



Ministério da  
Ciência e Tecnologia



INPE-15298-TDI/1350

**PREVISÃO DE DURAÇÃO DE NEVOEIRO NO  
AEROPORTO INTERNACIONAL DE SÃO PAULO  
UTILIZANDO ANÁLISE DE SOBREVIVÊNCIA**

Marcos Paulo de Souza

Dissertação de Mestrado do Curso de Pós-Graduação em Meteorologia, orientada  
pela Dr<sup>a</sup> Marley Cavalcanti de Lima Moscati, aprovada em 14 de abril de 2008

Registro do documento original:

<<http://urlib.net/sid.inpe.br/mtc-m17@80/2008/05.27.23.29>>

INPE  
São José dos Campos  
2008

# **Livros Grátis**

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

Milhares de livros grátis para download.

## **PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6911/6923

Fax: (012) 3945-6919

E-mail: [pubtc@sid.inpe.br](mailto:pubtc@sid.inpe.br)

## **CONSELHO DE EDITORAÇÃO:**

### **Presidente:**

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

### **Membros:**

Dr<sup>a</sup> Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Haroldo Fraga de Campos Velho - Centro de Tecnologias Especiais (CTE)

Dr<sup>a</sup> Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)

## **BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Jefferson Andrade Ancelmo - Serviço de Informação e Documentação (SID)

Simone A. Del-Ducca Barbedo - Serviço de Informação e Documentação (SID)

## **REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Marilúcia Santos Melo Cid - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva e Souza - Serviço de Informação e Documentação (SID)

## **EDITORAÇÃO ELETRÔNICA:**

Viveca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)



Ministério da  
Ciência e Tecnologia



INPE-15298-TDI/1350

**PREVISÃO DE DURAÇÃO DE NEVOEIRO NO  
AEROPORTO INTERNACIONAL DE SÃO PAULO  
UTILIZANDO ANÁLISE DE SOBREVIVÊNCIA**

Marcos Paulo de Souza

Dissertação de Mestrado do Curso de Pós-Graduação em Meteorologia, orientada  
pela Dr<sup>a</sup> Marley Cavalcanti de Lima Moscati, aprovada em 14 de abril de 2008

Registro do documento original:

<<http://urlib.net/sid.inpe.br/mtc-m17@80/2008/05.27.23.29>>

INPE  
São José dos Campos  
2008

Dados Internacionais de Catalogação na Publicação (CIP)

---

S89p Souza, Marcos Paulo.

Previsão de duração de nevoeiro no Aeroporto Internacional de São Paulo utilizando análise de sobrevivência/  
Marcos Paulo de Souza. – São José dos Campos: INPE, 2008.

173p. ; (INPE-15298-TDI/1350)

1. Nevoeiro. 2. Análise de sobrevivência. 3. Previsão.  
I. Título.

CDU (551.509.314)

---

Copyright © 2008 do MCT/INPE. Nenhuma parte desta publicação pode ser reproduzida, armazenada em um sistema de recuperação, ou transmitida sob qualquer forma ou por qualquer meio, eletrônico, mecânico, fotográfico, microfílmico, reprográfico ou outros, sem a permissão escrita da Editora, com exceção de qualquer material fornecido especificamente no propósito de ser entrado e executado num sistema computacional, para o uso exclusivo do leitor da obra.

Copyright © 2008 by MCT/INPE. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use of the reader of the work.

**Aprovado (a) pela Banca Examinadora  
em cumprimento ao requisito exigido para  
obtenção do Título de Mestre em  
Meteorologia**

**Dr. Prakki Satyamurty**



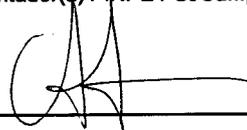
Presidente / INPE / SJCampos - SP

**Dr. Marley Cavalcante de Lima Moscati**



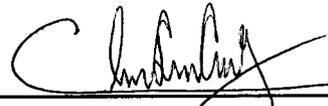
Orientador(a) / INPE / SJCampos - SP

**Dr. Clóvis Angeli Sansigolo**



Membro da Banca / INPE / SJCampos - SP

**Dr. Cleber Souza Corrêa**



Convidado(a) / ICEA/CTA / São José dos Campos - SP

**Aluno (a): Marcos Paulo de Souza**

**São José dos Campos, 14 de Abril de 2008**



A minha esposa Amanda, a meu filho,  
Vitor, e a meus pais,  
Aparecido e Eva  
Dedico



## **AGRADECIMENTOS**

A Deus.

À minha orientadora Dra. Marley Cavalcante de Lima Moscati, por sua dedicação e pela paciência que teve na orientação à distancia. Este trabalho só tornou-se possível graças à sua compreensão.

À minha Professora de Graduação Dra. Rosa Cruz Marques, pelo incentivo à continuar os estudos.

Ao Corpo Docente do INPE, especialmente ao Professor Clóvis Sansígolo, por sua atenção e indicação de algumas das referências bibliográficas.

Ao INPE/CPTEC e a INFRAERO pelo convênio firmado, que proporcionou o desenvolvimento deste trabalho. À CAPES pela bolsa de estudos.

A todos amigos que conquistei no decorrer do Curso, especialmente à Priscila Tavares, que me ajudou na formatação deste trabalho.

À Secretária da PGMET/INPE, Sra. Lílian e também ao Suporte Técnico de Informática, Sr. César, pela dedicação e eficiência.

Aos servidores da INFRAERO, Cristina Fogaccia, Marcelo Presti, Carlos Robles, Mufida, Márcia, Josi e Angela, pela receptividade e atenção.

A todos que contribuíram nas diversas fases desse trabalho, como a Dra. Marília Cristina de Souza Antunes dos Reis, que enviou cópia de sua dissertação de mestrado sobre análise de sobrevivência; a Dra. Suely Ruiz Giolo, que tirou as dúvidas surgidas ao ler o livro de Colosimo e Giolo (2006); as comunidades do R e do SAS, que ajudaram a resolver muitas dúvidas surgidas no uso desses softwares; ao Ms. Ney Ricardo Moscati, pela ajuda na etapa final do trabalho.

À Banca Examinadora, pelas sugestões e críticas.

Á minha esposa Amanda e a meu filho Vítor, que ainda não entende este agradecimento, mas um dia saberá que o pouco tempo dedicado a ele nos finais de semanas foi pensando em um futuro melhor.

A meus pais Aparecido e Eva, pelo incentivo nos estudos e pelo apoio incondicional.

A meus irmãos Ricardo e Fernanda.



## RESUMO

Nesse trabalho é desenvolvido um método de previsão de duração de nevoeiro, para todos os meses do ano, para o Aeroporto Internacional de São Paulo utilizando o método de Análise de Sobrevivência. Utilizaram-se séries temporais horárias de várias variáveis meteorológicas do Aeroporto Internacional de São Paulo, para o período de 1 de janeiro de 1988 até 31 de dezembro de 2005. As características climatológicas (1988-2003) típicas dos nevoeiros nessa região foram: 1) Formam-se preferencialmente nos meses de maio a agosto e a maior quantidade de horas com nevoeiro ocorre em junho (551,3 h, que representa 20% do total de horas mensal); 2) Formam-se preferencialmente entre as 20 HL e as 8 HL, verificando-se apenas quatro casos de nevoeiros que se formaram fora desse intervalo, sendo dois casos as 9 HL, um caso as 16 HL e um caso as 19 HL; 3) O pico de maior ocorrência de horas com nevoeiro ocorreu entre as 5 HL e 6 HL; 4) Há maior probabilidade de duração para os nevoeiros que começam mais cedo, que ocorrem nos meses de junho, julho, maio e agosto, que iniciam com direção do vento menor que  $150^\circ$  e próxima de  $90^\circ$ , que se iniciam com velocidade do vento menor que 8 nós e mediana de 5 nós, que apresentam pressão maior ou igual a 1018 hPa e mediana entre 1018 a 1025 hPa,  $T_d$  menor que  $16^\circ\text{C}$  e mediana entre  $13^\circ\text{C}$  e  $16^\circ\text{C}$  e  $T_w$  maior ou igual a  $17^\circ$  e mediana entre  $11^\circ\text{C}$  e  $17^\circ\text{C}$ . A aplicação da Análise de Sobrevivência para definir um método de previsão de duração de nevoeiro nesse aeroporto considerou 697 eventos de nevoeiros. O modelo que melhor se ajustou aos dados dos nevoeiros do Aeroporto Internacional de São Paulo, mediante Estatística de Teste (TRV) foi o Modelo de Weibull e os melhores preditores para a previsão da duração do nevoeiro foram as variáveis: mês de ocorrência do nevoeiro (V1), hora de início do nevoeiro (V2), direção do vento (V3), velocidade do vento (V4), visibilidade (V5),  $T_d$  (V7) e  $T_w$  (V8). A equação gerada com as estimativas dos parâmetros de Weibull apresentou resultados satisfatórios, sendo que de 47 casos de nevoeiros identificados no ano de 2004, apenas 9 casos tiveram o tempo de duração subestimado. Para melhorar o desempenho desse modelo, sugere-se que essa análise considere apenas os meses de maior tempo de duração de nevoeiros (maio, junho, julho e agosto).



## FORECAST OF DURATION OF FOGS INTERNATIONAL AIRPORT OF SÃO PAULO USING SURVIVAL ANALYSIS

### ABSTRACT

In this work, a duration fog forecast method for all months of the year to the International Airport of Sao Paulo is developed with the use of the Survival Analysis method. It was used time series of meteorological variables from the International Airport of Sao Paulo for the period from January 1, 1988 until December 31, 2005. The typical of fogs climatological characteristics (1988-2003) in this region were: 1) Formation mainly in the months from May to August and the largest amount of hours with fog occurs in June (551.3 h, which represents 20% of total hours monthly); 2) The fogs are formed mainly between 20 HL and 8 HL, verifying only four cases of fogs that were formed outside of this interval, being two cases at 9 HL, one case at 16 HL and one case at 19 HL; 3) The peak of bigger occurrence of hours with fog occurred between 5 HL and 6 HL; 4) It has greater probability of duration for the fogs that more early start, occur in June, July, May and August, that initiate with direction of the wind lesser than  $150^{\circ}$  and near to  $90^{\circ}$ , that initiate with speed of the wind lesser than 8 kts and average of 5 kts, that present pressure larger or like 1018 hPa and average between 1018 to 1025 hPa, Td lesser than  $16^{\circ}$  C and average between  $13^{\circ}$  C and  $16^{\circ}$  C and Tw larger or like  $17^{\circ}$  and average between  $11^{\circ}$  C and  $17^{\circ}$  C. The application of the Survival Analysis to define duration fog forecast method in this airport found 697 fog events. The model that best fit to the Sao Paulo International Airport fog data, using Statistical Test (TRV) was Weibull model and the best predictors to duration fog forecast were the variables: occurrence of fog months (V1), fog initials hours (V2), wind direction (V3), wind speed (V4), visibility (V5), Td (S7) and Tw (V8). The equation generated with Weibull parameters showed satisfactory results, considering 47 fog cases in 2004, only 9 cases underestimated the duration time. Finally, in order to improve the performance of this model, it is suggested that this analysis considers only greater duration time fog months (May, June, July and August).



## SUMÁRIO

Pág.

LISTA DE FIGURAS	
LISTA DE TABELAS	
LISTA DE SIGLAS E ABREVIATURAS .....	
<b>1 INTRODUÇÃO .....</b>	<b>25</b>
<b>2 REVISÃO BIBLIOGRÁFICA.....</b>	<b>31</b>
2.1 – CLASSIFICAÇÃO DE NEVOEIROS .....	31
2.2 – CARACTERÍSTICAS GERAIS DOS NEVOEIROS .....	35
2.3 – MÉTODOS DE PREVISÃO DE OCORRÊNCIA DE NEVOEIRO .....	40
2.4 – MÉTODOS DE PREVISÃO DE DURAÇÃO DE NEVOEIRO .....	43
2.5 – OCORRÊNCIA DE NEVOEIROS NO BRASIL E MÉTODOS DE PREVISÃO .....	44
<b>3 MATERIAIS E MÉTODOS .....</b>	<b>56</b>
3.1 – ÁREA DE ESTUDO.....	57
3.2 - DADOS .....	59
3.3 - METODOLOGIA .....	61
3.4 – APLICAÇÃO DOS DADOS DO AEROPORTO INTERNACIONAL DE SÃO PAULO ÀS TÉCNICAS DE ANÁLISE DE SOBREVIVÊNCIA.....	63
3.4.1 – ELABORAÇÃO DO MÉTODO DE PREVISÃO DE DURAÇÃO DE NEVOEIRO PARA O AEROPORTO INTERNACIONAL DE SÃO PAULO.....	65
<b>4 RESULTADOS E DISCUSSÕES.....</b>	<b>71</b>
4.1 – CLIMATOLOGIAS DE NEVOEIRO NO AEROPORTO INTERNACIONAL DE SÃO PAULO.....	71
a) Climatologia A .....	71
b) Climatologia B .....	75
4.2– CARACTERÍSTICAS TÍPICAS DOS PARÂMETROS METEOROLÓGICOS DO AEROPORTO INTERNACIONAL DE SÃO PAULO .....	77
4.3 - RESULTADOS DA ANÁLISE DE SOBREVIVÊNCIA.....	85
<b>5 CONCLUSÕES E CONSIDERAÇÕES FINAIS.....</b>	<b>99</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>103</b>
<b>BIBLIOGRAFIA CONSULTADA.....</b>	<b>109</b>

<b>APÊNDICE A</b> .....	<b>111</b>
SÉRIES TEMPORAIS CLIMATOLÓGICAS (1988 - 2005) MENSAIS DAS VARIÁVEIS METEOROLÓGICAS PARA O AEROPORTO INTERNACIONAL DE SÃO PAULO.....	111
<b>APÊNDICE B</b> .....	<b>115</b>
CONCEITO BÁSICOS SOBRE ANÁLISE DE SOBREVIVÊNCIA.....	115
B.1 - ASPECTOS GERAIS.....	116
B.1.1 - Cálculo de Estimadores.....	118
a) Estimadores Paramétricos.....	118
b) Estimadores Não-Paramétricos.....	124
c) Estimador Semi-paramétricos.....	126
B.1.2 - Etapas de um Estudo de Análise de Sobrevivência.....	127
B.2 - DESCRIÇÃO SUCINTA DA METODOLOGIA PROPOSTA POR REIS REIS (1995).....	127
<b>APÊNDICE C</b> .....	<b>135</b>
NOÇÕES SOBRE O AMBIENTE R E SUA UTILIZAÇÃO.....	135
C.1 - UTILIZAÇÃO DO R.....	135
<b>APÊNDICE D</b> .....	<b>145</b>
NOÇÕES BÁSICAS SOBRE O SISTEMA SAS E SUA UTILIZAÇÃO.....	145
D.1 - UTILIZAÇÃO DO PROGRAMA SAS.....	146
D.2 - SUBROTINA DOS SAS: PROC LIFEREG.....	156
<b>APÊNDICE E</b> .....	<b>153</b>
ALGUNS PROGRAMAS UTILIZADOS, ESCRITOS EM R.....	153
E.1 - CONSTRUÇÃO DO BOX PLOT PARA A VARIÁVEL "DIREÇÃO DO VENTO".....	153
E.2 - CONSTRUÇÃO DAS FIGURAS DO ESTIMADOR DE KAPLAN-MEIER.....	157
E.3 - ESTIMATIVAS DOS PARÂMETROS E TESTE (valor p) DO MODELO DE WEIBULL.....	159
E.4 - CONSTRUÇÃO DOS GRÁFICOS DOS RESÍDUOS.....	160
<b>APÊNDICE F</b> .....	<b>161</b>
ALGUNS PROGRAMAS UTILIZANDOS, ESCRITOS EM LINGUAGEM SAS.....	161
<b>APÊNDICE G</b> .....	<b>167</b>
VERIFICAÇÃO DO MODELO.....	167

## LISTA DE FIGURAS

2.1 – Exemplos de situações sinóticas associadas com eventos de nevoeiro de radiação forte em Albany (1970-1979), para os tipos 1, 2, 3, 4 e 5, respectivamente.....	37
2.2 – Climatologia (julho de 1966 a abril de 2002) de nevoeiros densos no Aeroporto Internacional de O'Hare. ....	39
2.3 - a) Total anual de horas de nevoeiro nas estações meteorológicas de Cumbica (1969–1984) e de Guarulhos (1985–1997); b) Total anual de horas de nevoeiro nos Aeroportos de Guarulhos e Congonhas (1972-1997); c) Freqüência média mensal de horas de nevoeiro dos Aeroportos de Guarulhos e Congonhas (1980–1997); d) Horários de início de nevoeiro no Aeródromo de Guarulhos (1993-1997), e) Freqüência de duração do nevoeiro no Aeródromo de Guarulhos, no período de 1993 a 1997 (continua). ....	45
2.4 – a) Ocorrência de nevoeiros (expressa em horas) em São José dos Campos, SP, para o período de 1974 a 1992; b) Distribuição horária média de maior ocorrência de nevoeiros; c) Distribuição horária média da T; d) Distribuição horária média de UR; e) Distribuição horária média da velocidade do vento. ....	48
2.5 - Fluxograma de decisão para previsão do nevoeiro no Aeroporto Internacional de São Paulo. ....	54
3.1 - Mapa da localização da pista do Aeroporto Internacional de São Paulo, em Guarulhos (indicado pelo círculo azul) e arredores. ....	58
3.2 - a) Topografia da região em torno da cidade de São Paulo; b) Topografia de uma sessão perpendicular à costa, na direção noroeste-sudeste (Figura 3.2 a), com indicação de algumas características geográficas. As siglas sr, fr, si, pn, cg, eg, cm, gr e iag indicam a localização das estações de São Roque, Franco da Rocha, Santa Isabel, Ponte Nova, Aeroporto de Congonhas, Embu-Guaçu, Campo de Marte (Aeroporto de Guerra), Aeroporto Internacional de São Paulo e Estação do IAG/USP, respectivamente. As siglas CRV, TRV, PRV, R e SM significam Crista da Cantareira, Vale do Rio Tietê, Vale do Rio Pinheiros, Reservatório de água e Serra do Mar, respectivamente. ....	59
3.3 – Esquema ilustrativo de um: a) <i>boxplot</i> ; b) <i>Plot</i> esquemático. ....	65
4.1 – Climatologia (1988-2003) da freqüência relativa (expressa em %) e da freqüência absoluta (indicadas pelos números acima de cada coluna) de totais mensais de horas com nevoeiro, no Aeroporto Internacioanl de São Paulo. ....	72
4.2 – Climatologia (1988-2003) de horas totais anuais de nevoeiros para o Aeroporto Internacional de São Paulo. ....	73

- 4.3 – Climatologia (1988-2003) de horas totais anuais de nevoeiros considerando apenas os meses de maio, junho, julho e agosto para o Aeroporto Internacional de São Paulo..... 73
- 4.4 – Climatologia (1988-2003) da frequência relativa (expressa em %) e da frequência absoluta (identificada pelos números acima das colunas) da ocorrência dos nevoeiros com a hora de início do nevoeiro, (em intervalos de 1 hora), considerando o período de 1988 a 2003. .... 74
- 4.5 - Climatologia B da frequência relativa (expressa em %) e da frequência absoluta (indicadas pelos números acima de cada coluna) de totais mensais de horas com nevoeiro, no Aeroporto Internacional de São Paulo.
- 4.6 - Climatologia (1988-2003) de horas totais anuais de nevoeiros para o Aeroporto Internacional de São Paulo.
- 4.7 – *Boxplot* relacionando as horas de início dos nevoeiros (intervalos de uma hora) com suas durações, considerando o período de 1988 a 2003. .... 77
- 4.8 – *Boxplot* para a relação dos meses de ocorrência dos nevoeiros com suas durações (expressas em intervalos de uma hora), considerando o período de 1988 a 2003. .... 78
- 4.9 – *Boxplot* para a direção do vento em Graus (em <sup>o</sup>), para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** (≥ 2 h e < 3 h), **3** (≥3 h e < 4 h), **4** (≥ 4 h e < 5 h), **5** (≥ 5 h e < 6 h), **6** (≥ 6 h e < 7 h), **7** (≥ 7 h e < 8 h), **8** (≥ 8 h e < 9 h), **9** (≥ 9 h e < 10 h), **11** (≥ 11 h e < 12 h), **12** (≥ 12 h e < 13 h), **13** (≥13 h e < 14h), e **14** (≥14 h e < 15h)..... 80
- 4.10 – *Boxplot* para a velocidade do vento (expressa em nós) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** (≥ 2 h e < 3 h), **3** (≥3 h e < 4 h), **4** (≥ 4 h e < 5 h), **5** (≥ 5 h e < 6 h), **6** (≥ 6 h e < 7 h), **7** (≥ 7 h e < 8 h), **8** (≥ 8 h e < 9 h), **9** (≥ 9 h e < 10 h), **11** (≥ 11 h e < 12 h), **12** (≥ 12 h e < 13 h), **13** (≥13 h e < 14h), e **14** (≥14 h e < 15h). .... 80
- 4.11 – *Boxplot* para a visibilidade (expressa em m) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** (≥ 2 h e < 3 h), **3** (≥3 h e < 4 h), **4** (≥ 4 h e < 5 h), **5** (≥ 5 h e < 6 h), **6** (≥ 6 h e < 7 h), **7** (≥ 7 h e < 8 h), **8** (≥ 8 h e < 9 h), **9** (≥ 9 h e < 10 h), **11** (≥ 11 h e < 12 h), **12** (≥ 12 h e < 13 h), **13** (≥13 h e < 14h), e **14** (≥14 h e < 15h)..... 81
- 4.12 – *Boxplot* para a pressão (expressa em hPa) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período

- de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 até 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** (≥ 2 h e < 3 h), **3** (≥3 h e < 4 h), **4** (≥ 4 h e < 5 h), **5** (≥ 5 h e < 6 h), **6** (≥ 6 h e < 7 h), **7** (≥ 7 h e < 8 h), **8** (≥ 8 h e < 9 h), **9** (≥ 9 h e < 10 h), **11** (≥ 11 h e < 12 h), **12** (≥ 12 h e < 13 h), **13** (≥13 h e < 14h), e **14** (≥14 h e < 15h)..... 81
- 4.13 – *Bosxplot* para a  $T_d$  para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 até 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** (≥ 2 h e < 3 h), **3** (≥3 h e < 4 h), **4** (≥ 4 h e < 5 h), **5** (≥ 5 h e < 6 h), **6** (≥ 6 h e < 7 h), **7** (≥ 7 h e < 8 h), **8** (≥ 8 h e < 9 h), **9** (≥ 9 h e < 10 h), **11** (≥ 11 h e < 12 h), **12** (≥ 12 h e < 13 h), **13** (≥13 h e < 14h), e **14** (≥14 h e < 15h)..... 82
- 4.14 – *Bosxplot* para a  $T_w$  para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 até 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** (≥ 2 h e < 3 h), **3** (≥3 h e < 4 h), **4** (≥ 4 h e < 5 h), **5** (≥ 5 h e < 6 h), **6** (≥ 6 h e < 7 h), **7** (≥ 7 h e < 8 h), **8** (≥ 8 h e < 9 h), **9** (≥ 9 h e < 10 h), **11** (≥ 11 h e < 12 h), **12** (≥ 12 h e < 13 h), **13** (≥13 h e < 14h), e **14** (≥14 h e < 15h)..... 82
- 4.15 – *Bosxplot* para a UR (em %) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 até 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** (≥ 2 h e < 3 h), **3** (≥3 h e < 4 h), **4** (≥ 4 h e < 5 h), **5** (≥ 5 h e < 6 h), **6** (≥ 6 h e < 7 h), **7** (≥ 7 h e < 8 h), **8** (≥ 8 h e < 9 h), **9** (≥ 9 h e < 10 h), **11** (≥ 11 h e < 12 h), **12** (≥ 12 h e < 13 h), **13** (≥13 h e < 14h), e **14** (≥14 h e < 15h)..... 83
- 4.16 – *Bosxplot* para a visibilidade média (expressa em m) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 até 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** (≥ 2 h e < 3 h), **3** (≥3 h e < 4 h), **4** (≥ 4 h e < 5 h), **5** (≥ 5 h e < 6 h), **6** (≥ 6 h e < 7 h), **7** (≥ 7 h e < 8 h), **8** (≥ 8 h e < 9 h), **9** (≥ 9 h e < 10 h), **11** (≥ 11 h e < 12 h), **12** (≥ 12 h e < 13 h), **13** (≥13 h e < 14h), e **14** (≥14 h e < 15h)..... 83
- 4.17 - Estimador de Kaplan-Meier da Função de Sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “mês de ocorrência”. ..... 87
- 4.18 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “hora de início do nevoeiro”. ..... 87

4.19 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “direção do vento”.....	88
4.20 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “velocidade do vento”.....	88
4.21 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “visibilidade”. ....	89
4.22 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “pressão”. ....	89
4.23 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável Td. ....	90
4.24 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável Tw.....	90
4.25 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável UR. ....	91
4.26 - Sobrevivência dos Resíduos de Cox-Snell estimadas pelo método de Kaplan-Meier e pelo Modelo Exponencial padrão (para o Modelo Final).....	96
4.27 - Curvas de sobrevivência estimadas (para o Modelo Final).....	96
A.1 - Direção do Vento em graus ( $^{\circ}$ ) para o período de 1988 à 2005.....	111
A.2 - Velocidade do Vento em Nós (KT) para o período de 1988 à 2005.....	111
A.3 - Visibilidade em Metros (M) para o período de 1988 à 2005.....	112
A.4 - Pressão em hectopascal (hPa) para o período de 1988 à 2005.....	112
A.5 - Td em Celsius ( $^{\circ}$ C) para o período de 1988 à 2005.....	113
A.6 - Tw em Celsius ( $^{\circ}$ C) para o período de 1988 à 2005.....	113
A.7 - UR em Porcentagem (%) para o período de 1988 à 2005.....	114

## LISTA DE TABELAS

1.1 - Valores mínimos de Teto e RVR para as várias CATs, estabelecidos como limites de segurança para pousos e decolagens em aeroportos.....	26
1.3 - Acidentes causados por condições meteorológicas adversas, para a categoria de operações doméstica, de bandeira e suplementares. ....	28
1.4: Frequência anual de nevoeiros para os principais aeroportos dos EUA. ....	28
3.1 – Relação das covariáveis utilizadas no modelo de previsão de duração de nevoeiro para o Aeroporto Internacional de São Paulo.....	66
4.1 - Teste de Logrank e de Wilcoxon.....	92
4.2 - Seleção das covariáveis.....	93
4.3 - Estimativas dos parâmetros do Modelo de Weibull.....	95
B.1 - Parâmetros associados com algumas distribuições de probabilidade.....	123
B.2 - Resultado do teste para o Modelo de Cox completo.....	130
B.3 - Resultado do teste para o Modelo de Cos considerando as variáveis significativas $(x_3$ e $x_4)$ .....	130
B.4 - Resultado do teste para o Modelo de Regressão Exponencial completo (considerando todas as variáveis).....	131
B.5 - Resultado do teste para o Modelo de Regressão Exponencial com as variáveis significativas $x_3$ e $x_4$ .....	131
B.6 - Divisão da covariável $x_4$ em apenas dois grupos para cada valor fixo de $x_3$ .....	132
C.1 - Principais comando básicos utilizados na Programação R.....	144
D1 - Estrutura DATA Step.....	147
D.2 Estrutura PROC Step.....	148
D.3 - Tipos de Censura.....	151
G.1 - Eventos de nevoeiros ocorridos em 2004 no Aeroporto Internacional de São Paulo.....	167
G.2 - Dicotomização dos dados referente a eventos de nevoeiros ocorridos em 2004.....	169
G.3 - Probabilidade dos nevoeiros ocorridos no Aeroporto Internacional de São Paulo em 2004 durarem além de um tempo $t$ , conhecida as covariáveis selecionadas no Modelo Final.....	171
G.4 - Verificação do modelo.....	173



## LISTA DE SIGLAS E ABREVIATURAS

APS	- Anticiclone do Pacífico Sul
BI	- Base de Inversão de temperatura
CAT	- Categoria de Operação
CGP	- Código Geral de Tempo
CPTEC	- Centro de Previsão de Tempo e Estudos Climáticos
CLP	- Camada Limite Planetária noturna
CMT	- Camada de Mistura Turbulenta
EUA	- Estados Unidos da América
FAR	- Federal Aviation Regulation (regulamento aeronáutico)
f.d.p	- função densidade e probabilidade
FSI	- <i>Fog Stability Index</i>
GrADS	-Grid Analysis and Display System
HL	- Hora Local
IAG	- Instituto de Astronomia, Geofísica e Ciências Atmosféricas.
ILS	- Instrument Landing System (Sistema de pouso por instrumento)
IME/USP	- Instituto de Matemática e Estatística/Universidade de São Paulo
INFRAERO	- Empresa Brasileira de Infra-Estrutura Aeroportuária
INPE	- Instituto Nacional de Pesquisas Espaciais
MAE 514	- Disciplina de Introdução a análise de sobrevivência do Instituto de Matemática e Estatística da Universidade de São Paulo.
METAR	- Meteorological Aerodrome Reports.
MM5	- Modelo de Mesoescala da <i>Pennsylvania State University/National Center for Atmospheric Research</i>
NCAR	- National Center for Atmospheric Research
NCEP	- National Center for Environmental Prediction
Q_CLR	- Quantidade de horas com céu claro
Q_HZBR	- Quantidade de horas de névoas seca ou úmida
PNM	- Pressão ao Nível médio do Mar
RAMS	- <i>Regional Atmospheric Modeling System</i>
RVR	- Runway Visual Range (Alcance Visual da Pista)
ROTAER	- Manual de Rotas Aéreas
T850	- Temperatura no nível de 850 hPa

TS	- Temperatura do ar da superfície
TDS	- Temperatura do Ponto de Orvalho
TRV	- Teste de Razão de Verossimilhança
$T - T_d$	- Depressão do ponto de orvalho
UFRJ	- Universidade Federal do Rio de Janeiro
UR	- Umidade Relativa
USP	- Universidade de São Paulo
UTC	- Universal Time Coordinate (hora no meridiano de Greenwich)
VASP	- Viação Aérea São Paulo
W850	- Velocidade do vento no nível de 850 hPa.
WMO	- World Meteorological Organization

## LISTA DE SÍMBOLOS

AIQ	- amplitude Inter-Quartis
$\alpha, \sigma, \mu, \kappa$	- Parâmetros de Forma
$d_j$	- número de falhas em um instante $t_j$
$\delta_i$	- Indicador de Falha ou Censura
$H_0$	- Hipóteses nula
$L(\theta)$	- Logaritmo de Verossimilhança
$N_j$	- número de eventos sob risco em um instante $t_j$
$\theta$	- Representação dos Parâmetros de Forma
P	- Pressão reduzida ao Nível médio do Mar (hPa)
$Q_1$	- Primeiro quartil
$Q_3$	- Terceiro quartil
$S(t)$	- Função de Sobrevivência
$\hat{S}(t)$	- Função de Sobrevivência do Estimador de Kaplan-Meier
$\tilde{S}(t)$	- Função de Sobrevivência do Estimador de Nelson Aalen
T	- Temperatura do ar ( $^{\circ}\text{C}$ )
$T_d$	- Temperatura do Ponto de Orvalho ( $^{\circ}\text{C}$ )
$T_F$	- Tempo de Falha
$t_j$	- tempo na j-ésima observação
$T_w$	- Temperatura do Bulbo Seco ( $^{\circ}\text{C}$ )
V1	- Mês de ocorrência dos nevoeiros
V2	- Hora de início dos nevoeiros (h)
V3	- Direção do Vento ( $^{\circ}$ )
V4	- Velocidade do Vento (kt)
V5	- Visibilidade (M)
V6	- Pressão reduzida ao Nível médio do Mar (hPa)
V7	- Temperatura do Ponto de Orvalho ( $^{\circ}\text{C}$ )
V8	- Temperatura do Bulbo Seco ( $^{\circ}\text{C}$ )
V9	- Umidade Relativa (%)

$\chi^2$	- Qui-quadrado
$\lambda(t)$	- Função Taxa de Falha
$\infty$	- Infinito
$x'$	- Representação de mais de uma covariável

## 1 INTRODUÇÃO

Nevoeiro (do inglês, *fog*) é um fenômeno de tempo de grande impacto em vista de suas conseqüências em vários campos de atividades, principalmente para a aviação e para todas as formas de transporte em superfície (MASON, 1982; DUYNKERKE, 1991). De acordo com a *World Meteorological Organization - WMO* (1966), o nevoeiro pode ser descrito como uma suspensão de gotículas d'água ou cristais de gelo na atmosfera que reduz a visibilidade horizontal na superfície a menos de 1 km. Segundo o Great Britain. Meteorological Office (1994), o nevoeiro é caracterizado quando a visibilidade for menor que 1.000 m e a Umidade Relativa (UR) próxima a 100%, com um mínimo de 95%. Condições em que a visibilidade é maior que 1.000 m e a UR menor que 100% caracteriza o fenômeno de névoa úmida.

Em aeroportos, a ocorrência de nevoeiro significa que vôos podem ser cancelados ou atrasados. Isto resulta em perdas financeiras substanciais para a indústria da aviação, devido modificações nos planejamentos de vôos das empresas aéreas, despesas extras com hotel e combustível, demasiada ocupação dos pátios dos aeroportos alternativos ocasionando a sobrecarga operacional nos sistemas destes aeroportos, entre outros, além de causar prejuízos aos passageiros (OLIVEIRA, 2002). Para que as companhias aéreas possam manter os vôos nos horários pré-determinados durante todo o ano, é necessário que o aeroporto seja capaz de operar as aeronaves nas fases de pouso e decolagem em condições de baixa visibilidade. Para isso, alguns aeroportos dispõem de um sistema de pouso e decolagem através da Categoria de Operação (CAT) em condições de baixa visibilidade, e o sistema para pouso por instrumento é denominado de *Instrument Landing System (ILS)*. A Tabela 1.1 apresenta as especificações para Teto e *Runway Visual Range (RVR)* mínimos para vários tipos de CAT, visando garantir a segurança da operação de pouso e decolagem em aeroportos. O RVR indica a distância (expressa em metros ou em pés) na qual o piloto de uma aeronave posicionado

no centro da pista consegue ver as marcações da pista, ou ainda a linha de centro de pista em condições de baixa visibilidade (MOSCATI, 2005).

Tabela 1.1 - Valores mínimos de Teto e RVR para as várias CATs, estabelecidos como limites de segurança para pousos e decolagens em aeroportos.

<b>Especificações</b>	<b>Teto (mínimo)</b>	<b>RVR (mínimo)</b>
CAT I	200 pés	800 m
CAT II	100 pés	400 m
CAT III A	0	200 m
CAT III B	0	100 m
CAT III C	0	0

Fonte: Adaptado de Moscati (2005).

Nos aeroportos do Brasil, a maior categoria de operação é a de CAT II, sendo que apenas o Aeroporto Internacional do Rio de Janeiro/Galeão e o Aeroporto de São Paulo/Guarulhos operam nessa categoria (ROTAER, 1999). Vale ressaltar que embora estes aeroportos possuam o sistema de CAT II, há momentos críticos, como aqueles no qual a altura da camada das nuvens e a visibilidade estão abaixo dos limites de segurança (conforme definido na Tabela 1.1), em que se deve evitar o pouso. Como exemplo, cita-se um caso recente em que o Aeroporto Internacional de São Paulo/Guarulhos ficou fechado no dia 09 de junho de 2007 das 2 h e 45 min às 9 h e 50 min devido a ocorrência do nevoeiro, que restringiu a visibilidade abaixo dos limites de segurança inviabilizando o emprego do sistema de CAT II (INFRAERO, 2007). Moscati (2005) alerta que não basta apenas o aeroporto estar equipado com o sistema de CAT II, além disso, a companhia aérea deve estar homologada de acordo com as especificações do órgão competente e realizar treinamentos periódicos com sua tripulação.

Também, o nevoeiro tem sido apontado como a causa direta de várias ocorrências de acidentes e incidentes aéreos e rodoviários em diversos lugares. Como exemplo, no Brasil, destaca-se o caso de nevoeiro ocorrido em 28 de janeiro de 1986 e que causou um acidente com o Boeing 737 da VASP (Prefixo PP-SME) no Aeroporto Internacional de São Paulo (CABRAL, 1998). Outro acidente que teve como causa direta a ocorrência de nevoeiro foi aquele

de 25 de maio de 1998, provocando o fechamento do Aeroporto Salgado Filho, em Porto Alegre/RS, desde às 1 h e 30 min até às 9 h, causando atrasos em 11 decolagens e 8 aterrissagens, e acidentes rodoviários em dez estradas federais (com registro de dois óbitos) e dois acidentes em estradas estaduais (PIVA E FEDOROVA, 1998 a).

Moscati (2005) apresentou uma revisão sobre as estatísticas que correlacionam acidentes aéreos com condições meteorológicas adversas, para o período de 1994 a 2003, usando a base de dados do *National Transport Safety Board*, considerando as categorias de operação da aeronave de aviação geral – FAR Parte 91 (Tabela 1.2) e de operações doméstica, de bandeira e suplementares (Tabelas 1.3). A análise das Tabelas 1.2 e 1.3 mostra que para essas duas categorias de operação cujas causas para o acidente aéreo foram condições meteorológica adversas, mais de 19 % e 2% são decorrentes de baixa visibilidade ou baixa altura da camada de nuvens, respectivamente. Isso sugere que erros na previsão da ocorrência ou duração de nevoeiro nos aeroportos podem ter um impacto considerável, tanto econômico como pela perda de vidas.

Tabela 1.2 – Acidentes aéreos causados por condições meteorológicas adversas, para categoria de operação de aviação geral (FAR Parte 121).

CONDIÇÕES METEOROLÓGICAS	TOTAL DE EVENTOS
Vento	2.514
<b>Visibilidade/Altura da camada de nuvens</b>	<b>977</b>
Turbulência	371
Gelo	350
Altitude barométrica	347
Precipitação	208
Tempestade	94
Windshear	59
Outros	7
<b>Total</b>	<b>4.927</b>

Fonte: Moscati (2005).

Tabela 1.3 - Acidentes causados por condições meteorológicas adversas, para a categoria de operações doméstica, de bandeira e suplementares.

CONDIÇÕES METEOROLÓGICAS	TOTAL DE EVENTOS
Turbulência	92
Vento	11
Precipitação	8
<b>Visibilidade/Altura da camada de nuvens</b>	<b>3</b>
Tempestade	3
Altitude barométrica	2
Windshear	2
Outros	2
Gelo	1
<b>Total</b>	<b>124</b>

Fonte: Moscati (2005).

O Aeroporto Internacional de São Paulo/Guarulhos, cuja frequência de pousos e decolagens é uma das maiores da América do Sul (OLIVEIRA, 2002), apresenta alta incidência de nevoeiros do tipo radiação (CABRAL, 1998). Nessa região, o nevoeiro de radiação ocorre preferencialmente no outono e no inverno austral, apresentando máximo em junho (CABRAL, 1998). Apenas para destacar, em 1997, a frequência de nevoeiro nesse aeroporto foi da ordem de 107 horas ano<sup>-1</sup>, valor este comparável a outros aeroportos dos Estados Unidos da América (EUA), como os de Washington, New Orleans e Detroit (Tabela 1.4).

Tabela 1.4: Frequência anual de nevoeiros para os principais aeroportos dos EUA.

Aeroportos	Frequência de nevoeiro	Aeroportos	Frequência de nevoeiro
Seattle	240	Baltimore	81
Portland, Oregon	214	Atlanta	74
Anchorage	194	New York (JFK)	74
Salt Lake City	139	Oakland	73
Los Angeles	138	Cincinnati (Covington)	72
Houston (Hobby)	132	Dayton	68
Washington, D.C.	116	Chicago (O' Hare)	60
New Orleans	105	Kansas City (Mid Continent)	60
Detroit (Metro)	105	Columbus	59
Hartford (Bradley)	97	San Francisco	58
Milwaukee	84	Brubank	57

Fonte: Adaptado de Beckwith (1985).

A formação de nevoeiro pode resultar de diferentes causas e como é um fenômeno da camada limite, a qual apresenta grande variabilidade em tempo e em espaço, é difícil incluir na rotina operacional dos aeroportos uma previsão de ocorrência, extensão, duração e intensidade do fenômeno (CROFT et al., 1997). Dessa forma, no desenvolvimento de procedimentos para a previsão de nevoeiro, vários aspectos são levados em consideração, tais como o tipo de nevoeiro (de radiação, advecção, frontal), a região geográfica em que eles se formam (terra ou oceano), os processos microfísicos envolvidos, entre outros (WILLETT, 1928; FISHER E CAPLAN, 1963; KOZIARA et al., 1983; TURTON E BROWN, 1987; MUSSON-GENON, 1987; TREMANT, 1989; XINMEI et al., 1990; LEIPPER, 1995; CROFT et al., 1997; TEIXEIRA E MIRANDA, 2001). Também, uma previsão de nevoeiro acurada requer conhecimento da climatologia do fenômeno e das condições meteorológicas locais (MEYER E LALA, 1990). Assim, apesar dos progressos recentes, até o momento não existe nenhum método objetivo de previsão de ocorrência, extensão, duração e intensidade de nevoeiro eficiente em sua totalidade.

No Brasil, poucos estudos foram realizados sobre a ocorrência de nevoeiros no Aeroporto Internacional de São Paulo (CABRAL, 1998; OLIVEIRA, 2002). Também, há poucos estudos sobre métodos de previsão de nevoeiros em aeroportos brasileiros, tanto de ocorrência (LIMA, 1982; OLIVEIRA E FEDOROVA, 1998 d; OLIVEIRA, 2002), quanto de duração (LIMA, 1982). Lima (1982) propôs um método gráfico para a previsão da duração do nevoeiro de radiação no Aeroporto Salgado Filho, em Porto Alegre/RS para os meses de abril a setembro. Porém, o método não se mostrou tão rigoroso, pois leva em consideração apenas uma variável (hora de início do nevoeiro) e a observação das condições de equilíbrio associadas às posições das frentes em relação a Porto Alegre entre 0 h e 6 h local.

Diante do exposto, o objetivo desse estudo é desenvolver um método de previsão de duração de nevoeiro para o Aeroporto Internacional de São Paulo/Guarulhos para todos os meses do ano, através da Análise de Sobrevivência (REIS, 1995), que utiliza observação meteorológica de superfície no estudo do tempo de duração do nevoeiro e que gera um produto de fácil aplicação na rotina operacional dos aeroportos. Para tal, propõe-se:

- Determinar a climatologia horária, diária, mensal e anual das variáveis meteorológicas de superfície do Aeroporto Internacional de São Paulo/Guarulhos;
  
- Aprimorar o método de previsão de duração de nevoeiro como proposto por Reis (1995), para as condições do Aeroporto Internacional de São Paulo/Guarulhos, baseado em Análise de Sobrevivência.
  
- Determinar:
  - O modelo estatístico que melhor se ajusta aos dados de nevoeiro para o Aeroporto Internacional de São Paulo.
  - As variáveis meteorológicas que devem ser consideradas em um modelo de previsão de duração de nevoeiro para o Aeroporto Internacional de São Paulo.
  - Desenvolver e verificar o modelo de previsão de duração de nevoeiros para o Aeroporto Internacional de São Paulo, baseado em Análise de Sobrevivência, para os casos de nevoeiro ocorridos em um ano específico (não incluído nos cálculos).

## **2 REVISÃO BIBLIOGRÁFICA**

Como poucos estudos foram realizados sobre a ocorrência de nevoeiros no Brasil, bem como os métodos de previsão de ocorrência, duração e extensão do fenômeno, apresenta-se inicialmente uma revisão bibliográfica abordando os aspectos mais relevantes para a caracterização do fenômeno, com ênfase sobre o nevoeiro do tipo radiação. Em seguida, enfatiza-se os estudos de nevoeiro no Brasil em todos os seus aspectos.

### **2.1 CLASSIFICAÇÃO DE NEVOEIROS**

Há várias classificações para nevoeiros, de acordo com as condições meteorológicas que os produzem, a região geográfica em que eles se formam (terra ou oceano) ou os processos que afetam sua formação ou deslocamento (WILLET, 1928; GEORGE, 1951; HANDBOOK OF METEOROLOGY, 1973). A classificação mais utilizada é a de Willett (1928), no qual os nevoeiros são agrupados em termos de processos de formação e características, em duas classes distintas, a saber:

a) Nevoeiros de massa de ar, que ocorrem dentro de uma dada massa de ar como resultado de processos meteorológicos que afetam a massa de ar por toda sua extensão horizontal. Os nevoeiros de massa de ar subdividem-se nos tipos advectivos, radiativo e advectivo-radiativo;

b) Nevoeiros frontais, que ocorrem imediatamente antes ou depois da zona frontal. Estes tipos de nevoeiros são de natureza transitória e se concentram em faixas estreitas, movendo-se com as frentes. Os nevoeiros frontais subdividem-se nos tipos pré-frontal, frontal e pós-frontal.

Nevoeiros do tipo advectivo são produzidos pelo transporte de ar quente sobre uma superfície fria, ou pelo transporte de ar frio sobre uma superfície quente, característico de regiões costeiras, e ocorrem praticamente durante todo o ano

(WALLACE E HOBBS, 2006). Nesse tipo, têm-se os nevoeiros associados às brisas marítima e terrestre, nevoeiro de ar marítimo, nevoeiro de ar tropical e o nevoeiro de vapor.

Na maioria dos casos de nevoeiro associado à brisa marítima, flutuações na direção do vento, geralmente de natureza diurna, faz com que o ar proveniente do continente aquecido seja resfriado ao passar sobre a superfície fria do oceano. Se os ventos apresentarem intensidades de moderada à forte, a turbulência pode provocar um resfriamento abrupto nas camadas inferiores e se formarão nuvens estratiformes sob a inversão turbulenta. Entretanto, se o vento for fraco, pode se formar uma superfície de nevoeiro denso sobre o oceano. Este tipo de nevoeiro também ocorre sobre grandes corpos d'água no continente, embora não se estendam muito continente adentro. Como exemplos, citam-se os nevoeiros que ocorrem durante a primavera e o início do verão na área costeira da Nova Inglaterra e nos Grandes Lagos, nos Estados Unidos da América (EUA). O nevoeiro de ar marítimo se dá pelo resfriamento do próprio ar marítimo sobre uma corrente marítima fria, podendo ocorrer em qualquer lugar do oceano onde houver significativa diferença de temperatura. Esse tipo de nevoeiro ocorre mais freqüentemente próximo ao continente, onde se localizam a maior parte das águas frias oceânicas presentes nas correntes costeiras. Regiões onde se observa esse tipo de nevoeiro no verão são as áreas costeiras do Peru, do Chile (CERECEDA E SCHEMENAUER, 1991), do noroeste e sudoeste da África, na costa oeste dos EUA (PILIÉ et al., 1979; LEIPPER, 1994). O nevoeiro de ar tropical está relacionado ao resfriamento gradativo do ar tropical (quente) à medida que ele se move sobre o oceano em direção aos pólos (frio). Esse tipo de nevoeiro pode ocorrer no inverno sobre os continentes, onde o gradiente latitudinal de temperatura pode ser maior do que aquele sobre os oceanos. Por outro lado, a turbulência sobre o continente é maior do que sobre o oceano devido a rugosidade da superfície, o que pode tornar mais difícil a condensação direta como nevoeiro de superfície, exceto em casos de vento fraco. Sobre o mar, o nevoeiro pode se manter com ventos intensos mais facilmente do que sobre o continente. Entretanto, no inverno, o

ar tropical marítimo em movimento sobre o continente está sujeito a fortes processos de resfriamento radiativo, o que pode se tornar mais importante do que o próprio resfriamento pela advecção latitudinal. Nessas circunstâncias é difícil classificar o nevoeiro como sendo de ar tropical ou do tipo radiativo. Esse tipo de nevoeiro ocorre, por exemplo, no Golfo do Alasca; na porção leste do Oceano Atlântico Norte e na Europa Ocidental, nos quais toda a intrusão de ar tropical resulta em formação de nevoeiro ou em nuvens estratiformes. Nevoeiros de vapor são formados quando o ar frio com baixa pressão de vapor passa sobre água relativamente quente. Em geral esses nevoeiros são rasos, da ordem de 15 m a 30 m, porém, espessos o bastante para interferir na navegação ou nos vôos sobre o mar. Ocorre também sobre rios, quando o ar foi resfriado por radiação. Essas ocorrências são comuns no outono, quando a água ainda está quente e o ar já está sendo resfriado, sendo verificado, por exemplo, no Lago Michigan/EUA.

O nevoeiro de radiação é freqüente em várias regiões, como por exemplo, na América do Sul (CABRAL, 1998; OLIVEIRA, 2002; LIMA, 1992), na Austrália (<http://www.auf.asn.au/meteorology/section3.html>), em Elkins-W. VA. (HOOVER, 1950), em New York (MEYER E LALA, 1990), entre outras. As condições necessárias para a formação do nevoeiro radiativo são encontradas em massas continentais em transição, ou seja, anticlones velhos que se tornam termicamente estáveis (WILLET, 1928), pois apresentam subsidência que não permite a formação de nuvens (PETERSEN, 1940). Segundo o *Handbook of Aviation Meteorology* (1994), as condições favoráveis para o nevoeiro de radiação são: 1) alta umidade relativa, de tal forma que é necessário um pequeno resfriamento para atingir a saturação; 2) pouca ou nenhuma nuvem, de forma que o calor é perdido pela radiação da superfície; 3) vento fraco, de forma que o resfriamento é confinado às camadas superficiais, mas suficiente para trazer turbulência. A intensidade do vento considerada mais favorável para a formação de nevoeiro é da ordem de 2 nós a 8 nós, embora possa haver nevoeiro mesmo com vento calmo. Os nevoeiros radiativos classificam-se nos tipos, de superfície e de “alta inversão”. O

nevoeiro de superfície é formado a partir de uma inversão térmica de superfície causada pelo resfriamento radiativo que ocorre durante uma única noite, sendo a inversão destruída durante o dia. Esse tipo de nevoeiro é raso, geralmente denso, ocorrendo em condições de céu claro. O termo nevoeiro de “alta inversão” significa que a inversão ocorre em uma camada mais espessa do que a do nevoeiro de superfície. Os nevoeiros radiativos do tipo alta inversão é basicamente um fenômeno de inverno, ocorre apenas sobre o continente e se forma quando o ar marítimo estaciona sobre um vale sob o domínio de um sistema de alta pressão persistente (TAYLOR, 1941, citado por HOLETS e SWANSON, 1981). Em alguns casos, este tipo de nevoeiro se torna uma nuvem estratiforme baixa durante o dia, alterando para condições de nevoeiro denso durante a noite. Regiões onde predominam nevoeiros radiativos do tipo alta inversão, durante o inverno, são o oeste dos EUA e a Europa Central (Byers, 1959) e no Vale Central da Califórnia (HOLETS E SWANSON, 1981). Holets e Swanson (1981) estudaram sete casos de nevoeiro de alta inversão no Vale Central da Califórnia verificando que, em geral, estes são caracterizados por condições meteorológicas uniforme temporalmente e espacialmente, apresentando velocidade do vento da ordem de  $5 \text{ ms}^{-1}$  ou menos, a camada de nevoeiro tem uma espessura de 100 m a 600 m, *lapse rate* adiabático úmido e temperatura máxima ambiente da ordem de  $10^0 \text{ C}$ .

O nevoeiro do tipo advectivo-radiativo se forma sobre o continente por resfriamento radiativo noturno de ar procedente do mar durante o dia. Ocorre principalmente no fim do verão e outono quando a água está relativamente mais quente e é, portanto, capaz de produzir alta temperatura do ponto de orvalho no ar sobrejacente e também quando as noites são longas o bastante para haver um resfriamento considerável.

Nos nevoeiros pré-frontais em frentes quentes, o efeito da precipitação em colunas estáveis de ar podem aumentar a  $T_d$  até que o nevoeiro seja formado sem resfriamento da camada de ar inferior. Estas condições são mais facilmente obedecidas no lado frio adiante de uma frente quente. No inverno,

massas de ar continental polar quando associadas com frentes quentes e precipitantes, geralmente apresentam nevoeiro ou nuvens estratiformes baixas por serem estáveis. Por outro lado, uma massa de ar marítima polar não é estável o bastante para permitir a formação de nevoeiro. Assim, a intensidade do vento é um fator importante.

Em nevoeiros pré-frontais em frentes frias também se formam nevoeiro através da umidade da precipitação frontal. Apenas frentes frias que se tornaram quase-estacionárias, usualmente orientadas na direção leste-oeste, apresentam extensas áreas de nevoeiro. Estas circunstâncias causam nevoeiro apenas se o ar frio for estável.

Existe uma variedade de formas nas quais os nevoeiros podem se formar temporariamente durante a passagem de uma frente em superfície. A mistura de massas de ar quente e frio na zona frontal pode produzir nevoeiro frontal se o vento for calmo e se ambas as massas estiverem perto da saturação antes da mistura. O súbito resfriamento do ar sobre a superfície úmida com a passagem de uma frente fria marcadamente precipitante pode causar um nevoeiro provisório ao longo da frente. No verão, especialmente em latitudes baixas, o resfriamento da superfície por evaporação de água de chuva pode ser tanto o resfriamento necessário quanto o suprimento de umidade necessário para a formação do nevoeiro. Basicamente, este tipo de nevoeiro se dá por abaixamento da base da nuvem durante a passagem da frente em condições extremamente úmidas.

## **2.2 – CARACTERÍSTICAS GERAIS DOS NEVOEIROS**

André e Marht (1982) ao estudarem os perfis verticais das localidades de Wangara e Voves nos Estados Unidos, verificaram que o nevoeiro é um fenômeno atmosférico que ocorre na baixa atmosfera e tem sua existência na Camada Limite Planetária noturna (CLP). Esta camada pode ser descrita como

estável ou turbulenta, implicando em uma discriminação de uma estrutura estratificada e turbulenta, no qual mostra a existência de um equilíbrio dinâmico, entre o resfriamento por radiação de onda longa (da superfície e da baixa atmosfera) e o termo de turbulência mecânica. A Camada de Mistura Turbulenta (CMT) que é observada nesta estrutura estratificada da CLP noturna apresenta o seguinte comportamento físico: em nevoeiros predominado pelo mecanismo de resfriamento por radiação de onda longa, a CMT varia entre 100 m e 300 m, bem abaixo da altura da inversão. Este perfil de resfriamento pode ser observado pela temperatura potencial apresentando um perfil exponencial. Em nevoeiros, nos casos no qual exista o aumento da turbulência mecânica (ventos fortes ou a existência de jatos de baixos níveis) a altura da CMT varia entre 500 m e 800 m ou até mais. Este perfil de resfriamento pode ser observado pela temperatura potencial apresentando um perfil linear. Nevoeiros destes tipos podem durar mais de doze horas, começando na noite anterior e terminando após o meio dia, no dia seguinte.

Meyer e Lala (1990) realizaram uma climatologia (1970-1979) de nevoeiro em Albany, Nova York, considerando apenas casos de nevoeiros intensos, definidos como aqueles em que a visibilidade é inferior ou igual a 400 m. Por estação do ano, os nevoeiros de radiação nessa região são mais prevalentes no outono boreal (com 65 % do total de casos), no verão (23,8 % dos casos), na primavera (8,4 %) e no inverno (2,8 %). Nenhum caso de nevoeiro de radiação ocorreu em dezembro. Em média, os nevoeiros duram de 3 h a 4 h. Analisando 143 eventos de nevoeiro de radiação em Albany, identificou-se 5 situações sinóticas associadas (Figura 1): 1) anticiclone, com centro de 1032 hPa (40 % de todos os casos), 2) crista principal, geralmente a nordeste da estação (~ 22 % dos casos), 3) crista pós-frontal (~ 16 % dos casos), 4) frente fraca ou dissipando localizada na vizinhança local (~12 %), 5) associado a um colo (~9 %). Condições favoráveis para a ocorrência de nevoeiro de radiação dos tipos 1 e 2, são aquelas em que o centro do anticiclone (ou o eixo da crista) em superfície relativo à localização da estação meteorológica, está localizado à leste ou sudeste da estação, e é desfavorável se o sistema estiver se

aproximando da estação de oeste. As direções médias do vento prevalecente durante nevoeiro são calmo (~40 % dos casos), de sul (~20 %), de norte (~15 %) e de nordeste (~13%). Essas direções se correlacionam bem com a localização do anticiclone.

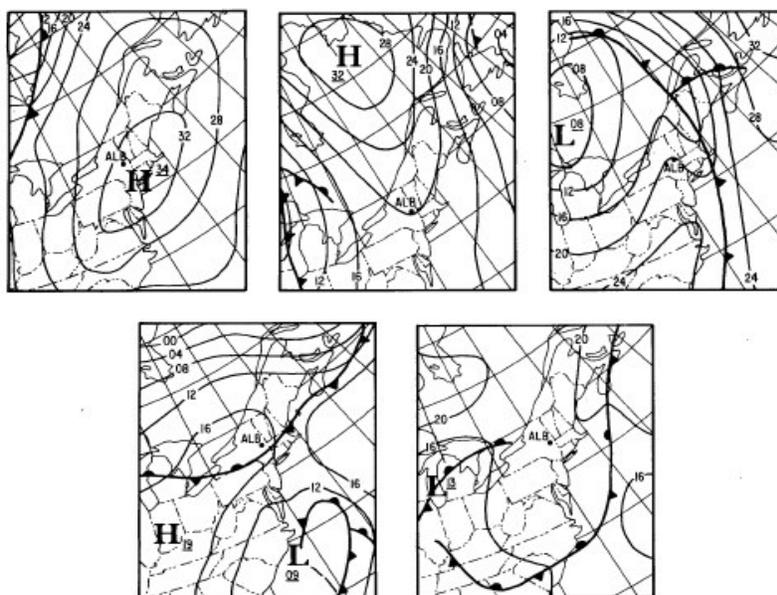


Figura 2.1 – Exemplos de situações sinóticas associadas com eventos de nevoeiro de radiação forte em Albany (1970-1979), para os tipos 1, 2, 3, 4 e 5, respectivamente.  
 FONTE: Meyer e Lala (1990).

Para o sul dos Estados Unidos (Jackson, Mississippi, Mobile, Alabama, New Orleans Moisant Airport, Louisiana), Croft (1997) fez um estudo climatológico (1948-1990) da frequência do nevoeiro. De 53 dias de ocorrência de nevoeiros, eliminou desse conjunto de dados os dias que a precipitação ocorreu na mesma hora do nevoeiro, em que os nevoeiros duraram menos de uma hora, e em que os nevoeiros tiveram poucos metros de extensão vertical. Com isso, restaram apenas 27 dias de nevoeiros, sendo 8 casos do tipo advectivo e 19 casos do tipo radioativo. Verificou que a maior concentração de ocorrência de nevoeiros dos tipos advecção e radiação aconteceram as 6 h (40 % dos casos), 7 h (27 % dos casos) e 8 h (30% dos casos). A maior ocorrência de nevoeiros ocorreu próximo ao litoral e durante o inverno. As temperaturas de ocorrência do nevoeiro variaram de  $-3^{\circ}\text{C}$  a  $22^{\circ}\text{C}$ . A altura média da camada

úmida foi da ordem de 350 m para o nevoeiro de radiação e 846 m para o nevoeiro de advecção. Também se observou a existência de uma camada muito seca acima da camada úmida (com depressão do ponto de orvalho ( $T - T_d$ ) em torno de 20°C), representando uma janela aberta para a radiação de onda longa. Para o nevoeiro de radiação, em 12 casos a  $T_d$  em superfície foi de 0,0°C, e nos sete casos restantes, foram de 1,0 °C, enquanto para o nevoeiro de advecção, em 7 casos a  $T_d$  foi de 0,0°C e com apenas um caso de  $T_d$  da ordem de 1,0 °C. Quanto à mudança vertical da direção do vento, observou-se que dos 19 casos de ocorrência do nevoeiro de radiação, 12 deles apresentaram giro horário (advecção fria no Hemisfério Norte - HN), 6 casos sem advecção e apenas em um caso houve giro anti-horário (advecção quente no HN). Dos 8 casos de ocorrência de nevoeiro de advecção, sete deles apresentaram giro anti-horário (advecção quente no HN) e em apenas um caso não houve advecção.

Altamirano (2001) realizou uma análise climatológica de 30 anos de dados (1968 a 1997) de nevoeiro para o Aeroporto Internacional Jorge Chavez Callao, em Lima, no Peru. Os principais resultados desse trabalho, foram: 1) Em 30 anos, foram identificados 329 eventos de nevoeiros; 2) o horário de maior ocorrência de nevoeiro foi entre as 22 Horas Local - HL (158 observações) e às 12 HL, com máximo ocorrendo às 7 HL (203 observações), enquanto no período das 13 HL até as 16 HL não se observou nenhum evento de nevoeiro; 3) a maior frequência média mensal de nevoeiros ocorreu entre os meses de março a junho, sendo o máximo observado em abril (3 casos/ano), e a menor frequência de ocorrência de nevoeiro foi no mês de novembro, com nenhum caso registrado; 4) a maior frequência anual de nevoeiro foi no ano de 1989 (41 casos), enquanto a menor frequência anual foi em 1969 e em 1992, ambos com um caso no ano; 4) a frequência média de duração dos nevoeiros foi de 4,1 horas.

Friedlein (2004) realizou uma climatologia (1995 até 2002) de nevoeiro denso no Aeroporto Internacional de O'Hare, Illinois/EUA, região que apresenta alta

incidência de nevoeiro dos tipos radiação e advecção. Nevoeiro denso foi definido pelo *National Weather Service* quando a visibilidade horizontal for igual ou menor do que 5/16 de uma milha. Os resultados principais desse estudo mostram que os nevoeiros densos apresentam maior frequência de ocorrência no inverno (55,8 % dos casos), seguidos do outono (26 % dos casos), da primavera (10,8 %) e do verão (7,4%). Durante todas as estações do ano, os nevoeiros intensos ocorrem mais frequentemente durante a madrugada e no início da manhã (Figura 2.2). Os nevoeiros densos do tipo advecção, de longa duração, são típicos de inverno e apresentam a maior quantidade de horas de nevoeiro.

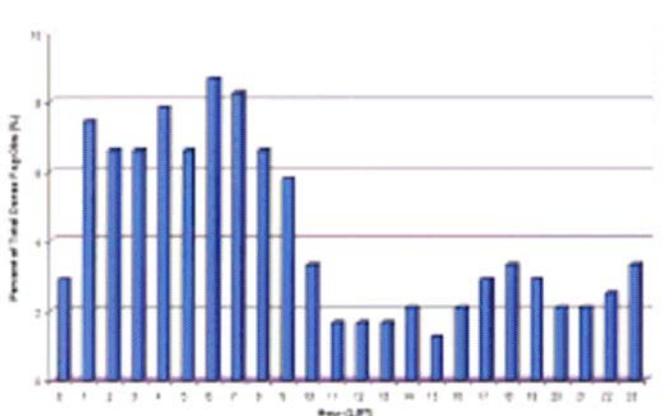


Figura 2.2 – Climatologia (julho de 1966 a abril de 2002) de nevoeiros densos no Aeroporto Internacional de O’Hare.  
 Fonte: Friedlein (2004).

Em relação ao processo de dissipação de nevoeiro, há dois processos envolvidos (CUESTA E MARTINEZ, 1993): por aquecimento na massa de ar com nevoeiro ou pela diminuição de vapor de água na massa de ar com nevoeiro. O aumento da temperatura da massa de ar é o processo mais freqüente na dissipação do nevoeiro e pode se dar pelos seguintes mecanismos enumerados por ordem de importância:

- i) A radiação solar produz um aumento na temperatura das gotículas que formam o nevoeiro ativando uma agitação turbulenta. De outro lado, aumenta a temperatura do solo, que por sua vez transporta

calor por radiação infravermelha à camada de ar mais próxima rompendo pouco a pouco a inversão.

- ii) Mistura por turbulência mecânica: O surgimento do vento provoca uma turbulência dinâmica, que força as camadas do nevoeiro a misturar-se com as camadas superiores, mais quentes e secas.
- iii) Aquecimento por compressão adiabática, originado por efeito da subsidência.
- iv) Transferência turbulenta de calor das camadas superiores. Também aqui um efeito dinâmico, vento na camada superior do nevoeiro, produz uma agitação pequena não adiabática, que transporta calor em direção ao solo.
- v) E, finalmente, advecção do nevoeiro sobre uma superfície mais quente. Este mecanismo exige que a condição dinâmica seja favorável para poder empurrar os bancos de nevoeiros (por exemplo, regime de brisas).

A dissipação do nevoeiro por diminuição da umidade na massa de ar com nevoeiro, ocorre devido a:

- vi) O fenômeno mais freqüente é o depósito de orvalho ou geada sobre o solo;
- vii) A transferência turbulenta em direção as camadas superiores (este processo, já citado anteriormente, tem valor duplicado, porque diminui a umidade e aumenta a temperatura).

### **2.3 – MÉTODOS DE PREVISÃO DE OCORRÊNCIA DE NEVOEIRO**

Hoover (1950) propôs um método de previsão de ocorrência de nevoeiro de radiação para Elkins, W. VA., região situada no Vale do Rio Tygart. Nesse método, foram consideradas as variações da velocidade e direção do vento

gradiente, a visibilidade, a depressão do ponto de orvalho ( $T_d$ ), e a temperatura do ar ( $T$ ), as quais foram combinadas estatisticamente dentro de um conjunto de cartas, através de um procedimento de correlação gráfica. Como principais resultados, têm-se: 1) visibilidade inferior a 1 milha é mais comum em julho, agosto e setembro; 2) maior frequência de nevoeiro ocorre com ventos de leste; 3) chance da visibilidade decrescendo a valores inferiores a 1 milha ocorre quando a velocidade do vento gradiente aumenta.

Duynkerke (1991) usou um modelo unidimensional que incorpora muitos processos físicos importantes (radiação, turbulência, vegetação e propriedades do solo) para prever a ocorrência de nevoeiro na Torre de Cabauw, Holanda. Das simulações com o modelo, verificou-se que o resfriamento radiativo para céu claro é importante para a formação do nevoeiro, e quando este processo é desconsiderado no modelo, o nevoeiro não se forma durante a noite.

Croft (1997) utilizou o modelo estatístico *Model Output Statistic* (MOS) para construir um modelo conceitual para previsão operacional de extensão, intensidade e duração de nevoeiros densos na região sul dos EUA (Jackson, Mississippi, Mobile, Alabama, New Orleans Moisant Airport, Louisiana). Para tal, inicialmente, definiu os melhores preditores estatísticos para cada localidade considerada, por meio de coeficientes de correlação parcial. Considerou nessa análise dados de  $T$ ,  $T_d$ , estação do ano em que ocorreu o nevoeiro, ocorrência de precipitação observadas durante as 24 horas antecedentes as observações de nevoeiros, altura da camada úmida e as características de vento (advecção fria ou quente) para a superfície de 700 hPa, dados de radiossondagem, entre outros parâmetros derivados. Ressaltou que para prever nevoeiro mais precisamente, em termos de ocorrência, extensão, intensidade e duração, é necessário conhecer e diagnosticar os processos termodinâmicos, cinemáticos e microfísicos dos nevoeiros. O modelo quantificou adequadamente os dois primeiros processos, mas ainda precisa de ajustes no processo de microfísica, em particular na física de nuvens operacional.

Altamirano (2001) construiu um modelo conceitual de previsão de ocorrência de nevoeiro para o Aeroporto Internacional Jorge Chavez Callao, em Lima, no Peru. Para tal, inicialmente analisou o comportamento das variáveis meteorológicas visibilidade, direção do vento, vento zonal e meridional,  $T$ ,  $T_d$ , temperatura da superfície do mar, pressão atmosférica e UR, antes, durante e após a ocorrência de nevoeiro nesse aeroporto, utilizando o método dos Eventos Médios. Este método consistiu em determinar as características sinóticas médias do nevoeiro e usar esta informação para prever sua formação, desenvolvimento e dissipação. Este método mostrou como condições sinóticas típicas antes da ocorrência do nevoeiro:

- i) Diminuição na temperatura da superfície do mar de  $0,4^{\circ}\text{C}/\text{horas}$ ;
- ii) Diminuição da velocidade do vento da componente sul, entre  $10 \text{ km h}^{-1}$  e  $2 \text{ km h}^{-1}$ ;
- iii) Diminuição da pressão atmosférica em superfície de  $1 \text{ hPa}/6\text{horas}$  e uma mudança na direção e intensidade do vento em superfície da componente oeste;
- iv) Aumento de  $T_d$  em  $0,8^{\circ}\text{C}/6 \text{ horas}$ .

Também, para o Modelo Conceitual de Previsão de Nevoeiros, utilizou-se de informações sinóticas de superfície, campos de e mapas de tempo prognosticados e reanalisados da América do Sul, que mostraram que a ocorrência de nevoeiros está associada com:

- i) Domínio inicial do Anticiclone do Pacífico Sul (APS) frente as costas peruanas;
- ii) Movimento do APS em direção a sudoeste e posterior enfraquecimento;
- iii) Um cavado em altura avançando em direção ao norte da América do Sul.

O Modelo foi validado com um evento de nevoeiro analisado do dia 12 de outubro de 1997.

Teixeira e Miranda (2001) propôs um modelo de previsão de ocorrência e intensidade de nevoeiro para o Aeroporto de Lisboa usando um modelo de camada limite unidimensional. Nessa região o nevoeiro de radiação é relativamente comum, embora também haja casos de nevoeiro associado com padrões de advecção particulares. O modelo usa esquemas de parametrização para turbulência, radiação, condensação, e tamanho de gotas, além de método de elementos finitos para a discretização vertical das equações. O modelo foi usado para simular oito eventos de nevoeiro que ocorreu nesse aeroporto. Os resultados mostraram que o modelo proposto foi capaz de prever razoavelmente bem o ciclo de vida do nevoeiro (início, crescimento e dissipação). Entretanto, ressaltaram que para fins operacionais, se faz necessário acoplar esse modelo com um esquema de superfície e forçantes de grande escala, por exemplo, com a saída de um modelo global ou um modelo de previsão numérica de tempo de área limitada, como feito por Musson-Genon (1989).

## **2.4 – MÉTODOS DE PREVISÃO DE DURAÇÃO DE NEVOEIRO**

Reis (1995) construiu um modelo de previsão de duração de nevoeiro para o Aeroporto da Portela em Lisboa, Portugal, utilizando a Análise de Sobrevivência e, em particular, os Modelos de Cox e o de Regressão Exponencial. Para isso, usou dados horários de visibilidade horária, no período de 1 de janeiro de 1970 a 31 de dezembro de 1989. Com estes dados, definiu o início e a dissipação do nevoeiro, considerando o tempo quando a visibilidade for inferior a 1.000 m e quando a visibilidade for superior 1.000 m, respectivamente. Elaborou um programa em linguagem PASCAL para identificar as situações de nevoeiros, filtrando os valores de visibilidades menores que 1000 m. Na determinação do Modelo de Cox que melhor se

ajustou aos dados do Aeroporto de Portela utilizou o pacote estatístico *Statistica*. No Apêndice B, apresenta-se resumidamente uma descrição do modelo proposto por Reis. Os resultados foram considerados satisfatórios, especialmente devido a pobre disponibilidade de dados (apenas de visibilidade horária). A relevância desse trabalho está em sua aplicação prática, principalmente para utilização por Centros Operacionais, uma vez que gera tabelas e gráficos de fácil consulta e que permite acompanhar a situação de nevoeiro durante toda a sua ocorrência, em tempo hábil. Por outro lado, a utilização do método de análise de sobrevivência se mostrou adequado a estudos que envolvam dados censurados.

Leipper (1995) utilizou o método de previsão de duração de nevoeiro *Leipper Inversion Based Statistics* (LIBS) para região da costa da Califórnia e áreas vizinhas. Esse método faz a relação da Base de Inversão de temperatura (BI) com a previsão do nevoeiro. Para uma previsão de 24 horas, essa relação é feita da seguinte maneira:  $BI \approx 0$ , há 28% de probabilidade do nevoeiro ocorrer entre 1 HL e 6 HL;  $0 < BI \leq 250$  m, há 59% de probabilidade do nevoeiro ocorrer entre 23 HL e 5 HL;  $250 \text{ m} < BI \leq 400$  m, há 28% de probabilidade do nevoeiro ocorrer entre as 2 HL e às 6 HL;  $400 \text{ m} < BI \leq 800$  m, há 94% de probabilidade de não ocorrer o nevoeiro.

## **2.5 – OCORRÊNCIA DE NEVOEIROS NO BRASIL E MÉTODOS DE PREVISÃO**

Cabral (1998) realizou um estudo climatológico dos nevoeiros no Aeroporto Internacional de São Paulo, considerando várias fontes de dados, no período de 1969 a 1997. Como principais resultados, destacam-se: i) Após o término das obras do aeroporto em 1984 (a construção desse aeroporto foi realizada no período de 1980 a 1984), houve uma significativa tendência decrescente nos totais anuais de horas de nevoeiro. O ano de 1997 registrou 107 horas ano<sup>-1</sup>, sendo considerado o menor valor da série analisada (Figura 2.3 a); ii) no início da construção do Aeroporto de Guarulhos, os totais anuais de horas de

nevoeiro em Guarulhos apresentaram uma queda abrupta, a qual pode ter sido influenciada pela pavimentação e pela retirada de vegetação da área (Figura 2.3 b); iii) maior freqüência mensal média de nevoeiro nos meses de outono e inverno (Figura 2.3 c); iv) maior concentração de ocorrência de nevoeiro entre 1 HL e 7 HL, com um pico entre 6 HL e 7 HL (Figura 2.3 d); v) freqüência de maior duração de nevoeiro entre uma e três horas, podendo persistir até 14 horas (Figura 2.3 e).

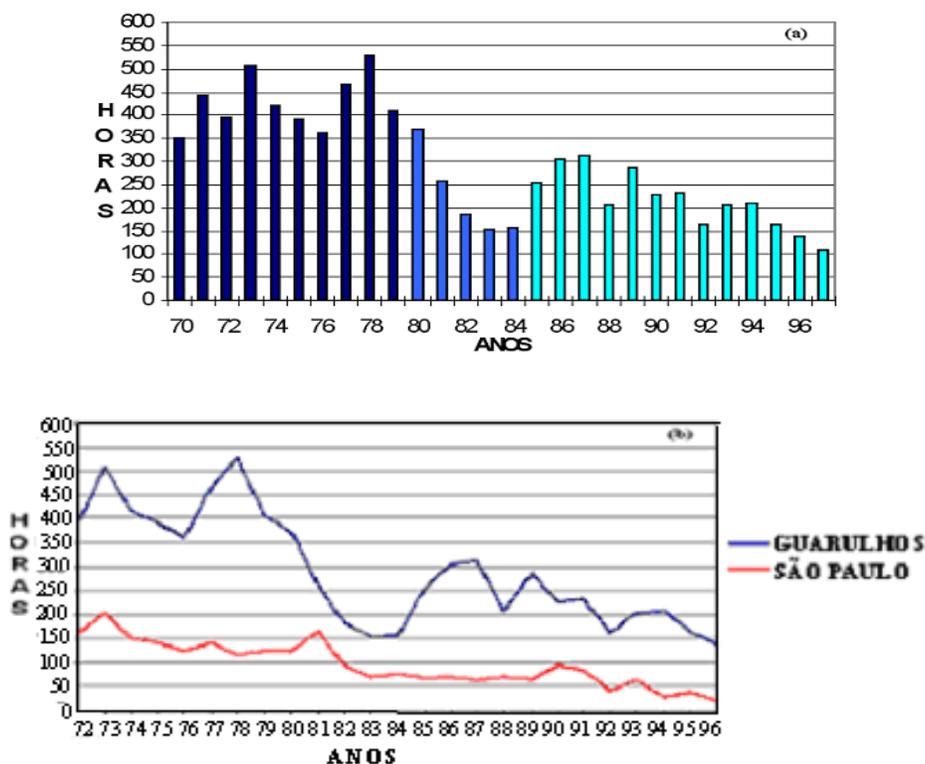


Figura 2.3 - a) Total anual de horas de nevoeiro nas estações meteorológicas de Cumbica (1969–1984) e de Guarulhos (1985–1997); b) Total anual de horas de nevoeiro nos Aeroportos de Guarulhos e Congonhas (1972-1997); c) Freqüência média mensal de horas de nevoeiro dos Aeroportos de Guarulhos e Congonhas (1980–1997); d) Horários de início de nevoeiro no Aeródromo de Guarulhos (1993-1997), e) Freqüência de duração do nevoeiro no Aeródromo de Guarulhos, no período de 1993 a 1997 (continua).  
 FONTE: Cabral (1998).

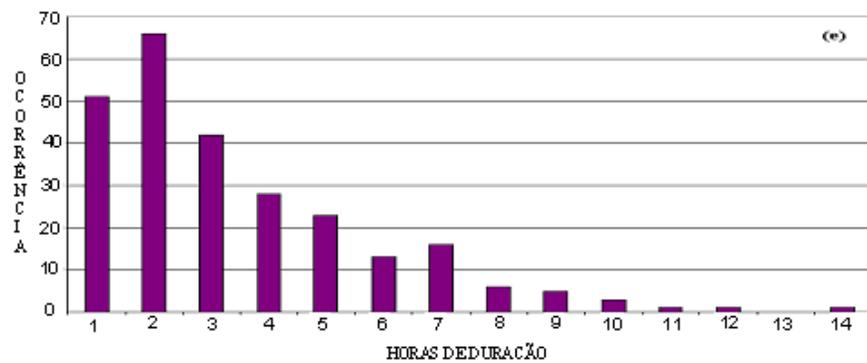
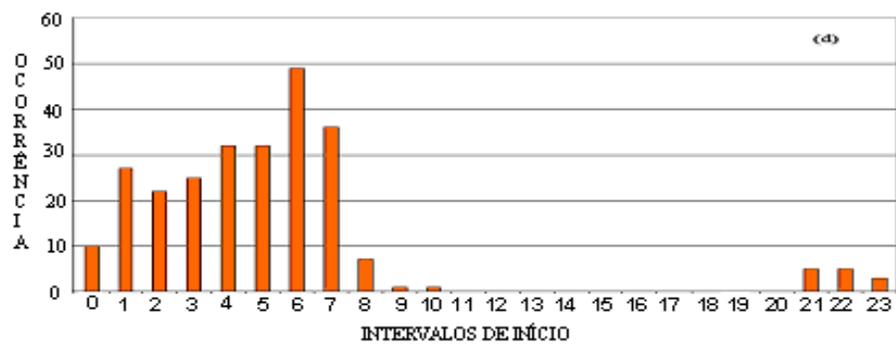
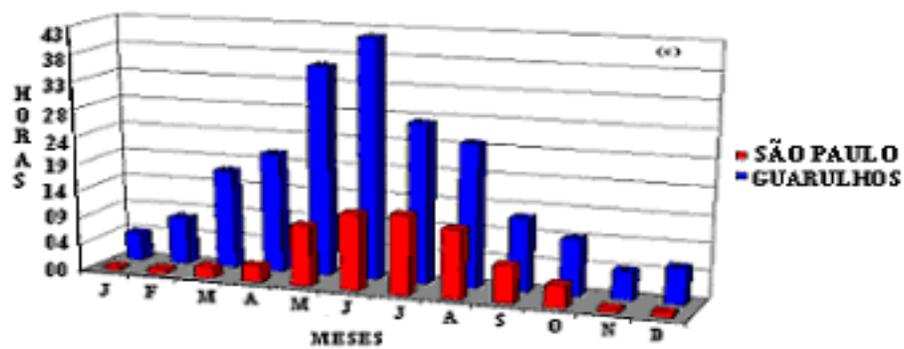
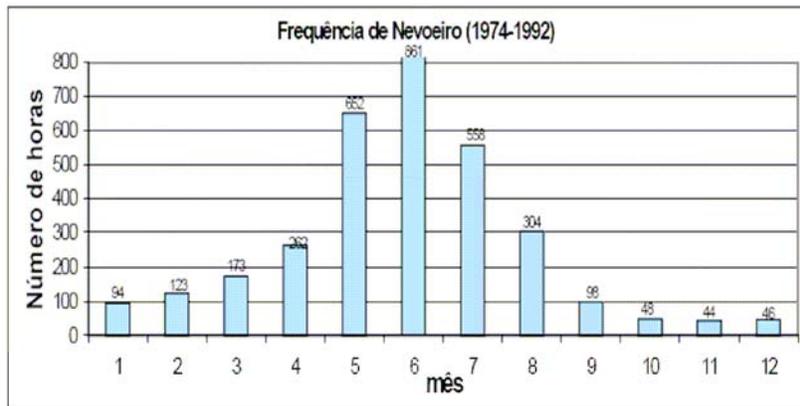


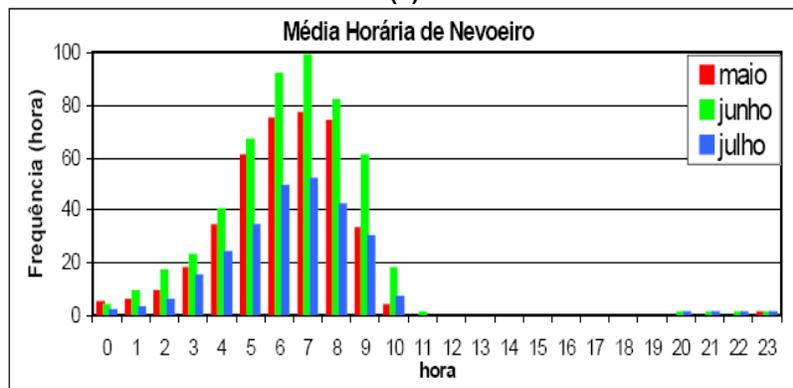
Figura 2.3 – Conclusão.

Perella e Sakuragi (1998) realizaram um estudo climatológico dos nevoeiros com dados do Aeroporto de São José dos Campos/SP, para o período de 1974 a 1992.

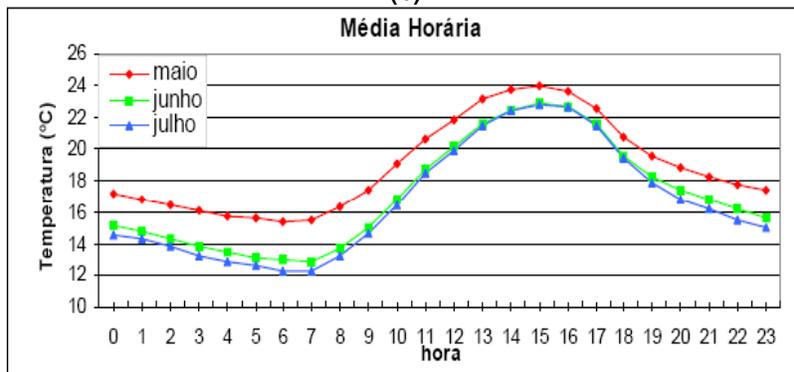
- i) O mês de maio apresentou maior ocorrência de nevoeiro do que o mês de julho (Figura 2.4 a), devido os ventos serem menos intensos nesse mês, embora julho apresente baixa temperatura e alta umidade;
- ii) Os horários de maior ocorrência de nevoeiro de radiação foram entre 5 HL e 9 HL, com máxima às 7 HL (Figura 2.4 b);
- iii) Na distribuição horária média da temperatura do ar para os meses de maio, junho, e julho, observou-se que os menores valores foram registrados entre 5 HL e 7 HL, sendo registrada no mês de julho a menor temperatura, da ordem de 12,3°C (Figura 2.4 c);
- iv) Na média horária, os maiores valores de UR ocorreram entre 4 HL e 7 HL durante os meses de maio, junho e julho, com máxima UR verificada as 7 HL no mês de junho, da ordem de 96,9% (Figura 2.4 d). Sugere-se que os fatores determinantes para o aumento da UR e a redução na duração dos nevoeiros foram a construção de represas, a industrialização e mudanças das propriedades do solo;
- v) A velocidade média horária dos ventos em superfície analisada para os meses de maio, junho e julho indicam fracas velocidades para as primeiras horas do dia, principalmente entre 3 HL e 7 HL. A velocidade do vento nos meses de maio e junho apresentaram valores menores do que o do mês de julho. O menor valor de velocidade do vento no mês de junho foi de 0,8 ms<sup>-1</sup> as 6 HL e no mês de maio foi de 0,9 ms<sup>-1</sup> as 3 HL (Figura 2.4 e).



(a)

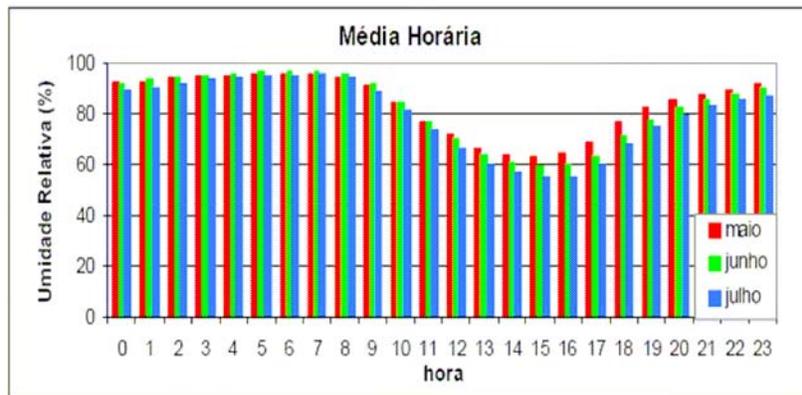


(b)

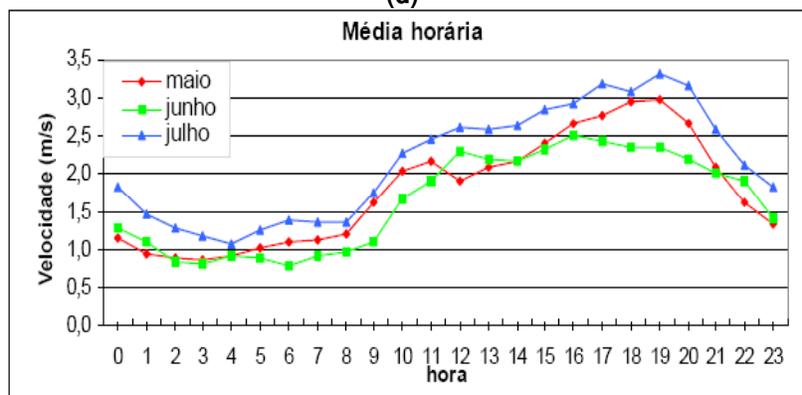


(c)

Figura 2.4 – a) Ocorrência de nevoeiros (expressa em horas) em São José dos Campos, SP, para o período de 1974 a 1992; b) Distribuição horária média de maior ocorrência de nevoeiros; c) Distribuição horária média da T; d) Distribuição horária média de UR; e) Distribuição horária média da velocidade do vento. FONTE: Perella e Sakuragi (1998).



(d)



(e)

Figura 2.4 – Conclusão.

Em relação aos métodos de previsão de nevoeiro, Lima (1982) propôs um algoritmo para previsão de ocorrência de nevoeiro de radiação para o Aeroporto de Porto Alegre, no Rio Grande do Sul. Verificou que para que ocorram os nevoeiros, o anticiclone em superfície deve estar à sudeste, leste ou nordeste da região em estudo. Com informações da carta sinótica de superfície das 1200 UTC e informações sobre a posição do sistema de alta pressão, o qual deve encontrar-se sobre a cidade de Porto Alegre, e das frentes, pode-se fazer uma previsão da ocorrência, ou não, de nevoeiro para o Aeroporto de Porto Alegre.

Machado (1993), através de um modelo de circulação de brisas vale-montanha utilizado para previsão de nevoeiro em São Paulo, verificou que:

- (i) Necessita de uma inversão de umidade em baixos níveis;

- (ii) Ocorre com mais frequência após um episódio de entrada de brisa marítima em São Paulo;
- (iii) Não ocorre quando o jato de baixos níveis se estabelece e quando há ventos de noroeste (traz ar seco e quente para a região);
- (iv) Tem maior ocorrência no final de outono e início do inverno.

Oliveira e Fedorova (1998 a, b, c e d) utilizaram cartas sinóticas para identificar os padrões de situações sinóticas favoráveis à ocorrência de nevoeiro em Pelotas/RS. Como situação sinótica favorável, tem-se que os nevoeiros se formam quando a cidade de Pelotas está à oeste/sudoeste do anticiclone em superfície com curvatura ciclônica, há valores elevados de UR, declínio na temperatura do ar e velocidades do vento baixa.

Piva e Fedorova (1998 a, b e c), analisando as condições sinóticas no nevoeiro de radiação para Porto Alegre, encontraram a maior frequência de localização do anticiclone sobre o estado, ou à leste da região. Em média, a pressão no núcleo do anticiclone foi da ordem de 1022,7 hPa (para os dias anteriores) e 1024,2 hPa (para os dias com nevoeiro). O anticiclone se intensificou em 45,5% dos casos, permaneceu com a mesma intensidade em 36,4% e enfraqueceu em 18,1% dos casos. Piva e Fedorova (1998 b) também analisaram os perfis verticais em Pelotas e encontraram dois tipos de inversão: Tipo I e Tipo II. A inversão do Tipo I apresentou uma altura média de 946 hPa, a diferença entre as temperaturas do topo e da base da inversão foi de 8,5°C, com inclinação (ângulo entre a distribuição de temperatura e a isóbara, no Diagrama Skew-T log P) média de 14,3°C e um gradiente vertical de temperatura de 1,6°C/100 m. Os déficits do ponto de orvalho na base da inversão foram pequenos, na média, da ordem de 1,2°C, no topo da inversão foi de 11,5°C e para o resto da atmosfera o déficit médio ficou em 15,5°C. A inversão do Tipo II apresentou uma altura média de 958 hPa, a diferença entre as temperatura do topo e da base da inversão foi de 1,5°C, em média de 22,6° e um gradiente vertical de temperatura de 0,5°C/100 m. A depressão do ponto de orvalho na base da inversão foi pequena, da ordem de 0,3°C e para o resto

da atmosfera ficou em 17,7°C. Os resultados obtidos por Piva e Fedorova (1998 b) concordaram parcialmente com Lima (1982), que também encontrou uma camada úmida superficial, porém, com uma altura menor e mostraram resultados similares aos resultados obtidos por Croft (1997).

Oliveira e Fedorova (1998, a-d) propuseram um método de previsão de nevoeiro em Pelotas. Analisaram 91 dias com nevoeiro, 89 dias anteriores a estes e igual número de dias sem nevoeiro e anteriores aos mesmos. Através de uma análise estatística de correlações lineares entre várias variáveis determinaram os melhores preditores. Analisaram as variáveis T, UR, velocidade do vento e pressão, em diferentes horários, no período de janeiro de 1995 até abril de 1996. Baseado nos resultados da análise da correlação, concluiu-se que: i) O horário das 10 HL foi utilizado para previsão do nevoeiro devido ao valor alto do coeficiente de correlação obtido em relação às demais variáveis; ii) A pressão atmosférica foi desconsiderada como preditor para a previsão de ocorrência de nevoeiro, tanto nos dias com nevoeiro como nos dias sem nevoeiro; iii) Todas as correlações foram positivas e os melhores coeficientes de correlação foram obtidos com duas ou três variáveis, ou seja, com a T e UR, T e velocidade do vento, UR e velocidade do vento.

Lima (1982) desenvolveu um algoritmo para previsão de nevoeiro de radiação para a cidade de Porto Alegre. Utilizou-se como preditores o vento (direção e velocidade) em 500 hPa, a T do ar, a depressão do ponto de orvalho ( $T - T_d$ ), a diferença de pressão em 24 horas e a nebulosidade do horário das 19 horas local, quando é elaborada a previsão de ocorrência ou ausência de nevoeiro. Nessa região, verificou-se que nenhum nevoeiro ocorre de abril a setembro, quando o valor máximo da pressão reduzida ao nível médio do mar no dia anterior for menor que 1013 hPa. Analisando-se dois perfis verticais de T e  $T_d$ , obteve-se em um dos perfis uma camada úmida ( $T - T_d \leq 6^\circ\text{C}$ ) abaixo de 750 hPa e uma camada seca ( $T - T_d > 6^\circ\text{C}$ ) que se estende de 750 hPa até a altura final da sondagem. Para o segundo perfil, a diferença em relação ao primeiro foi a presença de uma camada úmida ( $T - T_d \leq 6^\circ\text{C}$ ) acima da camada seca. O

modelo de situação sinótica favorável sugerida para a ocorrência de nevoeiro, caracteriza-se pela estabilidade da massa de ar, umidade elevada na camada abaixo de 750 mb ( $T - T_d \leq 6 \text{ }^\circ \text{C}$ ) e ar seco acima deste nível, o que é identificado com o uso do diagrama SKEW T log P. Para complementar essas informações, deve existir cobertura de nuvens durante o dia e céu claro e ventos fracos à noite, com possibilidade de ocorrência de nevoeiro quando a direção do vento variar de leste a sudeste com intensidade abaixo de  $1,4 \text{ ms}^{-1}$  e para direção oeste com intensidade inferior à  $0,6 \text{ ms}^{-1}$ . Também, Lima (1982) notando que os nevoeiros que se formavam nas primeiras horas da madrugada tinham maior duração dos que os se formavam nas primeiras horas da manhã, desenvolveu um diagrama que tinha como ordenada a hora de início dos nevoeiros e como abscissa as durações, em seguida, computou-se as observações e fez um ajuste dos pontos do diagrama, chegando-se, então, a um método analítico que emprega a seguinte equação:

$$d = - 1,4i + 10,8. \quad (1)$$

onde d representa a duração do nevoeiro e i o início do nevoeiro, ambos em décimos de horas. Tal equação deve ser empregada apenas para nevoeiros que começam antes das 7,7 h, pois um valor acima deste, resultaria numa hora de duração negativa. O método foi considerado pouco rigoroso, e sugeriu-se para uma previsão de duração de nevoeiro mais acurada, observar as condições de equilíbrio associadas às posições das frentes em superfície em relação a Porto Alegre, entre 0 HL e 6 HL. Também, para uma melhor previsão de nevoeiro, sugeriu-se analisar os casos de condições de equilíbrio para cada mês particular. Foram propostos três procedimentos para casos de condições de equilíbrio, a saber:

- 1) Estável – frentes em superfície abaixo da cidade de Vitória e abaixo da cidade de Bahia Blanca. Em seguida, usava-se o valor obtido no gráfico ou equação.

2) Mais Estável – Posições das frentes em:

Vitória – somar 0,7 horas

Caravelas – somar 1,5 horas

Salvador – somar 2,5 horas

3) Menos Estável – Posições das frentes em superfície em:

Bahia Blanca – subtrair 0,7 horas

Buenos Aires – subtrair 1,5 horas

Bagé – subtrair 2,5 horas

Oliveira (2002) elaborou um método estatístico baseado em equação de regressão linear múltipla, para a previsão de nevoeiro de radiação no Aeroporto Internacional de São Paulo, considerando parâmetros meteorológicos favoráveis à formação do nevoeiro de radiação, juntamente com o índice de estabilidade de nevoeiro *Fog Stability Index* (FSI). Foram analisadas três equações de regressão, escolhendo-se como a equação que melhor se ajustou aos dados, dada por:

$$VIS = 9415 - 345 * Q\_HZBR - 233 * Q\_CLR \dots \dots \dots (2)$$

onde VIS representa a visibilidade prevista em metros, das 0000 UTC; Q\_HZBR a quantidade de horas de névoa seca e/ou úmida e Q\_CLR a quantidade de horas de céu claro e/ou somente nuvens da camada mais alta acima de 25000 pés ocorridas, ambas no período de 12 horas passadas a contar de 0000 UTC, e extraída das mensagens METAR. Tal equação restringe-se apenas para o período de maio a agosto e pode ser aplicada em qualquer condição, pois as variáveis HZBR e Q\_CLR assumem valores amostrais de 0 a 12, o que resultará sempre num valor positivo de visibilidade prevista. Sua validação foi feita com o emprego de um arquivo de dados da

mensagem METAR do Aeroporto Internacional de São Paulo para o período de 1 de maio até 31 de agosto de 2001. Segundo o autor, foram obtidos resultados com um total de 80% de acertos. Para ampliar a análise da tomada de decisão sobre o alerta, ou não, de ocorrência de nevoeiro, utilizou-se um fluxograma de decisões (Figura 2.5). Este fluxograma conta com a verificação da presença, ou não, de instabilidade ou nebulosidade baixa, visto que o nevoeiro de radiação se forma diante de total estabilidade e predominância de céu claro, ou somente da presença de nuvens acima de 25.000 pés presentes no momento atual considerado.

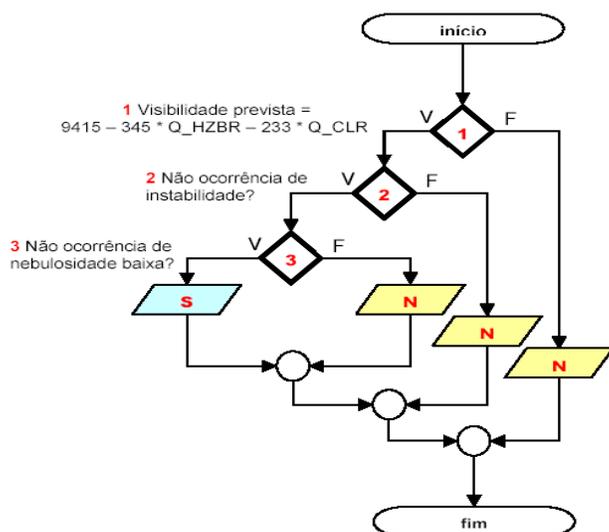


Figura 2.5 - Fluxograma de decisão para previsão do nevoeiro no Aeroporto Internacional de São Paulo.  
 FONTE: Oliveira (2002).

O fluxograma apresentado na Figura 2.5 deverá ser executado, considerando as seguintes etapas:

Passo 1: Se a visibilidade prevista for menor que 8.000 m vá para o passo 2, caso contrário, não prevê nevoeiro de radiação e pare de executar o fluxo.

Passo 2: Consiste na análise das mensagens METAR do Aeródromo de Guarulhos das 0000 UTC (21 horas, hora de Brasília) e de 12 horas passadas

para verificação da ocorrência ou não de instabilidade no período. Se não ocorreu chuva, chuveiro, pancada de chuva ou trovoadas vá para o passo 3, caso contrário não prevê nevoeiro de radiação e pare de executar o fluxo.

Passo 3: Consiste na análise do METAR do Aeródromo de Guarulhos das 0000 UTC (21 horas, hora de Brasília) para verificação da ocorrência ou não de nebulosidade baixa (abaixo de 2000 pés). Se não ocorreu nebulosidade baixa então prevê nevoeiro de radiação, caso contrário, não.

Segundo Reymann et al. (1988), o *Fog Stability Index* (FSI) foi criado e testado por Herr Harald Strauss, na Alemanha, em 1970, e pode ser obtido pela aplicação da seguinte equação:

$$FSI = 4*TS - 2*(T850 + TDS) + W850 \dots\dots\dots(3)$$

Onde TS, T850, TDS e W850, representam respectivamente, a temperatura do ar da superfície, a temperatura no nível de 850 hPa, a temperatura do ponto de orvalho na superfície, expressa em °C, e a velocidade do vento no nível de 850 hPa, expressa em nós. As variáveis deveriam ser obtidas da radiossondagem das 00Z e o FSI obtido indicaria a probabilidade de ocorrência de nevoeiro nas próximas 12 h, com baixa possibilidade de ocorrência de nevoeiro de radiação se o índice for maior que 55, moderada possibilidade de ocorrência se o índice for entre 31 e 55 e, finalmente, alta possibilidade de ocorrência quando o índice for menor que 31.

Há alguns modelos regionais no Brasil que têm elaborado uma previsão de nevoeiro para várias regiões do Brasil, podendo-se citar o ETA/CPTEC-INPE ([www.cptec.inpe.br](http://www.cptec.inpe.br)), o Modelo de Mesoescala da *Pennsylvania State University/National Center for Atmospheric Research* (MM5), disponível na Universidade Federal do Rio de Janeiro (UFRJ) e o *Regional Atmospheric Modeling System* (RAMS), disponível na Universidade de São Paulo (USP), (<http://www.master.iag.usp.br>). Recentemente, França (2008) realizou uma

avaliação da metodologia de previsão de nevoeiro utilizada no Modelo ETA/CPTEC-INPE, considerando as regiões Sul e Sudeste do Brasil para a validação dessa metodologia. Sob o aspecto regional, os resultados mostraram-se satisfatórios quando comparado com o produto de detecção de nevoeiros e stratus a partir de imagens GOES-12. Entretanto, sob o aspecto pontual, mostraram-se deficientes em detectar as ocorrências de baixa visibilidade horizontal e nevoeiros nos aeroportos estudados.

Muitos progressos têm sido feitos ao longo das últimas décadas para prever a ocorrência, intensidade, extensão e duração de nevoeiros em muitas regiões afetadas pelo fenômeno. Entretanto, as dificuldades são inúmeras, podendo-se destacar, a falta de conjuntos de dados adequados e suficientes (em especial sobre os oceanos), vários processos complexos envolvidos e que não são adequadamente parametrizados (em especial os processos microfísicos associados), há vários tipos de nevoeiros (radiativo, advectivo, advectivo-radiativo, frontal), de forma que nenhum modelo numérico consegue incluir todos os tipos, disponibilidade de ferramentas adequadas nas instituições onde o modelo será desenvolvido (por exemplo, para muitos teste usados em modelos de regressão, a disponibilidade de pacotes estatísticos como o SAS, apresentado em detalhes no Apêndice D), entre outros. Também, ressalta-se que uma utilidade primordial desses modelos de previsão de nevoeiro é que os resultados gerados sejam de fácil aplicação para uso em Centros Operacionais, os quais necessitam de informação rápida e simples, para a tomada de decisão para a divulgação de alertas aos órgãos que utilizam dessa informação, e para a sociedade em geral.

### **3 MATERIAIS E MÉTODOS**

### 3.1 – ÁREA DE ESTUDO

O Aeroporto Internacional de São Paulo localiza-se em uma planície no vale do Rio Baquirivu-Guaçu, afluente da margem direita do Alto do Tietê (ROTAER, 2005). A Figura 3.1 a mostra a localização desse aeroporto e arredores, destacando o Rio Tietê (próximo da cidade de Guarulhos) e a Serra da Cantareira (próximo de Mairiporã), a 1.100 m acima do nível do mar. As coordenadas geográficas do Aeródromo de Guarulhos são 23°26'06''S, 046°28'22''W e 749 m de altitude. A análise da Figura 3.2 a apresenta a topografia da região em torno do Aeroporto Internacional de Guarulhos, destacando a Serra do Mar, que atinge 800 m acima do nível do Mar, e o Vale do Rio Tietê, com altitude de 700 m acima do nível do mar. Na Figura 3.2 b, apresenta-se a localização do Aeroporto Internacional de Congonhas (23°37'32''S e 46°39'21''W), situado na área urbana do Município de São Paulo e posicionado em uma esplanada tabular suavizada no nível mais elevado das colinas pliocênicas (710 a 810 m), a sul-sudoeste do bairro do Jabaquara, 10 km ao sul da Praça da Sé (AB'SABER, 1975; CABRAL E OLIVEIRA, 1998).



Figura 3.1 - Mapa da localização da pista do Aeroporto Internacional de São Paulo, em Guarulhos (indicado pelo círculo azul) e arredores.  
FONTE: Google Earth.

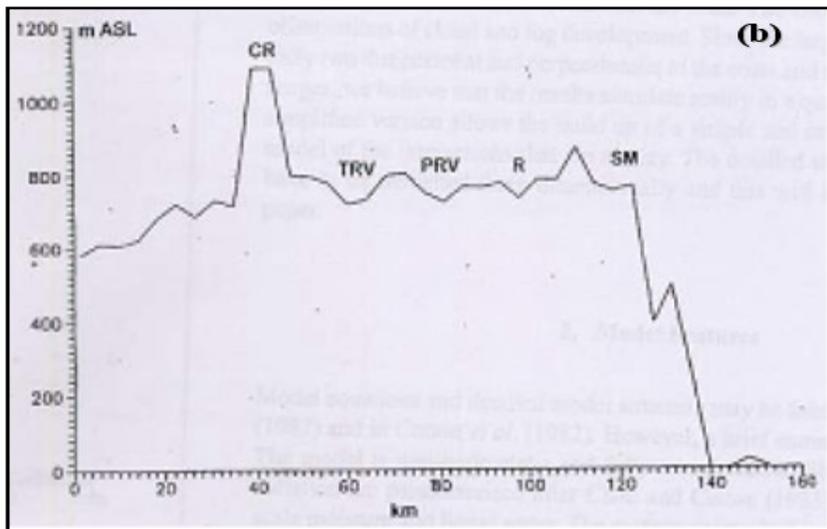
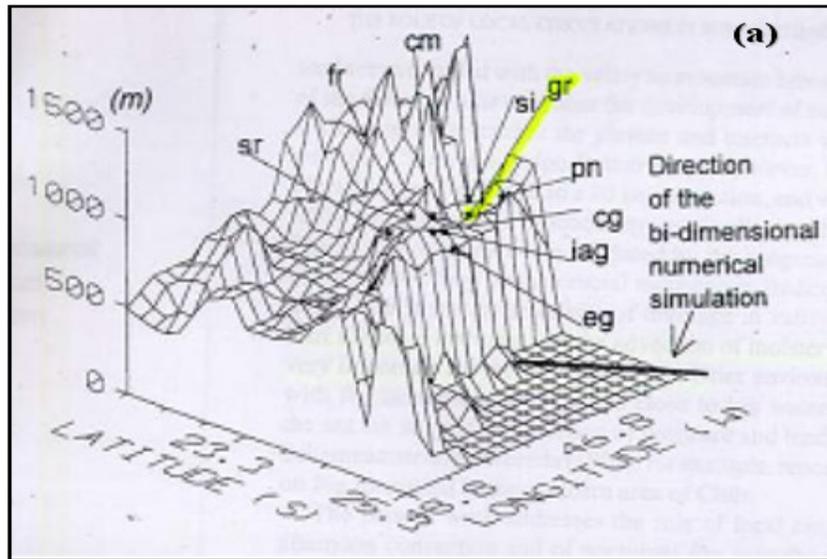


Figura 3.2 - a) Topografia da região em torno da cidade de São Paulo; b) Topografia de uma sessão perpendicular à costa, na direção noroeste-sudeste (Figura 3.2 a), com indicação de algumas características geográficas. As siglas sr, fr, si, pn, cg, eg, cm, gr e iag indicam a localização das estações de São Roque, Franco da Rocha, Santa Isabel, Ponte Nova, Aeroporto de Congonhas, Embu-Guaçu, Campo de Marte (Aeroporto de Guerra), Aeroporto Internacional de São Paulo e Estação do IAG/USP, respectivamente. As siglas CRV, TRV, PRV, R e SM significam Crista da Cantareira, Vale do Rio Tietê, Vale do Rio Pinheiros, Reservatório de água e Serra do Mar, respectivamente.  
 FONTE: Silva Dias e Machado (1997).

### 3.2 - DADOS

Foram utilizados dados horários dos Códigos IEPV 105-25 e IEPV 105-78 do Aeroporto Internacional de São Paulo. Esses dados foram obtidos através do convênio firmado entre a Empresa Brasileira de Infra-Estrutura Aeroportuária (INFRAERO) e o Instituto Nacional de Pesquisas Espaciais/Centro de Previsão de Tempo e Estudos Climáticos (INPE/CPTEC), em processo iniciado em 08/11/2005, assinado em 02/05/2007, e publicado no Diário Oficial da União em 18 de maio de 2007 (RD nº 01.01.010.0/2007). A vigência do convênio é de 02/05/2007 a 02/05/2009. Considerando que o conjunto de dados da INFRAERO anteriores a novembro de 1987 estava disponível apenas em papel, o que implicaria também em digitalizá-los e fazer o controle de qualidade, não foi utilizado nesse estudo devido o tempo curto para a elaboração dessa dissertação. Outra razão para a não utilização desse período de dados foi que em função da construção desse aeroporto (1982-1984), as séries temporais de todas as variáveis dessa estação meteorológica sofreram uma variação abrupta na quantidade de horas com nevoeiro, considerando o período anterior e posterior à construção do aeroporto, devido a pavimentação e a retirada da vegetação local (CABRAL, 1998). Isso poderia afetar os resultados com o modelo de previsão de duração de nevoeiro a ser elaborada.

Assim, pelo exposto acima, o período de dados horários utilizado foi de 1 de janeiro de 1988 até 31 de dezembro de 2005 e as variáveis disponíveis são: data e hora local, vento à superfície (direção e velocidade), visibilidade mínima (janeiro de 1988 a novembro 2004), visibilidade predominante (dezembro 2004 a dezembro 2005), Pressão reduzida ao Nível médio do Mar (P), temperatura do bulbo seco ( $T_w$ ), Temperatura do Ponto de Orvalho ( $T_d$ ), UR e precipitação. Vale ressaltar que apenas algumas dessas variáveis foram consideradas para a elaboração do método de previsão de duração de nevoeiro. A visibilidade mínima é definida como a menor visibilidade observada considerando todos os setores e é registrada em incrementos de 50 m até o valor de 500 m, em incrementos de 100 m de 500 até 5.000 m, e em incrementos de 1.000 m acima do valor de 5.000 m (ICA 105-7, 2000), enquanto a visibilidade

predominante é aquela que cobre pelo menos a metade do horizonte em setores contíguos, ou não (MCA 105-10, 2006).

O controle de qualidade utilizado consistiu em, inicialmente, uniformizar o formato dos dados diários, uma vez que no período de janeiro de 1988 a dezembro de 1999 estes se encontravam em planilhas Excel e, a partir desse ano até 2005, no formato do sistema Infomet. Assim, o novo conjunto de dados foi transformado em planilhas Excel. Em seguida, checaram-se todos os dados da planilha Excel com os formulários originais em papel e consistidos (controle de qualidade feito pela INFRAERO). Nesse processo, verificaram-se e corrigiram-se erros de digitação, fez-se o preenchimento de falhas (para dado ausente ou errado) e, em muitos casos, corrigiram-se os dados que estavam fora da ordem horária. Esse trabalho foi executado nas instalações da INFRAERO, sob a supervisão de um funcionário desse órgão. No Apêndice A apresentam-se as séries temporais climatológicas (1988-2005) mensais das variáveis selecionadas.

### **3.3 - METODOLOGIA**

Para a identificação de ocorrências de nevoeiro, o Aeroporto Internacional de São Paulo utiliza o Código Geral de Tempo (CGP), que define nevoeiro levando em conta apenas a visibilidade inferior a 1.000 m, e nas planilhas codifica os casos de nevoeiro pelo número 4. Assim, foram identificados nevoeiros com UR menor que 95% nos anos de 1988 (janeiro a junho), 1998 (em todo ano), 1999 (em todo ano) e 2000 (janeiro a novembro), indicando assim uma possível falha no instrumento de mediação. Sendo assim, construiu-se dois conjuntos de dados para análise dos eventos de nevoeiro a partir do conjunto original de dados da INFRAERO, separando-se: 1) os casos de nevoeiros como identificados pelo Aeroporto Internacional de São Paulo, 2) considerando-se a definição dada para nevoeiro no *Handbook of Aviation*

*Meteorology* (1994), onde foram selecionados os eventos em que os valores de visibilidade foram menores que 1.000 m e a UR próxima a 100%, com um mínimo de 95%. Nessa etapa, também se identificou e se eliminou situações com ocorrências de fumaça, chuva forte, chuvisco ou trovoada, que reduzem a visibilidade a valores inferiores a 1.000 m, baseada em informações contidas nas planilhas do aeroporto. A partir desses dois conjuntos de dados, fez-se duas climatologias de nevoeiro para o período de 1998 a 2003, sendo denominadas aqui de climatologia A e climatologia B, respectivamente. O objetivo da climatologia A foi estritamente para comparação com outros estudos já realizados para o Aeroporto Internacional de São Paulo (por exemplo, Cabral, 1998). Apenas o segundo conjunto de dados (e que gerou a climatologia B) será utilizado na Análise de Sobrevivência, por atender a definição de nevoeiro dada pela WMO (1966). A duração do nevoeiro foi definida como a diferença entre a hora de dissipação e a hora de início do nevoeiro (Reis, 1995; Cabral, 1998).

Há poucos trabalhos na literatura que, em ocorrências de nevoeiros consecutivos, apresenta uma definição para classificar o evento como um evento único. Para essa situação, Reis (1995) definiu como um evento único de nevoeiro aqueles separados por uma ou duas horas onde a visibilidade foi superior a 1.000 m em qualquer período do dia, ou quando esse intervalo não excedia a quatro horas no período da noite. Além disso, Reis (1995) ressaltou que para que um nevoeiro se dissipe e outro se forme, é necessário que as condições atmosféricas no período de tempo que os separa se modifiquem consideravelmente. Assim, no período da noite é necessário que haja ventos fortes para dissipar o nevoeiro instalado, e no período da noite/madrugada, que o vento seja fraco ou mesmo inexistente para que um novo nevoeiro se forme. No presente trabalho, para situações de nevoeiros consecutivos, considerou-se como um evento de nevoeiro único quando o intervalo de tempo entre a dissipação do primeiro evento e a formação do segundo evento não excedeu 2 horas em qualquer período do dia ou da noite e a visibilidade não ultrapassou 3 km no período. Tais especificações foram baseadas na análise dos dados de

visibilidade do Aeroporto Internacional de São Paulo, onde se observou que na maioria dos casos analisados o intervalo entre a dissipação do primeiro evento e a formação do segundo, não apresentou visibilidade maior que 3 km em um intervalo de até 2 horas.

A próxima etapa consistiu na identificação do dia e hora de início do nevoeiro e a hora de dissipação de cada nevoeiro. A seguir, calculou-se o total mensal das durações dos nevoeiros, as freqüências de horas totais mensais e anuais de nevoeiros e também, o número de ocorrências de nevoeiro em intervalos de 1 hora, os quais foram apresentados na forma de gráficos.

### **3.4 – APLICAÇÃO DOS DADOS DO AEROPORTO INTERNACIONAL DE SÃO PAULO ÀS TÉCNICAS DE ANÁLISE DE SOBREVIVÊNCIA**

No Apêndice B, apresenta-se uma breve descrição da técnica de Análise de Sobrevivência, descrevendo sucintamente a metodologia de análise de sobrevivência para construir um modelo de previsão de duração de nevoeiro proposta por Reis (1995). Nos Apêndices C e D, apresentam-se as noções básicas sobre o pacote estatístico R (COLOSIMO E GIOLO, 2006) e sobre o pacote estatístico SAS (CENAPAD/UNICAMP-SP, 2007), respectivamente, que servirão de ferramentas gráficas para a apresentação dos resultados.

Para a aplicação dos dados de nevoeiro à Análise de Sobrevivência como proposto por Reis (1995), optou-se por dividir os dados em estratos. Essa escolha foi permitida devido ao tamanho amostral ser relativamente grande (697 casos de nevoeiros). Para cada um dos estratos, uma covariável pode assumir dois ou mais valores. Assim, definiu-se os limites de valores das covariáveis em cada estrato com o uso de *box-and-whisher* or *boxplot* (Wilks, 1996; Verzani, 2002), uma representação diagramática para visualização de dispersão de conjunto de dados em termos de percentis (representado por uma

caixa). Em sua forma tradicional, o *boxplot* mostra as variações completas dos dados, destacando o valor mínimo (representa o quartil 0), o valor máximo (valor do topo, limite do 4<sup>o</sup> quartil), entre esses dois tem-se a mediana (2<sup>o</sup> quartil e por definição representa 50% do conjunto de dados. Quando a mediana está deslocada do ponto médio da caixa, significa que os dados não são distribuídos simetricamente), e o 1<sup>o</sup> e o 3<sup>o</sup> quartis (representam os 25% mais baixos e os 75% mais altos do conjunto de dados, respectivamente). A Figura 3.3a apresenta um exemplo esquemático de *boxplot*, onde se observa que as bases inferior e superior da caixa têm alturas correspondentes ao primeiro e terceiro quartis da distribuição ( $Q_1$  e  $Q_3$ , respectivamente) e a linha horizontal no meio da caixa corresponde à mediana. Numa versão modificada de um *boxplot* também se pode representar a variabilidade dos valores extremos através de um *plot* esquemático (Figura 3.3 b), que considera quatro limiares:  $Q_3+3$  AIQ (limite externo superior),  $Q_3+1,5$  AIQ (limite interno superior),  $Q_1-3$  AIQ (limite interno inferior) e  $Q_1-1,5$  AIQ (limite externo inferior), onde a Amplitude Inter-Quartis (AIQ) é definida por  $AIQ = Q_3 - Q_1$ . Os valores que estão dentro dos limites internos são representados pelas extensões das linhas pontilhadas e os valores entre os limites internos e externos são chamados de pontos extremos e são representados individualmente por um círculo. Nesse trabalho, preparou-se e analisou-se *boxplot* para cada faixa de duração dos nevoeiros para cada mês de todos os anos, para cada faixa de hora de início e para cada variável considerada.

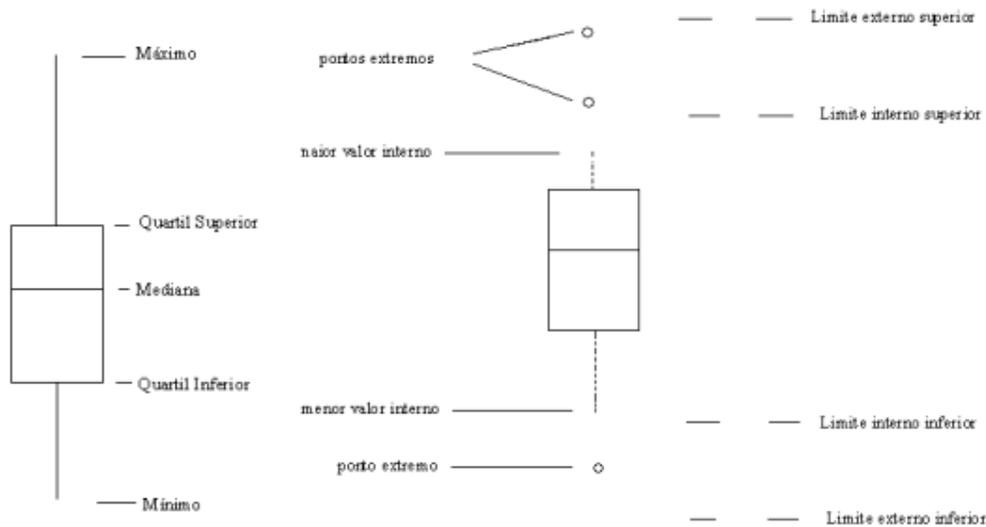


Figura 3.3 – Esquema ilustrativo de um: a) *boxplot*; b) *Plot* esquemático.  
 Fonte: Adaptado de Wilks (1996).

### 3.4.1 – ELABORAÇÃO DO MÉTODO DE PREVISÃO DE DURAÇÃO DE NEVOEIRO PARA O AEROPORTO INTERNACIONAL DE SÃO PAULO

Para elaboração do método de previsão de duração do nevoeiro para o Aeroporto Internacional de São Paulo, definiram-se as covariáveis utilizadas, discriminadas na Tabela 3.1. As valores dados às categorias foram definidos baseado no resultado dos *boxplots* de cada variável. Também, para a definição das categorias dos estratos de cada variável, optou-se pela dicotomização dos dados (isto é, transformou-se cada estrato em valores de zero e um). A seguir, construiu-se gráficos das estimativas da função de sobrevivência utilizando o estimador de Kaplan-Meier com o fim de modelar separadamente cada covariável com o tempo de duração do nevoeiro. Para definir as covariáveis que prosseguiram na análise, utilizou-se o critério de permanecer apenas com as que apresentem p-valor inferior a 0,25, no qual p-valor é a probabilidade de significância para qual uma hipótese nula  $H_0$  (testar a igualdade das curvas de sobrevivência) é rejeitada. Nesta etapa, utilizou-se os Teste de *Logrank* e

*Wilcoxon*. Uma interpretação para p-valor em termos de significância estatística é apresentada em Shimakura (2007), tal que:

$p \geq 0,10$  (não existe significância contra  $H_0$ )

$p < 0,10$  (fraca evidência contra  $H_0$ )

$p < 0,05$  (evidência significativa contra  $H_0$ )

$p < 0,01$  (evidência altamente significativa contra  $H_0$ )

$p < 0,0010$  (evidência muito altamente significativa contra  $H_0$ )

A escolha de um nível relativamente modesto de significância para rejeitar os valores com p-valor superior a 25%, é baseada em recomendações de Bendel e Afifi (1977) para regressão linear, de Costanza e Afifi (1979) para análise discriminante e de Mikey e Greenland (1989) para mudanças nos coeficientes do modelo de regressão logística, todos citados por Colosimo e Giolo (2006).

Tabela 3.1 – Relação das covariáveis utilizadas no modelo de previsão de duração de nevoeiro para o Aeroporto Internacional de São Paulo

<b>Código</b>	<b>Descrição</b>	<b>Categorias</b>
V1	Mês de ocorrência dos nevoeiros	0 se pertencente aos meses de maio, junho, julho ou agosto e 1 se pertencente aos outros meses do ano.
V2	Hora de início dos nevoeiros	0 se iniciou entre 20 h até as 2h e 59 min 1 se iniciou das 3 h e adiante
V3	Direção do Vento	0 se menor ou igual a $150^{\circ}$ e 1 se maior que $150^{\circ}$
V4	Velocidade do Vento	0 se menor ou igual a 8 nós e 1 se maior que 8 nós
V5	Visibilidade	0 se menor ou igual a 500 m ou 1 se maior que 500 m
V6	Pressão	0 se maior ou igual a 1018 hPa e 1 se menor que 1018 hPa
V7	Td	0 se menor ou igual a $16^{\circ}$ C e 1 se maior que $16^{\circ}$ C
V8	T <sub>w</sub>	0 se menor ou igual a $17^{\circ}$ C e 1 se maior que $17^{\circ}$ C
V9	UR	0 se igual a 100 % e 1 se menor que 100 %

Para a seleção das covariáveis que influenciam no tempo de dissipação do nevoeiro no Aeroporto Internacional de São Paulo, optou-se por utilizar a

metodologia proposta por Collet (1994), citado e descrita por Colosimo e Giolo (2006), e apresentada resumidamente a seguir:

1. Ajuste de todos os modelos contendo uma única covariável. Incluir todas as covariáveis que forem significativas ao nível de 0,10;
2. As covariáveis significativas no passo 1 são, então, ajustadas conjuntamente. Na presença de certas covariáveis, outras podem deixar de ser significativas. Conseqüentemente, ajustam-se modelos reduzidos, excluindo uma covariável de cada vez. Verificam-se as covariáveis que produzem um aumento estatisticamente significativo na estimativa da razão de verossilhança;
3. Ajusta-se um novo modelo com as covariáveis retidas no passo 2. Neste passo, as covariáveis excluídas no passo 2 retornam ao modelo para confirmar que elas não são estatisticamente significativas;
4. As eventuais covariáveis significativas no passo 3 são incluídas no modelo juntamente com aquelas do passo 2. Neste passo, retorna-se com as covariáveis excluídas no passo 1 para confirmar que elas não são estatisticamente significativas;
5. Ajusta-se um modelo incluindo-se as covariáveis significativas no passo 4. Neste passo é testado se alguma delas pode ser retirada do modelo;
6. Utilizando as covariáveis que sobreviveram ao passo 5, ajusta-se o modelo final para os efeitos principais. Para completar o procedimento, deve-se verificar a possibilidade de inclusão de termos de interação dupla entre as covariáveis incluídas no modelo. O modelo final fica determinado pelos efeitos principais identificados no passo 5 e os termos de interação significativos identificados neste passo.

Antigo (2005) ressalta que caso haja mais de uma possibilidade de significância dos termos de interação, um novo passo deve ser realizado. Assim, sugere o passo 7, tal que:

PASSO 7 - Este passo consistirá em comparar o modelo composto pelas variáveis significativas no passo 5 e todos os termos de interação que apresentaram significância, com os modelos que excluem, um a um, os termos de interação. O modelo final fica, então, definido pelas variáveis significativas no passo 5 e pelos termos de interação significativos identificados neste passo.

Cada passo da aplicação da estatística de teste será obtido utilizando-se o Teste de Razão de verossimilhanças com uma distribuição  $\chi^2$  de referência, com graus de liberdade igual ao número de termos excluídos (Colosimo e Giolo, 2006). Nesta etapa, os valores do logaritmo das funções de verossimilhança para a Distribuição Gama Generalizada foram obtidos com o auxílio do pacote estatístico SAS, uma vez que o pacote estatístico R ainda não disponibiliza esta função. Após a seleção conjunta das covariáveis significativas, também fez-se o Teste de Razão de Verossimilhança comparando-se o Modelo Gama generalizado com os Modelos de exponencial, Weibull e log-normal. O Modelo selecionado será ajustado no R com as covariáveis significativas que foram selecionadas com SAS.

Para determinar qual distribuição de probabilidade melhor se ajusta aos dados, utilizou-se o Modelo Gama Generalizado, que é de grande utilidade na discriminação entre modelos probabilísticos por incluir como casos especiais as distribuições Exponencial, de Weibull, Gama e Log-normal (Colosimo e Giolo, 2006). A adequação do modelo final será verificada pelo Teste de Razão de Verossimilhança, por meio dos resíduos Cox-Snell. Caso o modelo esteja bem ajustado a esse conjunto de dados, a distribuição dos resíduos na escala logarítmica deve estar próxima da normal padrão.

Como última etapa, será feita a verificação da equação de regressão apropriada para a previsão de duração de nevoeiro no aeroporto Internacional de São Paulo. Para tal, será utilizado os eventos de nevoeiro identificados no ano de 2004 (não inclusos nos cálculos) que serão dicotomizados e, em

seguida, aplicados na expressão para a Função de Sobrevivência do Modelo Final para cada intervalo de tempo considerado.



## **4 RESULTADOS E DISCUSSÕES**

Neste capítulo, apresenta-se os resultados referente às climatologias A e B, construídas para o Aeroporto Internacional de São Paulo, onde se considera as distintas definições de nevoeiro. Em seguida, apresenta-se e discutem-se as relações entre as ocorrência e durações dos nevoeiros e as várias variáveis meteorológicas consideradas nesse estudo, uma vez que o conjunto de dados é o mesmo. A seguir, por meio de Análise de Sobrevivência, apresenta-se duas técnicas para modelar a duração dos nevoeiros, uma não-paramétrica utilizando o estimador de Kaplan-Meier e a outra, paramétrica, utilizando o Modelo Gama Generalizada para definir qual o modelo que melhor se ajusta aos dados de nevoeiro do Aeroporto Internacional de São Paulo.

### **4.1 – CLIMATOLOGIAS DE NEVOEIRO NO AEROPORTO INTERNACIONAL DE SÃO PAULO**

#### **a) Climatologia A**

A Figura 4.1 apresenta a climatologia (1988 a 2003) de frequência relativa e frequência absoluta de totais mensais de horas com nevoeiro para o Aeroporto Internacional de São Paulo, onde nesse período foram identificados 951 casos de nevoeiro. Observa-se que os meses de maior ocorrência de horas de nevoeiro, em ordem decrescente, são os meses de junho (551,3 h), maio (408,2 h), julho (356,7 h) e agosto (352,4 h). O mês de maio apresentou a maior ocorrência de horas com nevoeiro, quando comparado aos meses de julho e agosto, estando em concordância com os resultados obtidos por Machado (1993) e Cabral (1998). Também se observou que os meses de menor ocorrência de horas de nevoeiro, em ordem crescente, são os meses de janeiro (66,4 h), dezembro (91,2 h) e fevereiro (92,1 h).

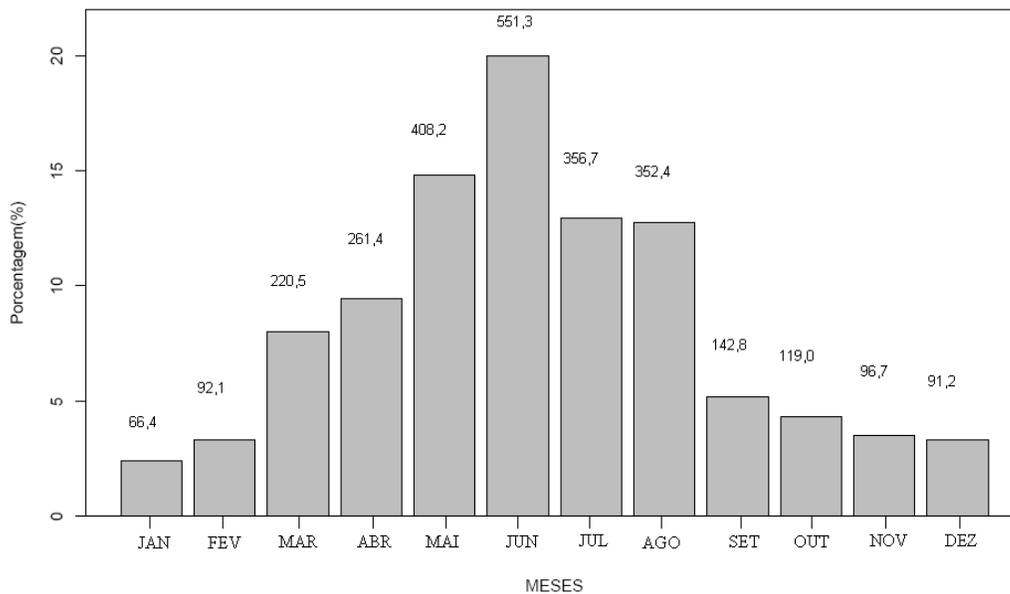


Figura 4.1 – Climatologia (1988-2003) da freqüência relativa (expressa em %) e da freqüência absoluta (indicadas pelos números acima de cada coluna) de totais mensais de horas com nevoeiro, no Aeroporto Internacioanl de São Paulo.

A Figura 4.2 apresenta a climatologia (1988-2003) do total anual de horas com nevoeiro no Aeroporto Internacional de São Paulo. A análise dessa figura mostra que a partir do ano de 1995 houve uma queda no total anual de horas com nevoeiro nessa região (161,42 h), e um aumento nos anos de 1999 (180,75 h) e 2003 (162,30 h). Para explicar o que causou esse aumento de horas de nevoeiro nesses anos, faz-se necessário realizar um estudo sinótico dos eventos de nevoeiro ocorrido nos meses de junho e agosto de 1999 e junho e maio de 2003. Como visto na Figura 4.3, nesses anos e nesses meses houve um aumento relativo de nevoeiros comparado com os anos a partir de 1996. O ano com maior quantidade de horas com nevoeiro foi 1989 (279,85 h) e o ano com menor quantidade de horas com nevoeiro foram 1998 (100,12 h), 2000 (135,60 h) e 2002 (106,72 h). Esses resultados concordam com aqueles obtidos por Cabral (1998).

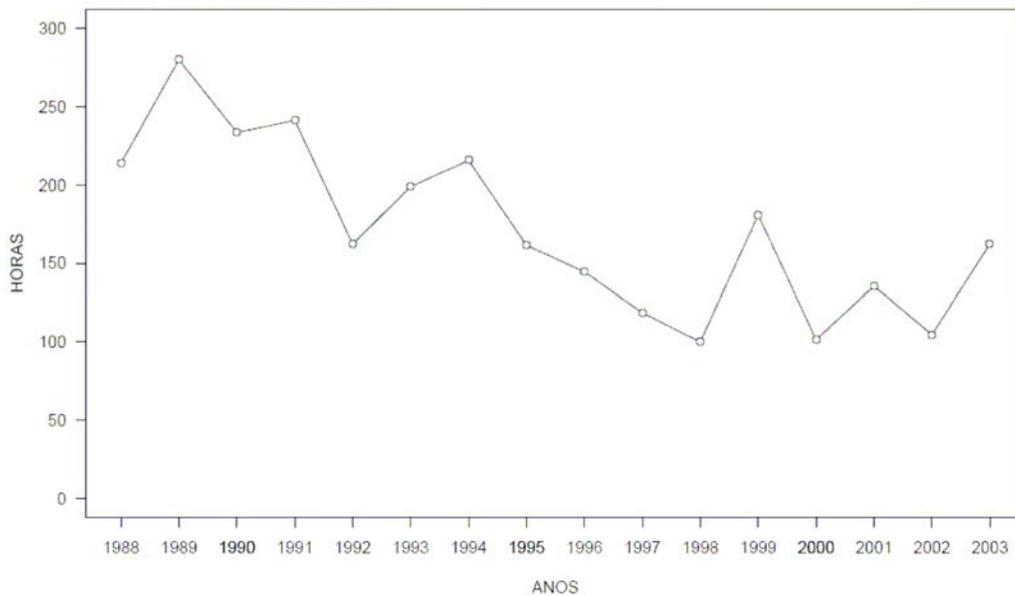


Figura 4.2 – Climatologia (1988-2003) de horas totais anuais de nevoeiros para o Aeroporto Internacional de São Paulo.

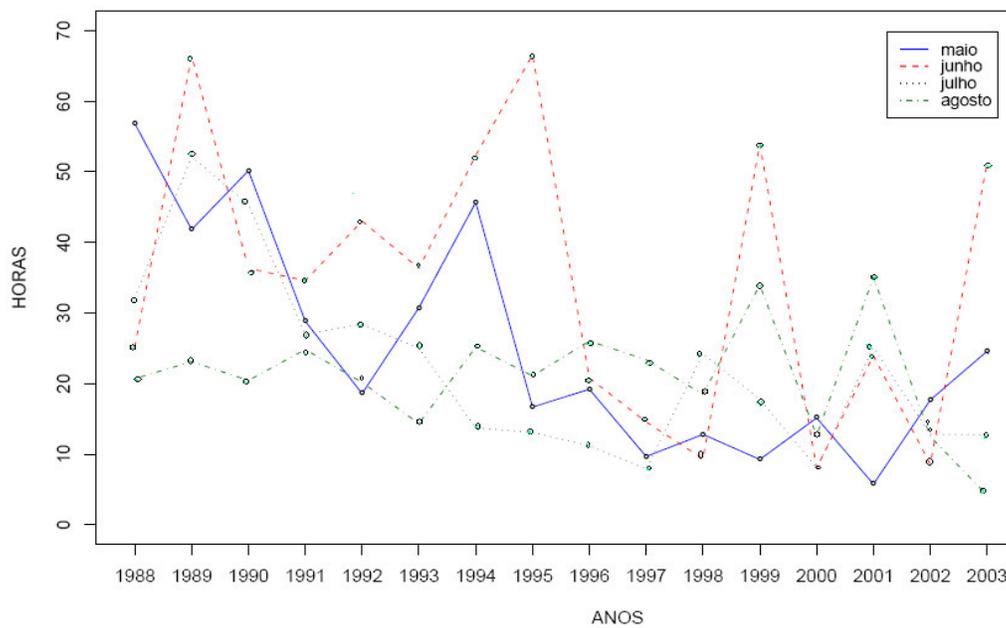


Figura 4.3 – Climatologia (1988-2003) de horas totais anuais de nevoeiros considerando apenas os meses de maio, junho, julho e agosto para o Aeroporto Internacional de São Paulo.

A Figura 4.4 mostra uma climatologia relacionando a hora do início dos nevoeiros no Aeroporto Internacional de São Paulo com o número de ocorrências observadas. Observa-se que os nevoeiros neste aeroporto se formam preferencialmente entre as 20 HL e as 8 HL. No período de 18 anos analisados, verificou-se apenas quatro casos de nevoeiros que se formaram fora desse intervalo, sendo dois casos as 9 HL, um caso as 16 HL e um caso as 19 HL. O pico de maior ocorrência de horas com nevoeiro ocorreu entre as 5 HL e 6 HL, com 179 e 175 casos, respectivamente. Esses resultados concordam parcialmente com Cabral (1998), que encontrou em ordem crescente maior pico de ocorrência de nevoeiro as 6 HL, 7HL e 5 HL. Uma explicação para essa diferença, é que os dados analisados por Cabral (1998) considera um período menor (1993-1997), e outra informação relevante é que, em média, o nascer do sol no Aeroporto Internacional de São Paulo ocorre entre 6 HL a 6,5 HL (CPTEC, 2008), inibindo assim a formação do nevoeiro.

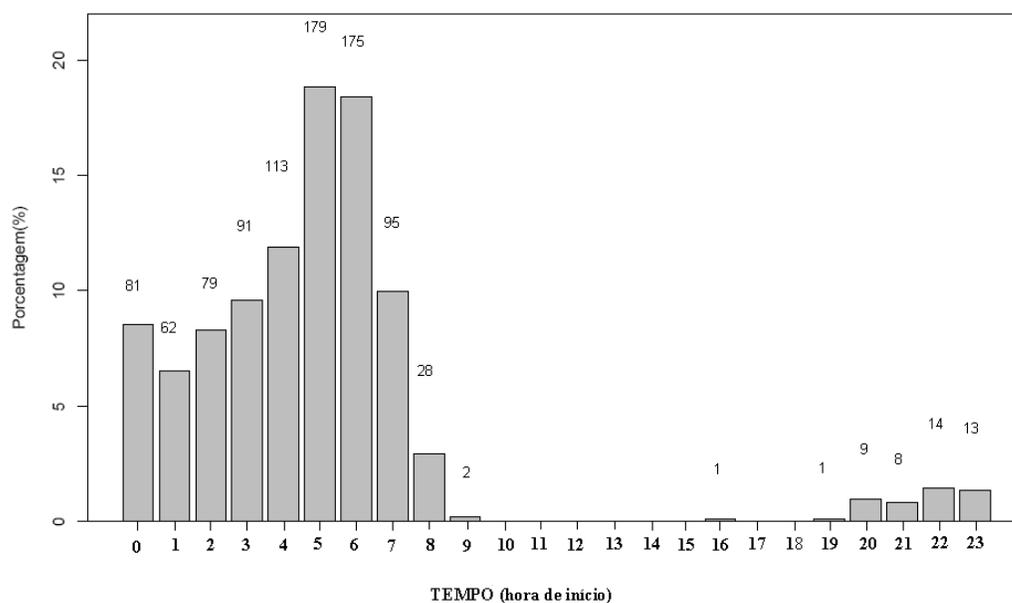


Figura 4.4 – Climatologia (1988-2003) da frequência relativa (expressa em %) e da frequência absoluta (identificada pelos números acima das colunas) da ocorrência dos nevoeiros com a hora de início do nevoeiro, (em intervalos de 1 hora), considerando o período de 1988 a 2003.

## b) Climatologia B

Na climatologia B foram identificados 697 casos de nevoeiros para o Aeroporto Internacional de São Paulo. Esse resultado é inferior ao número de casos identificados na climatologia A por dois motivos: 1) Não foram analisados os nevoeiros identificados pelo Aeroporto em que se tinha UR menor que 95 %, ficando excluído da análise os seis primeiros meses de 1988, os anos inteiros de 1998, 1999 e parte do ano de 2000; 2) Os nevoeiros que se dissiparam e em seguida formaram-se novamente num intervalo menor que 2 horas e não excedeu a visibilidade de 3 km, foi considerado um único nevoeiro. A Figura 4.5 apresenta as freqüências relativa e absoluta dos totais mensais de horas com nevoeiro para o Aeroporto Internacional de São Paulo, para a climatologia B, observa-se em confirmação com a Figura 4.1, que os meses de maior ocorrência de horas de nevoeiro, em ordem decrescente, são os meses de junho (480,3 h), maio (319,9 h), julho (284,9 h) e agosto (274,3 h).

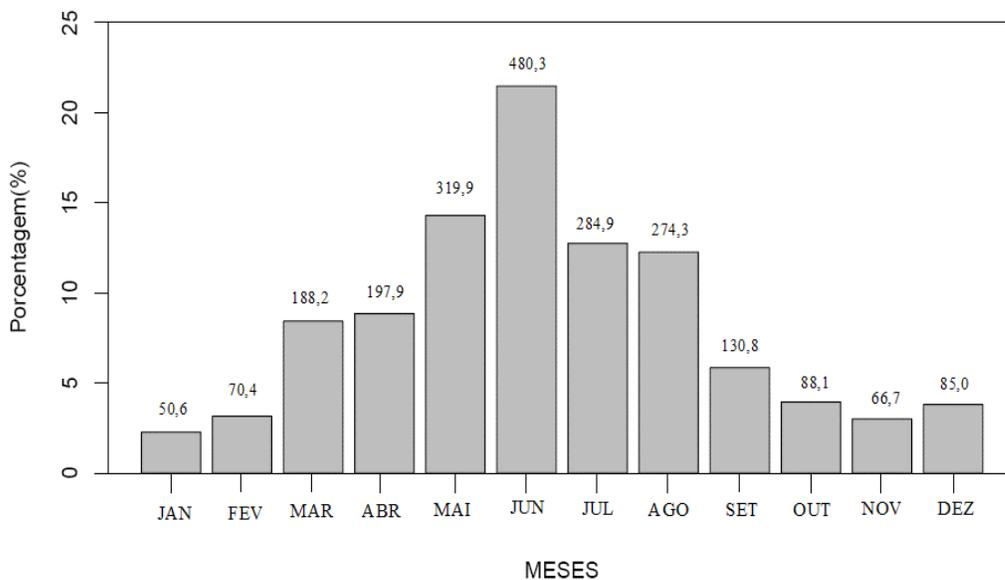


Figura 4.5 – Climatologia B da frequência relativa (expressa em %) e da frequência absoluta (indicadas pelos números acima de cada coluna) de totais mensais de horas com nevoeiro, no Aeroporto Internacional de São Paulo.

A Figura 4.6 apresenta a climatologia do total anual de horas com nevoeiro no Aeroporto Internacional de São Paulo para a climatologia B. A análise dessa figura mostra em confirmação com a Figura 4.2 que a partir do ano de 1995 houve uma queda no total anual de horas com nevoeiro nessa região e um aumento em 2003.

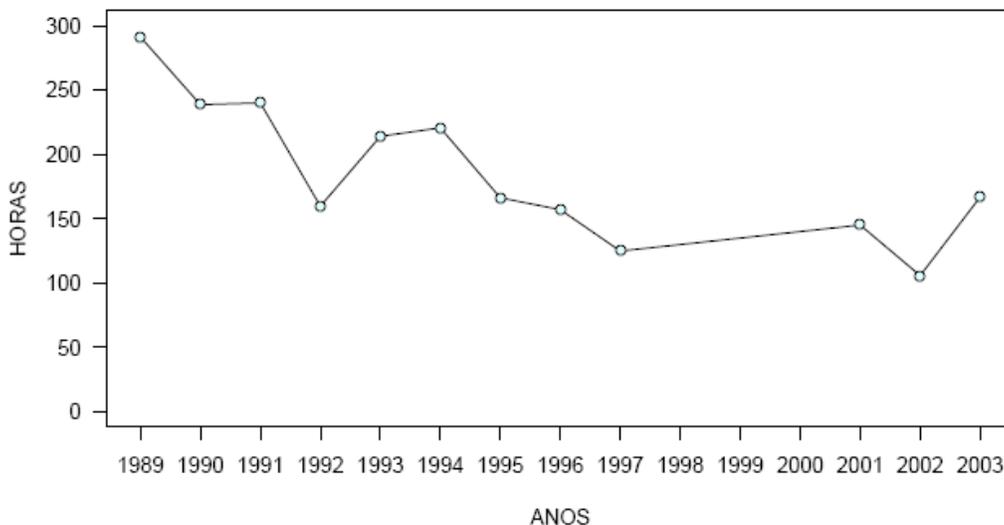


Figura 4.6 – Climatologia (1988-2003) de horas totais anuais de nevoeiros para o Aeroporto Internacional de São Paulo.

A Figura 4.7 mostra a climatologia (1988-2003) da relação da duração do nevoeiro com sua hora de início, onde observa-se que os nevoeiros que começam mais cedo tem maior probabilidade de durarem mais (medianas decrescentes), concordando com os resultados obtidos por Lima (1982) e Reis (1995). A análise da figura mostra que os nevoeiros que começaram às 20 HL tiveram duração de até 14 horas (houve apenas um caso, identificado em 18/06/1990). Por outro lado, os nevoeiros que começaram às 8 HL tiveram duração máxima de 2 h, este caso, foi representado por um ponto extremo, ou seja, é um caso destoante dos demais.

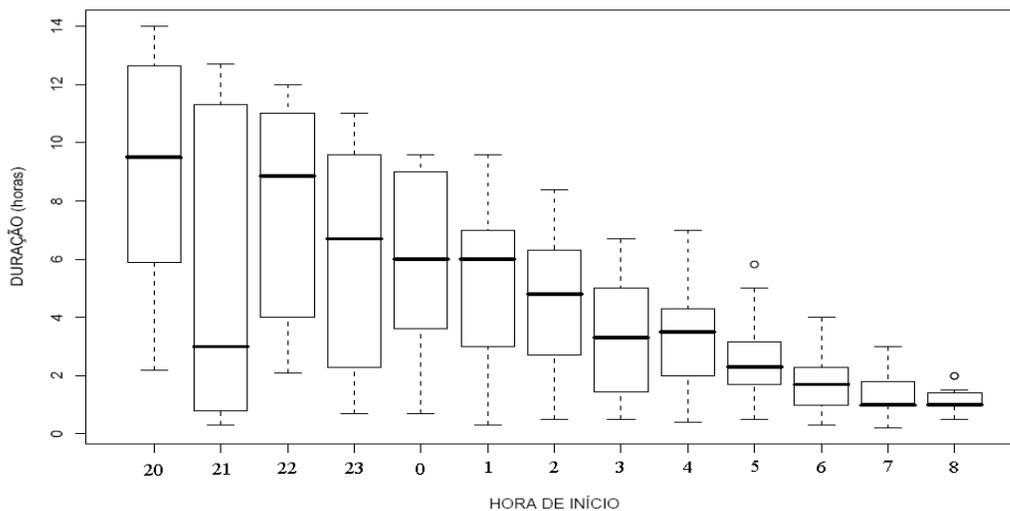


Figura 4.7 – *Boxplot* relacionando as horas de início dos neveiros (intervalos de uma hora) com suas durações, considerando o período de 1988 a 2003.

A análise da Figura 4.8 mostra a relação da duração dos neveiros com os meses de ocorrência, onde se observa que a mediana das durações dos neveiros para os meses de janeiro à dezembro está em torno de 2 h à quase 4 h de duração. O mês de junho foi o que apresentou neveiros com maiores tempos de durações, podendo persistir até 14 h, enquanto que os neveiros que ocorrem nos meses de dezembro, janeiro e fevereiro persistiram por até 7 h.

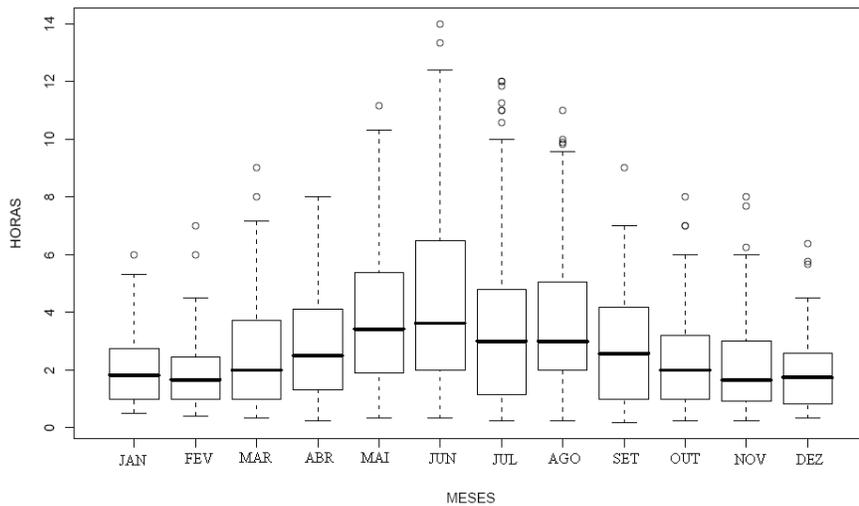


Figura 4.8 – *Boxplot* para a relação dos meses de ocorrência dos nevoeiros com suas durações (expressas em intervalos de uma hora), considerando o período de 1988 a 2003.

## 4.2 – CARACTERÍSTICAS TÍPICAS DOS PARÂMETROS METEOROLÓGICOS DO AEROPORTO INTERNACIONAL DE SÃO PAULO

As Figuras 4.9 até 4.16 apresentam, respectivamente, os Box-plots para os nevoeiros do Aeroporto Internacional de São Paulo, para as covariáveis: hora de início do nevoeiro, mês de ocorrência, direção do vento, visibilidade, pressão,  $T_d$ ,  $T_w$ , e UR. A Figura 4.9 mostra a relação dos intervalos de duração dos nevoeiros com a direção do vento. A análise dessa figura mostra que a maior parte dos casos de nevoeiros ocorrem quando a direção do vento está próxima de  $90^\circ$  (setor leste) e nenhum nevoeiro com duração maior que 8 h teve a direção do vento inicial maior que  $150^\circ$ . A Figura 4.10 mostra a relação dos intervalos de duração dos nevoeiros com a velocidade do vento em nós, observa-se que os nevoeiros ocorrem com velocidades medianas de 5 nós e que nenhum nevoeiro com duração maior que 7 h teve velocidade inicial maior que 10 nós. A Figura 4.11 mostra a relação da visibilidade inicial e as durações dos nevoeiros, observa-se que os dados com exceção do intervalo de 5 h se comportaram de maneira assimétrica (mediana deslocada do meio da caixa) e que para nevoeiros com duração maior que 10 h, a mediana foi inferior a 500

m. A Figura 4.12 mostra a relação da pressão inicial reduzida ao nível médio do mar com os intervalos de durações considerados, observa-se que os nevoeiros com duração maior que 8 h tiveram mediana maior que 1020 hPa. A Figura 4.13 mostra a relação da  $T_d$  e o intervalo de duração dos nevoeiros. A análise dessa figura mostra que os nevoeiros com duração a partir de 8 h tem mediana menor que 15° C, enquanto que para nevoeiros com duração menor que 7 h,  $T_d$  pode passar de 20°C. A Figura 4.14 mostra a relação da  $T_w$  com a duração do nevoeiro. Observa-se que a partir de 8 h de duração a temperatura inicial observada durante os nevoeiros não ultrapassam 20° C e que a mediana para os nevoeiros encontra-se em torno de 12° C a 17°C. A Figura 4.15 mostra a relação da duração do nevoeiro e a UR, observa-se que os dados se comportam de maneira assimétrica na maioria dos intervalos de durações e que neste caso, uma inferência de duração se torna mais difícil, dado que tanto um nevoeiro de curta quanto um de longa duração não aparenta diferenças a uma primeira vista. A Figura 4.16 mostra a relação da visibilidade média e as durações dos nevoeiros, onde observa-se que os nevoeiros que tem visibilidade média menores tem maior probabilidade de durarem mais (medianas decrescentes).

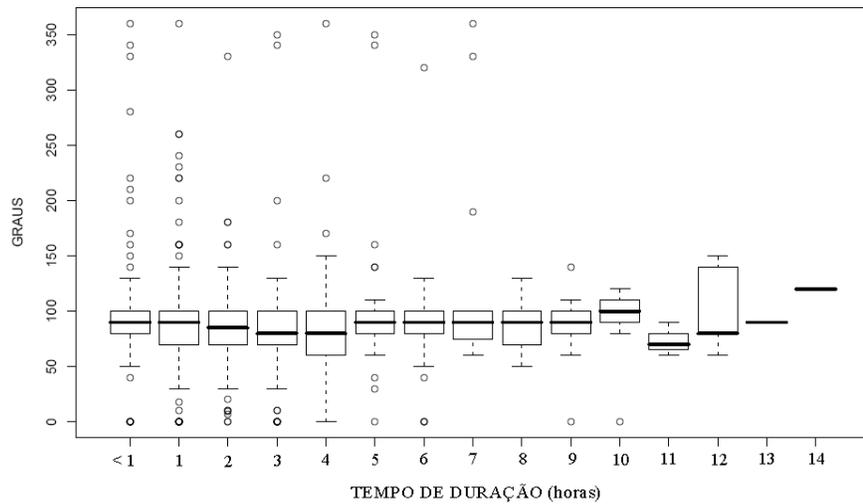


Figura 4.9 – *Boxplot* para a direção do vento em Graus (em °), para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** ( ≥ 2 h e < 3 h), **3** ( ≥ 3 h e < 4 h), **4** ( ≥ 4 h e < 5 h), **5** ( ≥ 5 h e < 6 h), **6** ( ≥ 6 h e < 7 h), **7** ( ≥ 7 h e < 8 h), **8** ( ≥ 8 h e < 9 h), **9** ( ≥ 9 h e < 10 h), **11** ( ≥ 11 h e < 12 h), **12** ( ≥ 12 h e < 13 h), **13** ( ≥ 13 h e < 14h), e **14** ( ≥ 14 h e < 15h).

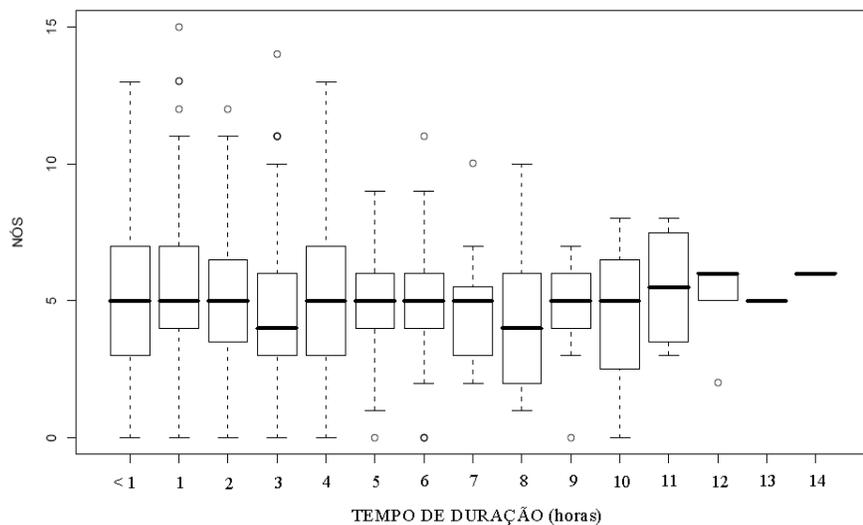


Figura 4.10 – *Boxplot* para a velocidade do vento (expressa em nós) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( ≥ 1 hora de duração e < 2h de duração), **2** ( ≥ 2 h e < 3 h), **3** ( ≥ 3 h e < 4 h), **4** ( ≥ 4 h e < 5 h), **5** ( ≥ 5 h e < 6 h), **6** ( ≥ 6 h e < 7 h), **7** ( ≥ 7 h e < 8 h), **8** ( ≥ 8 h e < 9 h), **9** ( ≥ 9 h e < 10 h), **11** ( ≥ 11 h e < 12 h), **12** ( ≥ 12 h e < 13 h), **13** ( ≥ 13 h e < 14h), e **14** ( ≥ 14 h e < 15h).

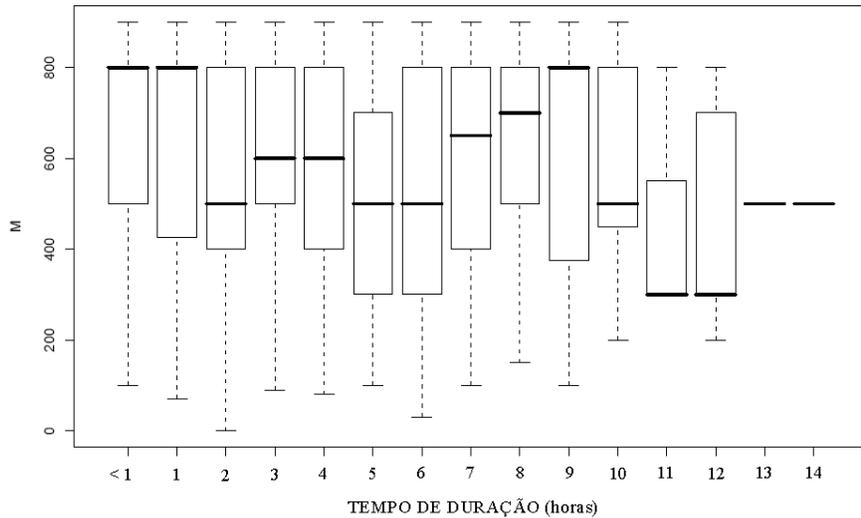


Figura 4.11 – *Boxplot* para a visibilidade (expressa em m) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( $\geq 1$  hora de duração e < 2h de duração), **2** ( $\geq 2$  h e < 3 h), **3** ( $\geq 3$  h e < 4 h), **4** ( $\geq 4$  h e < 5 h), **5** ( $\geq 5$  h e < 6 h), **6** ( $\geq 6$  h e < 7 h), **7** ( $\geq 7$  h e < 8 h), **8** ( $\geq 8$  h e < 9 h), **9** ( $\geq 9$  h e < 10 h), **11** ( $\geq 11$  h e < 12 h), **12** ( $\geq 12$  h e < 13 h), **13** ( $\geq 13$  h e < 14h), e **14** ( $\geq 14$  h e < 15h).

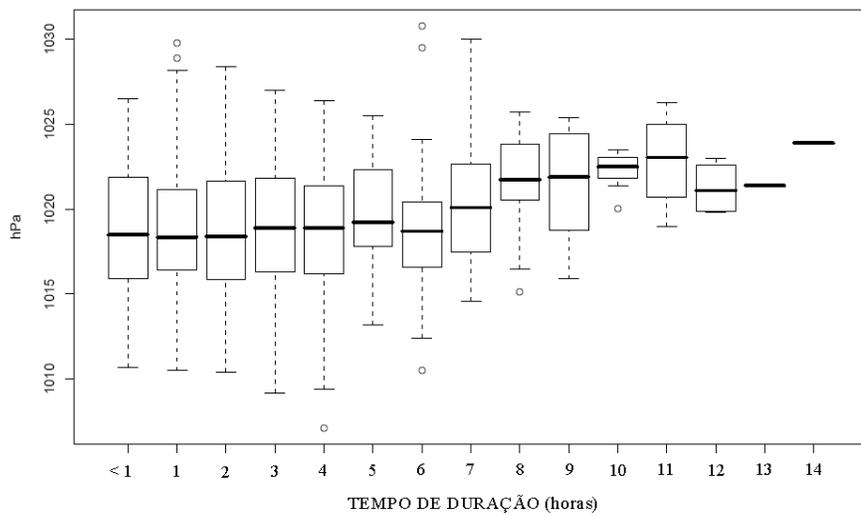


Figura 4.12 – *Boxplot* para a pressão (expressa em hPa) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( $\geq 1$  hora de duração e < 2h de duração), **2** ( $\geq 2$  h e < 3 h), **3** ( $\geq 3$  h e < 4 h), **4** ( $\geq 4$  h e < 5 h), **5** ( $\geq 5$  h e < 6 h), **6** ( $\geq 6$  h e < 7 h), **7** ( $\geq 7$  h e < 8 h), **8** ( $\geq 8$  h e < 9 h), **9** ( $\geq 9$  h e < 10 h), **11** ( $\geq 11$  h e < 12 h), **12** ( $\geq 12$  h e < 13 h), **13** ( $\geq 13$  h e < 14h), e **14** ( $\geq 14$  h e < 15h).

12 h e < 13 h), **13** ( $\geq 13$  h e < 14h), e **14** ( $\geq 14$  h e < 15h).

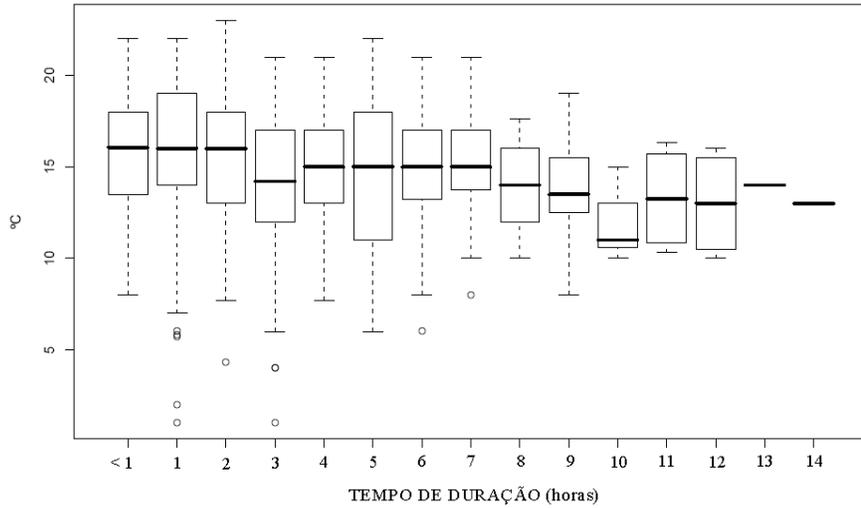


Figura 4.13 – *Bosxplot* para a  $T_d$  para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( $\geq 1$  hora de duração e < 2h de duração), **2** ( $\geq 2$  h e < 3 h), **3** ( $\geq 3$  h e < 4 h), **4** ( $\geq 4$  h e < 5 h), **5** ( $\geq 5$  h e < 6 h), **6** ( $\geq 6$  h e < 7 h), **7** ( $\geq 7$  h e < 8 h), **8** ( $\geq 8$  h e < 9 h), **9** ( $\geq 9$  h e < 10 h), **11** ( $\geq 11$  h e < 12 h), **12** ( $\geq 12$  h e < 13 h), **13** ( $\geq 13$  h e < 14h), e **14** ( $\geq 14$  h e < 15h).

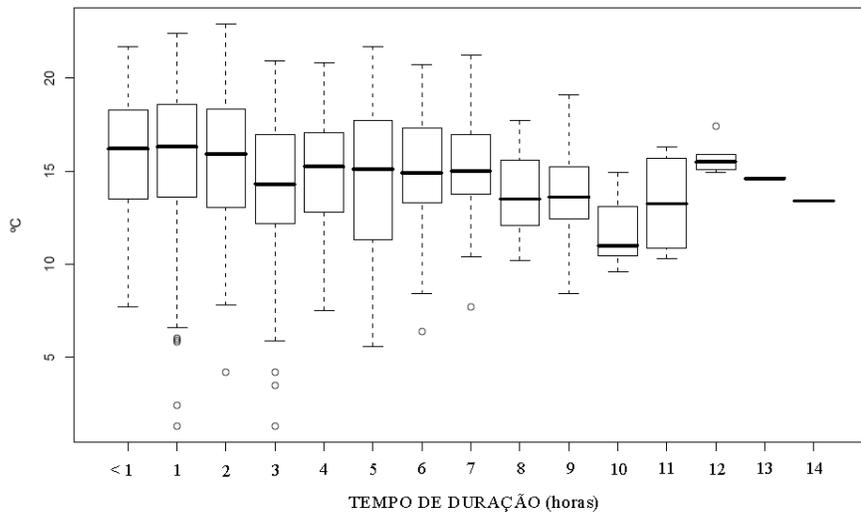


Figura 4.14 – *Bosxplot* para a  $T_w$  para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abscissa (<1 ate 14) representam, respectivamente : <1 hora de duração, **1** ( $\geq 1$  hora de duração e < 2h de duração), **2** ( $\geq 2$  h e < 3 h), **3** ( $\geq 3$  h e < 4 h), **4** ( $\geq 4$  h e < 5 h), **5** ( $\geq 5$  h e < 6 h), **6** ( $\geq 6$  h e < 7 h), **7** ( $\geq 7$  h e < 8 h), **8** ( $\geq 8$  h e < 9 h), **9** ( $\geq 9$  h e < 10 h), **11** ( $\geq 11$  h e < 12 h), **12** ( $\geq 12$  h e < 13 h), **13** ( $\geq 13$  h e < 14h), e **14** ( $\geq 14$  h e < 15h).

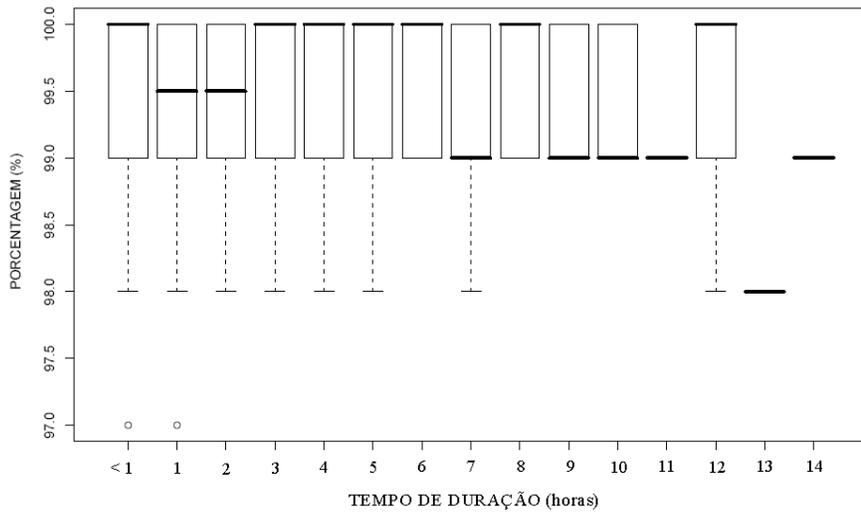


Figura 4.15 – *Bosxplot* para a UR (em %) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abcissa (<1 ate 14) representam, respectivamente : <1 hora de duração, 1 ( $\geq 1$  hora de duração e < 2h de duração), 2 ( $\geq 2$  h e < 3 h), 3 ( $\geq 3$  h e < 4 h), 4 ( $\geq 4$  h e < 5 h), 5 ( $\geq 5$  h e < 6 h), 6 ( $\geq 6$  h e < 7 h), 7 ( $\geq 7$  h e < 8 h), 8 ( $\geq 8$  h e < 9 h), 9 ( $\geq 9$  h e < 10 h), 11 ( $\geq 11$  h e < 12 h), 12 ( $\geq 12$  h e < 13 h), 13 ( $\geq 13$  h e < 14h), e 14 ( $\geq 14$  h e < 15h).

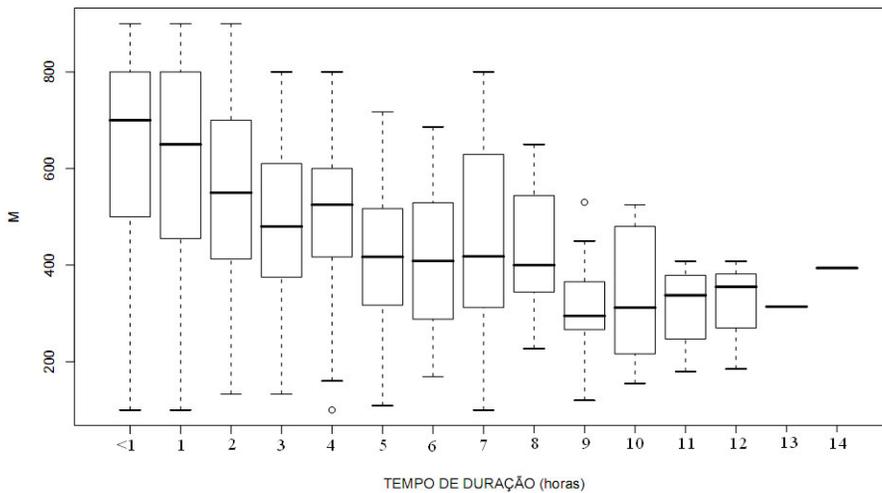


Figura 4.16 – *Bosxplot* para a visibilidade média (expressa em m) para os intervalos de durações (expresso em intervalos de uma hora), considerando o período de 1988 a 2003. Os tempos de duração marcados no eixo da abcissa (<1 ate 14) representam, respectivamente : <1 hora de duração, 1 ( $\geq 1$  hora de duração e < 2h de duração), 2 ( $\geq 2$  h e < 3 h), 3 ( $\geq 3$  h e < 4 h), 4 ( $\geq 4$  h e < 5 h), 5 ( $\geq 5$  h e < 6 h), 6 ( $\geq 6$  h e < 7 h), 7 ( $\geq 7$  h e < 8 h), 8 ( $\geq 8$  h e < 9 h), 9 ( $\geq 9$  h e < 10 h), 11 ( $\geq 11$  h e < 12 h), 12 ( $\geq 12$  h e < 13 h), 13 ( $\geq 13$  h e < 14h), e 14 ( $\geq 14$  h e < 15h).



### 4.3 - RESULTADOS DA ANÁLISE DE SOBREVIVÊNCIA

As Figuras 4.17 até 4.25 apresentam, respectivamente, o Estimador de Kaplan-Meier da Função de Sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para as covariáveis: mês de ocorrência (V1), hora de início do nevoeiro (V2), direção do vento (V3), velocidade do vento (V4), visibilidade (V5), pressão (V6),  $T_d$  (V7),  $T_w$  (V8) e UR (V9). Nessas figuras tem-se no eixo x a probabilidade de sobrevivência do nevoeiro, e no eixo y, o tempo de dissipação do nevoeiro. A análise da Figura 4.17 mostra que as curvas de sobrevivência estimadas não se cruzam (significa que há diferenças entre as curvas de sobrevivência estimadas), e que a curva estimada para os meses de maio, junho, julho e agosto estão acima da curva estimada para os meses de janeiro, fevereiro, março, abril, setembro, outubro, novembro e dezembro. Com isso, o tempo de sobrevivência dos meses de maio, junho, julho e agosto se mostram maiores. Por exemplo, para os nevoeiros que ocorrem nos meses de maio, junho, julho e agosto, o tempo de sobrevivência para eles durarem até quatro horas é de aproximadamente 40%, enquanto que para os nevoeiros dos meses de janeiro, fevereiro, março, abril, setembro, outubro, novembro e dezembro é de 20%. A Figura 4.18 mostra as curvas estimadas de sobrevivência para dois grupos horários considerados, das 20 h as 2 h (curva superior) e das 3 h em diante (curva inferior), onde observa-se que os nevoeiros que iniciam no período das 20 h às 2 h tem maior tempo de duração. Também, para os nevoeiros que se iniciaram depois das 3 h o tempo de duração máximo foi de aproximadamente 8 h, confirmando resultados da Figura 4.5, enquanto que para os nevoeiros que se iniciaram entre as 20 h as 2 h tiveram tempo de duração igual a 14 horas. A Figura 4.19 mostra as curvas de sobrevivência estimadas para a direção do vento, onde observa-se para a curva inferior ( $> 150^\circ$ ) os nevoeiros tem tempo de duração menor. A Figura 4.20 mostra as curvas de sobrevivência estimadas para a velocidade do vento, onde observa-se para a curva inferior ( $> 8$  nós) os nevoeiros tem tempo de duração menor. A Figura 4.21 mostra que para visibilidade menor ou igual a 500 m a curva de sobrevivência se mantém acima, indicando para esse grupo

maior tempo de sobrevivência. A Figura 4.22 mostra as curvas das sobrevivências estimadas para pressão maior ou igual a 1018 hPa e para pressão menor que 1018 hPa, onde observa-se a curva de sobrevivência estimada para os nevoeiros com pressões maiores que 1018 hPa, tem-se tempo de sobrevivência maior. As Figuras 4.23 e 4.24 mostram as curvas de sobrevivência estimada para a variável  $T_d$  e  $T_w$ , respectivamente. Observa-se que para  $T_d$  menor ou igual a 16° C o tempo de sobrevivência é maior que o tempo de sobrevivência para  $T_d$  maior que 16° C, enquanto para  $T_w$  o tempo de sobrevivência para temperaturas maiores que 17° C se mostrou menor do que para temperaturas menores que 17° C. Analisando-se a Figura 4.25, verifica-se que não há diferenças entre as curvas de sobrevivência para UR igual a 100% e UR menor que 100%, pois o tempo de duração para as duas curvas de sobrevivência estimada são iguais até o intervalo de duração de 2 h. Próximo de 5 h as curvas de UR se cruzam mostrando, assim, não haver diferença entre os tempos de sobrevivência para UR igual a 100% e UR menor que 100%.

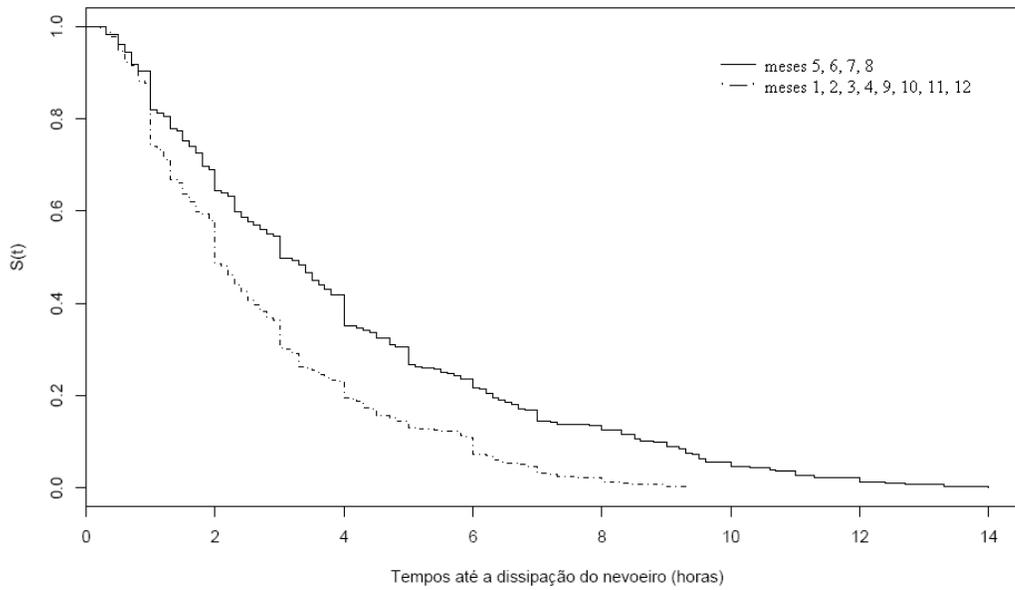


Figura 4.17 - Estimador de Kaplan-Meier da Função de Sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “mês de ocorrência”.

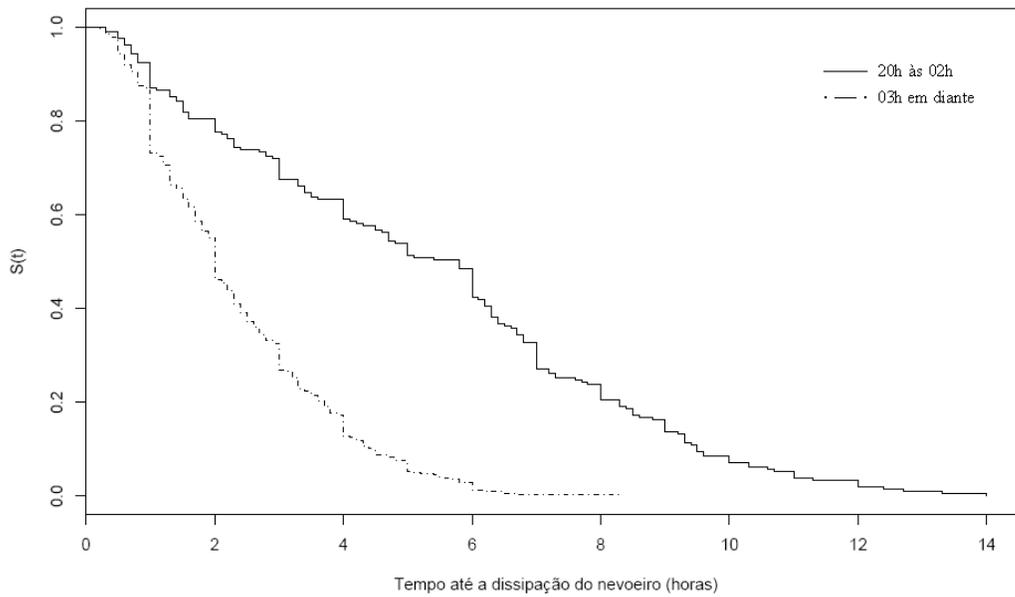


Figura 4.18 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “hora de início do nevoeiro”.

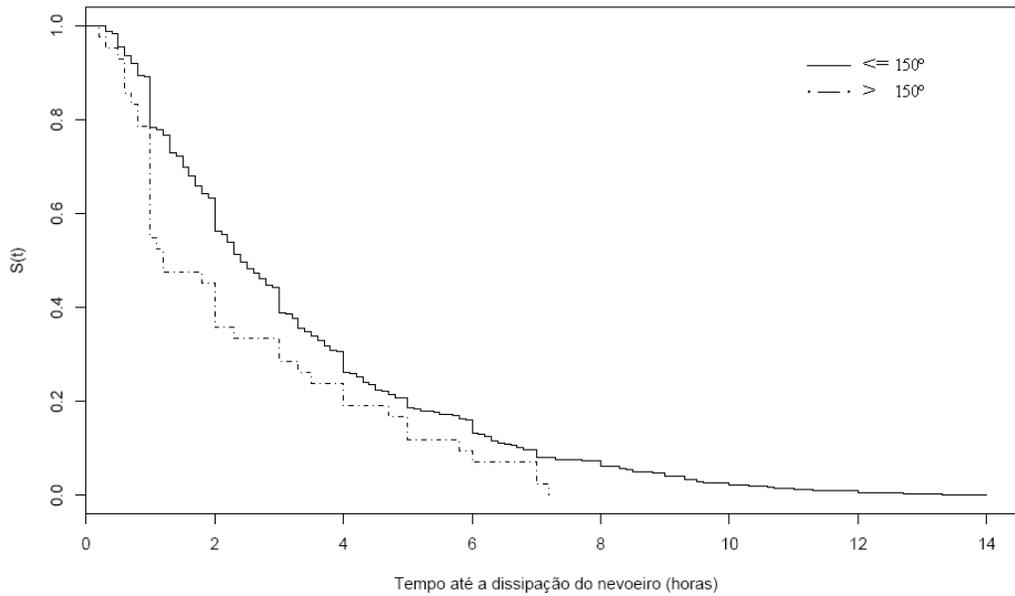


Figura 4.19 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “direção do vento”.

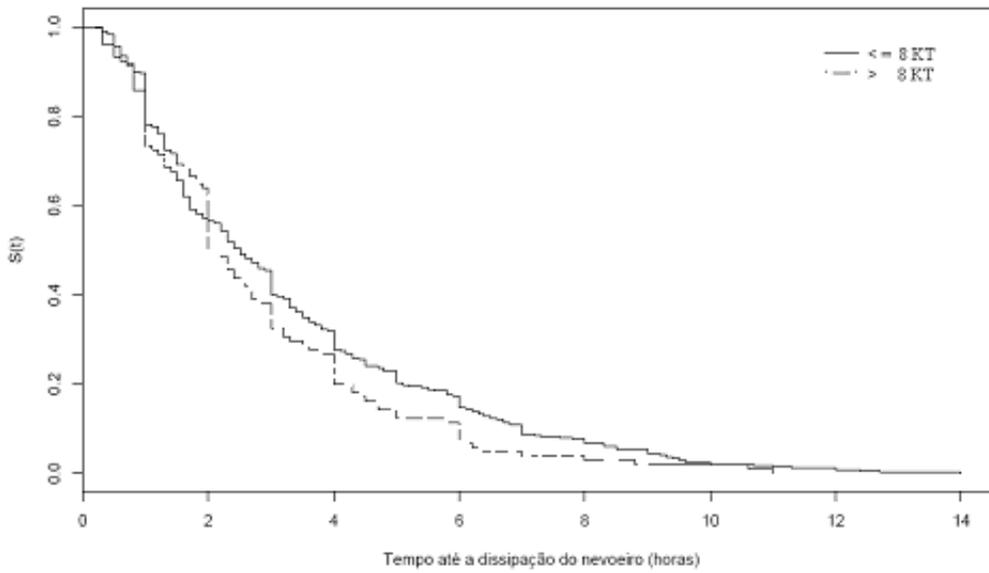


Figura 4.20 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “visibilidade”.

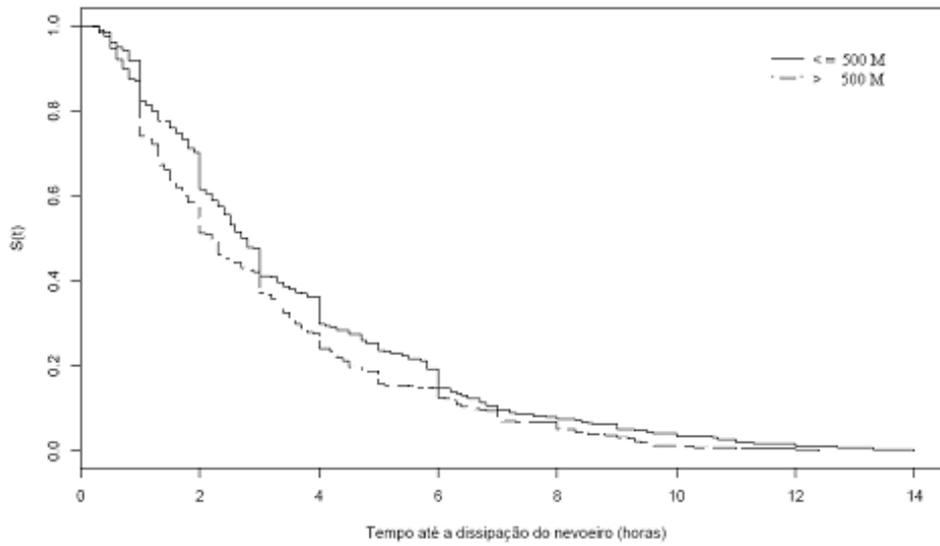


Figura 4.21 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “visibilidade”.

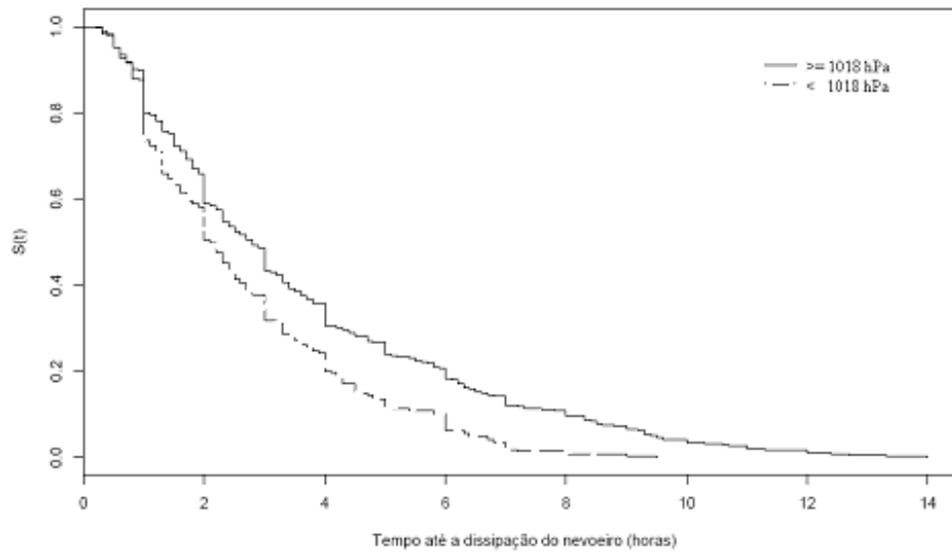


Figura 4.22 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável “pressão”.

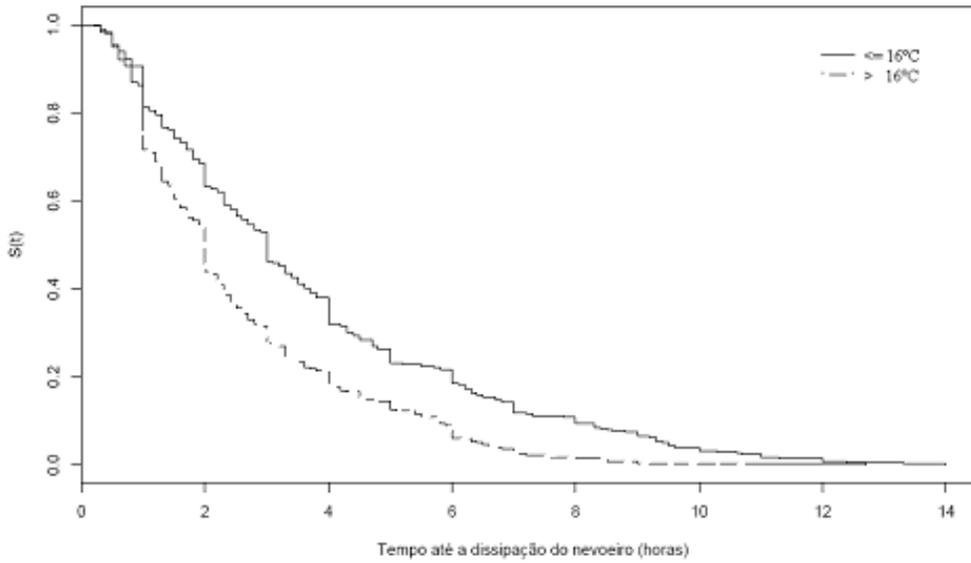


Figura 4.23 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável  $T_d$ .

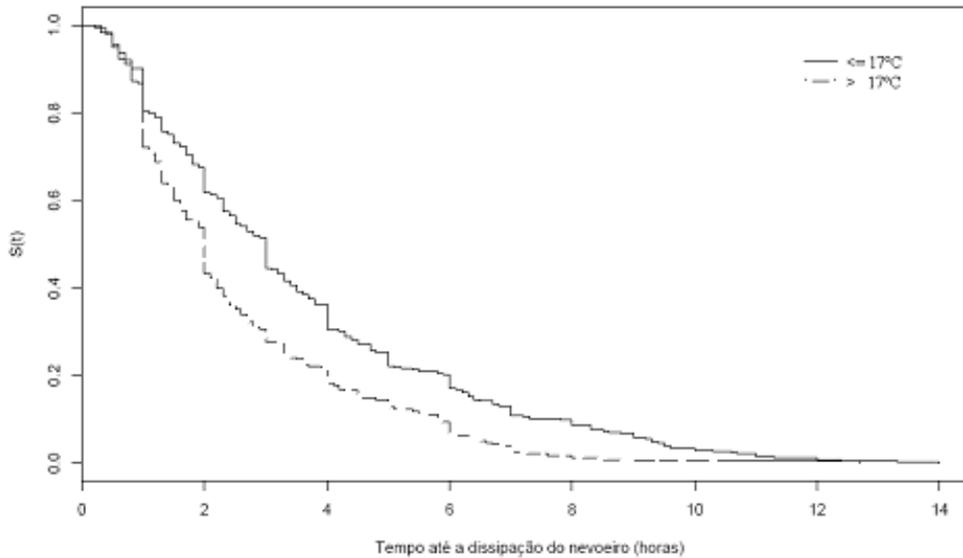


Figura 4.24 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável  $T_w$ .

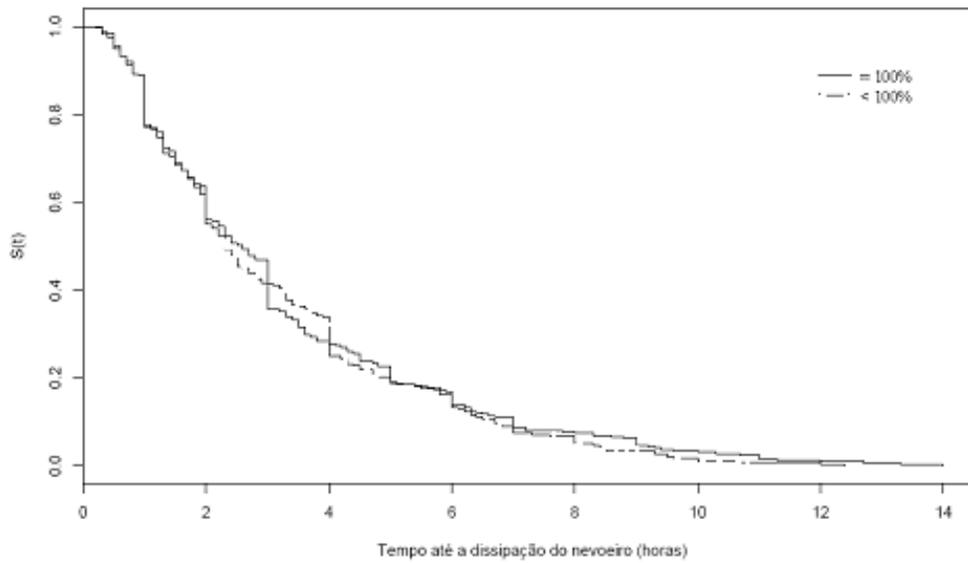


Figura 4.25 - Estimador de Kaplan-Meier da função de sobrevivência dos nevoeiros do Aeroporto Internacional de São Paulo para a covariável UR.

A Tabela 4.1 apresenta os testes de *logrank* e *Wilcoxon* obtidas para cada covariável apresentadas nas Figuras 4.17 a 4.25, observa-se que em concordância com a Figura 4.25 a única covariável que não prosseguirá na análise é a UR, pois esta apresenta p valor maior que 0,25 nos dois testes considerados. Cabe ainda ressaltar, que com um grau de liberdade o valor crítico de  $\chi^2$  correspondente a p valor 0,25 é 0,102.

Tabela 4.1 – Testes de *Logrank* e de *Wilcoxon*.

Covariável	Testes (valor p)	
	$\chi^2$ (logrank)	$\chi^2$ (Wilcoxon)
V1: Mês de ocorrência dos nevoeiros	46,9 ( $7,62 \times 10^{-12}$ )	28,3 ( $1,04 \times 10^{-07}$ )
V2: Hora de início dos nevoeiros	199 (0)	114 (0)
V3: Direção do Vento	5,4 (0,0196)	7,7 (0,00553)
V4: Velocidade do Vento	4,6 (0,0317)	3,4 (0,0667)
V5: Visibilidade	7,9 (0,00489)	8,9 (0,00287)
V6: Pressão	26,3 ( $2,85 \times 10^{-07}$ )	13,6 (0,000223)
V7: Td	38,6 ( $5,24 \times 10^{-10}$ )	31,1 ( $2,44 \times 10^{-08}$ )
V8: Tw	29,9 ( $4,53 \times 10^{-08}$ )	24,6 ( $6,93 \times 10^{-07}$ )
<b>V9: UR</b>	<b>0,0008 (0,914)</b>	<b>0,3 (0,572)</b>

A Tabela 4.2 mostra os passos de seleção das covariáveis feito com o Modelo Gama Generalizado. Embora as covariáveis pressão e  $T_w$  tenham se mostrado significativas quando testadas individualmente, nesta análise paramétrica onde se testa os efeitos conjuntos das covariáveis, elas não se mostraram significativas, pois apresentaram valor p maior que 0,10 (passo 2). Para confirmar a exclusão dessas covariáveis, elas foram testadas novamente no passo 3, onde mostraram também valor de p maior que 0,10 (Tabela 4.2). No passo 4, nenhum modelo foi analisado, pois, nesse passo, ajusta-se as covariáveis significativas do passo 3 com as covariáveis excluídas do passo 1. No passo 5, testa-se se pode ser retirada do estudo alguma covariável, onde se verificou que em nenhum caso se teve valor p maior que 0,10, o que mostra que todas as covariáveis testadas devem permanecer no estudo.

Tabela 4.2 – Seleção das covariáveis.

<i>Modelo</i>	<i>-2log L (θ)</i>	<i>(TRV)</i>	<i>Valor p</i>
<b>PASSO 1</b>			
Nulo	1702,821		
V1	1665,839	36,982	1,19224 x 10 <sup>-10</sup>
V2	1507,674	195,147	0
V3	1695,962	6,859	0,00881
V4	1698,942	3,879	0,04889
V5	1693,599	9,222	0,00239
V6	1684,230	18,591	1,61983 x 10 <sup>-06</sup>
V7	1669,653	33,168	8,45302 x 10 <sup>-9</sup>
V8	1677,331	25,49	4,44681 x 10 <sup>-7</sup>
<b>PASSO 2</b>			
V1+V2+V3+V4+V5+V6+V7+V8	1450,663		
V2+V3+V4+V5+V6+V7+V8	1449,764	9,101	0,002
V1+V3+V4+V5+V6+V7+V8	1638,000	187,337	0
V1+V2+V4+V5+V6+V7+V8	1457,597	6,934	0,008
V1+V2+V3+V5+V6+V7+V8	1453,748	3,083	0,079
V1+V2+V3+V4+V6+V7+V8	1456,964	6,301	0,012
<b>V1+V2+V3+V4+V5+V7+V8</b>	<b>1450,721</b>	<b>0,058</b>	<b>0,809</b>
V1+V2+V3+V4+V5+V6+V8	1454,878	4,215	0,040
<b>V1+V2+V3+V4+V5+V6+V7</b>	<b>1450,931</b>	<b>0,268</b>	<b>0,604</b>
<b>PASSO 3</b>			
V1+V2+V3+V4+V5+V7	1451,009		
<b>V1+V2+V3+V4+V5+V7+V6</b>	<b>1450,931</b>	<b>0,078</b>	<b>0,780</b>
<b>V1+V2+V3+V4+V5+V7+V8</b>	<b>1450,721</b>	<b>0,288</b>	<b>0,591</b>
<b>PASSO 4</b>			
Nenhuma covariável foi excluída no PASSO 1			
<b>PASSO 5</b>			
V1+V2+V3+V4+V5+V7	1451,009		
V2+V3+V4+V5+V7	1460,767	9,758	0,001
V1+V3+V4+V5+V7	1638,592	187,583	0
V1+V2+V4+V5+V7	1457,709	6,700	0,009
V1+V2+V3+V5+V7	1454,035	3,026	0,081
V1+V2+V3+V4+V7	1457,387	6,378	0,011
V1+V2+V3+V4+V5	1459,903	8,894	0,002

**PASSO 6**

V1+V2+V3+V4+V5+V7	1451.009		
V1+V2+V3+V4+V5+V7+V1*V2	1449.547	1,555	0,212
V1+V2+V3+V4+V5+V7+V1*V3	1448.353	2,962	0,103
V1+V2+V3+V4+V5+V7+V1*V4	1450.810	0,251	0,656
V1+V2+V3+V4+V5+V7+V1*V5	1450.676	0,386	0,531
<b>V1+V2+V3+V4+V5+V7+V1*V7</b>	<b>1448.010</b>	<b>2,94</b>	<b>0,083</b>
V1+V2+V3+V4+V5+V7+V2*V3	1450.969	0,013	0,950
V1+V2+V3+V4+V5+V7+V2*V4	1451.000	0,059	0,926
V1+V2+V3+V4+V5+V7+V2*V5	1450.773	0,31	0,627
V1+V2+V3+V4+V5+V7+V2*V7	1449.835	1,524	0,278
V1+V2+V3+V4+V5+V7+V3*V4	1449.432	1,446	0,209
V1+V2+V3+V4+V5+V7+V3*V5	1450.303	0,888	0,401
<b>V1+V2+V3+V4+V5+V7+V3*V7</b>	<b>1448.605</b>	<b>3,088</b>	<b>0,087</b>
V1+V2+V3+V4+V5+V7+V4*V5	1450.426	0,201	0,445
V1+V2+V3+V4+V5+V7+V4*V7	1449.724	1,285	0,256
V1+V2+V3+V4+V5+V7+V5*V7	1450,918	0,199	0,763

**PASSO 7**

V1+V2+V3+V4+V5+V7+V1*V7+V3*V7	1444,985		
<b>V1+V2+V3+V4+V5+V7+V1*V7</b>	<b>1448,085</b>	<b>3,100</b>	<b>0,078</b>
<b>V1+V2+V3+V4+V5+V7+V3*V7</b>	<b>1448,012</b>	<b>3,027</b>	<b>0,081</b>

**Etapa Final**

1) V1+V2+V3+V4+V5+V7+V1\*V7+V3\*V7 1444,985

**Modelo Final:**

1)V1+V2+V3+V4+V5+V7+V1\*V7+V3\*V7 1444,985

A Estatística de teste que compara o Modelo Gama Generalizada com os modelos Exponencial, log-normal e Weibull, apresentaram os seguintes resultados:

**Pág. 74 - Testes de Verossimilhança para o Modelo Final (Tabela 4.2)**

<i>Distribuição</i>	<i>-2log L (θ)</i>	<i>(TRV)</i>	<i>Valor p</i>
Gama Generalizada	1444,985		
Exponential	1702.927	257,942	0
Log normal	1534,164	89,179	0
<b>Weibull</b>	<b>1446.025</b>	<b>1,040</b>	<b>0,307</b>

A partir desses resultados, concluiu-se que o Modelo de Regressão de Weibull é o mais adequado (menor valor de TRV) para ajustar os tempos até a dissipação do nevoeiro do Aeroporto Internacional de São Paulo. A Tabela 4.3 apresenta as estimativas dos parâmetros desse modelo:

Tabela 4.3 a – Estimativas dos parâmetros do Modelo de Weibull (Modelo FINAL).

<b>COVARIÁVEL</b>	$\hat{\beta}$ (estimado)	<b>ERRO PADRÃO</b>	<b>VALOR p</b>
Constante	1,964	0,053	$4,05 \times 10^{-300}$
V1	-0,121	0,061	$4,75 \times 10^{-2}$
V2	-0,734	0,049	$9,54 \times 10^{-50}$
V3	-0,422	0,130	$1,26 \times 10^{-3}$
V4	-0,111	0,062	$7,26 \times 10^{-2}$
V5	-0,107	0,045	$1,69 \times 10^{-2}$
V7	0,016	0,126	$8,99 \times 10^{-1}$
V1*V7	-0,245	0,140	$8,10 \times 10^{-2}$
V3*V7	0,305	0,188	$1,05 \times 10^{-1}$
Escala	0,579	0,017	

Utilizou-se o Resíduo de Cox-Snell para confirmar a adequação do Modelo de Weibull, indicando que o modelo se encontra ajustado (distribuição semelhante a uma linha reta com inclinação igual a um) na Figura 4.24 e, equivalentemente, na Figura 4.23 (o Resíduo de Cox-Snell seguiu a distribuição exponencial padrão).

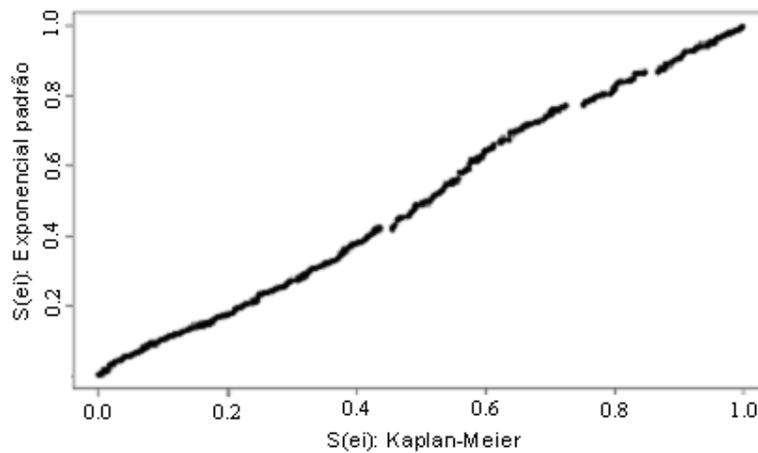


Fig. 4.26 – Sobrevivência dos Resíduos de Cox-Snell estimadas pelo método de Kaplan-Meier e pelo Modelo Exponencial padrão (para o Modelo Final).

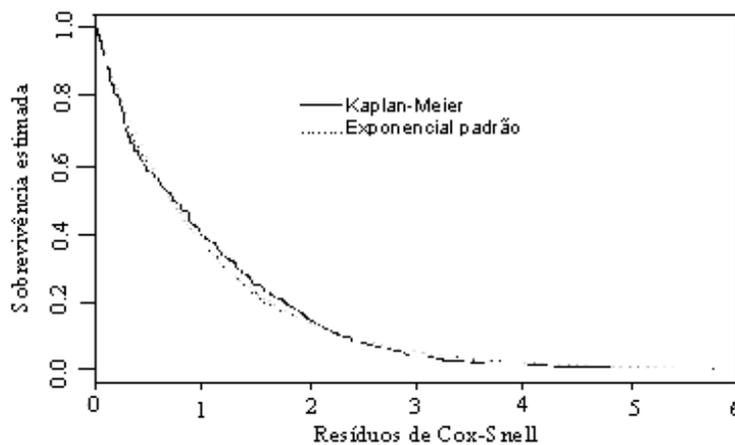


Fig. 4.27 – Curvas de sobrevivência estimadas (para o Modelo Final).

Utilizando as estimativas dos parâmetros do Modelo de Weibull e substituindo na expressão do Modelo de Regressão Weibull com t condicional a x, tem-se:

$$S(t|x) = \exp\left\{-\left[\frac{t}{\exp(1,964 - 0,121V1 - 0,734V2 - 0,421V3 - 0,111V4 - 0,107V5 + 0,016V7 - 0,245V1 \times V7 + 0,305V3 \times V7)}\right]^{1,725}\right\} \dots\dots\dots(4)$$

A expressão 4 dá a probabilidade de um dado nevoeiro durar além de um tempo t, conhecidas as covariáveis: mês de ocorrência, hora de início do nevoeiro, direção do vento, velocidade do vento, visibilidade e a temperatura do ponto de orvalho.

O Apêndice G apresenta os resultados da verificação do modelo de previsão de duração sugerido, considerando o ano de 2004. A Tabela G.1 apresenta a lista de eventos de nevoeiro identificados no ano de 2004 (não inclusos nos cálculos de Análise de Sobrevivência), a Tabela G.2 apresenta esses dados dicotomizados (de acordo com a Tabela 3.1) e a Tabela G.3, a probabilidade dos nevoeiros durarem além de um tempo t definido por cada intervalo de tempo de 0,5 h, 1 h, 2 h, 3 h, 4 h, 5 h, 6 h, 7 h, 8 h, 9 h, 10 h, 11 h, 12 h, 13 h. Ressalta-se, ainda, que os resultados apresentados na Tabela G.3, podem ser interpretados de outras duas maneiras, a saber: i) a probabilidade do nevoeiro ocorrer em um intervalos de tempo t1, t2, e para isso, calcula-se a diferença P1–P2; ii) a probabilidade de durar até t, e para isso, calcula-se 1 – P.

A Tabela G.4, apresenta os resultados da verificação do modelo, no qual apresenta-se o tempo de duração observado e a probabilidade de durar além de um tempo t imediatamente inferior ao observado. Nos resultados apresentados nesta tabela observou-se que dos 47 eventos de nevoeiros observados, 9 casos tiveram resultados com observações subestimadas. Isso pode ser explicado por ter-se construído a equação com todos os meses do

ano. Assim, sugere-se que o modelo seja aplicado apenas considerando os meses de maio, junho, julho e agosto.

## 5 CONCLUSÕES E CONSIDERAÇÕES FINAIS

O objetivo geral deste trabalho foi utilizar o método de Análise de Sobrevivência para elaborar um método de previsão de duração de nevoeiro para o Aeroporto Internacional de São Paulo. Para isso fez-se duas climatologias de nevoeiro (A e B), definidas usando limiares apenas para visibilidade e usando visibilidade e umidade relativa, respectivamente. Considerou-se o período de 1988-2003. Os principais resultados da climatologia A foram:

- i. Ocorreram 951 casos de nevoeiros de radiação e os meses de maior ocorrência foram junho, maio, julho e agosto, em ordem decrescente. Esses resultados estão em concordância com os obtidos por Cabral (1998) e Machado (1993);
- ii. Maior pico de ocorrência de nevoeiros as 5HL e 6 HL;
- iii. Uma queda no número de horas totais anuais de nevoeiro a partir do ano de 1995 e um relativo aumento nos anos de 1999 e 2003;

Os principais resultados da climatologia B foram:

- i. Ocorreram 697 casos de nevoeiro de radiação;
- ii. Em geral os nevoeiros duram de 2 h a quase 4h. Em junho os nevoeiros tiveram maior duração (um evento durou até 14h). Os nevoeiros que ocorrem em dezembro, janeiro e fevereiro tiveram duração máxima de 7h;
- iii. Há maior probabilidade de duração para os nevoeiros que começam mais cedo e que ocorrem nos meses de junho, julho, maio e agosto.

Da análise climatológica das várias variáveis meteorológicas consideradas, verificou-se que:

- i. Há maior probabilidade de duração para os nevoeiros que iniciam com direção do vento menor que  $150^\circ$  e próxima de  $90^\circ$ ;
- ii. Há maior probabilidade de duração para os nevoeiros que se iniciam com velocidade do vento menor que 8 nós e mediana de 5 nós;

- iii. Há maior probabilidade de duração para os nevoeiros com pressão maior ou igual a 1018 hPa e mediana entre 1018 a 1025 hPa;
- iv. Há maior probabilidade de duração para  $T_d$  menor que 16° C e mediana entre 13° C e 16 °C;
- v. Há maior probabilidade de duração para  $T_w$  maior ou igual a 17° e mediana entre 11 °C e 17° C.

As covariáveis finais ajustadas com o Modelo Gama Generalizada foram: mês de ocorrência do nevoeiro, hora de início do Nevoeiro, direção do vento, velocidade do vento, visibilidade e  $T_d$ . Para a escolha do modelo adequado, fez-se o Teste de Estatística (TRV) comparando o Modelo Gama Generalizado com os Modelos Exponencial, Weibull e Log-normal, no qual o Modelo de Weibull apresentou menor valor de (TRV). Sendo assim, este é o modelo que melhor se ajusta aos dados de nevoeiro do Aeroporto Internacional de São Paulo. Em seguida, ajustou-se o Modelo de Weibull com todas as covariáveis selecionadas pelo Modelo Gama Generalizado e com as estimativas de seus parâmetros, chegou-se na seguinte equação para previsão de duração de nevoeiro:

$$S(t|x) = \exp \left\{ - \left[ \frac{t}{\exp(1,964 - 0,121V1 - 0,734V2 - 0,421V3 - 0,111V4 - 0,107V5 + 0,016V7 - 0,245V1 \cdot x \cdot V7 + 0,305V3 \cdot x \cdot V7)} \right]^{1,725} \right\}$$

Embora o modelo tenha sido calibrado, a duração de alguns eventos de nevoeiro foram subestimados (de 47 casos ocorridos em 2004, 9 deles foram subestimados). Isso pode ter ocorrido por se considerar todos os meses do ano

na aplicação do método. Assim, sugere-se para trabalhos futuros refazer essa análise considerando apenas os meses de maio, junho, julho e agosto.

Também, sugere-se para trabalhos futuros:

- Incluir como covariável a camada de mistura turbulenta e altura da inversão dada pela temperatura potencial;
- Realizar a validação do modelo de usando o método de assimilação de dados;
- Aplicar o Modelo de Cox estratificado aos dados de nevoeiro do Aeroporto Internacional de São Paulo, que leva em consideração as variações das variáveis de acordo com o tempo de duração do nevoeiro;
- Realizar um estudo sinótico dos casos de nevoeiros de curta, média e longa duração, para entender os mecanismos de formação e manutenção dos nevoeiros no Aeroporto Internacional de São Paulo, aqui identificados.



## REFERÊNCIAS BIBLIOGRÁFICAS

ANDRÉ, J. C.; MAHRT, L. The nocturnal surface inversion and influence of clear-air radioactive cooling. **J. Atmos. Sci.**, v.39, n. 4, p. 864-878, 1981.

ALLISON, P. **Survival analysis using the SAS system: a practical guide.** Cary, NC: SAS Press, Nov. 1995. 304 p.

ALTAMIRANO, R. J. A. **Análisis climatico y sinóptico de las nieblas en el Aeropuerto Internacional Jorge Chávez Callao, Lima – Peru.** Dissertação de mestrado, Universidad Nacional Agraria La Molina, Dept. De Física y Meteorología, Faculdade de Ciências. La Molina, Ago. 2001. 126 p.

ANTIGO, M. F. **Desemprego na região metropolitana de Belo Horizonte: evidências de transição e permanência (1997 a 2001).** Dissertação de mestrado, Faculdade de Ciências Econômicas, Centro de Desenvolvimento e Planejamento Regional. Belo Horizonte, 2005, 113 p.

BECKWITH, W.B. Aviation. In: Houghton, D. D (ed.). **Handbook of Applied Meteorology.** New York: John Wiley & Sons, 1985. cap. 34, p. 953..

BENDEL, R. B.; AFIFI, A. A. Comparison of stopping rules in forward regression. **Journal of American Statistical Association**, v.. 72, p. 46-53, 1977.

BRASIL. MINISTÉRIO DA DEFESA. DEPARTAMENTO DE CONTROLE DO ESPAÇO AÉREO. **Portaria DECEA nº 01/SDOP, de 17 de janeiro de 2006.** Aprova a modificação do Manual de Estações Meteorológicas de Superfície. Brasília, 2006.

BRASIL. MINISTÉRIO DA DEFESA. DIRETORIA DE ELETRÔNICA E PROTEÇÃO AO VÔO. **Portaria DEPV nº 12/DIRPV, de 10 de fevereiro de 2000.** Aprova a Instrução de preenchimento do Formulário IEPV 105-78, Observação Meteorológica de Superfície. Brasília, 2000. (DEPV nº 028).

BRASIL. MINISTÉRIO DA DEFESA. DEPARTAMENTO DE CONTROLE DO ESPAÇO AÉRO. **Manual auxiliar de rotas aéreas**, 3. ed., de 04 de novembro de 1999.

BYERS, H. R. **General meteorology.** New York: 4. ed. MC Graw Hill Book Company, p. 480-510, 1959, 540 p.

CABRAL, E. Aspectos climatológicos da ocorrência de nevoeiro na área do aeroporto Internacional de São Paulo/Guarulhos. In: Encontro Nacional de

Geógrafos, 8., e Congresso Brasileiro de Meteorologia, 6., 1990. **Anais...** Salvador, BA: SBMeT, 1990.

CERECEDA, P.; SCHEMENAUER, R. S. The occurrence of fog in Chile. **Journal of Applied Meteorology**, vol. 30, p. 1097-1105, Aug. 1991.

CENAPAD/UNICAMP. **Introdução ao SAS – Programação 1: Data Step e PROC Step**. Versão mar. 2007. 113 p.

COLLET, D. **Modelling survival data in medical research**. New York, Chapman and Hall, 1994.

COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. São Paulo, Edgard Blücher, 1. ed., 2006.

COSTANZA, M. C.; AFIFI, A. A. Comparisons of stopping rules in forward stepwise discriminant analysis. **Journal of the American Statistical Association**, v. 74, p. 777-785, 1979.

CROFT, P. J.; PFOST, R.L.; MEDLIN, J.M.; JOHNSON, G.A. Fog forecasting for Southern region: a concept model approach. **Weather and Forecasting**, v. 12, p. 545-546, Sept. 1997.

DUYNKERKE, P. G. Radiation fog: a comparison of model simulation with detailed observations. **Monthly Weather Review**, v. 119, p. 324-341, Feb. 1991.

FISHER, E. L.; CAPLAN, P. An experiment in numerical prediction of fog and stratus. **Journal of the Atmospheric Science**, v. 20, p. 425-437, 1963.

FITZJARRALD, D. R.; LALA, G. G. Hudson Valley fog environments. **Journal of Applied Meteorology**, v. 28, p. 1303-1328, 1989.

FRANÇA, V. D. J. **Avaliação da metodologia de previsão de nevoeiro e visibilidade horizontal**. Dissertação de mestrado em Meteorologia. São José dos Campos, INPE, 2008, 169 p.

FRIEDLEIN, M. T. Dense fog climatology Chicago O'Hare International Airport, July 1996-April 2002. **Bulletin of the American Meteorological Society**, v. 85, p. 515-517, Apr. 2004.

GEORGE, J. J. **Fog. Compedium meteor**. Boston: Amer. meteor. Soc. p. 1179 – 1189, 1951.

GREAT BRITAIN. METEOROLOGICAL OFFICE. **Handbook of aviation meteorology**. 3. ed. London: HMSO, 1994. 401p. ISBN: 0114003653

BERRY Jr, F. A.; BOLLAY, E.; BEERS, N. R. **Handbook of meteorology**. New York: MacGraw-hill, 1945.

HOLETS, S.; SWANSON, R. N. High-inversion fog episodes in Central California. **Journal of Applied Meteorology**, v. 20, p. 890-899, Aug. 1981.

HOOVER, R. A. Forecasting radiation fog at Elkins, W. VA. **Monthly Weather Review**, v. 78, n. 5, May 1950.

INFRAERO, 2007.

KOZIARA, M. C.; RENARD, R. J.; THOMPSON, W. J. Estimating marine fog probability using a model output statistics scheme. **Monthly Weather Review**, v. 111, p. 2333-2340, Dec. 1983.

LA CUESTA, F. C. ; MARTINEZ, C. G-L. **Predicción de nieblas**. Madrid, Secretaria General Técnica, 1. ed., 1993

LEIPPER, D. F. Fog on the U.S. West Coast: a review. **Bulletin of the American Meteorological Society**, v. 75, n<sup>o</sup> 2, Feb. 1994.

LEIPPER, D.F. Fog forecasting objectively in the Califórnia coastal área using LIBS. **Weather and Forecasting**, v. 10, p. 741-762, Dec. 1995.

LIMA, J.S. **Previsão de ocorrência e duração de Nevoeiro em Porto Alegre**. método objetivo. São José dos Campos: Instituto de Proteção ao Vôo do Ministério da Aeronáutica, p. 1-18, 1982.

MACHADO, A. J. **As circulações locais associadas à formação de nevoeiro na grande São Paulo**. Dissertação de Mestrado em Meteorologia, Universidade de São Paulo, 1993, 143 p.

MAINDONALD, J. H.; **Using R for Data Analysis and Graphics**; introduction, code and commentary. Camberra: Australian National University, 2004.

MASON, J. The physics of radiation fog. **J. Meteorol. Soc. Jpn.** v. 60, p. 486-498, 1982.

MEDEIROS, A. M. S.; AB'SABER, A. N. **Introdução à geomorfologia de São José dos Campos (SP)**. São Paulo: Instituto de Geografia, USP Série Geomorfologia, n<sup>o</sup> 12, p. 2-4, 1969.

MEYER, M. B.; LALA, G. G. Climatological aspects of radiation fog occurrence at Albany, New York. **Journal of Climate**, v.3, p. 577-586, May 1990.

MIKEY, J.; GREENLAND, S. A study of the impact of confounder selection criteria on effect estimation. **American Journal of Epidemiology**, v. 129, p. 125-137, 1989.

MOSCATI, N. R. Acidentes Aéreos associados às condições meteorológicas adversas e melhoria contínua dos sistemas de aeronaves. **Boletim da Sociedade Brasileira de Meteorologia**, v. 29, nº 3, novembro de 2005.

MUSSON-GENON, L. Numerical simulation of a fog event with a one-dimensional boundary layer model. **Monthly Weather Review**, v. 115, p. 592-607, Feb. 1987.

OLIVEIRA, L. L. ; VIANELLO, R. L. ; FERREIRA, N. J. **Meteorologia Fundamental**. 1. ed. Erechim: EdIFAPES- Livraria e Editora, 2001. 430 p.

OLIVEIRA, G. A. **Método Estatístico no auxílio à previsão de nevoeiro para o aeródromo de Guarulhos**. Dissertação de Mestrado em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002. 105 p.

OLIVEIRA, V. M.; FEDEROVA, N. Condições para formação de nevoeiro em Pelotas: Parte I – Análise sinótica e estatística (UFPel). In: CONGRESSO BRASILEIRO DE METEOROLOGIA 10./ CONGRESSO DA FLISMET, 8., 1998, Brasília. **Anais...** Rio de Janeiro: SBMet, 1998 a.

OLIVEIRA, V. M.; FEDEROVA, N. Condições para formação de nevoeiro em Pelotas: Parte II – Análise estatística das correlações (UFPel). In: CONGRESSO BRASILEIRO DE METEOROLOGIA 10./ CONGRESSO DA FLISMET, 8., 1998, Brasília. **Anais...** Rio de Janeiro: SBMet, 1998 b.

OLIVEIRA, V. M., FEDEROVA, N. Condições para formação de nevoeiro em Pelotas: Parte III - Análise Termodinâmica(UFPel). In: CONGRESSO BRASILEIRO DE METEOROLOGIA 10./ CONGRESSO DA FLISMET, 8., 1998, Brasília. **Anais...** Rio de Janeiro: SBMet, 1998 c.

OLIVEIRA, V. M.; FEDEROVA, N. Condições para formação de nevoeiro em Pelotas: Parte IV – Método de previsão (UFPel). In: CONGRESSO BRASILEIRO DE METEOROLOGIA 10./ CONGRESSO DA FLISMET, 8., 1998, Brasília. **Anais...** Rio de Janeiro: SBMet, 1998 d.

PATRIOTA, A. G. **Análise de sobrevivência**. (Instituto de Matemática e Estatística da USP) 2006. disponível em:  
<<http://www.ime.usp.br/~patriota/as/as.html>> Acesso em: 10 dez. 2007. Notas de aula.

PERELLA, A. C. F., SAKURAGI, J. **Estudo de nevoeiro em São José dos Campos**. São José dos Campos: Universidade do Vale do Paraíba, Laboratório de Meteorologia, 1998.

PETERSSSEN, S. Stability and instability in relation to weather phenomena. In: \_\_\_\_\_. **Weather analysis an forecasting**. Mc-Graw Hill, New York, p. 110-125/130-131, 1940.

PILIÉ, R. J.; MACK, E. J.; ROGERS, C. W.; KATZ, U.; KOCCMOND, W.C. The formation of marine fog and the development of fog-stratus along the California Coast. **Journal of the Applied Meteorology**, v. 18, p. 1275-1286, 1979.

PIVA, E. D.; FEDOROVA, N. Parâmetros típicos para a ocorrência de nevoeiro de radiação. Parte I: Características em superfícies. (UFPel). In: CONGRESSO BRASILEIRO DE METEOROLOGIA 10./ CONGRESSO DA FLISMET, 8., 1998, Brasília. **Anais...** Rio de Janeiro: SBMet,. 1998 a.

PIVA, E. D.; FEDOROVA, N. Parâmetros típicos para a ocorrência de nevoeiro de radiação. Parte II: Características do perfil vertical. Duração do nevoeiro (UFPel). In: CONGRESSO BRASILEIRO DE METEOROLOGIA 10./ CONGRESSO DA FLISMET, 8., 1998, Brasília. **Anais...** Rio de Janeiro: SBMet, 1998 b.

PIVA, E. D.; FEDOROVA, N. Parâmetros típicos para a ocorrência de nevoeiro de radiação. Parte III: Características da radiação difusa. Análise sinótica. Esquema para análise do nevoeiro de radiação. (UFPel). In: CONGRESSO BRASILEIRO DE METEOROLOGIA 10./ CONGRESSO DA FLISMET, 8., 1998, Brasília. **Anais...** Rio de Janeiro: SBMet,. 1998 c.

REIS, M. C. S. A. **Nevoeiro no Aeroporto da Portela, Lisboa**. Dissertação de Mestrado, Curso de Pós-graduação em Probabilidade e Estatística, Faculdade de Ciências da Universidade de Lisboa, 1995, 127 p.

REYMANN, W. T.; BROWN, R.; CAUGHY, S. J.; GARLAND, J. A.; READINGS, C. J.: The Physics of radiation fog. Part I: a field study. **Quart. J. Roy. Meteorol. Soc.**, v 102, n. 432, p. 313-333, 1976.

SILVA DIAS, P. L.; MACHADO, A. J. The role of local circulations in summertime convective development and nocturnal fog in São Paulo, Brazil. **Boundary-Layer Meteorology**, v. 82, p. 135-157, 1997.

SANTOS, C. B. Características da formação de nevoeiro para Curitiba – Fase II. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 13., 2004, Fortaleza. **Anais...** São José dos Campos: INPE.CPTEC, , 2004. v. 12, 7 p.

SHIMAKURA, S. **Bioestatística A**. Notas de aula do Curso CE055, Departamento de Estatística da UFPR, 2007. 106 p.

TAYLOR, G. F. **Aeronautical Meteorology**. Pitman, 1941. 455 p.

TEIXEIRA, J.; MIRANDA, P. M. A. Fog prediction at Lisbon Airport using a one-dimensional boundary layer model. **Meteorological Applied**, v. 8, p. 497-505, 2001.

TREMANT, M. The forecasting of sea fog. **The Meteorological Magazine**, v. 118, no. 1401, p. 69-75, Apr. 1989.

TURTON, D.; BROWN, R. A comparison of a numerical model of radiation fog with detailed observations. **Quarterly Journal of the Royal Meteorological Society**, v. 113, p. 37-54, 1987.

TORGO, L. **Introdução à Programação em R**. Faculdade de Economia, Universidade de Porto, Grupo de Matemática e Informática. Out. 2006, 90 p.

TUTORIAL SAS, 2008. Disponível em: <http://www.sas.com>. Acesso em: 10 dez. 2007.

VERZANI, J. **SimpleR Using R for introductory statistics**. Notas de aula do Curso de Estatística, Aug. 2002. (disponível com John Verzani, e-mail [verzani@math.csi.cuny.edu](mailto:verzani@math.csi.cuny.edu)).

XINMEI, H.; LYONS, T. J.; PITTS, R. O. Fog formation at Perth Airport. **Australian Meteorological Magazine**, v. 38, p. 99-106, 1990.

WALLACE, J. M.; HOBBS, P. V. **Atmospheric Science – An introductory Survey**. 2. ed. Amsterdam: Elsevier, 2006. 483 p.

WILKS, D. S. **Statistical methods in the atmospheric sciences**. 2. ed. Amsterdam: Elsevier, , 1996. 627 p. (International Geophysics Series, vol. 91, 2006).

WILLET, H. C. Fog and haze, their causes, distribution, and forecasting. **Monthly Weather Review**, v. 56, n<sup>o</sup> 11, Nov. 1928.

WORLD METEOROLOGICAL ORGANIZATIONS (WMO) **International meteorological vocabulary**. World Meteorological Organization, Geneva, 1966. 275 p.

## BIBLIOGRAFIA CONSULTADA

Website de Alexandre Galvão Patriota: <http://www.ime.usp.br/~patriota/as/as.html>.

Website da comunidade R: [http://br.groups.yahoo.com/group/R\\_STAT/](http://br.groups.yahoo.com/group/R_STAT/).

Website da comunidade SAS: <http://br.groups.yahoo.com/group/SASbrasil/>.

Website do CPTEC/INPE: <http://www.cptec.inpe.br>

Website do IAG/USP – <http://master.iag.usp.br>

Website de Fernando Henrique Ferraz Pereira da Rosa: (<http://www.feferraz.net/br/listas.html#mae514>).

Website do Governo Australiano: <http://www.auf.asn.au/meteorology/section3.html>.

Website da Google Earth: <http://earth.google.com>

Website de Paulo Justiniano Ribeiro Junior: <http://www.est.ufpr.br/~paulojus/>.

Website do Programa de Computação Científica (PROCC/FIOCRUZ): <http://www.procc.fiocruz.br/~cursos/>

Website do R: <http://www.r-project.org>.

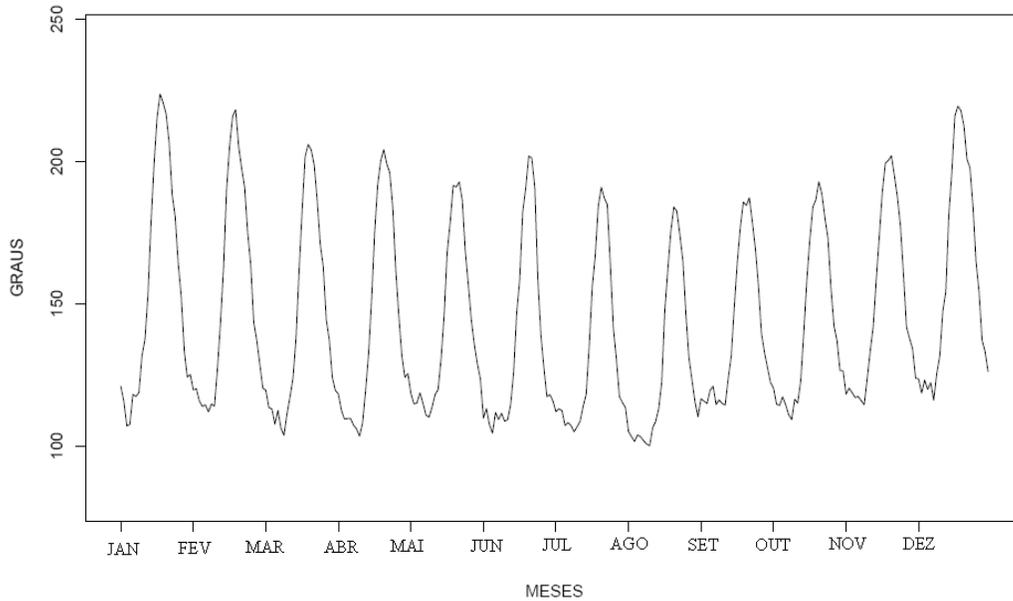
Website da SAS Business Intelligence Software and Predictive Analytics: <http://www.sas.com>.

Website SAS Fellowship Program: <http://www.sas.com/govedu/edu/programs/studentresources.html>.

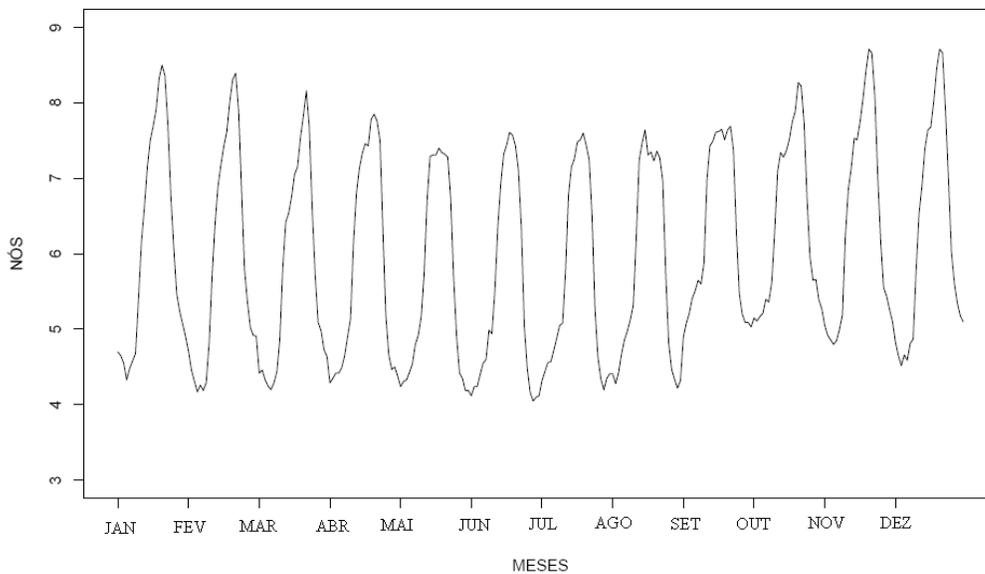


## APÊNDICE A

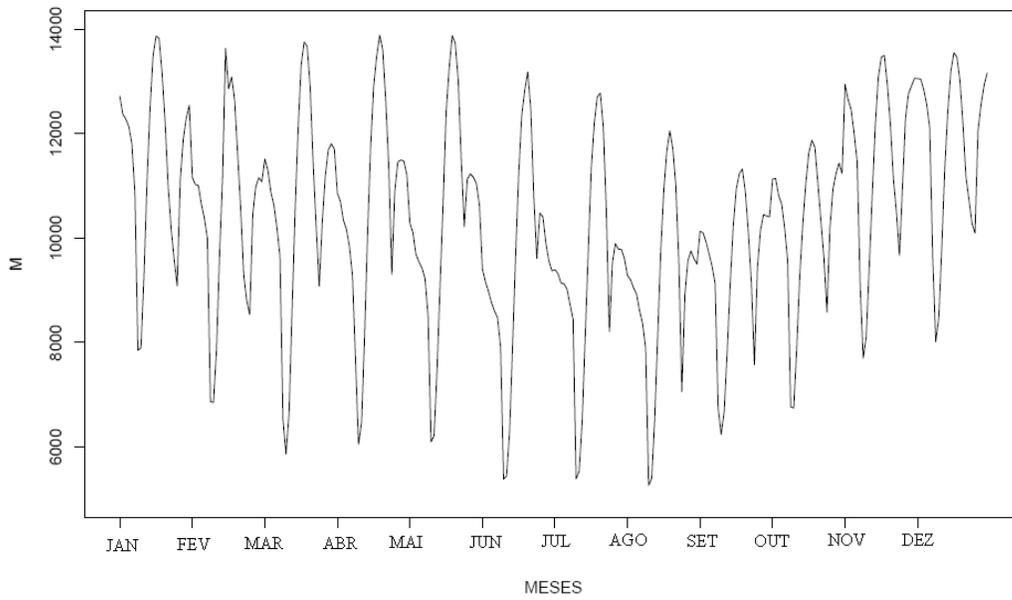
### SÉRIES TEMPORAIS CLIMATOLÓGICAS (1988 - 2005) MENSAIS DAS VARIÁVEIS METEOROLÓGICAS PARA O AEROPORTO INTERNACIONAL DE SÃO PAULO



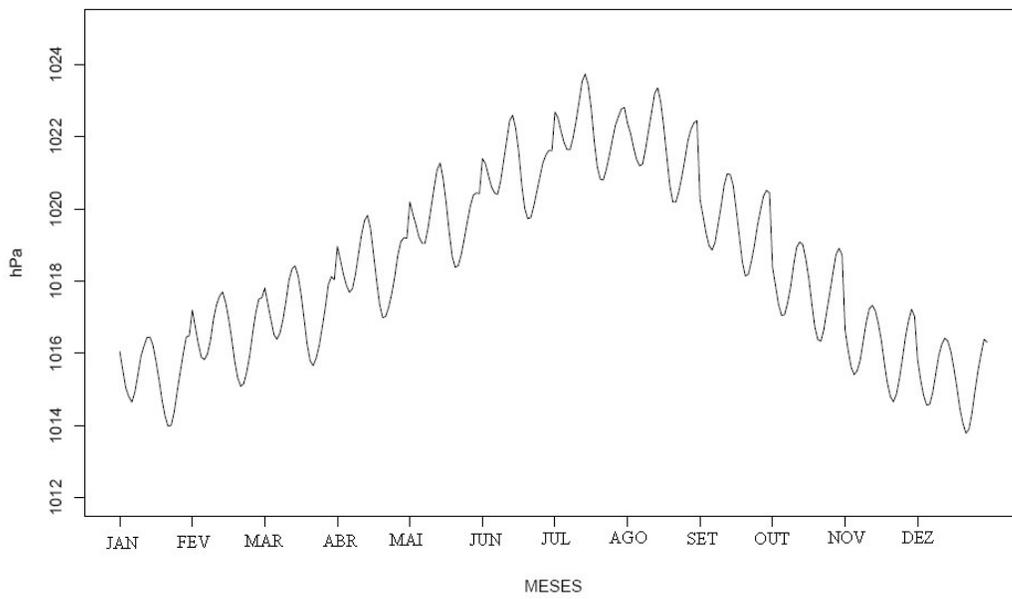
A.1 – Direção do Vento em graus ( ° )para o período de 1988 à 2005.



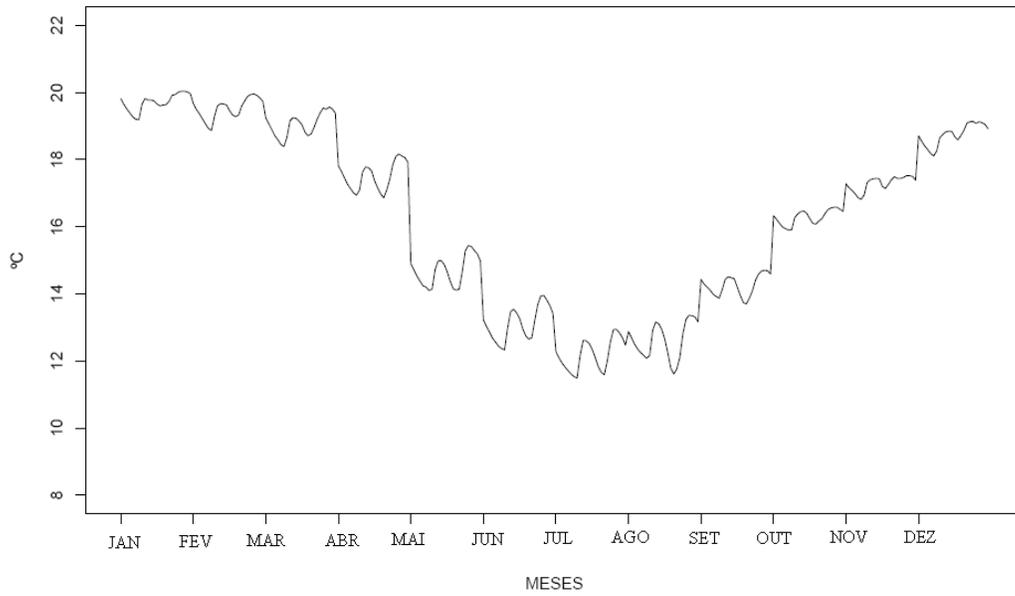
A.2 – Velocidade do Vento em Nós (KT) para o período de 1988 à 2005.



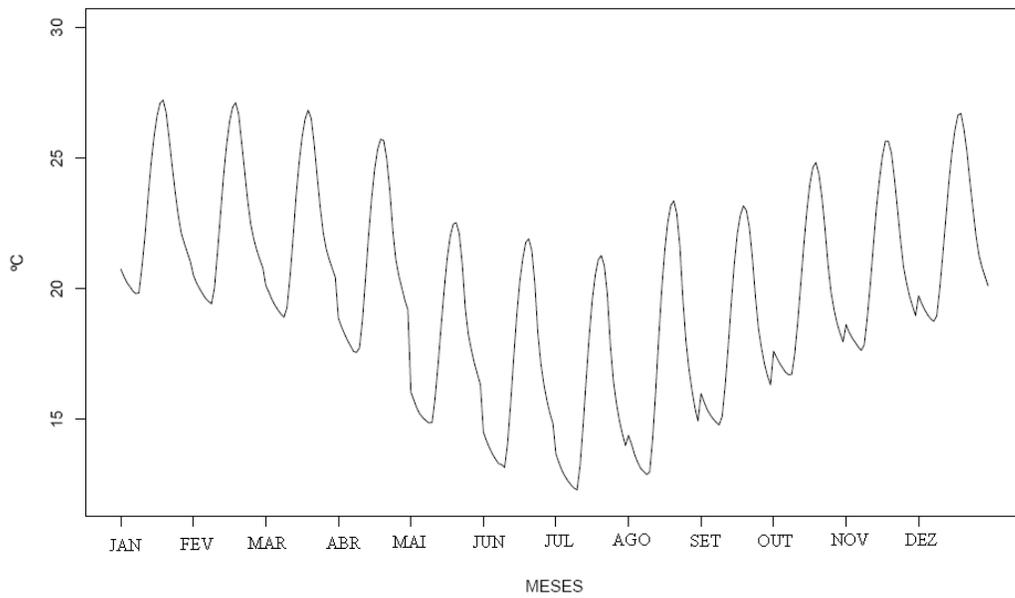
A.3 – Visibilidade em Metros (M) para o período de 1988 à 2005.



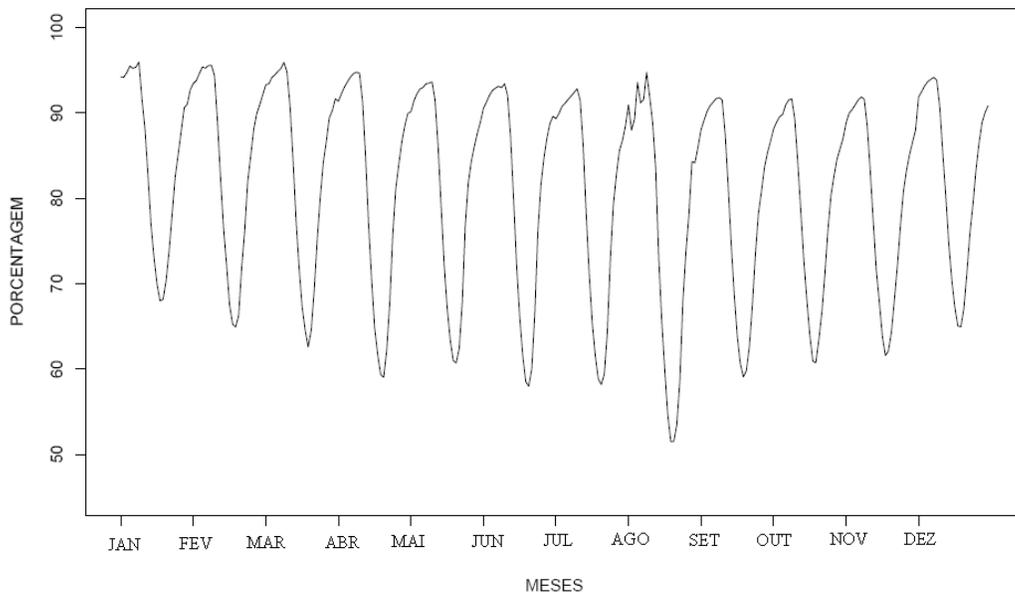
A.4 – Pressão em hectopascal (hPa) para o período de 1988 à 2005.



A.5 – Td em Celsius (°C) para o período de 1988 à 2005.



A.6 – Tw em Celsius (°C) para o período de 1988 à 2005.



A.7 – UR em Porcentagem (%) para o período de 1988 à 2005.

## APÊNDICE B

### CONCEITOS BÁSICOS SOBRE ANÁLISE DE SOBREVIVÊNCIA

Análise de sobrevivência é um conjunto de técnicas e modelos estatísticos usados para a análise de dados cuja variável resposta é o tempo até a ocorrência de um evento de interesse. O termo “Análise de Sobrevivência” se refere basicamente a situações médicas envolvendo dados censurados, isto é, observações incompletas ou parciais e, apesar disso, a informação é importante para a análise. Entretanto, as técnicas estatísticas utilizadas neste método são de ampla aplicação em outras áreas do conhecimento. O objetivo principal da Análise de Sobrevivência é caracterizar a distribuição do tempo de sobrevivência para uma dada população, para compará-lo entre diferentes grupos, ou para assegurar a dependência da variável tempo de sobrevivência com variáveis independentes (variáveis concomitantes, covariáveis ou fatores prognósticos).

Há um vasto material na *web* sobre Análise de Sobrevivência. Aqui, destaca-se a página do Programa de Computação Científica (PROCC/FIOCRUZ) <http://www.procc.fiocruz.br/~cursos/>, que disponibiliza para *download* as notas de aula do “Curso de Sobrevivência”, ministrado pela Profa. Marília Sá Carneiro e sua equipe; as Listas de Exercícios resolvidas do Curso “Introdução a Análise de Sobrevivência e Aplicações” (MAE514, ministrada no Instituto de Matemática Estatística (IME), da USP), que utiliza o Programa R para os cálculos estatísticos, disponível na página pessoal de Fernando Henrique Ferraz Pereira da Rosa (<http://www.feferraz.net/br/listas.html#mae514>). Essas listas de exercício ajudam o leitor que está iniciando na Programação R a interpretar os resultados gerados; Notas de aula do Prof. Alexandre Galvão Patriota, que ministra o Curso de Estatística no IME/USP intitulado “Análise de Sobrevivência”, disponível na página <http://www.ime.usp.br/~patriota/as/as.html>; entre outros.

Nesse Apêndice, descrevem-se sucintamente as noções básicas sobre o método de Análise de Sobrevivência, apresentadas em maiores detalhes em Colosimo e Giolo (2006). Vale destacar, também, o livro de Allison (1995), em inglês, que descreve de forma didática o método de Análise de Sobrevivência com aplicações para as ciências sociais e biomédicas utilizando o *software* SAS para o cálculo das estatísticas utilizadas. Para uma apresentação ampla e atualizada de vários modelos e técnicas probabilísticas, sugere-se consultar Wilks (1996). Por fim, descreve-se resumidamente a metodologia proposta por Reis (1995), referente à previsão de duração de nevoeiros usando a Análise de Sobrevivência, a qual serviu de base para a construção do método de previsão de duração de nevoeiro para o Aeroporto Internacional de São Paulo, desenvolvido no presente trabalho.

## **B.1 – ASPECTOS GERAIS**

Os conjuntos de dados de sobrevivência são caracterizados pelas variáveis: Tempo de Falha ( $T_F$ ), censura e truncamento, onde as duas primeiras variáveis constituem a resposta. O  $T_F$  é uma variável aleatória não-negativa, geralmente contínua, podendo ser considerado como o tempo medido a partir de um ponto referencial até a ocorrência de um evento de interesse. Para o caso do Aeroporto Internacional de São Paulo, o tempo de falha é a dissipação do nevoeiro. A censura é a observação parcial da resposta e ocorre em situações em que, por alguma razão, o acompanhamento do tempo de sobrevivência de um indivíduo foi interrompido por causas diferentes das estudadas. Há vários tipos de censuras (à direita, à esquerda e intervalar), e suas definições e uma explicação mais detalhada sobre o assunto são dadas em Reis (1995), Colosimo e Giolo (2006) e Marília et al. (2007). O truncamento, muitas vezes confundido com censura, é a observação parcial da resposta devido ao processo de seleção adotado, isto é, exclui-se do estudo indivíduos que ainda

não passaram por um certo evento. Por exemplo, no estudo de nevoeiro usando análise de sobrevivência, utilizam-se apenas os dados de visibilidades inferiores a 1.000 m, enquanto os dados com visibilidades maiores que 1.000 m são excluídos das análises em função da própria definição de nevoeiro (*Handbook of Aviation Meteorology*, 1994). Tem-se truncamento à direita e à esquerda, onde no truncamento à esquerda o tempo de sobrevida (conhecido antes de começar o estudo) considera indivíduos acompanhados até o desfecho, porém se desconhece os indivíduos que tiveram o desfecho antes do início do estudo, enquanto truncamento à direita, onde são incluídos apenas indivíduos que sofreram o evento (MARÍLIA et al., 2007).

O  $T_F$  geralmente é especificado pela sua Função de Sobrevivência ( $S(t)$ ) ou pela Função de Taxa de Falha ( $\lambda(t)$ ), ou risco. A primeira função é definida como sendo a probabilidade de um indivíduo não falhar até o tempo  $t$  ou, equivalentemente, a probabilidade de um elemento sobreviver ao tempo  $t$ . A segunda função especifica a probabilidade instantânea de um indivíduo sofrer o evento (falha) em um intervalo de tempo  $t$  e  $t + \Delta t$ , dado que ele sobreviveu até  $t$ , sendo expressas, respectivamente, por:

$$S(t) = P(T_F \geq t) \dots\dots\dots(1)$$

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T_F < t + \Delta t | T_F \geq t)}{\Delta t} \dots\dots\dots(2)$$

Onde P é a Probabilidade.

Na Equação 2, sendo  $\Delta t$  infinitamente pequeno,  $\lambda(t)$  expressa o risco instantâneo de ocorrência de um evento (falha), dado que até então o evento não tenha ocorrido. Distintas formulações para  $\lambda(t)$  são dadas em Reis (1995) e Colosimo e Giolo (2006).

### **B.1.1 – Cálculo de Estimadores**

Existem duas abordagens estatísticas para estimar a curva de  $(S(t))$ , uma delas considera testes paramétricos e a outra, testes não-paramétricos (Wilks, 2006). Os testes paramétricos são aqueles conduzidos em situações onde se conhece ou se assume que uma distribuição teórica particular é uma representação apropriada para os dados e/ou teste estatístico. Nos testes não-paramétricos, ou testes de distribuições livres, não se impõe um modelo teórico para as falhas e/ou censuras observadas. Para a escolha de um modelo adequado, tanto para testes paramétricos como não-paramétricos, é importante se ater às características de cada uma das distribuições, uma vez que cada distribuição de probabilidade pode gerar estimadores diferentes para a mesma quantidade desconhecida. Assim, a escolha de um modelo probabilístico adequado para descrever o tempo de falha deve ser feita com bastante cuidado para evitar erros grosseiros (COLOSIMO E GIOLO, 2006). Patriota (2006) ressalta que a escolha de uma dessas abordagens não significa a exclusão da outra e por isso, o uso das duas abordagens se faz essencial para garantir resultados fidedignos.

#### **a) Estimadores Paramétricos**

Estimadores paramétricos caracterizam-se pelo uso de covariáveis conjuntas na análise dos dados de sobrevivência. Para modelar as funções de sobrevivências, utilizam-se distribuições de probabilidades, onde se destacam as distribuições Exponencial, de Weibull, Log-normal, Gama e Gama Generalizada.

A função de densidade e probabilidade (*f.d.p.*) com Distribuição Exponencial para a variável aleatória  $T_F$ , e suas respectivas  $(S(t))$  e  $\lambda(t)$ , são dada por:

$$f(t) = \frac{1}{\alpha} \exp\left\{-\left(\frac{t}{\alpha}\right)\right\}, t \geq 0 \dots\dots\dots(3)$$

$$S(t) = \exp\left\{-\left(\frac{t}{\alpha}\right)\right\} \dots\dots\dots(4)$$

$$\lambda(t) = \frac{1}{\alpha}, \text{ para } t \geq 0 \dots\dots\dots(5)$$

Onde  $\alpha$  é o parâmetro de forma da Distribuição Exponencial.

Colosimo e Giolo (2006) alertam que para a distribuição Exponencial, tanto uma unidade velha (com maior tempo de duração) quanto uma unidade nova (com menor tempo de duração) que ainda não falharam, têm o mesmo risco de falhar em um futuro intervalo.

A Distribuição de Weibull para uma variável aleatória  $T_F$ , tem *f.d.p.*, e suas associadas funções  $S(t)$  e  $\lambda(t)$ , expressas por:

$$f(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1} \exp\left\{-\left(\frac{t}{\alpha}\right)^\gamma\right\}, t \geq 0 \dots\dots\dots(6)$$

$$S(t) = \exp\left\{-\left(\frac{t}{\alpha}\right)^\gamma\right\} \dots\dots\dots(7)$$

$$\lambda(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1}, \text{ para } t \geq 0, \alpha \text{ e } \gamma > 0 \dots\dots\dots(8)$$

onde  $\gamma$  e  $\alpha$  são os parâmetros de forma e de escala, respectivamente, sendo ambos positivos. A  $\lambda(t)$  da Distribuição de Weibull é monótona, isto é, ela é crescente se  $\gamma > 1$ , decrescente se  $\gamma < 1$  e constante se  $\gamma = 1$ .

A *f.d.p.* com Distribuição Log-normal para a variável aleatória  $T_F$ , e as funções  $S(t)$  e  $\lambda(t)$  (essas duas últimas funções não apresentam uma forma analítica explícita) são expressas, respectivamente, por:

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{\log(t) - \mu}{\sigma}\right)^2\right\} \text{ para } t > 0. \dots\dots\dots(9)$$

$$S(t) = \Phi\left(\frac{-\log(t) + \mu}{\sigma}\right) \dots\dots\dots(10)$$

$$\lambda(t) = \frac{f(t)}{S(t)} \dots\dots\dots(11)$$

Em que  $\mu$  é a média do logaritmo do tempo de falha,  $\sigma$  é o desvio-padrão e  $\Phi(\cdot)$  é a Função de Distribuição Acumulada de uma normal padrão. A função  $\lambda(t)$  para a Distribuição Log-normal é monótona. Reis (1995) alerta para o fato que uma propriedade pouco atrativa da Distribuição Log-normal em relação a sua  $\lambda(t)$  é que quando se tem  $t = 0$  ela toma o valor 0, cresce até um máximo, e decresce em seguida aproximando-se de zero quando  $t \rightarrow \infty$ , revelando-se como um modelo inadequado para elevados valores de  $t$ .

A *f.d.p.* da Distribuição Gama para a variável aleatória  $T_F$ , que é caracterizada pelos parâmetros  $k$  e  $\alpha$ , em que  $k > 0$  é chamado parâmetro de forma e  $\alpha > 0$  de parâmetro de escala, é expressa por:

$$f(t) = \frac{1}{\Gamma(k)\alpha^k} t^{k-1} \exp\left\{-\left(\frac{t}{\alpha}\right)\right\}, t > 0. \dots\dots\dots(12)$$

onde a função  $\Gamma(k)$  é definida por:

$$\Gamma(k) = \int_0^{\infty} x^{k-1} \exp\{-x\} dx \dots\dots\dots(13)$$

A função  $S(t)$  é dada por:

$$S(t) = \int_0^{\infty} \frac{1}{\Gamma(k)\alpha^k} u^{k-1} \exp\left\{-\left(\frac{u}{\alpha}\right)\right\} du \dots\dots\dots(14)$$

A função  $\lambda(t)$  da Distribuição Gama, é obtida através da relação:

$$\lambda(t) = \frac{f(t)}{S(t)} \dots\dots\dots(15)$$

O Modelo da Distribuição Gama Generalizada é caracterizado por apresentar três parâmetros ( $\gamma$ ,  $k$  e  $\alpha$ ), todos positivos, e incluir como casos particulares todos os casos de modelo apresentados: Exponencial ( $\gamma=k=1$ ), Weibull ( $k=1$ ), Gama ( $\gamma=1$ ) e Log-normal (quando  $k \rightarrow \infty$ ). Esta propriedade da Gama Generalizada faz com que a mesma seja de grande utilidade, por exemplo, na discriminação de modelos probabilísticos alternativos (COLOSIMO E GIOLO, 2006). A *f.d.p.* da Distribuição Gama Generalizada é dada por:

$$f(t) = \frac{\gamma}{\Gamma(k)\alpha^{\gamma k-1}} t^{\gamma k-1} \exp\left\{-\left(\frac{t}{\alpha}\right)^\gamma\right\}, t > 0. \dots\dots\dots(16)$$

Colosimo e Giolo (2006) apresentam ainda um Modelo de Regressão de Weibull com uma Função de Sobrevivência de t condicional a x:

$$S(t|x) = \exp\left\{-\left[\frac{t}{\exp(x'\beta)}\right]^{1/\sigma}\right\}.$$

onde  $x'$ , representa mais de uma covariável.....(17)

Os Modelos de Weibull, Exponencial, Log-normal e Gama Generalizado que foram apresentados anteriormente são caracterizados por quantidades desconhecidas denominadas parâmetros (representado por  $\theta$ ). Em Análise de Sobrevivência estas quantidades são estimadas pelo método da Máxima Verossimilhança ( $L(\theta)$ ), expressa por:

$$L(\theta) = \prod_{i=1}^n [\lambda(t_i; \theta)]^{\delta_i} S(t_i; \theta) \dots\dots\dots(18)$$

em que  $\delta_i$  é a variável indicadora de falha ou censura.

A Tabela B.1 sumaria os parâmetros para cada uma dessas distribuições de probabilidade.

Tabela B.1 - Parâmetros associados com algumas distribuições de probabilidade.

DISTRIBUIÇÕES	PARÂMETROS
Exponencial	$\alpha$
Weibull	$\gamma, \alpha$
Log-normal	$\gamma, \alpha$
Gama	$\gamma, \alpha$
Gama generalizada	$\gamma, k, \alpha$

A Equação 18 encontra o valor de  $\theta$  que maximiza a função  $L(\theta)$  ou equivalentemente o logaritmo de  $L(\theta)$ , tal que:

$$U(\theta) = \frac{\partial \log L(\theta)}{\partial \theta} = 0 \dots\dots\dots(19)$$

Para a escolha do melhor modelo que se ajusta a um conjunto de dados existem métodos que utilizam técnicas gráficas. Por exemplo, a comparação da Função de Sobrevivência do modelo proposto com o estimador de Kaplan-Meier, a linearização da Função de Sobrevivência do modelo proposto, que gera como resultado uma reta no gráfico caso o modelo seja adequado e, para modelos que utilizam mais de uma covariável, tem-se a análise de resíduos. Colosimo e Giolo (2006) alertam que técnicas gráficas devem ser utilizadas como meio de rejeitar modelos claramente inapropriados e não para provar que um particular modelo paramétrico é adequado. Uma forma de discriminar modelos que envolvem uma conclusão direta e não envolve qualquer componente subjetivo na sua interpretação é por meio de teste de hipótese. Neste teste, as hipóteses a serem testadas são: hipótese nula, definida por  $H_0$  (se o modelo de interesse é adequado) versus uma hipótese vaga  $H_1$  de que o modelo não é adequado. A estatística da razão de verossimilhanças (TRV) é dada por:

$$TRV = -2 \log \left[ \frac{L(\hat{\theta}_M)}{L(\hat{\theta}_G)} \right] = 2 \left[ \log L(\hat{\theta}_G) - \log L(\hat{\theta}_M) \right] \dots \dots \dots (20)$$

Para o cálculo da TRV deve-se obter os valores dos logaritmos da função de verossimilhança do modelo generalizado  $\left( \log L(\hat{\theta}_G) \right)$  e da função de verossimilhança do modelo de interesse  $\left( \log L(\hat{\theta}_M) \right)$ , em que sob  $H_0$ , tem aproximadamente uma distribuição  $\chi^2$  com graus de liberdade igual à diferença do número de parâmetros  $(\hat{\theta}_G$  e  $\hat{\theta}_M)$  dos modelos sendo comparados. Em Análise de Sobrevivência, este teste é realizado utilizando-se a Distribuição Gama Generalizada que apresenta os modelos Exponencial, de Weibull, Log-normal e Gama, como modelos encaixados (COLOSIMO E GIOLO, 2006).

**b) Estimadores Não-paramétricos**

Há diferentes estimadores não-paramétricos da Função de Sobrevivência, podendo-se citar, os de Nelson Aalen  $\left( S_{\tilde{t}} \right)$ , de Kaplan-Meier  $\left( S_{\hat{t}} \right)$  e Estimador da Tabela de Vida  $\left( S_{\hat{t}} \right)$ , entre outros. Dentre estes, o estimador de Kaplan-Meier tem se mostrado superior aos estimadores de Nelson-Aalen e Tabela de Vida, tanto para a função de sobrevivência em grandes amostras, como com amostras de pequeno e médio porte, sendo chamado de estimador limite-produto (Colosimo e Giolo, 2006), e definido por:

$$\hat{S}(t) = \prod_{j:t_j < t} \left(1 - \frac{d_j}{n_j}\right) \dots\dots\dots(21)$$

Onde:

- $t_1 < t_2 \dots < t_k$ , os  $k$  tempos distintos e ordenados de falha.
- $d_j$  o número de falhas em  $t_j$ ,
- $n_j$  o número de indivíduos sob risco em  $t_j$ , ou seja, os indivíduos que não falharam e não foram censurados ate o instante imediatamente anterior a  $t_j$ .

Onde  $j$  representa a  $j$ -ésima observação  $j = 1, \dots, k$ .

A relação entre  $\hat{S}(t)$  e  $\hat{\lambda}(t)$  é dada por:

$$\hat{\lambda}(t) = -\ln \hat{S}(t) \dots\dots\dots(22)$$

Um intervalo aproximado de  $100(1-\alpha)\%$  de confiança para  $\hat{S}(t)$  é dado por:

$$\hat{S}(t) = z_{\alpha/2} \sqrt{\widehat{Var}(\hat{S}(t))} \dots\dots\dots(23)$$

Entretanto, Colosimo e Giolo (2006) alertam que este intervalo pode apresentar valores negativos e maiores que 1. Nesses casos, o problema é resolvido utilizando uma transformação para  $S(t)$ , representada por  $\hat{U}(t)$  e expressa pela Equação 22, dada por:

$$\hat{U}(t) = \log[-\log(\hat{S}(t))] \dots\dots\dots(24)$$

Assim, um intervalo aproximado de  $100(1-\alpha)\%$  de confiança para  $\hat{S}(t)$ , sempre positivo e menor do que 1 é dado por:

$$[\hat{S}(t)]^{\exp\{\pm z_{\alpha/2} \sqrt{\widehat{Var}(\hat{U}(t))}\}} \dots\dots\dots(25)$$

Embora as técnicas não-paramétricas tenham a vantagem de fácil aplicação aos dados, há situações em que estas apresentam limitações quanto sua aplicação. Por exemplo, nos casos em que os dados apresentam um grande número de covariáveis e uma amostra relativamente pequena, o que resulta em estratos com poucos dados ou até mesmo sem nenhum, e também por não permitir a análise desses estratos conjuntamente.

**c) Estimador Semi-paramétrico**

O Modelo de Regressão de Cox, também conhecido como Modelo de Riscos Proporcionais, permite a análise de dados provenientes de estudos de tempo de vida em que a resposta é o tempo até a ocorrência de um evento de interesse, ajustado por covariáveis (Colosimo e Giolo, 2006). A expressão geral do Modelo de Regressão de Cox é dada por:

$$\lambda(t) = \lambda_0(t)g(x'\beta) \dots\dots\dots(26)$$

Este modelo é composto pelo produto de dois componentes, um não-paramétrico  $\lambda_0(t)$  e outro paramétrico  $g(x'\beta)$ . Onde:

$$g(x'\beta) = \exp\{x'\beta\} = \exp\{\beta_1x_1 + \dots + \beta_px_p\} \dots\dots\dots(27)$$

em que  $\beta$  é o vetor de parâmetros associado às covariáveis  $x$ .

Neste Modelo, a razão das funções de taxa de falha para os indivíduos  $i$  e  $j$  da equação 26 é constante no tempo, por isso é chamado de modelo de riscos proporcionais.

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \frac{\lambda_0(t)\exp\{x_i'\beta\}}{\lambda_0(t)\exp\{x_j'\beta\}} = \exp\{x_i'\beta - x_j'\beta\} \dots\dots\dots(28)$$

**B.1.2 - Etapas de um Estudo de Análise de Sobrevivência**

Os estudos de análise de sobrevivência pode ser divididos em três etapas (COLOSIMO E GIOLO, 2006), a saber:

- i) Formulação da hipótese de interesse: identificação de fatores que tenha relação com o tempo de falha;
- ii) Planejamento e coleta de dados: formatação dos dados para análise;
- iii) Análise estatística dos dados para testar a hipótese formulada.

**B.2 – DESCRIÇÃO SUCINTA DA METODOLOGIA PROPOSTA POR REIS (1995)**

Reis (1995) utilizou a Análise de Sobrevivência para construir um modelo de previsão da duração do nevoeiro para o Aeroporto da Portela em Lisboa, Portugal. Em particular, utilizou o Modelo de Cox e o de Regressão Exponencial. Para isso, considerou os valores horários de visibilidade para definir o início (visibilidade inferior a 1000 m) e a dissipação (visibilidade igual ou superior a 1000 m) do nevoeiro. O período dos dados utilizados foi de 1 de janeiro de 1970 a 31 de dezembro de 1989 e a variável disponível foi a visibilidade horária. Para identificar as situações de nevoeiros, elaborou um

programa escrito em PASCAL para filtrar os valores de visibilidades menores que 1000 m.

Na determinação do Modelo de Cox que melhor se ajustou aos dados, utilizou-se o pacote estatístico *Statistica*.

$$\lambda(t|x) = \lambda_0(t)g(x'\beta) \dots\dots\dots(29)$$

Para estimação do Parâmetro  $\beta$ , utilizou-se o método de máxima verossimilhança, equação 28, recorrendo-se ao método de Newton-Raphson para resolução das equações de verossimilhança.

$$L(\beta) = \prod_{i=1}^k \frac{\exp\{s_i'\beta\}}{\left[ \sum_{j \in R(t_i)} \exp\{x_j'\beta\} \right]^{d_i}} \dots\dots\dots(30)$$

onde  $R(t_i)$  é o conjunto do índices das observações sob risco no tempo  $t_i$  e  $s_i$ , o vetor formado pela soma das correspondentes p covariáveis para os indivíduos que falham no mesmo tempo t.

A significância geral do modelo utilizado foi avaliada usando a distribuição do  $\chi^2$  e através dos logaritmos da razão de verossimilhança do modelo nulo (quando todos os  $\beta$ 's forem nulos) e do modelo ajustado.

Utilizou-se também para testar a hipótese da nulidade dos efeitos das covariáveis a estatística da razão de verossimilhança (TRV), expressa por:

$$TRV = -2 \log \left[ \frac{L(\hat{\theta}_M)}{L(\hat{\theta}_G)} \right] = 2 \left[ \log L(\hat{\theta}_G) - \log L(\hat{\theta}_M) \right] \dots \dots \dots (31)$$

Na aplicação das covariáveis, foram consideradas:

- O mês de ocorrência do nevoeiro  $x_1$ ;
- A hora de início do nevoeiro  $x_2$ ;
- Valores agrupados da variável  $x_2$  em três grupos horários:
  - Grupo (1): das 19h às 4 h;
  - Grupo (2): das 5h às 8h e
  - Grupo (3): das 19h às 18h, que foi feito devido ao grande número de valores da covariável  $x_2$  fez-se agrupamentos dos seus valores, produzindo-se assim uma variável que toma um menor número de valores distintos;
- A visibilidade média  $x_4$ , que é a média dos valores de visibilidade registrada de cada nevoeiro até o momento da observação;
- Visibilidade inicial  $x_5$
- Censura  $x_6$ .

Na aplicação do Modelo de Cox aos dados utilizados, o primeiro modelo foi testado na versão completa, ou seja, foram consideradas todas as variáveis  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$ ,  $x_5$  e  $x_6$ , e feitas todas as combinações possíveis dessas covariáveis. Porém, apenas as covariáveis  $x_3$  e  $x_4$  apresentaram valores de estatística superior a dois em valor absoluto (valor de significância considerado).

No conjunto de dados utilizado não houve observações censuradas. Porém, foi necessária a definição da variável  $x_6$  para a utilização do *software* que tomou sempre o valor zero, que corresponde ao código para observação completa.

Tabela B.2 - Resultado do teste para o Modelo de Cox completo.

VARIÁVEIS	$\beta$	ERRO PADRÃO	t
$x_1$	0.000886	0.009709	0.09124
$x_2$	-0.010649	0.008844	-1.20404
$x_3$	0.3777839	0.069056	5.147245
$x_4$	0.001948	0.000263	7.40334
$x_5$	-0.000266	0.000207	-1.28682

FONTE: Reis (1995).

Segundo Reis (1995), do ponto de vista meteorológico, uma justificativa da significância das variáveis  $x_3$  e  $x_4$  em relação às outras variáveis é que  $x_1$  não apresentou significância pelo fato do nevoeiro de radiação apresentar características de duração independente do mês. No caso da variável  $x_3$ , esta foi mais significativa do que  $x_2$  devido o nevoeiro de radiação apresentar características de durações parecidas dentro do mesmo grupo horário, pelo menos no que se refere à radiação solar. Sendo assim, fez-se  $\beta_1 = \beta_2 = \beta_5 = 0$  (Tabela B.3).

Tabela B.3 - Resultado do teste para o Modelo de Cox considerando as variáveis significativas ( $x_3$  e  $x_4$ ).

VARIÁVEIS	$\beta$	ERRO PADRÃO	t
$x_3$	0.355573	0.062995	5.644475
$x_4$	0.001698	0.000192	8.848129

FONTE: Reis (1995).

Utilizou-se o Modelo de Regressão Exponencial para modelar a duração do nevoeiro no Aeroporto da Portela, tal que:

$$S(t|x) = \exp[-\exp(\log t - x\beta)] \dots \dots \dots (32)$$

E como na aplicação para o Modelo de Cox, ajustou-se o Modelo de Regressão Exponencial contendo todas as covariáveis escolhidas (modelo completo), conforme visto na Tabela B.4.

Tabela B.4 – Resultado do teste para o Modelo de Regressão Exponencial completo (considerando todas as variáveis).

VARIÁVEL	$\beta$	ERRO PADRÃO	t
$x_1$	-0.001185	0.009674	-0.12252
$x_2$	0.008021	0.008725	0.91933
$x_3$	-0.340035	0.067091	-5.06825
$x_4$	-0.001886	0.000250	-7.53108
$x_5$	0.000232	0.000205	1.13153
constante	2.863135	0.154944	00000000

FONTE: Reis (1995).

Em seguida, fez-se todas as combinações de covariáveis possíveis e igualmente como apresentou o Modelo de Cox, o melhor modelo foi aquele em que  $\beta_1 = \beta_2 = \beta_5 = 0$ . Portanto, as covariáveis regressoras com efeito significativo foram  $x_3$  e  $x_4$  (Tabela B.5).

Tabela B.5 - Resultado do teste para o Modelo de Regressão Exponencial com as variáveis significativas  $x_3$  e  $x_4$ .

VARIÁVEIS	B	ERRO PADRÃO	t
$x_3$	-0.326223	0.061894	-5.27064
$x_4$	-0.001678	0.000176	-9.55252
constante	2.910647	0.132228	22.01229

FONTE: Reis (1995).

Assim, o Modelo Exponencial mais adequado foi escrito por:

$$S(t|x) = \exp[-\exp(\log t - 2.910647 + 0.326223x_3 + 0.001678x_4)] \dots\dots\dots(33)$$

Esse Modelo apresentou a vantagem de determinar uma expressão analítica e se obter a probabilidade de um dado nevoeiro durar além de um tempo t, conhecidas a hora em que se iniciou (t), o grupo horário a que pertence ( $x_3$ ) e a visibilidade média ( $x_4$ ) registrada até o momento.

Para a verificação do modelo, utilizou-se o método dos resíduos, onde um comportamento linear dos pontos representados foi alcançado, mostrando que

o modelo foi adequado. Assim, das 615 observações de nevoeiros (período entre 1 de janeiro de 1970 a 31 de dezembro de 1989), 494 observações (período entre 1 de janeiro de 1970 a 31 de dezembro de 1985) foram utilizadas para a construção do modelo e 121 observações de nevoeiros (período entre 1 de janeiro de 1986 a 31 de dezembro de 1989), para sua validação.

Para a validação do modelo, comparou-se à Função de Sobrevivência Estimada com a Função de Sobrevivência Empírica para valores fixo das covariáveis. O fato de  $x_3$  tomar apenas três valores distintos simplificou a tarefa, já para a covariável  $x_4$  (variável contínua que toma valores de 0 a 1000) a situação foi mais complicada, optando-se então em dividi-la da seguinte maneira: para cada valor fixo de  $x_3$ , os valores possíveis para  $x_4$  em dois grupos, conforme visto na Tabela B.6.

Tabela B.6 – Divisão da covariável  $x_4$  em apenas dois grupos para cada valor fixo de  $x_3$ .

<b>Covariável</b>	<b>grupos considerados</b>					
$x_3$	<b>1</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>3</b>	<b>3</b>
$x_4$	<b>250</b>	<b>750</b>	<b>250</b>	<b>750</b>	<b>250</b>	<b>750</b>
Nº de obs. disponível	<b>17</b>	<b>18</b>	<b>32</b>	<b>29</b>	<b>7</b>	<b>18</b>

FONTE: Reis (1995).

Para o grupo horário 1 para os valores de visibilidades próximos a 250 m e 750 m, as curvas de sobrevivências empírica e estimada apresentaram-se muito próximas, indicando ótimos resultados. Para o grupo horário 2, para a visibilidade próxima de 250m, a função de sobrevivência estimada apresentou-se acima da curva de sobrevivência empírica, mostrando que para esse período do dia e para valores próximos a essa visibilidade, o modelo prever que os nevoeiros sejam mais persistentes do que efetivamente se verificou. Para o valor de visibilidade próximo a 750 m do grupo horário 2, as curvas também se apresentaram bem próximas, mostrando ótimos resultados. Por último, no grupo horário 3, para o valor de visibilidade próximo a 250 m, a validação do modelo ficou comprometida em função do número reduzido de observações disponíveis (7) para a função de sobrevivência empírica. Reis

(1995) considerou esses resultados satisfatórios, especialmente pela deficiência de informações disponíveis no que diz respeito as variáveis que são determinantes para a persistência do nevoeiro (intensidade do vento, temperatura do ar e temperatura do ponto de orvalho, entre outras). Acrescentou ainda que resultados melhores da validação do modelo poderiam ser esperados se tivesse sido possível dividir o domínio de  $x_4$  num maior número de subintervalos.



## APÊNDICE C

### NOÇÕES SOBRE O AMBIENTE R E SUA UTILIZAÇÃO

A linguagem de programação R permite a computação estatística e gráfica de alto nível, possui interface para outras linguagens, transferidor e armazenamento de dados, conjunto de operadores e coleção de ferramentas para análise de dados, entre outras facilidades (MAINDONALD, 2004). Este é um pacote de distribuição livre e de código aberto, obtido no endereço eletrônico <http://www.r-project.org>, onde estão disponibilizados uma série de documentos adicionais à familiarização, instalação do *software* e utilização da linguagem. Uma facilidade ao uso desse Programa é a *website* da comunidade R ([http://br.groups.yahoo.com/group/R\\_STAT/](http://br.groups.yahoo.com/group/R_STAT/)), onde se troca informações, dúvidas, material e discutem-se assuntos relacionados.

Há um vasto material na *web* sobre o *software* R e sua atualização. Aqui, destaca-se a *website* do Programa de Computação Científica (PROCC/FIOCRUZ) <http://www.procc.fiocruz.br/~cursos/>, que disponibiliza um material introdutório, escrito em português, sobre o *software* R, além de diversos cursos de estatística aplicados e que faz uso desse Programa; e a página do Prof. Paulo Justiniano Ribeiro Junior, do Departamento de Estatística da UFPR (<http://www.est.ufpr.br/~paulojus/>), que disponibiliza tutoriais sobre o R. Assim, nesse Apêndice, apresentam-se sucintamente as noções básicas sobre o *software* R, incluindo-se os programas escritos em R para o cálculo das estatísticas utilizadas.

#### C.1 – UTILIZAÇÃO DO R

A instalação do R vem com um conjunto de pacotes instalados. O R também possui o mecanismo *Rweb* que permite rodá-lo pela *web* sem que seja necessário tê-lo instalado no computador. Para isso, basta estar conectado à

internet e acessar a página do R e no *menu* à esquerda, clicar nos *links* RUIs e *Rweb*. Para instalação do R no computador, deve-se executar o download a partir da página do R, depois de instalado, para iniciá-lo, basta usar o ícone que normalmente está disponível no *desktop* do *windows*, ou então usar o item *Start* no menu. Na execução do R aparece na tela:

```
R version 2.4.1 (2006-12-18)
Copyright (C) 2006 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.

R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes do R em publicações.

Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

>
```

O sinal (>) é o *prompt* do R, é nele que serão introduzidos os comandos que o programa tem que executar.

A linguagem R é baseada em objetos que tem um nome associado e pode armazenar diferentes tipos de coisas, tais como números, textos, vetores, matrizes, etc. Por exemplo:

Para armazenar o valor 15 em um objeto com um nome dado de x, utiliza-se o símbolo (=) ou (<-), este último apenas nas versões mais antigas do R, embora a mais recente também o aceite.

```
|> x = 15
```

Para ver o conteúdo atribuído ao objeto *x* que agora está guardado na memória, basta digitar o nome do objeto no *prompt* do R e pressionar a tecla ENTER do computador:

```
|> x  
[1] 15
```

O símbolo [1] indica que o conteúdo exibido inicia-se com o primeiro elemento do objeto, que neste caso é o único.

Os números na linguagem R podem ser inseridos diretamente no *prompt* sem a necessidade de criação de um objeto.

```
|> 5  
[1] 5  
> 4  
[1] 4  
> 1  
[1] 1
```

As expressões aritméticas também podem ser resolvidas sem a necessidade de criação de um objeto, para isso, observam-se apenas os sinais a serem utilizados: soma (+), subtração (-), multiplicação (\*), divisão (/) e potenciação (^).

```
|> 5 + 10 + 25  
[1] 40
```

```
|> 6*3 + 4/2 - 1  
[1] 19
```

Os nomes de objetos também podem ser dados na forma de texto, devendo sempre iniciar com uma letra, podendo ter pontos ou números apenas nos finais ou entre as letras.

```
| > Umidade.Relativa = 97  
| > Umidade.Relativa  
| [1] 97
```

Também o programa distingue letras maiúsculas e minúsculas nos nomes dos objetos, exemplo:

```
| > umidade.rel原因ativa  
| Erro: objeto "umidade.rel原因ativa" não encontrado
```

Embora o nome dado ao objeto é o mesmo que está na memória, ocorreu o erro porque o objeto digitado está em letra minúscula.

Os vetores são uma estrutura de dados que permite armazenar um conjunto de valores do mesmo tipo e para criá-los é preciso utilizar a função **c()**, que indica ao R os elementos que formam o vetor, separando-os por vírgulas, como exemplificado abaixo, sendo os valores dentro do parêntese os valores de T e T<sub>d</sub>, respectivamente:

```
| > T = c(17.8, 18.3, 18.3, 17.7)  
| > Td = c(17.6, 18.1, 18.0, 17.4)
```

O R permite operações com vetores. Por exemplo, para achar a depressão do ponto de orvalho, que é a diferença de T e T<sub>d</sub>, aplica-se o sinal de subtração (-):

```
| > T - Td  
| [1] 0.2 0.2 0.3 0.3
```

As operações média (*mean*), mediana (*median*), soma (*sum*), máximos (*max*) e mínimos (*min*), o número de elementos (*length*), e os desvios-padrões (*sd*) dos vetores T e T<sub>d</sub>, podem ser calculados pelas seguintes comandos:

```
> mean(T)
```

```
[1] 18.025
```

```
> mean(Td)
```

```
[1] 17.775
```

```
> median(T)
```

```
[1] 18.05
```

```
> median(Td)
```

```
[1] 17.8
```

```
> sum(T)
```

```
[1] 72.1
```

```
> sum(Td)
```

```
[1] 71.1
```

```
> max(T)
```

```
[1] 18.3
```

```
> max(Td)
```

```
[1] 18.1
```

```
> min(T)
```

```
[1] 17.7
```

```
> min(Td)
```

```
[1] 17.4
```

```
> length(T)
```

```
[1] 4
```

```

> length(Td)
[1] 4

> sd(T)
[1] 0.3201562

> sd(Td)
[1] 0.3304038

```

As matrizes na linguagem R são vetores com uma propriedade especial que é a dimensão, pois ela armazena a informação em duas dimensões (Torgo, 2006). Como exemplo, tem-se a temperatura média mensal para os meses de janeiro a dezembro de um determinado ano:

```

> Temp = c(20.6, 19.3, 18.8, 18.1, 16.6, 13.0, 10.3, 13.1, 15.9, 17.4, 17.6, 19.7)
> Temp
[1] 20.6 19.3 18.8 18.1 16.6 13.0 10.3 13.1 15.9 17.4 17.6 19.7

```

Para dividi-las em dois semestres, será utilizada a função `dim(Temp) = c(6, 2)`, onde o número 6 representa as linhas (meses) e o número 2 representa as colunas (semestres):

```

> dim(Temp) = c(6,2)
> Temp
  [,1] [,2]
[1,] 20.6 10.3
[2,] 19.3 13.1
[3,] 18.8 15.9
[4,] 18.1 17.4
[5,] 16.6 17.6
[6,] 13.0 19.7

```

Ainda é possível nomear as colunas de acordo com o semestre, utilizando a função `colnames` e as linhas de acordo com os meses, utilizando a função `rownames`:

```
> rownames(Temp) = c("1ºmês", "2ºmês", "3ºmês", "4ºmês", "5ºmês", "6ºmês")
> colnames(Temp) = c("1ºsemestre", "2ºsemestre")
> Temp
```

	1ºsemestre	2ºsemestre
1ºmês	20.6	10.3
2ºmês	19.3	13.1
3ºmês	18.8	15.9
4ºmês	18.1	17.4
5ºmês	16.6	17.6
6ºmês	13.0	19.7

Inserindo-se outra matriz, para exemplificar as operações com as matrizes:

```
> Temp.Orvalho = c(18.2, 17.4, 16.7, 16.0, 14.6, 10.7, 9.6, 11.9, 13.8, 16.4, 16.4, 19.0)
> Temp.Orvalho
[1] 18.2 17.4 16.7 16.0 14.6 10.7 9.6 11.9 13.8 16.4 16.4 19.0
> dim(Temp.Orvalho) = c(6,2)
> Temp.Orvalho
[ ,1] [ ,2]
[1,] 18.2 9.6
[2,] 17.4 11.9
[3,] 16.7 13.8
[4,] 16.0 16.4
[5,] 14.6 16.4
[6,] 10.7 19.0

> rownames(Temp.Orvalho) = c("1ºmês", "2ºmês", "3ºmês", "4ºmês", "5ºmês", "6ºmês")
> colnames(Temp.Orvalho) = c("1ºsemestre", "2ºsemestre")
> Temp.Orvalho
```

	1ºsemestre	2ºsemestre
1ºmês	18.2	9.6
2ºmês	17.4	11.9
3ºmês	16.7	13.8

4ºmês	16.0	16.4
5ºmês	14.6	16.4
6ºmês	10.7	19.0

Pode-se obter a diferença entre as duas matrizes T (*Temp*) e T<sub>d</sub> (*Temp.Orvalho*), utilizando o sinal (-):

<i>&gt; Temp - Temp.Orvalho</i>		
	1ºsemestre	2ºsemestre
1ºmês	2.4	0.7
2ºmês	1.9	1.2
3ºmês	2.1	2.1
4ºmês	2.1	1.0
5ºmês	2.0	1.2
6ºmês	2.3	0.7

A multiplicação das matrizes é feita utilizando o sinal (\*):

<i>&gt; Temp*Temp.Orvalho</i>		
	1ºsemestre	2ºsemestre
1ºmês	374.92	98.88
2ºmês	335.82	155.89
3ºmês	313.96	219.42
4ºmês	289.60	285.36
5ºmês	242.36	288.64
6ºmês	139.10	374.30

A divisão das matrizes é feita utilizando o sinal (/):

<i>&gt; Temp/Temp.Orvalho</i>		
	1ºsemestre	2ºsemestre
1ºmês	1.131868	1.072917
2ºmês	1.109195	1.100840

3ºmês	1.125749	1.152174
4ºmês	1.131250	1.060976
5ºmês	1.136986	1.073171
6ºmês	1.214953	1.036842

Outro objeto do R que é normalmente usado para guardar tabelas de dados é o *Data Frames*, que é semelhante a uma matriz, porém, nas suas colunas podem ser inseridos nomes e podem conter dados de tipos diferentes. Exemplo:

```
> Dados = data.frame(Direçãodovento = c(80, 80, 110, 100, 90, 80),
+ velocidadevento = c(3, 3, 2, 7, 7, 9),
+ pressão = c(115, 115, 119, 161, 162, 162),
+ temperatura = c(18, 18, 18.4, 17.1, 17.6, 17.6),
+ umidaderelativa = c(100, 100, 99, 99, 99, 99),
+ condição = c("nevoeiro", "nevoeiro", "nevoeiro", "nevoeiro", "nevoeiro", "nevoeiro"))
> Dados
  Direçãodovento velocidadevento pressão temperatura umidaderelativa condição
1             80                3      115        18.0             100   nevoeiro
2             80                3      115        18.0             100   nevoeiro
3            110                2      119        18.4              99   nevoeiro
4            100                7      161        17.1              99   nevoeiro
5             90                7      162        17.6              99   nevoeiro
6             80                9      162        17.6              99   nevoeiro
```

A Tabela C.1 sumaria os principais comandos na linguagem R, extraído da publicação produzida pela Faculdade de Economia da Universidade do Porto, referente à Programação em R, em português. Essa publicação pode ser solicitada por e-mail a Luís Torgo (ltorgo@liacc.up.pt). Informações adicionais sobre os comandos e aplicações do R também pode ser obtidas em Verzani (2002), Maindonald (2004) e na *website* <http://cran.r-project.org>. Vale ressaltar que essa página é atualizada regularmente por vários usuários da linguagem.

Tabela C.1 - Principais comandos básicos utilizados na Programação R.

<b>Comandos Básicos</b>	
<i>Sair do programa</i>	<i>q ( )</i>
<i>Atribuir valor a um objeto</i>	<i>=</i>
<i>Carregar um pacote estatístico</i>	<i>Require (nome do pacote estatístico)</i>
<i>Obter ajuda para um comando</i>	<i>help(comando)</i>
<i>Obter ajuda em um pacote estatístico</i>	<i>help(nome do pacote estatístico)</i>
<i>Ls</i>	<i>Lista todos os objetos disponíveis</i>
<i>Rm</i>	<i>Remove um objeto</i>
<i>dim(x, y)</i>	<i>x é a dimensão da linha e y dimensão da coluna de uma matriz</i>
<i>Obter um comprimento de um vetor</i>	<i>length ( nome do vetor)</i>
<i>rep(x,n)</i>	<i>Repete o vetor x, n vezes</i>
<i>Rbind (a1, b1, c1)</i>	<i>Cola linha em uma matriz</i>
<i>cbind (a1, b1, c1)</i>	<i>Cola coluna em uma matriz</i>
<i>if (condition1) ... else ...</i>	<i>condicional</i>
<i>%/%, ^, %% , sqrt()</i>	<i>Inteiro de divisão, potenciação e raiz quadrada</i>
<i>max ( )</i>	<i>Maior variável</i>
<i>min ( )</i>	<i>Menor variável</i>
<i>Mean ( )</i>	<i>Média das variáveis</i>
<i>median ( )</i>	<i>Mediana das variáveis</i>
<i>sum ( )</i>	<i>Somatória das variáveis</i>
<i>summary(data.frame)</i>	<i>Estatísticas resumo</i>
<i>Rank(), sort()</i>	<i>posição e ordenação</i>
<i>ave(x1,y1)</i>	<i>Médias de x1 para cada grupo do fator y1</i>
<i>by()</i>	<i>Aplica uma função a dados para cada nível de um fator</i>
<i>apply(x1,n1,function1)</i>	<i>Aplica a função1 à linhas (n1=1) ou colunas (n1=2) de x1</i>
<i>tapply(list1,function1)</i>	<i>Aplica function1 a cada elemento de list1</i>
<i>Table()</i>	<i>Faz uma tabela</i>
<i>tabulate()</i>	<i>Tabulação de um vetor</i>
<i>t.test()</i>	<i>t-test</i>
<i>aov(), anova(), lm(), glm()</i>	<i>Modelos lineares, regressão e análise de variância</i>
<i>Chisq.test(matrix1)</i>	<i>Teste chi-quadrado a uma matriz matrix1</i>
<i>cor(x1)</i>	<i>Correlação</i>
<i>cor.test(x1,x2)</i>	<i>Testa correlação</i>
<i>binom.test()</i>	<i>Teste do sinal</i>
<i>fisher.test()</i>	<i>Teste exato de Fisher</i>
<i>friedman.test()</i>	<i>Teste de Friedman</i>

Fonte: Adaptado de Torgo (2006).

## APENDICE D

### NOÇÕES BÁSICAS SOBRE O SISTEMA SAS E SUA UTILIZAÇÃO

O SAS é um *software* integrado para análise de dados que proporcionam entre outras facilidades, recuperação de dados, gerenciamento de arquivos, análise estatística, acesso à banco de dados, geração de gráficos e relatórios, além de possuir versões em diversos idiomas e para cada perfil de usuário, sendo este sistema utilizado por 96 das 100 maiores empresas do mundo e com 40.000 instalações distribuídas em 110 países (CENAPAD/UNICAMP, 2007).

A SAS Business Intelligence Software and Predictive Analytics (<http://www.sas.com>) é a maior empresa de *software* de capital privado do mundo e disponibiliza informações relacionadas (como por exemplo, tutoriais sobre o *software* SAS, e outros), treinamento, produtos e serviços, além de suporte para dúvidas e consultas. Em 2006, a empresa lançou o “SAS Fellowship Program” (<http://www.sas.com/govedu/edu/programs/studentresources.html>), um Programa para os EUA, América Latina, Canadá e Caribe, que oferece ao estudante a oportunidade de acesso livre ao software SAS por seis meses, para fins acadêmicos ou de pesquisa. Para se candidatar ao Programa o candidato deve acessar a *website* da SAS ou enviar e-mail para [studentprograms@sas.com](mailto:studentprograms@sas.com). Uma outra facilidade ao uso desse Programa é a *website* da comunidade SAS (<http://br.groups.yahoo.com/group/SASbrasil/>), onde se troca informações, material e discutem-se assuntos relacionados. Também, um guia sobre vários aspectos do método e manipulação de dados de Análise de Sobrevivência com a utilização do SAS é o livro de Allison (1995), incluindo a listagem de programas estatísticos escritos na linguagem SAS e que resolvem os exercícios propostos no livro.

A seguir, apresentam-se sucintamente os aspectos básicos sobre o Sistema SAS e sua utilização, com ênfase sobre a Função Gama Generalizada,

utilizada nesse trabalho. Informações mais detalhadas são dadas nos tutoriais do SAS e no material disponível na *web*, citados anteriormente.

## **D.1 – UTILIZAÇÃO DO PROGRAMA SAS**

No *windows*, para acessar o SAS basta abrir a janela que geralmente fica na área de trabalho e executá-lo. O SAS possui um módulo específico para cada aplicação: SAS/STAT (módulo estatístico), SAS GRAPH (módulo gráfico), SAS/OR (módulo de análise e pesquisa operacional), SAS/QC (módulo para análise de controle de qualidade), SAS/ETC (Módulo de econometria), SAS/ASSIST (módulo de programação interativa e amigável), SAS/IML (módulo para análise e operações de matrizes), SAS/Access Módulo para acesso a banco de dados), SAS/CONNECT (módulo para conexão entre ambientes operacionais heterogêneos), SAS/AF (módulo para desenvolvimento de aplicações) e SAS/FSP (módulo para facilitar o acesso a arquivos com programação de telas), sendo o SAS/BASE obrigatório para toda instalação.

Os comandos utilizados no SAS são executados em estruturas, como por exemplo, DATA STEP (utilizada para criar um ou vários arquivos SAS), PROC STEP (utilizado para analisar e manipular os dados contidos num arquivo SAS), entre outras. As Tabelas D1 e D2 sumariam alguns tipos de comandos utilizados nas estruturas DATA Step e PROC STEP, respectivamente.

Tabela D1 - Estrutura DATA Step.

<b>ESTRUTURA DATA Step</b>	
<b>COMANDOS</b>	<b>ATRIBUIÇÕES</b>
DATA (Definição)	Sempre inicia um DATA SET e abre um novo arquivo DATA SET, vazio, em uma biblioteca SAS. <i>DATA &lt;bib&gt;.&lt;nome do arquivo&gt;;</i>
INFILE (Execução)	Identifica e abre um arquivo externo de dados (somente para leitura), ou seja, abre um arquivo que não está no formato SAS. Deve ser posicionado antes do comando INPUT e depois do comando DATA. <i>INFILE "&lt;arquivo de dados&gt;"[opções];</i> Opções: LRECL = (tamanho do registro); PAD (completa de branco os registros); MISSOVER (previne o SAS de ler um novo registro para completar os dados de uma variável); DLM = (especifica um delimitador entre os campos).
SET (Execução)	Abre um arquivo de dados DATA SET (arquivo que está no formato SAS), somente para leitura e que já existe na biblioteca do SAS. <i>SET &lt;bib&gt;.&lt;nome do arquivo&gt;;</i>
INPUT (Execução)	Descreve ao SAS como estão armazenados e organizados os dados em arquivos textos, que foram especificados pelo comando INFILE, ou pelos dados digitados após o comando CARDS. Arquivos abertos pelo comando INFILE sempre necessitarão de um comando INPUT para ler seu conteúdo. <i>INPUT &lt;variável&gt;[&lt;\$&gt;];</i> Onde <i>variável</i> indica o nome da variável que será associada a uma coluna de dados e \$ significa que a variável só irá conter dados alfanuméricos;
CARDS (Execução)	Indica ao SAS que os dados utilizados em um DATA Step estão digitados no programa, logo após o comando CARDS. Este comando é o último dentro de um DATA Step, antes do comando RUN; Utilizado junto com o comando INPUT, quando apenas se deseja efetuar alguns testes de execução do DATA Step, não necessitando de abrir um arquivo de leitura com o comando INFILE. <i>CARDS;</i>
LIBNAME	<i>LIBNAME &lt;biblioteca&gt;"&lt;diretório&gt;;</i> Onde biblioteca é uma palavra de no mínimo 1 e no máximo 8 caracteres, que identifica o primeiro nível em um nome de um arquivo e diretório é o nome de um diretório que já exista no seu ambiente. Comando opcional de configuração que define uma ou mais área de trabalho e armazenamento (diretórios), para os arquivos SAS (Dataset's), normalmente é colocado antes do comando DATA.
Atribuição (=)	<i>&lt;variável&gt; = &lt;expressão&gt;;</i> Caractere utilizado para criar novas variáveis e editar as já existentes em um Data Step.

FONTE: Adaptado de Cenapad/Unicamp (2007).

Tabela D.2 - Estrutura PROC Step.

<b>ESTRUTURA PROC Step</b>	
<b>COMANDOS</b>	<b>ATRIBUIÇÕES</b>
PROC	<p>Todos procedimentos iniciam-se com a palavra PROC seguida do nome específico do procedimento ou rotina desejada.</p> <p><i>PROC &lt; rotina &gt; [OPÇÕES];</i>  <i>[comando];</i>  <i>[comando];</i></p> <p><i>RUN;</i></p> <p><i>OPÇÕES:</i> (parâmetros opcionais de configuração para execução do procedimento); <i>comando</i> (comando de detalhadamente na execução do procedimento). <i>rotina</i> (palavra-chave que identifica o procedimento a ser executado);</p>
VAR	<p>Especifica as variáveis do arquivo SAS que serão processados pelo procedimento.</p> <p><i>VAR &lt; variável1 &gt; &lt; variável2 &gt; ...&lt;variáveln&gt;;</i></p>
BY	<p>Especifica uma ou mais variáveis que permitirão o processamento em grupos.</p> <p><i>BY &lt; variável1 &gt; &lt; variável2 &gt; ...&lt;variáveln&gt;;</i></p>
TITLE	<p>Especifica de 1 até 10 títulos para o relatório a ser gerado pelo procedimento.</p> <p><i>TITLE "&lt;cabeçalho&gt;;"</i>  <i>TITLE1 "&lt;cabeçalho&gt;;"</i>            ...  <i>TITLE10 "&lt;cabeçalho&gt;;"</i></p>
FOOTNOTE	<p>Especifica de 1 até 10 rodapés para o relatório a ser gerado pelo procedimento.</p> <p><i>FOOTNOTE "&lt;rodapé&gt;;"</i>  <i>FOOTNOTE1 "&lt;rodapé&gt;;"</i>            .....  <i>FOOTNOTE10 "&lt;rodapé&gt;;"</i></p>
PRINT	<p>Procedimento do SAS que imprime na saída padrão os dados contidos nos arquivos SAS.</p> <p><i>PROC PRINT [opções];</i>  <i>VAR &lt;lista de variáveis&gt;;</i>  <i>BY &lt; lista de variáveis&gt;;</i>  <i>SUM &lt;lista de variáveis&gt;;</i></p> <p>Opções: DATA (nome do arquivo SAS); DOUBLE (espacejamento duplo); NOOBS (suprime a coluna com o número de cada observação do arquivo); UNIFORM (formata todas as páginas uniformemente); LABEL (imprime os <i>label's</i> das variáveis); SPLIT (especifica um caractere do <i>label</i> para indicar a quebra da descrição). VAR especifica uma lista de variáveis que serão impressas, BY especifica a quebra ou agrupamento do relatório, por uma ou mais variáveis e SUM especifica uma ou mais variáveis numéricas que serão totalizadas.</p>
SORT	<p>Procedimento do SAS que reordena os registros dos arquivos por uma ou mais variáveis, em ordem ascendente ou descendente.</p> <p><i>PROC SORT [opções];</i>  <i>BY &lt;lista de variáveis&gt;;</i></p> <p>Opções: DATA = Nome do arquivo SAS para ser ordenado            OUT = Nome do arquivo SAS, aonde serão armazenados os dados ordenados. Se não for utilizada essa opção, a ordenação será feita em cima do arquivo original;            BY, especifica uma ou várias variáveis como chaves de ordenação.</p>

FONTE: Adaptado de Cenapad/Unicamp (2007).

Na estrutura DATA Step, os comando devem seguir uma ordem de colocação, pois a execução de um determinado comando pode depender da execução de um comando anterior. Na estrutura PROC os comando podem ser colocados em qualquer ordem dentro do comando PROC, visto que a execução de um determinado comando não depende da execução de um comando anterior. O Comando RUN é o comando de execução e dever ser utilizado no último procedimento do programa SAS.

Para dados serem analisados pelo procedimento PROC do SAS, estes devem estar armazenados em arquivos com estrutura DATA SET SAS. Os arquivos SAS podem ser temporários (biblioteca definida pelo SAS) ou permanentes (biblioteca definida pelo usuário). Todo comando SAS começa com uma palavra-chave, de identificação, e termina com um ponto e virgula (;).

Para exemplificar alguns comandos mostrados nas Tabelas D.1 e D2, segue um exemplo de aplicação que considera duas estações meteorológicas que registraram quatro temperaturas cada uma, referente a um determinado período, para calcular a soma e a média dessas temperaturas, faz-se:

```
LIBNAME T "C:\";  
DATA T. TEMP;  
    INFILE "C:ESTAÇÃO\TEMP.DAT";  
    INPUT (T1-T4)(4.);  
    TEMPSOMA = SUM (OF T1 – T4);  
    TEMPMED = MEAN(OF T1 – T4);  
RUN;
```

Arquivo: TEMP.DAT

1 2 3 4 5 6 7 8 9 10 11 12 13 14  
15 16

1	9	.	9	2	0	.	2	1	8	.	3	2	2	.	4
2	1	.	2	2	3	.	1	2	1	.	3	2	4	.	5

```
PROC PRINT DATA = T.TEMP DOUBLE NOOBS;  
RUN;
```

The SAS System

T1	T2	T3	T4	TEMPSOMA	TEMPMED
19.9	20.2	18.3	22.4	80.8	20.2000
21.2	23.1	21.3	24.5	90.1	22.5250

Logo, para a primeira estação meteorológica considerada, tem-se temperatura média de 20,2 °C, enquanto para a segunda estação meteorológica, tem-se temperatura média de 22,5°C.

## D.2 – SUBROTINA DO SAS: PROC LIFEREG

No SAS, o procedimento PROC LIFEREG ajusta modelos de tempo de regressão paramétricos a dados de tempo de falha, cuja censura pode ser à direita, à esquerda ou intervalar. Essa subrotina estima os parâmetros pela máxima verossimilhança usando o algoritmo de Newton Raphson, e estima os erros padrões dos parâmetros estimados do inverso da matriz de informação

observada (TUTORIAL SAS, 2008). Allison (1995), alerta que uma diferente síntese do Modelo é requerido se há diversos tipos de Censura, por exemplo:

Tabela D.3 – Tipos de Censuras

Observação	Inferior (lower)	Superior (upper)	Status
1	3,9	3,9	Não Censurado
2	7,2	...	Censura á direita
3	4,1	5,6	Censura Intervalar
4	...	2,0	Censura à esquerda
5	0	5,8	Excluído
6	3,2	1,9	Excluído

FONTE: Allison (1995).

Então a estimação para o Modelo de Weibull, fica da seguinte maneira:

Data;

Set recidlft;

/\* para casos não censurados: \*/

se arrest=1 and tempos ne.then do;

upper = tempos;

lower = tempos;

end;

/\* para casos de censura à esquerda: \*/

se arrest=1 and tempos .then do;

upper = 52;

lower = .;

end;

/\* para casos de censura à direita: \*/

se arrest=0 and then do;

upper = .;

```
lower = 52;
```

```
end;
```

```
proc lifereg;
```

```
model (lower, upper) = covariáveis / dist = WEIBULL;
```

```
run;
```

Outra forma de representar dados não censurados ou com censura à direita é considerar o valor para uma observação completa igual a 1, e escrever da seguinte maneira:

```
proc lifereg;
```

```
model tempo*cens(0)= /distribution = gamma;
```

```
run;
```

## APÊNDICE E

### ALGUNS PROGRAMAS UTILIZADOS, ESCRITOS EM R

#### E.1 – CONSTRUÇÃO DO BOX PLOT PARA A VARIÁVEL “DIREÇÃO DO VENTO”

#### E.1 – CONSTRUÇÃO DO BOX PLOT PARA A VARIÁVEL “DIREÇÃO DO VENTO”

**\* Leitura dos dados, arquivo em formato .csv ou txt.**

```
> dados = read.table("DV.csv", head=T, sep=":", dec=",")
```

```
> dados
```

**\* Inclusão dos dados, não reproduzidos na íntegra por questões de espaço.**

```
Dvmeia
```

```
1 220
```

```
2 130
```

```
.
```

```
. até
```

```
.
```

```
83 280
```

```
84 50
```

```
85 0
```

```
86 70
```

**\* Conversão dos dados para linguagem R**

```
> teste = c(dados)
```

```
> fix(teste)
```

**\* Dados gerados em R**

```
> Dvmeia = c(220, 130, 100, 100, 100, 70, 90, 70, 100, 100, 210, 90, 200, 70,  
+ 90, 100, 100, 100, 340, 90, 90, 80, 100, 110, 90, 90, 80, 100, 80, etc)
```

**\* Geracao de resultados para direção do vento menor que 1 hora**

```
> dados = read.table("DVuma.csv", head=T, sep=":", dec=",")
```

```
> fix(teste)
```

```
> dados = read.table("DVuma.csv", head=T, sep=":", dec=",")
```

```
> dados
```

```
  Dvuma  
1  100  
2   80  
.  
. até  
.  
194  90  
195  80  
196  70
```

\*

```
> fix(teste)
```

```
> teste = c(dados)
```

```
> fix(teste)
```

```
> Dvuma = c(100, 80, 90, 260, 100, 100, 150, 90, 110, 0, 90, 0, 110, 70, 10,  
140, 90, 100, 60, etc).
```

\*

```
> dados = read.table("DVduas.csv", head=T, sep=":", dec=",")
```

```
> teste = c(dados)
```

```
> fix(teste)
```

```
> Dvduas = c(120, 120, 80, 100, 80, 110, 90, 90, 90, 110, 90, 120, 80, 180, 100,  
130, 90, 50, 80, etc)
```

\*

```
> dados = read.table("DVtrês.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvtrês = c(100, 130, 90, 120, 80, 90, 100,90, 80, 120, etc)
```

\*

```
> dados = read.table("DVquatro.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvquatro = c(0, 120, 80, 90, 110, 90, 220,90, 100, 100, 90,etc)
```

\*

```
> dados = read.table("DVcinco.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvcinco = c(100, 90, 350, 70, 90, 100, 80, 90, 100, 340, 100, 90, 90, 90, 140,
80, 100, 30, 100, etc)
```

\*

```
> dados = read.table("DVseis.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvseis = c(120, 100, 70, 130, 100, 320, 90, 70, 100, 80, 60, 80, 100, 100, 80,
+110, 90, 100, etc)
```

```
> dados = read.table("DVsete.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvsete = c(90, 80, 100, 90, 90, 190, 60, 80, 90, 100, 100, 90, 70, 330, 80, 60,
+ 360, 70, 60, etc)
```

\*

```
> dados = read.table("DVsete.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvsete = c(90, 80, 100, 90, 90, 190, 60, 80, 90, 100, 100, 90, 70, 330, 80, 60,
+ 360, 70, 60,100)
```

\*

```
> dados = read.table("DVoito.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvoito = c(100, 100, 90, 90, 70, 130, 60,+ 80, 60, 70, 110, 50, 50, 80, 110,
+ 100, 100)
```

\*

```
> dados = read.table("DVnove.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvnove = c(80, 100, 90, 80, 90, 90, 100,100, 80, 0, 90, 80, 90, 110, 60, 140)
```

\*

```
> dados = read.table("DVdez.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvdez = c(120, 100, 80, 110, 110, 0, 100)
```

\*

```
> dados = read.table("DVonze.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvonze = c(60, 70, 90, 70)
```

\*

```

> dados = read.table("DVdoze.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvdoze = c(80, 140, 150, 80, 60)
> teste = c(dados)
> fix(teste)
> Dvdoze = c(80, 140, 150, 80, 60)

```

\*

```

> dados = read.table("DVtreze.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvtreze = 90

```

\*

```

> dados = read.table("DVcatorze.csv", head=T, sep=":", dec=",")
> teste = c(dados)
> fix(teste)
> Dvcatorze = 120

```

#### \* **Geração do boxplot**

```

> boxplot(Dvmeia, Dvuma, Dvduas, Dvtrês, Dvquatro, Dvcinco, Dvseis, Dvsete,
Dvoito, Dvnove, Dvdez, Dvonce, Dvdoze, Dvtreze, Dvcatorze, xlab = 'TEMPO
DE DURACAO (horas)', + ylab = 'GRAUS').

```

## **E.2 – CONSTRUÇÃO DAS FIGURAS DO ESTIMADOR DE KAPLAN-MEIER E TESTE (valor p)**

```

Estimadorkp = survfit(Surv(Tempos, Cens)~V1)
Summary(Estimadorkp)
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
Survdiff(Surv(Tempos, Cens)~V1, rho=1)

```

```
Plot(Estimadorkp, lty = c(1, 4), mark.time= F, xlab = "Tempo até a dissipação  
do nevoeiro em horas", ylab = "S(t)")
```

```
Estimadorkp = survfit(Surv(Tempos, Cens)~V2)
```

```
Summary(Estimadorkp)
```

```
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
```

```
Survdiff(Surv(Tempos, Cens)~V1, rho=1)
```

```
Plot(Estimadorkp, lty = c(1, 4), mark.time= F, xlab = "Tempo até a dissipação  
do nevoeiro em horas", ylab = "S(t)")
```

```
Estimadorkp = survfit(Surv(Tempos, Cens)~V3)
```

```
Summary(Estimadorkp)
```

```
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
```

```
Survdiff(Surv(Tempos, Cens)~V1, rho=1)
```

```
Plot(Estimadorkp, lty = c(1, 4), mark.time= F, xlab = "Tempo até a dissipação  
do nevoeiro em horas", ylab = "S(t)")
```

```
Estimadorkp = survfit(Surv(Tempos, Cens)~V4)
```

```
Summary(Estimadorkp)
```

```
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
```

```
Survdiff(Surv(Tempos, Cens)~V1, rho=1)
```

```
Plot(Estimadorkp, lty = c(1, 4), mark.time= F, xlab = "Tempo até a dissipação  
do nevoeiro em horas", ylab = "S(t)")
```

```
Estimadorkp = survfit(Surv(Tempos, Cens)~V5)
```

```
Summary(Estimadorkp)
```

```
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
```

```
Survdiff(Surv(Tempos, Cens)~V1, rho=1)
```

```
Plot(Estimadorkp, lty = c(1, 4), mark.time= F, xlab = "Tempo até a dissipação  
do nevoeiro em horas", ylab = "S(t)")
```

```
Estimadorkp = survfit(Surv(Tempos, Cens)~V6)
```

```
Summary(Estimadorkp)
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
Survdiff(Surv(Tempos, Cens)~V1, rho=1)
Plot(Estimadorkp, lty = c(1, 4), mark .time= F, xlab = "Tempo até a dissipação
do nevoeiro em horas", ylab = "S(t)")
```

```
Estimadorkp = survfit(Surv(Tempos, Cens)~V7)
Summary(Estimadorkp)
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
Survdiff(Surv(Tempos, Cens)~V1, rho=1)
Plot(Estimadorkp, lty = c(1, 4), mark .time= F, xlab = "Tempo até a dissipação
do nevoeiro em horas", ylab = "S(t)")
```

```
Estimadorkp = survfit(Surv(Tempos, Cens)~V8)
Summary(Estimadorkp)
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
Survdiff(Surv(Tempos, Cens)~V1, rho=1)
Plot(Estimadorkp, lty = c(1, 4), mark .time= F, xlab = "Tempo até a dissipação
do nevoeiro em horas", ylab = "S(t)")
```

```
Estimadorkp = survfit(Surv(Tempos, Cens)~V9)
Summary(Estimadorkp)
Survdiff(Surv(Tempos, Cens)~V1, rho=0)
Survdiff(Surv(Tempos, Cens)~V1, rho=1)
Plot(Estimadorkp, lty = c(1, 4), mark .time= F, xlab = "Tempo até a dissipação
do nevoeiro em horas", ylab = "S(t)")
```

### **E.3 – ESTIMATIVAS DOS PARÂMETROS DO MODELO DE REGRESSÃO DE WEIBULL**

```
ajustw = survreg(Surv(Tempos, Cens)~V1+V2+V3+V4+V5+V7, dist = 'weibull')
```

```
ajustw
summary(ajustw)
```

#### **E.4 – CONSTRUÇÃO DOS GRÁFICOS DOS RESÍDUOS**

```
res = (Tempos*exp(-xb))^1.72
ekm = survfit(Surv(res, Cens)~1)
summary(ekm)
par(mfrow=c(1,2))
plot(ekm, conf.int=F, lty=c(1,1))
ekm = survfit(Surv(res, Cens)~1, type=c("kaplan-meier"))
par(mfrow=c(1,2))

plot(ekm, conf.int=F, lty=c(1,1), xlab="resíduos", ylab="S(e) estimada")
res = sort(res)
exp1 = exp(-res)
lines(res, exp1, lty=3)
legend(2, 0.8, lty=c(1,3), c("Kaplan-Meier", "Exponencial(1)"), lwd=1, bty=
+"n", cex=0.7)
st= ekm$time
t=ekm$time
sexp1=exp(-t)

plot(st, sexp1, xlab="S(e) - Kaplan-Meier", ylab = "S(e) - Exponencial(1)",
pch=16)
ekm1 = survfit(Surv(res, Cens)~1)
t = ekm1$time
st=ekm1$surv
sexp = exp(-t)
par(mfrow=c(1,2))
```

## APÊNDICE F

### PROGRAMA ESCRITO EM LINGUAGEM SAS

```
data nevoeiro;
input id Tempos Cens V1 V2 V3 V4 V5 V6 V7 V8 V9;
V12 = V1*V2;
V13 = V1*V3;
V14 = V1*V4;
V15 = V1*V5;
V17 = V1*V7;
V23 = V2*V3;
V24 = V2*V4;
V25 = V2*V5;
V27 = V2*V7;
V34 = V3*V4;
V35 = V3*V5;
V37 = V3*V7;
V45 = V4*V5;
V57 = V5*V7;
cards;
```

#### **Ajuste para o modelo completo**

```
proc lifereg;
model tempo*cens(0)= /distribution = gamma;
run;
```

#### **Ajuste para cada covariável (PASSO 1)**

```
proc lifereg;
model tempo*cens(0)= V1 /distribution = gamma;
run;
```

```
proc lifereg;
model tempo*cens(0)= V2 /distribution = gamma;
run;
```

.

. até

.

.

```
proc lifereg;
model tempo*cens(0)= V8 /distribution = gamma;
run;
```

### **Ajuste conjunto das covariáveis (PASSO 2)**

```
proc lifereg;
model tempo*cens(0)= V1 V2 V3 V4 V5 V6 V7 V8 /distribution = gamma;
run;
```

```
proc lifereg;
model tempo*cens(0)= V2 V3 V4 V5 V6 V7 V8 /distribution = gamma;
run;
```

.

. até

.

```
proc lifereg;
model tempo*cens(0)= V1 V2 V3 V4 V5 V6 V7 /distribution = gamma;
run;
```

### **Ajuste conjunto das covariáveis (PASSO 3)**

```
proc lifereg;  
model tempo*cens(0)= V1 V2 V3 V4 V5 V7 /distribution = gamma;  
run;
```

```
proc lifereg;  
model tempo*cens(0)= V2 V3 V4 V5 V6 V7 /distribution = gamma;  
run;
```

```
proc lifereg;  
model tempo*cens(0)= V2 V3 V4 V5 V7 V8 /distribution = gamma;  
run;
```

### **Ajuste conjunto das covariáveis (PASSO 4)**

Nenhuma covariável foi excluída no passo 1

### **Ajuste conjunto das covariáveis (PASSO 5)**

```
proc lifereg;  
model tempo*cens(0)= V1 V2 V3 V4 V5 V7 /distribution = gamma;  
run;
```

```
proc lifereg;  
model tempo*cens(0)= V2 V3 V4 V5 V7 /distribution = gamma;  
run;
```

```
.  
.  
.
```

```
proc lifereg;  
model tempo*cens(0)= V1 V2 V3 V4 V5 /distribution = gamma;  
run;
```

### **Ajuste conjunto das covariáveis (PASSO 6)**

```
proc lifereg;  
model tempo*cens(0)= V1 V2 V3 V4 V5 V7 /distribution = gamma;  
run;
```

```
proc lifereg;  
model tempo*cens(0)= V1 V2 V3 V4 V5 V7 V12 /distribution = gamma;  
run;
```

```
.  
.  
.
```

```
proc lifereg;  
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V17 /distribution = gamma;  
run;
```

```
proc lifereg;  
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V23 /distribution = gamma;  
run;
```

```
.  
.  
.
```

```
proc lifereg;  
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V27 /distribution = gamma;  
run;
```

```
proc lifereg;  
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V34 /distribution = gamma;  
run;
```

```
.  
.  
.
```

```
proc lifereg;  
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V37 /distribution = gamma;
```

```

run;
proc lifereg;
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V45 /distribution = gamma;
run;
.
.
proc lifereg;
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V47 /distribution = gamma;
run;
proc lifereg;
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V57 /distribution = gamma;
run;

```

### **Ajuste conjunto das covariáveis (PASSO 7)**

#### **Ajuste conjunto das covariáveis (Modelo Final)**

```

proc lifereg;
model tempo*cens(0) = V1 V2 V3 V4 V5 V7 V17 V37 /distribution = gamma;
run;

```



## APÊNDICE G

### VERIFICAÇÃO DO MODELO

Tabela G.1 – Eventos de nevoeiros ocorridos em 2004 no Aeroporto Internacional de São Paulo.

ID	Mês ocorrência	Horário de início (h:min)	Direção do vento (graus)	Vel do vento ( $\text{ms}^{-1}$ )	Visibilidade ( m)	$T_d$ ( $^{\circ}$ C)	Tempo de duração (h)
	V1	V2	V3	V4	V5	V7	
1	09/01/04	5:25	80	3	400	18,0	0,92
2	11/01/04	4:40	00	0	800	15,8	3,33
3	14/02/04	6:35	100	7	800	17,0	1,00
4	17/02/04	3:45	80	2	900	17,3	0,75
5	27/02/04	6:00	140	1	500	13,7	2,67
6	29/02/04	6:45	80	6	400	13,8	1,58
7	08/03/04	4:28	80	2	600	17,0	3,87
8	19/03/04	6:00	90	8	800	16,8	1,00
9	25/03/04	7:00	00	0	800	12,3	1,58
10	30/03/04	2:15	60	9	500	16,8	5,75
11	31/03/04	2:00	110	2	400	16,6	1,00
12	03/04/04	3:30	80	4	800	18,1	3,17
13	05/04/04	5:35	90	2	600	17,0	2,58
14	07/04/04	3:40	80	4	500	17,0	3,83
15	08/04/04	4:40	80	3	800	18,5	4,33
16	11/04/04	7:10	110	2	500	17,0	0,83
17	30/04/04	5:00	100	9	800	14,2	3,33
18	02/05/04	3:40	00	0	800	14,8	3,33
19	13/05/04	7:00	40	2	900	15,8	1,50
20	17/05/04	4:10	110	1	800	7,7	2,08
21	17/05/04	21:10	130	2	400	12,6	2,25
22	18/05/04	0:00	00	0	300	13,1	0,42
23	30/05/04	8:00	70	6	500	13,5	1,00

24	06/06/04	6:20	00	0	800	5,8	1,67
25	08/06/04	4:00	100	6	600	9,0	3,67
26	15/06/04	20:20	90	6	800	14,0	13,25
27	17/06/04	0:00	90	6	100	12,3	9,38
28	20/06/04	3:00	110	9	900	12,8	2,42
29	21/06/04	5:50	100	6	900	11,5	3,17
30	24/06/04	5:40	100	7	500	12,2	3,67
31	29/06/04	1:00	100	3	300	14,5	7,67
32	01/07/04	5:48	100	2	800	13,2	2,95
33	04/07/04	8:25	280	3	700	15,9	0,75
34	07/07/04	3:35	80	5	900	15,0	5,38
35	13/07/04	7:00	80	9	900	12,5	1,00
36	14/07/04	4:10	90	3	600	13,7	2,33
37	26/07/04	4:30	110	3	900	8,7	2,17
38	03/08/04	7:42	50	4	400	14,1	0,80
39	25/08/04	5:43	80	5	800	14,0	1,95
40	02/09/04	6:00	80	6	800	14,5	2,00
41	06/09/04	4:10	100	4	400	13,8	4,08
42	07/09/04	6:00	80	9	800	15,3	2,00
43	21/09/04	7:00	0	0	800	16,7	0,58
44	29/09/04	3:35	90	3	900	16,0	1,42
45	16/10/04	3:00	80	5	800	18,2	1,67
46	27/10/04	3:00	70	4	800	16,5	0,50
47	03/11/04	4:40	90	5	500	17,2	2,00

ID – Identificador do evento.

Tabela G.2 - Dicotomização dos dados referente a eventos de nevoeiros ocorridos em 2004.

ID	V1	V2	V3	V4	V5	V7	Tempo de duração (h)
1	1	1	0	0	0	1	0,92
2	1	1	0	0	1	0	3,33
3	1	1	0	0	1	1	1,00
4	1	1	0	0	1	1	0,75
5	1	1	0	0	0	0	2,67
6	1	1	0	0	0	0	1,58
7	1	1	0	0	1	1	3,87
8	1	1	0	0	1	1	1,00
9	1	1	0	0	1	0	1,58
10	1	0	0	1	0	1	5,75
11	1	0	0	0	0	1	1,00
12	1	1	0	0	1	1	3,17
13	1	1	0	0	1	1	2,58
14	1	1	0	0	0	1	3,83
15	1	1	0	0	1	1	4,33
16	1	1	0	0	0	1	0,83
17	1	1	0	1	1	0	3,33
18	0	1	0	0	1	0	3,33
19	0	1	0	0	1	0	1,50
20	0	1	0	0	1	0	2,08
21	0	0	0	0	0	0	2,25
22	0	0	0	0	0	0	0,42
23	0	1	0	0	0	0	1,00
24	0	1	0	0	1	0	1,67
25	0	1	0	0	1	0	3,67
26	0	0	0	0	1	0	13,25
27	0	0	0	0	0	0	9,38
28	0	1	0	1	1	0	2,42

29	0	1	0	0	1	0	3,17
30	0	1	0	0	0	0	3,67
31	0	0	0	0	0	0	7,67
32	0	1	0	0	1	0	2,95
33	0	1	1	0	1	0	0,75
34	0	1	0	0	1	0	5,38
35	0	1	0	0	1	0	1,00
36	0	1	0	0	1	0	2,33
37	0	1	0	0	1	0	2,17
38	0	1	0	0	0	0	0,80
39	0	1	0	0	1	0	1,95
40	1	1	0	0	1	0	2,00
41	1	1	0	0	0	0	4,08
42	1	1	0	1	1	0	2,00
43	1	1	0	0	1	1	0,58
44	1	1	0	0	1	0	1,42
45	1	1	0	0	1	1	1,67
46	1	1	0	0	1	1	0,50
47	1	1	0	0	0	1	2,00

ID – identificador do evento.

Tabela G.3 - Probabilidades dos nevoeiros ocorridos no Aeroporto Internacional de São Paulo em 2004 durarem além de um tempo t, conhecida as covariáveis selecionadas no Modelo Final.

Combinação V1 + V2 + V3 + V4 + V5 + V7 + V17 + V37														
ID	Tempos (t)													
	0,5	1	2	3	4	5	6	7	8	9	10	11	12	13
1	94%	80%	48%	23%	9%	3%	1%	0%	0%	0%	0%	0%	0%	0%
2	95%	84%	56%	31%	14%	6%	2%	1%	0%	0%	0%	0%	0%	0%
3	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
4	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
5	96%	86%	61%	37%	20%	9%	4%	1%	0%	0%	0%	0%	0%	0%
6	96%	86%	61%	37%	20%	9%	4%	1%	0%	0%	0%	0%	0%	0%
7	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
8	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
9	95%	84%	56%	31%	14%	6%	2%	1%	0%	0%	0%	0%	0%	0%
10	98%	93%	78%	61%	44%	30%	19%	12%	7%	4%	2%	1%	0%	0%
11	98%	94%	82%	66%	51%	37%	26%	17%	11%	6%	4%	2%	1%	1%
12	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
13	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
14	94%	80%	48%	23%	9%	3%	1%	0%	0%	0%	0%	0%	0%	0%
15	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
16	94%	80%	48%	23%	9%	3%	1%	0%	0%	0%	0%	0%	0%	0%
17	94%	81%	49%	24%	10%	3%	1%	0%	0%	0%	0%	0%	0%	0%
18	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
19	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
20	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
21	99%	97%	89%	80%	69%	58%	48%	38%	30%	22%	17%	12%	9%	6%
22	99%	97%	89%	80%	69%	58%	48%	38%	30%	22%	17%	12%	9%	6%
23	96%	89%	67%	45%	27%	15%	7%	3%	1%	0%	0%	0%	0%	0%
24	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
25	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
26	99%	96%	87%	76%	64%	52%	41%	31%	23%	17%	12%	8%	5%	3%
27	99%	97%	89%	80%	69%	58%	48%	38%	30%	22%	17%	12%	9%	6%
28	95%	84%	56%	31%	15%	6%	2%	1%	0%	0%	0%	0%	0%	0%
29	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%

30	96%	89%	67%	45%	27%	15%	7%	3%	1%	0%	0%	0%	0%	0%
31	99%	97%	89%	80%	69%	58%	48%	38%	30%	22%	17%	12%	9%	6%
32	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
33	91%	74%	37%	14%	4%	1%	0%	0%	0%	0%	0%	0%	0%	0%
34	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
35	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
36	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
37	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
38	96%	89%	67%	45%	27%	15%	7%	3%	1%	0%	0%	0%	0%	0%
39	96%	87%	62%	38%	21%	10%	4%	2%	1%	0%	0%	0%	0%	0%
40	95%	84%	56%	31%	14%	6%	2%	1%	0%	0%	0%	0%	0%	0%
41	96%	86%	61%	37%	20%	9%	4%	1%	0%	0%	0%	0%	0%	0%
42	94%	81%	49%	24%	10%	3%	1%	0%	0%	0%	0%	0%	0%	0%
43	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
44	95%	84%	56%	31%	14%	6%	2%	1%	0%	0%	0%	0%	0%	0%
45	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
46	92%	77%	42%	17%	6%	1%	0%	0%	0%	0%	0%	0%	0%	0%
47	94%	80%	48%	23%	9%	3%	1%	0%	0%	0%	0%	0%	0%	0%

Tabela G.4– Verificação do modelo.

<b>ID</b>	<b>TEMPO DE DURAÇÃO</b>	<b>PROBABILIDADE (%)</b>	<b>NÍVEL DE ACERTO</b>
1	0,92	94 (além de 0,5h)	OK
2	3,33	56 (além de 3h)	OK
3	1,00	92 (além de 0,5h)	OK
4	0,75	92 (além de 0,5h)	OK
5	2,67	61 (além de 2h)	OK
6	1,58	86 (além de 1h)	OK
7	3,87	17 (além de 3h)	<b>NÃO</b>
8	1,00	92 (além de 0,5h)	OK
9	1,58	83 (além de 1h)	OK
10	5,75	30 (além de 5h)	OK
11	1,00	98 (além de 1h)	OK
12	3,17	17 (além de 3h)	<b>NÃO</b>
13	2,58	41 (além de 2h)	OK
14	3,83	23 (além de 3h)	<b>NÃO</b>
15	4,33	6 (além de 4h)	<b>NÃO</b>
16	0,83	93 (além de 0,5h)	OK
17	3,33	24 (além de 3h)	<b>NÃO</b>
18	3,33	38 (além de 3h)	OK
19	1,50	87 (além de 1h)	OK
20	2,08	62 (além de 2h)	OK
21	2,25	89 (além de 2h)	OK
22	0,42	98 (além de 0,5h)	OK
23	1,00	96 (além de 0,5h)	OK
24	1,67	87 (além de 1h)	OK
25	3,67	38 (além de 3h)	OK
26	13,25	3 (além de 3h)	<b>NÃO</b>
27	9,38	23 (além de 3h)	<b>NÃO</b>

28	2,42	56 (além de 2h)	OK
29	3,17	38 (além de 3h)	OK
30	3,67	45 (além de 3h)	OK
31	7,67	38 (além de 7h)	OK
32	2,95	62 (além de 2h)	OK
33	0,75	91 (além de 0,5h)	OK
34	5,38	10 (além de 5h)	<b>NÃO</b>
35	1,00	96 (além de 0,5h)	OK
36	2,33	62 (além de 2h)	OK
37	2,17	62 (além de 2h)	OK
38	0,80	96 (além de 0,5h)	OK
39	1,95	87 (além de 1h)	OK
40	2,00	84 (além de 1h)	OK
41	4,08	20 (além de 4h)	<b>NÃO</b>
42	2,00	80 (além de 1h)	OK
43	0,58	92 (além de 0,5h)	OK
44	1,42	84 (além de 1h)	OK
45	1,67	77 (além de 1h)	OK
46	0,50	92 (além de 0,5)	OK
47	2,00	48 (além de 2h)	OK

OK – considerou acerto > 25%

NÃO – considerou erro < 25%

## **PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE**

### **Teses e Dissertações (TDI)**

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

### **Manuais Técnicos (MAN)**

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

### **Notas Técnico-Científicas (NTC)**

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programa de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

### **Relatórios de Pesquisa (RPQ)**

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

### **Propostas e Relatórios de Projetos (PRP)**

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

### **Publicações Didáticas (PUD)**

Incluem apostilas, notas de aula e manuais didáticos.

### **Publicações Seriadas**

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

### **Programas de Computador (PDC)**

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. São aceitos tanto programas fonte quanto executáveis.

### **Pré-publicações (PRE)**

Todos os artigos publicados em periódicos, anais e como capítulos de livros.

# 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](#)