. Revista brasileira de meteorologia, São Paulo, v. 25, n. 2, pp. 185-204, 2010.

[65] REDEMETS. **Rede de Meteorologia do Comando da Aeronáutica**. Disponível em: <http://www.redemet.aer.mil.br>. Acesso em: 12 mai. 2010.

[66] RINDAT. **Rede Integrada Nacional de Detecção de Descargas Atmosféricas – RINDAT**. Disponível em: <http://www.rindat.com.br/index.html>. Acesso em 12 jun. 2011.

[67] SCHMITZ, J. E. **Cálculos de Estabilidade e Divisão de Fases por Meio de Redes Neurais Artificiais**. Tese de Doutorado, Faculdade de Engenharia Química da Universidade Estadual de Campinas, Campinas, São Paulo, 2006.

[68] SCOLAR, J., SILVA DIAS, M. A. F. “*Linhas de instabilidade pré-frontais e sua dependência de condições sinóticas*”. In: **II Congresso Brasileiro de Meteorologia**, SBMET, Anais... v.4, pp.88-102, 1982.

[69] SEVERO, D. L., GAN, M. A., FERREIRA, N. J. **Estudo de um caso de chuvas intensas associado a enchentes na Região do Vale do Itajaí, em Santa Catarina**, Anais do VIII Congresso Brasileiro de Meteorologia, Belo Horizonte / MG, SBMET ,v. 2, pp. 561-564, 1994.

[70] SIAS, E. E .K., SILVA DIAS, M. A. F. **Análise de um caso de precipitação extrema no inverno na cidade de São Paulo**, Anais do XII Congresso Brasileiro de Meteorologia, Foz do Iguaçu / PR, 2002.

[71] SILVA DIAS, M. A. F. **Índices de instabilidade para previsão de chuva e tempestades severas**, Universidade de São Paulo, SP, 2000.

[72] SOLOMON, S. “*The Physical Science Basis*”. In: Intergovernmental Panel on Climate Change (IPCC), p. 996, 2007.

[73] SOLOW, A. R. **On Testing for Change in Extreme Events**, *Climatic Change*, v. 42,

pp. 341-349, 1999.

[74] STENSRUD, D. J. *Importance of Low-Level Jets to Climate: A Review*, *Journal of Climate*, v. 9, pp. 1698-1711, 1996.

[75] STULL, R. B. *An Introduction to Boundary Layer Meteorology*, Dordrecht, Kluwer Ac. Publ., 1995.

[76] SUTTON, O. G. *Micrometeorology - A study of physical processes in the lowest layers of the Earth's atmosphere*, livro publicado por Krieger Publ., Huntington / NY, 1977.

[77] SUTTON, R., BARTO, A. G. *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, 1998.

[78] TEIXEIRA, M. S., SATYAMURTY, P. *Dynamical and synoptic characteristics of heavy rainfall episodes in southern Brazil*, *Monthly Weather Review*, v. 135, p. 598-617, 2007.

[79] THORPE, A.J., GUYMER T. H. *The Nocturnal Jet*, *Quarterly Journal of the Royal Meteorological Society*, v. 103, Issue 438, pp. 633–653, October 1977. Publicado on-line em set. 2007.

[80] VALENÇA, M. **Aplicando Redes Neurais – Um Guia Completo**, Editora Livro Rápido, Pernambuco, 2005.

[81] VITTE, A. C, GUERRA, A. J. T. **Reflexões sobre a geografia física no Brasil**, livro da editora Bertrand Brasil, Rio de Janeiro, 2004.

[82] VONK, E.; JAIN, L.C.; JOHNSON, R.P. *Automatic Generation of Neural Network Architecture Using Evolutionary Computation*. In: World Scientific Publishing Co. Pte.Ltd., 1997.

[83] WEATHER/WYOMING. **University of Wyoming**, College of Engineering Department of Atmospheric Science. Disponível em: <http://www.weather.uwyo.edu>. Acesso em 14 jan. 2011.

[84] WELLS, A. T., RODRIGUES, C. C. *Commercial aviation safety* , livro da ed. McGraw-Hill, New York, 2004.

[85] WILKS, S. D. – *Statistical Methods in the Atmospheric Sciences*. 2nd ed., New York, USA, Academic Press, 2006.

[86] WISNER, B., BLAIKIE, P., CANNON, T., DAVIS, I. *At Risk: Natural hazards, people's vulnerability and disasters*, 2nd ed., livro publicado por Routledge, Taylor & Francis Group, New York, 1994.

APÊNDICE I – Sobre o *script* feito no MATLAB para formatação dos dados de precipitação

Os dados de precipitação originais provindos da Geo-Rio tiveram de ser formatados a fim de atender a necessidade desta pesquisa em razão da compatibilidade entre data/hora dos demais dados presentes no trabalho. Para isso, utilizou-se de um “*script*” feito no programa MATLAB descrito da seguinte forma:

load('C:\Users\ALESSANA\UFRJ\MESTRADO\precipitacao.txt'); carrega o arquivo a ser trabalhado.

soh_dados_precipitacao=precipitacao(:,4:27); Exclui as colunas de a 3 que referem-se as datas para só trabalhar com os dados de precipitação, ou seja, da coluna 4 até a 27.

prec=soh_dados_precipitacao'; Transpondo a matriz de dados de precipitação, ou seja, linhas virão colunas e colunas virão linhas no novo arquivo.

total_elementos=numel(prec); Conta a quantidade de elementos da matriz prec.

Precipitacao1=reshape(prec,total_elementos,1); Coloca todas as colunas uma embaixo da outra na primeira coluna.

save

('C:\Users\ALESSANA\UFRJ\MESTRADO\precipitacao_estacao.txt','precipitacao', '-ascii'); Salva o arquivo precipitacao_estacao no formato .txt.

save

('C:\Users\ALESSANA\UFRJ\MESTRADO\precipitacao_estacao.xls','precipitacao', '-ascii'); Salva o arquivo precipitacao_estacao no formato .xls.

ANEXO A – Carta de Visibilidade



Figura A.1 – Imagem da carta de pontos de referência situados a menos de 2km para visibilidade do Departamento de Controle do Espaço Aéreo.

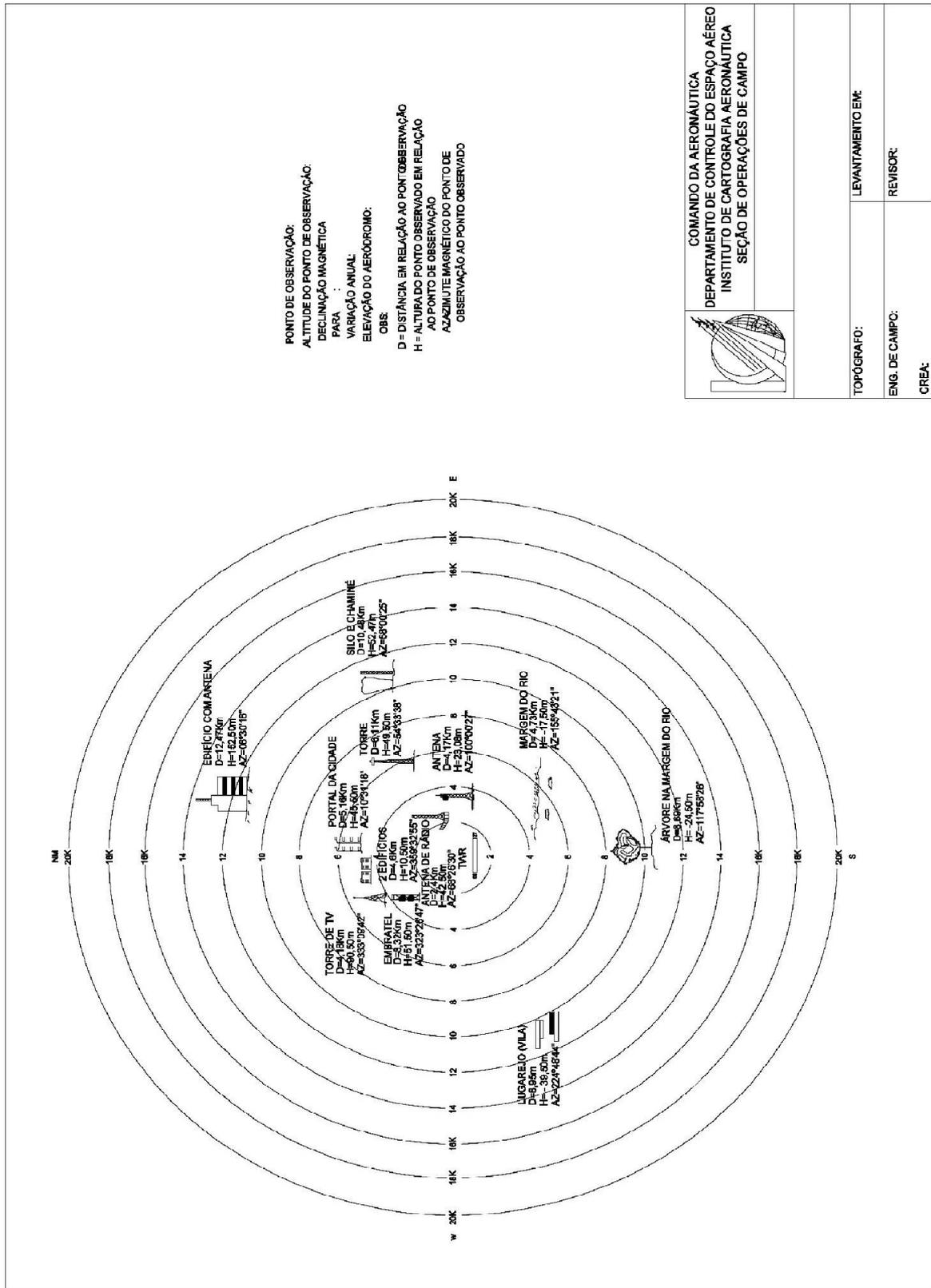


Figura A.2 – Imagem da carta de pontos de referência situados a mais de 2km para visibilidade do Departamento de Controle do Espaço Aéreo.

ANEXO B – Formatação dos dados de precipitação

Tabela B.1 – Extrato dos acumulados horários de precipitação para a estação pluviométrica da Ilha do Governador (dados em seu formato original).

Dia	Mês	Ano	1h	2h	3h	4h	5h	6h	7h	8h	9h	10h
1	1	2008	0	0	0	0	0	0	0	0	0	0
2	1	2008	0	0	0	0	0	0	0	0	0	0
3	1	2008	0	0	0	0	0	0	0	0	0	0
4	1	2008	0	0	0	0	0	0	0	0	0	0
5	1	2008	0	0	0	0	0	0	0	7.2	2.6	8.8
6	1	2008	1.0	0	.2	0	0	0	0	0	0	0
7	1	2008	0	0	0	0	0	0	0	0	0	0
8	1	2008	0	0	0	0	0	0	0	0	0	0
9	1	2008	0	0	0	0	0	0	0	0	0	0
10	1	2008	0	0	0	0	0	0	0	0	0	0
11	1	2008	0	0	0	0	0	0	0	0	0	0
12	1	2008	0	0	0	0	0	0	0	0	0	0
13	1	2008	0	0	0	0	0	0	0	0	0	0
14	1	2008	0	0	0	0	0	0	0	0	0	0

Tabela B.2 – Extrato dos dados de acúmulos horários de precipitação em sua disposição final para entrada da rede.

Ano	Mês	Dia	Hora	Ilha	Irajá	Penha	Piedade	Saúde
1997	1	1	0	0	0	0	0	0
1997	1	1	1	0	0	0	0	0
1997	1	1	2	0	0	0	0	0
1997	1	1	3	0	0	0	0	0
1997	1	1	4	0	0	0	0	0
1997	1	1	5	0	0	0	0	0
1997	1	1	6	0	0	0	0	0
1997	1	1	7	0	0	0	0	0
1997	1	1	8	0	0	0	0	0
1997	1	1	9	0	0	0	0	0
1997	1	1	10	0	0	0	0	0
1997	1	1	11	0	0	0	0	0
1997	1	1	12	0	0	0	0	0

ANEXO C – Variável Peso

Tabela C.1 – Extrato dos dados de altitude com os seus respectivos “pesos”.

ANO	MÊS	DIA	HORA	GEOPOBS1000	VWINDGRAD500	PESO
2007	4	3	14	148	0,8	0,833333333
2007	4	3	15	148	0,8	0,75
2007	4	3	16	148	0,8	0,666666667
2007	4	3	17	148	0,8	0,583333333
2007	4	3	18	148	0,8	0,5
2007	4	3	19	148	0,8	0,416666667
2007	4	3	20	148	0,8	0,333333333
2007	4	3	21	148	0,8	0,25
2007	4	3	22	148	0,8	0,166666667
2007	4	3	23	148	0,8	0,083333333
2007	4	4	0	99	-1,3	1
2007	4	4	1	99	-1,3	0,916666667
2007	4	4	2	99	-1,3	0,833333333
2007	4	4	3	99	-1,3	0,75
2007	4	4	4	99	-1,3	0,666666667
2007	4	4	5	99	-1,3	0,583333333
2007	4	4	6	99	-1,3	0,5
2007	4	4	7	99	-1,3	0,416666667
2007	4	4	8	99	-1,3	0,333333333
2007	4	4	9	99	-1,3	0,25
2007	4	4	10	99	-1,3	0,166666667
2007	4	4	11	99	-1,3	0,083333333
2007	4	4	12	109	-1,3	1
2007	4	4	13	109	-1,3	0,916666667
2007	4	4	14	109	-1,3	0,833333333

ANEXO D – Correlação

Tabela D.1 – Tabela de correlação entre as variáveis candidatas e a saída da RNA.

VARIÁVEL	CORREL_SAÍDA	VARIÁVEL	CORREL_SAÍDA
GEOPOBS1000	-0,07812	SATVAP700	0,040029
GEOPOBS850	-0,07636	SATVAP500	0,040015
GEOPOBS700	-0,04416	RH1000(%)	-0,03538
GEOPOBS500	-0,00182	RH850(%)	0,024574
GEOPCALC1000	-0,08916	RH700(%)	0,065655
GEOPCALC850	-0,07716	RH500(%)	0,040724
GEOPCALC700	-0,04908	RHGRAD1000	0,039432
GEOPCALC500	-0,0134	RHGRAD850	0,003303
T_1000 (C)	0,066553	RHGRAD700	0,004597
T_850 (C)	0,052207	RHGRAD500	-0,00496
T_700 (C)	0,041323	UWND1000(m/s)	0,016485
T_500 (C)	0,041043	UWND850(m/s)	0,018283
Td_1000(C)	0,037027	UWND700(m/s)	0,015528
Td_850(C)	0,063361	UWND500(m/s)	0,005228
Td_700(C)	0,071129	UWDGRAD1000	0,02275
Td_500(C)	0,041124	UWDGRAD850	0,004928
TEMPGRAD1000	-0,03103	UWDGRAD700	-0,00516
TEMPGRAD850	-0,00653	UWDGRAD500	-0,02348
TEMPGRAD700	-0,023	VWND1000(m/s)	0,006673
TEMPGRAD500	0,012648	VWND850(m/s)	-0,02057
PTEMP1000 (C)	0,066553	VWND700(m/s)	-0,03106
PTEMP850 (C)	0,052215	VWND500(m/s)	0,02469
PTEMP700 (C)	0,041262	VWNDGRAD1000	-0,00178
PTEMP500 (C)	0,040994	VWNDGRAD850	-0,00919
PTEMPGRA1000	-0,0311	VWNDGRAD700	0,001566
PTEMPGRAD850	-0,00505	WNDGRAD500	0,003884
PTEMPGRAD700	-0,02348	QNHsup	-0,07893
PTEMPGRAD500	0,011577	TMPsup	0,050981
VAPPRESS1000	0,036663	DEWsup	0,036034
VAPPRESS850	0,065066	Rhsup	-0,02899
VAPPRESS700	0,079938	DIRsup	0,020477
VAPPRESS500	0,045371	VELsup	0,082896
SATVAP1000	0,064158	VISsup	-0,00444
SATVAP850	0,052651	Tempo	0,129704

