

**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**PREVISÃO DE VENTOS E GERAÇÃO EÓLICA DO
SISTEMA NE: ANALISANDO DIVERSOS SÍTIOS E
BUSCANDO A MELHOR MODELAGEM ATRAVÉS
DA INTELIGÊNCIA ARTIFICIAL**

por

HUGO TAVARES VIEIRA GOUVEIA

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica da
Universidade Federal de Pernambuco como parte dos requisitos para a obtenção do grau de
Mestre em Engenharia Elétrica.

ORIENTADOR: RONALDO RIBEIRO BARBOSA DE AQUINO, D.Sc.

Recife, Dezembro de 2011.

Catalogação na fonte
Bibliotecária Margareth Malta, CRB-4 / 1198

G719p Gouveia, Hugo Tavares Vieira.

Previsão de ventos e geração eólica do sistema NE: analisando diversos sítios e buscando a melhor modelagem através da inteligência artificial / Hugo Tavares Vieira Gouveia. - Recife: O Autor, 2011.

xvi, 116 folhas, il., gráf., tabs.

Orientador: Prof. DSc. Ronaldo Ribeiro Barbosa de Aquino.

Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Elétrica, 2011.

Inclui Referências Bibliográficas e Apêndice.

1. Engenharia Elétrica. 2. Energia Eólica. 3. Inteligência artificial. 4. Lógica Fuzzy. 5. Previsão de ventos. 6. Redes neurais. 7. Análise de séries temporais. 8. Transformada Wavelet. I. Aquino, Ronaldo Ribeiro Barbosa de. (Orientador). II. Título.

UFPE

621.3 CDD (22. ed.)

BCTG/2012-089



Universidade Federal de Pernambuco

Pós-Graduação em Engenharia Elétrica

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE
DISSERTAÇÃO DO MESTRADO ACADÊMICO DE

HUGO TAVARES VIEIRA GOUVEIA

TÍTULO

**"PREVISÃO DE VENTOS E GERAÇÃO EÓLICA DO SISTEMA DO NE:
ANALISANDO DIVERSOS SÍTIOS E BUSCANDO A MELHOR
MODELAGEM ATRAVÉS DA INTELIGÊNCIA ARTIFICIAL"**

A comissão examinadora composta pelos professores: RONALDO RIBEIRO BARBOSA DE AQUINO, DEE/UFPE, MILDE MARIA DA SILVA LIRA, DEE/UFPE e BENEMAR ALENCAR DE SOUZA, DEE/UFCG sob a presidência do primeiro, consideram o candidato **HUGO TAVARES VIEIRA GOUVEIA**
APROVADO.

Recife, 22 de dezembro de 2011.

RAFAEL DUEIRE LINS
Coordenador do PPGEE

RONALDO RIBEIRO BARBOSA DE AQUINO
Orientador e Membro Titular Interno

BENEMAR ALENCAR DE SOUZA
Membro Titular Externo

MILDE MARIA DA SILVA LIRA
Membro Titular Externo

Dedico este trabalho à minha amada esposa Lilian, aos meus pais, Evandro e Maria Helena, às minhas irmãs, Marina e Elisa, e aos meus queridos avós, Edman e Berenice.

AGRADECIMENTOS

Agradeço a Deus pela saúde a mim concedida, bem como pelas condições intelectuais para o desenvolvimento deste trabalho.

À minha esposa Lilian, pela paciência e tolerância à minha falta de atenção durante as horas de estudo e pesquisa.

Ao Prof. Ronaldo Ribeiro Barbosa de Aquino, pela oportunidade de poder realizar o mestrado sob sua orientação, e também por todo o seu empenho, sabedoria, compreensão, segurança e, acima de tudo, agradeço pela amizade e companheirismo.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de estudos, sem a qual este trabalho não teria sido realizado.

Agradeço ainda, a todas as pessoas que de alguma forma contribuíram para a conclusão desta dissertação.

Resumo da Dissertação apresentada à UFPE como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica.

PREVISÃO DE VENTOS E GERAÇÃO EÓLICA DO SISTEMA NE: ANALISANDO DIVERSOS SÍTIOS E BUSCANDO A MELHOR MODELAGEM ATRAVÉS DA INTELIGÊNCIA ARTIFICIAL

Hugo Tavares Vieira Gouveia

Dezembro/2011

Orientador: Prof. Ronaldo Ribeiro Barbosa de Aquino, D. Sc.

Área de Concentração: Processamento de Energia.

Palavras-chave: Energia Eólica, Inteligência Artificial, Lógica Fuzzy, Previsão de Ventos, Redes Neurais, Análise de Séries Temporais, Transformada *Wavelet*.

Número de Páginas: 132.

RESUMO: A previsão de ventos é de extrema importância para auxiliar nos estudos de planejamento e programação da operação da geração eólica. Vários estudos já comprovaram que o potencial eólico brasileiro, principalmente no Nordeste, onde os ventos apresentam uma importante característica de complementaridade em relação às vazões do rio São Francisco, pode contribuir significativamente para o suprimento de energia elétrica. Entretanto, o uso das forças dos ventos para produção de energia elétrica produz alguns inconvenientes, tais como, incertezas na geração e a dificuldade no planejamento e operação do sistema elétrico. Este trabalho propõe e desenvolve modelos de previsões de velocidades médias horárias de ventos e geração eólica a partir de técnicas de Redes Neurais Artificiais, Lógica *Fuzzy* e Análise *Wavelet*. Os modelos foram ajustados para realizar previsões com passos variáveis de até vinte e quatro horas. Para as previsões realizadas com alguns dos modelos desenvolvidos, os ganhos em relação aos modelos de referência foram da ordem de 80% para as previsões com passo de uma hora. Os resultados demonstraram que a Análise *Wavelet* aliada às ferramentas de inteligência artificial fornecem previsões muito mais confiáveis do que aquelas obtidas com os modelos de referência, principalmente para as previsões com passos de 1 – 6 horas.

Abstract of Dissertation presented to UFPE as a partial fulfillment of the requirements for
the degree of Master in Electrical Engineering.

**WIND FORECASTING AND WIND POWER GENERATION
IN BRAZILIAN NORTHEAST: ANALYZING DIFFERENT
SITES AND LOOKING FOR THE BEST MODELING BASED
ON ARTIFICIAL INTELLIGENCE**

Hugo Tavares Vieira Gouveia

December/2011

Supervisor: Prof. Ronaldo Ribeiro Barbosa de Aquino, D. Sc.

Area of Concentration: Energy Processing.

Keywords: Wind Energy, Artificial Intelligence, Fuzzy Logic, Wind Forecasting, Neural Networks, Time Series Analysis, Wavelet Transforms.

Number of Pages: 132.

ABSTRACT: Wind forecasting is extremely important to assist in planning and programming studies for the operation of wind power generation. Several studies have shown that the Brazilian wind potential can contribute significantly to the supply of electricity, especially in the Northeast, where the winds have an important feature of complementarity in relation to the flows of the San Francisco River. However, the use of wind to generate electricity has some drawbacks, such as uncertainties in generation and some difficulty in planning and operation of the power system. This work proposes and develops models to forecast hourly average wind speeds and wind power generation based on techniques of Artificial Neural Networks, Fuzzy Logic and Wavelets. The models were adjusted for forecasting with variable steps up to twenty-four hours ahead. The gain of some models developed in relation to the reference models were approximately 80% for forecasts in a period of one hour ahead. The results showed that the wavelet analysis combined with artificial intelligence tools provide forecasts more reliable than those obtained with the reference models, especially for forecasts in a period of 1 to 6 hours ahead.

SUMÁRIO

Lista de Figuras	x
Lista de Tabelas	xvi
CAPÍTULO 1	1
1. Introdução	1
1.1 Caracterização do Problema	1
1.2 Objetivos	2
1.3 Estrutura da dissertação	3
CAPÍTULO 2	4
2. Potencial Eólico e Revisão Bibliográfica	4
2.1 Potencial Eólico no Nordeste Brasileiro	4
2.2 Geração de Energia Eólio-Elétrica	7
2.2.1 Relação entre a velocidade e a potência dos ventos	9
2.2.2 Potência extraída dos ventos	10
2.2.3 Curva de Potência	12
2.2.4 Densidade do ar e velocidade do vento em função da altura	13
2.2.5 Distribuição das velocidades dos ventos	15
2.2.6 Cálculo da energia elétrica gerada	18
2.3 Revisão Bibliográfica	19
CAPÍTULO 3	22
3. Inteligência Artificial e <i>Wavelets</i>	22
3.1 Redes Neurais Artificiais	22
3.1.1 O Neurônio Artificial	23
3.1.2 Funções de Ativação	24
3.1.3 Arquitetura da Rede Neural Artificial	25
3.1.4 Processo de Aprendizagem	26
3.1.5 Algoritmo <i>Backpropagation</i>	27
3.1.6 Algoritmo <i>Resilient Propagation</i>	29
3.1.7 Algoritmo <i>Levenberg-Marquardt</i>	31
3.2 Lógica <i>Fuzzy</i>	32
3.2.1 Sistema de inferência <i>Fuzzy</i>	33
3.2.2 Sistema Adaptativo de Inferência Neuro- <i>Fuzzy</i>	36
3.3 Análise <i>Wavelet</i>	38

3.3.1	Transformada <i>Wavelet</i>	39
3.3.2	Transformada <i>Wavelet</i> Contínua	40
3.3.3	Transformada <i>Wavelet</i> Discreta.....	41
3.3.4	Análise de Multiresolução	43
3.3.5	Filtragem em um estágio: aproximações e detalhes	43
3.3.6	Decomposição em múltiplos níveis.....	45
3.4	<i>Software</i> e parâmetros utilizados.....	46
CAPÍTULO 4		48
4.	Avaliação de Desempenho dos Modelos de Previsão	48
4.1	Séries Temporais	48
4.1.1	Previsão de Séries Temporais.....	48
4.1.2	Autocorrelação.....	49
4.2	Notações	50
4.3	Modelos de Referência	50
4.4	Definição dos Erros de Previsão.....	52
4.5	Critérios para Comparação dos Modelos.....	54
CAPÍTULO 5		56
5.	Desenvolvimento dos Modelos de Previsão	56
5.1	Séries de Velocidades Médias Horárias	56
5.1.1	Estatística Descritiva	56
5.2	Modelos de Previsão.....	60
5.2.1	Modelos <i>RNA(LM)</i> , <i>RNA(RP)</i> e <i>ANFIS</i>	60
5.2.2	Modelos <i>TWRNA(LM)</i> , <i>TWRNA(RP)</i> e <i>TWANFIS</i>	61
5.3	Procedimentos para o treinamento e ajuste dos modelos	62
5.3.1	Normalização dos dados	62
5.3.2	Partição das bases de dados	63
5.3.3	Determinação do número de neurônios da camada oculta	63
5.3.4	Escolha da melhor rede para cada modelo neural	66
5.3.5	Escolha do melhor Sistema de Inferência <i>Fuzzy</i>	69
CAPÍTULO 6		73
6.	Previsões e Comparações entre os Modelos.....	73
6.1	Previsões de velocidades para MACAU	73
6.2	Previsões de geração para MACAU.....	80

CAPÍTULO 7	88
7. Conclusões e Propostas para Trabalhos futuros	88
7.1 Conclusões.....	88
7.2 Propostas para trabalhos futuros.....	89
APÊNDICE A	90
A.1 Previsões para MOSSORÓ.....	90
A.2 Previsões para NATAL	101
Referências Bibliográficas.....	113

LISTA DE FIGURAS

Figura 2.1 – <i>Distribuição geral dos ventos. Fonte: [7].</i>	4
Figura 2.2 – <i>Modelo de relevo e vegetação do Brasil. Fonte: [7].</i>	5
Figura 2.3 – <i>Potencial eólico estimado para a região Nordeste do Brasil. Fonte: [7].</i>	7
Figura 2.4 – <i>Principais componentes de um aerogerador de velocidade de rotação constante. Fonte: [2].</i>	8
Figura 2.5 – <i>Esquema de funcionamento de um aerogerador de rotação variável. Fonte: [2].</i>	9
Figura 2.6 – <i>Eficiência do rotor em função da razão das velocidades.</i>	12
Figura 2.7 – <i>Curva de potência e curva do coeficiente de potência de um aerogerador.</i> ..	13
Figura 2.8 – <i>Função distribuição de probabilidades de Weibull com parâmetro de escala $c = 10$ m/s e parâmetros de forma $k = 1, 2$ e 3.</i>	16
Figura 2.9 – <i>Função distribuição de probabilidades de Weibull com parâmetro de forma $k = 3$ e parâmetros de escala c variando entre 4 e 12 m/s.</i>	17
Figura 2.10 – <i>Distribuições das velocidades ($k = 4$ e $c = 10$ m/s) e da energia gerada.</i> ...	18
Figura 3.1 – <i>Modelo de um neurônio artificial. Fonte: [6].</i>	23
Figura 3.2 – <i>Gráfico da função Sigmóide Logística.</i>	24
Figura 3.3 – <i>Gráfico da função Tangente Hiperbólica.</i>	25
Figura 3.4 – <i>Esquema de uma rede progressiva de três camadas.</i>	26
Figura 3.5 – <i>Estrutura do modelo ANFIS. Fonte: [28].</i>	37
Figura 3.6 – <i>Comparação entre uma senóide (esquerda) e uma wavelet (direita).</i>	39
Figura 3.7 – <i>Esquema de um plano tempo x frequência. Fonte: [6].</i>	40
Figura 3.8 – <i>Passo 2 da Transformada Wavelet Contínua.</i>	40
Figura 3.9 – <i>Passo 3 da Transformada Wavelet Contínua.</i>	40
Figura 3.10 – <i>Passo 4 da Transformada Wavelet Contínua.</i>	41
Figura 3.11 – <i>Deslocamento da wavelet na Transformada Wavelet Contínua.</i>	41

Figura 3.12 – Reticulado no plano escala x translação para a TWD. Fonte: [6].	42
Figura 3.13 – Processo básico de filtragem de sinais.	44
Figura 3.14 – Filtragem de sinais utilizando downsampling.	44
Figura 3.15 – Obtenção dos coeficientes da TWD através da filtragem em único estágio.	45
Figura 3.16 – Árvore de decomposição de um sinal em três níveis. Fonte: [6].	46
Figura 5.1 – Séries de velocidades e histogramas de MACAU, MOSSORÓ e NATAL.....	58
Figura 5.2 – Dia típico e comportamento sazonal de MACAU, MOSSORÓ e NATAL.....	59
Figura 5.3 – Coeficientes de autocorrelação de MACAU, MOSSORÓ e NATAL.....	59
Figura 5.4 – Entradas e saída dos modelos RNA(LM), RNA(RP) e ANFIS.	60
Figura 5.5 – Entradas e saída dos modelos TWRNA(LM), TWRNA(RP) e TWANFIS.....	61
Figura 5.6 – Número de neurônios na camada oculta.	64
Figura 5.7 – Tempos médios de treinamento dos modelos RNA(LM) e RNA(RP).	65
Figura 5.8 – Tempos médios de treinamento dos modelos TWRNA(LM) e TWRNA(RP). ..	65
Figura 5.9 – MAPE do conjunto de validação para as melhores arquiteturas das redes..	66
Figura 5.10 – Representação esquemática da matriz k -fold.	67
Figura 5.11 – MAPE do conjunto de teste para as melhores redes.....	69
Figura 5.12 – Raios dos clusters para os modelos ANFIS e TWANFIS.	70
Figura 5.13 – Tempos de treinamento dos modelos ANFIS.	71
Figura 5.14 – Tempos de treinamento dos modelos TWANFIS.	71
Figura 5.15 – MAPE do conjunto de teste para os melhores Sistemas de Inferência Fuzzy.	72
Figura 6.1 – MAE das previsões de velocidades em MACAU.....	73
Figura 6.2 – RMSE das previsões de velocidades em MACAU.....	74
Figura 6.3 – MAPE das previsões de velocidades em MACAU.	74
Figura 6.4 – Ganhos das previsões de velocidades em MACAU.....	75

Figura 6.5 – Erros quadráticos acumulados das previsões de velocidades em MACAU com passo de 1 hora.....	76
Figura 6.6 – Erros quadráticos acumulados das previsões de velocidades em MACAU com passo de 12 horas.	76
Figura 6.7 – Erros quadráticos acumulados das previsões de velocidades em MACAU com passo de 24 horas.	77
Figura 6.8 – Diagramas de dispersão das previsões de velocidades em MACAU com passo de 1 hora.....	78
Figura 6.9 – Diagramas de dispersão das previsões de velocidades em MACAU com passo de 12 horas.	78
Figura 6.10 – Diagramas de dispersão das previsões de velocidades em MACAU com passo de 24 horas.	79
Figura 6.11 – Curva de potência do aerogerador.....	80
Figura 6.12 – MAE das previsões de geração de MACAU.	81
Figura 6.13 – RMSE das previsões de geração de MACAU.	81
Figura 6.14 – Ganhos das previsões de geração de MACAU.	82
Figura 6.15 – Erros quadráticos acumulados das previsões de geração de MACAU com passo de 1 hora.....	82
Figura 6.16 – Erros quadráticos acumulados das previsões de geração de MACAU com passo de 12 horas.	83
Figura 6.17 – Erros quadráticos acumulados das previsões de geração de MACAU com passo de 24 horas.	83
Figura 6.18 – Diagramas de dispersão das previsões de geração de MACAU com passo de 1 hora.....	84
Figura 6.19 – Diagramas de dispersão das previsões de geração de MACAU com passo de 12 horas.	84
Figura 6.20 – Diagramas de dispersão das previsões de geração de MACAU com passo de 24 horas.	85
Figura 6.21 – Frequência percentual dos erros de geração de MACAU.....	86
Figura 6.22 – Previsões de geração de MACAU para o dia 01/01/2009.....	86
Figura 6.23 – Previsões de geração de MACAU para o dia 01/06/2009.....	87

Figura A.1 – <i>MAE das previsões de velocidades em MOSSORÓ</i>	90
Figura A.2 – <i>RMSE das previsões de velocidades em MOSSORÓ</i>	91
Figura A.3 – <i>MAPE das previsões de velocidades em MOSSORÓ</i>	91
Figura A.4 – <i>Ganhos das previsões de velocidades em MOSSORÓ</i>	92
Figura A.5 – <i>Erros quadráticos acumulados das previsões de velocidades em MOSSORÓ com passo de 1 hora</i>	92
Figura A.6 – <i>Erros quadráticos acumulados das previsões de velocidades em MOSSORÓ com passo de 12 horas</i>	93
Figura A.7 – <i>Erros quadráticos acumulados das previsões de velocidades em MOSSORÓ com passo de 24 horas</i>	93
Figura A.8 – <i>Diagramas de dispersão das previsões de velocidades em MOSSORÓ com passo de 1 hora</i>	94
Figura A.9 – <i>Diagramas de dispersão das previsões de velocidades em MOSSORÓ com passo de 12 horas</i>	94
Figura A.10 – <i>Diagramas de dispersão das previsões de velocidades em MOSSORÓ com passo de 24 horas</i>	95
Figura A.11 – <i>MAE das previsões de geração de MOSSORÓ</i>	95
Figura A.12 – <i>RMSE das previsões de geração de MOSSORÓ</i>	96
Figura A.13 – <i>Ganhos das previsões de geração de MOSSORÓ</i>	96
Figura A.14 – <i>Erros quadráticos acumulados das previsões de geração de MOSSORÓ com passo de 1 hora</i>	97
Figura A.15 – <i>Erros quadráticos acumulados das previsões de geração de MOSSORÓ com passo de 12 horas</i>	97
Figura A.16 – <i>Erros quadráticos acumulados das previsões de geração de MOSSORÓ com passo de 24 horas</i>	98
Figura A.17 – <i>Diagramas de dispersão das previsões de geração de MOSSORÓ com passo de 1 hora</i>	98
Figura A.18 – <i>Diagramas de dispersão das previsões de geração de MOSSORÓ com passo de 12 horas</i>	99
Figura A.19 – <i>Diagramas de dispersão das previsões de geração de MOSSORÓ com passo de 24 horas</i>	99

Figura A.20 – <i>Frequência percentual dos erros de geração de MOSSORÓ</i>	100
Figura A.21 – <i>Previsões de geração de MOSSORÓ para o dia 01/03/2009</i>	100
Figura A.22 – <i>Previsões de geração de MOSSORÓ para o dia 01/08/2009</i>	101
Figura A.23 – <i>MAE das previsões de velocidades em NATAL</i>	101
Figura A.24 – <i>RMSE das previsões de velocidades em NATAL</i>	102
Figura A.25 – <i>MAPE das previsões de velocidades em NATAL</i>	102
Figura A.26 – <i>Ganhos das previsões de velocidades em NATAL</i>	103
Figura A.27 – <i>Erros quadráticos acumulados das previsões de velocidades em NATAL com passo de 1 hora</i>	103
Figura A.28 – <i>Erros quadráticos acumulados das previsões de velocidades em NATAL com passo de 12 horas</i>	104
Figura A.29 – <i>Erros quadráticos acumulados das previsões de velocidades em NATAL com passo de 24 horas</i>	104
Figura A.30 – <i>Diagramas de dispersão das previsões de velocidades em NATAL com passo de 1 hora</i>	105
Figura A.31 – <i>Diagramas de dispersão das previsões de velocidades em NATAL com passo de 12 horas</i>	105
Figura A.32 – <i>Diagramas de dispersão das previsões de velocidades em NATAL com passo de 24 horas</i>	106
Figura A.33 – <i>MAE das previsões de geração de NATAL</i>	106
Figura A.34 – <i>RMSE das previsões de geração de NATAL</i>	107
Figura A.35 – <i>Ganhos das previsões de geração de NATAL</i>	107
Figura A.36 – <i>Erros quadráticos acumulados das previsões de geração de NATAL com passo de 1 hora</i>	108
Figura A.37 – <i>Erros quadráticos acumulados das previsões de geração de NATAL com passo de 12 horas</i>	108
Figura A.38 – <i>Erros quadráticos acumulados das previsões de geração de NATAL com passo de 24 horas</i>	109

Figura A.39 – Diagramas de dispersão das previsões de geração de NATAL com passo de 1 hora.....	109
Figura A.40 – Diagramas de dispersão das previsões de geração de NATAL com passo de 12 horas.....	110
Figura A.41 – Diagramas de dispersão das previsões de geração de NATAL com passo de 24 horas	110
Figura A.42 – Frequência percentual dos erros de geração de NATAL.....	111
Figura A.43 – Previsões de geração de NATAL para o dia 01/10/2009.....	111
Figura A.44 – Previsões de geração de NATAL para o dia 01/03/2010.....	112

LISTA DE TABELAS

Tabela 2.1 – <i>Coeficiente de rugosidade para diversos tipos de terreno. Fonte: [9].</i>	14
Tabela 5.1 – <i>Dados utilizados para o desenvolvimento dos modelos de previsão.</i>	57
Tabela 5.2 – <i>Estatísticas das séries de velocidades de MACAU, MOSSORÓ e NATAL.</i> ...	57
Tabela 5.3 – <i>Experimentos realizados no método de validação cruzada. Fonte: [6].</i>	68

CAPÍTULO 1

1. INTRODUÇÃO

Neste capítulo, o problema em estudo é contextualizado e caracterizado, os objetivos são descritos, e a estrutura da presente dissertação é apresentada.

1.1 Caracterização do Problema

Dentre todas as fontes alternativas de energia exploradas atualmente, a energia eólica é, sem dúvida, uma das mais bem sucedidas. Uma razão para este fato é a política de incentivo feita por vários países, assegurando a compra da energia elétrica produzida à partir das fontes eólicas, ainda que ela não ofereça preços competitivos. A Alemanha e a Dinamarca foram os primeiros países a adotar as políticas de incentivo ao desenvolvimento da geração eólio-elétrica, seguidos por diversos países, inclusive pelo Brasil, com a criação do “PROINFA”, Programa de Incentivo às Fontes Alternativas de Energia Elétrica [1].

O avanço tecnológico, tanto em pesquisa quanto em desenvolvimento, está fazendo com que ocorra uma rápida redução no custos de instalação e produção da energia elétrica à partir da energia dos ventos [1]. Este constante desenvolvimento tecnológico torna possível a concorrência das fontes eólicas com as tradicionais fontes de geração térmica, uma vez que a energia eólica proporciona a possibilidade de geração de energia elétrica em grandes blocos [2].

A energia disponibilizada pelos ventos é gratuita, logo, toda a energia elétrica gerada a partir de fontes eólicas é bem-vinda. Entretanto, a intermitência dos ventos é o grande desafio a ser enfrentado para que a energia eólica se torne uma fonte confiável para a produção de energia elétrica em grandes blocos. A inserção em larga escala de parques eólicos nos sistemas elétricos de transmissão e distribuição exige respostas para diversas questões, tais como, padrões para interconexão, qualidade de energia, capacidade dos sistemas de transmissão e suas futuras expansões, estabilidade e confiabilidade dos sistemas de potência, entre outras.

Com o aumento da capacidade de geração eólica instalada, os operadores do sistema elétrico devem saber como lidar com esse importante montante de energia flutuante. Portanto, uma área de pesquisa de extrema importância para o setor elétrico está relacionada com as previsões de curto prazo para a geração eólica. As escalas de tempo

envolvidas nas previsões de curto prazo são da ordem de dias (para o horizonte de previsão) e de minutos a horas (para o passo da previsão) [3].

A previsão de velocidades dos ventos desempenha um papel fundamental para enfrentar os desafios relacionados com a geração eólica. Estas previsões fazem parte da previsão do tempo há muitas décadas e são utilizadas para navegação, orientação de mísseis, controle de tráfego aéreo e lançamento de satélites [4]. Nas duas últimas décadas, o vento está sendo utilizado em larga escala para geração de energia elétrica e a previsão de velocidades dos ventos ganhou uma atenção especial. A previsão de ventos para estações meteorológicas é muito diferente da previsão de ventos para a geração de energia. Para os parques eólicos, a previsão de ventos no curto prazo devem ser precisas para que se mantenha a estabilidade na geração de energia elétrica [5].

Os sistemas de previsão de geração eólica, em países onde existe uma forte participação de energia eólica na matriz elétrica, representam hoje em dia uma grande parcela de investimentos em centros de despacho. Notadamente, podem ser citados os seguintes países: Espanha, Alemanha e Dinamarca. Nestes países, existem diversos centros de pesquisa trabalhando continuamente para o desenvolvimento de modelos para previsão de geração eólica em curto prazo.

A previsão de ventos e geração eólica é de extrema importância para auxiliar nos estudos de planejamento e programação da operação da geração do sistema hidrotérmico e eólico. Segundo [6], a previsão eficiente de ventos e geração eólica pode contribuir de forma positiva das seguintes maneiras: facilitando a comercialização no mercado de energia elétrica; subsidiando na solução do problema de otimização do despacho da geração do sistema hidrotérmico e eólico; e fornecendo dados para os sistemas de controles dos parques eólicos de geração.

1.2 Objetivos

O objetivo principal deste trabalho é desenvolver diversos modelos de previsão de velocidades de ventos e geração de energia elétrica a partir das previsões de velocidades utilizando-se a curva de potência dos aerogeradores. Para o desenvolvimento dos modelos de previsão, são utilizadas as Redes Neurais Artificiais (RNA), Lógica Fuzzy (LF) e Análise *Wavelet*.

Os modelos propostos realizam previsões das velocidades médias horárias com passos de previsão que variam de uma até vinte e quatro horas. Todos os modelos

desenvolvidos são univariados, ou seja, as variáveis de entrada e saída do modelo são apenas as velocidades dos ventos. A modelagem univariada foi escolhida devido à escassez de dados relacionados a geração eólica no Brasil.

Pretende-se, com este trabalho, contribuir para o desenvolvimento inicial de uma ferramenta computacional eficaz e confiável para a realização de previsões da geração eólio-elétrica de curto prazo. Para isto, optou-se por realizar o desenvolvimento dos modelos utilizando séries temporais de ventos da região Nordeste do Brasil, por ser uma região que apresenta condições de vento extremamente favoráveis.

1.3 Estrutura da dissertação

A estrutura desta dissertação é formada por sete capítulos. O primeiro deles é esta introdução. O segundo capítulo apresenta o potencial eólico da região Nordeste, trata da geração da energia elétrica a partir dos ventos e traz uma breve revisão bibliográfica sobre alguns modelos de previsão de ventos e geração eólica. No terceiro capítulo, são apresentadas as ferramentas de inteligência artificial utilizadas para o desenvolvimento dos modelos propostos nesta dissertação, e também as *wavelets*. O capítulo quatro trata dos critérios para avaliação dos modelos desenvolvidos. No quinto capítulo são apresentadas as topologias e os ajustes dos modelos propostos. O sexto capítulo traz os resultados das previsões realizadas com os melhores modelos desenvolvidos e os compara com resultados obtidos com modelos de referência. Com o sétimo capítulo a dissertação é concluída, fazendo um breve resumo dos resultados e contribuições dadas. Além disso, são apresentadas propostas para trabalhos futuros com os quais se possa dar continuidade às pesquisas no contexto aqui desenvolvido.

CAPÍTULO 2

2. POTENCIAL EÓLICO E REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta, na primeira seção, uma descrição do potencial eólico da região Nordeste do Brasil. Em seguida, na segunda seção, são apresentados os conceitos básicos sobre a geração de energia elétrica a partir da energia dos ventos. A terceira e última seção traz uma breve revisão bibliográfica sobre alguns modelos de previsão de ventos e geração eólica desenvolvidos no Brasil e no mundo.

2.1 Potencial Eólico no Nordeste Brasileiro

A distribuição geral dos ventos sobre o Brasil é controlada pelos aspectos da circulação geral planetária da atmosfera próxima, conforme se apresenta na Figura 2.1. Dentre esses aspectos, sobressaem os sistemas de alta pressão Anticiclone Subtropical do Atlântico Sul e do Atlântico Norte e a faixa de baixas pressões da Depressão Equatorial [7].

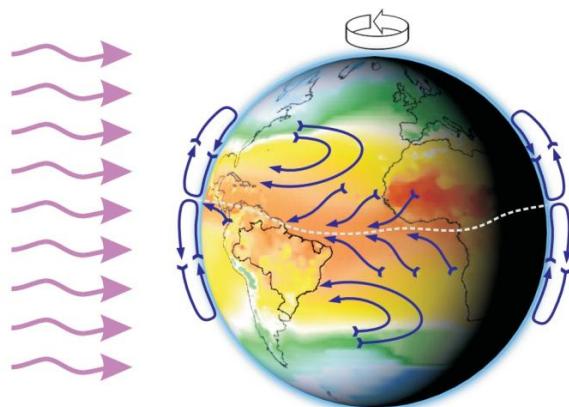


Figura 2.1 – Distribuição geral dos ventos. Fonte: [7].

A posição média da Depressão Equatorial estende-se de oeste a leste ao longo da região Norte do Brasil e sobre o Oceano Atlântico adjacente. Ela coincide com a localização e orientação da Bacia Amazônica, no centro da qual existe uma faixa persistente de baixas pressões. A Depressão Equatorial é geralmente uma zona de pequenos gradientes de pressão e ventos fracos. Ao norte da Depressão Equatorial os ventos são persistentes de leste a nordeste. Ao sul, os ventos são persistentes de leste a sudeste entre a Depressão Equatorial e o Anticiclone Subtropical Atlântico, o qual tem uma posição média anual próxima a 30° S, 25° W. Esse perfil geral de circulação atmosférica

induz ventos de leste ou nordeste sobre o território brasileiro ao norte da Bacia Amazônica e no litoral nordeste. Os ventos próximos à superfície são geralmente fracos ao longo da Depressão Equatorial, porém aumentam de intensidade ao norte e ao sul dessa faixa. A área entre a Depressão Equatorial e a latitude de 10° S é dominada pelos ventos alísios de leste a sudeste. Ao sul da latitude 10° S, até o extremo sul brasileiro, prevalecem os efeitos ditados pela dinâmica entre o centro de alta pressão Anticiclone Subtropical Atlântico, os deslocamentos de massas polares e a Depressão do Nordeste da Argentina – centro de baixas pressões a leste dos Andes [7].

Esse perfil geral de circulação atmosférica encontra variações significativas na mesoescala e na microescala, por diferenças em propriedades de superfícies, tais como geometria e altitude de terreno, vegetação (Figura 2.2) e distribuição de superfícies de terra e água. Esses fatores atuantes nas escalas menores podem resultar em condições de vento locais que se afastam significativamente do perfil geral da larga escala da circulação atmosférica [7].

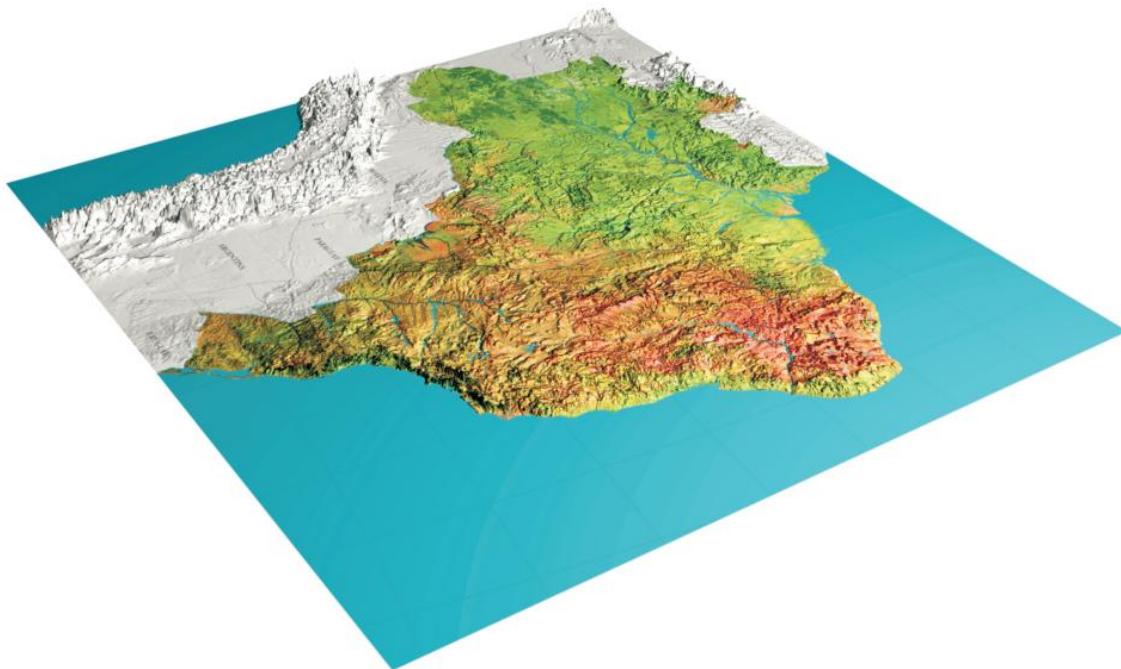


Figura 2.2 – Modelo de relevo e vegetação do Brasil. Fonte: [7].

A Zona Litorânea Norte-Nordeste é definida como a faixa costeira com cerca de 100km de largura, que se estende entre o extremo norte da costa do Amapá e o Cabo de São Roque, no Rio Grande do Norte. Nessa região, os ventos são controlados primariamente pelos alísios de leste e brisas terrestres e marinhas. Essa combinação das

brisas diurnas com os alísios de leste resulta em ventos médios anuais entre 5 m/s e 7,5 m/s na parte norte dessa região (litorais do Amapá e Pará) e entre 6 m/s a 9 m/s em sua parte sul, que abrange os litorais do Maranhão, Piauí, Ceará e Rio Grande do Norte. As velocidades são maiores na parte sul devido a dois principais fatores: (1) os ventos alísios geralmente tornam-se mais fortes à medida que se afastam da Depressão Equatorial; (2) as brisas marinhas são significativamente acentuadas ao sul dessa região em razão dos menores índices de vegetação e de umidade do solo, fazendo que a superfície do solo atinja temperaturas mais elevadas durante as horas de sol e, consequentemente, acentuando o contraste de temperaturas terra-mar e as brisas marinhas resultantes. As maiores velocidades médias anuais de vento ao longo dessa região estão ao norte do Cabo de São Roque, abrangendo os litorais do Rio Grande do Norte e Ceará, onde circulação de brisas marinhas é especialmente intensa e alinhada com os ventos alísios de leste-sudeste. Adicionalmente, ocorrem áreas em que os ventos são acentuados por bloqueios ao escoamento causados por montanhas na parte continental. Entretanto, o vento médio anual decresce rapidamente à medida que se desloca da costa para o interior, devido ao aumento de atrito e rugosidade de superfície e ao enfraquecimento da contribuição das brisas marinhas [7].

As Elevações Nordeste-Sudeste são definidas como as áreas de serras e chapadas que se estendem ao longo da costa brasileira, desde o Rio Grande do Norte até o Rio de Janeiro, a distâncias de até 1.000 km da costa. Velocidades médias anuais de 6,5 m/s até 8m/s devem ser encontradas nos cumes das maiores elevações da Chapada Diamantina e da Serra do Espinhaço. Essas áreas de maiores velocidades ocorrem em forma localizada, primariamente devido ao efeito de compressão vertical do escoamento predominante em larga escala, que é leste-nordeste, quando ultrapassa a barreira elevada das serras. Os ventos anuais mais intensos são geralmente encontrados nas maiores elevações, onde o efeito de compressão é mais acentuado. No entanto, o escoamento atmosférico é bastante complexo nessa região, existindo outras características locais com influência adicional, resultantes de uma combinação de fatores relacionados à topografia e ao terreno [7].

O potencial para a produção de energia elétrica através da fonte eólica, já identificado na Região Nordeste, pode dar uma contribuição significativa ao suprimento de energia elétrica da região, como complementar ou substituto às alternativas hidroelétricas e térmicas. Essa alternativa se tornou uma realidade na perspectiva atual da utilização de até 1.100 MW de geração eólica na região Nordeste, devido ao Programa de Incentivo às

Fontes Alternativas de Energia Elétrica (PROINFA) e à recente autorização de estudos dada pela ANEEL, de aproximadamente 4.800 MW na região [8].

Um dos estudos realizados pelo Centro de Referência para Energia Solar e Eólica – CRESESB/CEPEL, denominado de Atlas do Potencial Eólico Brasileiro [7], estimou um potencial eólico da ordem de 75,0 GW, conforme apresentado na Figura 2.3. Este potencial representa quase a metade de todo potencial estimado no Brasil por este estudo.

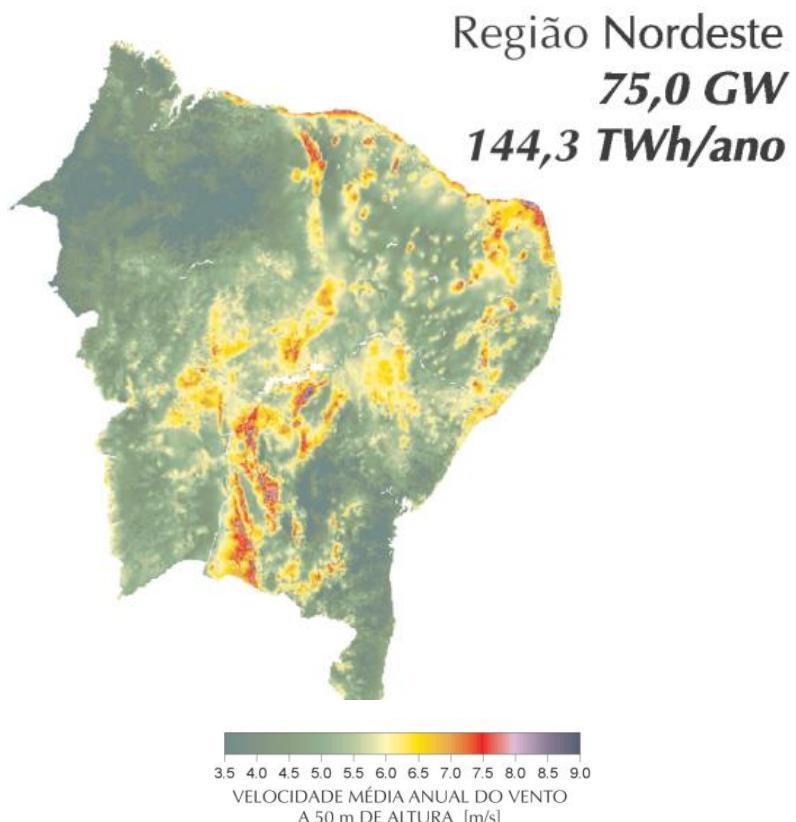


Figura 2.3 – Potencial eólico estimado para a região Nordeste do Brasil. Fonte: [7].

2.2 Geração de Energia Eólio-Elétrica

Um aerogerador é composto basicamente pela turbina eólica, que captura a energia cinética dos ventos e a transforma em energia mecânica em um eixo que está mecanicamente acoplado ao rotor de um gerador elétrico. A turbina é montada no alto de uma torre com o objetivo de aumentar a captação da energia dos ventos. De acordo com a capacidade de geração de energia elétrica desejada, são instalados diversos aerogeradores em um determinado local para a formação de um parque eólico. Obviamente, nos locais em

que a velocidade dos ventos é elevada e relativamente constante, a produção de energia será maior ao longo do ano [9].

Os aerogeradores disponíveis atualmente no mercado podem ser agrupados em dois grupos básicos. O primeiro grupo é composto pelos aerogeradores que operam com velocidade de rotação constante, ou seja, utilizam a filosofia “Dinamarquesa” [10]. Neste caso, o gerador é diretamente acoplado à rede elétrica utilizando somente um *soft-starter* para limitar a corrente durante a etapa de conexão. Já no segundo grupo, os aerogeradores operam com velocidade de rotação variável, ou seja, seus rotores podem girar em qualquer velocidade dentro da faixa admitida. Isto é possível, graças à inclusão de conversores eletrônicos de potência para o acoplamento do gerador elétrico com a rede elétrica, melhorando o rendimento na conversão da energia dos ventos [11].

Em todos os casos, um transformador elevador compatibiliza os níveis de tensão da geração do aerogerador com o nível de tensão da rede de distribuição, a qual normalmente opera entre 13,8 e 34,5 kV.

Um diagrama esquemático dos principais componentes de um aerogerador com filosofia construtiva “Dinamarquesa” é apresentado na Figura 2.4.

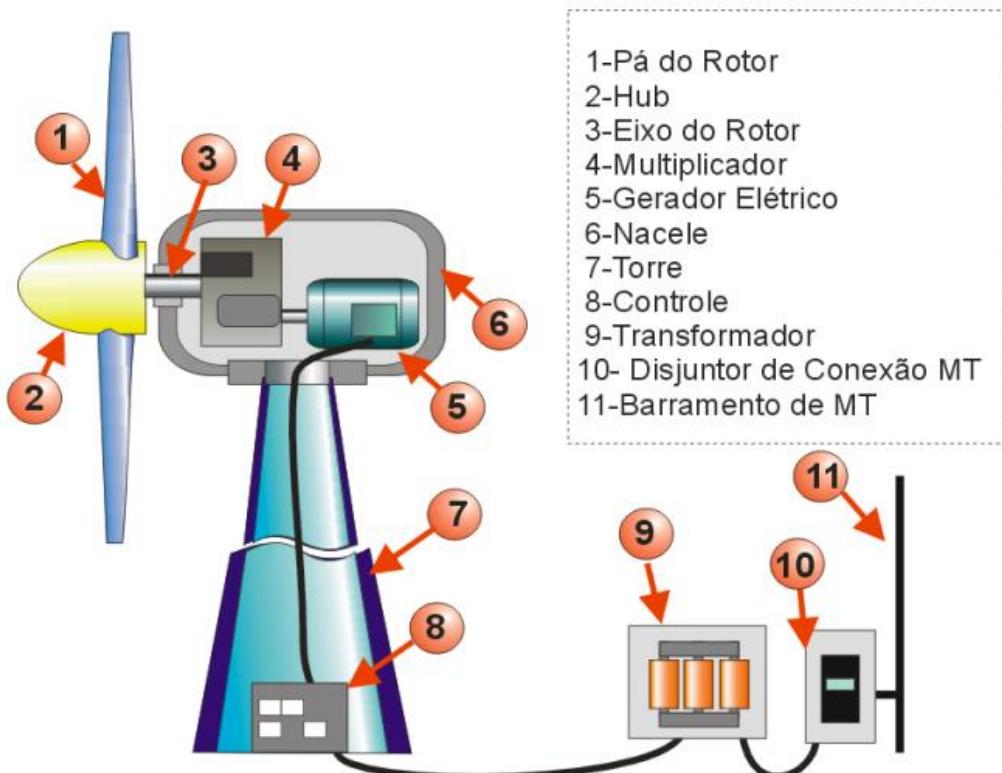


Figura 2.4 – Principais componentes de um aerogerador de velocidade de rotação constante. Fonte: [2].

Na Figura 2.5 apresenta-se um diagrama do esquema de funcionamento de um aerogerador de velocidade de rotação variável.

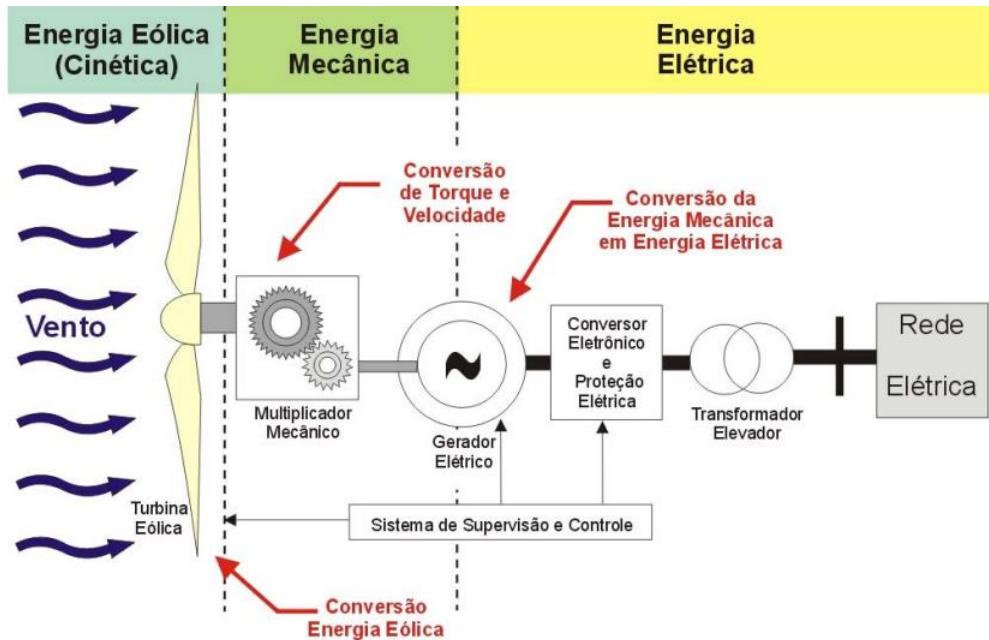


Figura 2.5 – Esquema de funcionamento de um aerogerador de rotação variável. Fonte: [2].

2.2.1 Relação entre a velocidade e a potência dos ventos

A energia cinética (em joules) de uma massa de ar m (kg) se movimentando com uma velocidade v (m/s) é dada pela seguinte equação:

$$E_c = \frac{1}{2}mv^2. \quad (2.1)$$

Sendo A a área varrida pelas pás da trubina (m^2) e ρ a densidade volumétrica do ar (kg/m^3), então a vazão mássica do ar em quilogramas por segundo é ρAv , e a potência mecânica do vento à montante do aerogerador é dada pela seguinte equação (em watts):

$$P_{mec} = \frac{1}{2}\rho Av^3. \quad (2.2)$$

A comparação do potencial de geração de dois ou mais locais candidatos à instalação de um parque eólico pode ser feita através da “potência específica do vento” expressa em watts por metro quadrado de área varrida pelas pás rotativas. Esta potência também é comumente denominada “densidade de potência do local”, e é dada pela seguinte equação [9]:

$$P_{esp} = \frac{1}{2} \rho v^3. \quad (2.3)$$

Observa-se que esta potência varia linearmente com a densidade do ar que varre as pás do aerogerador e com o cubo da velocidade do vento. Na prática, as pás não podem extrair toda a potência disponível no vento à montante, pois parte desta potência continua disponível na massa de ar que flui à jusante do aerogerador com uma velocidade menor.

2.2.2 Potência extraída dos ventos

A potência que realmente é extraída pelas pás do rotor é dada pela diferença entre as potências à montante e jusante do aerogerador. Ela pode ser obtida através da seguinte equação [9]:

$$P_0 = \frac{1}{2} \dot{m} (v^2 - v_0^2), \quad (2.4)$$

em que,

P_0 = potência mecânica no rotor (W);

\dot{m} = vazão mássica (kg/s);

v = velocidade do vento a montante das pás (m/s);

v_0 = velocidade do vento a jusante das pás (m/s).

Sob um ponto de vista macroscópico, a velocidade do ar é discontínua de v para v_0 no plano das pás do rotor, com um valor médio igual à média aritmética dessas velocidades. A vazão mássica poderá ser obtida multiplicando-se esta velocidade média

pela densidade do ar e pela área varrida pelas pás do aerogerador. Portanto, a potência mecânica disponível no eixo do rotor será dada por [9]:

$$P_0 = \frac{1}{2} \left[\rho A \frac{(v + v_0)}{2} \right] (v^2 - v_0^2). \quad (2.5)$$

A Equação (2.5) pode ser reescrita na seguinte forma:

$$P_0 = \frac{1}{2} \rho A v^3 \frac{\left(1 + \frac{v_0}{v}\right) \left[1 - \left(\frac{v_0}{v}\right)^2\right]}{2}. \quad (2.6)$$

A potência (em watts) extraída pelas pás de um aerogerador geralmente é expressa como uma fração da velocidade do vento a montante como segue [9]:

$$P_0 = \frac{1}{2} \rho A v^3 C_p, \quad (2.7)$$

em que,

$$C_p = \frac{\left(1 + \frac{v_0}{v}\right) \left[1 - \left(\frac{v_0}{v}\right)^2\right]}{2}. \quad (2.8)$$

Comparando as Equações (2.2) e (2.7), pode-se dizer que C_p é a fração da potência do vento à montante que é extraída pelas pás do rotor e fornecida ao gerador elétrico. O restante da potência é dissipada no vento à jusante. O fator C_p é denominado *coeficiente de potência do rotor* ou *eficiência do rotor*. Na Figura 2.6 mostra-se que C_p é uma função que possui valor máximo igual a 0,593 quando a razão v_0 / v é igual a um terço.

Na prática, os aerogeradores são projetados de modo que o máximo valor alcançável de C_p varia entre 0,4 e 0,5 para as turbinas modernas de alta velocidade, e entre 0,2 e 0,4 para as turbinas de baixa velocidade. Considerando-se 0,5 como um valor prático para a máxima eficiência do rotor, a máxima potência específica na saída da turbina (em watts por metro quadrado de área varrida pelas pás) será dada por [9]:

$$P_{\max} = \frac{1}{4} \rho v^3. \quad (2.9)$$

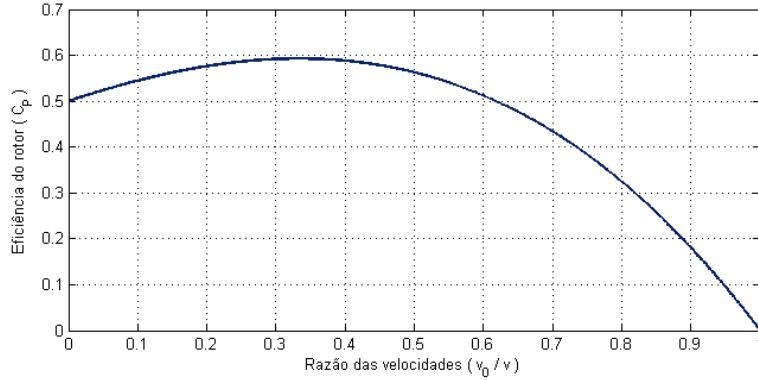


Figura 2.6 – Eficiência do rotor em função da razão das velocidades.

2.2.3 Curva de Potência

Na prática, o rendimento aerodinâmico das pás reduz ainda mais os valores teóricos obtidos para a eficiência do rotor. Para um aerogerador, existem ainda outras perdas, relacionadas com cada componente (rotor, transmissão, caixa multiplicadora e gerador). Além disto, o fato do rotor funcionar em uma faixa limitada de velocidade de vento também irá contribuir para reduzir a energia por ele captada, de acordo com a curva de potência do aerogerador. O coeficiente de potência real de um aerogerador é obtido através do produto da eficiência mecânica, eficiência elétrica e da eficiência aerodinâmica da turbina eólica. Todos estes três fatores dependem da velocidade do vento e da potência gerada [6].

A maneira mais simples de estimar a produção de energia elétrica através da geração eólica é utilizando a curva de potência. Esta curva relaciona a potência ativa fornecida pelo aerogerador com a velocidade do vento na altura do centro do eixo do rotor. Uma curva de potência típica de um aerogerador com potência nominal de 2.000 kW juntamente com a curva do coeficiente de potência são apresentadas na Figura 2.7.

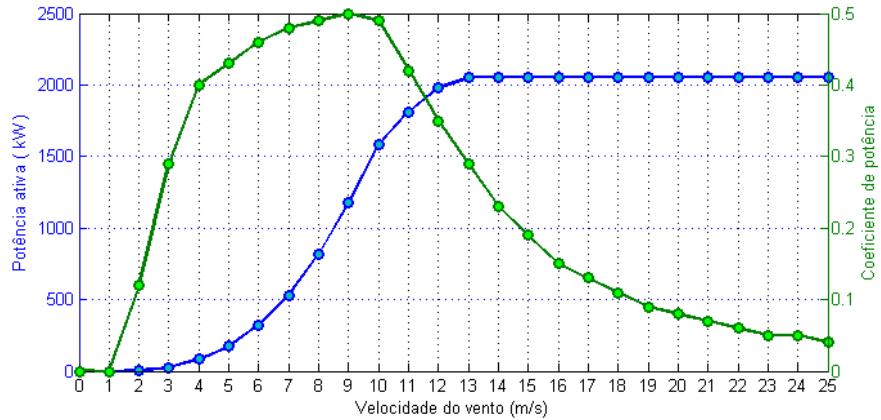


Figura 2.7 – Curva de potência e curva do coeficiente de potência de um aerogerador.

Observa-se, na Figura 2.7, que a potência ativa fornecida pelo gerador é nula para velocidades menores do que 2 m/s. Esta faixa de velocidades é denominada *zona de cut-in*. Para velocidades maiores do que 25 m/s (*zona de cut-out*) há um sistema de segurança que realiza o travamento mecânico da turbina para evitar danos causados pelos esforços mecânicos aos quais o aerogerador ficará submetido. A velocidade para a qual o aerogerador fornece a sua potência nominal é denominada *velocidade nominal*.

2.2.4 Densidade do ar e velocidade do vento em função da altura

A potência extraída dos ventos varia linearmente com a densidade do ar que “varre” as pás do aerogerador. A densidade do ar (ρ) varia com a pressão e a temperatura de acordo com a lei dos gases [9]:

$$\rho = \frac{P}{RT}, \quad (2.10)$$

em que,

P = pressão do ar; R = constante dos gases; T = temperatura na escala absoluta.

A densidade volumétrica do ar ao nível do mar ($P = 1$ atm e $T = 288$ K) é igual a $1,225 \text{ kg/m}^3$. As curvas de potência são obtidas para essas condições, portanto, ao utilizá-las para estimar a geração em uma determinada localidade, deve-se aplicar um fator de correção para levar em consideração as variações da densidade do ar para diferentes alturas e temperaturas.

A temperatura e a pressão variam com a altitude. O efeito combinado destas duas variáveis sobre a densidade do ar é dado pela seguinte equação (que é válida para uma elevação de até 6.000 m acima do nível do mar) [9]:

$$\rho = 1,225 e^{-\left\{ \frac{0,297 H_m}{3048} \right\}}, \quad (2.11)$$

em que, H_m = elevação do local (m).

O atrito entre a superfície terrestre e o vento tem como consequência um retardamento desse último, resultando numa variação de incremento da velocidade média do vento com a altura ao solo. O efeito da força de atrito vai-se desvanecendo até praticamente se anular a uma altura de aproximadamente 2.000 metros [8].

A variação da velocidade do vento depende, basicamente, da temperatura, da rugosidade, da topografia e dos obstáculos do local. De acordo com [9], a variação da velocidade de vento com a altura do solo pode ser expressa pela fórmula:

$$v(h) = v_1 (h/h_1)^\alpha, \quad (2.12)$$

em que:

v_1 – velocidade na altura h_1 (conhecida);

$v(h)$ – velocidade na altura h ;

h – altura para a velocidade $v(h)$;

h_1 – altura da velocidade v_1 ;

α – coeficiente dependente da natureza do terreno.

A Tabela 2.1 apresenta diversos valores do coeficiente de rugosidade (α) com relação ao tipo do terreno.

Tabela 2.1 – Coeficiente de rugosidade para diversos tipos de terreno. Fonte: [9].

TIPO DE TERRENO	COEFICIENTE (α)
Lago, oceano e solo liso	0,10
Grama	0,15
Cercas vivas e arbustos	0,20
Florestas	0,25
Pequenas cidades com poucas árvores e arbustos	0,30
Grandes cidades com altos edifícios e construções	0,40

2.2.5 Distribuição das velocidades dos ventos

A potência elétrica gerada por um aerogerador está relacionada diretamente com a velocidade do vento (elevada ao cubo), portanto, as velocidades são os dados mais críticos necessários para se avaliar o potencial energético de um local candidato. As velocidades e direções dos ventos não são constantes, sendo influenciadas pelo terreno, clima e pela altura em relação à superfície do solo. A velocidade do vento varia a cada minuto, hora, dia, estação, e até mesmo por ano. Com o objetivo de obter resultados mais precisos em relação à velocidade média anual de um determinado parque eólico, devem ser utilizados os dados coletados por um período de dez anos ou mais, desta forma, a avaliação do potencial energético fornecerá resultados mais precisos. No entanto, as medições de longo prazo são caras e a maioria dos projetos não pode esperar tanto tempo assim. Em tais situações, os dados de curto prazo, por exemplo, mais de um ano, são comparados com dados de longo prazo a partir de um local próximo para predizer a velocidade média anual do vento no local em questão [9].

Por ser influenciado pelo sol e pelas estações, o padrão dos ventos normalmente se repete ao longo do período de um ano. As variações da velocidade dos ventos durante o ano podem ser descritas por uma função de distribuição de probabilidades. A função que melhor descreve o comportamento da velocidade dos ventos é a de Weibull (h) com dois parâmetros: o de forma k , e o de escala c . A probabilidade da velocidade ser v durante qualquer intervalo de tempo é dada pela seguinte equação [9]:

$$h(v) = \left(\frac{k}{c}\right) \left(\frac{v}{c}\right)^{(k-1)} e^{-\left(\frac{v}{c}\right)^k}, \text{ para } 0 < v < \infty. \quad (2.13)$$

De acordo com a definição da função de probabilidade, a probabilidade de que a velocidade do vento esteja entre zero e infinito durante o período de tempo analisado é unitária, ou seja [9]:

$$\int_0^{\infty} h(v) dv = 1. \quad (2.14)$$

Como o período de estudo normalmente escolhido é de um ano, a função de distribuição de probabilidades pode ser expressa em função do número de horas no ano, de modo que [9]:

$$h = \frac{\text{número de horas / ano em que a velocidade está entre } v \text{ e } (v + \Delta v)}{\Delta v}. \quad (2.15)$$

A unidade de h é horas/ano por m/s, e a integral (2.14) agora se torna 8.760 (número total de horas em um ano) ao invés da unidade. Na Figura 2.8 ilustra-se o comportamento de h em função de v para três diferentes valores de k na Equação (2.13). A curva azul ($k = 1$) possui uma forte tendência para a esquerda, onde a maioria dos dias são sem vento ($v = 0$). A curva preta ($k = 3$) se parece mais com uma distribuição normal em forma de sino, em que alguns dias possuem altas velocidades de vento e outros possuem baixas velocidades. A curva vermelha ($k = 2$) é uma distribuição de velocidades de ventos típica encontrada na maioria dos sítios. Nesta distribuição, as velocidades são menores do que a média na maioria dos dias, enquanto que em alguns dias as velocidades são elevadas. O valor de k determina a forma da curva e, portanto, é denominado *parâmetro de forma*.

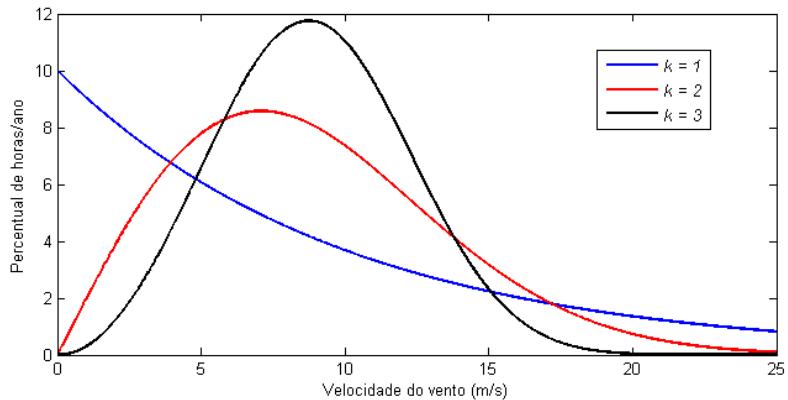


Figura 2.8 – Função distribuição de probabilidades de Weibull com parâmetro de escala $c = 10 \text{ m/s}$ e parâmetros de forma $k = 1, 2$ e 3 .

A distribuição de Weibull com $k = 1$ é denominada *distribuição exponencial*, sendo utilizada geralmente em estudos de confiabilidade. Quando $k = 2$, é denominada *distribuição de Rayleigh*. Para $k > 3$, ela se aproxima da *distribuição normal*, frequentemente denominada *Gaussian*.

As curvas de distribuição correspondentes a $k = 3$ com diferentes valores de c variando entre 4 e 12 metros por segundo são apresentadas na Figura 2.9. Para valores maiores de c , as curvas se deslocam para a direita (para as velocidades maiores). Ou seja, quanto maior o valor de c , maior será o número de dias com velocidades maiores de vento. Devido a esse deslocamento da distribuição de horas para uma escala de velocidades maiores, c é denominado *parâmetro de escala*.

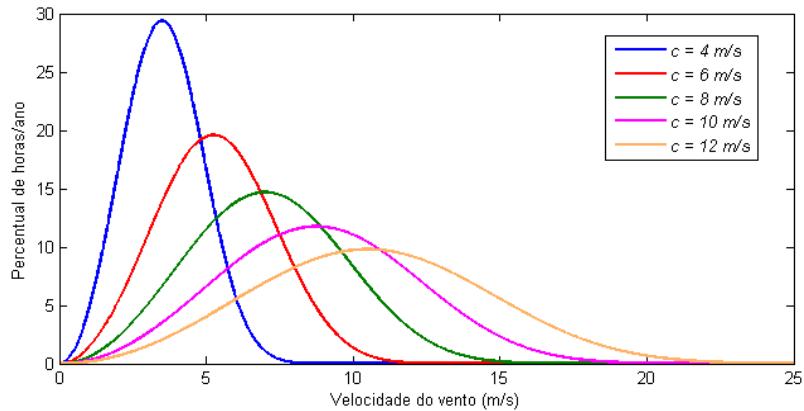


Figura 2.9 – *Função distribuição de probabilidades de Weibull com parâmetro de forma $k = 3$ e parâmetros de escala c variando entre 4 e 12 m/s.*

A velocidade média no período analisado é definida como a área total sob a curva $h - v$ integrada de $v = 0$ a ∞ e dividida pelo número total de horas do período (8.760 se o período for um ano). A velocidade média anual é, portanto, a velocidade média ponderada e é dada por [9]:

$$V_{med} = \frac{1}{8760} \int_0^{\infty} h(v) v dv. \quad (2.16)$$

A maioria das localidades possui parâmetro de escala variando entre 5 e 10 m/s, e parâmetro de forma variando entre 1,5 e 3,0. Para estes valores de c e k , a Equação (2.16) pode ser aproximada por [9]:

$$V_{med} = 0,90c. \quad (2.17)$$

2.2.6 Cálculo da energia elétrica gerada

Conhecendo-se o perfil de distribuição das velocidades dos ventos no local em que serão instalados os aerogeradores e a característica elétrica do aerogerador (curva de potência), pode-se definir a função de distribuição da energia e (MWh/ano/m/s) da seguinte maneira:

$$e(v) = h(v) \cdot P(v) . \quad (2.18)$$

As funções de distribuição das velocidades e da energia gerada são ilustradas na Figura 2.10. Os parâmetros de forma e de escala são dados na altura do eixo do aerogerador e a curva de potência utilizada é igual àquela da Figura 2.7.

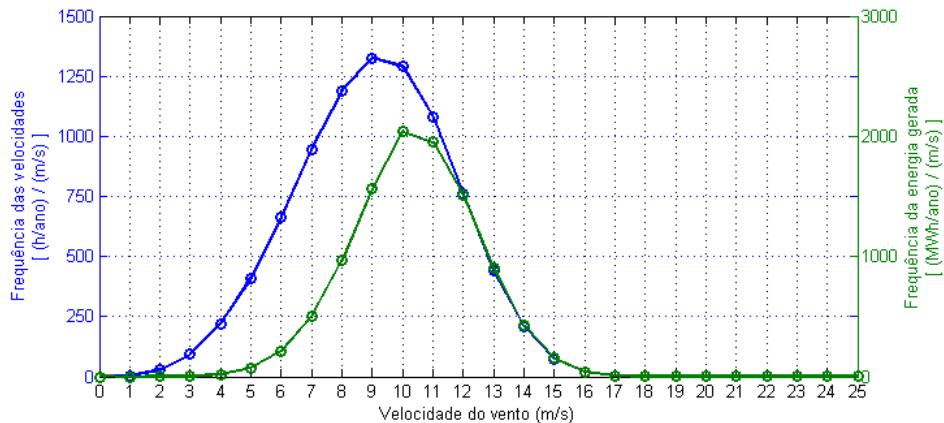


Figura 2.10 – Distribuições das velocidades ($k = 4$ e $c = 10$ m/s) e da energia gerada.

O valor esperado para a geração anual de energia elétrica (em MWh) é dado por:

$$E_a = \int_{v_{in}}^{v_{out}} e(v) dv , \quad (2.19)$$

em que, v_{in} = velocidade de “cut-in”; v_{out} = velocidade de “cut-out”.

A distribuição da velocidade média do vento é normalmente discreta, em classes de 1 m/s, portanto, o cálculo da energia gerada anualmente também pode ser calculado utilizando-se os valores discretos das distribuições, ou seja:

$$E_a = \sum_{v_{in}}^{v_{out}} e(v) . \quad (2.20)$$

O valor da energia anual calculado por (2.19) ou (2.20) considera a distribuição de ventos do local, assim como as potências calculadas para as respectivas velocidades do vento para a máquina em questão (através da curva de potência). Contudo, este cálculo da energia não considera as perdas na rede de distribuição do parque eólico, que provocarão um decréscimo na energia que será injetada na rede da concessionária.

2.3 Revisão Bibliográfica

Parte desta seção é baseada na referência [12], que fornece uma visão geral sobre os modelos de previsão de energia eólica com horizontes que variam desde alguns minutos até alguns dias à frente, para um único aerogerador e também para parques inteiros. Uma série de pesquisas em diversos artigos e periódicos foi realizada com o intuito de definir os modelos de previsão de ventos e geração eólica a serem utilizados no presente trabalho. Um resumo da literatura pesquisada será apresentado a seguir. Todos os modelos citados e propostos nesta dissertação terão como foco as previsões de curto prazo, com passos de previsão variando entre uma e vinte e quatro horas. Uma breve introdução para previsões de curto prazo também pode ser encontrada em [13].

Um dos maiores problemas da energia eólica, em comparação a eletricidade convencional, é sua dependência da volatilidade do vento. Esta dependência acontece em todas as escalas de tempo, mas duas delas são mais relevantes: uma é importante para o controle da turbina propriamente dita (de milissegundos a segundos), e a outra é importante para a integração da energia eólica na rede elétrica, sendo determinada pelas constantes de tempo da rede (de minutos a semanas).

Em geral, os modelos podem ou não envolver um modelo de previsão numérica do tempo (*Numerical Weather Prediction* – NWP). Normalmente, os modelos que utilizam NWP fornecem melhores previsões de séries temporais para horizontes maiores do que cerca de algumas horas (a partir de 3 – 6 horas), o que os fazem ser utilizados pelas concessionárias.

Existem dois tipos de modelagem diferentes em relação à previsão de curto prazo: a modelagem física e a estatística. Em alguns modelos, uma combinação das duas

modelagens é utilizada de modo a realizar previsões mais confiáveis. Em suma, com os modelos físicos se tenta utilizar variáveis físicas o maior tempo possível para chegar à melhor estimativa da velocidade do vento local antes de utilizar um modelo estatístico (*Model Output Statistics* – MOS) para reduzir o erro remanescente. A modelagem estatística tenta encontrar fortes relações entre os valores históricos da produção de energia elétrica (e de outros parâmetros meteorológicos) e as informações medidas em tempo real, recorrendo normalmente a técnicas recursivas.

Na implementação dos modelos estatísticos utilizam-se normalmente modelos do tipo “caixa preta”, por exemplo, Redes Neurais Artificiais (RNA). Alguns deles podem ser expressos analiticamente, outros não (é o caso das redes neurais).

Comparações utilizando modelos autoregressivos entre a previsão direta da energia eólica em relação às previsões da velocidade do vento, com subsequente conversão para energia eólica [14,15], demonstraram que o uso das previsões da velocidade do vento como variável explicativa é importante para horizontes de previsão de até 8–12 horas. Para horizontes maiores, o uso de previsões de velocidade como variável explicativa não oferece nenhuma vantagem em relação à previsão direta da energia eólica.

Em [16], mostra-se que ao utilizar modelos NWP é melhor aplicar um tratamento estatístico sobre as velocidades de ventos previstas do que sobre a potência final de saída.

Em [17], verificou-se melhorias na raiz do erro quadrático médio (*Root Mean Squared Error* – RMSE) para passos de previsão entre 1 e 10 minutos. As melhorias situam-se em torno de 10% sobre o *Modelo da Persistência*. Esta melhoria foi conseguida com uma topologia bastante simples, embora com estruturas mais complexas não houve melhoria significativa dos resultados. Uma limitação foi encontrada em eventos extremos que não estavam contidos no conjunto de dados usados para treinar a rede neural.

As diferenças entre as velocidades de ventos e as médias móveis foram utilizadas como dados de entrada em [18]. Para a mesma série temporal foram verificadas melhorias de até 13% em relação ao *Modelo da Persistência*, enquanto que a abordagem padrão de redes neurais obteve 9,5% de melhoria.

Em [19], foram utilizadas redes neurais e o modelo ARIMA (*Autoregressive Integrated Moving Average* – ARIMA) para a previsão das séries de velocidades dos ventos no Reino Unido e Grécia no horizonte de uma hora. Não foram obtidas melhorias significativas em relação ao *Modelo da Persistência* para ambas as localidades ao utilizar

as médias horárias das velocidades. Porém, ao utilizar as velocidades médias de intervalos de dez minutos, a melhoria foi de 10 – 20%.

Em [20], demonstrou-se que há melhorias ao aplicar a Transformada *Wavelet* (*Wavelet Transform*) às velocidades médias horárias antes de fornecê-las como entradas às redes neurais. Os horizontes de previsão utilizados foram iguais a 4 e 24 horas.

O artigo [21] sugere um protocolo padronizado para a avaliação dos sistemas de previsão de geração eólica de curto prazo. Também descreve alguns modelos de referência para a previsão, e argumenta-se que o uso do *Modelo da Persistência* como referência leva a conclusões ligeiramente equivocadas e mais otimistas sobre o desempenho dos modelos avaliados.

Em [22], um modelo *fuzzy* é sugerido para realizar a previsão da velocidade do vento e da energia elétrica produzida em um parque eólico. O modelo foi treinado usando um esquema de aprendizado baseado em algoritmos genéticos. O conjunto de treinamento incluia a velocidade do vento e dados de direção, medidos em locais vizinhos com até 30 km de distância dos grupos de aerogeradores. Foram apresentados os resultados das previsões entre 30 minutos e 4 horas. O modelo sugerido apresentou uma melhora média da ordem de 15 – 20% em comparação ao *Modelo da Persistência*.

Em [23], foi proposto um modelo híbrido que utiliza a Transformada *Wavelet*, *Particle Swarm Optimization* (PSO) e Lógica *Fuzzy*, para realizar a previsão da geração com horizontes de até 24 horas. As previsões foram realizadas para o ano de 2009 e os resultados foram comparados com outros 7 modelos (ARIMA, Redes Neurais, Redes Neurais + Lógica *Fuzzy*, *Persistência*, etc.). O modelo proposto foi o que obteve o melhor desempenho dentre todos os modelos comparados.

CAPÍTULO 3

3. INTELIGÊNCIA ARTIFICIAL E WAVELETS

Este capítulo apresenta os conceitos básicos necessários para o entendimento dos modelos de previsão que serão descritos no Capítulo 5. As duas primeiras seções do presente capítulo apresentam, as Redes Neurais Artificiais (RNA) e a Lógica *Fuzzy*, respectivamente. Na terceira seção, são apresentados os conceitos da análise de sinais com base nas *wavelets*. A quarta e última seção fala do *software* utilizado para o desenvolvimento dos modelos e apresenta os parâmetros necessários para realizar os ajustes dos mesmos.

3.1 Redes Neurais Artificiais

Uma Rede Neural Artificial é uma ferramenta computacional cuja estrutura é projetada com o objetivo de simular a maneira pela qual o cérebro humano funciona. Pelo fato de possuir analogia neurobiológica como fonte de inspiração, uma rede neural é tratada como uma “Ferramenta de Inteligência Artificial”.

As RNA se constituem em uma técnica de inteligência artificial cuja utilização prática está se tornando cada vez mais presente no nosso dia-a-dia. Aplicações realizadas com RNA têm apresentado desempenho satisfatório em diversas áreas de pesquisas, tais como: classificação, reconhecimento de padrões, aproximação de funções, processamento de séries temporais, otimização, etc [6].

Atualmente, as redes neurais apresentam-se como uma abordagem alternativa aos métodos estatísticos de previsão de séries temporais. O emprego desta técnica é atrativo em uma grande variedade de problemas que envolvem relacionamentos complexos entre as variáveis de entrada e de saída, uma vez que para utilização não se faz necessário o conhecimento prévio das relações matemáticas entre estas variáveis [6].

As RNA são sistemas paralelos distribuídos compostos por unidades de processamento, chamados de neurônios artificiais, que calculam determinadas funções matemáticas (normalmente não lineares). Esta forma de computação não-algorítmica é caracterizada por sistemas que, relembram a estrutura do cérebro humano. O grande apelo destes modelos está em sua capacidade de “aprender”, generalizar ou extrair regras automaticamente de conjuntos de dados complexos [24].

3.1.1 O Neurônio Artificial

Um neurônio artificial é uma unidade de processamento da informação que é fundamental para a operação de uma rede neural. O modelo de um neurônio é apresentado na Figura 3.1, nela se identificam três partes básicas: um conjunto de sinapses (ou elos de conexões); um somador de sinais; e uma função de ativação.

Cada sinapse é caracterizada por um peso, que pode possuir valores positivos ou negativos. O sinal de entrada de uma determinada sinapse é multiplicado pelo peso sináptico. O somador realiza a soma dos sinais das entradas, ponderados pelas respectivas sinapses. A função de ativação serve para restringir a amplitude do sinal de saída do neurônio artificial. Tipicamente, a amplitude do sinal de saída está restrita ao intervalo $[0,1]$.

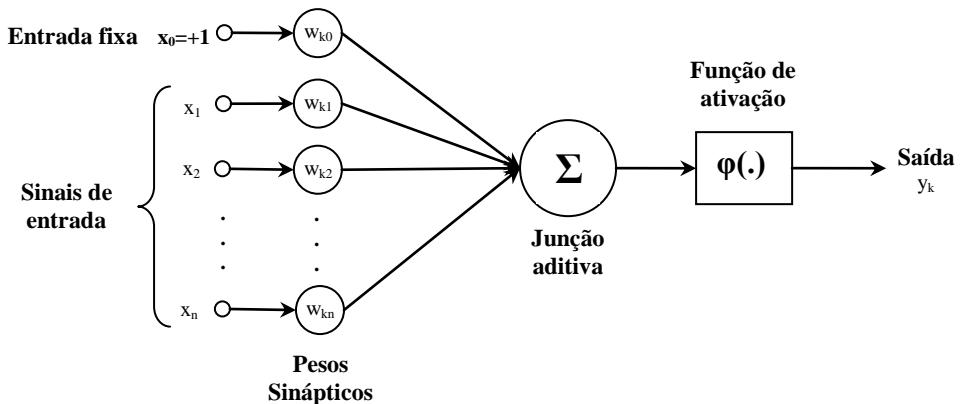


Figura 3.1 – Modelo de um neurônio artificial. Fonte:[6].

O modelo do neurônio artificial inclui também uma entrada fixa. O peso sináptico desta entrada fixa é chamado de *bias*. O *bias* tem o efeito de aumentar ou diminuir o valor do sinal de entrada da função de ativação.

Em termos matemáticos, um neurônio artificial k pode ser descrito através das seguintes equações:

$$v_k = \sum_{j=1}^n w_{kj} x_j + w_{k0}, \quad (3.1)$$

$$y_k = \varphi(v_k). \quad (3.2)$$

3.1.2 Funções de Ativação

Conforme visto na seção anterior, a função de ativação (função de transferência) define o sinal de saída do neurônio k em função do potencial de ativação v_k . Dentre as diversas funções de ativação conhecidas, são apresentadas nesta seção duas funções que serão utilizadas para o desenvolvimento dos modelos propostos no Capítulo 5 desta dissertação.

A primeira função a ser descrita é denominada Sigmóide Logística. Este tipo de função de transferência é o mais utilizado na construção de redes neurais. A Figura 3.2 é o gráfico deste tipo de função. O gráfico se assemelha a uma curva em “S”.

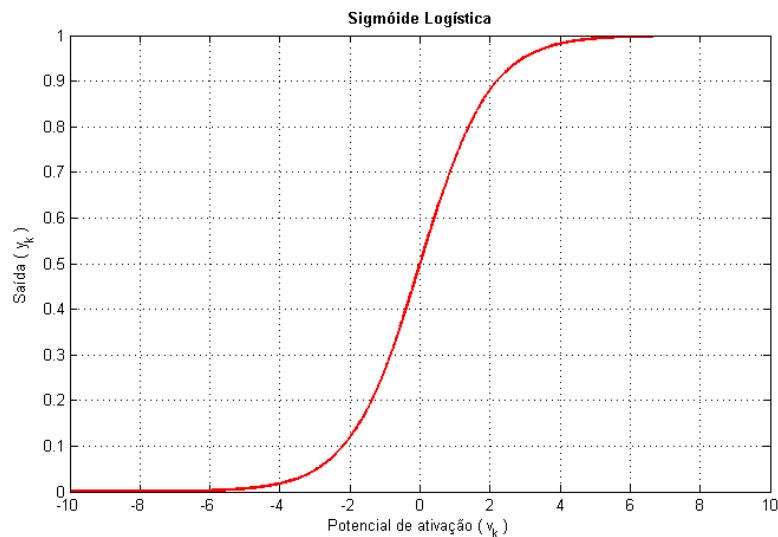


Figura 3.2 – Gráfico da função Sigmóide Logística.

Os valores de saída desta função pertencem ao intervalo $[0,1]$. Um outro aspecto interessante é que a função é diferenciável em todos os pontos do seu domínio. Ela tem a seguinte expressão:

$$\varphi(v_k) = \frac{1}{1 + \exp(-v_k)}. \quad (3.3)$$

Em alguns casos, é desejável que o sinal de saída da função de ativação do neurônio artificial seja definida no intervalo $[-1,1]$. Para permitir que a função de ativação assuma

valores negativos, utiliza-se a função sigmoidal Tangente Hiperbólica. Na Figura 3.3 apresenta-se um gráfico deste tipo de função.

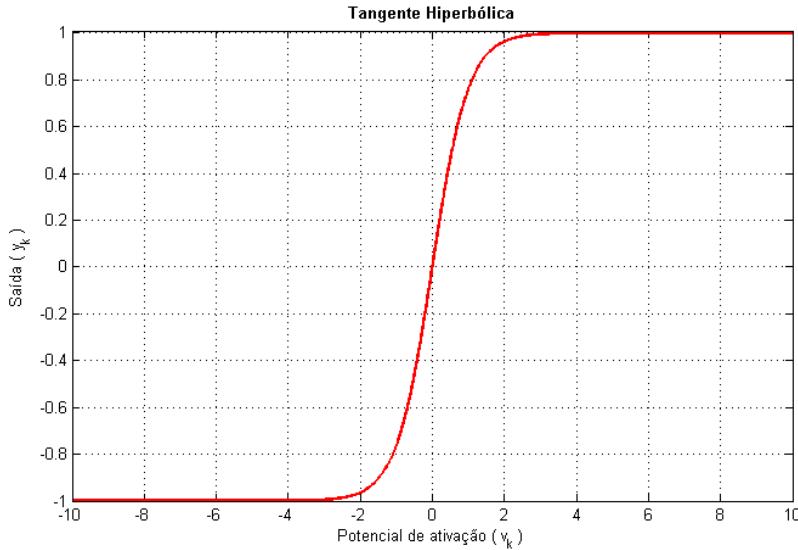


Figura 3.3 – Gráfico da função Tangente Hiperbólica.

Os valores de saída desta função pertencem ao intervalo $[-1,1]$ e ela também é diferenciável em todos os pontos do seu domínio, pois sua expressão é a seguinte:

$$\varphi(v_k) = \frac{2}{1 + \exp(-2v_k)} - 1. \quad (3.4)$$

3.1.3 Arquitetura da Rede Neural Artificial

A estrutura (arquitetura) de uma rede neural está intimamente relacionada ao algoritmo de aprendizagem utilizado para treiná-la. A arquitetura das redes neurais utilizadas para o desenvolvimento dos modelos de previsão do presente trabalho é do tipo *Multilayer Feedforward*, ou seja, redes progressivas de múltiplas camadas.

Em uma rede progressiva (*feedforward*), o sentido das conexões é sempre voltado para a camada de saída, ou seja, não há elos de realimentação de sinais entre as camadas. As redes progressivas de múltiplas camadas possuem, tipicamente, uma ou mais camadas intermediárias (ocultas) entre as camadas de entrada e saída. A adição de camadas intermediárias permite que a rede possa extrair dos sinais de entrada estatísticas de ordem superior.

O esquema de uma rede neural progressiva de três camadas é apresentada na Figura 3.4. A rede possui quatro nós na camada de entrada, seis neurônios na camada oculta e um neurônio na camada de saída. Esta rede é dita completamente conectada, pois cada nó da rede é conectado a todos os outros nós das camadas adjacentes. Esta é a arquitetura das redes neurais utilizadas nos modelos de previsão propostos no Capítulo 5.

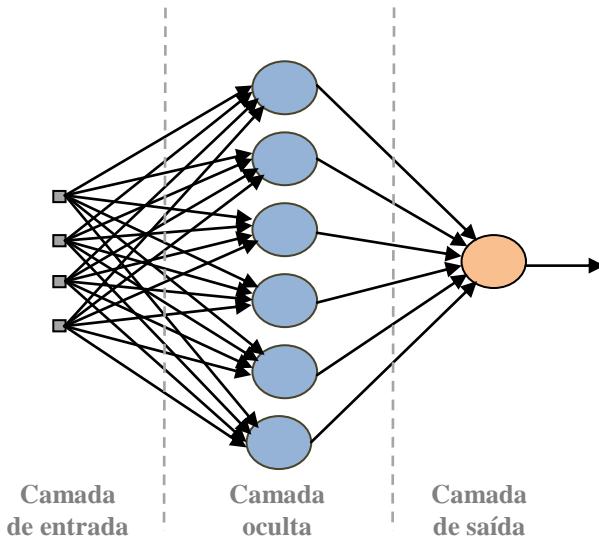


Figura 3.4 – Esquema de uma rede progressiva de três camadas.

Os nós da camada de entrada, também chamados de nós fonte, fornecem aos neurônios da camada oculta os sinais aplicados à entrada da rede neural. Os sinais das saídas dos neurônios da camada oculta são fornecidos às entradas do neurônio da camada de saída. O sinal de saída do neurônio desta última camada é a resposta global da rede ao padrão de ativação provido pelos nós da camada de entrada.

3.1.4 Processo de Aprendizagem

As redes neurais possuem a capacidade de adquirir conhecimento mediante um processo de aprendizagem. O conhecimento adquirido é armazenado nos parâmetros livres da rede, que são os pesos sinápticos, os *bias* e os parâmetros que definem as funções de transferência dos neurônios artificiais.

O procedimento utilizado para o processo de aprendizagem é chamado Algoritmo de Aprendizagem. A função deste algoritmo é modificar de forma adaptativa os parâmetros livres da rede para atingir um objetivo desejado. Em outras palavras, o processo de aprendizagem de uma rede *Multilayer Feedforward* é um problema de otimização não-

linear irrestrita cujo objetivo é minimizar a soma quadrática das diferenças entre as saídas desejadas e a resposta da rede.

3.1.5 Algoritmo *Backpropagation*

O algoritmo *backpropagation* é o algoritmo de aprendizagem mais utilizado para o treinamento das redes progressivas de múltiplas camadas. Sua popularidade é atribuída à relativa simplicidade de implementação, e também ao fato de ser um poderoso dispositivo para armazenar o conteúdo da informação através do ajuste dos pesos sinápticos da rede. Quando o conjunto de dados utilizado para treinar uma rede *Multilayer Feedforward* é grande o suficiente para ser representativo, o algoritmo *backpropagation* fornece à rede a capacidade de generalização.

O algoritmo *backpropagation* é composto por duas fases de treinamento. A primeira delas é a fase *forward*, na qual um padrão de dados é apresentado à rede e esta processa os dados, produzindo os sinais de saída (resposta). A segunda delas é a fase *backward*, que utiliza o erro obtido entre a resposta da fase *forward* e o resultado desejado (conhecido), para determinar os ajustes a serem feitos nos pesos das conexões sinápticas dos neurônios da rede. O algoritmo de retropropagação é apresentado a seguir.

Seja o sinal de erro na saída do neurônio j , na iteração k , definido pela Equação (3.5), em que, $d_j(k)$ é a saída desejada e $y_j(k)$ é a resposta apresentada na saída do neurônio.

$$e_j(k) = d_j(k) - y_j(k). \quad (3.5)$$

A função custo ε é a energia total do erro, obtida somando o quadrado dos erros de todos os neurônios da camada de saída. A função custo é dada pela seguinte equação:

$$\varepsilon(k) = \frac{1}{2} \sum_{i=1}^n e_i^2(k), \quad (3.6)$$

em que, n é o número de neurônios da camada de saída da rede e e_i é o erro do neurônio i , na iteração k .

O cálculo dos deltas para a correção dos pesos depende da posição da camada. Considere o neurônio da camada de saída s , sendo estimulado por um conjunto de

ativações produzido por uma camada com m neurônios à sua esquerda. O potencial de ativação do neurônio j é, portanto:

$$v_j(k) = \sum_{l=0}^m w_{jl}(k) y_l(k). \quad (3.7)$$

A função de ativação do neurônio da camada de saída define a saída do neurônio j na iteração k de acordo com a seguinte equação:

$$y_j(k) = \varphi_j(v_j(k)). \quad (3.8)$$

O algoritmo *backpropagation* aplica uma correção Δw_{jl} ao peso sináptico w_{jl} , que é proporcional à derivada parcial $\partial \varepsilon(k) / \partial w_{jl}(k)$. De acordo com a regra da cadeia (do cálculo diferencial), este gradiente é expresso como:

$$\frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} = \frac{\partial \varepsilon(k)}{\partial e_j(k)} \frac{\partial e_j(k)}{\partial y_j(k)} \frac{\partial y_j(k)}{\partial v_j(k)} \frac{\partial v_j(k)}{\partial w_{jl}(k)}. \quad (3.9)$$

Após manipulações algébricas, pode-se expressar o gradiente como:

$$\frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} = -e_j(k) \varphi_j'(v_j(k)) y_l(k). \quad (3.10)$$

A correção $\Delta w_{jl}(k)$ aplicada a $w_{jl}(k)$ é definida pela regra:

$$\Delta w_{jl}(k) = -\eta \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)}, \quad (3.11)$$

em que, η é a taxa de aprendizagem do algoritmo *backpropagation*. O uso do sinal negativo na Equação (3.11) indica a descida do gradiente no espaço de pesos, ou seja, busca uma direção para a mudança de pesos que reduz o valor de $\varepsilon(k)$.

Finalmente, a atualização dos pesos da rede é realizada de acordo com a seguinte equação:

$$w_{jl}(k+1) = w_{jl}(k) - \eta \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)}, \quad (3.12)$$

em que, $w_{jl}(k+1)$ é o valor atualizado do peso da conexão j do neurônio l .

3.1.6 Algoritmo *Resilient Propagation*

As redes multicamadas normalmente utilizam funções de ativação do tipo sigmoidal. Um problema que pode surgir durante o treinamento de uma rede *Multilayer feedforward* com funções sigmóides é que o gradiente pode ter valor muito pequeno e, consequentemente, as alterações nos parâmetros também serão muito pequenas, mesmo que estes parâmetros estejam longe de seus valores ótimos.

O objetivo do algoritmo de treinamento *Resilient Propagation*, proposto em [25], é eliminar os efeitos indesejados causados pelos valores das derivadas parciais. Neste algoritmo, apenas o sinal da derivada é utilizado para determinar a direção da atualização dos parâmetros, a magnitude da derivada não tem efeito sobre a atualização.

No processo de otimização do algoritmo *Resilient Propagation*, cada parâmetro é alterado individualmente através de um valor de atualização δ_{jl} . Este valor é definido conforme a seguinte equação:

$$\delta_{jl}(k) = \begin{cases} \alpha^+ \delta_{jl}(k-1), & \text{se } \frac{\partial \varepsilon(k-1)}{\partial w_{jl}(k-1)} \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} > 0 \\ \alpha^- \delta_{jl}(k-1), & \text{se } \frac{\partial \varepsilon(k-1)}{\partial w_{jl}(k-1)} \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} < 0 \\ \delta_{jl}(k-1), & \text{se } \frac{\partial \varepsilon(k-1)}{\partial w_{jl}(k-1)} \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} = 0 \end{cases}, \quad (3.13)$$

em que, $0 < \alpha^- < 1 < \alpha^+$.

A regra de adaptação do valor de atualização é a seguinte: sempre que a derivada parcial em relação ao peso correspondente w_{jl} muda de sinal, indicando que a última atualização foi muito grande e o algoritmo “pulou” um mínimo local, o valor de atualização δ_{jl} é reduzido por um fator α^- . Se a derivada mantém o sinal, o valor de

atualização é ligeiramente aumentado por um fator α^+ com o objetivo de acelerar o processo de convergência.

Após ser realizada a adaptação de δ_{jl} , a atualização dos pesos é realizada de acordo com uma regra bastante simples: se a derivada é positiva (indicando um aumento do erro), o respectivo peso é reduzido pelo seu valor de atualização, caso contrário, o peso será aumentado pelo seu valor de atualização (Equação (3.14)).

$$\Delta w_{jl}(k) = \begin{cases} -\delta_{jl}(k), & \text{se } \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} > 0 \\ +\delta_{jl}(k), & \text{se } \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} < 0 \\ 0, & \text{se } \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} = 0 \end{cases} \quad (3.14)$$

A atualização de cada parâmetro da rede é realizada de acordo com a seguinte equação:

$$w_{jl}(k+1) = w_{jl}(k) + \Delta w_{jl}(k). \quad (3.15)$$

Entretanto, há uma exceção: se a derivada parcial muda de sinal, ou seja, o passo anterior foi muito grande e o mínimo foi “perdido”, a atualização anterior do peso correspondente é revertida. A seguinte equação ilustra este procedimento:

$$\Delta w_{jl}(k) = -\Delta w_{jl}(k-1), \quad \text{se } \frac{\partial \varepsilon(k-1)}{\partial w_{jl}(k-1)} \frac{\partial \varepsilon(k)}{\partial w_{jl}(k)} < 0. \quad (3.16)$$

Com esta reversão da atualização, pode ser que a derivada mude de sinal novamente na próxima iteração. Para evitar uma “dupla punição do valor de atualização”, neste caso, não deve ser realizada a adaptação de δ_{jl} nesta última iteração. Na prática, isto pode ser feito atribuindo-se, na regra adaptativa de δ_{jl} , o valor zero à derivada do passo anterior.

Os valores de atualização e os pesos são modificados após a apresentação de todo o conjunto de padrões à rede, ou seja, após cada “época” de treinamento.

3.1.7 Algoritmo *Levenberg-Marquardt*

O algoritmo *Levenberg-Marquardt* é uma técnica iterativa de otimização utilizada para minimizar funções expressas como somas quadráticas de funções não-lineares. Este algoritmo pode ser visto como uma combinação do método do Gradiente Descendente e do método de Gauss-Newton. Quando a solução corrente se encontra distante da solução ótima, o algoritmo *Levenberg-Marquardt* se comporta como o método do Gradiente Descendente: lento, mas com convergência garantida. Quando a solução corrente está próxima à ótima, o algoritmo *Levenberg-Marquardt* se comporta como o método de Gauss-Newton.

Quando a forma da função objetivo é uma soma quadrática (como no caso do treinamento de redes progressivas de múltiplas camadas), então, a matriz Hessiana pode ser aproximada por

$$H = J^T J, \quad (3.17)$$

e o gradiente pode ser calculado como

$$g = J^T r, \quad (3.18)$$

em que, J é a matriz Jacobiana que contém as derivadas de primeira ordem dos erros da rede em relação aos parâmetros, e r é o vetor dos resíduos da rede (erros). A matriz Jacobiana pode ser calculada através de uma técnica padrão de retropropagação, que é bem mais simples do que calcular a matriz Hessiana [26].

O algoritmo *Levenberg-Marquardt* utiliza a aproximação da matriz Hessiana para realizar a atualização dos parâmetros da rede. A atualização é realizada de maneira similar ao método de Gauss-Newton, de acordo com a seguinte equação:

$$w_{k+1} = w_k - [J^T J + \mu I]^{-1} J^T r. \quad (3.19)$$

Quando o escalar μ é igual a zero, trata-se do método de Gauss-Newton. Quando μ é um valor grande, ele se torna o método do Gradiente Descendente com um passo pequeno. O método de Newton é mais rápido e preciso perto da solução ótima, portanto, o

objetivo do algoritmo *Levenberg-Marquardt* é se comportar como o método de Newton o mais rápido possível. Assim, μ é reduzido a cada iteração bem sucedida (diminuição da função objetivo) e só é aumentado se o valor da função objetivo para a próxima iteração aumentar.

O grande incoveniente do algoritmo *Levenberg-Marquardt* é que ele requer o armazenamento de algumas matrizes que podem ser muito grandes para determinados problemas. O tamanho da matriz Jacobiana é $Q \times n$, sendo Q o número de conjuntos de treinamento e n o número de parâmetros da rede. Porém, essa matriz não tem que ser calculada e armazenada como um todo. Por exemplo, pode-se dividir a matriz Jacobiana em duas submatrizes e posteriormente calcular a matriz Hessiana aproximada da seguinte forma:

$$H = J^T J = \begin{bmatrix} J_1^T & J_2^T \end{bmatrix} \begin{bmatrix} J_1 \\ J_2 \end{bmatrix} = J_1^T J_1 + J_2^T J_2. \quad (3.20)$$

Observe que não é necessário calcular a matriz Jacobiana completa de uma só vez. Pode-se calcular a aproximação da matriz Hessiana através de uma série de somas de submatrizes.

3.2 Lógica *Fuzzy*

A Lógica *Fuzzy*, ou lógica nebulosa, é baseada na teoria dos Conjuntos *Fuzzy*. Esta é uma generalização da teoria dos Conjuntos Tradicionais para resolver os paradoxos gerados à partir da classificação “verdadeiro ou falso” da Lógica Clássica. Tradicionalmente, uma proposição lógica tem dois extremos: ou “completamente verdadeiro” ou “completamente falso”. Entretanto, na Lógica *Fuzzy*, uma premissa varia em grau de verdade de 0 a 1, o que leva a ser parcialmente verdadeira ou parcialmente falsa.

A força da Lógica *Fuzzy* deriva da sua habilidade em inferir conclusões e gerar respostas baseadas em informações vagas, ambíguas e qualitativamente incompletas e imprecisas. Neste aspecto, os sistemas nebulosos têm habilidade de raciocinar de forma semelhante à dos humanos. Seu comportamento é representado de maneira muito simples e natural, levando à construção de sistemas comprehensíveis e de fácil manutenção.

Com a incorporação do conceito de “grau de verdade”, a teoria dos Conjuntos *Fuzzy* estende a teoria dos Conjuntos Tradicionais. Os grupos são rotulados

qualitativamente (usando termos linguísticos, tais como: quente, frio, grande, pequeno, etc.) e os elementos deste conjuntos são caracterizados variando-se o grau de pertinência (valor que indica o grau em que um elemento pertence a um conjunto). Por exemplo, um homem de 2,10 metros e um homem de 2,00 metros são membros do conjunto “alto”, embora o homem de 2,10 metros tenha um grau de pertinência maior neste conjunto.

A composição básica de um sistema *fuzzy* é formada pelos componentes descritos a seguir:

- I. *Fuzzificador* – Aplica a função de pertinência a um valor de entrada (valor real) e a saída será uma valor entre 0 e 1. Cada função de pertinência para uma dada variável de entrada é conhecida como uma variável lingüística;
- II. *Regras* – Um conjunto de regras do tipo SE-ENTÃO que são criadas pelos especialistas ou extraídas de dados numéricos;
- III. *Inferência* – É um sistema que mapeia conjuntos *fuzzy* de entradas em conjuntos *fuzzy* de saída, determinando como as regras são ativadas e combinadas. Os modelos de inferência mais conhecidos são: Modelo de Mamdani e Modelo de Takagi-Sugeno [27];
- IV. *Defuzzificador* – Quando se utiliza um sistema *fuzzy* o objetivo é encontrar um valor de saída real para o problema. Nesta etapa, após aplicação do sistema de inferência, transforma-se o valor de saída nebuloso em uma saída real.

3.2.1 Sistema de inferência *Fuzzy*

A estrutura básica de um sistema de inferência *fuzzy* (*Fuzzy Inference System* – FIS) consiste em três componentes conceituais: a base de regras, a qual contém a seleção de regras *fuzzy*, a base de dados, a qual define a função membro utilizada nas regras, e o mecanismo de raciocínio, o qual realiza o procedimento de inferência sobre as regras e fornece condições para derivar uma saída razoável ou uma conclusão [28].

- **Sistema de inferência *Mandani***

O modelo de inferência Mamdani foi um dos primeiros sistemas constituídos utilizando a teoria de conjuntos *fuzzy*, sendo proposto em 1975 por Ebrahim Mandani [29]. A regra de semântica tradicionalmente utilizada para o processamento de inferências com o modelo de Mamdani é chamada de inferência Máx-Mín, utilizando as operações de união e

de interseção entre conjuntos da mesma forma de Zadeh [29]. As regras de produção em um modelo de Mamdani possuem relações *fuzzy* tanto em seus antecedentes como em seus consequentes. O modelo Mandani é descrito a seguir.

Seja um sistema *fuzzy* composto de n regras, com uma das regras do tipo **Se** $x_1 = A_1$ **e** $x_2 = A_2$ **e** ... **e** $x_j = A_j$ **então** $y = B_i$, em que, x_k são as entradas do sistema, $A_1 \dots A_j$ são variáveis linguísticas definidas pelas funções de pertinência de entrada, y é a saída e B_i são as variáveis lingüísticas definidas pelas funções de pertinência de saída. O processo de inferência pode ser dividido em cinco etapas:

1^a etapa: Fuzzificação das Entradas

Nesta etapa, toma-se o valor de cada variável de entrada e determina-se o seu grau de pertinência para cada uma das regras, ou seja:

$$\mu_{A_1}^k(x_1), \mu_{A_2}^k(x_2), \dots, \mu_{A_j}^k(x_j), \quad k = 1, \dots, n. \quad (3.21)$$

2^a etapa: Aplicação do Operador *fuzzy*

Com as entradas fuzzificadas, sabe-se com qual grau cada parte do antecedente satisfaz cada regra. Precisa-se gerar o coeficiente de disparo de cada regra $D^{(k)}$, para isso, aplica-se o operador *fuzzy* presente no antecedente, por exemplo, o operador “e”.

Para aplicar o operador, o sistema de Mamdani utiliza a função “min”, conforme equação abaixo:

$$D^{(k)} = \min[\mu_{A_1}^k(x_1), \mu_{A_2}^k(x_2), \dots, \mu_{A_j}^k(x_j)]. \quad (3.22)$$

3^a etapa: Aplicação do método de implicação

O método de implicação é definido como a modelagem do consequente com base no coeficiente de disparo. A implicação se dá em todas as regras. O modelo de Mamdani utiliza a função “min”, que trunca a saída do conjunto *fuzzy*.

$$S^{(k)} = \min[D^k, \mu_{B_i}(y)]. \quad (3.23)$$

4^a etapa: Agregação das Saídas

Ao realizar o processo de implicação de cada regra, gera-se uma função de pertinência truncada para a saída da regra. Como o sistema possui n regras, devem-se ter n funções de pertinências truncadas que irão gerar a função de saída. Para isso, agregam-se todos os gráficos de cada uma das funções. No modelo Mamdani utiliza-se a função “max”.

$$\mu_B = \max[S^{(k)}]. \quad (3.24)$$

5^a etapa: Defuzzificação

Nesta etapa é realizada a conversão *fuzzy* – escalar, ou seja, transformam-se informações qualitativas em uma informação quantitativa. Os métodos mais utilizados para realizar esta conversão são os métodos do centro de massa e o método da média dos máximos.

- **Sistema de inferência *Takagi-Sugeno***

Um novo modelo de inferência baseado na teoria dos conjuntos *fuzzy* foi proposto na década de 80. Denominado de modelo Sugeno, modelo de inferência *fuzzy* paramétrico ou simplesmente modelo TSK. As pesquisas mostraram que este modelo conseguia respostas satisfatórias para problemas que fossem representados razoavelmente apenas pelas suas relações entrada e saída.

Diferente do modelo de Mamdani, os modelos de inferência do tipo TSK não utilizam funções de pertinência no consequente. As relações de saída são compostas de equações paramétricas que relacionam as entradas e saída do processo. O modelo TSK é descrito a seguir.

Seja um sistema *fuzzy* composto de n regras, com uma das regras do tipo **Se** $x_1 = A_1$ **e** $x_2 = A_2$ **e** ... **e** $x_j = A_j$ **então** $y = \phi(x_1, x_2, \dots, x_j)$. O processo de inferência do modelo TSK é similar ao modelo Mamdani com algumas alterações. Na segunda etapa do processo, quando se calcula o valor do grau de disparo, a função do operador normalmente não é a função “min”. Além de não existir etapa de defuzzificação.

A saída do modelo TSK é calculada como a média ponderada das saídas de cada uma das regras que compõem o sistema, onde os pesos são os coeficientes de disparo. Seja o coeficiente de disparo de cada regra dado por:

$$D^{(k)} = T[\mu_{A_1}^k(x_1), \mu_{A_2}^k(x_2), \dots, \mu_{A_j}^k(x_j)], \quad \text{onde} \quad k = 1, \dots, n. \quad (3.25)$$

A saída do sistema TSK será dada por:

$$z = \frac{\sum_{k=1}^n D^{(k)} \cdot \phi(x_1, x_2, \dots, x_j)}{\sum_{k=1}^n D^{(k)}}. \quad (3.26)$$

A eficiência do modelo de TSK está diretamente ligada à escolha dos parâmetros da função paramétrica de saída. Normalmente, os índices desta função são estimados seguindo algum índice de desempenho definido pelo usuário. A minimização do erro quadrático entre a saída do modelo de Sugeno e os dados de saída disponíveis é normalmente utilizada como medida de desempenho.

3.2.2 Sistema Adaptativo de Inferência Neuro-Fuzzy

O Sistema Adaptativo de Inferência Neuro-Fuzzy (*Adaptive Neuro Fuzzy Inference System* – ANFIS) é um sistema híbrido que utiliza de forma conjunta as vantagens das redes neurais artificiais (RNA) e da lógica *fuzzy*. Das redes neurais, utiliza-se a capacidade de aprendizagem, enquanto que da lógica *fuzzy* se utiliza a capacidade de interpretação.

O modelo ANFIS implementa uma base de regras *fuzzy* do tipo Takagi-Sugeno, ou seja, se um conjunto de condições antecedentes é satisfeito, então um conjunto de consequentes é inferido.

Como o modelo ANFIS utiliza somente funções deriváveis, torna-se viável a utilização de um algoritmo de aprendizado padrão da teoria de redes neurais artificiais. Para isso, uma combinação do algoritmo *backpropagation* e do método de estimação de mínimos quadrados é realizada. O algoritmo *backpropagation* é utilizado para o aprendizado dos antecedentes das regras *fuzzy*, isto é, as funções de pertinência, e a estimativa de mínimos quadrados é utilizada para determinar os coeficientes das combinações nos consequentes das regras *fuzzy*. A estrutura de um modelo ANFIS com duas entradas (x e y) e uma saída (f) é apresentada na Figura 3.5.

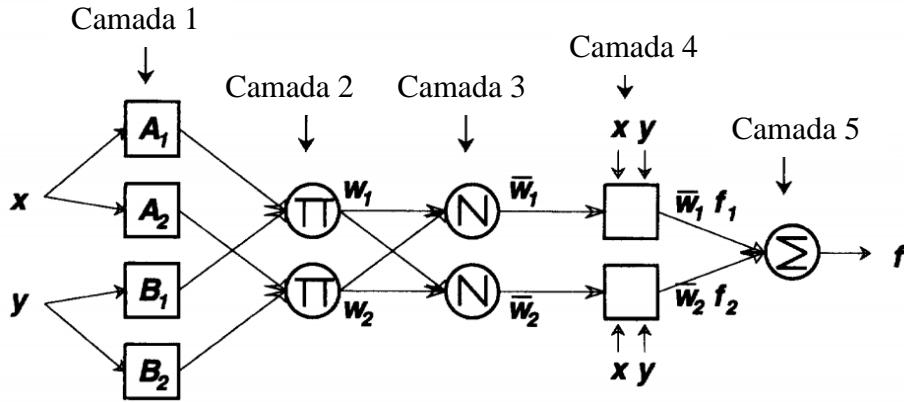


Figura 3.5 – Estrutura do modelo ANFIS. Fonte: [28].

No passo *forward*, os parâmetros das funções de pertinência são inicializados, e um vetor entradas-saída é apresentado. Calculam-se as saídas dos nós para cada camada do sistema, então os parâmetros do consequente são calculados a partir do método de mínimos quadrados. Depois de identificar os parâmetros do consequente, o erro é calculado como a diferença entre a saída do sistema e a saída desejada apresentada nos pares de treinamento.

No passo *backward*, os sinais do erro são propagados desde a saída na direção das entradas. O vetor gradiente é acumulado para cada dado de treinamento. No final do passo *backward* para todos os dados de treinamento, os parâmetros na camada 1 (os parâmetros das funções de pertinência) são atualizados pelo método do gradiente descendente. O processo de aprendizado termina quando é atingida a tolerância do erro médio quadrático ou o número máximo de épocas definido pelo usuário.

Com base na Figura 3.5, os passos que levam o ANFIS a uma adaptação, de acordo com [30], são apresentados a seguir.

Na primeira camada, cada unidade (A_1, A_2, B_1 e B_2) armazena os parâmetros para definir uma função de pertinência que representa um termo lingüístico como sendo:

$$O_i^1 = \mu_{C_i}(z), \quad (3.27)$$

em que,

z – é a entrada (x ou y) do nó i ;

C_i – é a classificação lingüística (A_i ou B_i) associada com a função do nó;

$\mu_{C_i}(z)$ – é a função de pertinência;

O_i^1 – é a saída do nó i da camada 1.

Na segunda camada, cada nó computa o peso w_i associado às funções de ativação, ou seja:

$$w_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y), \quad i = 1, 2. \quad (3.28)$$

A saída de cada nó da segunda camada representa o nível de ativação de uma regra.

Cada nó na terceira camada calcula o grau de desempenho relativo das i -ésimas regras para o somatório dos níveis de ativação dos nós, isto é:

$$\bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2. \quad (3.29)$$

Por conveniência, a saída da terceira camada pode ser chamada nível de ativação normalizado. Na quarta camada é calculado o produto da saída do nó i da terceira camada pela função f_i , $i = 1, 2$; ou seja,

$$O_i^4 = \bar{w}_i \cdot f_i = \bar{w}_i (p_i x + q_i y + r_i), \quad (3.30)$$

em que: $\{ p_i, q_i, r_i \}$ é o conjunto de parâmetros associado ao nó i .

Finalmente, a quinta camada é composta por um único nó que computa a saída do sistema como sendo o somatório de todos os sinais de entrada deste nó, isto é,

$$O_5^5 = f = \sum \bar{w}_i \cdot f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}. \quad (3.31)$$

3.3 Análise Wavelet

Uma *wavelet* é uma forma de onda com duração efetivamente limitada e possui valor médio nulo. Diferentemente das senóides, que formam a base da análise de Fourier e são suaves e simétricas, as *wavelets* tendem a ser irregulares e assimétricas.

A análise de Fourier consiste em representar determinado sinal por uma soma de senóides em diferentes frequências. De maneira similar, a análise *wavelet* decompõe um

determinado sinal como uma soma de *wavelets* deslocadas e em diferentes escalas da versão original da *wavelet* (*wavelet mother*). Na Figura 3.6 são apresentadas uma senóide e uma *wavelet* do tipo “*daubechies 10*”.

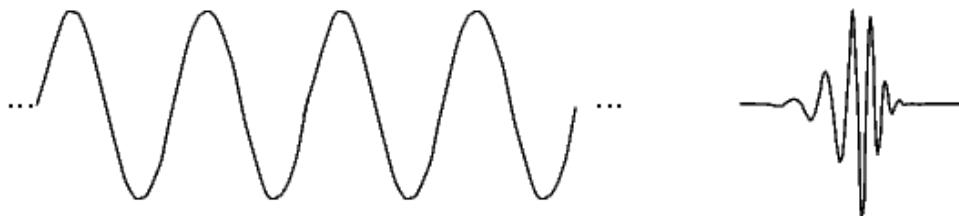


Figura 3.6 – Comparação entre uma senóide (esquerda) e uma wavelet (direita).

As *wavelets* são funções matemáticas que separam dados em suas diferentes componentes freqüenciais, e extraem cada componente com uma resolução adequada à sua escala. Elas têm vantagens em relação à análise de Fourier, pois esta última analisa o sinal como um todo, acarretando numa representação mais pobre para sinais que contêm descontinuidades e variações bruscas.

3.3.1 Transformada Wavelet

A Transformada Wavelet (TW) é uma transformada linear que pode ser utilizada na análise de sinais não estacionários para extrair informações das variações em freqüência desses sinais e para detectar suas estruturas temporalmente e/ou espacialmente localizadas [6]. Para a TW, os dados são representados via superposição de *wavelets* com diferentes posições e escalas, cujos coeficientes essencialmente quantificam a força da contribuição das *wavelets* naquelas posições e escalas.

A Figura 3.7 é de funções de base *wavelet* da família *daubechies*, os ladrilhos e o plano tempo \times frequência de cobertura. Para frequências mais altas, tem-se uma resolução alta no tempo e baixa na frequência. Já para frequências mais baixas, tem-se uma resolução baixa no tempo e alta na frequência. Isto se deve ao fato de que sinais com componentes em alta freqüência possuem rápidas alterações no domínio temporal, e sinais com componentes de baixa freqüência apresentam alterações mais lentas no domínio temporal.

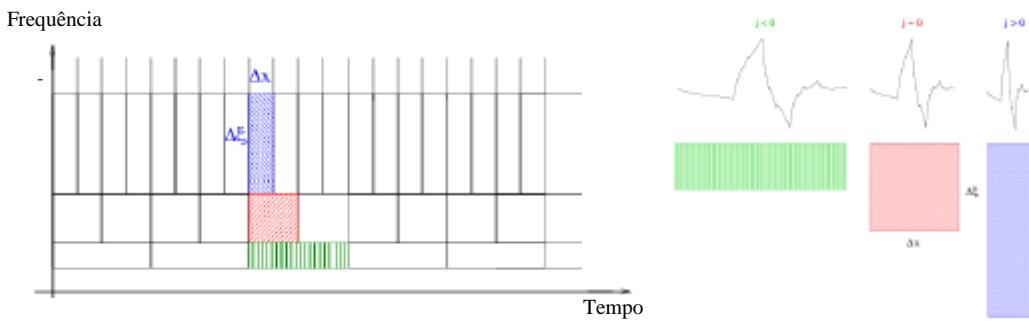


Figura 3.7 – Esquema de um plano tempo x frequência. Fonte: [6].

3.3.2 Transformada Wavelet Contínua

A Transformada Wavelet Contínua (TWC) é calculada realizando contínuas translações e mudanças de escala de uma função, *Wavelet Mother* (WM), sobre um sinal, calculando uma correlação entre eles. Para calcular a Transformada Wavelet Contínua o procedimento é o seguinte:

- 1) Definir a WM a ser utilizada e compará-la a parte inicial do sinal em análise;
- 2) Calcular a correlação (C) entre a parte inicial do sinal e a WM;

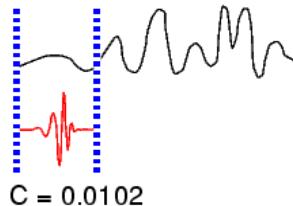


Figura 3.8 – Passo 2 da Transformada Wavelet Contínua.

- 3) Transladar a *wavelet* para direita e repetir os passos 1 a 3 até que tenha varrido todo o sinal;

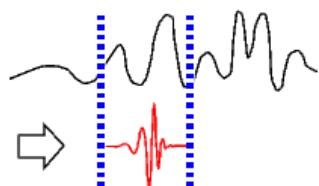


Figura 3.9 – Passo 3 da Transformada Wavelet Contínua.

- 4) Aumentar a escala da *wavelet* e repetir os passos de 1 a 3;

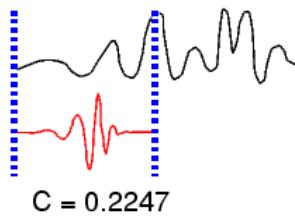


Figura 3.10 – Passo 4 da Transformada Wavelet Contínua.

- 5) Repetir os passos de 1 a 4 até que todas as escalas definidas sejam varridas.

Ao realizar o processamento computacional de sinais utilizando dados do mundo real, deve-se ter em mente que os cálculos serão executados a partir de um sinal discreto, ou seja, sobre um sinal que foi medido em intervalos discretos. Portanto, o que é "contínuo" sobre a TWC, e o que a distingue da Transformada Wavelet Discreta (que será discutida na seção seguinte), é o conjunto de escalas e posições em que a TWC opera.

Ao contrário da Transformada Wavelet Discreta, a TWC pode operar em qualquer escala, desde a escala do sinal original até uma escala máxima que pode ser determinada de acordo com a necessidade de detalhamento requerida na análise, e também de acordo com a potência computacional disponível. A TWC também é contínua em termos de deslocamento: durante a computação, a *wavelet* em análise é deslocada suavemente sobre o domínio completo da função analisada (Figura 3.11).

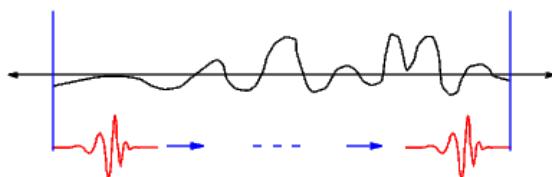


Figura 3.11 – Deslocamento da wavelet na Transformada Wavelet Contínua.

3.3.3 Transformada Wavelet Discreta

Realizar o cálculo dos coeficientes de correlação entre a *wavelet* e o sinal analisado a cada possível escala utilizada é extremamente oneroso e gera uma grande quantidade de dados, dos quais muitos são redundantes.

A Transformada Wavelet Discreta (TWD) foi desenvolvida com o intuito de proporcionar uma eficiência maior à análise *wavelet*. Diferentemente da TWC, a aplicação

da TWD não necessita que a *wavelet* seja transladada nem redimensionada continuamente, mas sim em intervalos discretos. Isto pode ser feito com uma pequena modificação na *Wavelet Contínua*, de acordo com a equação a seguir:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \Rightarrow \psi_{m,n}(t) = \frac{1}{\sqrt{|a_0^m|}} \psi\left(\frac{t-nb_0a_0^m}{a_0^m}\right), \quad (3.32)$$

em que, m e n são números inteiros, $a_0 > 1$ é um parâmetro de escala fixo, b_0 é o fator de deslocamento (que depende do fator de escala), e ψ representa a *wavelet mother*.

Os coeficientes no domínio da transformada correspondem a pontos em um reticulado bidimensional no plano escala x translação. A grade é indexada por dois inteiros m e n , sendo o primeiro associado aos passos na escala discreta e o segundo aos passos das translações discretas [6]. A Figura 3.12 ilustra o reticulado.

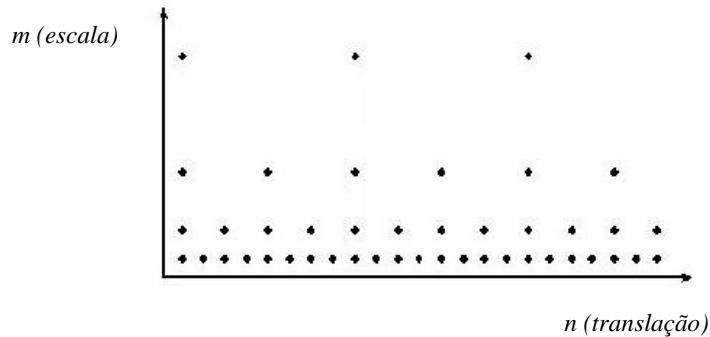


Figura 3.12 – Reticulado no plano escala x translação para a TWD. Fonte: [6].

Um caso particular amplamente utilizado para a aplicação da TWD é baseado na escolha das escalas e translações com base em potências de dois, as chamadas escalas e translações diádicas. A aplicação da TWD utilizando estes parâmetros é bastante eficiente e, normalmente, possui precisão suficiente para a realização da análise *wavelet* desejada.

De modo geral, as transformadas contínuas são primordialmente empregadas na dedução de propriedades das transformadas. Formas discretas são atraentes do ponto de vista de implementação e do ponto de vista computacional [6].

3.3.4 Análise de Multiresolução

Análise de multiresolução é uma técnica originada na área de processamentos de sinais, que no contexto das *wavelets*, constitui-se na forma padrão de construção das bases de *wavelets* e da implementação das transformadas *wavelets* ortonormais [6].

Uma forma de implementar computacionalmente a transformada *wavelet* diádica ortonormal pode ser obtida através do algoritmo baseado na representação multiresolução de sinais que, assim como a transformada *wavelet*, decompõe o sinal em escalas com diferentes resoluções no tempo e na freqüência [31].

As *wavelets* estão associadas a uma filtragem passa-faixa interativa na qual a banda passante dos filtros consecutivos adjacentes é a metade de seu antecessor. Entretanto, para se evitar um número infinito de filtros analisadores é usado um único filtro para baixas freqüências, quando a faixa de freqüência é suficientemente pequena [32].

A função escala (*Low Pass Filter* – LPF), denotada geralmente por $\phi(t)$, foi introduzida por Mallat [33]. O princípio fundamental é analisar o sinal através de uma combinação de uma função escala $\phi(t)$ (passa-baixa) e *wavelet* $\psi(t)$ (passa-faixa). Esta idéia é essencial na codificação em sub-bandas e na análise de multiresolução [6].

3.3.5 Filtragem em um estágio: aproximações e detalhes

Para muitos sinais, o conteúdo de baixa freqüência é a parte mais importante. É o que dá ao sinal a sua identidade. O conteúdo de alta freqüência, por outro lado, dá nuance ou o “tom”. Considere a voz humana. Se as componentes de alta freqüência forem retiradas, a voz soa diferente, mas ainda assim pode-se entender o que foi dito. No entanto, ao remover uma quantidade suficiente de componentes de baixa freqüência, o conteúdo restante soará como uma “gritaria” e será impossível entender o que foi dito.

Na análise *wavelet*, é comum se falar em aproximações e detalhes. As aproximações são as componentes do sinal que possuem alta escala e baixa freqüência. Os detalhes são as componentes do sinal que possuem baixa escala e alta freqüência. O processo de filtragem básico mais utilizado na análise de sinais consiste em projetar filtros passa-alta (*High-Pass*) e passa-baixa (*Low-Pass*), conforme a Figura 3.13.

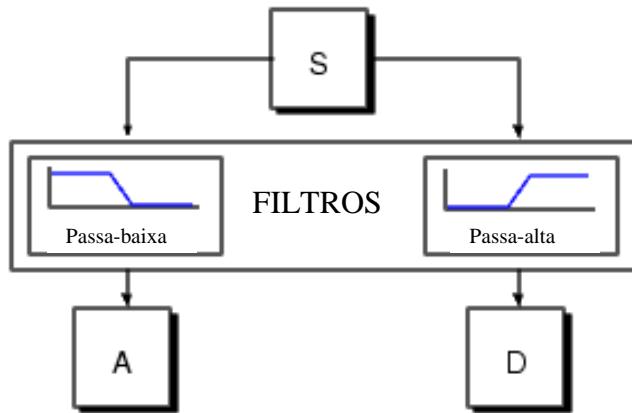


Figura 3.13 – Processo básico de filtragem de sinais.

Infelizmente, após a execução do processo básico de filtragem sobre um verdadeiro sinal digital, a quantidade de dados resultantes é igual ao dobro da quantidade de dados antes da filtragem. Por exemplo, suponha que o sinal original S é composto por 1.000 amostras de dados. Em seguida, cada um dos sinais resultantes terá 1.000 amostras, totalizando 2.000 dados.

Existe uma maneira mais sutil para realizar a decomposição utilizando *wavelets*. Ainda considerando o exemplo acima, para cada um dos sinais de saída, pode-se guardar apenas um de cada dois dados sequenciais para obter as informações mais relevantes a respeito destes sinais. Este processo é conhecido como *downsampling* (Figura 3.14). O processo à direita, que inclui o *downsampling*, produz os coeficientes da TWD.

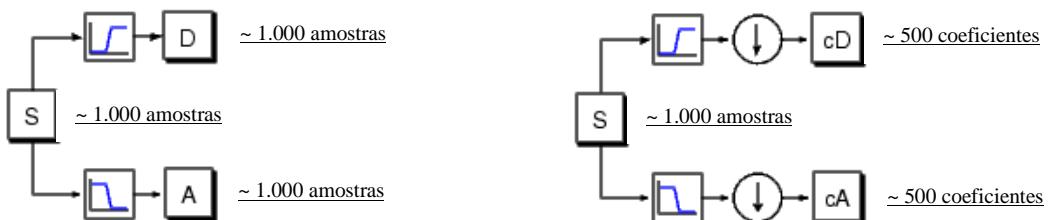


Figura 3.14 – Filtragem de sinais utilizando *downsampling*.

Na Figura 3.15 é esquematizada a filtragem em um estágio aplicando a TWD sobre uma senóide pura, distorcida pela adição de um ruído de alta frequência.

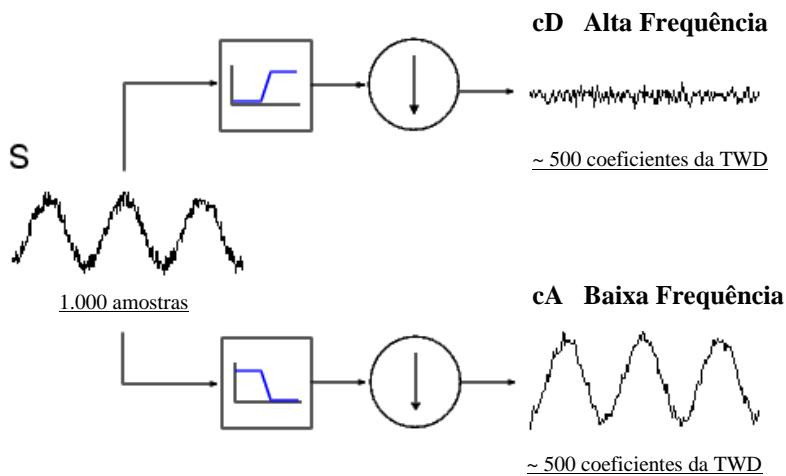


Figura 3.15 – Obtenção dos coeficientes da TWD através da filtragem em único estágio.

Na realidade, o tamanho real dos sinais de aproximação e detalhe é um pouco maior do que a metade do tamanho do sinal original. Isso se deve ao processo de filtragem, que é implementado através da convolução do sinal com um filtro.

3.3.6 Decomposição em múltiplos níveis

O processo de decomposição em múltiplos níveis pode ser realizado por sucessivas filtragens de único nível, de modo que um sinal original pode ser decomposto em vários sinais com menores resoluções.

Este processo de filtros consecutivos é conhecido como algoritmo piramidal. Este algoritmo possibilita obter “aproximações” e “detalhes” de um dado sinal de interesse. Uma aproximação é uma representação de baixa freqüência do sinal original, enquanto que um detalhe é a diferença entre duas representações sucessivas da aproximação do sinal original. Uma aproximação contém a tendência geral do sinal original, enquanto que um detalhe exibe os componentes de alta freqüência do sinal de entrada [6].

Na Figura 3.16 apresenta-se um exemplo de decomposição de um dado sinal S em três níveis de sua árvore de decomposição.

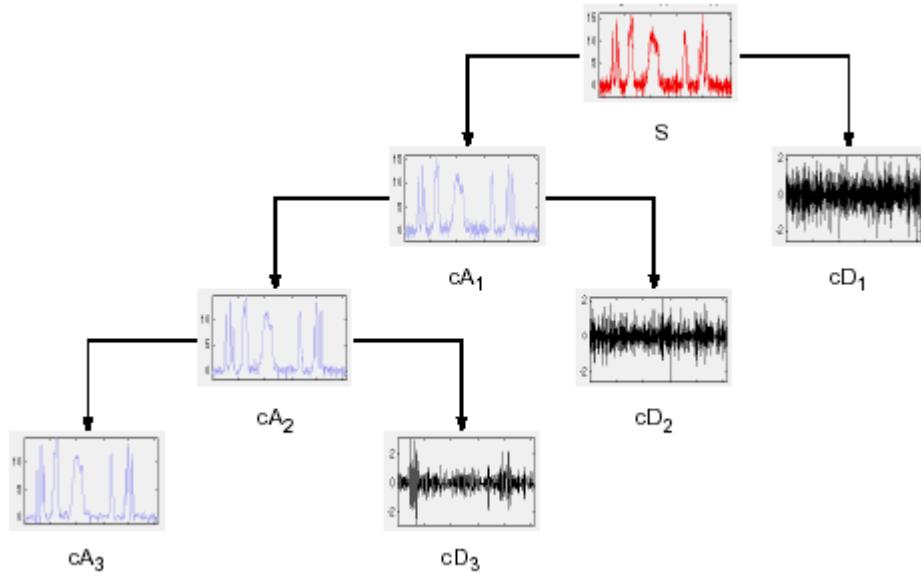


Figura 3.16 – Árvore de decomposição de um sinal em três níveis. Fonte: [6].

Por se tratar de um processo iterativo, em teoria, o processo de decomposição em múltiplos níveis pode ser mantido indefinidamente. Na realidade, a decomposição só pode prosseguir até que os detalhes individuais sejam constituídos de uma única amostra. Na prática, ao realizar-se a decomposição em múltiplos níveis, deve-se escolher um número adequado de níveis com base na natureza do sinal analisado ou em algum critério adequado.

3.4 Software e parâmetros utilizados

Para realizar a manipulação dos dados (armazenamento, tratamento estatístico, processamento de cálculos diversos, geração de gráficos) e o desenvolvimento dos modelos de previsão propostos, utilizou-se o *software* MATLAB® em sua versão 7.10.0.499 (R2010a). O processador de 64 bits utilizado foi o Intel(R) Core(TM)2 Duo CPU T6600 @ 2,20GHz 2,20GHz, com 4,00 GB de memória RAM, e com o sistema operacional Windows 7 – 64 bits.

Para criar as redes neurais e efetuar os treinamentos, foram utilizadas funções existentes no *toolbox Neural Network*. Os algoritmos de treinamento utilizados foram o *Resilient Propagation*, através da função *trainrp*, e o *Levenberg-Marquardt*, através da função *trainlm*. As redes neurais utilizadas foram do tipo *Multilayer Feedforward*, e os parâmetros utilizados para criar as redes e realizar o treinamento com os dois algoritmos foram os seguintes:

- net.divideFcn = 'divideind';
- net.divideParam.trainInd = ind_tr; % índices do conjunto de treinamento
- net.divideParam.valInd = ind_val; % índices do conjunto de validação
- net.divideParam.testInd = ind_tst; % índices do conjunto de teste
- net.inputs{1}.processFcns = { } ;
- net.inputs{1}.processParams = { } ;
- net.outputs{2}.processFcns = { } ;
- net.outputs{2}.processParams = { } ;
- net.performFcn = 'mse';
- net.trainFcn = 'trainlm'; % ou 'trainrp'
- net.trainParam.show = NaN;
- net.trainParam.showWindow = 0;
- net.trainParam.showCommandLine = 1;
- net.trainParam.epochs = 500;
- net.trainParam.goal = 0;
- net.trainParam.max_fail = 10.

Para os demais parâmetros necessários à realização dos treinamento e que não foram listados acima, foram utilizados os próprios valores *default* do *toolbox Neural Network*.

Para criar os sistemas de inferência *Fuzzy* e realizar o treinamento com o ANFIS, foram utilizadas funções existentes no *toolbox Fuzzy Logic*. Os sistemas de inferência foram criados a partir dos conjuntos de treinamento utilizando-se a função *genfis2*. A função *anfis* foi utilizada para adaptar as funções de pertinência dos sistemas de inferência gerados. Os conjuntos de treinamento e validação foram utilizados para realizar a adaptação, e o número de épocas utilizado foi igual a 100.

Em relação às *wavelets*, utilizou-se o *toolbox Wavelet* para aplicar a Transformada *Wavelet Discreta* na decomposição em múltiplos níveis de determinados sinais de entrada (velocidades médias horárias dos ventos). Dentre as bases *wavelets* testadas, aquelas que se demonstraram mais apropriadas à aplicação para os modelos desenvolvidos (ver Capítulo 5) foram as da família *Daubechies*, mais especificamente, as *wavelets* do tipo “*daubechies 10*”.

CAPÍTULO 4

4. AVALIAÇÃO DE DESEMPENHO DOS MODELOS DE PREVISÃO

Neste capítulo definem-se os critérios utilizados para avaliar os modelos de previsão propostos nesta dissertação. A seção inicial do capítulo apresenta alguns conceitos sobre séries temporais. A segunda seção traz as notações que são comumente utilizadas pela comunidade de previsão de geração eólica (ver [21]). Na terceira seção, são apresentados os modelos que servirão de referência para a comparação com as previsões dos modelos propostos. Na quarta seção são definidos os erros de previsão e, finalmente, na quinta seção, são apresentados os critérios de comparação do desempenho dos modelos.

4.1 Séries Temporais

Uma série temporal é um conjunto de observações de uma dada variável, ordenado segundo o parâmetro tempo, geralmente em intervalos equidistantes. Se Z_t representa o valor da variável aleatória Z no instante t , a série temporal pode ser denotada por Z_1, Z_2, \dots, Z_N , sendo N o tamanho da série ou o número de observações seriais da variável [34]. As séries temporais abordadas nesta dissertação são *discretas*, ou seja, séries cujo número de observações N é finito.

4.1.1 Previsão de Séries Temporais

A previsão de uma série temporal é simplesmente o estabelecimento dos valores futuros da série. Uma previsão é uma estimativa quantitativa (ou conjunto de estimativas) acerca da verossimilhança de eventos futuros baseados na informação atual e passada [34]. Uma característica dos modelos de previsão de séries temporais propostos nesta dissertação é que eles são *univariados*, ou seja, são fundamentados apenas na análise das observações da série de interesse para a especificação de algum modelo que descreva essas observações.

O horizonte de previsão é o comprimento de tempo, contado a partir de uma origem especificada, chamada origem das previsões, no sentido do futuro, para o qual as previsões devem ser determinadas. O horizonte de previsão irá variar de acordo com o propósito ou

uso final das previsões. Denomina-se o número de intervalos de tempo (períodos) para frente, a partir da origem das previsões, como o número de *passos* de uma previsão [34].

4.1.2 Autocorrelação

A autocorrelação serve para medirmos o comprimento da memória de um processo, ou seja, a extensão para a qual o valor tomado no tempo t depende daquele tomado no tempo $t - k$ [33]. A autocorrelação é uma medida que informa o quanto o valor de uma realização de uma variável aleatória é capaz de influenciar seus vizinhos, por exemplo, o quanto a existência de valor mais alto condiciona valores também altos de seus vizinhos. Por definição, o valor da autocorrelação está entre 1 (correlação perfeita) e -1 , o que significa anti-correlação perfeita. O valor 0 (zero) significa total ausência de correlação. A autocorrelação de uma dada variável se define pela distância, ou atraso com que se deseja medi-la. Quando essa distância é zero, tem-se o valor máximo 1, pois trata-se da variável correlacionada com ela mesma. Outros valores devem ser calculados caso a caso.

No presente trabalho, a série temporal discreta é um vetor contendo as velocidades médias horárias da respectiva estação anemométrica. Para o cálculo das autocorrelações utiliza-se a seguinte expressão:

$$r_k = \frac{\sum_{i=1}^{N-k} (v_i - \mu)(v_{i+k} - \mu)}{\sum_{i=1}^N (v_i - \mu)^2}, \quad (4.1)$$

em que,

v_i – é o i -ésimo elemento do vetor das velocidades;

μ – é a média do vetor das velocidades;

N – é o comprimento do vetor das velocidades;

k – é o deslocamento no tempo;

r_k – é a autocorrelação entre elementos do vetor das velocidades deslocados no tempo.

4.2 Notações

- P_{inst} : Potência instalada do parque eólico;
 P_{med} : Potência média gerada pelo parque eólico durante determinado período;
 $k = 1, 2, \dots, k_{\max}$: Passo da previsão (número de horas à frente);
 k_{\max} : Máximo passo da previsão;
 N : Número de dados utilizados para a avaliação do modelo;
 $v(t+k)$: Velocidade medida no instante $t+k$;
 $\hat{v}(t+k | t)$: Velocidade prevista na origem t para o instante $t+k$;
 $P(t+k)$: Potência medida no instante $t+k$;
 $\hat{P}(t+k | t)$: Potência prevista na origem t para o instante $t+k$;
 $e_v(t+k | t)$: Erro correspondente ao instante $t+k$ para a previsão da velocidade realizada na origem t ;
 $e_p(t+k | t)$: Erro correspondente ao instante $t+k$ para a previsão da potência realizada na origem t ;
 $\epsilon_{P_{inst}}(t+k | t)$: Erro da previsão de potência normalizado pela potência instalada;
 $\epsilon_{P_{med}}(t+k | t)$: Erro da previsão de potência normalizado pela potência média gerada.

Cabe ressaltar que, nesta dissertação, as potências são extraídas da curva de potência do aerogerador. Por exemplo, para obter o valor de $P(t+k)$, verifica-se na curva de potência do aerogerador qual é a potência gerada quando a velocidade do vento é $v(t+k)$.

4.3 Modelos de Referência

Os modelos de referência resultam de considerações simples e não exigem esforços de modelagem. Portanto, só é vantajoso desenvolver e implementar uma ferramenta avançada de previsão de geração eólica se ela for capaz de superar os modelos de

referência, ou seja, se a ferramenta fornecer resultados melhores do que aqueles obtidos com os modelos de referência [21]. Provavelmente, o modelo de referência mais comumente utilizado na previsão de energia eólica ou no campo meteorológico é o *Modelo da Persistência*. Este modelo simples assume que a medida no instante tempo $t + k$ é igual ao último valor medido (em t), ou seja,

$$\hat{v}_{PERS}(t+k|t) = v(t). \quad (4.2)$$

Apesar de sua aparente simplicidade, este modelo pode ser difícil de ser batido para os primeiros passos de previsão (em torno de 4 – 6 horas), pois a escala de mudanças na atmosfera é lenta [21]. Uma generalização do *Modelo da Persistência* é obtida ao substituir o último valor medido pela média dos últimos n valores medidos:

$$\hat{v}_{MM,n}(t+k|t) = \frac{1}{n} \sum_{i=0}^{n-1} v(t-i). \quad (4.3)$$

Às vezes, tais modelos são referidos como previsores de média móvel. Assintoticamente (quando n tende ao infinito), eles tendem à média global:

$$\hat{v}_0(t+k|t) = \overline{v(t)}. \quad (4.4)$$

Este último modelo também pode servir como um modelo de referência, mas como ele não é muito dinâmico, o seu desempenho pode ser ruim para horizontes de previsão curtos. No entanto, para horizontes mais longos, a sua habilidade de previsão é melhor do que a do *Modelo da Persistência* [21]. A fim de obter um melhor desempenho ao longo de toda a gama de horizontes de previsão, os autores propuseram a fusão dos dois modelos, que levou a um novo modelo de referência

$$\hat{v}_{NEWREF}(t+k|t) = \alpha_k v(t) + (1 - \alpha_k) \overline{v(t)}, \quad (4.5)$$

em que α_k é o coeficiente de autocorrelação entre $v(t)$ e $v(t+k)$. Os valores de $\overline{v(t)}$ e α_k devem ser estimados ou determinados a partir do conjunto de treinamento.

4.4 Definição dos Erros de Previsão

No campo de previsões de séries temporais em geral, o erro é definido como a diferença entre os valores medido e previsto. Para cada passo, os erros da previsão são definidos como:

$$e_v(t+k|t) = v(t+k) - \hat{v}(t+k|t), \quad (4.6)$$

$$e_p(t+k|t) = P(t+k) - \hat{P}(t+k|t). \quad (4.7)$$

É conveniente normalizar e_p em função da potência instalada com a finalidade de produzir resultados em valores percentuais para compará-los adequadamente com os resultados de outras localidades:

$$\epsilon_{P_{inst}}(t+k|t) = 100 \left(\frac{e_p(t+k|t)}{P_{inst}} \right). \quad (4.8)$$

Esta normalização pode levar a valores de erro muito baixos para parques com capacidade instalada elevada. Para obter valores mais conservadores, se sugere a normalização pela potência média gerada [35]:

$$\epsilon_{P_{med}}(t+k|t) = 100 \left(\frac{e_p(t+k|t)}{P_{med}} \right). \quad (4.9)$$

Qualquer erro de previsão pode ser decomposto como a soma de duas parcelas, sendo uma delas denominada erro sistemático (μ_e), e a outra, erro aleatório (ξ_e) [21].

$$e = \mu_e + \xi_e, \quad (4.10)$$

em que, μ_e é um valor constante, enquanto que ξ_e é uma variável aleatória cuja média é zero. O erro sistemático é igual ao valor médio do erro de previsão sobre todo o período de avaliação e é calculado para cada passo de previsão de acordo com a seguinte equação:

$$\hat{\mu}_e(k) = \overline{e(k)} = \frac{1}{N} \sum_{t=1}^N e(t+k | t). \quad (4.11)$$

Os quatro tipos básicos de erros utilizados nesta dissertação para medir o desempenho de um modelo de previsão são: o erro absoluto médio (*Mean Absolut Error – MAE*), o erro absoluto percentual médio (*Mean Absolut Percentage Error – MAPE*), o erro quadrático médio (*Mean Squared Error – MSE*) e a raiz do erro quadrático médio (*Root Mean Squared Error – RMSE*). As equações utilizadas para calculá-los são definidas a seguir:

$$MAE_v(k) = \frac{1}{N} \sum_{t=1}^N |e_v(t+k | t)|, \quad (4.12)$$

$$MAPE_v(k) = \frac{100}{N} \sum_{t=1}^N \frac{|e_v(t+k | t)|}{v(t+k)}, \quad (4.13)$$

$$MSE_v(k) = \frac{1}{N} \sum_{t=1}^N e_v(t+k | t)^2, \quad (4.14)$$

$$RMSE_v(k) = \sqrt{MSE_v(k)}. \quad (4.15)$$

As expressões do *MAE*, *MSE* e *RMSE* também se aplicam para as potências geradas. Não se deve utilizar o *MAPE* para as potências, pois $P(t+k)$ será nula se houver instantes nos quais $v(t+k)$ é menor do que a velocidade de *cut-in* do aerogerador.

Estatisticamente, os valores do erro médio e do *MAE* estão associados com o momento de primeira ordem do erro de previsão, e portanto, são medidas que estão relacionadas diretamente com a energia produzida. Os valores do *RMSE* estão associados com o momento de segunda ordem, e portanto, estão relacionados com a variância do modelo de previsão [21].

4.5 Critérios para Comparação dos Modelos

Ao propor um novo modelo de previsão, é muito importante destacar e quantificar os ganhos obtidos em relação aos modelos de referência [21]. A fórmula utilizada para calcular estes ganhos percentuais para cada passo de previsão é a seguinte:

$$G_{ref,CA}(k) = 100 \left(\frac{CA_{ref}(k) - CA(k)}{CA_{ref}(k)} \right), \quad (4.16)$$

em que,

$CA_{ref}(k)$ – é o critério de avaliação do modelo de referência;

$CA(k)$ – é o critério de avaliação do modelo proposto.

O critério de avaliação pode ser o *MAE*, *MAPE*, *MSE* ou *RMSE*. Obviamente, ao calcular o ganho, o mesmo critério deve ser utilizado para os dois modelos que estão sendo comparados.

Uma outra maneira de avaliar o desempenho dos modelos é o coeficiente de determinação R^2 . Para cada passo de previsão, calcula-se o valor de R^2 pela seguinte equação:

$$R^2(k) = 1 - \left(\frac{\sum_{t=1}^N e_v(t+k | t)^2}{\sum_{t=1}^N v(t+k)^2} \right). \quad (4.17)$$

Este coeficiente representa a habilidade que o modelo possui para explicar a variância dos dados. O valor de R^2 deve estar situado entre zero e um. Quanto mais próximo da unidade estiver o valor de R^2 , melhor será o modelo de previsão.

Há ainda diversas ferramentas que podem ser utilizadas para a análise exploratória e comparação das previsões obtidas com diferentes modelos. Algumas delas são mais adequadas à previsão de geração eólica, pois permitem uma visão mais profunda sobre o desempenho dos modelos que estão sendo analisados. Uma ferramenta útil é o gráfico dos erros quadráticos médios acumulados, pois pode-se analisar visualmente o comportamento do modelo de previsão ao longo de determinado período [21]. Também é importante traçar

os diagramas de dispersão dos valores medidos e as respectivas previsões. Nestes diagramas, quanto maior a proximidade dos pontos em relação à reta, maior será a correlação entre a série temporal medida e as previsões.

CAPÍTULO 5

5. DESENVOLVIMENTO DOS MODELOS DE PREVISÃO

O presente capítulo apresenta a metodologia utilizada para o desenvolvimento dos modelos de previsão. A seção inicial do capítulo apresenta uma estatística descritiva das séries de velocidades médias horárias utilizadas para realizar os ajustes dos modelos. A segunda seção apresenta os modelos propostos. Na terceira seção, são apresentados os procedimentos para o treinamento e ajustes destes modelos. Finalmente, na quarta seção, são definidos os melhores modelos de previsão para cada localidade estudada.

5.1 Séries de Velocidades Médias Horárias

As séries de velocidades utilizadas nesta dissertação são publicadas pelo Instituto Nacional de Meteorologia (INMET) na internet (<http://www.inmet.gov.br/>). As séries correspondem aos dados medidos nas estações meteorológicas de superfície automáticas situadas no estado do Rio Grande do Norte, nas cidades de Macau, Mossoró e Natal.

Uma estação meteorológica de superfície automática é composta de uma unidade de memória central (*datalogger*) ligada a vários sensores dos parâmetros meteorológicos (pressão atmosférica, temperatura e umidade relativa do ar, precipitação, radiação solar, direção e velocidade do vento, etc.) que integra os valores observados minuto a minuto e calcula o valor médio atualizando os dados automaticamente a cada hora. Os dados das estações automáticas estão disponíveis no site do INMET por apenas três meses.

Conforme dito anteriormente, os modelos de previsão de séries temporais propostos neste trabalho são univariados, logo, todas as variáveis de entrada e saída dos modelos de previsão serão apenas as velocidades médias horárias. As velocidades previstas correspondem à altura de 10 metros acima do nível do solo, pois os dados de velocidade das estações de superfície automática do INMET são medidos nesta altura.

5.1.1 Estatística Descritiva

A estatística descritiva é um ramo da estatística que trata da extração de informações contidas em conjuntos de dados. A apresentação destas informações pode ser feita utilizando-se a representação tabular, a representação por parâmetros e a

representação gráfica, que possibilita uma rápida visão geral dos dados. Objetivando a eliminação de erros capazes de provocar futuros enganos de apresentação e análise, deve-se realizar uma revisão crítica dos dados. Após a revisão, convém organizar os dados de maneira prática e racional, para que se obtenha um melhor entendimento do fenômeno estudado.

Para cada uma das estações de superfície automáticas, os dados (velocidades) foram armazenados em um vetor de comprimento igual ao número de amostras N , de modo a facilitar a análise e manipulação dos mesmos. O armazenamento dos dados em forma de vetores facilita a programação computacional, permitindo que os diversos cálculos e medidas estatísticas sejam realizados de forma rápida e eficaz. São apresentados, na Tabela 5.1, os períodos e o total de amostras (tamanho das séries) utilizadas para o desenvolvimento dos modelos de previsão de cada estação de superfície automática.

Tabela 5.1 – Dados utilizados para o desenvolvimento dos modelos de previsão.

Estação	Período	Total de Amostras (N)
MACAU	Janeiro a Dezembro – 2008	8.784
MOSSORÓ	Janeiro a Dezembro – 2008	8.784
NATAL	Janeiro a Dezembro – 2008	8.784

Na Figura 5.1 são apresentadas as séries de velocidades e os histogramas para as estações de MACAU, MOSSORÓ e NATAL, respectivamente. Os histogramas permitem que seja feita uma análise das faixas de velocidade para as quais ocorrem as maiores frequências de observações dos dados. Os valores mínimos, máximos, as médias e os desvios padrões das séries de velocidades das três estações são apresentados na Tabela 5.2.

Tabela 5.2 – Estatísticas das séries de velocidades de MACAU, MOSSORÓ e NATAL.

Estação	Mínima [m/s]	Máxima [m/s]	Média [m/s]	Desvio padrão [m/s]
MACAU	0,20	11,60	4,48	1,99
MOSSORÓ	0,10	9,40	3,32	1,95
NATAL	0,09	10,50	4,70	1,81

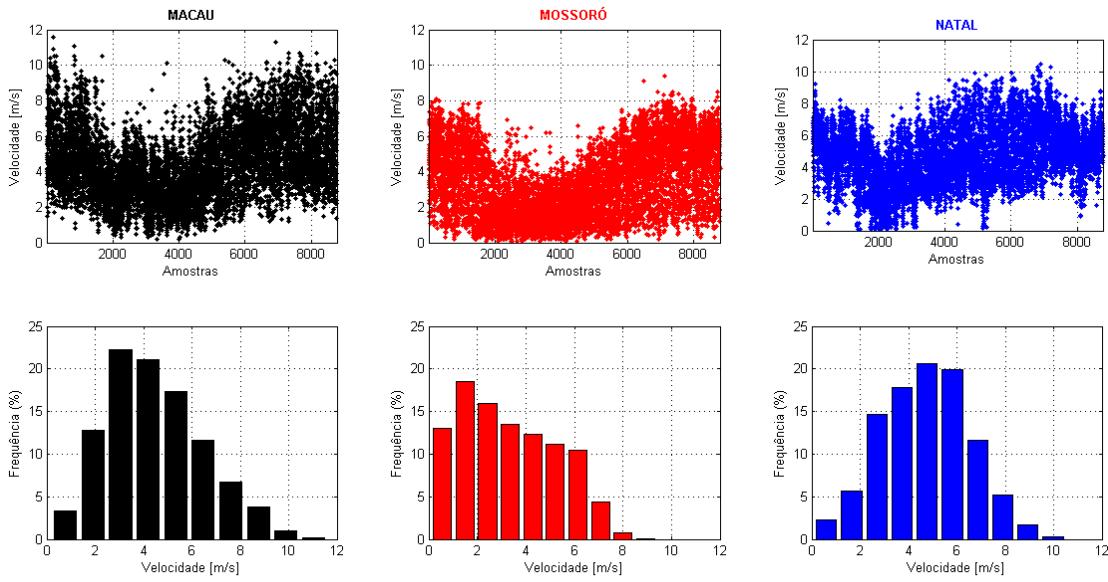


Figura 5.1 – Séries de velocidades e histogramas de MACAU, MOSSORÓ e NATAL.

Há uma pequena quantidade de velocidades menores ou iguais a 2 m/s para MACAU e NATAL, correspondendo a 10,19% e 7,17% dos dados, respectivamente. Já para MOSSORÓ, 33,40% das velocidades estão situadas nessa faixa. Outra característica das três estações é que a grande maioria das velocidades possui valor menor ou igual a 7 m/s. Para MACAU, 88,35% das velocidades são menores ou iguais a 7 m/s, e para MOSSORÓ e NATAL, estes valores percentuais correspondem a 97,32% e 90,08% dos dados, respectivamente.

Na Figura 5.2 são apresentados os dias típicos e as sazonalidades das velocidades dos ventos para as três localidades. A composição do dia típico de uma determinada localidade é obtida calculando-se a média aritmética das velocidades correspondentes para cada hora do dia ao longo de todo o ano em estudo. Observa-se claramente que a velocidade dos ventos possui uma variação cíclica diária nas três localidades. As velocidades diminuem ao longo do dia, atingindo seu valor mínimo entre 08:00 – 09:00 horas (*Coordinated Universal Time – UTC*). Posteriormente, as velocidades aumentam até atingirem seu valor máximo entre 15:00 – 19:00 horas (UTC). A sazonalidade pode ser visualizada nos gráficos das médias mensais das velocidades dos ventos. O comportamento sazonal dos ventos na região Nordeste é de extrema importância para a geração eólica, pois há a possibilidade de utilizar este tipo de geração como uma forma complementar à geração hidrelétrica, uma vez que nos períodos de poucas chuvas os ventos são mais favoráveis, e nos períodos úmidos os ventos são mais fracos.

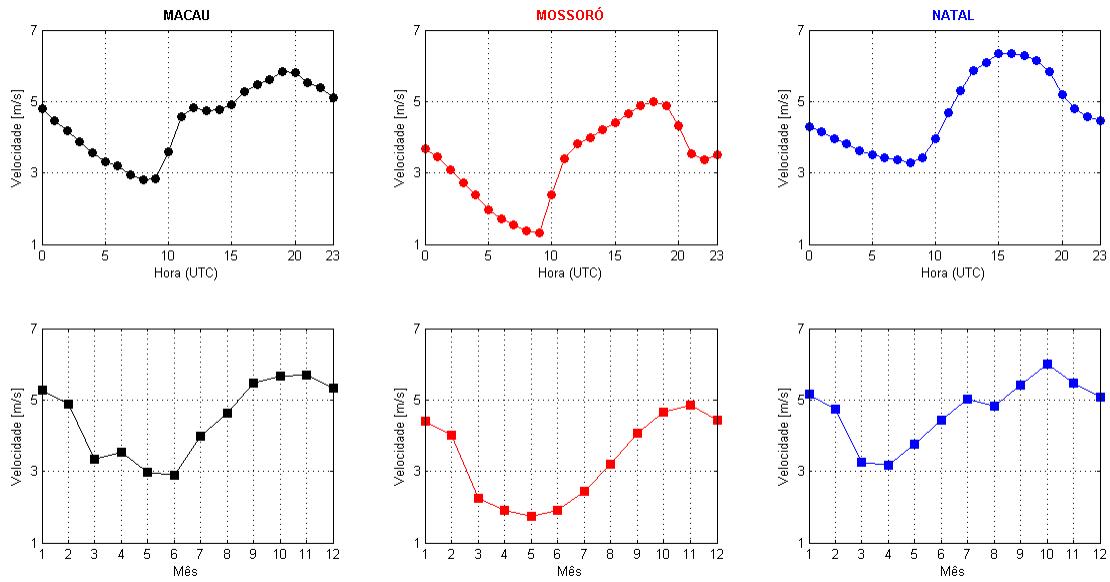


Figura 5.2 – Dia típico e comportamento sazonal de MACAU, MOSSORÓ e NATAL.

As autocorrelações das séries temporais abordadas podem ser analisadas na Figura 5.3, cujos gráficos são muito semelhantes. Há valores mínimos de autocorrelação para deslocamentos múltiplos de 12 horas, e valores máximos de autocorrelação para deslocamentos múltiplos de 24 horas. O máximo deslocamento utilizado para o cálculo da autocorrelação foi de 48 horas. Quando este deslocamento tende para um valor muito grande, a autocorrelação tende a zero.

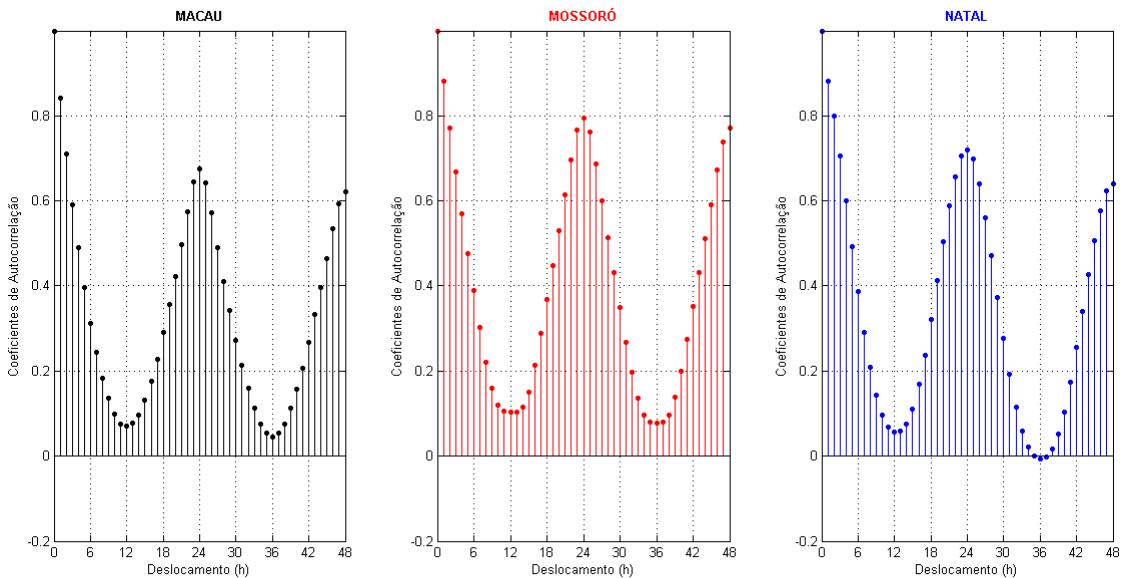


Figura 5.3 – Coeficientes de autocorrelação de MACAU, MOSSORÓ e NATAL.

5.2 Modelos de Previsão

Seis diferentes modelos para a previsão das velocidades médias horárias dos ventos são propostos nesta dissertação. Quatro deles utilizam as *Redes Neurais Artificiais* (RNA) do tipo *Multilayer Feedforward*, e os outros dois utilizam o *Adaptive Neuro-Fuzzy Inference System* (ANFIS). Estas ferramentas de inteligência artificial foram detalhadas no Capítulo 3.

5.2.1 Modelos RNA(LM), RNA(RP) e ANFIS

Estes três modelos possuem os mesmos padrões de entrada e saída, isto é, quatro entradas e uma saída. Os dados de entrada são as quatro últimas velocidades médias horárias e a saída corresponde à velocidade média horária prevista para o passo de previsão k . A Figura 5.4 é um esquema entradas-saída dos modelos.



Figura 5.4 – Entradas e saída dos modelos RNA(LM), RNA(RP) e ANFIS.

Os modelos *RNA(LM)* e *RNA(RP)* são formados por redes cujos algoritmos de treinamento são o *Levenberg-Marquardt* (LM) e o *Resilient Propagation* (RP), respectivamente. A arquitetura destas redes é formada por uma camada de entrada com quatro entradas, uma camada intermediária (oculta) e uma camada de saída com uma saída.

A quantidade de neurônios da camada oculta dos modelos neurais é determinada variando-se o número de neurônios desta camada, sendo selecionada a quantidade que fornecer o melhor desempenho durante os treinamentos. Os neurônios da camada oculta utilizam a função de ativação Tangente Hiperbólica. A camada de saída possui apenas um neurônio, pois há apenas uma saída. A função de ativação deste neurônio é a Sigmóide Logística.

O modelo *ANFIS* também possui quatro entradas e uma saída. A seleção do melhor sistema de inferência foi realizada a partir da técnica *subtractive clustering*. O comprimento do raio de influência de cada *cluster* é determinado variando-se o seu valor, sendo selecionado aquele comprimento que fornecer o melhor desempenho durante os treinamentos.

5.2.2 Modelos *TWRNA(LM)*, *TWRNA(RP)* e *TWANFIS*

As diferenças destes três modelos em relação aos anteriores são: o número de entradas e os tipos de sinais fornecidos às ferramentas de inteligência artificial. Os modelos *TWRNA(LM)*, *TWRNA(RP)* e *TWANFIS* possuem 16 entradas, definidas da seguinte maneira:

- Aplica-se ao vetor das velocidades a decomposição em três níveis da Transformada *Wavelet* Discreta;
- Montam-se quatro vetores de mesma ordem do vetor das velocidades, sendo eles: \mathbf{A}_3 – vetor de aproximação do 3º nível; \mathbf{D}_3 – vetor de detalhe do 3º nível; \mathbf{D}_2 – vetor de detalhe do 2º nível; \mathbf{D}_1 – vetor de detalhe do 1º nível.

Cada entrada $v(\tau)$ dos modelos anteriores é substituída pelas quatro entradas correspondentes $a_3(\tau), d_3(\tau), d_2(\tau)$ e $d_1(\tau)$, em que $\tau = t-3, t-2, t-1$ ou t . A Figura 5.5 é um esquema entradas-saída dos modelos.

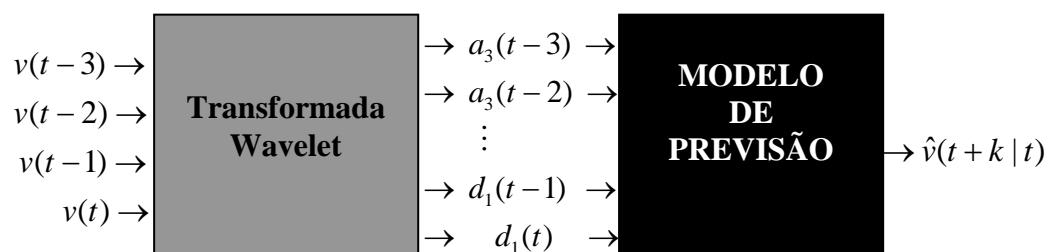


Figura 5.5 – Entradas e saída dos modelos *TWRNA(LM)*, *TWRNA(RP)* e *TWANFIS*.

5.3 Procedimentos para o treinamento e ajuste dos modelos

5.3.1 Normalização dos dados

Os dados de entrada dos modelos que utilizam as RNA devem ser normalizados. Segundo [6], a normalização é necessária para assegurar que todas as variáveis usadas nos modelos tenham igual atenção durante o treinamento. Além disto, os neurônios artificiais, geralmente, são compostos de funções de ativação que são limitadas. Assim, a normalização deve limitar os valores dos dados utilizados nos extremos das funções de ativação.

Para que os valores normalizados estejam contidos no intervalo $[0,1]$ a normalização é realizada empregando-se a seguinte expressão:

$$\tilde{x}(\tau) = \frac{x(\tau) - x_{\min}}{x_{\max} - x_{\min}}, \quad (5.1)$$

em que,

$\tilde{x}(\tau)$ – é o valor normalizado do dado de entrada correspondente ao instante τ ;

$x(\tau)$ – é o valor real do dado de entrada correspondente ao instante τ ;

x_{\min} – é o valor da menor componente do vetor ao qual pertence o dado de entrada;

x_{\max} – é o valor da maior componente do vetor ao qual pertence o dado de entrada.

Para os modelos *TWRNA(LM)* e *TWRNA(RP)*, devem ser realizadas duas normalizações distintas. Uma delas se aplica ao vetor \mathbf{A}_3 , a outra se aplica à concatenação dos vetores \mathbf{D}_3 , \mathbf{D}_2 e \mathbf{D}_1 . Para a normalização destes vetores, x_{\min} corresponde ao valor mínimo das três séries de detalhes concatenadas, e x_{\max} corresponde ao valor máximo.

O valor de saída dos quatro modelos que utilizam as RNA é normalizado, logo, os valores de x_{\min} e x_{\max} utilizados para o ajuste dos modelos devem ser armazenados para que se possa realizar a desnormalização de $\hat{v}(t+k | t)$.

5.3.2 Partição das bases de dados

Antes de realizar os treinamentos dos modelos, montou-se, para cada uma das séries de velocidades, e para cada passo de previsão (1 – 24 horas), uma matriz dos padrões de entradas e saída. Para cada linha destas matrizes de padrões, nas primeiras colunas estão as entradas e na última coluna a saída. Para os modelos *RNA(LM)*, *RNA(RP)* e *ANFIS*, estas matrizes possuem cinco colunas, enquanto que para os modelos *TWRNA(LM)*, *TWRNA(RP)* e *TWANFIS*, as matrizes têm dezessete colunas. No total, foram montadas 432 matrizes de padrões.

Os padrões de treinamento dos modelos que utilizam redes neurais devem ser normalizados, portanto, as matrizes dos padrões foram normalizadas antes de realizar os treinamentos dos modelos *RNA(LM)*, *RNA(RP)*, *TWRNA(LM)* e *TWRNA(RP)*.

A partir de cada uma das matrizes de padrões, foram criados os conjuntos de treinamento, com 60% das linhas, validação, com outros 30%, e teste, com os 10% restantes. Para os modelos que utilizam redes neurais, as matrizes normalizadas foram utilizadas para a criação destes conjuntos.

5.3.3 Determinação do número de neurônios da camada oculta

De acordo com a regra utilizada em [20], o número de neurônios na camada oculta é determinado por tentativas, sendo eleito o que corresponder ao melhor desempenho durante os treinamentos. Para cada um dos quatro modelos que utilizam redes neurais propostos nesta dissertação, o procedimento para determinar a quantidade de neurônios da camada oculta foi o seguinte:

- I. Criou-se uma rede com três neurônios na camada oculta;
- II. Atribuiu-se pesos aleatórios para todas as conexões da rede, realizou-se o treinamento através do algoritmo específico para o modelo (LM ou RP) e calculou-se o MAE, MSE, RMSE e MAPE para cada um dos conjuntos (treinamento, validação e teste);
- III. O passo II foi repetido dez vezes. Após a décima repetição, foram calculadas as médias do MAE, MSE, RMSE e MAPE, seguindo posteriormente ao próximo passo;
- IV. Foi adicionado mais um neurônio na camada intermediária, retornou-se ao passo II, e esse “loop” continuou até a rede possuir quinze neurônios na camada oculta;

V. Escolheu-se a arquitetura da rede que apresentou menor média do MAPE para o conjunto de validação.

Depois de realizados os passos descritos anteriormente, o número de neurônios na camada intermediária para cada um dos modelos e para cada uma das localidades consideradas (MACAU, MOSSORÓ e NATAL) estava determinado. As quantidades de neurônios para os modelos são apresentadas na Figura 5.6. Para cada uma das localidades, observa-se que os valores são bastante variados para os passos de previsão considerados. De uma maneira geral, pode-se dizer que os modelos que utilizam o algoritmo *Resilient Propagation* necessitam de uma quantidade maior de neurônios na camada oculta para fornecer melhores resultados de previsão.

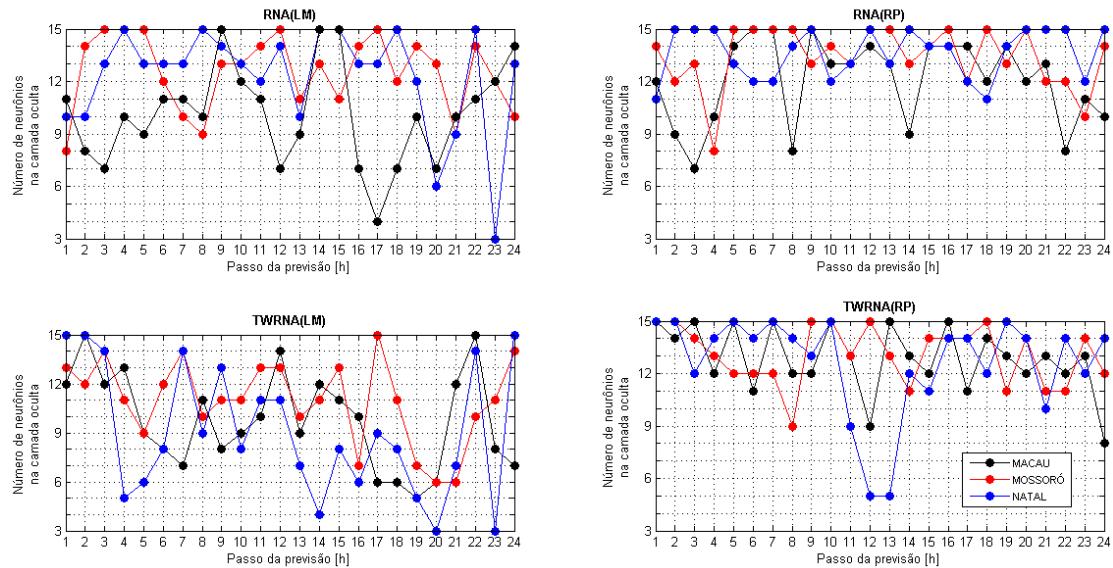


Figura 5.6 – Número de neurônios na camada oculta.

Os tempos médios para realizar uma inicialização dos pesos e o treinamento das redes são apresentados nas Figura 5.7 e Figura 5.8. Os tempos gastos pelos modelos *RNA(LM)* e *RNA(RP)* (para alguns passos) são apresentados na Figura 5.7. Observa-se que, em geral, o tempo gasto é maior com o aumento do número de neurônios na camada oculta, e reduz com o aumento do passo da previsão. O tipo de algoritmo de treinamento também influenciou significativamente nos tempos médios, sendo o *Resilient Propagation* aquele que exigiu mais tempo de treinamento.

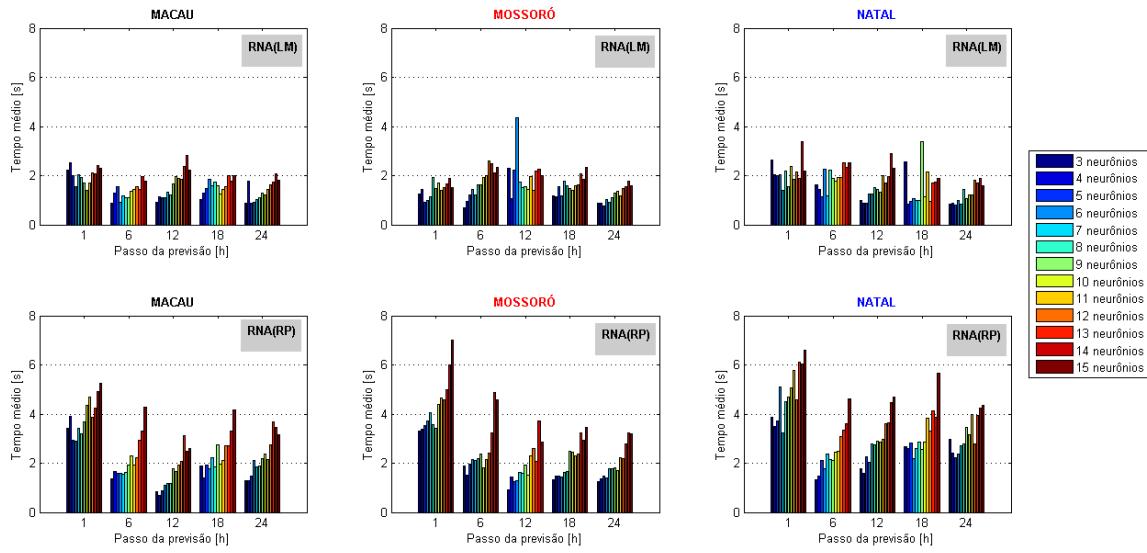


Figura 5.7 – Tempos médios de treinamento dos modelos RNA(LM) e RNA(RP).

Na Figura 5.8, são apresentados os tempos gastos pelos modelos TWRNA(LM) e TWRNA(RP) (para alguns passos). Por possuírem uma quantidade maior de entradas, estes modelos exigiram maiores tempos de treinamento. Novamente, observa-se que, de uma maneira geral, o tempo gasto aumenta com o aumento do número de neurônios na camada oculta, e reduz com o aumento do passo da previsão. Além disto, verifica-se também que o tipo de algoritmo de treinamento influenciou significativamente nos tempos médios, entretanto, para estes modelos, o algoritmo *Levenberg-Marquardt* exigiu maiores tempos de treinamento.

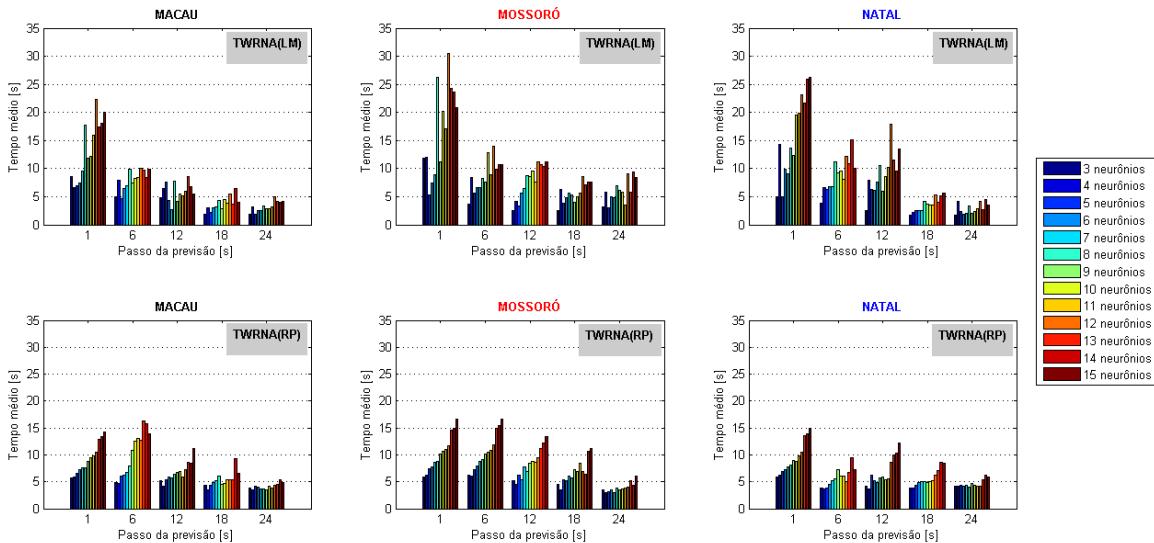


Figura 5.8 – Tempos médios de treinamento dos modelos TWRNA(LM) e TWRNA(RP).

Os gráficos comparativos do MAPE do conjunto de validação para as melhores arquiteturas são apresentados na Figura 5.9 para efeito de comparação. Na parte superior, estão traçados os gráficos para os modelos que utilizam o algoritmo *Levenberg-Marquardt* e na parte inferior, os gráficos para os modelos que utilizam o algoritmo *Resilient Propagation*. Observa-se que o desempenho dos modelos que utilizam a Transformada *Wavelet* é superior para praticamente todos os passos de previsão. Além disto, dentre os modelos que utilizam a TW, o algoritmo LM forneceu melhores previsões, principalmente para os passos mais curtos. Ainda em relação à Figura 5.9, observa-se que os valores do MAPE para MOSSORÓ são mais elevados. Isto se deve principalmente ao fato de que há um grande número de velocidades baixas para esta estação, o que contribui para o aumento do MAPE.

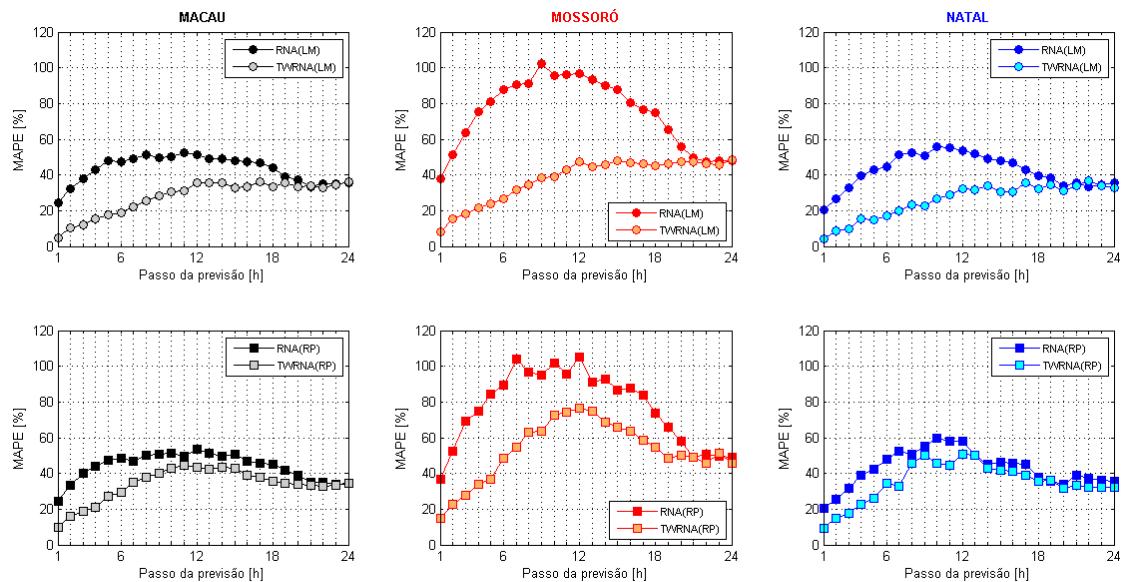


Figura 5.9 – MAPE do conjunto de validação para as melhores arquiteturas das redes.

5.3.4 Escolha da melhor rede para cada modelo neural

Após a determinação da quantidade de neurônios nas camadas ocultas, a qual fornece as arquiteturas finais das redes, falta ainda descobrir aquelas que se adaptem melhor ao problema abordado. Em outras palavras, falta definir quais redes possuem melhor capacidade de generalização para as previsões das velocidades de cada localidade e para cada um dos passos de previsão considerados.

A metodologia utilizada para a escolha da melhor rede foi baseada em [6]. Para realizar esta escolha, aplicou-se o método de validação cruzada múltipla, também conhecido como *k-fold cross-validation*, em que, k representa o número de partições geradas aleatoriamente a partir das matrizes dos padrões para treinar, testar e validar as redes. Nesse método, os padrões são divididos em k partições mutuamente exclusivas. A cada iteração do método, uma partição diferente é utilizada para testar o sistema e todas as outras ($k - 1$) partições são utilizadas para treinar e validar o treinamento das redes [6].

A partir de cada uma das matrizes dos padrões, para facilitar a implementação do método, montou-se uma matriz com “3 dimensões”, possibilitando a criação de dez experimentos por matriz. Para cada um dos experimentos, construíram-se os conjuntos de treinamento (seis partições), validação (três partições) e teste (uma partição). Pode-se observar, na Figura 5.10, como as matrizes *k-fold* foram montadas.

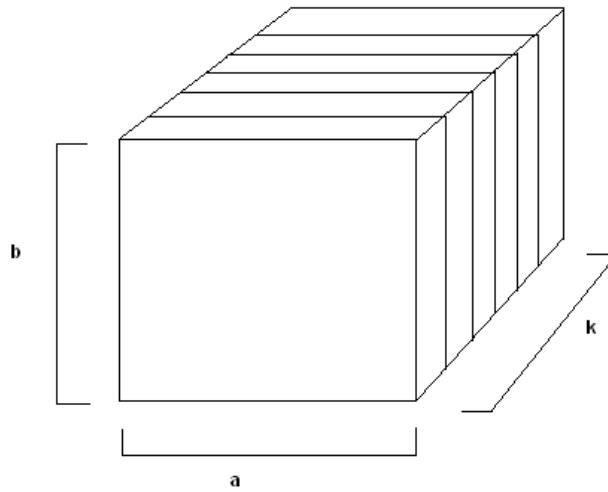


Figura 5.10 – Representação esquemática da matriz *k-fold*.

A montagem da matriz *k-fold* torna mais fácil a definição dos conjuntos de experimentos utilizados para realizar treinamentos com dados diferentes. Foram criados 10 experimentos, e para cada experimento, foram realizadas 10 inicializações dos pesos.

Na Tabela 5.3 descreve-se a formação dos conjuntos de treinamento, validação e teste, bem como identificam-se as redes, ou seja, a inicialização de cada experimento, que utilizaram os respectivos conjuntos.

Tabela 5.3 – Experimentos realizados no método de validação cruzada. Fonte: [6].

Experimento	Redes	Treinamento (partições)	Validação (partições)	Teste (partição)
1	1 – 10	5, 6, 7, 8, 9, 10	2, 3, 4	1
2	11 – 20	1, 6, 7, 8, 9, 10	3, 4, 5	2
3	21 – 30	1, 2, 7, 8, 9, 10	4, 5, 6	3
4	31 – 40	1, 2, 3, 8, 9, 10	5, 6, 7	4
5	41 – 50	1, 2, 3, 4, 9, 10	6, 7, 8	5
6	51 – 60	1, 2, 3, 4, 5, 10	7, 8, 9	6
7	61 – 70	1, 2, 3, 4, 5, 6	8, 9, 10	7
8	71 – 80	2, 3, 4, 5, 6, 7	9, 10, 1	8
9	81 – 90	3, 4, 5, 6, 7, 8	1, 2, 10	9
10	91 – 100	4, 5, 6, 7, 8, 9	1, 2, 3	10

Espera-se, com aplicação da técnica de validação cruzada, que os valores médios de MSE e MAPE, obtidos em cada experimento no conjunto de teste, sejam considerados como o resultado esperado para as redes *Multilayer Feedforward* [6].

A escolha do melhor modelo neural, para cada um dos modelos propostos, foi realizada a partir do experimento que forneceu o menor valor médio do MAPE das 10 inicializações para o conjunto de teste. Após a determinação deste experimento, escolheu-se a rede que apresentou o menor MAPE para o conjunto de teste do experimento determinado. Os gráficos comparativos do MAPE para as previsões dos modelos escolhidos após a validação cruzada são apresentados na Figura 5.11. As mesmas observações feitas para os gráficos da Figura 5.9 também se aplicam aos gráficos da Figura 5.11.

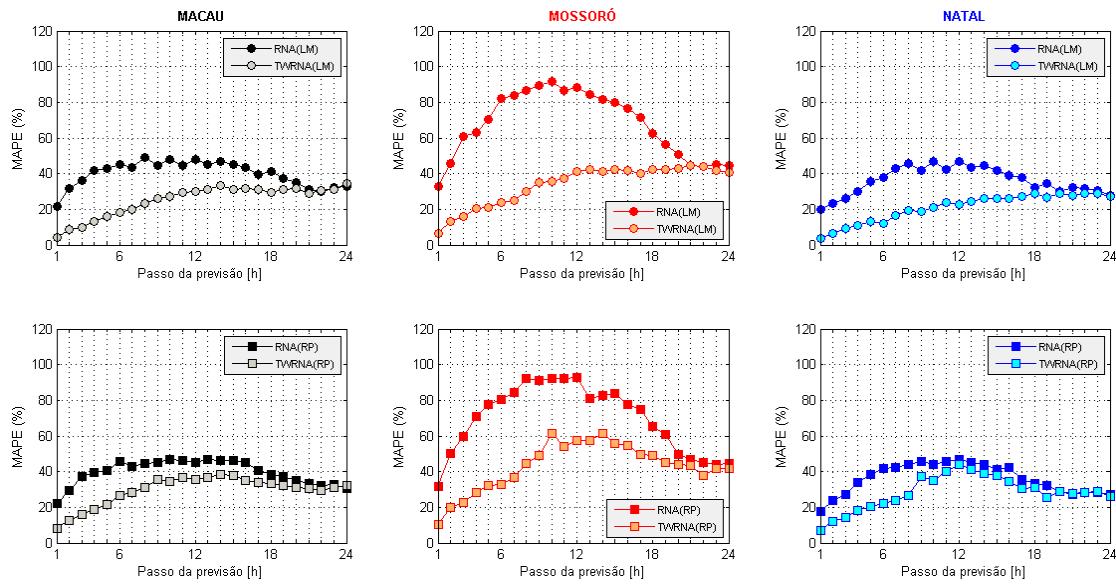


Figura 5.11 – MAPE do conjunto de teste para as melhores redes.

5.3.5 Escolha do melhor Sistema de Inferência Fuzzy

O melhor Sistema de Inferência Fuzzy (*Fuzzy Inference System – FIS*) foi escolhido adotando-se o seguinte procedimento:

- I. Utilizando a técnica *subtractive clustering*, a partir do conjunto de treinamento, gerou-se inicialmente o FIS com o tamanho do raio de influência igual a 0,3;
- II. Em seguida, as funções de pertinência do FIS foram adaptadas com o ANFIS. Os conjuntos de treinamento e validação foram utilizados para realizar a adaptação dessas funções;
- III. O conjunto de teste foi simulado e em seguida foram calculados o MAE, MSE, RMSE e MAPE;
- IV. Gerou-se um novo FIS incrementando o tamanho do raio em 0,1 e retornou-se ao passo II. Após calcular o MAE, MSE, RMSE e MAPE para o FIS com o raio igual a 0,7, passou-se para o passo V;
- V. Escolheu-se o melhor FIS de acordo com o menor MAPE para o conjunto de teste.

Este procedimento foi utilizado para MACAU, MOSSORÓ e NATAL, e para cada um dos modelos, *ANFIS* e *TWANFIS*.

Depois de realizados os passos descritos anteriormente, os melhores modelos *ANFIS* e *TWANFIS* para MACAU, MOSSORÓ e NATAL estavam determinados. Os raios de influências dos centros dos *clusters* para os modelos escolhidos são apresentados na Figura 5.12.

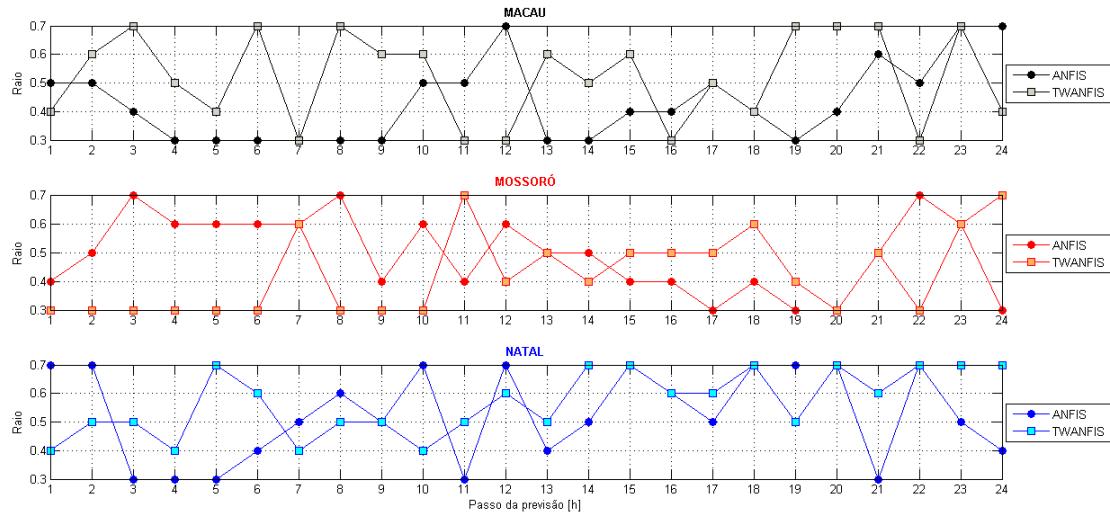


Figura 5.12 – Raios dos clusters para os modelos ANFIS e TWANFIS.

Os tempos gastos para realizar o treinamento dos modelos *ANFIS* e *TWANFIS* (para alguns passos) são apresentados nas Figura 5.13 e Figura 5.14. Os tempos gastos pelos modelos *ANFIS* são apresentados na Figura 5.13. Observa-se que o tempo gasto diminui com o aumento do raio. Para a estação de NATAL, os tempos gastos durante o treinamento com os raios 0,3 e 0,4 foram consideravelmente maiores.

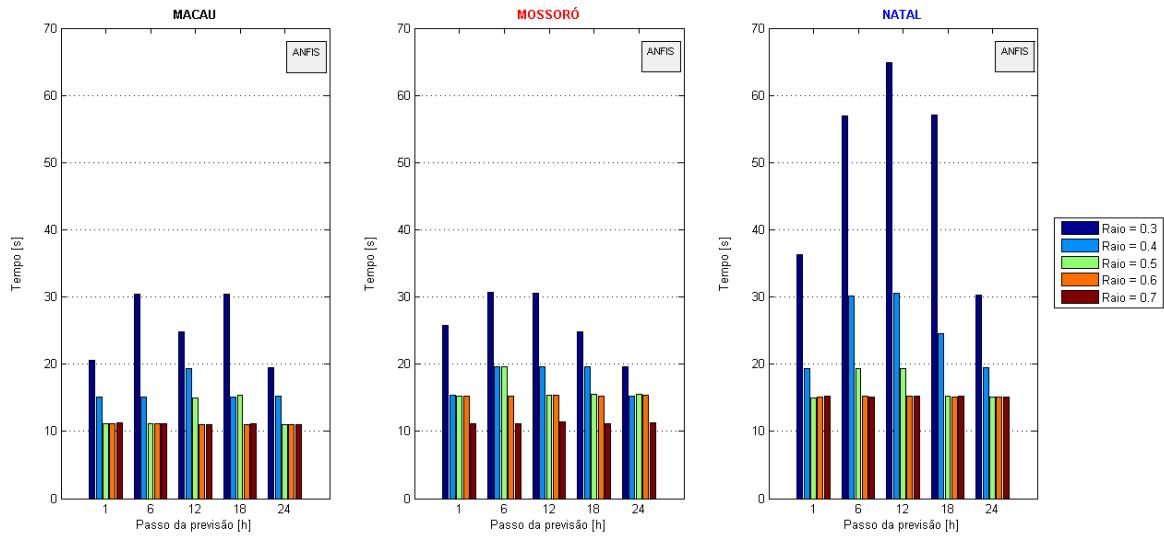


Figura 5.13 – Tempos de treinamento dos modelos ANFIS.

Na Figura 5.14, são apresentados os tempos gastos pelos modelos TWANFIS. Para estes modelos, os tempos sofreram um acréscimo significativo. Novamente, observa-se que o tempo gasto diminui com o aumento do raio. Para a estação de NATAL, os tempos gastos durante o treinamento com os raios 0,3 foram bastante elevados, chegando a atingir cerca de 8.200 segundos para o passo de previsão igual 12 horas.

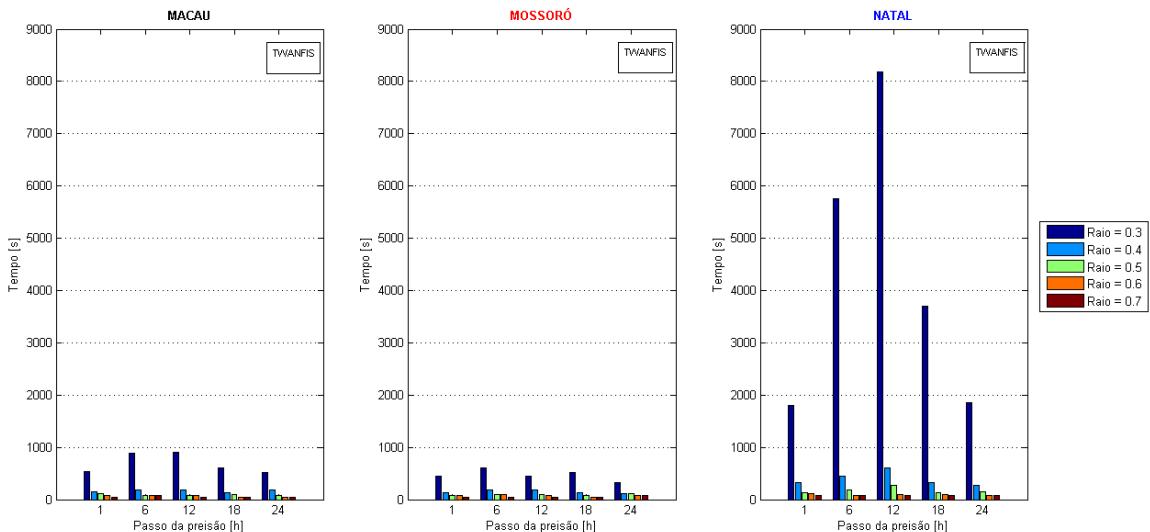


Figura 5.14 – Tempos de treinamento dos modelos TWANFIS.

Os gráficos do MAPE do conjunto de teste para os melhores modelos *ANFIS* e *TWANFIS* de MACAU, MOSSORÓ e NATAL são apresentados na Figura 5.15 para efeito de comparação. Observa-se que o desempenho dos modelos que utilizam a Transformada *Wavelet* é superior, principalmente para os passos de previsão mais curtos. Os valores do MAPE para MOSSORÓ são mais elevados, devido principalmente ao fato de que há um grande número de velocidades baixas para esta estação, o que contribui para o aumento do MAPE.

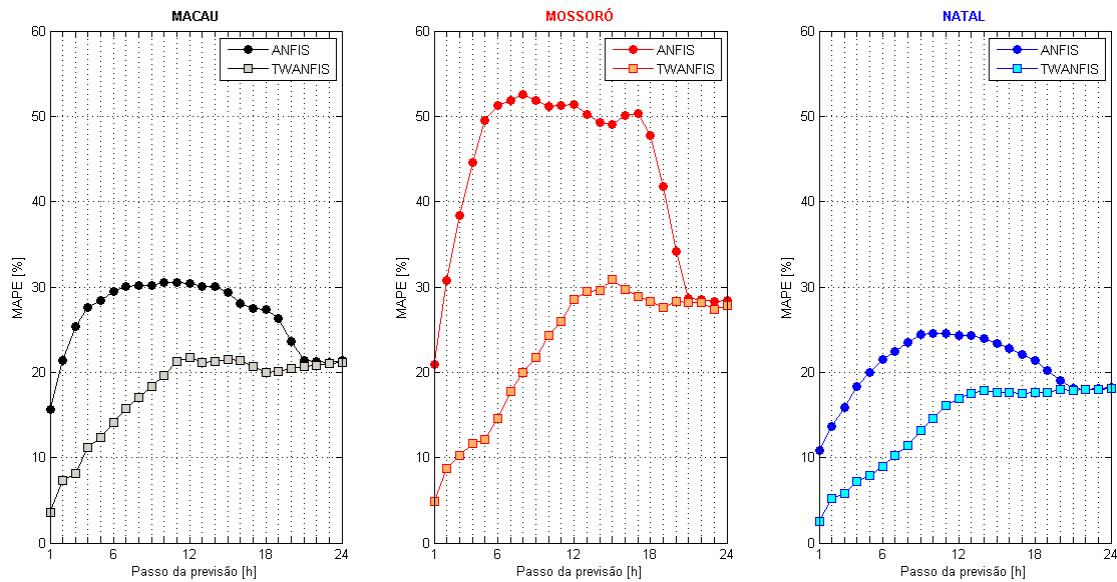


Figura 5.15 – MAPE do conjunto de teste para os melhores Sistemas de Inferência Fuzzy.

Como foi visto neste capítulo, os modelos de previsão propostos apresentaram bons desempenhos de previsão, principalmente para os passos mais curtos. No próximo capítulo serão apresentadas as comparações entre as previsões obtidas com os modelos *TWRNA(LM)* e *TWANFIS*, com as previsões obtidas com os modelos de referência apresentados no Capítulo 4. Além das previsões das velocidades, serão apresentadas as previsões de geração obtidas através da curva de potência do aerogerador.

CAPÍTULO 6

6. PREVISÕES E COMPARAÇÕES ENTRE OS MODELOS

Comparações dos desempenhos dos modelos *TWRNA(LM)* e *TWANFIS* com os desempenhos dos modelos de referência *PERSISTÊNCIA* e *NEWREF* para a estação de MACAU serão apresentadas agora. Os modelos *TWRNA(LM)* e *TWANFIS* foram escolhidos por terem apresentado, na maioria dos casos, as menores médias do MAPE para os conjuntos de teste.

O critério de seleção do período de previsão foi baseado na escolha de meses consecutivos que não apresentavam falhas nos dados de velocidades. O período escolhido para a avaliação dos modelos é formado pelas 8.016 velocidades médias horárias correspondentes à hora zero (UTC) do dia 01/01/2009 até a hora 23 (UTC) do dia 30/11/2009.

6.1 Previsões de velocidades para MACAU

A Figura 6.1 é o gráfico dos diferentes valores do MAE para os passos de previsão variando entre 1 e 24 horas.

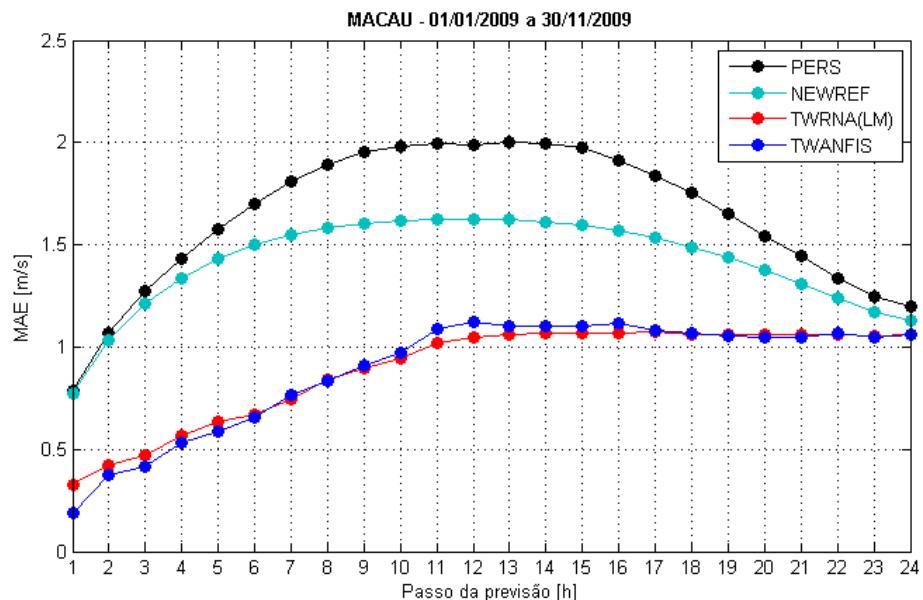


Figura 6.1 – MAE das previsões de velocidades em MACAU.

As Figura 6.2 e Figura 6.3 são dos valores do RMSE e MAPE para os passos de previsão variando de 1 a 24 horas.

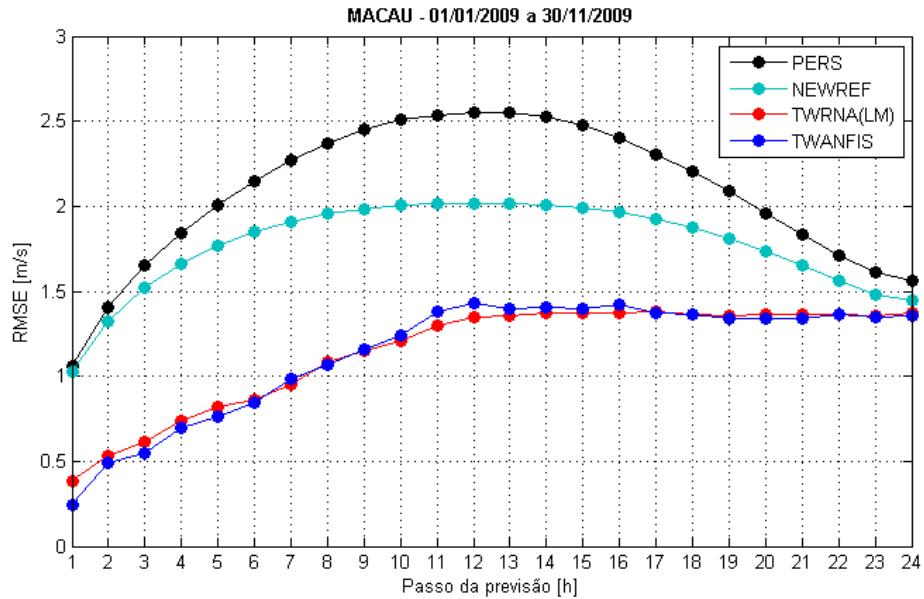


Figura 6.2 – RMSE das previsões de velocidades em MACAU.

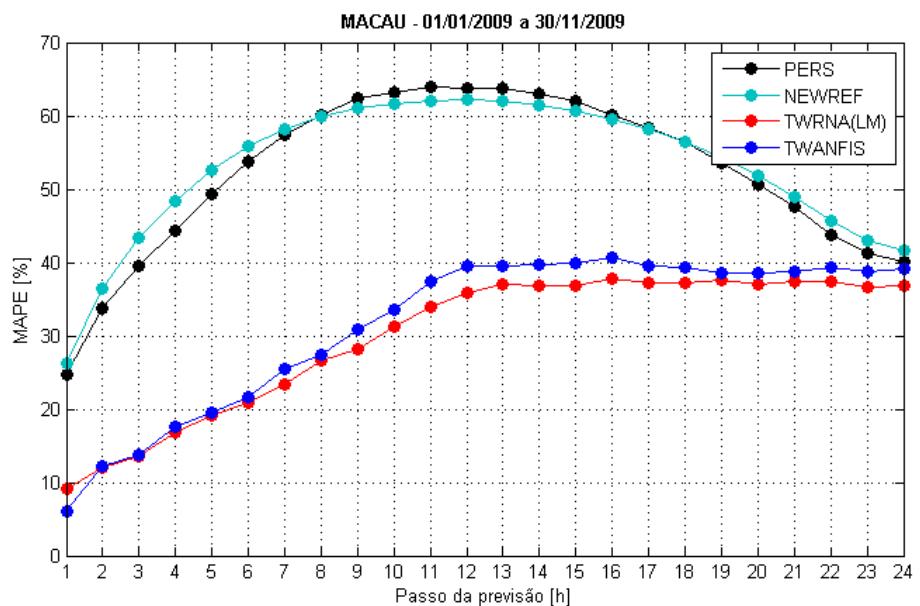


Figura 6.3 – MAPE das previsões de velocidades em MACAU.

Observa-se, nos gráficos do MAE, RMSE e MAPE, que os modelos *TWRNA(LM)* e *TWANFIS* fornecem previsões mais confiáveis do que os modelos de referência para todos

os passos de previsão considerados. Os ganhos desses dois modelos em relação ao modelo da *PERSISTÊNCIA* ficam ainda mais evidentes na Figura 6.4, bem como os ganhos do modelo *NEWREF*. No sentido da esquerda para a direita, podem ser observados os ganhos do MAE, RMSE e MAPE, para as previsões de velocidades do vento em MACAU. Os ganhos obtidos com os modelos propostos são consideráveis e, além disto, são maiores do que aqueles obtidos com o modelo *NEWREF* para todos os passos de previsão. Observa-se, no gráfico do $G_{PERS,MAPE}$, que apenas para passos entre 8 e 17 horas o modelo *NEWREF* possui um desempenho superior ao do modelo da *PERSISTÊNCIA*.

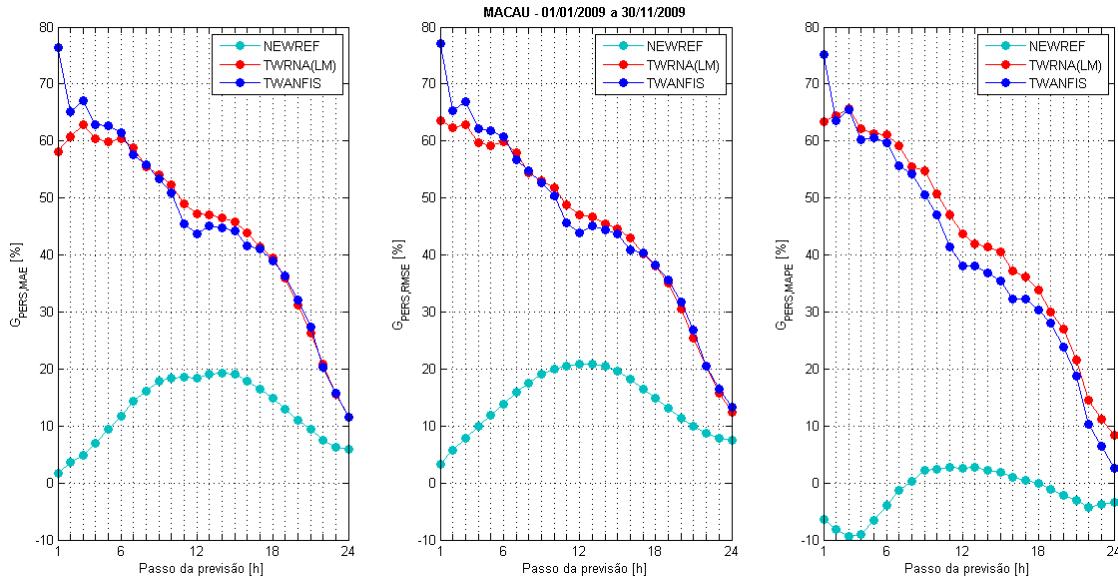


Figura 6.4 – Ganhos das previsões de velocidades em MACAU.

Os gráficos dos erros quadráticos acumulados para as previsões com passos iguais a uma hora, doze horas e vinte e quatro horas, respectivamente, são apresentados nas Figura 6.5, Figura 6.6 e Figura 6.7. A análise destes gráficos é bastante importante, pois pode-se visualizar de forma clara a evolução das previsões ao longo do tempo para um determinado modelo de previsão, ou seja, os gráficos dos erros quadráticos acumulados permitem que se analise de forma qualitativa o comportamento dos modelos de previsão. A inclinação (tendência) destes gráficos permite identificar a qualidade das previsões. Quanto menor for a inclinação, melhores serão as previsões obtidas com um determinado modelo de previsão ao longo de determinado período. Observa-se, na Figura 6.6, que os erros de previsão obtidos para o passo de 12 horas são mais acentuados nos períodos de fortes ventos (faixas

nas quais a inclinação dos gráficos é mais acentuada). Este comportamento é observado mais claramente nos modelos de referência.

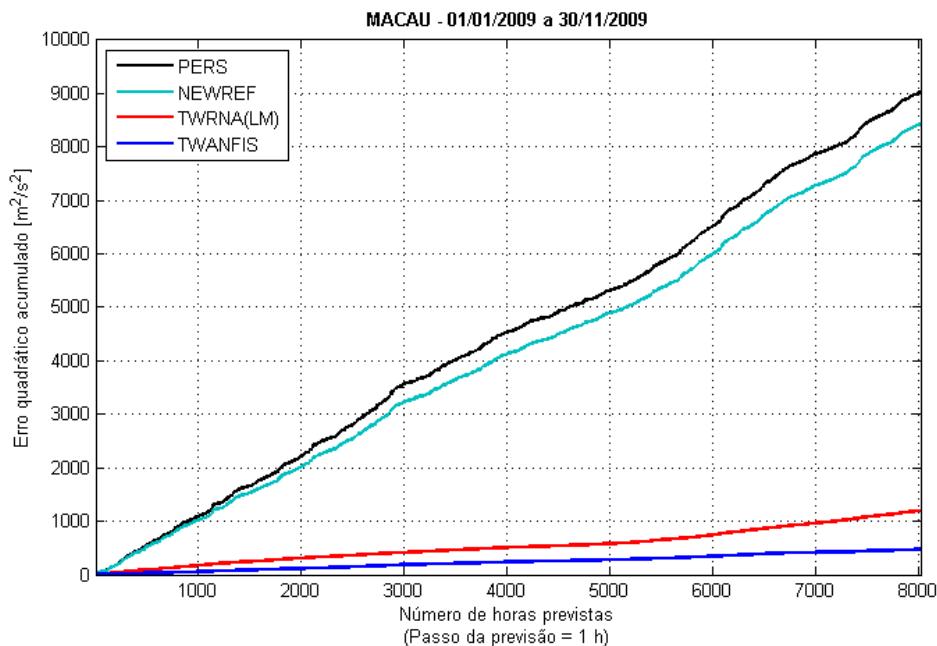


Figura 6.5 – Erros quadráticos acumulados das previsões de velocidades em MACAU com passo de 1 hora.

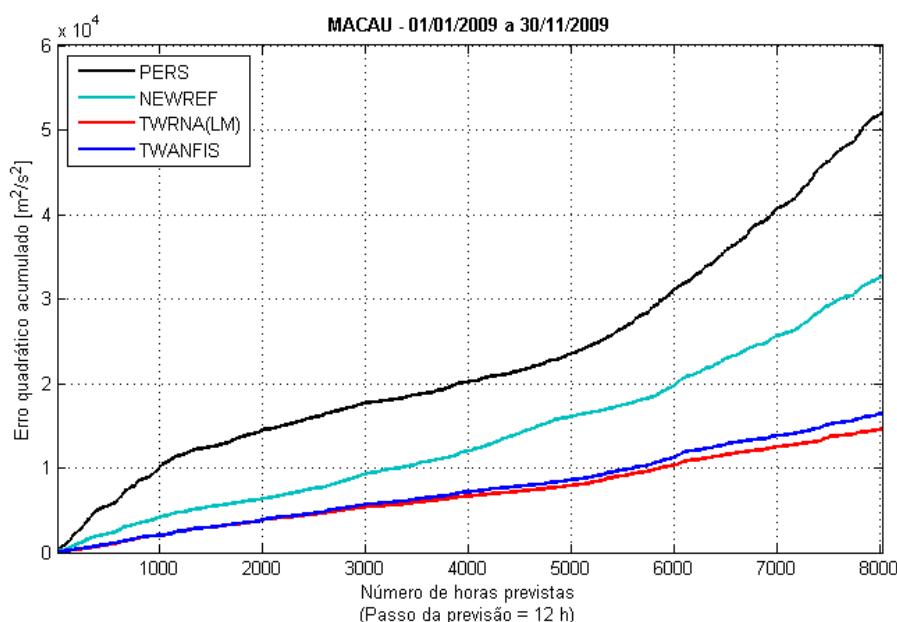


Figura 6.6 – Erros quadráticos acumulados das previsões de velocidades em MACAU com passo de 12 horas.

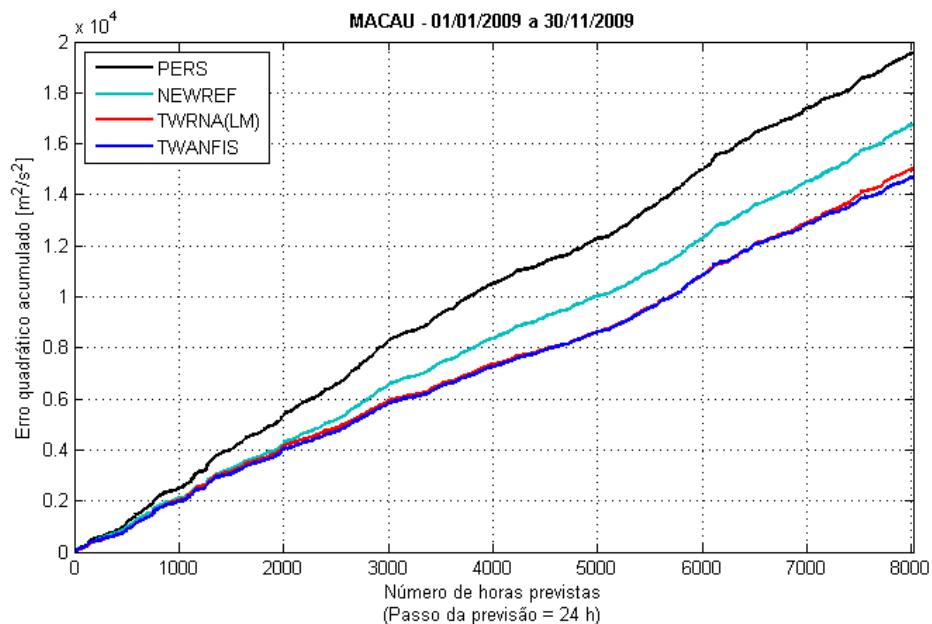


Figura 6.7 – Erros quadráticos acumulados das previsões de velocidades em MACAU com passo de 24 horas.

Os gráficos dos erros quadráticos acumulados confirmam que, para MACAU, os modelos *TWRNA(LM)* e *TWANFIS* fornecem, em média, previsões de velocidades de ventos muito mais confiáveis do que aquelas obtidas com os modelos de referência, principalmente as previsões com passos mais curtos.

Os diagramas de dispersão para os passos de previsão iguais a uma hora, doze horas e vinte e quatro horas são apresentados nas Figura 6.8, Figura 6.9 e Figura 6.10, respectivamente. Ao realizar uma análise visual destes diagramas (análise qualitativa), deve-se observar o quanto próximos da reta estão os pontos. Quanto mais próximos da reta estiverem estes pontos, melhor será o modelo de previsão. Juntamente com os diagramas de dispersão, são apresentados também os coeficientes de determinação R^2 para cada um dos modelos. O coeficiente de determinação é um indicador quantitativo da qualidade do modelo de previsão (quanto mais próximo da unidade, melhor a qualidade do modelo).

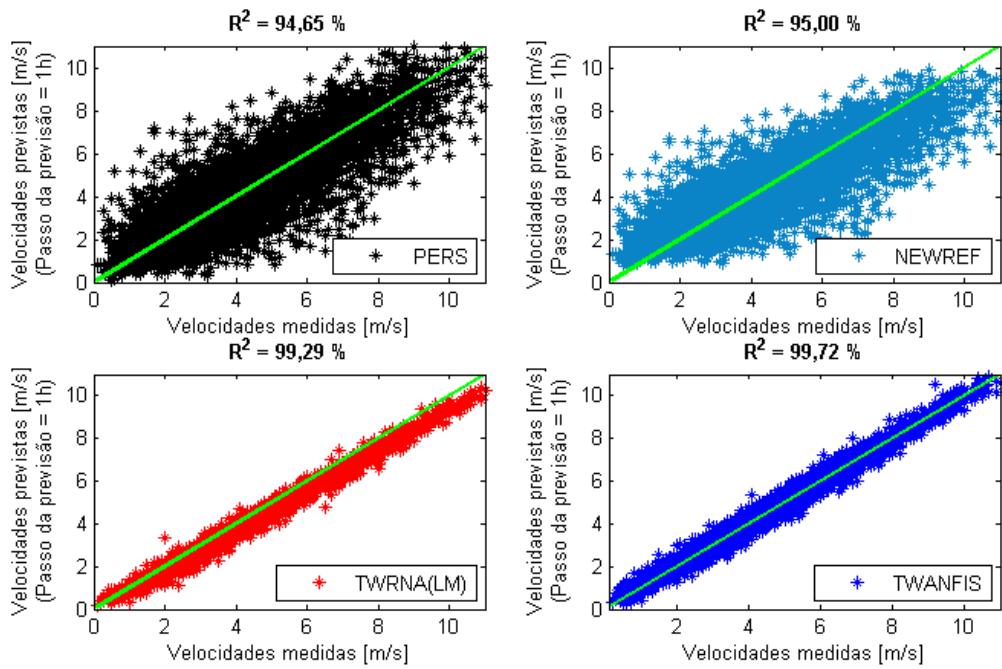


Figura 6.8 – Diagramas de dispersão das previsões de velocidades em MACAU com passo de 1 hora.

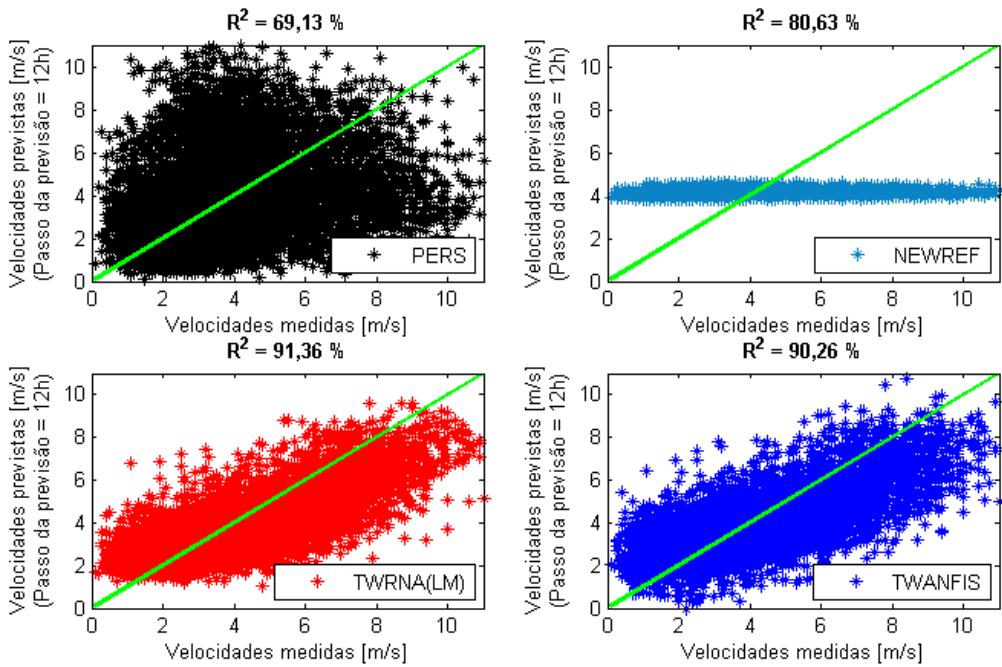


Figura 6.9 – Diagramas de dispersão das previsões de velocidades em MACAU com passo de 12 horas.

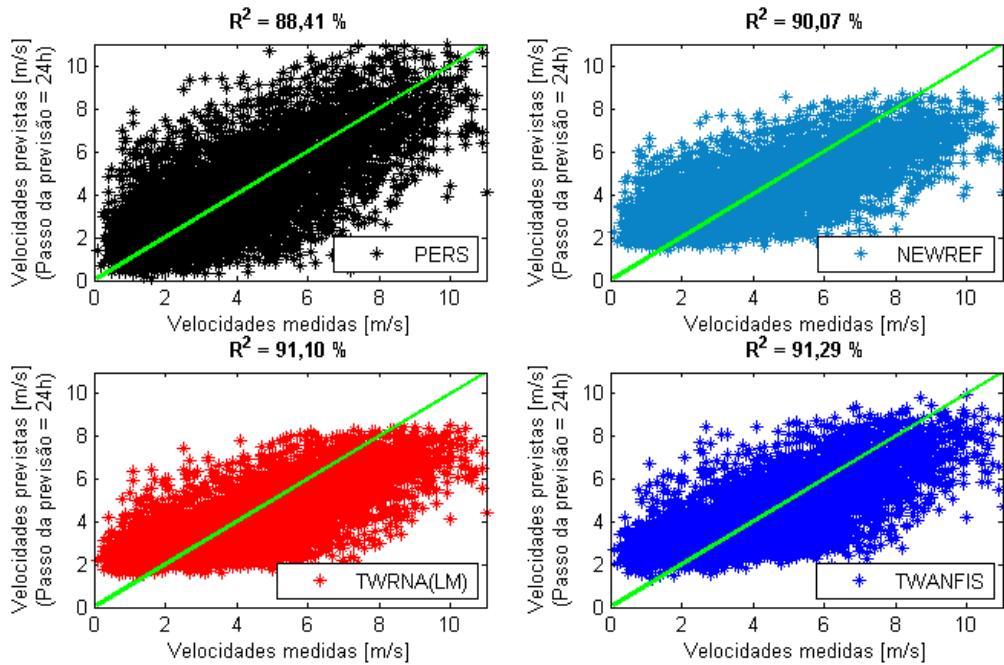


Figura 6.10 – Diagramas de dispersão das previsões de velocidades em MACAU com passo de 24 horas.

Os diagramas de dispersão e os coeficientes de determinação apresentados para as previsões de velocidades da estação de MACAU confirmam a qualidade dos modelos propostos, sobretudo para as previsões de velocidades com passo de uma hora. Para passos de 12 horas, observa-se que as previsões do modelo *NEWREF* são praticamente iguas à média, pois o coeficiente de autocorrelação para velocidades deslocadas de 12 horas é muito pequeno.

Todos os critérios de análises realizados para as previsões das velocidades de vento em MACAU refletiram o excelente desempenho dos modelos *TWRNA(LM)* e *TWANFIS* quando confrontados com os modelos de referência *PERSISTÊNCIA* e *NEWREF*. Observou-se que, para passos diferentes, a qualidade das previsões obtidas com os modelos de referência é fortemente influenciada pela autocorrelação da série temporal. Já para os modelos propostos, observou-se que a aplicação da Transformada *Wavelet* melhora significativamente as previsões para os passos de até 12 horas. A partir daí, pode-se dizer que o MAE, RMSE e MAPE dos previsores propostos praticamente não variam com o passo da previsão.

6.2 Previsões de geração para MACAU

A Figura 6.11 é a curva de potência utilizada para realizar as simulações da geração desta dissertação. Trata-se de um aerogerador com potência nominal de 2.300 kW e a altura do cubo é igual a 57 metros.

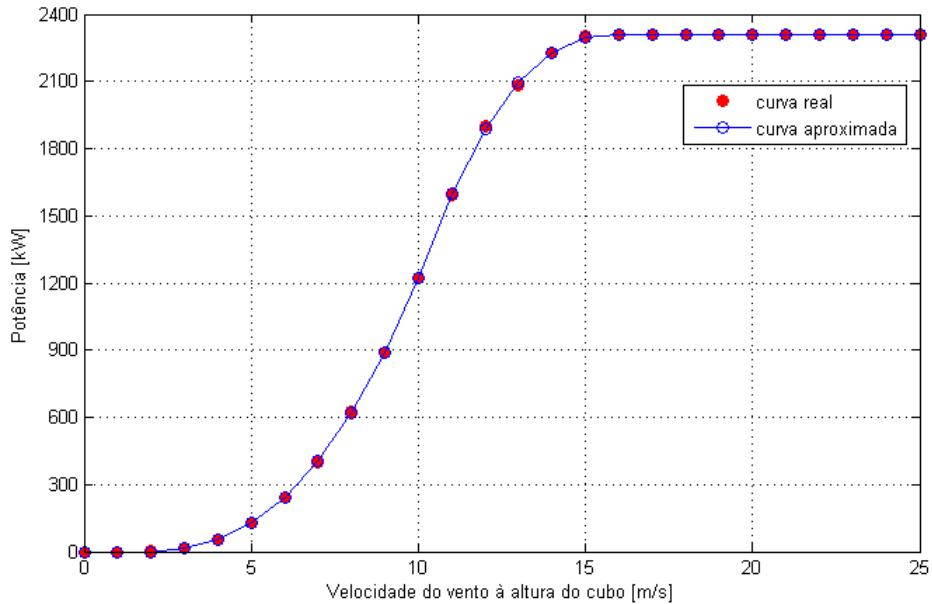


Figura 6.11 – Curva de potência do aerogerador.

Os fabricantes fornecem apenas alguns pontos da curva (pontos em vermelho), logo, para estimar a potência (em kW) fornecida para qualquer velocidade compreendida entre as velocidade de *cut-in* (igual a 2 m/s para o aerogerador considerado) e *cut-out* (igual a 25 m/s para o aerogerador considerado), realizou-se a parametrização da curva em quatro intervalos através da minimização do MSE (método dos mínimos quadrados). A parametrização utilizada é definida em (6.1).

$$\begin{aligned}
 P(v) &= 0 \quad \text{para } v \leq 2 \text{ m/s}; \\
 P(v) &= 0,93v^3 + 6,42v^2 - 40,18v + 51,89 \quad \text{para } 2 \text{ m/s} < v < 10 \text{ m/s}; \\
 P(v) &= 1,31v^3 - 87,17v^2 + 1774,24v - 9109,71 \quad \text{para } 10 \text{ m/s} \leq v < 16 \text{ m/s}; \\
 P(v) &= 2310 \quad \text{para } v \geq 16 \text{ m/s}.
 \end{aligned} \tag{6.1}$$

As velocidades previstas são referentes à altura de 10 metros, logo, para realizar as previsões de geração foi necessário utilizar (2.12) para converter as velocidades para a

altura do cubo do aerogerador, que é igual a 57 metros. O coeficiente de rugosidade do terreno (α) utilizado foi igual a 0,10.

Com exceção do MAPE, os mesmos critérios utilizados para a análise do desempenho das previsões de velocidades também são utilizados para as previsões de geração obtidas através da curva de potência do aerogerador. As Figura 6.12 e Figura 6.13 são do MAE e o RMSE das previsões de geração, respectivamente.

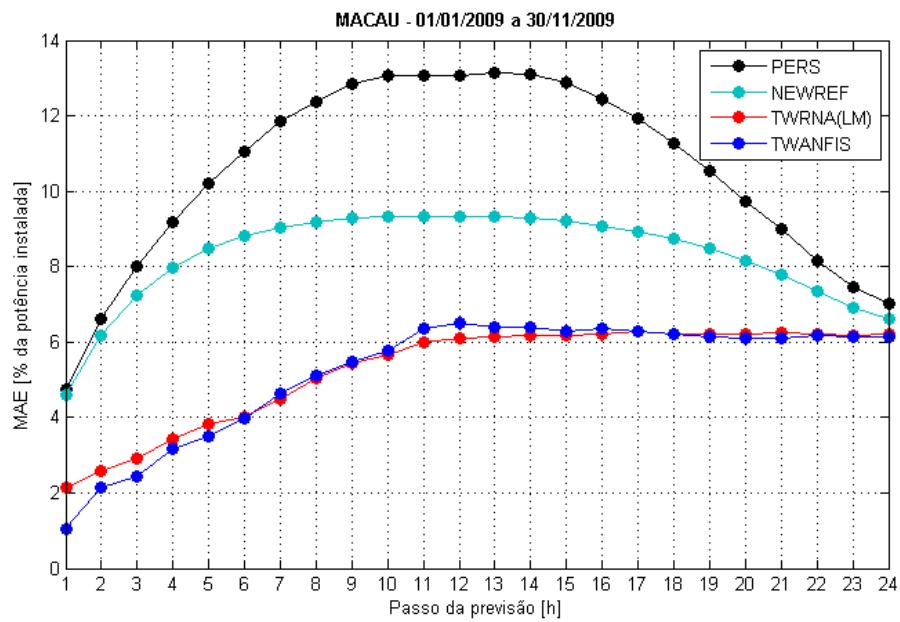


Figura 6.12 – MAE das previsões de geração de MACAU.

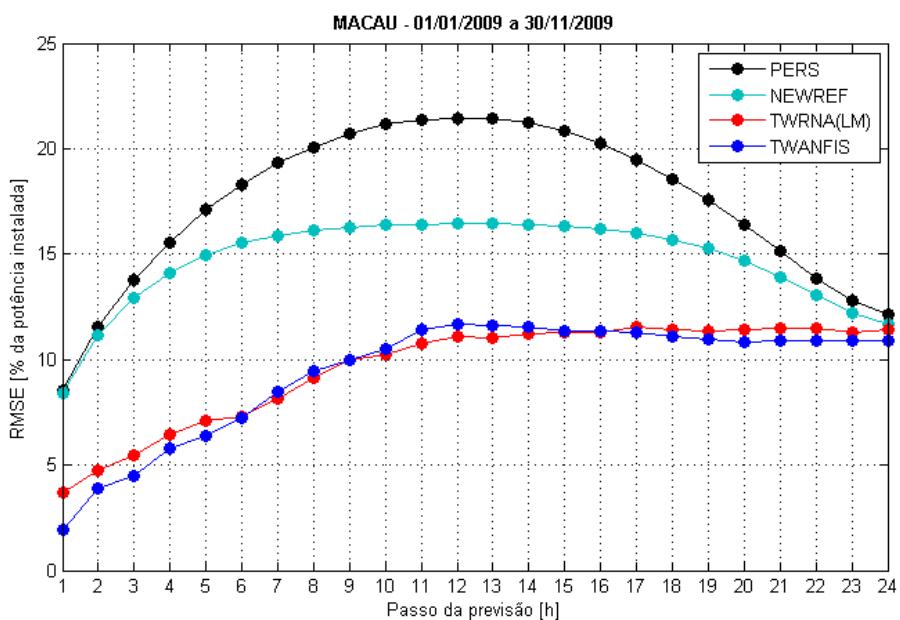


Figura 6.13 – RMSE das previsões de geração de MACAU.

Na Figura 6.14, são apresentados os ganhos das gerações obtidas através da curva de potência utilizando-se as velocidades previstas (convertidas para a altura do cubo do aerogerador) com os modelos *TWRNA(LM)*, *TWANFIS* e *NEWREF*, em relação ao modelo da *PERSISTÊNCIA*.

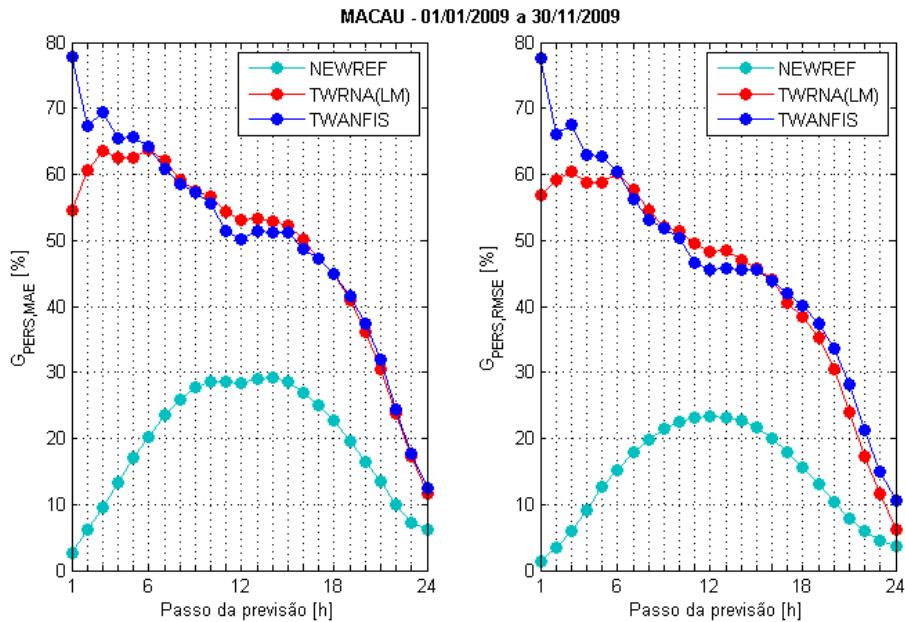


Figura 6.14 – Ganhos das previsões de geração de MACAU.

Os gráficos dos erros quadráticos acumulados para as previsões de geração com passo de uma hora são apresentados na Figura 6.15.

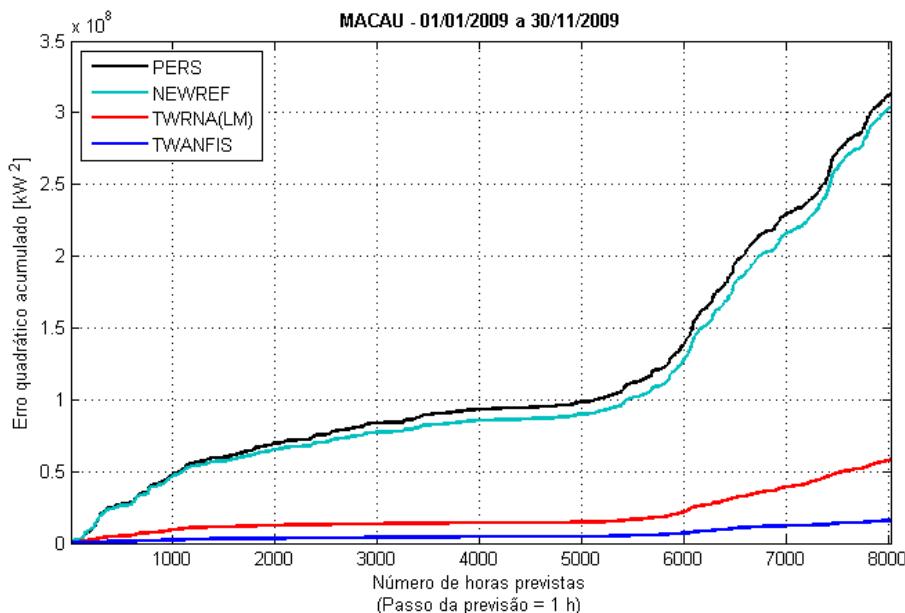


Figura 6.15 – Erros quadráticos acumulados das previsões de geração de MACAU com passo de 1 hora.

Nas Figura 6.16 e Figura 6.17 apresentam-se os gráficos dos erros quadráticos acumulados para as previsões de geração com passos 12 e 24 horas, respectivamente. Os gráficos dos erros quadráticos acumulados apresentam claramente o efeito que os erros de previsão de velocidades causam nas previsões de geração (aumento na inclinação dos gráficos), principalmente nos meses em que a média das velocidades é maior.

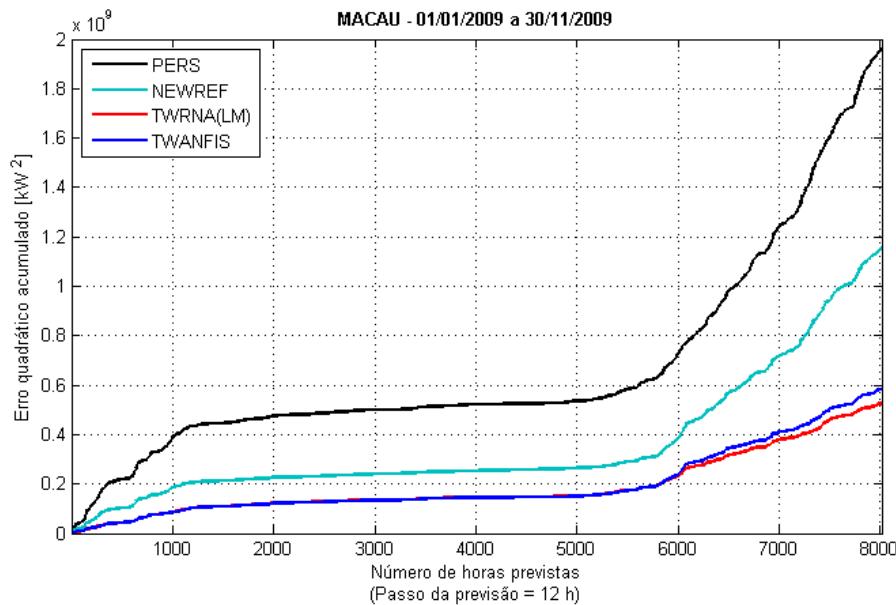


Figura 6.16 – Erros quadráticos acumulados das previsões de geração de MACAU com passo de 12 horas.

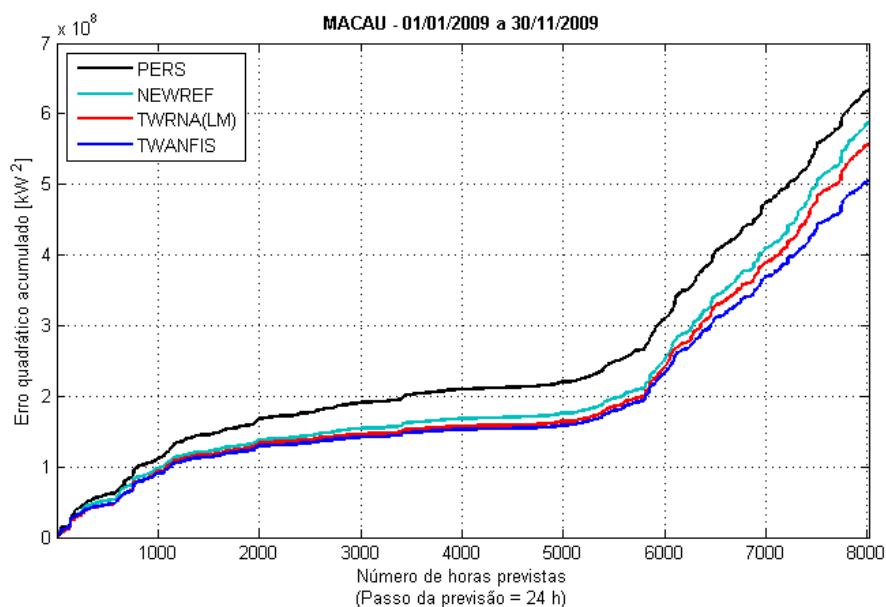


Figura 6.17 – Erros quadráticos acumulados das previsões de geração de MACAU com passo de 24 horas.

Os diagramas de dispersão e os coeficientes de determinação R^2 para as previsões de geração com passos iguais a 1, 12 e 24 horas são apresentados nas Figura 6.18, Figura 6.19 e Figura 6.20, respectivamente. Nestes diagramas, o eixo horizontal representa as gerações obtidas através da curva de potência utilizando-se as velocidades medidas.

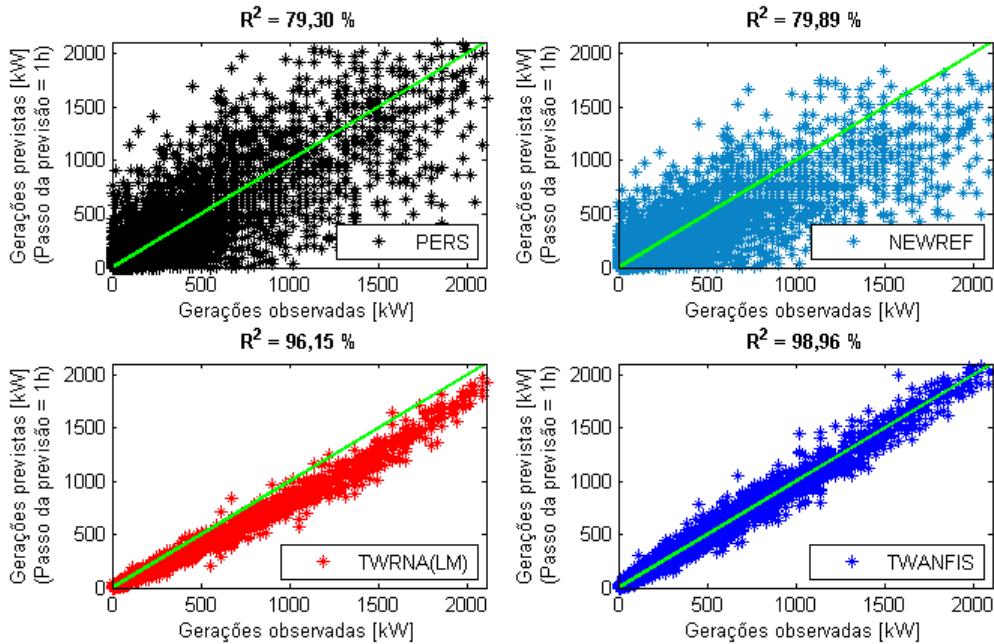


Figura 6.18 – Diagramas de dispersão das previsões de geração de MACAU com passo de 1 hora.

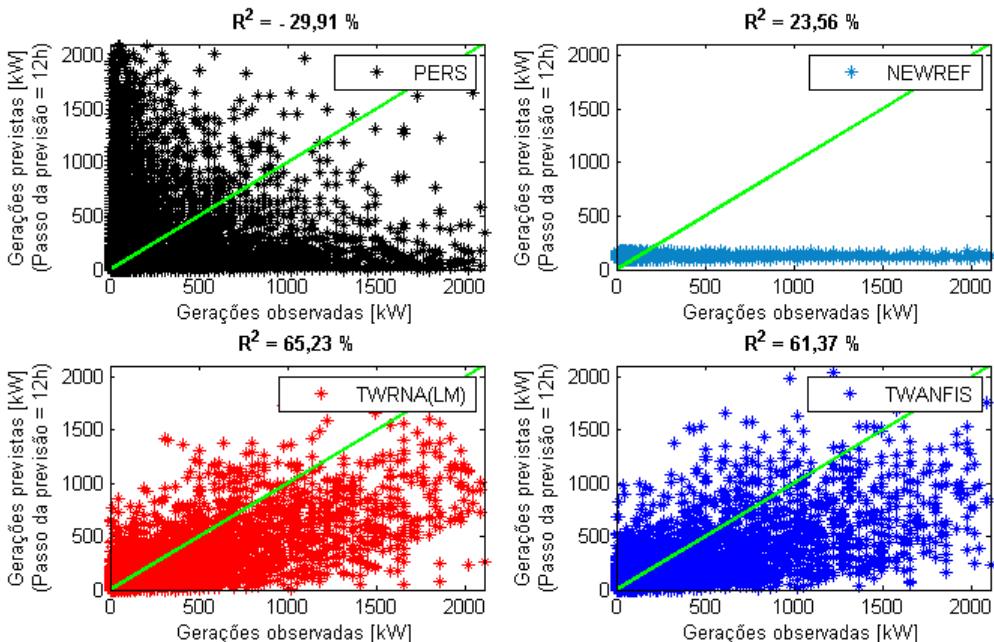


Figura 6.19 – Diagramas de dispersão das previsões de geração de MACAU com passo de 12 horas.

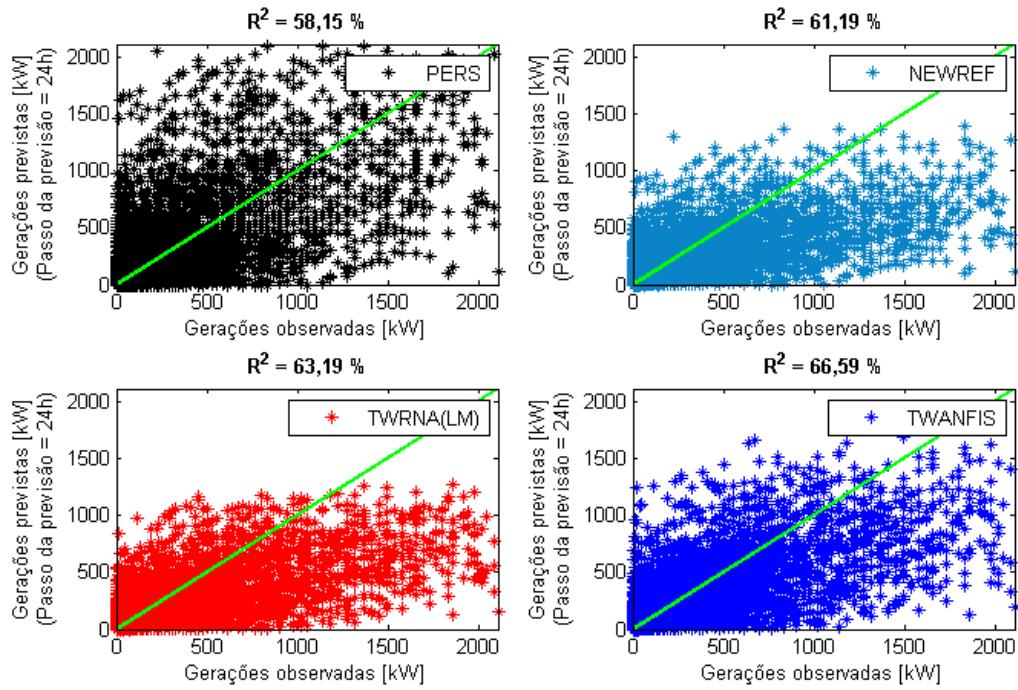


Figura 6.20 – Diagramas de dispersão das previsões de geração de MACAU com passo de 24 horas.

O coeficiente de determinação para as previsões de 12 horas com o modelo da *PERSISTÊNCIA* possui um valor negativo (ver Figura 6.19). Isto acontece quando este modelo é utilizado para realizar previsões com passos para os quais a autocorrelação da série temporal seja pequena. Nestes casos, a variância dos erros de previsão para o modelo da *PERSISTÊNCIA* se torna maior do que a média global dos dados observados.

As frequências percentuais dos erros de geração obtidos com os quatro modelos considerados neste capítulo são apresentadas na Figura 6.21. No gráfico à esquerda, pode-se visualizar o percentual dos erros de previsão positivos (quando o valor real é maior do que o valor previsto). Observa-se que a maioria dos erros das previsões obtidas com o modelo *TWRNA(LM)* são sempre positivos para todos os passos considerados, pois a frequência percentual é maior do que 50%. O comportamento inverso é observado para as previsões do modelo *NEWREF*. Para o modelo da *PERSISTÊNCIA*, a maioria dos erros de previsão são negativos, exceto para os passos iguais a 8 e 9 horas. Para o modelo *TWANFIS*, a maioria dos erros são negativos para os passos entre 8 e 13 horas. No gráfico à direita, observa-se a frequência dos erros de previsão nulos (previsões exatas). Apesar do modelo da *PERSISTÊNCIA* realizar um maior número de previsões exatas para passos maiores, este não é o melhor modelo de previsão.

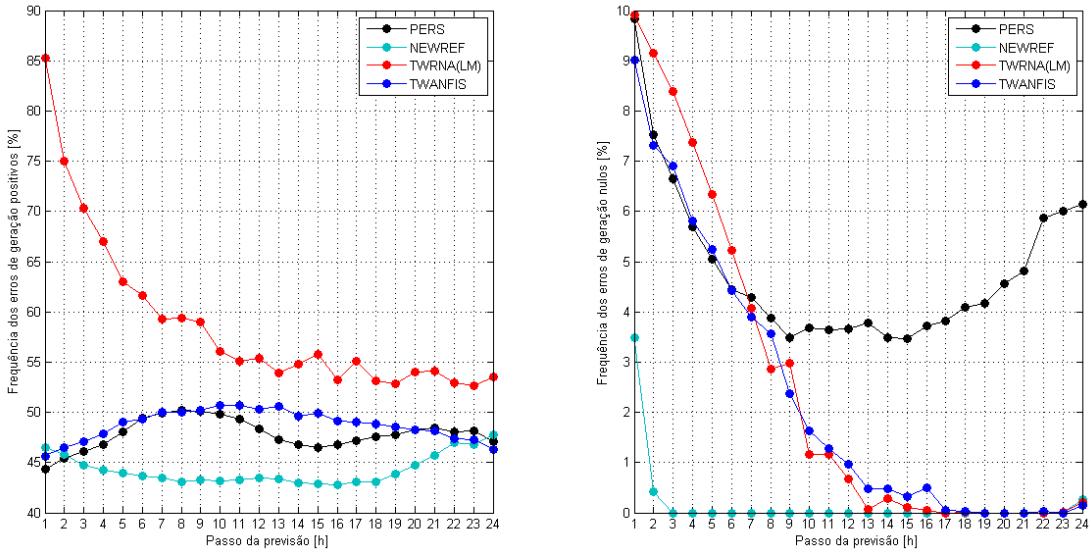
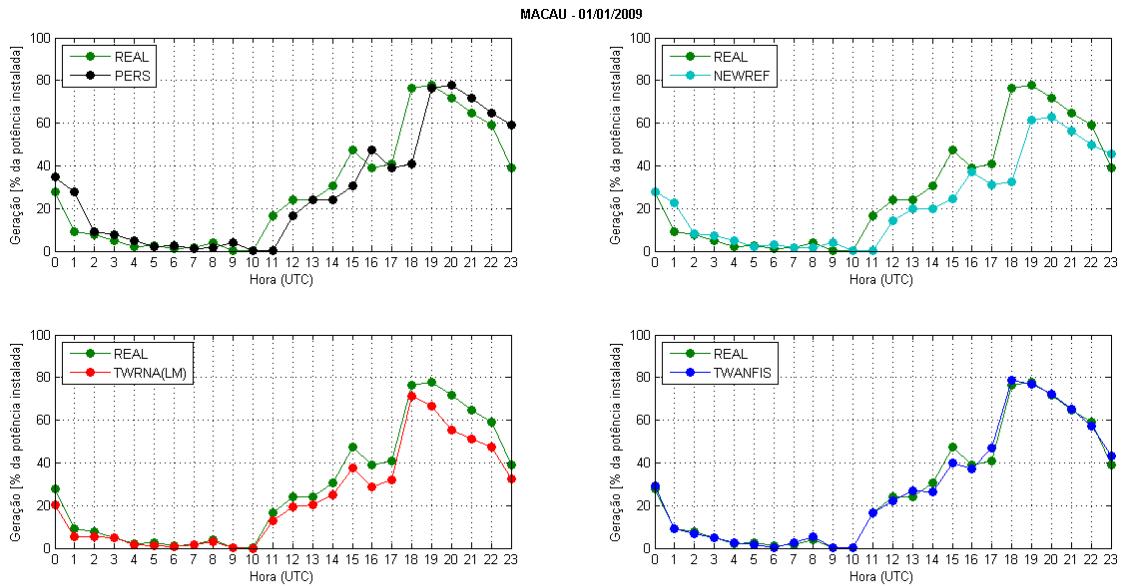
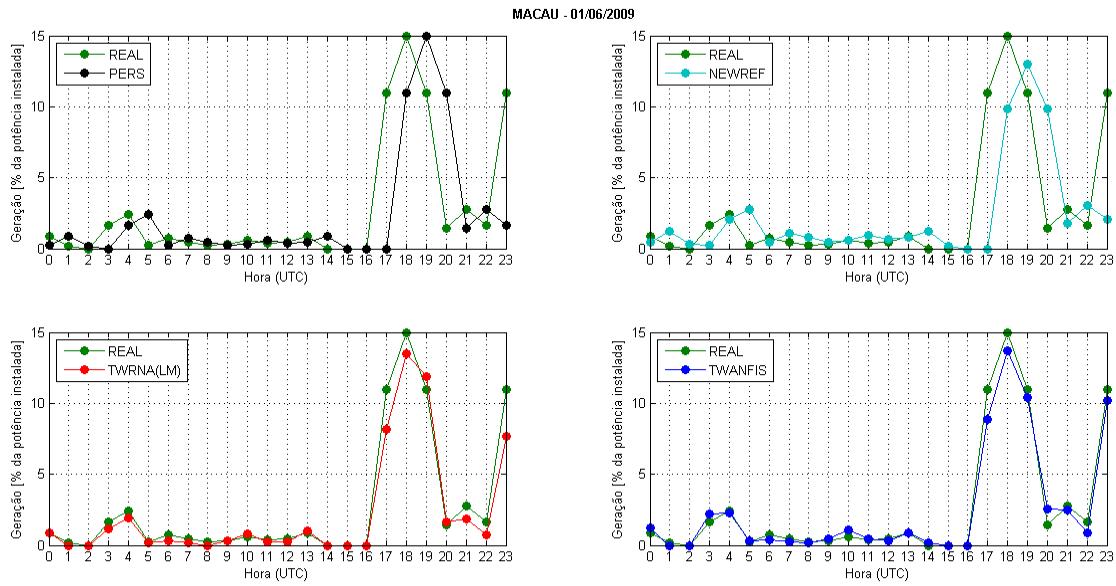


Figura 6.21 – Frequência percentual dos erros de geração de MACAU.

Para finalizar o capítulo, são apresentadas, nas Figura 6.22 e Figura 6.23, as curvas de geração obtidas a partir das previsões com passo de uma hora para dois dias distintos. Os gráficos da Figura 6.22 são referentes ao dia 01/01/2009, e os gráficos da Figura 6.23, referentes ao dia 01/06/2009. Optou-se pela escolha destes dois dias porque o primeiro deles se encaixa no período de fortes ventos, já o segundo, se encaixa no período de ventos mais fracos.



**Figura 6.22 – Previsões de geração de MACAU para o dia 01/01/2009.
(Passo de 1 hora).**



**Figura 6.23 – Previsões de geração de MACAU para o dia 01/06/2009.
(Passo de 1 hora).**

Todas as análises realizadas neste capítulo para a MACAU também foram aplicadas a MOSSORÓ e NATAL. As figuras apresentando os resultados das análises para estas duas localidades podem ser visualizadas no Apêndice A.

CAPÍTULO 7

7. CONCLUSÕES E PROPOSTAS PARA TRABALHOS FUTUROS

Neste capítulo são apresentadas as conclusões gerais do trabalho realizado, incluindo comentários sobre os modelos desenvolvidos e os resultados obtidos. Além disto, são identificados alguns aspectos relevantes que podem originar novos trabalhos de pesquisa.

7.1 Conclusões

Os modelos desenvolvidos basearam-se na análise de séries temporais de velocidades de vento através da inteligência computacional. Os horizontes de previsão apresentados enquadram-se no horizonte temporal de curto prazo, para previsões de até vinte e quatro horas e uma discretização em intervalos de uma hora. Entende-se que este é um horizonte adequado para subsidiar o planejamento da operação dos sistemas hidrotérmico e eólico, uma vez que a entrada em operação de uma usina termoelétrica precisa ser definida com antecedência e o tempo de partida varia de uma planta para outra.

Boas previsões foram obtidas com os modelos desenvolvidos para todos os 24 passos de previsão considerados, principalmente para aqueles passos mais curtos. Verificou-se que a qualidade das previsões é fortemente influenciada pela autocorrelação das séries temporais, tanto para os modelos de referência adotados quanto para os modelos que não utilizam as *wavelets*.

A decomposição das séries de velocidades empregando *wavelets* possibilitou a extração de informações relevantes sobre o comportamento cíclico e sazonal das velocidades dos ventos. Estas informações contidas nos sinais de aproximação e detalhes foram decisivas para a melhoria significativa das previsões com os modelos que utilizam estes sinais como entradas.

Com base nas análises realizadas, verificou-se que o comportamento dos erros de previsão com a variação dos passos de previsão foi bastante semelhante para MACAU, MOSSORÓ e NATAL. A metodologia adotada para o desenvolvimento dos modelos foi bastante adequada, o que garantiu previsões bastante confiáveis estatisticamente, ou seja, os modelos adquiriram capacidade de generalização sem se tornarem tendenciosos.

7.2 Propostas para trabalhos futuros

Mesmo os modelos de previsão desenvolvidos tendo dado bons resultados eles ainda podem ser aprimorados. Este processo de melhoria contínua deve ser realizado com o intuito de se obter previsões mais confiáveis, reduzindo assim os erros entre os valores previstos e reais. É fato que sempre existirá um erro entre os valores previsto e verificado, portanto, o desafio será sempre buscar minimizá-lo.

Como sugestão de trabalhos futuros voltados para o melhoramento dos modelos se apresentam:

- Em virtude do aumento dos dados relativos à geração eólica no Brasil, é oportuno desenvolver modelos de previsão que forneçam como saída a potência gerada. Desta forma, a curva de potência estaria incorporada aos parâmetros do próprio modelo (por exemplo, nos pesos sinápticos das redes neurais);
- Buscar uma base de dados mais extensa com o objetivo de desenvolver os modelos para previsões sazonais;
- Utilizar dados anemométricos medidos em alturas mais elevadas, de modo que a influência da rugosidade do terreno seja reduzida;
- Investigar a utilização de *wavelets* como funções de ativação para os neurônios das redes neurais;
- Desenvolver modelos multivariados, considerando outras séries temporais como dados de entrada;
- Estudar o funcionamento das redes neurais do tipo *Reservoir Computing* e pesquisar a possibilidade de desenvolver modelos híbridos.

APÊNDICE A

As figuras deste apêndice são referentes às previsões de velocidades de ventos e geração eólica para MOSSORÓ e NATAL. Para realizar as previsões de geração, utilizaram-se os mesmos parâmetros aplicados às gerações de MACAU, ou seja, mesma curva de potência, mesma altura do *hub* e mesmo coeficiente de rugosidade de terreno (ver Seção 6.2).

A.1 Previsões para MOSSORÓ

Para MOSSORÓ, o período escolhido para a avaliação dos modelos é formado pelas 4.416 velocidades médias horárias correspondentes à hora zero (UTC) do dia 01/03/2009 até a hora 23 (UTC) do dia 31/08/2009.

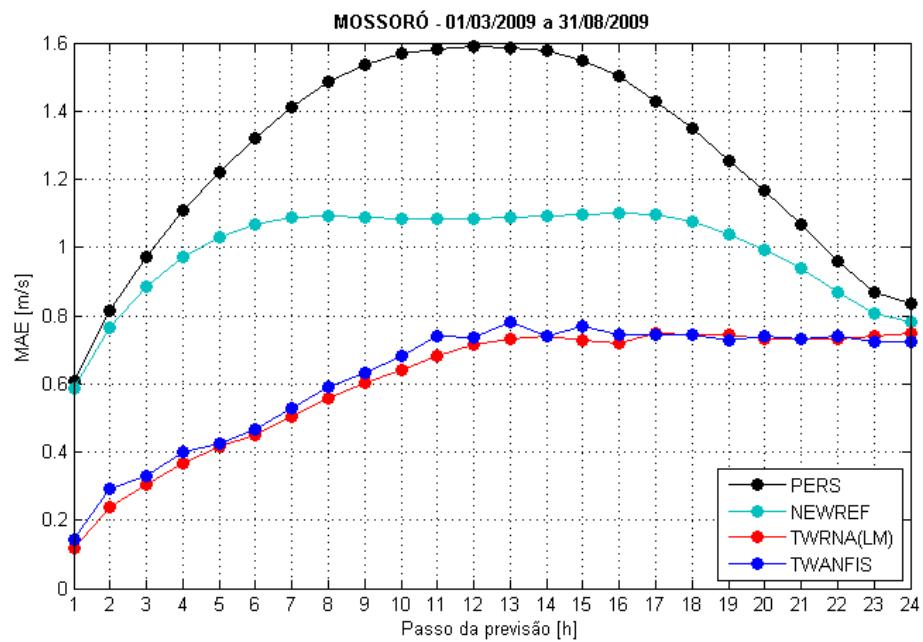


Figura A.1 – MAE das previsões de velocidades em MOSSORÓ.

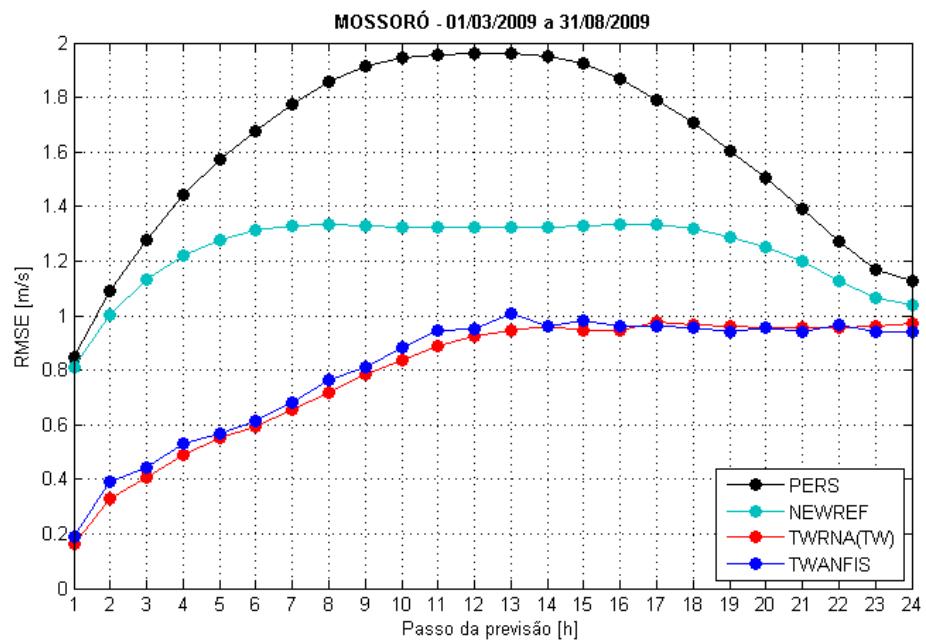


Figura A.2 – RMSE das previsões de velocidades em MOSSORÓ.

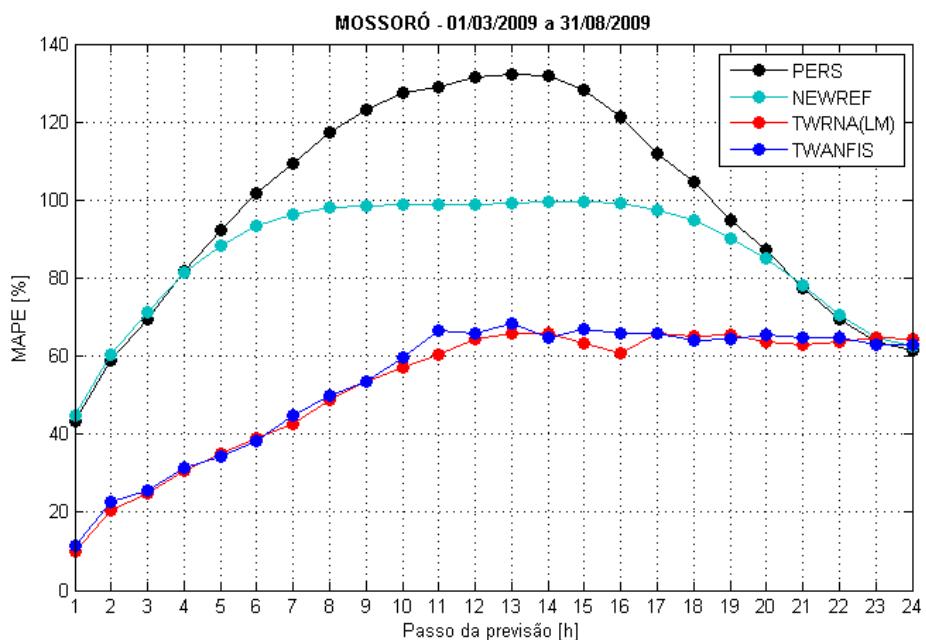


Figura A.3 – MAPE das previsões de velocidades em MOSSORÓ.

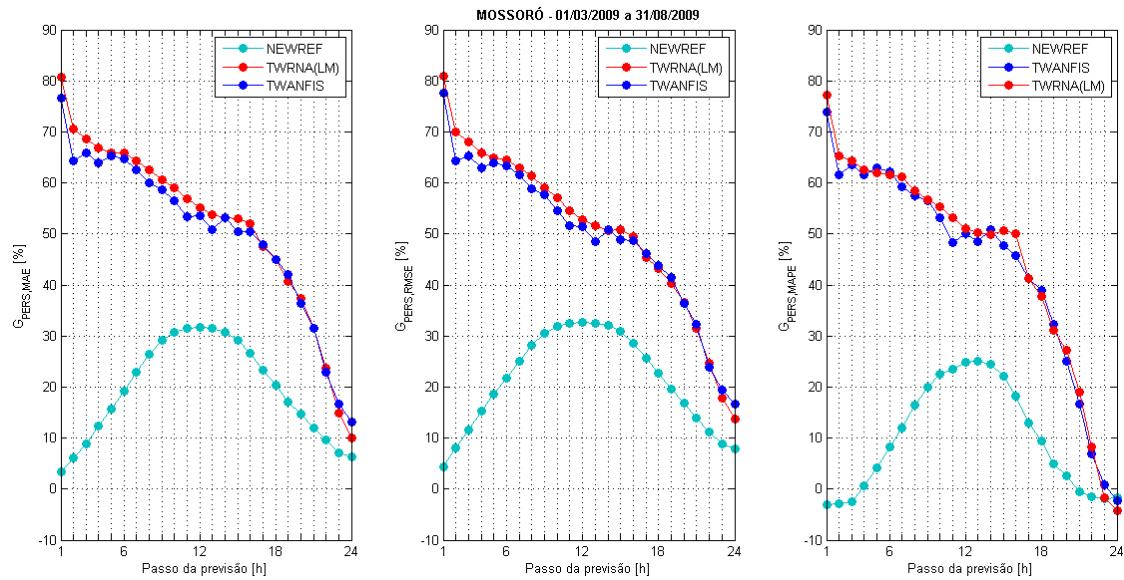


Figura A.4 – Ganhos das previsões de velocidades em MOSSORÓ.

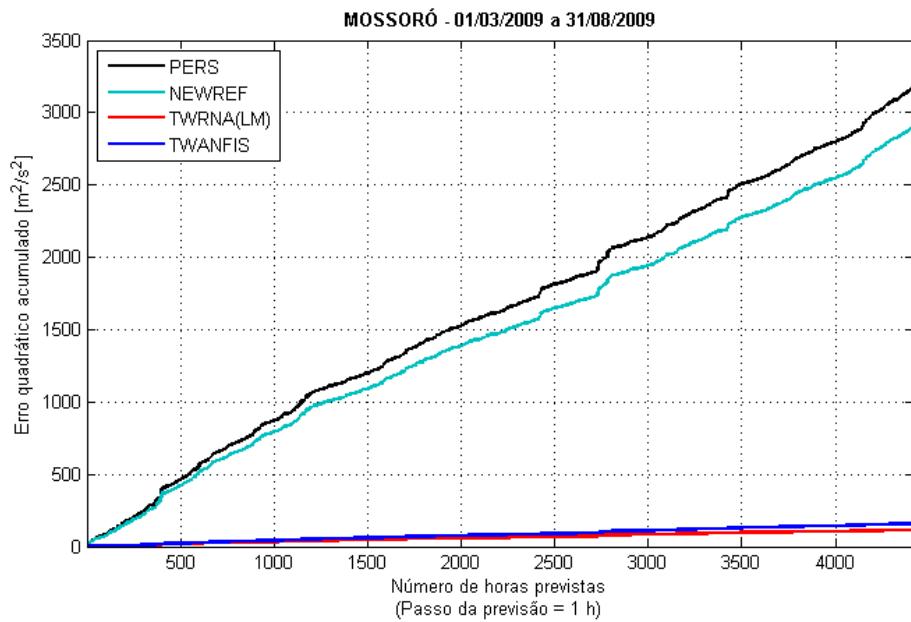


Figura A.5 – Erros quadráticos acumulados das previsões de velocidades em MOSSORÓ com passo de 1 hora.

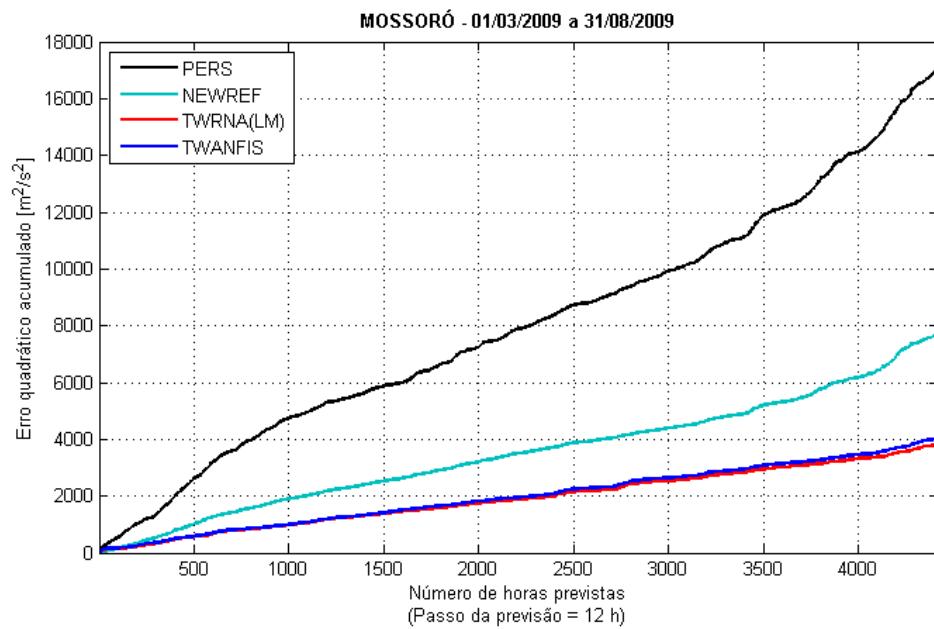


Figura A.6 – Erros quadráticos acumulados das previsões de velocidades em MOSSORÓ com passo de 12 horas.

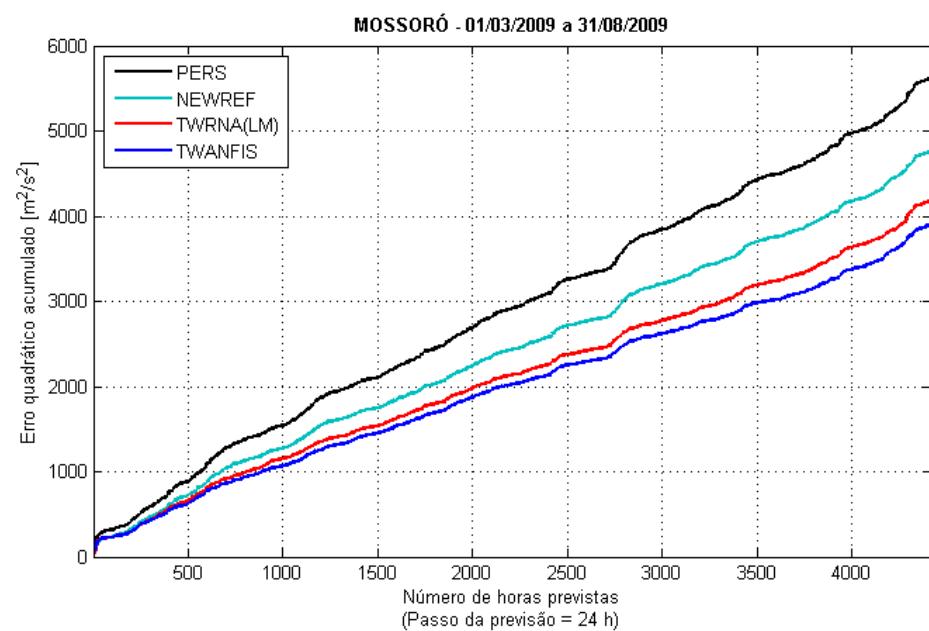


Figura A.7 – Erros quadráticos acumulados das previsões de velocidades em MOSSORÓ com passo de 24 horas.

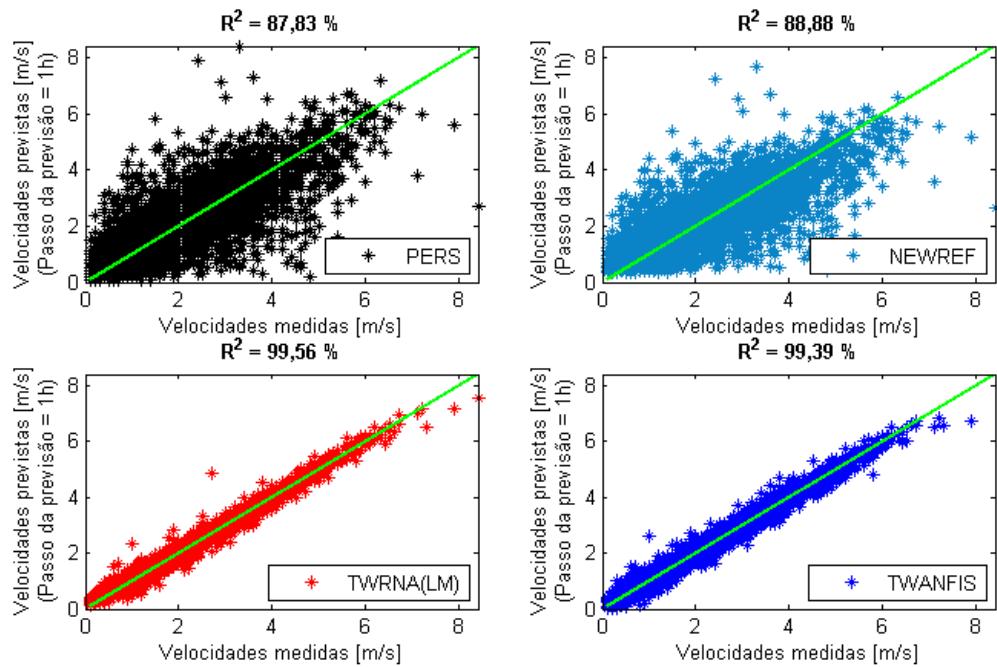


Figura A.8 – Diagramas de dispersão das previsões de velocidades em MOSSORÓ com passo de 1 hora.

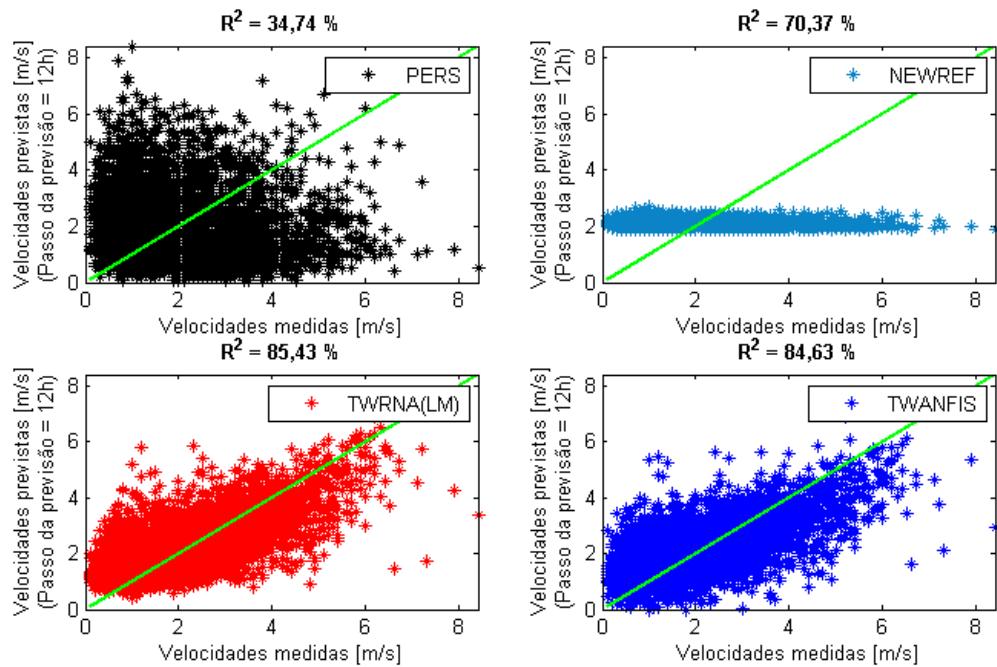


Figura A.9 – Diagramas de dispersão das previsões de velocidades em MOSSORÓ com passo de 12 horas.

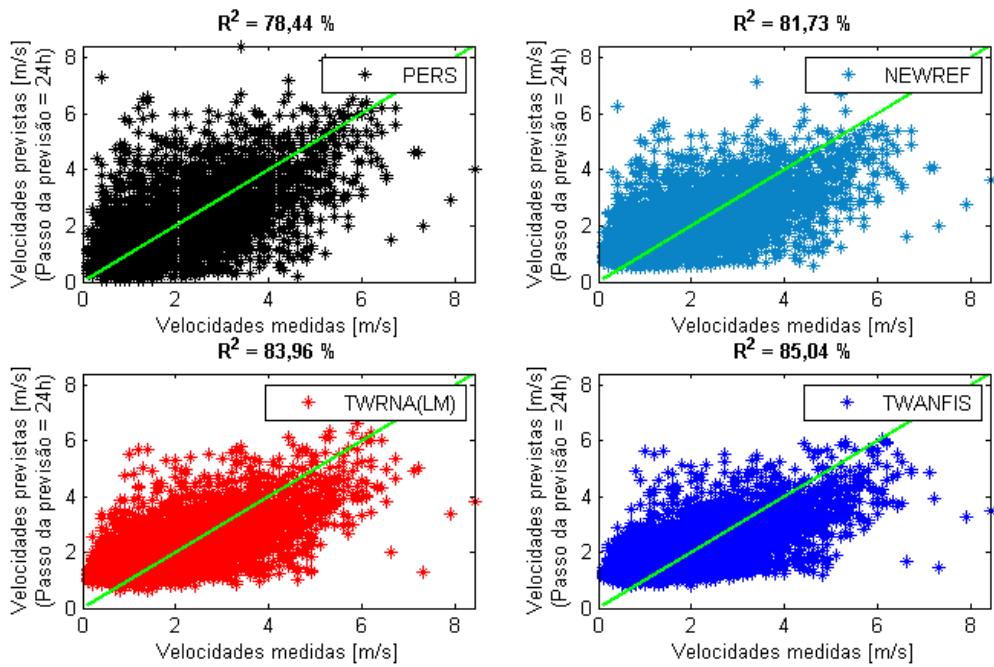


Figura A.10 – Diagramas de dispersão das previsões de velocidades em MOSSORÓ com passo de 24 horas.

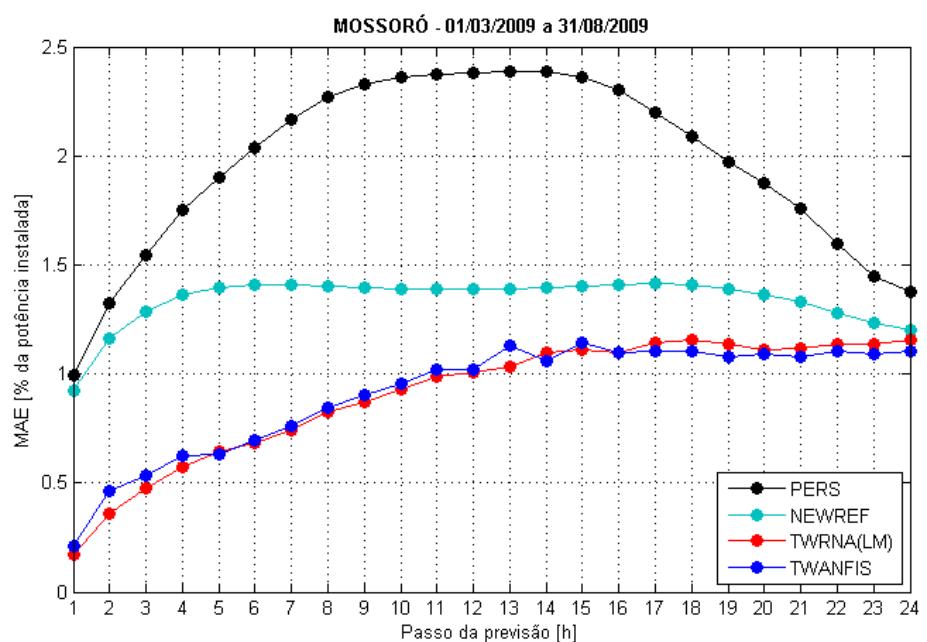


Figura A.11 – MAE das previsões de geração de MOSSORÓ.

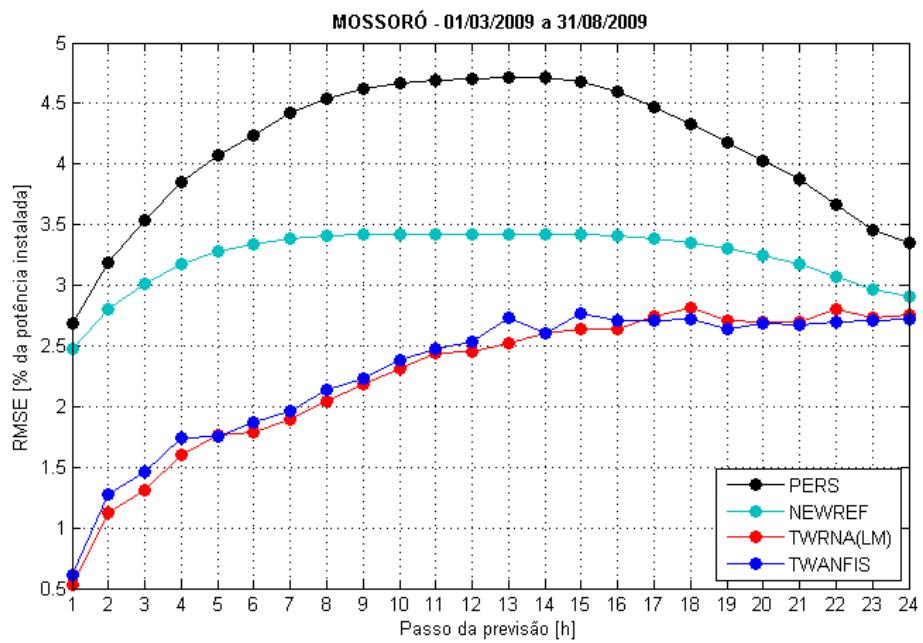


Figura A.12 – RMSE das previsões de geração de MOSSORÓ.

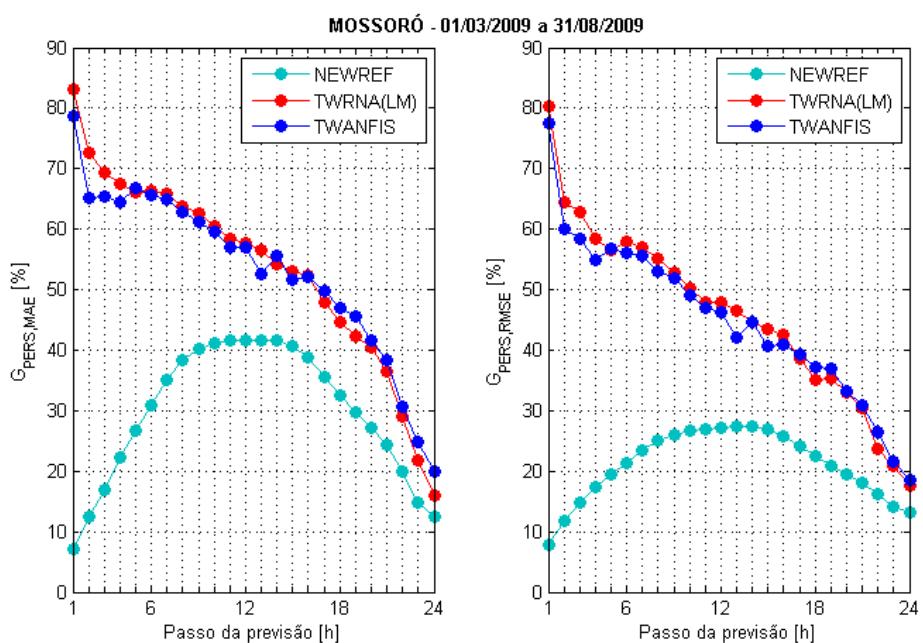


Figura A.13 – Ganhos das previsões de geração de MOSSORÓ.

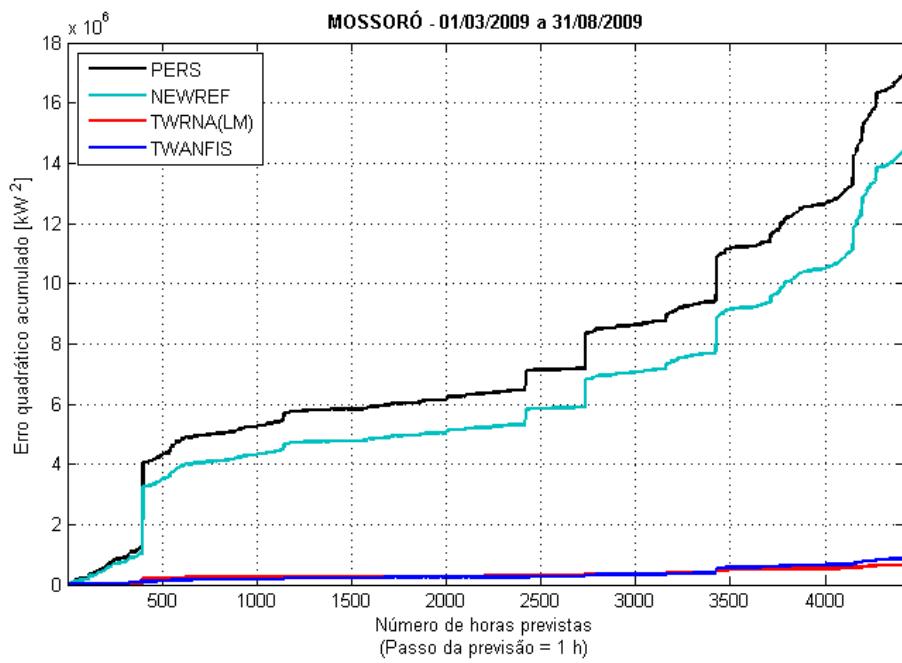


Figura A.14 – Erros quadráticos acumulados das previsões de geração de MOSSORÓ com passo de 1 hora.

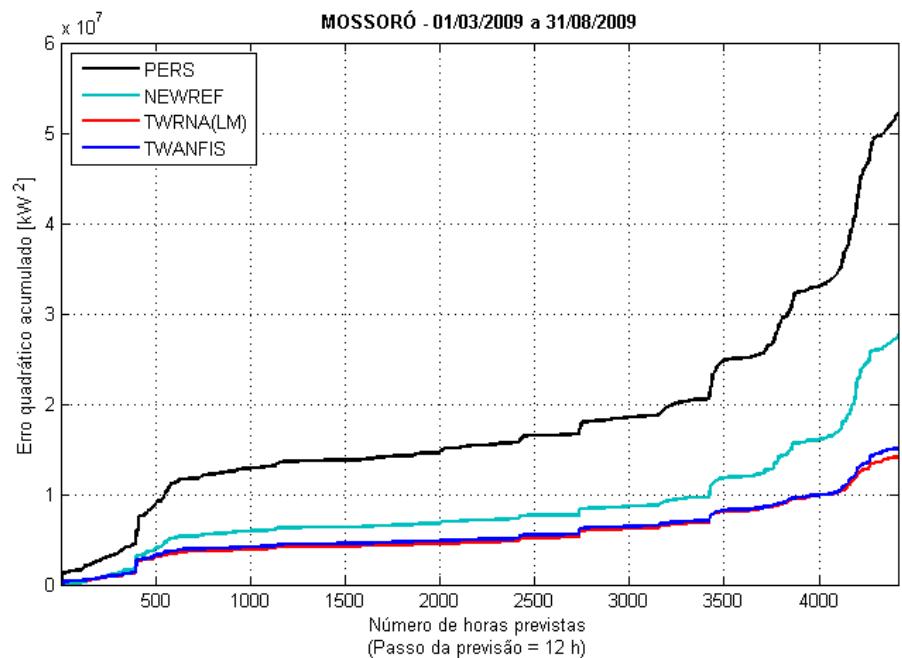


Figura A.15 – Erros quadráticos acumulados das previsões de geração de MOSSORÓ com passo de 12 horas.

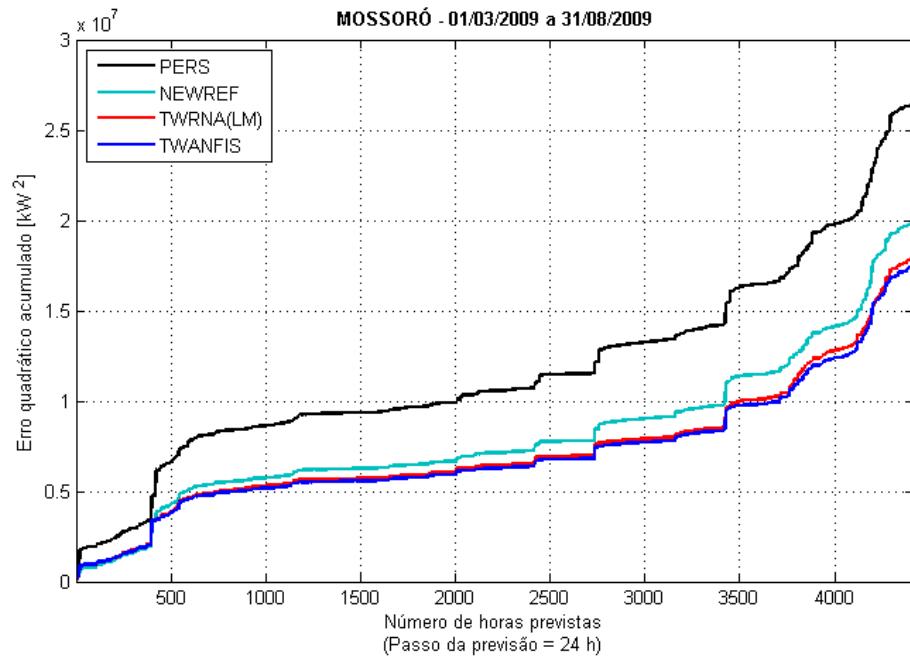


Figura A.16 – Erros quadráticos acumulados das previsões de geração de MOSSORÓ com passo de 24 horas.

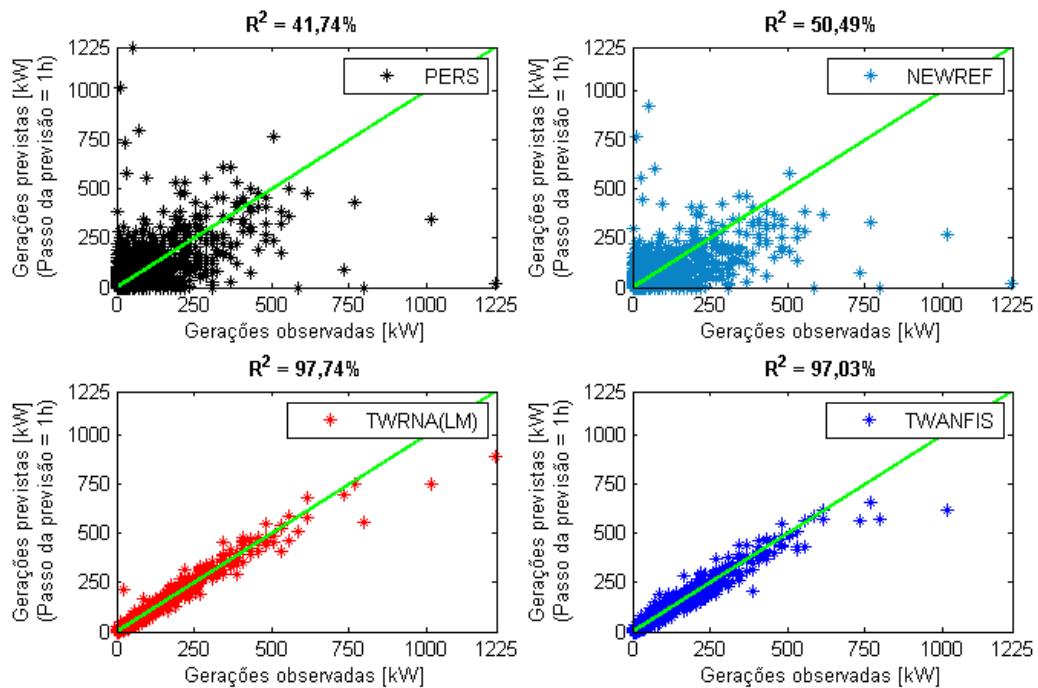


Figura A.17 – Diagramas de dispersão das previsões de geração de MOSSORÓ com passo de 1 hora.

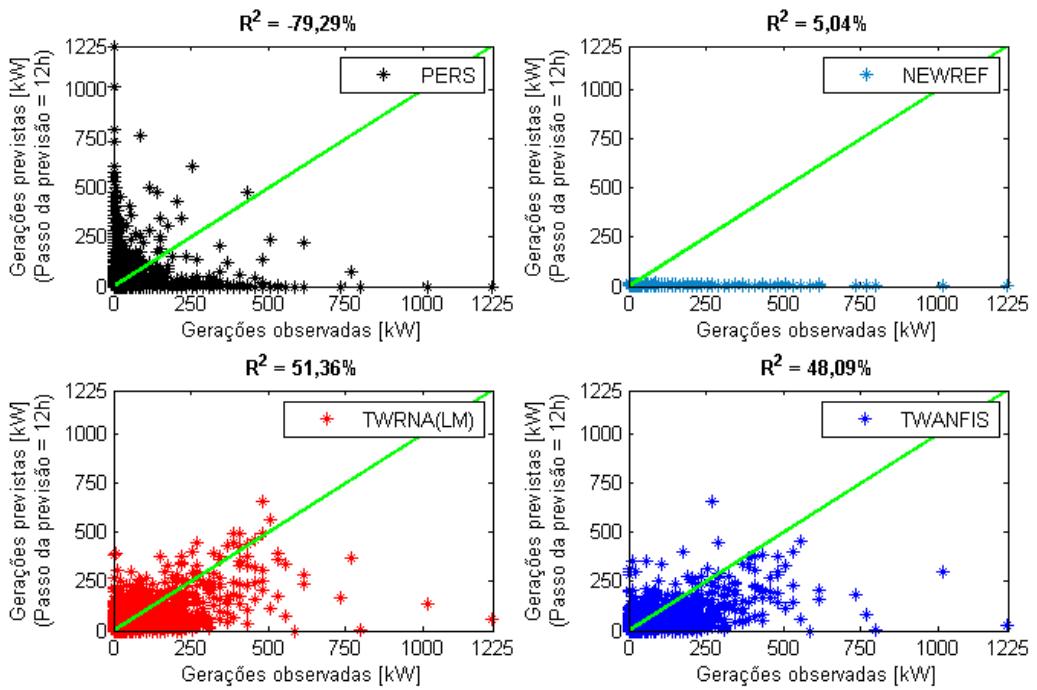


Figura A.18 – Diagramas de dispersão das previsões de geração de MOSSORÓ com passo de 12 horas.

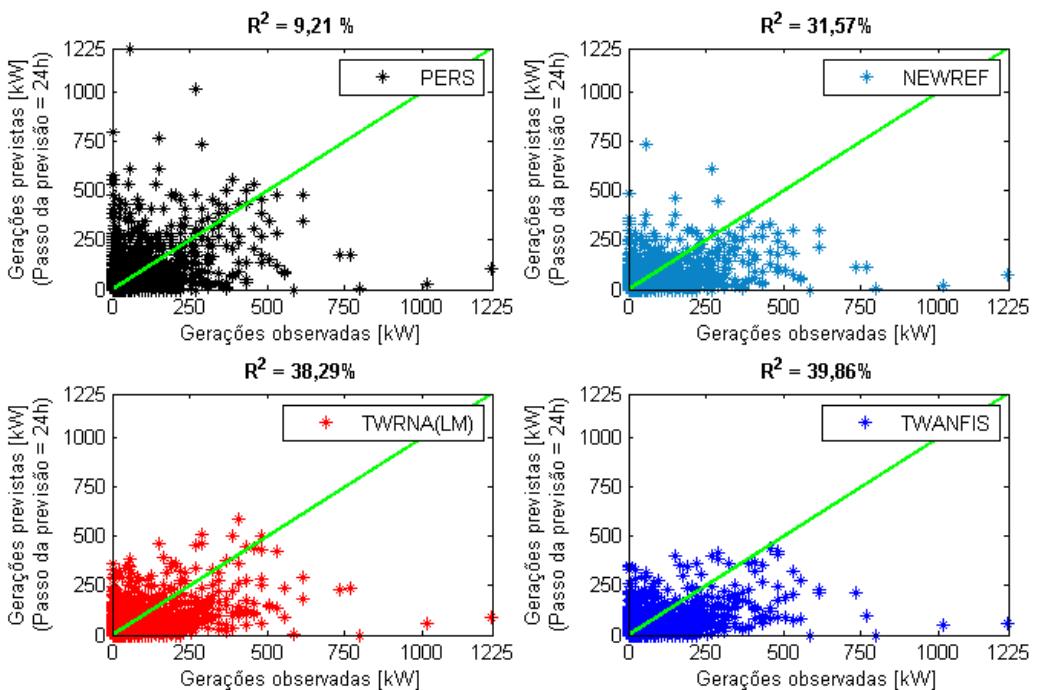


Figura A.19 – Diagramas de dispersão das previsões de geração de MOSSORÓ com passo de 24 horas.

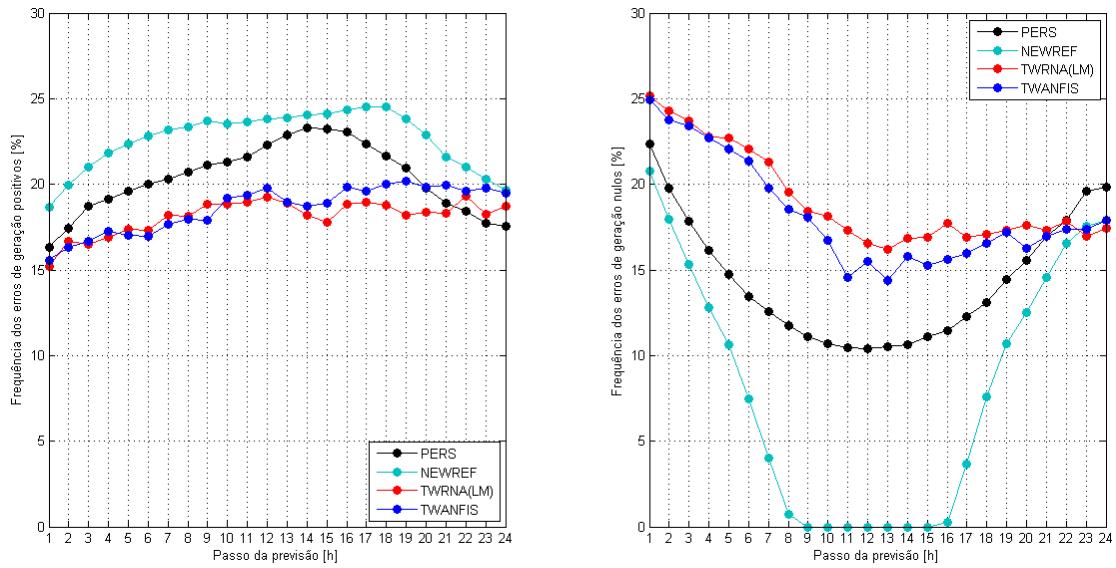
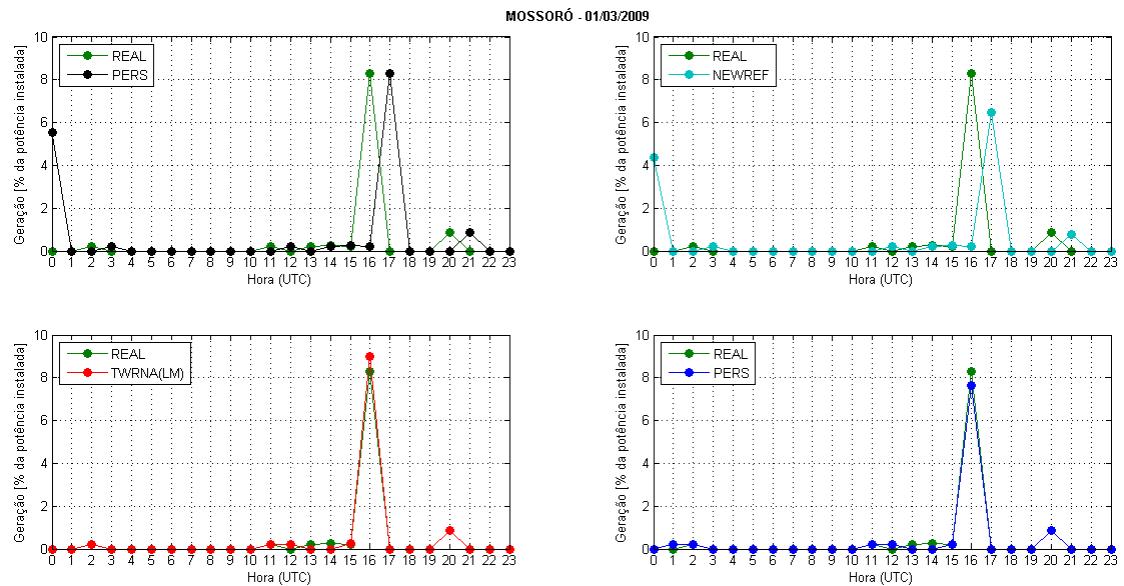
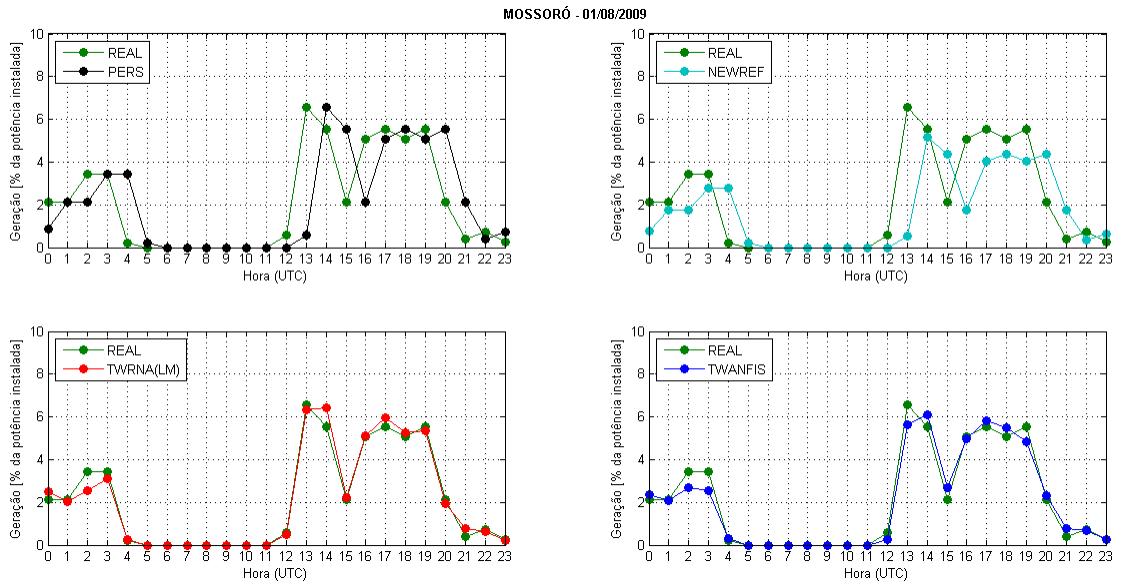


Figura A.20 – Frequência percentual dos erros de geração de MOSSORÓ.



**Figura A.21 – Previsões de geração de MOSSORÓ para o dia 01/03/2009.
(Passo de 1 hora).**



**Figura A.22 – Previsões de geração de MOSSORÓ para o dia 01/08/2009.
(Passo de 1 hora).**

A.2 Previsões para NATAL

Para NATAL, o período escolhido para a avaliação dos modelos é formado pelas 4.368 velocidades médias horárias correspondentes à hora zero (UTC) do dia 01/10/2009 até a hora 23 (UTC) do dia 31/03/2010.

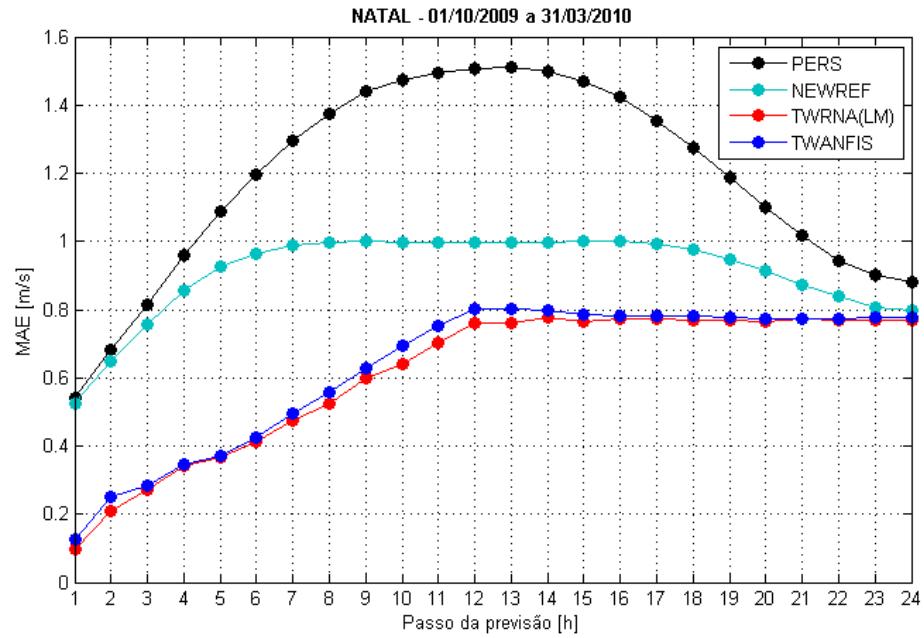


Figura A.23 – MAE das previsões de velocidades em NATAL.

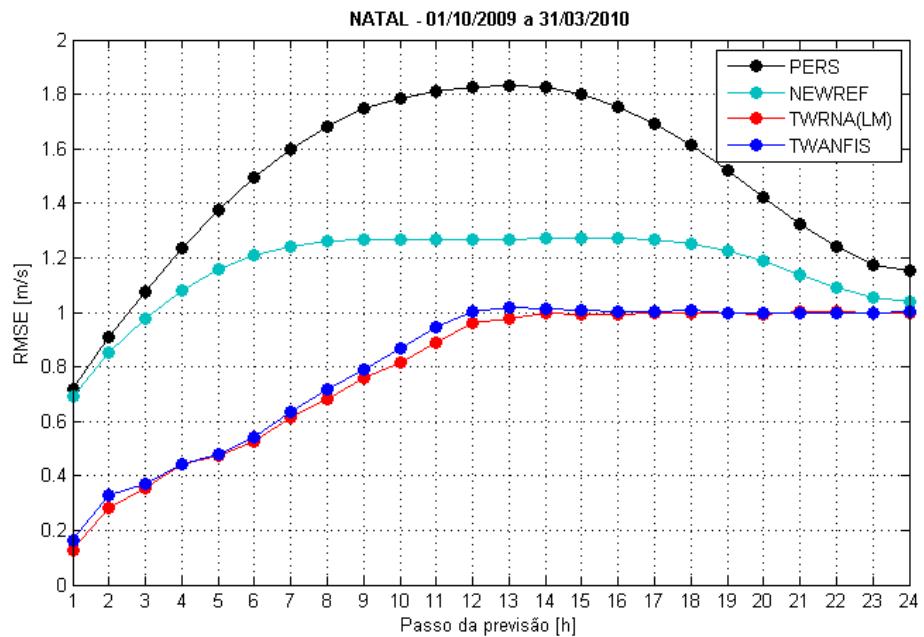


Figura A.24 – RMSE das previsões de velocidades em NATAL.

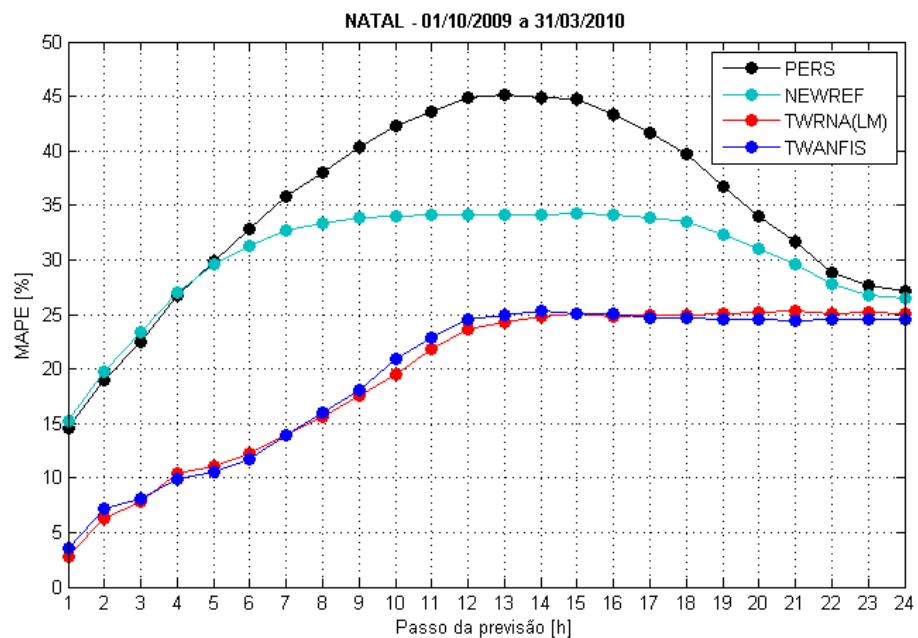


Figura A.25 – MAPE das previsões de velocidades em NATAL.

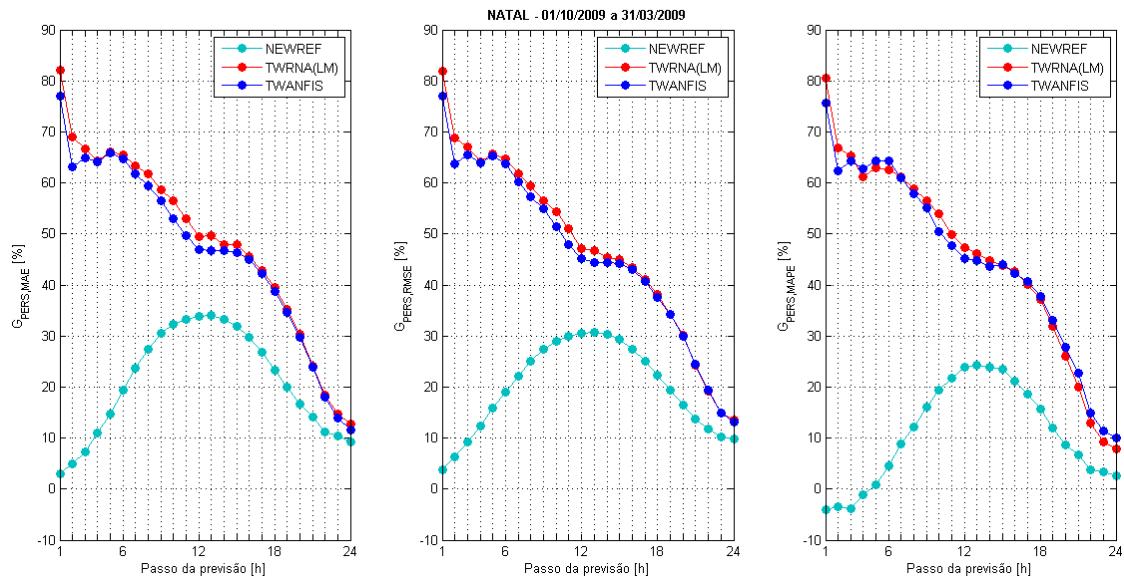


Figura A.26 – Ganhos das previsões de velocidades em NATAL.

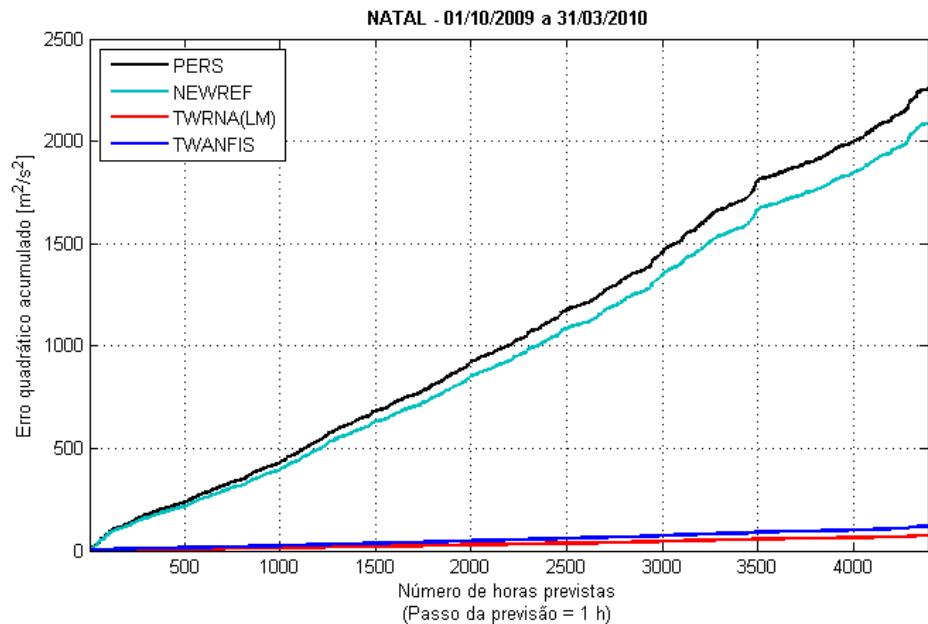
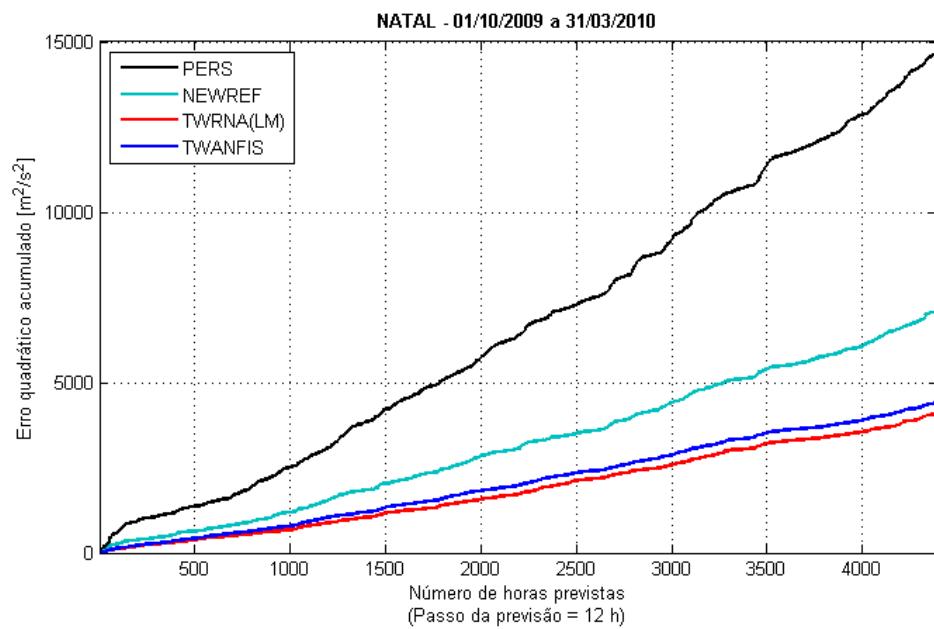
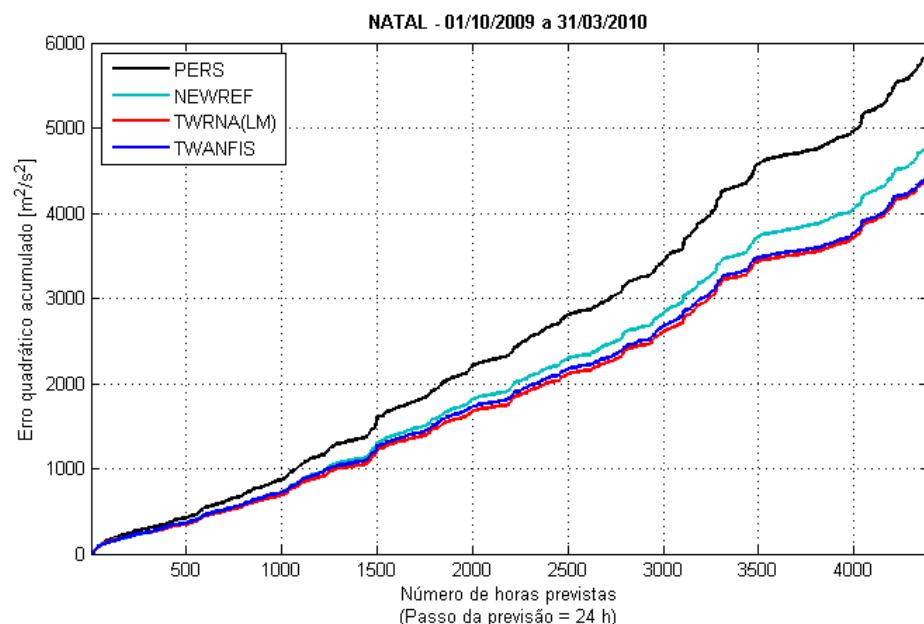


Figura A.27 – Erros quadráticos acumulados das previsões de velocidades em NATAL com passo de 1 hora.



**Figura A.28 – Erros quadráticos acumulados das previsões de velocidades em NATA
com passo de 12 horas.**



**Figura A.29 – Erros quadráticos acumulados das previsões de velocidades em NATA
com passo de 24 horas.**

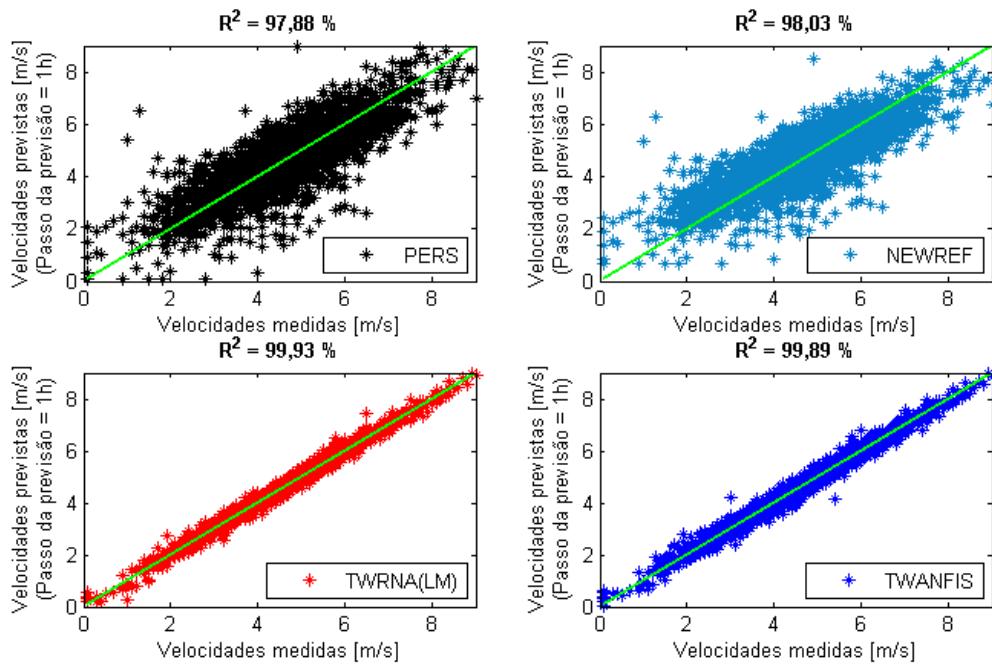


Figura A.30 – Diagramas de dispersão das previsões de velocidades em NATAL com passo de 1 hora.

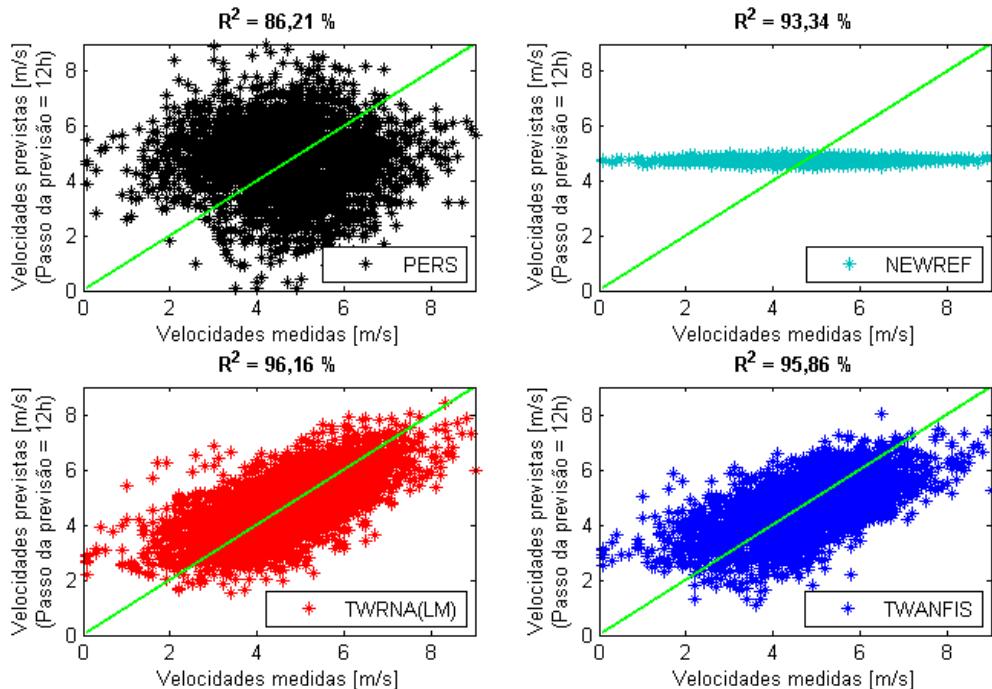


Figura A.31 – Diagramas de dispersão das previsões de velocidades em NATAL com passo de 12 horas.

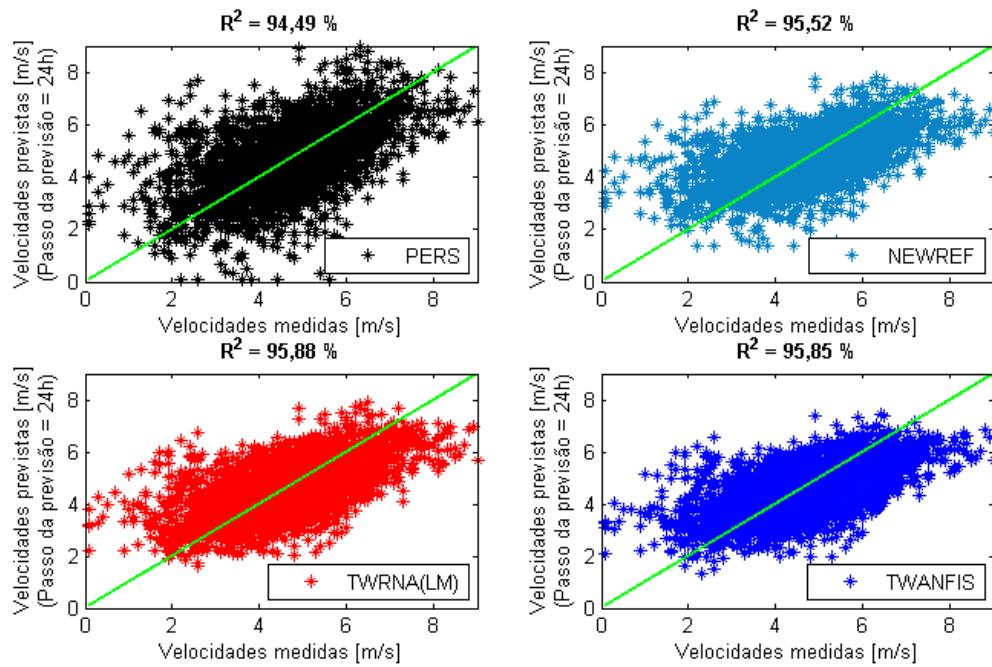


Figura A.32 – Diagramas de dispersão das previsões de velocidades em NATAL com passo de 24 horas.

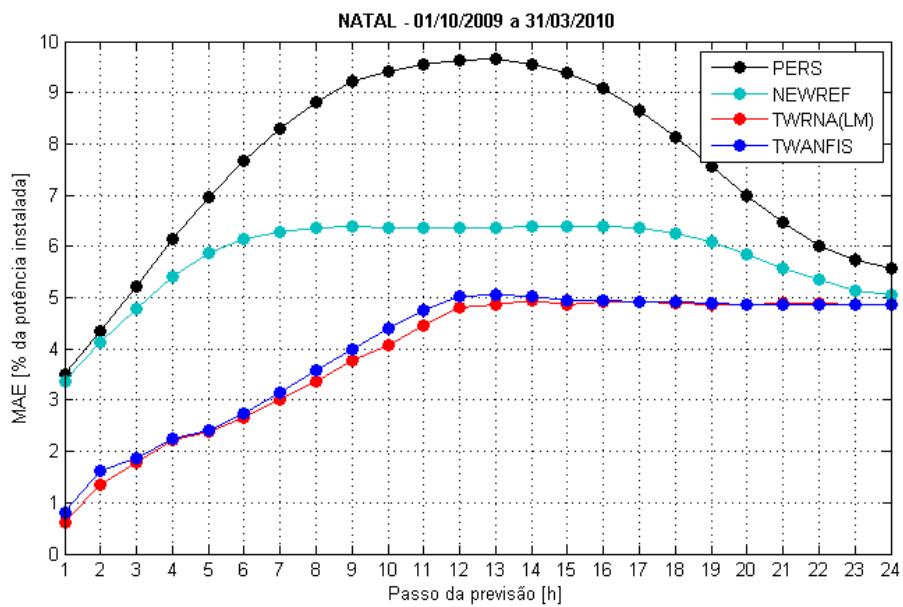


Figura A.33 – MAE das previsões de geração de NATAL.

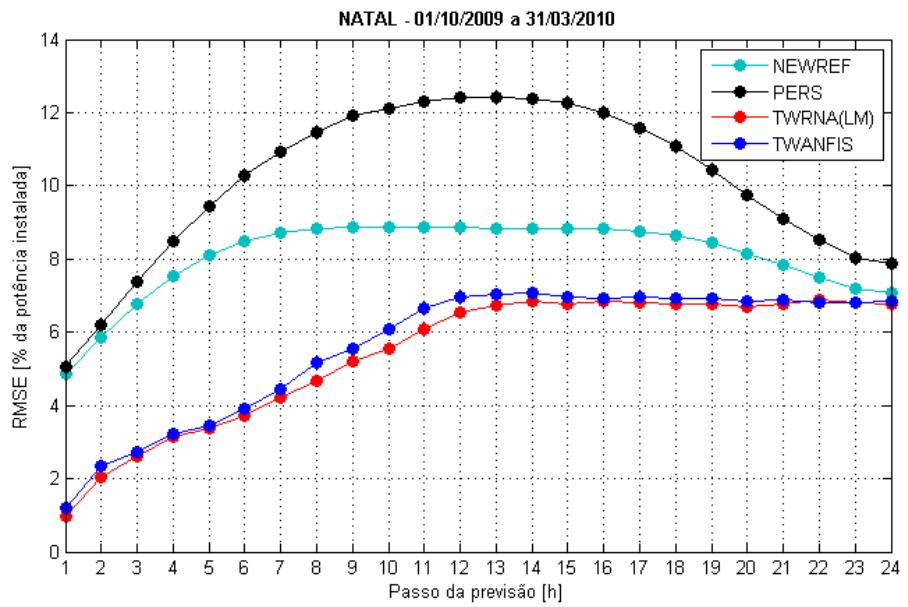


Figura A.34 – RMSE das previsões de geração de NATAL.

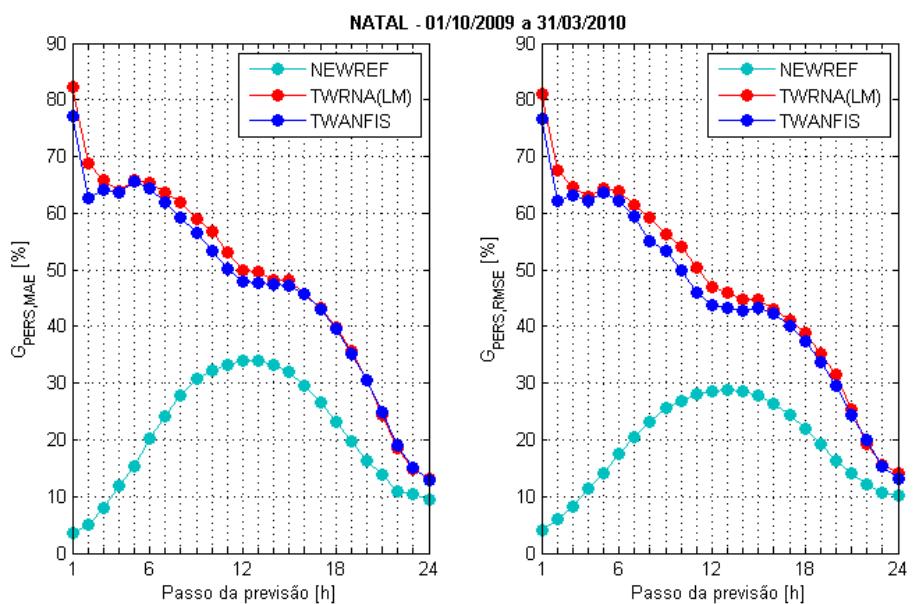


Figura A.35 – Ganhos das previsões de geração de NATAL.

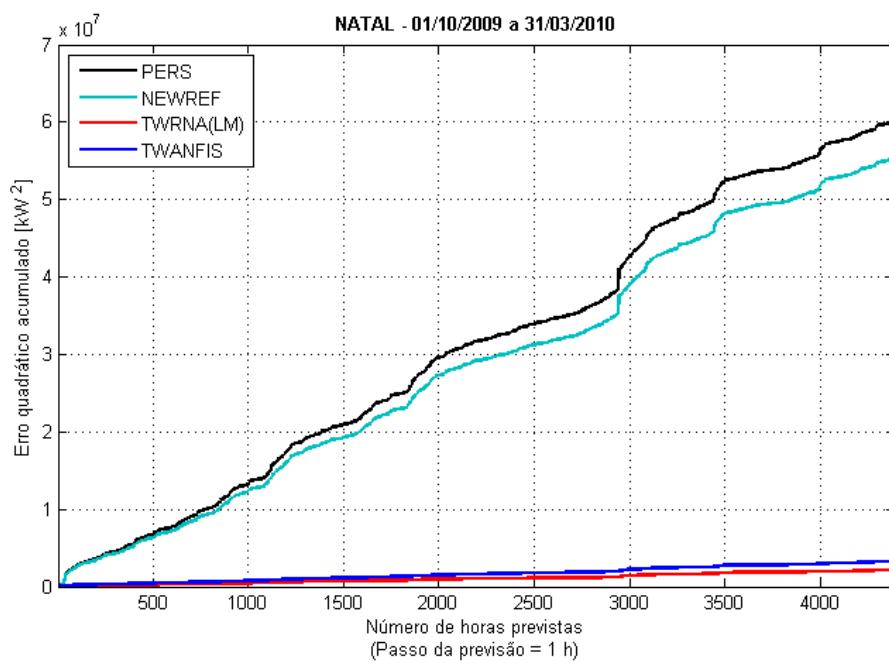


Figura A.36 – Erros quadráticos acumulados das previsões de geração de NATAL com passo de 1 hora.

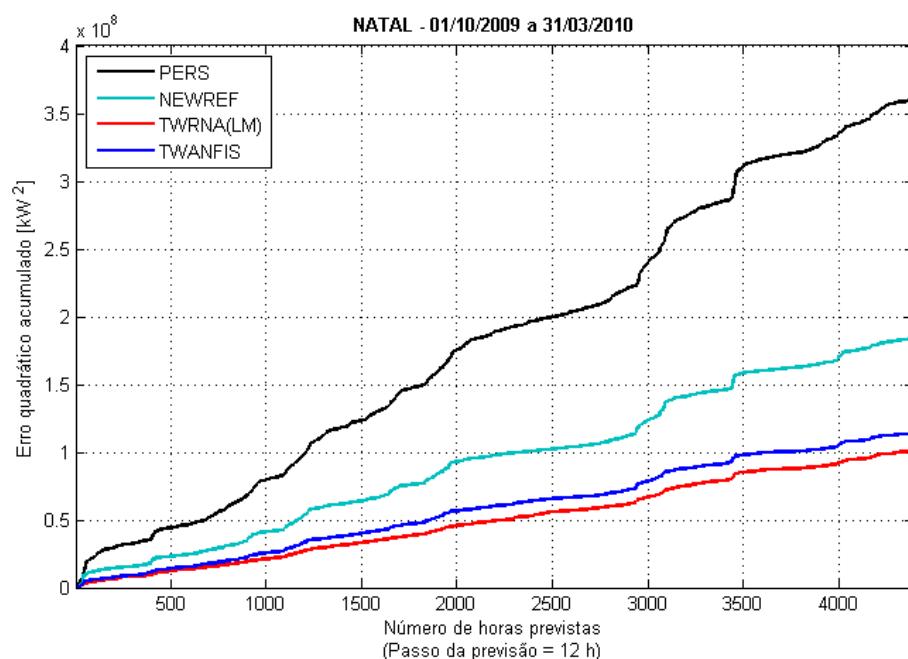


Figura A.37 – Erros quadráticos acumulados das previsões de geração de NATAL com passo de 12 horas.

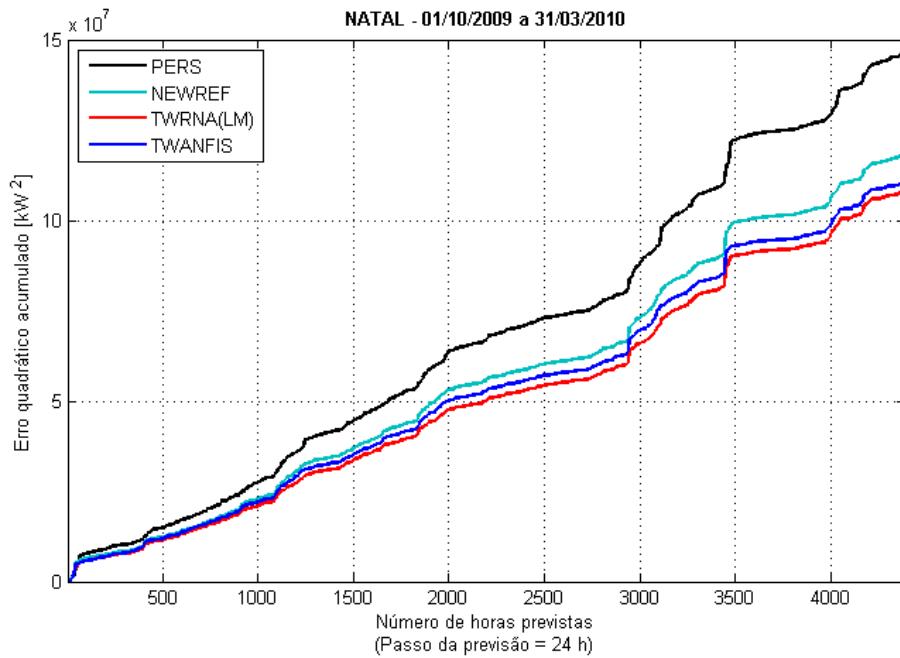


Figura A.38 – Erros quadráticos acumulados das previsões de geração de NATAL com passo de 24 horas.

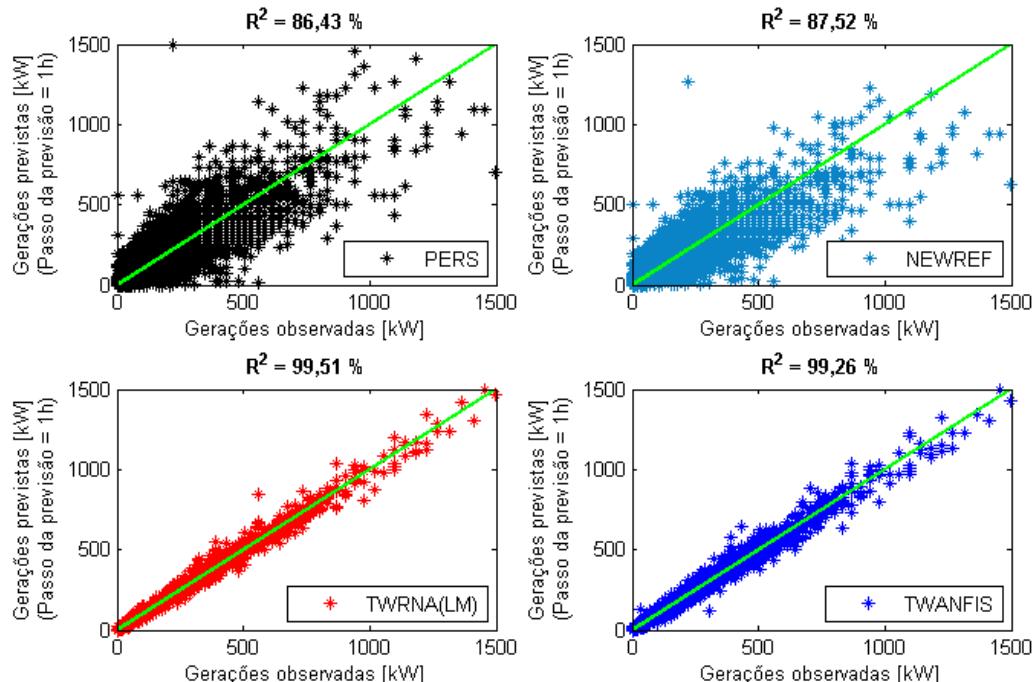


Figura A.39 – Diagramas de dispersão das previsões de geração de NATAL com passo de 1 hora.

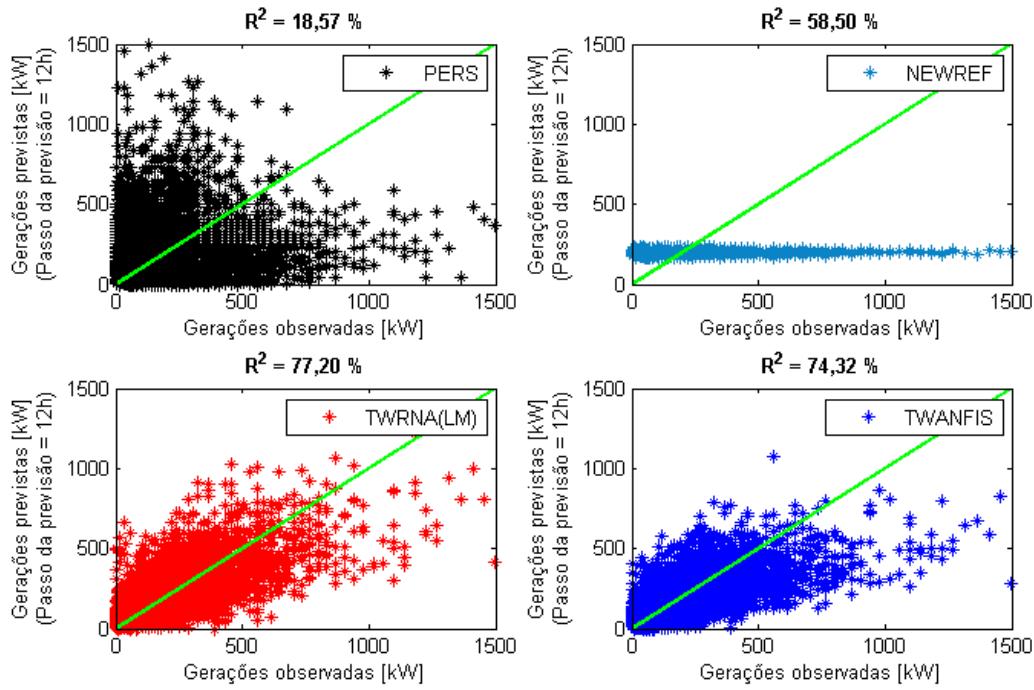


Figura A.40 – Diagramas de dispersão das previsões de geração de NATAL com passo de 12 horas.

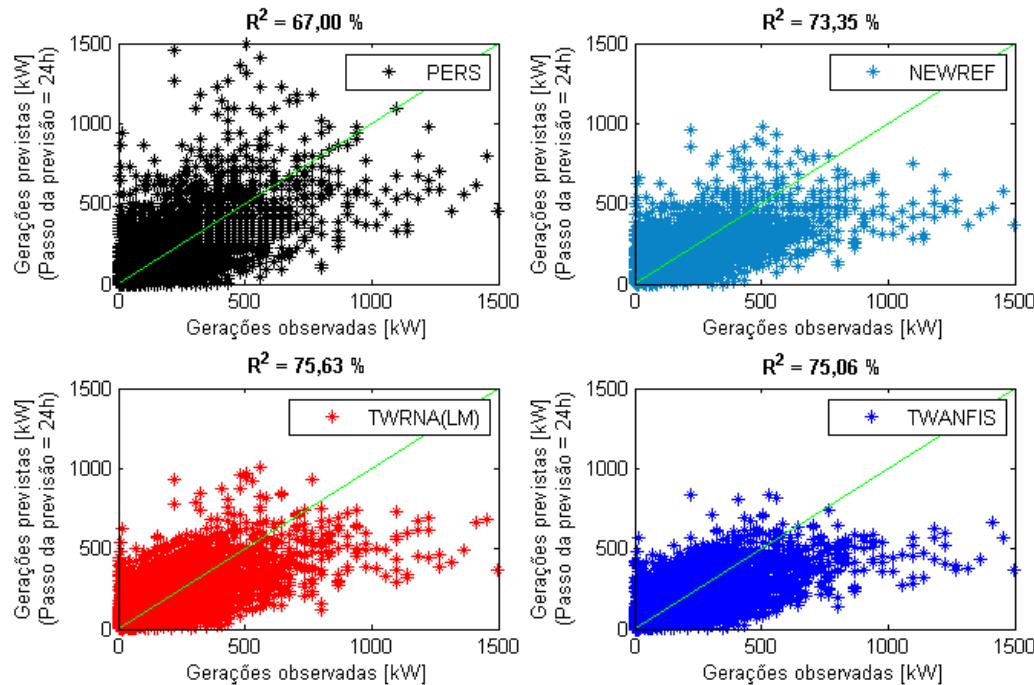


Figura A.41 – Diagramas de dispersão das previsões de geração de NATAL com passo de 24 horas.

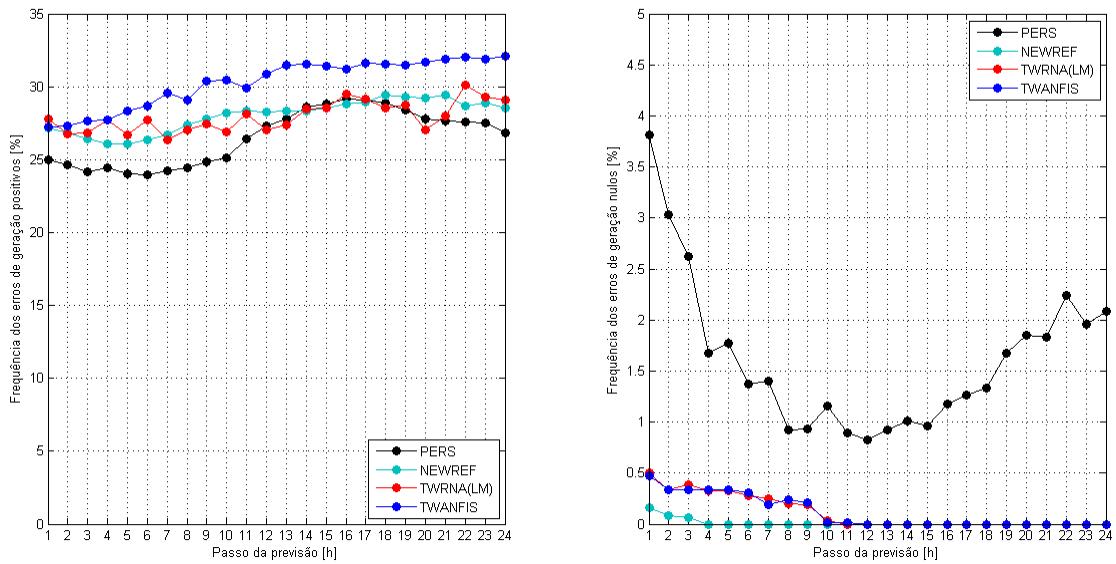
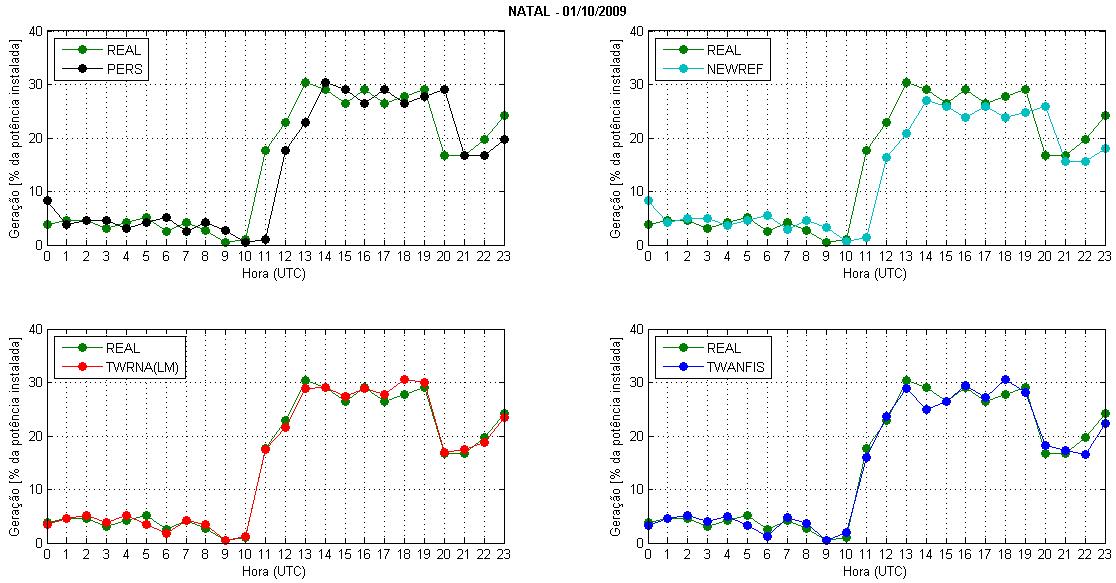
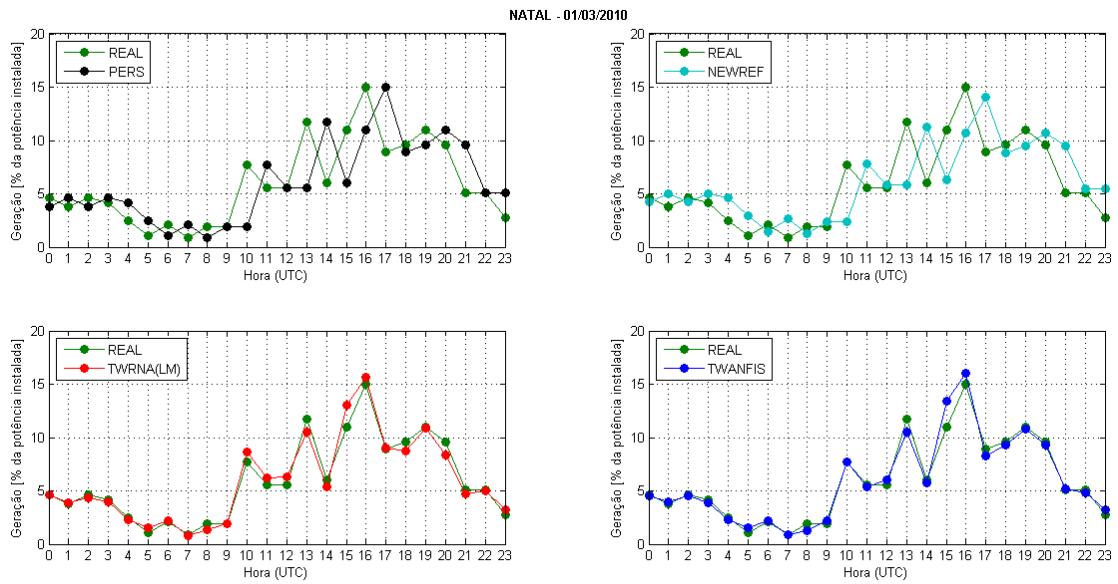


Figura A.42 – Frequência percentual dos erros de geração de NATAL.



**Figura A.43 – Previsões de geração de NATAL para o dia 01/10/2009.
(Passo de 1 hora).**



**Figura A.44 – Previsões de geração de NATAL para o dia 01/03/2010.
(Passo de 1 hora).**

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] LEITE, A. P. **MODELAGEM DE FAZENDAS EÓLICAS PARA ESTUDOS DE CONFIABILIDADE**. 2005. 151 p. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- [2] PAVINATTO, E. F. **FERRAMENTA PARA AUXÍLIO À ANÁLISE DE VIABILIDADE TÉCNICA DA CONEXÃO DE PARQUES EÓLICOS À REDE ELÉTRICA**. 2005. 165 p. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.
- [3] COSTA, A. et al. “A review on the young history of the wind power short-term prediction”. **Renewable & Sustainable Energy Reviews**, v. 12, n. 6, p.1725-1744, 2008.
- [4] SREEKLAHKSHMI, K.; KUMAR, P. R. “Performance evaluation of short term wind speed prediction techniques”. **International Journal of Computer Science and Network Security**, v. 8, n. 8, p.162-169, Aug. 2008.
- [5] BHASKAR, M.; JAIN, A.; VENKATA, S. N. “Wind speed forecasting: Present status”. In: **IEEE International Conference on Power System Technology (POWERCON)**, 2010, Hangzhou.
- [6] OLIVEIRA, J. B. **SISTEMAS INTELIGENTES E WAVELETS PARA PREVISÃO DE VENTO E GERAÇÃO EÓLICA**. 2008. 92 p. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal de Pernambuco, Recife, 2005.
- [7] AMARANTE, O. A. C. et al. **Atlas do Potencial Eólico Brasileiro**, Ministério de Minas e Energia, Eletrobrás, CEPEL, 2001. Disponível em: <http://www.cresesb.cepel.br/publicacoes/download/atlas_eolico/Atlas%20do%20Potencial%20Eolico%20Brasileiro.pdf>. Acesso em: 01 dez. 2011.
- [8] AQUINO, R. R. B. et al. “AVALIAÇÃO DA COMPLEMENTARIDADE ENTRE OS FLUXOS HIDROLÓGICO E EÓLICO NA REGIÃO NORDESTE”. In: **XIII Congresso Brasileiro de Energia**, 2010, Rio de Janeiro.
- [9] PATEL, M. R. **Wind and Solar Power Systems - Design, Analysis, and Operation**. 2. ed. Boca Raton: CRC Press, 2006.

- [10] MARQUES, J. et al. “A survey on Variable-Speed Wind Turbine System”. In: **6. Congresso Brasileiro de Eletrônica de Potência – COBEP’02**, v.1, p. 732-738, 2002.
- [11] CARLIN, P. W.; LAXSON, A. S.; MULJADI, E. B. **The History and State of the Art of Variable-Speed Wind Turbines Technology**, 2001, NREL. Disponível em: <<http://www.nrel.gov/docs/fy01osti/28607.pdf>>. Acesso em: 05 dec. 2011.
- [12] GIEBEL, G.; KARINIOTAKIS, G.; BROWNSWORD, R. **The State of the art in Short-Term Prediction of Wind Power – A Literature Overview**. Disponível em: <http://anemos.cma.fr/download/ANEMOS_D1.1_StateOfTheArt_v1.1.pdf>. Acesso em: 26 mar. 2010.
- [13] LANDBERG, L. et al. “Short-term Prediction – An Overview”. **Wind Energy** 6(3), p. 273-280, Jun. 2003. DOI 10.1002/we.96.
- [14] JENSEN, U.S.; PELGRUM, E.; MADSEN, H. “The Development of a Forecasting Model for the Prediction of Wind Power Production to be Used in Central Dispatch Centres”. **Proceedings of the EWEC '94**, Thessaloniki, p. 353-356, 1994.
- [15] ELSAM, **FINAL Report on EU JOULE II Project JOU-CT92-0083**, 1996.
- [16] GIEBEL, G. **On the Benefits of Distributed Generation of Wind Energy in Europe**. PhD thesis from the Carl von Ossietzky Universität. Fortschr.-Ber. VDI Reihe 6 Nr. 444. Düsseldorf, VDI Verlag 2001. ISBN 3-18-344406-2.
- [17] BEYER, H.G. et al. “Short Term Prediction of Wind Speed and Power Output of a Wind Turbine with Neural Networks”. **Proceedings of the EWEC '94**, Thessaloniki, p. 349-352, 1994.
- [18] ALEXIADIS, M.C. et al. “Short-Term Forecasting of Wind Speed and Related Electrical Power”. **Solar Energy**, n. 63, p. 61-68, 1998. DOI:10.1016/S0038-092X(98)00032-2
- [19] SFETSOS, A. “A novel approach for the forecasting of mean hourly wind speed time series”. **Renewable Energy**, n. 27, p. 163-174, 2001.
- [20] AQUINO, R. R. B. M. et al. “Application of wavelet and neural network models for wind speed and power generation forecasting in a brazilian experimental wind park”. In: **International Joint Conference on Neural Networks (IJCNN)**, 2009, Atlanta.

- [21] MADSEN, H. et al. **A Protocol for Standardizing the Performance Evaluation of Short-term Wind Power Prediction Models**. Disponível em: <http://anemos.cma.fr/download/ANEMOS_D2.3_EvaluationProtocol.pdf>. Acesso em: 26 mar. 2010.
- [22] DAMOUSIS, I. G. et al. “A Fuzzy Model for Wind Speed Prediction and Power Generation in Wind Parks Using Spatial Correlation”. **IEEE Transactions on Energy Conversion**, v. 19, n. 2, p. 352-361, Jun. 2009.
- [23] CATALÃO, J. P. S.; POUSINHO, H. M. I.; MENDES, V. M. F. “Hybrid Wavelet-PSO-ANFIS Approach for Short-Term Wind Power Forecasting in Portugal”. **IEEE Transactions on Sustainable Energy**, v. 2, n. 1, p. 50-59, Jan. 2011.
- [24] HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: BOOKMAN Editora S.A., 2001.
- [25] RIEDMILLER, M.; BRAUN, H. “A Direct Adaptive Method for Faster Backpropagation Learning: the RPROP Algorithm”. In: **IEEE International Conference on Neural Networks**, 1993, San Francisco.
- [26] HAGAN, M. T.; MENHAJ, M. B. “Training Feedforward Networks with the Marquardt Algorithm”. **IEEE Transactions on Neural Networks**, v. 5, n. 6, Nov. 1994.
- [27] KOSKO, B. **Fuzzy engineering**. ISBN 0-13-124991-6. New Jersey: Prentice Hall, 1997.
- [28] JANG, J.R., SUN, C. Neuro-Fuzzy Modeling and Control. **Proceedings of the IEEE**, v. 83, n. 3, Mar. 1995.
- [29] REZENDE, S. O. **Sistemas inteligentes: fundamentos e aplicações**. 2. ed. ISBN 85-204-1683-7. Barueri: Manole, 2005.
- [30] JANG, J.R. “ANFIS: Adaptive-Network-Based Fuzzy Inference System”. **IEEE Transactions on Electric Power Systems Research**, v. 23, n 3, p. 169-176, 1993.
- [31] SANTOSO, S.; POWER, E. J.; GRANDY, W. M. “Power Quality Disturbance Data Compression Using Wavelet Transform Methods”. **IEEE Transactions on Power Delivery**, v. 12, n. 3, p. 1250-1257, Jul. 1997.

- [32] LIRA, M. M. S. **WAVELETS NA COMPACTAÇÃO E PROCESSAMENTO DE SINAIS DE DISTÚBIOS EM SISTEMAS DE POTÊNCIA PARA CLASSIFICAÇÃO VIA REDES NEURAIS ARTIFICIAIS.** 118 p. Tese (Doutorado em Engenharia Elétrica) – Universidade Federal de Pernambuco, Recife, 2004.
- [33] MALLAT, S. “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation”. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 11, p. 674-693, July 1989.
- [34] SOUZA, R. C.; CAMARGO, M. E. **ANÁLISE E PREVISÃO DE SÉRIES TEMPORAIS. OS MODELOS ARIMA.** 2. ed. Rio de Janeiro: Gráfica e Editora Regional, 2004.
- [35] SOUSA, J. N. M. S. **Previsão da Produção Eléctrica em Parques Eólicos.** 2007. 221 p. Dissertação (Mestrado em Engenharia Eletrotécnica e de Computadores) – Faculdade de Engenharia da Universidade do Porto, Porto, 2007.