



**Júlia Moreira Miranda**

**Modelos de Predição da Velocidade e  
Direção do Vento Utilizando Redes Neurais  
Artificiais**

**Wind Speed and Direction Forecasting Models Using  
Artificial Neural Networks**

**Projeto de Graduação**

Projeto de Graduação apresentado ao Departamento de Engenharia Mecânica da PUC-Rio.

Orientador: Helon Vicente Hultmann Ayala

Área de Concentração: Identificação de Sistemas

Rio de Janeiro  
Julho de 2019

## Agradecimentos

À minha família e, em especial, a minha mãe Tereza Cristina Moreira, ao meu pai Renato Miranda, minha avó Ilma Miranda e a minha tia Ilma Gabriela Miranda, sempre me apoiando emocionalmente e financeiramente, além de sempre me motivar a estudar incansavelmente e acreditar no meu potencial.

Ao meu namorado, Richard Mendes, por todo apoio e compreensão durante os dias em que fui ausente, além de ser a pessoa que sempre me deu forças para continuar nos momentos de cansaço.

Ao meu orientador, Helon Vicente Hultmann Ayala, por toda a orientação e ajuda na confecção deste trabalho, além de todos os conselhos, apoio e confiança durante o período que fui sua aluna e orientanda.

Aos amigos que fiz na PUC, por todos os momentos e estudos compartilhados, por todo o apoio antes das provas e entrega de trabalhos, além dos momentos de distração e entretenimento ao longo desses 5 anos.

À todos os professores que tive ao longo da Graduação, pelos conhecimentos transmitidos e pela vivência diária, contribuindo para minha construção profissional e pessoal.

Ao professor Carlos Alberto de Almeida, pelo prazer de ter sido sua aluna ao longo de todo Ciclo Profissional, além de toda atenção, cuidado e paciência que sempre me dedicou, dentro e fora da sala.

À PUC-Rio, como instituição educacional, que me trouxe diversas experiências e amadurecimento ao longo dos 5 anos de Graduação. Pela oportunidade de monitorias, iniciação científica e estágios.

Finalmente, dedico este trabalho a Deus, que sempre me protegeu e me proporcionou chegar até aqui. Tenho certeza que Ele me guiou e me ajudou a traçar meu caminho, me conduzindo, orientando com as escolhas certas e me afastando de todo mal hoje e sempre.

## Resumo

### Modelos de Predição da Velocidade e Direção do Vento Utilizando Redes Neurais Artificiais

Nas Ciências Econômicas, assim como na Engenharia e nas Ciências Naturais, ocorrem fenômenos que dependem da observação de dados em intervalos de tempo, durante um período específico. Estas observações de valores são denominadas de Séries Temporais. As técnicas disponíveis para analisar estas observações são denominadas de Análise de Séries Temporais. Este conjunto de técnicas tem por objetivo construir um modelo com número adequado de parâmetros estimados de forma a ajustar o modelo à série temporal. A obtenção de um modelo adequado é de extrema importância, pois pode revelar algumas características da série temporal que ajudam na previsão de seus valores futuro, sendo esta habilidade de grande relevância prática, especialmente na geração de energia eólica. Diversas técnicas para realizar a previsão de séries temporais são utilizadas, porém todas possuem um erro de previsão. O desafio é então minimizar o erro de previsão, construindo um modelo que melhor se ajuste aos dados. Dessa forma, o objetivo deste trabalho é criar modelos de predição da velocidade e direção dos ventos, considerando que haja dados históricos. Esses modelos serão criados utilizando primeiramente Rede Neural Artificial (RNA), e depois o modelo NARMA (*Nonlinear Autoregressive Moving Average*) é empregado juntamente com a RNA para comparação de resultados. A previsão é efetuada para um e seis passos à frente. O uso apenas de RNA resultou em resultados satisfatórios para todos os casos, com exceção da predição para seis passos à frente da velocidade. Entretanto, quando foi empregado o modelo híbrido (NARMA+RNA), todos os resultados melhoraram consideravelmente. As amostras de séries temporais de vento para previsão são oriundas de Parques Eólicos no Rio Grande do Norte, de propriedade da Empresa Brookfield Energia Renovável, sendo coletadas na base de 10 minutos.

### Palavras-chave

Rede Neural Artificial; NARMA; Predição; Séries Temporais; Direção do vento; Velocidade do vento.

## **Abstract**

### **Wind Speed and Direction Forecasting Models Using Artificial Neural Networks**

In Economic Sciences, as well as in Engineering and Natural Sciences, phenomena occurs that depends on observing data at time intervals during a specific period. These observations of values are called Time Series. The techniques available to analyze these observations are called Time Series Analysis. This set of techniques aims to construct a model with adequate number of parameters estimated in order to fit the model to the time series. Obtaining a suitable model is extremely important because it can reveal some characteristics of the time series that help predict future values. This ability is of great practical relevance, especially in the generation of wind energy. Several techniques to perform the time series prediction are used, but all have a forecast error. The challenge is then to minimize forecast error by constructing a model that best fits the data. In this way, the objective of this work is to create prediction models of the speed and direction of the winds, considering that there are historical data. These models will first be created using Artificial Neural Network (RNA), and then the NARMA (Nonlinear Autoregressive Moving Average) model is used together with RNA for comparison of results. The forecast is made for one and six steps ahead. The use of RNA alone resulted in satisfactory results in all cases, with the exception of the prediction for six steps ahead of speed. However, when the hybrid model (NARMA + RNA) was used, all results improved considerably. The wind forecast series samples are from Wind Farms in Rio Grande do Norte, owned by the Brookfield Renewable Energy Company, and are collected at the base of 10 minutes.

## **Keywords**

Artificial Neural Network; NARMA; Forecasting; Time Series; Wind direction; Wind speed.

## Sumário

1	Introdução	8
1.1	Motivação	8
1.2	Objetivo	9
1.3	Organização do Trabalho	9
2	Revisão Bibliográfica	11
3	Métodos	14
3.1	Séries Temporais	14
3.2	Redes Neurais Artificiais	18
3.3	Modelo NARMA	21
3.4	Previsão de Séries com RNAs	23
3.5	Problema: Previsão de Séries de Vento	24
3.6	Métricas de Desempenho	25
3.7	Testes de Correlação para Modelos de Séries Temporais	25
4	Dados Disponíveis	27
4.1	Autocorrelação dos Dados Disponíveis	30
5	Resultados para Rede Neural Artificial	33
5.1	Busca de Arquiteturas para a Rede Neural Artificial	33
5.2	Transformações	54
6	Resultado com NARMA e RNA	56
6.1	Resultados para a Predição da Velocidade do Vento com NARMA + RNAs	57
6.2	Resultados para a Predição da Direção do Vento com NARMA + RNAs	62
7	Conclusão	67

## Lista de figuras

3.1	Arquitetura do Perceptron - fonte: própria.	19
4.1	Amostra utilizada da velocidade do vento do Gerador 1.	28
4.2	Frequência dos dados para a velocidade do vento.	28
4.3	Amostra utilizada da direção do vento do Gerador 1.	29
4.4	Frequência dos dados para a direção do vento.	30
4.5	Correlograma para Velocidade do Vento.	31
4.6	Correlograma para Direção do Vento.	32
5.1	Predição da velocidade do vento ( $k=1$ ) e erro utilizando RNA.	42
5.2	Predição da velocidade do vento ( $k=6$ ) e erro utilizando RNA.	42
5.3	Gráfico de $y(t)$ versus $\hat{y}(t)$ para velocidade utilizando RNA.	43
5.4	Testes de correlação para velocidade ( $k=1$ ) utilizando RNA.	44
5.5	Predição da direção do vento ( $k=1$ ) e erro utilizando RNA.	52
5.6	Predição da direção do vento ( $k=6$ ) e erro utilizando RNA.	52
5.7	Gráfico de $y(t)$ versus $\hat{y}(t)$ para direção utilizando RNA.	53
5.8	Testes de correlação para direção ( $k=1$ ) utilizando RNA.	54
6.1	Esquema utilizado para gerar resultados no Capítulo 6.	56
6.2	Predição da velocidade do vento e erro utilizando NARMA.	58
6.3	Gráfico de $y(t)$ versus $\hat{y}(t)$ para velocidade utilizando NARMA.	58
6.4	Testes de correlação para velocidade ( $k=1$ ) utilizando NARMA.	59
6.5	Predição da velocidade do vento e erro utilizando NARMA+RNA.	60
6.6	Gráfico de $y(t)$ versus $\hat{y}(t)$ para velocidade utilizando NARMA+RNA.	60
6.7	Predição da direção do vento e erro utilizando NARMA.	63
6.8	Gráfico de $y(t)$ versus $\hat{y}(t)$ para direção utilizando NARMA.	63
6.9	Testes de correlação para direção ( $k=1$ ) utilizando NARMA.	64
6.10	Predição da direção do vento e erro utilizando NARMA+RNA.	65
6.11	Gráfico de $y(t)$ versus $\hat{y}(t)$ para direção utilizando NARMA+RNA.	65

## Lista de tabelas

5.1	Resultados de coeficiente de determinação para velocidade obtidos com learning rate = $10^{-2}$ , para um passo à frente.	36
5.2	Resultados de coeficiente de determinação para velocidade obtidos com learning rate = $10^{-2}$ , para seis passos à frente.	37
5.3	Resultados de coeficiente de determinação para velocidade obtidos com learning rate = $10^{-3}$ , para um passo à frente.	38
5.4	Resultados de coeficiente de determinação para velocidade obtidos com learning rate = $10^{-3}$ , para seis passos à frente.	39
5.5	Resultados de coeficiente de determinação para velocidade obtidos com learning rate = $10^{-4}$ , para um passo à frente.	40
5.6	Resultados de coeficiente de determinação para velocidade obtidos com learning rate = $10^{-4}$ , para seis passos à frente.	41
5.7	Resultados de coeficiente de determinação para direção obtidos com learning rate = $10^{-2}$ , para um passo à frente.	46
5.8	Resultados de coeficiente de determinação para direção obtidos com learning rate = $10^{-2}$ , para seis passos à frente.	47
5.9	Resultados de coeficiente de determinação para direção obtidos com learning rate = $10^{-3}$ , para um passo à frente.	48
5.10	Resultados de coeficiente de determinação para direção obtidos com learning rate = $10^{-3}$ , para seis passos à frente.	49
5.11	Resultados de coeficiente de determinação para direção obtidos com learning rate = $10^{-4}$ , para um passo à frente.	50
5.12	Resultados de coeficiente de determinação para direção obtidos com learning rate = $10^{-4}$ , para seis passos à frente.	51
6.1	Resultados de coeficiente de determinação para velocidade obtidos com NARMA ( $ne=1$ e $nl=1$ ), para um passo à frente.	57
6.2	Resultados de coeficiente de determinação para direção obtidos com NARMA ( $ne=1$ e $nl=1$ ), para um passo à frente.	62

# 1

## Introdução

Nas últimas décadas a energia eólica tem conquistado amplo espaço na matriz energética de diversos países. A crescente preocupação com aquecimento global, poluição do ar, utilização de matérias-primas renováveis e crise energética resulta inevitavelmente em uma mudança de paradigmas em como produzir energia no mundo atual.

A energia eólica e outras fontes de energias sustentáveis surgem como uma excelente opção frente aos desafios que existem nos dias de hoje. Se por um lado é necessário produzir energia de forma barata e segura, por outro a utilização de combustíveis fósseis não é vista mais com os mesmos olhos de antigamente.

A geração de energia eólica, contudo, depende diretamente de um fator natural, o vento. Os valores de vento medidos em uma determinada região ao longo do tempo, entretanto, podem ser modelados como uma série temporal, que é qualquer conjunto de observações ordenadas no tempo em intervalos iguais.[1] Utilizando os conceitos e propriedades das séries temporais é possível analisar e estudar a aleatoriedade do vento, contribuindo para uma geração de energia mais rentável e segura.

### 1.1

#### Motivação

Um dos principais objetivos do estudo de séries temporais é construir modelos que demonstrem o comportamento e perfil do fenômeno estudado e, a partir desses dados, gerar previsões. Modelos de previsão são aplicados nas mais diversas áreas da Engenharia, Economia, Ciências Médicas, dentre outros, servindo como base para o planejamento, permitindo avaliar antecipadamente a demanda, projetar capacidade e necessidade de recursos, além de outras atividades essenciais.

Dado que o vento é um recurso natural variável no tempo, é necessário adotar medidas que ajudem a proporcionar reserva e armazenamento para compensar essa variabilidade. Dessa forma, as previsões das séries temporais de vento que consistem em utilizar o passado histórico da variável para projetar observações futuras, são uma grande arma para minimizar essa problemática. As previsões precisas de direção e velocidade do vento reduzem os encargos de desequilíbrio para os geradores, tornando esse tipo de energia mais competitiva no mercado.



Diversas técnicas para realizar a previsão de séries temporais são utilizadas, porém todas possuem um erro de previsão. O desafio é, então minimizar o erro de previsão, construindo um modelo que melhor se ajuste aos dados. A construção deste modelo não é uma tarefa simples e muitas vezes conduz a uma previsão pouco eficiente. Os métodos convencionais de previsão fornecem resultados precisos quando os dados estudados apresentam um comportamento linear, porém quando há um grau elevado de não linearidade, estes métodos passam a ser pouco eficientes. Neste sentido, pesquisas vêm sendo desenvolvidas, nas mais diversas áreas, para se obter modelos de previsão com o menor erro possível e melhores resultados.

Nesse cenário, o uso de Redes Neurais Artificiais se apresenta como uma ferramenta muito atrativa para solução desses casos. Redes Neurais Artificiais são modelos computacionais que buscam simular o processamento feito pelo cérebro humano e desempenhar uma determinada tarefa. Elas são compostas por estruturas denominadas neurônios, que são capazes de realizar conexões chamadas de sinapses e executar o processamento de informações. [2]

## 1.2 Objetivo

Usando técnicas de Redes Neurais Artificiais e do modelo NARMA é possível criar modelos de predição da velocidade e direção dos ventos, considerando que haja dados históricos. Dessa forma, o principal objetivo desse trabalho é construir um modelo de previsão que melhor se ajuste aos dados, reduzindo os erros de predição. Esse modelo utilizará como entrada um grande conjunto de dados reais medidos em intervalos de 10 minutos de sensores em 15 turbinas eólicas. Estes são originados do Sistema AMA (Acompanhamento de Medições Anemométricas), relativos aos Parques Eólicos no Rio Grande do Norte, de propriedade da Empresa Brookfield Energia Renovável.

## 1.3 Organização do Trabalho

Essa dissertação está dividida em 7 capítulos, iniciando por esta Introdução geral. No Capítulo 2 serão apresentados trabalhos da literatura relacionados a este trabalho e que contribuíram para o seu desenvolvimento, juntamente com uma breve descrição. Já no Capítulo 3 do trabalho serão apresentados o conceito e contextualização das séries temporais, além de breve teoria acerca de Redes Neurais Artificiais e modelo NARMA. No Capítulo 4 os dados disponíveis de velocidade e direção do vento serão descritos. No Capítulo 5 serão apresentados os resultados de predição da velocidade e direção do vento uti-

lizando Redes Neurais Artificiais. No Capítulo 6 serão listados os resultados alcançados utilizando RNAs juntamente com o modelo NARMA. Por fim, no Capítulo 7, serão apresentadas as conclusões e considerações finais, bem como ideias para trabalhos futuros.

## 2

### Revisão Bibliográfica

Para o desenvolvimento da previsão da velocidade e direção do vento, muitos estudos foram executados. Na referência [3] os autores implementaram Redes Neurais Artificiais (RNA) para predição da velocidade do vento a curto prazo em Mardin, Turquia. A rede neural multicamadas do tipo perceptron (*multilayer perceptron neural network* - MLP), foi usada para realizar a previsão por eles. Neste trabalho, diversas configurações de RNA com diferentes números de neurônios foram gerados e comparados através de medidas de erro, garantindo o desempenho e a precisão dos modelos escolhidos. Erro médio quadrático e coeficiente de correlação foram utilizados como parâmetros estatísticos para avaliar a rede nesse estudo, sendo os índices reportados respectivamente por 0,3780  $m/s$  e 0,9704  $m/s$ .

O artigo [4] apresentou métodos diferentes para realizar previsões da velocidade do vento a curto prazo e longo prazo. Este trabalho mostra duas metodologias híbridas inovadoras capazes de realizar previsões de velocidade do vento de curto e longo prazo a partir da junção matemática de dois modelos clássicos de séries temporais, a *Autoregressive Integrated Moving Average with Explanatory Variable* (ARIMAX) e os *Holt-Winters* (HW), ambos combinados com Redes Neurais Artificiais (RNAs). O primeiro modelo híbrido (ARIMAX e RNA) é feito das relações físicas entre pressão, temperatura e precipitação com a velocidade do vento, ou seja, este modelo é considerado multivariado. O segundo modelo híbrido (HW e RNA) é considerado como univariado, isto é, permitindo apenas entradas de velocidade do vento. Por meio de análise estatística de erro verifica-se que os modelos híbridos propostos oferecem ajustes perfeitos aos dados observados nas regiões de estudo, e assim, melhores comparações com as tradicionais da literatura. É possível encontrar nesta análise um erro percentual de 5,0% e um coeficiente de eficiência (Nash-Sutcliffe) de aproximadamente 0,96. A confirmação de precisão pelos modelos híbridos revela que eles fornecem séries temporais que são capazes de seguir os perfis de séries temporais observadas com semelhanças de valores máximos e mínimos entre as duas séries. Portanto, tornou-se um importante indicativo na representação das características de sazonalidade pelos modelos.

Já para a predição da direção do vento a curto prazo o artigo [5] propôs dois métodos baseados em Redes Neurais Artificiais (RNAs) e support vector machines (SVM). Neste artigo, a direção do vento medida nos últimos minutos é usada como entrada e a direção do vento para os próximos dois minutos

constitui a saída. Para ambos os métodos é otimizado o comprimento da média móvel usado para pré-processar os dados de entrada, o comprimento do vetor de entrada e, para a RNA apenas, o número de neurônios de cada camada. A previsão é avaliada olhando para o erro absoluto médio e um índice médio de efetividade, que avalia a porcentagem de vezes que a previsão é precisa o suficiente para prever a escolha tática correta. A previsão utilizando RNA com base na média do conjunto de dez redes mostra um erro absoluto médio maior e um índice médio de eficácia semelhante ao da previsão de SVM. No entanto, é mostrado que a precisão das previsões de RNA aumenta significativamente com o tamanho do conjunto. Portanto, aumentando o poder computacional, isso pode levar a uma melhor previsão.

Já no artigo [6], foram utilizados outros métodos de previsão de séries de vento. Neste estudo, três modelos de algoritmos de aprendizado de máquina são implementados para prever a velocidade do vento, a direção do vento e potência de saída de uma turbina eólica em Bushehr, Iran. O primeiro modelo é *multilayer perceptron neural network* (MLP) treinado com nove algoritmos de treinamento de dados. O segundo modelo é *support vector regression* com função de base radial (SVR-RBF). O terceiro modelo é o *adaptive neuro-fuzzy inference system* (ANFIS). Temperatura, pressão, tempo e umidade relativa da região são consideradas variáveis de entrada dos modelos. Modelos de previsão foram comparados com intervalos de 5 min, 10 min, 30 min e 1 h de velocidade do vento e sua direção. O estudo revelou que para os casos de dados da velocidade do vento em intervalos de 5 e 10 minutos, os valores previstos foram encontrados em excelente acordo com os dados reais dos três modelos desenvolvidos. Para outros intervalos de tempo, o modelo SVR-RBF funciona melhor que o MLP e ANFIS. Para a predição da direção do vento os modelos desenvolvidos mostraram desempenho inferior comparado com a previsão da velocidade do vento.

O artigo [7] apresenta de forma mais ampla a grande aplicabilidade das Redes Neurais Artificiais em sistema de energia eólica. Este artigo apresenta uma revisão exaustiva das redes neurais artificiais utilizadas em sistemas de energia eólica, identificando os métodos mais empregados em diferentes aplicações e demonstrando que as Redes Neurais Artificiais podem ser uma alternativa aos métodos convencionais em muitos casos. Os métodos são classificados e analisados em quatro grupos de acordo com a aplicação: previsão; otimização de design; detecção e diagnóstico de falhas; e controle ótimo. Associada a cada aplicação, o artigo faz uma análise sobre os pontos fortes e fracos de cada estrutura de RNA, além de descrever os principais desafios e lacunas tecnológicas em relação à aplicação de RNA às turbinas eólicas, de acordo com a revisão

da literatura. Uma tabela geral é fornecida para resumir as referências mais importantes de acordo com os grupos de aplicações e estudos de caso.

Este trabalho de pesquisa também faz um enfoque sobre o estudo dos problemas mais influentes e estudados nas aplicações de Redes Neurais Artificiais em cada um dos quatro grupos analisados. Para o grupo de predição, ele ressalta os estudos de velocidade e força do vento. Já para otimização de design a ênfase é dada para o design de turbinas e de parques eólicos, processos complexos dada a necessidade de considerar um grande número de variáveis. No grupo de detecção e diagnóstico de falhas são tratados os estudos de caixa de engrenagens e rolamentos; geradores; rotor, lâminas e hidráulico; e redução da taxa de alarmes falsos. Por último, no grupo de controle ótimo, os são abordados os temas de rastreamento de potência máxima; ângulo de *pitch*; velocidade; potência reativa; e conversor.

De acordo com o artigo [7] as RNAs demonstraram ser eficazes para prever a velocidade do vento e a produção de energia eólica. Os resultados mostrados na Seção 3 deste artigo determinam que as RNAs geralmente geram resultados com mais precisão do que outros métodos convencionais, por exemplo, métodos físicos ou estatísticos. No entanto, a precisão dos métodos baseados em RNA diminui significativamente para previsões de longo prazo. Além disso, o número de dados históricos requeridos pela RNA para previsão de longo prazo é alto. Um desafio neste campo é desenvolver métodos baseados em RNA flexíveis adaptados aos dados de entrada e ao horizonte de tempo desejado.

## 3

### Métodos

Neste capítulo será apresentado o conceito, contextualização e aplicabilidade das séries temporais, incluindo o tópico de previsão, ressaltando a importância de modelos de predição de séries temporais. Além disso, será introduzida a concepção de RNA e abordado o problema de previsão de séries de vento.

#### 3.1

##### Séries Temporais

Uma série temporal é um vetor  $\mathbf{Z}$  de observações ordenadas ao longo do tempo, em intervalos iguais [1],

$$\mathbf{Z} = \{Z(t), t = 1, 2, \dots, N\}$$

onde  $t$  é o índice de tempo e  $N$  é o número de observações.

As Séries Temporais podem ser contínuas ou discretas. No caso de um conjunto de dados com valores contínuos, este deve ser discretizado no tempo e com intervalos iguais

Dessa forma, este conjunto de dados difere dos demais por ser ordenado no tempo, ou seja, todos os valores são referentes ao mesmo fato, no entanto, em períodos distintos, sendo a ordem dos dados fundamental. Uma característica muito relevante dessa coleção de dados é que as observações vizinhas são dependentes e o interesse é analisar e modelar essa dependência.

Os métodos de análise para séries temporais auxiliam tanto na análise para a compreensão do objeto de estudo como fornecem instrumentais matemático-metodológicos para fazer projeções. Um dos principais objetivos do estudo de séries temporais é gerar previsões através da criação de modelos que demonstrem o comportamento do fenômeno estudado. Modelos de previsão são aplicados em diversas áreas da Engenharia, Economia, Ciências Médicas, dentre outros, servindo como base para o planejamento, permitindo avaliar antecipadamente a demanda, extração de informações, projeção da capacidade e necessidade de recursos, além de outras atividades. [8]

Existem diversos exemplos de séries temporais inseridas no cotidiano de qualquer cidadão, com os mais diferentes fins. Pode-se citar:

- Economia: preços diários das ações da empresa "X", taxa de desemprego em uma capital brasileira;

- Medicina: níveis de eletrocardiograma ou eletroencefalograma;
- Epidemiologia: casos semanais de sarampo, casos mensais de AIDS;
- Meteorologia: temperatura diária, registro de marés.

A análise de séries temporais é um importante instrumento no entendimento das mais diversas variáveis, seja dos casos semanais de sarampo até a taxa de desemprego mensal, como pode-se perceber pelos exemplos citados anteriormente. Dessa forma, obtida uma série temporal observada em  $N$  instantes de tempo, pode-se estar interessado em:

- Investigar o mecanismo que gerou a série temporal; por exemplo, analisando uma série de ventos, pode-se querer saber como o vento foi gerado;
- Fazer previsões de valores futuros da série; podendo ser a curto prazo, como o número de atendimentos em um Call Center esperados para o dia seguintes, ou a longo prazo, como para séries de natalidade, por exemplo;
- Apenas descrever o comportamento da série; verificando a existência de tendências, ciclos e variações sazonais, construção de gráficos e histogramas, etc;
- Procurar periodicidades relevantes nos dados.

Segundo a abordagem de componentes não observáveis, as séries temporais podem ser representadas como a combinação de quatro componentes [9]:

- Tendência: as componentes de tendência representam o comportamento e produzem mudanças graduais em longo prazo. São normalmente provocadas, por exemplo, pelo crescimento constante na população, mudança gradual de hábitos de consumo, no efeito da competição, ou por outros fatores que falham na tentativa de produzir mudanças repentinas, mas produzem variações graduais e regulares ao longo do tempo, afetando a variável de interesse em longo prazo;
- Cíclica: as componentes cíclicas são aquelas que provocam flutuações de subida e de queda nas séries, de forma suave e com certa periodicidade, ao longo da componente de tendência. Geralmente os efeitos cíclicos em uma série são causados por mudanças na demanda do produto, variações da economia como períodos de crescimento ou recessão e, em particular, pela inabilidade de se suprir as necessidades do consumidor;

- Sazonal: as componentes sazonais em uma série são aquelas oscilações de subida e de queda que sempre ocorrem em um determinado período do ano, do mês, da semana, do dia ou horário. A diferença essencial entre as componentes sazonais e cíclicas é que a primeira possui movimentos facilmente previsíveis, ocorrendo em intervalos regulares de tempo, por exemplo, ano a ano, mês a mês, semana a semana, ou mesmo dia a dia. Já os movimentos cíclicos tendem a ser irregulares, ocorrendo sobre um período de muitos anos;
- Erro: a quarta componente da série, chamada de componente de erro, apresenta movimentos ascendentes e descendentes da série após a ocorrência de um efeito de tendência, um efeito cíclico, ou de um efeito sazonal. Nas componentes de erro aparecem flutuações inexplicáveis de período curto, com deslocamento inesperado e geralmente são causadas, entre outros motivos, por eventos políticos e oscilações climáticas imprevisíveis.

De acordo com [1], existem dois enfoques usados na análise de séries temporais. Em ambos, o objetivo é o mesmo: a construção de modelos para as séries, com diferentes propósitos determinados. No primeiro enfoque, a análise é feito no domínio do tempo e os modelos propostos são modelos com número finito de parâmetros paramétricos, ou seja, paramétrico. No segundo, a análise é conduzida no domínio de frequências e os modelos propostos são modelos não-paramétricos.

Também de acordo com [1] pode-se classificar as séries temporais como: estacionárias, não estacionárias e ergódicas. Séries estacionárias são aquelas que se desenvolvem no tempo aleatoriamente ao redor de uma média constante, refletindo alguma forma de equilíbrio estável, ou seja, é aquela que possui uma tendência horizontal, não tem variações cíclicas ou sazonais, e o componente de erro desenvolve assim uma variância constante. Séries não estacionárias são aquelas em que se observa uma alteração na média se calculada para diferentes períodos da série temporal. Séries ergódicas são aquelas na qual apenas uma realização do processo estocástico é suficiente para se obter todas as estatísticas do mesmo.

### 3.1.1

#### Previsão de Séries Temporais

Com o crescimento e aumento da complexidade nos ramos econômicos, industriais e tecnológico, a necessidade de se efetuar previsões que auxiliem no planejamento empresarial torna-se cada vez mais importante. A previsão



de determinados fatores auxilia a tomada de decisões que poderão melhorar o desempenho das empresas ou até mesmo minimizar prejuízos. [10]

O histórico de uma observação pode ser utilizado na identificação de períodos de crescimento/decrescimento, sazonalidade e ainda para “prever” observações futuras. [11] escreve que a hipótese básica no uso das séries temporais é a de que os valores futuros das séries de demanda podem ser estimados com base nos valores passados da série existente sobre a própria demanda. Dessa forma, utiliza-se o passado histórico da variável para projetar observações futuras, fazendo uma previsão do comportamento nos próximos períodos.

Segundo [1], a finalidade dos métodos de predição é distinguir o padrão de qualquer ruído que possa estar contido nas observações, descobrindo o modelo da série histórica de dados e, posteriormente, projetá-lo para o futuro. Os modelos de previsão podem ser classificados em univariados, os quais têm a previsão dos valores futuros explicados somente pelos valores passados da própria série ou causais, os que levam em conta outras informações relevantes como influentes para a previsão de uma variável.

Existem muitos modelos de previsão baseados em séries temporais. Pode-se listar: Média Móvel Simples (MMS), Média Móvel Exponencialmente Ponderada com Tendência, Projeção de Tendência (PT), modelos de Decomposição Aditiva (DA) e multiplicativa (DM) e o modelo de Holt Winters (HW). Uma grande classe de modelos de previsão, tenta tratar ambas as causas de flutuações em séries de tempo e a das suavizações (média móvel e amortecimento exponencial). Técnicas específicas desse tipo assumem que os valores extremos da série representam a aleatoriedade e assim, por meio da suavização desses extremos, pode-se identificar o padrão básico.

Nos modelos MMS, a previsão para o período posterior é obtida por uma média aritmética dos  $n$  valores da demanda passados. Nos modelos de média móvel exponencialmente ponderada com tendência, também conhecidos como Double Exponential Smoothing (DSE), os dados mais recentes que integrarão a média passam a ter maior peso, com sua importância declinando exponencialmente à medida que esses dados vão se tornando ultrapassados. Nele, tanto a estimativa para o nível quanto para a tendência são suavizadas. Já os modelos de PT ajustam a linha de tendência dos dados matematicamente e os projeta para o futuro. Esses modelos de ajuste podem utilizar o método dos mínimos quadrados para ajustar a tendência de modo linear, quadrática, logística ou exponencial. O modelo DM estabelece que qualquer valor observado em uma série temporal é produto de quatro fatores de influência: componentes de tendência, componentes cíclicos, componentes irregulares e fatores sazonais.

Já os modelos DA, consideram qualquer valor observado como um somatório desses fatores. O modelo Holt Winters inclui a suavização da sazonalidade nos modelos DSE.

### 3.2

#### Redes Neurais Artificiais

As RNAs são estruturas complexas baseadas em neurônios biológicos. Essas estruturas fornecem uma boa solução para problemas que não podem ser analiticamente definidos.

Uma RNA consiste em neurônios que são unidades simples de processamento e conexões ponderadas entre eles. Os neurônios se unem por conexões chamadas sinapses, onde cada conexão é extremamente especializada e envia sinais de um neurônio para o outro. Inicialmente, a RNA recebe um conjunto de dados e começa um processo de treinamento para ajustar os pesos das interconexões entre os neurônios. O treinamento é dito supervisionado se a saída é conhecida, caso contrário, será nomeado treinamento não supervisionado.

O aprendizado dessas redes vem dos erros cometidos pelas mesmas. O resultado de saída delas é comparado com o valor desejado de saída (conhecido a priori). O erro cometido pelo sistema serve para que os pesos das ligações entre os neurônios artificiais sejam determinados novamente, até que o valor de saída se assemelhe com o desejado.

Um grande benefício da RNA é sua capacidade de generalizar. A generalização se refere ao fato de a rede neural produzir saídas adequadas para entradas que não estavam presentes durante o processo de aprendizagem (treinamento), tornando possível a resolução de problemas complexos de grande escala. Além disso, a RNA têm capacidade de adaptabilidade, sendo capaz de adaptar seus pesos sinápticos a modificações do meio ambiente. Dessa forma, uma rede treinada para operar em um ambiente específico pode ser facilmente retreinada para trabalhar com pequenas modificações nas condições de operação do ambiente.

O Perceptron mostrado na Figura 3.1 é a forma mais simples de Rede Neural Artificial. Nele existe apenas um neurônio que abrange todas as entradas, como pode ser observado na Figura 3.1. Cada entrada recebe um peso associado a ela, refletindo assim sua importância para a rede. Esse peso pode ser positivo (sinapse excitatória) ou negativo (sinapse inibitória). O valor denominado *net* é formado pela combinação linear das entradas com os pesos, mais um limiar, denominado bias. [2] Sendo assim, o valor denominado *net* é definido como:

$$u_k = \sum_{j=1}^m w_{kj}x_j + b_k \quad (3-1)$$

Onde  $w_{kj}$  são os pesos associados as entradas  $x_j$  e  $b_k$  o bias do neurônio  $k$ . Esse valor  $net$  será enviado para uma função de ativação, que será responsável por definir a saída do neurônio (valor  $y$ ).

O bias faz o papel de um fator de correção, aumentando ou diminuindo o valor de  $net$ . Essa alteração realizada tem por finalidade a aproximar a saída do neurônio ao valor esperado. A função de ativação pode ser a função degrau (neurônio de McCulloch e Pitts), função linear, função rampa e função sigmoidal. O domínio de  $y$  pode ser binário ( $y \in 0, 1$  ou  $y \in -1, 1$ ) ou contínuo ( $y \in \mathbb{R}$ ). [2] Dessa forma, a saída  $y$  será definida por:

$$y_k = \varphi(u_k) \quad (3-2)$$

Onde  $u_k$  foi definido na equação 3-1 e  $\varphi$  é uma função de ativação.

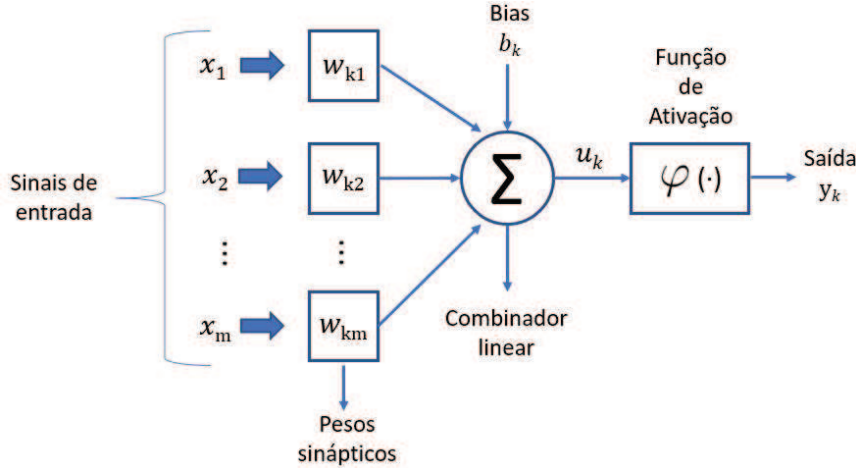


Figura 3.1: Arquitetura do Perceptron - fonte: própria.

No Perceptron mostrado na Figura 3.1 existe apenas um neurônio em uma única camada. De acordo com [12], um Perceptron de múltiplas camadas treinado com o algoritmo de retropropagação se manifesta como um esquema aninhado sigmóide, escrito na forma compacta para o caso de uma única saída:

$$F(x, w) = \varphi \left( \sum_k w_{ok} \varphi \left( \sum_j w_{kj} \varphi \left( \dots \varphi \left( \sum_i w_{li} x_i \right) \right) \right) \right) \quad (3-3)$$

onde  $\varphi$  é uma função de ativação sigmóide comum,  $w_{ok}$  é o peso sináptico de neurônio  $k$  na última camada oculta para o único neurônio de saída, e assim por diante para os outros pesos sinápticos, e  $x_i$  é o  $i$ -ésimo elemento do vetor de entrada  $x$ . O vetor peso  $w$  representa o conjunto inteiro de pesos sinápticos

ordenados por camada, por neurônios em uma camada e, então, por sinapses em um neurônio.

As RNAs podem ser utilizadas em diversos campos devido à suas inúmeras vantagens, como por exemplo: medicina, robótica, processamento de imagens, etc. As principais vantagens são:

- Aprendizagem adaptativa: Eles podem aprender a executar tarefas através de um processo de treinamento;
- Auto-organização: as RNAs podem criar sua própria estrutura para representar as informações através de um processo de treinamento;
- Tolerância a falhas: as RNAs ainda pode operar quando sua estrutura está danificada (tolerância à degradação) e distorcida ou incompleta quando os dados são ruidosos (tolerância a dados);
- Fácil implementação nos sistemas: existem chips especializados que podem facilitar a integração de RNAs nos sistemas.

Existem quatro variáveis básicas que caracterizam uma RNA: a topologia; o método de treinamento; o tipo de associação entre input e dados de saída e a apresentação da informação.

De acordo com [7], mais de 50 tipos de RNAs podem ser distinguidas, por exemplo: *multilayer perceptron* (MLP); *radial basis function neural network* (RBFNN); *backpropagation networks* (BPNN); *Wavelet neural network* (Wavelet NN); *self-organized-map NN* (SOM NN); *Recurrent NN*; *time delay NN*; *Hopfield network*; *auto-associative NN*; *convolutional NN*; *learning vector quantization networks*; *adaptive resonance theory* (ART) NN; *neuro-fuzzy networks*.

A seleção da Rede Neural mais adequada depende de fatores que podem ser classificados como:

- Fatores endógenos do problema: esses fatores estão relacionados ao natureza do problema, ou seja, aquelas características que são impostas o problema;
- Fatores exógenos do problema: aqueles parâmetros que são selecionados pela pessoa que resolve o problema. Por exemplo, o *timehorizon* de previsões, a robustez do método, a precisão da análise ou o custo computacional;
- Limitações: existem várias características que limitam a precisão de RNAs. As principais desvantagens são a quantidade de insumos necessários, *overtraining* das redes, os erros de extrapolação e as dificuldades de otimizar a rede.

### 3.3

#### Modelo NARMA

Muitos tipos de modelos não lineares têm sido propostos na literatura. O NARMA (*Nonlinear Autoregressive Moving Average*) é um deles, e trata-se de uma generalização não-linear do modelo ARMA. Dada uma série temporal de dados  $X_t$ , o modelo ARMA (*Autoregressive Moving Average*) é uma ferramenta para entender e, talvez, prever valores futuros nesta série. O modelo consiste em duas partes, uma parte auto-regressiva (AR) e uma parte de média móvel (MA). A parte AR envolve regressar a variável em seus próprios valores defasados, isto é, passados. A parte MA envolve modelar o termo de erro como uma combinação linear de termos de erro que ocorrem contemporaneamente e em vários momentos no passado. O modelo é geralmente referido como o modelo ARMA( $p, q$ ), em que  $p$  é a ordem da parte auto-regressiva e  $q$  é a ordem da parte de média móvel.

O modelo geral NARMA( $p, q$ ) pode ser definido por:

$$x(t) = e(t) + f(x(t-1)), \dots, x(t-p), e(t-1), \dots, e(t-q)) \quad (3-4)$$

onde  $e(t)$  é um ruído branco ( $\mu_e=0$  e variância  $\sigma_e^2$ ). Como no ARMA,  $p$  é a ordem da parte auto-regressiva e  $q$  é a ordem da parte de média móvel. O preditor NARMA correspondente é dado por:

$$\hat{x}(t) = e(t) + f(x(t-1)), \dots, x(t-p), e(t-1), \dots, e(t-q)) \quad (3-5)$$

e desde  $e(t-i)$ ,  $i = 1, \dots, q$  for aproximado usando  $\hat{e} = x(t) - \hat{x}(t)$ , tem-se:

$$\hat{x}(t) = e(t) + f(x(t-1)), \dots, x(t-p), \hat{e}(t-1), \dots, \hat{e}(t-q)) \quad (3-6)$$

$$\hat{e}(t-i) = x(t-i) - \hat{x}(t-i), \quad i = 1, \dots, q \quad (3-7)$$

O algoritmo FROLS é uma maneira eficaz de selecionar os termos em um modelo de parâmetros lineares, com o objetivo de contribuir para a minimização do erro. No entanto, o FROLS não é capaz de lidar diretamente com a estimativa dos termos MA nos modelos NARMA. Felizmente isso não é uma complicação, pois uma estratégia simples, mas eficaz, pode ser adotada para criar modelos NARMA baseados no algoritmo FROLS, empregando o Estratégia ELS-FROLS.

A estratégia ELS-FROLS consiste em 3 fases básicas: (i) seleção e estimação dos termos de processo, (ii) seleção e estimação dos termos de ruído,

(iii) modelagem estimativa dos parâmetros de termos selecionados. Os termos do processo contêm atrasos  $u(k)$ ,  $y(k)$  enquanto os termos de ruído têm pelo menos um componente  $e(k)$  defasada.

Assim, na fase (i), apenas os termos do processo são considerados e o processo FROLS é executado. No final desta fase, tem-se os termos do processo selecionados, sendo capaz de calcular uma previsão preliminar e assim inicializar os resíduos usados para os componentes de MA. Dessa forma, na fase (ii) prossegue-se com seleção de termos dos componentes de ruído com FROLS,  $e(k)$  foi obtida provisoriamente até ao final da fase (i). No final da fase (ii), sabe-se quais termos constituem o modelo final do NARMA, mas a estimação dos parâmetros ainda pode ser melhorada dado que o vetor residual foi uma estimativa até o final de fase (i). Então, na fase (iii), estima-se os parâmetros do modelo usando mínimos quadrados estendidos (ELS), que consiste em uma iteração para estimar parâmetros e atualizar o vetor residual.

- Passo 1: Definir os parâmetros para a estrutura do modelo: definir as ordens do modelo e grau de não linearidade;
- Passo 2: Seleção dos termos do processo: selecionar os termos do processo, de acordo com um limiar  $\rho_p$ , usando o algoritmo FROLS, e calcular um vetor de resíduos usando o modelo com os termos do processo;
- Passo 3: Seleção dos termos de ruído: selecionar os termos de ruído, de acordo com um limiar  $\rho_n$ , usando o algoritmo FROLS;
- Passo 4: Estimativa dos parâmetros para os termos selecionados: estimar os parâmetros do modelo para os termos selecionados. Executar para um número de iterações: estimar usando qualquer procedimento de mínimos quadrados o vetor de parâmetros do modelo selecionado e atualizar o vetor de resíduos.

O ELS na Etapa 4 requer normalmente poucas iterações, geralmente menos de 5, para convergir. O passo 4 deve parar quando o vetor de resíduos ou o vetor de parâmetros não mudam significativamente no processo. Também é importante para destacar que  $\rho_n$  deve ser menor que  $\rho_p$ , caso contrário nenhum termo MA é selecionado.  $\rho_p$  e  $\rho_n$  são, respectivamente, os *thresholds* para seleção dos termos de processo e ruído baseado na taxa de redução de erro.

### 3.4

#### Previsão de Séries com RNAs

Os dados de séries temporais são utilizados frequentemente como entradas para modelos de previsão baseados em RNAs. Entretanto, é necessário definir não só o horizonte de previsão mas também a janela a ser utilizada. O primeiro relaciona-se à quantidade de valores subsequentes que serão previstos; a segunda regulariza a quantidade de elementos anteriores ao que será previsto. O horizonte de previsão pode ser classificado como de curto, médio e longo prazo. Quanto maior o horizonte de previsão maior a chance de erros, logo, previsões de curto prazo possibilitam uma boa visão do futuro próximo com um baixo erro.

Quando se trata de previsão, a rede é aplicada iterativamente. Para a previsão um passo à frente, simplesmente usa-se as entradas históricas disponíveis. Para prever dois passos à frente é utilizada a previsão de uma etapa como entrada, juntamente com os dados históricos. Este processo continua até que se tenha computado todas as previsões necessárias.

A quantidade necessária de elementos anteriores ao que será previsto depende diretamente da ordem do modelo. Por exemplo, a saída  $\hat{y}_t$  para previsão um passo à frente pode ser descrita como:

$$\hat{y}_t = f(\mathbf{y}_{t-1}) \quad (3-8)$$

$$y_t = \hat{y}_t + \epsilon_t \quad (3-9)$$

onde  $f$  é a rede neural,  $\mathbf{y}_{t-1} = (y_{t-1}, y_{t-2}, \dots, y_{t-N})'$  é um vetor de entrada com  $N$  elementos, se a ordem do modelo for  $N$ , e  $\epsilon_t$  é o erro associado a previsão.

Para a previsão dois passos à frente ( $\hat{y}_{t+1}$ ), será utilizado a previsão de um passo à frente ( $\hat{y}_t$ ) juntamente com os dados históricos. Dessa forma, a saída  $\hat{y}_{t+1}$  pode ser descrita como:

$$\hat{y}_{t+1} = f(\mathbf{y}_t) \quad (3-10)$$

$$y_{t+1} = \hat{y}_{t+1} + \epsilon_{t+1} \quad (3-11)$$

onde  $f$  é a rede neural,  $\mathbf{y}_t = (\hat{y}_t, y_{t-1}, y_{t-2}, \dots, y_{t-N})'$  é um vetor de entrada com  $N+1$  elementos, se a ordem do modelo for  $N$ , e  $\epsilon_{t+1}$  é o erro associado a previsão.

Dessa forma, aplicando a rede iterativamente, para a previsão  $K$  passos à frente ( $\hat{y}_{t+(K-1)}$ ), serão utilizadas as previsões anteriores juntamente com os dados históricos. Dessa forma, a saída  $\hat{y}_{t+(K-1)}$  pode ser descrita como:

$$\hat{y}_{t+(K-1)} = f(\mathbf{y}_{t+(K-2)}) \quad (3-12)$$

$$y_{t+(K-1)} = \hat{y}_{t+(K-1)} + \epsilon_{t+(K-1)} \quad (3-13)$$

onde  $f$  é a rede neural,  $\mathbf{y}_{t+(K-2)} = (\hat{y}_{t+(K-2)}, \dots, \hat{y}_t, y_{t-1}, y_{t-2}, \dots, y_{t-N})'$  é um vetor de entrada com  $N+K-1$  elementos, se a ordem do modelo for  $N$ , e  $\epsilon_{t+(K-1)}$  é o erro associado a previsão.

### 3.5

#### Problema: Previsão de Séries de Vento

A previsão da produção de energia eólica é uma tarefa complexa, entretanto é crucial estabelecer um planejamento ideal para os fornecedores de energia, proprietários e operadores de parques eólicos, equipes de manutenção, etc. Por exemplo, os fornecedores de energia podem evitar a superprodução levando em conta os sistemas de armazenamento de energia ou coordenando a produção de energia eólica estimada e a procura; os operadores podem adotar estratégias para fazer ofertas nos mercados de eletricidade; as tarefas de manutenção podem ser programadas de acordo com as previsões, etc.

As RNAs demonstraram alta capacidade na modelagem de séries temporais para diferentes aplicações, sendo eficientes quando os processos físicos não são compreendidos ou são muito complexos. De acordo com o artigo [7] os principais parâmetros considerados para as previsões de séries de ventos são a velocidade e a produção de energia eólica.

A velocidade do vento é um parâmetro essencial para a operação dos parques eólicos. Os modelos mais importantes para a previsão da velocidade do vento são: os métodos físicos; os métodos estatísticos, onde o mais popular é o modelo ARIMA; os modelos inteligentes, onde os mais populares são baseados em RNAs; e os modelos de previsão híbrida, que combinam diferentes tipos de algoritmos.

A maioria dos estudos e métodos de pesquisa para previsão de velocidade do vento estão focados na previsão de muito curto (segundos-30 minutos) ou curto prazo (30 minutos-6 horas). Previsões de muito curto prazo são úteis para aplicações de controle de turbinas no intervalo de segundos, sendo o custo computacional um fator essencial para os modelos a serem usados. Já para previsões a longo prazo da velocidade do vento a literatura não é muito extensa devido à baixa precisão. Da mesma forma os métodos de RNA também são empregados principalmente para previsões de energia eólica de curto prazo porque a precisão diminui quando o horizonte de tempo é longo.



### 3.6

#### Métricas de Desempenho

Para avaliar a performance de previsões através de RNAs, existem diversos critérios. Neste presente trabalho, foi utilizado o critério do Coeficiente de Determinação, também chamado de  $R^2$ .

#### 3.6.1

##### Coeficiente de Determinação

Uma das formas de avaliar a qualidade do ajuste do modelo é através do coeficiente de determinação, também chamado de  $R^2$ . Basicamente, este coeficiente indica quanto o modelo foi capaz de explicar os dados coletados. O  $R^2$  varia entre 0 e 1, indicando, em percentagem, o quanto o modelo consegue explicar os valores observados. Quanto maior o  $R^2$ , mais explicativo é o modelo, melhor ele se ajusta à amostra.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3-14)$$

onde  $y_i$  é o valor observado,  $\bar{y}$  é a média das observações e  $\hat{y}_i$  é o valor estimado (previsão) de  $y_i$ .

### 3.7

#### Testes de Correlação para Modelos de Séries Temporais

A modelagem de séries temporais é um caso especial de identificação de entrada-saída onde o sinal de entrada é indisponível, não medido ou desconhecido. Os testes de validade para modelos de séries temporais são diferentes daqueles para modelos de entrada-saída, pois os testes envolvendo a entrada não podem ser usados. Segundo Billings e Tao [13], um grupo de testes que são eficazes para validação do modelo de séries temporais. Os três testes são descritos a seguir:

$$\begin{aligned} \phi_{\xi'\xi'}(\tau) &= \delta(\tau), \quad \forall \tau \\ \phi_{\xi'(\xi^2)'}(\tau) &= 0, \quad \forall \tau \\ \phi_{(\xi^2)'(\xi^2)'}(\tau) &= \delta(\tau), \quad \forall \tau \end{aligned} \quad (3-15)$$

onde  $\delta(\tau)$  é a função impulso unitário, também chamada delta de Dirac. Essa distribuição satisfaz:

$$\int_{-\infty}^{\infty} \delta(\tau) d\tau = 1 \text{ e } \delta(\tau) = 0, \text{ para todo } \tau \neq 0 \quad (3-16)$$

A função de correção  $\phi_{xy}(\tau)$  entre dois sinais  $x$  e  $y$  é definida como:

$$\begin{aligned}\phi_{xy}(\tau) &= \frac{\frac{1}{N} \sum_{k=1}^{N-\tau} [x(k) - \bar{x}][y(k + \tau) - \bar{y}]}{\sqrt{\frac{1}{N} \sum_{k=1}^N [x(k) - \bar{x}]^2} \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - \bar{y}]^2}} \\ \phi_{xy}(\tau) &= \frac{\sum_{k=1}^{N-\tau} [x(k) - \bar{x}][y(k + \tau) - \bar{y}]}{\sqrt{\sum_{k=1}^N [x(k) - \bar{x}]^2} \sqrt{\sum_{k=1}^N [y(k) - \bar{y}]^2}}\end{aligned}\tag{3-17}$$

Durante a realização da predição para um passo à frente, serão realizados os testes de correlação listados nas Equações 3-15.

## 4

### Dados Disponíveis

Os dados disponíveis que serão utilizados como entrada do modelo de predição da velocidade e direção do vento são originados do Sistema AMA (Acompanhamento de Medições Anemométricas), relativos aos Parques Eólicos no Rio Grande do Norte, de propriedade da Empresa Brookfield Energia Renovável. Este Parque Eólico conta com 15 *Wind Turbine Generator* (WTG) Vestas V100-2MW, totalizando 30 MW de potência instalada. As turbinas estão instaladas a 95 metros de altura, com diâmetro de pá de 100 metros.

Esses dados são referentes a um ano de medições (abril de 2017 a abril de 2018), coletados na base de 10 minutos. Para cada um dos 15 geradores, existem dados de velocidade do vento (em m/s) registrados no anemômetro instalado a 95 metros de altura de cada WTG e dados da direção da velocidade (em graus), também registrados a 95 metros de altura para cada WTG. Os dados escolhidos para serem utilizados no presente trabalho foi do gerador 1, de 01/06/2017 (12:40) até 18/06/2017 (14:10), consistindo em 2458 dados.

Os dados da velocidade variam de aproximadamente 3 a 15 m/s, durante os 17 dias de medição, tendo alguns picos e vales na amostra que se destacam, como é possível observar na Figura 4.1. De acordo com o histograma (Figura 4.2), a maioria dos valores da magnitude da velocidade estão concentrados no intervalo de 7 a 9 m/s.

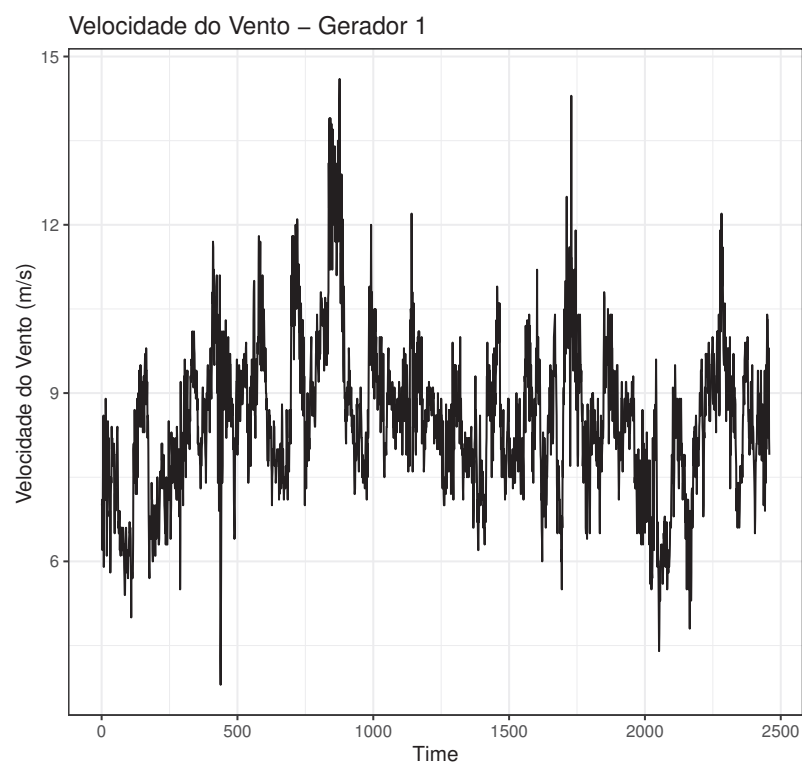


Figura 4.1: Amostra utilizada da velocidade do vento do Gerador 1.

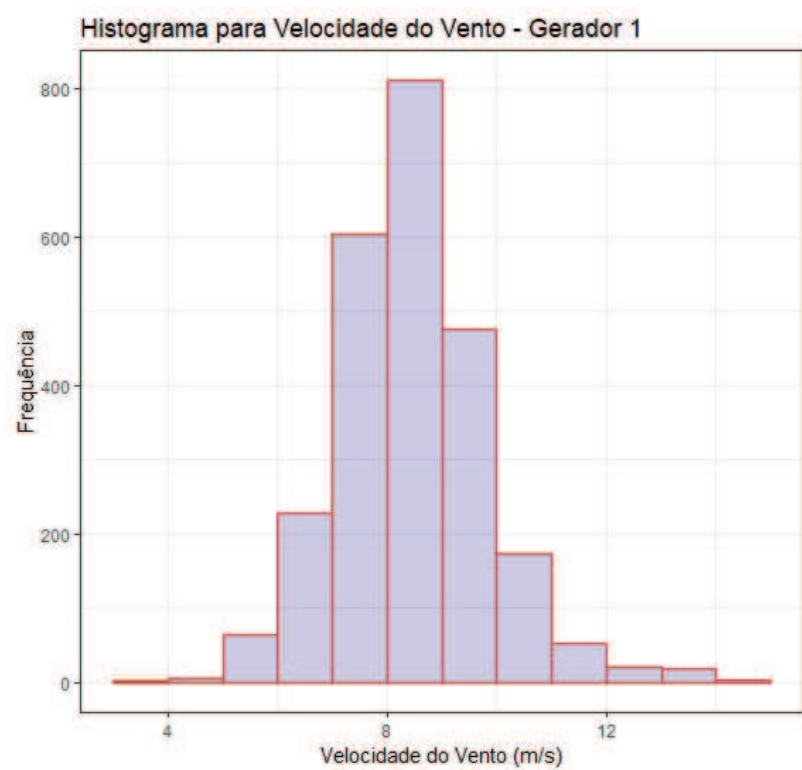


Figura 4.2: Frequência dos dados para a velocidade do vento.

Já os dados de direção do vento variam de aproximadamente  $100^\circ$  a  $200^\circ$ , durante os mesmos 17 dias de medição. Assim como foi possível observar nos dados da velocidade, os valores de direção do vento também apresentam alguns picos e vales, entretanto menos relevantes, como é possível observar na Figura 4.3. De acordo com o histograma (Figura 4.4), a maioria das medidas de direção estão concentradas no intervalo de  $150^\circ$  a  $180^\circ$ .

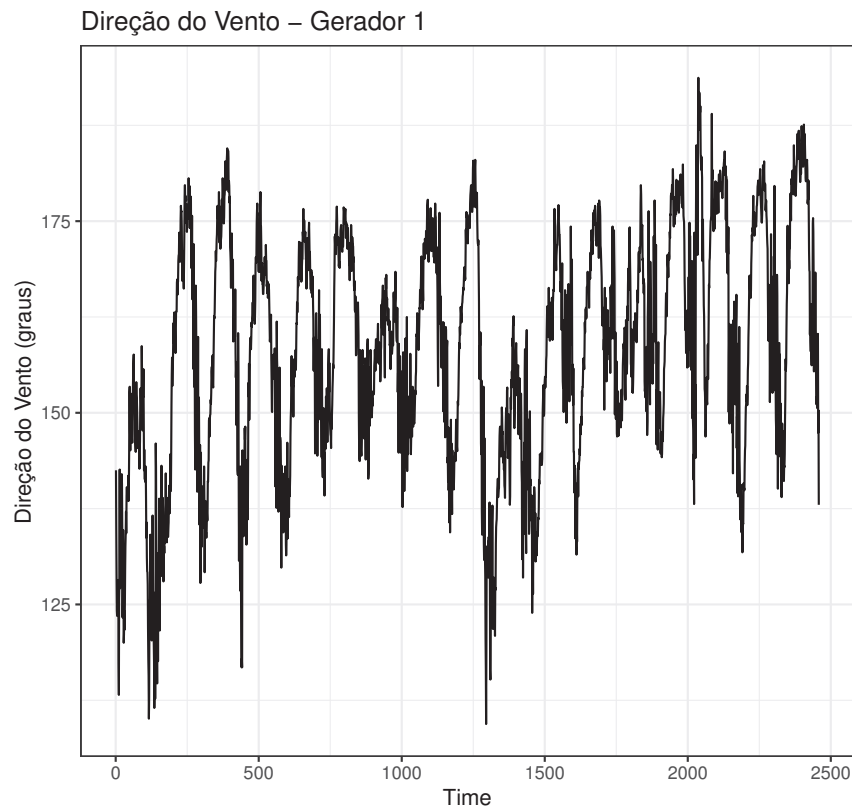


Figura 4.3: Amostra utilizada da direção do vento do Gerador 1.

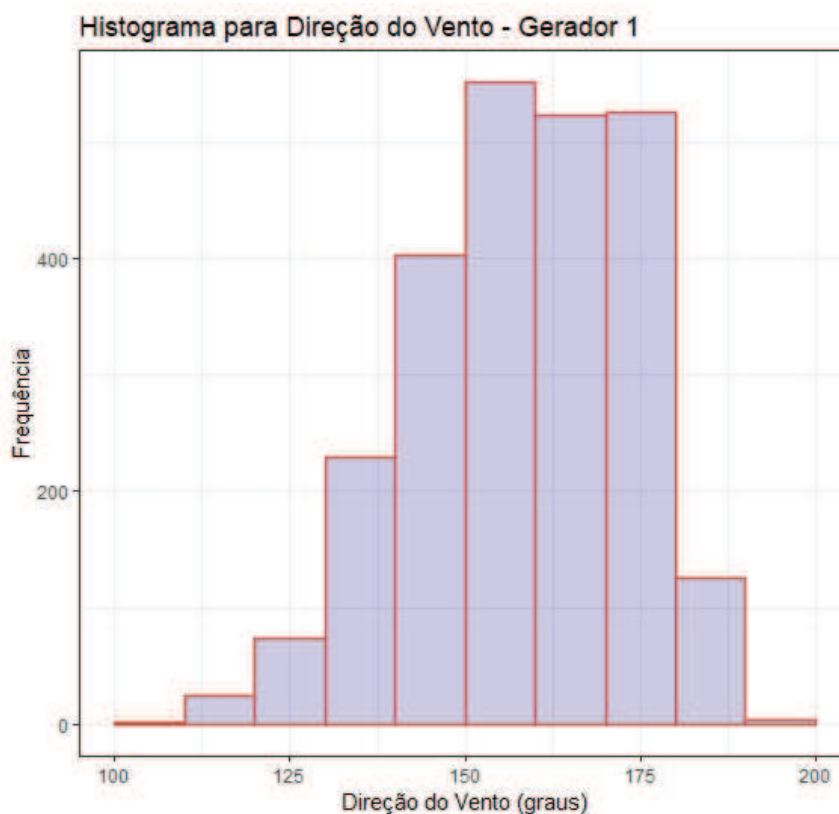


Figura 4.4: Frequência dos dados para a direção do vento.

#### 4.1

##### Autocorrelação dos Dados Disponíveis

Em uma série temporal, não há como desconsiderar a estrutura de dependência das observações. Dessa forma, é importante realizar uma análise da autocorrelação de uma série, para entender, principalmente sua aleatoriedade, já que várias técnicas partem dessa premissa. Autocorrelação é simplesmente a correlação entre uma série e ela mesma defasada. Ou seja, é a correlação entre os valores da série em um determinado período de tempo, e os valores da mesma série em um outro momento no tempo, relacionando uma observação num determinado instante com as observações passadas.

O correlograma é o gráfico utilizado em séries temporais para traçar as autocorrelações em diversas defasagens. A análise desse gráfico permite entender se a série é aleatória ou possui alguma tendência ou sazonalidade. No gráfico, o eixo vertical indica a autocorrelação e o horizontal a defasagem. A linha tracejada azul indica onde é significativamente diferente de zero. Quando todos os valores estiverem dentro do limite da linha tracejada azul a autocorrelação é igual a zero, indicando que a série é aleatória.

A seguir, nas Figuras 4.5 e 4.6 será possível observar o correlograma para velocidade e direção, respectivamente. Em ambos os casos, a defasagem (eixo horizontal denominado de *lag*) utilizada foi de 300.

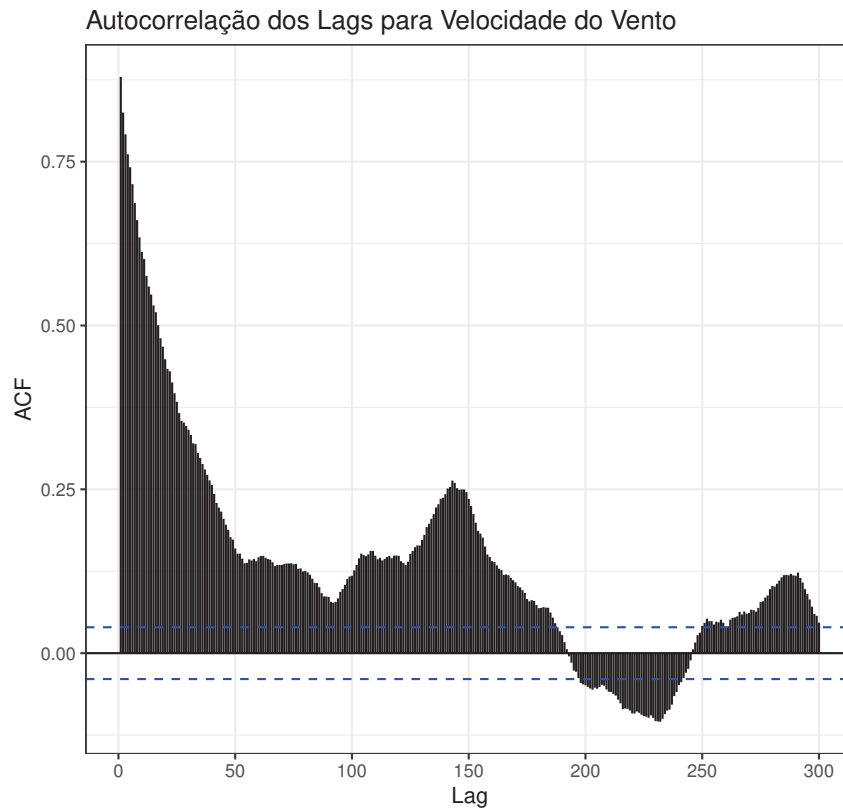


Figura 4.5: Correlograma para Velocidade do Vento.

Observando a Figura 4.5, conforme era esperado, o  $lag=0$  teve autocorrelação igual a 1, pois nesse caso está sendo realizada a correlação da série com ela mesma, sem nenhuma defasagem, constituindo duas séries iguais. Com o aumento da defasagem, a autocorrelação diminuiu, até que volta a subir novamente perto do  $lag=120$ . Em nenhuma defasagem observada o valor de autocorrelação está dentro da linha azul pontilhada, indicando a ausência de aleatoriedade. Entretanto, não é possível observar um padrão de sazonalidade para o aumento e diminuição da autocorrelação.

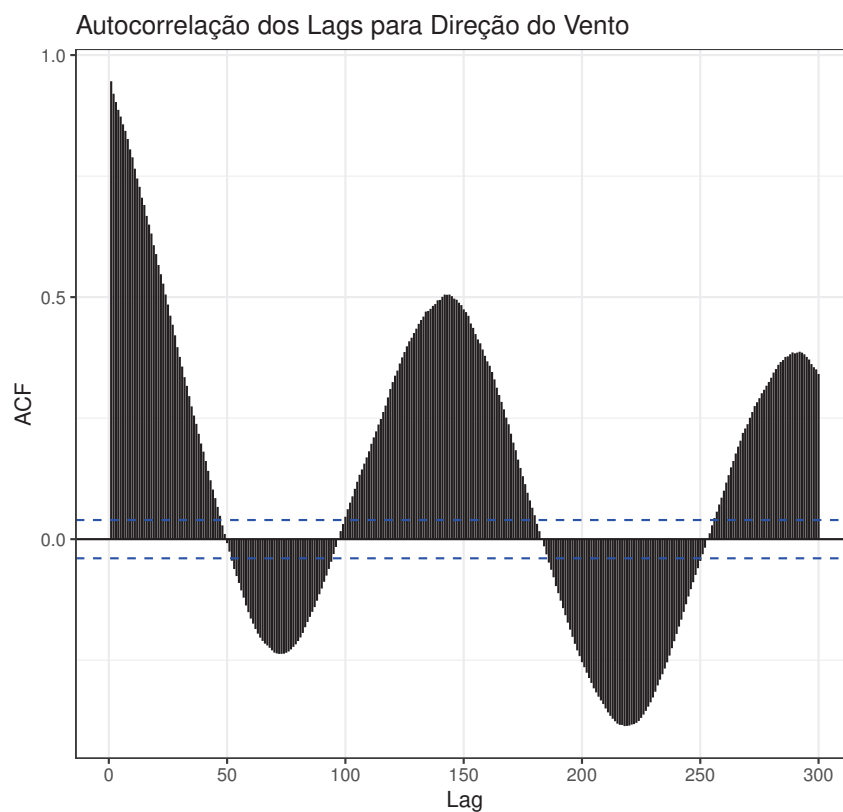


Figura 4.6: Correlograma para Direção do Vento.

Examinando a Figura 4.6, o  $lag=0$  também teve autocorrelação igual a 1, como era esperado e deveria acontecer. Diferentemente do correlograma da velocidade, o da direção apresenta uma sazonalidade, embora sem um padrão muito evidente. Contudo, é possível observar que a autocorrelação observada perto do  $lag=140$  é bem alta, indicando que a série contém muita informação para essa defasagem. Dessa forma, possivelmente se aumentasse a complexidade aumentando a ordem da RNA utilizada para predição, bons resultados seriam encontrados, já que a série tem esse comportamento.



## 5

### Resultados para Rede Neural Artificial

Neste capítulo, serão discutidos os resultados encontrados para a predição das séries temporais da velocidade e direção do vento utilizando RNAs. Será apresentado a busca de arquiteturas da rede, além de transformações utilizadas nos dados com o objetivo de melhorar os resultados alcançados.

#### 5.1

##### Busca de Arquiteturas para a Rede Neural Artificial

O funcionamento das RNAs depende de uma série de parâmetros que tem influência direta sobre sua eficiência e sua eficácia na resolução dos problemas a que se propõem resolver. Dessa forma, o objetivo dos resultados listados nessa seção é avaliar quais parâmetros colaboram para a melhor configuração da Rede Neural Artificial utilizada para a predição de séries de ventos.

A RNA utilizada recebe como *input* os seguintes parâmetros: ordem, número de camadas, número de neurônios em cada camada, função de ativação, *learning rate*, *epochs*, *batch size* e *verbose*. Sendo assim, existem diversas combinações possíveis para os parâmetros. Durante os testes preliminares, entretanto, os parâmetros alterados foram:

- Ordem [1:10]
- Número de camadas [1:5]
- Número de neurônios em cada camada [16,32,64]
- *Learning Rate* [ $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ ]

Os parâmetros que permaneceram fixos durante os testes, foram:

- Função de ativação: Relu
- *Epochs*: 100
- *Batch Size*: 2
- *Verbose*: 1

Vale ressaltar que em todos os resultados obtidos no Capítulo 5 deste trabalho, antes que os dados fossem apresentados à RNA, foi utilizado um processo de normalização, no qual o vetor com os dados de entrada foi normalizado ao longo de intervalo [0,1] através da seguinte operação:

$$y_n = \frac{\mathbf{y} - \min(\mathbf{y})}{\max(\mathbf{y}) - \min(\mathbf{y})} \quad (5-1)$$

Onde  $\mathbf{y}$  é o vetor de dados original sem normalização,  $y_n$  é o vetor de dados normalizados,  $\min(\mathbf{y})$  e  $\max(\mathbf{y})$  são o menor e o maior elemento, respectivamente. A normalização entre  $[0,1]$  é feita para facilitar o procedimento de otimização que ajusta os parâmetros da RNA. Nos resultados finais, contudo, os dados são exibidos em sua escala original.

Em todo o trabalho, as previsões sempre são realizadas um e seis passos à frente. Um passo à frente ( $k=1$ ) consiste numa previsão de 10 minutos, dado que os valores reais de velocidade e direção foram tomados de 10 em 10 minutos. Já seis passos à frente ( $k=6$ ) é uma previsão de 1 hora. De acordo com [7], um horizonte de até 30 minutos se enquadra numa predição de muito curto prazo, enquanto o range de 30 minutos-6 horas se enquadra em curto prazo.

### 5.1.1

#### Resultados para a Predição da Velocidade do Vento com RNAs

A seguir, para cada um dos três valores de *learning rate*, existe uma tabela com os Coeficientes de Determinação ( $R^2$ ) para a predição da velocidade, variando a ordem [1:10], o número de camadas [1:5] e o número de neurônios em cada camada [16,32,64], com a função de ativação *relu* e para um passo à frente ( $k=1$ ) ou seis passos à frente ( $k=6$ ). As tabelas seguem a seguinte ordem:

- Tabela 5.1  $\implies$  *Learning Rate* =  $10^{-2}$  e um passo à frente ( $k=1$ );
- Tabela 5.2  $\implies$  *Learning Rate* =  $10^{-2}$  e seis passos à frente ( $k=6$ );
- Tabela 5.3  $\implies$  *Learning Rate* =  $10^{-3}$  e um passo à frente ( $k=1$ );
- Tabela 5.4  $\implies$  *Learning Rate* =  $10^{-3}$  e seis passos à frente ( $k=6$ );
- Tabela 5.5  $\implies$  *Learning Rate* =  $10^{-4}$  e um passo à frente ( $k=1$ );
- Tabela 5.6  $\implies$  *Learning Rate* =  $10^{-4}$  e seis passos à frente ( $k=6$ ).

A observação das 6 tabelas acima listadas evidenciou a melhora dos  $R^2$  com o aumento da complexidade relacionada à diminuição do *learning rate*. Quando menor o *learning rate*, melhores foram o  $R^2$ . Dessa forma, provavelmente é possível alcançar resultados ainda melhores com o *learning rate* menor que  $10^{-4}$  ou aumentando-se a arquitetura da rede.

Após a realização dos testes preliminares, foi escolhido um conjunto de parâmetros que resultava em um dos melhores  $R^2$  para a predição da

velocidade. Neste caso, os parâmetros que resultaram no melhor resultado foram: *learning rate* =  $10^{-4}$ , ordem = 9 e 2 camadas com 64 neurônios cada (tabela 5.6).

Para esse conjunto de parâmetros escolhidos, foi plotado o gráfico da predição para um e seis passos à frente juntamente com o valor real e o gráfico do erro ( $y(t) - \hat{y}(t)$ ) - Figuras 5.1 e 5.2, além de um gráfico de  $y(t)$  versus  $\hat{y}(t)$  com a reta  $y = x$  (Figura 5.3), sendo possível observar a dispersão dos dados reais e estimados.

Coeficientes de Determinação para Velocidade - learning rate = $10^{-2}$						
Ordem	Neurônios	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,7515	0,5326	0,6527	0,2376	0,0337
2	16	0,7574	0,7070	0,5863	0,4505	0,0178
3	16	0,7565	0,6916	0,5815	0,0479	0,0570
4	16	0,7796	0,7673	0,7084	0,0025	0,0366
5	16	0,7819	0,7750	0,6808	0,0204	0,0176
6	16	0,7460	0,7652	0,7664	0,0258	0,5327
7	16	0,7550	0,6976	0,7447	0,0108	0,1613
8	16	0,7759	0,7736	0,7193	0,0794	0,0164
9	16	0,7156	0,7725	0,6761	0,0687	0,0212
10	16	0,7952	0,7609	0,7615	0,0510	0,1211
1	32	0,7528	0,7521	0,7491	0,2185	0,0574
2	32	0,7754	0,7523	0,6094	0,7258	0,0095
3	32	0,7522	0,7349	0,6043	0,0021	0,0032
4	32	0,7771	0,7676	0,7802	0,3846	0,0380
5	32	0,7497	0,7858	0,7873	0,4502	0,0193
6	32	0,7864	0,7861	0,6963	0,1253	0,0001
7	32	0,7865	0,7486	0,7854	0,1384	0,0027
8	32	0,7622	0,6861	0,7141	0,3411	0,0994
9	32	0,7407	0,7316	0,7419	0,5165	0,0483
10	32	0,7643	0,6655	0,7681	0,1923	0,0648
1	64	0,6018	0,7416	0,7275	0,7614	0,1832
2	64	0,7691	0,7575	0,7264	0,7157	0,0072
3	64	0,7892	0,6342	0,7500	0,7760	0,0053
4	64	0,7621	0,6408	0,7377	0,7699	0,0154
5	64	0,7780	0,6940	0,7436	0,7523	0,0016
6	64	0,7194	0,7796	0,7733	0,7393	0,0727
7	64	0,7826	0,7403	0,7625	0,7763	0,0081
8	64	0,7572	0,7204	0,7831	0,6680	0,1553
9	64	0,7942	0,7792	0,7813	0,7170	0,0382
10	64	0,7828	0,7891	0,6871	0,7714	0,1188

Tabela 5.1: Resultados de coeficiente de determinação para velocidade obtidos com learning rate =  $10^{-2}$ , para um passo à frente.

Coeficientes de Determinação para Velocidade - learning rate = $10^{-2}$						
Ordem	Neurônios	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,0088	-2,8803	-0,9797	0,1462	0,0383
2	16	0,0466	-2,1186	-6,4368	0,2349	0,0234
3	16	0,2270	0,5705	-8,8003	0,0472	0,0563
4	16	0,3588	0,3504	0,4955	0,0024	0,0011
5	16	0,3965	0,3526	0,0230	0,0201	0,0173
6	16	0,2939	0,2007	0,3983	0,0263	0,5770
7	16	0,1099	0,4858	0,0407	0,0105	0,1599
8	16	0,0915	0,2456	0,7624	0,0398	0,0160
9	16	0,7796	0,3074	0,0296	0,0696	0,0217
10	16	0,4892	0,3813	0,3093	0,0517	0,1201
1	32	0,1049	0,2545	0,1889	0,1363	0,0638
2	32	0,3479	0,1090	-2,3487	0,2797	0,0098
3	32	0,0994	0,0569	-5,1304	0,0023	0,0031
4	32	0,3491	0,1220	0,4281	0,2559	0,0073
5	32	0,2362	0,4230	0,4619	0,2541	0,0196
6	32	0,4123	0,4302	0,1971	0,0659	0,0002
7	32	0,4401	0,3176	0,4362	0,2360	0,0025
8	32	0,0072	-1,3864	-1,0897	0,1528	0,0984
9	32	0,1123	0,3654	0,2438	0,1611	0,0475
10	32	0,1954	-1,4382	0,2349	0,1150	0,0656
1	64	-4,0349	0,2293	0,2812	0,4306	0,1479
2	64	0,2596	0,0089	0,1620	0,2978	0,0069
3	64	0,4978	-2,3261	0,1683	0,3871	0,0294
4	64	0,1254	-5,0245	0,4444	0,4093	0,0156
5	64	0,3599	-1,2365	0,0062	0,3137	0,0016
6	64	0,5415	0,3513	0,3829	0,3722	0,0184
7	64	0,3734	0,2096	0,2632	0,4313	0,0084
8	64	0,0138	0,3918	0,4317	0,1910	0,1541
9	64	0,5194	0,3934	0,3743	0,2255	0,0481
10	64	0,4125	0,4792	-0,8376	-1,3369	0,0287

Tabela 5.2: Resultados de coeficiente de determinação para velocidade obtidos com learning rate =  $10^{-2}$ , para seis passos à frente.

Coeficientes de Determinação para Velocidade - learning rate = $10^{-3}$						
Ordem	Neurônios	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,7734	0,7512	0,7654	0,7247	0,7727
2	16	0,7758	0,7831	0,7862	0,7828	0,7863
3	16	0,7914	0,7916	0,7936	0,7883	0,7808
4	16	0,7923	0,7885	0,7531	0,7814	0,7647
5	16	0,7918	0,7950	0,7920	0,7437	0,7921
6	16	0,7874	0,7820	0,7825	0,7921	0,7917
7	16	0,7869	0,7933	0,7811	0,7830	0,7866
8	16	0,7959	0,7981	0,7937	0,7764	0,7706
9	16	0,7932	0,7688	0,7968	0,6963	0,7942
10	16	0,7931	0,7944	0,7918	0,7852	0,7505
1	32	0,7601	0,7737	0,7700	0,7739	0,7613
2	32	0,7477	0,7824	0,7634	0,7714	0,7804
3	32	0,7766	0,7920	0,7808	0,7512	0,7761
4	32	0,7736	0,7941	0,7908	0,7906	0,7866
5	32	0,7838	0,7949	0,7478	0,7923	0,7639
6	32	0,7939	0,7969	0,7950	0,7944	0,7797
7	32	0,7823	0,7823	0,7971	0,7543	0,7797
8	32	0,7638	0,7962	0,7660	0,7978	0,7870
9	32	0,7863	0,7983	0,7936	0,7975	0,7806
10	32	0,7911	0,7980	0,7980	0,7886	0,7591
1	64	0,7548	0,7707	0,7726	0,7506	0,7672
2	64	0,7688	0,7880	0,7863	0,7795	0,7193
3	64	0,7918	0,7868	0,7675	0,7904	0,7707
4	64	0,7907	0,7928	0,7878	0,7160	0,7924
5	64	0,7957	0,7962	0,7969	0,7696	0,7817
6	64	0,7921	0,7936	0,7924	0,7964	0,7699
7	64	0,7954	0,7934	0,7864	0,7828	0,7792
8	64	0,7975	0,7568	0,7933	0,7958	0,7808
9	64	0,7721	0,7970	0,7840	0,7970	0,7935
10	64	0,7788	0,7870	0,7988	0,8013	0,7791

Tabela 5.3: Resultados de coeficiente de determinação para velocidade obtidos com learning rate =  $10^{-3}$ , para um passo à frente.

Coeficientes de Determinação para Velocidade - learning rate = $10^{-3}$						
Ordem	Neurônios	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,4496	0,0063	0,2588	-0,3679	0,3945
2	16	0,3527	0,4229	0,4937	0,4368	0,5171
3	16	0,5210	0,5145	0,5162	0,5025	0,3197
4	16	0,5102	0,4622	-0,0312	0,3761	0,1489
5	16	0,4729	0,5349	0,4957	-0,1040	0,5340
6	16	0,4295	0,3246	0,3758	0,5082	0,5249
7	16	0,4196	0,4863	0,3214	0,3840	0,4838
8	16	0,5317	0,5367	0,5065	0,2808	0,3003
9	16	0,5039	0,1918	0,5175	-0,9321	0,5238
10	16	0,4892	0,5053	0,4297	0,4102	-0,0426
1	32	0,1818	0,4086	0,3309	0,4640	0,2061
2	32	0,1518	0,4411	0,0545	0,2085	0,4213
3	32	0,2485	0,5060	0,3287	-0,1145	0,3465
4	32	0,2007	0,5208	0,4661	0,4836	0,4742
5	32	0,4011	0,5280	-0,1422	0,4996	0,1450
6	32	0,5195	0,5371	0,5073	0,5297	0,4136
7	32	0,3531	0,3267	0,5143	0,0943	0,3489
8	32	0,1233	0,4773	0,0378	0,5388	0,4161
9	32	0,3904	0,5371	0,4323	0,5286	0,3994
10	32	0,4368	0,5336	0,5134	0,4897	0,2127
1	64	0,0730	0,4150	0,4129	0,0757	0,3691
2	64	0,0780	0,5089	0,4954	0,3848	-0,0991
3	64	0,4860	0,4375	0,1546	0,4684	0,3133
4	64	0,5122	0,5019	0,4202	-0,7318	0,5241
5	64	0,5384	0,5264	0,5208	0,0643	0,3842
6	64	0,4733	0,5170	0,4719	0,5283	0,3449
7	64	0,5318	0,4685	0,4101	0,3800	0,3805
8	64	0,5337	0,1019	0,4602	0,5193	0,4132
9	64	0,2137	0,4974	0,2943	0,4776	0,4960
10	64	0,2627	0,3495	0,5199	0,5211	0,3629

Tabela 5.4: Resultados de coeficiente de determinação para velocidade obtidos com learning rate =  $10^{-3}$ , para seis passos à frente.

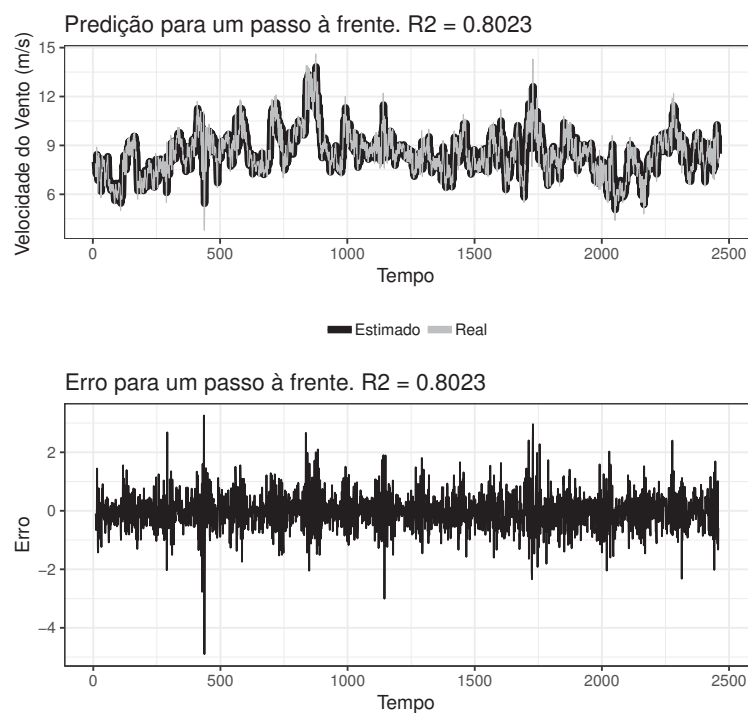
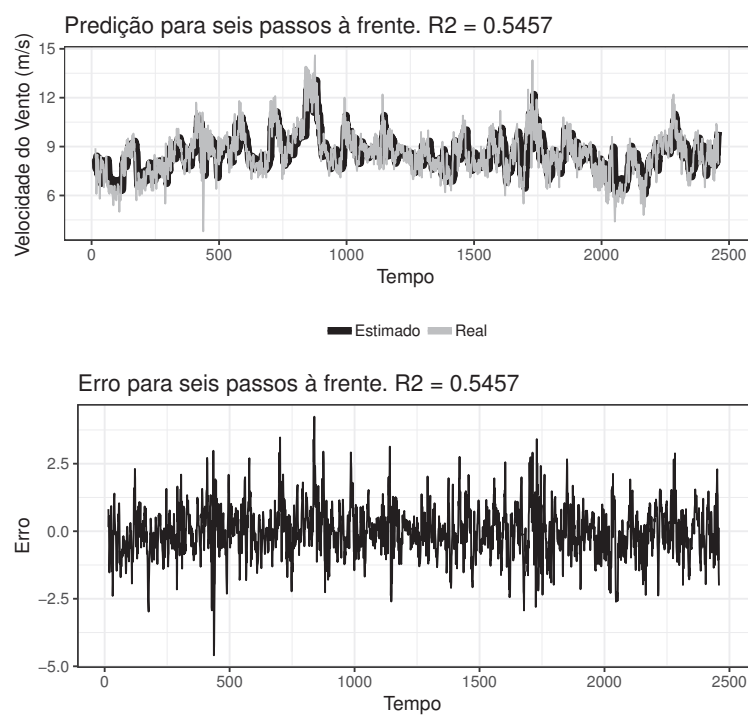
Coeficientes de Determinação para Velocidade - learning rate = $10^{-4}$						
Ordem	Neurônios	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,7741	0,7699	0,7705	0,7730	0,7608
2	16	0,7865	0,7866	0,7883	0,7839	0,7863
3	16	0,7874	0,7896	0,7821	0,7822	0,7923
4	16	0,7920	0,7927	0,7933	0,7908	0,7905
5	16	0,7907	0,7937	0,7900	0,7939	0,7864
6	16	0,7926	0,7868	0,7942	0,7929	0,7886
7	16	0,7932	0,7951	0,7931	0,7933	0,7895
8	16	0,7919	0,7961	0,7954	0,7830	0,7976
9	16	0,7947	0,7827	0,7956	0,7940	0,7821
10	16	0,7880	0,7971	0,7966	0,7949	0,7969
1	32	0,7740	0,7723	0,7726	0,7413	0,7726
2	32	0,7866	0,7869	0,7871	0,7867	0,7879
3	32	0,7917	0,7904	0,7936	0,7952	0,7915
4	32	0,7930	0,7935	0,7932	0,7801	0,7947
5	32	0,7912	0,7900	0,7924	0,7812	0,7871
6	32	0,7938	0,7946	0,7927	0,7871	0,7939
7	32	0,7959	0,7984	0,7998	0,7905	0,7907
8	32	0,7914	0,7982	0,7974	0,7739	0,8020
9	32	0,7967	0,7984	0,7994	0,8035	0,8019
10	32	0,7894	0,7995	0,7986	0,8015	0,7955
1	64	0,7732	0,7613	0,7736	0,7567	0,7707
2	64	0,7862	0,7816	0,7796	0,7876	0,7834
3	64	0,7927	0,7933	0,7894	0,7889	0,7887
4	64	0,7862	0,7932	0,7938	0,7968	0,7956
5	64	0,7949	0,7826	0,7962	0,7896	0,7933
6	64	0,7942	0,7985	0,7890	0,8024	0,7756
7	64	0,7901	0,7994	0,7657	0,7996	0,7837
8	64	0,7962	0,7960	0,8023	0,8050	0,7948
9	64	0,7854	0,8035	0,8023	0,8001	0,8082
10	64	0,7948	0,7980	0,7999	0,7912	0,8013

Tabela 5.5: Resultados de coeficiente de determinação para velocidade obtidos com learning rate =  $10^{-4}$ , para um passo à frente.



Coeficientes de Determinação para Velocidade - learning rate = $10^{-4}$						
Ordem	Neurônios	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,4594	0,3795	0,3761	0,4287	0,1829
2	16	0,4898	0,5018	0,5171	0,4473	0,5125
3	16	0,4889	0,5266	0,3613	0,3244	0,5287
4	16	0,5236	0,5328	0,5296	0,5236	0,4865
5	16	0,4995	0,5326	0,4247	0,5355	0,3988
6	16	0,5327	0,4347	0,5346	0,5301	0,4650
7	16	0,5351	0,5341	0,5377	0,5087	0,4380
8	16	0,5184	0,5369	0,5201	0,3910	0,5323
9	16	0,5294	0,3573	0,5134	0,5081	0,3475
10	16	0,4576	0,5304	0,5374	0,5081	0,5159
1	32	0,4542	0,4078	0,4707	-0,1462	0,3744
2	32	0,5125	0,4907	0,5249	0,4873	0,4971
3	32	0,5271	0,5061	0,5336	0,5292	0,5062
4	32	0,5307	0,5091	0,4968	0,2934	0,5381
5	32	0,5248	0,4717	0,4923	0,2783	0,3995
6	32	0,5344	0,5197	0,4573	0,4053	0,4922
7	32	0,5341	0,5423	0,5429	0,3969	0,3816
8	32	0,4843	0,5406	0,5114	0,2444	0,5423
9	32	0,5182	0,5341	0,5256	0,5255	0,5377
10	32	0,4702	0,5345	0,5249	0,5349	0,4870
1	64	0,4366	0,2007	0,4423	0,0921	0,3691
2	64	0,5012	0,4194	0,3751	0,4939	0,4031
3	64	0,5274	0,5049	0,4820	0,4510	0,4360
4	64	0,4348	0,4702	0,5026	0,5367	0,5318
5	64	0,5254	0,3376	0,4806	0,3692	0,4269
6	64	0,5292	0,5341	0,4063	0,5342	0,2093
7	64	0,4486	0,5400	-0,0790	0,4977	0,3320
8	64	0,5374	0,4885	0,4902	0,5309	0,4290
9	64	0,3549	0,5457	0,5297	0,4769	0,5434
10	64	0,4805	0,4579	0,4780	0,3421	0,4909

Tabela 5.6: Resultados de coeficiente de determinação para velocidade obtidos com learning rate =  $10^{-4}$ , para seis passos à frente.

Figura 5.1: Predição da velocidade do vento ( $k=1$ ) e erro utilizando RNA.Figura 5.2: Predição da velocidade do vento ( $k=6$ ) e erro utilizando RNA.

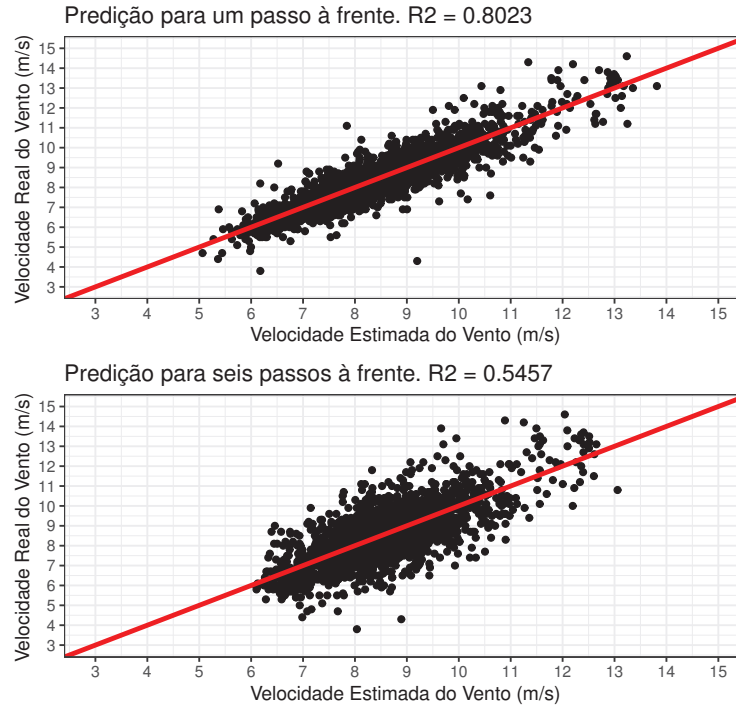


Figura 5.3: Gráfico de  $y(t)$  versus  $\hat{y}(t)$  para velocidade utilizando RNA.

As figuras anteriores são os melhores resultados alcançados para predição da série de velocidade um e seis passos à frente ( $Learning\ rate = 10^{-4}$ , ordem = 9 e 2 camadas com 64 neurônios cada). Na Figura 5.1 é possível observar que a Rede Neural realiza uma boa previsão da amostra de velocidade do vento, para  $k=1$ . Já para  $k=6$ , Figura 5.2, o resultado pode ser melhorado. Os maiores valores de erro são notados em instantes que a série mais varia, com um pico ou vale mais relevante. Na Figura 5.3 é possível notar a dispersão dos valores estimados em referência a reta  $y = x$ . Para  $k=6$  os pontos se distribuem mais distantes da reta  $y = x$  se comparado para  $k=1$ , confirmando o  $R^2$  inferior. Se os dados estimados fossem iguais aos reais, todos os pontos desse gráfico deveriam estar sobre a reta.

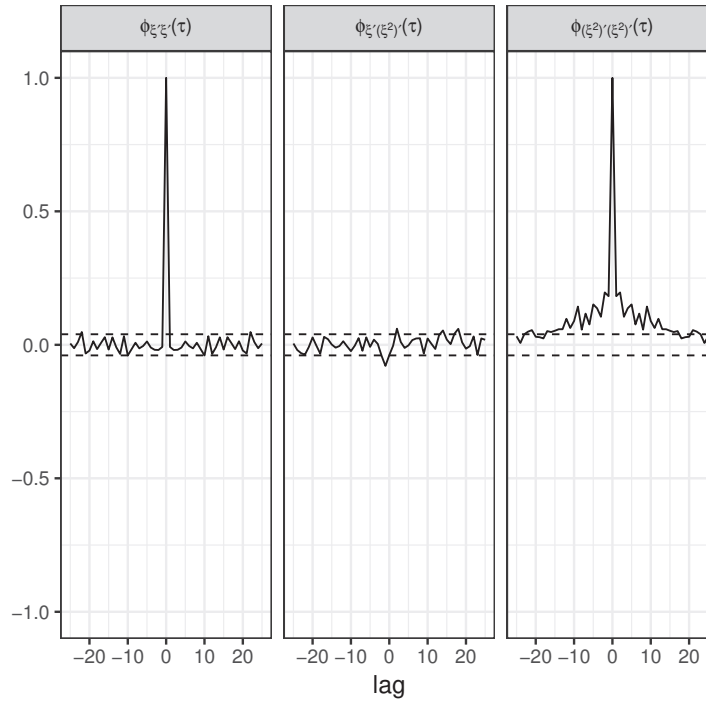


Figura 5.4: Testes de correlação para velocidade ( $k=1$ ) utilizando RNA.

Na Figura 5.4, é possível observar que apesar do  $R^2=0,8023$  associado a predição um passo à frente para a velocidade não ter sido tão alto, no primeiro teste de correlação o resultado obtido foi muito satisfatório, encontrando-se dentro das linhas pontilhadas (valores próximos a zero). Para o  $lag = 0$  (sem defasagem), a correlação obtida foi 1, o que está correto, devido ao fato de nesse caso a série estar sendo relacionada com ela mesmo. Para o segundo teste de correlação, o resultado obtido também foi satisfatório. Nesse caso, segundo a Equação 3-15, todos os valores devem ser iguais a zero (dentro da linha pontilhada). Contudo, o terceiro teste de correlação não foi satisfatório, ressaltando que algumas melhorias ainda podem ser feitas, em termos de predição, utilizando o mesmo conjunto de dados e alterando-se a arquitetura do modelo e sua criação.

### 5.1.2

#### Resultados para a Predição da Direção do Vento com RNAs

Da mesma forma que foi realizado para a predição da velocidade do vento com RNAs, a seguir, para cada um dos três valores de *learning rate*, existe uma tabela com os Coeficientes de Determinação ( $R^2$ ) para a predição da direção, variando a ordem [1:10], o número de camadas [1:5] e o número de neurônios em cada camada [16,32,64], com a função de ativação *relu* e para um passo à frente (k=1) ou seis passos à frente (k=6). As tabelas seguem a seguinte ordem:

- Tabela 5.7  $\implies$  *Learning Rate* =  $10^{-2}$  e um passo à frente (k=1);
- Tabela 5.8  $\implies$  *Learning Rate* =  $10^{-2}$  e seis passos à frente (k=6);
- Tabela 5.9  $\implies$  *Learning Rate* =  $10^{-3}$  e um passo à frente (k=1);
- Tabela 5.10  $\implies$  *Learning Rate* =  $10^{-3}$  e seis passos à frente (k=6);
- Tabela 5.11  $\implies$  *Learning Rate* =  $10^{-4}$  e um passo à frente (k=1);
- Tabela 5.12  $\implies$  *Learning Rate* =  $10^{-4}$  e seis passos à frente (k=6).

A observação das 6 tabelas acima listadas também evidenciou a melhora dos  $R^2$  com o aumento da complexidade relacionada à diminuição do *learning rate*, igualmente como foi constatado para a predição da velocidade. Quando menor o *learning rate*, melhores foram o  $R^2$ . Dessa forma, provavelmente é possível alcançar resultados ainda melhores com o *learning rate* menor que  $10^{-4}$  ou arquiteturas mais complexas.

Após a realização dos testes preliminares, foi escolhido um conjunto de parâmetros que resultava em um dos melhores  $R^2$  para a predição da direção. Neste caso, os parâmetros que resultaram no melhor resultado foram: *learning rate* =  $10^{-4}$ , ordem = 10 e 2 camadas com 64 neurônios cada (tabela 5.12).

Para esse conjunto de parâmetros escolhidos, foi plotado o gráfico da predição para um e seis passos à frente juntamente com o valor real e o gráfico do erro ( $y(t) - \hat{y}(t)$ ) - Figuras 5.5 e 5.6, além de um gráfico de  $y(t)$  versus  $\hat{y}(t)$  com a reta  $y = x$  (Figura 5.7), sendo possível observar a dispersão dos dados reais e estimados.

Coeficientes de Determinação para Direção - learning rate = $10^{-2}$						
Ordem	Neurônios	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,8575	0,8891	0,8803	0,5418	0,0267
2	16	0,8812	0,8945	0,8787	0,0465	0,0009
3	16	0,8705	0,8943	0,8465	0,0105	0,1013
4	16	0,9035	0,8916	0,8903	0,1012	0,0016
5	16	0,8929	0,8701	0,8620	0,0022	0,0127
6	16	0,9049	0,8509	0,9012	0,0152	0,0026
7	16	0,7771	0,8896	0,8336	0,0027	0,1806
8	16	0,9006	0,9023	0,7882	0,0103	0,0035
9	16	0,4121	0,8666	0,8741	0,0261	0,0327
10	16	0,8940	0,8530	0,8969	0,0057	0,0099
1	32	0,8954	0,8944	0,8825	0,1739	0,5014
2	32	0,8637	0,8906	0,8832	0,1443	0,5586
3	32	0,8821	0,8954	0,8620	0,3536	0,0058
4	32	0,8934	0,8965	0,8992	0,1981	0,0003
5	32	0,9036	0,8901	0,8934	0,5547	0,0342
6	32	0,8977	0,7724	0,9035	0,2214	0,0081
7	32	0,7568	0,8938	0,8976	0,0000	0,0141
8	32	0,8690	0,8659	0,9038	0,0188	0,0106
9	32	0,8831	0,8876	0,8829	0,3907	0,0002
10	32	0,8813	0,8924	0,8304	0,6209	0,0528
1	64	0,7127	0,8625	0,8084	0,8699	0,0580
2	64	0,8777	0,8986	0,8807	0,7730	0,0041
3	64	0,8995	0,8283	0,8998	0,6884	0,0151
4	64	0,8997	0,8907	0,8807	0,8843	0,0229
5	64	0,8998	0,8489	0,7900	0,8870	0,0128
6	64	0,9014	0,8902	0,8739	0,8001	0,0077
7	64	0,8500	0,8647	0,8828	0,8560	0,0168
8	64	0,6826	0,8692	0,8981	0,8914	0,0150
9	64	0,8949	0,8928	0,9036	0,8983	0,0295
10	64	0,9049	0,8992	0,8753	0,8955	0,0177

Tabela 5.7: Resultados de coeficiente de determinação para direção obtidos com learning rate =  $10^{-2}$ , para um passo à frente.

Coeficientes de Determinação para Direção - learning rate = $10^{-2}$						
Ordem	Neurônios	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,2238	0,5645	0,2657	0,4709	-0,0256
2	16	0,2468	0,6237	0,4663	0,0835	-0,0012
3	16	0,1448	0,5831	0,0443	0,0115	-0,1049
4	16	0,7389	0,5148	0,5626	0,0529	-0,0016
5	16	0,5620	0,3860	0,0979	0,0018	0,0079
6	16	0,7447	0,3930	0,7030	0,0166	-0,0022
7	16	0,8298	0,5610	-1,2120	0,0022	0,1411
8	16	0,6892	0,7436	-1,2487	0,0114	-0,0013
9	16	-1,1599	0,3645	0,1037	0,0278	-0,0347
10	16	0,6038	0,0664	0,6175	0,0063	-0,0109
1	32	0,6928	0,6996	0,4744	0,1299	0,4083
2	32	0,0323	0,5731	0,2456	0,1069	0,4398
3	32	0,3622	0,6561	0,3051	0,2780	-0,0052
4	32	0,5390	0,5728	0,5924	0,1547	-0,0001
5	32	0,7348	0,5037	0,6072	0,4770	-0,0362
6	32	0,6763	-1,8687	0,7415	0,1812	-0,0091
7	32	-1,9080	0,6639	0,6837	0,0001	-0,0155
8	32	0,3724	0,5138	0,7081	0,0177	-0,0097
9	32	0,4691	0,5358	0,6223	0,3296	-0,0001
10	32	0,4674	0,5987	0,0997	0,4298	-0,0512
1	64	-6,3046	0,0123	-3,2399	0,0741	-0,0603
2	64	0,3038	0,6435	0,4397	-8,4197	-0,0047
3	64	0,6448	0,4452	0,6701	-2,7912	-0,0164
4	64	0,7051	0,6054	0,5356	0,5966	-0,0257
5	64	0,6804	0,2618	0,4098	0,5082	-0,0139
6	64	0,7281	0,5327	0,3818	0,1151	0,0042
7	64	0,1681	0,0411	0,5224	0,1690	-0,0182
8	64	-4,3459	0,2926	0,6341	0,5722	-0,0140
9	64	0,6298	0,6106	0,6884	0,6634	-0,0494
10	64	0,7522	0,6898	0,2429	0,7038	0,0084

Tabela 5.8: Resultados de coeficiente de determinação para direção obtidos com learning rate =  $10^{-2}$ , para seis passos à frente.

Coeficientes de Determinação para Direção - learning rate = $10^{-3}$						
Ordem	Neurônios	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,8866	0,8892	0,8961	0,8941	0,8675
2	16	0,8976	0,8982	0,9028	0,8955	0,9030
3	16	0,9024	0,9057	0,9020	0,9058	0,9054
4	16	0,9066	0,9087	0,9086	0,9060	0,9061
5	16	0,9023	0,9057	0,8971	0,8875	0,9087
6	16	0,9005	0,9028	0,8722	0,9075	0,8986
7	16	0,9059	0,9080	0,9033	0,9093	0,9055
8	16	0,9061	0,9042	0,9079	0,9085	0,8997
9	16	0,9068	0,9071	0,9064	0,8871	0,9074
10	16	0,9068	0,9099	0,9016	0,9094	0,8980
1	32	0,8931	0,8954	0,8958	0,8901	0,8897
2	32	0,8914	0,9033	0,8988	0,8990	0,9015
3	32	0,9054	0,8724	0,9061	0,9071	0,8984
4	32	0,9062	0,8967	0,9072	0,9068	0,8961
5	32	0,9075	0,9095	0,9088	0,8935	0,9008
6	32	0,8858	0,9089	0,9089	0,9031	0,8988
7	32	0,8951	0,9025	0,8901	0,9089	0,8846
8	32	0,9066	0,9122	0,9125	0,9109	0,9050
9	32	0,9069	0,9067	0,9132	0,8986	0,8974
10	32	0,9065	0,9092	0,9075	0,9083	0,9037
1	64	0,8953	0,8965	0,8960	0,8946	0,8902
2	64	0,8989	0,8981	0,9022	0,8973	0,8825
3	64	0,9022	0,9053	0,9074	0,9073	0,9052
4	64	0,8971	0,9107	0,9089	0,9108	0,9087
5	64	0,8820	0,9057	0,8973	0,9086	0,9046
6	64	0,8650	0,9096	0,8915	0,9118	0,8407
7	64	0,8942	0,8895	0,8997	0,9108	0,9086
8	64	0,9071	0,8447	0,9131	0,9137	0,9054
9	64	0,8908	0,9105	0,9078	0,9133	0,9010
10	64	0,9072	0,9027	0,9081	0,9125	0,9053

Tabela 5.9: Resultados de coeficiente de determinação para direção obtidos com learning rate =  $10^{-3}$ , para um passo à frente.



Coeficientes de Determinação para Direção - learning rate = $10^{-3}$						
Ordem	Neurônios	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,4336	0,4937	0,7275	0,6991	0,0567
2	16	0,6395	0,6024	0,7276	0,5943	0,7485
3	16	0,7103	0,7497	0,6908	0,7309	0,7229
4	16	0,7594	0,7338	0,7276	0,7359	0,7172
5	16	0,6458	0,6552	0,4926	0,3484	0,7607
6	16	0,6890	0,6758	0,1319	0,7386	0,5137
7	16	0,7501	0,7565	0,5874	0,7595	0,7386
8	16	0,7388	0,6834	0,7654	0,7522	0,6427
9	16	0,7516	0,7350	0,6820	0,4110	0,7624
10	16	0,7584	0,7604	0,6057	0,7665	0,6023
1	32	0,6222	0,6738	0,7031	0,5612	0,5284
2	32	0,4902	0,7295	0,6830	0,6824	0,7111
3	32	0,7525	0,2646	0,7153	0,7589	0,6310
4	32	0,7614	0,4965	0,7171	0,7057	0,5918
5	32	0,7423	0,7563	0,7270	0,5459	0,6884
6	32	0,4235	0,6985	0,7361	0,6514	0,5696
7	32	0,5772	0,6243	0,3738	0,7451	0,4733
8	32	0,7599	0,7655	0,7455	0,7569	0,7407
9	32	0,7603	0,7310	0,7545	0,5517	0,6108
10	32	0,7338	0,7408	0,7036	0,7591	0,7532
1	64	0,6965	0,7285	0,7008	0,6753	0,5543
2	64	0,6718	0,6281	0,7183	0,6153	0,1808
3	64	0,6456	0,6853	0,7548	0,7509	0,7366
4	64	0,5160	0,7588	0,7394	0,7492	0,7617
5	64	0,4059	0,6892	0,5803	0,7369	0,7094
6	64	0,2466	0,7099	0,4379	0,7367	0,3721
7	64	0,4870	0,3900	0,4886	0,6615	0,7535
8	64	0,7389	0,3934	0,7318	0,7644	0,7112
9	64	0,4989	0,7575	0,6657	0,7473	0,6359
10	64	0,7270	0,5658	0,7069	0,7230	0,6798

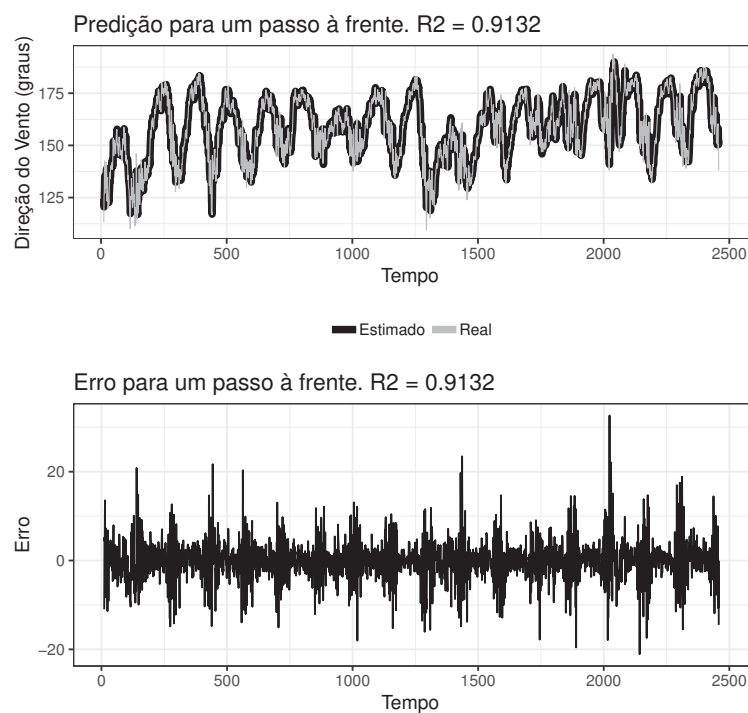
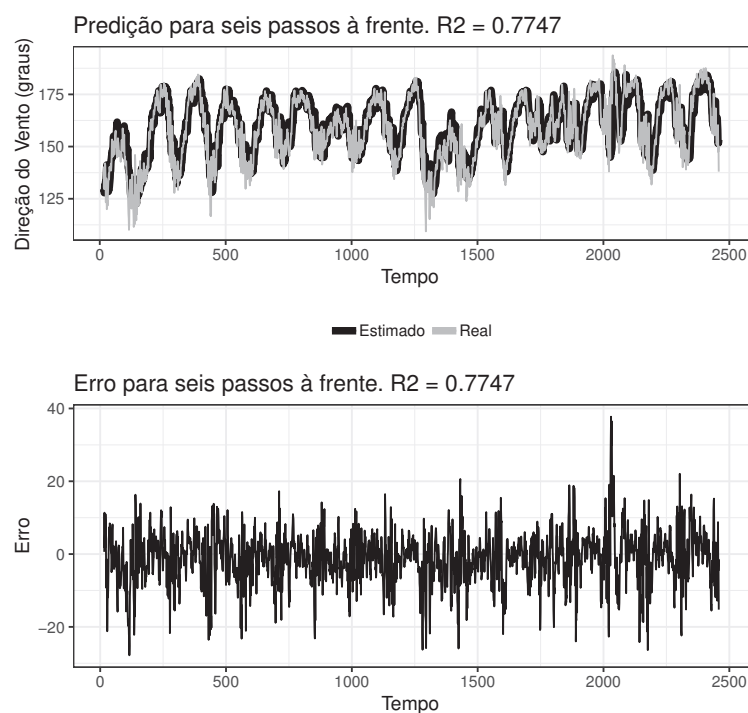
Tabela 5.10: Resultados de coeficiente de determinação para direção obtidos com learning rate =  $10^{-3}$ , para seis passos à frente.

Coeficientes de Determinação para Direção - learning rate = $10^{-4}$						
Ordem	Neurônios	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1	$R^2$ - k=1
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,8960	0,8940	0,8950	0,8957	0,8962
2	16	0,9021	0,9020	0,8980	0,8998	0,8944
3	16	0,9037	0,9056	0,9062	0,9054	0,9039
4	16	0,9039	0,9064	0,9038	0,9070	0,9070
5	16	0,9036	0,9077	0,9085	0,9037	0,9069
6	16	0,9047	0,9062	0,9070	0,9063	0,9069
7	16	0,9054	0,9067	0,9040	0,9067	0,8976
8	16	0,9047	0,9074	0,9064	0,9081	0,8929
9	16	0,9050	0,9072	0,9066	0,9062	0,9098
10	16	0,9050	0,9066	0,9059	0,9056	0,9090
1	32	0,8944	0,8961	0,8911	0,8951	0,8961
2	32	0,9032	0,9017	0,9008	0,8954	0,9039
3	32	0,9066	0,9060	0,9054	0,9068	0,9050
4	32	0,9063	0,9068	0,9083	0,9093	0,9092
5	32	0,9064	0,9074	0,9042	0,9086	0,9100
6	32	0,9072	0,9082	0,9074	0,9003	0,9075
7	32	0,9067	0,9107	0,9094	0,9090	0,9085
8	32	0,9072	0,9073	0,9087	0,9078	0,9019
9	32	0,9052	0,9082	0,9098	0,9090	0,9121
10	32	0,9052	0,9081	0,9120	0,9089	0,9108
1	64	0,8959	0,8965	0,8963	0,8961	0,8962
2	64	0,9028	0,9039	0,9012	0,9012	0,9038
3	64	0,9050	0,9070	0,9072	0,9083	0,9074
4	64	0,9051	0,9088	0,9121	0,9073	0,9119
5	64	0,9060	0,9047	0,9106	0,9088	0,9103
6	64	0,9073	0,9070	0,9135	0,9085	0,9134
7	64	0,9074	0,9109	0,9136	0,9132	0,9076
8	64	0,9086	0,9111	0,9114	0,9149	0,8923
9	64	0,9069	0,9081	0,9158	0,9180	0,9091
10	64	0,9083	0,9132	0,9158	0,9183	0,9144

Tabela 5.11: Resultados de coeficiente de determinação para direção obtidos com learning rate =  $10^{-4}$ , para um passo à frente.

Coeficientes de Determinação para Direção - learning rate = $10^{-4}$						
Ordem	Neurônios	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6	$R^2$ - k=6
		1 camada	2 camadas	3 camadas	4 camadas	5 camadas
1	16	0,7356	0,6839	0,6898	0,7211	0,7251
2	16	0,7465	0,7328	0,6339	0,6578	0,4997
3	16	0,7345	0,7618	0,7592	0,7463	0,6945
4	16	0,7148	0,7602	0,6784	0,7540	0,7416
5	16	0,7406	0,7608	0,7616	0,7237	0,7139
6	16	0,7509	0,7505	0,7614	0,7555	0,7459
7	16	0,7567	0,7645	0,6808	0,7128	0,5130
8	16	0,7484	0,7660	0,7346	0,7589	0,5551
9	16	0,7574	0,7667	0,7661	0,7473	0,7736
10	16	0,7548	0,7600	0,7365	0,7204	0,7682
1	32	0,6884	0,7325	0,5803	0,6736	0,7235
2	32	0,7470	0,7035	0,6607	0,5053	0,7461
3	32	0,7612	0,7548	0,7057	0,7327	0,7387
4	32	0,7596	0,7523	0,7627	0,7622	0,7579
5	32	0,7444	0,7582	0,6916	0,7218	0,7434
6	32	0,7605	0,7593	0,7019	0,5647	0,7507
7	32	0,7599	0,7660	0,7578	0,7612	0,7447
8	32	0,7647	0,7409	0,7420	0,7636	0,5974
9	32	0,7078	0,7434	0,7739	0,7699	0,7662
10	32	0,7434	0,7415	0,7667	0,7445	0,7084
1	64	0,7321	0,7238	0,7283	0,7089	0,7409
2	64	0,7413	0,7458	0,6900	0,6852	0,7382
3	64	0,7411	0,7493	0,7241	0,7543	0,7598
4	64	0,7423	0,7381	0,7673	0,7223	0,7493
5	64	0,7490	0,6780	0,7372	0,6846	0,7646
6	64	0,7444	0,7556	0,7179	0,6574	0,7378
7	64	0,7548	0,7648	0,7678	0,7385	0,5359
8	64	0,7705	0,7674	0,7146	0,7265	0,3766
9	64	0,7360	0,6996	0,7776	0,7792	0,6074
10	64	0,7667	0,7747	0,7681	0,7519	0,6716

Tabela 5.12: Resultados de coeficiente de determinação para direção obtidos com learning rate =  $10^{-4}$ , para seis passos à frente.

Figura 5.5: Predição da direção do vento ( $k=1$ ) e erro utilizando RNA.Figura 5.6: Predição da direção do vento ( $k=6$ ) e erro utilizando RNA.

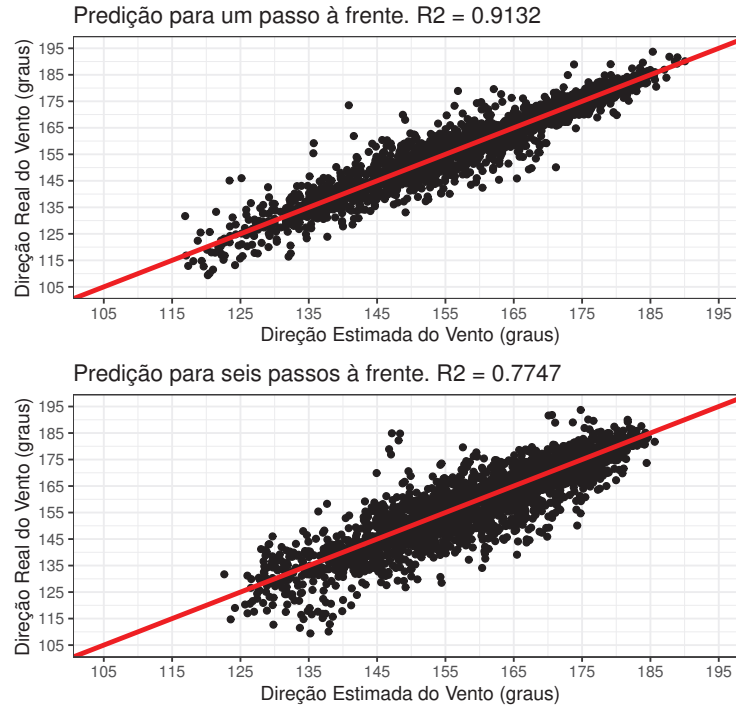


Figura 5.7: Gráfico de  $y(t)$  versus  $\hat{y}(t)$  para direção utilizando RNA.

As figuras anteriores são os melhores resultados alcançados para predição série de direção um e seis passos à frente ( $Learning\ rate = 10^{-4}$ , ordem = 10 e 2 camadas com 64 neurônios cada). Nas Figuras 5.5 e 5.6 é possível observar que a RNA é capaz de realizar uma boa previsão da amostra de direção do vento, mesmo para  $k=6$ . Os maiores valores de erro são notados em instantes que a série mais varia, com um pico ou vale mais relevante. Na Figura 5.7 é possível notar a dispersão dos valores estimados em referência a reta  $y = x$ . Se os dados estimados fossem iguais aos reais, todos os pontos desse gráfico deveriam estar sobre a reta.

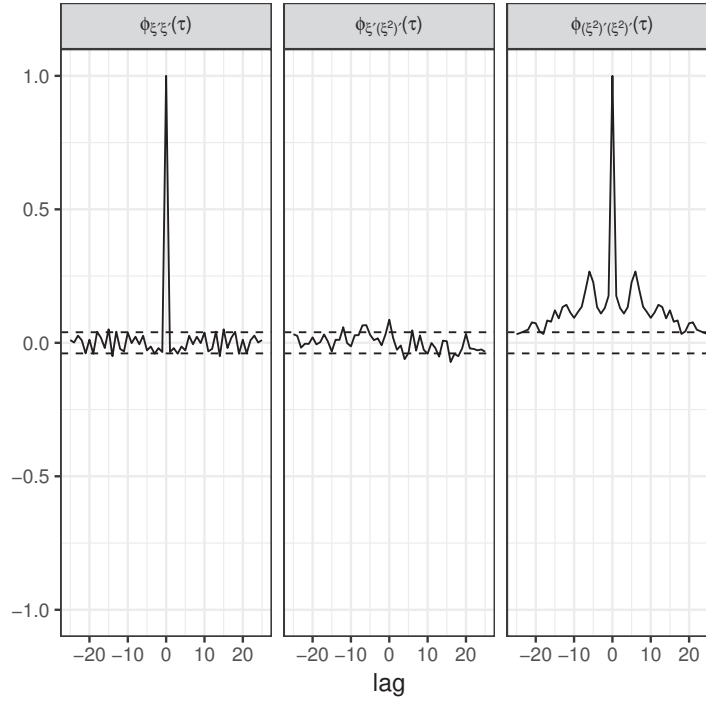


Figura 5.8: Testes de correlação para direção ( $k=1$ ) utilizando RNA.

Analisando a Figura 5.8, percebe-se que o primeiro e segundo testes de correlação obtiveram resultados satisfatórios. Contudo, o terceiro, não. Essa mesmo comportamento foi verificado na Figura 5.4, apesar do  $R^2$  associado a predição um passo à frente para a direção ser superior ao da velocidade.

## 5.2

### Transformações

Aplicar transformações aos dados históricos muitas vezes pode levar a uma tarefa de previsão mais simples. A finalidade dessas transformações é na tentativa de simplificar os padrões nos dados históricos removendo as fontes de variação conhecidas ou tornando o padrão mais consistente em todo o conjunto de dados. Padrões mais simples geralmente levam a previsões mais precisas. No caso do presente trabalho, foram realizadas transformações matemáticas que serão listadas e discutidas a seguir.

#### 5.2.1

##### Transformação Logarítmica

A primeira transformação matemática dos dados utilizada foi a transformação logarítmica. Se as observações originais forem denotadas como  $y_1, \dots, y_T$ , e as observações transformadas como  $w_1, \dots, w_t$ , então  $w_t = \log(y_t)$ .

Os logaritmos são úteis porque são interpretáveis: as alterações em um valor de log são alterações relativas (ou percentuais) na escala original. Portanto, se a base de log 10 for usada, então um aumento de 1 na escala de log corresponde a uma multiplicação de 10 na escala original.

Após testes utilizando dados onde se foi aplicada a transformação logarítmica e utilizando as arquiteturas de rede que obtiveram melhores resultados para velocidade (*learning rate* =  $10^{-4}$ , ordem = 9 e 2 camadas com 64 neurônios cada) e direção (*learning rate* =  $10^{-4}$ , ordem = 10 e 2 camadas com 64 neurônios cada) na seção anterior, os resultados encontrados não obtiveram melhora significativa.

### 5.2.2

#### Transformação Média = 0 e Desvio Padrão = 1

Com o objetivo de tornar os dados de entrada da Rede Neural com média=0 e desvio padrão=1, foi aplicada a seguinte transformação:

$$y' = \frac{y - \text{média}(\mathbf{y})}{\text{desvio padrão}(\mathbf{y})} \quad (5-2)$$

Onde  $y'$  será o vetor de dados transformado com média = 0 e desvio padrão = 1 e  $\mathbf{y}$  é o vetor de dados original, sem transformação.

Após testes onde se foi aplicada a transformação 5-2 e utilizando os mesmos parâmetros de arquitetura de rede do caso da transformação logarítmica, os resultados encontrados não obtiveram melhora significativa, novamente.

## 6

### Resultado com NARMA e RNA

Dado que os resultados alcançados e listados no Capítulo 5 deste trabalho, fazendo uso somente de Rede Neural Artificial para realizar a predição das séries de velocidade e direção do vento, ainda tinham margem para ser melhorados, uma opção usada foi utilizar o modelo NARMA juntamente com a RNA.

A metodologia adotada foi primeiramente realizar a predição de ambas as séries com NARMA, sempre fixando os parâmetros  $ne = 1$  e  $nl = 1$ , o que resulta em um modelo linear e por conseguinte numa busca por parâmetros mais simples. O erro  $(y(t) - \hat{y}(t))$  originário da primeira previsão era utilizado como entrada para a segunda predição, agora usando RNA. Por fim, é realizada a soma das duas predições. Esta prática é no geral referida de *model ensemble* na literatura [14]. O esquema adotado pode ser observado na Figura 6.1.

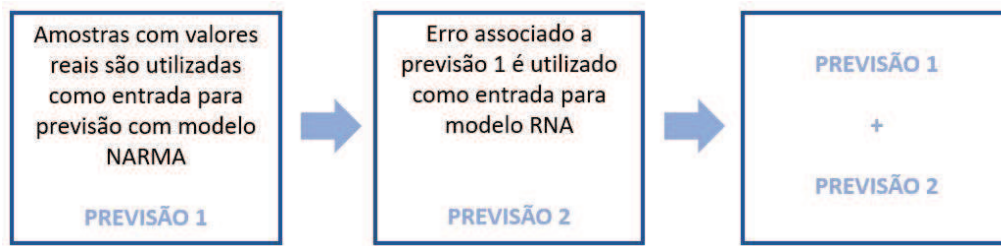


Figura 6.1: Esquema utilizado para gerar resultados no Capítulo 6.

O objetivo é que o modelo NARMA com os parâmetros  $ne = 1$  e  $nl = 1$  consiga prever com maior acurácia a parte linear das séries de vento. Isso se deve ao fato de quando os valores dos parâmetros  $ne$  e  $nl$  é fixado igual a 1, o NARMA trabalha como ARMA (*autoregressive-moving-average*). A RNA ficaria responsável por prever a parte não-linear, neste caso, o erro da previsão 1. Quanto maior o  $R^2$  da previsão do erro, melhor será o  $R^2$  da soma da previsão 1 com a 2. Se o erro conseguir ser previsto com  $R^2=1$ , a soma das previsões também terá  $R^2=1$  quando comparada com os valores reais, dado que:

$$y_t = \hat{y}_t + \epsilon \quad (6-1)$$

onde  $y_t$  são os valores reais (medidos),  $\hat{y}_t$  os valores estimados e  $\epsilon$  o erro associado, na previsão 1.



Vale ressaltar que tanto para a velocidade quanto para a direção do vento, os dados utilizados neste capítulo foram os mesmos empregados no Capítulo 5, e estão plotados nas Figuras 4.1 e 4.3.

## 6.1

### Resultados para a Predição da Velocidade do Vento com NARMA + RNAs

O primeiro passo, assim como no Capítulo 5, foi relizar testes afim de encontrar os parâmetros do modelo, neste caso, do NARMA. Como mencionado anteriormente, os valores de  $ne=1$  e  $nl=1$  foram fixos, assim como  $\rho_p = 5 \times 10^{-4}$ ,  $\rho_n = 6 \times 10^{-4}$ . O único parâmetro que variou foi o  $ny$ . A Tabela 6.1 traz os valores dos Coeficientes de Determinação ( $R^2$ ), para as ordens de [1:20].

Coeficientes de Determinação para Velocidade - NARMA		
Ordem	R2 - k=1	R2 - k=6
1	0,7900	0,5144
2	0,7925	0,4219
3	0,7923	0,4716
4	0,7923	0,3532
5	0,7929	0,5309
6	0,7928	0,4094
7	-Inf	-1,0249
8	0,7945	0,5316
<b>9</b>	<b>0,7947</b>	0,5354
10	Inf	-5,2947
11	0,7945	0,5344
12	0,7943	0,5336
13	-Inf	-4,5546
14	0,7937	0,5363
15	0,7942	0,5377
16	0,7943	0,5334
17	0,7945	0,5327
18	0,7939	0,4980
<b>19</b>	0,7939	<b>0,5377</b>
20	-Inf	0,1327

Tabela 6.1: Resultados de coeficiente de determinação para velocidade obtidos com NARMA ( $ne=1$  e  $nl=1$ ), para um passo à frente.

Os melhores resultados para a predição de um e seis passos à frente foi atingido com a ordem  $ny = 9$  e  $ny = 19$ , respectivamente. Os valores de  $R^2$  encontrados em ambos os casos foi bem similar ao encontrado no Capítulo 5, seção 5.1.1, nas Figuras 5.1 e 5.2 utilizando somente a RNA. Vale ressaltar que os valores infinitos para  $R^2$  encontrados significam que o modelo não produz previsões estáveis. A seguir, nas Figuras 6.2 e 6.3 estão plotados os gráficos utilizando o modelo NARMA, sendo essa a previsão 1.

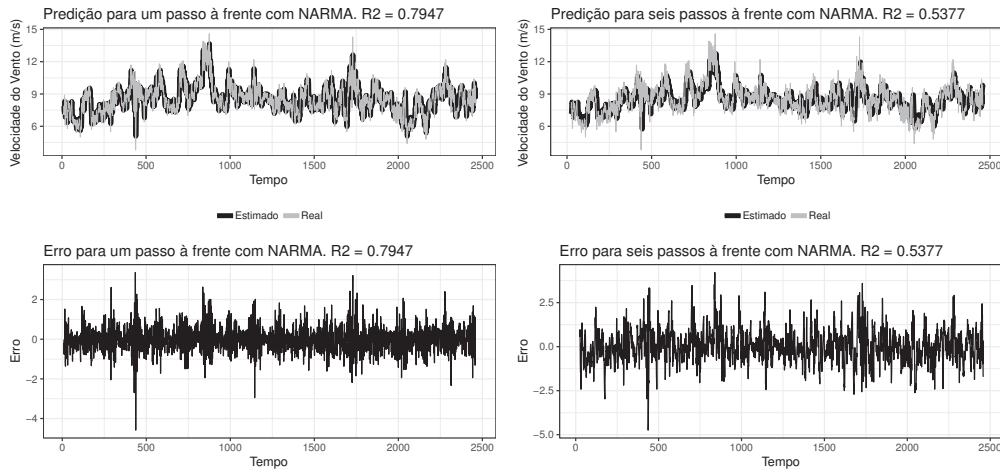
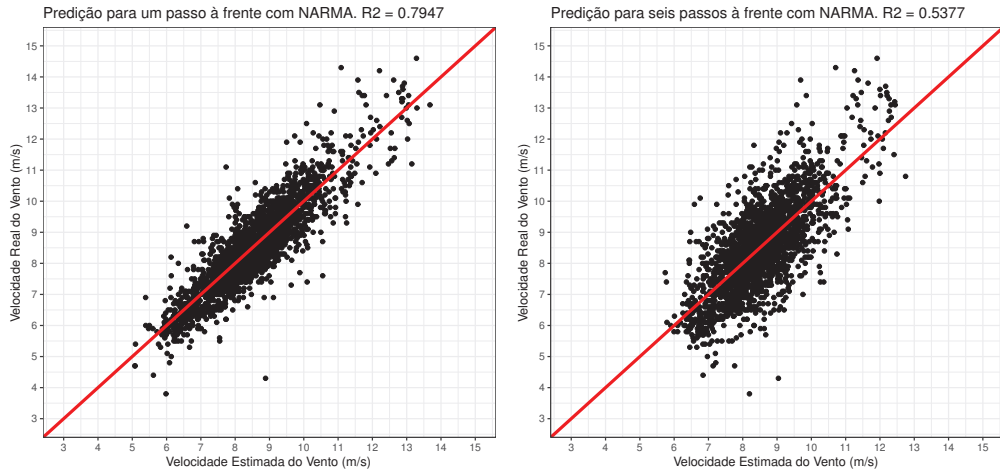
(6.2(a)) Para  $k=1$  com  $ny=9$ (6.2(b)) Para  $k=6$  com  $ny=19$ 

Figura 6.2: Predição da velocidade do vento e erro utilizando NARMA.

(6.3(a)) Para  $k=1$  com  $ny=9$ (6.3(b)) Para  $k=6$  com  $ny=19$ Figura 6.3: Gráfico de  $y(t)$  versus  $\hat{y}(t)$  para velocidade utilizando NARMA.

Os testes de correlação utilizando NARMA na Figura 6.4 foram bem

similares ao da Figura 5.4, onde foi utilizado apenas a RNA para previsão um passo à frente.

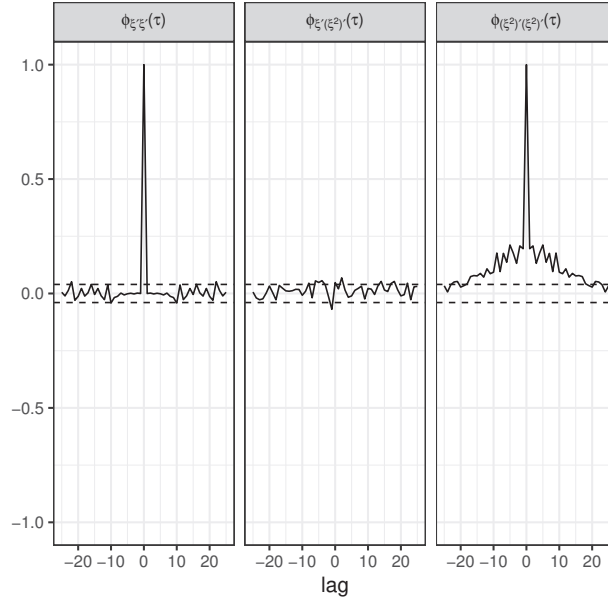


Figura 6.4: Testes de correlação para velocidade ( $k=1$ ) utilizando NARMA.

Após realizar a previsão 1, o próximo passo foi realizar a previsão 2 para um e seis passos à frente, onde os erros plotados nas Figuras 6.2 e 6.3 serão previstos utilizando RNA. Da mesma forma do Capítulo 5, foram realizados testes para escolha da Arquitetura da Rede Neural. Durante os testes preliminares, os parâmetros alterados foram:

- Ordem [1:30]
- Número de camadas [1:5]
- Número de neurônios em cada camada [16,32,64]

Os parâmetros que permaneceram fixos durante os testes, foram:

- *Learning Rate*:  $10^{-4}$
- Função de ativação: Relu
- *Epochs*: 100
- *Batch Size*: 2

A melhor arquitetura para a previsão do erro (previsão 2) com  $k=1$  foi ordem  $ny=18$ , 5 camadas com 64 neurônios cada, onde o  $R^2=0,9115$ . Já para  $k=6$ ,  $ny=24$ , 5 camadas com 64 neurônios cada, onde o  $R^2=0,6283$ .

Após a realização das previsões 1 e 2 para  $k=1$  e  $k=6$ , foi realizada a soma do resultado alcançado com o modelo NARMA e RNA. Para  $k=1$  o  $R^2=0,9819$ , sendo um resultado 22,39 % maior do que o atingido no Capítulo 5, onde somente a RNA foi utilizada. Já para  $k=6$  o  $R^2=0,8287$ , sendo um resultado 51,86 % maior.

As figuras a seguir foram plotadas como o resultado do modelo NARMA+RNA, para  $k=1$  e  $k=6$ .

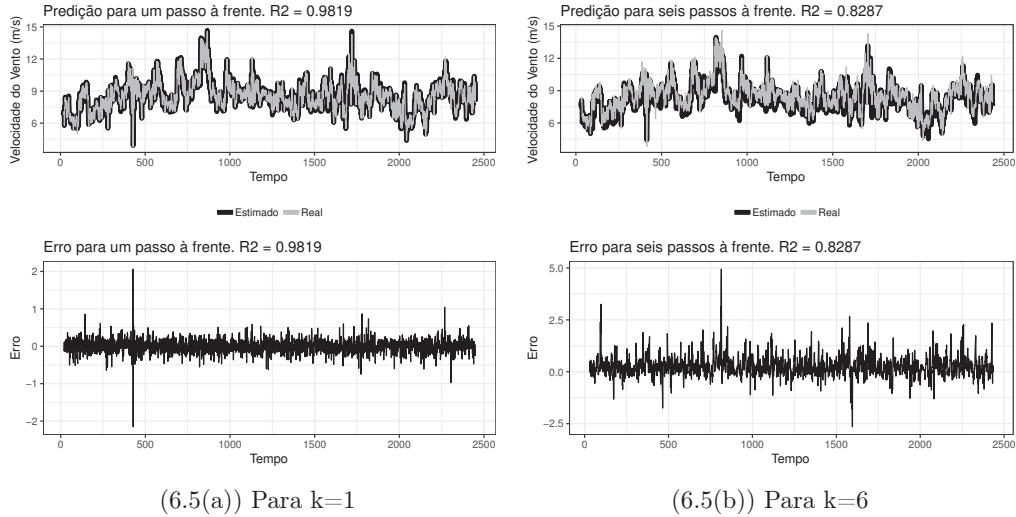


Figura 6.5: Predição da velocidade do vento e erro utilizando NARMA+RNA.

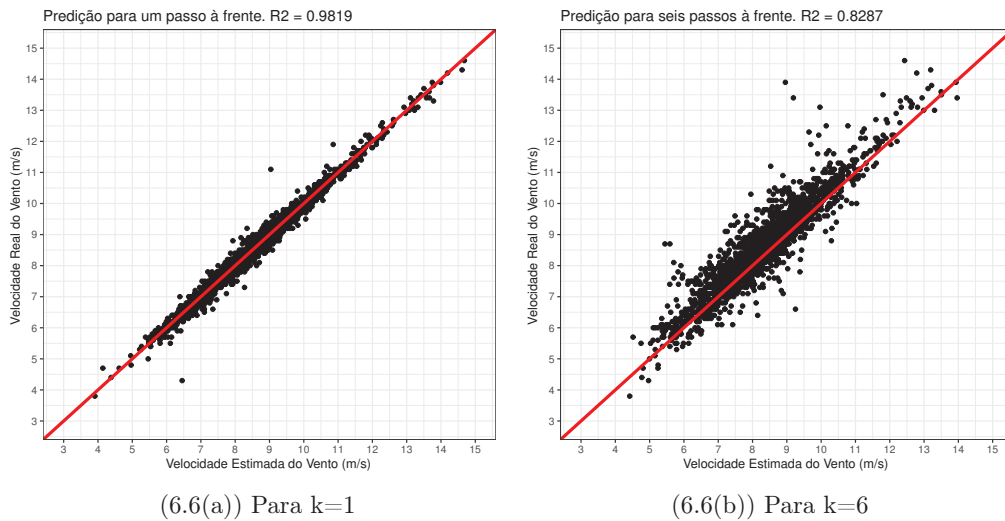


Figura 6.6: Gráfico de  $y(t)$  versus  $\hat{y}(t)$  para velocidade utilizando NARMA+RNA.

Comparando o erro da Figura 6.5 com o das Figuras 5.1 e 5.2 do Capítulo 5, seção 5.1.1, onde foi utilizado somente RNA na predição, percebe-se que o

erro é bem menor, com valores mais próximos de zero. Além disso, os pontos da Figura 6.6 tem uma dispersão muito menor em relação a reta  $y = x$  do que na Figura 5.3, ressaltando o melhor resultado quando é utilizado os dois modelos.

## 6.2

## Resultados para a Predição da Direção do Vento com NARMA + RNAs

Assim como foi feito para a velocidade na seção 6.1, o primeiro passo foi realizar testes afim de encontrar os parâmetros do modelo, neste caso, do NARMA. Como mencionado anteriormente, os valores de  $ne=1$  e  $nl=1$  foram fixos, assim como  $\rho_p = 5 \times 10^{-4}$ ,  $\rho_n = 6 \times 10^{-4}$ . O único parâmetro que variou foi o  $ny$ . A Tabela 6.2 traz os valores dos Coeficientes de Determinação ( $R^2$ ), para as ordens de [1:20].

Coeficientes de Determinação para Direção - NARMA		
Ordem	R2 - k=1	R2 - k=6
1	0,9040	0,7415
2	0,9050	0,6985
3	0,9051	0,7201
4	0,9049	0,7103
5	0,9048	0,7514
6	-Inf	-83,3859
7	0,9042	0,7505
8	0,9047	0,6732
<b>9</b>	<b>0,9048</b>	0,7517
10	0,9046	0,7383
11	0,9040	0,7570
12	0,9042	0,6439
13	0,9039	0,3166
14	0,9039	0,7556
15	0,9038	0,7560
<b>16</b>	0,9039	<b>0,7586</b>
17	0,9041	0,7560
18	0,9040	0,7573
19	0,9043	0,7582
20	0,9041	0,7578

Tabela 6.2: Resultados de coeficiente de determinação para direção obtidos com NARMA ( $ne=1$  e  $nl=1$ ), para um passo à frente.

Os melhores resultados para a predição de um e seis passos à frente foi atingido com a ordem  $n_y = 9$  e  $n_y = 16$ , respectivamente. Os valores de  $R^2$  encontrados em ambos os casos foi bem similar ao encontrado no Capítulo 5, seção 5.1.2, nas Figuras 5.5 e 5.6 utilizando somente a RNA. A seguir, nas Figuras estão plotados os gráficos utilizando o modelo NARMA, sendo essa a previsão 1.

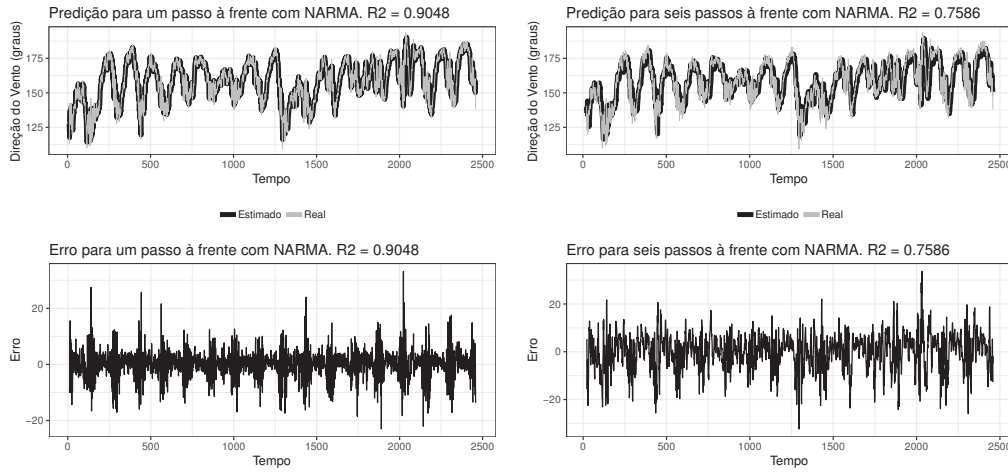
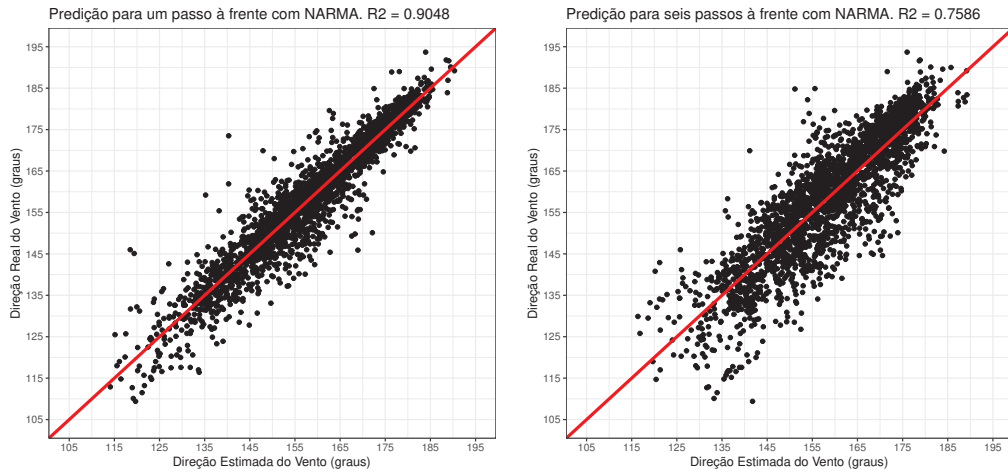
(6.7(a)) Para  $k=1$  com  $n_y=9$ (6.7(b)) Para  $k=6$  com  $n_y=16$ 

Figura 6.7: Predição da direção do vento e erro utilizando NARMA.

(6.8(a)) Para  $k=1$  com  $n_y=10$ (6.8(b)) Para  $k=6$  com  $n_y=16$ Figura 6.8: Gráfico de  $y(t)$  versus  $\hat{y}(t)$  para direção utilizando NARMA.

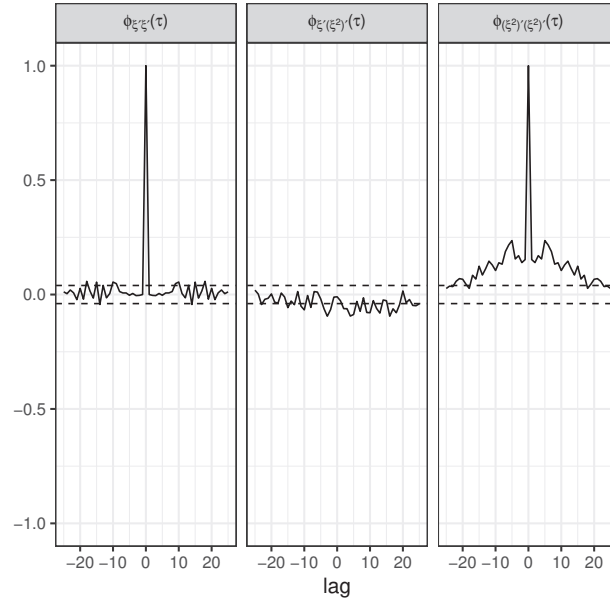


Figura 6.9: Testes de correlação para direção ( $k=1$ ) utilizando NARMA.

Da mesma forma que ocorreu para a velocidade, os testes de correlação para a direção utilizando NARMA (Figura 6.9) foram bem similares a previsão utilizando somente RNA (Figura 5.8).

Após realizar a previsão 1, o próximo passo foi realizar a previsão 2 para um e seis passos à frente, onde os erros plotados nas Figuras 6.7 e 6.8 serão previstos utilizando RNA. Da mesma forma do Capítulo 5, foram realizados testes para escolha da Arquitetura da Rede Neural. Durante os testes preliminares, os parâmetros alterados foram:

- Ordem [1:25]
- Número de camadas [1:5]
- Número de neurônios em cada camada [16,32,64]

Os parâmetros que permaneceram fixos durante os testes, foram:

- *Learning Rate*:  $10^{-4}$
- Função de ativação: Relu
- *Epochs*: 100
- *Batch Size*: 2

A melhor arquitetura para a previsão do erro (previsão 2) com  $k=1$  foi ordem  $ny=10$ , 5 camadas com 64 neurônios cada, onde o  $R^2=0,7924$ . Já para  $k=6$ ,  $ny=28$ , 5 camadas com 64 neurônios CADA, onde o  $R^2=0,7013$ .



Após a realização das previsões 1 e 2 para  $k=1$  e  $k=6$ , foi realizada a soma do resultado alcançado com o modelo NARMA e RNA. Para  $k=1$  o  $R^2=0,9802$ , sendo um resultado 7,34 % maior do que o atingido no Capítulo 5, onde somente a RNA foi utilizada. Já para  $k=6$  o  $R^2=0,9271$ , sendo um resultado 19,67 % maior.

As figuras a seguir foram plotadas como o resultado do modelo NARMA+RNA, para  $k=1$  e  $k=6$ .

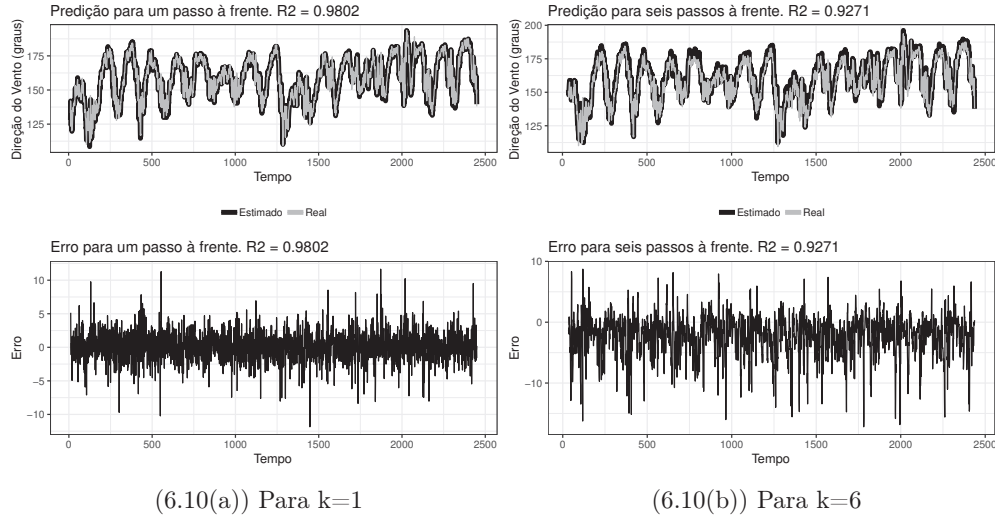


Figura 6.10: Predição da direção do vento e erro utilizando NARMA+RNA.

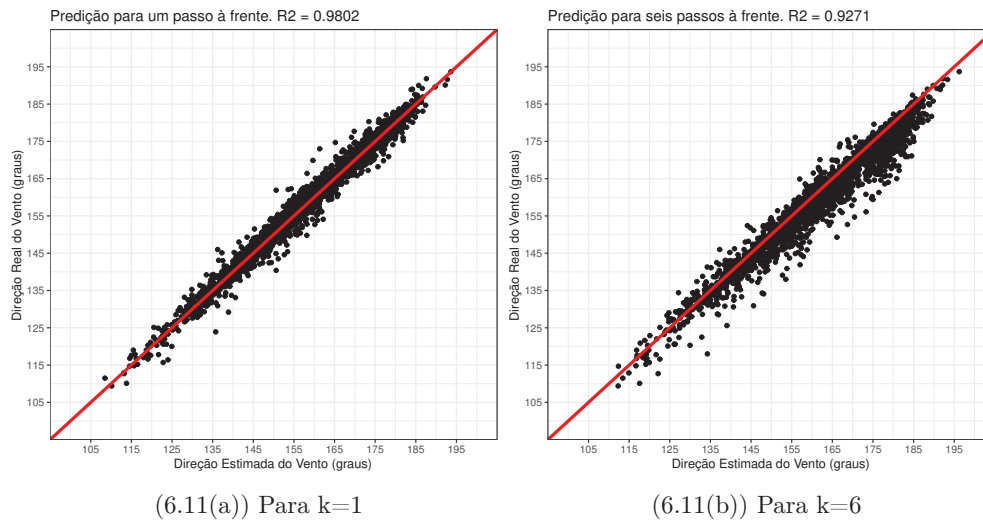


Figura 6.11: Gráfico de  $y(t)$  versus  $\hat{y}(t)$  para direção utilizando NARMA+RNA.

Da mesma forma que aconteceu para a velocidade, comparando o erro da Figura 6.10 com o das Figuras 5.5 e 5.6 do Capítulo 5, seção 5.1.2, onde

foi utilizado somente RNA na predição, percebe-se que o erro é bem menor, com valores máximos menores. Além disso, os pontos da Figura 6.11 tem uma dispersão muito menor em relação a reta  $y = x$  do que na Figura 5.7, ressaltando o melhor resultado quando é utilizado os dois modelos.

## 7

### Conclusão

Devido a crescente utilização de energia eólica e a aleatoriedade do vento, o principal objetivo desse trabalho foi construir um modelo de previsão de séries temporais que melhor se ajustasse aos dados disponíveis, minimizando o erro de previsão. O requisito de avaliação foi de previsão de 6 passos a frente, que no caso de amostragem de 10 minutos equivale a 1 hora, considerado suficiente para a aplicação. Os métodos convencionais de previsão fornecem resultados precisos quando os dados estudados apresentam um comportamento linear, porém quando há um grau de não linearidade, como nesse caso, estes métodos passam a ser pouco eficientes. Neste sentido, o uso de RNA se apresenta como uma ferramenta muito atrativa para solução desses casos. No presente trabalho, no Capítulo 6, também foi feito uso do modelo NARMA juntamente com a RNA, afim de aumentar a qualidade dos resultados alcançados.

Dessa forma, ao longo do trabalho, foram testadas diversos parâmetros de Arquitetura de Rede para os casos de aplicação da RNA e no caso do modelo NARMA, foram testadas algumas ordens em buscas dos melhores resultados possíveis, minimizando o erro de previsão. No caso da utilização somente de RNA, foi observado que para os dados utilizados, os resultados melhoravam com o aumento da complexidade da rede referente a diminuição do *learning rate*, tanto para a predição da velocidade quanto da direção. Entretanto, o aumento da complexidade demanda um maior tempo de computação do programa, realizando a estimação e a predição.

Os resultados obtidos primeiramente foram satisfatórios para a direção do vento, entretanto o mesmo não ocorreu para a velocidade, usando o horizonte de predição de 60 minutos ( $k=6$ ). Sendo assim, buscou-se refinar os resultados, utilizando o modelo NARMA primeiramente, e realizando a predição do erro com RNA, numa estratégia híbrida de previsão. Essa técnica adotada melhorou todos os resultados, para um e seis passos à frente, tanto da velocidade quanto da direção. O maior aumento de  $R^2$  foi para predição da velocidade com  $k=6$ , dado que tinha sido o pior resultado atingido anteriormente. Quantitativamente, os resultados de  $R^2$  melhoraram 22,39 % para 1 passo a frente e 51,86 % para 6 passos a frente no caso da magnitude, e 7,34 % para 1 passo a frente e 19,67 % para 6 passos a frente no caso da direção.

Em resumo, os resultados encontrados foram considerados satisfatórios por satisfazer os requisitos de projeto e prover previsão de 6 passos a frente (1 hora), em função do  $R^2$  obtido. Ainda assim há espaço para melhorias,

principalmente na escolha dos parâmetros da RNA, do modelo NARMA e na realização da previsão em um horizonte maior do que 60 minutos. Além disso, é possível que o aumento ainda maior da complexidade da rede, diminuindo o *learning rate*, atinja resultados melhores. Trabalhos futuros deverão testar diferentes arquiteturas, dado que os experimentos não apontaram uma limitação na melhora do resultado em função da complexidade.

Como forma de aprimorar e aumentar o escopo deste trabalho, as amostras disponíveis de velocidade e direção dos outros 14 geradores poderiam ser testadas. Dessa forma, seria possível estabelecer uma relação de todo parque eólico levando em conta a distribuição espacial dos geradores e a série de vento, bem como suas interdependências.

## Referências Bibliográficas

- [1] Pedro A Morettin and Clélia Toloí. *Análise de Séries Temporais*. Edgard Blucher Ltda, 2006.
- [2] Victor Henrique Gonçalves. *Previsão de séries temporais econômicas usando redes neurais caóticas*. PhD thesis, Universidade de São Paulo.
- [3] H Selcuk Nogay, Tahir Cetin Akinci, and Marija Eidukeviciute. Application of artificial neural networks for short term wind speed forecasting in mardin, turkey. *Journal of Energy in Southern Africa*, 23(4):2–7, 2012.
- [4] Henrique do Nascimento Camelo, Paulo Sérgio Lucio, João Bosco Verçosa Leal Junior, Paulo Cesar Marques de Carvalho, and Daniel von Glehn dos Santos. Innovative hybrid models for forecasting time series applied in wind generation based on the combination of time series models with artificial neural networks. *Energy*, 151:347–357, 2018.
- [5] F Tagliaferri, IM Viola, and RGJ Flay. Wind direction forecasting with artificial neural networks and support vector machines. *Ocean Engineering*, 97:65–73, 2015.
- [6] A Khosravi, RNN Koury, L Machado, and JJG Pabon. Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system. *Sustainable Energy Technologies and Assessments*, 25:146–160, 2018.
- [7] Alberto Pliego Marugán, Fausto Pedro García Márquez, Jesus María Pinar Perez, and Diego Ruiz-Hernández. A survey of artificial neural network in wind energy systems. *Applied energy*, 228:1822–1836, 2018.
- [8] Thaisa de Freitas. *Modelos e Aplicações para Séries Temporais Hierárquicas: abordagens de Reconciliação Ótima e Proporções de Previsão*. PhD thesis, PUC-Rio, 2016.
- [9] William Mendenhall, Terry Sincich, and Nancy S Boudreau. *A second course in statistics: regression analysis*, volume 5. Prentice Hall Upper Saddle River, NJ, 1996.
- [10] Luzia Vidal de Souza. Programação genética e combinação de preditores para previsão de séries temporais. 2006.
- [11] D. M. Moreira. Administração da produção e operações. 2001.

- [12] S. S. Haykin. *Neural networks and learning machines*. Prentice Hall, Upper Saddle River, 3rd edition, 2009.
- [13] SA Billings and QH Tao. Model validity tests for non-linear signal processing applications. *International Journal of Control*, 54(1):157–194, 1991.
- [14] Gabriel Trierweiler Ribeiro, Viviana Cocco Mariani, and Leandro dos Santos Coelho. Enhanced ensemble structures using wavelet neural networks applied to short-term load forecasting. *Engineering Applications of Artificial Intelligence*, 82:272–281, 2019.