



**COPPE/UFRJ**

APLICAÇÃO DE TÉCNICAS DE REDES NEURAIIS ARTIFICIAIS  
NA PREVISÃO DE CURTÍSSIMO PRAZO DA VISIBILIDADE E TETO PARA O  
AEROPORTO DE GUARULHOS – SP

Manoel Valdonel de Almeida

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Civil, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Civil.

Orientadores: Nelson Francisco Favilla Ebecken  
Gutemberg Borges França

Rio de Janeiro  
Outubro de 2009

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

APLICAÇÃO DE TÉCNICAS DE REDES NEURAIS ARTIFICIAIS  
NA PREVISÃO DE CURTISSIMO PRAZO DA VISIBILIDADE E TETO PARA O  
AEROPORTO DE GUARULHOS – SP

Manoel Valdonel de Almeida

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ  
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA  
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS  
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM  
ENGENHARIA CIVIL.

Aprovada por:

---

Prof. Nelson Francisco Favilla Ebecken, D.Sc.

---

Prof. Gutemberg Borges França, Ph.D.

---

Prof. Augusto José Pereira Filho, Ph.D.

---

Prof<sup>a</sup>. Beatriz de Souza Leite Pires Lima, D. Sc.

---

Prof. Luiz Pereira Calôba, Dr. Ing.

RIO DE JANEIRO, RJ - BRASIL  
OUTUBRO DE 2009

Almeida, Manoel Valdonel de

Aplicação de Técnicas de Redes Neurais Artificiais na Previsão de Curtíssimo Prazo da Visibilidade e Teto para o Aeroporto de Guarulhos – SP./Manoel Valdonel de Almeida - Rio de Janeiro: UFRJ/COPPE, 2009.

XXIV, 182 p. : il. ; 29,7 cm.

Orientadores: Nelson Francisco Favilla Ebecken  
Gutemberg Borges França.

Tese (doutorado) – UFRJ/ COPPE/ Programa de Engenharia Civil, 2009.

Referências Bibliográficas: p. 139-145.

1. Redes Neurais Artificiais 2. Previsão de curtíssimo prazo 3. Visibilidade 4. Nevoeiro 5. Teto I. Ebecken, Nelson Francisco Favilla *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Civil. III. Título.



*A Deus, o criador e Senhor de todas as coisas,  
à minha esposa, filhos e futuros netos.*

## **AGRADECIMENTOS**

Agradeço, em primeiro lugar a Deus, Senhor de todas as coisas.

Agradeço a todos que direta ou indiretamente contribuíram para o desenvolvimento e conclusão deste trabalho.

Quero agradecer especialmente aos meus orientadores Prof. Nelson Francisco Favilla Ebecken, e Prof. Gutemberg Borges França pelo apoio, orientações balizadas e pela paciência e incentivo com meus momentos de desânimos.

Também, quero expressar meus agradecimentos ao Comando da Aeronáutica, através das seguintes organizações Departamento de Controle do Espaço Aéreo – DECEA e do Instituto de Controle do Espaço Aéreo - ICEA pela gentileza em fornecer os dados de trabalho e os manuais relacionados ao assunto.

Não poderia deixar de agradecer aos companheiros do Laboratório de Meteorologia Aplicada – LMA que de certa forma, em algum momento da caminhada me forneceram alguma informação útil ao trabalho.

Não esquecendo, também, de agradecer aos amigos da Secretária Acadêmica do PEC pelas suas orientações, e gentileza.

Finalmente quero agradecer à minha família: minha querida esposa Cledionora, meus filhos João Ricardo, Valdonel Jr. E Ana Noemi, pelo amor, carinho e compreensão.

*Tudo tem seu tempo determinado, e  
há tempo para todo o propósito debaixo do céu.*

Eclesiastes. 3.1

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

APLICAÇÃO DE TÉCNICAS DE REDES NEURAIAS ARTIFICIAIS  
NA PREVISÃO DE CURTÍSSIMO PRAZO DA VISIBILIDADE E TETO PARA O  
AEROPORTO DE GUARULHOS - SP

Manoel Valdonel de Almeida

Outubro/2009

Orientadores: Nelson Francisco Favilla Ebecken  
Gutemberg Borges França.

Programa: Engenharia Civil

Um sistema de previsão de curtíssimo prazo (nowcasting) para previsão de teto de nuvens e visibilidade horizontal no Aeroporto de Guarulhos - SP foi desenvolvido e testado operacionalmente. O desenvolvimento do sistema foi baseado em 56 anos de dados meteorológicos da Estação Meteorológica de Superfície de Cumbica/Guarulhos. Três abordagens metodológicas foram testadas: uma com Regressão Linear Múltipla (MLR), duas com Redes Neurais Artificiais (ANN): Probabilistic Neural Network (PNN) e Network General Regression Neural Network (GRNN). Os resultados são apresentados e discutidos para as três abordagens. Em particular, os resultados com PNN indicaram um melhor desempenho, para a previsão de curtíssimo prazo (3 horas) do teto de nuvens e da visibilidade horizontal, do que as outras duas técnicas mencionadas. Em resumo, o desempenho do sistema pode ser considerado aceitável para ser implementado operacionalmente. As respostas do sistema de previsão para as médias de três horas dos acertos e dos alarmes falsos, respectivamente, são: a) 85% e 8% para visibilidade menor que 400 m, b) 92% e 6% para nevoeiro, c) 68% e 6% para teto de nuvem abaixo de 30 m, d) 77% e 5% para teto de nuvem abaixo de 60 m.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

FORECASTING SHORT-TERM OF VISIBILITY AND CEILING FOR AIRPORT  
GUARULHOS - SP BY APPLICATION OF TECHNIQUES OF ARTIFICIAL NEURAL  
NETWORKS

Manoel Valdonel de Almeida

October/2009

Advisors: Nelson Francisco Favilla Ebecken  
Gutemberg Borges França.

Department: Civil Engineering

Based on 56 years of hourly meteorological observation data, a nowcasting system was developed and operationally tested for forecasting cloud ceiling and horizontal visibility at Guarulhos Airport in the State of São Paulo, Brazil. Three methodological approaches were tested namely, Multiple Linear Regression (MLR), Artificial Neural Networks (ANN), specifically the Probabilistic Neural Network (PNN) and the Network General Regression Neural Network (GRNN). The results are presented and discussed for all approaches. In particular, PNN results have indicated better performance than the others aforementioned techniques for nowcasting (3 hours) of cloud ceiling and horizontal visibility. In summary, the system performance could be considered as reasonable one in order to be operationally implemented. The forecasting system responses of hit and false alarm average statistics of 3 hours are as follows: a) 85% and 8% for visibility less than 400 m, b) 92% and 6% for fog, c) 68% and 6% for cloud ceiling less than 100ft, d) 77% and 5% for cloud ceiling less than 200ft, respectively.

## SUMÁRIO

LISTA DE FIGURAS .....	xi
LISTA DE TABELAS .....	xviii
LISTA DE SIGLAS E ABREVIATURAS .....	xxiii
CAPÍTULO 1 – INTRODUÇÃO .....	1
1.1 MOTIVAÇÃO .....	1
1.2 JUSTIFICATIVAS .....	3
1.3 OBJETIVOS .....	6
1.4 ORGANIZAÇÃO DO TEXTO .....	8
CAPÍTULO 2 - FUNDAMENTAÇÃO .....	9
2.1 CONCEITOS BÁSICOS EM METEOROLOGIA .....	9
2.1.1 VISIBILIDADE .....	10
2.1.2 RESTRIÇÕES À VISIBILIDADE .....	14
2.1.3 NEVOEIRO .....	15
2.1.4 SAZONALIDADE .....	19
2.2 PREVISÃO METEOROLÓGICA .....	23
2.2.1 INTRODUÇÃO .....	23
2.2.2 ESCALAS DE PREVISÃO .....	24
2.2.3 MÉTODOS DE PREVISÃO METEOROLÓGICA .....	25
2.3 REDES NEURAIS ARTIFICIAIS .....	30
2.3.1 INTRODUÇÃO .....	30
2.3.2 APLICAÇÕES DAS REDES NEURAIS ARTIFICIAIS .....	33
2.3.3 NEURÔNIO ARTIFICIAL OU UNIDADE DE PROCESSAMENTO .....	40
2.3.4 PROCESSOS DE APRENDIZAGEM .....	42
2.3.5 TIPOS (OU CLASSES) DE REDES NEURAIS .....	44
CAPÍTULO 3 – DADOS E MÉTODOS .....	54
3.1 DADOS .....	54
3.1.1 INTRODUÇÃO .....	54
3.1.2 ANÁLISE PRELIMINAR DOS DADOS .....	61
3.1.2 PRÉ-PROCESSAMENTO DOS DADOS .....	65
3.1.3 CRIAÇÃO DE VARIÁVEIS DERIVADAS .....	68
3.1.4 ESCOLHA DAS VARIÁVEIS PREDITORAS .....	71
3.1.5 CORRELAÇÃO DAS VARIÁVEIS COM A VISIBILIDADE .....	74
3.2 MÉTODOS .....	82
3.2.1 INTRODUÇÃO .....	82
3.2.2 PASSOS DA METODOLOGIA .....	86
CAPÍTULO 4 – RESULTADOS E DISCUSSÃO .....	101

4.1 RESULTADOS PARA A IDÉIA INICIAL.....	102
4.2 DISCUSSÃO SOBRE OS RESULTADOS DA IDÉIA INICIAL .....	107
4.3 PROPOSTA DE NOVAS METAS .....	109
4.4 RESULTADOS COM A RLM.....	109
4.5 RESULTADOS COM REDE GRNN.....	116
4.6 RESULTADOS COM REDE PNN.....	123
4.7 COMPARAÇÃO DOS RESULTADOS COM OS TRÊS MÉTODOS .....	130
4.8 DEFINIÇÃO DO MODELO A SER OPERACIONALIZADO. ....	132
CAPÍTULO 5 – CONCLUSÕES E SUGESTÕES.....	135
5.1 CONCLUSÕES .....	135
5.2 SUGESTÕES.....	137
CAPÍTULO 6 – REFERÊNCIAS BIBLIOGRÁFICAS .....	139
GLOSÁRIO.....	146
APÊNDICE A – SOBRE O REGISTRO DOS DADOS .....	148
APÊNDICE B - DIA JULIANO. ....	158
APÊNDICE C – ESTATÍSTICAS DA RLM. ....	159
APÊNDICE D – OPERACIONALIZAÇÃO DO SISTEMA DE PREVISÃO DE CURTÍSSIMO PRAZO.....	171
ANEXO A – FORMULARIO IEPV 105-78 .....	173
ANEXO B - DESENVOLVIMENTO DA OPERACIONALIZAÇÃO DO SISTEMA DE PREVISÃO .....	174

## LISTA DE FIGURAS

- Figura 2.1 – Modelo de carta de visibilidade de uma hipotética estação de observação meteorológica, com pontos de referência situados a menos de 2 km. A carta é formada por círculos concêntricos com raios crescentes de 200 m, de zero até 2000 m, com indicações de edificações, arvoredos e outros pontos existentes no entorno da estação. A posição do responsável pela observação meteorológica é no ponto central indicado pela carta. .... 12
- Figura 2.2 - Modelo de carta de visibilidade de uma hipotética estação de observação meteorológica, com pontos de referência situados a partir de 2 km. A carta é formada por círculos concêntricos com raios crescentes de 2 km, de zero até 20 km com indicações de edificações, arvoredos e outros pontos existentes no entorno da estação. A posição do responsável pela observação meteorológica é no ponto central indicado pela carta. .... 13
- Figura 2.3 - Ciclo diurno da depressão ou diferença entre a temperatura do ar e a temperatura do ponto do orvalho, em um período de três dias, referentes ao mês de abril de 1998, da Estação de Guarulhos. .... 21
- Figura 2.4 – Semiciclo diurno da variação da pressão atmosférica, em um período de três dias, referentes ao mês de abril de 1998, da Estação de Guarulhos. .... 22
- Figura 2.5 - Ciclo diurno da temperatura do ar, em um período de seis semanas; com as setas em vermelhos, são destacadas as chegadas de sistemas frontais na Estação de Observação Meteorológica de Guarulhos, em 1986. .... 22
- Figura 2.6 - Recuperação de um padrão de referência a partir de uma entrada ruidosa. (Fonte: Von Zuben, 2003). .... 40
- Figura 2.7 – Modelo do Neurônio Artificial  $k$ , onde são destacados os elos de conexão  $xiw_{ki}$ , o combinador linear  $\Sigma$ , a função de ativação  $f$  e a saída  $y_k$ . .... 41
- Figura 2.8 – Exemplos de classificações de RNA: única camada (a), múltiplas camadas (b, c, d); acíclica ou alimentada adiante ou feedforward (a, b, c, d); cíclica ou recorrente ou feedback (e) (Fonte: Braga et al. 2000). .... 45
- Figura 2.9 – Diagrama de uma rede neural de três camadas. As variáveis de entrada, ocultas e de saída são representadas por nós. Os parâmetros adaptativos são representados pela ligação entre os nós. .... 46
- Figura 2.10: A arquitetura em cascata mostrando as unidades de entradas e de saída e duas unidades ocultas acrescentadas durante o processo. As linhas verticais somam todas as ativações de entrada. Os símbolos na parte superior esquerda da figura mostram os tipos de conexões e a função de ativação (Fonte: Chandra e Varghese, 2007). .... 53
- Figura 3.1 – Imagem das pistas do Aeroporto de Guarulhos, nela estão indicados seis sítios meteorológicos (SM) nos quais estão os instrumentos de medidas dos parâmetros meteorológicos. O SM-1 é, atualmente, o sitio principal. A obtenção da imagem foi na internet (Google Maps) e das informações sobre as localizações dos sítios pessoalmente com o ICEA. .... 55
- Figura 3. 2 – Horas anuais de nevoeiro de 1951 a 2007. No gráfico é mostrada a linha de tendência polinomial de sexta ordem com  $R^2$  igual a 0,7324. .... 56



Figura 3. 3 - (a) e (b) Composição colorida (RGB 7,4,2) do satélite Landsat5 TM de 1986 e 2006, respectivamente, que mostra a evolução da ocupação urbana na região da vizinhança do Aeroporto de Guarulhos-SP (Fonte: Freitas et. Al., 2008).58

Figura 3. 4 – Números de ocorrências de nevoeiro em cada hora do dia em cada uma das seis décadas discriminadas no gráfico para a estação meteorológica de Cumbica/Guarulhos. A última década foi extrapolada dos dados dos anos 2001 a 2007. .... 60

Figura 3. 5 – Temperatura média decenal das 7 h e das 19 h, para as décadas discriminadas no gráfico, para a estação meteorológica de Cumbica/Guarulhos. A última década foi extrapolada dos dados dos anos 2001 a 2007. .... 60

Figura 3. 6 – Umidade relativa média decenal das 7 h e das 19 h, para as décadas discriminadas no gráfico, para a estação meteorológica de Cumbica/Guarulhos. A última década foi extrapolada dos dados dos anos 2001 a 2007. .... 61

Figura 3.7 - Histograma de seis intervalos da visibilidade das observações meteorológicas horárias relativas ao período de 01 de janeiro de 1980 a 31 de março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. O primeiro intervalo, de 0-1600 m, com 8.951, é o da visibilidade baixa; os intervalos restantes são considerados visibilidade alta. Existe um total de 230.194 eventos meteorológicos. .... 62

Figura 3.8 - Histograma dos intervalos de visibilidade até 2500 m dos dados fornecidos, mostrando a “preferência” de registro de alguns valores de visibilidade. A amplitude de cada intervalo é de 100 m, com exceção do primeiro que igual a zero. As informações se referem ao período de janeiro de 1980 a março de 2007 da Estação de Observação Meteorológica de Cumbica/Guarulhos. .... 63

Figura 3. 9 – Histograma de seis intervalos de visibilidade relativos aos eventos meteorológicos do período de janeiro de 1951 a março de 2007. O primeiro intervalo, de 0-1600 m, com 21.057, é o da visibilidade baixa; os intervalos restantes são considerados visibilidade alta. O número total de eventos meteorológicos relativos à representação do histograma é 493.056 e se referem à Estação Meteorológica de Cumbica/Guarulhos. .... 63

Figura 3. 10 - Histograma dos intervalos da altura das nuvens (teto) até 300 m. A amplitude de cada intervalo é de 10 m. Percebe-se a “preferência” pelo registro de valores múltiplo de 30 m. As informações se referem ao período do ano de 1998 ao ano de 2007 da Estação Meteorológica de Guarulhos. .... 64

Figura 3.11 - Relação entre o dia juliano e o dia juliano modificado. Na abscissa está o dia juliano e nas ordenadas o valor correspondente do dia juliano modificado. .... 71

Figura 3.12 - Relação entre a hora e a hora modificada. Nas abscissas se encontra a hora do dia e nas ordenadas o valor correspondente da hora modificada. .... 71

Figura 3.13 – Fluxograma com os passos da metodologia empregada para o desenvolvimento do sistema de previsão de curtíssimo prazo para a visibilidade e teto para o Aeroporto de Guarulhos. .... 87

Figura 3.14 - Frequência e percentagem cumulativa dos intervalos de visibilidade do conjunto de treinamento visibT, onde todas as faixas de visibilidade estão bem representadas. Em torno de 20% dos valores de visibilidade estão abaixo de 1000 m. O total de eventos meteorológicos representados é 63.670 e são relativos ao período de janeiro de 1980 a julho de 2005. ....	88
Figura 3.15 - Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento visibB, onde são representados apenas visibilidade baixa (menor ou igual a 1600 m). Nota-se que em torno de 80% dos valores de visibilidade estão abaixo de 1000 m. O total de eventos meteorológicos representados é 14.948 e são relativos ao período de janeiro de 1980 a julho de 2005. ....	88
Figura 3.16 - Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento visib, onde são mostrados valores representativos de todas as faixas de visibilidade, mas predominando os valores de visibilidade baixa (menor ou igual a 1600 m). Nota-se que em torno de 50% dos valores de visibilidade estão abaixo de 1000 m. O total de eventos meteorológicos representados é 48.473 e são relativos ao período de janeiro de 1980 a julho de 2005. ....	89
Figura 3.17 - Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento TESTE00. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005. ....	90
Figura 3.18 - Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento TESTE01. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005. ...	90
Figura 3.19 – Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento TESTE02. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005. ....	91
Figura 3.20 – Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento TESTE03. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005. ....	91
Figura 3.21 – Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento TESTE04. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005. ....	92
Figura 3.22 – Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento TESTE05. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005. ....	92
Figura 3.23 – Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento TESTE06. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005. ....	93
Figura 3.24 – Frequência e percentagem cumulativa da visibilidade do conjunto de treinamento TESTE07. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005. ....	93
Figura 3.25 - Histograma das faixas de visibilidade do conjunto de validação para o preditando visibilidade na primeira hora futura (VISIB1), referente ao período de janeiro de 1980 a julho de 2005. É constituído de 56.489 eventos meteorológicos. ....	95

Figura 3.26 - Histograma dos valores da altura de nuvens da 1ª camada do conjunto total de dados fornecidos para o período de janeiro de 1980 a março de 2007. É constituído de 186.760 eventos meteorológicos. .... 95

Figura 3.27 - Histograma dos intervalos de altura da 1ª camada de nuvens do conjunto de validação para testar a capacidade de generalização do modelo. O período abrange de janeiro de 1980 a julho de 2005. É constituído de 55.000 eventos. .... 96

Figura 4.1 – Fluxograma indicando a seqüência de apresentação da discussão dos resultados iniciais obtidos, da proposta de novas metas, dos resultados com a RLM, a rede GRNN e a rede PNN, da avaliação dos resultados e da operacionalização do sistema..... 101

Figura 4.2 – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.441 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na primeira hora futura (VISIB1) em “< 1000 m” ou “≥ 1000 m”. .... 111

Figura 4.3 – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.441 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na primeira hora futura (VISIB1) em “< 1000 m” ou “≥ 1000 m”..... 112

Figura 4.4– Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 30.192 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na segunda hora futura (VISIB2) em “< 1000 m” ou “≥ 1000 m”..... 113

Figura 4.5 – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 30.192 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na segunda hora futura (VISIB2) em “< 1000 m” ou “≥ 1000 m”..... 114

Figura 4.6 – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.778 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na terceira hora futura (VISIB3) em “< 1000 m” ou “≥ 1000 m”. .... 115

Figura 4.7 – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.778 eventos meteorológicos relativos ao período de

janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na terceira hora futura (VISIB3) em “< 1000 m” ou “≥ 1000 m”..... 116

Figura 4.8 – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de GRNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.694 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na primeira hora futura (VISIB1) em “< 1000 m” ou “≥ 1000 m”. ..... 118

Figura 4.9 – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de GRNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.694 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na primeira hora futura (VISIB1) em “< 1000 m” ou “≥ 1000 m”..... 119

Figura 4.10 – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de GRNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 31.245 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na segunda hora futura (VISIB2) em “< 1000 m” ou “≥ 1000 m”. ..... 120

Figura 4. 11 – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de GRNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 31.245 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na segunda hora futura (VISIB2) em “< 1000 m” ou “≥ 1000 m”..... 121

Figura 4.12 – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de GRNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.908 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na terceira hora futura (VISIB3) em “< 1000 m” ou “≥ 1000 m”. ..... 122

Figura 4.13 – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de GRNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.908 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na terceira hora futura (VISIB3) em “< 1000 m” ou “≥ 1000 m”..... 123

Figura 4.14 – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de PNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.693 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de nevoeiro (SIM) e não-nevoeiro (NÃO) na primeira hora futura. .... 125

Figura 4.15 – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de PNN

(ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.693 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de neveiro (SIM) e não-neveiro (NÃO) na primeira hora futura. .... 126

Figura 4.16 – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de PNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 30.192 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de neveiro (SIM) e não-neveiro (NÃO) na segunda hora futura. .... 127

Figura 4.17– Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de PNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 30.192 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de neveiro (SIM) e não-neveiro (NÃO) na segunda hora futura..... 128

Figura 4.18 – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de PNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.908 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de neveiro (SIM) e não-neveiro (NÃO) na terceira hora futura. .... 129

Figura 4.19 – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de PNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.908 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de neveiro (SIM) e não-neveiro (NÃO) na terceira hora futura. .... 130

Figura 4.20 – Representação da tela de interface do modelo de previsão desenvolvido e instalado no ICEA/Aeroporto de Guarulhos. Na primeira parte da tela é mostrada a situação atual, na segunda parte a previsão para as três horas, e na última parte a estatística de acertos do sistema relativo aos dois últimos anos (agosto de 2005 a julho de 2007). .... 134

Figura A. 1- Imagem mostrando parte do arquivo “sbgr105251” com dados relativos a observações meteorológicas dos dias 1 e 2 de janeiro de 1951. Esses dados são oriundos do formulário IEPV-105-25. .... 148

Figura A.2 - Imagem de parte do arquivo SBGR\_8086.xls destacando o dia 01/01/1981, onde não houve observações meteorológicas ou foram perdidos os registros. .... 150

Figura A.3 - Imagem de parte do arquivo SBGR\_8086.xls, mostrando os dados básicos de trabalho com seus rótulos destacados na primeira linha. São destacados, também, as situações de ausência de nuvens em eventos meteorológicos existentes. .... 151

Figura A. 4 – Imagem do formulário de registro de observações meteorológicas IEPV 105-78..... 173

Figura D. 1 – Fluxograma mostrando as seis etapas operacionais do sistema, bem como o armazenamento dos dados transformados, das previsões realizadas e das estatísticas das previsões. .... 171

## LISTA DE TABELAS

Tabela 1.1 – Mínimos Operacionais para o Aeroporto de Guarulhos – SP. As seis primeiras linhas da tabela se referem às condições mínimas de visibilidade e teto de nuvens para pouso, nas duas condições de voo: VFR (Visual Flight Rules – regras de voo visual) e IFR (Instrument Flight Rules – regras de voo por instrumentos). Na última linha estão as condições mínimas para decolagem (fonte: DECEA)..... 7

Tabela 3. 1 – Estatística da regressão  $UR = f(BSECO, BUMIDO)$  realizada com um conjunto de 65.090 observações meteorológicas de horários sinóticos que contivessem as três variáveis UR, BSECO e BUMIDO, do período anterior a 1998. ..66

Tabela 3. 2 - Correspondência entre as tabelas de nuvens: Tabela 0500 do Manual de Códigos Meteorológicos do Comando da Aeronáutica (primeira coluna), usada anteriormente ao ano de 2005, a tabela criada para o formulário IEPV-105-78 (segunda coluna) e a Tabela de nuvens (terceira coluna), em relação à tabela da segunda coluna, foi incluída a sigla SKC (sky clear) indicativa da condição de céu claro..... 68

Tabela 3. 3 – Correlação entre as candidatas a variáveis de entrada e cada uma das três visibilidades futuras (VISIB1, VISIB2 E VISIB3) realizada com o conjunto de 7.248 eventos de visibilidade baixa (menor ou igual a 1600 m), relativo ao período de janeiro de 1980 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As informações estão ordenadas pelo valor absoluto do fator de correlação. .... 77

Tabela 3. 4 - Correlação entre as candidatas a variáveis de entrada e cada uma das três visibilidades futuras (VISIB1, VISIB2 E VISIB3) realizada com o conjunto de 187.708 eventos de visibilidade em todas as faixas, relativo ao período de janeiro de 1980 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As informações estão ordenadas pelo valor absoluto do fator de correlação. .... 78

Tabela 3. 5- Relação das variáveis preditoras. .... 79

Tabela 3. 6 - Correlações entre as 36 candidatas a variáveis de entrada e a visibilidade na primeira hora futura (VISIB1) realizadas com dois conjuntos de dados. O primeiro com 19.927 eventos de visibilidade baixa (menor ou igual a 1600 m) e o segundo com 50. 889 eventos de visibilidade em todas as faixas. Os dados se referem ao período de janeiro de 1951 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As correlações estão ordenadas pelo valor absoluto do fator de correlação. .... 80

Tabela 3. 7- Correlações entre as 36 candidatas a variáveis de entrada e a visibilidade na segunda hora futura (VISIB2) realizadas com dois conjuntos de dados. O primeiro com 19.792 eventos de visibilidade baixa (menor ou igual a 1600 m) e o segundo com 57. 627 eventos de visibilidade em todas as faixas. Os dados se referem ao período de janeiro de 1951 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As correlações estão ordenadas pelo valor absoluto do fator de correlação. .... 81

Tabela 3. 8- Correlações entre as 36 candidatas a variáveis de entrada e a visibilidade na terceira hora futura (VISIB3) realizadas com dois conjuntos de dados. O primeiro com 19.861 eventos de visibilidade baixa (menor ou igual a 1600 m) e o segundo com 57. 589 eventos de visibilidade em todas as faixas. Os dados se referem ao período de janeiro de 1951 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As correlações estão ordenadas pelo valor absoluto do fator de correlação. .... 82

Tabela 3. 9 (a) e (b) – A parte (a) da tabela está relacionada aos valores de visibilidade menores que 1000 m, e a parte (b) aos valores de 1000 m até 1600 m. Nas duas partes da tabela estão representados os números de ocorrências dos fenômenos meteorológicos (nevoeiro, névoa seca, chuva, chuvisco, e névoa úmida) em cada um dos intervalos de visibilidade da coluna da esquerda, bem como a percentagem total de ocorrência de cada um desses fenômenos em cada uma das partes da tabela. Os dados se referem ao período de 01 de janeiro de 1980 a 31 de março de 2007..... 85

Tabela 3. 10 - Mínimos Operacionais para o Aeroporto de Guarulhos – SP. Na segunda à sétima linha estão as informações relativas aos mínimos de visibilidade e teto de nuvens para pouso, nas duas condições de voo: VFR (Visual Flight Rules – regras de voo visual) e IFR (Instrument Flight Rules – regras de voo por instrumentos). A precisão instrumental nas quatro condições de voo IFR (NDB, VOR/DME, (ILS) CAT I, (ILS) CAT II) é inversamente proporcional ao valor da visibilidade e do teto. Na última linha estão as condições mínimas para decolagem (fonte: DECEA). ..... 85

Tabela 3. 11- Tabela de contingência 2x2 (Fonte: Wilks, 2006)..... 97

Tabela 3. 12– Relação do desempenho de um modelo de previsão em relação à sensibilidade e à especificidade..... 100

Tabela 4.1 – Resultados da aplicação da rede neural PNN, treinada pelo programa NeuroShell Classifier, sobre o conjunto de validação para a classificação nas cinco faixas de visibilidade conforme a idéia inicial, para a primeira hora futura. O conjunto de validação contém 56.489 observações meteorológicas referentes a janeiro de 1980 a julho de 2005. .... 103

Tabela 4.2 - Resultados da aplicação da rede neural PNN, treinada pelo programa NeuroShell Classifier, sobre o conjunto de validação para a classificação nas cinco faixas de visibilidade conforme a idéia inicial, para a segunda hora futura. O conjunto de validação contém 56.489 observações meteorológicas referentes a janeiro de 1980 a julho de 2005. .... 104

Tabela 4.3 - Resultados da aplicação da rede neural PNN, treinada pelo programa NeuroShell Classifier, sobre o conjunto de validação para a classificação nas cinco faixas de visibilidade conforme a idéia inicial, para a terceira hora futura. O conjunto de validação contém 56.489 observações meteorológicas referentes a janeiro de 1980 a julho de 2005. .... 104

Tabela 4.4 – Resumo relativo à sensibilidade e à especificidade, para os resultados relativos à previsão das cinco faixas de visibilidade, para as três horas futuras, conforme a idéia inicial. .... 105

Tabela 4.5 – Resultados da aplicação da rede neural PNN, treinada pelo programa NeuroShell Classifier, sobre o conjunto de validação para a classificação nas seis faixas de teto de nuvens conforme a idéia inicial, para a primeira hora futura. O conjunto de validação contém 55.000 observações meteorológicas referentes a janeiro de 1980 a julho de 2005. .... 105

Tabela 4.6 - Resultados da aplicação da rede neural PNN, treinada pelo programa NeuroShell Classifier, sobre o conjunto de validação para a classificação nas seis faixas de teto de nuvens conforme a idéia inicial, para a segunda hora futura. O conjunto de validação contém 55.000 observações meteorológicas referentes a janeiro de 1980 a julho de 2005. .... 106



Tabela 4.7 - Resultados da aplicação da rede neural PNN, treinada pelo programa NeuroShell Classifier, sobre o conjunto de validação para a classificação nas seis faixas de teto de nuvens conforme a idéia inicial, para a terceira hora futura. O conjunto de validação contém 55.000 observações meteorológicas referentes a janeiro de 1980 a julho de 2005. ....	106
Tabela 4.8 - Resumo relativo à sensibilidade e à especificidade, para os resultados relativos à previsão das seis faixas de teto de nuvens, para as três horas futuras, conforme a idéia inicial .....	107
Tabela 4.9 – Resumo dos melhores resultados das 72 opções de modelos treinados: (3 métodos) x (8 conjuntos de treinamentos) x (três variáveis de saída).....	131
Tabela A. 1- Correspondência entre as tabelas de nuvens: Tabela 0500 do Manual de Códigos Meteorológicos do Comando da Aeronáutica (primeira coluna), usada anteriormente ao ano de 2005, a tabela criada para o formulário IEPV-105-78 (segunda coluna) e a Tabela de nuvens (terceira coluna), em relação à tabela da segunda coluna, foi incluída a sigla SKC (sky clear) indicativa da condição de céu claro.....	149
Tabela A.2 – Relação das Condições Gerais de Tempo, ou do fenômeno meteorológico, que era utilizada para o preenchimento do formulário IEPV 105-25, até o ano de 1979.....	152
Tabela A.3 - Tabela 4677 do Manual de Códigos Meteorológicos (MCA 105-10) mostrando a relação das Condições Gerais de Tempo, ou do fenômeno meteorológico. Esta tabela passou a ser utilizada nos formulários de registros meteorológicos a partir de 1994: IEPV 105-13 e IEPV 105-78.....	153
Tabela B.1 - Quadro mostrando a relação entre dia gregoriano e dia Juliano. As células coloridas indicam o início/término das estações meteorológicas.....	158
Tabela C.1 - Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE00. ....	159
Tabela C.2 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE01. ....	159
Tabela C.3 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE02. ....	160
Tabela C.4 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE03. ....	160
Tabela C.5 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE04. ....	161
Tabela C.6 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE05. ....	161

Tabela C.7 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE06. ....	162
Tabela C. 8 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE07. ....	162
Tabela C.9 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE00. ....	163
Tabela C.10 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE01. ....	163
Tabela C.11 - Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE02. ....	164
Tabela C. 12 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE03. ....	164
Tabela C. 13 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE04. ....	165
Tabela C. 14 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE05. ....	165
Tabela C. 15 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE06. ....	166
Tabela C. 16 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE07. ....	166
Tabela C. 17 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE00. ....	167
Tabela C. 18 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE01. ....	167
Tabela C. 19 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE02. ....	168

Tabela C. 20 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE03. ....	168
Tabela C. 21 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE04. ....	169
Tabela C. 22 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE05. ....	169
Tabela C. 23 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE06. ....	170
Tabela C. 24 – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE07. ....	170

## LISTA DE SIGLAS E ABREVIATURAS

CEPTEC - Centro de Previsão de Tempo e Estudos Climáticos.

DECEA – Departamento de Controle do Espaço Aéreo.

DME - Distance Measuring Equipment.

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

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

ICEA - Instituto de Controle do Espaço Aéreo.

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

INFRAERO - Empresa Brasileira de Infra-Estrutura Aeroportuária.

INPE - Instituto Nacional de Pesquisa Espacial.

MOR - Meteorological Optical Range (Alcance Óptico Meteorológico).

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

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

NEMEF0 - Neural Meteorological Forecast.

NWS - National Weather Service (Serviço Meteorológico Nacional).

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.

RVR - Runway Visual Range - Alcance visual na pista.

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

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. É o sucessor do GMT (*Greenwich Mean Time*) Tempo Médio de Greenwich.

# CAPÍTULO 1 – INTRODUÇÃO

## 1.1 MOTIVAÇÃO

O ser humano por sua natureza inquieta está sempre em movimento, seja este por motivo de trabalho, de estudo, de lazer ou por qualquer outro que o impulsiona a ir de um lugar a outro, perto ou longe de seu local de origem. Mas, muitas vezes, a sua movimentação através de um dos vários meios de transporte terrestre, de navegação aquaviária ou aérea é atrasada ou mesmo impedida por condições meteorológicas adversas. Inegavelmente a navegação aérea sofre grandemente os efeitos das condições desfavoráveis do tempo.

A aviação, de um modo geral, é muito sensível às condições meteorológicas e suas variações, e quando essas condições são adversas elas acarretam muitas vezes perda de tempo e de recursos financeiros e, em alguns casos, a irrecuperável perda de vidas. A sensibilidade dessa variabilidade das condições meteorológicas se faz sentir mais agudamente no pouso e decolagem, a fase mais crítica do voo.

Vários estudos e análises sobre as causas e os efeitos dos atrasos de voo em aeroportos norte-americanos concluíram que os efeitos nocivos do tempo, em relação às condições adversas de visibilidade horizontal e nuvens baixas, provocam prejuízos da ordem de milhões de dólares anuais ((Allan *et al.*, 2001, Valdez, 2000 apud Hansen e Riordan, 2003). Não se tem conhecimento de um estudo semelhante para os aeroportos brasileiros, mas tudo leva a crer que em nosso país a situação seja muito parecida em relação às perdas de recursos financeiros devido às condições meteorológicas adversas.

O conhecimento das condições futuras do tempo no destino e nas alternativas de um voo são fatores importantíssimos para os tomadores de decisão em relação à

navegação aérea para um vôo o mais seguro possível, não somente em relação à segurança patrimonial, mas principalmente em relação à segurança de vidas.

A capacidade de previsão do tempo vem aumentando muito nos últimos tempos graças principalmente ao progresso da computação e suas múltiplas aplicações e utilidades, bem como da utilização de medidas de sensoriamento remoto, como, por exemplo, as de radares meteorológicos e de satélites, etc. No caso da modelagem numérica, esta tornou-se um importante instrumento científico para a evolução do progresso da previsão do tempo. De acordo com José Paulo Bonatti, apud Porto (2002), a modelagem numérica utilizada para fazer a previsão do tempo e clima, nada mais é do que Leis de Newton aplicadas aos movimentos da atmosfera. Estas equações tentam representar os fenômenos que acontecem nela, os processos físicos de formação de nuvens, chuvas, ventos e também a diferenciação de condições dos mares, solos e vegetações. Para isto são utilizados pontos de grades, pelos os quais o globo terrestre é dividido na horizontal, em retângulos ou quadrados, e na vertical, em camadas. Esse conjunto de pontos representa uma grade numérica. Existem várias dimensões de grades, que podem ser de acordo com o fenômeno que se quer representar.

Todo esse esforço computacional demanda grandes recursos. Por exemplo, para fazer o processamento de modelos complexos, como o que reproduz o sistema oceano-atmosfera, ou a previsão para médio prazo, com metodologias avançadas, o Centro de Previsão de Tempo e Estudos Climáticos – CPTEC faz uso atualmente dos supercomputadores SX-6 e SX-4 que possuem capacidade de processar até 768 bilhões e 16 bilhões de operações aritméticas por segundo, respectivamente (<http://www.cptec.inpe.br/ocptec/informacao.shtml>).

Mas, conforme Holton (1992), não importa quão fina seja a resolução de grade do modelo, haverá sempre movimentos cujas escalas sejam ainda menores para serem representadas propriamente nesse modelo numérico.

Apesar do imenso progresso ocorrido com a operacionalidade da *previsão numérica do tempo* nos últimos vinte anos, segundo Pinto (2002), o *ponto morto*, ou a condição de estagnação desta em relação à previsão nas primeiras poucas horas ainda permanece. Isto ocorre devido ao tempo necessário para a análise, a assimilação, o processamento dos dados, e a execução do processo da simulação numérica em si mesma. Isto cria espaço para *nowcasting* ou *previsão de curtíssimo prazo*, a qual apresenta capacidade de previsão para as primeiras poucas horas.

Além do baixo desempenho existente nas primeiras horas da previsão numérica, há ainda a questão da previsão *local*. Muitas vezes a grade não consegue “enxergar” as particularidades de uma determinada localidade. Isto explica as tentativas nos últimos anos para desenvolver métodos de previsão de curtíssimo prazo e, posteriormente, métodos que possam efetivamente combinar os prognósticos dados pela previsão de curtíssimo prazo e modelos de previsão numérica de sistemas de mesoescala.

## 1.2 JUSTIFICATIVAS

A aviação é um importante meio de transporte de recursos materiais e vidas não somente no Brasil como em todo mundo, e não apenas do ponto de vista comercial e de utilidade pública como também porque ela é um extraordinário fator de segurança e de vigilância. Através dela há o movimento de milhões e milhões de pessoas para diversas atividades, lazer, trabalho, estudo, em diversas horas do dia e épocas do ano. Por meio da aviação é feito o de transporte de milhares de toneladas dos mais diversos produtos e equipamentos.

Por isso, o sistema de transporte aéreo como um todo é muito grande e complexo e há o envolvimento de muito recurso financeiro em todo o seu contexto de funcionamento. Uma parte ou subsistema importantíssimo desse conjunto da aviação é o sistema de proteção ao vôo, que procura zelar pela segurança de vôo em todas as suas etapas, tanto no vôo em rota quanto na decolagem e pouso.



Sabe-se que precipitações com grande volume de chuva têm causado imensos transtornos à vida de populações brasileiras como enchentes, deslizamentos de barreiras e encostas, deixando milhares de desabrigados e dezenas de mortos; isto tem sido amplamente divulgado pelos diversos meios de comunicação. Este é um grande problema relativo às condições do tempo. E ele tem causado alguns acidentes aéreos. Mas no caso específico da aviação não é o mais perigoso.

Como já citado anteriormente, o pouso e a decolagem é a parte mais sensível do voo (muitos acidentes aéreos ocorrem nestas etapas da navegação aérea).

No Relatório Anual de 2007 do NCAR (National Center for Atmospheric Research) dos EUA é citado que condições adversas de teto e visibilidade criam sérios riscos para a segurança de voo da aviação em geral, bem como na diminuição da eficiência do fluxo do tráfego de operações comerciais; e que, em uma média anual, riscos de teto e visibilidade resultam de 35 acidentes de aviões, de 70 mortes e em perdas na ordem de 155 milhões de dólares. No mesmo relatório é dito que pesquisas do Research Applications Laboratory - RAL financiadas pelo Programa de Pesquisa Meteorológica para a Aviação da Federal Aviation Administration - FAA dirigidas diretamente à segurança e eficiência reportar-se ao desenvolvimento da melhoria de sistemas de suporte à decisão e informações do tempo sobre teto e visibilidade ([http://www.ral.ucar.edu/lar/2007/goal\\_2/priority\\_2.php](http://www.ral.ucar.edu/lar/2007/goal_2/priority_2.php)).

Grande parte dos acidentes aéreos acontece devido à restrição da visibilidade horizontal causadas por chuvas, névoas ou nevoeiro ou devido à restrição da visibilidade vertical causadas por nuvens baixas, ou por motivo de ventos fortes no momento do pouso ou decolagem. A redução da visibilidade por esses fenômenos meteorológicos, em grande parte, depende das condições meteorológicas locais.

Muito embora todos esses fatores citados sejam perigosos à aviação quanto ao pouso e decolagem, estatisticamente o fenômeno de maior ocorrência para a restrição de visibilidade é o nevoeiro (mais de 90%) e, pela própria definição, é aquele que está mais relacionado a uma visibilidade muito baixa: menos de 1000m. Certamente,

devido a isto, a maioria dos trabalhos sobre redução de visibilidade se restringe ao estudo de nevoeiro (Oliveira, 2002; Pinheiro, 2006; Forkel et al., 1987; Bott et al. 1990; Siebert et al., 1992; Gurka e Mosher, 2000; Von Glasow e Bott, 1999, apud Bott e Trautmann, 2002).

A comunidade científica que inclui, entre outros, meteorologistas, físicos, engenheiros, médicos, biólogos, e ambientalistas têm mostrado interesse em um melhor entendimento de nevoeiro por causa dos seus efeitos, direta ou indiretamente, sobre o cotidiano das pessoas (Murphy e Ehrendorfer, 1987; Murphy, 1994; NAV CANADA, 2002). A perda econômica total associada com o impacto da presença de nevoeiro na aviação, no transporte terrestre e marinho pode ser comparável àqueles causados por tornados ou, em alguns casos, tempestades de invernos ou furacões. O número de artigos que inclui a palavra “nevoeiro” em jornais da Sociedade Americana de Meteorologia é em torno de 4700 (Gultepe *et al.*, 2007).

O Aeroporto Internacional de São Paulo/Guarulhos ao longo de seus 23 anos de operações apresenta números impressionantes: mais de 230 milhões de pessoas embarcaram e desembarcaram pelos terminais de passageiros, através de mais de três milhões de operações de pousos e decolagens que transportaram cerca de oito milhões de toneladas de cargas (fonte: <http://www.infraero.gov.br/aero.php>).

Do Aeroporto de Guarulhos partem e chegam vôos procedentes e com destino a mais de 20 países e mais de uma centena de cidades nacionais e estrangeiras. Atualmente, o Aeroporto de Guarulhos opera com dois terminais com capacidade para movimentar 17 milhões de passageiros/ano. São 44 empresas aéreas nacionais e internacionais, regulares, de carga ou de fretamentos que utilizam esse complexo aeroportuário.

As empresas instaladas na comunidade aeroportuária empregam um total de 25,5 mil profissionais que mantêm o funcionamento desta importante infra-estrutura, durante 24 horas por dia. Entre esses profissionais está o pessoal ligado ao sistema de proteção ao vôo.

Nas pistas de Guarulhos acontecem em média diária quase 500 operações de pouso e decolagem (fonte: <http://www.infraero.gov.br/aero.php>).

Embora seja sabido que qualquer forma de previsão meteorológica específica para um determinado aeroporto não possa eliminar o mau tempo nesse aeroporto, ela ajuda a minorar os efeitos do mau tempo se for capaz de prever com boa precisão as condições de advertência com antecedência suficiente para ser incorporado no planejamento dos tomadores de decisão em relação à segurança de vôo.

Previsão de fenômenos meteorológicos é um assunto difícil principalmente devido à complexidade dos processos físicos envolvidos e à variabilidade dos fenômenos no espaço e no tempo. A técnica com RNA tem evoluído como um ramo da inteligência artificial e é hoje uma reconhecida ferramenta para modelar as complexidades embutidas em sistemas artificiais ou físicos. Devido a seu potencial para representar comportamento não-linear complexo como, por exemplo, a relação entre o tempo futuro e dados disponíveis que contribuem para certas condições meteorológicas, RNA estão crescentemente sendo usadas em muitas áreas da previsão meteorológica e da previsão da qualidade do ar. Alguns exemplos: McCann, 1992; Boznar et al., 1993, Marzban e Stumpf, 1996, 1998; Hsieh e Tang, 1998; Gardner e Dorling, 1998; Santamouris et al., 1999; Hall et al., 1999; Pasini et al., 2001; Lambert, 2001; Niskaa et al, 2004; Jianga, D., et al, 2004; Pereira Filho e Santos, 2006; Nascimento, 2006.

### 1.3 OBJETIVOS

Diante da constatação dos efeitos danosos das condições desfavoráveis do tempo para a aviação, principalmente na sensível etapa do pouso e decolagem e da conclusão de que imensos prejuízos são causados por essas condições meteorológicas adversas; e, considerando o fraco desempenho existente nas primeiras horas da *previsão numérica do tempo*, e que esta embute uma grande carga de recursos empregados; e, considerando, ainda, que *previsão de curtíssimo prazo* é

ainda um assunto palpitante no meio científico, explicado pelas muitos trabalhos em curso sobre o tema; e, considerando que o Aeroporto Internacional de Guarulhos é uma importante estrutura aeroportuária do País, o presente trabalho pretende trazer uma contribuição nesse sentido, o que é algo novo em nosso país. Este trabalho pretende responder as seguintes questões: a) É possível desenvolver um modelo de previsão que possa preencher essa lacuna deixada pelo pela modelagem numérica? b) Esse modelo será capaz de fornecer previsão localizada e em curtíssimo prazo para as faixas de visibilidade que satisfaçam os mínimos operacionais constantes da Tabela 1.1? c) De igual forma, para as faixas de teto baixo da mesma tabela? d) E que grau de precisão terá essa previsão? e) Se essa precisão não for satisfatória, há uma outra opção satisfatória para a problemática da previsão de curtíssimo prazo da visibilidade e teto previsão para o Aeroporto de Guarulhos? f) Havendo essa opção satisfatória, sua implementação operacional será viável? g) E a que custo se dará a operacionalização desse modelo no Aeroporto de Guarulhos?

**Tabela 1.1** – Mínimos Operacionais para o Aeroporto de Guarulhos – SP. As seis primeiras linhas da tabela se referem às condições mínimas de visibilidade e teto de nuvens para pouso, nas duas condições de voo: VFR (Visual Flight Rules – regras de voo visual) e IFR (Instrument Flight Rules – regras de voo por instrumentos). Na última linha estão as condições mínimas para decolagem (fonte: DECEA).

	Visibilidade (m)	Teto (ft)	Teto (m)
VFR	5000	1500	450
IFR	1200	700	210
	800	600	180
	800	200	60
	400	100	30
Decolagem	RVR = 200m; teto = 180m ou 600 ft		

Para responder a essas questões levantadas, este trabalho almeja desenvolver, portanto, um modelo de previsão de curtíssimo prazo para as visibilidades horizontal e vertical em condições adversas à navegação aérea, para o Aeroporto Internacional de Guarulhos.

- a) Portanto, o presente trabalho objetiva desenvolver um modelo de previsão de curtíssimo prazo das visibilidades horizontal e vertical (teto) restritas, por meio de métodos estatísticos com base no comportamento dos dados meteorológicos pretéritos para os fenômenos restritivos dos limites operacionais (teto e visibilidade) para o Aeroporto de Guarulhos. As técnicas a serem utilizadas são Regressão Linear Múltipla (RLM) e Redes Neurais Artificiais (RNA).

#### 1.4 ORGANIZAÇÃO DO TEXTO

O presente trabalho está organizado em seis capítulos, quatro apêndices e dois anexos. Neste presente capítulo, foram apresentadas as motivações e os objetivos, bem como as questões levantadas a serem respondidas ao final da tese. No capítulo 2, são apresentadas as fundamentações sobre o tema e o embasamento teórico sobre rede neurais artificiais bem como as revisões bibliográficas. No capítulo 3 são descritos os dados, seu tratamento e análise preliminar e a metodologia aplicada. No capítulo 4, os resultados obtidos nos vários experimentos realizados são apresentados e discutidos. No capítulo 5 são apresentadas as conclusões e sugestões. As referências bibliográficas utilizadas na presente tese estão no capítulo 6. Logo em seguida, é apresentado o glossário com uma relação de palavras ou expressões técnicas utilizadas. No Apêndice A são apresentados informações adicionais sobre os dados meteorológicos, seus registros e as tabelas usadas na confecção da observação meteorológica; no Apêndice B é mostrada a relação entre dia *juliano* e dia *gregoriano*; no Apêndice C são apresentadas as estatísticas da Regressão Linear Múltipla; no Apêndice D são fornecidos detalhes da operacionalização do sistema de previsão de curtíssimo; no Anexo A é apresentado o formulário de registros das observações meteorológicas atualmente em uso; no Anexo B é apresentado o desenvolvimento da operacionalização do sistema de previsão de curtíssimo.

## CAPÍTULO 2 - FUNDAMENTAÇÃO

### 2.1 CONCEITOS BÁSICOS EM METEOROLOGIA

*Meteorologia* em uma definição simplificada é a ciência que estuda a atmosfera terrestre nos aspectos do *tempo*, seus efeitos, sua previsão e sua estatística (climatologia). O *tempo* (condições meteorológicas) pode ser definido como o estado da atmosfera em determinado instante e lugar (próximo à superfície da Terra). O *clima* geralmente é definido como um “tempo médio”, ou seja, obtido das médias das observações meteorológicas durante um determinado período. As condições do tempo geralmente são descritas em termos de algumas quantidades ou propriedades básicas medidas regularmente. As mais importantes são:

- a temperatura do ar,
- a umidade do ar,
- a pressão do ar,
- a velocidade e direção do vento,
- tipo e intensidade do fenômeno meteorológico presente;
- tipo e quantidade de nuvens.

A atmosfera é uma camada relativamente fina de gases ( $N_2$ ,  $O_2$ , Ar e outros em menor proporção) e material particulado (aerossóis) que envolve a Terra. De fato, 99% da massa da atmosfera estão contidas numa camada de aproximadamente 0,25% do diâmetro da Terra (~32 km). A composição da Atmosfera é relativamente constante, no tempo e no espaço e o *vapor de água* é um dos mais variáveis gases na atmosfera. Nos trópicos úmidos e quentes, ele constitui até 4% do volume da baixa atmosfera, enquanto sobre os desertos e regiões polares constitui apenas algo um pouco menos de 1%. Contudo, sem vapor de água não há nuvens, chuva ou neve (Grimm, 1999).

Conforme Orlanski (1975), os movimentos atmosféricos ocorrem em diferentes escalas que variam de níveis moleculares a níveis planetários, com grandezas temporais correspondentes de frações de segundo até meses ou anos. Essas escalas de movimentos geralmente são classificadas em três categorias principais: micro, meso e macroescala (Orlanski, 1975). O estudo na menor dessas escalas geralmente está limitado aos fenômenos originados dentro da camada baixa da atmosfera que sofre influência da superfície terrestre. Esta camada é conhecida como *camada limite atmosférica* (CLA) ou *camada limite planetária* (CLP). Na microescala são estudados, entre outros assuntos, os fenômenos causadores das restrições à visibilidade horizontal e à vertical.

### 2.1.1 VISIBILIDADE

Segundo o capítulo 10 do Guia para Instrumentos Meteorológicos e Métodos de Observação, da Organização Meteorológica Mundial, (*Guide to Meteorological Instruments and Methods of Observations*, 1983) visibilidade é um fenômeno psicofísico complexo, intimamente ligado a fatores envolvidos na percepção humana. Sua estimativa é sujeita não apenas a variações na habilidade perceptiva e interpretativa individual, como também a características da fonte luminosa e dos fatores de transmissão. Assim, qualquer estimativa da visibilidade tende a ser subjetiva.

O conceito de visibilidade é largamente usado em meteorologia de duas formas distintas. Primeiro, é uma das variáveis identificadores das características de massas de ar; para isto a visibilidade deve ser representativa do *estado óptico* da atmosfera. Segundo, ela é uma variável operacional definido em termos do “alcance da visibilidade”.

O estado óptico da atmosfera depende diretamente das condições meteorológicas. Um conveniente parâmetro derivado do estado óptico da atmosfera é o *alcance óptico meteorológico* (MOR - Meteorological Optical Range). O MOR pode

ser medido, ou aproximadamente estimado, por instrumentos apropriados ou visualmente.

Segundo o Guia para Instrumentos Meteorológicos e Métodos de Observação, citado anteriormente, a visibilidade meteorológica pode ser definida como “a maior distância na qual um objeto escuro de dimensões convenientes pode ser visto e reconhecido contra o horizonte durante o dia ou poderia ser visto e reconhecido durante a noite se a iluminação aumentasse ao nível da luz normal do dia; ou a maior distância em que luzes de intensidade moderada pode ser vistas e identificadas, à noite”.

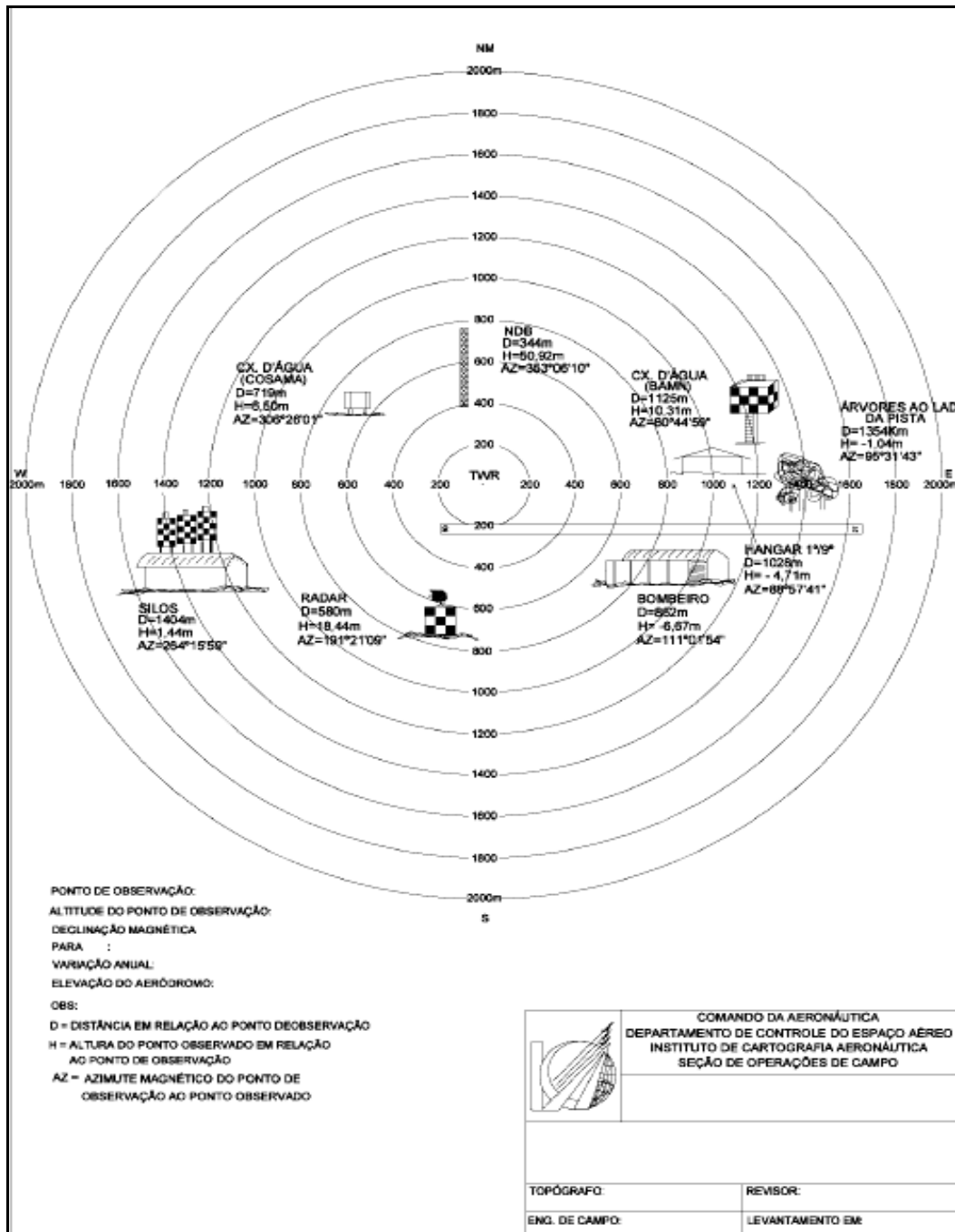
A seguir será feita uma explanação sobre o modo de obtenção da estimativa visual da visibilidade nas estações de observações meteorológicas de superfície do Comando da Aeronáutica. Nas Figuras 2.1 e 2.2 são apresentados os dois modelos de *carta de visibilidade* usados por essas estações. Estes modelos foram obtidos do MCA 105-2 - Manual de Estações Meteorológicas de Superfície, ano 2008, do Comando da Aeronáutica.

Essas cartas de visibilidade são utilizadas no momento da observação meteorológica pelo observador meteorologista, ou alguém substitutivo com treinamento adequado: esse operador se posiciona no ponto central da estação (geralmente próximo à torre de controle do aeroporto) que é o centro dos círculos indicados na carta. Os pontos referenciados nas cartas de visibilidade (edificações, arvoredos, morros, etc.) indicam a distância deles ao ponto central, as suas direções (azimute) e as suas alturas. Baseado nos pontos referenciados nas cartas e no seu treinamento específico, o operador estima a visibilidade horizontal da estação meteorológica.

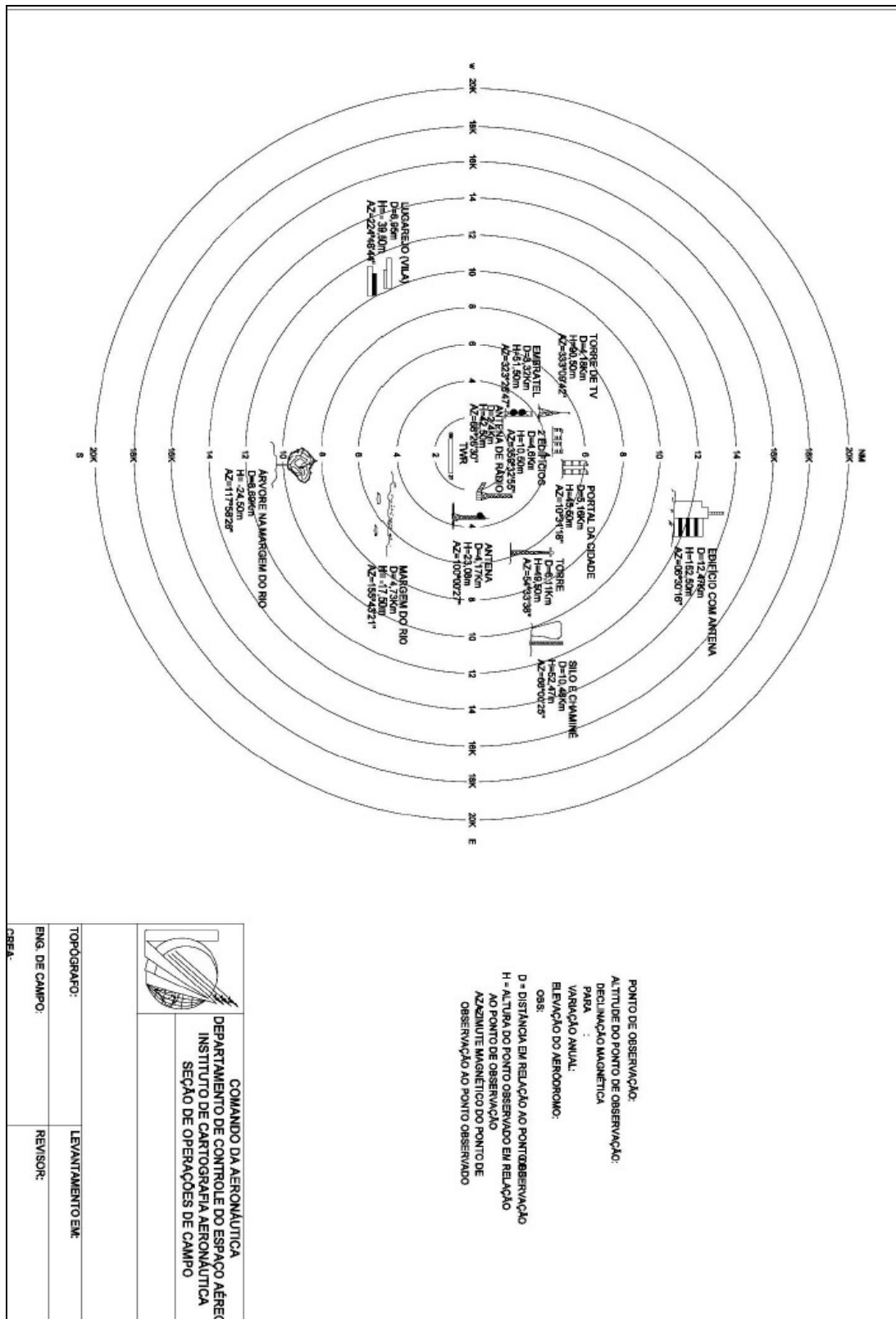
De uma forma similar o observador meteorologista estima os parâmetros relativos às nuvens. Para isto ele recebeu, também, treinamento específico para reconhecer os tipos de nuvens, suas alturas e as suas quantidades. Ele, no ambiente externo à estação, no tempo determinado para a confecção do boletim meteorológico,



olha para o céu e estima visualmente a quantidade, o tipo e a altura das nuvens presentes na estação.



**Figura 2.1** – Modelo de carta de visibilidade de uma hipotética estação de observação meteorológica, com pontos de referência situados a menos de 2 km. A carta é formada por círculos concêntricos com raios crescentes de 200m, de zero até 2000m, com indicações de edificações, arvoredos e outros pontos existentes no entorno da estação. A posição do responsável pela observação meteorológica é no ponto central indicado pela carta.



**Figura 2.2** - Modelo de carta de visibilidade de uma hipotética estação de observação meteorológica, com pontos de referência situados a partir de 2 km. A carta é formada por círculos concêntricos com raios crescentes de 2 km, de zero até 20 km com indicações de edificações, arvoredos e outros pontos existentes no entorno da estação. A posição do responsável pela observação meteorológica é no ponto central indicado pela carta.

## 2.1.2 RESTRIÇÕES À VISIBILIDADE

Em relação à restrição da visibilidade, antes se deve definir *meteoros*: são fenômenos meteorológicos observados na atmosfera, ou à superfície terrestre que podem resultar da suspensão, depósito ou precipitação de partículas sólidas ou líquidas, de natureza aquosa ou não. Alguns meteoros são, ainda, manifestações de natureza ótica ou elétrica. Conforme o Vocabulário Meteorológico Internacional, considerando a constituição e as condições de formação dos meteoros, eles são classificados em quatro grupos:

- hidrometeoros;
- litometeoros;
- fotometeoros;
- eletrometeoros.

Desses quatro grupos, os hidrometeoros e os litometeoros reduzem a visibilidade em uma determinada localidade onde eles estejam ocorrendo.

Ainda, segundo o Vocabulário Meteorológico Internacional:

- 1) hidrometeoro é um meteoro constituído por um aglomerado de partículas de água, na fase líquida ou sólida, em queda livre ou em suspensão na atmosfera, ou levantada da superfície terrestre pelo vento, ou depositadas sobre objetos, no solo ou na atmosfera livre. Os hidrometeoros mais comuns são: Chuva, Chuvisco, Orvalho, Garoa, Neve, Grãos de Neve, Pelotas de Gelo, Cristais de Gelo, Granizo, Nevoeiro, Névoa, Neve, Tornado ou Tromba d'água;
- 2) litometeoro é o fenômeno causado pela suspensão de partículas no ar, geralmente sólidas, de natureza não aquosa. Os litometeoros mais comuns são: Névoa Seca, Fumaça, Poeira ou de Areia, Turbilhão de Poeira ou de Areia, Cinza Vulcânica.

A redução da visibilidade por esses fenômenos meteorológicos, em grande parte, depende das condições locais de estabilidade atmosférica que pode intensificá-los ou enfraquecê-los. Um estado de estabilidade do ar, associado às condições apropriadas de umidade, pode favorecer a formação de nevoeiro, névoa úmida, nuvens baixas com precipitação leve, além de ajudar a manter a névoa seca e a fumaça em níveis baixos; situações essas que reduzem a visibilidade horizontal. Por outro lado, uma condição de ar instável, que produz correntes verticais, tende a colaborar na dissipação de nevoeiro, bem como na dispersão da névoa seca e da fumaça, mas pode favorecer a formação de sistemas convectivos com pancadas de chuvas e conseqüente redução da visibilidade.

Dentre os fenômenos citados anteriormente, o nevoeiro, pela própria definição, é aquele que está mais relacionado a uma visibilidade perigosamente reduzida (menos de 1000 m) para os vários meios de transportes. Em relação à aviação, para esta faixa de visibilidade há um grande potencial de ocorrência de acidentes. Provavelmente, devido a isto, a maioria dos trabalhos sobre redução de visibilidade se restringe ao estudo de nevoeiro (Forkel et al., 1987; Bott et al. 1990; Siebert et al., 1992; Bergot et Guédalia, 1994; Dal Piva et Fedorova, 1999; Von Glasow e Bott, 1999, apud Bott e Trautmann, 2002; Oliveira, 2002; Costa *et al.*, 2006; Pinheiro, 2006; Cermak e Bendix, 2007).

### 2.1.3 NEVOEIRO

O nevoeiro, a névoa e as nuvens “stratus” baixas sempre representam perigo à aviação. Segundo Oliveira (2002), o “vôo por instrumentos”, apresenta recursos técnicos que possibilitam, em grande parte, a solução das dificuldades de vôo. Os vários métodos e recursos desenvolvidos permitem o pouso com nuvens muito baixas e visibilidade horizontal restrita ou mesmo quase nula. Porém, mesmo com auxílio desses recursos, previsões mais aprimoradas das condições meteorológicas

constituem um fator imprescindível às atividades aeronáuticas. Ainda, para Oliveira (2002), no contexto dos problemas que mais implicam cuidados aos pousos e decolagens está o nevoeiro, que reduz a visibilidade horizontal do local de pouso, contribuindo em muito para a ocorrência de acidentes.

Dada a sua importância no contexto deste trabalho, serão apresentados a seguir alguns conceitos básicos sobre nevoeiro, como a definição, classificação, processos de formação.

Nevoeiro é um aglomerado de minúsculas gotículas de água ou cristais de gelo em suspensão próximo ou junto à superfície da Terra que reduzem a visibilidade horizontal, segundo convenção internacional, a menos de 1000m e estando a umidade relativa do ar a 80% ou mais.

Quando há restrição de visibilidade, mas a umidade relativa é menor do que 80%, o fenômeno chama-se *névoa seca* (ou *Fumaça* ou *Tempestade de Poeira* ou de *Areia*, se for o caso de um desses fenômenos estiver ocorrendo).

Quando a visibilidade estiver ainda reduzida, mas com valor igual ou superior a 1000 m, e a umidade relativa do ar ainda estiver acima ou igual a 80%, o fenômeno é denominado *névoa úmida*.

Portanto, a umidade relativa do ar, além da visibilidade, são variáveis determinante da definição do fenômeno restritor da visibilidade; e 1000 m, para a estimativa da visibilidade, e 80%, para a medida da umidade relativa do ar, são os valores limítrofes para a definição de nevoeiro.

O nevoeiro pode se formar quando o ar se torna saturado. Uma parcela de ar se torna saturada, a uma dada temperatura, quando a quantidade de vapor de água existente em seu interior atinge a capacidade máxima da parcela. O processo de saturação pode ocorrer de dois modos: a) por meio do resfriamento da parcela de ar ou b) pela adição de vapor d'água a essa parcela. Já o resfriamento da parcela de ar pode ser radiativo, advectivo ou por expansão (adiabático). Assim, classifica-se o tipo

de nevoeiro de acordo com o seu processo de formação. Essa classificação considera duas classes principais de nevoeiros (Binhua, 1985):

- a) nevoeiros formados por diminuição da temperatura de uma massa de ar:  
*de radiação, de advecção, e de radiação–advecção;*
- b) nevoeiros formados pelo acréscimo de vapor d'água: *frontais e de evaporação.*

O nevoeiro de radiação resulta do resfriamento radiativo da superfície e do ar adjacente. Este tipo de nevoeiro ocorre em noites de céu limpo, com ventos fracos e umidade relativa razoavelmente alta. Stull (1994) descreve que, à medida que o ar próximo ao solo se resfria, a ponto de se igualar à temperatura do ponto de orvalho, inicia-se o processo de formação. Neste momento inicial, o conteúdo de água líquida na base da camada do nevoeiro é maior do que o conteúdo no seu topo mais difuso. Se não há vento (vento calmo), não há mistura e a transferência de calor é apenas por condução. Como o ar é um mal condutor de calor, somente uma fina camada de ar próxima ao solo é resfriada até a saturação, situação que caracteriza o nevoeiro como *raso* (de 1 a 5 m de altura). Eventualmente, o nevoeiro se torna espesso o suficiente para que haja uma maior divergência de fluxo radiativo mais próximo ao seu topo do que na sua base. Neste momento, o resfriamento no topo do nevoeiro gera uma subsidência das camadas mais frias acima, que iniciam um processo de mistura "convectiva" na camada. Rapidamente, o nevoeiro se torna mais homogêneo na vertical, com um topo bem definido.

Conforme Grimm (1999), para um nevoeiro com maior extensão vertical, é necessária uma brisa leve de 3 a 4 kmh<sup>-1</sup>. O vento fraco produz mistura fraca que transfere calor para a superfície fria, fazendo com que uma camada maior se resfrie formando nevoeiro até uma altura maior (10 ou 30 m) sem dispersá-lo. Por outro lado, se o vento se torna mais forte, o ar úmido em níveis baixos se mistura com o ar mais seco acima, com isto a saturação não é atingida e não se desenvolve o nevoeiro de radiação.

O ar resfriado por radiação tende a escoar para áreas mais baixas. Como resultado, o nevoeiro de radiação é mais espesso nos vales do que nas elevações em volta. Normalmente o nevoeiro de radiação se dissipa em 1 a 3 horas após o nascer do sol. Frequentemente se diz que o nevoeiro se "levanta". O que ocorre é que o Sol ao aquecer a Terra ao amanhecer estará aquecendo inicialmente o ar mais próximo à superfície. Conseqüentemente, o nevoeiro evapora a partir da base, dando a impressão de levantamento.

O nevoeiro de advecção ocorre quando ar quente e úmido passa sobre uma superfície fria, resfriando-se por contato e também por mistura com o ar frio que estava sobre a superfície, até atingir a saturação. Portanto, a principal característica dos nevoeiros de advecção é o deslocamento de uma camada de ar quente sobre uma superfície mais fria. Conforme esse deslocamento, alguns tipos de nevoeiro advectivo podem ser citados, como, por exemplo, nevoeiro marítimo citados por Oliveira *et al.* (2001), no qual o ar quente proveniente de uma área marítima de superfície quente desloca-se para outra área marítima de superfície mais fria.

Para Grimm, (1999) certa quantidade de turbulência é necessária para que haja um desenvolvimento maior do nevoeiro. Assim, ventos entre 10 e 30 kmh<sup>-1</sup> são usualmente associados com nevoeiro de advecção. A turbulência não só facilita o resfriamento de uma camada mais profunda de ar, mas também leva o nevoeiro para alturas maiores. Diferentemente dos nevoeiros de radiação, nevoeiros de advecção são frequentemente mais profundos (300-600 m) e mais persistentes.

O nevoeiro orográfico é formado quando ar úmido é elevado sobre um terreno inclinado, como as encostas de colinas ou de montanhas. Devido ao movimento ascendente, o ar se expande e se resfria adiabaticamente. Se nesta situação, o ponto de orvalho é atingido, pode-se formar uma extensa camada de nevoeiro.

Outra situação, quando ar frio se move sobre água mais quente, a água evapora, aumentando a razão de mistura do ar (ou pressão de vapor) que, com suficiente evaporação, pode atingir a umidade relativa de 100%. A saturação ou equilíbrio ocorre

ainda que o ar frio que se move seja um pouco aquecido em contato com a água mais quente. O aumento na umidade relativa, causada pela rápida evaporação, compensa a diminuição da umidade relativa causada pelo aquecimento do ar pela água. Como o ar é aquecido por baixo, ele se torna instável e sobe, e o vapor d'água contido nesta parcela de ar se resfria condensando-se. O nevoeiro então aparece como correntes ascendentes que lembram fumaça ou "vapor". Por esta razão, o nevoeiro produzido quando ar frio entra em contato com água mais quente é denominado nevoeiro de vapor. Ocorre freqüentemente sobre lagos e rios no outono e início do inverno, quando a água pode ainda estar relativamente quente. O mesmo fenômeno também ocorre em dias frios sobre uma piscina externa aquecida. A saturação por adição de vapor pode ocorrer também por evaporação de chuva em ar frio próximo ao ponto de orvalho (nevoeiro frontal ou de precipitação). O nevoeiro de vapor é freqüentemente muito raso, pois quando sobe volta a evaporar no meio não saturado acima.

Existem outros fenômenos como névoa úmida, chuvisco, chuva leve, precipitação em pancada (forte e rápida) e, mais raramente, névoa seca e fumaça que podem reduzir a visibilidade a valores próximos de 1000m; mas as suas ocorrências são em pequena percentagem em relação à ocorrência de nevoeiro, conforme será mostrado ao longo do texto.

Os processos físicos relacionados a alguns desses fenômenos redutores de visibilidade citados anteriormente são bastante diferentes, o que aumenta a dificuldade de relacioná-los a um mesmo *predictando*: a visibilidade.

#### 2.1.4 SAZONALIDADE

*Sazonar quer dizer amadurecer, tornar-se maduro* (dicionário Aulete). Neste sentido, *sazonalidade* é a qualidade do que é sazonal, ou seja, referente à estação ou à época em que os grãos, os frutos estão maduros, prontos para a colheita. Num sentido mais geral, sazonalidade se refere à condição de repetição de um fenômeno em determinada estação ou período de tempo. De um ponto de vista estatístico,

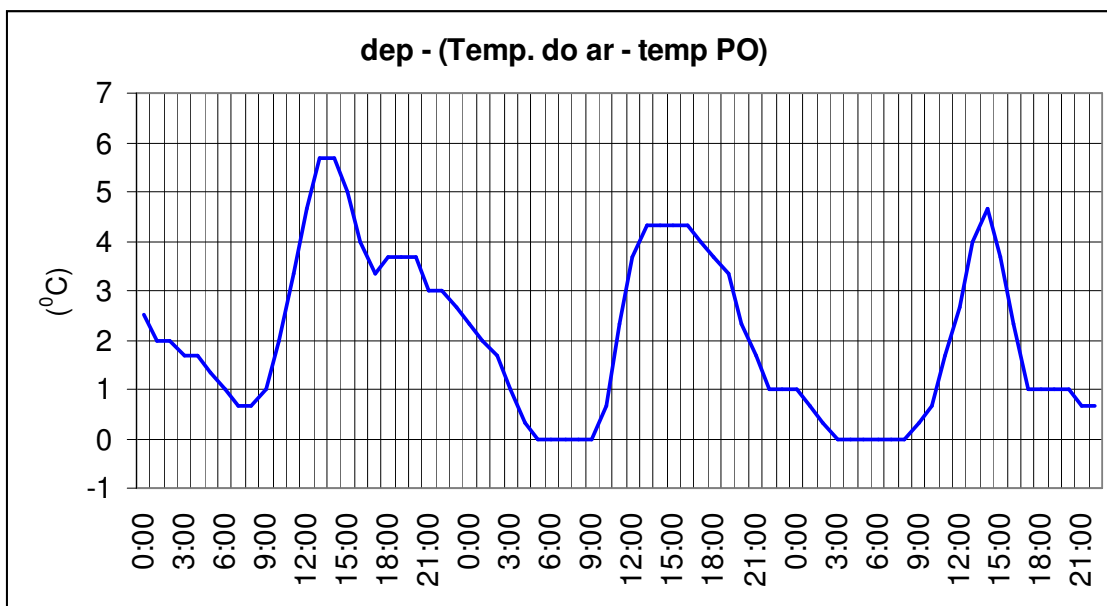


sazonalidade se refere às *variações* ou *ciclos* de padrões idênticos ou quase idênticos que se repetem durante o mesmo período de tempo (anual, mensal, diário, semi-diário, etc.). De um modo geral, na análise da alteração temporal das variáveis usadas em trabalhos de previsão, se percebe o mascaramento, pelo menos parcial, dos eventos transientes pela sazonalidade ou ciclos.

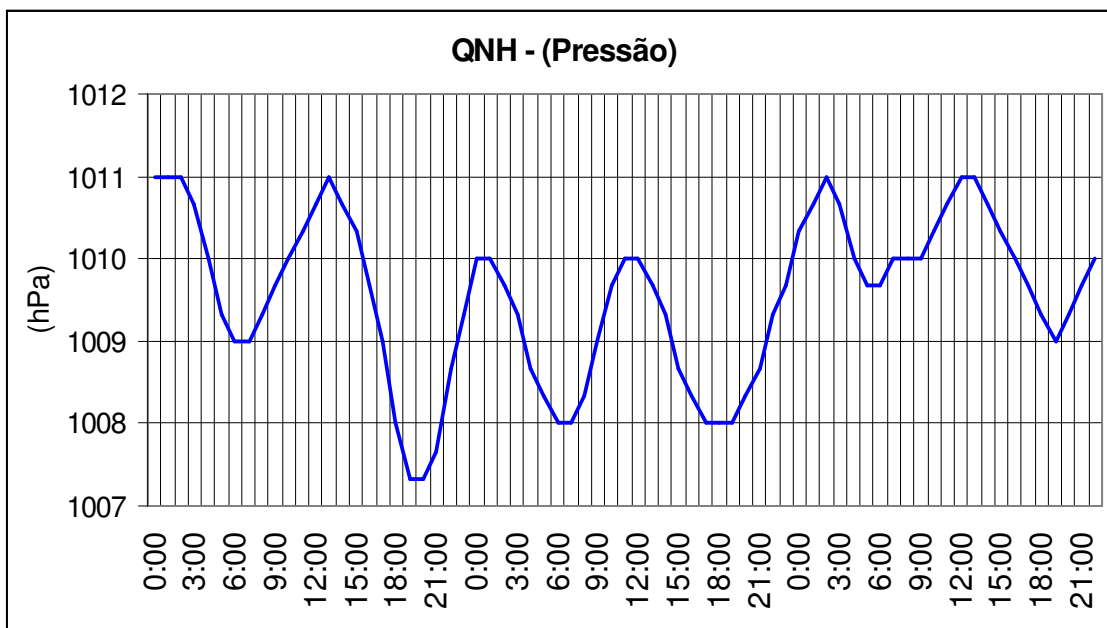
Em um trabalho sobre redes neurais e sazonalidade, Zhang e Min Qi (2005) levantam a questão da sazonalidade; eles citam que vários estudos empíricos sobre previsão de séries temporais sazonais com RNA produziram resultados confusos. Em alguns estudos foi encontrado que RNA são capazes de modelar sazonalidade diretamente, sendo desnecessária uma pré-dessazonalização, outros estudos concluíram justamente o contrário. Os autores, no trabalho citado, investigaram a efetividade do pré-processamento de dados, incluindo a dessazonalização e a exclusão da tendência, na performance da previsão e concluíram que: a) RNA não são capazes efetivamente de capturar as variações da tendência ou sazonalidade com dados brutos sem um pré-processamento; b) retirar a tendência e fazer a dessazonalização pode reduzir fortemente os erros na previsão. O trabalho deles corrobora o que Nelson et al. (1999) apud Zhang e Min Qi (2005) haviam concluído: RNA construídas sem a sazonalidade produziram previsões bem mais precisas que aquelas construídas com dados sazonais. Ainda, segundo Zhang e Min Qi, a utilidade do pré-processamento é que as variações da tendência e da sazonalidade de séries temporais sazonais e tendenciosas poderia responder pelo predomínio de sua variância total. Modelos que ignoram estes padrões sazonais ou tendenciosos resultariam em uma alta variância com fraca performance de previsão. Segundo esses autores, os processos de dessazonalização e eliminação da tendência removem essas grandes variações dos dados brutos, melhorando assim a precisão da modelagem.

Esses estudos mostram a importância do tema e que não deve ser descuidado o tratamento dado à questão da sazonalidade. Neste trabalho este assunto é levado em consideração. Como exemplos, nas Figuras 2.3, 2.4 e 2.5 são apresentados ciclos

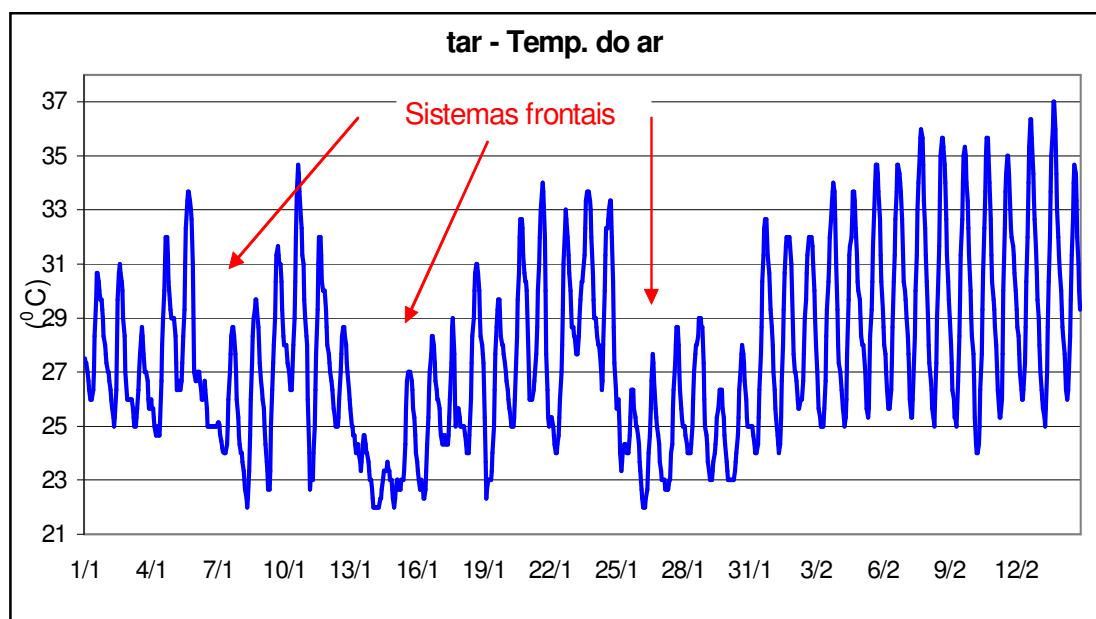
diurnos relativos a algumas variáveis meteorológicas. Na Figura 2.3 é mostrada o ciclo diurno da depressão ou diferença entre a temperatura do ar e a temperatura do ponto do orvalho. Esta depressão está inversamente correlacionada à umidade do ar. Na Figura 2.4 é apresentado o semiciclo diurno da variação da pressão atmosférica. Na Figura 2.5 são mostrados os ciclos diurnos da temperatura, em um período de seis semanas e são destacadas as chegadas de sistemas frontais na estação de observação meteorológica. Da Figura 2.5 se depreende que muito do efeito transiente dos sistemas frontais (como por exemplo, a queda da temperatura) pode ser mascarado pelo ciclo diurno.



**Figura 2.3** - Ciclo diurno da depressão ou diferença entre a temperatura do ar e a temperatura do ponto do orvalho, em um período de três dias, referentes ao mês de abril de 1998, da Estação de Guarulhos.



**Figura 2.4** – Semiciclo diurno da variação da pressão atmosférica, em um período de três dias, referentes ao mês de abril de 1998, da Estação de Guarulhos.



**Figura 2.5** - Ciclo diurno da temperatura do ar, em um período de seis semanas; com as setas em vermelhas, são destacadas as chegadas de sistemas frontais na Estação de Observação Meteorológica de Guarulhos, em 1986.

Como proceder para que a sazonalidade, ou mais especificamente, os ciclos diurnos não influenciem negativamente no treinamento do modelo? Concordando com Zhang e Min Qi, citados anteriormente nesta seção do trabalho, deve ser feita a dessazonalização, para que os efeitos transientes sejam percebidos? Ou deve ser

evitado ao mais máximo modificar os dados para não incorrer em perdas de informações? Como incorporar, então, essas informações dos ciclos ao modelo, para que o mesmo possa diferenciar, por exemplo, uma temperatura de 18°C causada pela advecção de uma massa de ar fria ou devido à hora do dia? No capítulo seguinte serão arrazoados mais detalhes sobre o assunto, ao ser discutida a escolha das variáveis preditoras.

## 2.2 PREVISÃO METEOROLÓGICA.

### 2.2.1 INTRODUÇÃO

O que é previsão?

Se for suposto que uma série temporal é representada pelos valores:

$$x(1), x(2), \dots, x(N)$$

Então, previsão é encontrar os valores futuros  $x(N+1)$ ,  $x(N+2)$ ,...

Previsões do tempo são altamente caóticas, com um grande número de variáveis que afetam as condições de tempo. Em uma série temporal não determinística, é comum que o erro de estimação cresça com o tempo, isto é, se não se consegue obter  $x(N+1)$  corretamente, existe uma alta probabilidade de que em  $x(N+2)$  aconteça um erro ainda maior (Wiberg e Kjellberg, 2001).

A previsão das condições do tempo, segundo Lorenz (1969), é feita sob duas abordagens básicas:

- a) Abordagem Dinâmica - baseada em equações da atmosfera, usa técnicas de elementos finitos, e é comumente chamada modelagem computacional ou previsão numérica de tempo.
- b) Abordagem Empírica - baseada na teoria estatística e, em geral, implicitamente, no "princípio analógico": situações de tempo semelhantes conduzem a resultados semelhantes.

Segundo Holton (1992), na abordagem dinâmica, não importa quão fina seja a resolução de grade, haverá sempre movimentos cujas escalas sejam ainda menores

para serem representadas propriamente no modelo numérico. Assim, há um nível inevitável de erro na determinação do estado inicial. A não linearidade e a instabilidade do escoamento atmosférico, inevitavelmente, causarão os pequenos erros inerentes aos dados iniciais que crescerão e afetarão gradualmente as escalas maiores, de tal maneira que o campo de escoamento atmosférico previsto, eventualmente evoluirá diferentemente do campo de escoamento real.

### 2.2.2 ESCALAS DE PREVISÃO

Previsões meteorológicas em escala sinótica em geral são preparadas e emitidas de quatro a seis horas após a observação meteorológica geradora. Estas previsões se estendem até em torno de sete dias no futuro, dependendo do serviço meteorológico particular que emite a previsão (Pasero e Moniaci, 2004). Previsões de um a dois dias são consideradas previsões sinóticas de *curto prazo*, de três ou mais dias são consideradas de *médio prazo*. Previsões de *mesoescala* possuem uma resolução mais alta e são relativamente mais recentes. Para Doswell, 1986 apud Pasero e Moniaci 2004, previsões de *mesoescala* tipicamente são válidas para um período que se estende de em torno de três horas após a observação meteorológica até um ou dois dias, cobrindo uma área menor (domínio do modelo) que as previsões sinóticas. Mas seu grande atributo é sua habilidade para fornecer uma resolução espacial e temporal maior, com alguns poucos quilômetros de resolução espacial. O intervalo importante entre as condições de tempo presente e o início da validade para uma previsão de mesoescala é o domínio de *nowcasting*. Em termos aproximados, o período de interesse para os previsores de *nowcasting* varia de umas poucas dezenas de minutos até três a seis horas

Para Pasero e Moniaci (2004) *nowcasting* significa uma previsão restrita às próximas três horas e em uma área muito limitada, e baseado nessa abordagem eles desenvolveram o sistema NEMEF0 (Neural Meteorological Forecast), no Instituto Politécnico de Milão, Itália. Esta abordagem não substitui a mensagem meteorológica

tradicional. Ela ajuda a identificar situações locais onde é importante prevenir eventos críticos tais como gelo, neve, nevoeiro. O sistema NEMEF0 usa dados amostrados a cada 15 minutos de uma estação meteorológica e procura prever a evolução desses dados para as próximas três horas. Modelos analíticos estimam esses valores futuros para avaliar o risco de gelo nas pistas ou estradas molhadas ou nevoeiro. Os dados usados pelo sistema são os mais largamente usados em mensagens meteorológicas, como *umidade relativa do ar, velocidade do vento, radiação solar, quantidade de precipitação, pressão atmosférica, temperatura do ar (a 2 m de altitude), temperatura do terreno (a 10 cm de profundidade)*.

Previsões numéricas de condições sinóticas e de mesoescala dependem de solução numérica de um conjunto acoplado de equações para energia, *momentum* e conservação de massa. Elas usam observações meteorológicas para especificar o estado inicial da atmosfera e auxiliar na definição das condições no limite do domínio do modelo. Observações meteorológicas também são assimiladas no processo de modelagem numérica, como uma maneira de harmonizar ou ajustar o modelo às mais recente mudanças no estado da atmosfera (Pasero e Moniaci, 2004).

Em contraste, métodos *nowcasting* dependem fortemente de observações locais das condições do tempo acopladas a uma variedade de formulações empíricas baseadas em regras, e poucas equações teóricas e alguns modelos numéricos (Browning, 1980 apud Pasero e Moniaci, 2004). Para estes autores, métodos de *nowcasting* variam largamente e são feitos sob medida para ajustar a aplicação que deles se faz necessário. Mas a essência de *nowcasting*, em meteorologia, é um conjunto definido de observações e medidas meteorológicas locais.

### 2.2.3 MÉTODOS DE PREVISÃO METEOROLÓGICA

#### A) MÉTODOS ESTATÍSTICOS

Conforme Wilks (2006), muito da meteorologia operacional tem uma base estatística. Como um sistema dinâmico não-linear, a atmosfera não é perfeitamente previsível em um sentido determinístico. Por isso, métodos estatísticos são úteis e certamente necessários como parte integrante de um empreendimento de previsão. Alguns métodos de previsão estatística operam sem informação oriunda de modelos de predição numérica de tempo. Tais métodos de previsão estatísticos puros são algumas vezes chamados de Clássicos. Os métodos estatísticos podem ser objetivos ou subjetivos. Método objetivo significa que para um determinado conjunto de dados de entrada o método sempre produz a mesma saída. Método subjetivo, por sua vez, produz a saída como uma probabilidade. Existem vários métodos de previsão estatística clássicos. Alguns desses métodos são: a) regressão linear; b) regressão não-linear.

Regressão Linear Múltipla (RLM) é a situação mais geral e comum em se tratando de regressão linear para obtenção de uma função de previsão. A regressão linear múltipla apresenta apenas uma variável dependente (Y), mas utiliza mais de uma variável independente (ou preditora). Informações mais completas sobre os fundamentos de regressão linear podem ser obtidas no capítulo 6 do livro de Wilks (2006). A equação de previsão, usando a regressão linear múltipla, é da forma:

$$Y = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i \quad (2.1)$$

onde,

$\beta_0, \beta_1, \dots, \beta_{p-1}$ , são parâmetros a serem determinados e são chamados de coeficientes de regressão. Em particular, o  $\beta_0$  é denominado constante da regressão.

$X_{i1}, X_{i2}, \dots, X_{i,p-1}$  são as variáveis preditoras;

$\varepsilon_i$  é um erro aleatório com média igual a zero e variância constante.

Cada uma das (p-1) variáveis preditoras tem seu próprio coeficiente. Os p parâmetros apresentados na equação (2.1) são obtidos minimizando a soma dos erros

ao quadrado. Os detalhes deste procedimento estão além do escopo deste trabalho. Na prática, estes cálculos são feitos por programas ou pacotes estatísticos.

Um dos métodos estatísticos não-lineares são as Redes Neurais Artificiais. Embora existam outros métodos estatísticos não-lineares, RNA geralmente são muito úteis porque, em adição ao fato de serem capazes de aproximar uma grande classe de funções, comparadas a muitos outros métodos, elas são menos inclinadas a sobreajustar os dados. (Sobre-ajustamento dos dados ocorre quando um método estatístico tem suficiente flexibilidade para ajustar aspectos dos dados que são puramente flutuações estatísticas). RNA serão vistas com mais detalhes mais adiante neste capítulo.

## B) MODELAGEM NUMÉRICA

A capacidade de previsão da ciência vem aumentando desde que o advento da mecânica newtoniana, em fins do século XVII, trouxe um aumento sem precedentes das suas possibilidades preditiva. Uma das características mais marcantes dessa revolução protagonizada pelo cientista inglês Isaac Newton foi a "matematização da física": toda afirmação física deveria ser exprimível por meio de equações matemáticas e as conclusões seriam obtidas por meio da resolução dessas equações e da manipulação de expressões também matemáticas (Diniz, 2002).

Isso deu à mecânica um caráter determinístico radical. Conforme Diniz (2002) se forem conhecidas a posição e a velocidade de todas as partículas de um sistema, em um dado instante, pode-se, em tese, conhecer o seu estado físico em qualquer instante posterior (ou anterior), bastando para isso resolver as equações. Mesmo que essas equações sejam por demais complexas, tornando a sua solução humanamente impossível, a previsão é possível em princípio.

Para Diniz (2002), é possível serem obtidas soluções aproximadas das equações, boas o suficiente para fazer previsões confiáveis. O conceito de aproximação é fundamental. O principal nesse caso, segundo Diniz (2002), é a



escolha de variáveis que têm maior influência no comportamento de um sistema, e a dispensa de outras de menor influência. Para uma escolha criteriosa, faz-se necessário um conhecimento qualitativo prévio sobre o sistema. Por exemplo, para o estudo do clima, são fundamentais variáveis como a velocidade e a direção dos ventos, a pressão atmosférica, umidade e a temperatura em diversos pontos da atmosfera.

Ainda, segundo Diniz (2002), se as variáveis forem bem escolhidas, obtém-se um conjunto de equações que representam o sistema em estudo com boa precisão. Além disso, muitas aproximações simplificam as equações, mas reduzem o grau de confiabilidade dos resultados. Poucas aproximações permitem maior precisão, mas tornam as equações mais complexas. Um sistema pode ser (e em geral é) tão complexo que, para se conseguir uma aproximação que produza equações tratáveis, a confiabilidade ficaria tão pequena que o resultado seria inútil. Com o aumento do poder computacional, serão necessárias cada vez menos aproximações e, assim, é possível se obter resultados cada vez mais precisos.

Há um segundo ponto a ser considerado: de um modo geral não se conhece completamente o estado de um sistema em um dado instante, pois para isto seria necessário se conhecer as várias variáveis em cada ponto da atmosfera. Isso tornaria o número de parâmetros observacionais praticamente infinitos. Por isso, são escolhidos criteriosamente alguns pontos. O conjunto desses pontos constitui o que é chamado de grade numérica.

Uma terceira questão importante é o modo pelo qual são resolvidas as equações. Para Diniz (2002), as equações praticamente nunca têm solução conhecida ou que possam ser resolvidas mediante as técnicas matemáticas conhecidas. Devem, portanto, ser resolvidas computacionalmente por métodos indiretos chamados métodos numéricos. Por esses métodos, as equações não são resolvidas no sentido próprio do termo: obtêm-se, por meios indiretos, soluções aproximadas, e o processo é repetido muitas vezes, de forma que as diversas "soluções" encontradas em cada

passo vão convergindo para algum valor. Muitas vezes são necessários milhões de passos para se obter uma solução confiável.

Após as equações serem resolvidas deve-se, então, interpretar os resultados. Neste ponto é fundamental, conforme Diniz (2002), o conceito de margem de erro. Essa margem será definida pelo grau de aproximação dos diversos passos acima: a escolha inicial das variáveis relevantes, a precisão da medida das variáveis observacionais, a precisão do método numérico utilizado, o tamanho da grade numérica, a adequação dessa grade (pode ser necessário que ela seja mais "fina" em certas regiões). É necessário, ao escolher as variáveis no início do processo, que se quantifique a provável influência no resultado final das variáveis desprezadas.

Hoje há inúmeros mecanismos de obtenção de informação, por isso ela se torna múltipla na sua natureza, variada no seu grau de precisão e extremamente complexa. Para o seu uso eficaz, normalmente é feita a sua conversão num sistema de representação numérico, que seja reconhecido pelos computadores. Dessa maneira são compostos os modelos matemáticos, que consistem num conjunto de registros e parâmetros que traduzem as características e atributos do mundo real. Evidentemente, um modelo é bem mais simples que a realidade, exatamente para que se possa manipulá-lo e fazer experimentações. Para Diniz (2002), à medida que surgem novas possibilidades de observação do sistema real, novos elementos vão sendo introduzidos na representação e o modelo vai se aperfeiçoando.

Porém, para Porto (2002), a distribuição desses dados é totalmente irregular e os modelos necessitam de uma distribuição regular para poder efetuar os cálculos. Há, então, uma etapa do trabalho (a interpolação do campo irregular para o regular) que demanda a técnica de análise objetiva. Isto é o que se denomina resolução do modelo. Há que se dividir o globo terrestre na horizontal, em retângulos ou quadrados, e na vertical, em camadas. A superfície da atmosfera é dividida em quadrados. No centro é preciso ter um valor que se supõe válido para todo o quadrado. Na vertical, os valores trabalhados são divididos como se fossem camadas. Em consequência disso, é feita a

transferência dos dados observados para uma grade do tipo tridimensional, num dado instante fixo. Isto é o que se chama de condição inicial do modelo, uma transferência daquelas variáveis observáveis e em disposições irregulares para um arranjo bem regular. Desta forma, o modelo consegue fazer a leitura.

Conforme Porto (2002), existe um sistema mundial de observação, que é coordenado pela Organização Meteorológica Mundial, chamado Sistema Global de Telecomunicação (GTS). Este sistema é distribuído ao redor do globo e as medições (de responsabilidade de cada país) são feitas em horários pré-definidos. Os dados são todos reunidos em centros regionais, onde estes levantamentos são feitos, sendo o centro mundial mais próximo o de Washington (outros centros mundiais estão localizados em Moscou e Melbourne). Os centros mundiais coletam os dados regionais, depois eles trocam informações entre si e devolvem os dados para os centros regionais. Então, todos têm acesso a essas informações globais, com as quais são preparados os dados que vão entrar nos modelos.

Para Pinto (2002), "Hoje não se trabalha mais sem modelo". Segundo ele, a previsão do tempo hoje é boa, mas ainda está longe do ideal. "É boa ao prever que há 100% de probabilidade que à tarde chova que haja raios, trovões, temporal, em algum lugar. Mas o grande problema é saber com uma precisão maior, onde vai acontecer isso". Completando a idéia do referido pesquisador pode-se perguntar: Será que o Aeroporto de Guarulhos vai está "fechado" por causa do nevoeiro, daqui a uma, duas horas? Isto, hoje, os modelos não conseguem responder. Isto, em parte, porque o tempo mínimo necessário para o modelo produzir a previsão é maior do que essas primeiras poucas horas. Este tipo de resposta, atualmente, está sendo procurada por meio da previsão de curtíssimo prazo ou *nowcasting*.

## 2.3 REDES NEURAS ARTIFICIAIS

### 2.3.1 INTRODUÇÃO

O termo “rede neural” tem sua origem nas tentativas de encontrar representações matemáticas de informações processadas em sistemas biológicos (Bishop, 2006, pg. 226).

A tecnologia de Redes Neurais procura imitar o processo de resolver problemas do cérebro. Assim como o ser humano aplica conhecimento adquirido de experiências passadas para resolver novos problemas ou situações, de igual modo uma rede neural trabalha com exemplos previamente resolvidos para construir um sistema de “neurônios” que tomem novas decisões ou façam classificações e previsões.

Redes Neurais procuram por padrões em conjuntos de dados de treinamento, aprende esses padrões e desenvolve a habilidade de corretamente classificar novos padrões ou fazer previsões e classificações. Redes neurais se destacam em problemas de diagnóstico, de tomada de decisão, de classificação e de outros problemas onde o reconhecimento de padrões é importante, mas respostas computacionais precisas não são tão essenciais.

Então, basicamente, Redes Neurais Artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neuronal de organismos inteligentes e que adquirem conhecimento por meio da experiência. O cérebro é formado por uma rede com uma grande quantidade de processadores simples, chamados de neurônios, interconectados por estruturas ramificadas, no caso os axônios e os dendritos. As sinapses fazem a ligação dos axônios e dendritos entre os neurônios. Apesar de nem todas as características das RNAs possuírem uma inspiração na biologia existem duas que possuem esta origem:

- O conhecimento adquirido por meio de um processo de aprendizado;
- O armazenamento deste conhecimento é feito entre os neurônios, no caso nas sinapses.

Outras características importantes das RNAs

- São modelos adaptativos que apresentam facilidade de treinamento;

- Apresentam a capacidade de representar domínios complexos (não lineares);
- Possuem a aptidão de generalização diante de informações incompletas;
- São algoritmos robustos;
- São capazes de fazer armazenamento associativo de informações;
- São capazes de processar informações espaço / temporais;
- Possui paralelismo intrínseco, o que lhe conferem rapidez no processamento.

A seguir são apresentadas algumas definições de redes neurais artificiais (RNA), extraídas da literatura:

- 1) Redes Neurais Artificiais (RNAs) são sistemas de processamento de informação formados pela interconexão de unidades simples de processamento, denominadas neurônios artificiais. Os neurônios artificiais recebem essa denominação porque foram originados a partir de um modelo matemático de um neurônio natural ou biológico (Von Zuben, 2003).
- 2) RNAs são modelos matemáticos que se assemelham às estruturas neurais biológicas e que têm capacidade computacional adquirida por meio de aprendizagem e generalização (Rezende, 2003, pág. 142).
- 3) Uma RNA é um processador constituído de unidades simples de processamento (neurônios), distribuídas paralelamente, e que têm a capacidade de armazenar conhecimento experimental e torná-lo disponível para o uso. Esse conhecimento é adquirido a partir do ambiente por meio de um processo de aprendizagem e é armazenado por meio de parâmetros ponderados (pesos sinápticos) (Haykin, 2002).
- 4) Redes neurais são algoritmos que buscam imitar, de forma simplificada, o cérebro no processamento de informações. São baseadas em processadores elementares chamados neurônios (Calôba, 2002).
- 5) Rede Neural Artificial é um sistema constituído por elementos de processamento interconectados, chamados neurônios, os quais estão

dispostos em camadas e são responsáveis pela não-linearidade e pela memória da rede (Valença, 2005, p.5-6).

- 6) RNAs são sistemas paralelos distribuídos compostos por unidades de processamento simples (nodo) que calculam 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. Em geral, estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo (Braga et al., 2000, p.1).

Na literatura atual há, também, certa diversidade de nomenclatura de redes neurais: *redes neurais*, *redes neurais artificiais*, *rede neural artificial*, *RNA* ou *RNAs*. Em relação a isto, neste trabalho, se adotará indistintamente qualquer um desses termos. Quanto à definição de redes neurais, há uma “simpatia” maior pela de Haykin, item 3 anterior:

*Uma RNA é um processador constituído de unidades simples de processamento (neurônios), distribuídas paralelamente, e que têm a capacidade de armazenar conhecimento experimental e torná-lo disponível para o uso. Esse conhecimento é adquirido a partir do ambiente por meio de um processo de aprendizagem e é armazenado por meio de parâmetros ponderados (pesos sinápticos).*

### 2.3.2 APLICAÇÕES DAS REDES NEURAIS ARTIFICIAIS

Para quê serve as redes neurais?

Segundo Luk et al. (2000), devido às suas funcionalidade e capacidade poderosas, RNA fornecem uma abordagem alternativa para muitos problemas de engenharia difíceis de serem resolvidos por abordagens convencionais. Previsão de fenômenos meteorológicos é um assunto difícil, principalmente devido à complexidade dos processos físicos envolvidos e à variabilidade do fenômeno no espaço e no tempo. E, para estes autores, a abordagem com RNA evoluiu como um ramo da

inteligência artificial e é agora uma reconhecida ferramenta para modelar complexidades embutidas em qualquer sistema artificial ou físico; e que a estrutura de modelagem com RNA tem sido crescentemente usada pela ciência e engenharia por causa da sua habilidade de modelar tanto sistema linear quanto não-linear sem a necessidade de serem feitas hipóteses adicionais como as que são implícitas na maioria das abordagens estatísticas tradicionais.

Redes Neurais Artificiais (RNA) são largamente utilizadas em diferentes campos de pesquisa e desenvolvimento. Redes neurais têm sido usadas com sucesso para modelar séries temporais em uma variedade de aplicações na ciência e na engenharia. O leque de aplicações usando RNA em ciências atmosféricas é grande. Alguns exemplos: Bremnes e Michaelide (2005) fez uso de redes neurais para a previsão probabilística de eventos extremos de visibilidade; Jianga (2004) mostrou o progresso do uso de RNA na previsão de um índice de poluição do ar; Osródka *et al.*,(2002) usou RNA para a previsão de curto-prazo da precipitação baseada em dados de radar meteorológico; Ramirez *at al.* (1999) fez uso de RNA, transformadas de Ondeletas e dados de imagens do satélite GOES-8 para estimativa de precipitação; Ramirez *at al.* (2005) trabalhou com técnicas de redes neurais para a previsão de chuva na região de São Paulo. Alguns trabalhos se referem, particularmente, à problemática da previsão do tempo para a aviação: previsão de curto prazo de visibilidade e teto de nuvens (Leyton e Fritsch, 2003; Marzban *et al.*, 2005 Oliveira, 2002; Bott e Trautmann, 2002).

Mais alguns exemplos do uso de RNA: aplicações para converter chuva em previsão de cheias de rios (Luk *et al.*, 2000; French *et al.*, 1992); Pereira Filho e Santos (2006) desenvolveram um modelo que usa Redes Neurais Artificiais, dados telemétricos e de radar meteorológico para simular e prever o nível e correnteza na Bacia do Rio Tamanduatei, em São Paulo.

Ainda, mais exemplos do uso de redes neurais artificiais: Santamouris *at al.* (1999) usou uma abordagem de redes neurais para modelagem de ilhas de calor em

áreas urbanas; Yao *at al.* (2000) trabalhou com transformadas de Ondeletas e redes neurais para a previsão de curto prazo de carga elétrica.

Moutinho (2005), com um banco de imagens contendo exemplos de *pele* e *não-pele* desenvolveu, baseado em redes neurais artificiais MLP, um sistema de detecção de pele, com isto diminuindo o espaço de busca para o posterior desenvolvimento de um sistema de localização de faces. Este sistema atingiu um índice de acertos de 95,6%.

Braga (2006), em sua dissertação de mestrado desenvolveu um sistema para realizar a previsão da cotação de fechamento do IBOVESPA (índice da Bolsa de Valores de São Paulo), utilizando estudos estatísticos, conceitos de séries temporais e redes neurais. O mesmo alcançou resultados interessantes e motivadores.

Härter e Campos Velho (2005) analisaram o desempenho de duas redes feedforward (perceptron de múltiplas camadas e função de base radial) e duas redes recorrentes (Elman e Jordan) na previsão operacional de tempo, usando o *sistema de Lorenz sob regime caótico* como problema teste.

Pasini, A., Pelino (2005) apresentaram uma análise da previsibilidade para várias regiões do “atrator do sistema 63 de Lorenz”, um modelo não linear simples que mimetiza alguns aspectos da atmosfera, como seu comportamento caótico e a presença de estados preferenciais ou “regimes”. Nesta estrutura, por meio de uma atividade de previsão no atrator, um *perceptron de múltiplas camadas* mostrou capacidade em reconhecer diferentes valores de previsibilidade em várias zonas do atrator. Esse artigo deve ser considerado como uma tentativa preliminar no uso de redes neurais para avaliação da previsibilidade.

Em resumo, a aplicação de redes neurais se estende, hoje, a quase todos os campos da atividade humana. A seguir é apresentada uma relação de algumas áreas de aplicação de redes neurais, a qual foi montada a partir de consultas a diversos centros de estudos e pesquisas nacionais e estrangeiros. É claro que esta relação não



pretende apresentar todas as áreas de aplicação de RNA, mas apenas mostrar o potencial de uso de redes neurais em áreas bem diversificadas:

#### Aplicações médicas

- Acompanhamento de Sinais Fisiológicos;
- Previsão de Custos de Tratamento;
- Diagnóstico de Doenças;
- Previsão de Risco de Malignidade em Anormalidades em Exames Radiológicos.
- Previsão de Respostas de Longo Prazo às Infecções por meio de Medidas de Curto Prazo

#### Aplicações em Ciências Atmosféricas e Ambientais

- Predição de Modelagem Hidrometeorológica;
- Melhoria na qualidade de água potável;
- Previsão do Nível de Marés;
- Previsão da Qualidade do Ar;
- Previsão Meteorológica (de precipitação, de nevoeiro, eventos extremos, como tempestades, tornados, etc.);
- Previsão das Condições Meteorológicas para Pouso e Decolagem em Aeroportos;
- Previsão da Temperatura da Superfície do Mar;
- Previsão de Irradiação Solar.

#### Aplicações em Finanças, Negócios e na Indústria:

- Monitoramento da Qualidade de Produtos;
- Identificação Química com Nariz Artificial;
- Previsão de Vendas de Grande Volume de Bens de Consumo;
- Otimizando de Estratégia de Preços;
- Predizendo a Resposta de Consumidor;
- Selecionando Metas de Auditoria;

- Programação de Horários do Quadro de Funcionários;
- Seleção de Mercados-alvo;
- Indústria petrolífera;
- Otimização de Processos de Manufatura;
- Seleção de Empregados;
- Previsão de Bolsa de Valores;
- Classificação de Candidatos.

Na continuidade da resposta à pergunta “para que servem as redes neurais artificiais?”, será dada ênfase à funcionalidade, em detrimento dos aspectos conceituais. Não será analisado o potencial das ferramentas computacionais inspiradas em redes neurais artificiais, e sim serão considerados problemas genéricos e de grande interesse prático e a possibilidade de suas resoluções via redes neurais artificiais.

De um modo geral, conforme Von Zeuban (2003), problemas tratados com sucesso com redes neurais artificiais normalmente são aqueles de *análise multivariada* de dados, e que esteja em uma das três situações básicas descritas a seguir:

- a) Os dados correspondem a relações de entrada-saída de fenômenos observados, dos quais se podem extrair amostras.
- b) A amostra descreve apenas uma lista de atributos característicos de um evento ou objeto. Não há relações expressas de entrada-saída entre variáveis independentes e variáveis dependentes.
- c) Os dados correspondem a amostras ruidosas ou com conteúdo parcial de memórias previamente armazenadas.

A informação contida nas amostras, na primeira situação, pode advir de relações de *causa e efeito*, *atributos e classes*, *sintomas e diagnósticos*, etc. Normalmente, estas relações expressam associações não-lineares entre as variáveis independentes (vetor de entrada) e as variáveis dependentes (vetor de saída).

As informações que se pretende extrair, na segunda situação, podem ser: encontrar agrupamentos de pontos; os tipos de atributos de cada agrupamento, ou seja, o que caracteriza cada agrupamento; a possibilidade de classificar uma nova amostra, de acordo com seus atributos, em um desses agrupamentos.

Na terceira situação, o objetivo é reconhecer uma informação ruidosa e/ou incompleta, que é apresentada, por meio de informações armazenadas; ou seja, recuperar a memória armazenada a partir de uma variante ruidosa ou incompleta presente na amostra.

Em seguida serão apresentados alguns exemplos, adaptados de Von Zeuban (2003), onde é possível a utilização de redes neurais artificiais (RNA).

### **Exemplo 1: Aproximação de Funções**

Uma função é, na verdade, um mapeamento único do domínio para o contradomínio, ele representa uma função da forma:  $f: \mathbb{R}^n \rightarrow \mathbb{R}^1$ . Esta função toma  $n$  variáveis de entrada como argumentos e mapeia em uma variável de saída. A questão, então, é descobrir qual é a relação funcional entre as variáveis de entrada e a variável de saída, restrita a uma região fechada e limitada do domínio, denominada de região de operação.

### **Exemplo 2: Classificação de Padrões**

A partir de amostras de entrada-saída contendo, no vetor de entrada, uma lista de atributos numéricos e, na saída, a classe a que a amostra pertence obter um classificador de padrões que, quando novos vetores de entrada sejam apresentados, eles sejam classificados em uma das classes disponíveis.

Exemplos: classificação de textos, de imagens, de sons e de qualquer outro tipo de objetos, eventos ou pessoas que possam ser descritas por atributos numéricos.

### **Exemplo 3: Identificação de Sistemas Dinâmicos**

A partir de amostras de entrada-saída de um sistema dinâmico contendo, no vetor de entrada, as entradas e o estado atual do sistema e no vetor de saída o novo estado do sistema, obter um modelo para o sistema dinâmico, capaz de reproduzir a evolução temporal de seu estado.

#### **Exemplo 4: Simulação de Atividades Humanas (ou de ações animais)**

A partir de amostras de entrada-saída contendo, no vetor de entrada, um conjunto de estímulos atuantes em um ser humano e, no vetor de saída, as respectivas ações resultantes desses estímulos, obter um modelo que simule o comportamento do ser humano. Exemplos: Sistemas de diagnóstico médico; nariz artificial para reconhecimento de drogas nos aeroportos e aplicações que envolvem segurança no trabalho.

#### **Exemplo 5: Predição de Séries Temporais**

A partir de amostras da evolução de uma série temporal, obter um preditor capaz de estimar valores futuros da própria série a partir do comportamento passado e atual da referida série temporal. Exemplos, produzir preditores para:

- ocorrência de abalos sísmicos;
- eventos financeiros, como cotações em bolsas de valores;
- eventos hidrometeorológicos, como vazão de rios e represas, ocorrência de tempestades, redução da visibilidade horizontal.

#### **Exemplo 6: Agrupamento de Dados ou Categorização**

A partir de um conjunto de amostras em um espaço multidimensional, em que as coordenadas representam atributos numéricos e cada amostra vai corresponder a um ponto em um espaço multidimensional, possivelmente de elevada dimensão, deve-se responder as seguintes perguntas:

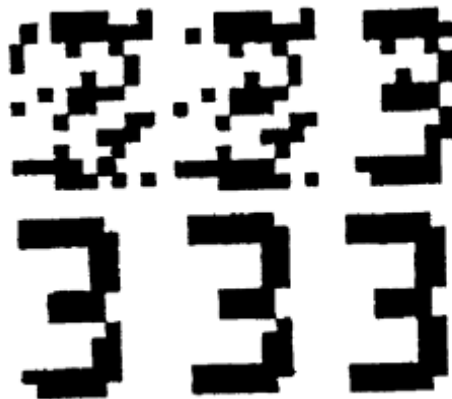
- a) existem agrupamentos ou categorias nesta distribuição de pontos?

- b) caso afirmativo, quantos são eles?
- c) para cada agrupamento, existem sub-agrupamentos?
- d) o que caracteriza cada agrupamento?
- e) é possível associar novas amostras aos agrupamentos existentes?

Exemplo: identificação de classes de usuários de cartão de crédito.

### Exemplo 7: Memória Endereçável por Conteúdo

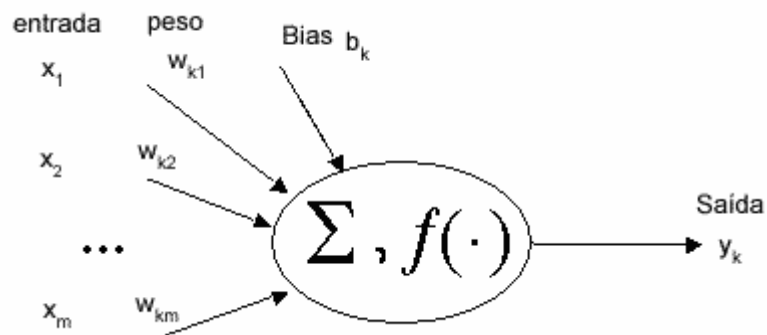
Dados  $n$  padrões de referência, que normalmente correspondem a vetores de atributos, quando confrontado com qualquer novo vetor de atributos na entrada, apontar qual é o padrão de referência associado. De um modo geral, nesse novo vetor de entradas os dados correspondem a amostras ruidosas ou com conteúdo parcial dos padrões de referência. O objetivo, portanto, é recuperar a memória armazenada a partir de sua variante ruidosa ou incompleta presente na amostra. Na Figura 2.6 é apresentada uma imagem mostrando a recuperação de um padrão de referência a partir de uma entrada ruidosa.



**Figura 2.6** - Recuperação de um padrão de referência a partir de uma entrada ruidosa. (Fonte: Von Zuben, 2003).

#### 2.3.3 NEURÔNIO ARTIFICIAL OU UNIDADE DE PROCESSAMENTO

Os neurônios são as unidades de processamento das Redes Neurais e, portanto, são fundamentais para o seu funcionamento. Na Figura 2.7 é mostrado o esquema de um neurônio  $k$  qualquer.



**Figura 2.7** – Modelo do Neurônio Artificial  $k$ , onde são destacados os elos de conexão  $x_i w_{ki}$ , o combinador linear  $\Sigma$ , a função de ativação  $f$  e a saída  $y_k$ .

Os elementos básicos de um neurônio artificial são:

1. **Conjunto de sinapses ou elos de conexão**, onde cada elo de conexão é caracterizado por um estímulo de entrada ( $\mathbf{x}$ ) e por um peso ( $\mathbf{w}$ ).
2. **Combinador Linear** ( $\Sigma$ ): efetua o somatório dos sinais ponderados de entrada do neurônio.
3. **Função de ativação**: Produz a saída do neurônio que é calculada da seguinte forma:

$$y_k = f\left(\sum_{i=1}^m w_{ki} x_i + b\right)$$

onde:

$w_{ki}$  são os  $m$  pesos das conexões do neurônio  $k$ ;

$x_i$  é o valor de cada um dos  $m$  estímulos que chegam ao neurônio  $k$ ;

$b$  é o valor do *bias* que será somado ao valor do combinador linear;

$y_k$  é a saída do neurônio  $k$ .

Embora cada neurônio ou elemento processador execute uma função simples, a RNA, que pode possuir até várias dezenas ou algumas centenas deles, tem, como um todo, a capacidade computacional para a resolução de problemas complexos.

#### 2.3.4 PROCESSOS DE APRENDIZAGEM

Segundo Haykin (2003):

*Aprendizagem é um processo pelo qual os parâmetros livres (ou pesos) de uma rede neural são adaptados por meio de estímulos do ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira como ocorre a modificação dos parâmetros.*

Essa definição implica que a rede:

- 1) *é estimulada* pelo ambiente;
- 2) *sofre modificações* nos seus parâmetros livres como resultado desse estímulo.

Em resumo pode-se dizer que a essência da aprendizagem é “Espera-se que a rede se torne bem treinada de modo que aprenda o suficiente sobre o passado para generalizar bem o futuro” (Haykin, 2003).

De um modo geral, as principais características das Redes Neurais Artificiais são: não-linearidade, mapeamento de entrada-saída, aprendizagem, processamento paralelo, tolerância a falhas, generalização. Mas, a característica de importância fundamental de uma rede neural é a sua habilidade de *aprender* a partir de seu ambiente. O processo de aprendizado é iterativo e por meio dele a RNA deve *melhorar* o seu desempenho gradativamente à medida que interage com o meio ambiente, o que deve resultar em melhor *generalização*.

Existem dois grandes paradigmas de aprendizagem:

- a) Aprendizagem supervisionada, quando existe um conjunto de exemplos de entrada-saída fornecido pelo ambiente externo que orienta o treinamento da rede.
- b) Aprendizagem não-supervisionada; neste caso não há pares de entrada-saída definidos para orientar o treinamento.

Baseados nos dois grandes paradigmas de aprendizagem citados, existe um grande número de regras de aprendizagem as quais, no final, são casos particulares dos dois grandes paradigmas. A seguir são mencionadas algumas dessas regras:

- a) *Aprendizagem por correção de erro* (supervisionada); durante o aprendizado é aplicada uma seqüência de ajustes corretivos aos pesos sinápticos, procurando minimizar o erro ou a diferença entre a saída gerada pela rede e a saída desejada.
- b) *Aprendizagem por reforço* (supervisionada); neste caso, embora haja apenas dados de entrada, há um agente externo (supervisor) que retorna um sinal de reforço ou de penalidade a cada saída da rede.
- c) *Aprendizagem competitiva* (não-supervisionada); nesta regra, os neurônios competem entre si, para responder a um determinado estímulo. Neste processo há um conjunto de neurônios diferenciados por seus pesos, que respondem de forma diferente a um estímulo; e há um mecanismo que permita a competição entre os neurônios, onde apenas um possa responder a uma classe de estímulos de entrada: *o vencedor leva tudo*.
- d) *Aprendizagem hebbiana* (não-supervisionada) é baseada nos seguintes postulados:
  - 1. Se a ativação de um neurônio tiver grande influência na ativação de um outro neurônio, a ligação entre estes dois neurônios deve ser reforçada;
  - 2. Se a ativação de um neurônio não influencia a ativação de um segundo neurônio, a ligação entre estes dois neurônios deve ser enfraquecida ou até mesmo removida;



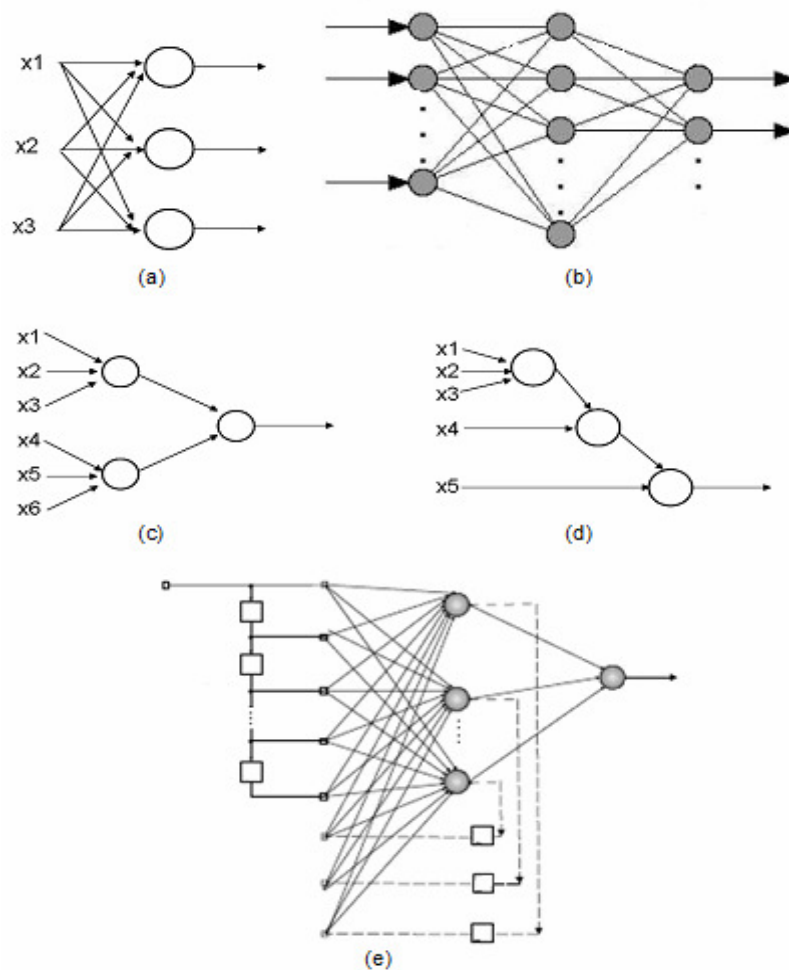
- e) *Aprendizagem baseada em memória (não-supervisionada)*; neste processo, padrões são armazenados, e posteriormente, recuperados, na presença de estímulos a eles relacionados. Neste caso, a memória, além de ser recuperável por meio de estímulos, também deve influenciar as ações.

A fase de aprendizagem consiste de um processo iterativo de ajuste dos parâmetros (pesos) da rede. Um conjunto preestabelecido de procedimentos bem-definidos para ajustar os parâmetros de uma RNA durante o aprendizado chama-se *algoritmo de aprendizagem*. Como é de se esperar, como há um grande número de regras de aprendizagem há, também, um vasto número de algoritmos de aprendizagem. Cada um deles mais apropriado a um determinado tipo de rede e/ou problema.

### 2.3.5 TIPOS (OU CLASSES) DE REDES NEURAIIS

De um modo geral, as redes neurais podem ser classificadas quanto:

- a) Ao número de camadas:
- única camada (Figura 2.8 a);
  - múltiplas camadas (Figura 2.8 b, c, d);
- b) às conexões:
- Acíclica ou Alimentada Adiante ou *feedforward* (Figura 2.8 a, b, c, d);
  - Cíclica ou Recorrente ou *feedback* (Figura 2.8 e).



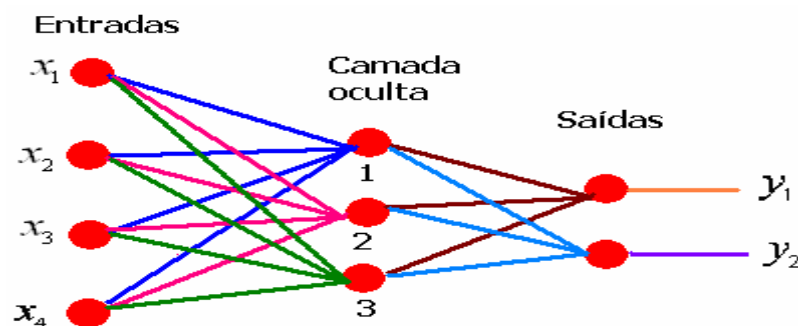
**Figura 2.8** – Exemplos de classificações de RNA: única camada (a), múltiplas camadas (b, c, d); acíclica ou alimentada adiante ou *feedforward* (a, b, c, d); cíclica ou recorrente ou *feedback* (e) (Fonte: Braga *et al.* 2000).

Há um grande número de tipos de Redes Neurais Artificiais que abrange estas classificações (quanto ao número de camadas e às conexões). A seguir são relacionados alguns tipos de rede neural encontrados na literatura.

1. Multilayer Perceptron (MLP) ou *múltiplas camadas e alimentada Adiante (feedforward)*. Tipicamente, esse tipo de rede consiste de um conjunto de unidades sensoriais que constituem a *camada de entrada*; uma ou mais *camadas ocultas* e uma *camada de saída* de nós computacionais. O sinal de entrada se propaga para frente por meio da rede, camada por camada (Haykin, 2002).

2. Redes auto-organizativas ou *self-organizing* - Baseadas na aprendizagem competitiva (o *neurônio vencedor leva tudo*). Exemplo é a rede de Kohonen - inspirada nos mapas corticais (Braga *et al.*, 2000).
3. Rede RBF ou *radial-basis function*. É uma rede de uma ou mais camadas que utiliza *funções de base radial* nos nós da camada oculta (Braga *et al.* 2000).
4. Redes construtivas. Neste tipo, a arquitetura de rede, inicialmente pequena, cresce de tamanho durante o treinamento. Algoritmos como este recebe o nome de algoritmos construtivos (*growing*) (Braga *et al.*, 2000).
5. Redes de Jordan e Elman. São redes alimentadas adiante parcialmente recorrentes. Na de *Elman*, a realimentação se dá entre neurônios de uma mesma camada oculta; já na de *Jordan*, a realimentação se dá da saída da rede para as entradas (Braga *et al.*, 2000).
6. Redes de Hopifeld. Redes recorrentes auto-associativas. Também conhecida como Memória Associativa, as Redes Hopfield armazenam padrões que são recuperados a partir de estímulos de entrada (Haykin, 2002).
7. Redes híbridas. Esta classe combina dois ou mais tipos de subsistemas, sendo um deles uma RNA. Exemplo: redes neurais e métodos estatísticos; redes neurais e algoritmo genético (Braga *et al.*).

A representação de um exemplo de rede de *múltiplas camadas e alimentada adiante* é mostrada na Figura 2.9.



**Figura 2.9** – Diagrama de uma rede neural de três camadas. As variáveis de entrada, ocultas e de saída são representadas por nós. Os parâmetros adaptativos são representados pela ligação entre os nós.

Matematicamente, a rede representada na Figura 2.9 pode ser escrita como:

$$y_1(\mathbf{x}, \mathbf{w}) = \sigma \left( \sum_{j=0}^M w_{1j}^{(2)} f \left( \sum_{i=0}^D w_{ji}^{(1)} x_i \right) \right)$$

$$y_2(\mathbf{x}, \mathbf{w}) = \sigma \left( \sum_{j=0}^M w_{2j}^{(2)} f \left( \sum_{i=0}^D w_{ji}^{(1)} x_i \right) \right).$$

onde,

- a) Os sobrescritos (1) e (2) indicam respectivamente, a primeira camada (entrada) e a segunda camada (oculta).
- b) Os parâmetros  $w$  com os índices  $i$  e  $j$  iguais a “0” são *bias*, os quais permitem um deslocamento fixo nos dados. Para englobar os *bias* foram criadas as variáveis fictícias  $x_0 = 1$  e  $f_0(\cdot) = 1$ .
- c) Os parâmetros  $w$  restantes (com índices diferentes de “0”) são os *pesos* ou *parâmetros adaptativos* ou *coeficientes* a serem determinados.
- d) As funções  $f$  e  $\sigma$  são funções diferenciáveis, não-lineares chamadas, respectivamente, de *função de ativação* aplicada aos neurônios da camada oculta e de *função de ativação* aplicada aos neurônios de saída.

Assim, o modelo de rede neural analisado aqui é simplesmente uma função não-linear multivariada de um conjunto de variáveis de entrada  $\{x_j\}$  para um conjunto de variáveis de saída  $\{y_k\}$  controlada por um vetor  $\mathbf{w}$  de parâmetros ajustáveis.

A busca pela melhor configuração possível dos parâmetros ajustáveis, configuração que representa o armazenamento de conhecimento obtido a partir do ambiente (variáveis de entrada) é feita, como já citado anteriormente, por meio de um processo de *aprendizagem*.

Se após tiver sido atingida essa melhor configuração, a rede consegue dar boa resposta a novos estímulos (novas entradas) então se diz que a rede *generaliza* bem e está apta a disponibilizar o conhecimento adquirido para o uso operacional.

Em um procedimento de *aprendizagem por correção de erro*, considera-se que  $y_k(n)$  seja a resposta de um neurônio de saída  $k$  e seja  $d_k(n)$  a *resposta desejada*. Consequentemente é produzido um *signal de erro* representado por  $e_k(n)$ :

$$e_k(n) = d_k(n) - y_k(n)$$

A finalidade do processo é diminuir  $e_k(n)$  a um valor mínimo possível, este procedimento de correção de erro é feito por meio de um algoritmo de aprendizagem. Um dos mais conhecidos algoritmos para a rede de *múltiplas camadas e alimentada adiante* é o *back-propagation* ou algoritmo de retropropagação (Braga *et al.*, 2000).

A seguir são fornecidas algumas informações adicionais sobre as redes RBF e duas de suas variantes *Probabilistic Neural Networks* - PNN e *Generalized Regression Neural Networks* – GRNN e sobre o *Algoritmo de Correlação em Cascata* tendo em vista a utilização dessas ferramentas neste trabalho.

Uma *rede de função de base radial* (RBF, *radial-basis function*), em sua forma mais básica consiste de três camadas com procedimentos totalmente diferentes. A camada de entrada é constituída por nós de fonte (unidades sensoriais) que conectam a rede ao seu ambiente. A segunda camada, em geral, a *única* oculta da rede, aplica uma transformação não-linear do espaço de entrada para o espaço oculto, que na maioria das aplicações é de alta dimensionalidade. A camada de saída é linear, fornecendo a resposta da rede ao padrão (sinal) de ativação aplicado à camada de entrada (Haykin, 2002).

Funções radiais representam uma classe de funções cujo valor diminui ou aumenta em relação à distância de um ponto central. Diferentes funções de base radial têm sido utilizadas em redes RBF. Conforme Braga *et al.* (2000), as mais comuns são:

- Função gaussiana:  $f(u) = \exp\left(-\frac{v^2}{2\sigma^2}\right)$ ;
- Função multiquadrática:  $f(u) = \sqrt{v^2 + \sigma^2}$ ;
- Função *thin-plate-spline*:  $f(u) = v^2 \log(v)$ ;

onde:

$v = \|x - \mu\|$  é a distância euclidiana ( $x$  é o vetor de entrada e  $\mu$  representa o centro da função radial);

$\sigma$  representa a largura da função radial.

As funções radiais apresentadas possuem dois parâmetros: o centro  $\mu$  e o raio  $\sigma$ . A distância euclidiana  $\|x - \mu\|$  do vetor de entrada  $x$  ao vetor centro  $\mu_i$  serve de entrada para a função, que retorna o valor de ativação da unidade intermediária. Segundo Braga *et al.* (2000), a resposta gerada por um dado neurônio de saída  $\eta_j$  da rede é definida pela equação:

$$y_j = \sum_{i=1}^K w_{ji} \phi(\|x - \mu_i\|) + w_{j0}.$$

## REDES NEURAIS DE REGRESSÃO GENERALIZADA – GRNN

A GRNN foi inicialmente proposta por Specht (1991). Ela é uma rede neural de três camadas com um neurônio na camada intermediária para cada padrão de treinamento. O número de neurônios na camada de entrada é o número de entradas do problema e o número de neurônios na camada de saída é o número de saídas. O treinamento da GRNN é realizado pela comparação de um padrão de entrada no espaço dimensional de todos os padrões de treinamento para determinar quão longe o padrão de entrada está dos demais padrões. A saída é uma predição proporcional à distância do padrão de entrada a todos os outros padrões (Walter, 2003).

Marquez e Hill, 1993 *apud* Walter, 2003 compararam a capacidade de aproximação das redes *backpropagation* (BPN) e GRNN utilizando dados de 14 funções sob três níveis de ruído.

Prado *et al.*, 2001 *apud* Walter, 2003 propuseram a utilização de informações georeferenciadas e redes neurais como uma ferramenta de análise e gestão para uma empresa de distribuição de energia. A modelagem das séries e sua subsequente predição foram realizadas através de redes GRNN. O resultado de predição da demanda comparado com uma região geográfica planejada através de métodos usuais levou a um erro menor que 5%.

Em Gordon, 1995 *apud* Walter, 2003, modelos de redes neurais são apresentados para prever a temperatura da superfície do mar com um ano de antecedência baseado na série temporal derivada de um anel de corais do Pacífico. Duas diferentes redes neurais foram utilizadas e comparadas, um modelo *feedforward backpropagation* e uma GRNN utilizando um algoritmo genético para adaptar dinamicamente seus fatores de suavidade.

## REDES NEURAIIS PROBABILÍSTICAS – PNN

A Rede Neural Probabilística foi inicialmente proposta por Specht (1990). Ela é uma rede neural de arquitetura direta, multi-camadas com mapeamento não linear da entrada para a saída. A classificação realizada por esta rede é baseada na teoria de decisão Bayesiana e ela realiza a estimação da função de densidade de probabilidade (Ciarelli, 2008).

As Redes Neurais Probabilísticas (PNNs) ganharam este nome devido suas raízes se basearem na Teoria da Probabilidade. Specht, 1990 *apud* Moura Jr. *et al.* 2002 mostrou que pela organização do fluxo de operações em camadas, e assegurando operações primitivas para neurônios individuais em cada camada, o algoritmo pode ser utilizado para atuar como uma rede feedforward de quatro camadas com função de ativação exponencial. Para Moura Jr. *et al.* 2002, a vantagem

quanto ao tradicional Backpropagation consiste em que a PNN possui um aprendizado mais rápido, classificando tão bem ou melhor que uma Backpropagation. As PNNs, com o número suficiente de dados, podem convergir para o Classificador Bayesiano que é definido como ótimo.

Moura Jr. *et. al.* (2002) utilizaram PNN para simular e comprovar o processo de tomada de decisão em um servidor de internet. Eles acharam que as redes PNN é um tipo apropriado de RBF utilizado para problemas de classificação. Os resultados mostraram que foram resolvidos os problemas de decisão de servidores nos casos testados.

Schmitz (2006) em sua tese de doutorado testou as redes neurais PNN para a resolução do problema da estabilidade e divisão de fases na simulação de processos de separação de sistemas complexos. Ele encontrou que as PNN apresentaram uma excelente precisão.

Costa e Souza Filho, 2009 utilizaram um sistema de redes neurais artificiais para reconhecimento de padrões de solos. Eles encontraram que o sistema de classificação com rede PNN apresentou o melhor resultado.

## ALGORITMO DE CORRELAÇÃO EM CASCATA

O algoritmo de correlação em cascata - Cascade Correlation (CC) foi desenvolvido por Fahlman *et al.* (1991). Ele faz crescer dinamicamente o número de neurônios e de camadas internas de uma rede neural durante o treinamento. A rede começa encontrando *relações lineares* entre entrada e saída. Valores ponderados são assinalados às ligações entre os neurônios de entrada e de saída. Após essas relações terem sido encontradas, novos neurônios são adicionados, então, à camada oculta de modo que *relações não-lineares* possam assim serem encontradas.

O algoritmo CC combina duas idéias:

- A primeira é a arquitetura em cascata, na qual neurônio ocultos são adicionados um de cada vez. Quando um novo neurônio intermediário é

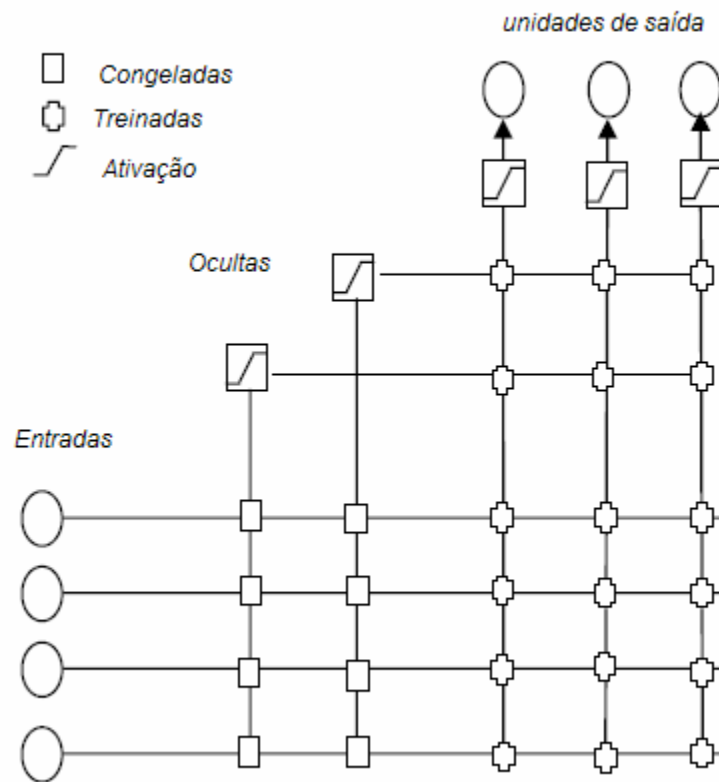


inserido, os pesos de suas entradas (conexões) são congelados. Este neurônio começa então a influenciar permanentemente nas operações da rede, sendo utilizado para detectar novas características.

- A segunda é o seu algoritmo de aprendizagem, o qual cria e instala novas unidades escondidas. Para cada nova unidade escondida, o algoritmo tenta maximizar a magnitude da correlação entre a saída da nova unidade e o sinal de erro residual da rede.

Em outras palavras, o critério de seleção do neurônio candidato é a correlação que este neurônio tem com a saída da rede. Quanto maior a correlação, mais influência este neurônio candidato tem no resultado da rede. Portanto, os pesos das conexões entre os neurônios candidatos e as camadas de entrada e intermediárias devem ser definidos de maneira a maximizar a correlação entre esses neurônios e a saída. Assim, o neurônio candidato que apresentar o maior valor de correlação será incluído na rede como uma nova camada intermediária, sendo conectado a todas as outras camadas da rede. (Vargas *et al.* 1998).

A arquitetura em cascata é ilustrada na Figura 2.10 que mostra as unidades de entradas e de saída, duas unidades ocultas acrescentadas durante o processo. As linhas verticais somam todas as ativações de entrada. Os símbolos na parte superior esquerda da figura mostram os tipos de conexões e a função de ativação. De um modo geral, o número de entradas e de saídas é determinado pelo tipo de problema. Cada entrada é conectada a cada unidade de saída por uma conexão com peso ajustável.



**Figura 2.10:** A arquitetura em cascata mostrando as unidades de entradas e de saída e duas unidades ocultas acrescentadas durante o processo. As linhas verticais somam todas as ativações de entrada. Os símbolos na parte superior esquerda da figura mostram os tipos de conexões e a função de ativação (Fonte: Chandra e Varghese, 2007).

## CAPÍTULO 3 – DADOS E MÉTODOS

### 3.1 DADOS

#### 3.1.1 INTRODUÇÃO

Os dados utilizados na pesquisa, fornecidos pelo Comando da Aeronáutica, através do ICEA (Instituto de Controle do Espaço Aéreo), são observações meteorológicas de superfície da Estação Meteorológica de Cumbica/Guarulhos.

A Estação Meteorológica de Guarulhos passou a existir em 1985, substituindo a Estação Meteorológica de Cumbica situada na Base Aérea de Cumbica. Na Figura 3.1 é apresentada uma imagem atual das pistas do Aeroporto de Guarulhos, nessa imagem observam-se os caracteres indicativos de seis sítios meteorológicos. Antes do ano de 1985, a Estação de Observação Meteorológica estava situada na Base Aérea de Cumbica próxima ao atual Sítio Meteorológicos 2 (SM-2), conforme a Figura 3.1. Depois, a Estação foi transferida para o Aeroporto de Guarulhos, com o seu ajardinado meteorológico localizado atualmente no Sítio Meteorológico 1 (SM-1) onde está concentrada a maioria dos instrumentos de medida usados para registros das informações meteorológicas. Portanto, os dados meteorológicos utilizados, relativos à Guarulhos, se referem ao SM-1 e, os relativos à Cumbica, se referem ao SM-2.

Uma curiosidade relativa ao Aeroporto de Guarulhos é sobre o seu, ainda popularmente conhecido, nome *Cumbica* que em tupi-guarani quer dizer *nuvem baixa* ou *nevoeiro* ([http://www.servicos.hd1.com.br/ventonw/aero\\_climasp.htm](http://www.servicos.hd1.com.br/ventonw/aero_climasp.htm)). Realmente, pela análise que será apresentada adiante, neste capítulo, sabe-se que ocorria uma grande quantidade de horas anuais de nevoeiro em algumas décadas passadas.

Os dados meteorológicos utilizados abrangem o período de janeiro de 1951 a março de 2007. Por questões técnicas, eles não foram fornecidos todos de uma única vez. Inicialmente foram disponibilizados os dados referentes ao período de janeiro de

1980 a março de 2007. Posteriormente, os dados mais antigos: de janeiro de 1951 a dezembro de 1979.



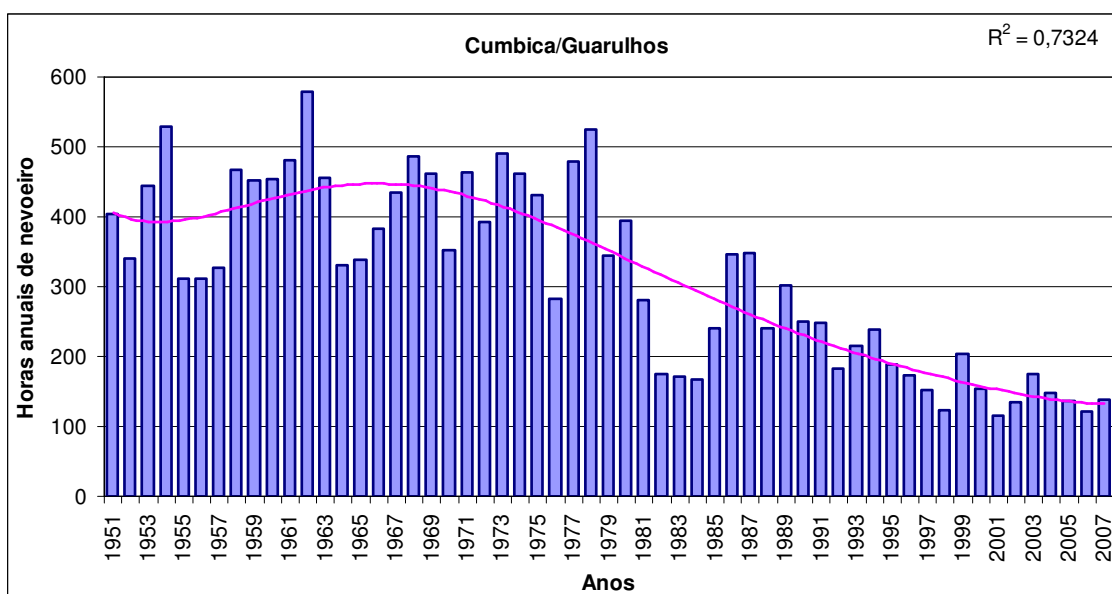
**Figura 3.1** – Imagem das pistas do Aeroporto de Guarulhos, nela estão indicados seis sítios meteorológicos (SM) nos quais estão os instrumentos de medidas dos parâmetros meteorológicos. O SM-1 é, atualmente, o sitio principal. A obtenção da imagem foi na internet (*Google Maps*) e das informações sobre as localizações dos sítios pessoalmente com o ICEA.



Como já citado no Capítulo 1, estatisticamente o fenômeno de maior ocorrência para a restrição de visibilidade é o nevoeiro (mais de 90%) e que, devido a isto, a maioria dos trabalhos sobre redução de visibilidade se restringe ao estudo de nevoeiro. Em sintonia com este enfoque, será apresentada a seguir uma análise sobre o assunto em relação à Cumbica/Guarulhos.

Na Figura 3.2, é apresentado um gráfico com as horas anuais de nevoeiro dos anos de 1951 a 2007. Foi adicionada ao gráfico a linha de tendência de sexta ordem com  $R^2$  igual 0,7324. Das informações do gráfico, observa-se que o valor médio das horas de nevoeiro é, em torno de:

- 420 horas, para os anos de 1951 a 1980;
- 210 horas, para os anos de 1981 a 2007;
- 146 horas, para a década de 1997 a 2006.



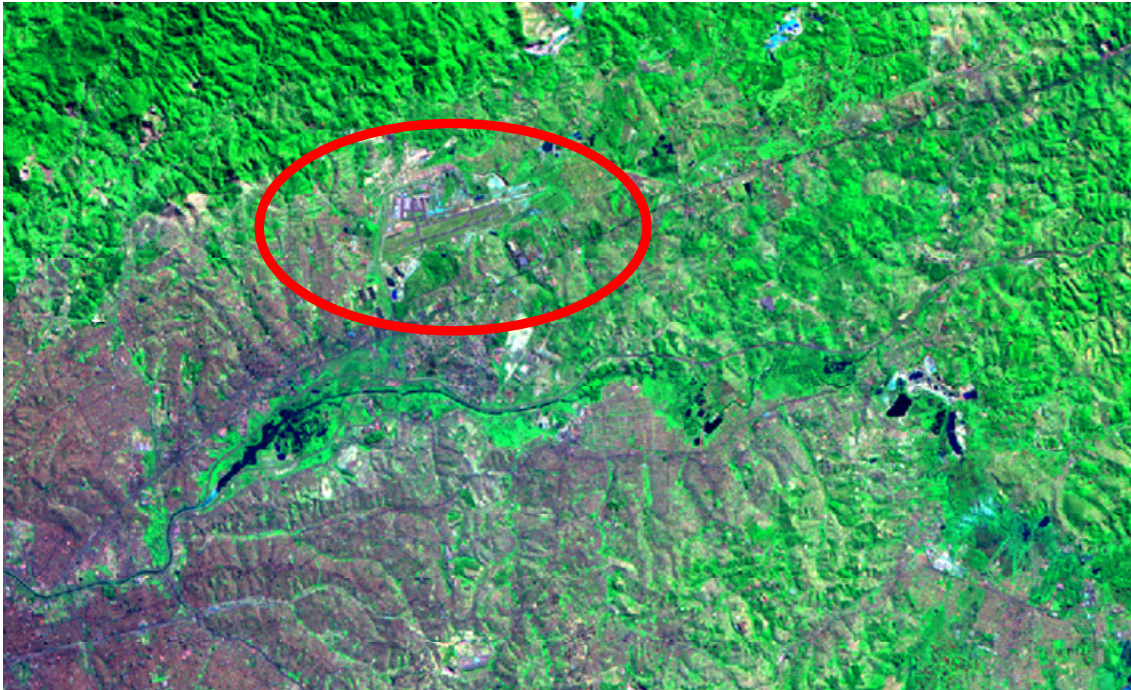
**Figura 3. 2** – Horas anuais de nevoeiro de 1951 a 2007. No gráfico é mostrada a linha de tendência polinomial de sexta ordem com  $R^2$  igual a 0,7324.

Pela Figura 3.2, observa-se nitidamente a tendência de queda do número de horas anuais de nevoeiro ao longo do tempo no período estudado. Percebe-se, também, que a linha de tendência, concomitantemente com a indicação de queda,

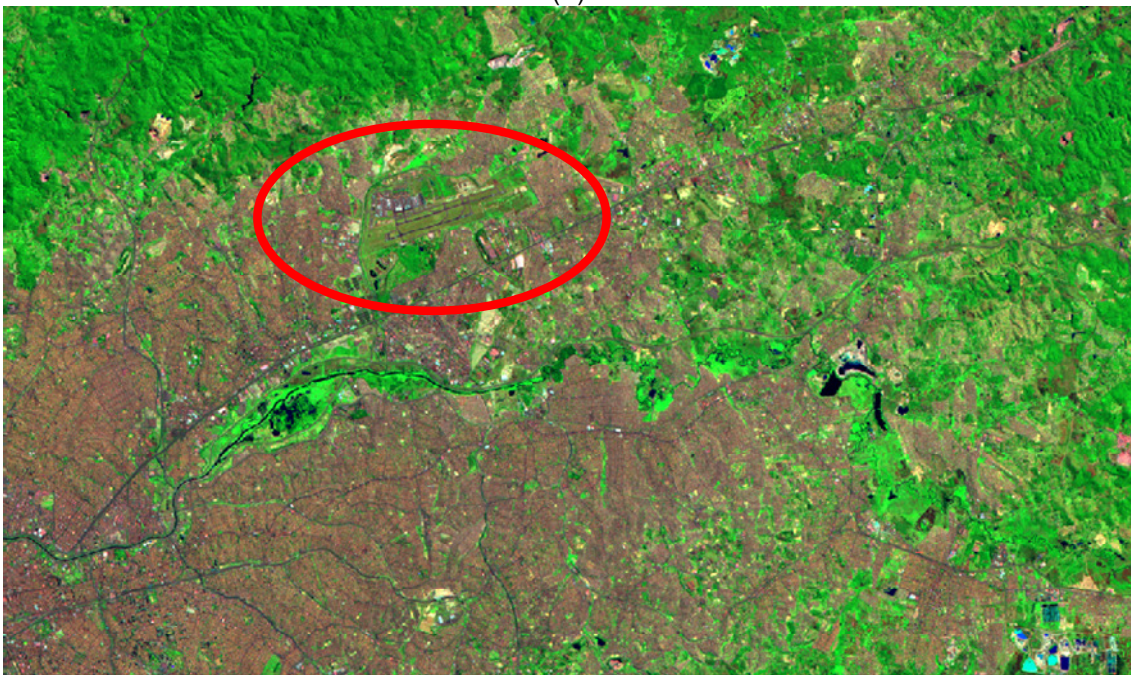
parece indicar a existência de um ciclo de período de aproximadamente de três décadas. Porém um estudo sobre este assunto está além do escopo deste trabalho.

Na análise da ocorrência de nevoeiro no Aeroporto de Guarulhos, Freitas et. al., (2008) mostraram que houve uma evolução urbana em torno de 50% na vizinhança do aeroporto, no período de 1986 a 2006, o qual coincide com o período posterior à inauguração do aeroporto. As Figuras 3.3 (a) e (b), as quais são partes dessa análise, são apresentadas a seguir. Estas figuras apresentam duas composições coloridas (RGB 7,4,2) do satélite Landsat5 TM referentes aos anos 1986 e 2006, respectivamente, que mostram a evolução da ocupação urbana no entorno do Aeroporto de Guarulhos-SP.

A análise das Figuras 3.2, 3.3 (a) e (b) mostra que possivelmente ocorreu mudança do microclima em torno da Estação de Observação Meteorológica de Cumbica/Guarulhos devido à urbanização, a qual provavelmente criou ilhas de calor nas proximidades do aeroporto. Isto, certamente, refletiu na decrescente média das horas anuais de nevoeiro desde o ano de 1951 ao ano de 2007. Essa média parece ter se estabilizado atualmente em torno de 150 horas/ano.



(a)



(b)

**Figura 3. 3** - (a) e (b) Composição colorida (RGB 7,4,2) do satélite Landsat5 TM de 1986 e 2006, respectivamente, que mostra a evolução da ocupação urbana na região da vizinhança do Aeroporto de Guarulhos-SP (Fonte: Freitas *et. Al.*, 2008).

Diante dos resultados desta análise, poderia se perguntar: valeria a pena trabalhar com os dados mais antigos, já que possivelmente houve alteração no

microclima da localidade? Não seria mais conveniente trabalhar apenas com os dados mais recentes que refletiriam uma situação meteorológica mais atual?

Como parte da resposta a essas indagações, são apresentadas a seguir as Figuras 3.4, 3.5 e 3.6. Na Figura 3.4 são mostrados os números de ocorrências de nevoeiro em cada hora do dia em cada uma das décadas 51-60, 61-70, 71-80, 81-90, 91-00 e 01-10. A última década foi extrapolada dos dados dos anos 2001 a 2007. Observa-se que a forma das seis curvas mostradas no gráfico da Figura 3.4 é praticamente a mesma: maior número de ocorrência da formação de nevoeiro acontece em torno de 6 h a 7 h da manhã e praticamente não há formação de nevoeiro entre 10 h e 20 h.

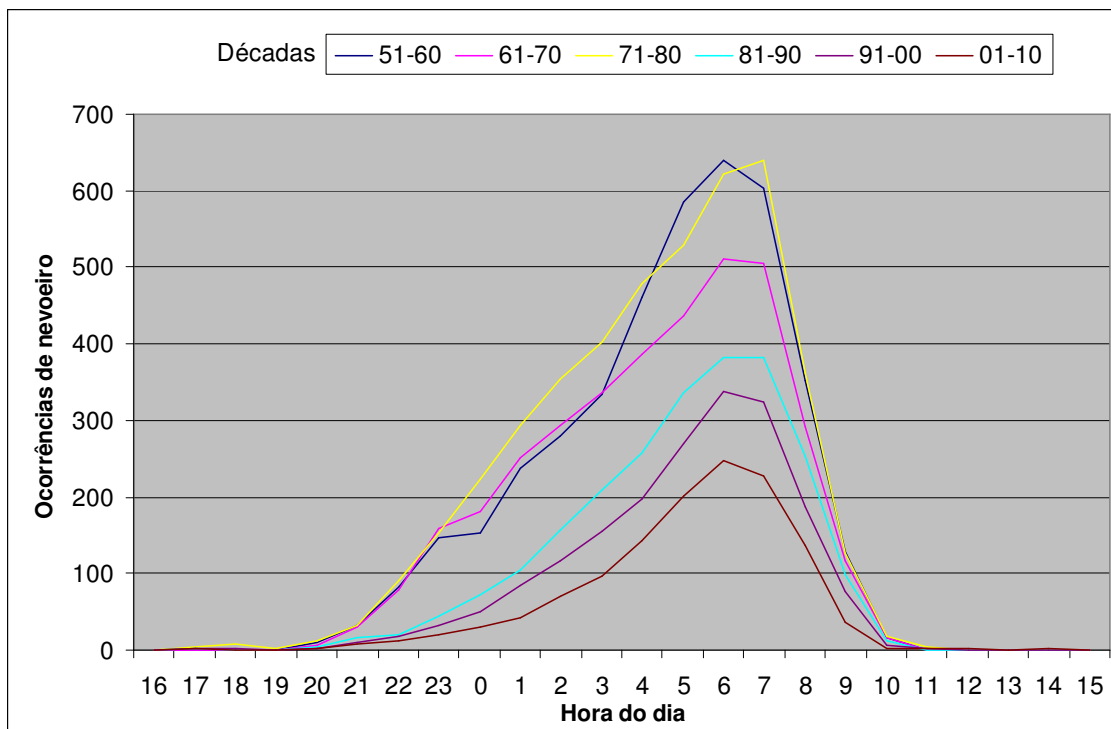
Na Figura 3.5 é apresentada a temperatura média decenal das 7 h e das 19 h para o período de estudo. Nota-se um acréscimo médio de 1 °C na temperatura média das 19 h e de 2 °C na temperatura média das 7h. Na Figura 3.6 é apresentada a umidade relativa (UR) média decenal das 7 h e das 19 h para o período de estudo. Neste caso, praticamente somente há alteração na média da UR das 19 h: há uma queda em torno de 5%.

O que tudo isto quer dizer?

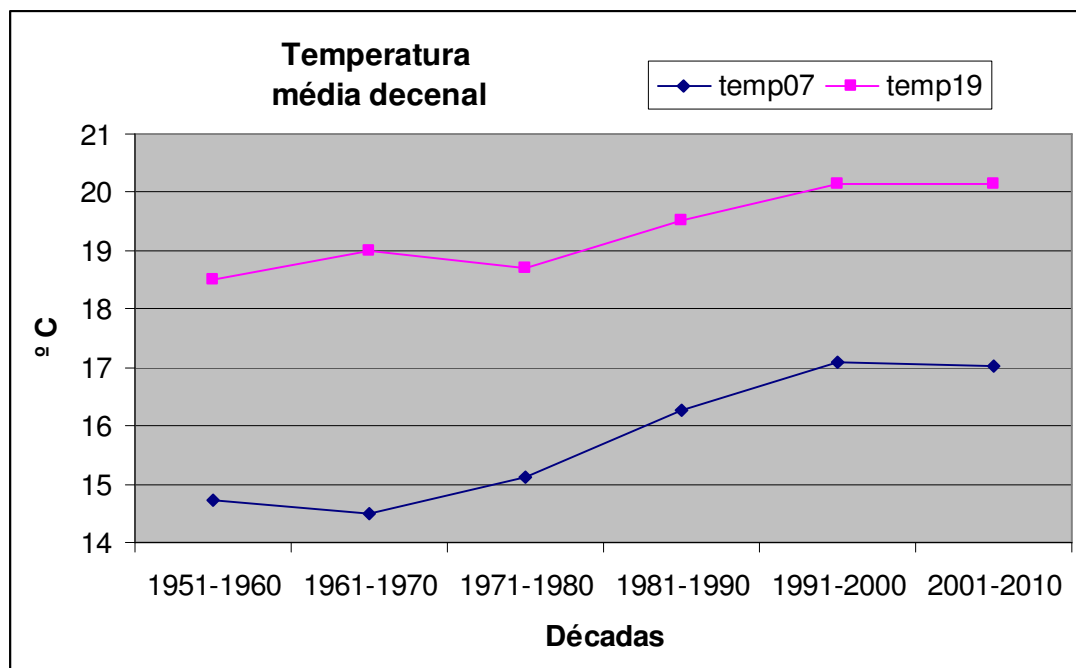
O que já era esperado: o processo de formação de nevoeiro em si mesmo não mudou: *“O nevoeiro pode se formar quando o ar se torna saturado. Isto ocorre, a uma dada temperatura, quando a quantidade de vapor de água existente no interior de uma parcela atinge a capacidade máxima; esta pode ser atingida de dois modos: por meio do resfriamento ou pela adição de vapor d’água a essa parcela”*. O que ocorreu, então, foi que, devido às circunstâncias ambientais: aumento médio das temperaturas e diminuição da umidade relativa do ar no início da noite (certamente causados pelo aumento da urbanização citada anteriormente) há menos exemplos de nevoeiro numa amostra atual de observações meteorológicas do que em uma amostra mais antiga. Mas, deste ponto de vista, uma observação meteorológica dos anos 60, por exemplo,



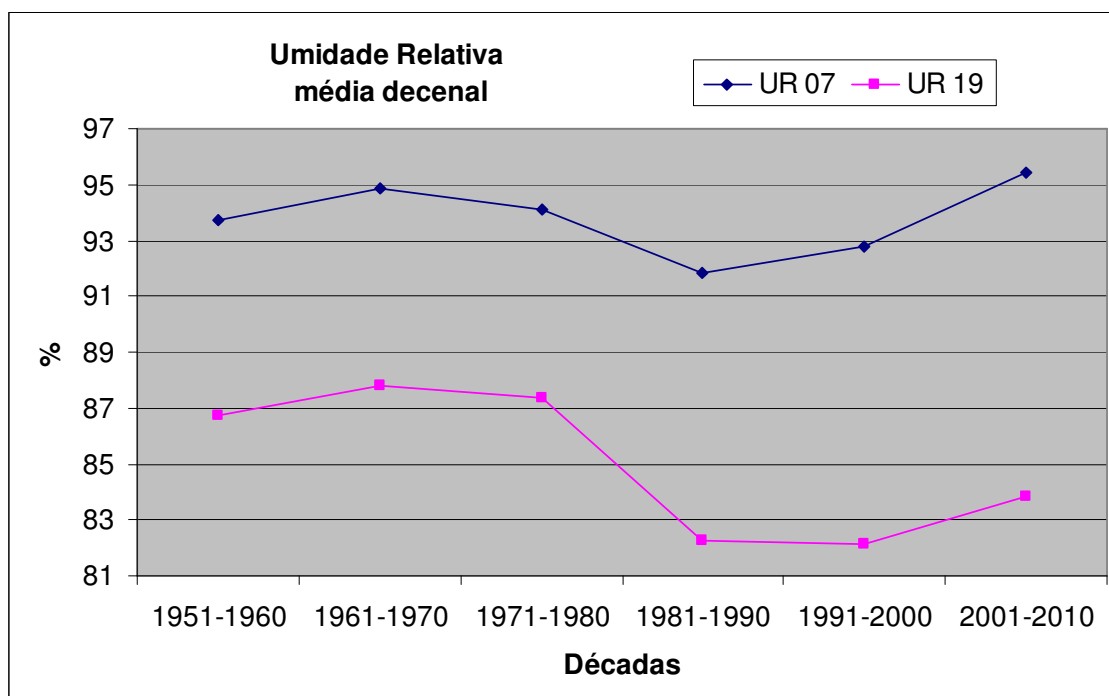
é tão boa quanto uma dos dias atuais para o treinamento do modelo de previsão de nevoeiro ou da visibilidade baixa.



**Figura 3. 4** – Números de ocorrências de nevoeiro em cada hora do dia em cada uma das seis décadas discriminadas no gráfico para a estação meteorológica de Cumbica/Guarulhos. A última década foi extrapolada dos dados dos anos 2001 a 2007.



**Figura 3. 5** – Temperatura média decenal das 7 h e das 19 h, para as décadas discriminadas no gráfico, para a estação meteorológica de Cumbica/Guarulhos. A última década foi extrapolada dos dados dos anos 2001 a 2007.



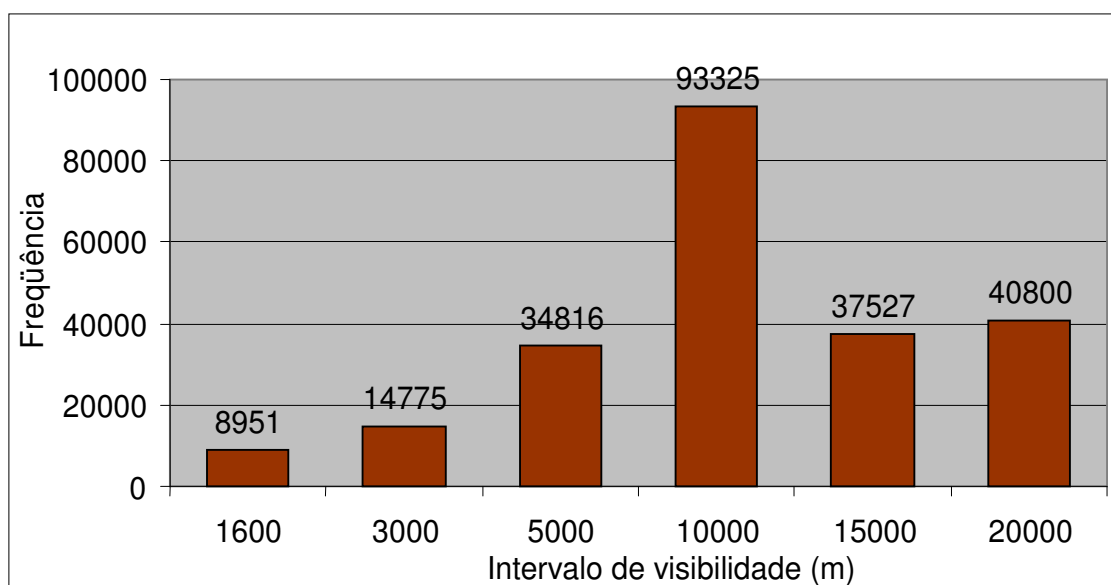
**Figura 3. 6** – Umidade relativa média decenal das 7 h e das 19 h, para as décadas discriminadas no gráfico, para a estação meteorológica de Cumbica/Guarulhos. A última década foi extrapolada dos dados dos anos 2001 a 2007.

### 3.1.2 ANÁLISE PRELIMINAR DOS DADOS

Com a primeira porção ou primeiro conjunto dos dados recebidos (janeiro de 1980 a março de 2007) e considerando-se o interesse da aviação em geral e o objetivo deste trabalho em relação à questão da baixa visibilidade e para efeitos de visualização geral dos valores da visibilidade, a amplitude total dos registros da estimativa da visibilidade, 0-20km, foi dividida em seis intervalos: 0-1600, 1601-3000, 3001-5000, 5001-10000, 10001-15000, 15001-20000m. O primeiro intervalo é o das medidas ou estimativas dos valores de visibilidades mais baixos.

Esse primeiro conjunto de dados fornecidos possui um total de 230.194 observações meteorológicas horárias. O histograma com os intervalos das visibilidades desse conjunto de dados está apresentado na Figura 3.7. Observa-se que ocorreram 8951 eventos meteorológicos com a visibilidade igual ou inferior a 1600m. Esse valor corresponde a apenas 3,9% da quantidade de observações existentes no período citado.

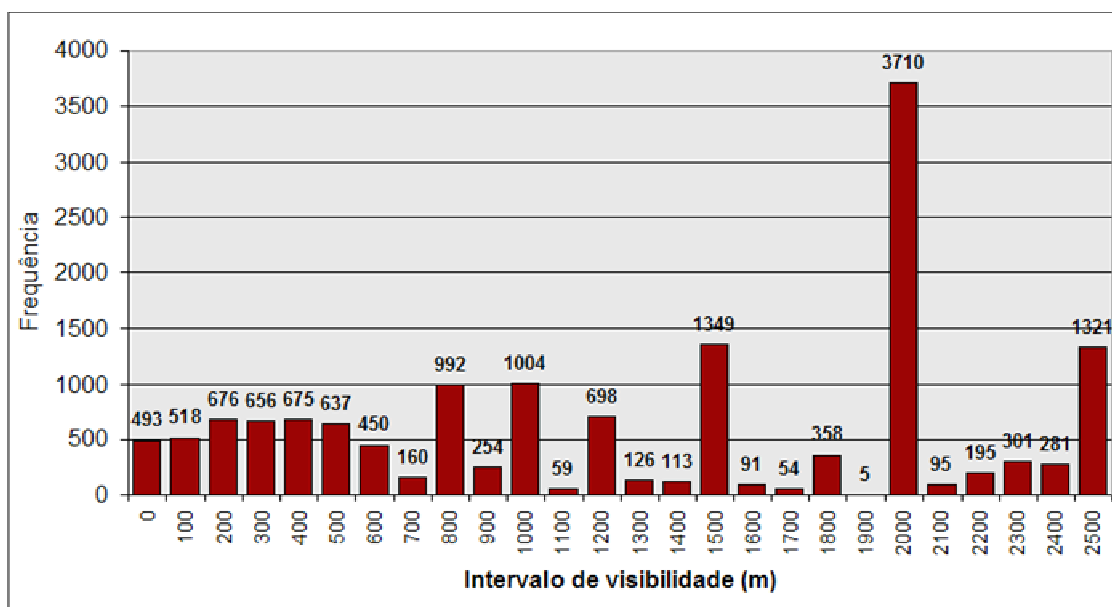
Em seguida, são apresentados minudenciosamente o primeiro e parte do segundo intervalo de visibilidade da Figura 3.7 no histograma da Figura 3.8. Analisando as freqüências dos intervalos de visibilidades mostrados na Figura 3.8 observa-se uma menor freqüência de registros de valores para a visibilidade de 700, 900, 1100, 1300, 1400, 1600, 1700, 1900, 2100 e 2200m, em contrapartida observa-se uma preferência pelo registro dos valores 800, 1000, 1200, 1500, 2000 e 2500m, especialmente a forte preferência pelo registro de 2000m.



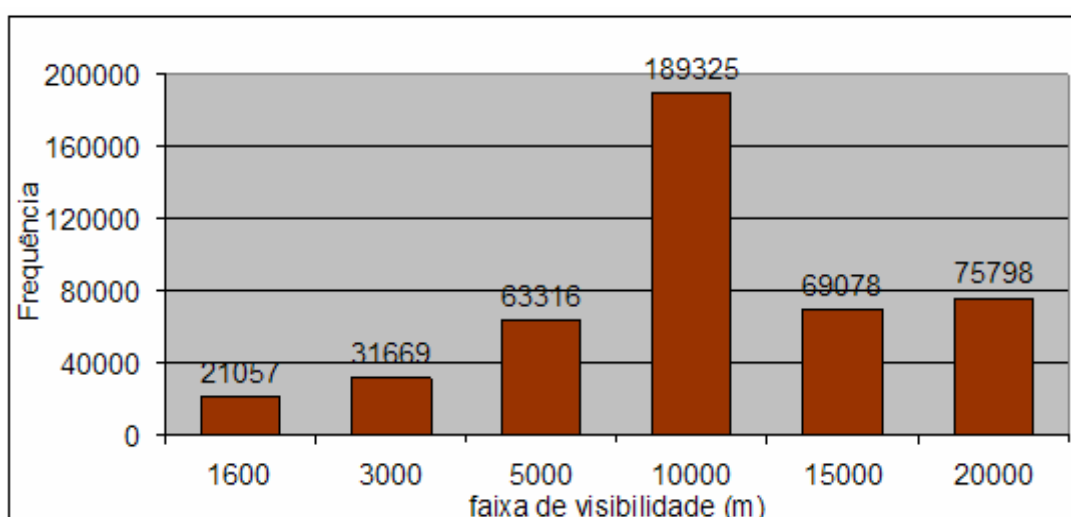
**Figura 3.7** - Histograma de seis intervalos da visibilidade das observações meteorológicas horárias relativas ao período de 01 de janeiro de 1980 a 31 de março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. O primeiro intervalo, de 0-1600m, com 8.951, é o da visibilidade baixa; os intervalos restantes são considerados visibilidade alta. Existe um total de 230.194 eventos meteorológicos.

Após o recebimento da segunda porção dos dados (de janeiro de 1951 a dezembro de 1979), foi feita uma análise semelhante com os dados meteorológicos do período todo (de janeiro 1951 a março de 2007). Este período corresponde a 20.544 dias que totalizariam 493.056 observações meteorológicas horárias. Mas devido a problemas técnicos com os dados, foram fornecidas efetivamente 450.243 observações meteorológicas horárias. O histograma com os intervalos das visibilidades desse conjunto de dados está apresentado na Figura 3.9. Observa-se que desse total de observações meteorológicas em apenas 21.057 delas ocorreu

visibilidade menor ou igual a 1600m, ou seja, aproximadamente 4,7% do total. Esta percentagem é um pouco maior que a de 3,9% citada em relação à Figura 3.7. Isto está coerente com a análise feita anteriormente na qual é mostrada que em amostras mais antigas há mais exemplos de nevoeiro ou visibilidade mais baixa.

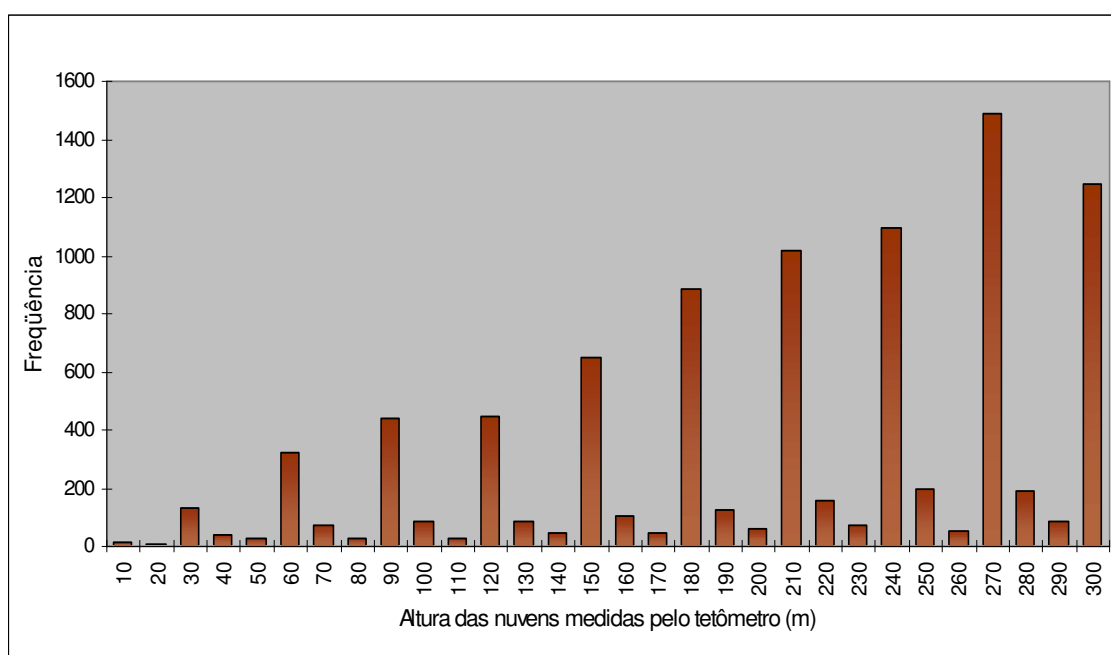


**Figura 3.8** - Histograma dos intervalos de visibilidade até 2500m dos dados fornecidos, mostrando a “preferência” de registro de alguns valores de visibilidade. A amplitude de cada intervalo é de 100m, com exceção do primeiro que igual a zero. As informações se referem ao período de janeiro de 1980 a março de 2007 da Estação de Observação Meteorológica de Cumbica/Guarulhos.



**Figura 3.9** – Histograma de seis intervalos de visibilidade relativos aos eventos meteorológicos do período de janeiro de 1951 a março de 2007. O primeiro intervalo, de 0-1600m, com 21.057, é o da visibilidade baixa; os intervalos restantes são considerados visibilidade alta. O número total de eventos meteorológicos relativos à representação do histograma é 493.056 e se referem à Estação Meteorológica de Cumbica/Guarulhos.

De forma semelhante à análise feita com a visibilidade, foi realizada uma análise para os valores registrados para a altura das nuvens da primeira camada, geralmente nuvens baixas. Na Figura 3.10 é apresentado o histograma das alturas de nuvens medidas por tetômetro, referentes aos anos de 1998 a 2007. Pela análise da Figura 3.10, nota-se, também, que há registros de valores preferenciais. Observa-se a “preferência” pelos valores múltiplos de 30 m: 30, 60, 90, 120, 150, 180, 210, 240, 270 e 300m que correspondem, respectivamente, a 100, 200, 300, 400, 500, 600, 700, 800, 900 e 1000 ft.



**Figura 3. 10** - Histograma dos intervalos da altura das nuvens (teto) até 300m. A amplitude de cada intervalo é de 10m. Percebe-se a “preferência” pelo registro de valores múltiplo de 30 m. As informações se referem ao período do ano de 1998 ao ano de 2007 da Estação Meteorológica de Guarulhos.

Considerando a existência de registros valores preferenciais de visibilidade e de teto, como já inferida da análise das Figuras 3.8 e 3.10 e considerando o fato de que a natureza não deve ter esse comportamento “preferencial”, pode-se afirmar que há uma forte indicação que os registros dos valores da estimativa de visibilidade horizontal e teto de nuvens são tendenciosos.

Ressalta-se que durante o período abrangido pelos dados foram utilizados três distintos formulários de registro dos dados observacionais (formulários IEPV-105-25,

IEPV-105-13 e IEPV-105-78, uma imagem deste último está apresentada no Anexo A). A adoção de cada novo formulário apresentava diferentes tipos de registro tanto no modo de preenchimento quanto nos formatos, tipos de dados e diferentes tabelas de dados (por exemplo, de nuvens, dos fenômenos meteorológicos...). Isto resultou em um razoável trabalho de adequação para a uniformização desses dados para que os mesmos pudessem ser usados como parâmetros de trabalho no modelo de previsão a ser desenvolvido. Os períodos de vigência dos formulários são os seguintes:

- a) IEPV 105-25 – anterior ao ano de 1980 até 1993/94;
- b) IEPV 105-13, de 1994/95 até o final de 1997;
- c) IEPV 105-78, a partir de 1998 até os dias atuais.

No apêndice A são apresentadas algumas informações mais detalhadas sobre os registros dos dados meteorológicos.

### 3.1.2 PRÉ-PROCESSAMENTO DOS DADOS

Os dados mais antigos, anteriores ao ano de 1998, entrada em vigor do formulário IEPV-105-78, continham a UR (umidade relativa) apenas nos horários sinóticos (00, 03, 06, 09, 12, 15, 18, 21 UTC). Foram adotados dois procedimentos para o preenchimento das UR não medidas: a) foi ajustada uma função de regressão com as variáveis “temperatura do bulbo seco” (PSECO) e “temperatura do bulbo úmido” (BUMIDO), b) foi usada a equação de Clausius-Clapeyron com as variáveis PSECO e a “temperatura do ponto de orvalho” (PO). A necessidade de se fazer os dois procedimentos se deve ao fato de que em alguns anos havia apenas uma das duas variáveis BUMIDO e PO. Esses procedimentos são descritos a seguir:

#### **a) UR como função da temperatura do bulbo seco e do bulbo úmido.**

$$UR = f (BSECO, BUMIDO).$$

Usando um conjunto de 65.090 observações meteorológicas de horários sinóticos que contivessem as três variáveis UR, BSECO e BUMIDO, do período anterior a 1998, foi ajustada a seguinte função:

$$UR = 93,922 - 6,638BSECO + 6,857BUMIDO$$

Na Tabela 3.1 são apresentadas as estatísticas desta regressão. Observa-se que há um ótimo ajustamento, com  $R^2$  quase igual a 1.

**Tabela 3. 1** – Estatística da regressão  $UR = f(BSECO, BUMIDO)$  realizada com um conjunto de 65.090 observações meteorológicas de horários sinóticos que contivessem as três variáveis UR, BSECO e BUMIDO, do período anterior a 1998.

Estatística de regressão	
R múltiplo	0,99
R-Quadrado	0,97
R-quadrado ajustado	0,97
Erro padrão	2,6985
Observações	65090

ANOVA					
	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	2	17604162,09	8802081,04	1208720,06	
Resíduo	65087	473973,31			0
Total	65089	18078135,4	7,28		

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	93,922	0,059	1594,031	0	93,8062	94,0372	93,8062	94,0372
BSECO	-6,638	0,004	-1500,03	0	-6,6472	-6,6298	-6,6472	-6,6298
BUMIDO	6,857	0,006	1076,304	0	6,8450	6,8700	6,8450	6,8700

### b) UR como função da temperatura do bulbo seco e do ponto de orvalho.

Segundo (Rogers, 1975), a umidade relativa pode ser calculada a partir da equação de Clausius-Clapeyron:

$$UR = \frac{6,112 \frac{17,67 \cdot PO}{PO + 243,5}}{6,112 \frac{17,67 \cdot PSECO}{BSECO + 243,5}} \times 100$$

Após o trabalho de uniformização e complementação dos dados fornecidos, os dados básicos de trabalho passaram a ser os seguintes:

- DIA, MÊS, ANO, HORA: data-hora do evento;
- DIRVENTO, VELVENTO: direção do vento, em dezenas de graus, em relação ao norte geográfico; e a velocidade do vento em “nós” (*kt*);
- VISIB: visibilidade horizontal, em dezenas de metros;

- d) CGT: Condições Gerais de Tempo, condições meteorológicas atuais;
- e) QTDNUVEM1, TIPONUVEM1, ALTNUVEM1: quantidade, em oitavos da abóbada celeste, da cobertura de nuvens; tipo de nuvens; altura da nuvem, em dezenas de metros; tudo da primeira camada de nuvens;
- f) QTDNUVEM2, TIPONUVEM2, ALTNUVEM2: as mesmas definições do item “e”, relativas à segunda camada de nuvens;
- g) QTDNUVEM3, TIPONUVEM3, ALTNUVEM3 as mesmas definições do item “e”, relativas à terceira camada de nuvens;
- h) QTDNUVEM4, TIPONUVEM4, ALTNUVEM4: as mesmas definições do item “e”, relativas à quarta camada de nuvens;
- i) TEMPAR: temperatura do ar na superfície, em graus Celsius, com aproximação em décimos;
- j) UR: umidade relativa do ar, em percentagem;
- k) QNH; pressão atmosférica, em hectoPascal, com aproximação em décimos.

Ao analisar os dados, nota-se que existem situações onde o evento meteorológico não é ausente – a observação meteorológica foi realizada a contento, mas alguns campos são deixados “em branco” – pelas normas de preenchimento do formulário como, por exemplo, os campos relativos às nuvens. Nesses casos, pode ocorrer situação de não estarem presentes todas as várias camadas de nuvens ou até não se ter nuvem nenhuma – céu claro. Para que todas as células da planilha eletrônica ficassem preenchidas, foi adotado o procedimento a seguir.

Os dados de nuvens não existentes de eventos presentes foram preenchidos da seguinte forma:

QTDNUVEM (quantidade de nuvens, nas 4 camadas) = “0”;

TIPONUVEM (tipo de nuvens, nas 4 camadas) = “0” (igual a SKC – céu claro; ver a terceira coluna da Tabela 3.2);



ALTNUVEM – igual a 999 (a ausência de nuvens foi registrada como se fosse uma nuvem muito alta; no caso, conforme a tabela de altura de nuvens, 999 significa acima de 10.000m).

**Tabela 3. 2** - Correspondência entre as tabelas de nuvens: Tabela 0500 do Manual de Códigos Meteorológicos do Comando da Aeronáutica (primeira coluna), usada anteriormente ao ano de 2005, a tabela criada para o formulário IEPV-105-78 (segunda coluna) e a Tabela de nuvens (terceira coluna), em relação à tabela da segunda coluna, foi incluída a sigla SKC (*sky clear*) indicativa da condição de céu claro.

Tabela 0500 (Manual de Códigos)	Tabela do formulário IEPV 105-78 (2005)	Tabela de nuvens	
0 Cirrus... CI	0 Cirrocumulus/Cirrus CC/CI	0	CI/CC/SKC
1 Cirrocumulus... CC	1 Cirrostratus... CS	1	CS
2 Cirrostratus... CS	2 Altocumulus... AC	2	AC
3 Altocumulus... AC	3 Altostratus... AS	3	AS
4 Altostratus... AS	4 Nimbostratus... NS	4	NS
5 Nimbostratus... NS	5 Stratocumulus... SC	5	SC
6 Stratocumulus... SC	6 Stratus... ST	6	ST
7 Stratus... ST	7 Cumulus... CU	7	CU
8 Cumulus... CU	8 Grandes Cumulus... TCU	8	TCU
9 Cumulonimbus... CB	9 Cumulonimbus... CB	9	CB

### 3.1.3 CRIAÇÃO DE VARIÁVEIS DERIVADAS

As doze variáveis de nuvens (QTDNUVEM1, TIPONUVE1, ALTNUVEM1, QTDNUVEM2, TIPONUVE2, ALTNUVEM2 QTDNUVEM3, TIPONUVE3, ALTNUVEM3, QTDNUVEM4, TIPONUVE4, ALTNUVEM4) foram transformadas e reduzidas a quatro novas variáveis: *TIPO*, *Q1*, *Q2* e *Q3*, da seguinte forma:

1. A variável *TIPO* foi gerada buscando em TIPONUVE1, TIPONUVE2, TIPONUVE3, TIPONUVE4, nessa ordem, o maior valor registrado (correspondente ao tipo de nuvem da terceira coluna da

Tabela 3.2). A razão para isto é fornecer ao modelo a informação da existência de nuvens convectivas (CB, TCU, CU, AC) ou nuvens estratiformes (ST, AS) , respectivamente os números 9, 8, 7, 2, e os números 6, 3 da coluna três da Tabela 3.2, no momento da observação meteorológica. Isto traduz, em uma forma indireta, as condições de instabilidade/estabilidade e de umidade em níveis baixo/médio da atmosfera local no momento da observação.

2. As variáveis  $Q1$ ,  $Q2$  e  $Q3$  foram criadas somando, os valores de QTDNUVEM1, QTDNUVEM2, QTDNUVEM3, QTDNUVEM4, que existissem dentro das seguintes faixas de altura:

$$Q1 \leq 210\text{m}$$

$$210\text{m} < Q2 < 2000\text{m};$$

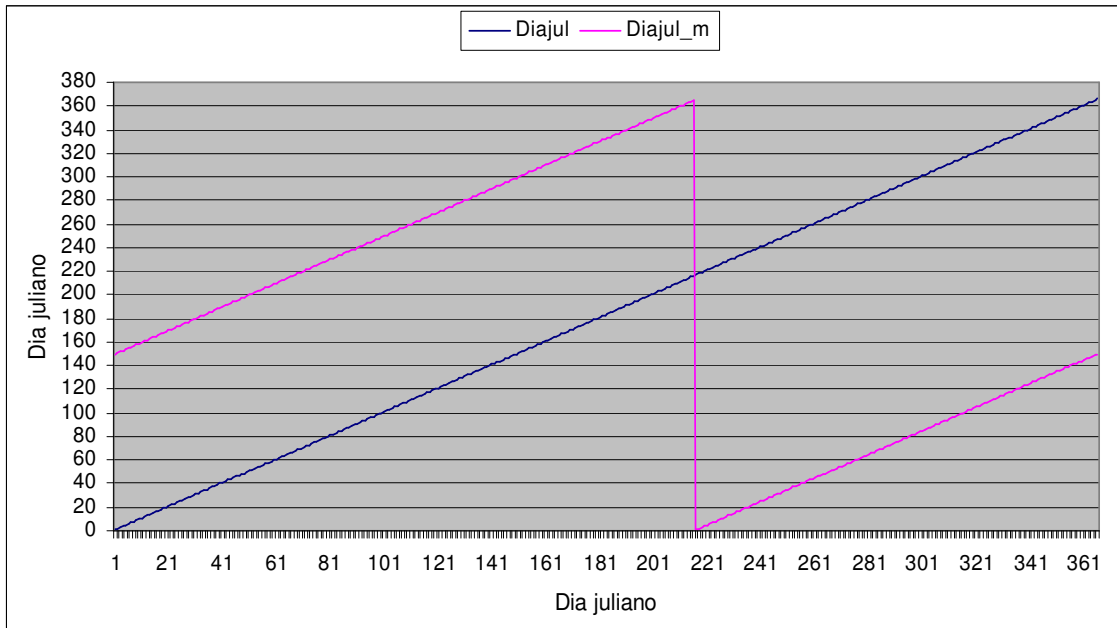
$$2000\text{m} \leq Q3 < 5000\text{m}.$$

Como no item anterior, essas variáveis relacionadas às quantidades de nuvens dão informações indiretas sobre a quantidade de umidade na atmosfera abaixo do nível de 5000m. As variáveis  $Q1$  e  $Q2$  consideram as nuvens baixas, e  $Q3$  as nuvens médias, dentro da Classificação Internacional de Nuvens.

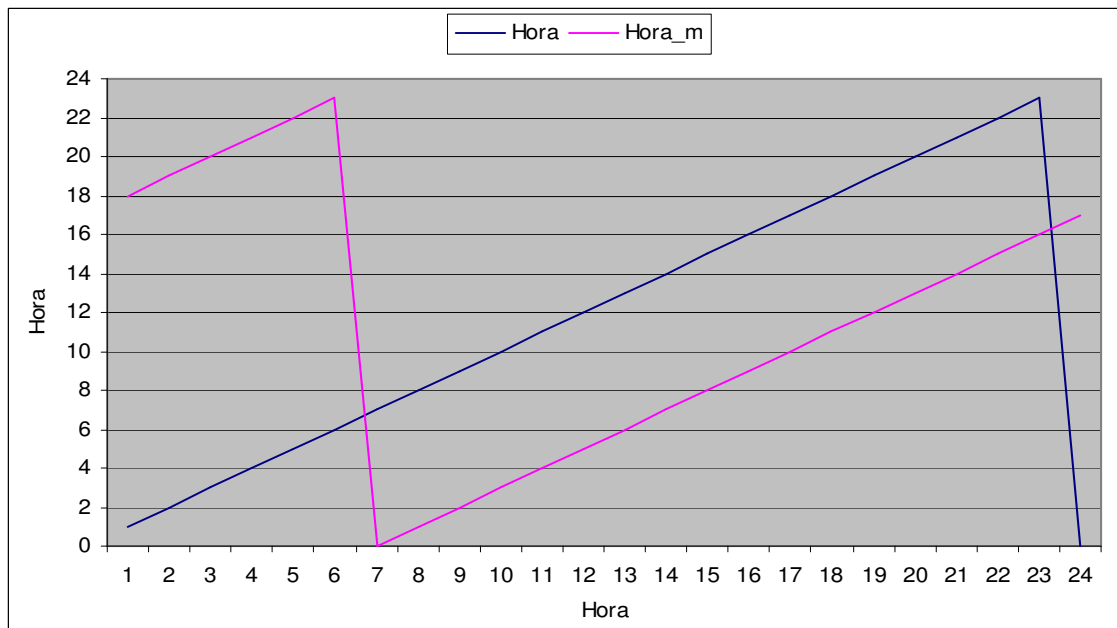
3. Foi criada a variável 6H\_SKC que representa o número de horas de céu claro durante as últimas seis horas. Para isto, foi considerada como céu claro a condição de 2/8 ou menos da abóbada celeste encoberta por nuvens no céu. A criação desta variável foi motivada pela intenção de fornecer ao modelo informações indiretas da radiação solar recebida, em horas de céu claro, durante o dia, e a radiação terrestre perdida nas horas de céu claro, durante a noite.

Os dados de data e hora foram transformados em duas novas variáveis, da seguinte forma:

4. A variável ANO não foi levada em consideração, as variáveis DIA e MÊS foram combinadas para formar a variável *dia juliano* (ver Apêndice B). Posteriormente, esta foi transformada na variável DIAJULm (dia juliano modificado) que começa no dia 4 ou 5 de agosto (conforme o ano seja bissexto ou não) com o valor inicial 1. A Figura 3.11 apresenta a relação entre os dois tipos de *dia*. A justificativa para a criação desta variável está relacionada à questão da sazonalidade conforme foi discutido na seção 2.1.3 SAZONALIDADE. A escolha dos dias 4 ou 5 de agosto para o início da contagem dos dias coincide com a metade do inverno, no hemisfério sul. A partir daí começa a haver um aquecimento gradual da superfície até meado de março. A idéia foi associar o valor crescente do início do calendário modificado ao crescimento gradual do aquecimento da Terra, no hemisfério sul.
5. A hora da observação HORA foi transformada na variável HORAm (hora modificada). Essa variável começa com o valor “0” às 7:00 hs. Como no item anterior, a justificativa para a criação desta variável está relacionada à questão da sazonalidade. A escolha das 7:00 hs para o início da contagem da hora modificada coincide, em geral, com o início do aquecimento do dia. A partir daí começa a haver um aquecimento gradual da superfície até mais ou menos às 17:00 hs. Novamente, a idéia foi associar o valor crescente da hora modificada ao aquecimento gradual do dia. A Figura 3.12 apresenta a relação entre essas variáveis citadas.



**Figura 3.11** - Relação entre o dia juliano e o dia juliano modificado. Na abscissa está o dia juliano e nas ordenadas o valor correspondente do dia juliano modificado.



**Figura 3.12** - Relação entre a hora e a hora modificada. Nas abscissas se encontra a hora do dia e nas ordenadas o valor correspondente da hora modificada.

### 3.1.4 ESCOLHA DAS VARIÁVEIS PREDITORAS

## INTRODUÇÃO

O desempenho de um modelo de previsão meteorológica é muito dependente das variáveis de entradas (variáveis preditoras). Por isso é recompensador investir no estudo da escolha dessas variáveis. A seguir, portanto, serão apresentadas algumas análises e justificativas relacionadas.

O *tempo* pode ser definido como o estado da atmosfera em determinado instante e lugar. As condições de tempo locais são descritas em termos de parâmetros ou quantidades básicas medidas regularmente, como, por exemplo:

- temperatura do ar,
- umidade do ar,
- pressão do ar,
- velocidade e direção do vento,
- o tipo e quantidade de nuvens,
- tipo de fenômeno meteorológico.

## A TEMPERATURA DO AR

A temperatura do ar é dependente dos fatores:

a) da radiação que, mantidas a latitude e superfície de estudo, depende da hora do dia, do dia do ano e da cobertura de nuvens. Por isto, a temperatura do ar é geralmente:

- maior em janeiro que em julho (hemisfério sul);
- maior durante o dia que à noite;
- durante o dia, maior sob céu claro;
- durante a noite, menor sob céu claro e seco.

Por isso, é recomendável que sejam utilizadas variáveis que levem em consideração essa sazonalidade. Nesse caso foram criadas as já citadas variáveis HORAm e DIAJULm.

b) da advecção de massa de ar que pode compensar ou mesmo sobrepor-se à influência da radiação sobre a temperatura.

Também, aqui, é recomendável que sejam usadas variáveis que contenham a indicação dessa situação transiente - a advecção da massa de ar. Para tanto foram criadas variáveis que contivessem informações do passado: *informações meteorológicas de horas anteriores*. Estas variáveis serão discriminadas adiante.

## UMIDADE RELATIVA

Variações da umidade relativa dependem, basicamente, do acréscimo ou diminuição de vapor d'água ao ar, ou da variação da temperatura tanto pelos movimentos horizontais ou verticais de massa de ar como pelo ciclo diurno. A variável usada no trabalho são os valores da umidade relativa anotados nos formulários de registros, em percentagem.

## PRESSÃO DO AR

Os dados referentes à pressão do ar utilizados são os oriundos dos registros originais do parâmetro ajuste do altímetro, em hectopascal com aproximação de décimos, sem modificações.

## VELOCIDADE E DIREÇÃO DO VENTO

Os dados referentes ao vento utilizados são a direção, em graus em relação ao norte verdadeiro, e a velocidade, em *kt* (nós). Um nó é igual 1 milha náutica/hora ou 1852 metro/hora ou aproximadamente 0,51 metro/segundo. Os dados usados são oriundos dos registros originais, sem modificações.

## TIPO E QUANTIDADE DE NUVENS

O perfil vertical de temperatura e, portanto, a estabilidade atmosférica varia significativamente com a hora do dia, e com o dia do ano (da estação do ano). A

estabilidade atmosférica influencia o tempo ao afetar o movimento vertical do ar. Ar estável enfraquece o movimento vertical e ar instável provoca e favorece o movimento vertical, a convecção, o resfriamento por expansão e desenvolvimento de nuvens. Nuvens com desenvolvimento vertical estão relacionadas com ar instável. *Nuvens em camadas geralmente indicam um ar estável* na altitude em que elas ocorrem.

Para que as possíveis variáveis preditoras tivessem alguma indicação da condição de estabilidade da atmosfera, foram, para isto, criadas variáveis relativas a nuvens como um parâmetro indireto da estabilidade/instabilidade atmosférica: as já citadas variáveis TIPO, Q1, Q2 e Q3.

O resfriamento radiativo da superfície da Terra ocorre principalmente à noite, sob céu limpo. A camada de ar superficial é então resfriada por contato com a superfície mais fria e uma inversão superficial de temperatura se desenvolve. Este processo é o grande responsável pela formação de orvalho e alguns tipos de nevoeiros (Justificativa para a variável “*número de horas de céu claro durante as últimas seis horas*”).

A motivação desta análise sobre os fatores que afetam o estado da atmosfera e as condições locais de tempo é a justificativa da escolha das variáveis de entrada do modelo adotadas neste projeto. A pretensão é que elas, as variáveis, contivessem a maior quantidade possível de informações para ser fornecido ao treinamento do modelo de previsão.

### 3.1.5 CORRELAÇÃO DAS VARIÁVEIS COM A VISIBILIDADE

As variáveis dependentes ou os preditandos são seis: as três relativas à visibilidade futura (VISIB1, VISIB2 e VISIB3) e as três relativas à altura da base da camada de nuvem mais baixa (ALT1, ALT2 e ALT3). Para se ter uma idéia sobre a força de cada possível variável preditora, foi feita uma correlação entre as candidatas a variáveis preditoras e a visibilidade. A correlação entre as candidatas a variáveis de entrada e a visibilidade foi feita com dois conjuntos de dados, um com visibilidade

baixa (menor ou igual a 1600m) e o segundo com todas as faixas de visibilidade. A finalidade de se ter usado dois conjuntos foi para verificar se a correlação das variáveis candidatas com o preditando dependia da faixa de valores do próprio preditando.

Esta correlação foi realizada para dois períodos de tempos diferentes: a) período de janeiro de 1980 a março de 2007; este período está relacionado à primeira a idéia inicial, que será detalhada no item 3.2.1; b) período de janeiro de 1951 a março de 2007. A seguir serão detalhadas as correlações para esses dois períodos:

a) PERÍODO DE JANEIRO DE 1980 A MARÇO DE 2007

A correlação entre as candidatas a variáveis de entrada e cada uma das três visibilidades futuras (VISIB1, VISIB2 E VISIB3) foi feita com dois conjuntos de dados. O primeiro conjunto contendo 7.248 eventos de visibilidade baixa (menor ou igual a 1600m) e o segundo, com todas as faixas de visibilidade, contendo 187.708 eventos.

Como já dito no parágrafo anterior, em cada um dos dois casos, foram realizadas três correlações entre as possíveis variáveis preditoras e as visibilidades futuras da primeira, segunda e terceira hora, que são as variáveis dependentes ou os preditandos: VISIB1, VISIB2, e VISIB3, respectivamente. Nesses seis casos, a correlação foi realizada com 31 variáveis candidatas: 14 da hora atual, 10 da hora anterior, 6 de 24 horas passadas e uma de 12 horas passadas. Estas variáveis estão relacionadas a seguir:

VARIÁVEIS DA HORA ATUAL (14):

DIRVENT

VELVEN

VISIB

CGT

TIPO

Q1, Q2, Q3



TEMPAR

QNH

UR

6H\_SKC

HORAm

DIAJULm

VARIÁVEIS DA HORA ANTERIOR (10):

DIRVENT\_1h

VELVEN\_1h

VISIB\_1h

TEMPAR\_1h

QNH\_1h

UR\_1h

TIPO\_1h

Q1\_1h, Q2\_1h, Q3\_1h

VARIÁVEL DE 12 HORAS ANTERIORES (1)

QNH\_12h

VARIÁVEIS DE 24 HORAS ANTERIORES (6)

DIRVENT\_24h

VELVEN\_24h

VISIB\_24h

TEMPAR\_24h

QNH\_24h

UR\_24h

Os resultados estão apresentados nas Tabelas 3.3 e 3.4. Na Tabela 3.3, estão os resultados relativos ao conjunto de baixa visibilidade (menor ou igual 1600m) e na Tabela 3.4, os resultados relativos ao conjunto com todas as faixas de visibilidade.

**Tabela 3. 3** – Correlação entre as candidatas a variáveis de entrada e cada uma das três visibilidades futuras (VISIB1, VISIB2 E VISIB3) realizada com o conjunto de 7.248 eventos de visibilidade baixa (menor ou igual a 1600m), relativo ao período de janeiro de 1980 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As informações estão ordenadas pelo valor absoluto do fator de correlação.

VARIÁVEL	VISIB1	VARIÁVEL	VISIB2	VARIÁVEL	VISIB3
Q2	0,311	Q2	0,287	Q2	0,245
Q2_1h	0,286	HORAm	-0,274	HORAm	-0,241
VISIB	0,277	Q2_1h	0,255	UR	-0,230
TEMPAR	0,259	TEMPAR	0,253	TEMPAR	0,228
TEMPAR_1h	0,252	TEMPAR_1h	0,246	UR_1h	-0,225
Q3	0,250	UR_1h	-0,244	Q2_1h	0,225
UR_24h	-0,246	UR	-0,239	TEMPAR_1h	0,215
HORAm	-0,240	UR_24h	-0,233	UR_24h	-0,212
UR_1h	-0,239	VISIB	0,224	TEMPAR_24h	0,194
UR	-0,233	Q3	0,224	Q3	0,190
TEMPAR_24h	0,229	TEMPAR_24h	0,216	6H_SKC	-0,179
Q3_1h	0,224	Q3_1h	0,199	Q3_1h	0,170
VISIB_1h	0,223	6H_SKC	-0,193	VISIB	0,163
6H_SKC	-0,192	VISIB_1h	0,189	DIAJULm	-0,152
DIRVENT	0,151	DIRVENT_1h	0,158	VISIB_1h	0,143
DIAJULm	-0,145	DIAJULm	-0,146	Q1	-0,107
Q1	-0,144	DIRVENT	0,142	QNH_1h	-0,104
DIRVENT_1h	0,142	QNH_1h	-0,117	DIRVENT	0,103
Q1_1h	-0,109	QNH_24h	-0,114	QNH_24h	-0,102
DIRVENT_24h	0,108	VELVEN_1h	0,112	QNH	-0,096
VELVEN	0,107	Q1	-0,108	TIPO_1h	0,095
QNH_24h	-0,103	QNH	-0,105	QNH_12h	-0,085
QNH_1h	-0,103	DIRVENT_24h	0,102	DIRVENT_1h	0,084
VELVEN_1h	0,098	VELVEN	0,098	Q1_1h	-0,081
QNH	-0,093	QNH_12h	-0,086	CGT	-0,080
VISIB_24h	0,081	VISIB_24h	0,082	DIRVENT_24h	0,074
VELVEN_24h	0,070	Q1_1h	-0,072	TIPO	0,067
QNH_12h	-0,070	TIPO_1h	0,071	VISIB_24h	0,063
TIPO	-0,055	VELVEN_24h	0,070	VELVEN_1h	0,058
CGT	0,036	CGT	-0,055	VELVEN	0,055
TIPO_1h	0,019	TIPO	0,020	VELVEN_24h	0,041

**Tabela 3. 4 -** Correlação entre as candidatas a variáveis de entrada e cada uma das três visibilidades futuras (VISIB1, VISIB2 E VISIB3) realizada com o conjunto de 187.708 eventos de visibilidade em todas as faixas, relativo ao período de janeiro de 1980 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As informações estão ordenadas pelo valor absoluto do fator de correlação.

VARIÁVEL	VISIB1	VARIÁVEL	VISIB2	VARIÁVEL	VISIB3
VISIB	0,861	VISIB	0,719	VISIB	0,590
VISIB_1h	0,721	VISIB_1h	0,592	VISIB_1h	0,483
CGT	-0,582	CGT	-0,485	CGT	-0,397
UR	-0,331	UR	-0,279	<b>HORAm</b>	-0,226
Q1	-0,315	Q1	-0,271	Q1	-0,224
VISIB_24h	0,297	VISIB_24h	0,250	UR	-0,223
TEMPAR	0,284	TEMPAR	0,249	TEMPAR	0,207
UR_1h	-0,279	Q1_1h	-0,224	VISIB_24h	0,198
Q1_1h	-0,271	UR_1h	-0,223	DIRVENT	0,176
TEMPAR_1h	0,249	TEMPAR_1h	0,208	Q1_1h	-0,176
DIRVENT	0,215	DIRVENT	0,197	VELVEN	0,171
TEMPAR_24h	0,209	TEMPAR_24h	0,190	UR_1h	-0,169
DIRVENT_1h	0,198	<b>HORAm</b>	-0,188	TEMPAR_1h	0,165
VELVEN	0,182	VELVEN	0,180	TEMPAR_24h	0,163
VELVEN_1h	0,180	DIRVENT_1h	0,176	VELVEN_1h	0,161
UR_24h	-0,158	VELVEN_1h	0,172	DIRVENT_1h	0,157
DIRVENT_24h	0,114	UR_24h	-0,131	VELVEN_24h	0,096
<b>HORAm</b>	-0,114	VELVEN_24h	0,101	UR_24h	-0,096
VELVEN_24h	0,105	DIRVENT_24h	0,100	DIRVENT_24h	0,087
QNH	-0,069	QNH	-0,057	QNH_1h	-0,052
QNH_24h	-0,058	QNH_1h	-0,052	QNH	-0,052
QNH_1h	-0,057	QNH_24h	-0,049	QNH_24h	-0,045
QNH_12h	-0,049	QNH_12h	-0,044	QNH_12h	-0,044
<b>6H_SKC</b>	0,044	TIPO	-0,031	<b>Q3_1h</b>	0,031
TIPO	-0,038	<b>Q3_1h</b>	0,028	<b>Q3</b>	0,028
TIPO_1h	-0,031	<b>6H_SKC</b>	0,026	TIPO	-0,022
<b>Q3_1h</b>	0,023	<b>Q3</b>	0,023	<b>DIAJULm</b>	-0,021
<b>DIAJULm</b>	-0,021	TIPO_1h	-0,022	<b>6H_SKC</b>	0,010
<b>Q3</b>	0,015	<b>DIAJULm</b>	-0,021	TIPO_1h	-0,009
<b>Q2</b>	-0,006	<b>Q2_1h</b>	0,005	<b>Q2_1h</b>	0,008
<b>Q2_1h</b>	0,001	<b>Q2</b>	0,001	<b>Q2</b>	0,005

VARIÁVEIS ESCOLHIDAS:

Após análises e testes, foram abandonadas as variáveis candidatas QNH\_12h, QNH\_1h, VELVEN\_24h, VELVEN\_1h ou por alta correlação com outra variável independente ou por baixa correlação com as variáveis dependentes. Foram mantidas as 27 variáveis restantes. Portanto, as variáveis de entrada para o treinamento do modelo, para o período de janeiro de 1980 a março de 2007, são:

- DIRVENTO\_24h, VISIB\_24h, TEMPAR\_24h, UR\_24h, QNH\_24h;

- DIRVENTO\_1h, VISIB\_1h, TIPO\_1h, Q1\_1h, Q2\_1h, Q3\_1h, TEMPAR\_1h, UR\_1h;
- DIRVENTO, VELVENTO, VISIB, CGT, TIPO, Q1, Q2, Q3, 6H\_SKC, TEMPAR, UR, QNH, DIAJULm, HORAm.

b) PERÍODO DE JANEIRO DE 1951 A MARÇO DE 2007

Essa correlação foi feita, também, com dois conjuntos de dados entre as candidatas a variáveis de entrada e cada uma das três visibilidades futuras (VISIB1, VISIB2 E VISIB3). O primeiro conjunto contém visibilidade baixa (menor ou igual a 1600m) e o segundo, com todas as faixas de visibilidade.

Também, aqui, em cada um dos dois casos (visibilidade baixa e todas as visibilidades) foram realizadas três correlações entre as possíveis variáveis preditoras e as visibilidades futuras da primeira, segunda e terceira hora. Nesses seis casos, a correlação foi realizada com 36 variáveis candidatas: 14 da hora atual, 11 da hora anterior, 11 das 24 horas passadas. Estas variáveis estão relacionadas na Tabela 3.5.

**Tabela 3. 5-** Relação das variáveis preditoras.

Variáveis Preditoras		
24 horas anteriores	Hora anterior	Hora atual
DIRVENT_24h	DIRVENT_1h	DIRVENT
VELVEN_24h	VELVEN_1h	VELVEN
VISIB_24h	VISIB_1h	VISIB
CGT_24h	CGT_1h	CGT
TIPO_24h	TIPO_1h	TIPO
Q1_24h	Q1_1h	Q1
Q2_24h	Q2_1h	Q2
Q3_24h	Q3_1h	Q3
TEMPAR_24h	TEMPAR_1h	TEMPAR
QNH_24h	QNH_1h	QNH
UR_24h	UR_1h	UR
-	-	6H_SKC
-	-	HORAm
-	-	DIAJULm

As diferenças da correlação realizada neste item em relação à feita no item “a” anterior são: a) o número das variáveis independentes. Observa-se da Tabela 3.5 que, com exceção das variáveis 6H\_SKC, HORAm e DIAJULm, de todas as variáveis da hora atual foram usadas as correspondentes da hora anterior e de 24 horas anteriores; b) foram incluídos dados de anos anteriores a 1980.

Os resultados das correlações realizadas estão apresentados nas Tabelas 3.6, 3.7 e 3.8, correspondentes, respectivamente, a primeira, segunda e terceira hora futura. Em cada tabela, a coluna da esquerda refere-se ao conjunto com baixa visibilidade (menor ou igual 1600m) e a da direita, ao conjunto com todas as faixas de visibilidade.

**Tabela 3. 6** - Correlações entre as 36 candidatas a variáveis de entrada e a visibilidade na primeira hora futura (VISIB1) realizadas com dois conjuntos de dados. O primeiro com 19.927 eventos de visibilidade baixa (menor ou igual a 1600 m) e o segundo com 50. 889 eventos de visibilidade em todas as faixas. Os dados se referem ao período de janeiro de 1951 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As correlações estão ordenadas pelo valor absoluto do fator de correlação.

Visib baixa (19927 eventos)		Visib geral (50889 eventos)	
Variável	fator	Variável	fator
VISIB	0.8575	VISIB	0.8382
VISIB_1h	0.7011	VISIB_1h	0.6706
UR	-0.6508	UR	-0.4749
UR_1h	-0.6089	CGT	-0.4718
TEMPAR	0.6021	UR_1h	-0.4025
TEMPAR_1h	0.5737	CGT_1h	-0.3977
TEMPAR_24h	0.5480	TEMPAR	0.3961
UR_24h	-0.5281	TEMPAR_1h	0.3463
CGT	-0.5165	Q1	-0.3453
DIRVENTO	0.5030	TEMPAR_24h	0.3143
DIRVENTO_1h	0.4866	VISIB_24h	0.3091
CGT_1h	-0.4601	UR_24h	-0.2988
Q2	0.4248	Q1_1h	-0.2876
VELVENTO_1h	0.4145	VELVENTO_1h	0.2834
VELVENTO	0.4062	VELVENTO	0.2811
Hora_m	-0.3955	DIRVENTO	0.2494
Q2_1h	0.3889	CGT_24h	-0.2407
Q1	-0.3718	DIRVENTO_1h	0.2282
VISIB_24h	0.3627	Hora_m	-0.2264
Q3	0.3478	VELVENTO_24h	0.1977
DIRVENTO_24h	0.3395	Q1_24h	-0.1835
CGT_24h	-0.3313	DIRVENTO_24h	0.1437
Q3_1h	0.3235	QNH	-0.1244
Q1_1h	-0.3052	Diajul_m	-0.1242
VELVENTO_24h	0.2822	QNH_24h	-0.1077
6H_SKC	-0.2819	Q2	0.1036
QNH_24h	-0.2632	QNH_1h	-0.1006
QNH	-0.2566	Q2_1h	0.1005
QNH_1h	-0.2478	Q2_24h	0.0842
Q1_24h	-0.2227	Q3	0.0614
Diajul_m	-0.2188	Q3_1h	0.0614
Q3_24h	0.1832	Q3_24h	0.0461
Q2_24h	0.1370	6H_SKC	-0.0393
TIPO_1h	0.0798	TIPO_24h	0.0120
TIPO	0.0387	TIPO	-0.0073
TIPO_24h	-0.0088	TIPO_1h	0.0051

**Tabela 3. 7-** Correlações entre as 36 candidatas a variáveis de entrada e a visibilidade na segunda hora futura (VISIB2) realizadas com dois conjuntos de dados. O primeiro com 19.792 eventos de visibilidade baixa (menor ou igual a 1600 m) e o segundo com 57. 627 eventos de visibilidade em todas as faixas. Os dados se referem ao período de janeiro de 1951 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As correlações estão ordenadas pelo valor absoluto do fator de correlação.

<i>Visib baixa (19792 eventos)</i>		<i>Visib geral (57627 eventos)</i>	
<i>Variável</i>	<i>fator</i>	<i>Variável</i>	<i>fator</i>
VISIB	0,6850	VISIB	0,7349
UR	-0,5916	VISIB_1h	0,6064
UR_1h	-0,5426	CGT	-0,4846
TEMPAR	0,5365	CGT_1h	-0,4055
VISIB_1h	0,5346	UR	-0,3382
Hora_m	-0,5170	TEMPAR	0,3189
TEMPAR_1h	0,4967	VELVENTO	0,2813
TEMPAR_24h	0,4882	UR_1h	-0,2754
UR_24h	-0,4779	VELVENTO_1h	0,2711
DIRVENTO	0,4434	TEMPAR_1h	0,2697
CGT	-0,4365	VISIB_24h	0,2692
DIRVENTO_1h	0,4125	Q1	-0,2653
VELVENTO	0,4093	TEMPAR_24h	0,2542
VELVENTO_1h	0,3920	Hora_m	-0,2416
CGT_1h	-0,3760	Q1_1h	-0,2195
Q2	0,3619	VELVENTO_24h	0,2147
Q2_1h	0,3229	UR_24h	-0,2034
Q3	0,2985	CGT_24h	-0,1945
Q1	-0,2937	DIRVENTO	0,1702
VELVENTO_24h	0,2772	DIRVENTO_1h	0,1460
DIRVENTO_24h	0,2757	QNH	-0,1424
Q3_1h	0,2737	Q1_24h	-0,1414
VISIB_24h	0,2590	QNH_1h	-0,1327
6H_SKC	-0,2457	QNH_24h	-0,1187
QNH_24h	-0,2323	TIPO	-0,0678
Q1_1h	-0,2323	TIPO_1h	-0,0656
QNH_1h	-0,2269	DIRVENTO_24h	0,0499
QNH	-0,2233	Q3_24h	0,0491
CGT_24h	-0,2211	6H_SKC	0,0385
Q1_24h	-0,1792	Q2_24h	0,0382
Q3_24h	0,1665	Q3_1h	0,0334
Diajul_m	-0,1655	Q3	0,0271
TIPO_1h	0,1130	Diajul_m	-0,0174
Q2_24h	0,0928	TIPO_24h	-0,0163
TIPO	0,0780	Q2	-0,0163
TIPO_24h	0,0241	Q2_1h	-0,0133

**Tabela 3. 8-** Correlações entre as 36 candidatas a variáveis de entrada e a visibilidade na terceira hora futura (VISIB3) realizadas com dois conjuntos de dados. O primeiro com 19.861 eventos de visibilidade baixa (menor ou igual a 1600 m) e o segundo com 57. 589 eventos de visibilidade em todas as faixas. Os dados se referem ao período de janeiro de 1951 a março de 2007 da Estação Meteorológica de Cumbica/Guarulhos. As correlações estão ordenadas pelo valor absoluto do fator de correlação.

<i>Visib baixa (19861 eventos)</i>		<i>Visib geral (57589 eventos)</i>	
<b>Variável</b>	<b>fator</b>	<b>Variável</b>	<b>fator</b>
UR	-0,5821	VISIB	0,4505
Hora_m	-0,5675	VISIB_1h	0,3278
VISIB	0,5318	Hora_m	-0,2769
TEMPAR	0,5280	CGT	-0,2307
UR_1h	-0,5167	TEMPAR	0,2279
TEMPAR_24h	0,4862	TEMPAR_24h	0,1933
TEMPAR_1h	0,4786	Q1	-0,1868
VELVENTO	0,4776	UR	-0,1832
UR_24h	-0,4740	TEMPAR_1h	0,1732
VELVENTO_1h	0,4613	CGT_1h	-0,1722
DIRVENTO	0,4564	DIRVENTO	0,1718
DIRVENTO_1h	0,4254	VELVENTO	0,1579
VISIB_1h	0,4037	DIRVENTO_1h	0,1536
Q2	0,3823	VELVENTO_1h	0,1472
CGT	-0,3748	Q1_1h	-0,1310
VELVENTO_24h	0,3667	UR_1h	-0,1229
Q2_1h	0,3448	VISIB_24h	0,1153
DIRVENTO_24h	0,3308	DIRVENTO_24h	0,1125
6H_SKC	-0,3028	VELVENTO_24h	0,1016
QNH_1h	-0,3022	UR_24h	-0,0952
CGT_1h	-0,3022	Q1_24h	-0,0875
QNH_24h	-0,2990	CGT_24h	-0,0824
QNH	-0,2904	QNH_1h	-0,0791
Diajul_m	-0,2834	QNH_24h	-0,0724
Q3	0,2677	QNH	-0,0706
Q3_1h	0,2494	6H_SKC	-0,0625
Q1	-0,2276	Q2_1h	0,0623
TIPO_1h	0,2040	Q2	0,0619
CGT_24h	-0,1853	TIPO_1h	0,0608
TIPO	0,1819	Q2_24h	0,0594
VISIB_24h	0,1751	TIPO_24h	0,0475
Q3_24h	0,1743	TIPO	0,0455
Q1_1h	-0,1692	Diajul_m	-0,0420
Q2_24h	0,1563	Q3	0,0378
Q1_24h	-0,1355	Q3_1h	0,0350
TIPO_24h	0,1054	Q3_24h	0,0345

## 3.2 MÉTODOS

### 3.2.1 INTRODUÇÃO

A metodologia do trabalho terá a meta de previsão da visibilidade e teto para as três primeiras horas futuras, a partir de dados meteorológicos definidos na seção anterior.

## IDÉIA INICIAL

A pretensão de desenvolver um sistema de previsão para as faixas de visibilidade e de teto discriminadas a seguir, cada uma para as três primeiras horas futuras, é chamada aqui de *idéia inicial*. Esta foi desenvolvida com a primeira parte dos dados recebidos (de janeiro 1980 a março de 2007) e está em consonância com a Tabela 3.10 - Mínimos Operacionais para o Aeroporto de Guarulhos – SP:

### I. Faixas de visibilidade:

- 1 -  $vis < 400$  m;
- 2 -  $400 \text{ m} \leq vis < 800$  m;
- 3 -  $800 \text{ m} \leq vis < 1200$  m;
- 4 -  $1200 \text{ m} \leq vis < 1600$  m.
- 5 -  $vis \geq 1600$  m (visibilidade alta).

### II. Faixas de teto (altura de nuvens):

- $teto < 30$  m;
- $30 \text{ m} \leq teto < 60$  m;
- $60 \text{ m} \leq teto < 90$  m;
- $90 \text{ m} \leq teto < 180$  m;
- $180 \text{ m} \leq teto < 210$  m;
- $teto \geq 210$  m.

Com base na análise dos dados de trabalho e para estabelecer as relações entre as faixas de visibilidade e o fenômeno restritor da mesma no Aeroporto de Guarulhos, criou-se a Tabela 3.9. As faixas ou intervalos de visibilidade dessa tabela estão relacionados às faixas de visibilidade da *idéia inicial*.

Na primeira parte da Tabela 3.9, são apresentadas as três primeiras faixas de visibilidade (0-399, 400-799, 800-999m) e os relacionados números de ocorrência dos



fenômenos restritores de visibilidade (nevoeiro, névoa seca, chuva e chuveiro). Da mesma forma, na segunda parte da Tabela 3.9 são apresentados os dois últimos intervalos de visibilidade (1000-1199, 1200-1599m) e seus números de ocorrências dos fenômenos restritores da visibilidade (neste caso, nevoeiro, por sua própria definição, é substituído por névoa úmida). No final de cada parte da Tabela 3.9, é apresentada a percentagem de ocorrência do fenômeno em relação à respectiva amplitude da visibilidade (0-999m e 1000-1599m, respectivamente).

Pelos dados mostrados na Tabela 3.9 (a), pode-se concluir que o nevoeiro é responsável por aproximadamente 98% da restrição da visibilidade inferior a 1000m, não apenas pela sua própria definição, mas pela pequena ocorrência dos outros fenômenos restritores. Chama atenção aqui o fato de que as percentagens dos outros fenômenos restritores da visibilidade menor que 1000m (névoa seca, chuva e chuveiro) são quase que traços, em média menor que 1%.

Esclarecendo as siglas que aparecem nessa Tabela 3.10, existem duas condições de voo: VFR (Visual Flight Rules – regras de voo visual) e IFR (Instrument Flight Rules – regras de voo por instrumentos). No voo IFR, há uma orientação do voo por um instrumento que pode ser NDB, VOR/DME ou ILS. A sigla RVR (Runway Visual Range) está relacionada ao alcance visual da pista medido por equipamento. Para mais detalhes sobre estes termos, refira-se ao Glossário. Analisando conjuntamente as Tabelas 3.9 e 3.10, fica claro que o nevoeiro é o principal agente restritor da visibilidade para os mínimos operacionais IFR nas condições VOR/DME, ILS (CAT I e II).

**Tabela 3. 9 (a) e (b)** – A parte (a) da tabela está relacionada aos valores de visibilidade menores que 1000m, e a parte (b) aos valores de 1000 m até 1600 m. Nas duas partes da tabela estão representados os números de ocorrências dos fenômenos meteorológicos (nevoeiro, névoa seca, chuva, chuvisco, e névoa úmida) em cada um dos intervalos de visibilidade da coluna da esquerda, bem como a percentagem total de ocorrência de cada um desses fenômenos em cada uma das partes da tabela. Os dados se referem ao período de 01 de janeiro de 1980 a 31 de março de 2007.

(a)

Visibilidade (m)	Nevoeiro	Nevoa seca	Chuva	Chuvisco
0-399	2362	1	19	3
400-799	1866	9	33	12
800-999	1188	9	47	5
Percentagem	97,5%	0,3%	1,8%	0,4%

(b)

Visibilidade (m)	Nevoa úmida	Nevoa seca	Chuva	Chuvisco
1000-1199	848	55	124	43
1200-1599	1798	124	287	118
Percentagem	77,9%	5,3%	12,1%	4,7%

**Tabela 3. 10** - Mínimos Operacionais para o Aeroporto de Guarulhos – SP. Na segunda à sétima linha estão as informações relativas aos mínimos de visibilidade e teto de nuvens para pouso, nas duas condições de voo: VFR (Visual Flight Rules – regras de voo visual) e IFR (Instrument Flight Rules – regras de voo por instrumentos). A precisão instrumental nas quatro condições de voo IFR (NDB, VOR/DME, (ILS) CAT I, (ILS) CAT II) é inversamente proporcional ao valor da visibilidade e do teto. Na última linha estão as condições mínimas para decolagem (fonte: DECEA).

		Visibilidade (m)	Teto (ft)	Teto (m)
VFR		5000	1500	450
IFR	NDB	1600	700	210
	VOR/DME	1200	600	180
	(ILS) CAT I	800	200	60
	(ILS) CAT II	400	100	30
Decolagem		RVR = 200m; teto = 180m ou 600 ft		

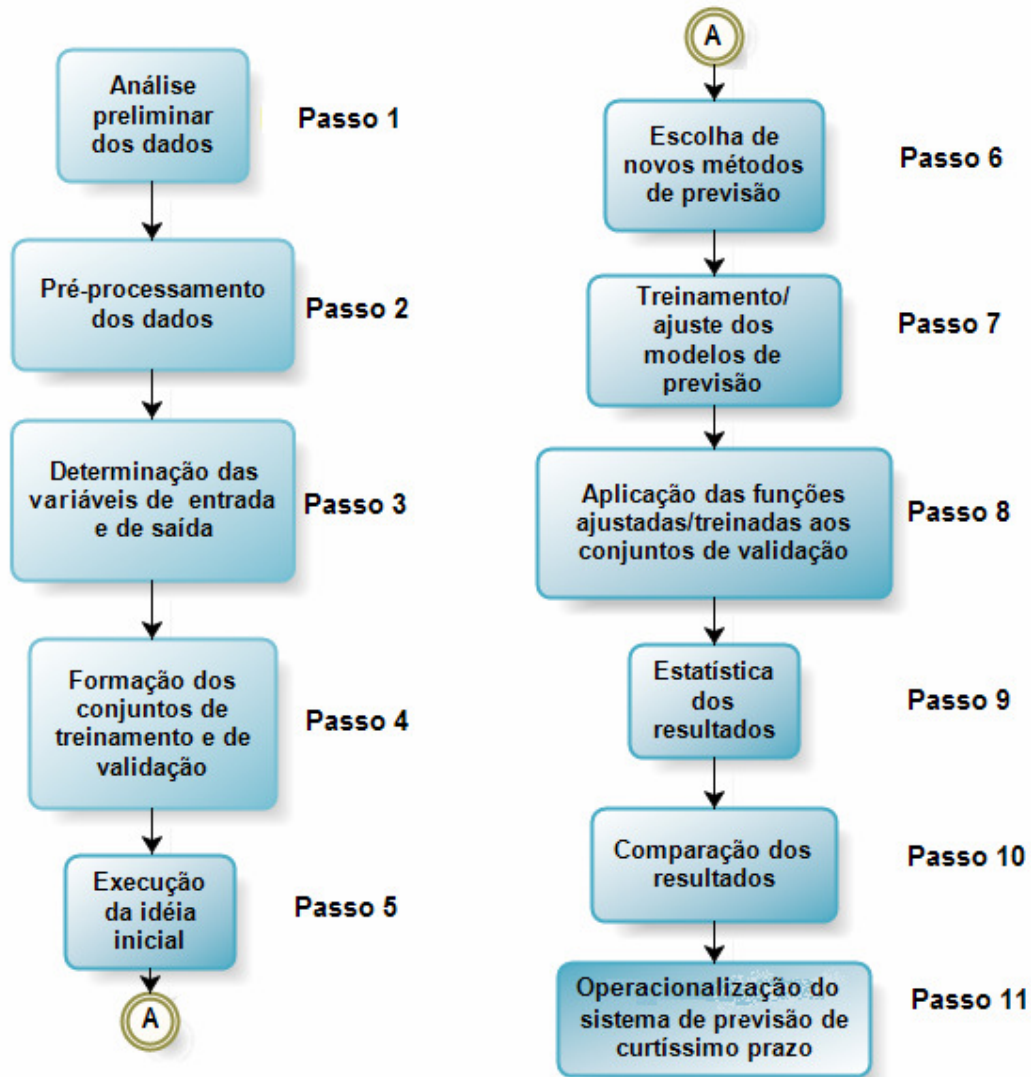
## FERRAMENTAS COMPUTACIONAIS

Foram utilizados os programas de rede neural *NeuroShell Predictor* e *NeuroShell Classifier* da empresa *Ward Systems Group, Inc.* O *NeuroShell Classifier* usa as Redes Neurais Probabilísticas - PNN para problemas de classificação e o *NeuroShell Predictor* usa as Redes Neurais de Regressão Geral - GRNN para resolver problemas de reconhecimento de padrões e de previsão.

Essas redes são treinadas pelo algoritmo *TurboProp2*, desenvolvido pela *Ward Systems*, o qual é uma variante do algoritmo *Correlação em Cascata* desenvolvido por *Fahlman et al.* (1991). Durante o treinamento da rede, o algoritmo *TurboProp2* aumenta paulatinamente o número de neurônios e de camadas internas da rede. Desta forma, a rede que está sendo treinada começa com as *relações lineares* entre entradas e saídas, onde pesos são assinalados às ligações entre a entrada e a saída e, após as relações lineares terem sido encontradas, novos neurônios são adicionados, então, à camada oculta de forma que possam ser encontradas as *relações não-lineares*.

### 3.2.2 PASSOS DA METODOLOGIA

A seguir será detalhada a metodologia desenvolvida para alcançar o objetivo proposto no capítulo 1: um modelo de previsão de curtíssimo prazo (3 horas) para visibilidade horizontal e para teto de nuvens, em condições adversas à navegação aérea, para o Aeroporto Internacional de Guarulhos. Esta metodologia está esquematizada através do fluxograma da Figura 3.13, onde são mostrados os passos do desenvolvimento do sistema de previsão de curtíssimo prazo.



**Figura 3.13** – Fluxograma com os passos da metodologia empregada para o desenvolvimento do sistema de previsão de curtíssimo prazo para a visibilidade e teto para o Aeroporto de Guarulhos.

A seguir, a metodologia esquematizada no fluxograma da Figura 3.13 é detalhada em cada um dos passos:

Passo 1: Análise preliminar dos dados (apresentada no item 3.1.1);

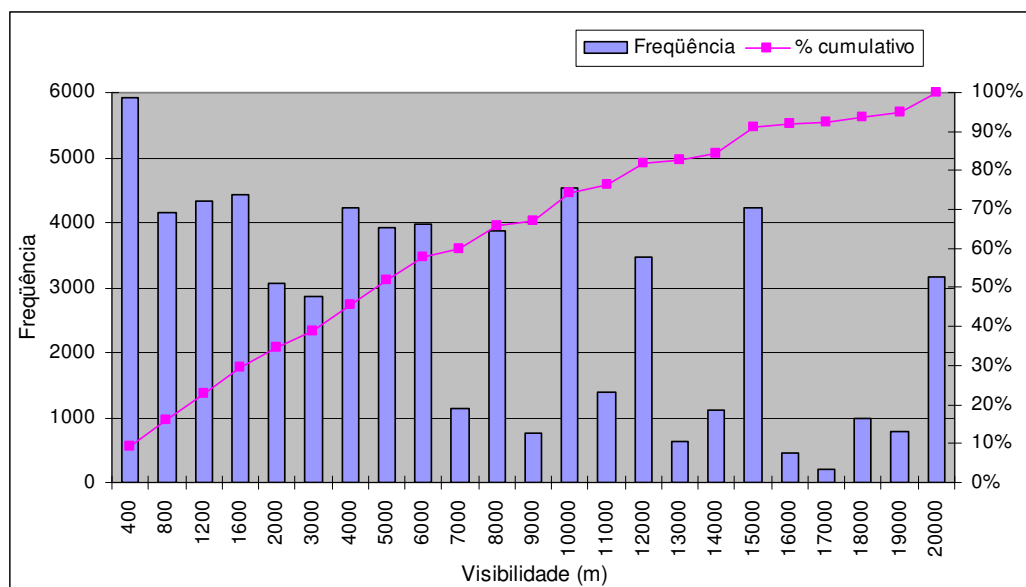
Passo 2: Pré-processamento dos dados (conforme descrito no item 3.1.2);

Passo 3: Determinação das variáveis de entrada e de saída (conforme item 3.1.3);

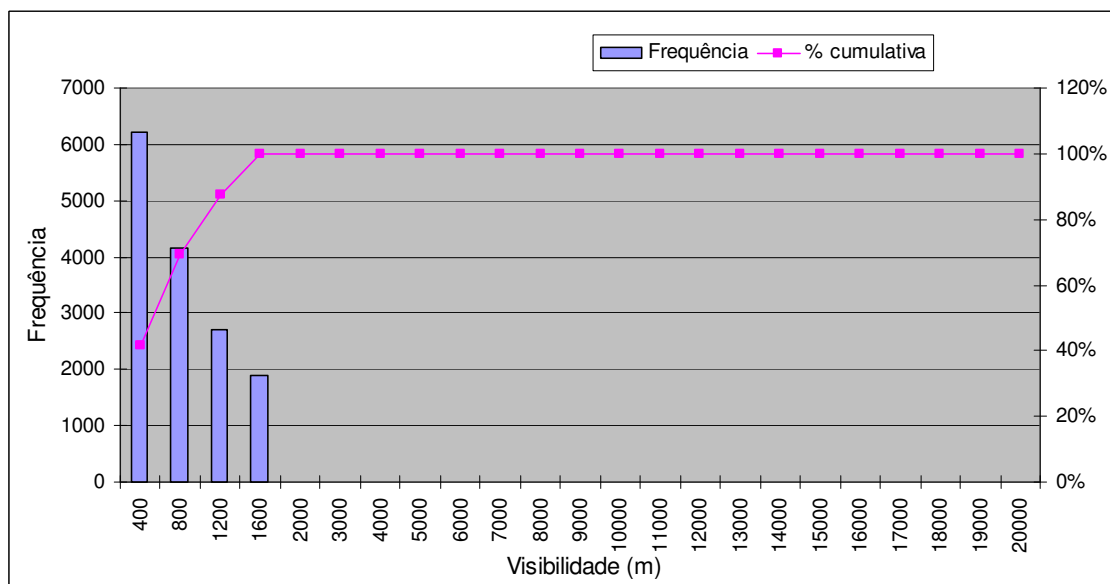
Passo 4: Formação dos conjuntos de treinamento e de validação.

Para a execução da ideia inicial, foram criados três conjuntos de treinamentos, *visibT*, *visibB* e *visib* com os seguintes números de eventos

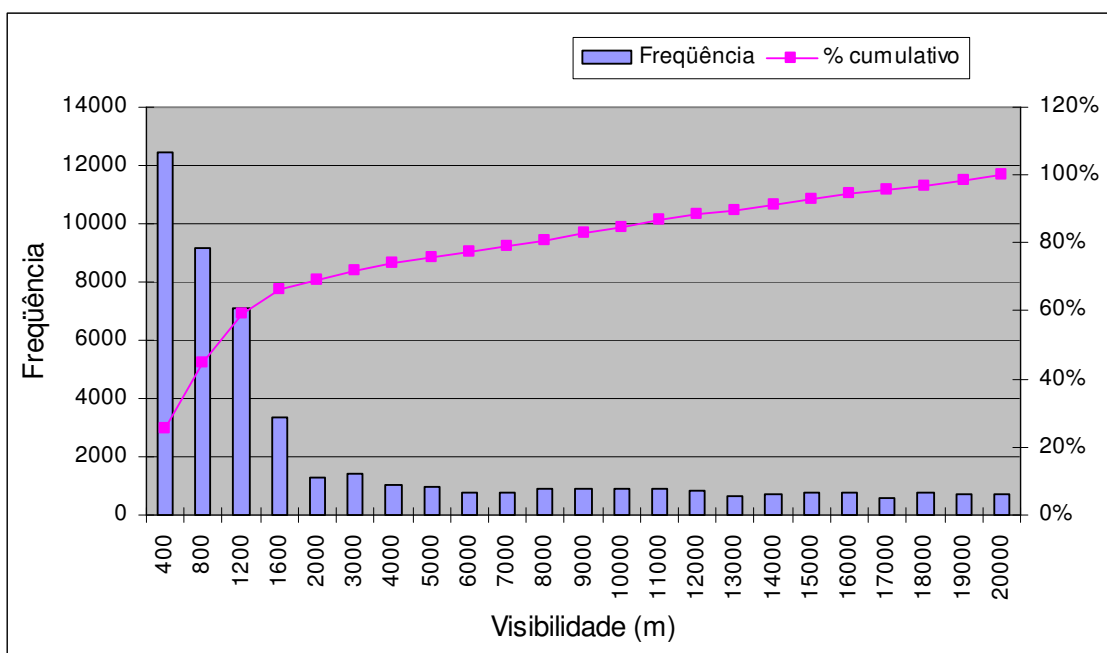
meteorológicos, respectivamente: 63.670, 14.948 e 48.473. Estes conjuntos estão detalhados a seguir nas Figuras 3.14, 3.15 e 3.16.



**Figura 3.14** - Frequência e porcentagem cumulativa dos intervalos de visibilidade do conjunto de treinamento *visibT*, onde todas as faixas de visibilidade estão bem representadas. Em torno de 20% dos valores de visibilidade estão abaixo de 1000m. O total de eventos meteorológicos representados é 63.670 e são relativos ao período de janeiro de 1980 a julho de 2005.

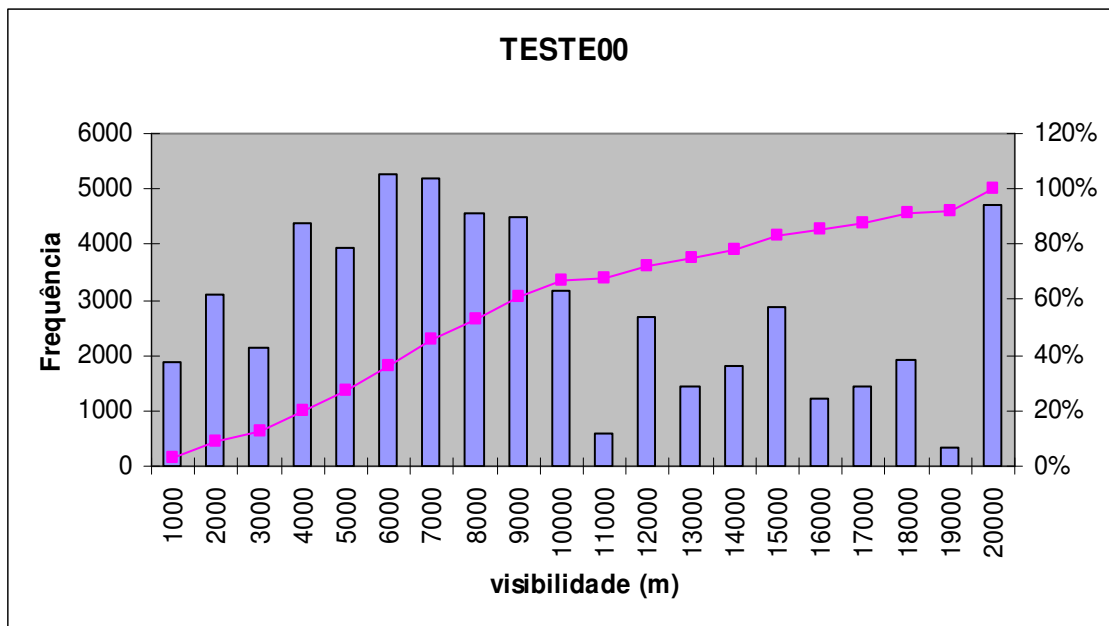


**Figura 3.15** - Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento *visibB*, onde são representados apenas visibilidade baixa (menor ou igual a 1600m). Nota-se que em torno de 80% dos valores de visibilidade estão abaixo de 1000m. O total de eventos meteorológicos representados é 14.948 e são relativos ao período de janeiro de 1980 a julho de 2005.

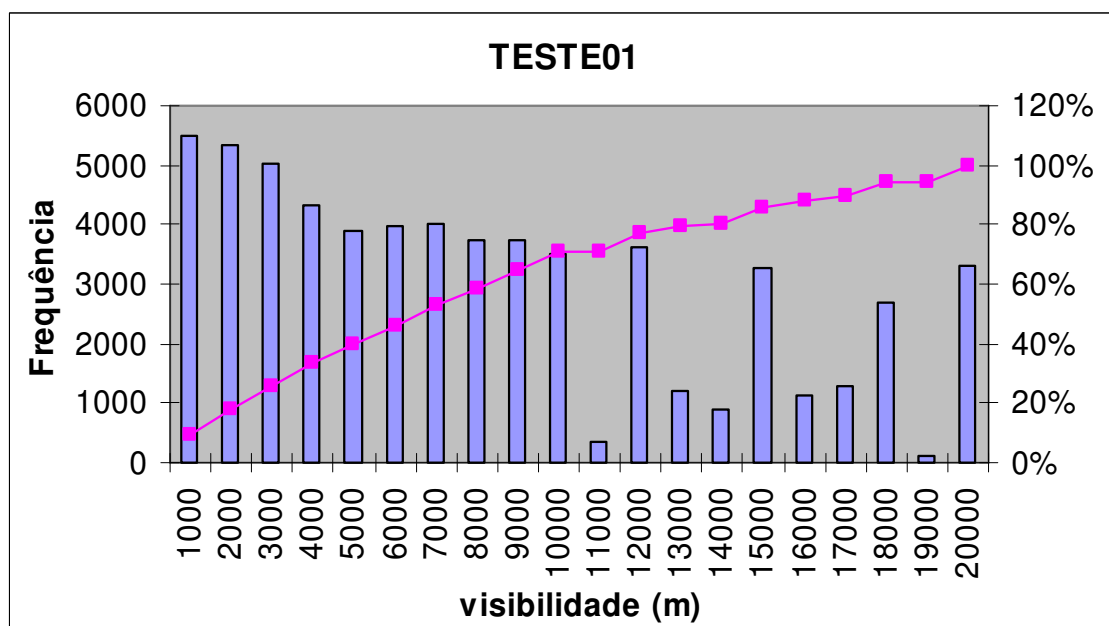


**Figura 3.16** - Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento *visib*, onde são mostrados valores representativos de todas as faixas de visibilidade, mas predominando os valores de visibilidade baixa (menor ou igual a 1600m). Nota-se que em torno de 50% dos valores de visibilidade estão abaixo de 1000m. O total de eventos meteorológicos representados é 48.473 e são relativos ao período de janeiro de 1980 a julho de 2005.

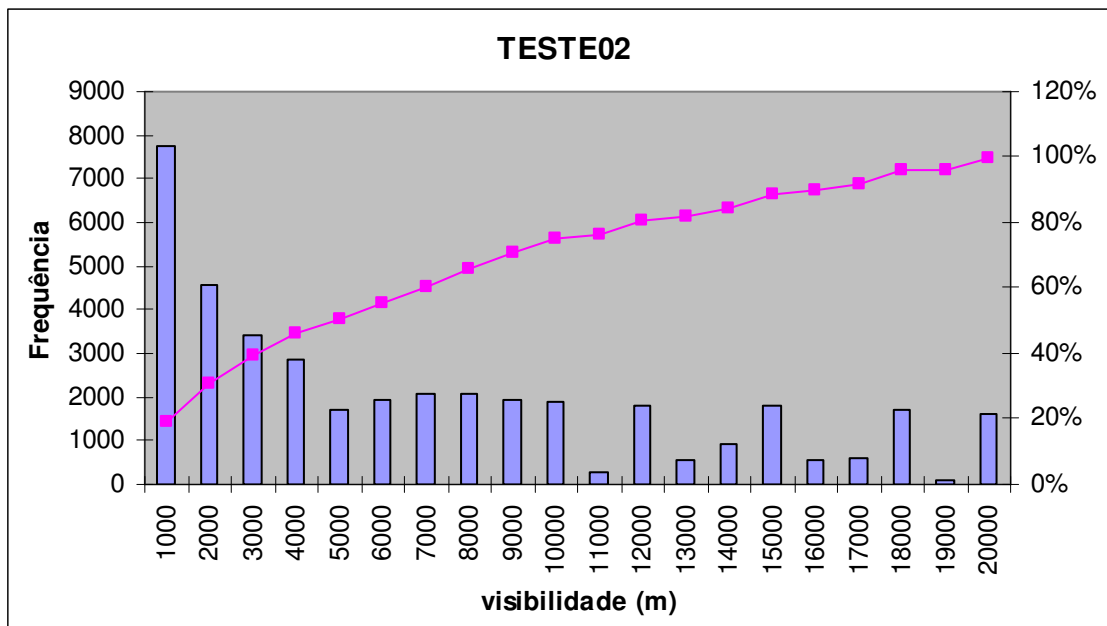
Foram criados, para o período de janeiro de 1951 a março de julho de 2007, oito conjuntos de treinamentos, chamados de TESTE00, TESTE01, TESTE02, TESTE03, TESTE04, TESTE05, TESTE06, TESTE07. O diferencial destes conjuntos é a proporção crescente da frequência das faixas ou intervalos de visibilidade mais baixa. A finalidade deste procedimento é mostrar que não somente a escolha das variáveis preditoras é importante para o desempenho do modelo final como também a proporção das informações (faixas de visibilidade) contidas no conjunto de treinamento.



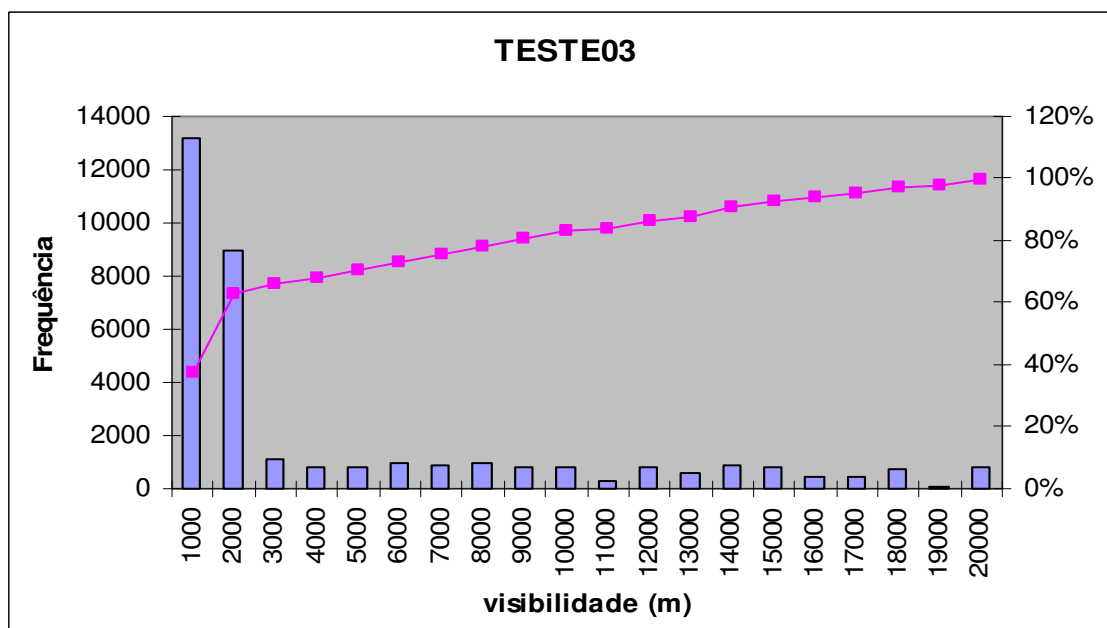
**Figura 3.17** - Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento TESTE00. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005.



**Figura 3.18** - Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento TESTE01. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005..

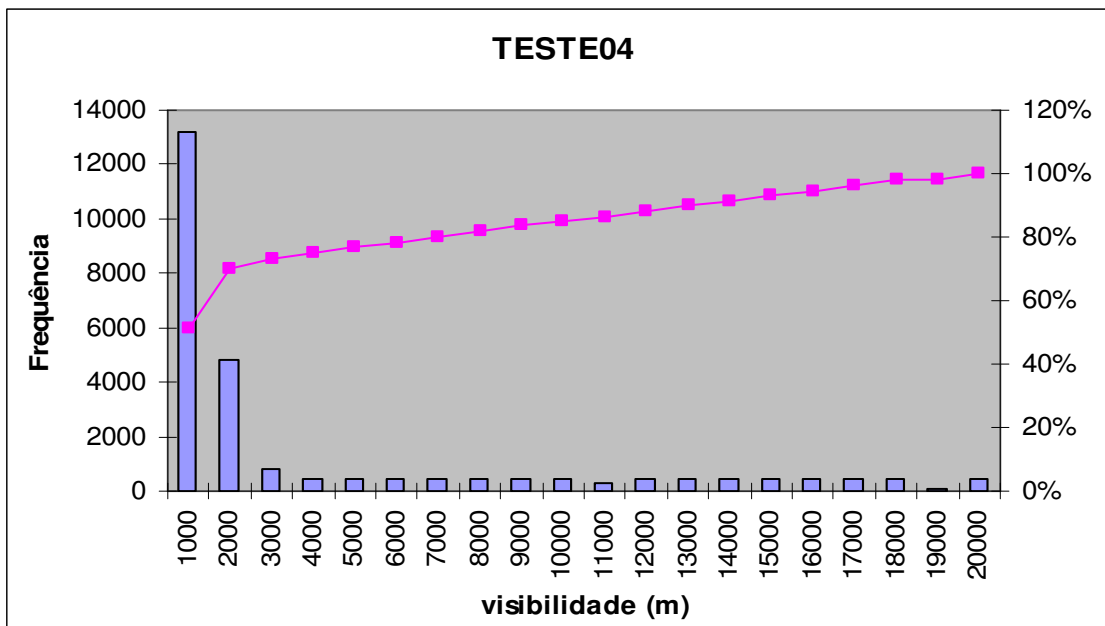


**Figura 3.19** – Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento TESTE02. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005.

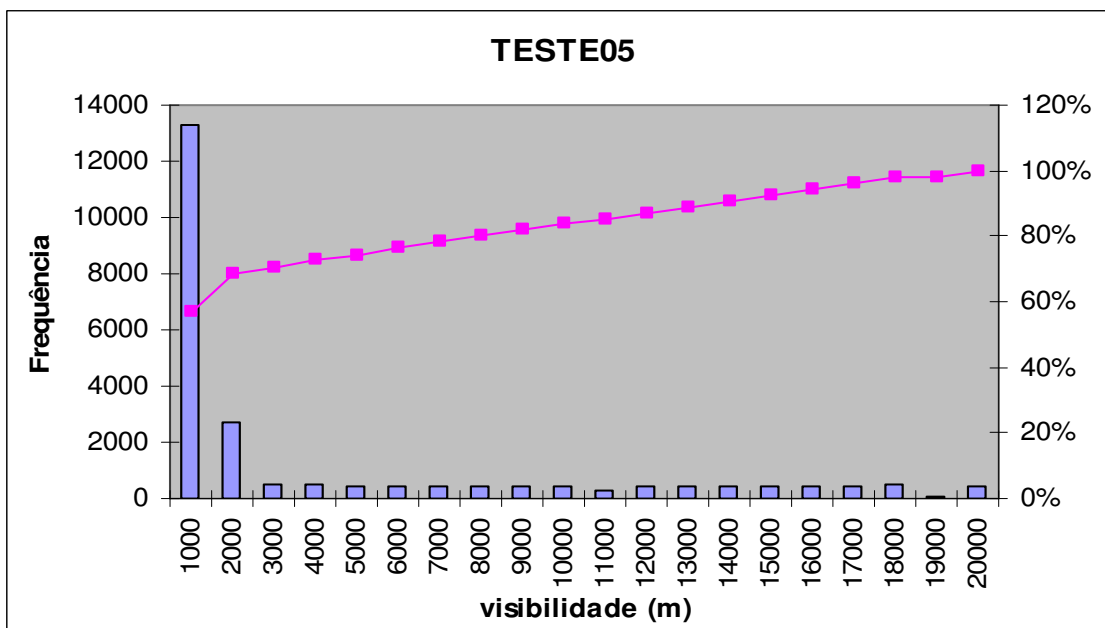


**Figura 3.20** – Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento TESTE03. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005.

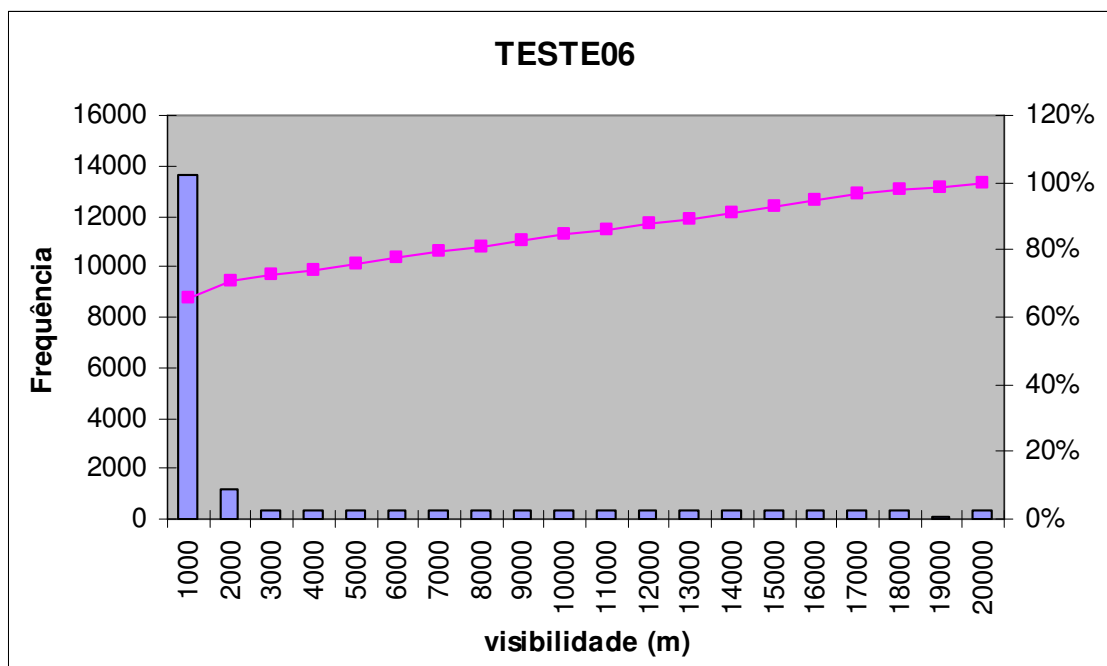




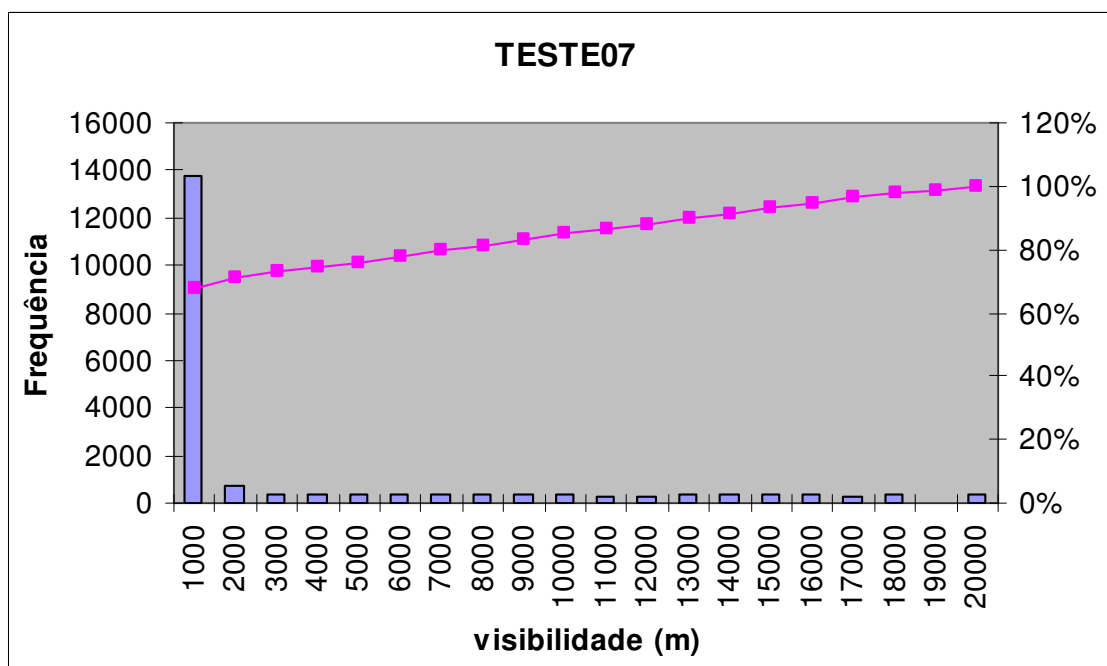
**Figura 3.21** – Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento TESTE04. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005.



**Figura 3.22** – Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento TESTE05. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005.



**Figura 3.23** – Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento TESTE06. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005.

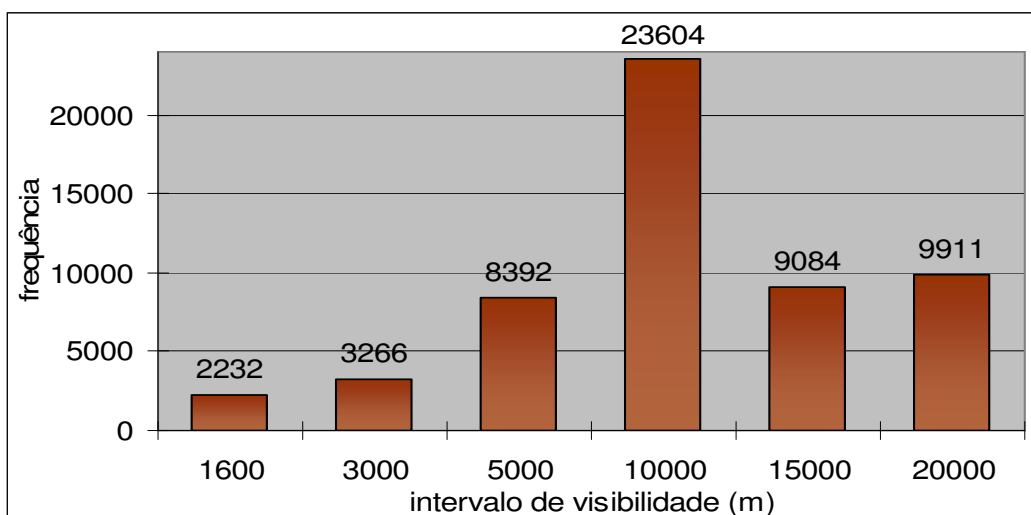


**Figura 3.24** – Frequência e porcentagem cumulativa da visibilidade do conjunto de treinamento TESTE07. Os eventos meteorológicos são relativos ao período de janeiro de 1951 a julho de 2005.

Para avaliar o desempenho dos modelos desenvolvidos, foram criados conjuntos de validação para serem aplicados aos modelos treinados/ajustados. Para

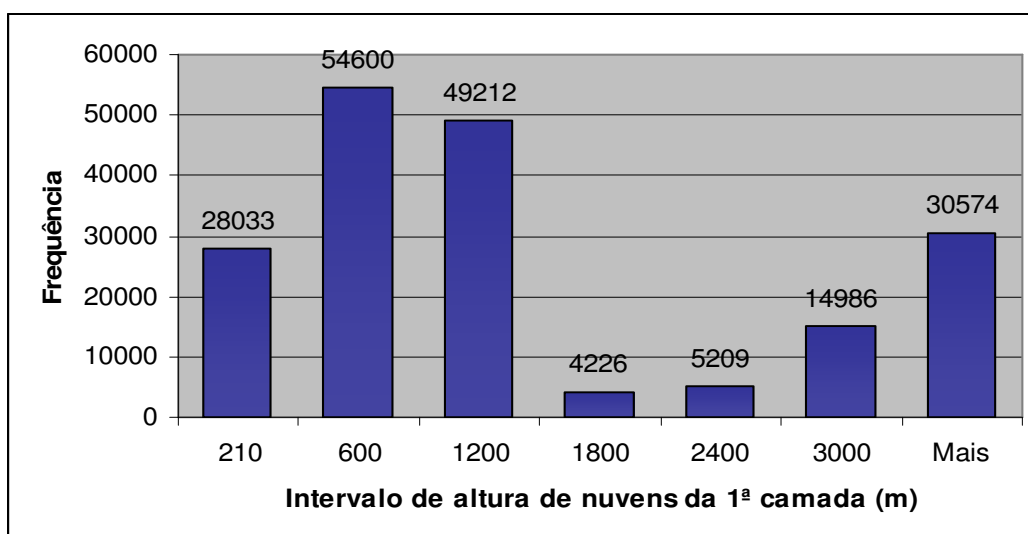
todos os conjuntos de validação usados para os dois períodos citados no item 3.1.5: de janeiro de 1980 a março de 2007 (idéia inicial) e de janeiro de 1951 a março de 2007 (novas metas), foram seguidos os procedimentos:

- a) foram usados dados de até julho de 2005; os dados a partir de agosto de 2005 foram usados para compor a operacionalização do sistema de previsão, conforme o item 4.8 adiante;
- b) todos os conjuntos de validação, como é próprio da sua definição, ficaram à parte, sem nenhuma participação no treinamento, até serem aplicados nos modelos finais treinados;
- c) todos os conjuntos de validação seguiram o modelo representado na Figura 3.25. Nessa figura é apresentado o histograma das faixas de visibilidades do conjunto de validação constituído de 56.489 eventos meteorológicos, referentes ao período de janeiro de 1980 a julho de 2005. Este conjunto refere-se ao primeiro preditando: visibilidade na primeira hora futura - VISIB1. Observa-se que esse conjunto é uma *imagem*, em menor escala, do conjunto total dos dados do período para estudo representado na Figura 3.9. A intenção de criar essa “imagem” é que o conjunto de validação represente o máximo possível a situação real, e com isto apresente a melhor avaliação possível do desempenho do modelo quanto à generalização (comportamento diante de novas informações). Em resumo, todos os conjuntos de validação, para o preditando visibilidade deverão ter em média as percentagens 4%, 6%, 15%, 42%, 16% e 17%, correspondentes aos percentuais de cada uma das seis faixas de visibilidade em relação à quantidade total de eventos meteorológicos. Isto tudo conforme o exemplo do conjunto de validação mostrado na Figura 3.25.

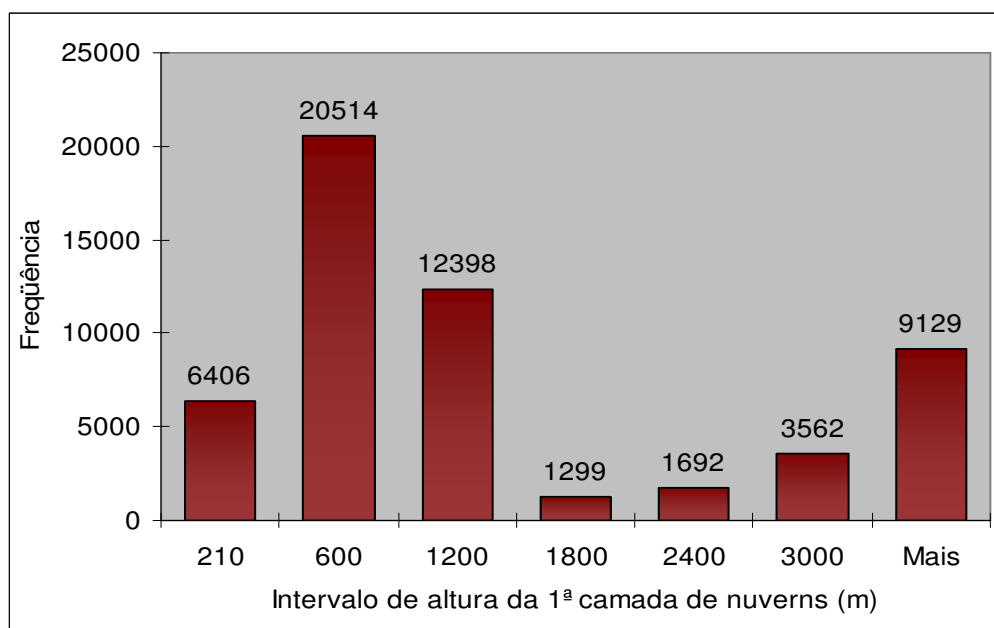


**Figura 3.25** - Histograma das faixas de visibilidade do conjunto de validação para o preditando visibilidade na primeira hora futura (VISIB1), referente ao período de janeiro de 1980 a julho de 2005. É constituído de 56.489 eventos meteorológicos.

De forma semelhante foi feito para os valores de altura das nuvens. Na Figura 3.27 é apresentado o histograma das faixas de altura das nuvens da primeira camada, com 55.000 eventos meteorológicos, relativo ao conjunto de validação para a previsão da variável ALT1 (altura das nuvens da primeira camada para a primeira hora futura). Observa-se que esse conjunto é uma *imagem*, em menor escala, do conjunto total dos dados representado na Figura 3.26, que possui 186.760 eventos meteorológicos.



**Figura 3.26** - Histograma dos valores da altura de nuvens da 1ª camada do conjunto total de dados fornecidos para o período de janeiro de 1980 a março de 2007. É constituído de 186.760 eventos meteorológicos.



**Figura 3.27** - Histograma dos intervalos de altura da 1ª camada de nuvens do conjunto de validação para testar a capacidade de generalização do modelo. O período abrange de janeiro de 1980 a julho de 2005. É constituído de 55.000 eventos.

Passo 5: Execução da idéia inicial.

Neste passo, foi utilizada a ferramenta computacional *NeuroShell Classifier* para o treinamento e a validação dos conjuntos definidos no Passo 4. Aqui é verificada a capacidade do modelo treinado/avaliado em satisfazer a idéia inicial de prever as cinco faixas de visibilidades e as seis faixas de teto nas três horas futuras discriminadas.

Passo 6: Escolha de novos métodos de previsão:

- a) Diante dos fracos resultados para a idéia inicial, serão definidas novas metas de previsão;
- b) As novas metas serão primeiramente treinadas com uma metodologia de uso de Regressão Linear Múltipla (RLM);
- c) Em seguida, será usado um procedimento usando redes GRNN;
- d) Por fim, será usado um procedimento com redes PNN utilizando todas as 36 variáveis preditoras;

Passo 7: Treinamento / Ajuste dos modelos de previsão. Neste passo cada método ou modelo será treinado/ajustado usando os oito conjuntos de treinamentos discriminados, para cada uma das três horas futuras;

Passo 8: Em seguida cada um dos modelos treinados / ajustados no Passo 7 será aplicado ao mesmo conjunto de validação;

Passo 9: Após cada conjunto treinado/ajustado ser aplicado sobre o conjunto de validação, será feita a estatística dos resultados;

Passo 10: Neste passo ser feita avaliação e comparação dos resultados:

- a) a avaliação do desempenho dos diversos treinamentos será feita através de medidas de avaliação derivadas de tabelas de contingência semelhantes à Tabela 3.11.
- b) será criada uma tabela com o resumo das avaliações feitas neste subitem “a” anterior para a comparação do desempenho dos modelos;

**Tabela 3. 11-** Tabela de contingência 2x2 (Fonte: Wilks, 2006).

		Observado		
		evento	não-evento	
Previsto	evento	a	b	a+b
	não-evento	c	d	c+d
		a+c	b+d	n=a+b+c+d

Na tabela 3.11, os significados das letras são:

- a é o número de acertos do evento;
- b é o número de eventos previstos, mas que não foram observados;
- c é o número de eventos que foram observados e não foram previstos;
- d é o número de acertos do não-evento;

- (a+c) é o total de eventos observados;
- (b+d) é o total de não-eventos observados;
- (a+b) é o total de eventos previstos;
- (c+d) é o total de não-eventos previstos
- $n$  é o tamanho do conjunto ou amostra.

As medidas de avaliação do desempenho do modelo ou as estatísticas de previsão utilizadas neste trabalho, derivadas da Tabela 3.11, são:

POD – Probabilidade de Detecção ou Previsão Correta do Evento (ou *verdadeiro-positivo*); fornece a taxa de previsão correta do evento desejado (por exemplo, visibilidade-baixa ou nevoeiro):

$$POD = \frac{a}{a + c}$$

POFD – Probabilidade de Falsa Detecção ou alarme falso (ou *falso-positivo*). Corresponde à razão do número de alarme falso dividido pelo total de não-eventos observados:

$$POFD = \frac{b}{b + d}$$

CSI – Índice Crítico de Sucesso é o número correto de previsão do evento dividido pelo número total de ocasiões nas quais o evento foi previsto e/ou observado. Pode ser visto como a proporção correta da quantidade que está sendo previsto, após remover as previsões corretas do não-evento:

$$CSI = \frac{a}{a + b + c}$$

B – Bias é comparação da previsão com a observação. É a razão entre o número de previsões do evento pelo número total de eventos observados. Uma previsão não tendenciosa exibe  $B = 1$ . Um valor maior que 1 indica uma previsão superestimada e menor que 1, uma previsão subestimada.

$$B = \frac{a + b}{a + c}$$

SA – Soma dos Acertos. Está relacionada às somas das taxas de acertos de eventos e de não-eventos, ou a soma das taxas relativas à *sensibilidade* e à *especificidade*. Em outras palavras:

$$SA = \frac{a}{a + c} + \frac{d}{b + d}$$

Estas medidas de verificação são baseadas em Wilks (2006).

*Sensibilidade* é a habilidade de o modelo classificar corretamente o evento desejado ou o percentual de *verdadeiro-positivo*. *Especificidade* é a habilidade de o modelo classificar corretamente o não-evento ou o percentual de *verdadeiro-negativo*. Quanto maiores forem conjuntamente os dois valores, melhor será o modelo de previsão. Ter uma sensibilidade alta e uma especificidade baixa significa que o modelo emite muito *alarme falso* (ou *falso-positivo*) ou, o modelo superestima. Porém, ter uma sensibilidade baixa e uma especificidade alta significa que o modelo erra muito (muito *falso-negativo*) ou subestima. Na Tabela 3.12 é apresentada uma visão do desempenho de um modelo de previsão em relação à categoria das medidas da sensibilidade e da especificidade.

Passo 11: Operacionalização do sistema de previsão de curtíssimo prazo.

Após terem sido executados os treinamentos e as validações dos modelos de previsão e analisados os seus desempenhos e tendo sido determinando o melhor



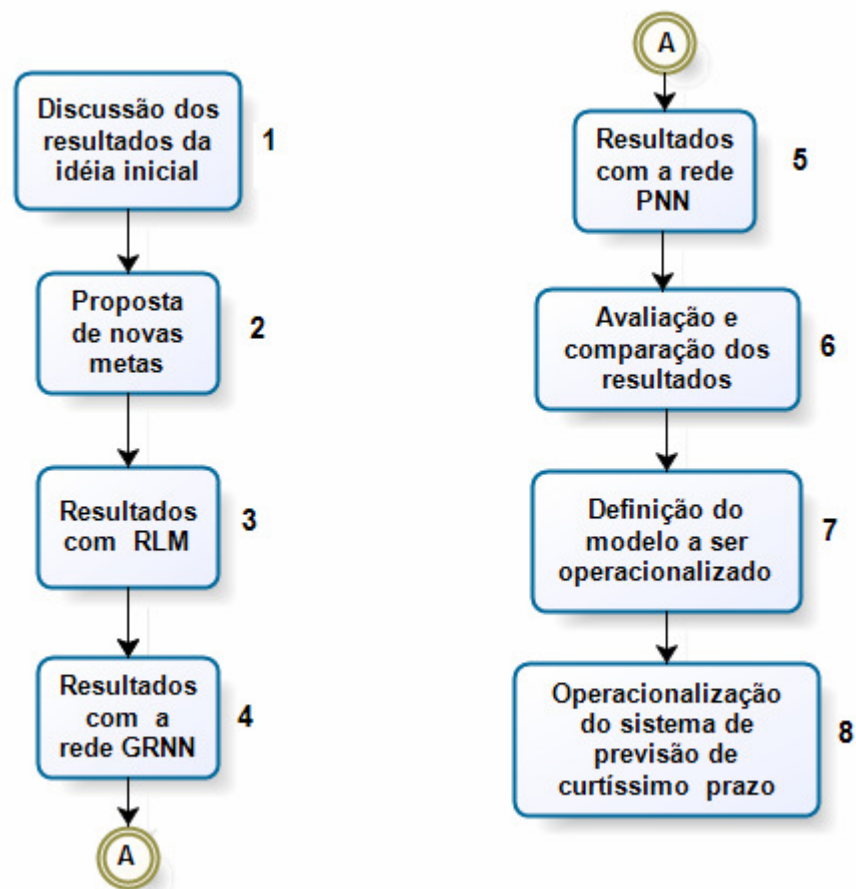
deles, o código ou função ajustada desse modelo, chamado aqui de *executável*, será, então, embutido no modelo operacional a ser implantado pelo Comando da Aeronáutica no Aeroporto de Guarulhos. As etapas operacionais do sistema de previsão desenvolvido são apresentadas no apêndice D e o desenvolvimento da operacionalização desse sistema está no Anexo B.

**Tabela 3. 12**– Relação do desempenho de um modelo de previsão em relação à sensibilidade e à especificidade.

<b>Desempenho de um modelo de previsão em relação à Sensibilidade e à Especificidade</b>		
Sensibilidade	Especificidade	Desempenho
alta	alta	ótimo
alta	baixa	superestimação
baixa	alta	subestimação
baixa	baixa	péssimo

## CAPÍTULO 4 – RESULTADOS E DISCUSSÃO

Seguindo a metodologia apresentada no item 3.2.2, os resultados obtidos e suas avaliações, a proposta de novas metas, a definição do modelo a ser operacionalizado e a operacionalização são apresentados seguindo o fluxograma da Figura 4.1.



**Figura 4.1** – Fluxograma indicando a seqüência de apresentação da discussão dos resultados iniciais obtidos, da proposta de novas metas, dos resultados com a RLM, a rede GRNN e a rede PNN, da avaliação dos resultados e da operacionalização do sistema.

A seguir são detalhadas as fases do fluxograma da Figura 4.1:

1. Nesta fase, são mostrados e discutidos os resultados conforme a idéia inicial: previsão de ocorrência da visibilidade e do teto nas faixas que mais afetam as operações de pouso e decolagem no Aeroporto de Guarulhos,

consoante à Tabela 3.10. Adiante, no item 4.2, são apresentados esses resultados;

2. São apresentadas as novas metas de previsão de curtíssimo prazo:
  - a) prever a primeira faixa de visibilidade: 400m;
  - b) prever a ocorrência de nevoeiro;
  - c) prever a primeira faixa de teto (menor que 30m);
  - d) prever segunda faixa de teto (menor que 60m)

Ressalta-se que essas novas metas representam um grande auxílio nas necessidades de previsão de visibilidade para o voo IFR nos modos VOR/DME e ILS (CAT I e CAT II), e de previsão de teto no modo ILS (CAT I e CAT II), conforme a Tabela 3.10.

3. São apresentados os resultados das previsões para as novas metas, através do método de Regressão Linear Múltipla;
4. São apresentados os resultados obtidos com o uso de redes GRNN;
5. São mostrados os resultados através do método de redes PNN;
6. É realizada a comparação dos resultados obtidos pelos três métodos;
7. Definição do modelo a ser operacionalizado.

#### 4.1 RESULTADOS PARA A IDÉIA INICIAL

Primeiramente serão apresentados os resultados da previsão da visibilidade considerando as 5 faixas de visibilidade, conforme definidas no capítulo 3:

- 1 – menor que 400m;
- 2 – maior ou igual 400 e menor que 800 m;
- 3 - maior ou igual a 800 e menor que 1200 m;
- 4 - maior ou igual a 1200 e menor que 1600 m.
- 5 - maior ou igual a 1600m (visibilidade alta).

Para a classificação das cinco faixas de visibilidade, para as três horas futuras, foram utilizadas as redes PNN treinadas com o programa *NeuroShell Classifier* por meio do algoritmo *Cascade Correlation* (CC). As redes atingiram as configurações de 137, 150 e 144 neurônios nas camadas ocultas, respectivamente. Os resultados da aplicação das redes neurais treinadas para sobre o conjunto de validação com 56.489 eventos são apresentados nas Tabelas 4.1, 4.2 e 4.3, mostradas a seguir.

De forma semelhante, para a classificação das seis faixas de teto de nuvens, para as três horas futuras, foram utilizadas as redes PNN treinadas com o programa *NeuroShell Classifier* por meio do algoritmo *Cascade Correlation* (CC). As redes treinadas atingiram as configurações de 133, 150 e 135 neurônios nas camadas ocultas, respectivamente. Os resultados da aplicação das redes neurais treinadas sobre o conjunto de validação de 55.000 eventos meteorológicos são apresentados nas Tabelas 4.5, 4.6 e 4.7.

**Tabela 4.1** – Resultados da aplicação da rede neural PNN, treinada pelo programa *NeuroShell Classifier*, sobre o conjunto de validação para a classificação nas cinco faixas de visibilidade conforme a idéia inicial, para a primeira hora futura. O conjunto de validação contém 56.489 observações meteorológicas referentes a janeiro de 1980 a julho de 2005.

		Observado					
		Faixa "1"	Faixa "2"	Faixa "3"	Faixa "4"	Faixa "5"	Total
Previsto	Faixa "1"	509	409	409	346	1557	3230
	Faixa "2"	23	47	58	50	267	445
	Faixa "3"	26	39	86	99	526	776
	Faixa "4"	1	1	5	9	57	73
	Faixa "5"	62	127	323	412	51041	51965
	Total	621	623	881	916	53448	56489
	Verd-pos.	0,8196	0,0754	0,0976	0,0098	0,9550	
	Falso-pos.	0,0487	0,0071	0,0124	0,0012	0,3038	
	Verd-neg.	0,9513	0,9929	0,9876	0,9988	0,6962	
	Falso-neg.	0,1804	0,9246	0,9024	0,9902	0,0450	
	Sensibilidade	81,96%	7,54%	9,76%	0,98%	95,50%	
	Especificidade	95,13%	99,29%	98,76%	99,88%	69,62%	

**Tabela 4.2** - Resultados da aplicação da rede neural PNN, treinada pelo programa *NeuroShell Classifier*, sobre o conjunto de validação para a classificação nas cinco faixas de visibilidade conforme a idéia inicial, para a segunda hora futura. O conjunto de validação contém 56.489 observações meteorológicas referentes a janeiro de 1980 a julho de 2005.

		Observado					Total
		Faixa "1"	Faixa "2"	Faixa "3"	Faixa "4"	Faixa "5"	
Previsto	Faixa "1"	470	218	320	277	3790	5075
	Faixa "2"	11	13	20	19	355	418
	Faixa "3"	0	0	0	0	10	10
	Faixa "4"	0	0	0	0	3	3
	Faixa "5"	125	185	358	409	49906	50983
	Total	606	416	698	705	54064	56489
	Verd-pos.	0,7756	0,0313	0,0000	0,0000	0,9231	
	Falso-pos.	0,0824	0,0072	0,0002	0,0001	0,4441	
	Verd-neg.	0,9176	0,9928	0,9998	0,9999	0,5559	
	Falso-neg.	0,2244	0,9688	1,0000	1,0000	0,0769	
	Sensibilidade	77,56%	3,13%	0,00%	0,00%	92,31%	
	Especificidade	91,76%	99,28%	99,98%	99,99%	55,59%	

**Tabela 4.3** - Resultados da aplicação da rede neural PNN, treinada pelo programa *NeuroShell Classifier*, sobre o conjunto de validação para a classificação nas cinco faixas de visibilidade conforme a idéia inicial, para a terceira hora futura. O conjunto de validação contém 56.489 observações meteorológicas referentes a janeiro de 1980 a julho de 2005.

		Observado					Total
		Faixa "1"	Faixa "2"	Faixa "3"	Faixa "4"	Faixa "5"	
Previsto	Faixa "1"	532	303	232	181	4894	6142
	Faixa "2"	38	48	21	20	964	1091
	Faixa "3"	6	3	4	4	116	133
	Faixa "4"	0	0	0	0	3	3
	Faixa "5"	168	217	237	218	48280	49120
	Total	744	571	494	423	54257	56489
	Verd-pos.	0,7151	0,0841	0,0081	0,0000	0,8898	
	Falso-pos.	0,1006	0,0187	0,0023	0,0001	0,3763	
	Verd-neg.	0,8994	0,9813	0,9977	0,9999	0,6237	
	Falso-neg.	0,2849	0,9159	0,9919	1,0000	0,1102	
	Sensibilidade	71,51%	8,41%	0,81%	0,00%	88,98%	
	Especificidade	89,94%	98,13%	99,77%	99,99%	62,37%	

Na tabela 4.4 é apresentado um resumo dos resultados, relativos à sensibilidade e especificidade, apresentados nas Tabelas 4.1, 4.2 e 4.3. Observa-se que apenas os resultados, para a primeira faixa - visibilidade menor que 400 m, nas três horas futuras, apresentam bons desempenhos, com os valores de sensibilidade e especificidade conjuntamente maiores 71%. Os resultados para a faixa "5" –

visibilidades maiores que 1600m, para as três horas, mostram boa sensibilidade: acima de 88%, mas a especificidade apenas razoável: não chega a 70%. As outras três faixas, "2", "3" e "4", mostram péssimos resultados, pois a sensibilidade não chega a 10%, para as três horas.

**Tabela 4.4** – Resumo relativo à sensibilidade e à especificidade, para os resultados relativos à previsão das cinco faixas de visibilidade, para as três horas futuras, conforme a idéia inicial.

		Faixa "1"	Faixa "2"	Faixa "3"	Faixa "4"	Faixa "5"
1ª hora	Sensibilidade	81,96%	7,54%	9,76%	0,98%	95,50%
	Especificidade	95,13%	99,29%	98,76%	99,88%	69,62%
2ª hora	Sensibilidade	77,56%	3,13%	0,0%	0,0%	92,31%
	Especificidade	91,76%	99,28%	99,98%	99,99%	55,59%
3ª hora	Sensibilidade	71,51%	8,41%	0,81%	0,0%	88,98%
	Especificidade	89,94%	98,13%	99,77%	99,99%	62,37%

**Tabela 4.5** – Resultados da aplicação da rede neural PNN, treinada pelo programa *NeuroShell Classifier*, sobre o conjunto de validação para a classificação nas seis faixas de teto de nuvens conforme a idéia inicial, para a primeira hora futura. O conjunto de validação contém 55.000 observações meteorológicas referentes a janeiro de 1980 a julho de 2005.

		Observado						Total
		Faixa "1"	Faixa "2"	Faixa "3"	Faixa "4"	Faixa "5"	Faixa "6"	
Previsto	Faixa "1"	692	277	225	599	46	1037	2876
	Faixa "2"	93	136	180	706	61	1039	2215
	Faixa "3"	1	4	16	175	7	419	622
	Faixa "4"	5	14	68	704	51	1193	2035
	Faixa "5"	0	1	0	0	0	2	3
	Faixa "6"	41	59	89	904	100	46056	47249
	Total	832	491	578	3088	265	49746	55000
	Verd-pos.	0,8317	0,3	0,0277	0,2280	0,0	0,9258	
	Falso-pos.	0,0403	0,0381	0,0111	0,0256	0,0001	0,2271	
	Verd-neg.	0,9597	0,9619	0,9889	0,9744	0,9999	0,7729	
	Falso-neg.	0,1683	0,7	0,9723	0,8	1,0	0,0742	
	Sensibilidade	83,17%	27,70%	2,77%	22,80%	0%	92,58%	
	Especificidade	95,97%	96,19%	98,89%	97,44%	99,99%	77,29%	

**Tabela 4.6** - Resultados da aplicação da rede neural PNN, treinada pelo programa *NeuroShell Classifier*, sobre o conjunto de validação para a classificação nas seis faixas de teto de nuvens conforme a idéia inicial, para a segunda hora futura. O conjunto de validação contém 55.000 observações meteorológicas referentes a janeiro de 1980 a julho de 2005.

		Observado						Total
		Faixa "1"	Faixa "2"	Faixa "3"	Faixa "4"	Faixa "5"	Faixa "6"	
Previsto	Faixa "1"	623	278	230	569	33	1647	3380
	Faixa "2"	26	44	41	131	11	277	530
	Faixa "3"	17	28	54	354	12	555	1020
	Faixa "4"	25	49	75	556	39	940	1684
	Faixa "5"	0	0	0	1	0	5	6
	Faixa "6"	111	169	199	1617	96	46188	48380
	Total	802	568	599	3228	191	49612	55000
	Verd-pos.	0,7768	0,0775	0,0902	0,1722	0,0	0,931	
	Falso-pos.	0,0509	0,0089	0,0178	0,0218	0,0001	0,4068	
	Verd-neg.	0,9491	0,9911	0,9822	0,9782	0,9999	0,5932	
	Falso-neg.	0,2232	0,9225	0,9098	0,8278	1,0	0,069	
	Sensibilidade	77,68%	7,75%	9,02%	17,22%	0,00%	93,10%	
	Especificidade	94,91%	99,11%	98,22%	97,82%	99,99%	59,32%	

**Tabela 4.7** - Resultados da aplicação da rede neural PNN, treinada pelo programa *NeuroShell Classifier*, sobre o conjunto de validação para a classificação nas seis faixas de teto de nuvens conforme a idéia inicial, para a terceira hora futura. O conjunto de validação contém 55.000 observações meteorológicas referentes a janeiro de 1980 a julho de 2005.

		Observado						Total
		Faixa "1"	Faixa "2"	Faixa "3"	Faixa "4"	Faixa "5"	Faixa "6"	
Previsto	Faixa "1"	622	337	246	645	33	4402	6285
	Faixa "2"	38	33	34	140	5	554	804
	Faixa "3"	1	18	13	138	6	346	522
	Faixa "4"	14	19	31	143	3	608	818
	Faixa "5"	0	0	0	4	0	61	65
	Faixa "6"	163	279	341	2411	91	43221	46506
	Total	838	686	665	3481	138	49192	55000
	Verd-pos.	0,7422	0,0481	0,0195	0,0411	0,0	0,879	
	Falso-pos.	0,1046	0,0142	0,0094	0,0131	0,0012	0,5656	
	Verd-neg.	0,8954	0,9858	0,9906	0,9869	0,9988	0,4344	
	Falso-neg.	0,2578	0,9519	0,9805	0,9589	1,0	0,121	
	Sensibilidade	74,22%	4,81%	1,95%	4,11%	0,00%	87,86%	
	Especificidade	89,54%	98,58%	99,06%	98,69%	99,88%	43,44%	

O resumo dos resultados apresentados nas Tabelas 4.5, 4.6 e 4.7, relativos às três horas futuras, é mostrado na tabela 4.8, onde se percebe que:

- a) O melhor resultado para teto é da primeira faixa, com sensibilidades e especificidades, conjuntamente maiores que 74%.



- b) Para as faixas 2, 3, 4 e 5 existem resultados fracos, todos com valor de sensibilidade abaixo de 28%.
- c) Para a sexta faixa, há um resultado razoável apenas para a primeira hora. Para as outras duas horas, os resultados são fracos, pois apresentam uma baixa especificidade (abaixo de 60%).

**Tabela 4.8** - Resumo relativo à sensibilidade e à especificidade, para os resultados relativos à previsão das seis faixas de teto de nuvens, para as três horas futuras, conforme a idéia inicial

		Faixa "1"	Faixa "2"	Faixa "3"	Faixa "4"	Faixa "5"	Faixa "6"
1ª hora	Sensibilidade	83,17%	27,70%	2,77%	22,80%	0,0%	92,58%
	Especificidade	95,97%	96,19%	98,89%	97,44%	99,99%	77,29%
2ª hora	Sensibilidade	77,68%	7,75%	9,02%	17,22%	0,0%	93,10%
	Especificidade	94,91%	99,11%	98,22%	97,82%	99,99%	59,32%
3ª hora	Sensibilidade	74,22%	4,81%	1,95%	4,11%	0,0%	87,86%
	Especificidade	89,54%	98,58%	99,06%	98,69%	99,88%	43,44%

#### 4.2 DISCUSSÃO SOBRE OS RESULTADOS DA IDÉIA INICIAL

Os resultados, considerando a idéia inicial de classificar as cinco faixas de visibilidade, mostram que as previsões para as faixas 2, 3 e 4 de visibilidade apresentam fracos desempenhos (ver Tabela 4.4). O desempenho para a faixa 5 já é um pouco melhor, mas para essa faixa de visibilidade (maior ou igual a 1600m) não há um interesse maior para a proteção do voo tendo em vista o pouso e a decolagem.

De igual forma, em relação ao teto, as previsões para as faixas 2, 3, 4 e 5 apresentam resultados fracos com valores de sensibilidade menores que 28% (ver Tabela 4.8). Uma possível justificativa para este comportamento dos desempenhos para as diferentes faixas de visibilidade é dada a seguir.

Como em qualquer modelo de previsão, as variáveis de entrada possuem uma importância muito grande no desempenho final do modelo (Valença, 2005).

O objetivo principal do aprendizado em Redes Neurais é a obtenção de modelos com boa capacidade de generalização (previsão a partir de dados novos não usados na aprendizagem) tendo como base um conjunto de dados de treinamento.



O treinamento de redes neurais está normalmente associado à capacidade de as mesmas adaptarem os seus parâmetros ou coeficientes como consequência de sua interação com o meio externo ou às variáveis de entrada (Rezende, 2003). Este ajuste dos coeficientes é feito através de pares de entrada e saída  $(\mathbf{x}, \mathbf{y}_d)$  do conjunto de dados. Essa adaptação dos coeficientes ou pesos modifica a saída  $\mathbf{y}$  (saída do modelo) de forma que a diferença entre  $\mathbf{y}$  e  $\mathbf{y}_d$ , ou seja, o erro diminua a cada iteração, convergindo para um valor desejável.

A questão levantada aqui é o problema da subjetividade dos dados de visibilidade e teto conforme mostrado nas Figuras 3.8 e 3.10 do capítulo 3. Em relação à visibilidade, por exemplo, muitos dos valores registrados como 2000m devem ser na realidade 1700m, 1800m, 1900m, 2100m, 2200m, ou 2300m. O que isto significa? Isto significa que se estará informando à rede, por exemplo, que a visibilidade é 2000m com as condições meteorológicas de 1700m. Em outras palavras, em lugar de se ter uma relação biunívoca dos pares  $(\mathbf{x}, \mathbf{y}_d)$ , teremos diferentes “ $\mathbf{x}$ ” para o mesmo  $\mathbf{y}_d$ . Isto certamente prejudica grandemente o aprendizado da rede e o consequente desempenho preditivo do modelo.

Observa-se, na já citada Figura 3.8, que há uma incerteza maior nos valores de visibilidade entre 600m e 2000 m, dentro das faixas de interesse para a previsão. Porém, os valores abaixo de 600m apresentam um comportamento mais uniforme, parecendo indicar, com isto, uma confiabilidade maior quanto aos seus registros. Então, parece que devido a essa incerteza nos valores registrados, conforme inferido das Figuras 3.8 e 3.10, a rede não consegue discriminar bem as faixas 2, 3 e 4 para a visibilidade e as faixas 2, 3, 4, e 5 para o teto.

Resolver essa problemática da subjetividade dos dados em relação às variáveis visibilidade e teto não é elementar, pois todo o conjunto de dados está contaminado por ela. Esta subjetividade representa uma prática rotineira de décadas na confecção das observações meteorológicas.

Como o modelo não conseguiu responder de forma satisfatória a pretensão original da previsão das faixas de visibilidade e de teto, foram definidas novas metas que são apresentadas a seguir.

#### 4.3 PROPOSTA DE NOVAS METAS

Diante dos fracos resultados para algumas faixas da idéia inicial, as novas metas de previsão são:

- a) visibilidade abaixo de 400 m;
- b) nevoeiro (visibilidade abaixo de 1000 m);
- c) teto abaixo de 30 m;
- d) teto abaixo de 60 m.

Ressalta-se aqui que essas novas metas representam um grande auxílio nas necessidades de previsão de visibilidade para o voo IFR nos modos VOR/DME e ILS (CAT I e CAT II), e de previsão de teto no modo ILS ( CAT I e CAT II), conforme a Tabela 3.10.

#### 4.4 RESULTADOS COM A RLM

Analisando a correlação entre as variáveis, foi feita opção de usar para a RLM as quatorze seguintes variáveis independente:

1. *DVISIB\_24h*
2. *DUR\_24h*
3. *DVISIB\_1h*
4. *DTEMPAR\_1h*
5. *DQNH\_1h*
6. *DIRVENTO*
7. *VELVENTO*
8. *VISIB*
9. *Q1*
10. *TEMPAR*
11. *QNH*
12. *VISIB*

13. Diajul\_m

14. Hora\_m

As variáveis dependentes são a visibilidade nas três horas futuras: VISIB1, VISIB2 E VISIB3.

Foram ajustadas vinte e quatro funções de regressão linear múltipla, concernentes aos oito conjuntos de treinamento (TESTE00, TESTE01, TESTE02, TESTE03, TESTE04, TESTE05, TESTE06, TESTE07), para cada uma das três variáveis dependentes (visibilidade nas três horas futuras). As estatísticas desses ajustes estão apresentadas no Apêndice C.

Após os ajustes, cada uma dessas funções é aplicada ao mesmo conjunto de validação de 19.441 observações meteorológicas relativas ao período de janeiro de 1951 a julho de 2005. Os resultados são então avaliados por meio de uma tabela de contingência 2x2, onde o *evento* é a previsão da visibilidade “< 1000 m” e o *não-evento* é a previsão da visibilidade “≥ 1000 m”. Os resultados para cada uma das três horas futuras, relativos às variáveis VISIB1, VISIB2 e VISIB3, são apresentados a seguir, da seguinte forma:

- primeiro, em um quadro contendo as tabelas de contingência 2x2 com os resultados da aplicação dos oito modelos de RLM sobre o conjunto de validação;
- segundo, em um gráfico com as medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de RLM sobre o conjunto de validação.

RESULTADOS PARA A PRIMEIRA HORA

Teste 00		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	341	117	458
	V>=1000	429	18554	18983
	Total	770	18671	19441
	verd-pos.	0,4429	0,9937	
	falso-pos.	0,0063	0,5571	
	verd-neg.	0,9937	0,4429	
	falso-neg.	0,5571	0,0063	
	sensibilidade	44,29%	99,37%	
	especificidad	99,37%	44,29%	

Teste 01		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	404	36	440
	V>=1000	366	18635	19001
	Total	770	18671	19441
	verd-pos.	0,5247	0,9981	
	falso-pos.	0,0019	0,4753	
	verd-neg.	0,9981	0,5247	
	falso-neg.	0,4753	0,0019	
	sensibilidade	52,47%	99,81%	
	especificidad	99,81%	52,47%	

Teste 02		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	480	106	586
	V>=1000	290	18565	18855
	Total	770	18671	19441
	verd-pos.	0,6234	0,9943	
	falso-pos.	0,0057	0,3766	
	verd-neg.	0,9943	0,6234	
	falso-neg.	0,3766	0,0057	
	sensibilidade	62,34%	99,43%	
	especificidad	99,43%	62,34%	

Teste 03		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	538	142	680
	V>=1000	232	18529	18761
	Total	770	18671	19441
	verd-pos.	0,6987	0,9924	
	falso-pos.	0,0076	0,3013	
	verd-neg.	0,9924	0,6987	
	falso-neg.	0,3013	0,0076	
	sensibilidade	69,87%	99,24%	
	especificidad	99,24%	69,87%	

Teste 04		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	553	151	704
	V>=1000	217	18520	18737
	Total	770	18671	19441
	verd-pos.	0,7182	0,9919	
	falso-pos.	0,0081	0,2818	
	verd-neg.	0,9919	0,7182	
	falso-neg.	0,2818	0,0081	
	sensibilidade	71,82%	99,19%	
	especificidad	99,19%	71,82%	

Teste 05		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	554	151	705
	V>=1000	216	18520	18736
	Total	770	18671	19441
	verd-pos.	0,7195	0,9919	
	falso-pos.	0,0081	0,2805	
	verd-neg.	0,9919	0,7195	
	falso-neg.	0,2805	0,0081	
	sensibilidade	71,95%	99,19%	
	especificidad	99,19%	71,95%	

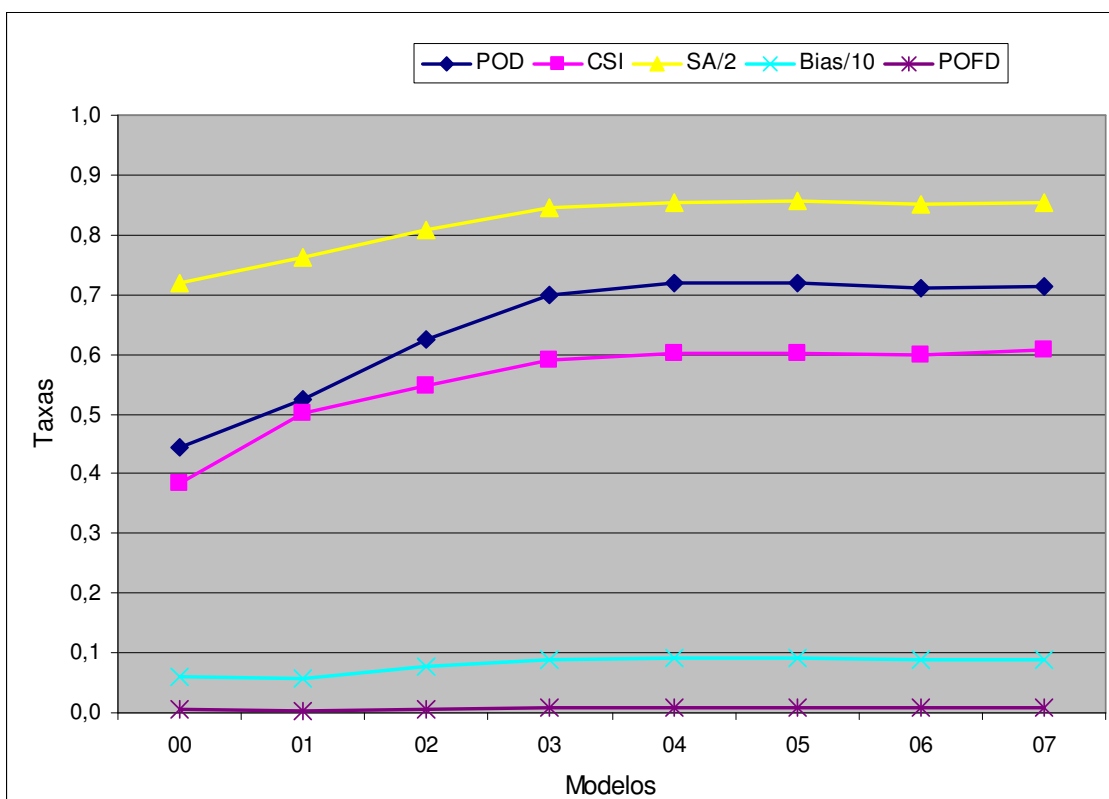
  

Teste 06		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	548	143	691
	V>=1000	222	18528	18750
	Total	770	18671	19441
	verd-pos.	0,7117	0,9923	
	falso-pos.	0,0077	0,2883	
	verd-neg.	0,9923	0,7117	
	falso-neg.	0,2883	0,0077	
	sensibilidade	71,17%	99,23%	
	especificidad	99,23%	71,17%	

Teste 07		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	549	135	684
	V>=1000	221	18536	18757
	Total	770	18671	19441
	verd-pos.	0,7130	0,9928	
	falso-pos.	0,0072	0,2870	
	verd-neg.	0,9928	0,7130	
	falso-neg.	0,2870	0,0072	
	sensibilidade	71,30%	99,28%	
	especificidad	99,28%	71,30%	

**Figura 4.2** – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.441 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na primeira hora futura (VISIB1) em “< 1000 m” ou “≥ 1000 m”.



**Figura 4.3** – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e POFD sobre os resultados da aplicação de oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.441 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na primeira hora futura (VISIB1) em “< 1000 m” ou “≥ 1000 m”.

#### RESULTADOS PARA A SEGUNDA HORA



<b>Teste 00</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	203	342	545
	V>=1000	850	28797	29647
	Total	1053	29139	30192
	verd-pos.	0,1928	0,9883	
	falso-pos.	0,0117	0,8072	
	verd-neg.	0,9883	0,1928	
	falso-neg.	0,8072	0,0117	
	sensibilidade:	19,28%	98,83%	
	especificidade:	98,83%	19,28%	

<b>Teste 01</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	419	545	964
	V>=1000	634	28594	29228
	Total	1053	29139	30192
	verd-pos.	0,3979	0,9813	
	falso-pos.	0,0187	0,6021	
	verd-neg.	0,9813	0,3979	
	falso-neg.	0,6021	0,0187	
	sensibilidade:	39,79%	98,13%	
	especificidade:	98,13%	39,79%	

<b>Teste 02</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	534	675	1209
	V>=1000	519	28464	28983
	Total	1053	29139	30192
	verd-pos.	0,5071	0,9768	
	falso-pos.	0,0232	0,4929	
	verd-neg.	0,9768	0,5071	
	falso-neg.	0,4929	0,0232	
	sensibilidade:	50,71%	97,68%	
	especificidade:	97,68%	50,71%	

<b>Teste 03</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	672	903	1575
	V>=1000	381	28236	28617
	Total	1053	29139	30192
	verd-pos.	0,6382	0,9690	
	falso-pos.	0,0310	0,3618	
	verd-neg.	0,9690	0,6382	
	falso-neg.	0,3618	0,0310	
	sensibilidade:	63,82%	96,90%	
	especificidade:	96,90%	63,82%	

<b>Teste 04</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	680	1180	1860
	V>=1000	373	27959	28332
	Total	1053	29139	30192
	verd-pos.	0,6458	0,9595	
	falso-pos.	0,0405	0,3542	
	verd-neg.	0,9595	0,6458	
	falso-neg.	0,3542	0,0405	
	sensibilidade:	64,58%	95,95%	
	especificidade:	95,95%	64,58%	

<b>Teste 05</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	681	1446	2127
	V>=1000	372	27693	28065
	Total	1053	29139	30192
	verd-pos.	0,6467	0,9504	
	falso-pos.	0,0496	0,3533	
	verd-neg.	0,9504	0,6467	
	falso-neg.	0,3533	0,0496	
	sensibilidade:	64,67%	95,04%	
	especificidade:	95,04%	64,67%	

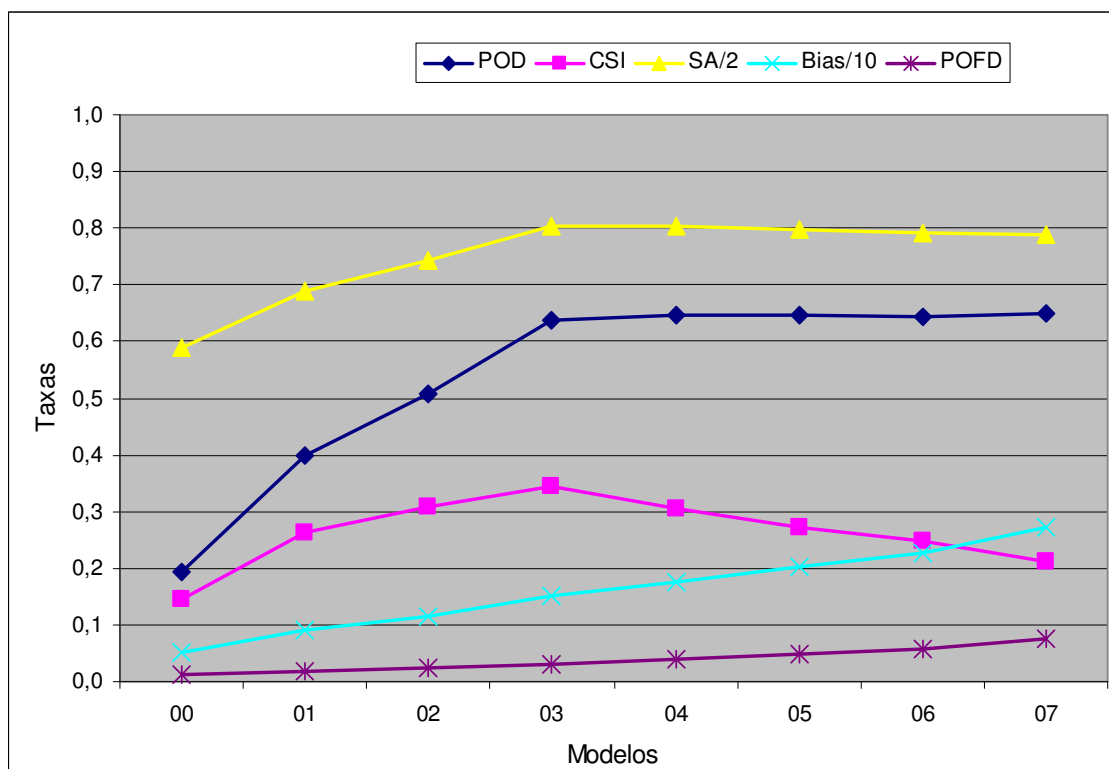
  

<b>Teste 06</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	678	1693	2371
	V>=1000	375	27446	27821
	Total	1053	29139	30192
	verd-pos.	0,6439	0,9419	
	falso-pos.	0,0581	0,3561	
	verd-neg.	0,9419	0,6439	
	falso-neg.	0,3561	0,0581	
	sensibilidade:	64,39%	94,19%	
	especificidade:	94,19%	64,39%	

<b>Teste 07</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	684	2176	2860
	V>=1000	369	26963	27332
	Total	1053	29139	30192
	verd-pos.	0,6496	0,9253	
	falso-pos.	0,0747	0,3504	
	verd-neg.	0,9253	0,6496	
	falso-neg.	0,3504	0,0747	
	sensibilidade:	64,96%	92,53%	
	especificidade:	92,53%	64,96%	

**Figura 4.4**– Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 30.192 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na segunda hora futura (VISIB2) em “< 1000 m” ou “≥ 1000 m”.



**Figura 4.5** – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e POFD sobre os resultados da aplicação de oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 30.192 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na segunda hora futura (VISIB2) em “< 1000 m” ou “≥ 1000 m”.

#### RESULTADOS PARA A TERCEIRA HORA

<b>Teste 00</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	71	15	86
	V>=1000	982	40710	41692
Total		1053	40725	41778
verd-pos.		0,0674	0,9996	
falso-pos.		0,0004	0,9326	
verd-neg.		0,9996	0,0674	
falso-neg.		0,9326	0,0004	
sensibilidade:		6,74%	99,96%	
especificidade:		99,96%	6,74%	

<b>Teste 01</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	141	78	219
	V>=1000	912	40647	41559
Total		1053	40725	41778
verd-pos.		0,1339	0,9981	
falso-pos.		0,0019	0,8661	
verd-neg.		0,9981	0,1339	
falso-neg.		0,8661	0,0019	
sensibilidade:		13,39%	99,81%	
especificidade:		99,81%	13,39%	

<b>Teste 02</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	262	250	512
	V>=1000	791	40475	41266
Total		1053	40725	41778
verd-pos.		0,2488	0,9939	
falso-pos.		0,0061	0,7512	
verd-neg.		0,9939	0,2488	
falso-neg.		0,7512	0,0061	
sensibilidade:		24,88%	99,39%	
especificidade:		99,39%	24,88%	

<b>Teste 03</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	427	675	1102
	V>=1000	626	40050	40676
Total		1053	40725	41778
verd-pos.		0,4055	0,9834	
falso-pos.		0,0166	0,5945	
verd-neg.		0,9834	0,4055	
falso-neg.		0,5945	0,0166	
sensibilidade:		40,55%	98,34%	
especificidade:		98,34%	40,55%	

<b>Teste 04</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	459	851	1310
	V>=1000	594	39874	40468
Total		1053	40725	41778
verd-pos.		0,4359	0,9791	
falso-pos.		0,0209	0,5641	
verd-neg.		0,9791	0,4359	
falso-neg.		0,5641	0,0209	
sensibilidade:		43,59%	97,91%	
especificidade:		97,91%	43,59%	

<b>Teste 05</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	465	865	1330
	V>=1000	588	39860	40448
Total		1053	40725	41778
verd-pos.		0,4416	0,9788	
falso-pos.		0,0212	0,5584	
verd-neg.		0,9788	0,4416	
falso-neg.		0,5584	0,0212	
sensibilidade:		44,16%	97,88%	
especificidade:		97,88%	44,16%	

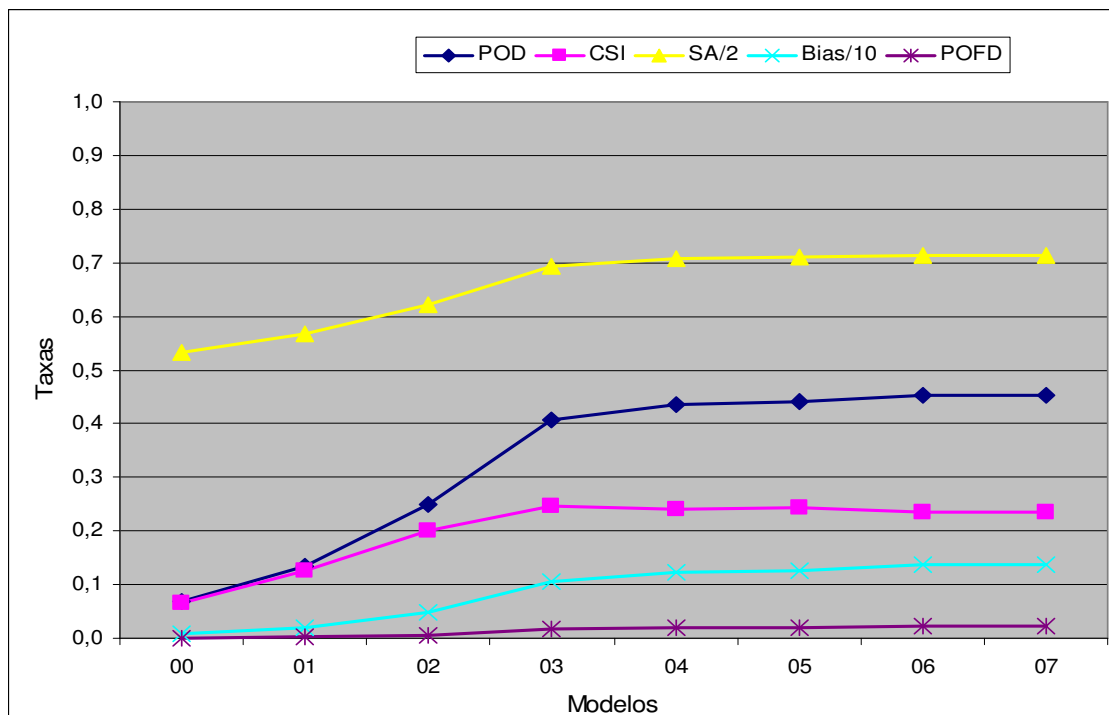
<b>Teste 06</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	476	967	1443
	V>=1000	577	39758	40335
Total		1053	40725	41778
verd-pos.		0,4520	0,9763	
falso-pos.		0,0237	0,5480	
verd-neg.		0,9763	0,4520	
falso-neg.		0,5480	0,0237	
sensibilidade:		45,20%	97,63%	
especificidade:		97,63%	45,20%	

<b>Teste 07</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	477	978	1455
	V>=1000	576	39747	40323
Total		1053	40725	41778
verd-pos.		0,4530	0,9760	
falso-pos.		0,0240	0,5470	
verd-neg.		0,9760	0,4530	
falso-neg.		0,5470	0,0240	
sensibilidade:		45,30%	97,60%	
especificidade:		97,60%	45,30%	

**Figura 4.6** – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.778 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na terceira hora futura (VISIB3) em “< 1000 m” ou “≥ 1000 m”.





**Figura 4.7** – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e POFD sobre os resultados da aplicação de oito modelos de RLM (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.778 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na terceira hora futura (VISIB3) em “< 1000 m” ou “≥ 1000 m”.

#### 4.5 RESULTADOS COM REDE GRNN

Para estimativa das visibilidades, para as três horas futuras, foram utilizadas as redes GRNN treinadas com o programa *NeuroShell Predictor* por meio do algoritmo *Cascade Correlation (CC)*. As redes treinadas com os oito conjuntos de treinamentos TESTE00, TESTE01, TESTE02, TESTE03, TESTE04, TESTE05, TESTE06, TESTE07, atingiram as seguintes configurações de neurônios nas camadas ocultas, respectivamente:

- para a primeira hora: 70, 24, 72, 140, 2, 2, 2 e 2;
- para a segunda hora: 17, 108, 26, 141, 122, 19, 26 e 53;
- para a terceira hora: 148, 95, 148, 141, 141, 31, 7 e 139.

Após os treinamentos, cada uma dessas redes é aplicada ao mesmo conjunto de validação de 19.694 observações meteorológicas relativas ao período de janeiro de 1951 a julho de 2005. Os resultados são avaliados por meio de uma tabela de contingência 2x2, onde o *evento* é a previsão da visibilidade “< 1000 m” e o *não-evento* é a previsão da visibilidade “≥ 1000 m”. Os resultados para cada uma das três horas futuras, relativos às variáveis VISIB1, VISIB2 e VISIB3, são apresentados a seguir, da seguinte forma:

- primeiro, em um quadro contendo as tabelas de contingência 2x2 com os resultados da aplicação das oito redes neurais GRNN treinadas sobre o conjunto de validação;
- segundo, em um gráfico com as medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos GRNN sobre o conjunto de validação.

## RESULTADOS PARA A PRIMEIRA HORA

<b>Teste 00</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	314	27	341
	V>=1000	461	18892	19353
	Total	775	18919	19694
	verd-pos.	0,4052	0,9986	
	falso-pos.	0,0014	0,5948	
	verd-neg.	0,9986	0,4052	
	falso-neg.	0,5948	0,0014	
	sensibilidade:	40,52%	99,86%	
	especificidad:	99,86%	40,52%	

<b>Teste 01</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	464	64	528
	V>=1000	311	18855	19166
	Total	775	18919	19694
	verd-pos.	0,5987	0,9966	
	falso-pos.	0,0034	0,4013	
	verd-neg.	0,9966	0,5987	
	falso-neg.	0,4013	0,0034	
	sensibilidade:	59,87%	99,66%	
	especificidad:	99,66%	59,87%	

<b>Teste 02</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	544	82	626
	V>=1000	231	18837	19068
	Total	775	18919	19694
	verd-pos.	0,7019	0,9957	
	falso-pos.	0,0043	0,2981	
	verd-neg.	0,9957	0,7019	
	falso-neg.	0,2981	0,0043	
	sensibilidade:	70,19%	99,57%	
	especificidad:	99,57%	70,19%	

<b>Teste 03</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	580	102	682
	V>=1000	195	18817	19012
	Total	775	18919	19694
	verd-pos.	0,7484	0,9946	
	falso-pos.	0,0054	0,2516	
	verd-neg.	0,9946	0,7484	
	falso-neg.	0,2516	0,0054	
	sensibilidade:	74,84%	99,46%	
	especificidad:	99,46%	74,84%	

<b>Teste 04</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	585	114	699
	V>=1000	190	18805	18995
	Total	775	18919	19694
	verd-pos.	0,7548	0,9940	
	falso-pos.	0,0060	0,2452	
	verd-neg.	0,9940	0,7548	
	falso-neg.	0,2452	0,0060	
	sensibilidade:	75,48%	99,40%	
	especificidad:	99,40%	75,48%	

<b>Teste 05</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	588	112	700
	V>=1000	187	18807	18994
	Total	775	18919	19694
	verd-pos.	0,7587	0,9941	
	falso-pos.	0,0059	0,2413	
	verd-neg.	0,9941	0,7587	
	falso-neg.	0,2413	0,0059	
	sensibilidade:	75,87%	99,41%	
	especificidad:	99,41%	75,87%	

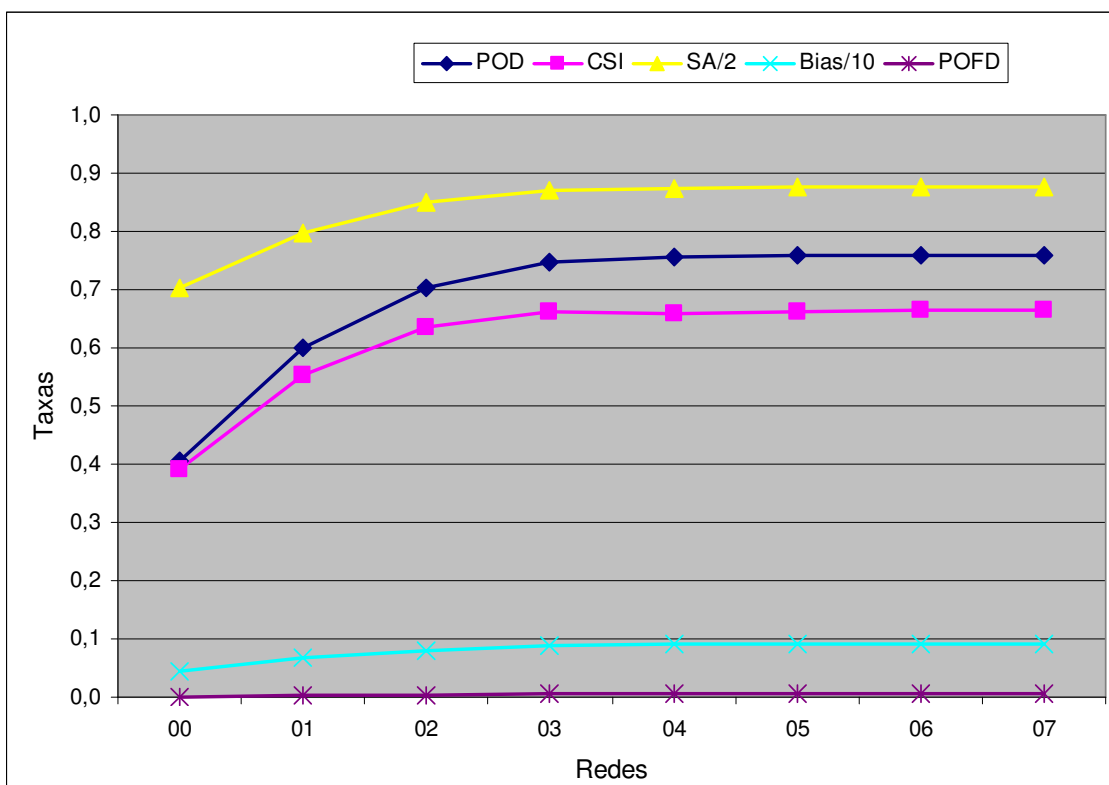
  

<b>Teste 06</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	588	108	696
	V>=1000	187	18811	18998
	Total	775	18919	19694
	verd-pos.	0,7587	0,9943	
	falso-pos.	0,0057	0,2413	
	verd-neg.	0,9943	0,7587	
	falso-neg.	0,2413	0,0057	
	sensibilidade:	75,87%	99,43%	
	especificidad:	99,43%	75,87%	

<b>Teste 07</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	588	108	696
	V>=1000	187	18811	18998
	Total	775	18919	19694
	verd-pos.	0,7587	0,9943	
	falso-pos.	0,0057	0,2413	
	verd-neg.	0,9943	0,7587	
	falso-neg.	0,2413	0,0057	
	sensibilidade:	75,87%	99,43%	
	especificidad:	99,43%	75,87%	

**Figura 4.8** – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de GRNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.694 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na primeira hora futura (VISIB1) em “< 1000 m” ou “≥ 1000 m”.



**Figura 4.9** – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de GRNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.694 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na primeira hora futura (VISIB1) em “< 1000 m” ou “≥ 1000 m”.

#### RESULTADOS PARA A SEGUNDA HORA



<b>Teste 00</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	264	1203	1467
	V>=1000	789	28989	29778
Total		1053	30192	31245
verd-pos.		0,2507	0,9602	
falso-pos.		0,0398	0,7493	
verd-neg.		0,9602	0,2507	
falso-neg.		0,7493	0,0398	
sensibilidade		25,07%	96,02%	
especificidade		96,02%	25,07%	

<b>Teste 01</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	485	1388	1873
	V>=1000	568	28804	29372
Total		1053	30192	31245
verd-pos.		0,4606	0,9540	
falso-pos.		0,0460	0,5394	
verd-neg.		0,9540	0,4606	
falso-neg.		0,5394	0,0460	
sensibilidade		46,06%	95,40%	
especificidade		95,40%	46,06%	

<b>Teste 02</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	578	1506	2084
	V>=1000	475	28686	29161
Total		1053	30192	31245
verd-pos.		0,5489	0,9501	
falso-pos.		0,0499	0,4511	
verd-neg.		0,9501	0,5489	
falso-neg.		0,4511	0,0499	
sensibilidade		54,89%	95,01%	
especificidade		95,01%	54,89%	

<b>Teste 03</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	667	1694	2361
	V>=1000	386	28498	28884
Total		1053	30192	31245
verd-pos.		0,6334	0,9439	
falso-pos.		0,0561	0,3666	
verd-neg.		0,9439	0,6334	
falso-neg.		0,3666	0,0561	
sensibilidade		63,34%	94,39%	
especificidade		94,39%	63,34%	

<b>Teste 04</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	690	1752	2442
	V>=1000	363	28440	28803
Total		1053	30192	31245
verd-pos.		0,6553	0,9420	
falso-pos.		0,0580	0,3447	
verd-neg.		0,9420	0,6553	
falso-neg.		0,3447	0,0580	
sensibilidade		65,53%	94,20%	
especificidade		94,20%	65,53%	

<b>Teste 05</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	691	1742	2433
	V>=1000	362	28450	28812
Total		1053	30192	31245
verd-pos.		0,6562	0,9423	
falso-pos.		0,0577	0,3438	
verd-neg.		0,9423	0,6562	
falso-neg.		0,3438	0,0577	
sensibilidade		65,62%	94,23%	
especificidade		94,23%	65,62%	

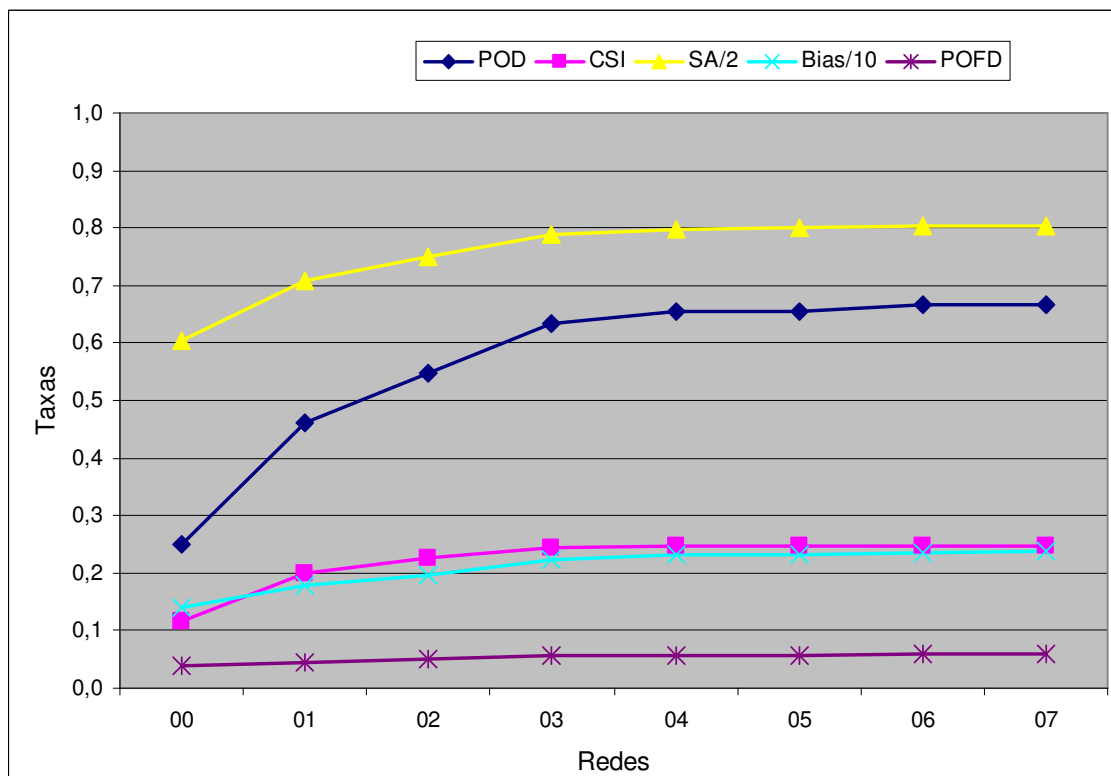
  

<b>Teste 06</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	701	1786	2487
	V>=1000	352	28406	28758
Total		1053	30192	31245
verd-pos.		0,6657	0,9408	
falso-pos.		0,0592	0,3343	
verd-neg.		0,9408	0,6657	
falso-neg.		0,3343	0,0592	
sensibilidade		66,57%	94,08%	
especificidade		94,08%	66,57%	

<b>Teste 07</b>		Observado		
		V<1000	V>=1000	Total
Previsto	V<1000	702	1798	2500
	V>=1000	351	28394	28745
Total		1053	30192	31245
verd-pos.		0,6667	0,9404	
falso-pos.		0,0596	0,3333	
verd-neg.		0,9404	0,6667	
falso-neg.		0,3333	0,0596	
sensibilidade		66,67%	94,04%	
especificidade		94,04%	66,67%	

**Figura 4.10** – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de GRNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 31.245 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na segunda hora futura (VISIB2) em “< 1000 m” ou “≥ 1000 m”.



**Figura 4. 11** – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos de GRNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 31.245 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na segunda hora futura (VISIB2) em “< 1000 m” ou “≥ 1000 m”.

#### RESULTADOS PARA A TERCEIRA HORA

Teste 00		Observado		Total
		V<1000	V>=1000	
Previsto	V<1000	65	32	97
	V>=1000	988	40823	41811
Total		1053	40855	41908
verd-pos.		0,0617	0,9992	
falso-pos.		0,0008	0,9383	
verd-neg.		0,9992	0,0617	
falso-neg.		0,9383	0,0008	
sensibilidade		6,17%	99,92%	
especificidade		99,92%	6,17%	

Teste 01		Observado		Total
		V<1000	V>=1000	
Previsto	V<1000	242	156	398
	V>=1000	811	40699	41510
Total		1053	40855	41908
verd-pos.		0,2298	0,9962	
falso-pos.		0,0038	0,7702	
verd-neg.		0,9962	0,2298	
falso-neg.		0,7702	0,0038	
sensibilidade		22,98%	99,62%	
especificidade		99,62%	22,98%	

Teste 02		Observado		Total
		V<1000	V>=1000	
Previsto	V<1000	350	331	681
	V>=1000	703	40524	41227
Total		1053	40855	41908
verd-pos.		0,3324	0,9919	
falso-pos.		0,0081	0,6676	
verd-neg.		0,9919	0,3324	
falso-neg.		0,6676	0,0081	
sensibilidade		33,24%	99,19%	
especificidade		99,19%	33,24%	

Teste 03		Observado		Total
		V<1000	V>=1000	
Previsto	V<1000	470	765	1235
	V>=1000	583	40090	40673
Total		1053	40855	41908
verd-pos.		0,4463	0,9813	
falso-pos.		0,0187	0,5537	
verd-neg.		0,9813	0,4463	
falso-neg.		0,5537	0,0187	
sensibilidade		44,63%	98,13%	
especificidade		98,13%	44,63%	

Teste 04		Observado		Total
		V<1000	V>=1000	
Previsto	V<1000	477	934	1411
	V>=1000	576	39921	40497
Total		1053	40855	41908
verd-pos.		0,4530	0,9771	
falso-pos.		0,0229	0,5470	
verd-neg.		0,9771	0,4530	
falso-neg.		0,5470	0,0229	
sensibilidade		45,30%	97,71%	
especificidade		97,71%	45,30%	

Teste 05		Observado		Total
		V<1000	V>=1000	
Previsto	V<1000	482	949	1431
	V>=1000	571	39906	40477
Total		1053	40855	41908
verd-pos.		0,4577	0,9768	
falso-pos.		0,0232	0,5423	
verd-neg.		0,9768	0,4577	
falso-neg.		0,5423	0,0232	
sensibilidade		45,77%	97,68%	
especificidade		97,68%	45,77%	

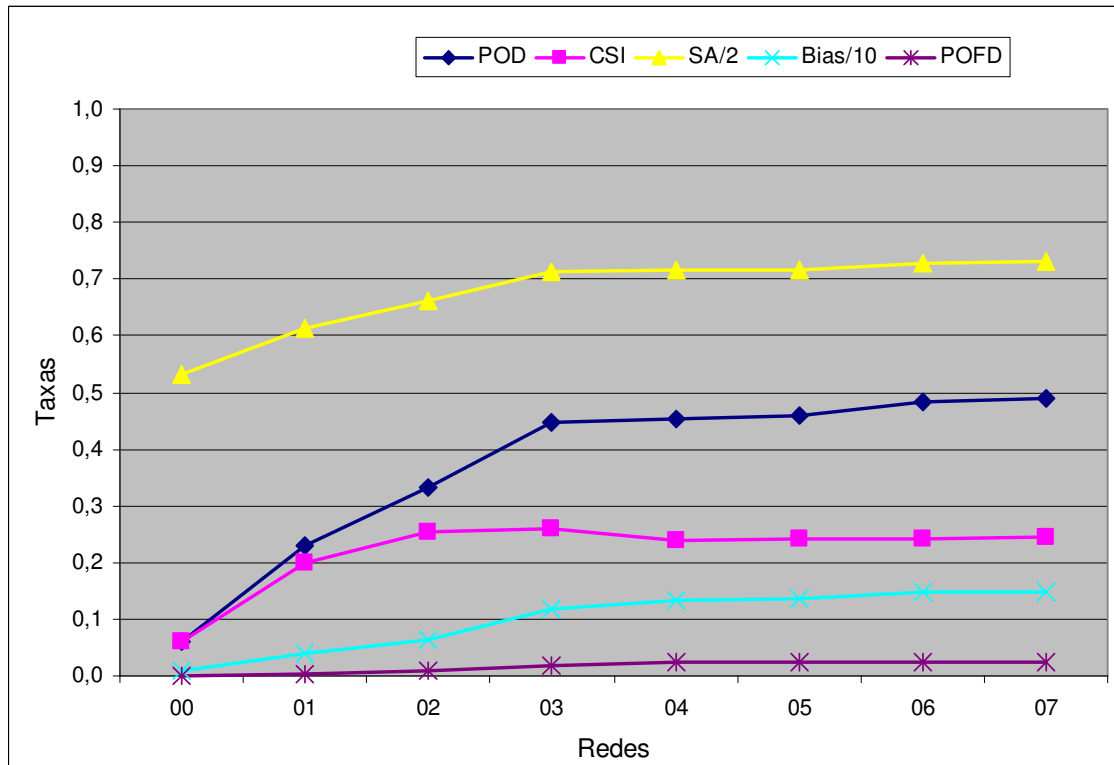
Teste 06		Observado		Total
		V<1000	V>=1000	
Previsto	V<1000	509	1043	1552
	V>=1000	544	39812	40356
Total		1053	40855	41908
verd-pos.		0,4834	0,9745	
falso-pos.		0,0255	0,5166	
verd-neg.		0,9745	0,4834	
falso-neg.		0,5166	0,0255	
sensibilidade		48,34%	97,45%	
especificidade		97,45%	48,34%	

Teste 07		Observado		Total
		V<1000	V>=1000	
Previsto	V<1000	515	1043	1558
	V>=1000	538	39812	40350
Total		1053	40855	41908
verd-pos.		0,4891	0,9745	
falso-pos.		0,0255	0,5109	
verd-neg.		0,9745	0,4891	
falso-neg.		0,5109	0,0255	
sensibilidade		48,91%	97,45%	
especificidade		97,45%	48,91%	

**Figura 4.12** – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de GRNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.908 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na terceira hora futura (VISIB3) em “< 1000 m” ou “≥ 1000 m”.





**Figura 4.13** – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e POFD sobre os resultados da aplicação de oito modelos de GRNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.908 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação da visibilidade na terceira hora futura (VISIB3) em “< 1000 m” ou “≥ 1000 m”.

#### 4.6 RESULTADOS COM REDE PNN

Para a classificação de nevoeiro, para as três horas futuras, foram utilizadas as redes PNN treinadas com o programa *NeuroShell Classifier* por meio do algoritmo *Cascade Correlation (CC)*. As redes treinadas com os oitos conjuntos de treinamentos TESTE00, TESTE01, TESTE02, TESTE03, TESTE04, TESTE05, TESTE06, TESTE07, atingiram as seguintes configurações de neurônios nas camadas ocultas, respectivamente:

- para a primeira hora: 47, 135, 35, 143, 90, 128, 127 e 125;
- para a segunda hora: 146, 150, 150, 111, 136, 130, 121 e 150;
- para a terceira hora: 57, 104, 146, 150, 146, 116, 148 e 23.



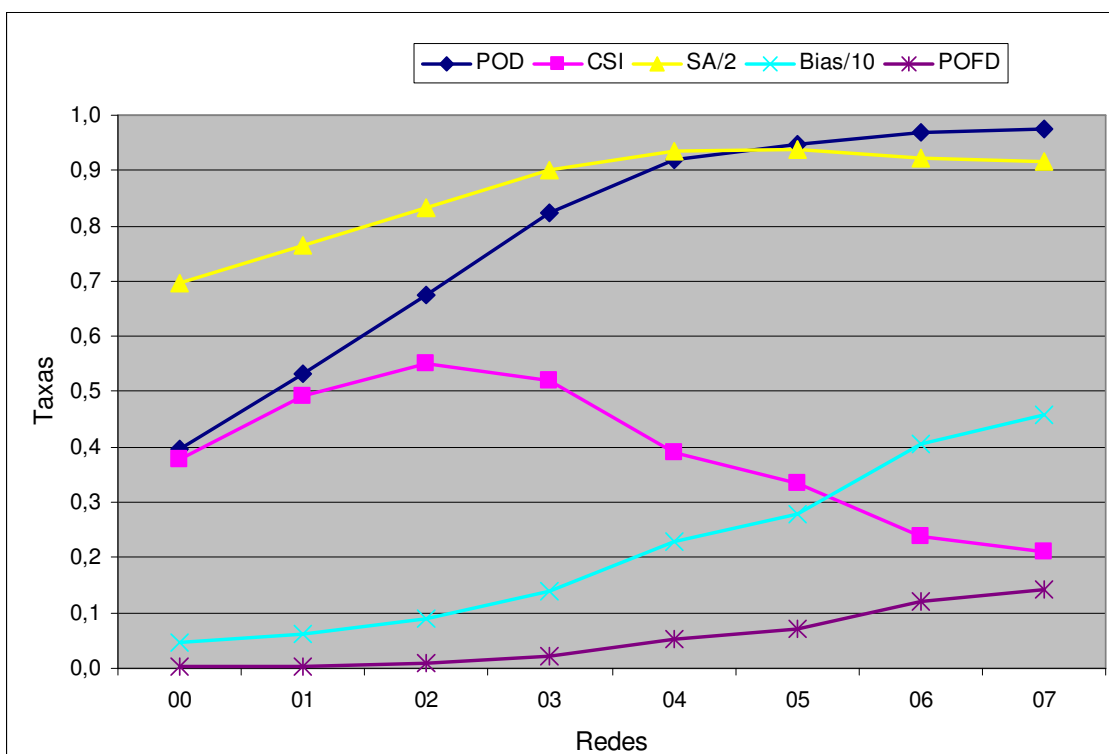
Após os treinamentos, cada uma dessas redes é aplicada ao mesmo conjunto de validação de 19.693 observações meteorológicas relativas ao período de janeiro de 1951 a julho de 2005. Os resultados são avaliados por meio de uma tabela de contingência 2x2, onde o *evento* é a previsão de “nevoeiro” e o *não-evento* é a previsão de “não-nevoeiro”. Os resultados para cada uma das três horas futuras são apresentados a seguir, da seguinte forma:

- primeiro, em um quadro contendo as tabelas de contingência 2x2 com os resultados da aplicação das oito redes neurais PNN treinadas sobre o conjunto de validação;
- segundo, em um gráfico com as medidas de avaliação POD, CSI, SA/2, Bias/10 e PFOD sobre os resultados da aplicação de oito modelos PNN sobre o conjunto de validação.

## RESULTADOS PARA A PRIMEIRA HORA - NEVOEIRO

Rede 00		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	18909	449	19358
	"SIM"	39	296	335
	Total	18948	745	19693
	verd-pos.	0,9979	0,3973	
	falso-pos.	0,6027	0,0021	
	verd-neg.	0,3973	0,9979	
	falso-neg.	0,0021	0,6027	
	sensibilidade	99,79%	39,73%	
	especificidade	39,73%	99,79%	
Rede 01		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	18889	349	19238
	"SIM"	59	396	455
	Total	18948	745	19693
	verd-pos.	0,9969	0,5315	
	falso-pos.	0,4685	0,0031	
	verd-neg.	0,5315	0,9969	
	falso-neg.	0,0031	0,4685	
	sensibilidade	99,69%	53,15%	
	especificidade	53,15%	99,69%	
Rede 02		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	18782	243	19025
	"SIM"	166	502	668
	Total	18948	745	19693
	verd-pos.	0,9912	0,6738	
	falso-pos.	0,3262	0,0088	
	verd-neg.	0,6738	0,9912	
	falso-neg.	0,0088	0,3262	
	sensibilidade	99,12%	67,38%	
	especificidade	67,38%	99,12%	
Rede 03		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	18514	132	18646
	"SIM"	434	613	1047
	Total	18948	745	19693
	verd-pos.	0,9771	0,8228	
	falso-pos.	0,1772	0,0229	
	verd-neg.	0,8228	0,9771	
	falso-neg.	0,0229	0,1772	
	sensibilidade	97,71%	82,28%	
	especificidade	82,28%	97,71%	
Rede 04		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	17929	59	17988
	"SIM"	1019	686	1705
	Total	18948	745	19693
	verd-pos.	0,9462	0,9208	
	falso-pos.	0,0792	0,0538	
	verd-neg.	0,9208	0,9462	
	falso-neg.	0,0538	0,0792	
	sensibilidade	94,62%	92,08%	
	especificidade	92,08%	94,62%	
Rede 05		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	17583	40	17623
	"SIM"	1365	705	2070
	Total	18948	745	19693
	verd-pos.	0,928	0,9463	
	falso-pos.	0,0537	0,072	
	verd-neg.	0,9463	0,928	
	falso-neg.	0,072	0,0537	
	sensibilidade	92,80%	94,63%	
	especificidade	94,63%	92,80%	
Rede 06		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	16657	24	16681
	"SIM"	2291	721	3012
	Total	18948	745	19693
	verd-pos.	0,8791	0,9678	
	falso-pos.	0,0322	0,1209	
	verd-neg.	0,9678	0,8791	
	falso-neg.	0,1209	0,0322	
	sensibilidade	87,91%	96,78%	
	especificidade	96,78%	87,91%	
Rede 07		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	16259	18	16277
	"SIM"	2689	727	3416
	Total	18948	745	19693
	verd-pos.	0,8581	0,9758	
	falso-pos.	0,0242	0,1419	
	verd-neg.	0,9758	0,8581	
	falso-neg.	0,1419	0,0242	
	sensibilidade	85,81%	97,58%	
	especificidade	97,58%	85,81%	

**Figura 4.14** – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de PNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.693 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de nevoeiro (SIM) e não-nevoeiro (NÃO) na primeira hora futura.

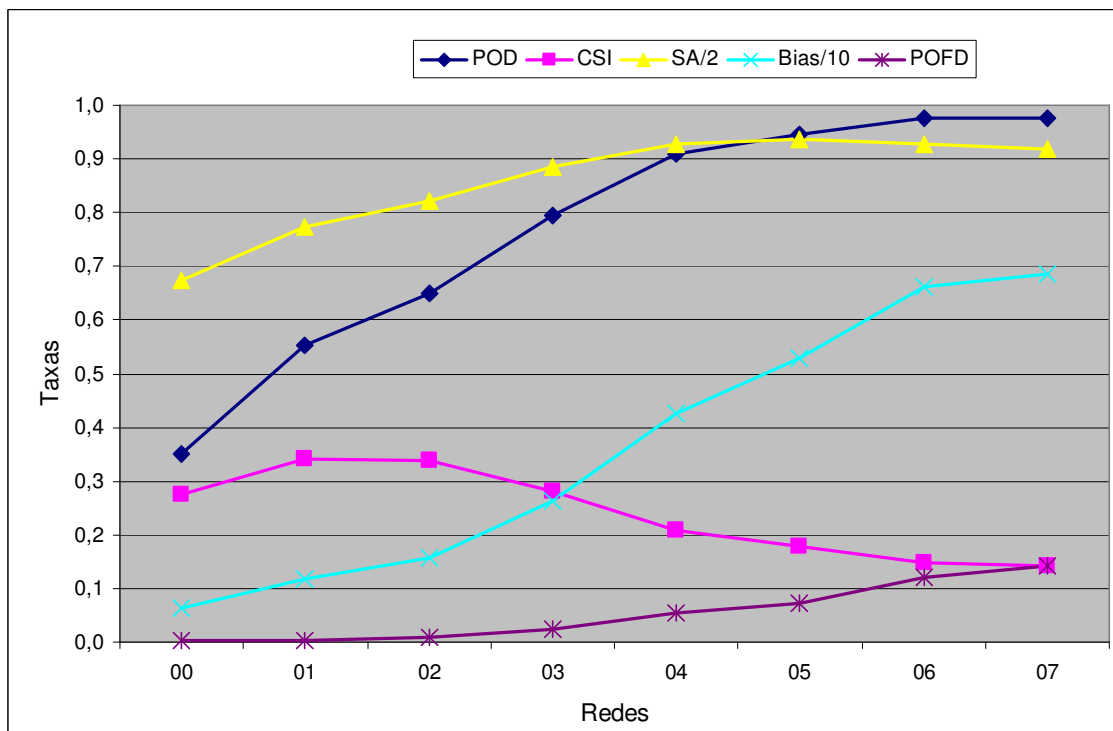


**Figura 4.15** – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e POFD sobre os resultados da aplicação de oito modelos de PNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 19.693 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de nevoeiro (SIM) e não-nevoeiro (NÃO) na primeira hora futura.

#### RESULTADOS PARA A SEGUNDA HORA – NEVOEIRO

Rede 00		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	28852	685	29537
	"SIM"	287	368	655
	Total	29139	1053	30192
	verd-pos.	0,9902	0,3495	
	falso-pos.	0,6505	0,0098	
	verd-neg.	0,3495	0,9902	
	falso-neg.	0,0098	0,6505	
	sensibilidade	99,02%	34,95%	
	especificidade	34,95%	99,02%	
Rede 01		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	28484	472	28956
	"SIM"	655	581	1236
	Total	29139	1053	30192
	verd-pos.	0,9775	0,5518	
	falso-pos.	0,4482	0,0225	
	verd-neg.	0,5518	0,9775	
	falso-neg.	0,0225	0,4482	
	sensibilidade	97,75%	55,18%	
	especificidade	55,18%	97,75%	
Rede 02		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	28177	369	28546
	"SIM"	962	684	1646
	Total	29139	1053	30192
	verd-pos.	0,967	0,6496	
	falso-pos.	0,3504	0,033	
	verd-neg.	0,6496	0,967	
	falso-neg.	0,033	0,3504	
	sensibilidade	96,70%	64,96%	
	especificidade	64,96%	96,70%	
Rede 03		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	27222	217	27439
	"SIM"	1917	836	2753
	Total	29139	1053	30192
	verd-pos.	0,9342	0,7939	
	falso-pos.	0,2061	0,0658	
	verd-neg.	0,7939	0,9342	
	falso-neg.	0,0658	0,2061	
	sensibilidade	93,42%	79,39%	
	especificidade	79,39%	93,42%	
Rede 04		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	25618	95	25713
	"SIM"	3521	958	4479
	Total	29139	1053	30192
	verd-pos.	0,8792	0,9098	
	falso-pos.	0,0902	0,1208	
	verd-neg.	0,9098	0,8792	
	falso-neg.	0,1208	0,0902	
	sensibilidade	87,92%	90,98%	
	especificidade	90,98%	87,92%	
Rede 05		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	24565	57	24622
	"SIM"	4573	996	5569
	Total	29138	1053	30191
	verd-pos.	0,8431	0,9459	
	falso-pos.	0,0541	0,1569	
	verd-neg.	0,9459	0,8431	
	falso-neg.	0,1569	0,0541	
	sensibilidade	84,31%	94,59%	
	especificidade	94,59%	84,31%	
Rede 06		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	23187	26	23213
	"SIM"	5952	1027	6979
	Total	29139	1053	30192
	verd-pos.	0,7957	0,9753	
	falso-pos.	0,0247	0,2043	
	verd-neg.	0,9753	0,7957	
	falso-neg.	0,2043	0,0247	
	sensibilidade	79,57%	97,53%	
	especificidade	97,53%	79,57%	
Rede 07		Observado		
		"NÃO"	"SIM"	Total
Previsto	"NÃO"	22950	25	22975
	"SIM"	6189	1028	7217
	Total	29139	1053	30192
	verd-pos.	0,7876	0,9763	
	falso-pos.	0,0237	0,2124	
	verd-neg.	0,9763	0,7876	
	falso-neg.	0,2124	0,0237	
	sensibilidade	78,76%	97,63%	
	especificidade	97,63%	78,76%	

**Figura 4.16** – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de PNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 30.192 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de nevoeiro (SIM) e não-nevoeiro (NÃO) na segunda hora futura.



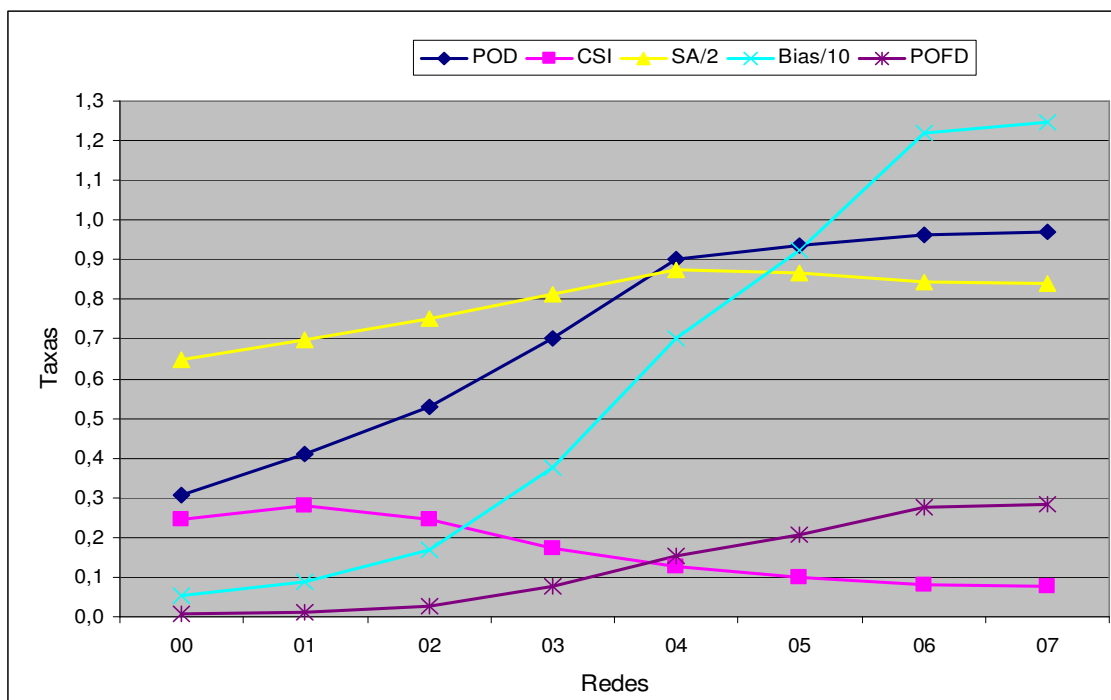
**Figura 4.17**– Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e POFD sobre os resultados da aplicação de oito modelos de PNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 30.192 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de nevoeiro (SIM) e não-nevoeiro (NÃO) na segunda hora futura.

#### RESULTADOS PARA A TERCEIRA HORA – NEVOEIRO



		Rede 00		Observado		Total
		"NÃO"	"SIM"			Total
Previsto	"NÃO"	40644	704			41348
	"SIM"	250	310			560
	Total	40894	1014			41908
verd-pos.		0,9939	0,3057			
falso-pos.		0,6943	0,0061			
verd-neg.		0,3057	0,9939			
falso-neg.		0,0061	0,6943			
sensibilidade		99,39%	30,57%			
especificidade		30,57%	99,39%			
		Rede 01		Observado		Total
		"NÃO"	"SIM"			Total
Previsto	"NÃO"	40413	598			41011
	"SIM"	481	416			897
	Total	40894	1014			41908
verd-pos.		0,9882	0,4103			
falso-pos.		0,5897	0,0118			
verd-neg.		0,4103	0,9882			
falso-neg.		0,0118	0,5897			
sensibilidade		98,82%	41,03%			
especificidade		41,03%	98,82%			
		Rede 02		Observado		Total
		"NÃO"	"SIM"			Total
Previsto	"NÃO"	39729	476			40205
	"SIM"	1165	538			1703
	Total	40894	1014			41908
verd-pos.		0,9715	0,5306			
falso-pos.		0,4694	0,0285			
verd-neg.		0,5306	0,9715			
falso-neg.		0,0285	0,4694			
sensibilidade		97,15%	53,06%			
especificidade		53,06%	97,15%			
		Rede 03		Observado		Total
		"NÃO"	"SIM"			Total
Previsto	"NÃO"	37806	303			38109
	"SIM"	3088	711			3799
	Total	40894	1014			41908
verd-pos.		0,9245	0,7012			
falso-pos.		0,2988	0,0755			
verd-neg.		0,7012	0,9245			
falso-neg.		0,0755	0,2988			
sensibilidade		92,45%	70,12%			
especificidade		70,12%	92,45%			
		Rede 04		Observado		Total
		"NÃO"	"SIM"			Total
Previsto	"NÃO"	34677	99			34776
	"SIM"	6217	915			7132
	Total	40894	1014			41908
verd-pos.		0,848	0,9024			
falso-pos.		0,0976	0,152			
verd-neg.		0,9024	0,848			
falso-neg.		0,152	0,0976			
sensibilidade		84,80%	90,24%			
especificidade		90,24%	84,80%			
		Rede 05		Observado		Total
		"NÃO"	"SIM"			Total
Previsto	"NÃO"	32480	64			32544
	"SIM"	8414	950			9364
	Total	40894	1014			41908
verd-pos.		0,7942	0,9369			
falso-pos.		0,0631	0,2058			
verd-neg.		0,9369	0,7942			
falso-neg.		0,2058	0,0631			
sensibilidade		79,42%	93,69%			
especificidade		93,69%	79,42%			
		Rede 06		Observado		Total
		"NÃO"	"SIM"			Total
Previsto	"NÃO"	29524	37			29561
	"SIM"	11370	977			12347
	Total	40894	1014			41908
verd-pos.		0,722	0,9635			
falso-pos.		0,0365	0,278			
verd-neg.		0,9635	0,722			
falso-neg.		0,278	0,0365			
sensibilidade		72,20%	96,35%			
especificidade		96,35%	72,20%			
		Rede 07		Observado		Total
		"NÃO"	"SIM"			Total
Previsto	"NÃO"	29225	32			29257
	"SIM"	11669	982			12651
	Total	40894	1014			41908
verd-pos.		0,7147	0,9684			
falso-pos.		0,0316	0,2853			
verd-neg.		0,9684	0,7147			
falso-neg.		0,2853	0,0316			
sensibilidade		71,47%	96,84%			
especificidade		96,84%	71,47%			

**Figura 4.18** – Quadro contendo as tabelas de contingência relativas aos resultados da aplicação dos oito modelos de PNN (treinadas com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.908 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de nevoeiro (SIM) e não-nevoeiro (NÃO) na terceira hora futura.



**Figura 4.19** – Representação gráfica das medidas de avaliação POD, CSI, SA/2, Bias/10 e POFD sobre os resultados da aplicação de oito modelos de PNN (ajustados com os oito conjuntos de treinamento discriminados no texto) sobre o conjunto de validação de 41.908 eventos meteorológicos relativos ao período de janeiro de 1951 a julho de 2005. Esses resultados se referem à classificação de nevoeiro (SIM) e não-nevoeiro (NÃO) na terceira hora futura.

#### 4.7 COMPARAÇÃO DOS RESULTADOS COM OS TRÊS MÉTODOS

A seguir é feita uma análise comparativa dos resultados alcançados com os três métodos RLM, GRNN e PNN. Na tabela 4.9 é apresentado um resumo dos melhores resultados das 72 opções de modelos treinados: foram três métodos com oito conjuntos de treinamentos e três variáveis de saída.

**Tabela 4.9** – Resumo dos melhores resultados das 72 opções de modelos treinados: (3 métodos) x (8 conjuntos de treinamentos) x (três variáveis de saída).

<b>1º HORA</b>						
	Redes/Ajuste	POD	CSI	SA/2	Bias	POFD
RLM	03	70%	0,59	0,85	0,88	0,8%
GRNN	03	75%	0,66	0,87	0,88	0,5%
PNN	03	82%	0,52	0,90	1,41	2,3%
PNN	04	92%	0,39	0,93	2,29	5,4%

<b>2º HORA</b>						
	Redes/Ajuste	POD	CSI	SA/2	Bias	POFD
RLM	03	64%	0,34	0,80	1,50	3,1%
GRNN	03	63%	0,24	0,79	2,24	5,6%
PNN	02	65%	0,34	0,82	1,56	0,9%
PNN	03	79%	0,28	0,89	2,61	2,3%

<b>3º HORA</b>						
	Redes/Ajuste	POD	CSI	SA/2	Bias	POFD
RLM	04	44%	0,24	0,71	1,24	2,1%
GRNN	03	45%	0,26	0,71	1,17	1,9%
PNN	02	53%	0,25	0,75	1,68	2,9%
PNN	03	70%	0,17	0,81	3,75	7,6%

Analisando os resultados, as seguintes conclusões podem ser extraídas:

- Para a primeira hora, a previsão de visibilidade menor que 1000m ou nevoeiro tem melhor desempenho com os modelos treinados com a rede PNN e com os conjuntos de treinamento 03 e 04. Se houver uma maior preocupação com o valor do *bias* e com o nível de alarme falso, das duas redes a melhor é a 03; mas, se for considerado que a percentagem de 5,4% da POFD (alarme falso) do modelo 04 não é tão relevante nessa situação, teremos, então, uma sensibilidade ou percentagem de acerto de 92%, o que é um excelente resultado. Os modelos com RLM e GRNN não apresentara diferenças muito significativas entre si, e apresentam desempenho abaixo dos modelos treinados com rede PNN.
- Para a segunda hora, VISIB2 ou nevoeiro, novamente os modelos PNN apresentam melhor desempenho que os modelo RLM e GRNN, estes dois apresentam resultados muito próximos entre si.



- Na terceira hora, observa-se que há uma queda nos desempenho de todos os modelos. Inclusive, os modelos RLM e GRNN mostram desempenho abaixo de 50%. Porém, os modelos com PNN continuam com melhor desempenho. Embora com um aumento no número de alarme falso (7,6%), a sensibilidade ou POD continua alta (70%).

Uma palavra deve se dita aqui sobre o aparente fraco desempenho dos modelos levando-se em consideração os baixos valores do índice CSI mostrados na Tabela 4.9. Foi demonstrado por Schaefer (1990) que o CSI não é uma medida imparcial da habilidade de previsão de um modelo e que, entretanto, é um indicador válido do valor relativo de diferentes técnicas de previsão quando são aplicadas ao mesmo ambiente. Gerapetritis e Pelissier (2004) dizem que o CSI não fornece informação de verificação única, pois é uma função da taxa de alarme falso e da taxa de acerto (POD).

De fato, em um conjunto de dados em que o percentual dos não-eventos (por exemplo, visibilidade alta) é muito maior que o percentual dos eventos (por exemplo, visibilidade baixa, nevoeiro), como é o presente caso (em torno de 96% e 4%, respectivamente), o valor numérico dos alarmes falsos, mesmo que percentualmente baixo, pode reduzir fortemente o CSI se esse valor for alto relativamente aos valores dos acertos e dos erros, conforme a equação a seguir.

$$CSI = \frac{\text{acertos}}{\text{acertos} + \text{alarmes falsos} + \text{erros de previsão}}$$

#### 4.8 DEFINIÇÃO DO MODELO A SER OPERACIONALIZADO.

Essa discussão do item anterior foi na direção de encontrar um modelo de previsão que apresentasse a melhor taxa de acerto juntamente com pequena taxa de alarme falso, em outras palavras alta sensibilidade e alta especificidade. Foi visto que de um modo geral, os modelos PNN apresentaram melhor desempenho. E, neste caso, os melhores resultados foram aqueles mostrados na Tabela 4.9. Mas em

termos práticos, do ponto de vista operacional, onde um determinado nível de alarme falso seja aceitável, ou seja, poder-se-ia então ser usado outro modelo que tenha uma maior sensibilidade e uma especificidade um pouco menor. Em vista disto, por exemplo, analisando os quadros das Figuras 4.14, 4.16 e 4.18, é possível escolher um outro modelo de maior sensibilidade:

	Rede PNN	Sensibilidade	Bias	Alarme falso
a) 1ª hora	06	97%	4,04	12,1%
b) 2ª hora	06	98%	6,63	12,1%
c) 3ª hora	04	90%	7,03	15,2%

Pelos resultados mostrados se percebe que a RNA PNN consegue um bom desempenho em prever a classificação do nevoeiro para a primeira, segunda e terceira horas futuras.

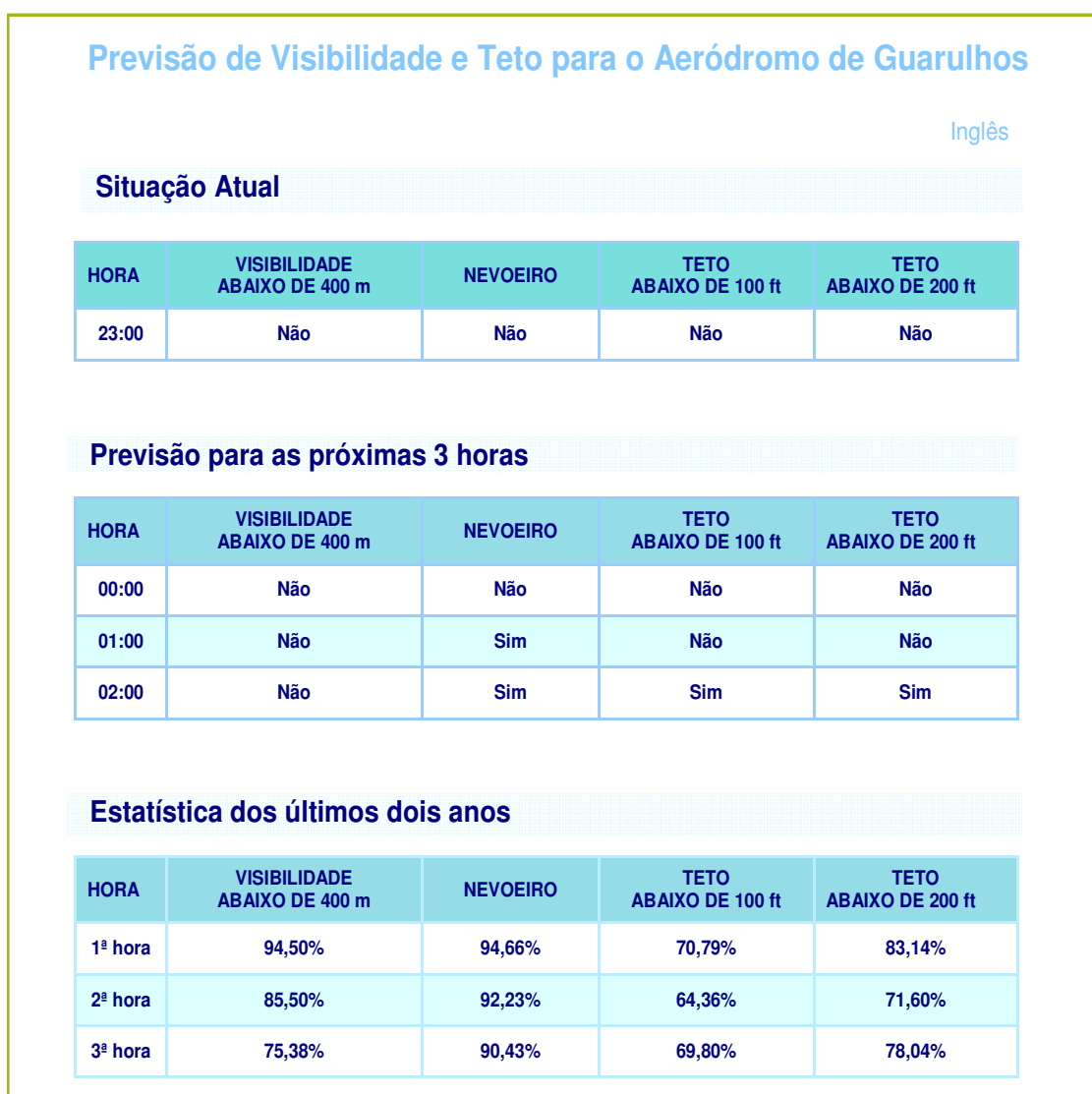
Diante desses bons resultados para a previsão de ocorrência de nevoeiro e para a previsão de classificação da faixa 1 da visibilidade – menor que 400 m, o modelo final a ser implantado para a previsão de visibilidade no aeroporto de Guarulhos terá as seguintes metas:

- a. A previsão de nevoeiro nas três horas futuras.
- b. A previsão da visibilidade menor que 400 m, nas três horas futuras.

Diante dos estudos semelhantes realizados para a previsão do teto e dos resultados obtidos, o padrão a ser implantado para a previsão de teto nas três horas futuras no Aeroporto de Guarulhos será o seguinte:

- a. Teto abaixo de 30m (ou 100ft);
- b. Teto abaixo de 60m (ou 200ft).

Após a definição do modelo e tendo sido gerado os códigos executáveis, foi desenvolvida a estrutura de operacionalização do sistema (vide Anexo B). A Figura 4.20 mostra a representação da tela final de interface do sistema com o usuário, apresentando resultados operacionalizados com dados de agosto de 2005 a julho de 2007.



**Figura 4.20** – Representação da tela de interface do modelo de previsão desenvolvido e instalado no ICEA/Aeroporto de Guarulhos. Na primeira parte da tela é mostrada a situação atual, na segunda parte a previsão para as três horas, e na última parte a estatística de acertos do sistema relativo aos dois últimos anos (agosto de 2005 a julho de 2007).

## CAPÍTULO 5 – CONCLUSÕES E SUGESTÕES

### 5.1 CONCLUSÕES

Da constatação dos efeitos danosos das condições meteorológicas adversas à aviação, principalmente no pouso e decolagem, e de que imensos prejuízos são advindos disto, o desenvolvimento e implementação de um sistema de previsão de curtíssimo prazo da visibilidade restrita e do teto baixo para a importante estrutura aeroportuária do Aeroporto de Guarulhos foi o objetivo e a contribuição deste trabalho. Adicionalmente, também, o objetivo foi responder às questões formuladas no item 1.3 desta tese.

Conclui-se deste trabalho:

- Para a primeira hora, a previsão de visibilidade menor que 1000m ou nevoeiro tem melhor desempenho com os modelos treinados com a rede PNN e com os conjuntos de treinamento TESTE03 e TESTE04. Os modelos com RLM e GRNN não apresentam diferenças muito significativas entre si, e apresentam desempenho abaixo dos modelos treinados com rede PNN.
- Para a segunda hora, para a previsão da visibilidade menor que 1000m ou nevoeiro, novamente os modelos PNN apresentam melhor desempenho que os modelos RLM e GRNN, e estes dois apresentam resultados muito próximos entre si.
- Na terceira hora, observa-se que há uma queda nos desempenho de todos os modelos. Os modelos RLM e GRNN mostram desempenhos abaixo de 50%. Porém, os modelos com PNN continuam com melhor desempenho.

Ao responder as questões levantadas, citadas anteriormente, este trabalho:

- a. Desenvolveu um sistema de previsão da visibilidade abaixo de 400m e abaixo de 1000m (nevoeiro) e do teto (nuvens baixas) abaixo de 30m e

abaixo de 60m, para as três primeiras horas futuras para um aeroporto individualizado, no caso o Aeroporto de Guarulhos-SP, algo que os atuais modelos numéricos não são capazes e que é inovador em nosso país;

b. Mostrou que o sistema desenvolvido não foi capaz de responder a contento na totalidade a questão quanto às previsões conforme a idéia inicial. Mas trouxe uma importante contribuição ao fornecer, para as três primeiras horas futuras, para o Aeroporto de Guarulhos:

- I. A previsão da visibilidade horizontal abaixo de 400m e a previsão de nevoeiro (visibilidade abaixo de 1000m);
- II. A previsão de teto abaixo de 30m e de 60m.

Esses valores correspondem a parte mais crítica da Tabela de Mínimos Operacionais para o Aeroporto de Guarulhos.

c. Indicou pelos resultados apresentados que:

I. Consegue prever o nevoeiro com o desempenho:

1. Para primeira hora, com as opções:

Sensibilidade	Especificidade
82%	98%
92%	95%
95%	93%

2. Para a segunda hora, com as opções:

Sensibilidade	Especificidade
65%	99%
79%	98%
91%	95%

3. Para a terceira hora, com as opções:

Sensibilidade	Especificidade
53%	97%
70%	92%
90%	85%

II. Consegue prever o teto abaixo de 30m com as performances:

	Sensibilidade	Especificidade
1ª hora	83%	96%
2ª hora	78%	95%
3ª hora	74%	90%

- d. Mostrou que a precisão apresentada no item anterior mostra-se aceitável, não somente pelo bom desempenho apresentado como também por que não há presentemente outra opção mais satisfatória para a problemática da previsão de curtíssimo prazo da visibilidade e teto previsão para o Aeroporto de Guarulhos.
- e. Além disso, mostrou que esse sistema de previsão desenvolvido tem um custo baixíssimo não somente na sua implementação como também na sua operacionalização. Após a implementação do executável desenvolvido (a rede PNN treinada), será necessário um PC (computador pessoal), para entrada dos dados em tempo real, interconectado via *intranet* a um *servidor*, também um PC, instalado no ICEA, em São José dos Campos e monitores de visualização aonde se fizer necessário.

## 5.2 SUGESTÕES

São sugeridas a seguir algumas recomendações surgidas a partir deste trabalho de tese:

- 1) O sistema desenvolvido não foi capaz de responder a contento na totalidade a questão quanto às previsões conforme a idéia inicial. Isto, possivelmente, deve-se, pelo menos em parte, a que os dados oriundos dos registros de visibilidade e teto mostraram uma tendenciosidade para alguns valores preferenciais. Isto acarreta a conseqüências de não de se ter uma relação

biunívoca de uma determinada condição meteorológica e um determinado valor de visibilidade ou de teto, o que resulta em um aprendizado deficiente do modelo. Portanto, é feita aqui a sugestão de que seja tomada uma ação no sentido de que os registros da visibilidade e teto sejam os mais próximos possíveis da realidade tendo em vista que, além de sua importância como boletim informativo para fins aeronáuticos, os mesmos comporão um futuro banco de dados meteorológicos. Banco este de importância capital para pesquisa e desenvolvimento.

- 2) A variável *TIPO* (que indica o tipo de nuvem) e as variáveis *Q1*, *Q2* e *Q3* (quantidade de nuvens em três camadas) foram usadas no treinamento para fornecer ao modelo informações indiretas das condições de instabilidade/estabilidade e da umidade em níveis baixo/médio da atmosfera local. Por isso, é feita aqui a sugestão para trabalhos futuros da inclusão de informações de ar superior (por exemplo, radiossondagens, perfis atmosféricos de modelos numéricos de mesoescala).
- 3) A criação da variável *6H\_SKC* (horas de céu claro nas últimas seis horas) pretendeu fornecer ao modelo informações indiretas da radiação solar recebida em horas de céu claro, durante o dia, e a radiação terrestre perdida nas horas de céu claro, durante a noite. Talvez, para trabalhos futuros, fosse interessante a instalação de um aparelho medidor de radiação solar recebida e da radiação terrestre emitida nas vizinhanças do aeroporto, a fim de se ter essa informação direta.
- 4) O uso das informações do *dia* e *mês* do ano (no formato do dia *juliano*) está relacionado à questão da sazonalidade. Como sugestão, para trabalhos futuros, redes neurais especializadas para as diferentes estações ou períodos do ano poderiam ser desenvolvidas. E, também, em vista da sazonalidade, talvez, algum trabalho incluindo alguma informação sobre anos de “el-niño” ou “la-niña”.

## CAPÍTULO 6 – REFERÊNCIAS BIBLIOGRÁFICAS

ALLAN, S. S., GADDY, S. G., EVANS, J. E. **Delay Causality and Reduction at the New York City Airports Using Terminal Weather Information Systems.** *Lincoln Laboratory, Massachusetts Institute of Technology*, <http://www.ll.mit.edu/AviationWeather/atc-291.pdf>, downloaded May 1, 2001.

**Assessment of Aerodrome Forecast Accuracy Improvement - Final Report.** In: NAV CANADA, May 2002.

BERGOT, T., GUÉDALIA, D., 1994. **Numerical Forecasting of Radiation Fog. Part I: Numerical Model and Sensitivity Tests**”, *Mon. Wea. Rev.*, v.122, pp. 1218-1230.

BINHUA, W. **Sea Fog.** Beijing, China Ocean Press, 1985, 330 pp.

BISHOP, C. M. **Pattern Recognition and Machine Learning.** New York, NY, USA: Springer Science-Business Media, 2006 (Printed in Singapore).

BOTT, A., TRAUTMANN, T. **PAFOG—A New Efficient Forecast Model of Radiation Fog and Low-Level Stratiform Clouds.** *Atmospheric Research* 64, pp. 191-203, 2002.

BOTT, A., SIEVERS, U., ZDUNKOWSKI, W. **A Radiation Fog Model with a Detailed Treatment of the Interaction between Radiative Transfer and Fog Microphysics.** *Journal of the Atmospheric Sciences*, 47, pp. 2153– 2166, 1990.

BOZNAR, M., LESJAK, M., MLAKAR, P. **A Neural Network-Based Method for Short-Term Predictions of Ambient SO<sub>2</sub> Concentrations in Highly Polluted Industrial Areas of Complex Terrain.** *Atmosphere Environment*, 27B, pp. 221-230, 1993.

BRAGA, A. P., CARVALHO, A. P. L., LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações.** Rio de Janeiro: LTC Editora, 2000.

BRAGA, P. C. A. S. **Previsão do IBOVESPA utilizando Modelos Híbridos.** Dissertação para a obtenção do Grau de Mestre em Ciências em Engenharia Elétrica, COPPE/UFRJ, abril 2006.

BREMNES, J. B.; MICHAELIDES, S. C. “*Probabilistic Forecast of Rare Visibility Events using Neural Networks*”. In: **Symposium on Nowcasting and Very Short Range Forecasting**, Toulouse, France, September 2005.



CABRAL, E. **Climatologia dos nevoeiros no Aeroporto de Guarulhos**. SIPAER, São Paulo, n. 57, p. 262-268, 1998.

CALÔBA, L. P. “*Introdução ao Uso de Redes Neurais na Modelagem de Sistemas Dinâmicos e Séries Temporais*”. In: **Minicurso no XIV Congresso Brasileiro de Automática**, Natal, 2002.

CERMAK, J., BENDIX, J. **A novel approach to fog/low stratus detection using Meteosat 8 data**. Atmospheric Research, 2007.

CIARELLI, P. M. **Rede Neural Probabilística para a Classificação de Atividades Econômicas**. Dissertação de mestrado, Universidade Federal do Espírito Santo, ES, 2008.

SCHAEFER, J. T. **The Critical Success Index as an Indicator of Warning Skill**. Weather and Forecasting, vol 5, pag 570-573, dec-1990.

CHANDRA, B., VARGHESE, P. P. “*Applications of Cascade Correlation Neural Networks for Cipher System Identification*”. In: **Proceedings of World Academy of Science, Engineering and Technology**, vol. 20, April 2007.

COSTA, F.H. S., SOUZA FILHO, C. R. “*Aplicação de Redes Neurais Artificiais para Reconhecimento de Padrões em Solos*”. In: **Anais XIV Simpósio Brasileiro de Sensoriamento Remoto**, pp. 5177-5183, Natal, abril de 2009.

COSTA, S.B., et al. “*Fog forecast for the international airport of Maceió, Brazil using artificial neural network*”. In: **Proc. 8th ICSHMO**, pp.1741-1750, Foz do Iguaçu, Abril 2006.

DAL PIVA, E., FEDEROVA, N. “*Um estudo sobre a formação de nevoeiro de radiação em Porto Alegre*”. In: **Revista Brasileira de Meteorologia**, v. 14, n.2, pp. 47-62, 1999.

DOSWELL, C. A., BROOKS, H.E. “*Budget-Cutting and the Value of Weather Services*”. In: **Weather and Forecasting**, v. 13, pp. 206-212.

FORKEL, R., SIEVERS, U., ZDUNKOWSKI, W. “*Fog Modeling with a New Treatment of the Chemical Equilibrium Condition*”. In: **Beitraege zur Physik der Atmosphaere**, nr. 60, pp. 340-360, 1987.

FREITAS, J. H. V., ALMEIDA, M. V., FRANÇA, G. B. “*Análise da Ocorrência de Nevoeiro no Aeroporto de Guarulhos*”. In: **XXX Jornada Giulio Massarani de Iniciação Científica, Artística e Cultural**, Rio de Janeiro: UFRJ, novembro de 2008.

GARDNER, M. W., DORLING, S. R. “*Artificial Neural Networks - a Review of Applications in the Atmospheric Sciences*”. In: **Atmospheric Environment**, nr. 32, pp. 2627-2636, 1998.

GERAPETRITIS, H., PELISSIER, J. M. “*The Critical Success Index and Warning Strategy*”. In: **17th Conference on Probability and Statistics in the Atmospheric Sciences**, Seattle, Washington, 2004.

GRIMM, A. M. **Meteorologia Básica - Notas de Aula**. Primeira versão eletrônica – set/1999 (<http://fisica.ufpr.br/grimm/aposmeteo/index.htm> 06/09/07).

**Guide to Meteorological Instruments and Methods of Observations**. Fifth edition, Geneva, Switzerland, WMO, 1983.

GULTEPE, I., PAGOWSKI, M., REID, J.”*Using Surface Data to Validate a Satellite Based Fog Detection Schem*”. In: **Weather and Forecasting**, nr. 22, pp. 444-456, 2007.

GURKA, J. J., MOSHER, F. R. “*Steps to Improve Ceiling and Visibility Forecasts for Aviation*”. In: **Preprints, 9th Conference on Aviation, Range, and Aerospace Meteorology**. Orlando, FL, 11-15. *American Meteorological Society*, pp. 327-330, September 2000.

HANSEN, K. B., RIORDAN, D. “*Fuzzy Case-based Prediction of Cloud Ceiling and Visibility*”. In: **3rd Conference on Artificial Intelligence Applications to the Environmental Science**. *American Meteorological Society*, 2003.

HAYKIN, S. **Redes Neurais Princípios e Prática**. 2ª ed., Porto Alegre: Editora Bookman, 2002.

HALL, T., BROOKS, H. E., DOSWELL, C. A. “*Precipitation Forecasting Using a Neural Network*”. In: **Weather Forecasting**, nr. 14, pp. 338-345, 1999.

HÄRTER, F. P., CAMPOS VELHO, H. F. “*Redes Neurais Recorrentes e Feedforward Treinadas com Correlação Cruzada Aplicadas à Assimilação de Dados em Dinâmica Não-Linear*”. In: **Revista Brasileira de Meteorologia**, v. 20, nr.3, pp. 411-420, 2005.

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

HSIEH, W. W., TANG, B. “*Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography*”. In: **Bulletin of the American Meteorological Society**, nr. 79, pp. 1855-1870, 1998.

JIANGA, D. *et al.* "Progress in Developing an ANN Model for Air Pollution Index Forecast". In: **Atmospheric Environment**, nr. 38, pp. 7055-7064, 2004.

LAMBERT, W. C. **Statistical Short-Range Forecast Guidance for Cloud Ceilings over the Shuttle Landing Facility**. For Kennedy Space Center, under contract NAS10-96018 - NASA, Office of Management Scientific and Technical Information Program, 2001

LEYTON, S.M., FRITSCH, J.M. "Short-term Probabilistic Forecasts of Ceiling and Visibility utilizing High-density Surface Weather Observations". In: **Weather and Forecasting**, nr. 18, pp. 891-902, 2003.

LUK, K.C., BALL, J.E., SHARMA, A.. "A Study of Optimal Model Lag and Spatial Inputs to Artificial Neural Network for Rainfall Forecasting". In: **Journal of Hydrology**, nr. 227, pp. 56-65, 2000.

**Manual on the Global Observing System**. Vol. 1, appendix p.8, WMO - World Meteorological Organization, 2002.

MARZBAN, C., STUMPF, G. F. "A Neural Network for Tornado Prediction Based on Doppler Radar-Derived Attributes". In: **Journal Applied Meteorology**, nr. 35, pp. 617-626, 1996.

MARZBAN, C., STUMPF, G. F. "A Neural Network for Damaging Wind Prediction". In: **Weather Forecasting**, nr. 13, pp. 151-163, 1998.

MARZBAN, C., LEYTON, S., Colman, B. "Ceiling and Visibility Forecasts via Neural Nets". In: **Technical Report no. 490**, University of Washington, September 2005.

McCANN, D. W. "A Neural Network Short-Term Forecast of Significant Thunderstorms". In: **Weather Forecasting**, nr. 7, pp. 525-534, 1992.

MOURA JR, J. R. V., MARTINS, G., JACOMINE, L. *et al.* "Redes Neurais Probabilísticas e Aplicação em Otimização de Servidores de Internet". In: **47ª Reunião Anual Região Brasileira da Sociedade Internacional de Biometria**, Rio Claro-SP, 2002.

MOUTINHO, A. M. **Identificação de Padrões Faciais usando Redes Neurais Artificiais**. Dissertação de mestrado em Informática do Núcleo de Computação Eletrônica, do Instituto de Matemática da Universidade Federal do Rio de Janeiro, 2005.

MURPHY, A.H. **Assessing the Economic Value of Weather Forecasts: An Overview of Methods, Results and Issues**. Meteorological Institute, University of Hamburg, Germany in *Meteorology Applications*, p. 69-73, 1994.

MURPHY, A.H., EHRENDORFER, M.. **On the Relationship between the Accuracy and Value of Forecasts in the Cost-Loss Ration Situation**. Department of Atmospheric Sciences, Oregon State University, in *American Meteorology Society*, p. 243-251, September 1987.

NASCIMENTO, E. M. **Redes Neurais Artificiais: Uma Aplicação No Estudo da Poluição Atmosférica e seus Efeitos Adversos à Saúde**. Dissertação de mestrado em Ciências em Engenharia de Produção, COPPE-UFRJ, Rio de Janeiro, 2006.

NISKAA, H. et al. "*Evolving the Neural Network Model for Forecasting Air Pollution Time Series*". In: **Engineering Applications of Artificial Intelligence**, nr. 17 p. 159-167, 2004.

OLIVEIRA, A. O. **Método Estatístico no Auxílio à Previsão de Nevoeiro para o Aeródromo de Guarulhos**. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, 2002.

ORLANSKI, I. "*A Rational Subdivision of Scales for Atmospheric Processes*". In: **Bulletin of the American Meteorological Society**, 56(5), pp. 527-530, 1975.

OSRÓDKA, K., SZTURC, J., JURCZYK, A. **Using Neural Network to Precipitation Nowcasting based on Weather Radar Data**. WWRP QPF Conference, UK, The Royal Meteorology Society, September 2002.

PASERO, E; MONIACI, W. "*Artificial Neural Networks for Meteorology Nowcast*". In: **IEEE International Conference on Computational Intelligence Systems and Applications**, Boston, USA, July 2004.

PASINI, A., PELINO, V. "*Can We Estimate Atmospheric Predictability by Performance of Neural Network Forecasting? The Toy Case Studies of Unforced and Forced Lorenz Models*". In: **IEEE International Conference on Computational Intelligence for Measurement Systems and Applications**, Italy, July 2005.

PASINI, A., PELINO, V., POTESTÀ, S. "*A Neural networks Model for Visibility Nowcasting from Surface Observations: Results and Sensitivity to Physical Input Variables*". In: **Journal of Geophysical Research**, vol. 106, D14, July 2001.

PEREIRA FILHO, A. J., SANTOS, C. C. "*Modeling a Densely Urbanized Watershed with an Artificial Neural Network, Weather Radar and Telemetric Data*". In: **Journal of Hydrology**, nr. 317, pp. 31-48, 2006.

PINHEIRO, F. R. **Previsão de Eventos de Nevoeiro Utilizando o Modelo Cobel**. Dissertação de mestrado, Engenharia Civil, COPPE-UFRJ, Rio de Janeiro, 2006.

PINTO, H. S. Artigo na **ComCiência - Revista Eletrônica de Jornalismo Científico**. 2002.

RAMÍREZ, M. C. V. ,CAMPOS VELHO, H. F., FERREIRA, N. J. “*Artificial Neural Network Technique for Rainfall Forecasting Applied to the São Paulo Region*”. In: **Journal of Hydrology**, v. 301, Issues 1-4, pp. 146-162, January 2005.

RAMÍREZ, M. C. V., LI WEIGANG, FERREIRA, N. J. “*Aplicação de Redes Neurais Artificiais e Transformada de Ondeletas para Estimativas de Precipitação usando imagens do satélite GOES-8: Resultados Preliminares*”. In: **Proceedings of the IV Brazilian Conference on Neural Networks - IV Congresso Brasileiro de Redes Neurais**, pp. 252-256, São José dos Campos, julho de 1999.

REZENDE, S. O. **Sistemas Inteligentes – Fundamentos e Aplicações**. São Paulo: Ed. Manole, 2003.

ROGERS, R.R, YAU, M.K. **A Short Course in cloud physics**. 3 ed In: *Int. Series in Nat. Phil.*, Pergamon Press, Oxford, 1989.

SANTAMOURIS, M. et al. “*A Neural Network Approach for Modeling the Heat Island Phenomenon in Urban Areas during the Summer Period*”. In: **Geophysics Research Letters**, nr. 26, pp. 337-340, 1999.

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.

S. J. YAO et al. “*Wavelet Transform and Neural Networks for Short-Term Electrical Load Forecasting*”. In: **Energy Conversion & Management**, nr. 41, pp. 1975-1988, 2000.

SIEBERT, J., BOTT, A., ZDUNKOWSKI, W. “*Influence of a Vegetation-Soil Model on the Simulation of Radiation Fog*”. In: **Beitraege zur Physik der Atmosphaere**, nr. 65, pp. 93-106, 1992.

SILVA, L. C. F. **Inteligência Computacional para Predição de Produção de Reservatórios de Petróleo**. Tese de doutorado em Ciências em Engenharia Civil, COPPE-UFRJ, Rio de Janeiro, 2006.

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

SPECHT, D. F. “*Probabilistic Neural Networks*”. In: **IEEE Transactions on Neural Networks**, vol. 3, pp.109-118, 1990.

SPECHT, D. F. “*A General Regression Neural Network*”. In: **IEEE Transactions on Neural Networks**, vol. 2(6), pp. 568-576, 1991.

VALENÇA, M. **Aplicando Redes Neurais – Um Guia Completo**. Pernambuco: Elógica, 2005.

VALDEZ, J. **NWS Reinventing Goals for 2000**. National Weather Service, National Oceanic and Atmospheric Administration, United States Department of Commerce. Downloaded <http://www.nws.noaa.gov/npr5.html>, 2000.

VARGAS, E. C., SOUSA, H. C., CARVALHO, A. C. P. L. F. “*Reconhecimento de Alvos utilizando Redes Neurais Construtivas*”. In: **Anais IX Simpósio Brasileiro de Sensoriamento Remoto**, pp. 1247-1256, Santos-SP, setembro 1998.

“*Vocabulário Meteorológico Internacional*”. In: **Curso de Geografia**, UFJF, Juiz de Fora, MG.

VON GLASOW, R., BOTT, A. “*Interaction of Radiation Fog with Tall Vegetation*”. In: **Atmospheric Environment**, nr. 33, pp. 1333-1346, 1999.

VON ZUBEN, F. J. “*Uma Caricatura Funcional de Redes Neurais Artificiais*”. In: **Revista da Sociedade Brasileira de Redes Neurais**, vol. 1 (2), pp. 77-87, 2003.

WALTER, I. A. “*Revisão Bibliográfica sobre General Regression Neural Network (GRNN)*”. In: **Curso IA353 - Redes Neurais**. Departamento de Engenharia de Computação e Automação Industrial da Faculdade de Eng. Elétrica e de Computação da Universidade Estadual de Campinas, Abril 2003.

WIBERG, B., KJELBERG, J. **Project: Artificial Neural Networks in Weather Prediction**. Artificiella neuronnät DV1, VT, Stockholm, Sweden 2001.

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

ZHANG, G. P.; MIN Qi. “*Neural Network Forecasting for Seasonal and Trend Time Series*”. In: **European Journal of Operational Research**, nr.160, pp. 501-514, 2005.

## GLOSSÁRIO

Coeficiente de Determinação Múltipla (R) - É um indicador estatístico geralmente aplicado em análise de regressão múltipla. Ele compara a precisão do modelo em estudo com a de um modelo de referência trivial (como, por exemplo, a simples média dos dados).

Cumulunimbus - Abreviatura CB. Tipo de nuvem excepcionalmente densa e de grande desenvolvimento vertical. Sua precipitação geralmente é em forma de aguaceiro. É responsável pela formação de relâmpagos, trovões e tempestades e, em alguns casos, tornados.

DME - Distance Measuring Equipment – Assim como o VOR, são procedimentos que exigem ainda ampla atuação do piloto na condução do voo.

IFR- Instrument Flight Rules – regras de voo por instrumentos. No voo IFR, há uma orientação do voo por um instrumento que pode ser NDB, VOR/DME ou ILS.

ILS - Instrument Landing System – Sistema de Pouso por Instrumentos. É o mais preciso procedimento de pouso. Há uma menor interferência do piloto na condução do voo; o ILS pode ser de três categorias: CAT I, II e III.

Tabela de contingência (ou *Matriz de erro* Ou *matriz de confusão*) - é uma matriz quadrada de números que expressam a quantidade de unidades amostrais (associada a uma dada categoria durante um processo de classificação) e a categoria real a que pertencem essas unidades (Rosenfield e Fitzpatrick, 1986).

METAR - *Meteorological Terminal Air Report* - Mensagem Meteorológica para Terminal Aéreo. É uma mensagem meteorológica regular para aviação; faz parte de códigos internacionais padrões de observações meteorológicas horárias.

Nó (kt) - Medida de velocidade. Um nó é igual uma milha náutica/hora ou 1852m/h ou aproximadamente 0,51m/s.

Stratus - Tipo de nuvem geralmente em camada uniforme e cinzenta, mas, às vezes em forma esgarçada, situada em baixa altura (algumas vezes parece um nevoeiro que se elevou do solo).

Subsidência - Movimento lento, descendente, de uma massa de ar na atmosfera, do que resulta na transferência de suas características para uma área maior.

Temperatura do ponto de orvalho - É a temperatura em que o ar, nas condições atuais, deve ser resfriado para ficar saturado (100% de umidade).

Tetômetro – É um sensor para a estimativa da altura de nuvem que, em geral, emprega tecnologia LIDAR (*Light Detecting and Ranging*).

UTC - Universal Coordinated Time. Hora mundial, ou Tempo Coordenado Universal. É também conhecido com Hora Média de Greenwich (GMT - Greenwich Mean Time) ou, no jargão da aviação, hora "Zulu".

VOR - Very High Frequency Omnidirectional Range - O VOR opera nas frequências de 108.0 MHz à 117.9 MHz. Ele é utilizado como balizamento na navegação em rota, e também nas aproximações, como auxílio básico de procedimentos.

Visibilômetro – conjunto composto de projetor, receptor de luminância: tem a finalidade de fornecer os valores de visibilidade ao longo da(s) pista(s), baseada na transmissividade horizontal da atmosfera ( MCA 105-2 “Manual de Estações Meteorológicas de Superfície”).



## APÊNDICE A – SOBRE O REGISTRO DOS DADOS

Os dados utilizados na pesquisa foram extraídos do Banco de Dados do Instituto de Controle do Espaço Aéreo - ICEA do Comando da Aeronáutica, e estavam inicialmente no formato *texto* (TXT). Todos os dados foram, então, passados para planilhas Excel, onde cada linha representa um evento meteorológico ou observação meteorológica. Como total de linhas de dados é muito grande (mais de 400.000) e como há uma limitação do número de linhas na planilha ( $2^{16}$  ou 65.536 linhas), os dados foram organizados em várias planilhas. Na Figura A.1 é mostrada a imagem de parte de um dos arquivos originais dos dados, “sbgr105251”, relativa a observações meteorológicas do dia primeiro de janeiro de 1951. Esses dados foram oriundos do formulário IEPV-105-25.

NSINOTICO	DIA	MES	ANO	HORA	MINUTO	TOTNUVEM	DIRVENTO	VELVENTO	VISIB	CGT
83778	1	1	1951	2	0	8	10	2	1200	4
83778	1	1	1951	3	0	8	11	3	1200	4
83778	1	1	1951	4	0	8	0	0	800	4
83778	1	1	1951	5	0	8	0	0	500	4
83778	1	1	1951	6	0	8	18	2	800	6
83778	1	1	1951	7	0	8	0	0	1000	1
83778	1	1	1951	8	0	8	0	0	2000	0
83778	1	1	1951	9	0	8	0	0	2000	0
83778	1	1	1951	10	0	8	17	10	2000	0
83778	1	1	1951	11	0	6	18	7	2000	0
83778	1	1	1951	12	0	8	18	4	2000	0
83778	1	1	1951	13	0	8	14	7	2000	0
83778	1	1	1951	14	0	8	18	10	2000	0
83778	1	1	1951	15	0	8	16	12	2000	0
83778	1	1	1951	16	0	8	18	3	2000	0
83778	1	1	1951	17	0	8	18	6	2000	0
83778	1	1	1951	18	0	7	18	7	2000	0
83778	1	1	1951	19	0	8	16	4	2000	0
83778	1	1	1951	20	0	8	16	3	1400	1
83778	1	1	1951	21	0	8	0	0	1400	4
83778	1	1	1951	22	0	8	0	0	1400	4
83778	1	1	1951	23	0	8	16	7	1400	4
83778	2	1	1951	0	0	8	16	10	1400	4
83778	2	1	1951	1	0	8	16	8	1400	4
83778	2	1	1951	2	0	8	16	8	1400	4

217295 linhas selecionadas

**Figura A. 1-** Imagem mostrando parte do arquivo “sbgr105251” com dados relativos a observações meteorológicas dos dias 1 e 2 de janeiro de 1951. Esses dados são oriundos do formulário IEPV-105-25.

A tabela de nuvens usada anteriormente ao ano de 2005 era a Tabela 0500 do Manual de Códigos Meteorológicos do Comando da Aeronáutica (primeira coluna da Tabela A.1), a partir de então passou a ser usada uma tabela criada para o formulário IEPV-105-78 (segunda coluna da Tabela A.1), a qual foi adaptada e adotada para este trabalho de pesquisa como “Tabela de nuvens” (terceira coluna da Tabela A.1). A única diferença entre essas duas tabelas (as das segunda e terceira colunas) está relacionada à correspondência do número “0” com o tipo de nuvens. Na “Tabela de nuvens” foi incluída a sigla SKC (*sky clear*) indicativa da condição de céu claro.

**Tabela A. 1-** Correspondência entre as tabelas de nuvens: Tabela 0500 do Manual de Códigos Meteorológicos do Comando da Aeronáutica (primeira coluna), usada anteriormente ao ano de 2005, a tabela criada para o formulário IEPV-105-78 (segunda coluna) e a Tabela de nuvens (terceira coluna), em relação à tabela da segunda coluna, foi incluída a sigla SKC (*sky clear*) indicativa da condição de céu claro.

Tabela 0500 (Manual de Códigos)	Tabela do formulário IEPV 105-78 (2005)	Tabela de nuvens	
0 Cirrus... CI	0 Cirrocumulus/Cirrus CC/CI	0	CI/CC/SKC
1 Cirrocumulus... CC	1 Cirrostratus... CS	1	CS
2 Cirrostratus... CS	2 Altocumulus... AC	2	AC
3 Altocumulus... AC	3 Altostratus... AS	3	AS
4 Altostratus... AS	4 Nimbostratus... NS	4	NS
5 Nimbostratus... NS	5 Stratocumulus... SC	5	SC
6 Stratocumulus... SC	6 Stratus... ST	6	ST
7 Stratus... ST	7 Cumulus... CU	7	CU
8 Cumulus... CU	8 Grandes Cumulus... TCU	8	TCU
9 Cumulonimbus... CB	9 Cumulonimbus... CB	9	CB

Como um procedimento de pré-processamento dos dados, foram inseridas linhas, nas planilhas dos arquivos de dados, para os eventos meteorológicos faltosos. As células, dessas linhas da planilha, relativas à hora, ao dia, ao mês e ao ano foram preenchidas com seus valores respectivos. Foram mantidas vazias as células

restantes de cada linha. A finalidade desse procedimento foi estabelecer toda a seqüência temporal do período dos dados. Deve-se esclarecer aqui que este é apenas um procedimento de organização dos dados. Quando futuramente forem determinados os conjuntos de treinamento e de validação, todas as linhas com dados faltosos serão eliminadas. Esse procedimento é exemplificado através da Figura A.2, onde é mostrada a imagem de parte do arquivo *SBGR\_8086.xls* onde é destacada a ausência de observações meteorológicas no dia 01 de janeiro de 1981.

	A	B	C	D	E	F	G	H	I	J	K	
1	DIA	MES	ANO	HORA	DIRVENTO	VELVENTO	VISIB	CGT	QTDNUVEM1	TIPONUVEM1	ALTNUVEM1	QTE
8779	31	12	1980	17	18	14	1500	2	4	7	90	
8780	31	12	1980	18	18	10	2000	0	3	5	90	
8781	31	12	1980	19	20	10	2000	0	2	5	90	
8782	31	12	1980	20	0	0	1400	2				
8783	31	12	1980	21	15	4	1000	2	3	5	60	
8784	31	12	1980	22	10	4	800	2	6	5	60	
8785	31	12	1980	23	12	8	800	4	6	5	60	
8786	1	1	1981	0								
8787	1	1	1981	1								
8788	1	1	1981	2								
8789	1	1	1981	3								
8790	1	1	1981	4								
8791	1	1	1981	5								
8792	1	1	1981	6								
8793	1	1	1981	7								
8794	1	1	1981	8								
8795	1	1	1981	9								
8796	1	1	1981	10								
8797	1	1	1981	11								
8798	1	1	1981	12								
8799	1	1	1981	13								
8800	1	1	1981	14								
8801	1	1	1981	15								
8802	1	1	1981	16								
8803	1	1	1981	17								
8804	1	1	1981	18								
8805	1	1	1981	19								
8806	1	1	1981	20								
8807	1	1	1981	21								
8808	1	1	1981	22								
8809	1	1	1981	23								

**Figura A.2** - Imagem de parte do arquivo *SBGR\_8086.xls* destacando o dia 01/01/1981, onde não houve observações meteorológicas ou foram perdidos os registros.

Esses dados básicos, nessa seqüência apresentada, estão representados na Figura A.3, relativa à imagem de parte do arquivo *SBGR\_8086.xls*. Nessa figura, representando parte de uma planilha Excel, os rótulos dos dados básicos citados

estão destacados na primeira linha. Destacam-se, também, as situações de ausência de nuvens em eventos meteorológicos existentes.

1	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	DIA	MES	ANO	HORA	DIRVENTO	VELVENTO	VISIB	CGT	QTDNUVEM1	TIPONUVEM1	ALTNUVEM1	QTDNUVEM2	TIPONUVEM2	ALTNUVEM2	QTDNUVEM3	TIPONUVEM3	ALTNUVEM3	TEMPAR	UR	QNH
2	1	1	1980	0	18	7	500	6	8	5	30							19	95	1011,1
3	1	1	1980	1	18	8	600	4	8	5	30							17	94	1011,5
4	1	1	1980	2	15	3	600	5	8	5	30							16	94	1010,9
5	1	1	1980	3	0	0	600	4	8	5	30							16	95	1010,5
6	1	1	1980	4	18	6	800	4	8	5	30							16	94	1010,7
7	1	1	1980	5	19	6	400	5	8	5	30							16	97	1011
8	1	1	1980	6	18	6	300	5	8	5	30							16	100	1011,9
9	1	1	1980	7	12	4	300	4	8	5	30							16	97	1012,8
10	1	1	1980	8	20	6	400	5	8	5	20							15,5	97	1013,2
11	1	1	1980	9	18	7	400	5	6	6	20	8	5	60				15,5	95	1013,9
12	1	1	1980	10	18	9	800	4	8	5	30							17	87	1014
13	1	1	1980	11	18	9	800	4	8	5	40							18	84	1014,1
14	1	1	1980	12	22	5	800	2	8	5	45							18,5	78	1013,9
15	1	1	1980	13	18	5	800	2	8	5	45							19	81	1013,5
16	1	1	1980	14	18	4	800	2	8	5	50							19	81	1013,1
17	1	1	1980	15	12	6	600	2	8	5	50							19,5	74	1012,5
18	1	1	1980	16	18	4	600	2	8	5	50							19	81	1012,4
19	1	1	1980	17	18	8	600	4	8	5	40							18	84	1012,7
20	1	1	1980	18	18	6	400	4	8	5	35							17	86	1013,1
21	1	1	1980	19	15	6	200	4	8	5	30							16,5	87	1013,7
22	1	1	1980	20	15	4	300	4	8	5	30							16,5	87	1014,5
23	1	1	1980	21	13	6	400	4	7	5	30							16,5	86	1015,1
24	1	1	1980	22	0	0	1000	4	3	5	60							15	94	1015,6
25	1	1	1980	23	0	0	1000	4	3	5	60							14	94	1015,6
26	2	1	1980	0	0	0	800	4	6	5	45							14,5	95	1015,4
27	2	1	1980	1	0	0	1000	4										13	93	1014,7
28	2	1	1980	2	0	0	1000	4										13	93	1014,5
29	2	1	1980	3	9	4	600	4	6	5	40							13	95	1014,3
30	2	1	1980	4	9	3	500	4	6	5	40							13	93	1014,4

**Figura A.3** - Imagem de parte do arquivo SBGR\_8086.xls, mostrando os dados básicos de trabalho com seus rótulos destacados na primeira linha. São destacados, também, as situações de ausência de nuvens em eventos meteorológicos existentes.

Nas tabelas A.2 e A.3 são mostrados os formatos dos registros relativos às condições do tempo ou fenômeno meteorológico. A Tabela A.2, das *Condições Gerais de Tempo*, se refere ao formulário IEPV 105-25, que vigorou anteriormente ao ano de 1980 até 1993/1994. Nesse caso, fica clara a simplicidade dessa tabela (com apenas 10 linhas) em contraste com a complexidade da Tabela A.3, que representa a Tabela 4677 do Manual de Códigos Meteorológicos do Comando da Aeronáutica. Esta tabela é apresentada na forma completa com 100 linhas descritivas dos fenômenos meteorológicos. A Tabela 4677 passou a ser utilizada nos formulários de registros meteorológicos seguintes, a partir de 1994: IEPV 105-13 e IEPV 105-78.

**Tabela A.2** – Relação das *Condições Gerais de Tempo*, ou do fenômeno meteorológico, que era utilizada para o preenchimento do formulário IEPV 105-25, até o ano de 1979.

<b>Fenômeno Meteorológico</b>	<b>Nº</b>
Nada a relatar	0
Precipitação à vista	1
Névoa seca ou fumaça	2
Areia ou poeira	3
Névoa úmida ou nevoeiro	4
Chuvisco	5
Chuva	6
Neve	7
Trovoada ou relâmpago	8
Granizo	9

**Tabela A.3** - Tabela 4677 do Manual de Códigos Meteorológicos (MCA 105-10) mostrando a relação das *Condições Gerais de Tempo*, ou do fenômeno meteorológico. Esta tabela passou a ser utilizada nos formulários de registros meteorológicos a partir de 1994: IEPV 105-13 e IEPV 105-78.

1º JUL 2001		MCA 105-10			
TABELA 4677					
ww - Tempo presente informado de uma estação meteorológica dotada de pessoal					
ww = 00-49		<i>Sem precipitação na estação na hora da observação</i>			
ww = 00-19		Sem precipitação, nevoeiro, nevoeiro glacial (exceto para 11 e 12), tempestade de areia, tempestade de poeira, neve soprada baixa ou em suspensão na estação, na hora da observação ou, exceto para 09 e 17, durante a hora precedente			
<b>Nº do Código</b>					
Nenhum névoa, exceto fotonevoeiro	{	00	Nenhum desenvolvimento de nuvens foi observado ou não pode ser observado	} Variação característica do estado do céu, durante a hora passada	
		01	Nuvens geralmente se dissipando ou se tornando menos espessas		
		02	Estado do céu invariável no conjunto		
		03	Nuvens geralmente em formação ou em desenvolvimento		
Névoa seca, areia, poeira, ou fumaça	{	04	Visibilidade reduzida por fumaça; por exemplo, de queimadas ou incêndios de florestas, fumaça de origem industrial ou cinzas vulcânicas	}	
		05	Névoa seca		
		06	Poeira generalizada, em suspensão no ar, com grande extensão, não levantada pelo vento, perto da estação ou arredores, na hora da observação		
		07	Poeira ou areia levantada pelo vento na/ou perto da estação, na hora da observação, mas sem remoinhos de poeira ou de areia definidos, e sem tempestade de poeira ou de areia à vista; ou no caso de navios, borrifos do mar em suspensão		
		08	Remoinho(s) de poeira ou de areia bem definido(s), observados na estação ou arredores, durante a hora precedente ou na hora da observação, mas sem tempestade de poeira ou de areia		
		09	Tempestade de poeira ou de areia à vista na hora da observação, ou na estação, durante a hora precedente		
		10	Névoa úmida		
		11	Em bancos		} Camada fina de nevoeiro ou de nevoeiro gelado na estação, quando em terra sua espessura não é superior a 2 metros e no mar a 10 metros
		12	Mais ou menos contínua		
					13

**Tabela A. 3 – Continuação.**

1º JUL 2001		MCA 105-10
14	Precipitação à vista, que não chega ao solo ou à superfície do mar	
15	Precipitação à vista, que chega ao solo ou à superfície do mar, mas longe da estação, isto é, estimada ser a mais de 5 km	
16	Precipitação à vista, que chega ao solo ou à superfície do mar próximo à estação, mas não na estação	
17	Trovoada, sem precipitação, na hora da observação	
18	Forte aguaceiro com trovoada e vento forte	} Na Estação ou a distância, durante a hora precedente ou no momento da observação
19	Nuvem(ns) fúmil (tornado ou tromba d'água)	
<hr/>		
ww = 20-29	Precipitação, nevoeiro, nevoeiro glacial ou trovoada na estação durante a hora precedente, mas não na hora da observação	
<hr/>		
20	Chuvisco (não congelante) ou neve granulada	} em pancadas
21	Chuva (não congelante)	
22	Neve	
23	Chuva e neve ou grãos de gelo	
24	Chuvisco ou chuva congelante	
25	Pancada(s) de chuva	
26	Pancada(s) de neve, ou de chuva e neve	
27	Pancada(s) de granizo*, ou de chuva e granizo*	
28	Nevoeiro ou nevoeiro gelado	
29	Trovoada (com ou sem precipitação)	
* Granizo, granizo pequeno, neve granulada		
<hr/>		
ww = 30-39	Tempestade de poeira, tempestade de areia, neve soprada ou em suspensão	
<hr/>		
30	} Tempestade de poeira ou areia, leve ou moderada	- Diminuindo durante a hora precedente
31		- Sem mudança apreciável durante a hora precedente
32		- Começou ou aumentou durante a hora precedente
33	} Tempestade forte de poeira ou areia	- Diminuindo durante a hora precedente
34		- Sem mudança apreciável durante a hora precedente
35		- Começou ou aumentou durante a hora precedente
36	Neve flutuando, leve ou moderada	} Geralmente baixa (abaixo do nível dos olhos)
37	Neve flutuando, forte	

**Tabela A. 3 – Continuação.**

	1º JUL 2001		MCA 905-10
38	Neve soprada, leve ou moderada	} <i>Garimpo alta (acima do nível dos olhos)</i>	
39	Neve soprada, forte		
<hr/>			
<b>ww = 40-49</b>	<b>Nevoeiro ou nevoeiro glacial na hora da observação</b>		
<hr/>			
40	Nevoeiro ou nevoeiro glacial à distância na hora da observação, mas não na estação durante a hora precedente; o nevoeiro ou nevoeiro glacial se estende a um nível superior ao do observador		
41	Nevoeiro ou nevoeiro glacial em bancos		
42	Nevoeiro ou nevoeiro glacial, céu visível	} Tornou-se mais fraco durante a hora precedente	
43	Nevoeiro ou nevoeiro glacial, céu invisível		
44	Nevoeiro ou nevoeiro glacial, céu visível	} Sem mudança apreciável durante a hora precedente	
45	Nevoeiro ou nevoeiro glacial, céu invisível		
46	Nevoeiro ou nevoeiro glacial, céu visível	} Começou ou tornou-se mais forte durante a hora precedente	
47	Nevoeiro ou nevoeiro glacial, céu invisível		
48	Nevoeiro, depositando escarcha, céu visível		
49	Nevoeiro, depositando escarcha, céu invisível		
<hr/>			
<b>ww = 50-99</b>	<b>Precipitação na estação na hora da observação</b>		
<hr/>			
<b>ww = 50-59</b>	<b>Chuvisco</b>		
<hr/>			
50	Chuvisco, sem congelação, intermitente	} Leve na hora da observação	
51	Chuvisco, sem congelação, contínuo		
52	Chuvisco, sem congelação, intermitente	} <b>Medio na hora da observação</b>	
53	Chuvisco, sem congelação, contínuo		
54	Chuvisco, sem congelação, intermitente	} Forte (denso) na hora da observação	
55	Chuvisco, sem congelação, contínuo		
56	Chuvisco, com congelação, leve		
57	Chuvisco, com congelação, moderado ou forte (denso)		
58	Chuvisco e chuva, leve		
59	<b>Chuvisco e chuva, moderada ou forte</b>		
<hr/>			



**Tabela A.3 – Continuação.**

1º JUL 2001		MCA 105-10
<b>ww = 60-69      Chuva</b>		
60	Chuva, sem congelação, intermitente	} Leve na hora da observação
61	Chuva, sem congelação, contínua	
62	Chuva, sem congelação, intermitente	} <b>ada na hora da observação</b>
63	Chuva, sem congelação, contínua	
64	Chuva, sem congelação, intermitente	} Forte na hora da observação
65	Chuva, sem congelação, contínua	
66	Chuva, com congelação, leve	
67	Chuva, com congelação, moderada ou forte	
68	Chuva ou chuveiro e neve, leve	
69	<b>Chuva ou chuveiro e neve, moderado ou forte</b>	
<b>ww = 70-79      Precipitação sólida, não em pancadas</b>		
70	Queda de flocos de neve, intermitente	} Leve na hora da observação
71	Queda de flocos de neve, contínua	
72	Queda de flocos de neve, intermitente	} <b>ada na hora da observação</b>
73	Queda de flocos de neve, contínua	
74	Queda de flocos de neve, intermitente	} Forte na hora da observação
75	Queda de flocos de neve, contínua	
76	Aglúlas de gelo (com ou sem nevoeiro)	
77	Grãos de neve (com ou sem nevoeiro)	
78	Cristais de neve estrelados, isolados (com ou sem nevoeiro)	
79	Grãos de gelo	
<b>ww = 80-99      Precipitação em forma de pancadas, ou precipitação com trovoadas ou logo depois de uma trovoadas</b>		
80	Pancada(s) de chuva, leve(s)	
81	Pancada(s) de chuva, moderada(s) ou forte(s)	
82	Pancada(s) de chuva, muito forte(s)	
83	Pancada(s) de chuva e neve misturadas, leve(s)	

Tabela A.3 – Continuação.

1º JUL 2001		MCA 905-10
84	Pancada(s) de chuva e neve misturadas, moderada(s) ou forte(s)	
85	Pancada(s) de neve, leve(s)	
86	Pancada(s) de neve, moderada(s) ou forte(s)	
87	} Pancada(s) de neve granulada ou de granizo pequeno, com ou sem chuva ou chuva e neve misturadas	} -leve -moderada ou forte
88		
89	} Pancada(s) de granizo, com ou sem chuva ou chuva e neve misturadas, não associada(s) com trovão	} -leve -moderada ou forte
90		
91	<b>Chuva leve na hora da observação</b>	
92	Chuva moderada ou forte na hora da observação	
93	Queda leve de neve, ou chuva e neve misturadas ou granizo*, na hora da observação	} Trovoada durante a hora precedente, mas não na hora da observação
94	Queda moderada ou forte de neve ou chuva e neve misturadas ou granizo*, na hora da observação	
95	Trovoada, leve ou moderada, sem granizo*, mas com chuva e/ou neve na hora da observação	} Trovoada na hora da observação
96	Trovoada leve ou moderada, com granizo*, na hora da observação	
97	Trovoada, forte, sem granizo*, mas com chuva e/ou neve na hora da observação	
98	Trovoada combinada com tempestade de poeira ou de areia, na hora da observação	
99	Trovoada, forte, com granizo*, na hora da observação	
* Granizo, granizo pequeno, neve granulada		

## APÊNDICE B - DIA JULIANO.

Na Tabela B.1 é apresentado um quadro mostrando a relação entre dia do calendário Gregoriano e o dia do calendário Juliano. As células coloridas indicam o início/término das estações meteorológicas.

**Tabela B.1** - Quadro mostrando a relação entre dia gregoriano e dia Juliano. As células coloridas indicam o início/término das estações meteorológicas.

	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1	1	32	60	91	121	152	182	213	243	274	305	335
2	2	33	61	92	122	153	183	214	244	275	306	336
3	3	34	62	93	123	154	184	215	245	276	307	337
4	4	35	63	94	124	155	185	216	246	277	308	338
5	5	36	64	95	125	156	186	217	247	278	309	339
6	6	37	65	96	126	157	187	218	248	279	310	340
7	7	38	66	97	127	158	188	219	249	280	311	341
8	8	39	67	98	128	159	189	220	250	281	312	342
9	9	40	68	99	129	160	190	221	251	282	313	343
10	10	41	69	100	130	161	191	222	252	283	314	344
11	11	42	70	101	131	162	192	223	253	284	315	345
12	12	43	71	102	132	163	193	224	254	285	316	346
13	13	44	72	103	133	164	194	225	255	286	317	347
14	14	45	73	104	134	165	195	226	256	287	318	348
15	15	46	74	105	135	166	196	227	257	288	319	349
16	16	47	75	106	136	167	197	228	258	289	320	350
17	17	48	76	107	137	168	198	229	259	290	321	351
18	18	49	77	108	138	169	199	230	260	291	322	352
19	19	50	78	109	139	170	200	231	261	292	323	353
20	20	51	79	110	140	171	201	232	262	293	324	354
21	21	52	80	111	141	172	202	233	263	294	325	355
22	22	53	81	112	142	173	203	234	264	295	326	356
23	23	54	82	113	143	174	204	235	265	296	327	357
24	24	55	83	114	144	175	205	236	266	297	328	358
25	25	56	84	115	145	176	206	237	267	298	329	359
26	26	57	85	116	146	177	207	238	268	299	330	360
27	27	58	86	117	147	178	208	239	269	300	331	361
28	28	59	87	118	148	179	209	240	270	301	332	362
29	29		88	119	149	180	210	241	271	302	333	363
30	30		89	120	150	181	211	242	272	303	334	364
31	31		90		151		212	243		304		365

## APÊNDICE C – ESTATÍSTICAS DA RLM.

A seguir são apresentadas as estatísticas relativas aos ajustes das 24 funções de Regressão Linear Múltipla (oitos conjuntos de ajustamentos TESTE00, TESTE01, TESTE03, TESTE04, TESTE05, TESTE06 e TESTE07 em cada uma das três horas futuras). As variáveis dependentes são as visibilidades nas três horas futuras: VISIB1, VISIB2 e VISIB3.

### A) PRIMEIRA HORA –VISIB1

**Tabela C.1** - Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE00.

TESTE00								
<i>Estatística de regressão</i>								
R múltiplo	0,84							
R-Quadrado	0,70							
R-quadrado ajustado	0,70							
Erro padrão	2703,17							
Observações	27401							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-42305,37	5510,03	-7,68	0,00	-53105,30	-31505,43	-53105,30	-31505,43
DVISIB_24h	0,03	0,00	8,76	0,00	0,02	0,03	0,02	0,03
DUR_24h	21,33	2,00	10,64	0,00	17,40	25,26	17,40	25,26
DVISIB_1h	-0,09	0,01	-15,41	0,00	-0,10	-0,08	-0,10	-0,08
DTEMPAR_1h	-325,70	17,83	-18,26	0,00	-360,65	-290,74	-360,65	-290,74
DQNH_1h	-7,61	28,69	-0,27	0,79	-63,85	48,63	-63,85	48,63
DIRVENTO	28,89	2,50	11,54	0,00	23,99	33,80	23,99	33,80
VELVENTO	48,66	4,46	10,91	0,00	39,91	57,40	39,91	57,40
VISIB	0,66	0,00	136,57	0,00	0,65	0,66	0,65	0,66
Q1	-49,10	4,92	-9,97	0,00	-58,75	-39,45	-58,75	-39,45
TEMPAR	94,63	4,75	19,93	0,00	85,32	103,94	85,32	103,94
QNH	41,36	5,36	7,71	0,00	30,84	51,87	30,84	51,87
Diajul_m	-1,39	0,16	-8,75	0,00	-1,70	-1,08	-1,70	-1,08
Hora_m	-47,56	2,48	-19,17	0,00	-52,42	-42,70	-52,42	-42,70

**Tabela C.2** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE01.

TESTE01								
<i>Estatística de regressão</i>								
R múltiplo	0,86							
R-Quadrado	0,75							
R-quadrado ajustado	0,75							
Erro padrão	2918,45							
Observações	61714							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-14308,41	3863,45	-3,70	0,00	-21880,79	-6736,03	-21880,79	-6736,03
DVISIB_24h	0,03	0,00	13,07	0,00	0,03	0,04	0,03	0,04
DUR_24h	5,19	1,08	4,80	0,00	3,07	7,31	3,07	7,31
DVISIB_1h	-0,07	0,00	-16,50	0,00	-0,08	-0,06	-0,08	-0,06
DTEMPAR_1h	-209,36	10,94	-19,14	0,00	-230,80	-187,92	-230,80	-187,92
DQNH_1h	-223,22	19,27	-11,58	0,00	-260,99	-185,45	-260,99	-185,45
DIRVENTO	3,14	1,43	2,20	0,03	0,35	5,94	0,35	5,94
VELVENTO	50,10	2,23	22,47	0,00	45,73	54,47	45,73	54,47
VISIB	0,85	0,00	268,37	0,00	0,84	0,85	0,84	0,85
Q1	-72,64	5,11	-14,21	0,00	-82,66	-62,62	-82,66	-62,62
TEMPAR	-0,52	3,54	-0,15	0,88	-7,45	6,41	-7,45	6,41
QNH	15,86	3,75	4,22	0,00	8,50	23,22	8,50	23,22
Diajul_m	-1,37	0,11	-11,99	0,00	-1,60	-1,15	-1,60	-1,15
Hora_m	-60,69	1,96	-31,00	0,00	-64,53	-56,86	-64,53	-56,86

**Tabela C.3** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE02.

TESTE02								
<i>Estatística de regressão</i>								
R múltiplo	0,87							
R-Quadrado	0,76							
R-quadrado ajustado	0,76							
Erro padrão	2914,80							
Observações	40195							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-23415,05	4773,41	-4,91	0,00	-32771,04	-14059,06	-32771,04	-14059,06
DVISIB_24h	-0,08	0,00	-31,92	0,00	-0,09	-0,08	-0,09	-0,08
DUR_24h	7,07	1,39	5,08	0,00	4,34	9,81	4,34	9,81
DVISIB_1h	-0,08	0,00	-16,64	0,00	-0,09	-0,07	-0,09	-0,07
DTEMPAR_1h	-218,49	13,28	-16,45	0,00	-244,52	-192,45	-244,52	-192,45
DQNH_1h	158,91	23,38	6,80	0,00	113,08	204,74	113,08	204,74
DIRVENTO	9,19	1,79	5,13	0,00	5,68	12,70	5,68	12,70
VELVENTO	57,93	3,02	19,18	0,00	52,01	63,85	52,01	63,85
VISIB	0,75	0,00	217,61	0,00	0,75	0,76	0,75	0,76
Q1	-53,24	5,49	-9,71	0,00	-63,99	-42,49	-63,99	-42,49
TEMPAR	49,91	4,24	11,76	0,00	41,59	58,22	41,59	58,22
QNH	24,52	4,64	5,28	0,00	15,42	33,61	15,42	33,61
Diajul_m	-1,55	0,14	-10,95	0,00	-1,82	-1,27	-1,82	-1,27
Hora_m	-65,97	2,18	-30,27	0,00	-70,24	-61,70	-70,24	-61,70

**Tabela C.4** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE03.

TESTE03								
<i>Estatística de regressão</i>								
R múltiplo	0,87							
R-Quadrado	0,76							
R-quadrado ajustado	0,76							
Erro padrão	2771,96							
Observações	35184							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-34365,07	4911,53	-7,00	0,00	-43991,83	-24738,32	-43991,83	-24738,32
DVISIB_24h	-0,07	0,00	-26,90	0,00	-0,08	-0,07	-0,08	-0,07
DUR_24h	12,46	1,64	7,60	0,00	9,24	15,67	9,24	15,67
DVISIB_1h	-0,12	0,01	-23,90	0,00	-0,13	-0,11	-0,13	-0,11
DTEMPAR_1h	-274,72	14,82	-18,54	0,00	-303,77	-245,67	-303,77	-245,67
DQNH_1h	85,82	25,16	3,41	0,00	36,50	135,14	36,50	135,14
DIRVENTO	20,65	2,07	9,98	0,00	16,60	24,71	16,60	24,71
VELVENTO	58,99	3,89	15,16	0,00	51,36	66,62	51,36	66,62
VISIB	0,69	0,00	180,76	0,00	0,68	0,70	0,68	0,70
Q1	-42,11	4,65	-9,05	0,00	-51,23	-32,99	-51,23	-32,99
TEMPAR	86,42	4,22	20,48	0,00	78,15	94,69	78,15	94,69
QNH	34,50	4,78	7,22	0,00	25,13	43,86	25,13	43,86
Diajul_m	-1,65	0,14	-11,57	0,00	-1,93	-1,37	-1,93	-1,37
Hora_m	-61,87	2,18	-28,41	0,00	-66,14	-57,60	-66,14	-57,60

**Tabela C.5** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE04.

TESTE04								
<i>Estatística de regressão</i>								
R múltiplo	0,87							
R-Quadrado	0,75							
R-quadrado ajustado	0,75							
Erro padrão	2687,56							
Observações	25643							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-27135,88	5656,19	-4,80	0,00	-38222,33	-16049,43	-38222,33	-16049,43
DVISIB_24h	-0,05	0,00	-17,17	0,00	-0,06	-0,05	-0,06	-0,05
DUR_24h	10,10	2,02	4,99	0,00	6,13	14,07	6,13	14,07
DVISIB_1h	-0,16	0,01	-27,74	0,00	-0,17	-0,15	-0,17	-0,15
DTEMPAR_1h	-271,76	17,54	-15,49	0,00	-306,14	-237,37	-306,14	-237,37
DQNH_1h	-1,68	28,21	-0,06	0,95	-56,97	53,61	-56,97	53,61
DIRVENTO	24,23	2,49	9,74	0,00	19,36	29,11	19,36	29,11
VELVENTO	52,77	4,58	11,51	0,00	43,79	61,76	43,79	61,76
VISIB	0,66	0,00	147,75	0,00	0,65	0,67	0,65	0,67
Q1	-36,90	5,08	-7,27	0,00	-46,85	-26,96	-46,85	-26,96
TEMPAR	98,10	4,82	20,35	0,00	88,64	107,55	88,64	107,55
QNH	27,13	5,51	4,93	0,00	16,34	37,92	16,34	37,92
Diajul_m	-1,71	0,16	-10,38	0,00	-2,03	-1,39	-2,03	-1,39
Hora_m	-59,41	2,49	-23,83	0,00	-64,30	-54,53	-64,30	-54,53

**Tabela C.6** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE05.

TESTE05								
<i>Estatística de regressão</i>								
R múltiplo	0,88							
R-Quadrado	0,78							
R-quadrado ajustado	0,77							
Erro padrão	2684,41							
Observações	23289							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-26438,06	5914,52	-4,47	0,00	-38030,91	-14845,21	-38030,91	-14845,21
DVISIB_24h	-0,05	0,00	-14,59	0,00	-0,05	-0,04	-0,05	-0,04
DUR_24h	7,54	2,15	3,50	0,00	3,32	11,76	3,32	11,76
DVISIB_1h	-0,19	0,01	-31,77	0,00	-0,20	-0,18	-0,20	-0,18
DTEMPAR_1h	-250,23	18,65	-13,42	0,00	-286,78	-213,68	-286,78	-213,68
DQNH_1h	-67,35	29,81	-2,26	0,02	-125,78	-8,93	-125,78	-8,93
DIRVENTO	28,22	2,61	10,83	0,00	23,11	33,33	23,11	33,33
VELVENTO	51,91	4,77	10,88	0,00	42,56	61,26	42,56	61,26
VISIB	0,68	0,00	144,16	0,00	0,67	0,69	0,67	0,69
Q1	-36,40	5,35	-6,80	0,00	-46,89	-25,91	-46,89	-25,91
TEMPAR	102,31	5,03	20,34	0,00	92,46	112,17	92,46	112,17
QNH	26,41	5,76	4,59	0,00	15,13	37,69	15,13	37,69
Diajul_m	-1,71	0,17	-9,89	0,00	-2,05	-1,37	-2,05	-1,37
Hora_m	-62,13	2,66	-23,34	0,00	-67,35	-56,91	-67,35	-56,91

**Tabela C.7** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE06.

TESTE06								
<i>Estatística de regressão</i>								
R múltiplo	0,88							
R-Quadrado	0,78							
R-quadrado ajustado	0,78							
Erro padrão	2642,12							
Observações	20891							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-30461,25	6180,75	-4,93	0,00	-42576,01	-18346,50	-42576,01	-18346,50
DVISIB_24h	-0,04	0,00	-12,58	0,00	-0,05	-0,03	-0,05	-0,03
DUR_24h	7,16	2,33	3,07	0,00	2,58	11,73	2,58	11,73
DVISIB_1h	-0,19	0,01	-31,42	0,00	-0,21	-0,18	-0,21	-0,18
DTEMPAR_1h	-255,47	19,89	-12,84	0,00	-294,46	-216,49	-294,46	-216,49
DQNH_1h	-85,49	31,29	-2,73	0,01	-146,82	-24,15	-146,82	-24,15
DIRVENTO	33,54	2,78	12,06	0,00	28,09	38,99	28,09	38,99
VELVENTO	56,58	5,05	11,21	0,00	46,69	66,48	46,69	66,48
VISIB	0,66	0,00	133,38	0,00	0,65	0,67	0,65	0,67
Q1	-45,85	5,49	-8,35	0,00	-56,61	-35,08	-56,61	-35,08
TEMPAR	113,59	5,25	21,65	0,00	103,31	123,88	103,31	123,88
QNH	30,21	6,02	5,02	0,00	18,42	42,01	18,42	42,01
Diajul_m	-1,60	0,18	-8,77	0,00	-1,96	-1,24	-1,96	-1,24
Hora_m	-62,42	2,79	-22,34	0,00	-67,89	-56,94	-67,89	-56,94

**Tabela C. 8** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB1, feita com o conjunto de treinamento TESTE07.

TESTE07								
<i>Estatística de regressão</i>								
R múltiplo	0,88							
R-Quadrado	0,77							
R-quadrado ajustado	0,77							
Erro padrão	2634,34							
Observações	20420							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-28571,51	6233,56	-4,58	0,00	-40789,78	-16353,24	-40789,78	-16353,24
DVISIB_24h	-0,04	0,00	-12,07	0,00	-0,05	-0,03	-0,05	-0,03
DUR_24h	7,66	2,37	3,24	0,00	3,03	12,30	3,03	12,30
DVISIB_1h	-0,19	0,01	-30,17	0,00	-0,20	-0,17	-0,20	-0,17
DTEMPAR_1h	-238,55	20,18	-11,82	0,00	-278,11	-199,00	-278,11	-199,00
DQNH_1h	-89,46	31,50	-2,84	0,00	-151,20	-27,73	-151,20	-27,73
DIRVENTO	34,84	2,82	12,36	0,00	29,31	40,36	29,31	40,36
VELVENTO	55,42	5,11	10,85	0,00	45,40	65,43	45,40	65,43
VISIB	0,66	0,00	132,04	0,00	0,65	0,67	0,65	0,67
Q1	-45,52	5,52	-8,24	0,00	-56,35	-34,69	-56,35	-34,69
TEMPAR	114,44	5,29	21,64	0,00	104,08	124,81	104,08	124,81
QNH	28,34	6,07	4,67	0,00	16,45	40,24	16,45	40,24
Diajul_m	-1,59	0,18	-8,66	0,00	-1,95	-1,23	-1,95	-1,23
Hora_m	-63,12	2,80	-22,50	0,00	-68,61	-57,62	-68,61	-57,62

## B) SEGUNDA HORA – VISIB2

**Tabela C.9** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE00.

TESTE00								
<i>Estatística de regressão</i>								
R múltiplo	0,79							
R-Quadrado	0,63							
R-quadrado ajustado	0,63							
Erro padrão	3528,95							
Observações	57627							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	65346,95	4748,45	13,76	0,00	56039,96	74653,93	56039,96	74653,93
DVISIB_24h	-0,13	0,00	-51,53	0,00	-0,14	-0,13	-0,14	-0,13
DUR_24h	1,29	1,28	1,01	0,31	-1,21	3,80	-1,21	3,80
DVISIB_1h	-0,15	0,01	-28,80	0,00	-0,16	-0,14	-0,16	-0,14
DTEMPAR_1h	-352,08	12,18	-28,91	0,00	-375,95	-328,21	-375,95	-328,21
DQNH_1h	6,84	23,18	0,29	0,77	-38,60	52,27	-38,60	52,27
DIRVENTO	-21,10	1,65	-12,78	0,00	-24,33	-17,86	-24,33	-17,86
VELVENTO	86,78	3,12	27,86	0,00	80,68	92,89	80,68	92,89
VISIB	0,64	0,00	210,51	0,00	0,63	0,64	0,63	0,64
Q1	-45,82	7,49	-6,12	0,00	-60,49	-31,14	-60,49	-31,14
TEMPAR	-19,38	4,15	-4,67	0,00	-27,52	-11,24	-27,52	-11,24
QNH	-58,93	4,61	-12,77	0,00	-67,98	-49,89	-67,98	-49,89
Diajul_m	-0,08	0,15	-0,55	0,58	-0,37	0,21	-0,37	0,21
Hora_m	-140,63	2,43	-57,77	0,00	-145,40	-135,86	-145,40	-135,86

**Tabela C.10** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE01.

TESTE01								
<i>Estatística de regressão</i>								
R múltiplo	0,81							
R-Quadrado	0,65							
R-quadrado ajustado	0,65							
Erro padrão	3598,82							
Observações	61335							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	37288,25	4706,31	7,92	0,00	28063,87	46512,63	28063,87	46512,63
DVISIB_24h	-0,14	0,00	-54,25	0,00	-0,14	-0,13	-0,14	-0,13
DUR_24h	1,14	1,31	0,87	0,38	-1,42	3,70	-1,42	3,70
DVISIB_1h	-0,16	0,01	-29,68	0,00	-0,17	-0,15	-0,17	-0,15
DTEMPAR_1h	-430,99	12,35	-34,90	0,00	-455,20	-406,78	-455,20	-406,78
DQNH_1h	-65,97	23,07	-2,86	0,00	-111,20	-20,74	-111,20	-20,74
DIRVENTO	-5,90	1,70	-3,46	0,00	-9,24	-2,56	-9,24	-2,56
VELVENTO	67,51	3,31	20,37	0,00	61,01	74,00	61,01	74,00
VISIB	0,64	0,00	209,84	0,00	0,63	0,64	0,63	0,64
Q1	-57,27	6,41	-8,94	0,00	-69,83	-44,71	-69,83	-44,71
TEMPAR	8,94	4,08	2,19	0,03	0,93	16,94	0,93	16,94
QNH	-32,17	4,57	-7,03	0,00	-41,14	-23,21	-41,14	-23,21
Diajul_m	-0,51	0,14	-3,61	0,00	-0,79	-0,23	-0,79	-0,23
Hora_m	-150,82	2,38	-63,44	0,00	-155,47	-146,16	-155,47	-146,16



**Tabela C.11** - Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE02.

TESTE02								
<i>Estatística de regressão</i>								
R múltiplo	0,80							
R-Quadrado	0,63							
R-quadrado ajustado	0,63							
Erro padrão	3575,37							
Observações	41945							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	16819,03	5689,40	2,96	0,00	5667,69	27970,37	5667,69	27970,37
DVISIB_24h	-0,11	0,00	-35,42	0,00	-0,11	-0,10	-0,11	-0,10
DUR_24h	5,87	1,65	3,55	0,00	2,63	9,12	2,63	9,12
DVISIB_1h	-0,22	0,01	-34,13	0,00	-0,23	-0,20	-0,23	-0,20
DTEMPAR_1h	-438,77	15,49	-28,33	0,00	-469,12	-408,41	-469,12	-408,41
DQNH_1h	-175,41	29,08	-6,03	0,00	-232,40	-118,42	-232,40	-118,42
DIRVENTO	4,19	2,22	1,89	0,06	-0,15	8,54	-0,15	8,54
VELVENTO	73,48	4,36	16,87	0,00	64,94	82,02	64,94	82,02
VISIB	0,58	0,00	152,55	0,00	0,57	0,59	0,57	0,59
Q1	-53,06	6,88	-7,71	0,00	-66,54	-39,57	-66,54	-39,57
TEMPAR	40,84	4,96	8,23	0,00	31,11	50,56	31,11	50,56
QNH	-12,78	5,53	-2,31	0,02	-23,62	-1,94	-23,62	-1,94
Diajul_m	-1,66	0,17	-9,70	0,00	-2,00	-1,33	-2,00	-1,33
Hora_m	-142,24	2,88	-49,35	0,00	-147,89	-136,59	-147,89	-136,59

**Tabela C. 12** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE03.

TESTE03								
<i>Estatística de regressão</i>								
R múltiplo	0,78							
R-Quadrado	0,61							
R-quadrado ajustado	0,61							
Erro padrão	3491,85							
Observações	34754							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	5277,46	6086,46	0,87	0,39	-6652,20	17207,11	-6652,20	17207,11
DVISIB_24h	-0,09	0,00	-26,27	0,00	-0,09	-0,08	-0,09	-0,08
DUR_24h	11,05	2,05	5,39	0,00	7,03	15,07	7,03	15,07
DVISIB_1h	-0,21	0,01	-31,76	0,00	-0,22	-0,20	-0,22	-0,20
DTEMPAR_1h	-510,18	18,64	-27,37	0,00	-546,71	-473,64	-546,71	-473,64
DQNH_1h	-291,90	32,63	-8,95	0,00	-355,85	-227,95	-355,85	-227,95
DIRVENTO	18,83	2,63	7,15	0,00	13,66	23,99	13,66	23,99
VELVENTO	127,25	4,81	26,45	0,00	117,82	136,69	117,82	136,69
VISIB	0,47	0,00	107,53	0,00	0,46	0,47	0,46	0,47
Q1	-48,27	6,24	-7,73	0,00	-60,51	-36,04	-60,51	-36,04
TEMPAR	79,69	5,41	14,74	0,00	69,09	90,29	69,09	90,29
QNH	-2,72	5,92	-0,46	0,65	-14,31	8,88	-14,31	8,88
Diajul_m	-1,46	0,18	-8,02	0,00	-1,81	-1,10	-1,81	-1,10
Hora_m	-133,55	3,23	-41,34	0,00	-139,89	-127,22	-139,89	-127,22

**Tabela C. 13** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE04.

TESTE04								
<i>Estatística de regressão</i>								
R múltiplo	0,77							
R-Quadrado	0,60							
R-quadrado ajustado	0,60							
Erro padrão	3355,45							
Observações	25563							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-15868,61	6950,31	-2,28	0,02	-29491,60	-2245,61	-29491,60	-2245,61
DVISIB_24h	-0,06	0,00	-16,99	0,00	-0,07	-0,05	-0,07	-0,05
DUR_24h	16,69	2,57	6,50	0,00	11,66	21,73	11,66	21,73
DVISIB_1h	-0,20	0,01	-27,33	0,00	-0,22	-0,19	-0,22	-0,19
DTEMPAR_1h	-536,05	22,40	-23,94	0,00	-579,94	-492,15	-579,94	-492,15
DQNH_1h	-287,27	36,75	-7,82	0,00	-359,30	-215,25	-359,30	-215,25
DIRVENTO	27,24	3,14	8,69	0,00	21,10	33,39	21,10	33,39
VELVENTO	125,24	5,86	21,36	0,00	113,75	136,73	113,75	136,73
VISIB	0,42	0,01	83,30	0,00	0,41	0,43	0,41	0,43
Q1	-56,01	6,72	-8,34	0,00	-69,18	-42,84	-69,18	-42,84
TEMPAR	108,10	6,11	17,70	0,00	96,13	120,07	96,13	120,07
QNH	17,66	6,76	2,61	0,01	4,41	30,90	4,41	30,90
Diajul_m	-1,25	0,21	-5,95	0,00	-1,66	-0,84	-1,66	-0,84
Hora_m	-139,61	3,80	-36,73	0,00	-147,06	-132,16	-147,06	-132,16

**Tabela C. 14** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE05.

TESTE05								
<i>Estatística de regressão</i>								
R múltiplo	0,79							
R-Quadrado	0,62							
R-quadrado ajustado	0,62							
Erro padrão	3396,70							
Observações	22920							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-4048,05	7454,12	-0,54	0,59	-18658,63	10562,53	-18658,63	10562,53
DVISIB_24h	-0,06	0,00	-15,42	0,00	-0,07	-0,05	-0,07	-0,05
DUR_24h	11,88	2,75	4,32	0,00	6,50	17,27	6,50	17,27
DVISIB_1h	-0,20	0,01	-26,04	0,00	-0,22	-0,19	-0,22	-0,19
DTEMPAR_1h	-510,29	24,25	-21,05	0,00	-557,82	-462,77	-557,82	-462,77
DQNH_1h	-392,90	39,13	-10,04	0,00	-469,60	-316,20	-469,60	-316,20
DIRVENTO	27,80	3,36	8,27	0,00	21,21	34,39	21,21	34,39
VELVENTO	135,73	6,35	21,36	0,00	123,28	148,19	123,28	148,19
VISIB	0,44	0,01	81,16	0,00	0,43	0,45	0,43	0,45
Q1	-54,24	7,24	-7,49	0,00	-68,43	-40,05	-68,43	-40,05
TEMPAR	103,84	6,56	15,82	0,00	90,98	116,71	90,98	116,71
QNH	6,31	7,25	0,87	0,38	-7,89	20,51	-7,89	20,51
Diajul_m	-1,28	0,22	-5,72	0,00	-1,72	-0,84	-1,72	-0,84
Hora_m	-152,71	4,22	-36,17	0,00	-160,99	-144,44	-160,99	-144,44

**Tabela C. 15** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE06.

TESTE06								
<i>Estatística de regressão</i>								
R múltiplo	0,79							
R-Quadrado	0,62							
R-quadrado ajustado	0,62							
Erro padrão	3323,65							
Observações	20040							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-5809,01	7800,76	-0,74	0,46	-21099,13	9481,12	-21099,13	9481,12
DVISIB_24h	-0,04	0,00	-10,43	0,00	-0,05	-0,03	-0,05	-0,03
DUR_24h	21,27	3,04	6,99	0,00	15,30	27,23	15,30	27,23
DVISIB_1h	-0,20	0,01	-25,17	0,00	-0,22	-0,19	-0,22	-0,19
DTEMPAR_1h	-489,97	26,06	-18,80	0,00	-541,05	-438,89	-541,05	-438,89
DQNH_1h	-363,99	41,16	-8,84	0,00	-444,66	-283,32	-444,66	-283,32
DIRVENTO	40,02	3,66	10,94	0,00	32,84	47,19	32,84	47,19
VELVENTO	121,41	7,01	17,33	0,00	107,68	135,15	107,68	135,15
VISIB	0,41	0,01	71,96	0,00	0,40	0,42	0,40	0,42
Q1	-56,73	7,44	-7,63	0,00	-71,31	-42,15	-71,31	-42,15
TEMPAR	113,78	6,96	16,34	0,00	100,13	127,42	100,13	127,42
QNH	8,18	7,58	1,08	0,28	-6,68	23,04	-6,68	23,04
Diajul_m	-1,68	0,23	-7,16	0,00	-2,14	-1,22	-2,14	-1,22
Hora_m	-165,40	4,53	-36,50	0,00	-174,28	-156,52	-174,28	-156,52

**Tabela C. 16** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB2, feita com o conjunto de treinamento TESTE07.

TESTE07								
<i>Estatística de regressão</i>								
R múltiplo	0,79							
R-Quadrado	0,62							
R-quadrado ajustado	0,62							
Erro padrão	3327,24							
Observações	19792							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	-10043,18	7871,08	-1,28	0,20	-25471,16	5384,81	-25471,16	5384,81
DVISIB_24h	-0,04	0,00	-10,16	0,00	-0,05	-0,03	-0,05	-0,03
DUR_24h	25,19	3,10	8,14	0,00	19,12	31,26	19,12	31,26
DVISIB_1h	-0,21	0,01	-25,52	0,00	-0,22	-0,19	-0,22	-0,19
DTEMPAR_1h	-499,87	26,59	-18,80	0,00	-551,99	-447,75	-551,99	-447,75
DQNH_1h	-353,57	40,81	-8,66	0,00	-433,57	-273,57	-433,57	-273,57
DIRVENTO	38,54	3,70	10,40	0,00	31,28	45,80	31,28	45,80
VELVENTO	118,30	7,05	16,79	0,00	104,49	132,11	104,49	132,11
VISIB	0,41	0,01	70,66	0,00	0,40	0,42	0,40	0,42
Q1	-56,38	7,46	-7,55	0,00	-71,01	-41,75	-71,01	-41,75
TEMPAR	117,66	7,05	16,70	0,00	103,84	131,47	103,84	131,47
QNH	12,36	7,65	1,62	0,11	-2,62	27,35	-2,62	27,35
Diajul_m	-1,70	0,24	-7,20	0,00	-2,17	-1,24	-2,17	-1,24
Hora_m	-169,68	4,57	-37,14	0,00	-178,63	-160,72	-178,63	-160,72

### C) TERCEIRA HORA – VISIB3

**Tabela C. 17** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE00.

TESTE00								
<i>Estatística de regressão</i>								
R múltiplo	0,55							
R-Quadrado	0,30							
R-quadrado ajustado	0,30							
Erro padrão	4687,90							
Observações	57591							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	67423,10	6528,30	10,33	0,00	54627,60	80218,60	54627,60	80218,60
DVISIB_24h	0,00	0,00	1,25	0,21	0,00	0,01	0,00	0,01
DUR_24h	5,69	1,76	3,22	0,00	2,23	9,14	2,23	9,14
DVISIB_1h	-0,07	0,01	-10,06	0,00	-0,08	-0,06	-0,08	-0,06
DTEMPAR_1h	-533,34	17,67	-30,18	0,00	-567,98	-498,70	-567,98	-498,70
DQNH_1h	-1150,14	30,56	-37,64	0,00	-1210,03	-1090,24	-1210,03	-1090,24
DIRVENTO	31,97	2,35	13,60	0,00	27,36	36,57	27,36	36,57
VELVENTO	40,54	3,10	13,06	0,00	34,45	46,62	34,45	46,62
VISIB	0,44	0,01	88,02	0,00	0,43	0,45	0,43	0,45
Q1	-90,14	9,82	-9,18	0,00	-109,38	-70,89	-109,38	-70,89
TEMPAR	-41,99	5,80	-7,24	0,00	-53,36	-30,62	-53,36	-30,62
QNH	-59,01	6,35	-9,30	0,00	-71,45	-46,57	-71,45	-46,57
Diajul_m	-0,92	0,19	-4,72	0,00	-1,30	-0,54	-1,30	-0,54
Hora_m	-156,98	3,49	-45,04	0,00	-163,81	-150,14	-163,81	-150,14

**Tabela C. 18** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE01.

TESTE01								
<i>Estatística de regressão</i>								
R múltiplo	0,61							
R-Quadrado	0,37							
R-quadrado ajustado	0,37							
Erro padrão	4739,79							
Observações	61400							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	58338,28	6414,40	9,09	0,00	45766,05	70910,51	45766,05	70910,51
DVISIB_24h	0,02	0,00	4,52	0,00	0,01	0,02	0,01	0,02
DUR_24h	7,75	1,78	4,35	0,00	4,26	11,24	4,26	11,24
DVISIB_1h	-0,06	0,01	-9,33	0,00	-0,08	-0,05	-0,08	-0,05
DTEMPAR_1h	-548,19	17,79	-30,82	0,00	-583,05	-513,33	-583,05	-513,33
DQNH_1h	-1223,34	29,71	-41,18	0,00	-1281,56	-1165,11	-1281,56	-1165,11
DIRVENTO	37,17	2,36	15,75	0,00	32,55	41,80	32,55	41,80
VELVENTO	50,36	3,08	16,34	0,00	44,32	56,39	44,32	56,39
VISIB	0,47	0,00	96,63	0,00	0,46	0,48	0,46	0,48
Q1	-100,29	8,59	-11,67	0,00	-117,13	-83,45	-117,13	-83,45
TEMPAR	-26,14	5,70	-4,59	0,00	-37,31	-14,98	-37,31	-14,98
QNH	-50,90	6,24	-8,16	0,00	-63,12	-38,68	-63,12	-38,68
Diajul_m	-1,83	0,19	-9,82	0,00	-2,20	-1,47	-2,20	-1,47
Hora_m	-183,89	3,45	-53,37	0,00	-190,65	-177,14	-190,65	-177,14

**Tabela C. 19** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE02.

TESTE02								
<i>Estatística de regressão</i>								
R múltiplo	0,65							
R-Quadrado	0,42							
R-quadrado ajustado	0,42							
Erro padrão	4655,64							
Observações	40633							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	76862,27	7771,83	9,89	0,00	61629,31	92095,24	61629,31	92095,24
DVISIB_24h	0,01	0,00	1,93	0,05	0,00	0,02	0,00	0,02
DUR_24h	11,47	2,20	5,21	0,00	7,16	15,78	7,16	15,78
DVISIB_1h	-0,08	0,01	-10,27	0,00	-0,10	-0,07	-0,10	-0,07
DTEMPAR_1h	-585,85	21,59	-27,13	0,00	-628,16	-543,53	-628,16	-543,53
DQNH_1h	-1055,92	35,06	-30,12	0,00	-1124,64	-987,21	-1124,64	-987,21
DIRVENTO	50,86	2,88	17,64	0,00	45,21	56,52	45,21	56,52
VELVENTO	70,77	3,71	19,06	0,00	63,49	78,05	63,49	78,05
VISIB	0,45	0,01	77,93	0,00	0,44	0,47	0,44	0,47
Q1	-84,44	9,38	-9,00	0,00	-102,83	-66,06	-102,83	-66,06
TEMPAR	-33,00	6,83	-4,83	0,00	-46,39	-19,62	-46,39	-19,62
QNH	-69,20	7,56	-9,16	0,00	-84,02	-54,39	-84,02	-54,39
Diajul_m	-2,54	0,22	-11,29	0,00	-2,98	-2,10	-2,98	-2,10
Hora_m	-202,11	4,19	-48,19	0,00	-210,33	-193,89	-210,33	-193,89

**Tabela C. 20** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE03.

TESTE03								
<i>Estatística de regressão</i>								
R múltiplo	0,69							
R-Quadrado	0,47							
R-quadrado ajustado	0,47							
Erro padrão	4096,19							
Observações	34443							
	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	77990,02	7449,23	10,47	0,00	63389,29	92590,76	63389,29	92590,76
DVISIB_24h	-0,02	0,00	-4,13	0,00	-0,03	-0,01	-0,03	-0,01
DUR_24h	24,69	2,44	10,10	0,00	19,90	29,47	19,90	29,47
DVISIB_1h	-0,08	0,01	-10,10	0,00	-0,10	-0,06	-0,10	-0,06
DTEMPAR_1h	-748,14	21,88	-34,19	0,00	-791,04	-705,25	-791,04	-705,25
DQNH_1h	-704,58	37,30	-18,89	0,00	-777,69	-631,47	-777,69	-631,47
DIRVENTO	44,90	3,03	14,83	0,00	38,96	50,84	38,96	50,84
VELVENTO	163,47	5,34	30,63	0,00	153,01	173,93	153,01	173,93
VISIB	0,30	0,01	55,43	0,00	0,29	0,31	0,29	0,31
Q1	-57,32	7,72	-7,43	0,00	-72,44	-42,19	-72,44	-42,19
TEMPAR	11,08	6,63	1,67	0,09	-1,92	24,08	-1,92	24,08
QNH	-71,51	7,24	-9,88	0,00	-85,70	-57,32	-85,70	-57,32
Diajul_m	-2,38	0,22	-11,04	0,00	-2,81	-1,96	-2,81	-1,96
Hora_m	-206,53	4,44	-46,54	0,00	-215,23	-197,84	-215,23	-197,84

**Tabela C. 21** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE04.

TESTE04								
<i>Estatística de regressão</i>								
R múltiplo	0,70							
R-Quadrado	0,49							
R-quadrado ajustado	0,49							
Erro padrão	3896,20							
Observações	25784							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	59423,73	8267,64	7,19	0,00	43218,69	75628,76	43218,69	75628,76
DVISIB_24h	-0,02	0,00	-3,24	0,00	-0,02	-0,01	-0,02	-0,01
DUR_24h	33,13	2,84	11,69	0,00	27,58	38,69	27,58	38,69
DVISIB_1h	-0,09	0,01	-10,11	0,00	-0,11	-0,07	-0,11	-0,07
DTEMPAR_1h	-733,38	24,88	-29,48	0,00	-782,14	-684,62	-782,14	-684,62
DQNH_1h	-617,14	41,17	-14,99	0,00	-697,83	-536,44	-697,83	-536,44
DIRVENTO	52,27	3,45	15,16	0,00	45,51	59,02	45,51	59,02
VELVENTO	147,72	5,85	25,26	0,00	136,26	159,18	136,26	159,18
VISIB	0,29	0,01	48,12	0,00	0,28	0,30	0,28	0,30
Q1	-34,91	8,37	-4,17	0,00	-51,31	-18,52	-51,31	-18,52
TEMPAR	26,22	7,37	3,56	0,00	11,78	40,66	11,78	40,66
QNH	-53,50	8,03	-6,66	0,00	-69,24	-37,75	-69,24	-37,75
Diajul_m	-3,55	0,24	-14,81	0,00	-4,02	-3,08	-4,02	-3,08
Hora_m	-205,38	5,12	-40,10	0,00	-215,42	-195,34	-215,42	-195,34

**Tabela C. 22** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE05.

TESTE05								
<i>Estatística de regressão</i>								
R múltiplo	0,72							
R-Quadrado	0,52							
R-quadrado ajustado	0,52							
Erro padrão	3874,64							
Observações	23207							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	64150,85	8701,98	7,37	0,00	47094,39	81207,31	47094,39	81207,31
DVISIB_24h	-0,02	0,00	-4,73	0,00	-0,03	-0,01	-0,03	-0,01
DUR_24h	29,59	2,99	9,89	0,00	23,73	35,46	23,73	35,46
DVISIB_1h	-0,09	0,01	-9,41	0,00	-0,10	-0,07	-0,10	-0,07
DTEMPAR_1h	-737,98	26,22	-28,14	0,00	-789,38	-686,59	-789,38	-686,59
DQNH_1h	-598,07	42,37	-14,12	0,00	-681,12	-515,02	-681,12	-515,02
DIRVENTO	55,58	3,65	15,22	0,00	48,42	62,74	48,42	62,74
VELVENTO	150,47	6,23	24,14	0,00	138,25	162,69	138,25	162,69
VISIB	0,29	0,01	45,40	0,00	0,27	0,30	0,27	0,30
Q1	-31,62	8,74	-3,62	0,00	-48,76	-14,49	-48,76	-14,49
TEMPAR	36,47	7,82	4,66	0,00	21,14	51,79	21,14	51,79
QNH	-57,99	8,45	-6,86	0,00	-74,56	-41,42	-74,56	-41,42
Diajul_m	-3,25	0,25	-12,86	0,00	-3,74	-2,75	-3,74	-2,75
Hora_m	-224,28	5,46	-41,09	0,00	-234,97	-213,58	-234,97	-213,58

**Tabela C. 23** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE06.

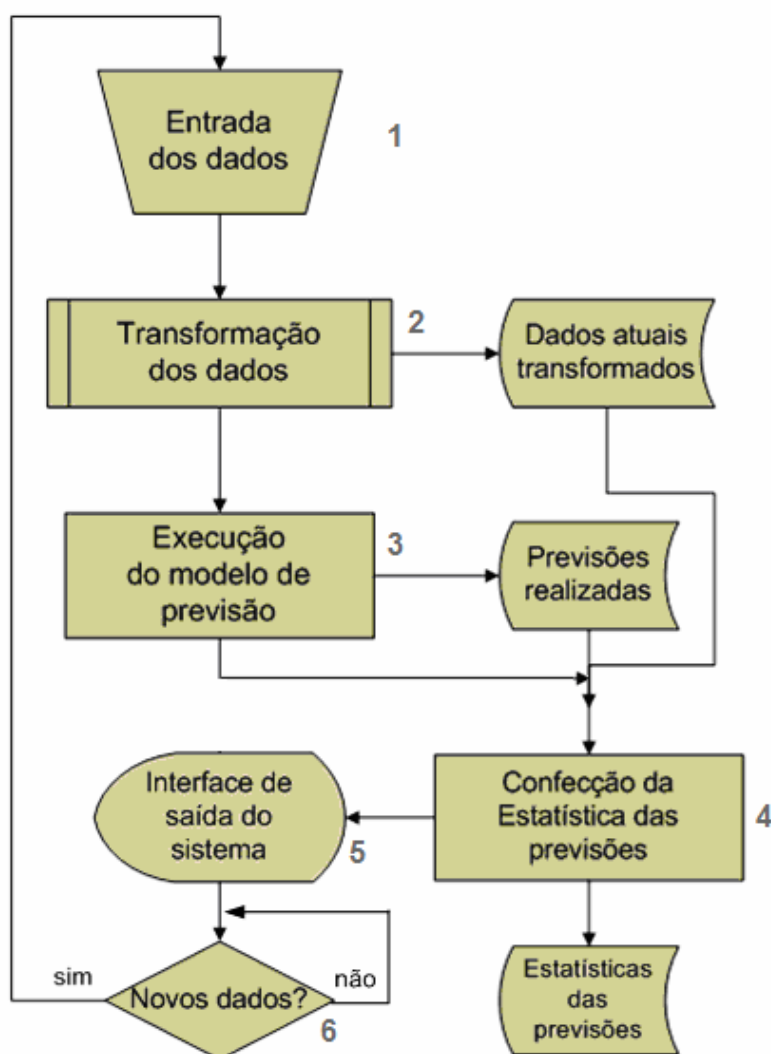
TESTE06								
<i>Estatística de regressão</i>								
R múltiplo	0,73							
R-Quadrado	0,53							
R-quadrado ajustado	0,53							
Erro padrão	3650,57							
Observações	20277							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	68141,80	8854,52	7,70	0,00	50786,22	85497,38	50786,22	85497,38
DVISIB_24h	-0,02	0,00	-4,88	0,00	-0,03	-0,01	-0,03	-0,01
DUR_24h	33,42	3,26	10,25	0,00	27,03	39,81	27,03	39,81
DVISIB_1h	-0,09	0,01	-9,43	0,00	-0,10	-0,07	-0,10	-0,07
DTEMPAR_1h	-778,71	27,94	-27,87	0,00	-833,47	-723,95	-833,47	-723,95
DQNH_1h	-563,96	45,60	-12,37	0,00	-653,34	-474,58	-653,34	-474,58
DIRVENTO	59,45	3,86	15,42	0,00	51,89	67,01	51,89	67,01
VELVENTO	152,73	6,69	22,82	0,00	139,61	165,85	139,61	165,85
VISIB	0,23	0,01	36,00	0,00	0,22	0,24	0,22	0,24
Q1	-42,88	8,66	-4,95	0,00	-59,85	-25,90	-59,85	-25,90
TEMPAR	64,45	8,08	7,98	0,00	48,61	80,29	48,61	80,29
QNH	-61,91	8,60	-7,20	0,00	-78,76	-45,05	-78,76	-45,05
Diajul_m	-3,62	0,26	-13,71	0,00	-4,13	-3,10	-4,13	-3,10
Hora_m	-233,70	5,88	-39,73	0,00	-245,23	-222,17	-245,23	-222,17

**Tabela C. 24** – Estatísticas e os coeficientes da regressão linear múltipla relacionados ao ajuste de regressão para a variável dependente VISIB3, feita com o conjunto de treinamento TESTE07.

TESTE07								
<i>Estatística de regressão</i>								
R múltiplo	0,74							
R-Quadrado	0,54							
R-quadrado ajustado	0,54							
Erro padrão	3648,89							
Observações	20002							
	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95,0%</i>	<i>Superior 95,0%</i>
Interseção	70208,51	8928,03	7,86	0,00	52708,83	87708,19	52708,83	87708,19
DVISIB_24h	-0,03	0,00	-5,12	0,00	-0,03	-0,02	-0,03	-0,02
DUR_24h	32,04	3,28	9,76	0,00	25,60	38,47	25,60	38,47
DVISIB_1h	-0,09	0,01	-9,63	0,00	-0,11	-0,07	-0,11	-0,07
DTEMPAR_1h	-780,79	28,12	-27,77	0,00	-835,90	-725,68	-835,90	-725,68
DQNH_1h	-559,41	45,77	-12,22	0,00	-649,13	-469,70	-649,13	-469,70
DIRVENTO	58,21	3,89	14,97	0,00	50,59	65,84	50,59	65,84
VELVENTO	162,04	6,81	23,80	0,00	148,70	175,39	148,70	175,39
VISIB	0,23	0,01	35,74	0,00	0,21	0,24	0,21	0,24
Q1	-40,72	8,75	-4,66	0,00	-57,86	-23,58	-57,86	-23,58
TEMPAR	63,99	8,14	7,86	0,00	48,03	79,96	48,03	79,96
QNH	-63,87	8,67	-7,37	0,00	-80,86	-46,87	-80,86	-46,87
Diajul_m	-3,66	0,27	-13,77	0,00	-4,19	-3,14	-4,19	-3,14
Hora_m	-237,27	5,95	-39,85	0,00	-248,94	-225,60	-248,94	-225,60

## APÊNDICE D – OPERACIONALIZAÇÃO DO SISTEMA DE PREVISÃO DE CURTÍSSIMO PRAZO.

Após terem sido executados os treinamentos e as validações dos modelos de previsão e analisados os seus desempenhos e tendo sido determinado o melhor deles, o código ou função ajustada desse modelo, chamado aqui de *executável*, foi, então, embutido no modelo operacional implantado pelo Comando da Aeronáutica no Aeroporto de Guarulhos. As etapas operacionais do sistema são apresentadas adiante seguindo a seqüência do fluxograma da Figura D.1. Detalhes do *desenvolvimento* da operacionalização do sistema são apresentados no anexo B.



**Figura D. 1** – Fluxograma mostrando as seis etapas operacionais do sistema, bem como o armazenamento dos dados transformados, das previsões realizadas e das estatísticas das previsões.



Os procedimentos apresentados no fluxograma da Figura D.1 são detalhados a seguir:

Etapa 1: Aqui é realizada a digitação, pelo operador meteorologista, dos dados da observação meteorológica (horária ou especial); esses dados manterão o formato do formulário de registro de observação meteorológica em uso pelo operador (atualmente é o formulário IEPV 105-78);

Etapa 2: Nesta etapa, uma rotina realizará a transformação dos dados, oriundos da etapa anterior, para colocá-los no formato de entrada do modelo. Por exemplo: cálculo do número de horas de céu claro, transformação das variáveis de tempo, diferença da pressão de 24h, diferença em relação à hora anterior da visibilidade, adequação das camadas de nuvens, etc. Esses dados transformados serão armazenados para fins de avaliação de desempenho do modelo;

Etapa 3: Nesta etapa o modelo é executado. Os resultados são armazenados para fins estatísticos, para avaliação do desempenho do modelo;

Etapa 4: Nesta etapa são confeccionadas as estatísticas das previsões realizadas;

Etapa 5: Aqui é gerada a interface de saída do sistema de previsão de curtíssimo prazo.

Etapa 6: A cada intervalo de tempo pré-determinado será executada uma rotina para verificar se há a inclusão de novos dados; caso positiva, será reiniciada a rotina de execução do modelo. Esse intervalo de tempo, a ser determinado pelos operadores do sistema e deve ser menor que 60 minutos, buscando alguma observação meteorológica especial que possa ter sido feita entre as observações meteorológicas horárias regulares.



## ANEXO B - DESENVOLVIMENTO DA OPERACIONALIZAÇÃO DO SISTEMA DE PREVISÃO

### A) INTRODUÇÃO

A operacionalização do sistema de previsão de visibilidade e teto para o aeródromo de Guarulhos foi desenvolvido em conjunto com a analista de sistema Bianca Ruivo. Utiliza-se dos seguintes recursos:

- linguagem de programação PHP – versão 5.2.6
- servidor Web Apache – versão 2.0
- linguagem de programação Shell Script
- linguagem C++ (uso do compilador Microsoft Visual C++)
- banco de dados MySQL - versão 5.0.67
- biblioteca ADODB (acesso aos bancos de dados)
- linguagem de marcação HTML/XHTML
- folhas de estilo CSS
- linguagem de programação Java script – framework mootools e prototype
- tecnologia AJAX

### B) ESTRUTURA DOS BANCOS DE DADOS

#### a) BANCO BDFORM10578

##### ▪ **Tabela form10578\_entrada**

Esta tabela armazena os dados digitados na interface que representa o formulário IEPV 105-78 (com exceção do campo id, que é chave primária auto incremental e do campo processado, que é uma *flag* para marcar se o registro já foi processado ou não). Os campos dela estão listados abaixo:

- id: inteiro - chave primária - auto incremental
- dia: string de 2 caracteres

- mês: string de 2 caracteres
- ano: string de 4 caracteres
- hora: string de 2 caracteres
- minuto: string de 2 caracteres
- dirvento: inteiro
- velvento: inteiro
- visib\_min: inteiro
- cg1: inteiro
- qtdnuvem1: inteiro
- tipnuvem1: inteiro
- altnuvem1: inteiro
- qtdnuvem2: inteiro
- tipnuvem2: inteiro
- altnuvem2: inteiro
- qtdnuvem3: inteiro
- tipnuvem3: inteiro
- altnuvem3: inteiro
- qtdnuvem4: inteiro
- tipnuvem4: inteiro
- altnuvem4: inteiro
- bseco: double
- ur: inteiro
- qnh: double
- processado: boolean - default sendo o valor 0.

- **Tabela histórico**

Esta tabela armazena o nome do usuário que preencheu o formulário IEPV 105-78 e em qual data isto ocorreu. Os campos dela estão listados abaixo:

- id: inteiro – chave primária – auto incremental;
- usuário: caractere (10);
- data: timestamp – default sendo a data atual do sistema;

- **Tabela usuários**

Esta tabela armazena os usuários que terão acesso ao preenchimento da interface do formulário IEPV 105-78. Os campos dela estão listados abaixo:

- id: inteiro – chave primária – auto incremental
- nome: string de 50 caracteres
- login: string de 10 caracteres
- senha: string de 15 caracteres

b) BANCO BDRNA

- **Tabela dados\_xml\_form10578**

Esta tabela é uma cópia idêntica da tabela form10578\_entrada (banco bdform10578), cujos campos já foram listados anteriormente.

- **Tabela rna\_entrada**

Esta tabela armazena os dados que servirão de entrada para o programa executável da RNA. Os campos dela estão listados abaixo:

- id: inteiro – chave primária – auto incremental
- dia: string de 2 caracteres
- mês: string de 2 caracteres
- ano: string de 4 caracteres
- hora: string de 2 caracteres
- minuto: string de 2 caracteres
- dirvent\_24h: inteiro
- visib\_24h: inteiro
- tempair\_24h: double

- ur\_24h: inteiro
- qnh\_24h: inteiro
- dirvent\_1h: inteiro
- visib\_1h: inteiro
- tipo\_1h: inteiro
- q1\_1h: inteiro
- q2\_1h: inteiro
- q3\_1h: inteiro
- tempair\_1h: double
- ur\_1h: inteiro
- dirvent\_atual: inteiro
- velvent\_atual: inteiro
- visib\_atual: inteiro
- cgt\_atual: inteiro
- tipo\_atual: inteiro
- q1\_atual: inteiro
- q2\_atual: inteiro
- q3\_atual: inteiro
- 6h\_skc\_atual: inteiro
- tempair\_atual: double
- ur\_atual: inteiro
- qnh\_atual: double
- diajulm: inteiro
- horam: inteiro

- **Tabela rna\_saida**

Esta tabela armazena os resultados retornados (em forma de arquivo texto, com exceção dos campos id, id\_hora\_atual, processado e proc\_estatistica) pelo programa executável da RNA . Os campos dela estão listados abaixo:

- id: inteiro – chave primária – auto incremental
- dia: string de 2 caracteres
- mês: string de 2 caracteres
- ano: string de 4 caracteres
- hora: string de 2 caracteres
- minuto: string de 2 caracteres
- tipo\_hora: inteiro
- faixa\_visib: inteiro
- faixa\_nevoeiro: inteiro
- teto: inteiro
- qteto: inteiro
- id\_hora\_atual: inteiro
- processado: boolean – default sendo o valor 0 e representa se o registro já foi mostrado na interface de saída
- proc\_estatistica: boolean – default sendo o valor 0 e representa se o registro já foi considerado no cálculo das estatísticas.

- **Tabela estatísticas**

Esta tabela armazena as estatísticas dos últimos 2 anos . Os campos dela estão listados abaixo:

- id: inteiro – chave primária – auto incremental
- tipo\_hora: inteiro
- flag\_visib: smallint(6)
- flag\_nevoeiro: smallint(6)
- flag\_teto: smallint(6)
- Qobservado: smallint(6)
- Altobservado: smallint(6)

- data: datetime

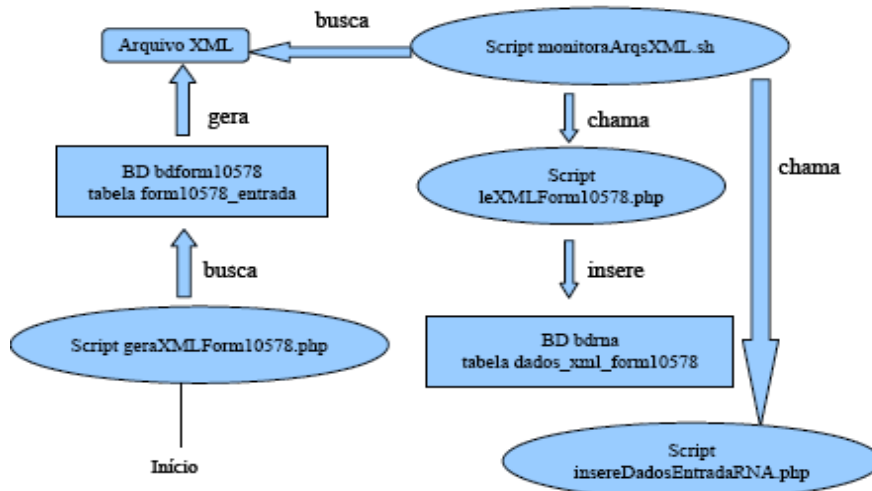
### C) INTERFACES

Neste tópico, serão abordadas as interfaces desenvolvidas.

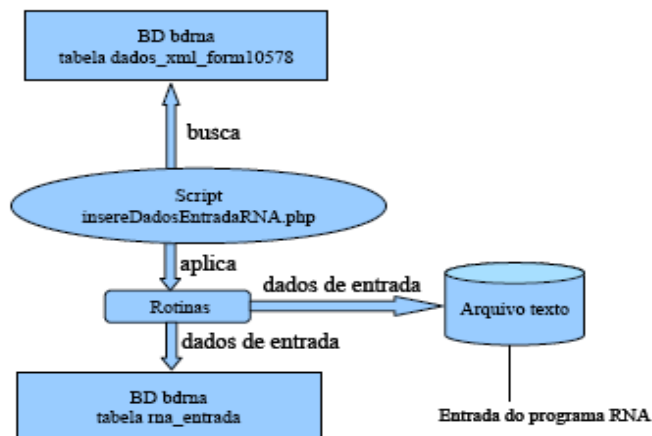
#### a) INTERFACE DE ENTRADA – FORMULÁRIO IEPV 105-78



#### b) INTERFACE FORMULÁRIO IEPV IEPV 105-78 – BANCO BDRNA

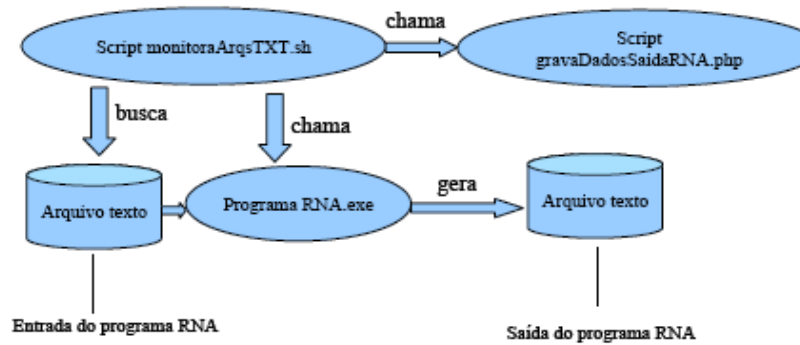


#### c) INTERFACE BANCO BDRNA – PROGRAMA RNA

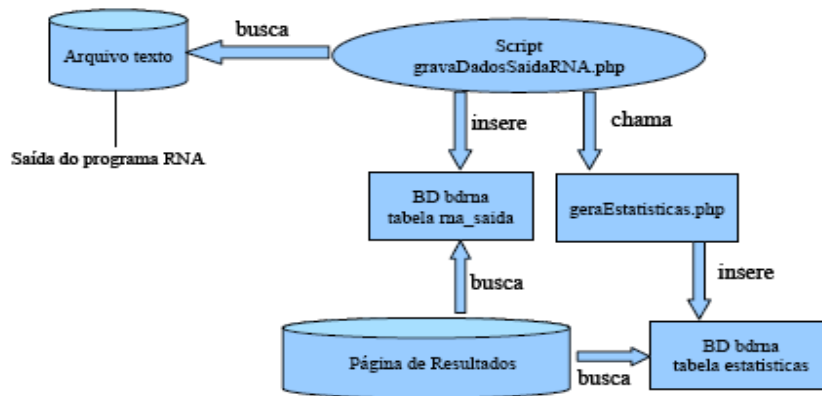




d) INTERFACE PROGRAMA RNA



e) INTERFACE BDRNA – PÁGINA DE RESULTADOS



f) PÁGINA DE RESULTADOS

**Previsão de Visibilidade e Teto para o Aeródromo de Guarulhos**

Inglês

**Situação Atual**

HORA	VISIBILIDADE ABAIXO DE 400 m	NEVOEIRO	TETO ABAIXO DE 100 ft	TETO ABAIXO DE 200 ft
23:00	Não	Não	Não	Não

**Previsão para as próximas 3 horas**

HORA	VISIBILIDADE ABAIXO DE 400 m	NEVOEIRO	TETO ABAIXO DE 100 ft	TETO ABAIXO DE 200 ft
00:00	Não	Não	Não	Não
01:00	Não	Sim	Não	Não
02:00	Não	Sim	Sim	Sim

**Estatística dos últimos dois anos**

HORA	VISIBILIDADE ABAIXO DE 400 m	NEVOEIRO	TETO ABAIXO DE 100 ft	TETO ABAIXO DE 200 ft
1ª hora	94,50%	94,66%	70,79%	83,14%
2ª hora	85,50%	92,23%	64,36%	71,60%
3ª hora	75,38%	90,43%	69,80%	78,04%

## D) LINGUAGENS E CÓDIGOS UTILIZADOS

- linguagem PHP – biblioteca ADODB
- linguagem Shell Script
- linguagem C++
- tecnologia AJAX
- linguagem HTML/XHTML
- linguagem Java script – frameworks mootools e prototype
- s folhas de estilo CSS

## E) OPERACIONALIZAÇÃO DO SISTEMA

Para operacionalizar o sistema, os seguintes passos são necessários:

- Alterar o conteúdo da variável `$diretorio_trabalho` localizada no arquivo `config.php` (diretório `/home/rna/Aeronautica/scripts/configs`). Esta variável deverá conter o caminho completo onde o sistema foi instalado.
- Alterar o conteúdo das variáveis `$servidor`, `$usuario`, `$senha` e `$bd` contidas no arquivo `config_bd_rna.php` (diretório `/home/rna/Aeronautica/scripts/configs`). Estas variáveis definem o acesso ao banco de dados da RNA.
- Alterar o conteúdo das variáveis `$servidor_form`, `$usuario_form`, `$senha_form` e `$bd_form` contidas no arquivo `config_bd_form10578.php` (diretório `/home/rna/Aeronautica/scripts/configs`). Estas variáveis definem o acesso ao banco de dados do formulário IEPV 105-78.
- Entrar no diretório `/home/rna/Aeronautica/scripts` e executar o script `monitoraArqsXML.sh` (comando `./monitoraArqsXML.sh`).
- Entrar no diretório `/home/rna/Aeronautica/scripts` e executar o script `monitoraArqsTXT.sh` (comando `./monitoraArqsTXT.sh`).

*Observação!* O script *monitoraArqsTXT.sh* deve ser executado depois do script *monitoraArqsXML.sh*!

- Colocar o script *roda\_sistema.sh* (diretório *home/rna/Aeronautica/scripts*) no crontab para ser executado de 10 em 10 minutos, por exemplo.
- Alterações na Interface de Resultados: alterar o conteúdo das variáveis *\$servidor*, *\$usuario*, *\$senha* e *\$bd* contidas no arquivo *config\_bd\_rna.php* localizado no diretório */home/rna/Aeronautica/Interface-saida/php/configs*. Estas variáveis definem o acesso ao banco de dados da RNA.
- Alterações na Interface de Entrada: alterar o conteúdo das variáveis *\$servidor*, *\$usuario*, *\$senha* e *\$bd* contidas no arquivo *conecta\_bd.php* localizado no diretório */home/rna/Aeronautica/Interface-entrada/php/*. Estas variáveis definem o acesso ao banco de dados do formulário IEPV 105-78.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)