



Ministério da  
**Ciência e Tecnologia**



# **REFINAMENTO ESTATÍSTICO DAS PREVISÕES DE VENTO DO MODELO ETA APLICADO AO SETOR EÓLIO-ELÉTRICO DO NORDESTE BRASILEIRO**

André Rodrigues Gonçalves

INPE  
São José dos Campos  
2011





Ministério da  
**Ciência e Tecnologia**



## **REFINAMENTO ESTATÍSTICO DAS PREVISÕES DE VENTO DO MODELO ETA APLICADO AO SETOR EÓLIO-ELÉTRICO DO NORDESTE BRASILEIRO**

André Rodrigues Gonçalves

Dissertação de Mestrado do Curso de Pós-Graduação em Meteorologia,  
orientada pelo Dr. Enio Bueno Pereira.

INPE  
São José dos Campos  
2011

## Ficha será revisada pelo SID.

### Dados Internacionais de Catalogação na Publicação

---

Cutter      Gonçalves, André Rodrigues.  
REFINAMENTO DAS PREVISÕES DE VENTO DO MODELO ETA  
UTILIZANDO REDES NEURAIAS ARTIFICIAIS / Nome Completo do  
Autor(es). - São José  
dos Campos: INPE, 2011.  
i + 0p. ; (aa/bb/cc/dd-TDI)

Mestrado em Meteorologia Instituto  
Nacional de Pesquisas Espaciais, São José dos Campos, 2011.  
Orientador: Enio Bueno Pereira.

1. Previsão de Vento. 2. Energia Eólica. 3. Refinamento Estatístico.  
4. Camada Limite Atmosférica. 5. Redes Neurais Artificiais.  
I. Título.

Copyright AAAA 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, reprográfico, de microfilmagem ou outros, sem a permissão escrita do INPE, 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 AAAA 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, recording, microfilming or otherwise, without written permission from the INPE, 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.

“O que você fizer será insignificante, mas é da maior importância que o faça”

*Mahatma Gandhi, pacifista indiano.*

“Se o homem não sabe a que porto se dirige, nenhum vento lhe será favorável”

*Sêneca, filósofo latino.*

“Sempre que ensinares, ensina também a duvidares do que ensinas”

*José Ortega y Gasset, filósofo espanhol.*



*Aos meus pais, João Bosco e Irley,  
e em memória de meus avós, Prof. Arlindo e Dona Zizinha.*



## AGRADECIMENTOS

Enfim, o mestrado se encerra. Depois de tudo pronto, o passado não aparenta tão árduo, quanto realmente foi. É sempre assim quando nos dedicamos a algo que vale a pena. Neste momento é importante recordar e agradecer a todos que nos apoiaram nesta empreitada. Felizmente são muitos, embora seja muita pretensão pensar que me lembrarei de todos.

Agradeço primeiramente a Deus pelo presente maior, a vida, me proporcionando saúde para lutar por tudo que almejo. À toda minha família pelo apoio incondicional desde o início do curso, em especial à Dona Irley, a melhor mãe do mundo, ao meu pai JB, pelo exemplo de caráter, aos meus irmãos Vanessa e Renan pelo carinho constante. Agradeço à minha avó, Dona Ivone, por cuidar tão bem de seu neto, à Tatati e Laís por fazer os meus dias mais descontraídos e, claro, ao mais novo membro da família, que distraiu e alegrou os momentos mais difíceis, meu sobrinho João Pedro.

Agradeço também aos amigos que me incentivaram a seguir este caminho, ao pessoal de Sanca e em especial à Aninha pelos conselhos sensatos. Aos colegas do mestrado pela amizade, companheirismo e boas risadas mesmo nos momentos mais tensos (CTF). Aos bons professores que tive, por mostrarem o lado estimulante de se fazer ciência e aos colegas de laboratório pela receptividade, cooperação, entusiasmo e principalmente pelas boas amizades, que tornaram nosso espaço uma verdadeira comunidade.

Agradeço ao INPE, à pos-graduação em meteorologia e à CAPES pelos recursos e infraestrutura disponibilizados sem os quais este trabalho não seria possível. Fundamental também foi a contribuição do Prof. Roberto Lyra, ao disponibilizar os dados anemométricos de Alagoas para a realização deste trabalho, a quem deixo também meus agradecimentos. Quero ainda registrar o meu respeito e admiração pelos meus orientadores, Enio e Fernando, que além de me proporcionar todo o suporte necessário à pesquisa, se mostraram sempre dispostos a conversar sobre os mais variados assuntos.

Um agradecimento especial deixo à Mári, pelo carinho e companheirismo ao longo deste último ano. Ela sabe melhor que ninguém o que a conclusão desta etapa representa para mim.

À todos que me apoiaram de alguma forma e que esqueci de mencionar, os meus sinceros agradecimentos. É curioso como é gratificante agradecer às pessoas !

## RESUMO

A capacidade eólica instalada vem crescendo consistentemente em diversos países nos últimos anos. Devido a seu caráter altamente variável no espaço e no tempo, um dos grandes desafios impostos por esta fonte energética está em prever a potência eólica disponível em um momento futuro, de modo a otimizar o despacho de energia, aumentar a segurança e a competitividade do setor. Apesar dos inúmeros estudos efetuados no exterior, pouco se avançou nesta área no Brasil, onde a capacidade eólica instalada vem atingindo patamares significativos, predominantemente na região Nordeste. Este trabalho se propõe a desenvolver um modelo de refinamento estatístico para previsão de vento a curto-prazo para a região Nordeste com o uso de regressões lineares múltiplas e redes neurais artificiais alimentadas por saídas do modelo de previsão do tempo Eta, utilizado pelo CPTEC/INPE. A previsão de vento próximo à superfície é discutida sob aspectos observacionais e de modelagem. É feita uma ampla avaliação dos possíveis preditores do modelo Eta para a previsão de ventos sobre o nordeste brasileiro. Testes de sensibilidade suportam o processo de configuração do modelo em redes neurais. Uma metodologia de refinamento estatístico das previsões numéricas de vento é proposta. Seu desempenho é comparado às saídas originais do modelo Eta e ao uso de regressões lineares múltiplas para previsões de vento a 50 m e de potência eólica a 80 m. Tanto o refinamento por regressões lineares múltiplas quanto por redes neurais artificiais se mostram superiores em relação às saídas do Eta. A técnica de redes neurais se mostra ligeiramente superior às regressões quando comparadas aos valores observados para 8 estações anemométricas da região. As correlações obtidas para o vento variam de 0,75 a 0,90 e o RMSE entre 0,93 m/s e 1,39 m/s. Para a potência eólica o RMSE varia entre 10,3% a 17,2% da potência eólica disponível. O ganho de potência garantida para uma probabilidade de 90% chega a 35% da potência eólica disponível em relação às previsões do Eta não refinadas.



# STATISTICAL DOWNSCALING OF ETA MODEL WIND FORECASTS APPLIED TO WIND POWER GENERATION AT NORTHEASTERN BRAZIL

## ABSTRACT

*The installed wind power capacity is increasing significantly worldwide. Due to wind high variability in space and time, one of the main challenges is to forecast the available wind power in the time frame from hours to days, in order to facilitate scheduled maintenance, prevent from extreme wind conditions and optimize the energy dispatch by electrical managing institutions, increasing wind energy competitiveness in the energy market. Despite several studies carried abroad, there are few advances in this subject in Brazil, where the installed wind power capacity is achieving relevant thresholds, predominantly in the northeastern region. This study aims to develop a methodology for short-term wind forecast from the outputs of CPTEC/INPE Eta- numerical weather predictions, based on multiple linear regressions and artificial neural networks. It is presented a discussion over the observational and modeling aspects of near surface wind fields. Several Eta derived predictors are evaluated for wind forecasts over northeastern Brazil. Sensitivity tests are carried out to configure the statistical model and a methodology is proposed for wind downscaling. Its performance is compared to Eta raw outputs and to linear regression models for wind and wind power forecasts at 50 m and 80 m respectively. Both the neural network and the regression model performed significantly better than the Eta outputs. The neural network downscaling performed slightly better than the regression when compared to 8 wind stations from northeastern region. The correlation coefficients obtained for wind varied between 0,75 and 0,90 and the RMSE from 0,93 m/s and 1,39 m/s. For wind power estimates the RMSE remains between 10,3% and 17,2% of the available wind power. The guaranteed wind power gain achieves up to 35% of the available power for a 90% error probability when compared to Eta outputs.*



## LISTA DE FIGURAS

		<u>Pág</u>
Figura 1.1	Evolução da potência eólica instalada mundial na última década	2
Figura 1.2	Matrizes Energética e Elétrica Brasileiras	3
Figura 1.3	Evolução da Intensidade Energética e Elétrica no Brasil	4
Figura 1.4	GEE emitidos pelas diversas fontes energéticas no Brasil	6
Figura 1.5	Mapa do potencial eólico brasileiro e do sistema interligado nacional SIN	7
Figura 1.6	Evolução da capacidade eólica instalada no Brasil	8
Figura 1.7	Contribuição da energia eólica na matriz elétrica de cada país ao final de 2010	10
Figura 2.1	A evolução diária da Camada Limite Atmosférica	16
Figura 2.2	Escalonamento dos fenômenos meteorológicos	17
Figura 2.3	Espectro de frequências dos processos atmosféricos e 'gap' espectral	18
Figura 2.4	Efeitos da estabilidade sobre os perfis de vento na CLS	23
Figura 2.5	Mecanismos de circulação de vale e montanha; ventos anabáticos e ventos catabáticos	24
Figura 2.6	Escoamento bidimensional sobre montanhas em atmosfera estável	25
Figura 2.7	Perfis verticais de velocidade ilustrando a aceleração do escoamento sobre topos de montanhas referentes ao experimento de Askervein	26
Figura 2.8	Esquema de desenvolvimento de camada limite interna (CLI) na transição entre superfícies com características distintas.	27
Figura 2.9	Sobreposição de CLIs na composição do perfil vertical de vento e temperatura em regiões heterogêneas	28
Figura 2.10	Componentes da circulação de brisa oceânica ou lacustre	29
Figura 3.1	Representação do espaçamento da grade E de Arakawa	30
Figura 3.2	Representação da topografia pela coordenada Eta	31
Figura 3.3	Fluxograma descrevendo as fontes potenciais de erro nos modelos de PNT	39
Figura 4.1	Modelo de neurônio de McCulloch e Pitts	41

Figura 4.2	Representação de uma rede tipo MLP	42
Figura 4.3	Funções de ativação degrau; linear; sigmoidal: 1) logística; 2) tangente hiperbólica	43
Figura 4.4	Modelo atual de neurônio perceptron	44
Figura 5.1	Mapa com a localização das torres anemométricas	49
Figura 5.2	Climatologia de precipitação sobre as torres de SJCA e ROTE	50
Figura 5.3	Ciclo anual do vento a 50 m em Petrolina-PE comparado à capacidade do reservatório da usina hidroelétrica de Sobradinho – BA	51
Figura 5.4	Distribuição da precipitação ao longo dos dados anemométricos disponíveis	52
Figura 5.5	Séries anemométricas e rosa dos ventos para cada torre anemométrica empregada neste estudo	53
Figura 5.6	Domínio da grade do modelo Eta/PREVENTO e topografia relativa na resolução de 5 km	55
Figura 6.1	Ondas senoidais introduzidas no conjunto de dados	62
Figura 6.2	Relação entre os valores de $Z_0$ e erro na determinação da rugosidade onde $z_0$ (50%) corresponde a mediana da distribuição. Método adotado consiste na identificação das posições de mínimo erro	64
Figura 6.3	Aerogerador de 2,1MW utilizado nas estimativas e sua curva de potência	72
Figura 7.1	Comparação entre mapas de magnitude do vento a 10m vento a 10m antes e após o processo de recorte, conversão e assimilação	74
Figura 7.2	Variáveis suspeitas durante a qualificação dos dados do modelo Eta/PV. Rótulos representam as taxas médias de falha para cada variável	75
Figura 7.3	Porcentagem dos dados anemométricos desqualificados por estação para cada nível	76
Figura 7.4	Fração de dados anemométricos reparados no nível mais elevado e respectivos erros (MAE e BIAS) cometidos	77
Figura 7.5	Séries temporais de magnitude observada do vento para as estações de PETR (50 m) e PALM (70 m) após reparos	78
Figura 7.6	Validação da sincronização entre os dados modelados e observados para todas as estações	78

Figura 7.7	Ciclos diários de intensidade e direção do vento ao longo do ano para SJCA e ROTE	83
Figura 7.8	Valores obtidos para a rugosidade $Z_0$ ao redor de cada torre em escala logarítmica	84
Figura 7.9	Correlações máximas obtidas para cada variável e exemplo de limiar adotado para definir os subconjuntos de preditores para SJCA e correlações máximas obtidas para ROTE	86
Figura 7.10	Defasagens de máxima correlação para cada variável em SJCA e ROTE	87
Figura 7.11	Comparação entre as dispersões obtidas entre o vento observado e o modelado pelo Eta/PV e as saídas da REGP para o vento a 50 m	88
Figura 7.12	Comparação entre as distribuições de probabilidade acumulada de RMSE e R e contra uma distribuição normal padrão para quatro subconjuntos de N preditores ( $N_{pred}$ ) da estação de SJCA	90
Figura 7.13	RMSE e R em função do número de preditores de cada subconjunto SJCA e ROTE	91
Figura 7.14	RMSE em função da razão entre neurônios e preditores para SJCA e ROTE	93
Figura 7.15	Testes de sensibilidade para cada configuração de RNA para SJCA e ROTE	94
Figura 7.16	Séries temporais de anomalias escalonadas para SJCA e ROTE	96
Figura 7.17	Dispersão final obtida entre o vento observado e o modelado pela RNA a 50 m para SJCA e ROTE	97
Figura 7.18	Dispersão final obtida entre o vento observado e o modelado pelo Eta/PV, REGPe RNA a 50 m para as demais estações	98
Figura 7.19	Comparação entre as correlações obtidas entre as anomalias de vento a 50 m para as configurações RNA_LIT e RNA_INT para todas as estações	99
Figura 7.20	Comparação entre as correlações R entre vento observado e modelado a 50 m pelo Eta/PV, REGP e RNA para todas as estações	100
Figura 7.21	Comparação entre os RMSE a 50 m pelos modelos Eta/PV, REGP e RNA para todas as estações	100
Figura 7.22	Comparação entre as séries temporais de Eta/PV, REGP e RNA e o vento observado a 50 m para (a) PETR; (b)	104

	SJCA; (c) TRFO; (d) ROTE; (e) MAGI; (f) GIRP; (g) AGUB; (h) PALM	
Figura 7.23	Comparação entre as séries temporais de potência eólica (kW) dadas pelo Eta/PV, REGP e RNA em relação à potência teórica observada a 80 m para (a) PETR; (b) SJCA; (c) TRFO; (d) ROTE; (e) MAGI; (f) GIRP; (g) AGUB; (h) PALM	114
Figura 7.24	Comparação entre os erros absolutos de potência eólica (kW) dadas pelo Eta/PV, REGP e RNA em relação à potência teórica observada a 80 m para (a) PETR; (b) SJCA; (c) TRFO; (d) ROTE; (e) MAGI; (f) GIRP; (g) AGUB; (h) PALM	118
Figura 7.25	Histograma da distribuição do erro entre a potência eólica prevista e disponível pelo Eta/PV, REGP e RNA para cada estação	120
Figura 7.26	Funções de densidade de probabilidade acumulada da diferença entre a potência eólica prevista e disponível pelo Eta/PV, REGP e RNA para cada estação	122
Figura 7.27	MAE obtido para as estimativas de potência eólica por estação	125
Figura 7.28	RMSE obtido para as estimativas de potência eólica por estação	125

## LISTA DE TABELAS

		Pág
Tabela 2.1	Comprimento de rugosidade típico para diversos tipos de superfície.	13
Tabela 5.1	Características das torres anemométricas.	48
Tabela 5.2	Descrição das variáveis Eta/PV.	55
Tabela 6.2	Experimentos referentes às rodadas do modelo Eta/PV.	58
Tabela 6.3	Critérios para o controle de qualidade dos dados anemométricos	62
Tabela 6.4	Limiares utilizados para os valores de correlação na pré-seleção dos preditores	65
Tabela 6.5	Testes de sensibilidade realizados sobre a rede neural	69
Tabela 7.1	Número de preditores por subconjunto antes e após REGP	87
Tabela 7.2	Resultados dos testes de Kolmogorov-Smirnov entre os subconjuntos de preditores adjacentes para SJCA e ROTE	90
Tabela 7.3	Resultados dos testes de Kolmogorov-Smirnov para as diferentes configurações (testes de sensibilidade) para SJCA e ROTE	93
Tabela 7.4	Testes de sensibilidade aprovados pelo KSTest para SJCA e ROTE.	95
Tabela 7.5	Resumo dos resultados obtidos pela modelagem por REGP e RNA para cada estação anemométrica	100
Tabela 7.6	Preditores selecionados para SJCA e respectivos impactos sobre o vento modelado pela RNA	105
Tabela 7.7	Preditores selecionados para ROTE e respectivos impactos sobre o vento modelado pela RNA. As defasagens mostradas significam atraso na previsão quando negativas	107
Tabela 7.8	Receita adicional por aerogerador devido à comercialização do excedente de energia garantida	123



## SUMÁRIO

	<u>Pág.</u>
<b>1 INTRODUÇÃO</b> .....	<b>1</b>
1.1. <b>Motivação</b> .....	<b>1</b>
1.2. <b>A Matriz Energética</b> .....	<b>2</b>
1.3. <b>A Energia Eólica</b> .....	<b>5</b>
1.4. <b>A Previsão de Ventos e a Geração Elétrica</b> .....	<b>9</b>
1.5. <b>Objetivos</b> .....	<b>10</b>
<b>2 A CAMADA LIMITE ATMOSFÉRICA</b> .....	<b>12</b>
2.1. <b>Estrutura vertical</b> .....	<b>12</b>
2.1.1. Sub-camada Laminar ou Viscosa .....	12
2.1.2. Camada Limite Superficial .....	12
2.1.3. Camada Limite Convectiva ou de Mistura .....	14
2.1.4. Camada Limite Estável ou Noturna (CLE) .....	14
2.1.5. Zona de Entranhamento (ZE) .....	14
2.2. <b>Turbulência Atmosférica</b> .....	<b>15</b>
2.3. <b>Equacionamento</b> .....	<b>18</b>
2.4. <b>Fluxos Verticais e Instabilidade</b> .....	<b>19</b>
2.5. <b>Teoria da Similaridade e Perfis Verticais</b> .....	<b>20</b>
2.6.1. Circulações de Vale e Montanha .....	23
2.6.2. Escoamento sobre Montanhas .....	23
2.6.3. Camada Limite Interna.....	25
2.6.4. Circulações de Brisa .....	27
<b>3 A PREVISÃO NUMÉRICA DE VENTO EM SUPERFÍCIE</b> .....	<b>29</b>
3.1. <b>O Modelo Eta</b> .....	<b>29</b>
3.2. <b>A Modelagem do Vento em Superfície</b> .....	<b>31</b>
3.3. <b>Aplicações de Modelagem Numérica do Vento</b> .....	<b>32</b>
<b>4 REDES NEURAIS ARTIFICIAIS</b> .....	<b>39</b>
4.1. <b>Histórico</b> .....	<b>40</b>
4.2. <b>Arquitetura</b> .....	<b>41</b>
4.3. <b>Aplicações</b> .....	<b>43</b>
<b>5 DADOS</b> .....	<b>47</b>
5.1. <b>Dados Observacionais</b> .....	<b>47</b>

5.2.	Dados do Modelo Eta/PREVENTO.....	53
<b>6</b>	<b>METODOLOGIA.....</b>	<b>56</b>
<b>6.2.</b>	<b>Tratamento dos Dados.....</b>	<b>57</b>
	6.2.1.Dados do Modelo.....	57
	6.2.2.Dados Anemométricos.....	61
	6.2.3.Sincronização dos Dados.....	64
<b>6.3.</b>	<b>Seleção dos Preditores.....</b>	<b>64</b>
	6.3.1.Definição dos Subconjuntos por Limiar de Correlação.....	64
	6.3.2.Regressão Linear Passo-a-Passo.....	65
	6.3.3.Treinamento por Redes Neurais Artificiais.....	66
<b>6.4.</b>	<b>Desenvolvimento do Modelo em RNAs.....</b>	<b>67</b>
	6.4.1.Definição do Número de Neurônios Ocultos.....	67
	6.4.2.Configuração dos Testes de Sensibilidade.....	68
<b>6.5.</b>	<b>Avaliação dos Modelos Obtidos.....</b>	<b>70</b>
	6.5.1.Desempenho na Previsão de Ventos.....	70
	6.5.2.Desempenho na Previsão de Potência Eólica.....	71
<b>7</b>	<b>RESULTADOS.....</b>	<b>73</b>
<b>7.1.</b>	<b>Verificação da Assimilação e do Tratamento dos Dados.....</b>	<b>73</b>
<b>7.2.</b>	<b>Climatologia.....</b>	<b>77</b>
	7.2.1.Ciclos Diários.....	78
	7.2.2.Determinação do Comprimento de Rugosidade.....	83
<b>7.3.</b>	<b>Preditores Selecionados.....</b>	<b>84</b>
<b>7.4.</b>	<b>Neurônios na Camada Oculta.....</b>	<b>91</b>
<b>7.5.</b>	<b>Testes de Sensibilidade.....</b>	<b>92</b>
<b>7.6.</b>	<b>Avaliação das Previsões.....</b>	<b>96</b>
<b>7.7.</b>	<b>Avaliação dos Pesos dos Preditores.....</b>	<b>103</b>
<b>7.8.</b>	<b>Impactos sobre a Potência Eólica Disponível.....</b>	<b>111</b>
<b>8</b>	<b>CONCLUSÃO.....</b>	<b>126</b>
<b>9</b>	<b>SUGESTÕES DE TRABALHOS FUTUROS.....</b>	<b>130</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>132</b>

# 1 INTRODUÇÃO

## 1.1. Motivação

É notável o avanço alcançado pelos modelos de Previsão Numérica de Tempo (PNT) nas últimas décadas em todo o mundo. Além da pesquisa extensiva na área de modelagem atmosférica, a melhor organização dos dados observacionais e o aumento da capacidade computacional podem ser colocados como fatores decisivos para este sucesso. Em contrapartida, o progresso tecnológico e o amadurecimento da sociedade impõem novos desafios à meteorologia, particularmente com relação à disponibilidade de recursos energéticos. Neste contexto, a qualidade das previsões de vento próximo à superfície têm se mostrado um fator determinante para o desenvolvimento da energia eólica em diversas nações, no momento em que esta se configura como a alternativa mais sólida de expansão da matriz energética mundial. A Figura 1.1 ilustra o forte crescimento do setor eólico na última década. O objeto deste trabalho é justamente a previsão de vento próximo à superfície, sendo esta uma área de atuação da meteorologia aplicada.

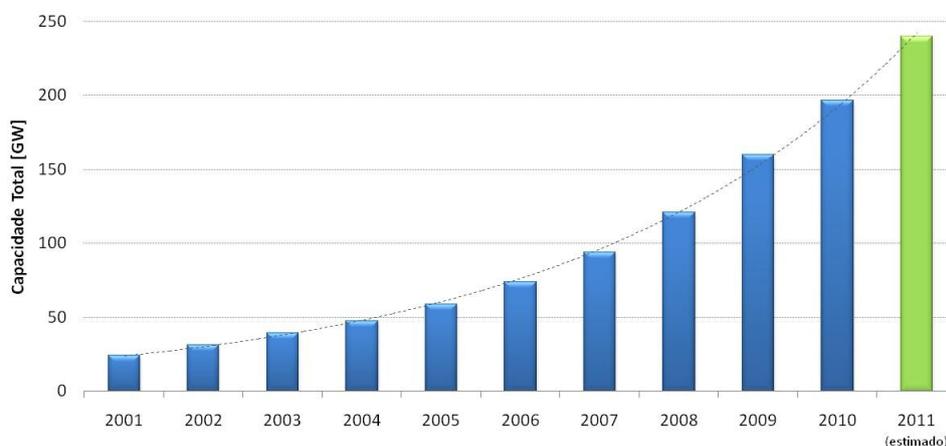


Figura 1.1 – Evolução da potência eólica instalada mundial na última década.

Fonte: GWEC (2010).

## 1.2. A Matriz Energética

A evolução da sociedade, da pré-história à era moderna, sempre esteve ligada a evolução da demanda dos recursos energéticos. O mundo passa atualmente por um período de transição, onde o uso predominante de combustíveis fósseis como fonte energética se mostra insustentável a médio e longo prazo, tornando evidente a necessidade de diversificação de sua matriz energética. Esta matriz energética é composta pela soma de todas as formas de energia consumidas por um país, contabilizando o consumo de carvão mineral e vegetal, derivados do petróleo, gás natural, lenha, energia hidráulica, solar, eólica, da biomassa, dentre outras. Já a matriz elétrica contabiliza apenas as fontes energéticas utilizadas para conversão em eletricidade, conforme ilustra a Figura 1.2.

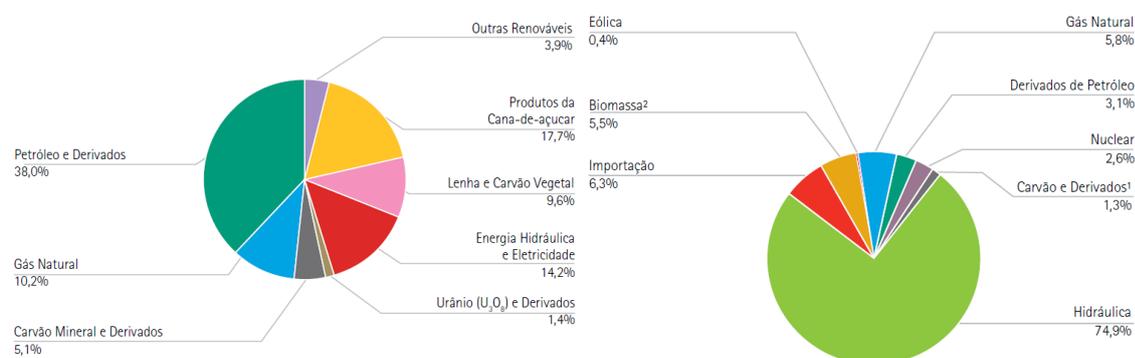


Figura 1.2 – Matriz Energética Brasileira (e); Matriz Elétrica Brasileira (d); Ambos referentes à oferta interna de energia no ano de 2010.

Fonte: Balanço Energético Nacional (Empresa de Pesquisa Energética, 2011)

A demanda energética está ligada ao crescimento populacional, ao crescimento econômico e à melhoria da qualidade de vida de uma população, de modo que o desenvolvimento dos países requer um aumento do consumo de energia per capita. Já foi demonstrado que existe uma relação explícita entre demanda energética ( $E$ ) e o crescimento do Produto Interno Bruto (PIB) de uma nação, conforme a Equação 1.1, onde se define a intensidade energética ( $I$ ) como a energia necessária para se produzir uma unidade monetária de PIB (Goldemberg e Villanueva, 2003).

$$E = I \times PIB \quad (1.1)$$

De maneira análoga, pode-se determinar uma intensidade elétrica, associada exclusivamente à demanda de energia elétrica, conforme ilustra a Figura 1.3. No Brasil, segundo dados do Balanço Energético Nacional (BEN 2010) e do Plano Decenal de Expansão de Energia (PDE 2019), a demanda de energia elétrica vêm crescendo em torno de 5% ao ano e projeções indicam que este ritmo deve ser mantido nos próximos 10 anos. Estes mesmos dados mostram que tanto a intensidade energética quanto a elétrica, que estão ligadas à eficiência no uso da energia, se apresentam praticamente estagnadas no Brasil. Com base nestas informações fica clara a necessidade de adoção de uma política mais agressiva quanto à eficiência e uso racional da energia, além de realizar um planejamento energético confiável por meio da diversificação da matriz energética. Estas ações podem contribuir de forma efetiva para ambos os aspectos mencionados e favorecem o desenvolvimento sócio-econômico de regiões do país com grande potencial para geração de energia renovável. Além de tudo, a menor dependência em relação à energia hidráulica contribui para a segurança energética nacional. Interrupções, ou “apagões” ocorridos no passado já demonstraram seu grande impacto sobre a economia e a sociedade como um todo.

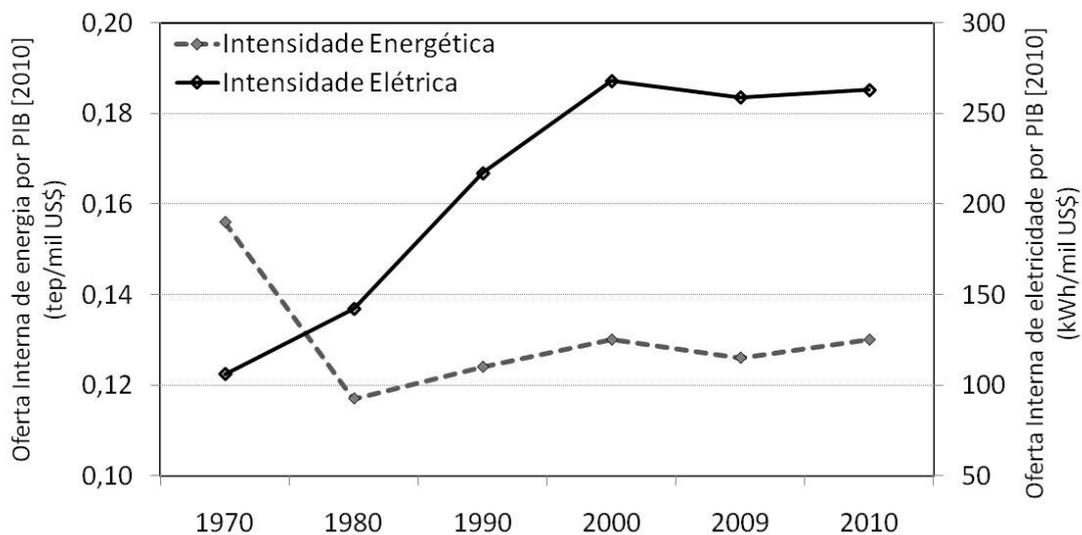


Figura 1.3 – Evolução da Intensidade Energética e Elétrica no Brasil.

Fonte: Balanço Energético Nacional (Empresa de Pesquisa Energética, 2011).

A expansão da oferta de energia é um assunto delicado, que envolve não somente o custo de investimento, mas também o custo ambiental de cada fonte o qual é difícil de ser medido. Fontes provenientes de combustíveis fósseis aumentam a concentração de gases do efeito estufa (GEE) na atmosfera, principalmente CO<sub>2</sub>, considerado o responsável pelas mudanças climáticas globais de origem antrópica (IPCC, 2007). Por ser este um assunto estratégico, têm ganhado uma importância cada vez maior no cenário internacional, originando fóruns de discussão científica e política como o IPCC (*Intergovernmental Panel on Climate Change*), que constitui um esforço da comunidade científica internacional no prognóstico de cenários climáticos futuros e a COP (Conferência das Nações Unidas para Mudanças Climáticas). O último deles, a COP16 realizada em Cancún, no México, demonstrou que ainda se está longe de um acordo internacional efetivo para a redução de emissões de GEE.

O Brasil é um país privilegiado sob o aspecto energético, pois seu relevo, hidrografia e clima tropical permitem um aproveitamento das diversas formas de energia originárias do Sol, como a hidráulica, a solar, a eólica e da biomassa, sendo estas, fontes de energia renováveis (Pereira e Lima, 2008). Um conceito amplamente utilizado para definir energia renovável, a caracteriza como toda fonte onde sua taxa de renovação natural é maior do que sua taxa de utilização, o que inclui a energia das marés e geotérmica além das citadas anteriormente.

Atualmente, o Brasil possui uma matriz de geração elétrica relativamente limpa, sustentada pelas usinas hidráulicas (74,9%), no entanto ao considerar sua matriz energética total, ainda se consome uma grande quantidade de energia de origem fóssil (53,3%), proveniente principalmente no setor de transportes, conforme ilustrado anteriormente pela Figura 1.2. Por fim, cabe ressaltar que apesar da matriz energética relativamente renovável, o Brasil está entre os países que mais emitem CO<sub>2</sub> no mundo, devido à alta taxa de desmatamento associadas às queimadas na região amazônica, responsáveis por aproximadamente 76% das emissões no país conforme ilustra a Figura 1.4. (Ministério da Ciência e Tecnologia, 2011).

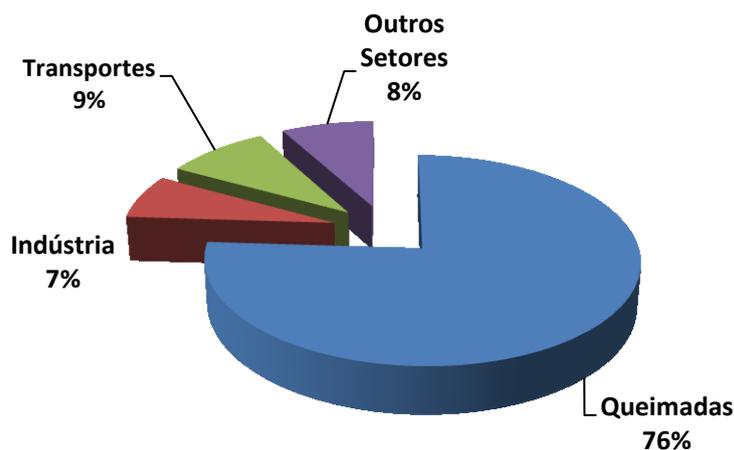


Figura 1.4 – GEE emitidos pelas diversas fontes energéticas no Brasil.

Fonte: Ministério da Ciência e Tecnologia, 2009.

Apesar do enorme potencial hidroelétrico brasileiro, estimado em 260GW, (ELETROBRAS, 2003), as grandes bacias das regiões Centro-Sul e Nordeste do Brasil foram exploradas extensivamente no último século, o que faz com que grandes usinas hidroelétricas sejam factíveis apenas no Centro-Norte do país. Nesta região o custo ambiental é elevado pela imensa área alagada, que geralmente engloba florestas, áreas de preservação, terras indígenas e povoadas. Isto aliado à baixa densidade de potência destas usinas, eleva o custo e o risco destes empreendimentos, ao ponto de se tornar necessário uma forte contrapartida de investimento público e uma flexibilização incondicional da legislação ambiental para garantir a sua viabilidade. Dado este cenário, a expansão energética brasileira na última década foi suportada por fontes de impacto ambiental gradativamente maior, seja por tais mega-usinas hidroelétricas na região norte, seja por usinas termoelétricas a gás natural, que apesar de mais eficientes que o carvão ou óleo diesel ainda emitem uma quantidade considerável de gases do efeito estufa durante sua operação.

### 1.3. A Energia Eólica

O Brasil possui um potencial eólico estimado de 143GW (Camargo *et al.*, 2001), o qual deve ser revisado em breve para valores ainda maiores ao se considerar o aumento do porte médio dos aerogeradores e a possibilidade de exploração do

potencial offshore (Pimenta *et al.* 2008). Ainda, projeções de cenários climáticos futuros indicam uma alta probabilidade de aumento deste potencial no decorrer deste século (Pes, 2010), o que coloca o Brasil em uma posição privilegiada no cenário mundial de recursos eólicos. Tal abundância se verifica principalmente nas regiões Nordeste e Sul, com predominância das regiões costeiras, o que se configura como uma vantagem para a transmissão elétrica, dada a maior densidade de linhas de transmissão e proximidade de grandes centros urbanos, que aliado à densa malha viária local repercute favoravelmente no custo dos projetos eólicos. A Figura 1.5 ilustra esta relação entre o potencial eólico e o Sistema Interligado Nacional (SIN)

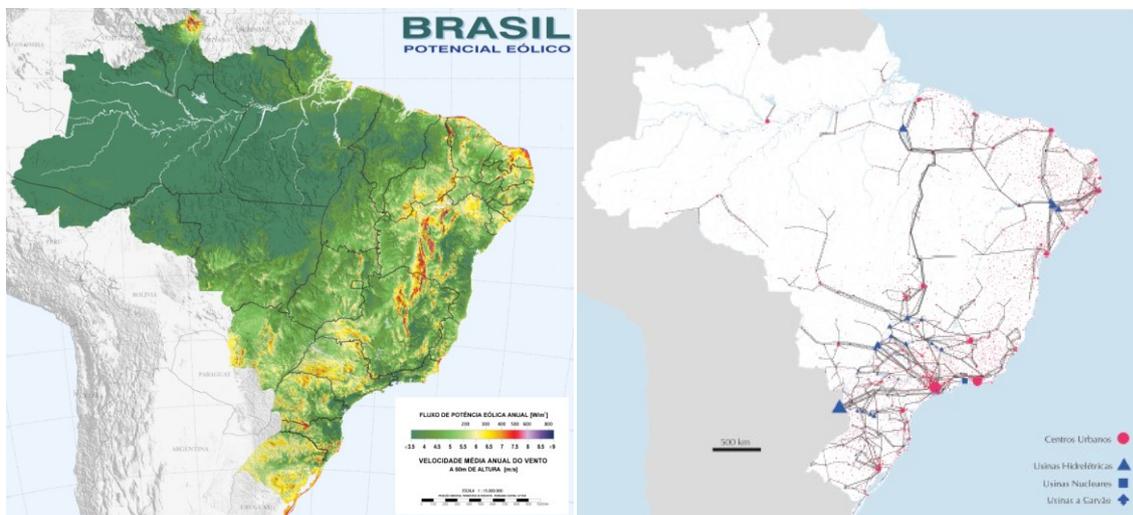


Figura 1.5 – Mapa do potencial eólico brasileiro (e); e do sistema interligado nacional SIN (d).

Fonte: Amarante *et al.*( 2001).

No Nordeste brasileiro ocorre ainda uma característica vantajosa que é a complementaridade hidro-eólica, uma vez que as vazões dos rios e o regime de ventos possuem um ciclo sazonal alternado, garantindo uma maior autonomia de geração da região e reduzindo as perdas por transmissão de energia a longas distâncias (Amarante *et al.*, 2001). A disponibilidade destas duas fontes energéticas aliado ao despacho complementar das usinas termoelétricas permite o ajuste à demanda de eletricidade, dando maior segurança à geração elétrica na região. Este gerenciamento só é possível devido ao Sistema Interligado Nacional (SIN), uma malha de transmissão elétrica que permite o remanejamento de cargas entre suas

subestações, o que é feito pelo Operador Nacional do Sistema elétrico (ONS), uma empresa de gestão autônoma criada pelo governo especificamente para este fim.

Neste contexto o governo federal instituiu, a partir de 2004, o PROINFA, que tem por objetivo aumentar em 3.300 MW a participação de usinas eólicas, pequenas centrais hidroelétricas (PCHs) e de biomassa na matriz de geração elétrica brasileira, o que corresponderia a cerca de 3% da capacidade de geração total instalada no país. Conforme ilustra a Figura 1.6, este programa marcou o início da expansão da energia eólica no Brasil e seus resultados justificam a perspectiva de um forte crescimento no setor nos próximos anos, acompanhando a tendência mundial (Dutra e Szklo, 2007). Além da energia contratada pelo PROINFA, outros empreendimentos eólicos têm sido contratados em leilões de energia organizados regularmente pelo governo, os quais somados à capacidade eólica atual alcançarão cerca de 4000 MW instalados até o final de 2013 (Martins e Pereira, 2011).

No leilão realizado em Julho de 2011, cerca de 48% de toda a energia contratada foram de plantas eólicas, enquanto 26% correspondem ao gás natural, 14% à biomassa e 12% à hídrica. O fato mais surpreendente foi que preço médio do MWh eólico ficou abaixo das termelétricas a gás, algo inédito no mercado de energia mundial, o que comprova mais uma vez a maturidade e competitividade desta tecnologia para geração de eletricidade.

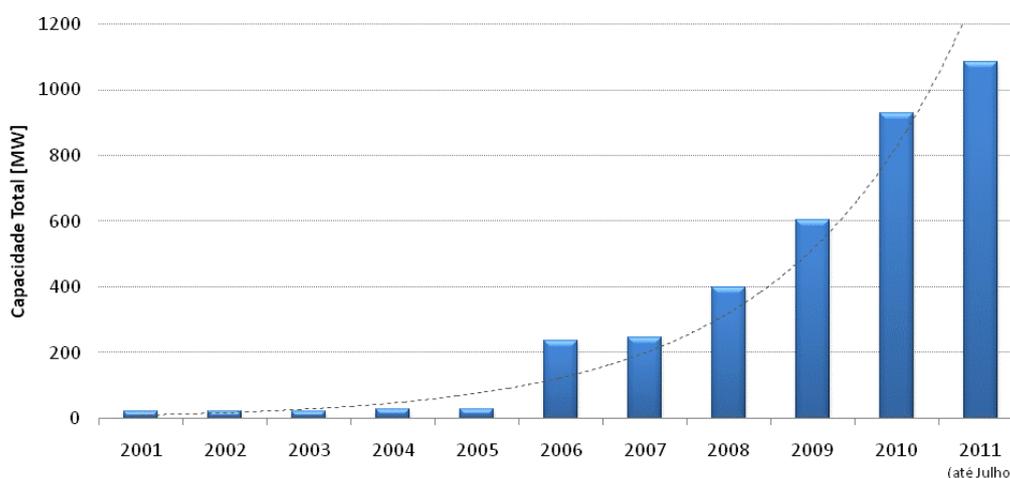


Figura 1.6 – Evolução da capacidade eólica instalada no Brasil.

Fonte: ANEEL, 2011.

Conforme demonstrado anteriormente, têm-se observado uma inserção crescente da geração eólico-elétrica na matriz energética mundial, e estudos indicam que este crescimento continuará nos próximos anos (WWEA, 2009; GWEC,2010). Isto se justifica principalmente pelo estágio de maturidade tecnológica alcançado pela indústria eólica, que possibilitou a redução dos custos da eletricidade produzida, aliado à pressão internacional para mitigar as possíveis causas do aquecimento global. Um fato interessante a ser ressaltado ocorreu na Europa entre 2008 e 2009, quando pelo segundo ano consecutivo a capacidade de geração eólico-elétrica instalada anualmente superou todas as outras fontes de energia elétrica, chegando a 39% da total, enquanto o gás natural representou 26% e a fotovoltaica 16% (EWEA,2009).

Entretanto, a geração eólica apresenta uma variabilidade elevada associada à dinâmica da atmosfera, tornando necessário que se desenvolvam métodos mais confiáveis para a previsão de curto prazo da potência disponível. Estas ferramentas além de permitirem um melhor gerenciamento do sistema elétrico interligado, auxiliam também na manutenção programada dos aerogeradores, reduzindo os riscos de danos por ventos extremos e tornando a energia eólica mais competitiva no mercado de eletricidade mundial (Campbell, 2007; Monteiro *et al.*, 2009; Lerner *et al.* 2009). Devido a esta incerteza associada à determinação da potência eólica disponível, existe uma discussão na comunidade internacional sobre os limites seguros para a penetração da energia eólica na matriz energética de um país (Landberg *et al.* 2003). Algumas economias já experimentam uma inserção próxima dos 20%, o que torna o aprimoramento das ferramentas de previsão um fator determinante para a segurança energética destas nações.

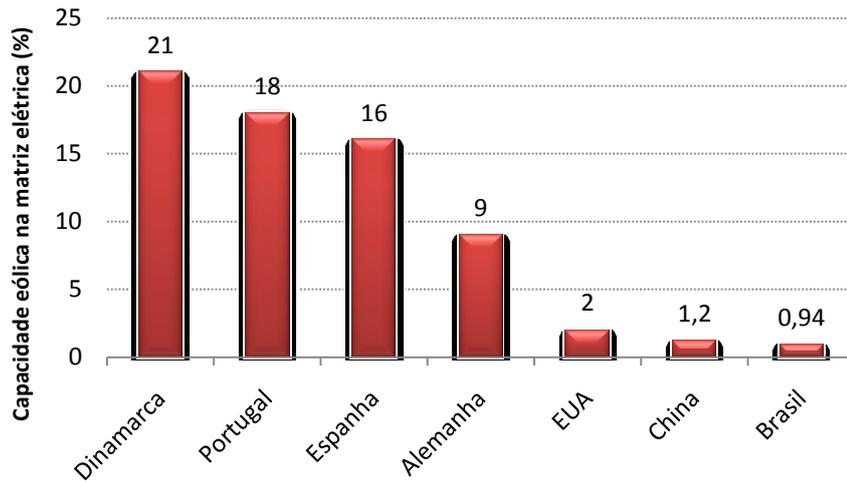


Figura 1.7 – Contribuição da energia eólica na matriz elétrica de cada país ao final de 2010.

Fonte: EWEA, 2009; Brasil. ANEEL 2011

#### 1.4. A Previsão de Ventos e a Geração Elétrica

A previsão de potência eólica disponível depende basicamente da previsão do vento à altura do aerogerador, situado entre 50 e 120 m acima da superfície, uma vez que as incertezas no processo de geração em si são menores e bem conhecidas. Desprezando-se as perdas relativas à instalação da usina, tem-se que esta potência  $P$  é proporcional ao cubo da velocidade do vento, conforme a Equação 1.2 (Burton *et al.*, 2001), evidenciado o impacto da qualidade das previsões de vento na energia final gerada (Tsikalakis *et al.*, 2008).

$$P = \frac{1}{2} \cdot C_p \cdot A \cdot \rho \cdot U^3 \quad (1.2)$$

Onde  $P$  é a potência instantânea, em Watts, e  $C_p$  é o coeficiente adimensional de potência do aerogerador que é função também da velocidade do vento.

Uma descrição mais detalhada da dinâmica atmosférica associada ao aproveitamento da energia eólica pode ser encontrada em Martins *et al.*, (2008).

O vento tem implicações também na capacidade de despacho de energia elétrica pelas linhas de transmissão, pois é um dos principais parâmetros meteorológicos que controlam a capacidade de dispersão do calor gerado devido ao aquecimento

destas linhas pelo Efeito Joule. Adicionalmente, previsões de vento melhores podem trazer outros benefícios à sociedade, principalmente para aeroportos, atividades esportivas e no lazer em geral. Além disso, dada a perspectiva de que a capacidade eólico-elétrica no Brasil instalada chegue a 3% do total por volta de 2014, representantes da EPE (Empresa de Planejamento Energético) e da ONS, órgãos de planejamento e operação do sistema elétrico ligados ao governo, já se pronunciaram em algumas oportunidades quanto à necessidade de uma ferramenta nacional confiável de previsão de potência eólica disponível a curto-prazo (Lopes, 2010).

No refinamento das previsões de vento duas abordagens são possíveis: a modelagem física (ou dinâmica) em microescala (Lange, 2006), ou modelagens estatísticas diversas (Fan *et al.*, 2009). Comumente opta-se por uma combinação entre ambos para um melhor desempenho das previsões (Ernst *et al.*, 2007; Lange, 2008). Para uma discussão mais aprofundada sobre as vantagens de cada método de previsão existem alguns trabalhos que tratam o tema em detalhes (Wu e Hong, 2007; Thor, 2008; Monteiro *et. al*, 2009;)

O refinamento através de modelagem física de microescala se faz necessário principalmente em terrenos complexos e, apesar de apresentar como vantagem a capacidade de generalização, seu custo computacional é alto para um modelo operacional, além de seu desempenho não dispensar o posterior ajuste por modelos estatísticos. Por esta razão optou-se por uma abordagem estatística, uma vez que esta apresenta a perspectiva de maiores ganhos de acurácia em relação à abordagem dinâmica para um ajuste local.

## **1.5. Objetivos**

Este trabalho se propõe a desenvolver uma metodologia para o refinamento da previsão de curto prazo do vento próximo à superfície baseada em Redes Neurais Artificiais (RNAs), a partir de saídas operacionais do modelo de PNT Eta do CPTEC. O intuito é que estas previsões sejam adequadas ao clima e território brasileiro, especificamente ao Nordeste, permitindo assim que as usinas eólicas adquiram maior competitividade e que o sistema elétrico seja gerenciado de forma otimizada.

A metodologia empregada neste estudo consiste no uso de uma técnica de ajuste estatístico local, neste caso RNAs, para a redução do erro entre as saídas de vento observadas e estimadas pelo modelo Eta para diversas estações no Nordeste brasileiro. A altura do solo adotada para estas comparações foi de 50 m, por ser este o nível mais comum das medições realizadas pelas torres anemométricas. Foram selecionadas variáveis do modelo mais correlacionadas ao erro das previsões de vento, as quais são utilizadas como entrada para a RNA. Após o período de treinamento, que consiste de um processo iterativo de ajuste dos coeficientes do modelo, os resultados obtidos demonstram reconhecer algum padrão nos erros das previsões, fornecendo saídas otimizadas de vento no nível de 50 m. Por se tratar de um ajuste local, isto foi feito para diversas estações, onde foi avaliado também o desempenho, e a capacidade de generalização de cada modelo obtido.

Esta dissertação foi estruturada de modo que no Capítulo 2 é feita uma revisão sobre a Camada Limite Atmosférica (CLA), abordando suas características, turbulência, dificuldades relacionadas à previsão de variáveis em seu interior, avaliando os efeitos do relevo e rugosidade sobre a dinâmica dos ventos. Em seguida, no Capítulo 3 é detalhado o modelo de PNT Eta do CPTEC, assim como suas parametrizações e princípios incorporados para a previsão do vento na camada limite superficial. Posteriormente, no Capítulo 4, são introduzidos os conceitos sobre Redes Neurais Artificiais, enfatizando suas potencialidades, e aplicabilidade para este estudo. No Capítulo 5 são descritos os dados utilizados. O capítulo 6 descreve a metodologia empregada, englobando a assimilação, processamento e qualificação dos dados; a seleção de preditores; a configuração e treinamento da rede neural; e o método de avaliação das previsões. Já no Capítulo 7 são apresentados os resultados obtidos, abordando os testes efetuados na definição da arquitetura do modelo e a validação dos resultados para cada uma das estações. Dada a interdisciplinaridade do tema proposto, a revisão bibliográfica foi distribuída ao longo de cada capítulo objetivando uma melhor contextualização dos trabalhos citados.

## **2 A CAMADA LIMITE ATMOSFÉRICA**

A troposfera, porção inferior da nossa atmosfera onde ocorrem os principais fenômenos meteorológicos, é limitada em sua base pela superfície terrestre. Pode-se distinguir, no entanto uma camada inferior da troposfera onde os efeitos da superfície são sentidos em uma escala de tempo de horas, à qual dá-se o nome de Camada Limite Atmosférica (CLA) (Stull, 1988). Sua espessura varia de centenas de metros a quase 2 km, dependendo das condições de tempo e da região. O seu limite superior é definido pela presença de uma inversão térmica, denominada “zona de entranhamento”, extremamente estável, que dificulta o transporte de grandezas escalares como umidade, calor e poluentes para a camada superior denominada atmosfera livre. Uma característica importante da CLA é a presença de turbulência, responsável pelo rápido fluxo vertical destes escalares em seu interior e também por dificultar bastante a modelagem física desta camada.

### **2.1. Estrutura vertical**

A CLA possui uma estrutura vertical que se modifica ao longo do dia devido ao ciclo diário de aquecimento e resfriamento da superfície, sendo que esta pode ser dividida da seguinte maneira:

#### **2.1.1. Sub-camada Laminar ou Viscosa**

Caracterizada pela ausência de turbulência e pela grande tensão de cisalhamento devido a viscosidade molecular. Sua espessura é da ordem de milímetros e depende apenas da velocidade do vento e do tipo de superfície

#### **2.1.2. Camada Limite Superficial**

A Camada Limite Superficial (CLS) é caracterizada pelo forte gradiente vertical de vento, umidade e temperatura, o que faz com que os fluxos de calor, momento e umidade sejam também maiores (Arya, 2001). No entanto estes fluxos podem ser considerados aproximadamente constantes, variando em até 10% para toda a camada (Stull, 1988). Sua espessura dependerá da estabilidade atmosférica, da intensidade do vento e da superfície em que se desenvolve, podendo chegar no máximo, a 200 m. É nesta a porção da CLA onde efeito da fricção com a superfície

e a presença de obstáculos exerce maior influência sobre o escoamento. De maneira geral, seu perfil de velocidade pode ser aproximado por uma função logarítmica conforme descrito pela Equação 2.1, tendo em mente que este ajuste varia ao longo do dia em função de alterações no escoamento e nos fluxos.

$$U = \frac{u^*}{k} \cdot \ln\left(\frac{Z}{Z_0}\right) \quad (2.1)$$

Onde  $k$  é a constante de Von-Karman,  $u^*$  a velocidade de fricção e  $Z_0$  o comprimento de rugosidade de momento.

O comprimento de rugosidade  $Z_0$  é uma medida da turbulência gerada pelos elementos de rugosidade da superfície, e na prática representa a altura até a qual o vento deve ser considerado nulo. Em tese é também uma maneira de se medir a capacidade de uma determinada superfície de absorver momento, sendo bastante utilizada em formulações para CLS. Seu valor é estimado empiricamente através de um ajuste de curvas pela Equação 2.1.

Tabela 2.1: Comprimento de rugosidade típico para diversos tipos de superfície.

<b>Tipo de Superfície</b>	<b>Comprimento de Rugosidade (m)</b>
gelo	$10^{-5}$
água	$10^{-4} - 10^{-3}$
neve ondulada	0,002
grama baixa	0,005
grama alta	0,02
plantações de cereais	0,05
arbustos	0,2
florestas	1 – 2
subúrbios da cidade	0,5 – 2

Fonte: Adaptado de Foken (2008).

### **2.1.3. Camada Limite Convectiva ou de Mistura**

Acima da CLS, durante o dia se desenvolve a CLC, que consiste na região onde a turbulência (predominantemente térmica) é responsável por misturar bem os escalares como temperatura e umidade, por isso também chamada de Camada de Mistura. Sua formação ocorre tipicamente durante o dia, quando o aquecimento da superfície faz com que a turbulência térmica alcance níveis cada vez mais elevados até que a Camada Limite Estável (Noturna) seja completamente erodida e alcance a camada residual, atingindo seu máximo desenvolvimento em torno das 17:00 local. Sua espessura depende principalmente dos fluxos de calor sensível e latente, variando de centenas de metros a 2 km. A camada residual é simplesmente o que restou da própria CLC do dia anterior, uma vez que a Camada Limite Noturna geralmente não atinge o topo da CLA. Existem casos onde a CLC pode se formar a noite devido à turbulência mecânica gerada por jatos noturnos.

### **2.1.4. Camada Limite Estável ou Noturna (CLE)**

Esta camada se desenvolve durante a noite a partir do resfriamento radiativo da atmosfera e da superfície, que resfria a camada de ar mais próxima dando origem a uma camada bastante estável que atinge seu máximo desenvolvimento no fim da madrugada. Seu limite inferior pode ser considerado a superfície, pois se confunde com a CLS noturna e sua altura é bastante variável, oscilando entre 100 m e 500 m.

### **2.1.5. Zona de Entranhamento (ZE)**

Região bastante estável de interface entre a CLC e a atmosfera livre, demarcando o limite superior da CLA. Seu topo é definido como a altura alcançada pela corrente da corrente termal mais intensa, (conhecido como 'overshoot') e por isso varia bastante em função da intensidade das termais em cada região. Sua espessura varia entre 50m e 200m.

A Figura 2.1 ilustra o desenvolvimento típico de algumas destas camadas ao longo de um dia.

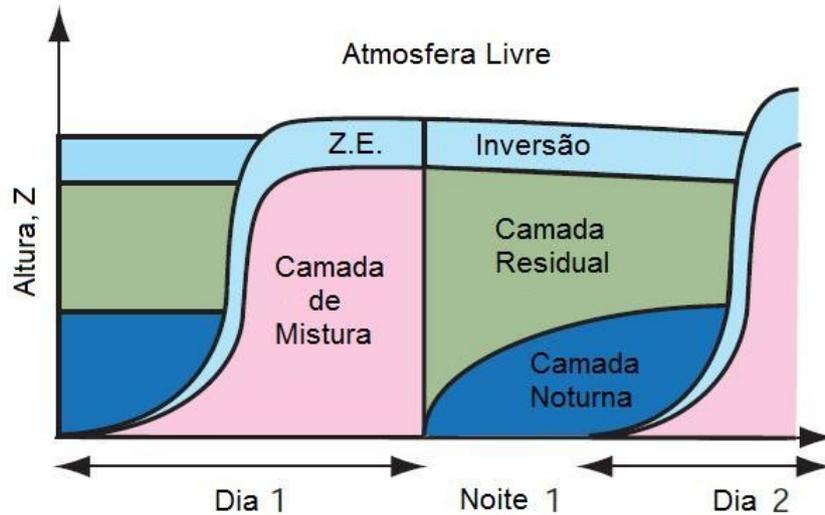


Figura 2.1 – A evolução diária da Camada Limite Atmosférica.  
 Fonte: Adaptado de Stull (1988).

## 2.2. Turbulência Atmosférica

As escalas de tempo e comprimento são utilizadas para caracterizar os fenômenos estudados na meteorologia. Há algum tempo foi proposta uma divisão entre os diversos processos em função de suas escalas características conforme ilustrado na Figura 2.2, onde há um escalonamento dos fenômenos até o limite da turbulência isotrópica (Orlanski, 1975). Pode-se notar que para que um fenômeno seja definido como de microescala, sua dinâmica deve ser regida por processos internos e restritos à CLA, os quais possuem uma escala de tempo inferior a 1 hora, e comprimento inferior a 2 km.

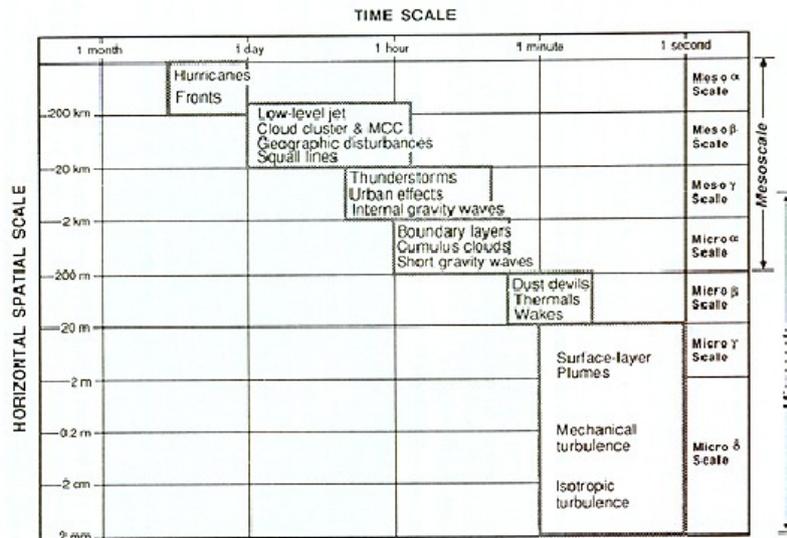


Figura 2.2 – Escalonamento dos fenômenos meteorológicos.  
Fonte: Orlandi (1975).

A turbulência consiste de um processo inerente à CLA, onde a heterogeneidade do terreno associada à modulação das forçantes externas, como vento e radiação solar, dão origem a uma sobreposição de movimentos oscilatórios denominados vórtices. Estes vórtices representam a busca da atmosfera por uma condição de equilíbrio, a qual nunca é alcançada dada a rapidez com que as mudanças ocorrem dentro da CLA.

A turbulência pode ter origem mecânica, térmica ou inercial. A primeira é originada pelo cisalhamento do vento, enquanto a segunda é devido à instabilidade atmosférica. Já a inercial pode ser considerada como consequência das duas primeiras, uma vez que vórtices maiores dão origem a vórtices menores em suas bordas, e assim sucessivamente até o limite da turbulência molecular, que é dissipada na forma de calor. Este processo é conhecido como “efeito cascata”, conforme ilustrado pela Figura 2.3.

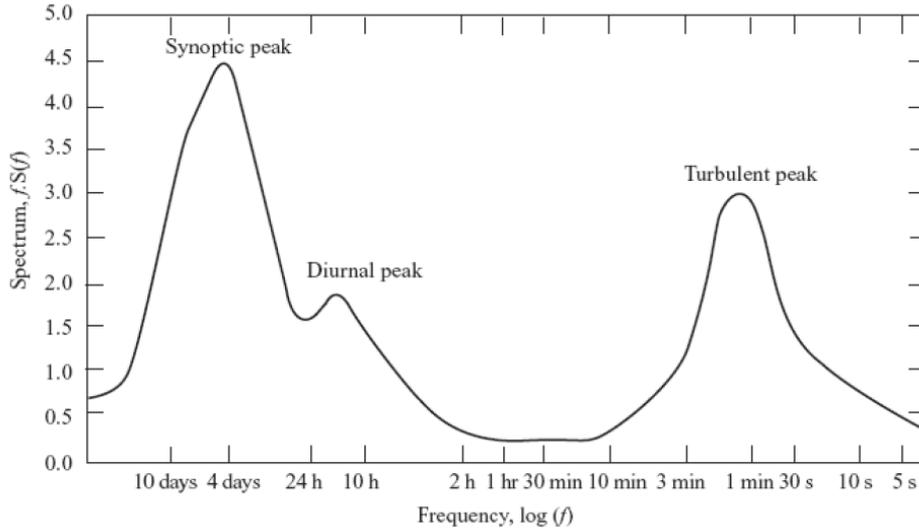


Figura 2.3 – Espectro de freqüências dos processos atmosféricos e ‘gap’ espectral.

Fonte: Van der Hoven, (1957).

Nota-se na Figura 2.3 a existência de uma lacuna entre os períodos de 10 minutos a 10 horas, o que evidencia que há uma separação de escalas entre os fenômenos atmosféricos, sendo a porção à esquerda caracterizada pelo escoamento médio e a da direita tratada como turbulência. Por este motivo medidas anemométricas que visam o escoamento médio devem ser armazenadas em médias de no mínimo 10 minutos, para filtrar os efeitos da turbulência, a qual é comumente armazenada na forma de variância da intensidade do vento neste período. Define-se então o conceito de intensidade de turbulência ( $I$ ) como sendo a razão entre a variância da velocidade do vento ( $\sigma$ ) e seu valor médio ( $\bar{U}$ ), conforme a Equação 2.2.

$$I = \sigma / \bar{U} \quad (2.2)$$

Uma vez que o processo de dissipação de energia pela cascata de turbulência é contínuo, para que esta se mantenha é necessário que energia turbulenta seja gerada também continuamente, de modo que a variação energia cinética turbulenta (TKE) pode ser descrita pela Equação 2.3.

$$d(TKE)/dt = S + B - D + Tr \quad (2.3)$$

Onde  $S$  é termo de geração de turbulência pelo cisalhamento do vento,  $B$  a geração pela instabilidade térmica,  $D$  a dissipação pela viscosidade molecular e  $Tr$  o transporte de turbulência entre as regiões (Arya, 2001).

### 2.3. Equacionamento

Devido à grande oscilação temporal dos escalares no interior da CLA por causa da turbulência, na sua modelagem matemática adota-se a técnica de decomposição de Reynolds, que consiste na divisão das variáveis meteorológicas em componentes médias ( $\bar{\phi}$ ) e turbulentas ( $\phi'$ ) conforme a Equação 2.4, onde  $\phi$  é um escalar qualquer. Admite-se que as componentes médias obedecem às equações primitivas que governam os movimentos do fluido, compostas pela Equação da continuidade (2.5), Equação de conservação de movimento (Navier-Stokes) (2.6), ambas em sua forma vetorial, e pela conservação de energia (2.7). Outra aproximação adotada é a de Boussinesq, que consiste em considerar a densidade constante em todas as equações com exceção do termo de flutuação, permitindo assim a convecção. Ao final, a componente turbulenta introduz termos não-lineares e não-desprezíveis no interior da CLA, fazendo com que o número de variáveis seja maior do que o número de equações e impedindo uma solução explícita para o sistema (Holton, 2004). Este é conhecido como “problema do não-fechamento da turbulência” e é um dos problemas físicos ainda não compreendidos completamente pela ciência. Posteriormente será descrito como os modelos numéricos tratam para encontrar uma solução aproximada destas equações dentro da CLA, ou seja, fazer o “fechamento da turbulência”.

$$\phi = \bar{\phi} + \phi' \quad (2.4)$$

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{U}) = 0 \quad (2.5)$$

$$\frac{D\vec{U}}{Dt} = -2\vec{\Omega} \times \vec{U} - \frac{1}{\rho} \vec{\nabla} p + \vec{g} + \vec{F}_r \quad (2.6)$$

$$c_p \frac{D \ln \theta}{Dt} = \frac{J}{T} \quad (2.7)$$

Onde  $\vec{\Omega}$  representa a velocidade angular de rotação da terra,  $\vec{F}_r$  representa a força de atrito com superfície,  $J$  representa o aquecimento diabático e  $c_p$  o calor específico do ar a pressão constante

## 2.4. Fluxos Verticais e Instabilidade

Dentre as componentes turbulentas originadas pela decomposição de Reynolds, as mais importantes no estudo da dinâmica da CLA são as covariâncias, ou fluxos turbulentos. Estes podem ser divididos em fluxo de calor sensível ( $Q_h$ ), de calor latente ( $Q_l$ ) e de momento ( $Q_m$ ), conforme as equações 2.8, 2.9 e 2.10, respectivamente (Foken, 2008). Estes fluxos indicam de que forma se dá a transporte vertical de momento horizontal e calor e estão relacionados às características da CLA como altura, instabilidade e turbulência.

$$Q_h = \rho c_p \overline{\theta' w'} \quad (2.8)$$

$$Q_l = \rho \lambda \overline{q' w'} \quad (2.9)$$

$$Q_m = u^{*2} = \frac{\tau_0}{\rho} = \sqrt{(\overline{u' w'})^2 + (\overline{v' w'})^2} \quad (2.10)$$

Onde  $\lambda$  é o calor latente da água,  $u^*$  a velocidade de fricção,  $\tau_0$  a tensão de cisalhamento do escoamento, e as componentes turbulentas  $\theta'$  da temperatura potencial,  $q'$  da umidade específica e  $u'$ ,  $v'$ ,  $w'$  das componentes cartesianas do vento em x, y e z respectivamente.

Além destes, a estabilidade ( $s$ ) da CLS exerce forte influência nos perfis verticais de temperatura e momento, como será visto a seguir, e pode ser determinada conforme a Equação 2.11.

$$s = \left(\frac{g}{T_v}\right) \left(\frac{\partial \theta_v}{\partial z}\right) p \quad (2.11)$$

Onde  $\theta_v$  é a temperatura potencial virtual,  $T_v$  a temperatura virtual e  $g$  a aceleração da gravidade.

Existem alguns métodos para se avaliar a turbulência de um fluido estabelecendo relações para o número de Richardson ( $Ri$ ) ou para número de Reynolds ( $Re$ ). O primeiro é dado pela Equação 2.12 e representa uma relação entre a turbulência térmica e a mecânica, de modo que o escoamento pode ser considerado turbulento

para  $Ri < Ri_{crítico}$ , onde  $Ri_{crítico} = 0,25$ . Já o número de Reynolds dado pela Equação 2.13 representa uma medida entre a força de inércia e a força viscosa em fluido, sendo que o escoamento pode ser considerado turbulento para  $Re > Re_{crítico}$ . O valor de  $Re_{crítico}$  para a atmosfera varia entre  $10^3$  e  $10^4$  em função do relevo e rugosidade local.

$$Ri = \frac{\left(\frac{g}{\theta}\right) (\overline{w'\theta'})}{(\overline{u'w'}) \left(\frac{\partial \bar{u}}{\partial z}\right)} \quad (2.12)$$

$$Re = \frac{U \cdot D}{\mu} \quad (2.13)$$

Onde  $\mu$  é a viscosidade cinemática e D corresponde ao comprimento característico para a camada limite atmosférica.

## 2.5. Teoria da Similaridade e Perfis Verticais

Na busca por funções universais para os perfis verticais de momento e temperatura, foi desenvolvida a teoria da similaridade de Monin e Obhukov (1954), que consiste de uma maneira de tratar de forma adimensional algumas propriedades do escoamento. Desenvolvida a partir do Teorema Pi de Buckingham de análise dimensional, tal teoria é válida apenas para a CLS e pressupõe que os fluxos verticais são constantes ao longo desta camada. Seu intuito é determinar os perfis verticais para qualquer local e qualquer condição da CLS a partir dos valores dos fluxos de superfície. No entanto verificou-se que esta apresenta problemas para ventos fracos ou alta rugosidade superficial, sendo portanto não-universal (Foken, 2006).

Sua aplicação consiste em obter o comprimento de Obhukov ( $L$ ) dado pela Equação 2.14, tal que o coeficiente adimensional para o gradiente de *momentum* ( $c_m$ ), e o coeficiente adimensional para o gradiente térmico ( $c_h$ ), sejam funções  $\varphi_1$  e  $\varphi_2$  de  $z/L$  apenas, conforme as Equações 2.15 e 2.16. Tais funções são obtidas empiricamente (parametrizadas), e quando integradas fornecem expressões para o

perfil vertical de temperatura potencial e momento, enquanto o módulo de  $L$  fornece a espessura da camada onde os efeitos de fricção e cisalhamento são importantes na CLS. Temos ainda, que seu comprimento é uma medida da estabilidade, sendo que  $L > 0$  indica estabilidade e  $L < 0$  indica instabilidade.

$$L = -\frac{u_*^3}{k\left(\frac{g}{T_0}\right)\left(\frac{q}{c_p\rho}\right)} \quad (2.14)$$

$$c_m = \left(\frac{kz}{u_*}\right)\frac{\partial\bar{v}}{\partial z} = \varphi_1(z/L) \quad (2.15)$$

$$c_h = \left(\frac{z}{T_*}\right)\frac{\partial\bar{T}}{\partial z} = \varphi_2(z/L) \quad (2.16)$$

Sendo

$$T_* = -\frac{\overline{w'T'}}{u_*} \quad (2.17)$$

Onde,  $T'$  é a flutuação da temperatura,  $T_0$  a temperatura média,  $w'$  é a flutuação da velocidade vertical,  $k$  é a constante de Von Karman e  $c_p$  o calor específico a pressão constante.

A estabilidade pode ser avaliada também pelo parâmetro  $\zeta$ , descrito na Equação 2.18, presente na teoria da similaridade para a CLS, a qual classifica a condição de estabilidade atmosférica (Monin e Yaglom, 1971).

$$\zeta = z/L \quad (2.18)$$

A análise do sinal de  $\zeta$ , indica que:

- a)  $\zeta = 0$ , a camada é neutra;
- b)  $\zeta < 0$ , a camada é instável;
- c)  $\zeta > 0$ , a camada é estável;

Esta teoria é largamente empregada em modelos meteorológicos de mesoescala para a determinação dos perfis de temperatura e vento próximos à superfície. A Figura 2.4 ilustra a variação dos perfis verticais de vento em função da estabilidade da atmosfera na CLS.

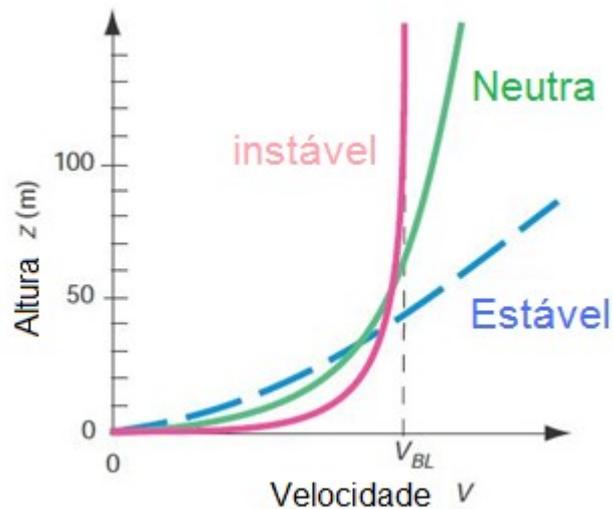


Figura 2.4 – Efeitos da estabilidade sobre os perfis de vento na CLS.

Fonte: Adaptado de Wallace & Hobbs (2001)

## 2.6. Influências Geográficas sobre o Vento

As teorias tradicionalmente empregadas na modelagem da camada limite atmosférica pressupõem que a superfície terrestre seja razoavelmente homogênea e que mudanças ocorram de maneira gradativa. Tais condições podem ser encontradas eventualmente em regiões extensas e pouco acidentadas como planícies ou sobre porções do oceano, mas não é o caso de grande parte das regiões de estudo (Arya, 2001). Fatores como orografia complexa, variações abruptas de rugosidade, regiões costeiras, grandes lagos e centros urbanos são comuns e alteram significativamente as características da camada limite, originando inclusive circulações específicas do local e exigindo que seus efeitos sejam tratados de maneira mais detalhada.

### 2.6.1. Circulações de Vale e Montanha

Dentre os efeitos induzidos pela orografia estão as circulações de vale e montanha. Durante o dia quando as encostas da montanha são aquecidas pelo Sol, o ar ascende ao longo das encostas até as cristas, originando o vento catabático. Por vezes, se há umidade suficiente observa-se a formação de cumulus ao longo da crista da montanha. Pela noite ocorre o inverso, o ar resfriado radiativamente tende a descer pelas encostas e se acumular nos fundos dos vales, gerando uma inversão mais intensa que pode ocasionar neblina. Em vales mais abruptos esta circulação das encostas pode ter um efeito adicional, ocasionando ventos subindo ao longo do vale pelo dia (ventos de vale) e descendo pela noite, denominados (ventos de montanha). Tais mecanismos estão ilustrados na Figura 2.5.

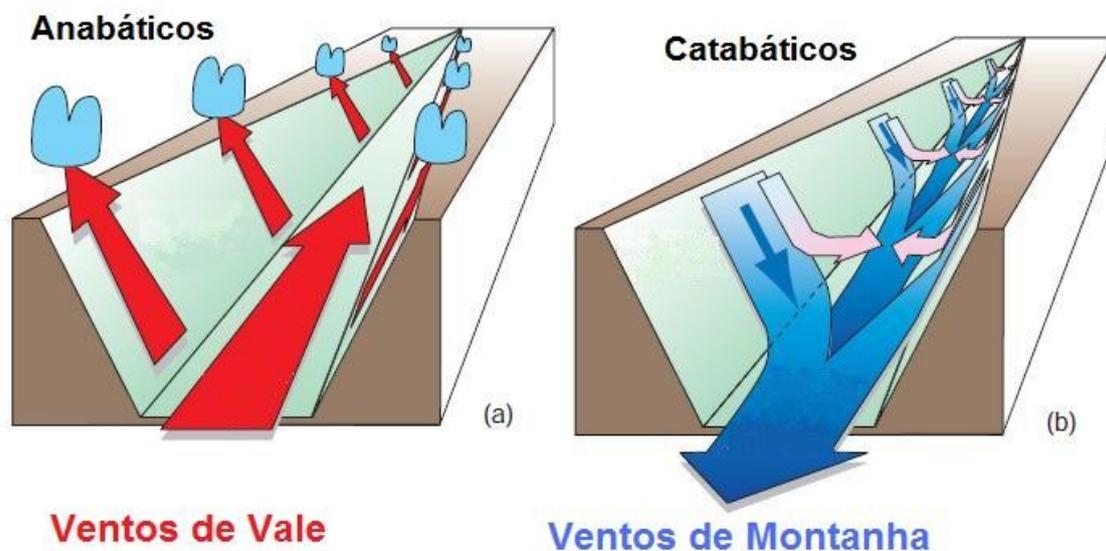


Figura 2.5 – Mecanismos de circulação de vale e montanha; ventos anabáticos (a) e ventos catabáticos (b).

Fonte: Adaptado de Wallace & Hobbs (2006).

### 2.6.2. Escoamento sobre Montanhas

Quando a intensidade dos ventos é maior devido à forçantes sinóticas, tais circulações locais são sobrepostas fazendo-se imperceptíveis. Nestes casos a orografia passa a agir como uma forçante dinâmica sobre os ventos, podendo induzir diversos fenômenos a jusante das montanhas como ondas de gravidade, esteiras de vórtices, nuvens lenticulares e tempestades de vento. Estes fenômenos

podem ter efeito por uma distância de até dez vezes a altura da montanha (Arya, 2001) e são fortemente influenciados pela intensidade do vento e pela estabilidade atmosférica. Nesse sentido, um parâmetro largamente utilizado na caracterização de escoamentos sobre montanhas é *número de Froude* ( $Fr$ ), que consiste na relação entre as forças de inércia e de gravidade, conforme Eq.2.19.

$$Fr = \frac{U}{N \cdot S} \quad (2.19)$$

Onde  $N$  é a frequência de Brunt-Vaisälá,  $U$  a velocidade do escoamento e  $S$  a escala vertical da montanha.

Conforme ilustra a Figura 2.6, para  $Fr < 1$  têm-se um escoamento pouco turbulento a jusante da montanha, enquanto para  $Fr > 1$  pode-se observar efeitos mais intensos como esteiras de vórtices, recirculações e tempestades de vento. No caso de  $Fr = 1$  pode haver ressonância entre a frequência de Brunt-Vaisälá e a frequência induzida pela montanha, causando uma amplificação da onda de montanha a jusante com a presença de rotores, grandes oscilações na intensidade do vento e possivelmente nuvens lenticulares.

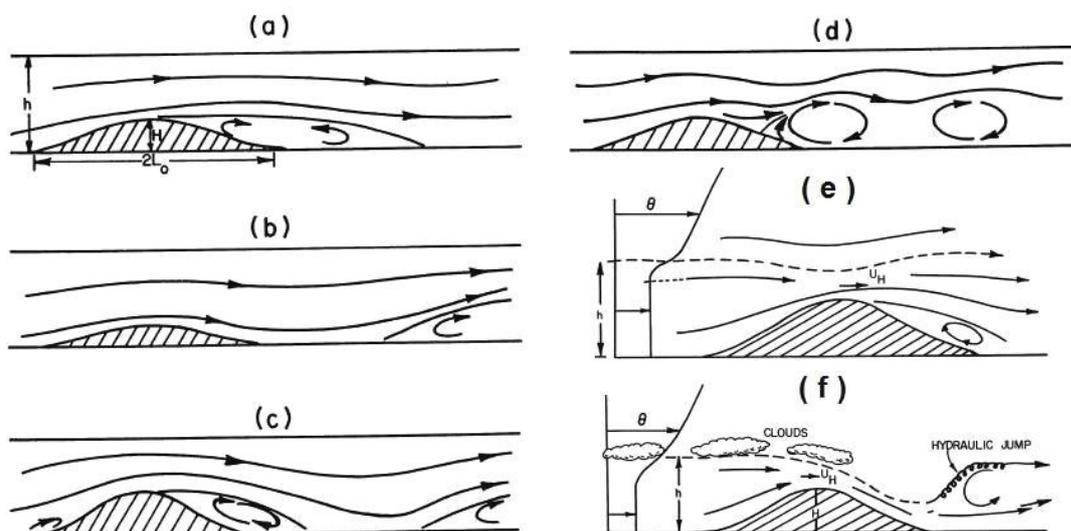


Figura 2.6 – Escoamento bidimensional sobre montanhas em atmosfera estável. Mecanismos de descolamento e formação de ondas de montanha e vórtices. Para grandes inclinações e  $Fr > 1$  (a); para inclinações pequenas e  $Fr < 1$  (b); para inclinações moderadas e  $Fr > 1$  (c); para inclinações pequenas e  $Fr < 1$  (d); para inversão próxima do topo (ZE) e  $Fr < 1$  (e); para inversão próxima do topo (ZE) e  $Fr > 1$  (f).

Fonte: Arya, (2001), adaptado de Hunt & Simpson (1982).

Outro efeito causado pelo relevo sobre os perfis verticais de vento é a aceleração do escoamento que ocorre no topo de morros e em canais devido ao efeito *Bernoulli*. Tal aceleração já foi medida em diversas campanhas, sendo função do formato da montanha, da estabilidade e da intensidade do vento a montante como ilustra a Figura 2.7, referente ao experimento de Askervein (Teunissen *et al.*, 1987). Em seu máximo, no topo, a velocidade pode chegar até a três vezes o valor do escoamento livre (Arya, 2001). Cabe ressaltar, entretanto que o efeito de uma cadeia de montanhas pode ser completamente distinto, uma vez que após algum tempo os perfis verticais podem adquirir uma nova configuração onde as acelerações deixam de ser significativas (Stull, 1988).

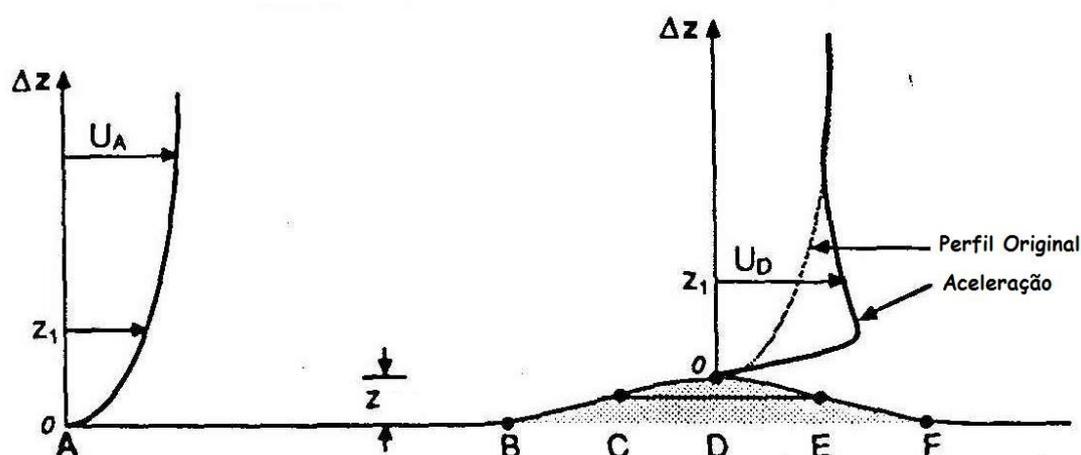


Figura 2.7 – Perfis verticais de velocidade ilustrando a aceleração do escoamento sobre topos de montanhas referentes ao experimento de Askervein.  
Fonte: Adaptado de Stull (1988)

### 2.6.3. Camada Limite Interna

As mudanças abruptas de rugosidade e temperatura exercem forte influência sobre os perfis verticais e demandam um maior detalhamento. Como a superfície terrestre não é homogênea, na medida em que o vento passa por regiões com características distintas, seu perfil de velocidade e temperatura adquire gradativamente as propriedades da superfície adjacente. Este processo culmina com o desenvolvimento de uma *Camada Limite Interna* (CLI), que cresce na medida em que o escoamento se afasta da borda de transição, como ilustra a Figura 2.8.

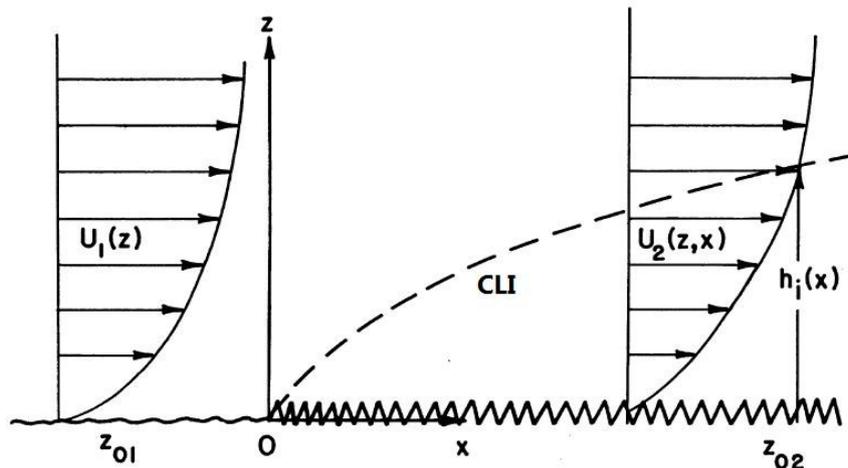


Figura 2.8 – Esquema de desenvolvimento de camada limite interna (CLI) na transição entre superfícies com características distintas.  
 Fonte: Adaptado de Arya, (2001).

No caso de uma variação na rugosidade, o perfil logarítmico do vento passa a ser então uma composição entre os perfis anterior e posterior à transição, onde a altura  $z_i$ , referente à CLI, passa a ser uma função não só da distância da borda, mas também das rugosidades anterior e posterior e da estabilidade da atmosfera (Stull, 1988). Como estas mudanças podem ser freqüentes em algumas regiões, o perfil vertical pode se tornar complexo, formado por uma sobreposição de CLIs, conforme ilustrado na Figura 2.9.

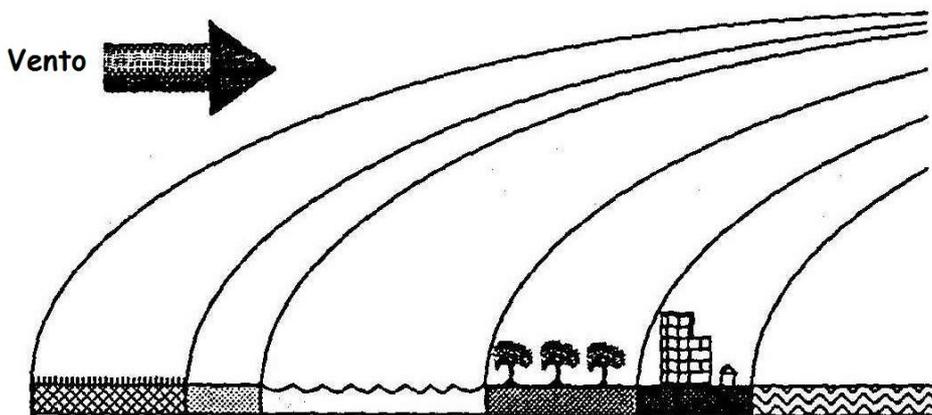


Figura 2.9 – Sobreposição de CLIs na composição do perfil vertical de vento e temperatura em regiões heterogêneas.  
 Fonte: Stull (1988).

Têm-se ainda que a transição de uma região de baixa rugosidade para regiões de vegetação mais alta e densa causa uma absorção ainda maior de momento devido

à circulação abaixo do dossel. Este efeito pode ser contabilizado pela elevação ( $d$ ) do plano aerodinâmico efetivo na equação do perfil logarítmico do vento (Eq. 2.20), de modo a compensar o momento absorvido abaixo do dossel.

$$U = \frac{u^*}{k} \cdot \ln\left(\frac{Z-d}{Z_0}\right) \quad (2.20)$$

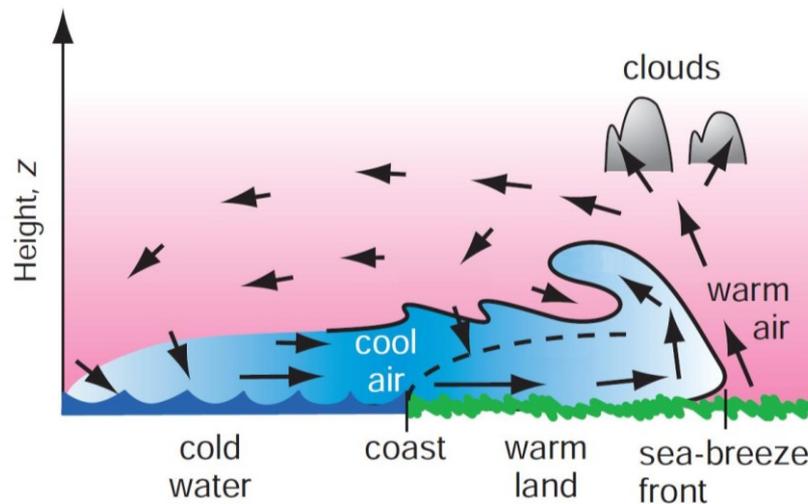
Onde as variáveis foram definidas anteriormente na Eq. 2.1

De maneira análoga, quando o escoamento passa por uma fronteira de transição brusca de temperatura, inicia-se uma CLI térmica, que desenvolve-se escoamento abaixo alterando a estabilidade da atmosfera. Um caso típico se dá quando há advecção continental fria sobre lagos ou oceanos mais quentes. Nestes casos forma-se uma camada interna instável (convectiva) que pode causar tempestades severas a jusante. Outro exemplo interessante ocorre nas bordas das grandes cidades, onde as diferenças de albedo, de disponibilidade de umidade e de geração de calor e poluentes entre a zona rural e urbana ocasionam as *Ilhas de Calor* (Arya, 2001). O escoamento ao adentrar a cidade tende a se tornar mais instável, intensificando as tempestades no local. O calor armazenado pela cidade é tão grande que pode sustentar uma camada limite convectiva inclusive durante a noite (Stull, 1988).

#### **2.6.4. Circulações de Brisa**

O fenômeno que possivelmente apresenta o efeito mais substancial nos ventos de uma região são as circulações de brisa, marítima ou lacustre. Em regiões tropicais, ou no verão em latitudes médias, a intensa insolação diária gera grandes contrastes de temperatura entre as superfícies aquáticas e terrestres, causada basicamente pela baixa capacidade térmica do solo. Durante o dia a superfície terrestre tende a ficar mais quente que a água, aquecendo a camada de ar imediatamente acima, o que causa a inclinação das isóbaras próximas à superfície. Este gradiente de pressão produz uma circulação local da água para o continente em baixos níveis, denominada *brisa marítima* (ou lacustre), que é compensada por um movimento em sentido oposto em níveis mais elevados. Durante a noite a dinâmica se inverte originando a *brisa terrestre*. Na brisa marítima, por se tratar de uma advecção mais fria sobre uma região quente, podem ocorrer *frentes de brisa*, similares às frentes

frias, como mostra a Figura 2.10. Nestes casos se há umidade suficiente sobre o continente e instabilidade atmosférica, estas frentes evoluem para tempestades (Wallace & Hobbs, 2006).



Figuras 2.10 – Componentes da circulação de brisa oceânica ou lacustre.  
Fonte: Adaptado de Wallace & Hobbs (2006)

Apesar de conhecidos, alguns dos efeitos geográficos acima descritos são de difícil modelagem, pois não são compreendidos totalmente. Dependem por vezes de uma combinação de fatores como velocidade e direção do vento, cisalhamento vertical, estabilidade, rugosidade e forma do relevo, além de se apresentarem sobrepostos e possivelmente ocorrerem em escalas inferiores à capacidade de representação dos modelos de PNT. Todavia, é inegável que uma melhor compreensão de seus efeitos pode ser útil na interpretação de dados observacionais assim como na discussão de possíveis desvios em relação a resultados esperados.

### 3 A PREVISÃO NUMÉRICA DE VENTO EM SUPERFÍCIE

#### 3.1. O Modelo Eta

O modelo de previsão numérica de tempo Eta é um modelo de área limitada desenvolvido inicialmente na Iugoslávia, na Universidade de Belgrado na década de 1970, entrando em operação posteriormente nos Estados Unidos, no *National Center for Environmental Prediction* – NCEP na década de 1980 após algumas modificações (Black, 1994).

Este modelo utiliza grade horizontal do tipo E de Arakawa (Arakawa e Lamb, 1977), conforme descrito na Figura 3.1, onde v representa os pontos onde são calculadas as componentes do vento horizontal e h são os pontos de massa, onde se calcula as demais variáveis do modelo como temperatura, umidade, velocidade vertical, dentre outras. A resolução horizontal é dada pela distância d entre dois pontos h ou v. Esta grade utiliza coordenadas esféricas, tendo o centro de seu domínio transladado para o plano do equador para minimizar a convergência entre os meridianos em grandes latitudes (Chou, 1996).

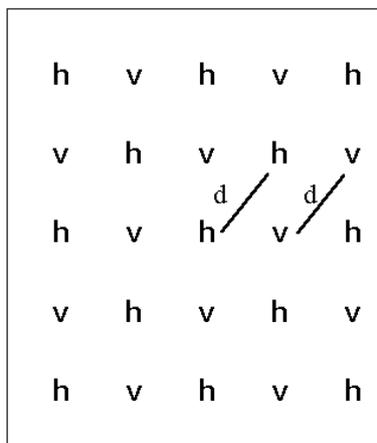


Figura 3.1 – Representação do espaçamento da grade E de Arakawa.

Fonte: Chou, (1996); adaptado de Arakawa e Lamb, (1977)

A principal característica deste modelo é a introdução da coordenada vertical eta ( $\eta$ ) definida pela Equação 3.1, em substituição à coordenada sigma ( $\sigma$ ) visando a redução dos erros nas derivadas horizontais sobre relevos montanhosos (Messinger, 1984).

$$\eta = \left( \frac{p - p_T}{p_{sfc} - p_T} \right) \cdot \left( \frac{p_{ref}(z_{sfc}) - p_T}{p_{ref}(0) - p_T} \right) \quad (3.1)$$

Nesta equação,  $p$  é a pressão local,  $p_{sfc}$  a pressão na superfície,  $p_{ref}$  a pressão de referência para uma determinada altura e  $p_T$  a pressão no topo do domínio.

Desta maneira o terreno passa a ser representado sob a forma de degraus discretos, onde o topo coincide com a interface do relevo. O nível de cada degrau é determinado pelo método das silhuetas, descrito em mais detalhes por Rozante, (2001), que pondera a área da grade na determinação de um nível representativo, que é então ajustado ao nível vertical discreto mais próximo. Deste modo os pontos de massa, temperatura e vento horizontal são dispostos no meio das camadas enquanto a velocidade vertical, a energia cinética turbulenta e o geopotencial são localizados nas bases da camada. A Figura 3.2 ilustra essa discretização.

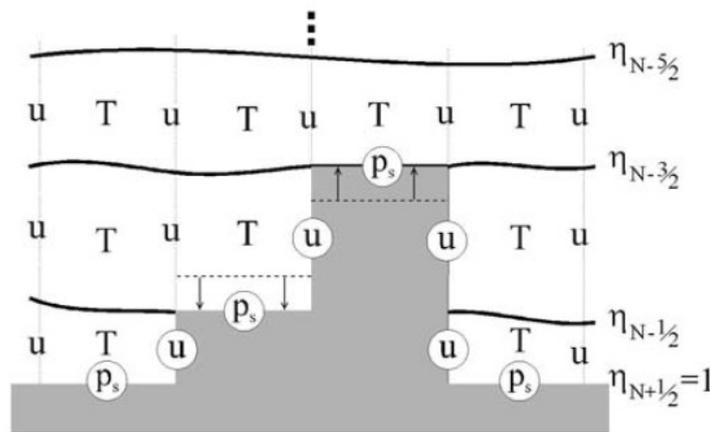


Figura 3.2 – Representação da topografia pela coordenada Eta

Fonte: Adaptado de Ničković et al. (1998).

O modelo Eta vêm sendo utilizado operacionalmente pelo CPTEC desde 1996, com o intuito de fornecer previsões de curto prazo para o Brasil. Seu domínio engloba praticamente toda a América do Sul e atualmente as condições iniciais são fornecidas pelas análises do GFS/NOAA e de contorno pelo Modelo de Circulação Global do CPTEC T213L42 (CPTEC, 2010). A temperatura da superfície do mar assimilada é a observada na semana anterior ao processamento, considerada constante no intervalo de integração. A umidade do solo é obtida a partir da climatologia anual. Suas variáveis prognósticas são temperatura do ar,

componentes zonal e meridional do vento, umidade específica, pressão à superfície e energia cinética turbulenta e hidrometeoros. A partir destas são derivadas as demais variáveis de saída do modelo.

Desde sua operacionalização o modelo Eta/CPTEC vêm sendo modificado para uma melhor performance nas previsões sobre o Brasil. Sua resolução horizontal que era inicialmente de 40 km passou para 20 km em 2006 e deve ser melhorada em breve. Atualmente o modelo roda duas vezes ao dia às 00UTC e 12UTC onde são disponibilizadas saídas a cada 6 horas para um horizonte de previsão de 7 dias. Estas saídas constituem um total de 58 variáveis sendo 9 dispostas em 19 níveis de pressão.

### 3.2. A Modelagem do Vento em Superfície

Com relação ao pacote físico, o Eta atualmente incorpora esquema de superfície Noah como 13 classes de vegetação e 4 camadas no solo (Mitchell, 2001) e fechamento de turbulência de Mellor-Yamada 2.5 (Mellor e Yamada, 1982). Na camada superficial é adotada a teoria de similaridade de Monin-Obukhov (Monin e Obukhov, 1954) com funções de estabilidade de Paulson (Paulson, 1970). Esta aproximação consiste de um conjunto de expressões analíticas que especificam o gradiente adimensional da velocidade do vento e o gradiente de temperatura potencial a partir das funções de estabilidade para o *momentum* ( $\Psi_m$ ) e para o calor ( $\Psi_h$ ). Estas funções foram definidas conforme as Equações 3.2 e 3.3 (Sun e Mahrt, 1995) onde os coeficientes de troca de *momento* e calor são definidos a partir das Equações 3.5 e 3.6. É importante ressaltar que considera-se a existência de dois comprimentos de rugosidade os quais representam a absorção de momento ( $Z_{0m}$ ) e calor ( $Z_{0h}$ ) pela superfície (Equação 3.7) (Chen *et al.*, 1997)

$$\Psi_m = \begin{cases} -5\zeta & , \quad 0 < \zeta < 1 \\ 2\ln\left(\frac{1+x}{2}\right) + \ln\left(\frac{1+x^2}{2}\right) - 2\tan^{-1}(x) + \frac{\pi}{2} & , \quad -5 < \zeta < 0 \end{cases} \quad (3.2)$$

$$\Psi_h = \begin{cases} -5\zeta & , \quad 0 < \zeta < 1 \\ 2\ln\left(\frac{1+x^2}{2}\right) & , \quad -5 < \zeta < 0 \end{cases} \quad (3.3)$$

$$x = (1 - \gamma\zeta)^{1/4} \quad \text{onde } \gamma = 16 \quad (3.4)$$

$$c_d = \frac{k^2/R}{\left[ \ln\left(\frac{Z}{Z_{0m}}\right) - \psi_m\left(\frac{Z}{L}\right) + \psi_m\left(\frac{Z_{0m}}{L}\right) \right]^2} \quad (3.5)$$

$$c_h = \frac{k^2/R}{\left[ \ln\left(\frac{Z}{Z_{0m}}\right) - \psi_m\left(\frac{Z}{L}\right) + \psi_m\left(\frac{Z_{0m}}{L}\right) \right] \cdot \left[ \ln\left(\frac{Z}{Z_{0h}}\right) - \psi_h\left(\frac{Z}{L}\right) + \psi_h\left(\frac{Z_{0h}}{L}\right) \right]} \quad (3.6)$$

$$Z_{0h} = \frac{D_h}{ku_*} \quad (3.7)$$

Onde:

$$D_h = \frac{k_d}{\rho_a c_{p,d}} \quad (3.8)$$

$D_h$  é o coeficiente de difusão térmica molecular,  $k_d$  é a condutividade térmica do ar seco e  $c_{p,d}$  é o calor específico do ar seco a pressão constante,  $z$  é altura de referência,  $R$  é a razão entre os coeficientes de momento e calor no limite neutro e  $L$  é o comprimento de Monin-Obukhov definido pela Equação 2.14.

### 3.3. Aplicações de Modelagem Numérica do Vento

No Brasil, pode-se destacar o trabalho de Lyra (2008), que efetuou testes de sensibilidade no modelo Eta/CPTEC com intuito de se avaliar o ganho de acurácia das previsões de vento. Foram avaliadas mudanças na resoluções horizontal e vertical do modelo, na relação dos comprimentos de rugosidade de calor e momento, nas funções de estabilidade de Paulson, no esquema de parametrização da CLS, além da influência da altura da CLA e da substituição da cobertura de solo existente. Os melhores resultados levaram a erro médio (ME) de -0,98 m/s, e erro quadrático médio (RMSE) de 2,08 m/s e coeficientes de correlação de 0.79 a 10 m do solo. Os resultados mostraram pequena sensibilidade aos parâmetros alterados com exceção da altura da CLA, que gerou diferença significativa entre os ventos no oceano e no continente. A baixa resolução das informações de temperatura da superfície do mar e de umidade pode ter limitado a sensibilidade dos experimentos.

Utilizando um modelo diferente, o RAMS, De Maria *et al.*, (2008) também realizou testes de sensibilidade para previsões de vento na costa do Ceará. Seus resultados mostraram ganhos para a resolução horizontal de 1km e fechamento de turbulência de Smagorinsky, onde a comparação com medidas de velocidade a 10 m levou a ME em torno de -1,6 m/s, RMSE de 2,03 m/s e coeficientes de correlação de até 0,86.

Previsões de vento a 50m com o modelo WRF foram feitas para a Romênia por Dica *et al* (2009). utilizando uma grade de 3.3 km. A comparação entre as previsões e as observações para 9 estações levaram a RMSE entre 1,87 m/s e 2,23 m/s.

Lazic *et al.* (2009) utilizou um refinamento com grades de 3,5 km para avaliação das previsões de vento do modelo Eta sobre uma ilha da Suécia com resolução de três horas. Para uma série de quatro anos de dados onde apenas o verão foi considerado, as comparações com as observações levaram a ME de 0,48 m/s, RMSE de 1,38 m/s e coeficientes de correlação de 0,79 a 10m do solo.

Também utilizando uma versão do modelo Eta, além do modelo RAMS, Louka *et al.* (2006) realizou previsões de vento de alta resolução sobre a ilha de Creta, na Grécia. Após tratar as saídas utilizando um filtro de Kalman obteve RMSE em torno de 2,04 m/s. O autor comenta que o aumento da resolução além de 6km nem sempre leva a uma melhora nas previsões.

### **3.4. Limitações da Modelagem Numérica do Vento**

Os modelos de PNT, apesar de todo o avanço alcançado nos últimos trinta anos, ainda estão sujeitos a diversas limitações, as quais possuem origens distintas. O uso adequado das previsões geradas demanda a compreensão destas limitações, possibilitando inferir prováveis fontes de erro e avaliar a incerteza associada a seus resultados. A seguir serão discutidas algumas das principais simplificações e pressupostos embutidos na formulação dos modelos, com o intuito de enriquecer a discussão dos resultados deste trabalho.

A primeira etapa e talvez a mais relevante com respeito à introdução de erros na modelagem é a determinação da condição inicial. Isto é feito através de um processo de assimilação de dados, o qual consiste de um método de diagnóstico

dos campos iniciais a partir de uma rede de observações, que inclui estações de superfície, sondagens, dados de aeronaves e de sensoriamento remoto por satélites. No entanto, além dos dados serem insuficientes para cobrir toda a variabilidade espacial da atmosfera, tais observações apresentam erros inerentes às medições, o que faz com que existam desvios mesmo nos campos de grande escala. Os métodos de assimilação têm por objetivo a remoção de erros sistemáticos e a redução do impacto dos erros aleatórios sobre as previsões do modelo, mas dada a dinâmica não-linear da atmosfera os erros tendem a se amplificar com o tempo, impondo um limite de previsibilidade determinística aos modelos que atualmente está em torno de 2 semanas (Wallace & Hobbs, 2006).

Da mesma forma, como os modelos de mesoescala geralmente possuem área limitada, é necessário que se forneça condições de contorno ao longo das integrações. Estas condições são fornecidas por modelos globais, que estão sujeitos aos mesmos erros citados anteriormente, além de poderem eventualmente induzir erros, uma vez que as equações diagnósticas podem diferir entre o modelo global e o de mesoescala. Para contornar este problema a inicialização dos modelos adota métodos para reduzir desvios entre os condições de contorno e os campos diagnosticados internamente o que inclui técnicas de *relaxação* e *'nudging'*.

Os movimentos convectivos na atmosfera sempre se colocaram como um desafio particular para a modelagem numérica, pois as análises de escala podem diferir bastante do escoamento médio nestas regiões. Nestes casos, torna-se conveniente adotar aproximações específicas na equação da continuidade (Eq.2.5) para cada fenômeno modelado, o que, em movimentos convectivos consiste na separação entre convecção rasa e convecção profunda.

Para a convecção profunda, que é o caso mais geral, admite-se que: i) a profundidade das circulações atmosféricas possuem escala vertical próxima da escala de variação de densidade na atmosfera ( $H_\alpha \cong 8 \text{ km}$ ); ii) a advecção e divergência local de massa estão em equilíbrio e não causam variações significativas de densidade; iii) a velocidade de grupo das ondas são da mesma ordem que a dos escoamentos atmosféricos. Estes requisitos têm como consequência básica a eliminação das ondas sonoras como possíveis soluções,

aumentando a estabilidade dos métodos numéricos de integração. Tal abordagem é também conhecida como *aproximação anelástica* (Pielke, 2001)

Na convecção rasa, mais condições são necessárias, dentre elas: iv) a escala vertical de variações de densidade dentro da CLA deve ser bem menor que a escala de variação de densidade na atmosfera ( $L_x \ll H_\alpha$ ); v) as perturbações de densidade, temperatura e pressão são muito menores que seus valores médios. vi) a atmosfera local deve ser, em média, estaticamente estável; vii) os movimentos verticais devem ser limitados pelo termo de flutuação; (Mahrt, 1986). Esta aproximação ignora não apenas as ondas sonoras, mas também as variações espaciais de densidade, sendo também chamada de *aproximação incompressível*.

Com relação à equação de conservação da energia termodinâmica (Eq. 2.7), uma análise de escala não é possível devido à complexidade dos termos fonte de aquecimento (J), fruto de processos diabáticos e radiativos, que são normalmente parametrizados nos modelos (Pielke, 2001).

Já na equação de conservação do movimento vertical, possivelmente a mais delicada das aproximações consiste em desprezar o termo de aceleração, de forma que a distribuição de pressão na vertical seja dada pelo *equilíbrio hidrostático*. Esta formulação, apesar de eficaz para movimentos de escala sinótica, se degrada na medida em que os fenômenos de interesse possuem uma menor escala de circulação ( $L_x$ ). Pode-se estimar que uma condição razoável para sua validade é dada por ( $L_x \leq H_\alpha$ ), ou seja, pela análise de escala, para fenômenos com escala inferior a 8 km a aproximação não-hidrostática passa a ser mais adequada (Pielke, 2001). Considerando que o espaçamento de grade do modelo deve ser pelo menos a metade da escala da circulação de interesse, isto corresponderia a grades inferiores a 4 km. Este limite é bastante controverso e depende também de outros fatores. Alguns resultados sugerem que a aproximação hidrostática se degrada na medida em que: a instabilidade aumenta, os fluxos verticais aumentam; e o vento se intensifica (Pielke, 2001).

Outra aproximação comum em modelos meteorológicos consiste em considerar que a perturbação na densidade local é muito pequena quando comparada ao seu valor médio ( $\rho'/\bar{\rho} \ll 1$ ). Com isto é possível desprezar este termo nas componentes

inerciais da equação de conservação do movimento vertical e mantê-lo apenas no termo de flutuação, o que é conhecido como *aproximação de Boussinesq*. Tal abordagem pressupõe que o escoamento é incompressível e não-divergente no plano horizontal, portanto as variações de densidade são consequência apenas das perturbações de temperatura e pressão, conforme a equação dos gases ideais. Ou seja, um pré-requisito para a validade desta aproximação é que as condições para convecção rasa sejam satisfeitas (Stull, 1988).

No equacionamento da conservação do movimento horizontal, poucas simplificações adicionais podem ser feitas. Uma delas consiste em desprezar a advecção vertical de escalares pela velocidade vertical média (ascendência/subsidência). Apesar de esta velocidade vertical ser pelo menos uma ordem de grandeza inferior ao vento horizontal, os gradientes verticais são mais intensos, o que faz com que os termos possuam aproximadamente a mesma ordem de grandeza. Nos fluxos horizontais a simplificação mais comum consiste em adotar homogeneidade horizontal, desprezando a advecção por fluxos turbulentos, que se torna mais significativa com o aumento da resolução espacial dos modelos (Stull, 1988).

Conforme observado anteriormente, o sistema de equações da dinâmica atmosférica (Eq. 2.5 - Eq.2.7) não possui solução explícita e apresenta mais incógnitas do que equações, tornando-se indeterminado, o que demanda a adoção de aproximações para algumas variáveis. Estas aproximações, por representarem fenômenos físicos, podem ser obtidas como funções empíricas de outras variáveis, ao que se dá o nome de *parametrização*. Esta técnica se faz necessária porque os fenômenos abordados são muito complexos para serem equacionados, e/ou porque ocorrem numa escala de tempo e espaço inferior à resolução do modelo. Dentre as variáveis comumente parametrizadas estão o aquecimento diabático e o aquecimento radiativo, na equação da energia termodinâmica, e os fluxos turbulentos de momento e temperatura nas equações de conservação do movimento. Tais parametrizações possuem uma escala espacial e temporal de validade para a qual foram desenvolvidas, portanto é de se esperar que na medida em que a resolução espacial aumenta, e conseqüentemente o intervalo de

integração diminui, a variância das grandezas sub-grade aumente, o que gradativamente reduz a acurácia das parametrizações.

Na CLS, que é a porção de atmosfera de maior interesse para a previsão de ventos, a parametrização adotada foi tratada em detalhes no item 3.2 deste capítulo e está sujeita à degradação por fatores como ventos fracos, topografia complexa ou alta rugosidade (Foken, 2006), além dos já citados anteriormente. Outras grandezas diretamente relacionadas à dinâmica da CLA e que são parametrizadas nos modelos de PNT são: a altura da CLC, a altura da CLE, os fluxos verticais de calor e movimento na zona de entranhamento e a altura da camada limite interna. Conforme exposto acima, a CLA é porção da atmosfera mais sensível às parametrizações, as quais possuem uma complexidade considerável dada a grande diversidade de características do conjunto superfície – atmosfera. Efeitos não-lineares são comuns entre estas variáveis, o que faz com que todas as parametrizações disponíveis atualmente nos modelos de PNT considerem homogeneidade horizontal na escala sub-grade. Esta simplificação talvez seja a principal fonte de erro nas grandezas próximas à superfície e mostra que esta é uma área onde ainda há muito a se avançar nas pesquisas (Pielke,2001).

Por fim, devido a restrições inerentes aos métodos numéricos de integração empregados nos modelos de PNT, os termos de advecção, gradiente de pressão e difusão horizontal só podem ser bem representados para escalas superiores a quatro vezes o espaçamento de grade do modelo. Esta deve ser também a escala espacial mínima de validade das parametrizações embutidas nos modelos, caso contrário a solução numérica pode introduzir erros significativos (Pielke,2001). Particularmente em relação ao modelo Eta, alguns estudos de caso indicam que o formato de degrau adotado para representação do relevo induz erros consideráveis no vento a baixos níveis a jusante de montanhas. Comparações com modelos que utilizam coordenada sigma (fiel ao relevo) suportam esta hipótese. Isto ocorre porque o Eta gera um descolamento fictício do fluxo após a montanha, e ainda falha em reproduzir a intensidade das ondas induzidas pela mesma. Foi mostrado que esta deficiência está presente já em grades de 32 km de resolução e tende a se amplificar na medida em que a grade é refinada (Gallus, 2000).

Avaliando em conjunto as possíveis fontes de erro citadas acima, pode-se ter uma idéia melhor das limitações dos modelos de PNT. A magnitude aproximada de cada uma das incertezas é de difícil quantificação, pois depende de uma série de fatores como resolução do modelo, fenômenos de interesse e aspectos geográficos da região, o que demanda testes que sensibilidade sejam efetuados para cada caso. Para auxiliar na visualização destas potenciais fontes de erro a Figura 33 ilustra-os em um fluxograma aproximado do processo de modelagem.

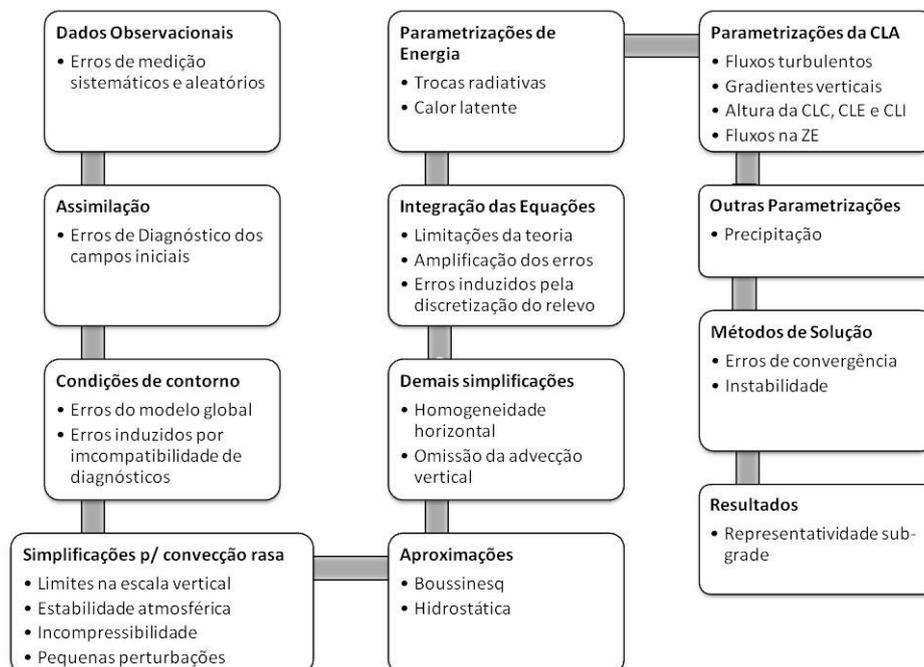


Figura 3.3 – Fluxograma descrevendo as fontes potenciais de erro nos modelos de PNT.

#### **4 REDES NEURAIS ARTIFICIAIS**

As Redes Neurais Artificiais (RNAs) surgiram como uma ferramenta para se extrair relações complexas de um conjunto de variáveis, possibilitando ir além das regressões lineares múltiplas. Trata-se de uma técnica de computação não-algorítmica capaz de modelar ou reconhecer padrões a partir de um conjunto de dados apresentados, possuindo inclusive capacidade de generalização para domínios até então desconhecidos (Braga, 1998).

Em resumo, tal método pode ser compreendido como um caso particular de regressão não-linear, sendo que sua aplicação segue uma metodologia onde as variáveis são organizadas e combinadas obedecendo um padrão pré-estabelecido de redes, conhecido como redes neurais. Posteriormente um processo iterativo ajusta os coeficientes desta rede, de forma a minimizar uma determinada função de custo, geralmente associada ao erro do modelo. A este processo dá-se o nome de treinamento. Ao final, o modelo obtido é uma função do valor inicial dos coeficientes e do algoritmo adotado para o treinamento, existindo portanto um grau de aleatoriedade que permite, dependendo da complexidade do fenômeno, que se chegue a soluções distintas para um mesmo problema.

Se corretamente empregada, esta técnica possui a capacidade de representar fenômenos físicos complexos e por isso ultimamente vêm sendo largamente utilizada nas geociências, propiciando grandes avanços na área de modelagem (Mendes e Marengo, 2009).

Do ponto de vista de sua arquitetura, as RNAs são compostas por um conjunto de nodos ou neurônios dispostos em camadas, os quais aplicam funções matemáticas geralmente não-lineares sobre os dados de entrada. Estas entradas são ponderadas por pesos, os quais são ajustados de modo a minimizar o erro cometido pela rede no cálculo do resultado final. Seu desenvolvimento passa necessariamente por uma fase de treinamento ou aprendizagem e outra de validação, podendo ser considerado neste sentido uma técnica de ajuste estatístico. Vários fatores influenciam no desempenho da RNA dentre estes a seleção das entradas, o tipo de neurônio, o tipo de função de ativação, o número de camadas e do tipo de algoritmo de treinamento seriam os mais importantes. Testes com

diferentes configurações são geralmente empregados para a definição da melhor arquitetura para um determinado problema (Haykin, 1994).

#### 4.1. Histórico

O conceito de redes neurais artificiais surgiu através de McCulloch e Pitts (1943), onde foi descrito um modelo simples de neurônio com  $n$  entradas  $X_i$  ajustadas por pesos  $W_i$  onde o neurônio computa a soma das entradas  $X_i W_i$  e compara a um valor limiar  $\theta$  para avaliar se a sinapse é ou não ativada, conforme Figura 4.1. Sua saída era, portanto, binária e este primeiro modelo apresentou diversas restrições em sua utilização. Posteriormente vieram os trabalhos de Hebb (1949), demonstrando a capacidade de adaptação das redes neurais e Rosenblatt (1958) que introduziu o modelo de neurônio *perceptron de uma camada*, dotado de pesos ajustáveis além de um algoritmo de treinamento, os quais tiveram grande contribuição para o progresso desta ciência.

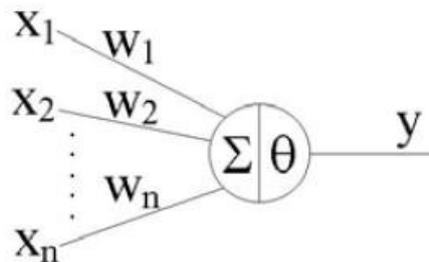


Figura 4.1 – Modelo de neurônio de McCulloch e Pitts.

Fonte: McCulloch e Pitts (1943).

Entretanto o estudo publicado por Minsky e Papert (1969) teve grande repercussão no meio científico, servindo como fator desestimulante na evolução das redes neurais. Eles demonstraram as limitações associadas ao modelo de *perceptrons* dando ênfase à incapacidade de se obter soluções para um espaço amostral onde as classes não são linearmente separáveis. Somente a partir de meados da década de 1980 com os trabalhos de Hopfield (1982) e Rumelhart *et al.* (1986), demonstrando as aplicações das RNAs na área da física e com o desenvolvimento do algoritmo “*backpropagation*” é que as pesquisas foram retomadas consistentemente, sendo hoje empregada nos mais diversos campos da ciência.

## 4.2. Arquitetura

Conforme mencionado anteriormente, a definição da arquitetura da RNA afeta substancialmente seu desempenho, pois restringe o tipo de problema que pode ser tratado pela rede. Segundo Cybenko (1989), a capacidade de solucionar problemas não linearmente separáveis por uma rede de perceptrons só é possível através da utilização de pelo menos uma camada intermediária ou oculta, permitindo a aproximação de qualquer função contínua. Com duas ou mais camadas ocultas é possível aproximar qualquer função. Por este motivo será abordado com mais ênfase as redes do tipo *Multi-Layer Perceptrons – MLP* uma vez que esta é a arquitetura mais empregada na resolução de problemas físicos, inclusive na meteorologia (Mendes e Marengo 2009, Guarnieiri, 2006; Martins *et al.*, 2011). A Figura 4.2 ilustra uma rede MLP com uma camada oculta.

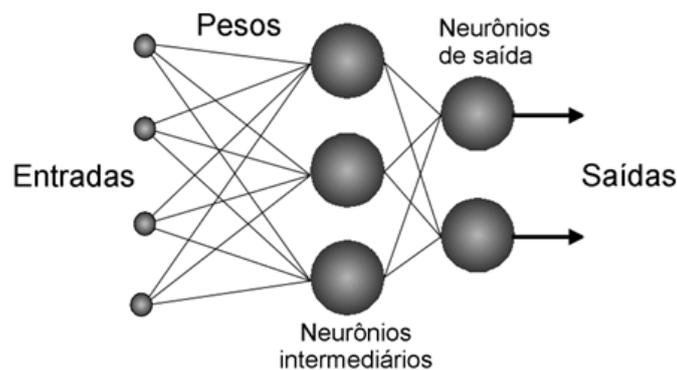


Figura 4.2 – Representação de uma rede tipo MLP.

As redes neurais podem possuir retroalimentação, ou seja, sua saída é utilizada como entrada para a própria rede, sendo estas denominadas *recorrentes*. As redes que não possuem este mecanismo são denominadas *não-recorrentes*. Rede recorrentes são indicadas quando há a necessidade de processamento temporal entre as variáveis de entrada.

A função de ativação representa o tipo de comparação que é feito com as entradas do neurônio para se determinar sua saída. No caso do neurônio de McCulloch e Pitts foi utilizada a função degrau, onde a saída é binária dependendo do valor das entradas. Entretanto existem outros tipos de funções passíveis de ser implementadas nos neurônios, como ilustra a Figura 4.3.

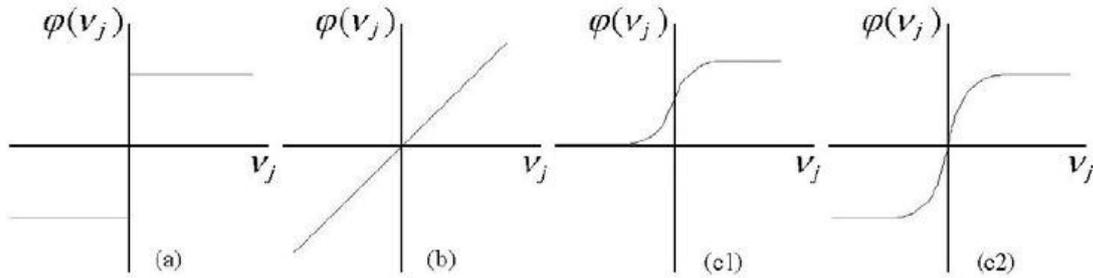


Figura 4.3 – Funções de ativação; a) degrau; b) linear; c) sigmoidal: 1) logística; 2) tangente hiperbólica.

Fonte: Braga, (1998)

As funções sigmoidais são as mais utilizadas em redes do tipo MLP em razão do uso da técnica de treinamento tipo “*backpropagation*” proposta por Rumelhart *et al.* (1986) na qual o ajuste dos pesos se dá pelo gradiente descendente de erro. Para utilização deste algoritmo, é necessário o uso de funções de ativação contínuas, diferenciáveis e de preferência não-decrescentes e não-lineares para permitir que as camadas ocultas interpretem as não-linearidades na rede (Braga, 1998). Dentre as funções sigmoidais mais utilizadas pode-se destacar a Função Logística e a Função Tangente Hiperbólica, descritas pelas Equações 4.1 e 4.2, respectivamente.

$$\varphi(v_j) = \frac{1}{1 + \exp(-\alpha v_j)} \quad (4.1)$$

$$\varphi(v_j) = \tanh\left(\frac{v_j}{2}\right) = \frac{1 - \exp(-v_j)}{1 + \exp(-v_j)} \quad (4.2)$$

Os neurônios que constituem uma rede MLP atualmente possuem uma arquitetura que incorpora as principais funcionalidades desenvolvidas, podendo gerar qualquer saída linear, conforme ilustra a Figura 4.4. Sua saída  $y$  é definida pela Equação 4.3, onde  $\varphi$  é a função de ativação,  $u$  é a somatório dos pesos ponderados por cada sinapse e  $b_j$  é definido como o limiar, ou viés do neurônio. A soma da saída da função de ativação com o viés define o nível de atividade interna do neurônio  $v_j$ , que é dado pela diferença  $u - b_j$ .

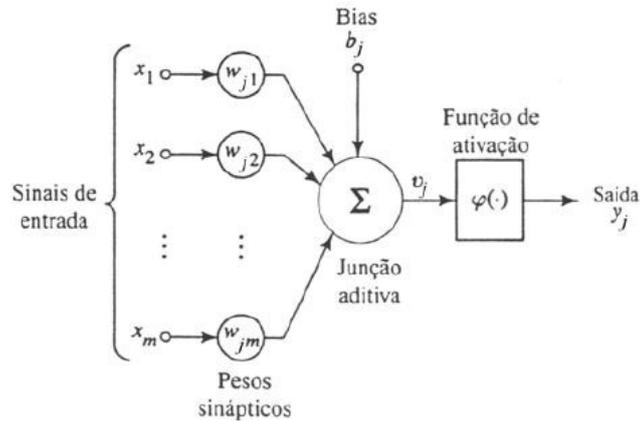


Figura 4.4 – Modelo atual de neurônio perceptron.

Fonte: Adaptado de Haykin, (2001)

$$y_j = \varphi(v_j) \quad (4.3)$$

Onde

$$v_j = \sum_{i=1}^m w_{ij}x_i + b_j \quad (4.4)$$

Por fim, a topologia de uma MLP precisa ser definida de uma forma que otimize o aprendizado e o desempenho da rede. Muitas camadas ou muitos neurônios em cada camada aumentam a flexibilidade da rede, ou seja, esta têm maior capacidade de representar as variações das informações que alimentam a rede (dados de entrada). Uma rede muito flexível além de aumentar o custo computacional, passa a modelar o próprio ruído contido nos dados, correndo o risco de perder sua capacidade de generalização e “decorar” as saídas, o que é caracterizado como “*overfitting*”, ou *super-ajuste*. Por outro lado uma rede com uma topologia pobre, apesar de mais facilmente treinada pode apresentar-se rígida demais para simular os processos a que se propõe. Este problema é conhecido na literatura como “dilema do viés e da variância” (Geman *et al.*, 1992) e envolve a obtenção do melhor compromisso estas duas opções.

### 4.3. Aplicações

Dadas as potencialidades da utilização de RNAs para modelagem de processos físicos, têm se observado um aumento no número de trabalhos na área de

geociências empregando esta técnica. Mendes e Marengo (2009) utilizaram uma rede tipo MLP para o refinamento das saídas de precipitação de modelos climáticos do IPCC sobre a região amazônica, obtendo resultados bem superiores em relação a modelos autoregressivos. Já com relação à previsão de curto prazo Guarnieri (2006) comparou RNAs e regressões múltiplas para estimativas de radiação solar a partir do modelo Eta. Seus resultados demonstraram que há um ganho significativo em relação às saídas do modelo, mas a performance das regressões e da RNA foram similares. Uma justificativa seria de que os processos físicos relacionados à determinação da radiação não apresentariam grandes não-linearidades que justificassem um melhor desempenho das RNAs.

No campo de estimativas de vento em superfície existe uma expectativa de que a capacidade de representar processos não-lineares das RNAs seja determinante, uma vez que estes exercem forte influência em fenômenos micrometeorológicos. No Brasil, Dalmaz (2007) utilizou RNAs no refinamento de previsões de vento para o estado de SC onde obteve RMSE entre 0,7 a 2,2 m/s. O fato de ter utilizado como previsores apenas vento e temperatura com resolução temporal de 6 horas do modelo Eta/CPTEC 40x40km, pode ter limitado o desempenho da RNA.

Utilizando apenas séries temporais de vento e potência de um aerogerador instalado em Olinda, Aquino *et. al.*, (2009) desenvolveu um modelo de previsão baseado na decomposição das séries por transformada de ondaletas, onde as componentes espectrais foram utilizadas como entrada para uma rede neural. As previsões para um horizonte de 24 horas indicaram um erro médio absoluto de 27,6% para o vento e de 16,9% da potência eólica em relação à capacidade instalada. Nesta mesma linha, a partir apenas de séries temporais de vento, Ferreira *et. al* (2008), utilizou um novo tipo de rede neural recorrente chamada *Echo State Network* (ESN), para realizar previsões de vento para duas estações anemométricas do nordeste brasileiro. Seus resultados se mostraram promissores e apontaram um erro médio absoluto (MAE) de 0,71 m/s e 0,57 m/s, para Belo Jardim e Olinda respectivamente, em um horizonte de 24 horas.

Ainda com relação às estimativas de vento, no exterior existem diversos trabalhos publicados utilizando RNAs, muitos dos quais se tornaram pacotes comerciais para utilização na indústria (Monteiro *et al.*, 2009). Dentre os artigos científicos pode-se citar Sanz *et al.*, (2008) que acoplou uma rede MLP às saídas do modelo MM5 e

obteve erro médio absoluto (MAE) de 1,80 m/s na velocidade do vento. Ramirez-Rosado (2009) comparou dois modelos de previsão de potência eólica de curto-prazo baseados em redes tipo MLP acopladas a saídas de modelos meteorológicos, onde ambos tiveram desempenho parecido chegando a RMSE da ordem de 16% da potência disponível num horizonte de 24 a 48 horas. Há também trabalhos que utilizaram outros tipos de RNAs como Sideratos e Hatzigryriou, (2007) que utilizaram Funções de Base Radiais (RBF) a partir de saídas do modelo meteorológico HIRLAM para prever a potência eólica disponível para uma determinada usina. Seu modelo foi subdividido em três classes de velocidade de vento e os resultados levaram a RMSE de 17% e MAE de 12% da potência eólica disponível para um horizonte de 24 a 48 horas.

Guenard *et al.*, (2007), apresentou os resultado de um grande projeto de cooperação entre instituições européias para desenvolvimento de uma plataforma para previsão de curto-prazo de potência eólica chamado ANEMOS. Esta plataforma utiliza uma combinação entre modelos físicos e estatísticos baseados em redes neurais para gerar suas previsões. Os resultados apresentados mostraram um RMSE normalizado de 12% a 20% da potência total instalada para cada aerogerador para um horizonte de 12 a 36 h.

Outra abordagem utilizando redes neurais foi descrita por Cali *et al.*, (2006). Em seu trabalho foram definidos como preditores a temperatura, umidade, pressão, intensidade e direção do vento em superfície, fornecidas por um *ensemble* de previsões numéricas com 75 membros. Foi utilizada uma RNA do tipo MLP com três camadas ocultas para cada membro, onde suas saídas foram posteriormente combinadas por outra RNA com uma camada oculta para finalmente gerar as previsões de potência eólica. As comparações para duas fazendas eólicas na Alemanha levaram um RMSE normalizado médio de 11,1% da capacidade total instalada em cada fazenda.

Também há registro de aplicações distintas na previsão de ventos, conforme Sailor *et. al* (2000), que utilizou redes neurais para o refinamento de modelos climáticos globais, obtendo valores de tendências futuras do vento para diversas estações no Texas, EUA.

Existem outros estudos similares, mas que fogem ao escopo deste texto por apresentar uma relação fraca ou pouco esclarecida com os modelos meteorológicos. Em geral as referências aqui colocadas utilizam poucas variáveis destes modelos como entrada para as RNAs, o que indica haver um potencial ainda a ser explorado nesta área.

Comparado à modelagem física tradicional, as RNAs apresentam como vantagem o custo computacional extremamente baixo depois de treinada a rede, enquanto as desvantagens seriam a especificidade, ou, baixa capacidade de generalização e a necessidade de se obter séries de dados observados para a geração do modelo. O conhecimento sobre os processos que influenciam o regime de ventos na CLA é essencial na aplicação desta técnica, uma vez que a correta seleção dos preditores é fundamental para uma boa performance do modelo (Mori e Umezawa, 2009).

## **5 DADOS**

No desenvolvimento deste trabalho foram utilizados dois tipos de dados: observacionais e de modelo. Os primeiros são oriundos de medições em torres anemométricas e os seguintes a partir dos pontos de grade do modelo Eta/PREVENTO, doravante denominado Eta/PV. O projeto PREVENTO foi fruto de uma cooperação entre CEPEL e o CPTEC/INPE, financiado pela ANEEL, que visou a melhoria das previsões de vento para a análise de inserção de sistemas eólicos na matriz de geração da CHESF. Este projeto teve início em 2004, sendo que de Agosto/2005 até Abril/2008 foram geradas previsões de vento operacionalmente pelo CPTEC, específicas para determinada região do Nordeste brasileiro.

### **5.1. Dados Observacionais**

Os dados observacionais disponíveis para este trabalho são oriundos do projeto SONDA, além de dados cedidos pela UFAL referentes ao mapeamento eólico do estado de Alagoas (Costa, 2009). Todos os dados foram coletados por torres anemométricas com diversos níveis de medição e são constituídos por séries de velocidade e direção do vento, além de registros de temperatura conforme descrito na Tabela 5.1. Estes dados possuem resolução temporal de 10 minutos, mas foram convertidos em médias horárias devido à limitação na resolução dos preditores, originários dos modelos meteorológicos.

Adotou-se o Nordeste brasileiro como zona alvo pela maior disponibilidade de torres anemométricas, e por se tratar da região de maior concentração de projetos eólicos no Brasil. A distribuição geográfica das torres pelo território abrange regiões com características de relevo e vegetação diversas, de modo que 2 torres se localizam no litoral e outras 6 torres no interior do continente como ilustra a Figura 5.1.

Tabela 5.1. Características das torres anemométricas.

Nro	ESTAÇÃO	SIGLA	ALTITUDE	ALTURA DA TORRE	NÍVEIS DE MEDIÇÃO
1	Petrolina - PE <sup>1</sup>	PETR	387 m	50 m	10/25/50 m
2	São João do Cariri - PB <sup>1</sup>	SJCA	486 m	50 m	25/50 m
3	Triunfo - PE <sup>1</sup>	TRFO	1123 m	50 m	25/50 m
4	Roteiro - AL	ROTE	60 m	50 m	30/50 m
5	Maragogi - AL	MAGI	50 m	50 m	30/50 m
6	Girau do Ponciano - AL	GIRP	410 m	50 m	30/50 m
7	Agua Branca - AL	AGUB	718 m	50 m	30/50 m
8	Palmeira dos Índios- AL	PALM	649 m	100 m	30/70/100 m

<sup>1</sup> Estações do Projeto SONDA

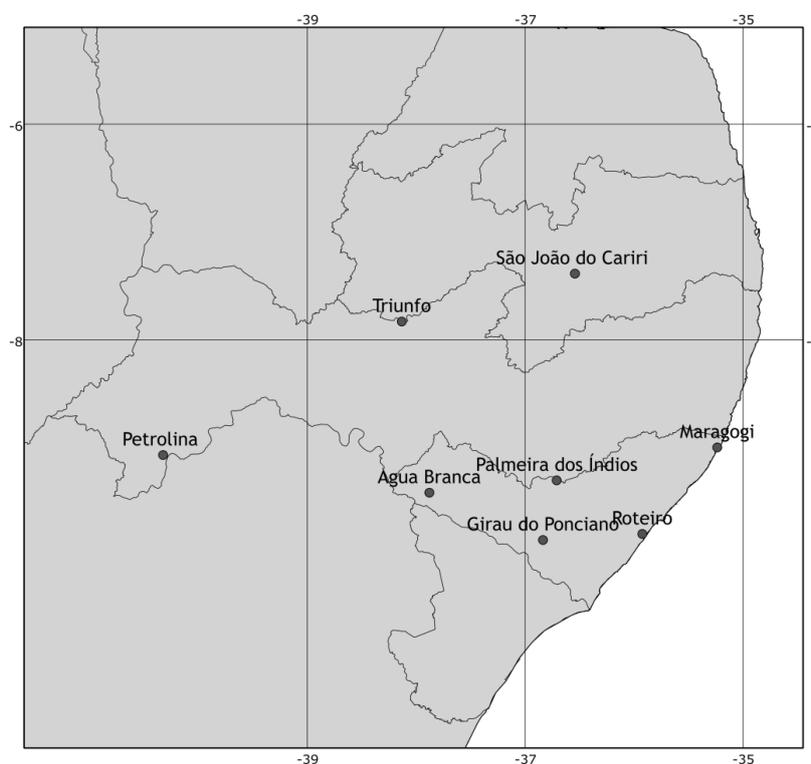


Figura 5.1 - Mapa com a localização das torres anemométricas

O período dos dados varia em cada torre, no entanto, para o desenvolvimento dos modelos RNA foram utilizados dados coletados entre Agosto/2005 e Abril/2008

devido à compatibilidade com as saídas disponíveis do modelo Eta/PV. Neste período de interesse, cada torre possui pelo menos 8 meses completos e ininterruptos, possibilitando que os modelos RNA desenvolvidos identifiquem as variações sazonais do vento.

A qualificação dos dados coletados nas 8 torres foi feita de acordo com a procedimento estabelecido pelo projeto SONDA (Martins *et al.*, 2007). Os dados anemométricos coletados após Abril/2008 foram utilizados no cálculo das climatologias diárias a fim de dar maior representatividade as mesmas.

Como as variações sazonais na precipitação podem ter influência sobre a modelagem estatística do vento, é preciso prezar pelo balanceamento entre os períodos seco e chuvoso no refinamento do modelo. Para isso foram obtidas junto ao Instituto nacional de Meteorologia (INMET) as climatologias de precipitação para os locais onde se encontram as torres anemométricas. A Figura 5.2 ilustra este regime de precipitação para as torres de SJCA e ROTE, instaladas no interior e no litoral, respectivamente.

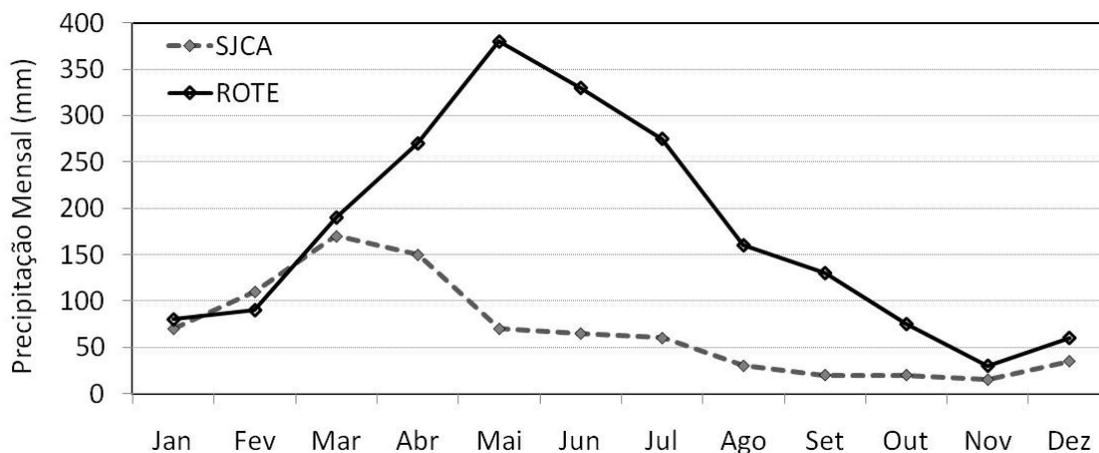


Figura 5.2 – Climatologia de precipitação sobre as torres de SJCA e ROTE.

Fonte: Brasil. INMET 2011

A região de estudo possui acumulados médios anuais de precipitação variando de 600 mm no sertão nordestino a 3000 mm na faixa litorânea. Os principais sistemas atmosféricos atuantes são a Zona de Convergência Intertropical, os Vórtices Cilíndricos de Altos Níveis, os Distúrbios Ondulatórios de Leste (DOL) e as incursões de frentes frias, esta última restrita ao sul do nordeste. Estes sistemas atuam como forçantes dinâmicas, organizando a convecção sobre o continente, entretanto fenômenos locais como a convecção profunda relacionada à orografia e

instabilidade termodinâmica, ainda respondem por uma boa parte da precipitação nesta região (Molion e Bernardo, 2000).

Uma característica já mencionada anteriormente em relação à climatologia de chuvas do nordeste é a sua complementaridade com o regime de ventos (Amarante *et al.*, 2001). Neste trabalho foi feita uma comparação entre estes regimes para a região de Petrolina – PE. A Figura 5.3 ilustra o ciclo anual do vento a 50 m, comparando-o à capacidade do reservatório da usina hidroelétrica de Sobradinho – BA, que está a menos de 50 km de distância da torre anemométrica. Fica evidente que esta complementaridade é válida também para este local.

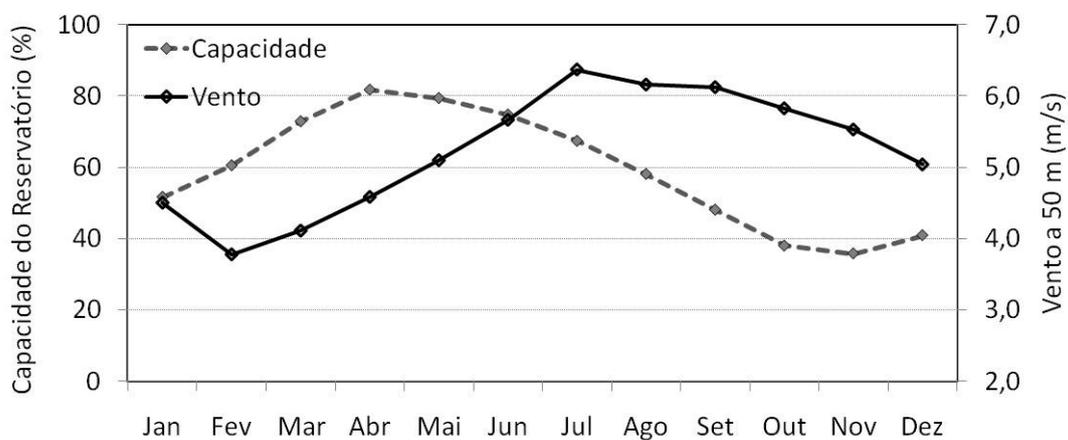


Figura 5.3 – Ciclo anual do vento a 50 m em Petrolina-PE comparado à capacidade do reservatório da usina hidroelétrica de Sobradinho – BA.

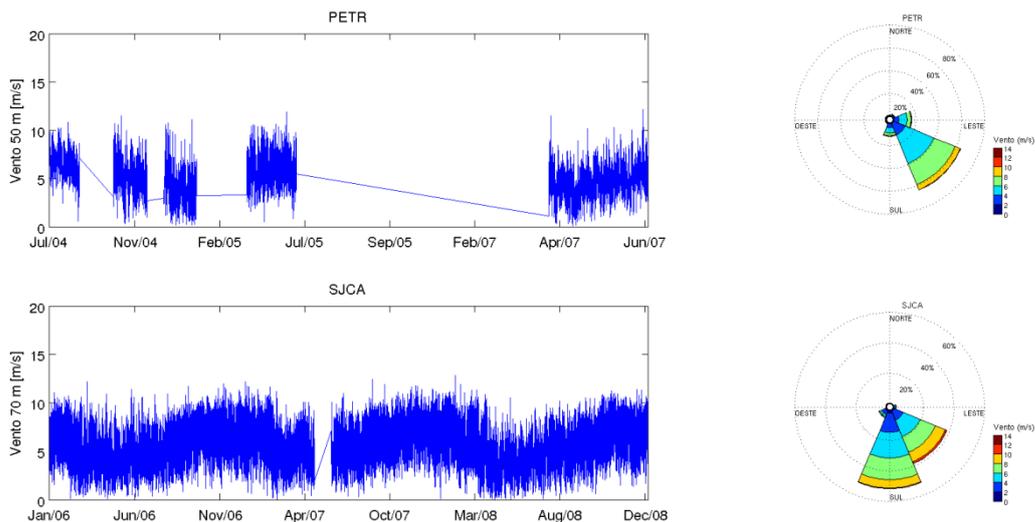
Fonte: Brasil. ANEEL 2011

A Figura 5.4 ilustra de que forma a precipitação se distribui ao longo dos dados anemométricos disponíveis. Como o volume anual de chuva varia muito entre as localidades, para separar os meses mais chuvosos foi feita a comparação entre cada mês e o valor médio mensal obtido caso a precipitação anual fosse uniformemente distribuída.

Distribuição dos dados anemométricos e sua relação com a precipitação																								
Dados	Nro Meses	2005					2006					2007				2008								
		A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A		
PETR	20											<	<	<	<									
SJCA	27											<	<	<	<	<	<							
TRFO	17																							
ROTE	9	LEGENDA:																						
MAGI	8	< < Média Mensal																						
GIRP	8	> Média Mensal																						
AGUB	8	> Média Mensal + 50%																						
PALM	9	* Dados Modelo Eta/PV																						
Eta/PV	33	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*		

Figura 5.4 – Distribuição da precipitação ao longo dos dados anemométricos disponíveis.

Além da qualificação, os dados anemométricos passaram por uma inspeção visual para detectar qualquer erro mais grosseiro que eventualmente não tenham sido detectados pelo processo de controle de qualidade adotado. Este procedimento é recomendável sempre que possível, até como uma maneira de familiarizar com o comportamento do vento em uma determinada região. Na Figura 5.5 são mostradas as séries de dados usadas neste estudo. Maiores detalhes sobre e fotos dos arredores de cada torre podem ser encontrados em Costa (2007) e no *website* do projeto SONDA (<[www.ccst.inpe.br/sonda/](http://www.ccst.inpe.br/sonda/)>).



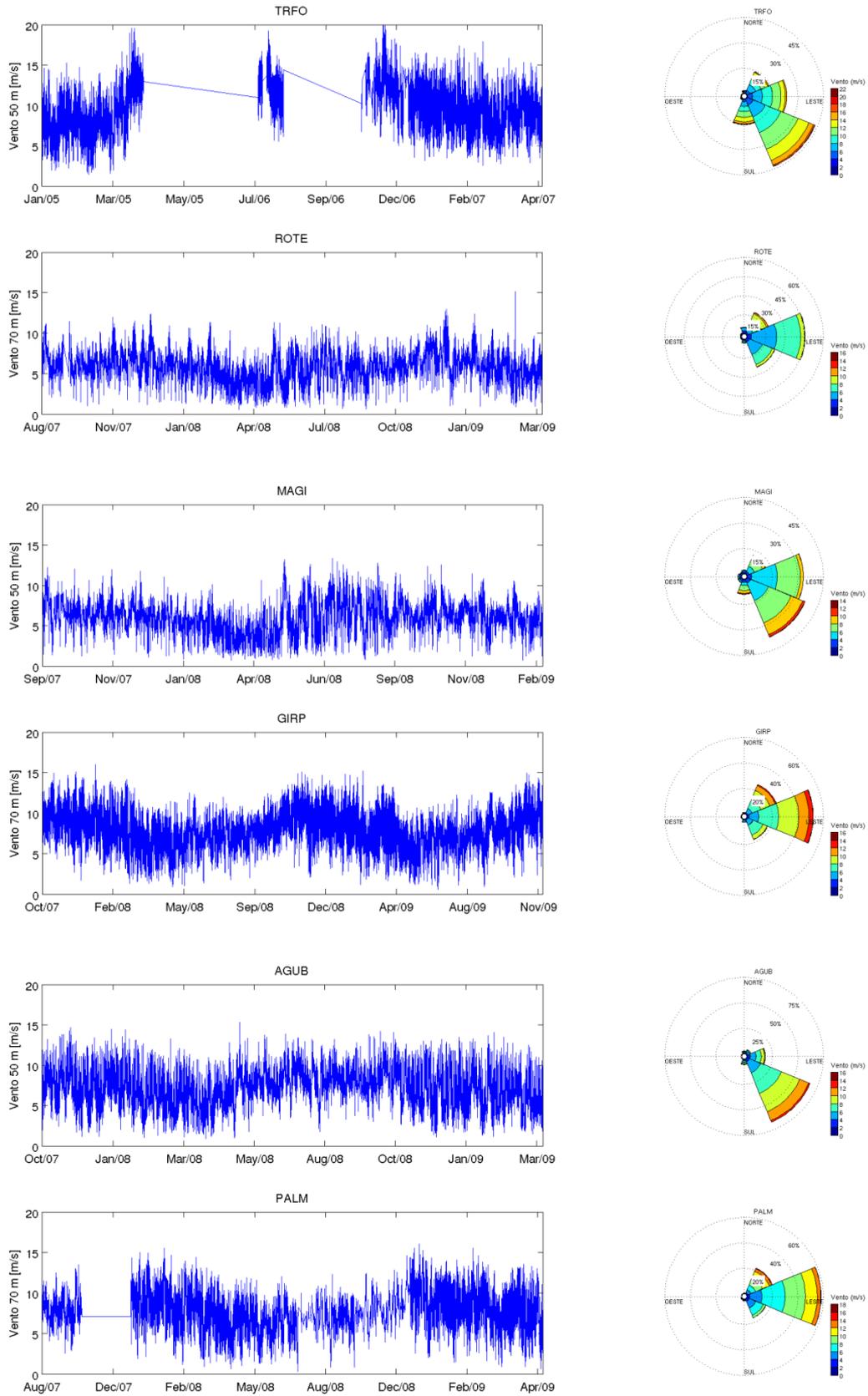


Figura 5.5 – Séries anemométricas e rosa dos ventos para cada torre anemométrica empregada neste estudo.

## 5.2. Dados do Modelo Eta/PREVENTO

O modelo Eta/PV consiste de uma versão do modelo Eta denominada *Eta Workstation*. Esta versão possui o mesmo código meteorológico da versão operacional, no entanto utiliza um processo de assimilação mais simplificado e uma instalação compatível com diversas máquinas, sendo assim mais utilizada para pesquisas (NCEP, 2011). Para este projeto foi adotada uma resolução horizontal menor (5 km), saídas de vento dedicadas a 10 m e 50 m de altura além de resolução temporal de uma hora. Através deste modelo foram geradas previsões de vento operacionais pelo CPTEC específicas para uma região do nordeste brasileiro. A opção por utilizar as saídas do Eta/PV, em detrimento ao Eta 20 x 20 km, que está operacional atualmente no CPTEC, ocorre porque a existência de saídas horárias é fundamental para o refinamento dos campos de vento dada a grande variabilidade em seu ciclo diário, o que não se verifica para o Eta operacional (somente saídas a cada 6 horas). Outro ponto favorável é que a próxima versão do Eta a entrar em operação, deverá incorporar uma resolução espacial e temporal menor, próximas ao Eta/PV, o que torna a metodologia desenvolvida neste estudo atual e válida para aplicação futura.

Para a redução do espaço em disco o domínio inicial do modelo foi recortado para uma região de interesse, ficando restrito entres as latitudes 7° S e 11° S e longitudes 34° W a 41° W. A Figura 5.6 ilustra o domínio do modelo, seu recorte e o mapa de relevo utilizado nas rodadas.

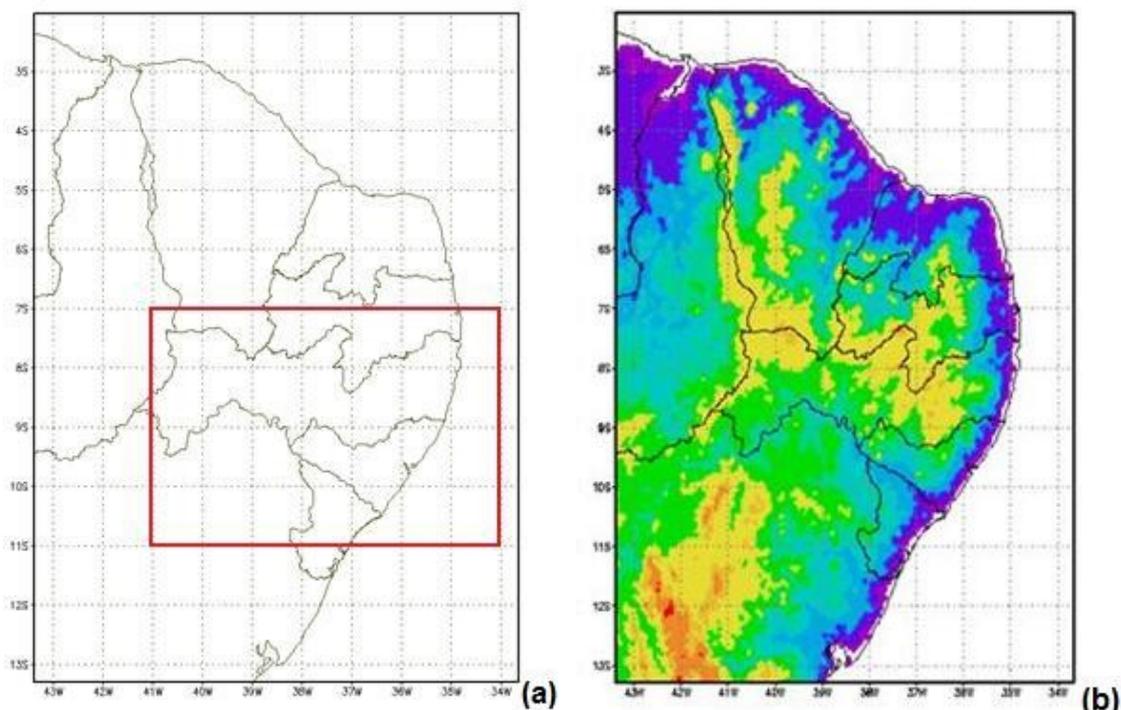


Figura 5.6 - Domínio da grade do modelo Eta/PREVENTO com região de recorte em vermelho (a); com topografia relativa na resolução de 5 km (b). Sentido de escalas do azul para o amarelo indicam aumento altitude.

Cada saída do Eta/PV é constituída por 58 variáveis sendo 9 dispostas em 19 níveis de pressão, com rodadas efetuadas diariamente às 00UTC para um horizonte de previsão de 72 horas, totalizando um grande volume de dados. Estas saídas foram recortadas para um domínio menor, que cobre a região de interesse, sendo que as primeiras 12 horas de previsão foram descartadas, pois sofrem fortes ajustes dos campos de condição inicial. Sendo assim somente as previsões para o intervalo de 12 a 36 horas a partir de cada rodada foram utilizadas. Posteriormente, para cada torre foram extraídas as saídas correspondentes à grade na qual esta se encontra, sem nenhum tipo de ponderação em função de sua proximidade com grades adjacentes. Na Tabela 5.2 estão listadas as variáveis de saída do modelo para cada ponto de grade.

Tabela 5.2 - Descrição das variáveis Eta/PV.

Nro	Variável	Níveis	Unidade	Descrição
1	apcpsfc	Superfície	kg/m <sup>2</sup>	Precipitação total
2	cdsfc	Superfície	non-dim	Coeficiente de arrasto superficial
3	dlwrfsfc	Superfície	W/m <sup>2</sup>	Onda longa Descendente
4	dpt2m	Superfície	K	Ponto de Orvalho 2m
5	dswrfsfc	Superfície	W/m <sup>2</sup>	Onda Curta Descendente
6	lcdclcl	Superfície	%	Fração de cobertura de nuvens baixas
7	lhtflsfc	Superfície	W/m <sup>2</sup>	Fluxo calor latente
8	msletmsl	Superfície	Pa	Pressão media nível do mar (Eta)
9	mstavdlr	Superfície	%	Disponibilidade de umidade
10	pressfc	Superfície	Pa	Pressão em superfície
11	sfexcsfc	Superfície	(kg/m <sup>3</sup> ). (m/s)	Coeficiente troca em superfície
12	shtflsfc	Superfície	W/m <sup>2</sup>	Fluxo de calor sensível
13	spfh10m	Superfície	kg/kg	Umidade especifica a 10m
14	tcdclm	Superfície	%	Cobertura total de nuvens
15	tmp2m	Superfície	K	Temperatura a 2m
16	uflxsfc	Superfície	N/m <sup>2</sup>	Fluxo zonal de momentum
17	ugrd10m	Superfície	m/s	Vento zonal a 10 m
18	vflxsfc	Superfície	N/m <sup>2</sup>	Fluxo meridional de momentum
19	vgrd10m	Superfície	m/s	Vento meridional a 10 m
20	hgtprs	19	gpm	Altura geopotencial da camada
21	spfhprs	19	kg/kg	Umidade especifica da camada
22	tmpprs	19	K	Temperatura da camada
23	ugrdprs	19	m/s	Vento zonal na camada
24	vgrdprs	19	m/s	Vento meridional na camada

## 6 METODOLOGIA

A metodologia empregada no desenvolvimento deste trabalho foi subdividida em tarefas específicas para facilitar o entendimento do processo. A cronologia proposta abrange: *Pré-processamento e Assimilação, Tratamento dos Dados, Seleção dos Preditores, Desenvolvimento do Modelo em RNAs e Avaliação dos Modelos Obtidos*

### 6.1. Pré-processamento e Assimilação

Tanto os dados observados quanto os modelados foram pré-processados para adquirir a forma necessária para serem usados no desenvolvimento do modelo de RNA.

Os dados do Eta/PREVENTO se encontravam em formato binário, conhecido como GRIB (*Gridded Binary*), um padrão adotado pela Organização Meteorológica Mundial (OMM). Apesar de este formato apresentar uma boa compactação, ocupando pouca memória, sua manipulação é trabalhosa, pois cada rodada do modelo fica diluída ao longo de vários arquivos binários e descritores que precisam ser combinados, o que prejudica a intercambiabilidade com outros softwares. Já o formato de dados NetCDF (*Network Common Data Form*) é auto-descritivo e independente da máquina, criado com a proposta de facilitar o acesso e intercâmbio de dados científicos vetoriais, bastando para isso a instalação de uma biblioteca de software. Como o volume de dados não foi um fator limitante para este trabalho optou-se por convertê-los para NetCDF.

Para tanto os dados foram recortados utilizando o código LATS4D do software GrADS (*Grid Analysis and Display System*), limitando o domínio espacial das saídas para o menor possível que englobasse todas as estações, o número de níveis em apenas 8, e o intervalo de previsão restrito entre +12 a +36 horas. Como o número de arquivos é muito grande, o recorte e a conversão para o formato NetCDF foi programada em linguagem SHELL. Apesar de o NetCDF ser um formato de menor compactação o volume de dados foi reduzido de 700 GB para 120 GB aproximadamente. A leitura destes arquivos, assim como o restante do trabalho, foi executada empregando o software MATLAB®.

Os dados do modelo foram lidos e concatenados, de modo a formar uma série temporal para cada estação de acordo com a grade correspondente a sua posição geográfica, sem nenhum tipo de interpolação quanto às grades adjacentes. Foram

assimiladas todas as variáveis de superfície e 5 variáveis por camada para 7 níveis de pressão. Os níveis incluídos foram 1000, 950, 925, 900 e 850 hPa, por estarem relacionados aos gradientes verticais e à dinâmica da CLA, e 500 e 200 hPa, por apresentarem uma relação com o vento geostrófico e principalmente com as forçantes dinâmicas de alta troposfera, como os distúrbios ondulatórios e VCANs. Portanto o conjunto total de dados do modelo assimilados para cada estação foi de 54 variáveis.

Os dados anemométricos se encontravam em arquivos mensais em formato texto de extensão CSV (*Comma Separated Value*), com pelo menos dois níveis de medição de intensidade do vento conforme mostrado anteriormente na Tabela 6.1. Tais valores foram lidos e concatenados de maneira a formarem uma série temporal por estação. Posteriormente a resolução temporal foi alterada de dez minutos para uma hora, através de médias móveis de 30 minutos, para remover oscilações de escala temporal inferior às de interesse deste trabalho. No caso dos dados da rede SONDA, foram assimilados também os arquivos de qualificação de cada estação para posterior tratamento das séries.

## **6.2. Tratamento dos Dados**

### **6.2.1. Dados do Modelo**

A primeira etapa realizada no tratamento foi a qualificação. Os dados de modelo, embora pareça desnecessário, precisaram também passar por um processo de verificação quanto à consistência das variáveis e/ou valores espúrios. Para o tipo de modelagem em questão, mesmo uma pequena quantidade de valores espúrios pode introduzir erros consideráveis no resultado final. Foram removidas as variáveis inconsistentes e mantidas apenas aquelas cuja porcentagem de valores espúrios foi inferior a 1%, os quais foram substituídos pelo valor médio da variável. As rodadas do Eta/PV estavam divididas em 3 experimentos, conforme descrito pela Tabela 6.2. Para o desenvolvimento deste trabalho foi utilizado apenas o experimento Exp\_3 pelo maior volume de dados.

Tabela 6.2 – Experimentos referentes às rodadas do modelo Eta/PV.

Experimento	Período	Resolução Horizontal
Exp_1	37 dias	5 km x 5 km
Exp_2	5 meses	5 km x 5 km
Exp_3	25 meses	5 km x 5 km

Como um dos grandes desafios deste trabalho é o de encontrar os melhores preditores para o modelo em RNA, tornou-se interessante que fossem construídas novas variáveis a partir das existentes, ou adicionadas variáveis externas que acrescentem alguma informação relevante na determinação dos campos de vento. Neste processo foram criadas primeiramente as variáveis de magnitude e direção do vento a 10m. A seguir, como a dinâmica atmosférica é regida por um sistema de equações diferenciais, espera-se que relações significativas possam existir entre a variação temporal de alguma variável e os valores de vento. Nesse sentido foram criadas variáveis de tendência horária  $\phi_{gh}$  e tendência diária  $\phi_{gd}$  para cada variável do modelo de forma que:

$$\phi_{gh} = \phi(t) - \phi(t - 1) \quad (6.1)$$

$$\phi_{gd} = \phi(t) - \phi(t - 24) \quad (6.2)$$

Onde  $\phi$  representa uma variável qualquer e  $t$  o tempo em horas.

Cabe ressaltar que como as redes empregadas neste estudo são não-recorrentes, não representando efeitos dinâmicos, faz-se necessário que qualquer relação temporal de interesse seja fornecida como entrada. Em compensação tais redes são mais simples de se treinar e indiferentes a descontinuidades nas séries temporais. Redes dinâmicas multivariadas ainda estão na fronteira do conhecimento da computação aplicada sendo difícil encontrar softwares e algoritmos de treinamento que permitam sua implementação de maneira eficiente.

Ainda como uma forma de auxiliar a RNA no reconhecimento de relações não-lineares, optou-se por incluir como preditores combinações também não-lineares entre as variáveis. Como as possibilidades seriam inúmeras, optou-se por obter relações similares aos termos da componente horizontal da equação de Navier-

Stokes (N-S) (Equações 2.5 a 2.7). Para tanto algumas aproximações foram necessárias e são descritas resumidamente a seguir.

Como o objetivo é modelar a magnitude do vento e não a direção, por conveniência adota-se a direção  $x$  sendo paralela ao eixo longitudinal do vento, de modo que a componente em  $y$ , denominada  $\hat{v}$  seja nula, e a componente  $u$  seja igual à magnitude do vento  $U$ , conforme descrito pela Eq. 6.3. Ainda, admite-se homogeneidade lateral, de modo que toda derivada em relação à  $y$  também seja nula, conforme Equação 6.4. Por fim, para eliminar derivadas espaciais, que são inviáveis neste caso, pois nada se sabe sobre as grades adjacentes, admite-se que no intervalo de uma hora entre uma amostra e outra o vento pode ser considerado em regime permanente, de modo que tais derivadas são aproximadas conforme a Equação 6.5.

$$\hat{v} = 0 \xrightarrow{\text{leva a}} \hat{u} = U \quad (6.3)$$

$$\frac{\partial \phi}{\partial y} = 0 \quad (6.4)$$

$$\partial x = \bar{U} \cdot \partial t \xrightarrow{\text{leva a}} \frac{\partial \phi}{\partial x} = \frac{1}{\bar{U}} \cdot \frac{\partial \phi}{\partial t} \quad (6.5)$$

Onde  $\phi$  representa uma variável qualquer,  $\bar{U}$  a magnitude média do escoamento entre uma amostra e outra e  $t$  o tempo. As variáveis indicadas por  $\hat{u}$  e  $\hat{v}$  representam as componentes do vento neste novo eixo rotacionado.

A partir deste ponto, é possível obter sinais proporcionais a cada termo da equação de N-S utilizando o método das diferenças finitas para aproximar as derivadas, conforme exposto nas Equações 6.6 a 6.10. Quando necessárias outras aproximações foram efetuadas para viabilizar a obtenção dos termos, mas seu detalhamento foge ao escopo deste capítulo.

A advecção turbulenta de momento na direção do escoamento ( $uxturb$ ) pode ser aproximada pela Equação 6.6.

$$uxturb = \frac{1}{\bar{U}(\Delta t)} \cdot \frac{U(t)^2 - U(t-1)^2}{\Delta t} \quad (6.6)$$

A advecção média de momento na direção do escoamento ( $uxmean$ ) pode ser aproximada pela Equação 6.7.

$$uxmean = \frac{1}{\bar{U}(\Delta t)} \cdot \frac{U(t)^2}{\Delta t} \quad (6.7)$$

O primeiro termo do fluxo vertical turbulento médio de momento ( $uwsdy1$ ) pode ser aproximado pela Equação 6.8.

$$uwsdy1 = v_{10m}(t) - v_{500}(t) \quad (6.8)$$

O segundo termo do fluxo vertical turbulento médio de momento ( $uwsdy2$ ) pode ser aproximado pela Equação 6.9.

$$uwsdy2 = \left[ \frac{U_{1000}(t) - U_{950}(t)}{h_{1000}(t) - h_{950}(t)} \right]^2 \quad (6.9)$$

O gradiente horizontal de pressão na direção do escoamento ( $dprsdx$ ) pode ser aproximada pela Equação 6.10.

$$dprsdx = \frac{1}{\bar{U}(\Delta t)} \cdot \frac{p(t) - p(t-1)}{\Delta t} \quad (6.10)$$

A absorção de momento pelos efeitos viscosos ( $uxvisc$ ) pode ser aproximada pela Equação 6.11.

$$uxvisc = \left[ \frac{1}{\bar{U}(\Delta t)} \cdot \frac{U(t) - U(t-1)}{\Delta t} \right]^2 \quad (6.11)$$

Outra aproximação para o fluxo vertical turbulento médio de momento em função do cisalhamento vertical ( $uwflux$ ) é dada pela Equação 6.12.

$$uwflux = \left[ \left( \frac{U_{925}(t) - U_{950}(t)}{\ln \left( \frac{h_{925}(t) - \bar{h}}{h_{950}(t) - \bar{h}} \right)} \right)^2 - \left( \frac{U_{950}(t) - U_{10m}(t)}{\ln \left( \frac{h_{950}(t) - \bar{h}}{h_{10m}(t) - \bar{h}} \right)} \right)^2 \right] \cdot \left( \frac{1}{h_{925}(t) - \bar{h}} \right) \quad (6.12)$$

Onde nas equações acima  $h_p$  representa a altura geopotencial no nível de pressão atmosférica  $p$ , e  $\bar{h}$  é o valor médio entre as duas camadas.

Para as variáveis  $uxmean$ ,  $dprsdx$  e  $uxvisc$ , foram geradas além dos fluxos horários, também os fluxos médios em 24 horas.

Finalmente, sendo o ciclo diário a oscilação mais intensa nas séries de vento, uma maneira encontrada para captar qualquer fenômeno cíclico com este período se deu pela introdução de 12 variáveis sintéticas representadas por ondas senoidais com 1 hora de defasagem entre elas, conforme ilustrado na Figura 6.1. Como não importa

se as correlações com o vento são positivas ou negativas, 12 horas são suficientes para cobrir todo o intervalo de possíveis ciclos ao longo do dia.

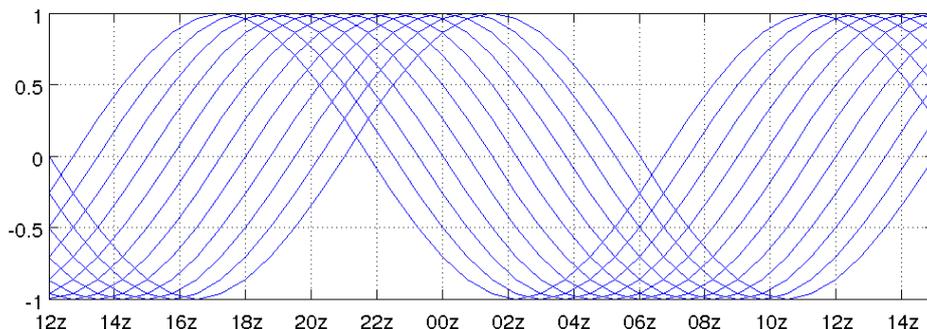


Figura 6.1 – Ondas senoidais introduzidas no conjunto de dados

Após estas manipulações o conjunto final de dados passou a ter 189 possíveis preditores para cada estação. O Apêndice A traz uma lista detalhada com todos eles.

### 6.2.2. Dados Anemométricos

Os dados oriundos das torres anemométricas foram qualificados adotando os critérios já utilizados pela rede SONDA, além de outros critérios mais restritivos. Estes critérios adicionais foram obtidos empiricamente através de testes sobre o conjunto de dados, e se fizeram necessários para garantir que nenhum valor incoerente viesse a prejudicar o treinamento da RNA. Quanto aos dados SONDA, uma vez que estes já passaram pelo processo de qualificação interno, foram aplicados apenas os critérios adicionais. A Tabela 6.3 ilustra os critérios utilizados, onde  $V$  representa o vento medido para cada um dos níveis,  $\Delta_n$  intervalos de  $n$  horas e  $z$  a altura de medição tal que  $z_2 > z_1$ . Ao final desta etapa obteve-se um arquivo de qualificação para cada estação, onde os dados reprovados em qualquer um dos testes foram marcados como suspeitos.

Tabela 6.3 – Critérios para o controle de qualidade dos dados anemométricos.

Nro	Critérios / Valores em (m/s)	Descrição	Origem
1	$Mín(V) > 0$	Limite Inferior	SONDA
2	$Máx(V) < 25$	Limite Superior	SONDA
3	$Máx[V(\Delta_3)] - Mín[V(\Delta_3)] > 0,1$	Variabilidade em 3 h	SONDA
4	$Máx[V(\Delta_{24})] - Mín[V(\Delta_{24})] > 0,5$	Variabilidade em 24 h	SONDA
5	$Mín[V(\Delta_{24})] < 13$	Limite superior para o valor mínimo em 24 h	Adicional
6	$Mediana[V(\Delta_{24})] < 18$	Limite superior para o valor médio em 24 h	Adicional
7	$Mediana[V(\Delta_{24})] < 15$	Limite Superior para o valor médio em 72 h	Adicional
8	$V(z1) - V(z2) < 0,5$	Limite para inversão de perfil	Adicional

Em seguida foram excluídos os dados anemométricos que se mostraram suspeitos em pelo menos dois níveis de observação. Para os casos onde apenas a velocidade no nível mais alto era considerada suspeita, esta foi inferida utilizando-se funções pré-definidas para o ajuste do perfil vertical, obtidas a partir dos dados não-suspeitos da torre aplicados à Equação 2.1. É possível mostrar que para uma atmosfera neutra, deverá existir uma relação linear entre a velocidade de fricção  $u^*$  e a velocidade do vento  $U$ , conforme a Equação 6.13. A partir da Equação 2.1, chega-se que coeficiente  $A$  depende apenas da altura de medição  $Z$  e da rugosidade local  $Z_0$ , conforme a Equação 6.14.

$$u^* = A \cdot U(z) \quad (6.13)$$

$$A = k \cdot \frac{1}{\ln\left(\frac{Z}{Z_0}\right)} \quad (6.14)$$

Onde  $k$  é a constante de Von-Karman.

A rugosidade local pode ser calculada pela Equação 6.15 para cada amostra, e seu valor final determinado através da análise de seu histograma para todos os ajustes de perfil efetuados. Uma alternativa seria definir  $Z_0$  como sendo a mediana dos valores, o que é uma aproximação bastante robusta, pois corresponde ao valor de 50% da função de densidade de probabilidade acumulada (CDF) não-paramétrica para a distribuição em questão. Outra opção seria a partir das Equações 6.13 e 6.14 adotar o valor de rugosidade que gera um viés nulo no nível mais elevado a através da auto-validação do perfil. Entretanto, como o objetivo é utilizar os valores extrapolados no treinamento da RNA, e esta utiliza como critério erro quadrático médio, resolveu-se adotar o valor de rugosidade que corresponde ao valor mínimo do RMSE na auto-validação do perfil. A Figura 6.2 ilustra este procedimento para uma das estações. Não foi considerada a variação da rugosidade em função de alterações sazonais na vegetação.

$$z_0 = EXP \left[ \ln(z_1) - U(z_1) \cdot \left( \frac{\ln\left(\frac{z_2}{z_1}\right)}{U(z_2) - U(z_1)} \right) \right] \quad (6.15)$$

Onde  $Z_1$  e  $Z_2$  são os níveis de medição utilizados.

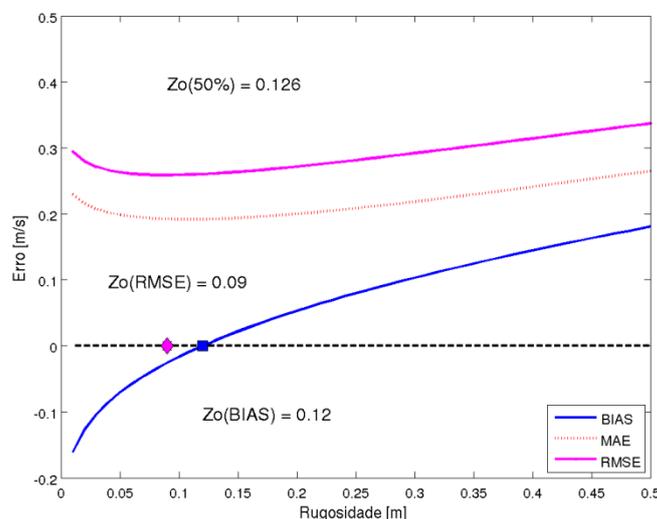


Figura 6.2 – Relação entre os valores de  $Z_0$  e erro na determinação da rugosidade onde  $z_0(50\%)$  corresponde a mediana da distribuição. Método adotado consiste na identificação das posições de mínimo erro.

Além disso, ao considerar a CLS como neutra, em média, espera-se que os erros não sejam significativos, pois conforme observado por Roballo & Fisch (2008) os

ventos mais intensos típicos nestas regiões tendem a neutralizar a CLS devido à grande turbulência mecânica gerada. Para mensurar o erro cometido, o método proposto foi validado através da comparação entre valores reais e estimados, onde chegou-se a valores de RMSE inferiores 0.3 m/s, o que é adequado para a modelagem em questão. Esta etapa do processo foi denominada reparação dos dados.

### **6.2.3. Sincronização dos Dados**

A última etapa do processo de tratamento foi o acoplamento entre dados modelados e dados anemométricos de forma sincronizada no tempo. Cabe ressaltar que algumas estações estavam referenciadas em hora local enquanto o modelo em horário UTC. Além disso foi preciso levar em consideração o padrão de data numérica adotada por cada formato de arquivo para evitar incompatibilidades. Após a sincronização o número de amostras por estação ficou reduzido, sendo limitado pela quantidade de dados anemométricos disponíveis.

## **6.3. Seleção dos Preditores**

Conforme discutido anteriormente a seleção dos preditores é uma das etapas mais sensíveis no desenvolvimento de modelos de RNAs e exige conhecimento sobre os processos físicos presentes na CLA. Parte-se de um grande número de previsores em potencial que agregam uma complexidade desnecessária à rede e precisa ser reduzido. A seleção com base em correlações lineares é uma técnica bastante consolidada quando os fenômenos envolvidos não apresentam grandes não-linearidades. Neste caso, entretanto, um cuidado adicional se faz necessário para preservar estas possíveis relações.

### **6.3.1. Definição dos Subconjuntos por Limiar de Correlação**

Nesta etapa definiram-se duas estações anemométricas para efetuar os testes necessários. Escolheu-se SJCA e ROTE pela boa qualidade dos dados e pela diversidade geográfica, já que a primeira se encontra a mais de 200 km do litoral e a

segunda a poucos metros da costa. Calculou-se as climatologias diárias de cada variável e a partir destas foram geradas as anomalias normalizadas.

Calculou-se a correlação cruzada entre as variáveis e eliminou-se aquelas consideradas redundantes (coeficiente de correlação superior a 0,99). Foram calculadas correlações de Pearson, Kendall e Spearman entre os preditores e o alvo, dado pela anomalia normalizada do vento observado a 50m. Neste processo foram testadas defasagens de -3 horas a +3 horas entre preditores e alvo. Todos os tipos de correlações utilizados são limitados entre -1 e 1, sendo que a de Kendall e Spearman possuem alguma sensibilidade a relações não-lineares.

Identificou-se para cada preditor qual a defasagem que leva à maior correlação. Criou-se então 8 subconjuntos de preditores a partir de limiares inferiores para os módulos das correlações. De forma conservadora, estes módulos variaram de 0,01 a 0,15 conforme mostra a Tabela 6.4.

Tabela 6.4 – Limiares utilizados para os valores de correlação na pré-seleção dos preditores

Limiar 1	Limiar 2	Limiar 3	Limiar 4	Limiar 5	Limiar 6	Limiar 7	Limiar 8
0,01	0,03	0,05	0,07	0,09	0,11	0,13	0,15

### 6.3.2. Regressão Linear Passo-a-Passo

Cada conjunto de preditores foi usado como entrada para uma regressão linear passo-a-passo (Wilks, 2006), em que o critério de significância mínima de entrada foi de 0,05 e a significância máxima para saída foi de 0,10. A regressão linear passo-a-passo é uma técnica de regressão multivariada em que preditores são adicionados ou removidos gradativamente ao modelo na medida em que atendam ao critério de entrada, ou deixem de atender ao critério de saída. Estes critérios estão ligados ao poder explanatório adicional que cada preditor agrega ao modelo. Mesmo para relações lineares sua eficácia não é absoluta, pois a ordem em que os preditores são adicionados é importante, e testar todas as combinações possíveis muitas vezes torna-se inviável. Apesar disso é uma forma sistemática e bem fundamentada de se efetuar uma seleção de variáveis, e desde que se admita que bons preditores lineares têm uma maior probabilidade de serem bons preditores

não-lineares, seu emprego pode ser estendido para a tarefa em questão. Cabe ressaltar que para efetuar regressões não-lineares multivariadas é preciso pré-definir o tipo de função que será utilizada (polinômiais, integrais, diferenciais, homogêneas, etc..) pois os sistemas de equações são indeterminados, o que dificulta bastante a aplicação desta técnica para fenômenos cujo equacionamento é muito complexo, como neste caso.

### **6.3.3. Treinamento por Redes Neurais Artificiais**

As saídas obtidas pela regressão definem 8 subconjuntos de preditores, os quais foram utilizados como entradas para as RNAs. Cada preditor é previamente defasado em função de sua posição de máxima correlação. Foram empregadas RNAs do tipo *perceptrons multi-camadas* (MLP) devido à sua capacidade de aproximação de qualquer função contínua aliada à capacidade de generalização. Como nesta etapa a melhor configuração da RNA ainda não foi definida, adota-se uma configuração padrão, com uma camada oculta, funções de transferência do tipo sigmoidais na camada intermediária e linear na camada de saída. O número de neurônios foi variado entre 0,5 e 1,5 vezes o número de preditores. Durante o treinamento foi utilizado o método proposto por Reed (1993) onde os dados foram divididos em conjunto de treinamento e de validação, para se evitar o *super-ajuste*. Cada estação ficou com pelo menos 5000 amostras para treinamento e 2000 para validação, onde foi observada a sazonalidade garantindo-se a proporcionalidade entre meses secos e chuvosos. Foram efetuados 200 treinamentos pelo algoritmo de retropropagação de Levenberg-Marquardt (TLM) para cada subconjunto de cada estação, e calculados os erros quadráticos médios (RMSE) e coeficientes de correlação (CC) entre a saída da RNA e seu respectivo alvo no conjunto de validação. Os treinamentos eram realizados de forma iterativa, onde os conjuntos de treinamento e validação eram divididos aleatoriamente respeitando a proporção definida para cada estação. O conjunto de validação variou entre 50% a 60% dos dados por estação, em função do tamanho da série disponível. Os erros e correlações foram avaliados sobre toda a série de dados.

Sobre cada subconjunto de erros e correlações foi aplicado o teste de Kolmogorov-Smirnov de duas amostras, comparando os subconjuntos de resultados adjacentes.

Este teste possui a vantagem de ser não-paramétrico e permite inferir para certo nível de significância se duas amostras podem ser oriundas da mesma distribuição de probabilidades (Wilks, 2006). Com isso foi possível identificar até que ponto é vantajoso aumentar o número de preditores, ou seja, progredir para o próximo subconjunto. Com isso definiu-se o melhor subconjunto de preditores para cada estação anemométrica.

Para avaliar a possibilidade de adotar técnicas paramétricas no desenvolvimento do trabalho foi aplicado um teste de normalidade das distribuições de RMSE e CC pela comparação entre as distribuições acumuladas de probabilidade, normal e amostrada, utilizando-se intervalos de confiança com 5% de significância. Caso a curva normal esteja contida no intervalo de confiança da amostra, testes paramétricos são apropriados.

#### **6.4. Desenvolvimento do Modelo em RNAs**

A definição da arquitetura possui um componente empírico relevante, portanto foram realizados testes de sensibilidade com configurações distintas para as duas torres anemométricas. Admite-se que as melhores configurações para estas também desempenharão melhor para as outras torres.

##### **6.4.1. Definição do Número de Neurônios Ocultos**

Uma vez que o número de preditores para a RNA já foi definido, o próximo passo foi determinar o número adequado de neurônios na camada oculta. Quanto maior este número, maior a flexibilidade da rede para se ajustar aos resultados, no entanto maior também é o número de mínimos locais, maior a dificuldade de treinamento e maior o risco de super-ajuste da rede. Para averiguar como o número ideal de neurônios varia na medida em que se altera a quantidade de preditores, foram testados 6 subconjuntos de preditores para cada estação, com os neurônios da camada oculta variando de 1,5 a 4 vezes o número de preditores. Os outros dois subconjuntos de preditores referentes ao Limiar 01 e Limiar 02 não foram utilizados, pois acarretariam um tempo excessivo de processamento devido ao número maior de preditores. Após 100 treinamentos para cada subconjunto, os melhores valores

de RMSE e CC analisados em função do número de preditores. O intervalo ideal é aquele em que um aumento no número de neurônios não acarreta uma melhora significativa nos resultados.

#### 6.4.2. Configuração dos Testes de Sensibilidade

Posteriormente, uma série de testes de sensibilidade foi efetuada para aperfeiçoar os principais parâmetros da rede. O subconjunto de preditores utilizado foi o de Limiar 5 e o número de neurônios ocultos foi restringido entre 1 e 2 vezes o número de preditores. Estes testes estão descritos pela Tabela 6.5 e inclui variações na defasagem máxima entre preditor e alvo, na função de entrada da RNA, no tratamento inicial dos preditores, no tipo de climatologia, no algoritmo de treinamento e no alvo final da rede.

A função de entrada da RNA pode ser do tipo 'Mín/Máx', quando os preditores são escalonados na entrada, ou do tipo 'std', quando são normalizados pelo desvio padrão. O tratamento inicial dos preditores pode variar entre o uso de anomalias absolutas, anomalias normalizadas ou anomalias escalonadas. É importante diferenciar o escalonamento efetuado nas entradas da RNA do efetuado no tratamento dos preditores. Apesar de ambos limitarem os valores mínimos e máximos entre -1 e 1, no tratamento dos preditores a média das anomalias é preservada em zero, evitando um desbalanceamento do sinal, conforme indica a Equação 6.16. A climatologia diária usada para o cálculo das anomalias pode ser única para todo o ano ou mensal. Os algoritmos de treinamento variam entre o TLM, o de retropropagação por gradientes conjugados (TCG) e o de retropropagação por regularização bayesiana (TBR). O alvo pode variar entre a anomalia do vento observado (Vobs), ou o pela anomalia do erro entre o vento previsto (Vp) e o observado.

$$\phi'_{esc} = \begin{pmatrix} se \phi' > 0 & \phi' / \phi'_{máx} \\ se \phi' < 0 & \phi' / \phi'_{mín} \end{pmatrix} \quad (6.16)$$

Onde  $\phi'$  é a anomalia de uma variável qualquer e os sub-índices representam seus valor máximo, mínimo e escalonado.

Tabela 6.5 – Testes de sensibilidade realizados sobre a rede neural.

	<b>Defas. Máxima</b>	<b>Função Entrada</b>	<b>Tratamento Preditores</b>	<b>Climatologia</b>	<b>Algoritmo Treinam.</b>	<b>Alvo</b>
<b>Controle</b>	+/- 3h	Mín/Máx	Anom Norm	Diária	TLM	Vobs 50m
<b>Teste 1</b>	+/- 6h	Mín/Máx	Anom Norm	Diária	TLM	Vobs 50m
<b>Teste 2</b>	+/- 12h	Mín/Máx	Anom Norm	Diária	TLM	Vobs 50m
<b>Teste 3</b>	+/- 3h	Std	Anom Norm	Diária	TLM	Vobs 50m
<b>Teste 4</b>	+/- 3h	Mín/Máx	Anom Absol	Diária	TLM	Vobs 50m
<b>Teste 5</b>	+/- 3h	Mín/Máx	Anom Escal	Diária	TLM	Vobs 50m
<b>Teste 6</b>	+/- 3h	Mín/Máx	Anom Norm	Diária / Mês	TLM	Vobs 50m
<b>Teste 7</b>	+/- 3h	Mín/Máx	Anom Norm	Diária	TCG	Vobs 50m
<b>Teste 8</b>	+/- 3h	Mín/Máx	Anom Norm	Diária	TBR	Vobs 50m
<b>Teste 9</b>	+/- 3h	Mín/Máx	Anom Norm	Diária	TLM	Erro V50m

Foram realizados 100 treinamentos para cada configuração e para ambas as torres, com exceção do Teste 9, para o qual foram realizados apenas 20 testes pois o algoritmo de treinamento utilizado demanda um tempo de processamento muito superior aos demais. Cada teste altera apenas um parâmetro da RNA e seus resultados foram comparados ao grupo Controle para se avaliar a influência de cada alteração separadamente.

Da mesma forma que na seleção de preditores, os resultados foram comparados através de um teste de Kolmogorov-Smirnov com 5% significância para duas amostras, sendo uma destas sempre a Controle.

## 6.5. Avaliação dos Modelos Obtidos

As melhores configurações de RNA obtidas para cada uma das torres foram aplicadas sobre as demais estações. A configuração obtida para SJCA foi denominada INT, como referência ao interior do continente, e a configuração obtida para ROTE foi chamada LIT, em referência ao litoral. As redes foram treinadas para cada uma das 8 torres e a avaliação final foi efetuada sobre todo o conjunto de dados disponível.

### 6.5.1. Desempenho na Previsão de Ventos

Os dados modelados e observados foram dispostos em forma de dispersão analisando-se o desempenho de cada um dos modelos, Eta/PV, regressão passo-a-passo (REGP) e redes neurais (RNA). Foram calculados o coeficiente de correlação ( $R$ ), o erro médio (ME) e a raiz do erro quadrático médio (RMSE) para o vento no nível de 50 m. Para a torre de PALM o alvo de comparação foi o nível de 70 m. As Equações 6.17, 6.18 e 6.19 descrevem o método de cálculo de cada um destes índices, onde  $P$  são dados previstos e  $O$  os dados observados. Foi gerado também o gráfico em escala temporal para estas séries de dados, possibilitando a visualização do desempenho de cada modelo.

O desempenho do modelo de previsão por RNAs foi avaliado também através do cálculo do *Skill*, que é uma medida do ganho de um modelo em relação a outro para um determinado índice estatístico, conforme descrito na Equação 6.20. No caso foi avaliado a redução do RMSE do modelo RNA em relação ao Eta/PV.

$$R = \frac{\sum_{i=1}^N (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^N (P_i - \bar{P})^2} \cdot \sqrt{\sum_{i=1}^N (O_i - \bar{O})^2}} \quad (6.17)$$

$$ME = \frac{1}{N} \sum_{i=1}^N (P_i - O_i) \quad (6.18)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2} \quad (6.19)$$

$$Skill(RMSE, Eta) = \frac{RMSE_{RNA} - RMSE_{Eta}}{RMSE_{Eta}} \quad (6.20)$$

### 6.5.2. Desempenho na Previsão de Potência Eólica

Outra etapa interessante realizada foi a avaliação do peso de cada preditor no modelo final obtido pela RNA. Esta é uma informação valiosa para o aprimoramento das técnicas de refinamento estatístico assim como para orientar modificações nos modelos de mesoescala.

No que se refere às possíveis aplicações, as previsões de vento fornecidas pelo modelo Eta/PV e pelo refinamento por RNAs foram utilizados em conjunto com a curva de potência de um aerogerador de 2,1 MW, uma faixa de potência que vêm sendo largamente empregada nos parques eólicos brasileiros, com o intuito de avaliar os erros nas estimativas de energia produzida no horizonte +12 a +36 horas. A curva de potência empregada está ilustrada na Figura 6.3.

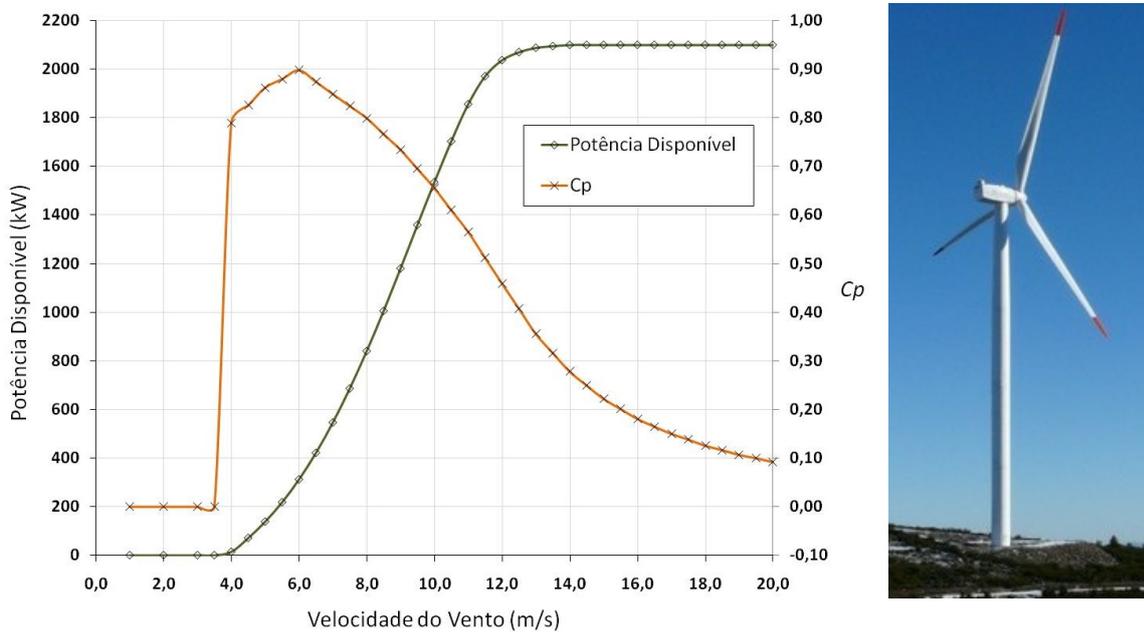


Figura 6.3 – Aerogerador de 2,1MW utilizado nas estimativas e sua curva de potência.

Fonte: <[www.thewindpower.net](http://www.thewindpower.net)>

Do ponto de vista microeconômico, a partir destas informações foi possível calcular o impacto que uma previsão melhor pode acarretar sobre a receita de um empreendimento eólico em um cenário de mercado livre de energia. No mercado livre, que já vigora hoje para outras fontes de energia como a hidráulica e a termoeletrica, o excedente de energia em relação ao contratado pode ser comercializado no curto-prazo, onde os preços variam em função da oferta e demanda do sistema elétrico (CCEE, 2011).

Já com relação ao gerenciamento do sistema elétrico nacional, a redução do erro nas estimativas de energia eólica disponível a curto-prazo para uma região pode auxiliar no gerenciamento do parque gerador brasileiro. Foi feita uma estimativa do ganho de confiabilidade de energia produzida, em MWh, com base na capacidade eólica atual instalada no nordeste brasileiro. Este ganho é uma consequência da redução das incertezas (variância) associadas à nova distribuição de probabilidade dos erros provenientes da RNA.

## 7 RESULTADOS

A seguir são descritos os resultados encontrados durante cada etapa de desenvolvimento do trabalho. O conceito de resultados aqui exposto compreende não somente aqueles que têm uma influência direta no modelo final de rede neural, mas também aqueles que foram obtidos de maneira complementar, seja para validar cada etapa do processo, seja para agregar informações que possam ser de alguma utilidade para trabalhos futuros.

### 7.1. Verificação da Assimilação e do Tratamento dos Dados

Primeiramente, após o processo de recorte e conversão dos dados do modelo Eta/PV, preocupou-se em eliminar a possibilidade de qualquer incompatibilidade de codificação ou sistema operacional que viesse degradar os dados. Para averiguar sua consistência espacial e temporal foram gerados dois mapas para o mesmo domínio e instante, sendo que o primeiro foi obtido a partir das saídas originais do Eta/PV através do software GrADS, e o segundo pelo MATLAB® a partir das saídas convertidas e assimiladas. A Figura 7.1 mostra que os dados estão íntegros após o processamento.

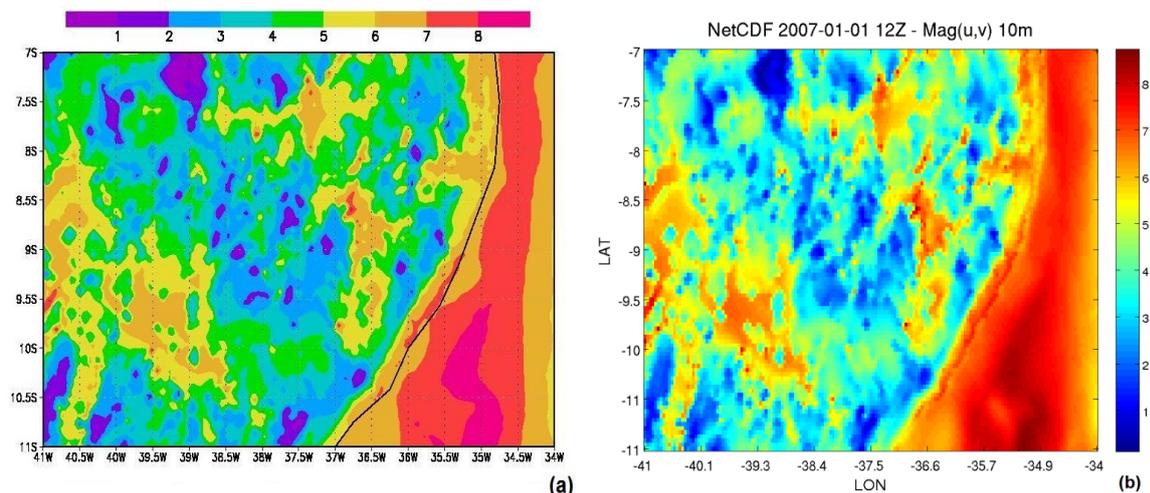


Figura 6.1 – Comparação entre mapas de magnitude do vento a 10m vento a 10m antes (a) e após (b) o processo de recorte, conversão e assimilação.

Na etapa seguinte, com a qualificação aplicada sobre os dados do modelo Eta/PV eliminou-se variáveis espúrias do conjunto de preditores. A Figura 7.2 mostra a

porcentagem de dados espúrios por variável para as estações de SJCA e ROTE. Entre as variáveis excluídas estão o coeficiente de arrasto superficial (*cdfsfc*), a disponibilidade de umidade em superfície (*mstavdl*), o coeficiente de troca em superfície (*sfexcsfc*) e os fluxos zonal e meridional de momentum em superfície (*uflxsfc* e *vflxsfc*), que não foram calculados em instante algum, possivelmente por opção de configuração inicial do modelo. As variáveis de fluxo de calor latente em superfície (*lhtflsfc*) e de vento zonal na camada de 850 hPa (*ugrdprs850*) tiveram uma taxa de falha bem discreta e foram mantidas no modelo após terem seus dados suspeitos substituídos pelas respectivas médias. Inspeccionando com detalhe as variáveis *lhtflsfc* e *mstavdlr* percebe-se que a taxa de falha é diferente entre os pontos de grade, o que sugere que pode haver alguma instabilidade numérica induzida por fatores locais. Os critérios para esta qualificação foram os limites máximos e mínimos fisicamente possíveis para cada variável, obtidos empiricamente com base na análise do conjunto de saídas do modelo. Dados falhos indicavam valores não-factiveis, tipicamente -999, sendo facilmente identificados. Para taxas de falha inferiores a 2% estes valores foram substituídos pela média simples da variável para cada estação.

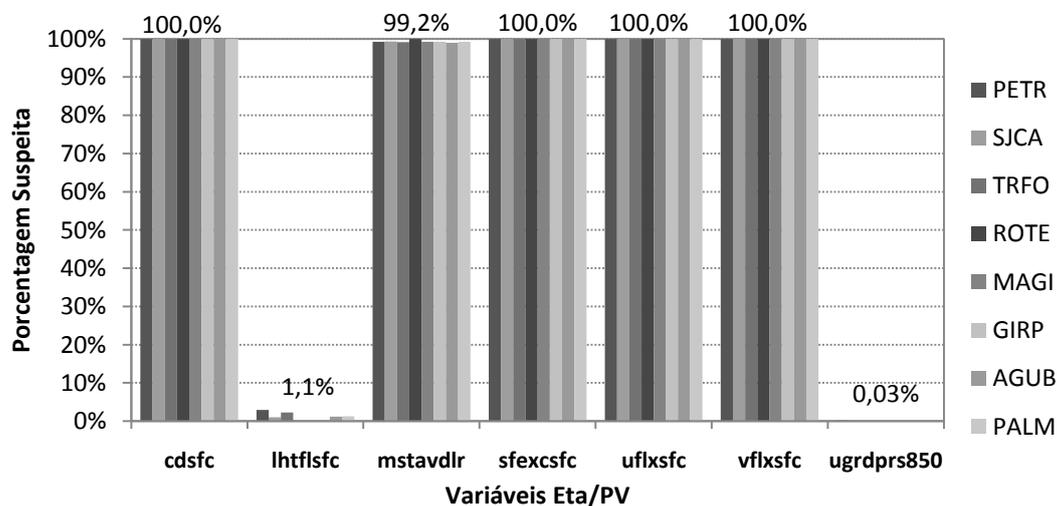


Figura 7.2 – Variáveis suspeitas durante a qualificação dos dados do modelo Eta/PV. Rótulos representam as taxas médias de falha para cada variável.

Quanto aos dados anemométricos sua qualificação foi imprescindível, pois foram detectadas diversas fontes de erro cujo diagnóstico é difícil por inspeção visual.

Neste caso o algoritmo utilizado foi baseado nos critérios do projeto SONDA, acrescido de critérios adicionais como já mencionado anteriormente. A Figura 7.3 ilustra a porcentagem de dados suspeitos para cada torre anemométrica e para cada nível. A altura dos níveis para as estações varia conforme descrito anteriormente pela Tabela 6.1.

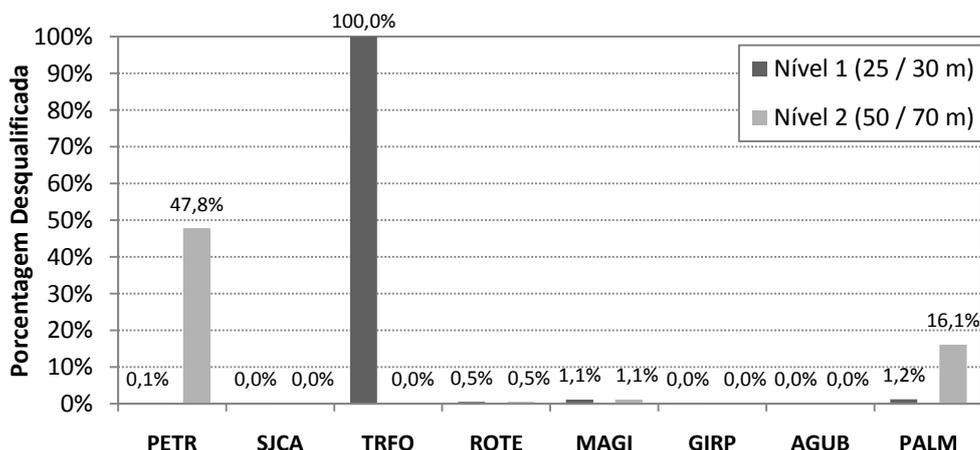


Figura 7.3 – Porcentagem dos dados anemométricos desqualificados por estação para cada nível.

Como o descarte dos dados anemométricos suspeitos é uma medida extrema, principalmente quando o tamanho das séries é um limitante para o trabalho, optou-se por recuperar os dados a 50 m a partir do nível inferior sempre que possível. Esta reparação foi auditada para se ter uma idéia dos erros cometidos com os ajustes verticais. A porcentagem de dados corrigidos a 50 m, a rugosidade e o erro médio absoluto (MAE) para cada estação são mostrados na Figura 7.3. Os valores dos erros cometidos foram aceitáveis e não comprometem o desenvolvimento do modelo em RNAs, uma vez que nestas extrapolações verticais as correlações praticamente não se alteram. A magnitude do BIAS ficou abaixo de 0,1 m/s e o MAE ficou limitado a 0,65 m/s. Nota-se que as estações de PETR e PALM possuem maiores desvios nos perfis verticais, seja pela presença de obstáculos, pela localização geográfica ou por um desajuste dos sensores. Em PETR sabe-se que há obstáculos no entorno da torre, enquanto em PALM o porte da vegetação e o relevo levemente acidentado no entorno podem ter contribuído para estes desvios.

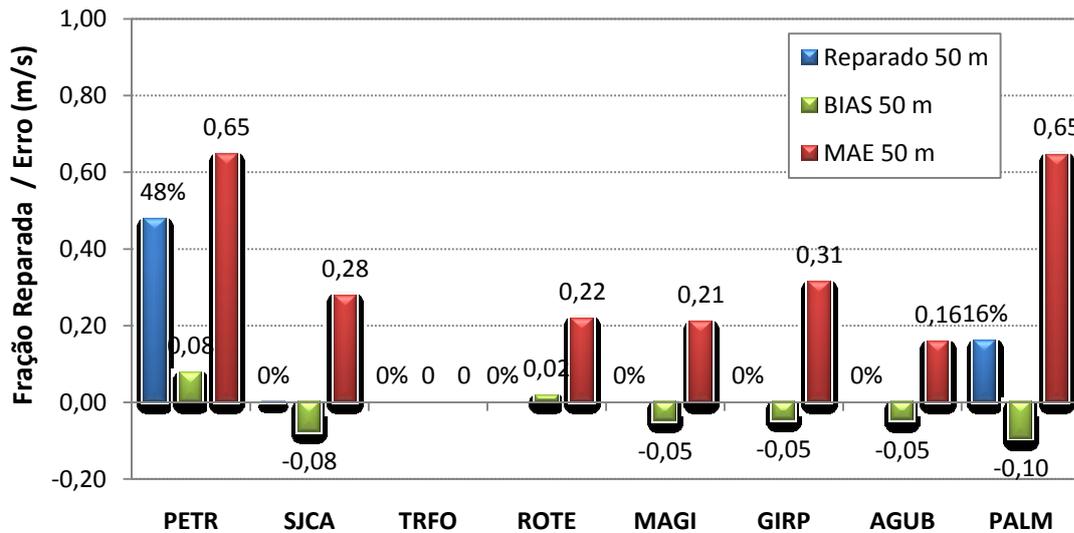
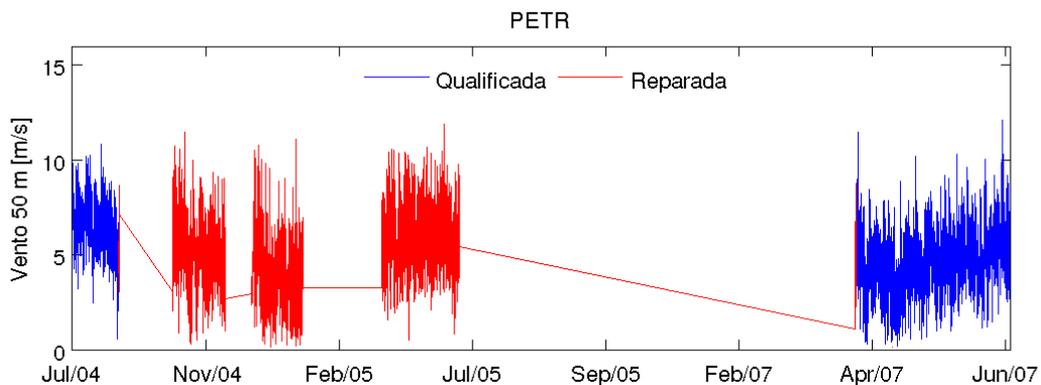


Figura 7.4 – Fração de dados anemométricos reparados no nível mais elevado e respectivos erros (MAE e BIAS) cometidos. Para a estação PALM o nível reparado foi o de 70 m.

As séries de dados qualificadas e reparadas foram analisadas visualmente quanto a sua consistência. Percebe-se que as estações do projeto SONDA possuem séries mais longas, embora descontinuidades ocorram com certa freqüência. As séries de Alagoas são curtas, porém bastante consistentes. A Figura 7.5 ilustra as séries de PETR e PALM onde os reparos estão destacados. Percebe-se que o padrão dos dados praticamente não se altera. É importante ressaltar que de toda forma estas séries representam bem a variabilidade local do vento a 50 m, portanto para todos os efeitos admite-se que estas sejam as séries efetivamente observadas.



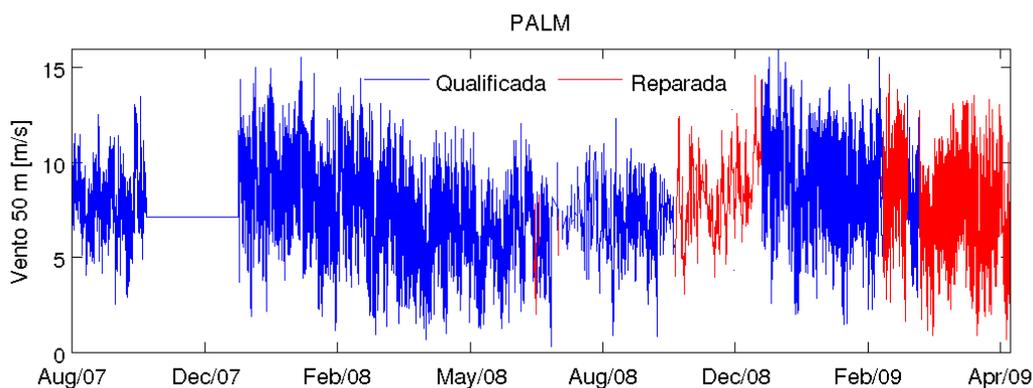


Figura 7.5 – Séries temporais de magnitude observada do vento para as estações de PETR (50 m) e PALM (70 m) após reparos.

Na sequência foi realizada a sincronização entre os dados modelados e observados, o que demandou certa cautela, pois alguns arquivos de dados anemométricos estavam em hora local e em um padrão de data numérica diferente dos modelos. A fim de validar este processo foram geradas figuras correlacionando a data do modelo com a data das torres para os períodos onde há dados anemométricos disponíveis. Estes resultados estão expostos pela Figura 7.6.

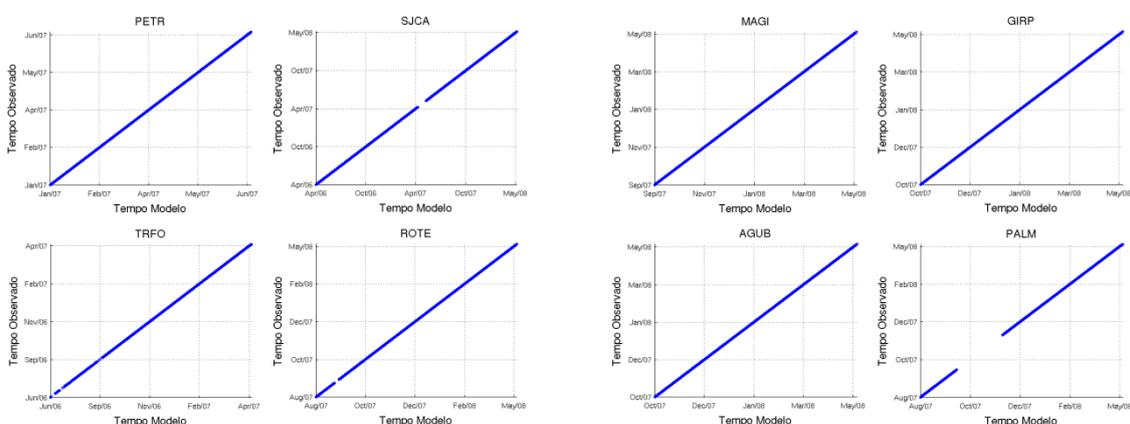


Figura 7.6 – Validação da sincronização entre os dados modelados e observados para todas as estações.

## 7.2. Climatologia

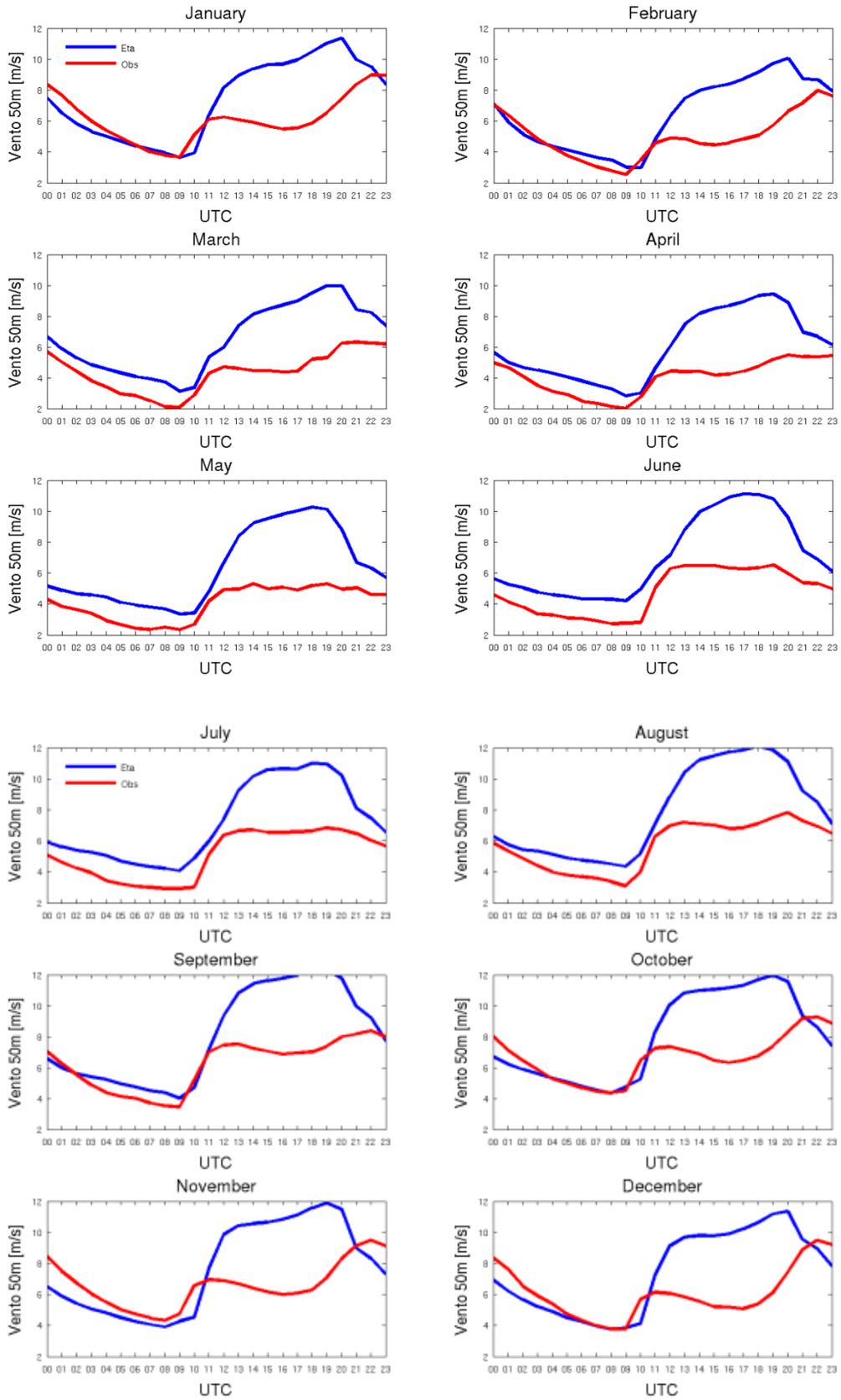
Foram removidas do conjunto de dados tanto as variáveis do Eta/PV não aprovadas na qualificação quanto as linhas de dados anemométricos suspeitos e não passíveis de reparos. Foram calculadas as climatologias diárias para o modelo e para as observações. Tendo em vista que após a sincronização o período de dados

anemométricos fica limitado pelo período do Eta/PV, a climatologia observada foi calculada utilizando todo o conjunto de dados anemométricos disponíveis, independentemente da sincronização.

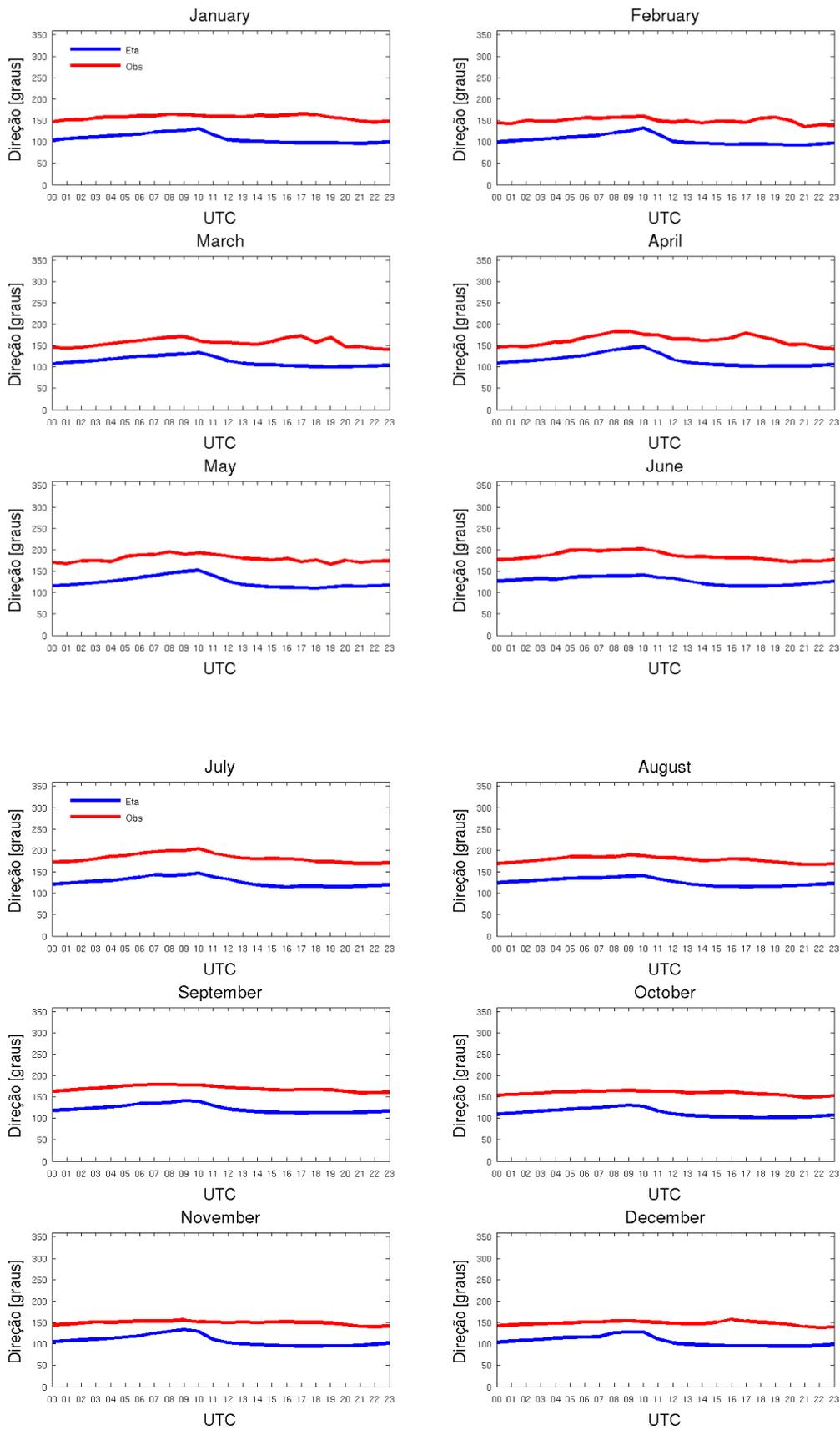
### **7.2.1. Ciclos Diários**

Na Figura 7.7 estão expostos os ciclos diários de intensidade e direção do vento para cada mês em SJCA e ROTE. Estas foram as duas estações utilizadas para os testes seguintes de desenvolvimento do modelo em RNAs. Este tipo de climatologia diária, dividido mensalmente, foi utilizado em um dos testes de sensibilidade do modelo em RNAs como será visto adiante, no entanto, o padrão para as demais rodadas foi de uma única climatologia diária válida para todo o ano.

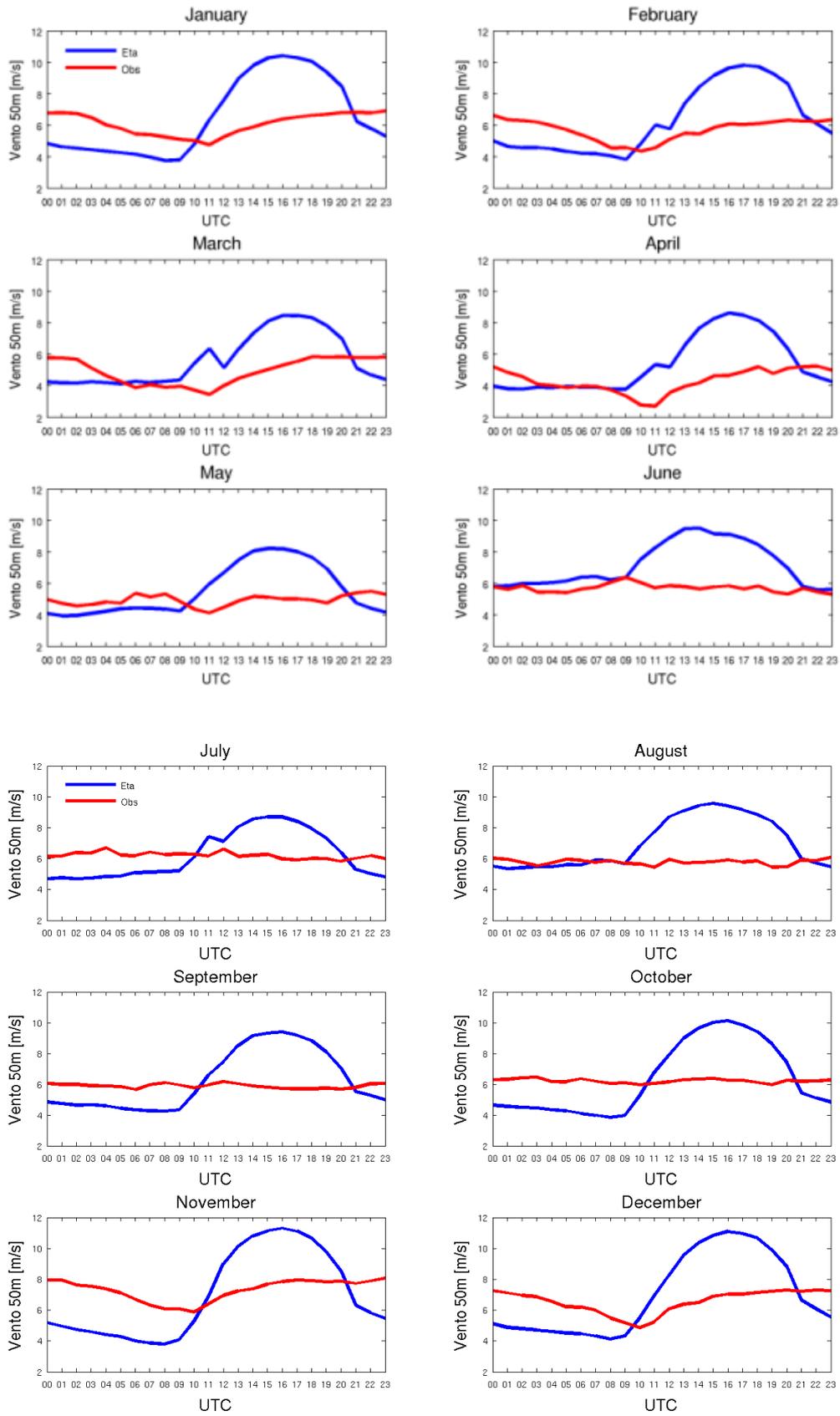
Observando a Figura 7.7 nota-se que o Eta/PV em média, detecta com precisão o instante em que o vento se intensifica em SJCA apesar de não representar bem as magnitudes, superestimando os máximos diários. Outro ponto interessante é a existência de frequências mais altas de oscilação da magnitude e direção do vento observado nos meses chuvosos em SJCA, possivelmente um efeito da convecção profunda. Este mesmo comportamento não é aparente nos dados do Eta/PV. Para ROTE percebe-se uma diferença maior no ciclo diário, onde o Eta/PV tipicamente superestima a brisa marítima e não detecta as oscilações na direção do vento nos meses chuvosos.



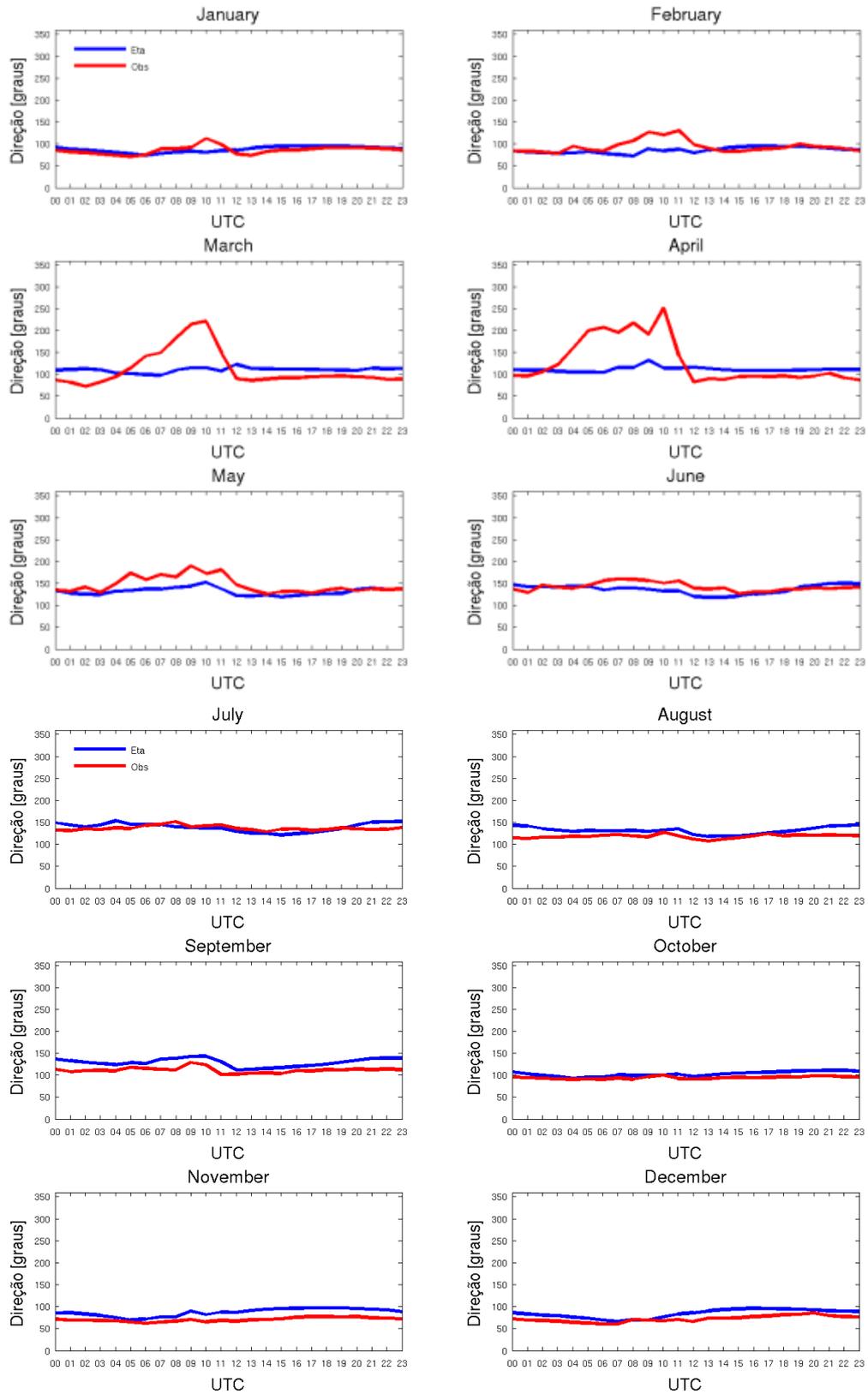
(a)



(b)



(c)



(d)

Figura 7.7 – Ciclos diários de intensidade e direção do vento ao longo do ano para SJCA (a) e (b); e ROTE (c) e (d).

### 7.2.2. Determinação do Comprimento de Rugosidade

A comparação entre os ciclos de vento a 50 m só foi possível devido à extrapolação dos valores do modelo Eta de 10 m para 50 m. Como no modelo só há um nível de altura do vento, que está a 10 m, utilizou-se dos dados anemométricos para determinar a rugosidade local e a partir daí calcular o vento no nível de 50 m. Para que esta técnica funcione adequadamente é preciso que o relevo e a vegetação local sejam razoavelmente homogêneos o que nem sempre é verdade. A técnica empregada consistiu no método de minimização do erro quadrático médio obtido pela auto-validação do perfil, conforme já explicado na metodologia.

A Figura 7.8 mostra os valores obtidos para as rugosidades ao redor de cada torre em escala logarítmica. O cálculo não foi dividido por setores por conveniência, uma vez que o vento predominante não varia muito de direção nesta região.

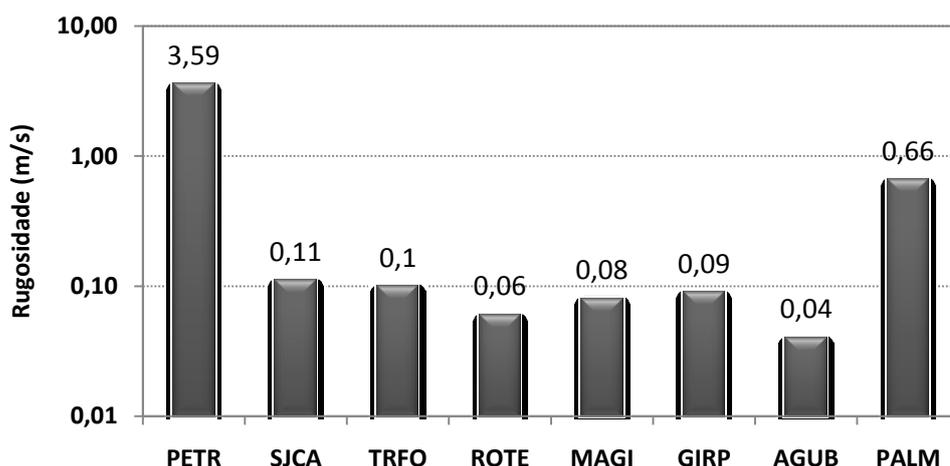


Figura 7.8– Valores obtidos para a rugosidade  $Z_0$  ao redor de cada torre em escala logarítmica.

Analisando os valores obtidos e as imagens disponíveis para a vegetação no entorno de cada estação, nota-se uma coerência nos resultados. SJCA está em um descampado com arbustos e inclinação suave; TRFO não possui dados a 25 m portanto adotou-se 0,2; ROTE e MAGI estão próximos do litoral; GIRP e AGUB aparentam estar cercados por vegetação rasteira; restaram PETR e PALM que apresentam obstáculos em seu entorno que justificam os altos valores encontrados, conforme já mencionado anteriormente.

### 7.3. Preditores Seleccionados

A partir das anomalias de cada variável calculou-se a correlação cruzada, onde variáveis redundantes, um total de 15, foram eliminadas. Foram obtidas as correlações máximas para cada variável em função das possíveis defasagens com o alvo. Tais defasagens a princípio foram limitadas entre +/- 1 hora. A Figura 7.9 ilustra a máxima correlação encontrada para cada variável para as duas estações em questão. Percebe-se que em média as correlações são inferiores a 0,5 o que a princípio podem ser considerados valores baixos. No entanto é preciso ter em mente que estas são correlações entre anomalias, onde o ciclo diário, que é o maior responsável pela variância destas séries, foi subtraído. Isto significa que para as séries absolutas estas correlações podem se amplificar. Outra ressalva a ser feita é quanto à precariedade deste método na detecção de não-linearidades, conforme discutido no capítulo anterior.

Na definição dos preditores foram adotados 8 limiares de correlação os quais deram origem aos respectivos subconjuntos de potenciais preditores. A função destes limiares é efetuar uma primeira filtragem nas variáveis. Estes subconjuntos foram posteriormente refinados através de uma regressão linear passo-a-passo, que tem como principal objetivo eliminar preditores incapazes de incrementar a variância explicada dos predictandos, ou seja, preditores linearmente dependentes de outros.

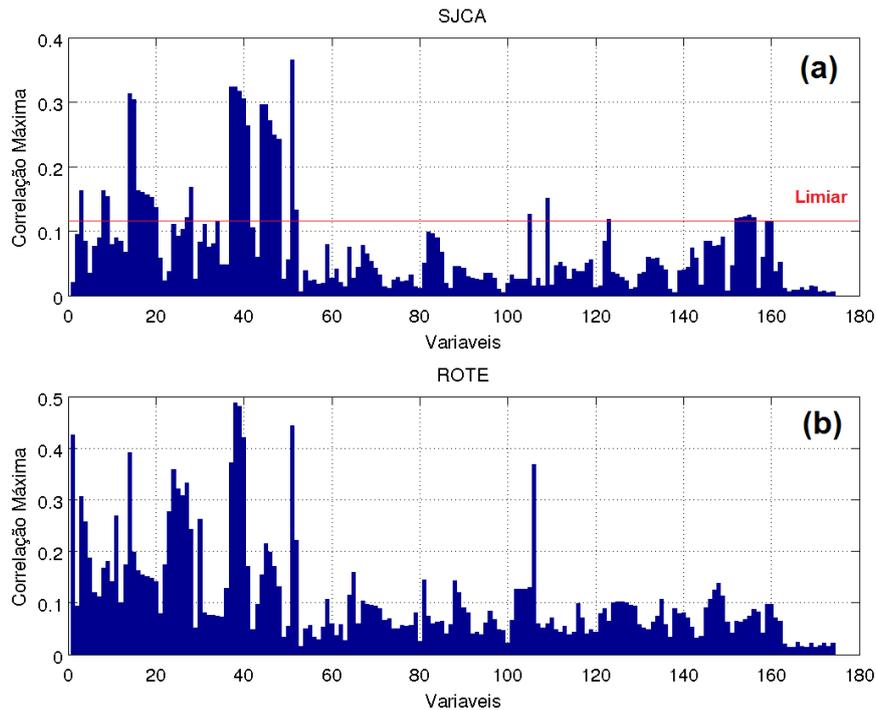


Figura 7.9 – Correlações máximas obtidas para cada variável e exemplo de limiar adotado para definir os subconjuntos de preditores para SJCA (a); correlações máximas obtidas para ROTE (b).

As correlações mostradas anteriormente representam as máximas entre diferentes defasagens. No entanto, uma forma de eliminar a influência do ruído na definição das defasagens foi aplicá-las somente se significassem um aumento superior a 10% na correlação para cada variável. A Figura 7.10 mostra como ficaram as defasagens para um dos subconjuntos de preditores para SJCA (+/- 1 h) e ROTE (+/-11h). Nota-se que para ROTE a maior parte das variáveis tendem a se atrasar em relação à previsão, o que não se torna um problema desde que se tenha saídas do modelo para um horizonte posterior ao da previsão desejada.

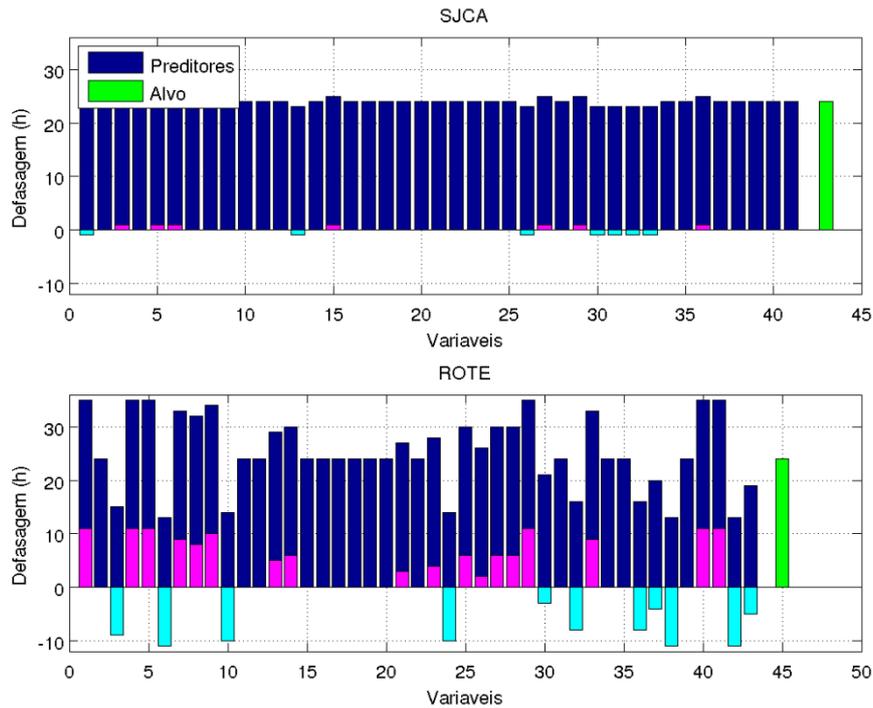


Figura 7.10 – Defasagens de máxima correlação para cada variável em SJCA e ROTE. O zero representa o momento da previsão e as barras azuis a posição no tempo dos 24 valores utilizados para prever o vento (alvo).

Nas duas primeiras etapas de seleção foram adotados critérios gradativamente mais restritivos (limiares) para que a análise fosse conservadora, uma vez que as potencialidades não-lineares de cada preditor são desconhecidas. A última etapa da seleção, porém, é determinante, pois os subconjuntos preditores são treinados em uma rede neural, onde têm seu desempenho comparado. Para um número de treinamentos suficientemente grande pode-se identificar estatisticamente quais subconjuntos representam uma melhora efetiva para o modelo em RNAs. A Tabela 7.1 lista o número de preditores obtidos por subconjunto após a aplicação dos limiares e da regressão passo-a-passo (REGP). Na sequência, a Figura 7.11 compara a dispersão anterior e posterior à REGP para as estações de SJCA e ROTE. Cabe ressaltar que para todos os subconjuntos o vento modelado pelas REGPs foi incluído como um preditor adicional.

Percebe-se que o número de preditores candidatos é reduzido significativamente com o uso de REGP. Uma característica interessante é que quanto maior o número de preditores fornecidos à REGP, maior o número de preditores selecionados, o que

mostra que mesmo para regressões lineares estes preditores agregam valor. A comparação das dispersões com o vento observado entre Eta/PV e REGP também evidenciam a potencialidade desta técnica. Para REGP o viés foi praticamente eliminado e as correlações aumentaram significativamente para ROTE, de 0,34 para 0,70, e modestamente para SJCA, de 0,58 para 0,63. Adiante será visto que esta dificuldade em modelar SJCA é freqüente.

Tabela 7.1 - Número de preditores por subconjunto antes e após REGP.

Número Preditores	Lim_1	Lim_2	Lim_3	Lim_4	Lim_5	Lim_6	Lim_7	Lim_8
<b>SJCA - Inicial</b>	144	124	103	82	66	52	42	37
Após REGP	54	49	37	35	28	25	22	17
<b>ROTE - Inicial</b>	151	145	120	86	63	47	38	32
Após REGP	56	47	34	28	26	25	21	18

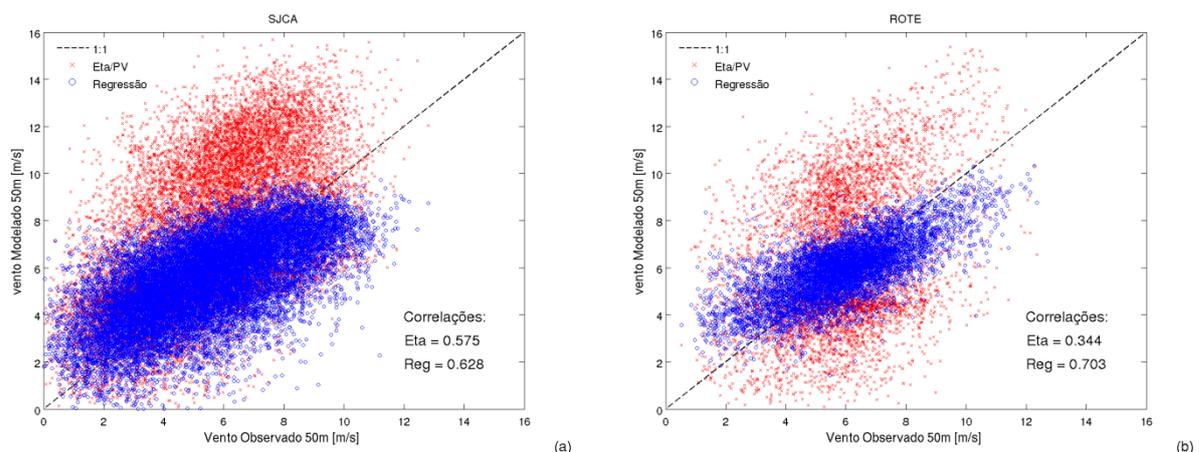
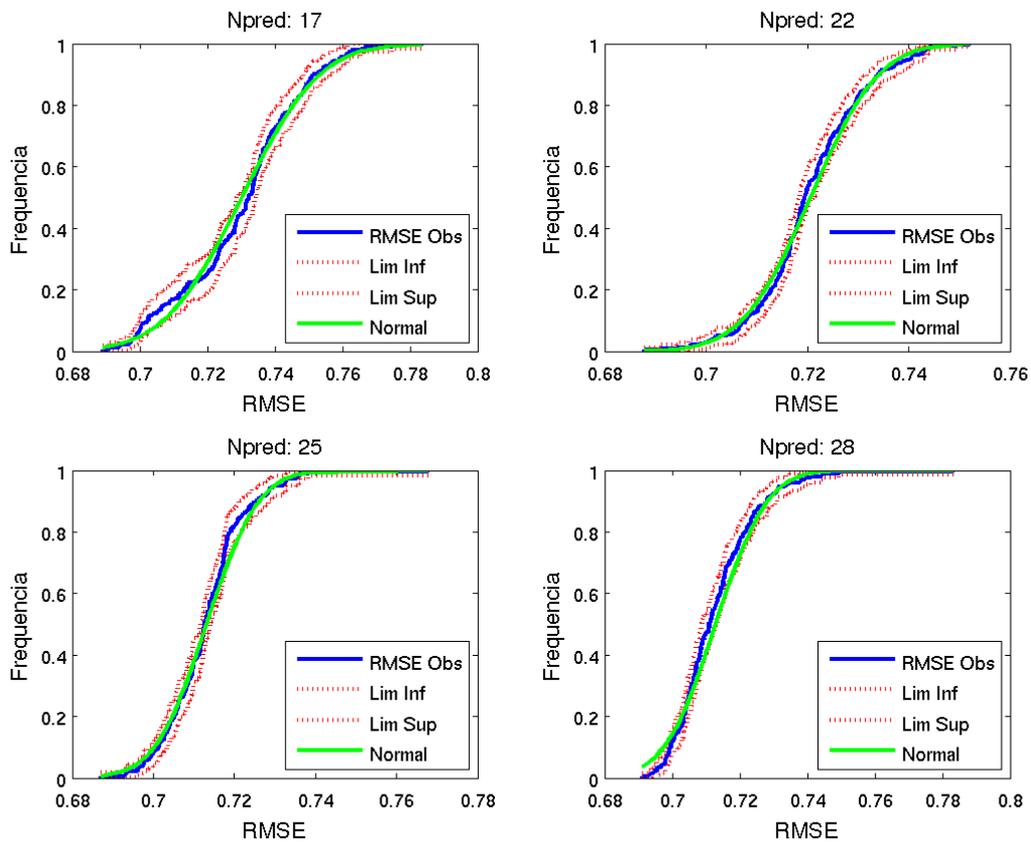


Figura 7.11 – Comparação entre as dispersões obtidas entre o vento observado e o modelado pelo Eta/PV (vermelho) e as saídas da REGP (azul) para o vento a 50 m.

Após a realização de 200 treinamentos para cada subconjunto de preditores da Tabela 7.1 utilizando uma rede neural com configuração preliminar foram obtidas distribuições de correlações ( $R$ ) e raiz do erro quadrático médio (RMSE). Estas distribuições foram analisadas para que se pudesse identificar qual seu padrão de comportamento para cada subconjunto. Na Figura 7.12 é mostrada uma comparação entre os intervalos de confiança para 5% de significância das distribuições de probabilidade acumulada de  $R$  e RMSE contra uma distribuição

normal padrão. Nota-se que apesar de algumas regiões não sobrepostas, em geral existe uma compatibilidade entre as curvas, o que indica certo grau de aleatoriedade da solução. Isto pode ser um sinal de que o treinamento da RNA está convergindo sistematicamente em torno de soluções semelhantes, as quais seriam função apenas das condições iniciais dos pesos e vieses, que são aleatórias.



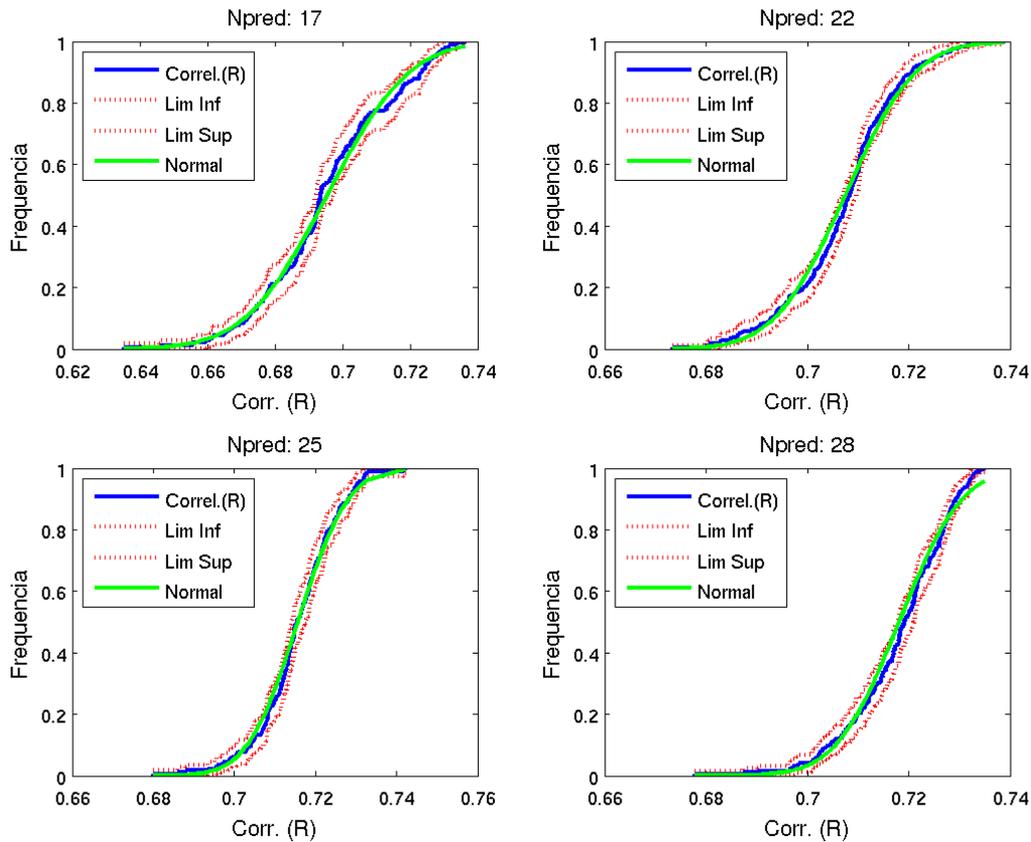


Figura 7.12 – Comparação entre as distribuições de probabilidade acumulada de RMSE e  $R$  e contra uma distribuição normal padrão para quatro subconjuntos de  $N$  preditores ( $N_{pred}$ ) da estação de SJCA.

Estas distribuições de  $R$  e RMSE são mostradas em um gráfico em função do número de preditores dos respectivos subconjuntos, conforme ilustra a Figura 7.13. Para que se pudesse identificar aquelas que são estatisticamente melhores foi efetuada uma comparação entre as distribuições adjacentes utilizando o teste de Kolmogorov-Smirnov (KSTest) para duas amostras, adotando estatística de uma banda em 5% de significância. A hipótese nula do KSTest é de que as amostras são iguais. No caso do RMSE foi testada a hipótese alternativa de que a amostra seguinte é menor que a atual, enquanto para  $R$  naturalmente foi testado o oposto. A Tabela 7.2 lista os valores obtidos para o teste, sendo que 1 indica que a hipótese alternativa é verdadeira, enquanto 0 indica que as amostras são iguais.

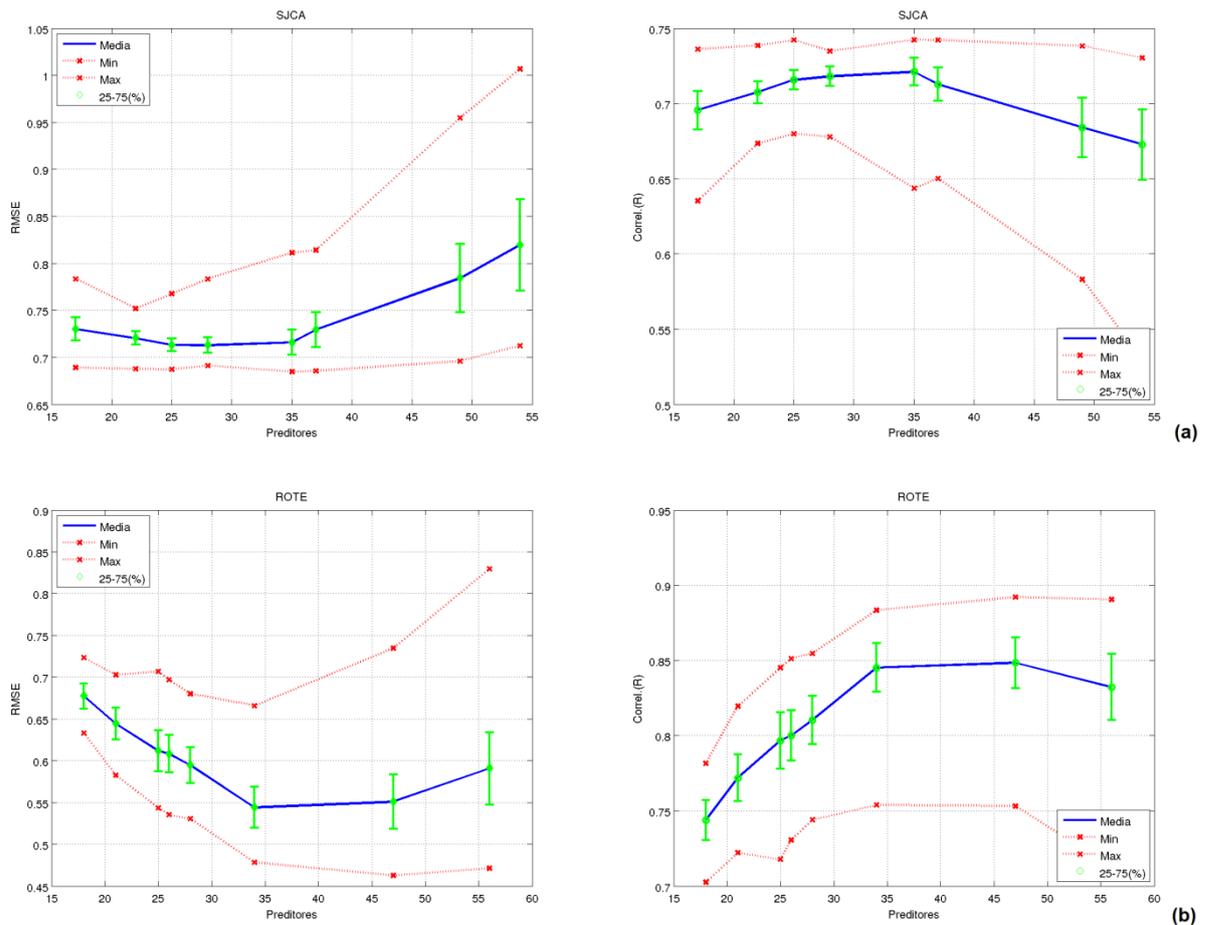


Figura 7.13 – RMSE e R em função do número de preditores de cada subconjunto SJCA (a); e ROTE (b).

Tabela 7.2 – Resultados dos testes de Kolmogorov-Smirnov entre os subconjuntos de preditores adjacentes para SJCA e ROTE.

Subconjuntos		8/7	7/6	6/5	5/4	4/3	3/2	2/1
SJCA	RMSE	1	1	1	1	0	0	0
	R	1	1	1	1	0	0	0
ROTE	RMSE	1	1	1	1	1	0	0
	R	1	1	1	1	1	0	0

A leitura da Tabela 7.2 deve ser feita de tal forma que para uma comparação entre os subconjuntos X/Y, o valor 1 indica que os parâmetros RMSE ou R de Y são melhores que os de X. A partir dos resultados da Tabela 7.2 define-se o subconjunto de preditores mais adequado para cada uma das estações. Para SJCA nota-se que um aumento no número de preditores leva a uma melhora significativa no modelo até o subconjunto correspondente ao Limiar 4. Já para ROTE esta melhora avança até o subconjunto de Limiar 3. Apesar dos limiares distintos obtidos para as estações o número de preditores ao final desta etapa ficou em 35 e 34, para SJCA e

ROTE respectivamente. Há diferenças entre os preditores de cada estação e isto é compreensível, pois como estas estações estão em locais geograficamente distintos (litoral e sertão), possuem uma dinâmica atmosférica diferenciada que permite correlações diferentes para algumas variáveis. Como este número final de preditores ainda vai sofrer pequenas alterações devido aos testes de sensibilidade, a comparação entre os conjuntos será feita somente ao final do capítulo.

#### **7.4. Neurônios na Camada Oculta**

Uma vez definidos os preditores procedeu-se para a definição do número ideal de neurônios na camada oculta. Para dar maior representatividade para este teste foram realizados treinamentos utilizando 5 subconjuntos de preditores para dar maior robustez à análise. Os subconjuntos com maior número de preditores não foram utilizados para se reduzir o tempo de processamento da análise. O desempenho da rede foi avaliado variando-se o número de neurônios de 1,5 a 4 vezes o número de preditores, onde foram realizados 50 testes para cada configuração. Os resultados foram colocados em um gráfico como função da razão entre neurônios ocultos e preditores conforme ilustra a Figura 7.14.

Percebe-se que para ROTE existe uma tendência de redução do RMSE na medida em que se aumenta o número de neurônios ocultos até aproximadamente a razão de 2,5. Para SJCA esta tendência não é evidente e o erro não se altera, e eventualmente até cresce com o aumento da razão. Este comportamento é atípico e precisa ser mais bem estudado para se identificar as restrições que impedem a redução do erro. No entanto, cabe ressaltar que na medida em que se aumenta o número de neurônios ocultos, por um lado aumenta-se a flexibilidade da rede, porém esta se torna mais difícil de treinar. Isto é compreensível, pois o aumento do número de pesos sinápticos, aumenta também o número de mínimos relativos, e portanto mais iterações são necessárias para que a rede convirja aproximadamente para uma mesma solução. Neste sentido SJCA está em desvantagem pois é também a estação com a mais longa série de treinamento.

Outro fator que pode estar limitando o desempenho da rede para SJCA é a própria complexidade dos fenômenos atmosféricos locais, que pode levar a uma incapacidade de prognóstico do Eta/PV para algumas escalas de circulação relevantes. Com isso a RNA empregada não possui recursos suficientes para extrair

mais informações dos preditores. Uma alternativa seria o emprego de redes neurais com mais de uma camada oculta e/ou recorrentes, mas que estão fora do escopo deste trabalho. Com base nestes resultados, optou-se por trabalhar sempre com um intervalo de neurônios ocultos entre 2 e 3 vezes o número de preditores, poupando tempo para realizar mais iterações e assim garantir a convergência do erro.

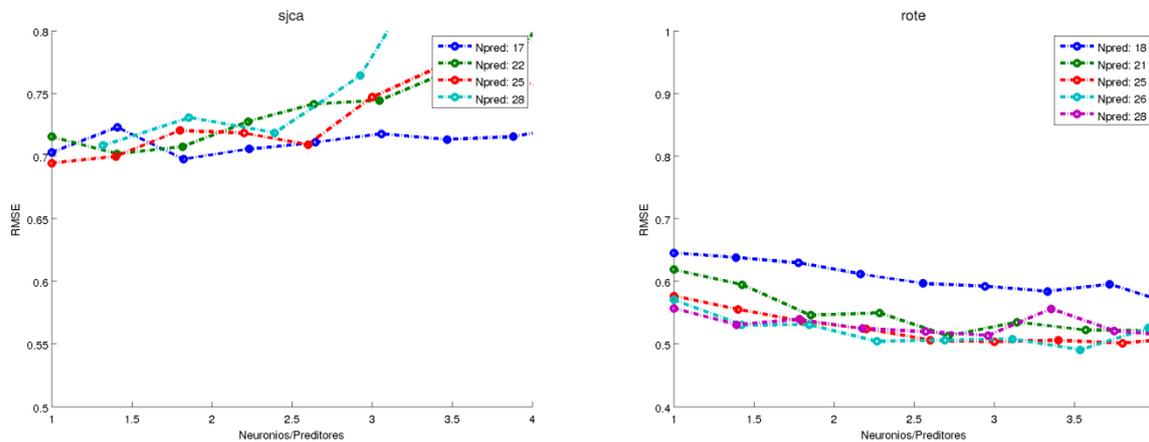


Figura 7.14 – RMSE em função da razão entre neurônios e preditores para SJCA e ROTE.

## 7.5. Testes de Sensibilidade

Conforme ilustrado anteriormente na Tabela 6.4, foram realizados testes de sensibilidade para identificar quais opções de pré-processamento e configuração da RNA trazem ganhos de desempenho para o modelo. Após 100 treinamentos para cada configuração os valores de RMSE e  $R$  para SJCA e ROTE foram colocados em um gráfico tipo *boxplot* com intervalos de 25% a 75% das respectivas distribuições, como ilustra a Figura 7.15. O número de neurônios ocultos para estes testes foi limitado a 2 vezes o número de preditores para redução do tempo de processamento.

Da mesma forma que na seleção de preditores, adotou-se o teste de Kolmogorov-Smirnov para duas amostras, utilizando-se estatística de uma banda apenas em 5% de significância. A diferença neste caso é que cada configuração foi comparada ao grupo controle para se avaliar o ganho relativo de cada modificação. A hipótese nula do KSTest é de que as amostras são iguais e a alternativa é de que a segunda amostra é melhor (maior ou menor, em função de RMSE ou  $R$ ). A Tabela 7.3 lista os valores obtidos para o teste, lembrando que 0 indica que a hipótese nula é verdadeira, enquanto 1 indica que a alternativa é verdadeira.

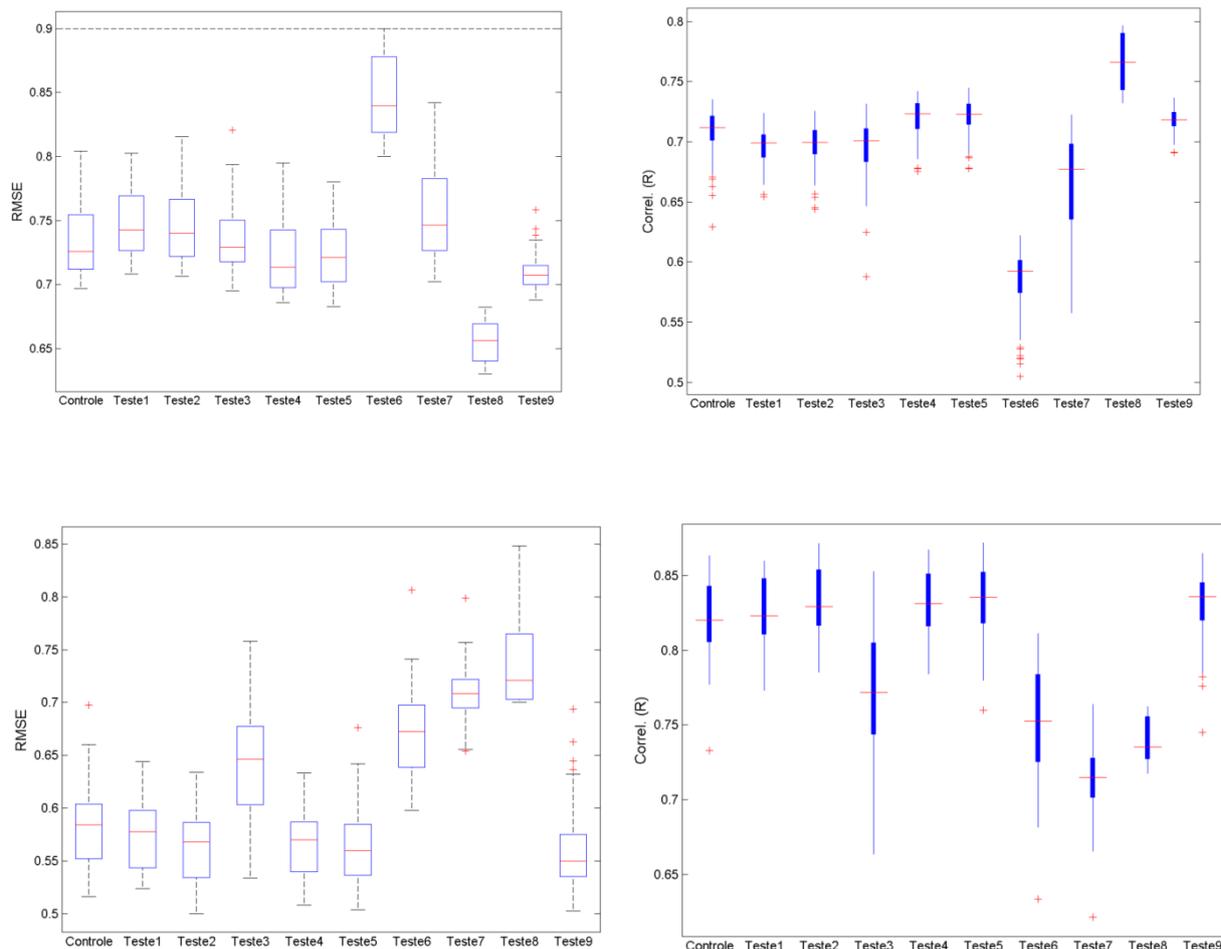


Figura 7.15 – Testes de sensibilidade para cada configuração de RNA para SJCA e ROTE, acima e abaixo respectivamente.

Tabela 7.3 – Resultados dos testes de Kolmogorov-Smirnov para as diferentes configurações (testes de sensibilidade) para SJCA e ROTE.

KSTEST		Tst1	Tst2	Tst3	Tst4	Tst5	Tst6	Tst7	Tst8	Tst9
SJCA	RMSE	0	0	0	1	1	0	0	1	1
	R	0	0	0	1	1	0	0	1	1
ROTE	RMSE	0	1	0	1	1	0	0	0	1
	R	0	1	0	1	1	0	0	0	1

Os resultados mostram que as alterações propostas nos Testes 4, 5 e 9 foram benéficas para ambas as estações. A diferença entre elas reside apenas no Teste 2, que foi melhor para ROTE enquanto o Teste 8 foi melhor para SJCA. A Tabela 7.4 detalha quais modificações foram introduzidas em cada um destes testes. Percebe-se que o uso de anomalias não-normalizadas e escalonadas no tratamento dos preditores melhora o desempenho da rede. Outra conclusão importante é que com a

utilização do erro da previsão ( $V_{prev} - V_{obs}$ ) como alvo, as correlações tendem a aumentar, o que está de acordo com o encontrado por Chou *et al.* (2007) para previsões de temperatura.

A primeira diferença se deu na defasagem máxima entre preditor e peditando. Uma defasagem maior, de +/- 12 horas, foi benéfica para ROTE. Esta pode estar associada a algum padrão sinótico de escala superior a 12 horas que têm uma influência importante sobre o vento durante sua aproximação ou afastamento. Uma análise mais detalhada das defasagens combinadas ao peso de cada preditor no modelo final da RNA pode trazer evidências do tipo de fenômeno que pode estar relacionado.

Por fim, a outra diferença restante se deu em relação algoritmo de treinamento utilizado. Embora o algoritmo TLM apresente uma convergência mais rápida para todos os casos, o algoritmo TBR teve um desempenho superior para SJCA. Conforme exposto anteriormente o algoritmo TBR utiliza regularização bayesiana, evitando que os pesos e bias assumam valores muito elevados e contribuindo para a generalização da RNA. A estação de SJCA é a que possui a mais longa série de treinamento o que indica que problemas de generalização não são esperados. A razão provável deste resultado pode estar relacionada ao processo de treinamento adotado pelo TBR, que por ser regularizado a princípio não adota uma parcela de validação para o critério de parada antecipada, podendo eventualmente “decorar” alguns padrões mesmo para uma série de dados extensa como a de SJCA. Como uma maneira de contornar esta deficiência, nos treinamentos posteriores pelo algoritmo TBR uma parcela dos dados foi separada para uma avaliação independente dos erros finais.

Um aspecto interessante destes resultados é a ausência de benefício com o uso do Teste 6. Neste teste as anomalias foram calculadas com base em ciclos diários mensais, e portanto representam apenas oscilações intra-sazonais, o que a princípio pode representar um avanço já que as previsões são de curto-prazo. Porém, de acordo com os resultados o prejuízo ao desempenho pela ausência do sinal sazonal é mais relevante do que uma eventual melhoria introduzida pelo teste. Uma alternativa seria combinar ambos, introduzindo uma variável sintética para representar o ciclo anual, assim como foi feito para o ciclo diário.

Tabela 7.4 – Testes de sensibilidade aprovados pelo KSTest para SJCA e ROTE.

	Aprovado para:	Defasagem Máxima	Tratamento Preditores	Algoritmo Treinamento	Alvo da RNA
<b>Controle</b>		<b>+/- 3h</b>	<b>Anom Norm.</b>	<b>TLM</b>	<b>Vobs 50m</b>
Teste 2	ROTE	<b>+/- 12h</b>	Controle	Controle	Controle
Teste 4	SJCA/ROTE	Controle	<b>Anom Absol.</b>	Controle	Controle
Teste 5	SJCA/ROTE	Controle	<b>Anom Escal.</b>	Controle	Controle
Teste 8	SJCA	Controle	Controle	<b>TBR</b>	Controle
Teste 9	SJCA/ROTE	Controle	Controle	Controle	<b>Vprev - Vobs 50m</b>

Um resultado esperado foi a melhoria com o Teste 5. Na assimilação dos dados pelas redes neurais, é conveniente que estes sejam escalonados entre -1 e 1, pois este é o domínio original das funções de transferência embutidas em cada neurônio. A princípio valores não-escalonados demandam um ajuste maior dos vieses internos à rede, retardando a convergência dos erros. A configuração de RNA utilizada faz automaticamente o escalonamento das variáveis, no entanto não preserva sua média em zero. A melhoria observada foi o resultado de se usar anomalias com média nula preservada. A Figura 7.16 ilustra este escalonamento.

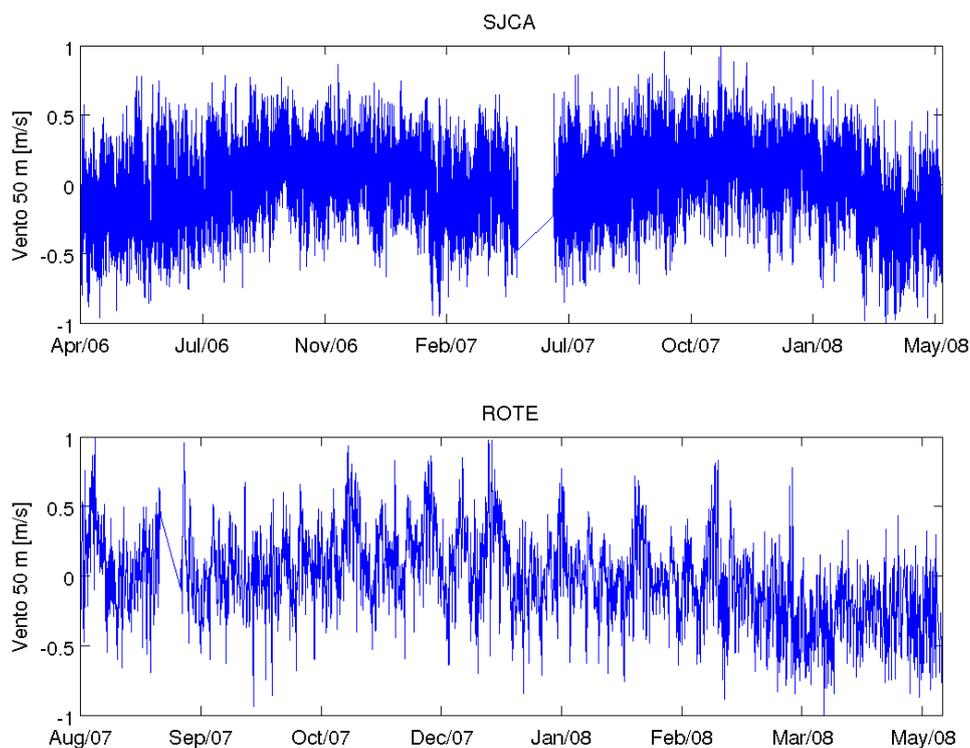


Figura 7.16 – Séries temporais de anomalias escalonadas para SJCA e ROTE.

## 7.6. Avaliação das Previsões

zz Uma vez que foram definidos os preditores, a arquitetura final e a configuração da RNA, pode-se dizer que a metodologia de refinamento estatístico do vento por RNAs está completa. Sua aplicação se deu a princípio sobre as duas estações utilizadas no seu desenvolvimento, para as quais foram realizados 100 treinamentos em cada. A Figura 7.17 ilustra a dispersão final obtida entre o vento observado e o modelado pela RNA a 50 m, e sua comparação com as dispersões anteriores.

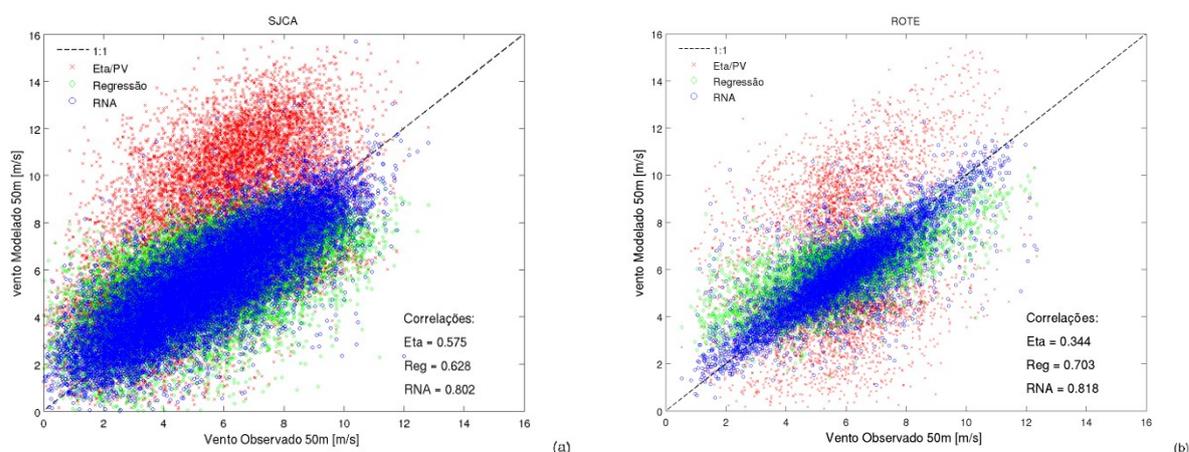


Figura 7.17 – Dispersão final obtida entre o vento observado e o modelado pela RNA a 50 m para SJCA e ROTE.

Como pode ser observado, as dispersões finais obtidas para a RNA têm um aspecto mais esbelto do que as obtidas anteriormente para o ETA/PV e REGP. O valor de  $R$  melhorou significativamente o que indica que a metodologia funciona bem para estas estações, embora o desempenho para SJCA não tenha sido tão bom quanto para ROTE. Na etapa seguinte aplicou-se a metodologia para as outras 6 estações, efetuando-se outros 100 treinamentos para cada uma delas.

Todavia, como foram desenvolvidas duas configurações possíveis de RNA, uma para ROTE, de agora em diante denominada RNA\_LIT (como referência ao litoral), e a outra para SJCA, denominada RNA\_INT (em referência ao interior), foi preciso testar ambas. A Figura 7.18 ilustra os melhores resultados obtidos para as demais estações.

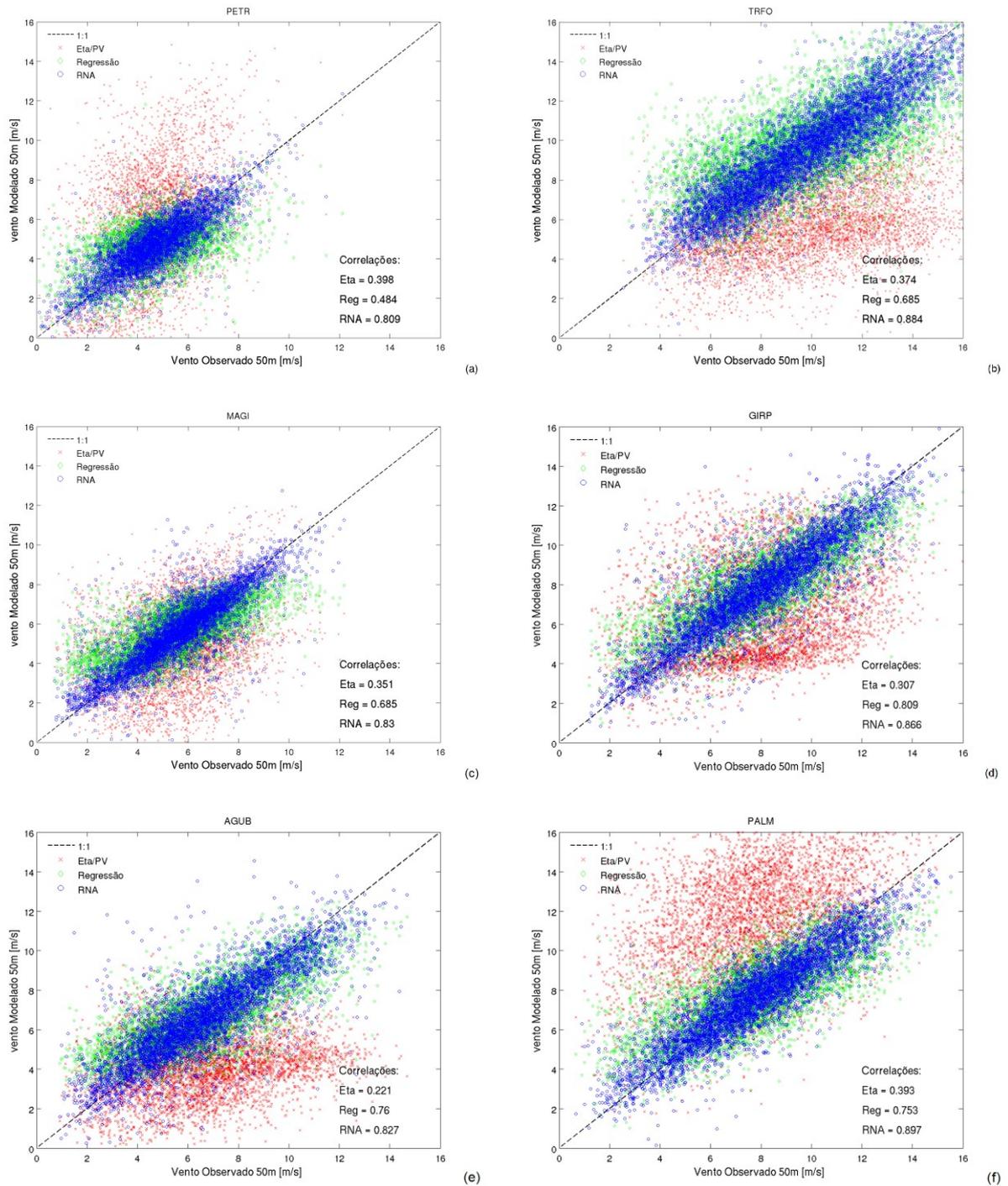


Figura 7.18 – Dispersão final obtida entre o vento observado e o modelado pelo Eta/PV, REGPe RNA a 50 m para as demais estações. PETR (a); TRFO (b); MAGI (c); GIRP (d); AGUB (e); PALM (f).

Os resultados apresentados mostram que as RNAs agregam uma quantidade significativa de variância explicada aos modelos estatísticos em relação às REGPs para todas as estações analisadas. Cabe lembrar ainda que estas correlações foram obtidas a partir da avaliação de todo o período de dados. Têm-se ainda que a comparação entre o desempenho da configuração RNA\_LIT e RNA\_INT mostrou

que a RNA\_LIT é superior para as demais estações, inclusive para a SJCA, como ilustra a Figura 7.19. Uma razão para isso pode ser porque nos testes de sensibilidade foram feitas alterações individuais, enquanto no modelo final estas alterações foram sobrepostas. Com isso é possível que na nova configuração da RNA\_LIT a melhoria tenha se amplificado e superado a RNA\_INT. Além disso, nos resultados finais, o desempenho para o algoritmo de treinamento TBR foi aferido em relação a um conjunto de validação independente, o que não ocorreu nos testes de sensibilidade.

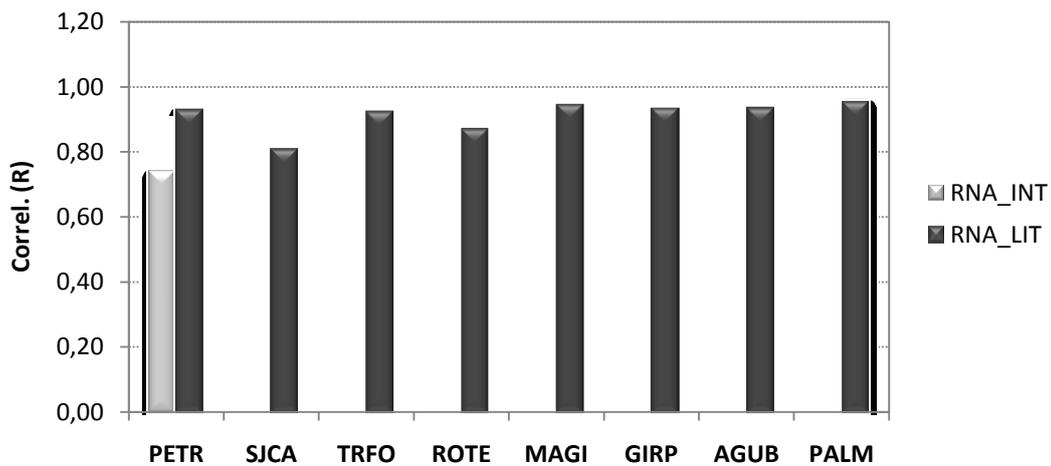


Figura 7.19 – Comparação entre as correlações obtidas entre as anomalias de vento a 50 m para as configurações RNA\_LIT e RNA\_INT para todas as estações.

Para uma visualização compacta dos resultados obtidos para cada um dos modelos e para cada estação foram gerados gráficos comparativos para  $R$  e RMSE conforme ilustram as Figuras 7.20 e 7.21 respectivamente. Com isto fica evidente que o desempenho das RNAs é superior ao das REGPs para todas as estações. Em alguns casos esta melhoria é modesta, como nas estações de ROTE e GIRP, enquanto para os outros esta é mais significativa. É importante ressaltar a melhora obtida para a estação de SJCA, que é substancial já que possui a mais longa série de dados com quase 3 anos. A correlação máxima foi 0,90 para a estação de PALM enquanto os menor RMSE foi obtido para PETR em aproximadamente 0,95 m/s.

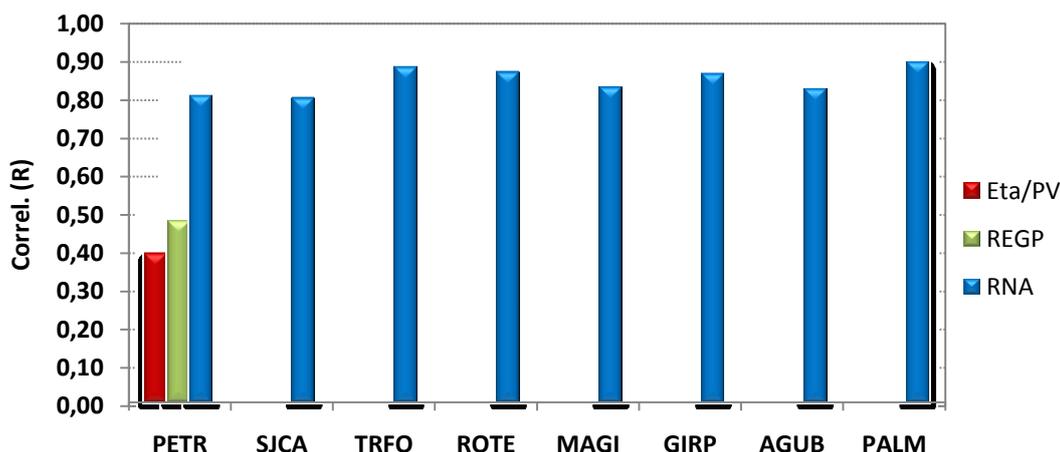


Figura 7.20 – Comparação entre as correlações  $R$  entre vento observado e modelado a 50 m pelo Eta/PV, REGP e RNA para todas as estações.

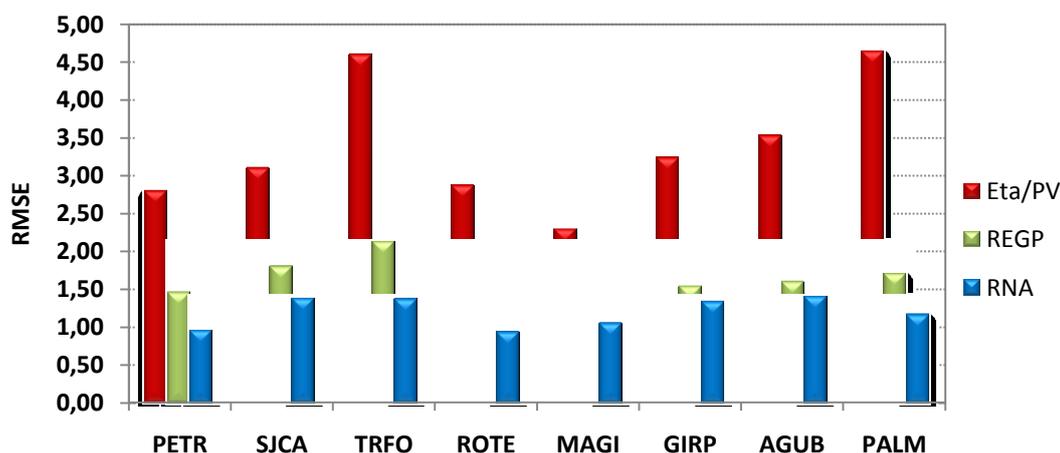


Figura 7.21 – Comparação entre os RMSE a 50 m pelos modelos Eta/PV, REGP e RNA para todas as estações.

Como uma forma de quantificar o incremento de desempenho introduzido pelas REGPs e RNAs na modelagem estatística do vento foi calculado o índice de desempenho ou *Skill Score* da RNA em relação aos modelos Eta/PV e REGP. O resumo destes resultados está apresentado na Tabela 7.5 incluindo os valores do desvio RMSE, da correlação  $R$  e do viés (BIAS) para cada estação. O cálculo destes índices foi efetuado utilizando-se toda a série de dados disponível.

A partir desta avaliação final dos resultados é possível determinar o quão efetivo é o método proposto na modelagem do vento para diversos locais do nordeste brasileiro. Têm-se ainda a possibilidade de inferir de que forma o tamanho das

séries anemométricas utilizadas e também características locais, podem afetar o desempenho do método proposto.

Os melhores desempenhos foram registrados para PALM, TRFO, ROTE e GIRP, com valores de R chegando à 0,90. O menor valor de RMSE, no entanto foi registrado para ROTE, possivelmente devido ao ciclo diário pouco acentuado da região. O modelo Eta/PV teve seu melhor desempenho para SJCA, onde a correlação chegou a 0,58. Percebe-se que os valores de SKILL (RMSE) em relação ao Eta/PV têm pequena dispersão, estando entre 55% e 70%, o que sugere que quanto maior o erro do modelo Eta, maior também será o erro da RNA. O inverso ocorre para o SKILL (R) em relação ao Eta/PV, onde há grande dispersão dos valores. Isto é justificável se admitirmos que as correlações R tendem a um limite de previsibilidade para o tipo de modelagem adotado entre 0,80 e 0,90. Portanto quanto melhor são os resultados do Eta/PV menor é o ganho relativo possível para as RNAs.

Os maiores valores de SKILL (RMSE) e SKILL (R) em relação à REGP ocorreram para as estações de PETR, TRFO e PALM que possuem o entorno mais heterogêneo, o que demonstra a aplicabilidade das RNAs a locais de topografia acidentada e vegetação irregular.

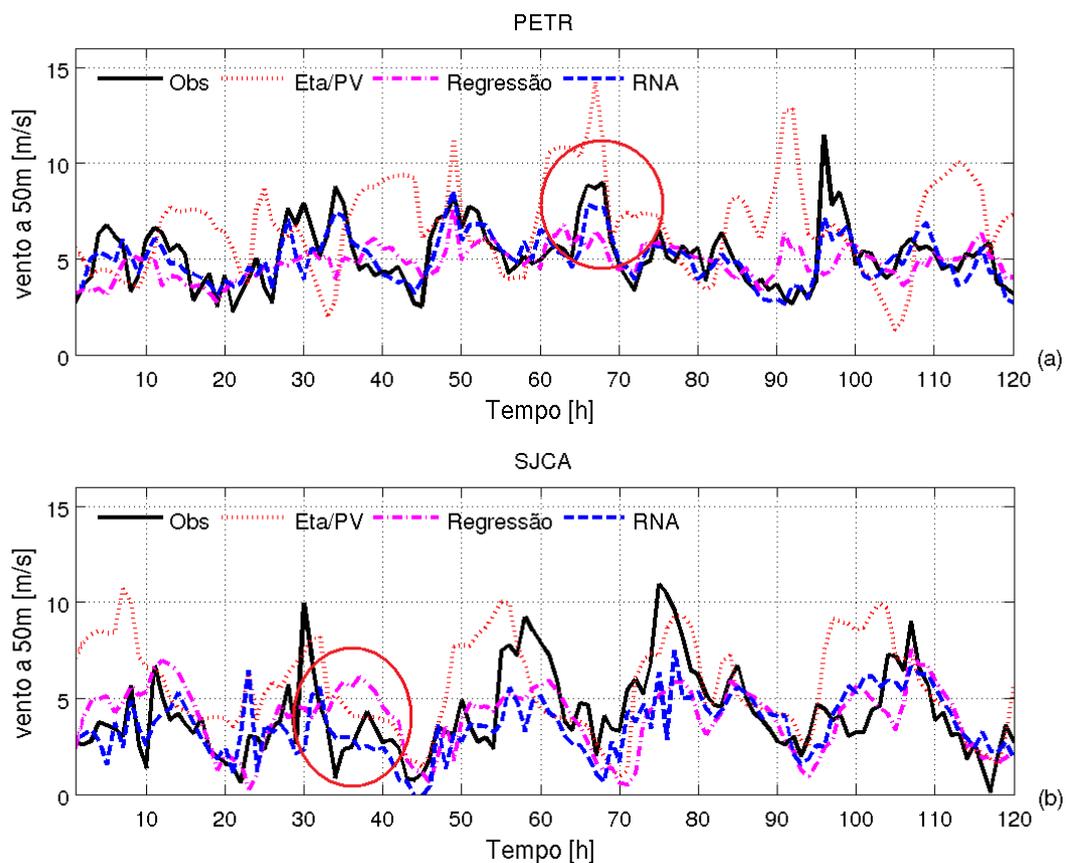
Tabela 7.5 – Resumo dos resultados obtidos pela modelagem por REGP e RNA para cada estação anemométrica.

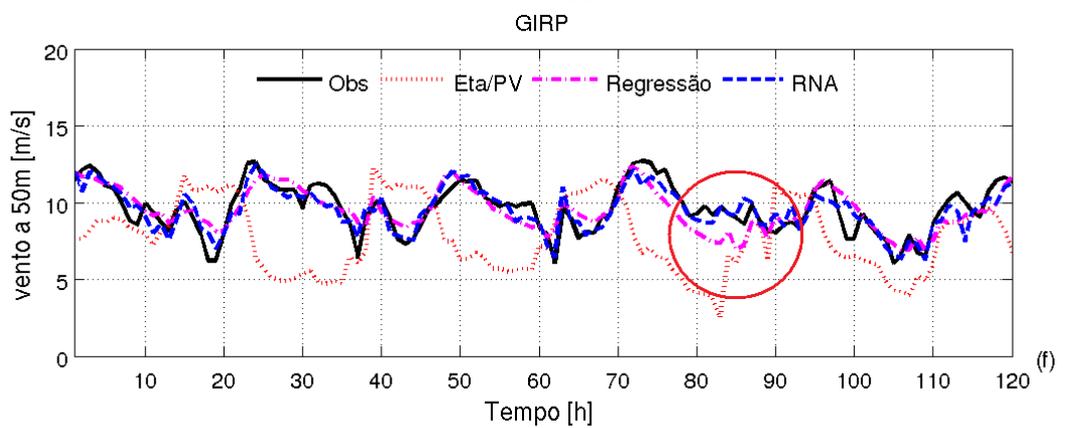
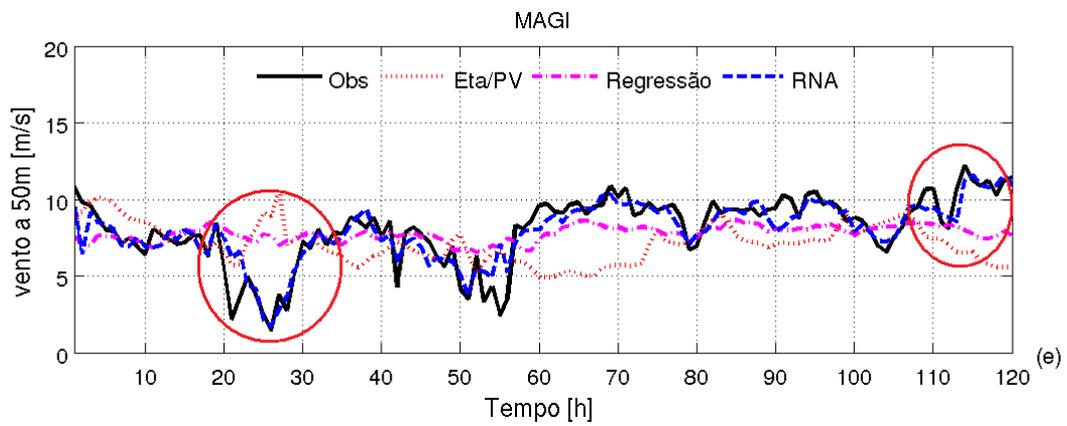
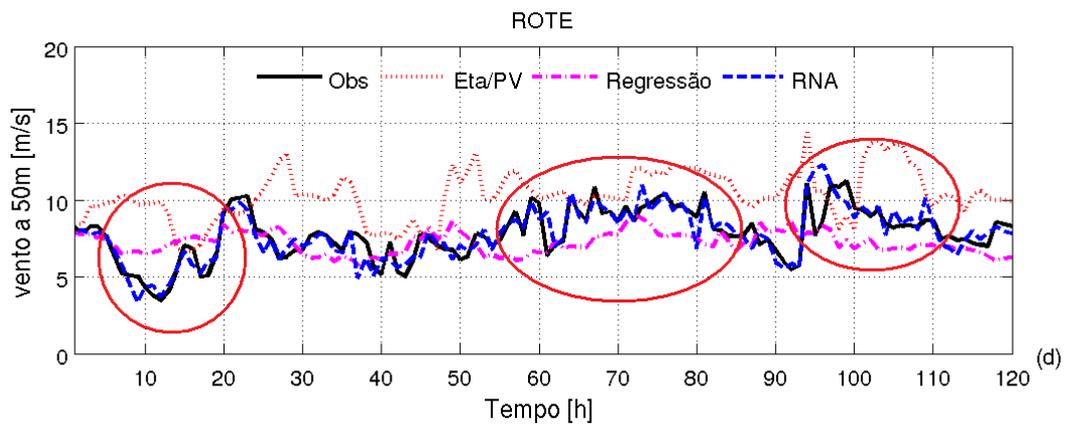
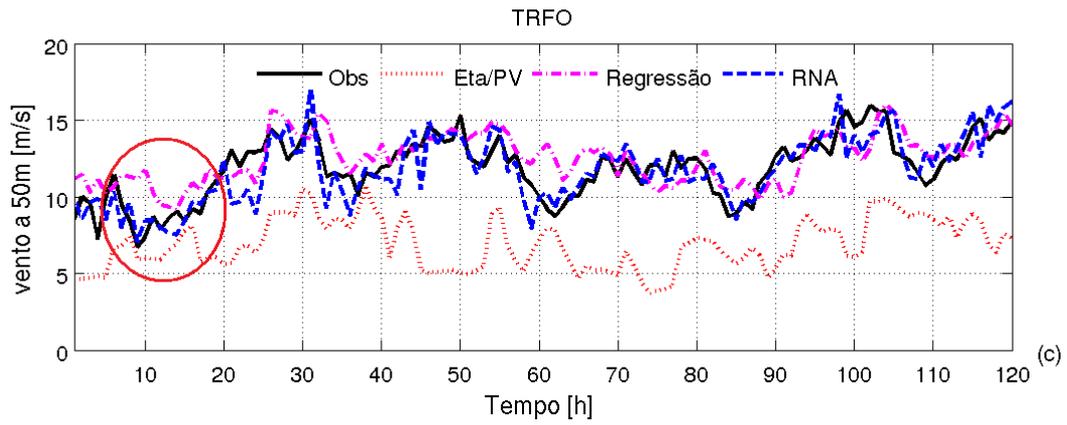
	RMSE (m/s)			R			BIAS (m/s)			SKILL (RMSE)		SKILL (R)	
	Eta	REG	RNA	Eta	REG	RNA	Eta	REG	RNA	Eta	REG	Eta	REG
<b>PETR</b>	2,80	1,46	0,95	0,40	0,48	0,81	-1,13	0,00	-0,01	-66%	-35%	103%	67%
<b>SJCA</b>	3,10	1,79	1,37	0,58	0,63	0,80	-1,70	-0,04	0,00	-56%	-24%	39%	28%
<b>TRFO</b>	4,59	2,11	1,36	0,37	0,68	0,88	3,64	0,06	0,17	-70%	-36%	136%	29%
<b>ROTE</b>	2,87	1,29	0,93	0,34	0,70	0,87	-0,43	-0,03	0,07	-68%	-28%	153%	24%
<b>MAGI</b>	2,30	1,33	1,04	0,35	0,68	0,83	0,67	-0,04	0,04	-55%	-22%	137%	21%
<b>GIRP</b>	3,23	1,53	1,33	0,31	0,81	0,87	1,28	0,04	0,15	-59%	-14%	182%	7%
<b>AGUB</b>	3,52	1,59	1,39	0,22	0,76	0,83	2,36	-0,07	0,07	-61%	-13%	273%	9%
<b>PALM</b>	4,63	1,70	1,15	0,39	0,75	0,90	-3,47	-0,02	0,15	-75%	-32%	128%	19%

Em geral os valores de erro e correlação obtidos para o modelo em RNAs foi considerado satisfatório quando comparados à literatura. Os valores para RMSE de

trabalhos similares oscila entre 1,38 m/s e 2,08 m/s e as correlações entre 0,79 e 0,87. Pode-se concluir que as correlações estiveram sempre acima da expectativa e os maiores valores de RMSE ficaram no limite inferior dos valores da literatura, o que pode ser considerado um bom resultado. Entretanto, as particularidades de cada trabalho devem ser consideradas uma vez que há fatores que influenciam os resultados como o tamanho das séries anemométricas, sua resolução temporal e o método de modelagem e avaliação dos erros empregado.

Por fim, para visualizar com maiores detalhes em uma escala temporal em quais momentos a RNA se comporta melhor em relação à REGP e ao Eta/PV, foram geradas figuras das séries temporais de vento a 50 m onde as diferentes previsões são comparadas para um horizonte de 5 dias consecutivos. A Figura 7.22 ilustra estas séries para as 8 estações, onde os períodos de maior interesse estão circulados em vermelho. Nota-se pela figura que o desempenho das RNAs se diferencia dos demais principalmente para os valores extremos, mínimos e máximos, e para variações bruscas de vento.





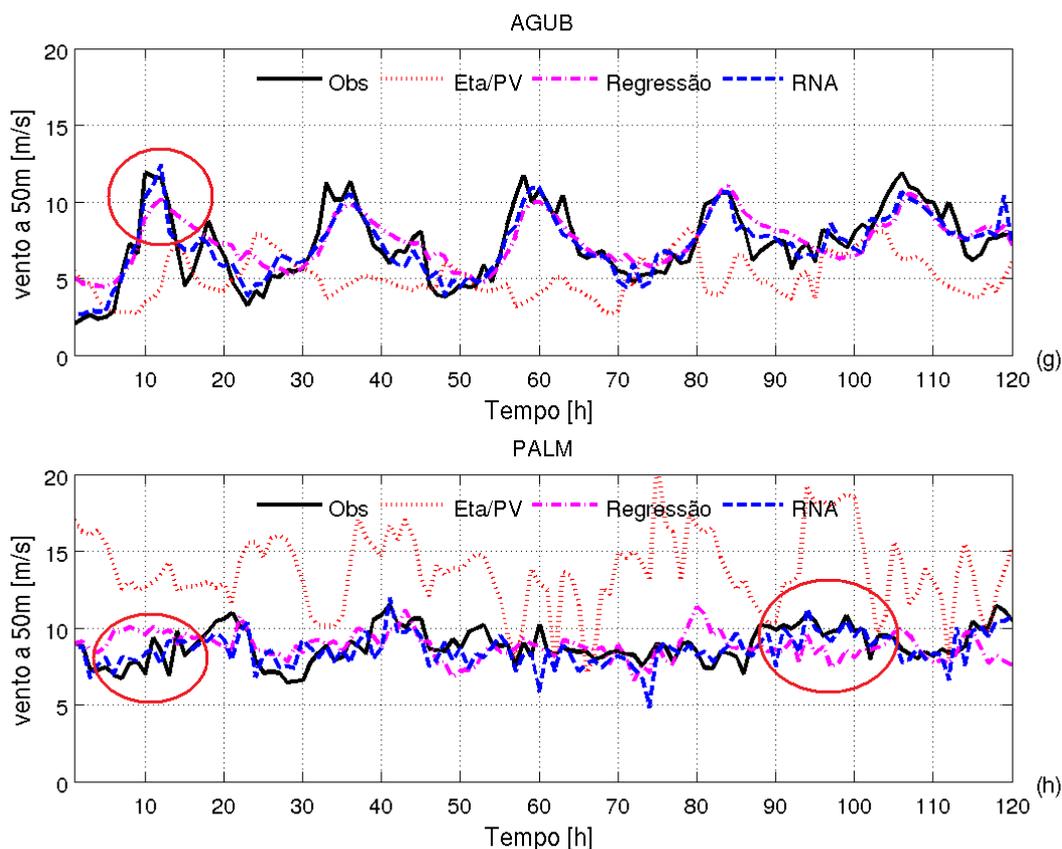


Figura 7.22 – Comparação entre as séries temporais de Eta/PV, REGP e RNA e o vento observado a 50 m para (a) PETR; (b) SJCA; (c) TRFO; (d) ROTE; (e) MAGI; (f) GIRP; (g) AGUB; (h) PALM. Regiões de interesse destacadas em vermelho.

### 7.7. Avaliação dos Pesos dos Preditores

Uma análise complementar que pode trazer benefícios quanto à compreensão das magnitudes dos fenômenos envolvidos na modelagem estatística do vento é a determinação da influência de cada preditor sobre o resultado final, o que pode ser entendido como seu peso final no modelo de RNAs. Cada preditor da RNA foi levemente deslocado do seu valor original para duas posições, uma em +10% (Delta+) e outra em -10% (Delta-) em relação suas anomalias máximas, onde as saídas do vento obtidas foram comparadas à saída original. Com isso admite-se que o viés apresentado pela nova saída de vento representa sua sensibilidade em relação àquele preditor, em valores percentuais. Cabe ressaltar que este foi deslocado para mais e para menos porque a sensibilidade pode ser diferente para

um dos lados uma vez que o modelo é não-linear. Repetindo-se este passo para todos os preditores têm-se o mapeamento dos pesos de cada preditor no modelo final. Estas informações podem ser úteis inclusive para orientar melhorias no tratamento do vento em modelos de mesoescala.

Os preditores selecionados para as estações de SJCA e ROTE durante a configuração da RNA estão listados pelas Tabelas 7.6 e Tabela 7.7 respectivamente. Estes foram os mesmos preditores utilizados pelas demais estações, os de ROTE para MAGI, por ser no litoral, e os de SJCA para as demais estações do interior. Nestas tabelas estão descritos o nível de cada preditor, a escala de tendência quando for o caso, a defasagem de máxima correlação e os respectivos impactos sobre o vento modelado pelas RNAs.

Tabela 7.6 – Preditores selecionados para SJCA e respectivos impactos sobre o vento modelado pela RNA. As defasagens mostradas significam atraso na previsão quando negativas.

Nro	Variável	Nível (hPa)	Tendência	Defasagem (h)	Descrição	Delta(-)	Delta(+)
1	apcpsfc	sfc		1	Precipitação total	0,1%	0,0%
2	dlwrfsfc	sfc		0	Onda longa Descendente	0,0%	-0,1%
3	dswrfsfc	sfc		-1	Onda Curta Descendente	-1,0%	1,0%
4	lhtflsfc	sfc		0	Fluxo calor latente	-2,7%	2,5%
5	shtflsfc	sfc		-1	Fluxo de calor sensível	3,3%	-3,1%
6	tmp2m	<b>sfc</b>		-1	<b>Temperatura a 2m</b>	<b>-6,5%</b>	<b>6,0%</b>
7	ugrd10m	sfc		0	Vento zonal a 10 m	3,0%	-2,5%
8	vgrd10m	<b>sfc</b>		0	<b>Vento meridional a 10 m</b>	<b>6,7%</b>	<b>-6,9%</b>
9	hgtprs500	500		0	Altura geopotencial da camada	0,4%	-0,5%
10	hgtprs200	200		0	Altura geopotencial da camada	0,9%	-0,8%
11	spfhprs950	950		0	Umidade especifica da camada	1,2%	-1,2%
12	spfhprs500	500		0	Umidade especifica da camada	2,9%	-2,8%
13	spfhprs200	<b>200</b>		1	<b>Umidade especifica da camada</b>	<b>-14,9%</b>	<b>11,6%</b>
14	tmpprs1000	1000		0	Temperatura da camada	0,7%	-0,8%
15	tmpprs900	900		-1	Temperatura da camada	3,9%	-4,1%
16	tmpprs850	<b>850</b>		0	<b>Temperatura da camada</b>	<b>6,8%</b>	<b>-6,4%</b>
17	tmpprs200	200		0	Temperatura da camada	0,6%	-0,6%
18	ugrdprs1000	1000		0	Vento zonal na camada	0,5%	-0,6%
19	ugrdprs925	925		0	Vento zonal na camada	1,1%	-1,2%
20	ugrdprs900	900		0	Vento zonal na camada	-2,1%	2,1%
21	ugrdprs850	850		0	Vento zonal na camada	0,5%	-0,4%
22	vgrdprs925	<b>925</b>		0	<b>Vento meridional na camada</b>	<b>-4,0%</b>	<b>4,2%</b>

<b>23</b>	vgrdprs900	900		0	Vento meridional na camada	4,3%	-3,4%
<b>24</b>	vgrdprs850	850		0	Vento meridional na camada	-0,1%	0,1%
<b>25</b>	wndmag10	<b>sfc</b>		0	<b>Magnitude do vento a 10 m</b>	<b>-7,9%</b>	<b>7,6%</b>
<b>26</b>	ghdlwrfsfc	sfc	horário	1	Onda longa Descendente	0,1%	-0,1%
<b>27</b>	ghdswrfsfc	sfc	horário	-1	Onda Curta Descendente	-0,3%	0,3%
<b>28</b>	ghlcdclcl	sfc	horário	0	Fr. de cobert. nuvens baixas	0,6%	-0,5%
<b>29</b>	ghshtflsfc	sfc	horário	-1	Fluxo de calor sensível	0,7%	-0,7%
<b>30</b>	ghtmp2m	sfc	horário	1	Temperatura a 2m	1,2%	-1,2%
<b>31</b>	ghtmpprs950	950	horário	1	Temperatura da camada	0,6%	-0,5%
<b>32</b>	ghtmpprs925	925	horário	1	Temperatura da camada	-1,1%	1,1%
<b>33</b>	ghtmpprs850	850	horário	1	Temperatura da camada	-1,7%	1,6%
<b>34</b>	uxvisc	sfc		0	Fluxo de momentum fricção	-0,8%	0,8%
<b>35</b>	uwflux	sfc		0	Fluxo de momentum vertical	1,3%	-1,3%
<b>36</b>	gdapcpsfc	sfc	diário	-1	Precipitação total	-0,8%	0,8%
<b>37</b>	gdlcdclcl	sfc	diário	0	Fr. de cobert. nuvensbaixas	-1,7%	2,1%
<b>38</b>	gdvgrdprs950	950	diário	0	Vento meridional na camada	-1,0%	1,0%
<b>39</b>	gdwndmag10	sfc	diário	0	Magnitude do vento a 10 m	2,3%	-2,2%
<b>40</b>	uxdvgd	sfc	diário	0	Fluxo de momentum vertical	-1,5%	1,5%
<b>41</b>	wave4	sfc		0	Onda senoidal com máx. às 10Z	1,4%	-1,2%

Tabela 7.7 – Preditores selecionados para ROTE e respectivos impactos sobre o vento modelado pela RNA. As defasagens mostradas significam atraso na previsão quando negativas.

Nro	Variável	Nível (hPa)	Tendência	Defasagem (h)	Descrição	Delta(-)	Delta(+)
1	msletmsl	sfc		-11	<b>Pressão media nível do mar (Eta)</b>	<b>-8,5%</b>	<b>8,6%</b>
2	spfh10m	sfc		0	Umidade especifica a 10m	-5,1%	4,9%
3	tcdclm	sfc		9	<b>Cobertura total de nuvens</b>	<b>-44,9%</b>	<b>41,3%</b>
4	hgtprs1000	<b>1000</b>		-11	<b>Altura geopotencial camada</b>	<b>-27,6%</b>	<b>27,4%</b>
5	hgtprs900	<b>900</b>		<b>-11</b>	<b>Altura geopotencial da camada</b>	<b>24,5%</b>	<b>-24,6%</b>
6	hgtprs200	200		11	Altura geopotencial da camada	2,6%	-2,3%
7	spfhprs850	850		-9	Umidade especifica da camada	2,6%	-2,8%
8	tmpprs1000	1000		-8	Temperatura da camada	-0,1%	0,2%
9	tmpprs950	950		-10	Temperatura da camada	1,3%	-1,9%
10	tmpprs200	200		10	Temperatura da camada	6,6%	-6,6%
11	ugrd10m	sfc		0	Vento zonal a 10 m	-4,2%	3,8%
12	ugrdprs900	900		0	Vento zonal na camada	1,1%	-1,5%
13	ugrdprs850	850		-5	Vento zonal na camada	-3,3%	3,3%
14	ugrdprs200	200		-6	Vento zonal na camada	-0,9%	0,9%
15	ugrdprs1000	<b>1000</b>		0	<b>Vento zonal na camada</b>	<b>-14,6%</b>	<b>14,6%</b>
16	vgrdprs1000	1000		0	Vento meridional na camada	-5,5%	8,4%
17	vgrdprs950	<b>950</b>		0	<b>Vento meridional camada</b>	<b>-28,6%</b>	<b>29,2%</b>
18	vgrdprs925	<b>925</b>		0	<b>Vento meridional camada</b>	<b>33,2%</b>	<b>-32,5%</b>
19	vgrdprs900	<b>900</b>		0	<b>Vento meridional camada</b>	<b>20,2%</b>	<b>-22,0%</b>
20	vgrdprs850	850		0	Vento meridional camada	3,6%	-3,8%
21	vgrdprs200	<b>200</b>		<b>-3</b>	<b>Vento meridional camada</b>	<b>-12,1%</b>	<b>12,5%</b>
22	wndmag10	sfc		0	<b>Magnitude do vento a 10 m</b>	<b>-41,9%</b>	<b>42,1%</b>

<b>23</b>	wnddir10	<b>sfc</b>		<b>-4</b>	<b>Direção do vento a 10 m</b>	<b>6,5%</b>	<b>-6,5%</b>
<b>24</b>	ghspfh10m	sfc	horário	10	Umidade especifica a 10m	0,4%	-0,4%
<b>25</b>	ghhgtprs500	500	diário	-6	Altura geopotencial camada	-1,2%	1,1%
<b>26</b>	ghspfhprs925	925	horário	-2	Umidade especifica camada	0,2%	-0,1%
<b>27</b>	ghtmpprs950	950	horário	-6	Temperatura da camada	-1,1%	1,1%
<b>28</b>	uwsdy2	sfc		-6	Fluxo momento reg. perman.	2,6%	-2,6%
<b>29</b>	gdmsletmsl	<b>sfc</b>	<b>diário</b>	<b>-11</b>	<b>Pressão media nível do mar</b>	<b>-5,2%</b>	<b>5,2%</b>
<b>30</b>	gdugrd10m	sfc		3	Vento meridional a 10 m	4,1%	-4,7%
<b>31</b>	gdhgtprs200	<b>200</b>	<b>diário</b>	0	<b>Altura geopotencial camada</b>	<b>26,0%</b>	<b>-26,6%</b>
<b>32</b>	gdspfhprs1000	<b>1000</b>	<b>diário</b>	<b>8</b>	<b>Umidade especifica camada</b>	<b>12,8%</b>	<b>-12,7%</b>
<b>33</b>	gdspfhprs850	850	diário	-9	Umidade especifica camada	0,4%	-0,4%
<b>34</b>	gdspfhprs500	500	diário	0	Umidade especifica camada	-5,0%	4,9%
<b>35</b>	gdtmppers1000	<b>1000</b>	<b>diário</b>	<b>0</b>	<b>Temperatura da camada</b>	<b>9,7%</b>	<b>-9,7%</b>
<b>36</b>	gdtmppers950	950	diário	8	Temperatura da camada	0,9%	-0,7%
<b>37</b>	gdugrdprs950	<b>950</b>	<b>diário</b>	4	<b>Vento zonal na camada</b>	<b>9,5%</b>	<b>-9,1%</b>
<b>38</b>	gdugrdprs900	900	diário	11	Vento zonal na camada	1,6%	-1,7%
<b>39</b>	gdugrdprs850	850	diário	0	Vento zonal na camada	-4,3%	4,3%
<b>40</b>	gdvgrdprs1000	1000	diário	-11	Vento meridional camada	-0,6%	0,2%
<b>41</b>	gdvgrdprs850	<b>850</b>	<b>diário</b>	<b>-11</b>	<b>Vento meridional camada</b>	<b>10,7%</b>	<b>-10,4%</b>
<b>42</b>	gdvgrdprs500	500	diário	11	Vento meridional camada	2,8%	-2,7%
<b>43</b>	gdwndmag10	sfc	diário	5	Magnitude do vento a 10 m	-4,1%	4,2%

Durante a etapa de seleção de preditores, chegou-se para SJCA a 41 preditores enquanto para ROTE chegou-se a 43, um número bastante próximo. Lembrando que para as duas estações foi adicionado ainda o vento modelado pela REGP ao conjunto de preditores, o qual não está listado pelas tabelas acima. Percebe-se que os conjuntos de preditores são bastante diferentes entre si, o que é até certo ponto compreensível devido à diversidade entre o litoral e o sertão. Há uma variação grande entre as defasagens máximas definidas para cada estação (+/- 1 h e +/- 11 h), o que pode justificar a diferença nos preditores. Não se pode negar todavia que há a deficiência do método de seleção, uma vez que a REGP é sensível à ordem em que os preditores são acrescentados, enquanto a RNA pode modelar algum ruído do sinal que não tem relação com os fenômenos de interesse. Por isso, somente a análise dos pesos dos preditores e sua comparação entre as estações pode trazer informações mais confiáveis sobre as variáveis mais relevantes para a modelagem estatística do vento.

Analisando os resultados da Tabela 7.6 e Tabela 7.7 percebe-se que a magnitude do impacto dos preditores em SJCA é bem menor do que para ROTE. Isto pode ter como causa a série mais curta disponível para treinamento em ROTE, que permite uma maior flexibilidade para o ajuste dos pesos dos preditores. A RNA para SJCA já aparenta ser bem mais estável o que a leva a ter uma maior capacidade de generalização.

Investigando as variáveis de maior impacto destacadas nas tabelas, pode-se notar para SJCA uma forte influência da umidade em altos níveis (-14,9% / +11,6%), afetando positivamente o vento, ou seja, maior umidade em altos níveis maior a intensidade do vento em superfície. Pela defasagem de +1 h esta deve estar associada à intensificação do vento pela presença de sistemas convectivos intensos. Não é possível distinguir se estes sistemas estão associados a forçantes dinâmicas, como VCANs ou ZCIT, devido ao baixo impacto dos sinais de altura geopotencial e temperatura em níveis mais elevados. Do mesmo modo nota-se a influência positiva da temperatura a 2 m

(-6,5% / +6,0%) o que sugere novamente a intensificação do vento por circulações locais de origem térmica. Um aspecto intrigante destes resultados é a inversão que há na influência do vento meridional em diferentes níveis. Enquanto em 925 hPa o vento meridional afeta positivamente a magnitude do vento (-4,0% / +4,0%), para o vento a 10 m esta influência é inversa (6,7% / -6,9%). Como estas duas variáveis estão altamente correlacionadas, uma hipótese seria que a influência de ambas está se somando no interior da RNA e a diferença entre elas fornece um sinal de cisalhamento do vento meridional.

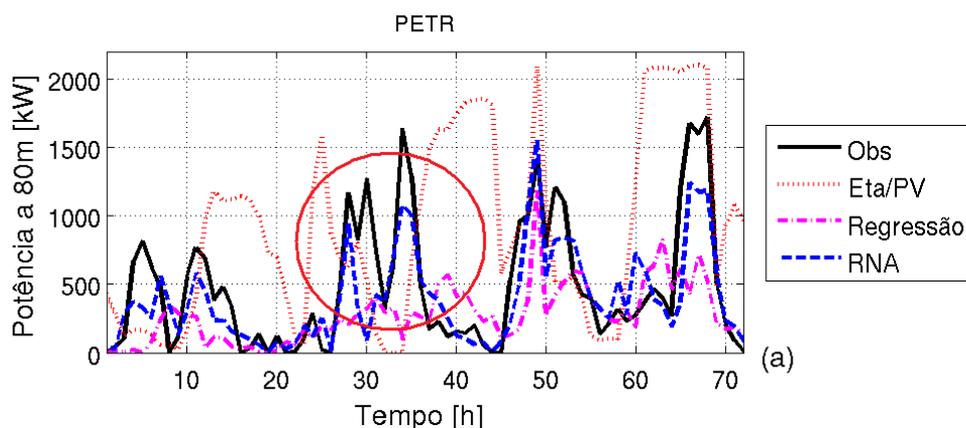
Já com relação à ROTE, nota-se um forte impacto positivo da cobertura de nuvens (-44,9% / 41,3%) o qual não aparenta estar associado à desintensificação da brisa marítima, mas sim a algum contexto sinótico uma vez que maior nebulosidade causa um aumento do vento. Além disso a defasagem indica que o vento se intensifica em média 9 h antes do aumento de nebulosidade o que deve estar relacionado à aproximação de algum sistema sinótico. É evidente ainda uma associação inversa entre a altura geopotencial na camada em 200 hPa e o vento (26,0% / -26,6%), que deve estar associado à chegada de um VCAN, ou ao afastamento da ZCIT. Neste caso, como a defasagem não indica nenhum aumento da correlação numa escala de +/- 11 h, acredita-se que a segunda opção é a mais provável. Observa-se ainda a alternância de influência do vento meridional em baixos níveis, conforme observado para SJCA e também a alternância de influência da altura geopotencial em baixos níveis. Como já foi discutido, estas por serem altamente correlacionadas se cancelam no interior da rede onde apenas o resíduo se torna relevante. Por fim, têm-se a forte influência positiva do vento previsto a 10 m, como já era esperado, e uma intensificação do vento com o aumento do gradiente diário de pressão em superfície. Um fator interessante é dado pela variação da direção do vento, que indica ventos mais fracos quando vindos de sul.

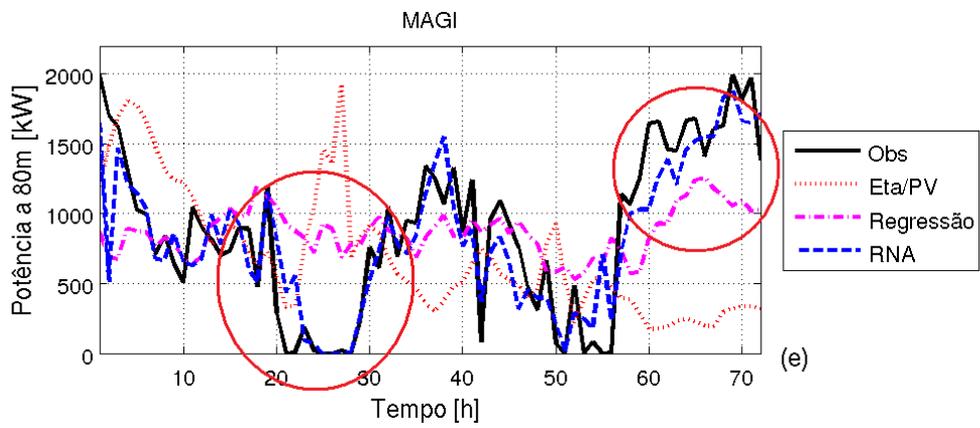
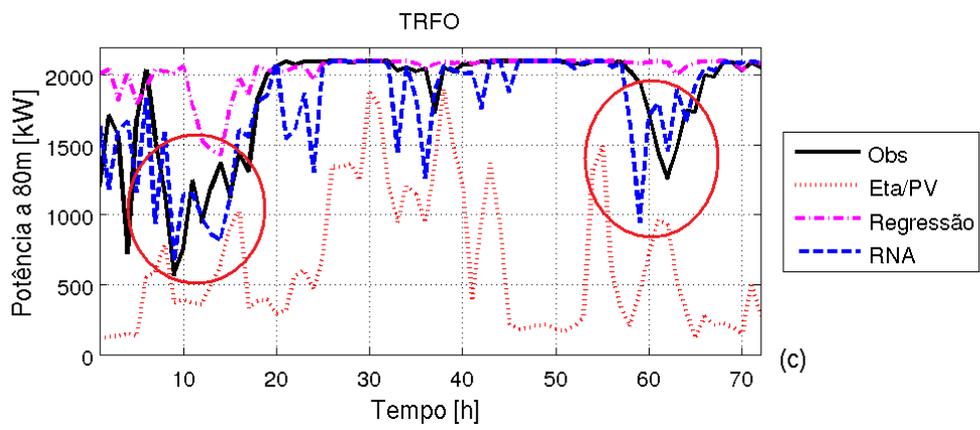
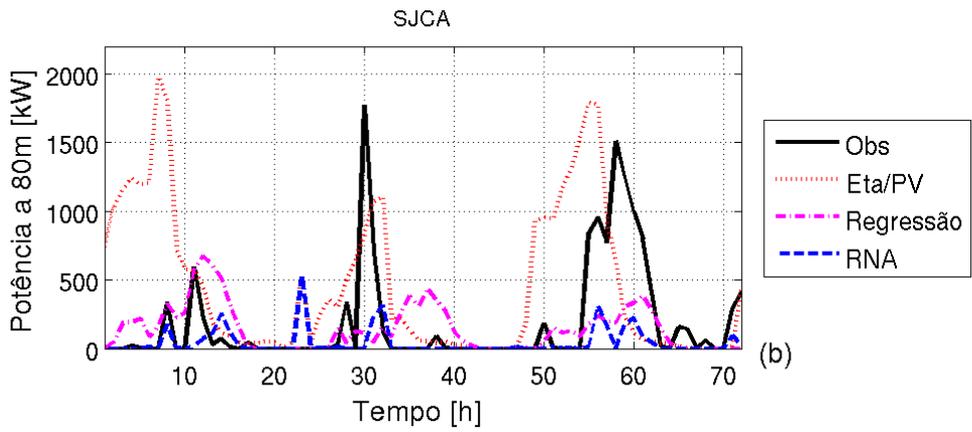
Estes resultados, apesar de rudimentares já trazem informações relevantes sobre a relação dos fenômenos atmosféricos com o vento. Análises mais

aprofundadas e criteriosas são possíveis, o que evidencia a potencialidade das RNAs também para o diagnóstico de mecanismos físicos complexos. É importante considerar nestes casos que a RNA não possui recurso similar ao das regressões passo-a-passo, uma vez que não identifica preditores redundantes. Eventualmente os modelos em RNAs podem fazer uso de preditores altamente correlacionados, porém em regiões distintas de sua topologia. Cabe a quem faz a configuração da rede neural encontrar formas de deixar o modelo o mais simples e eficiente possível.

### 7.8. Impactos sobre a Potência Eólica Disponível

As previsões de vento obtidas por cada modelo foram aplicadas sobre uma curva de potência de uma aerogerador de 2,1 MW. Para isso foi necessário extrapolar a magnitude do vento para a altura de cubo referente a este aerogerador, que é de 80 m. A extrapolação foi feita utilizando o perfil logarítmico do vento e o valor das rugosidades já calculado anteriormente. Foram geradas séries de potência eólica disponível com resolução horária para todo o período de dados das estações. A Figura 7.23 mostra a comparação entre a potência eólica disponível prevista por cada um dos modelos para um horizonte de 72 h.





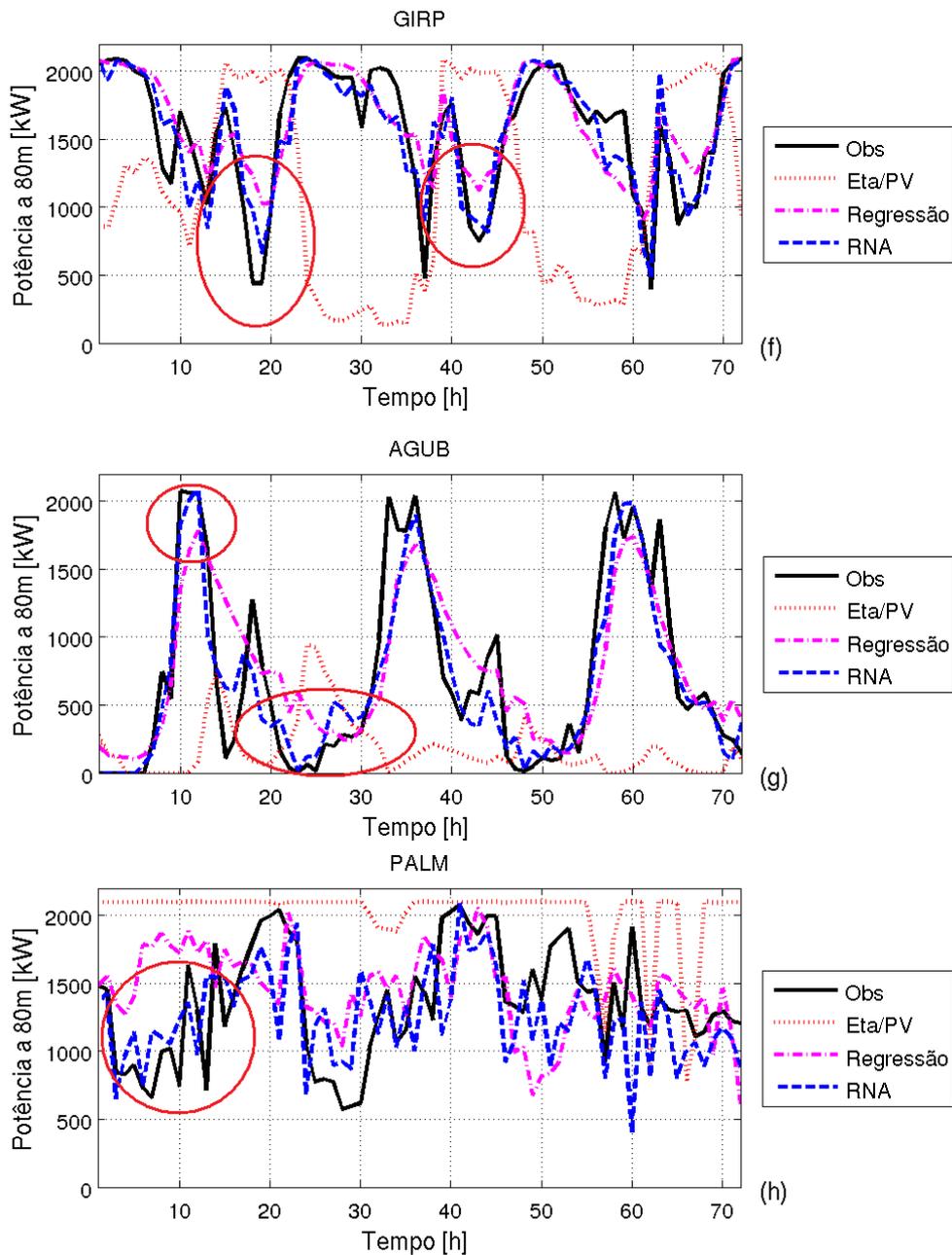


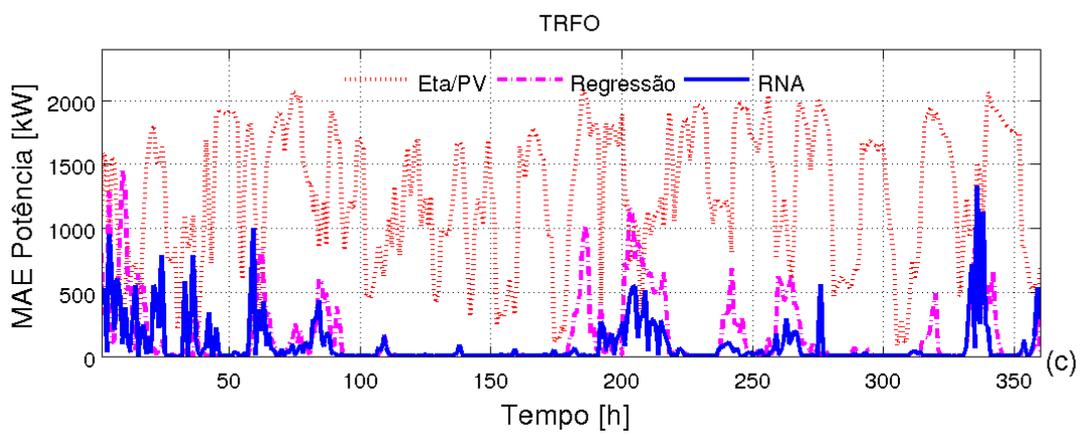
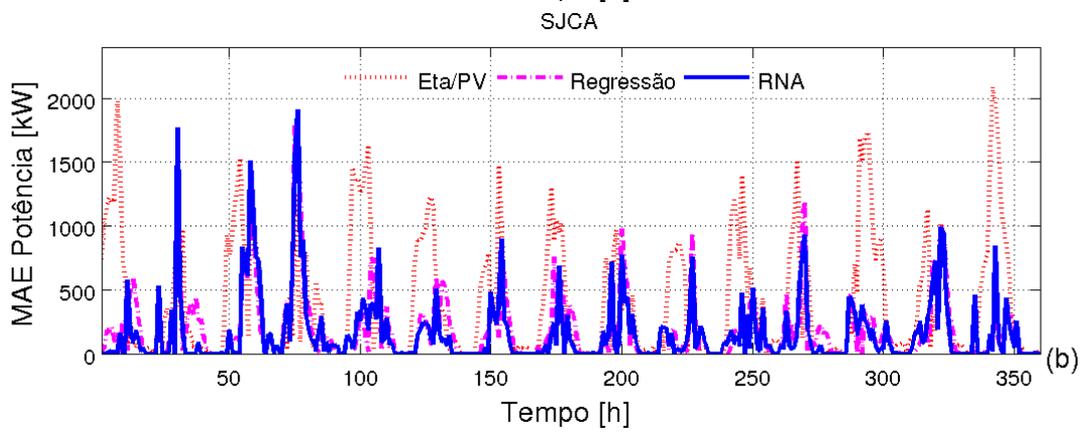
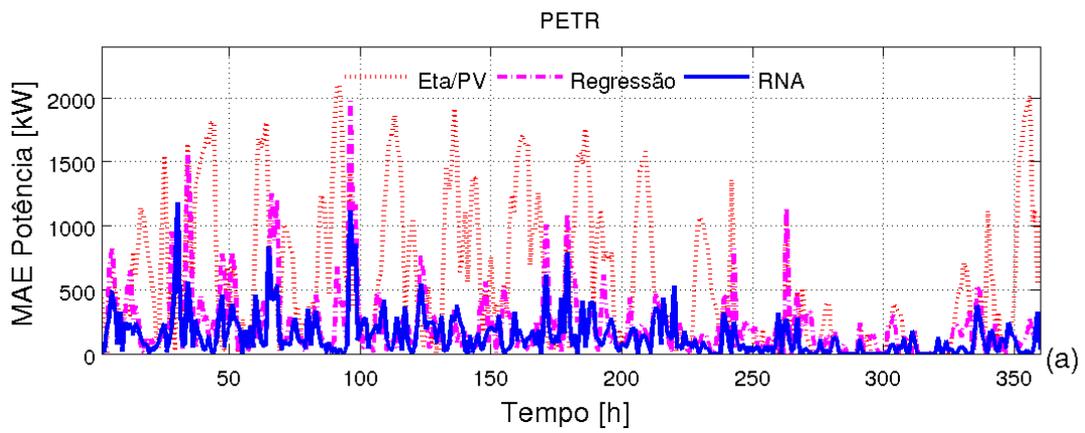
Figura 7.23 – Comparação entre as séries temporais de potência eólica (kW) dadas pelo Eta/PV, REGP e RNA em relação à potência teórica observada a 80 m para (a) PETR; (b) SJCA; (c) TRFO; (d) ROTE; (e) MAGI; (f) GIRP; (g) AGUB; (h) PALM. Áreas de interesse em destaque.

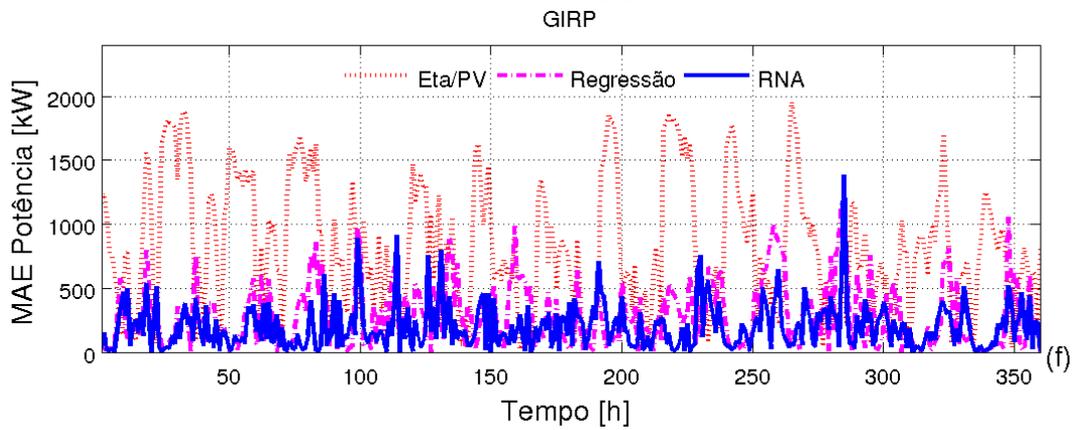
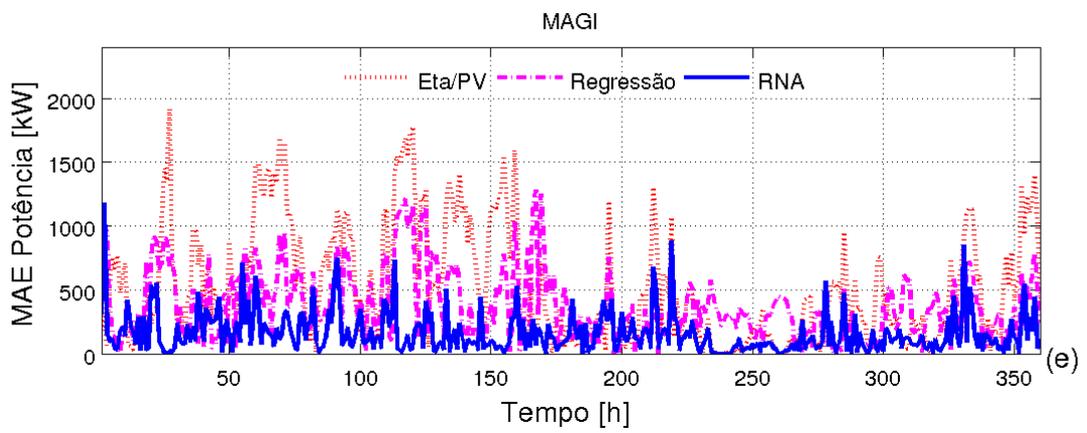
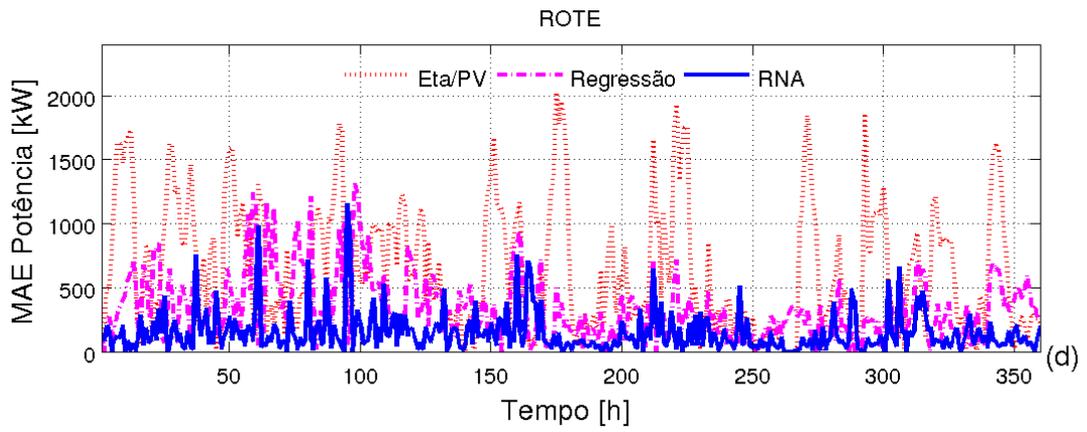
Nota-se que para a potência eólica, a RNA representa a série observada um pouco melhor do que a REGP, enquanto ambas apresentam um desempenho

bem superior ao Eta/PV. Este ganho das RNAs varia em função da amplitude média do ciclo diário do vento em cada local, uma vez que a relação entre o vento e a potência não é linear, como descrito anteriormente pela Equação 6.21. Esta variação ocorre porque a RNA foi obtida pela minimização do erro quadrático médio do vento, portanto é compreensível que a conversão em potência eólica degrade seu desempenho. O ideal seria utilizar como critério de treinamento da RNA o erro quadrático da potência eólica disponível, ou pelo menos um polinômio que represente o grau de não-linearidade típico da curva de potência de um aerogerador. Eventualmente, a simples inclusão de um filtro para eliminar a influência do erro para as regiões tabulares da curva de potência, leia-se velocidades inferiores à de partida ( $\sim 4$  m/s) ou superiores à de máxima potência ( $\sim 14$  m/s), já traria benefícios para a modelagem estatística da potência eólica.

Neste contexto, ao avaliarmos o desempenho da RNA entre as estações de SJCA e ROTE, fica evidente o que foi exposto acima. ROTE possui um ciclo diário muito suave, o que favorece a previsibilidade da RNA, enquanto SJCA possui um ciclo bastante acentuado, o que degrada bastante as previsões. Para as outras estações, o desempenho das RNAs é satisfatório, superando frequentemente o desempenho das REGPs, principalmente para oscilações de frequência mais alta. Cabe ressaltar que ainda assim a REGP foi capaz de eliminar a maior parte do erro quando comparada ao Eta/PV, mostrando o quão essencial é o uso de um refinamento estatístico nas previsões de vento e potência eólica.

A partir destas estimativas foi calculado o erro instantâneo e o erro absoluto para cada um dos modelos. O erro absoluto é um parâmetro mais adequado para avaliar visualmente de que forma o erro se comporta com o tempo. A Figura 7.24 traz as séries temporais de erro absoluto para a potência eólica durante 15 dias consecutivos.





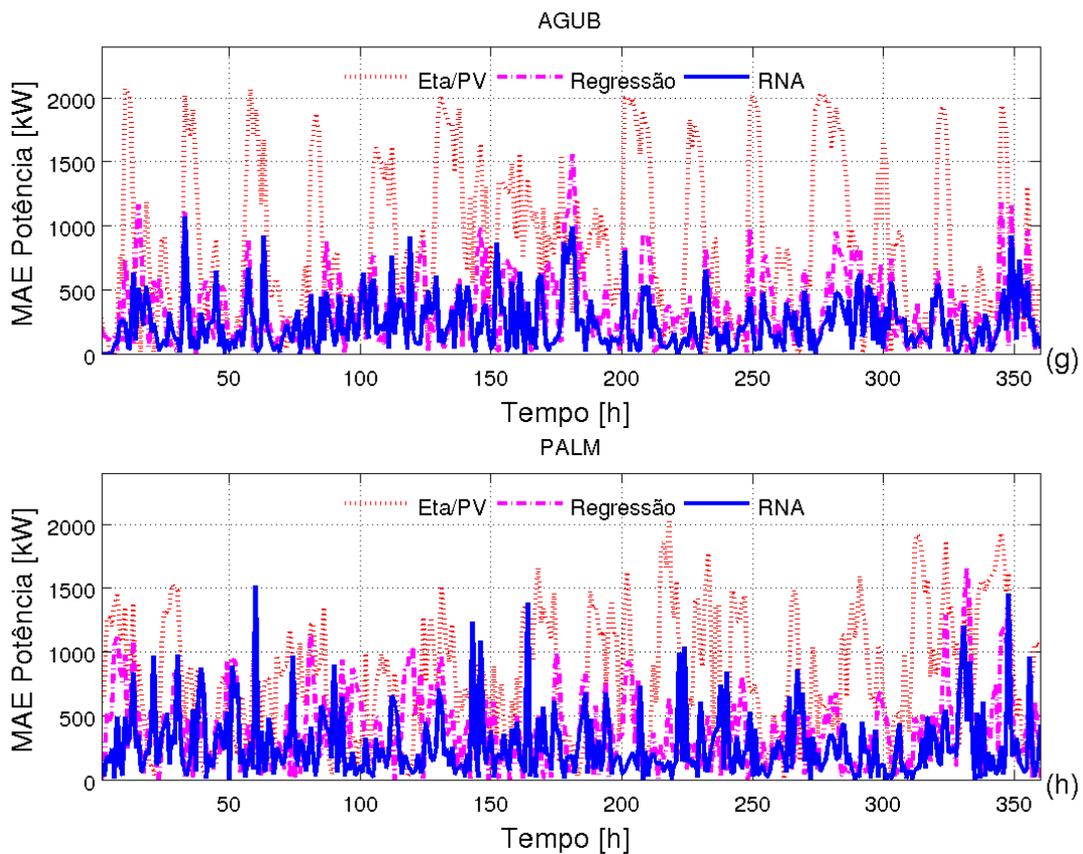
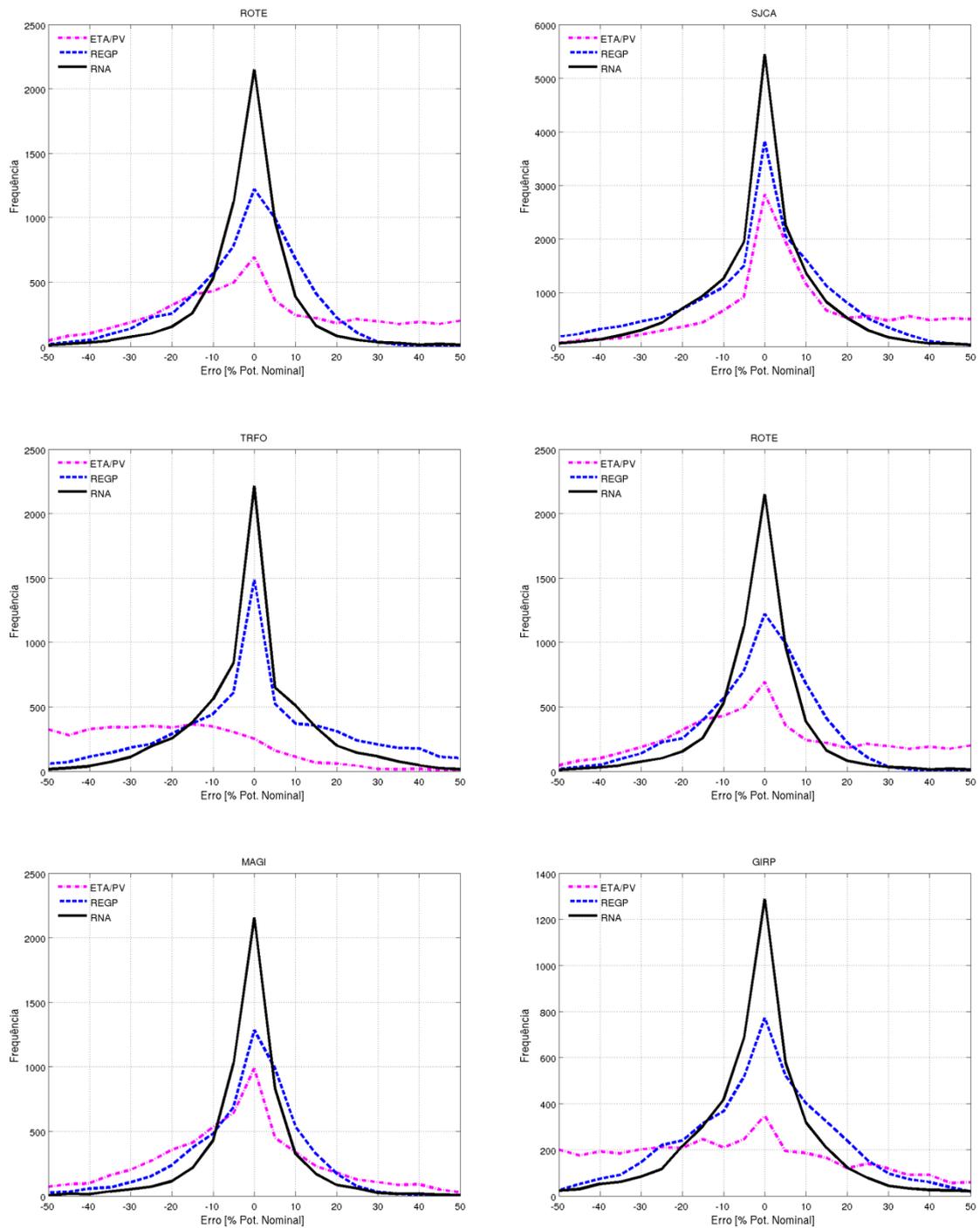


Figura 7.24 – Comparação entre os erros absolutos de potência eólica (kW) dadas pelo Eta/PV, REGP e RNA em relação à potência teórica observada a 80 m para (a) PETR; (b) SJCA; (c) TRFO; (d) ROTE; (e) MAGI; (f) GIRP; (g) AGUB; (h) PALM.

Analisando os gráficos acima, percebe-se claramente um ciclo diário do erro associado aos períodos de máxima intensidade do vento. Isto suporta o que já foi mencionado acima de que erros baixos podem não significar um modelo melhor, mas sim uma condição de vento fraca no local. O correto seria comparar erros de modelagem entre locais de densidade de potência eólica parecidas.

A etapa seguinte consistiu no cálculo do ganho de confiabilidade em termos de energia garantida no prazo de 24 horas com o uso de um modelo mais refinado. Foram calculados os histogramas para as distribuições de erro de cada estação. Como já era esperado, o aspecto dos histogramas se torna mais

esbelto com a utilização de um modelo mais refinado, como ilustra a Figura 7.25.



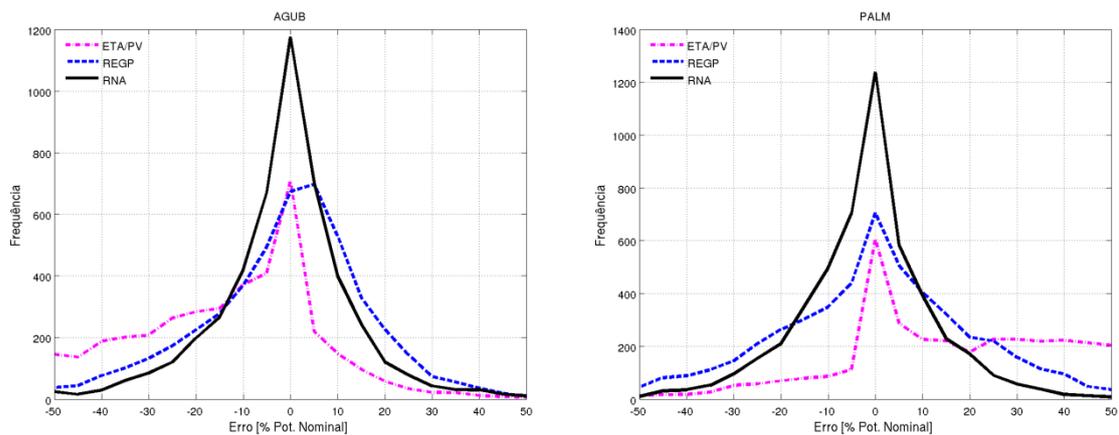
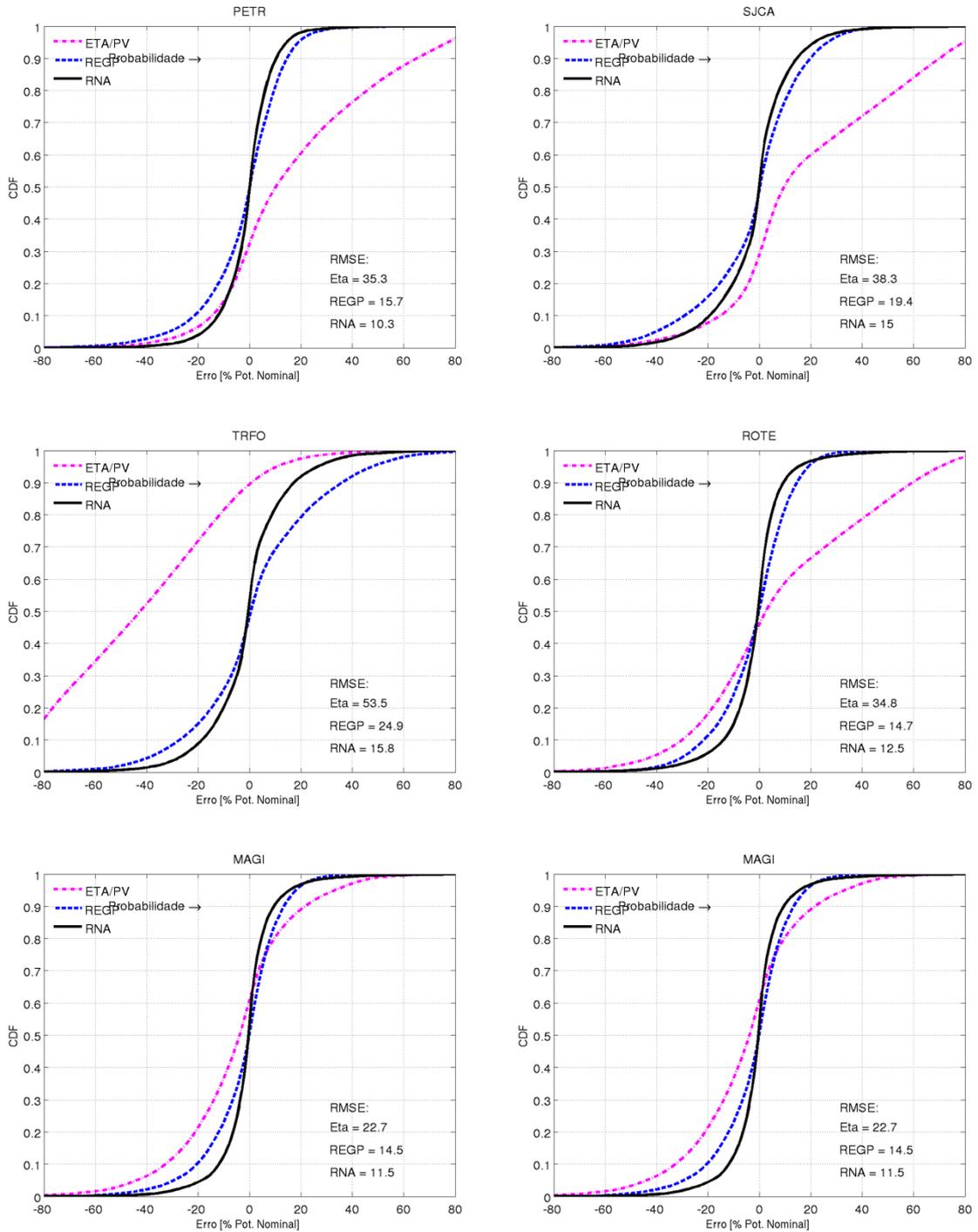


Figura 7.25 – Histograma da distribuição do erro entre a potência eólica prevista e disponível pelo Eta/PV, REGP e RNA para cada estação. Valores negativos indicam que a previsão subestimou a potência eólica, em pontos percentuais da potência nominal do aerogerador.

Os histogramas acima indicam qual a frequência do erro para cada intervalo onde as RNAs se mostraram mais precisas que os demais modelos. A simetria encontrada tanto nas curvas para as RNAs quanto para as REGPs indica que a média do erro está próxima de zero, ou seja, possuem viés muito baixo. Isto não se verifica para as previsões do Eta/PV, a qual apresenta histogramas bastante assimétricos em geral. Isto deverá ser considerado ao se calcular as probabilidades de erro de cada modelo, como será visto mais adiante. Uma vez mais nota-se o efeito dos “degraus” na curva de potência do aerogerador, que faz com que o erro nulo tenha uma frequência maior que o esperado para distribuições associadas a processos físicos comuns, ficando evidente pelo pico no centro dos histogramas.

Estas distribuições foram ajustadas por um suavizador para distribuições não-paramétricas do tipo ‘kernel’ normalizado, a partir das quais foram obtidas as funções de densidade de probabilidade acumulada ou CDFs para cada modelo em cada estação. Estas funções nos permitem calcular qual a probabilidade de ocorrência de um determinado valor de erro para cada modelo utilizado e

serão úteis no cálculo da potência eólica garantida para uma probabilidade específica. A Figura 7.26 ilustra as curvas CDF obtidas.



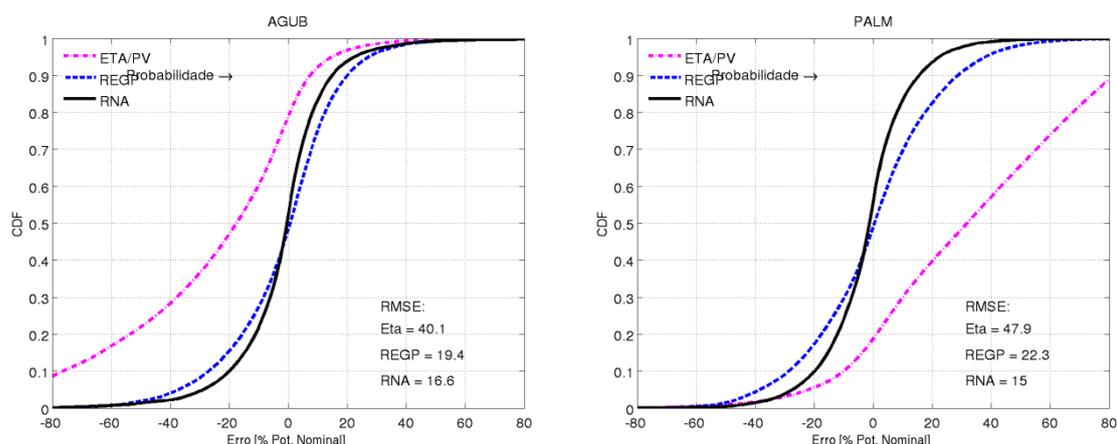


Figura 7.26 – Funções de densidade de probabilidade acumulada da diferença entre a potência eólica prevista e disponível pelo Eta/PV, REGP e RNA para cada estação. Valores negativos indicam que a previsão subestimou a potência eólica, em pontos percentuais da potência nominal do aerogerador. Eixo horizontal limitado em +/- 50%. Indicados adicionalmente os valores de MAE para os modelos.

A partir destas curvas pode-se identificar as estações que tendem a ter sua potência superestimada pelo Eta/PV como PETR, SJCA, ROTE e PALM assim como aquelas que tendem a ser subestimadas como TRFO, GIRP e AGUB. A probabilidade de ocorrência chega a ser de mais de 80% para superestimativas em PALM e mais de 90% para sub-estimativas em TRFO. Já os modelos REGP e RNA são bastante equilibrados, porém as RNAs apresentam menores erros, como já visualizado pelos valores de RMSE indicado nas figuras. Quando se fala em potência eólica garantida, considera-se esta a potência mínima esperada para uma determinada probabilidade de ocorrência  $p$ . Conhecendo-se as funções de probabilidade acumulada de erro das previsões é possível obter a potência mínima esperada ao longo do dia. Estas funções podem ser divididas por horário do dia, por horizonte de previsão ou simplesmente consideradas constantes ao longo do tempo. Neste caso, como

estão sendo avaliadas diversas estações simultaneamente e o objetivo não é operacional, mas sim obter uma ordem de grandeza destes valores optou-se por utilizar uma CDF constante ao longo do tempo para cada estação.

Adotou-se o valor de probabilidade  $P$  de 90%, ou seja, representa o limite inferior de uma previsão para o qual a probabilidade de ocorrência de um valor igual ou superior é de 90%. Fazendo-se a diferença entre os limites de cada previsão, é calculado o ganho efetivo em termos de potência (kW) disponível para cada aerogerador instalado com uma resolução horária. Uma consideração importante é que o viés de cada modelo deve ser igual para que esta comparação seja efetiva. Como as RNAs e as REGPs possuem um viés quase nulo enquanto o ETA/PV apresenta viés significativo, foi preciso remover o viés do Eta/PV para dar prosseguimento à análise.

Em uma primeira estimativa, admite-se que um operador de parque eólico decida comercializar este excedente de potência no mercado de curto-prazo. Considerando um valor aproximado de R\$ 45,00 por MWh foi calculada a receita adicional que poderia ser obtida com o aumento da energia garantida entre a RNA e o Eta/PV e entre a RNA e a REGP. É importante ressaltar que este é um exemplo hipotético onde as peculiaridades do processo de contratação e liquidação de energia não estão sendo consideradas. A Tabela 7.8 ilustra estes valores para cada parque em totais mensais. Percebe-se que as quantias são relevantes mesmo quando comparados RNAs às previsões pela REGP o que mostra que esta pode ser uma ferramenta importante para a viabilidade econômica de parques eólicos em determinadas regiões do Brasil

Tabela 7.8 – Receita adicional por aerogerador devido à comercialização do excedente de energia garantida.

Estação	Energia adicional diária (kWh)		Receita Bruta Mensal	
	RNA/Eta	RNA/REG	RNA/Eta	RNA/REG
<b>PETR</b>	18314	2868	R\$ 24.723,92	R\$ 3.871,51
<b>SJCA</b>	17389	3122	R\$ 23.474,55	R\$ 4.215,06
<b>TRFO</b>	12800	7898	R\$ 17.279,98	R\$ 10.662,72
<b>ROTE</b>	19940	2906	R\$ 26.918,69	R\$ 3.923,23
<b>MAGI</b>	8279	2704	R\$ 11.176,22	R\$ 3.650,23
<b>GIRP</b>	17782	3411	R\$ 24.005,09	R\$ 4.605,09
<b>AGUB</b>	8189	2606	R\$ 11.054,94	R\$ 3.518,77

Outra aplicação direta destes resultados seria na previsão de energia disponível em escala regional. Utilizando o desempenho médio da RNA para as estações analisadas e considerando a potência eólica instalada no nordeste brasileiro em torno de 1 GW, assumindo aerogeradores similares, pode se estimar o ganho nas previsões horárias em termos de potência garantida para a região. A comparação das RNAs ao modelo Eta/PV levou a uma diferença média de potência disponível de 293 MW enquanto em relação aos modelos REGP a diferença foi de 77 MW. Esta informação poderia auxiliar a ONS no gerenciamento da geração e dos despachos de carga, reduzindo perdas por transmissão e o tempo de operação das usinas termoelétricas. Na medida em que a capacidade eólica instalada cresce, ferramentas deste tipo ganham uma importância maior.

Para efeito de comparação foram calculados os valores de MAE e RMSE para cada modelo. As Figuras 7.25 e 7.26 ilustram estes valores onde nota-se a RNA se sempre abaixo dos demais.

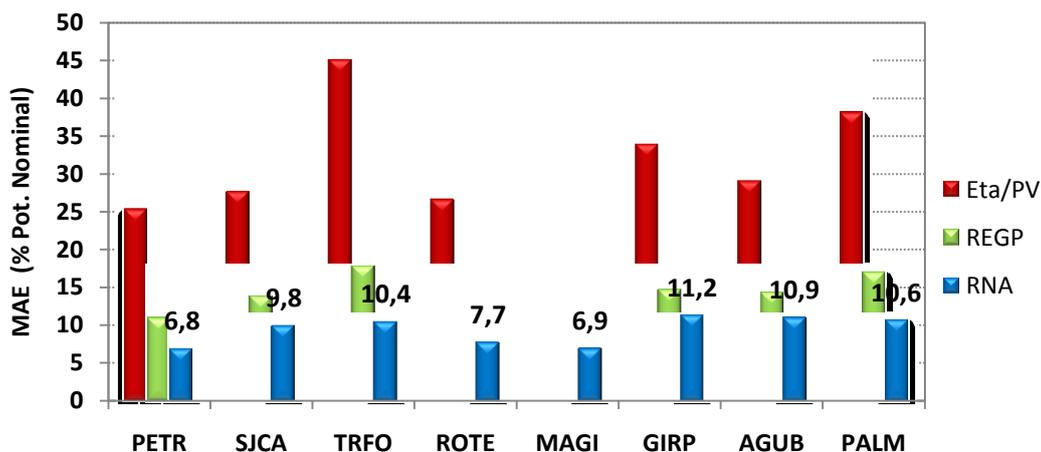


Figura 7.25 – MAE obtido para as estimativas de potência eólica por estação. Rótulos para valores das RNAs.

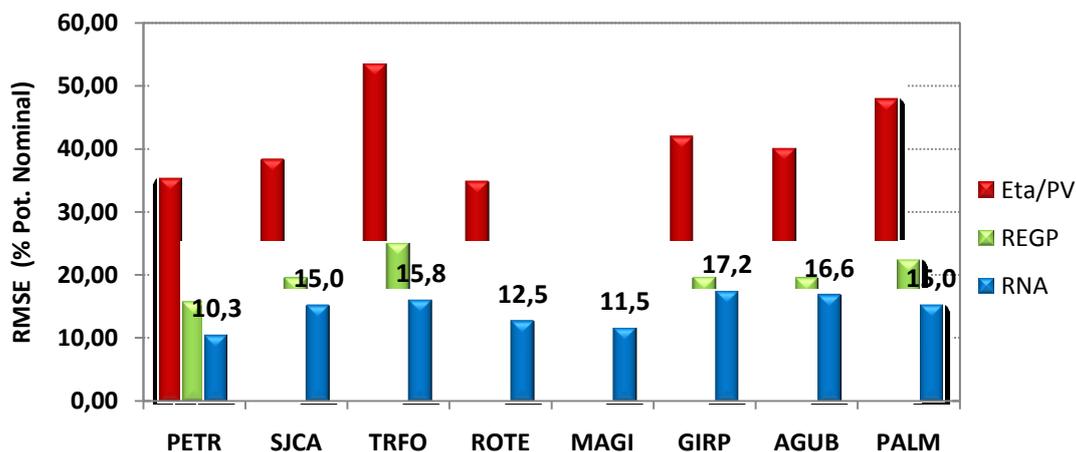


Figura 7.26 – RMSE obtido para as estimativas de potência eólica por estação. Rótulos para valores das RNAs.

A partir destes valores para os erros médios foi possível fazer uma comparação com os resultados encontrados na literatura. Para o erro normalizado de potência eólica disponível os valores de MAE oscilam entre 12% a 17,6% enquanto para o RMSE entre 11,1% a 20%. Os valores encontrados neste trabalho variam bastante entre as estações, sendo que para algumas o MAE foi inferior a 7% como PETR e ROTE enquanto os melhores valores de RMSE ficaram entre 10,3% e 12,5%. A média das estações ficou em 9,3% para MAE e 14,2% para RMSE, o que está coerente com a literatura.

Sem dúvida as características das séries de vento e potência utilizadas em cada trabalho é diferente o que torna a comparação direta não conclusiva. Neste trabalho as séries foram avaliadas integralmente, porém o tamanho das séries diferem entre si. A princípio séries menores podem favorecer os modelos de REGP e RNA apresentados acima. A determinação do tamanho mínimo das séries anemométricas para garantir uma boa capacidade de generalização para os modelos estatísticos demanda um estudo mais abrangente

De um modo geral os resultados obtidos neste estudo foram satisfatórios e atenderam a seu principal objetivo que era o desenvolvimento e validação de uma metodologia de refinamento estatístico do vento aplicado ao setor eólico-elétrico do nordeste brasileiro.

## 8 CONCLUSÃO

O presente trabalho investigou vários aspectos da modelagem estatística do vento, compreendendo desde o pré-processamento das variáveis, a seleção de preditores, as regressões lineares múltiplas, a definição da arquitetura e treinamento de redes neurais e posterior avaliação dos modelos e de seu impacto sobre a potência eólica.

Os desempenhos obtidos tanto pelas redes neurais quanto pelas regressões lineares múltiplas foram capazes de remover praticamente todo o erro sistemático do modelo Eta/PV e ainda elevar significativamente as correlações para todas as estações analisadas. Isto evidenciou a importância do acoplamento de modelos estatísticos às saídas dos modelos de PNT para obter previsões de grandezas de grande variabilidade espacial, como é o caso do vento próximo à superfície.

O desempenho das redes neurais artificiais em relação às regressões lineares múltiplas se mostrou relevante, muito embora a diferença para algumas estações tenha sido pequena.

Na previsão do vento a 50 m o viés do modelo Eta/PV variou de -0,43 m/s para ROTE a -3,47 m/s para PALM enquanto as REGPs e RNAs mantiveram-no abaixo de 0,15 m/s para todas as estações.

As correlações do Eta/PV foram em geral modestas, se mantendo sempre abaixo de 0,6, enquanto as REGPs e RNAs elevaram as correlações para a faixa de 0,7 a 0,9. A máxima correlação para as REGPs foi de 0,76, obtido para AGUB enquanto a máxima para as RNAs foi de 0,90, obtido para PALM.

O RMSE sofreu grande redução com o uso do ajuste estatístico, pois enquanto o Eta/PV indicava valores entre 2,30 m/s a 4,63 m/s, as REGPs levaram à faixa de 1,33 m/s a 2,11 m/s, e as RNAs reduziram a valores entre 0,95 m/s e 1,39 m/s.

Nas estimativas de potência eólica as diferenças entre os modelos foram reduzidas, principalmente porque a curva de potência do aerogerador atua

como um filtro, eliminando os erros para valores muito baixos, restritos pela velocidade de corte, ou muito altos, restritos pela velocidade de potência nominal.

A função de custo adotada como critério de treinamento para a rede neural não contempla estes aspectos da curva de potência, mostrando que o desempenho das previsões de potência eólica se degrada para locais onde o ciclo diário é bastante acentuado, como é o caso de SJCA.

Os valores para o RMSE da potência eólica encontrados ficaram entre 12% e 18% da potência eólica nominal, o que está dentro da faixa encontrada na literatura que chega no mínimo a 11,1%.

As estimativas de probabilidade de potência eólica disponível para um horizonte de curto-prazo se mostraram consistentes, evidenciando ganhos de até 35% da potência nominal horária com o uso de modelos estatísticos mais sofisticados em relação ao Eta/PV.

Quanto ao tratamento e pré-processamento das variáveis de entrada foi observado que critérios adicionais aos da rede SONDA foram necessários para eliminar erros dos dados anemométricos.

Mostrou-se que os dados do modelo Eta/PV estão sujeitos a erros (valores espúrios) que variam entre os pontos de grade.

Foi desenvolvido e avaliado um método de extrapolação vertical onde admite-se uma atmosfera em média estável e calcula-se a rugosidade pela minimização do erro no nível mais elevado. Deste modo pode-se eliminar o viés ou reduzir o MAE a valores inferiores a 0,65 m/s.

Com a metodologia adotada para a seleção dos preditores mostrou-se uma nova maneira de contornar a deficiência das correlações em diagnosticar relações não-lineares, adotando-se de uma seleção em três etapas: a primeira por limiares de correlação; a seguinte por regressões lineares múltiplas passo-a-passo, que eliminam redundâncias lineares; e por fim treinamentos em RNAs por subconjunto de preditores. A comparação entre as famílias de solução de

cada subconjunto foi feita por testes de kolmogorov-Smirnov para duas amostras, levando a resultados conclusivos para o melhor conjunto de preditores.

A configuração das redes neurais apresenta uma grande diversidade de opções, neste sentido os testes de sensibilidade se mostraram uma forma criteriosa de otimização de alguns parâmetros, devendo ser utilizado sempre que possível.

A otimização do número de neurônios ocultos não se mostrou conclusiva, pois apesar de se evidenciar uma convergência do erro para a estação de ROTE, o mesmo não ocorreu para SJCA, o que mostra que há outras restrições ao desempenho do modelo.

A investigação do peso de cada preditor no modelo final de rede neural sugeriu relações interessantes com sistemas sinóticos e fenômenos locais, mostrando a importância de se utilizar variáveis em níveis mais elevados para a previsão do vento.

A inclusão de preditores sintéticos, derivados de outras variáveis, trouxe um incremento de desempenho, uma vez que alguns foram selecionados como preditores no modelo final. No entanto a análise de seu peso não mostrou influência significativa nas saídas das RNAs.

Os testes entre as configurações de RNA obtidas para o interior (SJCA) e litoral (ROTE) não evidenciaram um padrão de comportamento semelhante para estações localizadas mais próximas, o que mostra que a complexidade intrínseca do modelo em RNAs faz com que variações de pequena escala tenham forte influência sobre seu desempenho.

Com base no que foi exposto acima, o principal objetivo do trabalho que era o desenvolvimento de uma metodologia para o refinamento estatístico do vento do modelo Eta foi atingido. Considera-se que pela complexidade tanto dos fenômenos que influenciam o vento em superfície quanto da técnica de modelagem por RNAs, este é um campo de estudos bastante vasto, e que as

análises realizadas foram compatíveis com o escopo de um mestrado. Com isso, espera-se que os resultados apresentados acima sirvam de base para a continuidade das pesquisas nesta área.

## **9 SUGESTÕES DE TRABALHOS FUTUROS**

### **Utilização de arquitetura de RNA com processamento temporal.**

Durante o presente trabalho identificou-se que a RNA empregada muitas vezes não possui recursos suficientes para extrair relações temporais relevantes entre os preditores. Uma alternativa seria o emprego de redes neurais recorrentes, as quais embora mais difíceis de treinar, podem extrair uma infinidade de informações a partir de poucos preditores.

### **Utilização de preditores provenientes de previsão numérica por conjunto (*Ensemble*)**

Previsões por conjunto possibilitam inferir o grau de previsibilidade da atmosfera para um determinado horizonte de previsão, reduzindo as incertezas. A princípio esta poderia ser uma boa forma de adicionar confiabilidade às saídas das redes neurais. Este trabalho passaria necessariamente por uma redefinição dos critérios para inclusão de preditores, a fim de evitar o acréscimo de informação redundante à RNA.

### **Treinamento de redes neurais sazonais**

Sabe-se que os sistemas de tempo atuantes em determinada região geralmente variam ao longo do ano. Neste trabalho as redes foram treinadas considerando-se um período único ao longo do ano. Uma opção seria treinar as redes separadamente para cada época, definindo também os respectivos grupos preditores mais adequados.

### **Utilização de uma curva de potência como saída da RNA**

Conforme discutido no texto acima, a conversão da previsão de vento em previsão de potência eólica degrada seu desempenho devido à não linearidade da curva de potência. Para previsões de potência eólica o ideal seria utilizar uma função de custo para o treinamento das RNAs que seja similar a esta curva de potência, fazendo com que a otimização da RNA seja aplicada diretamente à potência eólica disponível.

### **Separação das condições sinóticas por análise de agrupamento (*cluster*)**

A situação sinótica é provavelmente a principal forçante sobre os campos de vento em superfície. Uma análise interessante seria a separação das condições sinóticas por análise de cluster, valendo-se de isolinhas de geopotencial, para então treinar a rede neural especificamente para cada agrupamento de condições, o que facilitaria bastante o reconhecimento de padrões por parte da RNA.

### **Refinamento de modelos climáticos**

Partindo-se do pressuposto que as relações entre as variáveis atmosféricas e o vento em superfície se preservam para uma escala longa de tempo, modelos em RNAs treinados para curto-prazo a princípio funcionariam como funções de diagnóstico do vento em um determinado local quando acoplados às saídas de modelos climáticos de grande escala. Esta seria uma forma de se avaliar impactos de mudanças climáticas sobre o vento local.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Aquino, R. R. B.; Lira, M. M. S.; Oliveira, J. B.; Carvalho Jr, M. A.; Neto, O. N., Almeida, G. J., Application of Wavelet and Neural Network Models for Wind Speed and Power Generation Forecasting in a Brazilian Experimental Wind Park. **International Joint Conference on Neural Networks**, Proceedings of..., Atlanta, Georgia, USA, June, 2009.
- Amarante, O.C., Brower, M., Zack, J., Sá, A. **Atlas do Potencial Eólico Brasileiro**. CEPEL/ELETRORBRÁS, Brasília, 2001,[A].
- Amarante, O.C., Schultz, D., Bittencourt, R., Rocha, N. Wind-Hydro Complementary Seasonal Regimes in Brazil, **DEWI Magazine**, n. 19, Agosto, 2001, Wilhelmshaven, Germany,[B].
- Arakawa, A.; Lamb, V.R. Computational design of the basic dynamical process of the UCLA general circulation model. **Methods in Computational Physics**, v. 17, p. 173-265, 1977.
- Arya, S. P. **Introduction to Micrometeorology**, San Diego: Academic Press, 2001, 303 p. ISBN 9780120644902.
- Black, T. L. The new NMC mesoscale ETA-model: description and forecast examples, **Weather Forecasting**, vol.9, n.2, p. 265-278, 1994.
- Braga, A. P.; Carvalho A. C. P. F.; Ludermir, T. B. **Fundamentos de redes neurais artificiais**, 11a Escola de Computação / Universidade Federal do Rio de Janeiro, 1998, 246 p.
- Brasil. Centrais Elétricas Brasileiras - ELETRORBRÁS. Sistema de informação do potencial hidrelétrico brasileiro - SIPOT. Rio de Janeiro, abr. 2003.
- Brasil. Empresa de Pesquisas Energéticas - EPE, **Balanco Energético Nacional 2011 – Ano base 2010: Resultados Preliminares**. Rio de Janeiro, EPE, 2011. Disponível em: <<http://ben.epe.gov.br/>>. Acesso em: 12 ago 2011
- Brasil. Empresa de Pesquisas Energéticas - EPE, **PDE 2019 - Plano Decenal de Expansão de Energia**. Brasília: MME/EPE, 2010. Disponível em: <<http://www.epe.gov.br/>>. Acesso em: 12 ago 2011.
- Brasil. **Instituto Nacional de Meteorologia - INMET**. Disponível em: <<http://www.inmet.gov.br/>>. Acesso em: 19 set 2011.
- Brasil. Ministério de Minas e Energia - MME, **PROINFA**. Disponível em: <<http://www.mme.gov.br/programas/proinfa> >. Acesso em: 12 ago 2011.

- Brasil. **Agência Nacional de Energia Elétrica - ANEEL**. Disponível em: <<http://www.aneel.gov.br>>. Acesso em: 16 ago 2011.
- Brasil. Ministério de Ciência e Tecnologia - MCT. **Inventário Brasileiro das Emissões e Remoções Antrópicas – Informações Gerais e Valores Preliminares**, 24 de Novembro de 2009. Disponível em: <http://www.senado.gov.br>. Acesso em: 03 set 2011.
- Burton, T., Sharpe, D., Jenkins, N., Bossanyi, E. **Wind Energy Handbook**, John Willey & Sons, 2001, 617 p.
- Cali, U.; Jursa, R.; Kurt, M.; Lange, B.; Möhrlen, C.; Jørgensen, J.U.; Ernst, B. Artificial Neural Network Based Wind Power Forecasting Using a Multi-Scheme Ensemble Prediction Model, **European Wind Energy Conference – EWEC**, Proceedings..., Bruxelas, Março, 2008.
- Campbell, P.R.J. Short-Term Wind Energy Forecasting, **IEEE Canada Electrical Power Conference**, p. 342-346, 2007
- Câmara de Comercialização de Energia Elétrica – CCEE, Disponível em: <http://www.ccee.org.br/>, acessado em 28/09/2011.
- Costa, G. B. **Análise Numérica Espacial e Temporal do Vento no Estado de Alagoas**, Dissertação de Mestrado, UFAL, Maceió, 2009.
- Chou, S. C., 1996. Modelo Regional Eta. **Climanálise**, Edição comemorativa de 10 anos. Disponível em: <<http://www.cptec.inpe.br/products/climanalise/>>.
- Chen, F., Janjic Z. Mitchell, K. Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. **Boundary-Layer Meteorology**, v. 85, n. 3, p. 391-421, 1997.
- Chou, S.C.; Souza, C.R.; Gomes, J.L.; Evangelista, E.F.D.; Osório, C.; Cataldi, M. refinamento estatístico das previsões horárias de temperatura a 2 m do modelo Eta em estações do nordeste do Brasil, **Revista Brasileira de Meteorologia**, v.22, n.3, p.287-296, 2007.
- Centro de Previsão do Tempo e Estudos Climáticos – CPTEC/INPE, **Eta Model**. Disponível em: <http://etamodel.cptec.inpe.br>>
- Costa, 2009. **Análise espacial e temporal do vento no estado de Alagoas**, Dissertação de Mestrado, UFAL, Maceió.
- Cybenko, G. **Continuous valued neural networks with two hidden layers are sufficient**. Medford, MA: Department of Computer Science, Tufts University, 1988. Technical Report.

- Dalmaz, A. **Estudo do potencial eólico e previsão de ventos para geração de eletricidade em Santa Catarina**, Dissertação de Mestrado, UFSC, Florianópolis, 2007.
- De Maria, P. H., Costa, A. A., Sombra, S. S. Modelagem numérica em alta resolução para previsão de geração eólica no Ceará. **Revista Brasileira de Meteorologia**, v.23, n.4, p. 477-489, 2008.
- Dica, C., Dica, C. I., Vasiliu, D.,Comanescu, G., Ungureanu, M. Wind Power Short-Term Forecasting System, **IEEE Bucharest Power Tech Conference**, Proceedings..., Bucharest, Romania, 2009 .
- Dutra, R. M. e Szklo, A. S. Incentive policies for promoting wind power production in Brazil: Scenarios for the Alternative Energy Sources Incentive Program (PROINFA) under the New Brazilian electric power sector regulation, **Renewable Energy**, v. 33, p. 65-76, 2007.
- Ernst, B., Oakleaf, B., Ahlstron, M. L., Lange, M., Moehrlen, C., Lange, B., Focken, U., Rohrig, K. Predicting the Wind, **IEEE Power & Energy magazine**, v.5, n.6, p. 78-89, 2007.
- European Wind Energy Association - EWEA, **Wind in power 2009**: European statistics, Fev,2009. Disponível em: < [www.ewea.org](http://www.ewea.org)>. Acesso em: 09 set 2011.
- Fahlman, S. E. **An empirical study of learning speed in backpropagation networks**. Technical Report , Carnegie Mellow University, 1988.
- Fan, S., Liau, J. R., Yokoyama, R., Chen, L., Lee, W.J. Forecasting the Wind Generation Using a Two-Stage Network Based on Meteorological Information, **IEEE Transactions on Energy Conversion**, v. 24, n.2, 2009.
- Ferreira, A.; Ludermir, T.B.; Aquino, R. R. B.; Lira, M. M. S.; Neto, O. N., Investigating the use of Reservoir Computing for Forecasting Hourly Wind Speed in Short-Term, **International Joint Conference on Neural Networks**, Proceedings..., p.1649-1656, 2008.
- Foken, T. 50 years of the Monin-Obukhov similarity theory, **Boundary-Layer Meteorology**, vol.119, p. 431-447, 2006.
- Foken, T. **Micrometeorology**. Springer, 2008, 320 p.
- Foresee, F. D., Hagan, M. T. Gauss-Newton Approximation to Bayesian Learning, **IEEE 1997 International Joint Conference on Neural Networks**, Proceedings...,p.1930-1935, 1997.
- Gallus, W.A. The Impact of Step Orography on Flow in the Eta Model: Two Contrasting Examples, **Weather and Forecasting**, v.15, p.630-637, 2000.

- Guarnieri, R. A., 2008. **Emprego de redes neurais artificiais e regressão linear múltipla no refinamento das previsões de radiação solar do modelo ETA**, Dissertação de Mestrado, PGMET, INPE, São José dos Campos.
- Geman, S., Bienenstock, E., Doursat, R. Neural networks and the bias-variance dilemma. **Neural Computation**, v.4, p.1-58, 1992.
- Goldemberg, J., Villanueva, L. D. **Energia, meio ambiente e desenvolvimento**. São Paulo: EDUSP, 2003, 225p. ISBN 8531404525.
- Guenard, V., Kariniotakis, G., Marti, I. ANEMOS Advanced Wind Power Forecasting. Operational Challenges and On-line Performance, **European Wind Energy Conference – EWEC**, Proceedings..., Milan, Italy, 2007
- Haykin, S. **Neural networks: a comprehensive foundation**, Macmillan College Publishing Company, 1994, 696 p.
- Hagan, M.T.; Menhaj, M.B. Training feedforward networks with the Marquardt algorithm, **Neural Networks**, IEEE Transactions on ..., v.5, p.989 – 993, Nov, 1994.
- Hebb, D. O. **The organization of behavior**. New York: Wiley, 1949.
- Holton, J. **An Introduction to Dynamic Meteorology**. Fourth edition. San Diego: Elsevier Academic Press, 2004. 535 p. ISBN 0123540151.
- Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. **Proceedings of the National Academy of Sciences of the United States of America**, v. 79, n. 8, p. 2554-2558, 1982.
- Hunt, J.C.R.; Simpson, J.E., Atmospheric boundary layer over nonhomogeneous terrain, **Engineering Meteorology**, E.J. Plate Ed., p.269-318, Amsterdam, 1982.
- IPCC. **Climate Change 2007: Synthesis Report**. Contribution of Working Groups I, II and III to the Fourth Assessment, Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R.K and Reisinger, A. (eds.)]. IPCC, 2007, Genebra, Suíça, 104 p, ISBN 92-9169-122-4. Disponível em: [www.ipcc.ch](http://www.ipcc.ch). Acesso em: 06 set 2011.
- Janjić, Z. L. The step-mountain Eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes, **Monthly Weather Review**, vol. 122, n.5, p. 927-945, 1994.
- Landberg, L., Giebel, G., Nielsen, A., Nielsen, T., Madsen, H. Short-term Prediction - An Overview, **Wind Energy**, Vol.6, p. 273-280, 2003.

- Lange, M., Focken, U. **Physical Approach to Short-Term Wind Power Prediction**, Springer, 2006, 208 p.
- Lange, M., Focken, U. New developments in wind energy forecasting. **IEEE Power and energy society general meeting** – Conversion and delivery of electrical energy in the 21st century, Proceedings..., Pittsburgh, PA, p.1-8, 2008.
- Lazic, L., Pejanovic, Zivkovic, M. Wind forecasts for wind power generation using the Eta model, **Renewable Energy**, v. 35, n.1, p.1236-1243, 2009.
- Lerner, J., Grundmeyer, M., Garvert, M. The Importance of Wind Forecasting, **Renewable Energy Focus**, p.64-66, Março/Abril, 2009.
- Lyra, A. A. **Testes de sensibilidade e avaliação do modelo Eta para a região nordeste**, Dissertação de Mestrado, PGMET, INPE, São José dos Campos, 2009.
- Lopes, J. C., Palestra, **I Workshop sobre o desenvolvimento de tecnologia nacional para o levantamento de potencial eólico**, INPE, São José dos Campos, Abril, 2010.
- Louka, P., Kallos, G., Galanis, G., Katsafados, P., Pytharoulis, I., Resolution issues on NWP for wind power applications – Downscaling, **European Wind Energy Conference - EWEC, Proceedings**, Athens, Greece, 2006.
- Marquardt, D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. **SIAM Journal on Applied Mathematics**, n.11, v.2, p. 431–441, 1963.
- Martins, F. R., Guarnieri, R. A., Chagas, R. C., Mantelli, S. L., Pereira, E. B., Andrade, E., Thomaz, C. Projeto SONDA – Rede nacional de estações para coleta de dados meteorológicos aplicados ao setor de energia, **I CBENS - Congresso Brasileiro de Energia Solar**, Fortaleza, CE, 2007.
- Martins, F. R., Guarnieri, R. A., Pereira, E. B. O aproveitamento da energia eólica, **Revista Brasileira de Ensino de Física**, v.30, n.1, p.1304, 2008.
- Martins, F. R., Pereira, E. B. Enhancing information for solar and wind energy technology deployment in Brazil, **Energy Policy**, v.39, p.4378-4390, 2011.
- McCulloch e Pitts (1943) McCulloch, W. S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, v. 5, p. 115-133, 1943.
- Mendes, D., Marengo, J. A., 2009. Temporal downscaling: a comparison between artificial neural network and autocorrelation techniques over the

Amazon Basin in present and future climate change scenarios, **Theoretical and Applied Climatology**, publicado on-line em ago/2009.

Mellor, G. L., Yamada, T. Development of a turbulence closure model for geophysical fluid problems. **Review of Geophysics and Space Physics**, v. 20, p. 851- 875, 1982.

Mesinger, F. A blocking technique for representation of mountains in atmospheric models. **Rivista di Meteorologia Aeronautica**, v. 44, n.1-4, p. 195-202, 1984.

Minsky, M.; Papert, S. **Perceptrons**: an introduction to computational geometry. Massachusetts: MIT Press, 1969.

Mitchell, K., 2001: **The Community NOAA Land Surface Model (LSM)**. Disponível em: <[ftp.ncep.noaa.gov/pub/gcp/ldas/noahls/ver\\_2.2](ftp.ncep.noaa.gov/pub/gcp/ldas/noahls/ver_2.2)> Acesso em: 24 ago 2011.

Molion, L. C. B.; Bernardo, S.O. **Dinâmica das chuvas sobre o Nordeste brasileiro**, In: XI CONGRESSO BRASILEIRO DE METEOROLOGIA. Anais... , Sociedade Brasileira de Meteorologia – SBMET, Rio de Janeiro, 2000.

Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., Conzelmann, G., 2009. **Wind Power Forecasting: State-of-the-Art 2009**, Technical Report, Argonne National Laboratory - U.S. Department of Energy, Argonne, EUA, 2009.

Monin, A. S., Obhukov, A. M. Basic laws of turbulent mixing in the ground layer of the atmosphere. **Trudy Geofiz**, inst. AN SSSR, v. 24, n. 151, p. 163-187, 1954.

Monin, A. S., Yaglom, A. M. **Statistical fluid mechanics**: mechanics of turbulence. Cambridge, Massachusetts: MIT Press, p.769, 1971.

Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., Conzelmann, G. **Wind Power Forecasting: State-of-the-Art 2009**, Relatório, Argonne National Laboratory, U.S. Department of Energy, 6 de Novembro, 2009.

Mori, H. Umezawa, Y. Application of NBTree to Selection of Meteorological Variables in Wind Speed Prediction, **Transmission & Distribution Conference & Exposition: Asia and Pacific**, Seoul, Korea, 2009.

NCEP, Workstation Eta Homepage, disponível em [http://www.emc.ncep.noaa.gov/mmb/wrkstn\\_eta/](http://www.emc.ncep.noaa.gov/mmb/wrkstn_eta/), consulta realizada em 25/09/2011.

- Nicković, S., Mihailović, D., Rajković, B. **Scientific documentation of the Eta model**. Unofficial documentation for WMO publication. Geneva, Switzerland, 1998.
- Orlanski, I., A rational subdivision of scales for atmospheric process. **Bull. of the American Meteorological Society**, v.56, p.529-530,1975.
- Pereira, E.B.,Lima, J.H.G. **Solar and Wind Energy Resource Assessment in Brazil – SWERA**, MCT/INPE, São José dos Campos, SP, 2008.
- Pes, M. P. **Estudo do impacto das mudanças climáticas no potencial eólico no estado do Rio Grande do Sul para os períodos de 2010 a 2040 e 2070 a 2100**, Dissertação de Mestrado, PGMET, INPE, São José dos Campos, 2010.
- Pimenta, F., Kempton, W., Garvine, R. Combining meteorological stations and satellite data to evaluate the offshore wind power resource of Southeastern Brazil, **Renewable Energy**, v.33, p.2375-2387, 2008.
- Ramirez-Rosado,I. J., Jimenez, L., Monteiro, C., Sousa, J., Bessa, R. Comparison of two new short-term wind-power forecasting systems. **Renewable Energy**, v.34, p.1848-1854, 2009.
- Reed, R. Pruning algorithms – A survey. **IEEE Transactions on Neural Networks**, v.4, n.5, p. 740-746, 1993.
- Roballo, S. T., Fisch, G. Escoamento atmosférico no centro de Lançamento de Alcântara: Parte I-Aspectos Observacionais. **Rev. Bras. de Meteor.**, v.23, n.4, p.510-519, 2008.
- Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychol. Rev.**, v. 65, p. 386-408, 1958.
- Rozante, J. R. **O impacto do aninhamento no modelo Eta nas previsões de tempo**.. Dissertação de Mestrado, PGMET, INPE, São José dos Campos, 2001.
- Rumelhart, D. E.; Hinton G. E.; Williams, R. J. Learning representations by backpropagation errors. **Nature**, v. 323, p. 533-536, 1986.
- Sailor, D.J.; Hu, T.; Li, X.; Rosen J.N. A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change. **Renewable Energy**, n.19, p. 359-378, 2000.
- Sanz, S.S., Bellido, A. P., Garcia, E. O., Figueras, A.P., Prieto, L., Paredes, D., Correoso, F. Short-term Wind Speed Prediction by Hybridizing Global and Mesoscale Forecasting Models with Artificial Neural Networks. **8th**

- International Conference on Hybrid Intelligent Systems**, Proceedings..., p.608-612, 2008
- Sideratos, G., Hatziargyriou, N. An Advanced Statistical Method for Wind Power Forecast. **IEEE Transactions on Power Systems**, v. 22, n.1, p. 258-265, 2007.
- SONDA - Sistema de Organização Nacional de Dados Ambientais. Disponível em: <<http://sonda.cptec.inpe.br/>>
- Stull, R. B. **An introduction to boundary layer meteorology**. Dordrecht: Kluwer Academic Publishers, 1988.
- Sun, J., Mahrt, L. Determination of surface fluxes from the surface radiative temperature. **Journal of the Atmospheric Sciences**, v. 52, n. 8, p. 1096-1104, 1995.
- Teunissen, H.W.; Shokr, M.E.; Bowen, A.J. ;Wood, C.J.;Green,D.W.R, The Askervein Hill Project: wind tunnel simulations at three length scales, *Boundary Layer Meteorology*, n.40, p.1-29, 1987.
- Thor, S.E. **Optimal Use of Available Information in Wind Power Forecasting**, IEA Joint Assembly Symposium, Relatório, CENER, Madrid, Setembro, 2008.
- Tsikalakis, A.G. Hatziargyriou, N.D., Katsigiannis, Y.A.,Georgilakis, P.S. Impact of Wind Power Forecasting Error Bias on the Economic Operation of Autonomous Power Systems, **Wind Energy**, n.12, p.315–331, 2009.
- Van der Hoven, I., Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour. **Journal of Meteorology**, v.14, p.160-164 1957.
- Wallace, J. M.; Hobbs, P. V. **Atmospheric science: an introductory survey**. San Diego: Academic Press, 2001. 483p.
- Wilks, D. S. **Statistical Methods in the Atmospheric Sciences**, Elsevier, 2006, 649 p.
- World Wind Energy Association - WWEA. **World Wind Energy Report 2009**. Disponível em: <<http://www.wwindea.org/>>. Acesso em: 15 set 2011.
- Wu, Y. K., Hong, J. S. A literature review of wind forecasting, **IEEE Power Tech 2007**, Proceedings..., p. 504-509, Lausanne, 2007.

