



Pós-Graduação em Ciência da Computação

Rômulo Calado Pantaleão Camara

**Um Neurônio Artificial Morfológico-Linear com Aprendizagem baseada em  
Gradiente Descendente para Previsão de Séries Temporais Financeiras em Baixa  
Frequência**



Universidade Federal de Pernambuco  
posgraduacao@cin.ufpe.br  
<http://cin.ufpe.br/~posgraduacao>

Recife  
2017

Rômulo Calado Pantaleão Camara

**Um Neurônio Artificial Morfológico-Linear com Aprendizagem baseada em  
Gradiente Descendente para Previsão de Séries Temporais Financeiras em Baixa  
Frequência**

Trabalho apresentado ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutor em Ciência da Computação.

Universidade Federal de Pernambuco

Orientador: Cristiano Coelho de Araújo  
Coorientador: Ricardo de Andrade Araújo

Recife  
2017

Catálogo na fonte  
Bibliotecário Jefferson Luiz Alves Nazareno CRB4-1758

C172n Camara, Rômulo Calado Pantaleão  
Um neurônio artificial morfológico-linear com aprendizagem baseada em gradiente descendente para previsão de séries temporais financeiras em baixa frequência / Rômulo Calado Pantaleão Camara. – 2017.  
176 f.: il., fig., tab.

Orientador: Cristiano Coelho de Araújo.  
Tese (Doutorado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, Recife, 2017.  
Inclui referências e apêndices.

1. Inteligência artificial. 2. Previsão de séries temporais. I. Araújo, Cristiano Coelho de (orientador). II. Título.

006.3

CDD (23. ed.)

UFPE- MEI 2018-090

**Rômulo Calado Pantaleão Camara**

**Um Neurônio Artificial Morfológico-linear com Aprendizagem Baseada em Gradiente Descendente para Previsão de Séries Temporais Financeiras em Baixa Frequência**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação

Aprovado em: 12/06/2017

---

**Orientador: Prof. Dr. Cristianoo Coêlho Araújo**

**BANCA EXAMINADORA**

---

Prof. Dr. Geber Lisboa Ramalho  
Centro de Informática / UFPE

---

Prof. Dr. Ricardo Bastos Cavalcante Prudêncio  
Centro de Informática / UFPE

---

Prof. Dr. Adriano Lorena Inacio de Oliveira  
Centro de Informática / UFPE

---

Profa. Dra. Nadia Nedjah  
Departamento de Engenharia Eletrônica e Telecomunicações/UERJ

---

Prof. Dr. Waslon Terlizzie Araújo Lopes  
Departamento de Engenharia Elétrica / UFPB

## DEDICATÓRIA

*Dedico esse trabalho a meu irmão Ronaldo Calado Pantaleão Camara (In Memoriam), a Taise Ribeiro Moraes (In Memoriam) a minha família e aos professores que me ajudaram e incentivaram até o final desse longo caminho.*

## AGRADECIMENTOS

Inicialmente, gostaria de agradecer a Deus por me proporcionar vida, saúde, força, capacidade, felicidade e curiosidade de aprender cada vez mais. Aos meus pais, Aparecida e Raimundo, os maiores incentivadores do caminho que segui e todo o esforço vencido para eu ter uma educação de qualidade. Nunca esquecerei os almoços de todos os dias que minha mãe fazia, nem as buscas no ponto de ônibus tarde da noite que meu pai fez sem reclamar. Muito obrigado por tudo que fizeram para mim!

À minha esposa Laise Moraes por todo amor, carinho, compreensão e paciência com alguns momentos durante essa jornada de doutorado. Ao meu filho Gabriel, que está diariamente ao meu lado, provendo alegria, muito amor e aprendizado. Ao meu irmão Rafael pela paciência e atenção nos momentos que precisei, e ao meu irmão Ronaldo, que comprou meu primeiro computador na graduação e me cobrava como retribuição resultados positivos na Universidade. Tenho certeza que ele está feliz com mais essa conquista. Aos tios, tias, primos e primas pelos momentos de alegria, suavizando minha rotina durante todo o doutorado. Agradecimento especial a Rodrigo Costa Calado, pelo auxílio e incentivo nos dias próximos a minha defesa de tese.

Não posso esquecer de agradecer imensamente a colaboração do aluno e grande amigo Anderson Matias que me ajudou desde o início das pesquisas desenvolvidas até o final deste trabalho. Anderson você contribuiu como ninguém para este sucesso aprendendo e ensinando como um excelente pesquisador. Parabéns pelo seu empenho e muito obrigado por toda ajuda!

Um agradecimento especial também aos meus orientadores, Prof. Dr. Cristiano Araújo e Prof. Dr. Ricardo Araújo, por todas as suas contribuições e sugestões que enriqueceram o desenvolvimento deste trabalho. Pessoas simples mas com uma tremenda capacidade de inovar e pesquisar. Tenham certeza que nosso trabalho continuará!

Muito obrigado a toda a equipe da Secretaria de Pós-Graduação Centro de Informática, em especial a Maria Lilia Pinheiro de Freitas e Maria do Socorro Chaves de Oliveira, que sempre estavam de prontidão. Vocês também fazem parte dessa conquista!

A todos vocês - e àqueles que não me lembrei - meu fraterno abraço e cordial obrigado!

*“O destino conduz os que consentem, e arrasta os que resistem.”*  
*(Lucius Annaeus Seneca)*

## RESUMO

Atualmente existe um grande interesse da sociedade em encontrar meios de prever o futuro do mercado de ações para otimizar o processo de tomada de decisão visando a maximização do lucro em seus investimentos. Esta Tese apresenta um novo modelo de neurônio artificial morfológico-linear, denominado de Perceptron Morfológico Crescente Geral (Perceptron Morfológico Crescente Geral - General Increasing Morphological Perceptron (GIMP)), para previsão de séries temporais financeiras em baixa-frequência (diária, semanal e quinzenal). O neurônio GIMP é composto por uma combinação balanceada entre um módulo linear e um módulo não-linear crescente. Além disso, para o projeto do modelo proposto, é apresentado um processo de aprendizagem baseado em gradiente descendente com ajuste automático de fase temporal. Também é realizada uma análise experimental utilizando um conjunto de séries temporais financeiras provenientes do mercado de ações brasileiro, em baixa-frequência e os resultados obtidos foram analisados, utilizando um conjunto relevante de medidas de desempenho, e comparados aqueles obtidos utilizando modelos clássicos da literatura de previsão de séries temporais financeiras. Os resultados mostraram ganhos de desempenho de previsão de mais de 200% em relação aos modelos clássicos estudados.

**Palavras-chaves:** Previsão de Séries Temporais. Mercado de Ações em Baixa Frequência. Neurônios Artificiais. Morfologia Matemática. Gradiente Descendente. Ajuste de Fase Temporal.

## ABSTRACT

There is a great interest in society to find ways to predict the future of the stock market to optimize the decision-making process in order to maximize profit on its investments. This Doctoral Thesis presents a new morphological-linear artificial neuron model, called the General Increasing Morphological Perceptron (GIMP), for low-frequency (daily, weekly and biweekly) financial time series forecasting. The neuron GIMP is composed of a balanced combination between a linear module and an increasing nonlinear module. In addition to that, for the design of the proposed model, we present a descending gradient-based learning process with automatic time phase adjustment. Furthermore, an experimental analysis is conducted using relevant financial time series from the Brazilian stock market and the results is analyzed, according to a relevant set of performance measures, and compared to those obtained using classical models in the literature of financial time series prediction. The experimental results showed gains of more than 200% when compared with other classical models.

**Key-words:** Time Series Prediction. Low-Frequency Stock Market. Artificial Neurons. Morphological Mathematics. Descending Gradient. Time Phase Adjustment.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Estrutura de conexão de uma rede neural artificial. . . . .	43
Figura 2 – Fluxo de treinamento de uma rede neural artificial. . . . .	46
Figura 3 – Arquitetura da rede NARXTAKENS. . . . .	49
Figura 4 – Exemplo de uma máquina de vetor de suporte. . . . .	54
Figura 5 – Funções de perda utilizadas em regressão de vetor de suporte. . . . .	56
Figura 6 – Lagplot de baixa ordem da série temporal PETR4 na frequência diária. . . . .	64
Figura 7 – Lagplot de alta ordem da série temporal PETR4 na frequência diária. . . . .	65
Figura 8 – Gráfico de autocorrelação dos dados da PETR4 na frequência diária. . . . .	65
Figura 9 – Gráfico de autocorrelação parcial dos dados da PETR4 na frequência diária. . . . .	66
Figura 10 – Lagplot de baixa ordem da série temporal PETR4 na frequência semanal. . . . .	66
Figura 11 – Lagplot de alta ordem da série temporal PETR4 na frequência semanal. . . . .	67
Figura 12 – Gráfico de autocorrelação dos dados da PETR4 na frequência semanal. . . . .	67
Figura 13 – Gráfico de autocorrelação parcial dos dados da PETR4 na frequência semanal. . . . .	68
Figura 14 – Lagplot de baixa ordem da série temporal PETR4 na frequência quin- zenal. . . . .	68
Figura 15 – Lagplot de alta ordem da série temporal PETR4 na frequência quinzenal. . . . .	69
Figura 16 – Gráfico de autocorrelação dos dados da PETR4 na frequência quinzenal. . . . .	69
Figura 17 – Gráfico de autocorrelação parcial dos dados da PETR4 na frequência quinzenal. . . . .	70
Figura 18 – Lagplot de baixa ordem da série temporal VALE5 na frequência diária. . . . .	71
Figura 19 – Lagplot de alta ordem da série temporal VALE5 na frequência diária. . . . .	71
Figura 20 – Gráfico de autocorrelação dos dados da VALE5 na frequência diária. . . . .	72
Figura 21 – Gráfico de autocorrelação parcial dos dados da VALE5 na frequência diária. . . . .	72
Figura 22 – Lagplot de baixa ordem da série temporal VALE5 na frequência semanal. . . . .	73
Figura 23 – Lagplot de alta ordem da série temporal VALE5 na frequência semanal. . . . .	73
Figura 24 – Gráfico de autocorrelação dos dados da VALE5 na frequência semanal. . . . .	74
Figura 25 – Gráfico de autocorrelação parcial dos dados da VALE5 na frequência semanal. . . . .	74
Figura 26 – Lagplot de baixa ordem da série temporal VALE5 na frequência quin- zenal. . . . .	75
Figura 27 – Lagplot de alta ordem da série temporal VALE5 na frequência quinzenal. . . . .	75
Figura 28 – Gráfico de autocorrelação dos dados da VALE5 na frequência quinzenal. . . . .	76

Figura 29 – Gráfico de autocorrelação parcial dos dados da VALE5 na frequência quinzenal. . . . .	76
Figura 30 – Lagplot de baixa ordem da série temporal CSNA3 na frequência diária. . . . .	77
Figura 31 – Lagplot de alta ordem da série temporal CSNA3 na frequência diária. . . . .	77
Figura 32 – Gráfico de autocorrelação dos dados da CSNA3 na frequência diária. . . . .	78
Figura 33 – Gráfico de autocorrelação parcial dos dados da CSNA3 na frequência diária. . . . .	78
Figura 34 – Lagplot de baixa ordem da série temporal CSNA3 na frequência semanal. . . . .	79
Figura 35 – Lagplot de alta ordem da série temporal CSNA3 na frequência semanal. . . . .	79
Figura 36 – Gráfico de autocorrelação dos dados da CSNA3 na frequência semanal. . . . .	80
Figura 37 – Gráfico de autocorrelação parcial dos dados da CSNA3 na frequência semanal. . . . .	80
Figura 38 – Lagplot de baixa ordem da série temporal CSNA3 na frequência quinzenal. . . . .	81
Figura 39 – Lagplot de alta ordem da série temporal CSNA3 na frequência quinzenal. . . . .	81
Figura 40 – Gráfico de autocorrelação dos dados da CSNA3 na frequência quinzenal. . . . .	82
Figura 41 – Gráfico de autocorrelação parcial dos dados da CSNA3 na frequência quinzenal. . . . .	82
Figura 42 – Lagplot de baixa ordem da série temporal BBDC4 na frequência diária. . . . .	83
Figura 43 – Lagplot de alta ordem da série temporal BBDC4 na frequência diária. . . . .	83
Figura 44 – Gráfico de autocorrelação dos dados da BBDC4 na frequência diária. . . . .	84
Figura 45 – Gráfico de autocorrelação parcial dos dados da BBDC4 na frequência diária. . . . .	84
Figura 46 – Lagplot de baixa ordem da série temporal BBDC4 na frequência semanal. . . . .	85
Figura 47 – Lagplot de alta ordem da série temporal BBDC4 na frequência semanal. . . . .	85
Figura 48 – Gráfico de autocorrelação dos dados da BBDC4 na frequência semanal. . . . .	86
Figura 49 – Gráfico de autocorrelação parcial dos dados da BBDC4 na frequência semanal. . . . .	86
Figura 50 – Lagplot de baixa ordem da série temporal BBDC4 na frequência quinzenal. . . . .	87
Figura 51 – Lagplot de alta ordem da série temporal BBDC4 na frequência quinzenal. . . . .	87
Figura 52 – Gráfico de autocorrelação dos dados da BBDC4 na frequência quinzenal. . . . .	88
Figura 53 – Gráfico de autocorrelação parcial dos dados da BBDC4 na frequência quinzenal. . . . .	88
Figura 54 – Arquitetura do modelo GIMP. . . . .	95
Figura 55 – Procedimento de ajuste de fase automático. . . . .	103
Figura 56 – Gráfico de previsão da série temporal PETR4 diária (conjunto de teste). . . . .	112
Figura 57 – Gráfico de previsão da série temporal PETR4 semanal (conjunto de teste). . . . .	114

Figura 58 – Gráfico de previsão da série temporal PETR4 quinzenal (conjunto de teste). . . . .	115
Figura 59 – Gráfico de previsão da série temporal VALE5 diária (conjunto de teste). . . . .	119
Figura 60 – Gráfico de previsão da série temporal VALE5 semanal (conjunto de teste). . . . .	121
Figura 61 – Gráfico de previsão da série temporal VALE5 quinzenal (conjunto de teste). . . . .	123
Figura 62 – Gráfico de previsão da série temporal CSNA3 diária (conjunto de teste). . . . .	125
Figura 63 – Gráfico de previsão da série temporal CSNA3 semanal (conjunto de teste). . . . .	127
Figura 64 – Gráfico de previsão da série temporal CSNA3 quinzenal (conjunto de teste). . . . .	128
Figura 65 – Gráfico de previsão da série temporal BBDC4 diária (conjunto de teste). . . . .	130
Figura 66 – Gráfico de previsão da série temporal BBDC4 semanal (conjunto de teste). . . . .	134
Figura 67 – Gráfico de previsão da série temporal BBDC4 quinzenal (conjunto de teste). . . . .	136
Figura 68 – Estrutura do nó de julgamento do GNP-CVaR. . . . .	162
Figura 69 – Estrutura do nó de julgamento do <i>Genetic Network Programming (GNP) and Mean Conditional Value-at-Risk Model (GNP-CVaR)</i> . . . . .	163
Figura 70 – Estrutura do modelo <i>Genetic Network Programming (GNP) and Mean Conditional Value-at-Risk Model (GNP-CVaR)</i> . . . . .	163
Figura 71 – Estrutura do modelo <i>Genetic Network Programming (GNP) and Mean Conditional Value-at-Risk Model (GNP-CVaR)</i> . . . . .	164
Figura 72 – Estrutura do modelo <i>Multi-Objective Evolutionary Algorithms</i> . . . . .	165
Figura 73 – Estrutura do modelo MGA-MLP. . . . .	168
Figura 74 – Estrutura do modelo PSO-MLP. . . . .	171
Figura 75 – Estrutura do modelo TAEF. . . . .	173
Figura 76 – Estrutura do modelo SHIF. . . . .	175
Figura 77 – Estrutura do modelo DELP. . . . .	176

## LISTA DE TABELAS

Tabela 1 – Diferença entre redes MLP e RBF. . . . .	53
Tabela 2 – Comparação entre Modelagens para Séries. . . . .	61
Tabela 3 – Estatísticas PETR4 na frequência diária. . . . .	65
Tabela 4 – Estatísticas PETR4 na frequência semanal. . . . .	67
Tabela 5 – Estatísticas PETR4 na frequência quinzenal. . . . .	69
Tabela 6 – Estatísticas VALE5 na frequência diária. . . . .	73
Tabela 7 – Estatísticas VALE5 na frequência semanal. . . . .	75
Tabela 8 – Estatísticas VALE5 na frequência quinzenal. . . . .	77
Tabela 9 – Estatísticas CSNA3 na frequência diária. . . . .	79
Tabela 10 – Estatísticas CSNA3 na frequência semanal. . . . .	81
Tabela 11 – Estatísticas CSNA3 na frequência quinzenal. . . . .	83
Tabela 12 – Estatísticas BBDC4 na frequência diária. . . . .	85
Tabela 13 – Estatísticas BBDC4 na frequência semanal. . . . .	87
Tabela 14 – Estatísticas BBDC4 na frequência quinzenal. . . . .	89
Tabela 15 – Desempenho preditivo para a série PETR4 - Frequência Diária (conjunto de teste). . . . .	111
Tabela 16 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série PETR4 diário (conjunto teste). . . . .	112
Tabela 17 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal PETR4 Diário . . . . .	112
Tabela 18 – Desempenho preditivo para a série PETR4 - Frequência Semanal (conjunto de teste). . . . .	113
Tabela 19 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série PETR4 semanal (conjunto teste). . . . .	114
Tabela 20 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal PETR4 Semanal (conjunto de teste) . . . . .	115
Tabela 21 – Desempenho preditivo para a série PETR4 - Frequência Quinzenal (conjunto de teste). . . . .	116
Tabela 22 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série PETR4 quinzenal (conjunto teste). . . . .	117
Tabela 23 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal PETR4 quinzenal(conjunto de teste) . . . . .	117
Tabela 24 – Desempenho preditivo para a série VALE5 - Frequência Diária (conjunto de teste). . . . .	118
Tabela 25 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série VALE5 diário (conjunto teste). . . . .	119

Tabela 26 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal VALE5 Diário . . . . .	119
Tabela 27 – Desempenho preditivo para a série VALE5 - Frequência Semanal (conjunto de teste). . . . .	120
Tabela 28 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série VALE5 semanal (conjunto teste). . . . .	121
Tabela 29 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal VALE5 Semanal (conjunto de teste) . . . . .	121
Tabela 30 – Desempenho preditivo para a série VALE5 - Frequência Quinzenal (conjunto de teste). . . . .	122
Tabela 31 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série VALE5 quinzenal (conjunto teste). . . . .	123
Tabela 32 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal VALE5 quinzenal (conjunto de teste) . . . . .	123
Tabela 33 – Desempenho preditivo para a série CSNA3 - Frequência Diária (conjunto de teste). . . . .	124
Tabela 34 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série CSNA3 diária (conjunto teste). . . . .	125
Tabela 35 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal CNSA3 Diário (conjunto de teste) . . . . .	125
Tabela 36 – Desempenho preditivo para a série CSNA3 - Frequência Semanal (conjunto de teste). . . . .	126
Tabela 37 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série CSNA3 semanal (conjunto teste). . . . .	127
Tabela 38 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal CNSA3 Semanal . . . . .	128
Tabela 39 – Desempenho preditivo para a série CSNA3 - Frequência Quinzenal (conjunto de teste). . . . .	129
Tabela 40 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série CSNA3 quinzenal (conjunto teste). . . . .	130
Tabela 41 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal CNSA3 quinzenal (conjunto de teste) . . . . .	130
Tabela 42 – Desempenho preditivo para a série BBDC4 - Frequência Diária (conjunto de teste). . . . .	131
Tabela 43 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série BBDC4 diária (conjunto teste). . . . .	132
Tabela 44 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal BBDC4 diária (conjunto de teste) . . . . .	132

Tabela 45 – Desempenho preditivo para a série BBDC4 - Frequência Semanal (conjunto de teste). . . . .	133
Tabela 46 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série BBDC4 semanal (conjunto teste). . . . .	134
Tabela 47 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal BBDC4 semanal (conjunto de teste) . . . . .	134
Tabela 48 – Desempenho preditivo para a série BBDC4 - Frequência Quinzenal (conjunto de teste). . . . .	135
Tabela 49 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série BBDC4 quinzenal (conjunto teste). . . . .	136
Tabela 50 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal BBDC4 quinzenal (conjunto de teste) . . . . .	137
Tabela 51 – Média dos valores de $\lambda$ e $\theta$ para os experimento analisados. . . . .	139

## LISTA DE ABREVIATURAS E SIGLAS

**ACF** Função de Autocorrelação - autocorrelation function.

**AI** Inteligência Artificial - Artificial Intelligence.

**AMI** Informação Mútua Média - Average Mutual Information.

**AR** Modelo Autorregressivo - Autoregressive.

**ARIMA** Modelo Autorregressivo Integrado de Médias Móveis - Autoregressive Integrated Moving Average.

**ARMA** Modelo Autorregressivo de Médias Móveis - Autoregressive Moving Average.

**ARV** Variância Relativa Média - Average Relative Variance.

**BBDC4** Banco Bradesco SA.

**CSNA3** Ações Ordinárias da Empresa Companhia Siderúrgica Nacional.

**EF** Função de Avaliação - Evaluation Function.

**FAC** Função de Autocorrelação - Autocorrelation Function.

**GIMP** Perceptron Morfológico Crescente Geral - General Increasing Morphological Perceptron.

**LOGREG** Regressão Logística - Logistic Regression.

**MA** Médias Móveis - Moving Average.

**MAPE** Erro Médio Percentual Absoluto - Mean Absolute Percentage Error.

**MSE** Erro Médio Quadrático - Mean Squared Error.

**PACF** Função de Autocorrelação Parcial - Partial Autocorrelation Function.

**PETR4** Ações Preferenciais da Empresa Petrobras.

**POCID** Previsão de Mudança na Direção - Prediction of Change in Direction.

**RNA** Redes Neurais Artificiais.

**RWD** Dilema do Passeio Aleatório - Random Walk Dilemma.

**THEIL** Estatística U de Theil - U of Theil statistics.

**VALE5** Ações Preferenciais da Empresa Vale SA.

**VRT** Teste de Razão de Variância - Variance Ratio Test.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>20</b>
1.1	INTRODUÇÃO	20
1.2	MOTIVAÇÃO E JUSTIFICATIVA	23
1.3	OBJETIVOS E CONTRIBUIÇÕES	26
1.4	ESTRUTURA DA TESE	28
<b>2</b>	<b>SÉRIE TEMPORAL E O PROBLEMA DE PREVISÃO</b>	<b>30</b>
2.1	SÉRIES TEMPORAIS	30
<b>2.1.1</b>	<b>Análise de Séries Temporais</b>	<b>31</b>
2.2	PROBLEMA DA PREVISÃO DE SÉRIES TEMPORAIS	32
2.3	DILEMA DO PASSEIO ALEATÓRIO	33
<b>2.3.1</b>	<b>O conceito do Ajuste de Fase Temporal</b>	<b>35</b>
2.4	RESUMO DO CAPÍTULO	36
<b>3</b>	<b>O ESTADO DA ARTE</b>	<b>37</b>
3.1	INTRODUÇÃO	37
3.2	MODELOS ESTATÍSTICOS	38
<b>3.2.1</b>	<b>Modelos Lineares</b>	<b>38</b>
3.3	MODELOS DE REDES NEURAS ARTIFICIAIS	42
<b>3.3.1</b>	<b>Processo de treinamento de uma Rede Neural Artificial utilizada para Previsão de Séries Temporais Financeiras</b>	<b>44</b>
<b>3.3.2</b>	<b>Rede de <i>Perceptron</i> Multicamadas</b>	<b>44</b>
3.4	MODELO AUTORREGRESSIVO NÃO-LINEAR COM ENTRADA EXÓGENA ( <i>NONLINEAR AUTOREGRESSIVE EXOGENOUS MODEL</i> ) – <i>NARX</i>	47
3.5	REGRESSÃO LOGÍSTICA	50
3.6	REDE DE FUNÇÃO DE BASE RADIAL	51
3.7	MODELO DE REGRESSÃO DE VETOR SUPORTE	53
<b>3.7.1</b>	<b>Regressão de Vetor de Suporte Linear</b>	<b>55</b>
3.7.1.1	Regressão linear com função de erro $\varepsilon$ -insensitive	56
3.8	REGRESSÃO LINEAR COM FUNÇÃO DE ERRO QUADRÁTICA	58
<b>3.8.1</b>	<b>Regressão de Vetor de Suporte Não Linear</b>	<b>58</b>
3.8.1.1	Kernel Polinomial	59
3.8.1.2	Kernel Função de Base Radial	60
3.9	RESUMO DO CAPÍTULO	61
<b>4</b>	<b>ANÁLISE DAS SÉRIES TEMPORAIS FINANCEIRAS</b>	<b>62</b>

4.1	ANÁLISE DAS SÉRIES TEMPORAIS FINANCEIRAS . . . . .	62
4.1.1	<b>Análise do ativo PETR4 na frequência diária . . . . .</b>	<b>64</b>
4.1.2	<b>Análise do ativo PETR4 na frequência semanal . . . . .</b>	<b>66</b>
4.1.3	<b>Análise do ativo PETR4 na frequências quinzenal . . . . .</b>	<b>68</b>
4.1.4	<b>Análise do ativo VALE5 na frequência diária . . . . .</b>	<b>71</b>
4.1.5	<b>Análise do ativo VALE5 na frequência semanal. . . . .</b>	<b>72</b>
4.1.6	<b>Análise do ativo VALE5 na frequências quinzenal . . . . .</b>	<b>74</b>
4.1.7	<b>Análise do ativo CSNA3 na frequência diária . . . . .</b>	<b>77</b>
4.1.8	<b>Análise do ativo CSNA3 na frequência semanal . . . . .</b>	<b>79</b>
4.1.9	<b>Análise do ativo CSNA3 na frequências quinzenal . . . . .</b>	<b>81</b>
4.1.10	<b>Análise do ativo BBDC4 na frequência diária . . . . .</b>	<b>83</b>
4.1.11	<b>Análise do ativo BBDC4 na frequência semanal . . . . .</b>	<b>85</b>
4.1.12	<b>Análise do ativo BBDC4 na frequências quinzenal . . . . .</b>	<b>87</b>
4.1.13	<b>Considerações analíticas sobre as séries estudadas . . . . .</b>	<b>89</b>
4.2	RESUMO DO CAPÍTULO . . . . .	90
<b>5</b>	<b>O MODELO PROPOSTO . . . . .</b>	<b>91</b>
5.1	INTRODUÇÃO . . . . .	91
5.2	FUNDAMENTAÇÃO TEÓRICA RELEVANTE . . . . .	92
5.2.1	<b>Teoria do Reticulado Completo e a Morfologia Matemática . . . . .</b>	<b>92</b>
5.3	DEFINIÇÃO DO MODELO GIMP . . . . .	95
5.3.1	<b>Fundamentos do Processo de Aprendizagem . . . . .</b>	<b>97</b>
5.3.2	<b>Proposta do Algoritmo de Treinamento com o Ajuste de Fase Tem- poral . . . . .</b>	<b>100</b>
5.3.3	<b>Utilização do Ajuste de Fase Temporal no modelo proposto . . . . .</b>	<b>103</b>
5.4	RESUMO DO CAPÍTULO . . . . .	103
<b>6</b>	<b>SIMULAÇÕES E RESULTADOS EXPERIMENTAIS . . . . .</b>	<b>104</b>
6.1	METODOLOGIA . . . . .	104
6.2	MEDIDAS PARA DESEMPENHO DE PREVISÃO . . . . .	108
6.3	RESULTADOS . . . . .	110
6.3.1	<b>PETR4 - Frequência Diária . . . . .</b>	<b>110</b>
6.3.2	<b>PETR4 - Frequência Semanal . . . . .</b>	<b>112</b>
6.3.3	<b>PETR4 - Frequência Quinzenal . . . . .</b>	<b>114</b>
6.3.4	<b>VALE5 - Frequência Diária . . . . .</b>	<b>115</b>
6.3.5	<b>VALE5 - Frequência Semanal . . . . .</b>	<b>117</b>
6.3.6	<b>VALE5 - Frequência Quinzenal . . . . .</b>	<b>121</b>
6.3.7	<b>CSNA3 - Frequência Diária . . . . .</b>	<b>123</b>
6.3.8	<b>CSNA3 - Frequência Semanal . . . . .</b>	<b>126</b>
6.3.9	<b>CSNA3 - Frequência Quinzenal . . . . .</b>	<b>127</b>

6.3.10	<b>BBDC4 - Frequência Diária</b> . . . . .	<b>128</b>
6.3.11	<b>BBDC4 - Frequência Semanal</b> . . . . .	<b>132</b>
6.3.12	<b>BBDC4 - Frequência Quinzenal</b> . . . . .	<b>134</b>
6.3.13	<b>Considerações</b> . . . . .	<b>136</b>
6.3.13.1	Comparação com outros modelos do Estado da Arte PETR4 diário e Vale5 diário . . . . .	138
6.3.13.2	Análise dos valores de $\lambda$ e $\theta$ . . . . .	139
6.4	<b>RESUMO</b> . . . . .	<b>140</b>
<b>7</b>	<b>CONCLUSÃO</b> . . . . .	<b>141</b>
7.1	<b>TRABALHOS FUTUROS</b> . . . . .	<b>143</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>145</b>
	<b>APÊNDICES</b>	<b>155</b>
	<b>APÊNDICE A – OUTROS MODELOS PARA PREVISÃO DE SÉRIAS TEMPORAIS</b> . . . . .	<b>156</b>

# 1 INTRODUÇÃO

## 1.1 INTRODUÇÃO

A sociedade contemporânea tem sido caracterizada como uma sociedade da informação, devido ao papel central que a informação assume através das novas Tecnologias de Informação e Comunicação (TIC). Atualmente, a utilização das TICs e o processamento das informações de forma organizada geram dados muito importantes para a análise econômica, social, política e filosófica mundial.

No cenário econômico, pode-se observar que de todos os setores industriais o nicho de empresas que trabalham com informação é o que mais cresce no mundo (CUMMING; JOHAN; SCHWEIZER, 2017), sendo considerado um dos setores mais importantes por algumas razões, tais como: transformar informações em bens e serviços, produzir mídia (filmes, novelas, livros e periódicos - formas mais utilizadas de informações pelo público em geral), desenvolver o processamento de informações que influenciaram vários serviços (bancário, jurídico, vendas no varejo e até o mercado financeiro mundial).

Em particular o mercado financeiro está utilizando cada vez mais informações econômicas para definir as tomadas de decisões de compra e venda de um ativo. No mercado financeiro, o investidor está utilizando informações para apoiar a tomada de decisões e na rapidez com que ele é capaz de tomá-las e assim obter mais lucros em seus investimentos (CAVALCANTE et al., 2016). A utilização de plataformas de negociação eletrônica pode auxiliar esses investidores, visto que através delas podem ser desenvolvidas ferramentas de apoio a decisão, como indicadores e robôs de negociação. Atualmente 43% das negociações da BOVESPA são através de Acesso direto ao mercado DMA e 70% dessas são rônôs. Isso mostra o crescimento desse meio de negociação e lavam pesquisadores a pesquisarem ferramentas para esse tipo de sistema.

Informações econômicas são utilizadas para prever valores futuros de uma série temporal financeira, tanto em baixa frequência (ARAÚJO; SUSSNER, 2010) e (ARAÚJO, 2012b), quanto em operações em alta frequência (ARAÚJO, 2016). Observou-se nos trabalhos atuais, a busca por trabalhar em baixa frequência devido à facilidade de identificação do comportamento da série temporal financeira.

Neste contexto, a análise de séries temporais tem o intuito de analisar as características da série para construção modelos de previsão que sejam capazes de gerar valores futuros da série, com base nos valores passados. Para tal, diversos modelos de previsão têm sido abordados na literatura para prever séries temporais financeiras (NEKOUKAR; BEHESHTI, 2010; ADEODATO et al., 2011; ASADI et al., 2012; SHAHRABI; HADAVANDI; ASADI, 2013; ASADI; TAVAKOLI; HEJAZI, 2012; ADHIKARI; AGRAWAL, 2013; ARAÚJO, 2016), figurando como um desafio recorrente para a comunidade científica no geral, uma vez que prever o

futuro é a chave para o sucesso de uma tomada de decisão (ARAÚJO, 2016).

De acordo com (ARAÚJO; SUSSNER, 2010) e (ARAÚJO, 2012b), modelos de previsão são responsáveis pela construção de mapeamentos capazes de estimar o futuro a partir do passado e do presente. Independentemente do sistema a ser utilizado, este possui algum tipo de imprecisão, seja por falha na construção do sistema ou pela complexidade do fenômeno temporal.

A magnitude destas imprecisões depende da escolha propriamente dita do tipo de sistema de previsão e do seu custo, em termos de recursos computacionais, para geração das previsões. Araújo (ARAÚJO, 2012b) também afirma que devido ao processo de previsão, não é possível eliminar totalmente as imprecisões, sendo necessário sua inclusão no processo de tomada da decisão, sendo esta definida por:

$$DA = DP + E, \quad (1.1)$$

em que  $DA$  é a decisão atual,  $DP$  é a decisão assumindo que a previsão está correta, e  $E$  é a compensação devido à imprecisão no sistema de previsão.

Pode-se observar na literatura que os procedimentos, para a realização do processo de previsão podem ser divididos, de forma genérica, em duas grandes categorias (BOX et al., 2015):

1. Automáticos – procedimentos aplicados diretamente aos dados históricos de um fenômeno temporal para geração de previsão, a partir de um algoritmo determinístico que pode ser executado numa arquitetura de Von Neumann, ou seja, as observações do tempo passado de um dado fenômeno temporal, organizado como uma série temporal, são utilizadas para construir um mapeamento capaz de prever as observações no tempo futuro do fenômeno temporal, sem a necessidade de intervenção humana;
2. Não automáticos – procedimento que exige intervenção humana especializada no problema, sendo necessário intervenções na análise, construção e utilização prática no preditor de um fenômeno temporal.

Esta tese está relacionada ao desenvolvimento de modelos automáticos (sem intervenção humana) para previsão de séries temporais, na tentativa de prever um valor futuro de um determinado fenômeno temporal baseado nos valores passados e presentes de uma série temporal financeira. Porém, mesmo o processo de previsão sendo um sistema automático, faz-se necessário a organização dos dados de entrada (normalização, retirada de ruídos), procedimento feito neste estudo para todas as séries temporais financeiras analisadas.

Para construção do modelo proposto foram analisados os principais modelos clássicos existentes na literatura. Constatou-se que tais modelos têm desempenho preditivo muito bom quando aplicados a séries temporais não financeiras. Porém, a maioria não alcança o

mesmo desempenho quando trata-se de previsão de séries temporais financeiras. Este fato foi discutido nos trabalhos de (ARAÚJO, 2016) e (FERREIRA, 2006), e será o problema abordado neste trabalho.

Modelos matemáticos lineares e não-lineares já foram utilizados para estudar esse tipo de série temporal. O modelo *Autoregressive Integrated Moving Average* (ARIMA) (BOX et al., 2015) é um dos modelos mais utilizados para o estudo dessas séries; todavia, alguns estudos (FERREIRA, 2006; ARAÚJO, 2016) mostram que para séries temporais financeiras de baixa frequência, existe uma componente com comportamento não linear.

Alguns modelos foram criados com o intuito de prever melhor um valor futuro de uma série temporal financeira, buscando mapear tanto a componente linear quanto a componente não linear da função. Os modelos Dilatation-Erosion Linear Perceptron (ARAÚJO, 2012b), a Rede *Nonlinear Autoregressive Network with Exogenous Input* (NARX) (MENEZES JR.; BARRETO, 2008), a *Multilayer Perceptron* (MLP) (BRAGA; CARVALHO; LUDERMIR, 2000; FERREIRA, 2006), o *Time-Delay Added Evolutionary Forecasting* (TAEF) (FERREIRA, 2006) e o *Swarm-Based Hybrid Intelligent Forecasting* (SHIF) (ARAÚJO, 2007) são exemplos que evidenciam a existência de uma componente não-linear em séries temporais financeiras, em frequência diária.

É importante ressaltar que foram criadas metodologias híbridas baseadas em redes neurais, utilizando técnicas de gradiente descendente e técnicas evolucionárias, que buscam escolher parâmetros estruturais e arquiteturas subótimas dos modelos de previsão, bem como os melhores retardos temporais para caracterizar corretamente o fenômeno gerador das séries temporais financeiras. Contudo, tal técnica pode ser ineficaz com relação ao custo computacional, gerando um aumento de custo para a construção do preditor, sem implicar no aumento do desempenho preditivo do sistema (FERREIRA; VASCONCELOS; ADEODATO, 2008). Exemplos dessas redes são: a *Multilayer Perceptron com uso de Algoritmo Genético Modificado* (MLP-MGA) (FERREIRA, 2006), a *Multilayer Perceptron com Otimizador de Enxame de Partículas* (PSO-MLP) (ARAÚJO, 2012b).

No entanto, a literatura mostra que mesmo com modelos de previsão mais sofisticados, surge uma limitação para o caso de previsão de séries financeiras, o que chamamos de Dilema do Passeio Aleatório (RWD) (MALKIEL, 2007; SITTE; SITTE, 2002). Isso ocorre porque as séries temporais financeiras são definidas por alguns pesquisadores (ARAÚJO, 2016; ARAÚJO; SUSSNER, 2010; ADEODATO et al., 2011; SALGADO et al., 2010; SITTE; SITTE, 2002) como uma série Random Walk, ou seja, tem um comportamento puramente aleatório e quando se utiliza qualquer modelo linear de previsão de séries temporais em geral, o resultado cai no problema do Dilema do Passeio Aleatório com baixo desempenho preditivo. Esse dilema é caracterizado pelo atraso de um passo a frente em relação ao valores da série temporal (MALKIEL, 2007; SITTE; SITTE, 2002).

Nos trabalhos de (MALKIEL, 2007) e (SITTE; SITTE, 2002), argumenta-se que devido ao fato acima exposto as séries temporais financeiras são definidas por um processo pura-

mente aleatório, embora não exista prova matemática que sustente tal argumento. Artigos mais recentes (FERREIRA; VASCONCELOS; ADEODATO, 2008) mostram que há a possibilidade de uma componente não linear subdominante fazer parte da função de mapeamento de uma série temporal financeira de baixa frequência. Com base nesse fator, Ferreira e Adeodato apresentaram um procedimento chamado de ajuste de fase (FERREIRA; VASCONCELOS; ADEODATO, 2008), utilizando uma rede MLP construída por meio da combinação entre um algoritmo evolucionário (LEUNG et al., 2003b) e técnicas baseadas em gradiente descendente e métodos de segunda ordem (WIDROW; HOFF et al., 1960; BATTITI, 1992; HAGAN; MENHAJ, 1994; MØLLER, 1993; RIEDMILLER; BRAUN, 1993). Ferreira (FERREIRA; VASCONCELOS; ADEODATO, 2008) demonstra de forma prática a resolução do problema do Dilema do Passeio Aleatório para o caso particular em frequência diária. A literatura mostra que outros autores utilizaram essa mesma técnica em outros trabalhos sobre séries temporais financeiras (ARAÚJO, 2007; SALGADO et al., 2010; ARAÚJO; OLIVEIRA; MEIRA, 2015).

Dessa forma, é importante investigar as peculiaridades de uma série temporal financeira para tentar superar o problema do passeio aleatório com um alto desempenho de previsão. Neste sentido, diversos trabalhos tem empregado o conceito de ajuste de fase para modelos de neurônios morfológicos (ARAÚJO, 2010a; ARAÚJO, 2010b; ARAÚJO, 2010c; ARAÚJO, 2010d; ARAÚJO; FERREIRA, 2013; ARAÚJO, 2011a; ARAÚJO, 2011b; ARAÚJO; SOARES; OLIVEIRA, 2012; ARAÚJO, 2012a; SUSSNER; VALLE, 2007; SUSSNER; ESMI, 2011). Nestes trabalhos, o uso da morfologia matemática se mostrou uma boa prática, devido aos resultados obtidos nos experimentos desenvolvidos, reforçando a ideia de que o uso de um modelo de neurônio baseado na morfologia matemática otimiza os resultados apresentados por modelos clássicos.

## 1.2 MOTIVAÇÃO E JUSTIFICATIVA

Atualmente, qualquer informação pode conter dados importantes que influenciam a sociedade de forma direta e instantaneamente, assim como a área econômica. O volume de informações é muito grande, sendo estas armazenadas, analisadas e utilizadas na tentativa de definir padrões econômicos relacionados a transações bancárias, dados de compra e venda de produtos, dados de custos de manutenção de uma empresa, pesquisa de opiniões de clientes até dados de usuários de redes sociais (RAO; SRIVASTAVA, 2013; ARAÚJO; OLIVEIRA; MEIRA, 2015).

Os referidos dados trazem informações relevantes que podem ser utilizadas na tomada de decisão de uma indústria ou comércio, pois quanto mais essas empresas conhecerem seus clientes, maior será a possibilidade de adequar seus produtos ao público em geral, ampliando, conseqüentemente, a possibilidade de venda e lucro (JOHANSON; VAHLNE, 1977).

No mercado financeiro, as informações também são de suma importância, pois informações sobre o Produto Interno Bruto, Produção Industrial, Corrupção, Índice de Emprego de um país pode influenciar diretamente o mercado nacional e até mundial. Devido à conexão entre os mercados mundiais, qualquer informação que afete um país como a China, pode afetar indiretamente e imediatamente a economia do Brasil, visto que este é um grande exportador de Minério de Ferro que é utilizado para a produção industrial chinesa (RODRIGUEZ; BLAZQUEZ; SANTISO, 2008).

A ausência de informação para uma pessoa que pretende investir no mercado financeiro pode lhe causar um enorme prejuízo financeiro. De acordo com a Forex-money-management, 90% (noventa por cento) dos novos investidores no mercado de Forex vai à falência, e um dos principais motivos é a falta de informação sobre como funciona tal mercado (MANAGEMENT, 2016).

Isto posto, torna-se inquestionável que o estudo das informações tem um grande valor para quem pretende investir no mercado financeiro. Informações do passado e do presente podem conter dados que auxiliam no entendimento do mercado. De acordo com (NIU; WANG, 2014), a previsão de séries temporais financeiras continua sendo um problema difícil de ser solucionado, uma vez que nestas séries temporais é possível encontrar diversas características complexas em seu fenômeno gerador.

Neste cenário, o estudo de soluções para previsão de séries temporais financeiras revela-se de extrema importância, pois intrínseco a estes dados pode-se observar características como irregularidades, volatilidade, tendências e ruídos. Cavalcanti *et al.* (CAVALCANTE *et al.*, 2016) afirmam que não há um método bem estabelecido para a construção de um sistema de negociação inteligente bem sucedido. Contudo, o trabalho de Araújo (ARAÚJO, 2016) pode ser considerado um exemplo do uso de um sistema de previsão de séries temporais para o desenvolvimento de um sistema de negociação inteligente.

Diversos trabalhos, provenientes de várias áreas do conhecimento humano, têm sido estudados na tentativa de desenvolver modelos de previsão que mapeiem o comportamento futuro de um fenômeno temporal (ARAÚJO, 2012b; FERREIRA, 2006; NETO, 2012). Nas áreas de matemática e estatística, modelos lineares e não lineares foram propostos com o intuito de solucionar o problema de previsão. Um dos modelos lineares estatísticos mais conhecidos da literatura é o *Autoregressive Integrated Moving Average* (ARIMA). Porém, outros modelos estatísticos lineares e não lineares também foram desenvolvidos anteriormente, com a mesma finalidade de prever fenômeno temporal financeiro. O ARIMA foi modelado em 1976 pelos estatísticos George Box e Gwilym Jenkins, sendo por este motivo também conhecido como Modelo de Box-Jenkins (BOX *et al.*, 2015; GREENE, 1995). Por ser um modelo linear, sofre limitações quanto ao uso no estudo de séries temporais financeiras, pois estas também são governadas por fenômenos não lineares (ARAÚJO, 2016; ARAÚJO, 2012b). Clements (CLEMETS; FRANCES; SWANSON, 2004) afirma que não há justificativa para suposição de linearidade no fenômeno gerador de séries temporais financeiras.

A fim de superar as limitações do ARIMA, outros modelos estatísticos não lineares têm sido desenvolvidos para a previsão de séries temporais. Dentre eles, pode-se citar os modelos heterocedásticos condicionais autorregressivos (*autoregressive conditional heteroskedasticity*) (ENGLE, 1982), modelos bilineares (RAO; GABR, 2012), modelos de limiar autorregressivo (FERREIRA, 2006; ARAÚJO, 2012b), modelos de estados gerais dependentes (FERREIRA, 2006) e modelos autorregressivo exponenciais (BILLARD, 2014).

De acordo com Ferreira (FERREIRA, 2006), cada um destes modelos estatísticos não lineares tem suas vantagens e desvantagens. Embora o objetivo principal da inclusão da não linearidade nos modelos de previsão seja o aumento da capacidade de descrição dos fenômenos não lineares, os resultados práticos com esses modelos são insatisfatórios e podem ser comparados com modelos puramente lineares. Além deste fato, pode-se observar que o aumento da complexidade de um sistema não-linear impede o desenvolvimento de um sistema puramente automático de previsão, pois é necessária a atuação de um especialista no problema para validar as previsões geradas (ARAÚJO, 2012b) e o desempenho preditivo se assimila a um preditor linear (CLEMENTS; FRANCES; SWANSON, 2004).

Na perspectiva de previsão de séries temporais utilizando modelos não lineares estatísticos, pode-se observar estudos atuais que tentam desenvolver modelos mistos com desempenho melhorado (KÖMM; KÜSTERS, 2015). Outro caminho para sanar tal problema seria o uso de Inteligência Artificial, mais especificamente o uso de redes neurais artificiais.

As Redes Neurais têm sido aplicadas com sucesso para a modelagem não linear de alguns tipos de séries temporais, inclusive as financeiras (ARAÚJO; OLIVEIRA; MEIRA, 2015; CECEN; UGUR, 2005; RATHER; AGARWAL; SASTRY, 2015; CHONG; LAM, 2010; CLEMENTS; FRANCES; SWANSON, 2004; ADEODATO et al., 2011; ASADI et al., 2012; JÚNIOR; BARRETO, 2013; MENEZES JR.; BARRETO, 2008). Comparados com modelos puramente estatísticos não lineares, a complexidade de desenvolvimento é menor. Todavia, de acordo com (ARAÚJO, 2012b), as redes neurais necessitam de um conjunto de parâmetros bem definido: a topologia, a quantidade de unidades de processamento na camada escondida, o algoritmo de treinamento e a quantidade de observações - todos esse valores são exemplos de parâmetros que necessitam ser otimizados.

Com o objetivo de auxiliar as redes neurais a gerarem os melhores parâmetros para o treinamento, outras técnicas de Inteligência Artificial são observadas no estado da arte: algoritmos Genéticos, Programação Genética, Algoritmos Evolucionários, *Particle Swarm Optimization* (PSO) (LEUNG et al., 2003b; HANSEN; MÜLLER; KOUMOUTSAKOS, 2003; KAR; NAYAK; NAYAK, 2016; SILVA et al., 2016). A combinação dessas técnicas de Inteligência Artificial cria novas metodologias híbridas, que são utilizadas para análise e estudo de séries temporais.

Com fulcro no Dilema do Passeio Aleatório, alguns pesquisadores argumentam que séries temporais financeiras são geradas de forma aleatória e não podem ser previstas (ARAÚJO, 2016). Contudo, trabalhos recentes mostram que as séries temporais financeiras

são geradas por componentes não lineares subdominantes combinadas a uma componente linear dominante. Este fenômeno é observado quando a ordem dos retardos temporais analisados tem um curto espaçamento entre si. No momento em que o espaçamento entre os valores começa a aumentar, a componente não linear se torna dominante e a componente linear se torna subdominante. Araújo (ARAÚJO; OLIVEIRA; MEIRA, 2015) mostrou, através do conceito de *Lagplot*, a autodependência entre os valores de uma série temporal.

No trabalho de Ferreira (FERREIRA, 2006) foi proposto um modelo de ajuste de fase para modelos híbridos, com foco na determinação de retardos temporais de uma rede perceptron multicamadas (MLP) através de um algoritmo genético, partindo do pressuposto de que havia uma não linearidade nas séries temporais financeiras. Também foi proposto um procedimento de ajuste de fase que funciona para um modelo híbrido que combina Algoritmo Genético (AG), Rede MLP e algoritmo gradiente descendente.

Algumas limitações foram encontradas no trabalho de Ferreira. A primeira delas foi que o sistema híbrido com ajuste de fases funcionaria para séries financeiras com frequência diária. Não foi desenvolvido estudo sobre frequências usuais de séries temporais financeiras (F-frequências): um minuto, cinco minutos, quinze minutos, trinta minutos, uma hora, quatro horas, uma semana e um mês. Dessa forma, não há indícios no trabalho de Ferreira (FERREIRA, 2006) que consiga generalizar o conceito de ajuste de fases proposto para séries temporais com frequências diferentes da frequência diária. Outra limitação é o uso unicamente de uma rede MLP, que não é um tipo de rede neural artificial bem utilizada para problemas de memória de longo prazo (ARAÚJO, 2012b; MENEZES JR.; BARRETO, 2008). A terceira limitação foi a ausência da análise do erro preditivo para tentar identificar se o ruído existente no modelo é um ruído branco, estudo que constata se um modelo de previsão de série temporal extrai todas as características do fenômeno gerador da série temporal.

Identificou-se a importância em investigar as peculiaridades de uma série temporal financeira para tentar superar o problema do passeio aleatório com um alto desempenho de previsão, visto que não foi encontrado na literatura nenhuma prova matemática que defina as séries temporais financeiras como séries aleatórias além de ter ficado evidente que grandes esforços ainda são necessários para o desenvolvimento de modelos de neurônios com o uso da Inteligência Artificial, visando gerar modelos ótimos que utilizem componentes lineares e não-lineares que superem o Dilema do Passeio Aleatório. Outro fator relevante observado foi a falta de um modelo que pudesse generalizar o fenômeno temporal de qualquer série temporal financeira, devido ao seu fenômeno gerador ser bastante diversificado.

### 1.3 OBJETIVOS E CONTRIBUIÇÕES

O trabalho proposto tem como principal objetivo desenvolver um modelo de previsão de séries temporais financeiras em baixa frequência, que supere o Dilema do Passeio Aleatório

com alto desempenho preditivo, baseado num estudo das características do fenômeno gerador de cada série. Utilizará uma combinação balanceada entre operadores lineares e operadores não-lineares, tendo a capacidade de determinar automaticamente o percentual balanceado de cada operador (linear e não-linear).

Especificamente, o estudo das séries temporais e a proposição de um modelo proposto possui os seguintes objetivos:

1. Estudar as características dos fenômenos gerados de séries temporais financeiras nas frequências diária, semanal e quinzenal;
2. Propor um modelo de previsão híbrido com operadores balanceados utilizando os conceitos da Morfologia Matemática para uma função crescente;
3. Desenvolver um modelo de previsão morfológico-linear que dê suporte ao relacionamento de curto e longo prazo;
4. propor um algoritmo de treinamento baseado em gradiente descendente com ajuste de fase temporal automático que supere o Dilema do Passeio Aleatório de uma série temporal financeira e o problema de encontrar mínimos locais;
5. Realizar de uma análise experimental entre o modelo proposto gerado e os sistemas estudados na literatura, a partir de gráficos e métricas consolidadas no estado da arte.

De forma a realizar o intento, foi definido um modelo híbrido visando a junção dos pontos positivos dos trabalhos estudados, focando num comportamento crescentes e na utilização da morfologia matemática como principal teoria de suporte ao modelo desenvolvido (ARAÚJO, 2012b; ARAÚJO; OLIVEIRA; MEIRA, 2015; ARAÚJO, 2016), através da teoria da decomposição de Matheron (MATHERON, 1975), diferenciando este trabalho de todos encontrados no Estado da Arte.

No estudo de previsão de séries temporais baseada em dados históricos, é feita uma análise de dados passados e presentes (janela de tempo) a fim de determinar pontos futuros. Tal janela é formada por um conjunto de pontos da série, chamados de retardos temporais. Juntamente com os dados pré-processados, essas janelas devem conter informações suficientes para a determinação dos acontecimentos futuros. A escolha dos retardos temporais utilizados é de fundamental importância para uma melhor caracterização da série. Desta forma, é necessário buscar a melhor janela de tempo para a previsão que está sendo desenvolvida.

Vislumbrando a melhoraria do desempenho analisado no estado da arte, é proposta a modelagem e implementação de um modelo híbrido com comportamento crescente *General Increasing Morphological Perceptron* (GIMP) desenvolvido para dar suporte ao relacionamento de curto (passado próximo) e longo (passado distante) prazo, utilizando a

morfologia matemática e a decomposição de Matheron (MATHERON, 1975) para séries temporais financeiras nas frequências diária, semanal, e quizenal. Espera-se que com a decomposição de Matheron os resultados obtidos sejam melhores que os modelos estudados. Os experimentos serão realizados em diferentes frequências pois modelos encontrados na literatura não há um estudo em que avalie a hipótese da mudança de frequência possa modificar o comportamento da série temporal financeira. No capítulo 5 será detalhado o modelo proposto e todos os conceitos teóricos intrínsecos. O algoritmo de ajuste de fase será baseado no modelo proposto por (FERREIRA, 2006).

Para o desenvolvimento de uma análise experimental, o modelo proposto será comparado com outros métodos de previsão de séries temporais financeiras relevantes da literatura, através de medidas utilizadas em trabalhos da área (FERREIRA, 2006; ARAÚJO, 2012b; ARAÚJO, 2016). Neste sentido, foram utilizadas as seguintes medidas para avaliar o desempenho de previsão: erro médio quadrático (*mean squared error*, MSE), erro médio percentual absoluto (*mean absolute percentage error*, MAPE), estatística u de theil (*u of theil statistics*, THEIL), previsão de mudança na direção (*prediction of change in direction*, POCID), variância relativa média (*average relative variance*, ARV), e função de avaliação (*evaluation function*, EF). O resultado desta comparação mostrará o desempenho do modelo proposto, frente aos modelos estudados no Estado da Arte. O estudo em frequências diferentes tem o propósito de identificar se o modelo GIMP terá o comportamento desejado nas baixas frequências analisadas.

## 1.4 ESTRUTURA DA TESE

Para orientar o leitor deste trabalho, abaixo será mostrado a estrutura organizacional dos capítulos:

O Capítulo 1 introduz conceitos básicos sobre previsão de séries temporais financeiras, mostrando a importância do uso da informação para quem trabalha com o mercado financeiro, bem como os problemas e desafios encontrados na área e quais são os objetivos e contribuições deste trabalho. O capítulo 2 apresenta o conceito de séries temporais de baixa frequência, evidenciando a relevância, suas aplicações, o problema de previsão, definindo as nomenclaturas. No capítulo 3 será apresentado e analisado o estado da arte. No capítulo 4, é feita uma análise minuciosa das séries temporais financeiras com os principais ativos da BM&FBOVESPA em F-frequência (baixa frequência), observando a existência ou não da componente subdominante não linear no ativo estudado.

O modelo híbrido proposto na tese, principal foco deste trabalho, será descrito no capítulo 5, detalhando todo o funcionamento do sistema, os fundamentos para solucionar o problema do Dilema do Passeio Aleatório e como é realizado o processo de treinamento do modelo para resolver o problema de previsão de séries temporais financeiras. Seguindo com o trabalho, no capítulo 6 será desenvolvida uma análise de desempenho e experimental do modelo híbrido, e uma análise comparativa entre o sistema e outros modelos estudados

---

na literatura que deu suporte a este trabalho. A partir desses dados, será confirmada ou não a hipótese de o modelo híbrido proposto ser mais eficiente que os outros estudados. Finalizando o trabalho, no capítulo 7 há o relato, de forma conclusiva, dos resultados obtidos frente aos objetivos propostos, ressaltando as contribuições aqui realizadas e as considerações finais, que poderão ser ponto de partida para futuros trabalhos.

## 2 SÉRIE TEMPORAL E O PROBLEMA DE PREVISÃO

No presente capítulo serão apresentadas as definições formais de uma série temporal e do problema de previsão, basilares deste trabalho, com o intuito de facilitar o entendimento do conteúdo subsequente.

### 2.1 SÉRIES TEMPORAIS

De acordo com (SINGH, 2016)(SINGH, 1992), uma série temporal é definida como uma coleção de eventos ou tarefas obtidas num determinado tempo. Por sua vez, (BOX et al., 2015) define série temporal como uma sequência de amostras de um dado fenômeno, espaçadas num período de tempo. Portanto, pode-se definir série temporal financeira como um conjunto de amostras de um determinado ativo financeiro gerado num intervalo de tempo determinado. Este conjunto de dados pode ser obtido através de observações de um evento, que pode ser contínuo ou discreto. Os pontos de uma amostra devem seguir certas regras:

- Devem ser gerados por sucessivas medições, feitas ao longo de um intervalo de tempo;
- O intervalo de tempo deve ser contínuo;
- A distância entre duas amostras, neste período, deve ter o mesmo intervalo de tempo;
- Cada unidade de tempo no intervalo tem no máximo uma observação (uma amostra).

Como exemplos de séries temporais comumente observadas, pode-se citar uma sequência de bens produzidos por uma indústria mensal, uma série semanal de acidentes de motocicleta, o crescimento diário de um fungo e valor do fechamento diário de uma ação na bolsa de valores. Todos os acontecimentos mencionados são analisados com base no tempo, e sua frequência é constante: diária, semanal, mensal, dentre outras.

Considerando uma série temporal discreta no tempo com intervalos  $T$  entre cada uma das amostras, e  $\Delta t$  sendo o tempo que o evento foi observado, pode-se calcular o índice do evento como sendo:

$$U = \frac{T}{\Delta t} \quad (2.1)$$

Sabendo qual é o evento de uma série temporal, matematicamente uma série temporal pode ser definida por:

$$X = \{x_t \in \mathbb{R} \mid t = 1, 2, 3, 4 \dots U\} \quad (2.2)$$

Onde  $t$  é um índice temporal e  $U$  é o evento da série temporal.

Dessa forma,  $X$  representa uma sequência de observações que obedece uma ordem temporal definida por um espaço de tempo  $t$ , definindo a granularidade das observações. O intervalo de amostragem, definido pelo tempo entre duas observações consecutivas, é comumente determinado pela disponibilidade dos dados, almejando a obtenção de uma quantidade mínima de intervalos possíveis, de forma a representar mais precisamente o fenômeno temporal em questão (ARAÚJO, 2012b).

### 2.1.1 Análise de Séries Temporais

Uma característica importante das séries temporais é que, tipicamente, observações adjacentes são dependentes. Essa dependência entre observações é de interesse prático e a análise de séries temporais tem como objetivo desenvolver técnicas para o estudo destas dependências. Contudo, ao se analisar uma série temporal pode-se identificar algumas características (ARAÚJO, 2012b) (FERREIRA, 2006):

- Constantes – processamentos que permanecem constantes sobre um certo nível todo o tempo, com variações de período a período devido a causas aleatórias;
- Ciclos – representados por movimentações cíclicas na observação do fenômeno temporal, sem uma definição clara de sua frequência;
- Sazonalidades – similar à componente cíclica, embora as flutuações possuam caráter sazonal, isto é, com frequência bem definida;
- Tendência – representada pelo incremento ou decremento gradual sistemático das observações do fenômeno temporal;
- Flutuações Aleatórias – representados pelas imprecisões ou erros na observação do fenômeno temporal (ruído inerente ao processo de observação).

O processo de análise de uma série temporal pode ser executado através de algumas ferramentas matemáticas, procurando identificar cada uma dessas características, a fim de obter uma boa estimativa da contribuição de cada uma relacionada ao todo.

O primeiro método que pode-se utilizar para a análise de uma série temporal é um **gráfico de linhas**. A partir de análise do gráfico, pode-se observar o comportamento de uma função e analisar as características de série temporal. Para identificar a dependência entre os dados de uma série temporal, tem-se o segundo método, a autocorrelação, que verifica se há uma relação entre os dados de uma série temporal em diferentes pontos no tempo. Com esta ferramenta busca-se encontrar padrões repetidos, como a presença de um sinal periódico obscurecido pelo ruído.

De acordo com (ARAÚJO, 2012b) (FERREIRA, 2006) (NETO, 2012), deve-se seguir alguns critérios no processo de análise de séries temporais:

- **Planejamento:** os dados coletados de uma série temporal devem ser planejados, tanto os intervalos de amostragem quanto o evento que será utilizado;
- **Estacionariedade:** deve-se analisar como uma série temporal se desenvolve aleatoriamente no tempo em torno de uma média, e se o desvio padrão é constante (estacionariedade). Na prática um processo estocástico e dito estacionário de segunda ordem quando:

$$E[y_t] = \mu \forall t, \quad (2.3)$$

$$E[(y_t - \mu)^2] = \gamma_0 = \sigma^2 \forall t, \quad (2.4)$$

$$E[(y_t - \mu)(y_{t-k} - \mu)] = \gamma_k \forall t, \quad (2.5)$$

Na prática, as séries temporais apresentam uma não estacionariedade, e a amostra deve ser cuidadosamente analisada;

- **Transformações:** é observado na literatura que a maioria dos processos estatísticos utilizados para previsão de séries temporais supõe que estas são estacionárias. No entanto, a prática mostra o contrário e, por este motivo, é necessário desenvolver algumas transformações nos dados, na tentativa de tornar a série estacionária;
- **Irregularidades:** séries temporais podem conter problemas no evento devido a problemas na coleta dos dados, seja por defeito do equipamento ou por erro de quem está coletando os dados. Caso as irregularidades sejam mínimas, poderão ser sanadas mediante o uso de técnicas de interpolação de dados, mas se forem significantes, o método de interpolação de dados não eliminará o problema, de modo que a série temporal não poderá ser utilizada;
- **Outliers:** são valores que podem integrar a série temporal e divergem da maioria dos dados do evento. Tais valores podem ser erros no processo de coleta, ou podem ser dados reais que podem vir a identificar uma nova característica ou padrão da série temporal, sendo muito importante verificar a natureza de tais acontecimentos e tratá-los de forma devida.

## 2.2 PROBLEMA DA PREVISÃO DE SÉRIES TEMPORAIS

A previsão de séries temporais é definida como a técnica utilizada para identificação de padrões em valores passados e presentes de uma série temporal, para gerar os valores futuros desta.

A previsão de uma série pode ser analisada no domínio espectral (domínio da frequência) ou no domínio temporal. Na literatura, a abordagem mais encontrada é no domínio temporal A, mediante uso de um conjunto de pontos pertencente à série, que são coletados em períodos com mesmo intervalo discreto de tempo. O conjunto de pontos do passado

ou presente é denominado de janela temporal, sendo cada ponto da janela temporal denominado retardo temporal (*lag*).

Um fator crucial para um bom desempenho na previsão é a escolha correta dos retardos temporais, pois os valores relevantes à série, quando relacionados, geram estruturas de relacionamento entre os dados históricos de tal maneira a tornar possível a caracterização das leis que governam o fenômeno gerador da série (BOX et al., 2015).

Ferreira afirma que o relacionamento dos dados constitui um espaço de fase  $d$ -dimensional, onde  $d$  é a dimensão mínima capaz de representar tais relacionamentos. Portanto, um espaço de fase  $d$ -dimensional pode ser construído de tal modo que contenha toda informação referente a uma série temporal. (TAKENS, 1981) enunciou um conjunto de quatro teoremas que provam que se  $d$  é suficientemente grande, tal espaço de fase é homeomórfico<sup>1</sup> ao espaço de fase original que gerou a série temporal.

Assim como em (MENEZES JR.; BARRETO, 2008), os teoremas de Takens serão a justificativa teórica que dará suporte a possibilidade de construção de um espaço de estados com a utilização dos corretos retardos temporais e que, de forma secundária, também garantirá que se o espaço for corretamente reconstruído, a dinâmica deste será topologicamente idêntica a dinâmica do espaço de estado original.

De acordo com Box-Jenkins (BOX et al., 2015), pode-se utilizar as funções de autocorrelação e autocorrelação parcial para identificar as dependências relevantes entre as observações, se for identificado apenas relacionamento lineares entre as observações do fenômeno temporal. Caso haja um relacionamento não linear entre as observações, a análise e definição automática dos retardos temporais se tornam complexas e devem ser utilizados outros métodos de previsão.

Alguns métodos são encontrados na literatura e se baseiam no cálculo de probabilidades condicionais, considerando que séries temporais sejam representadas por

$$X_t = f(X_{t-1}, X_{t-2}, X_{t-3}, \dots, X_{t-d}) + r_t \quad (2.6)$$

Onde  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  representa uma função que realiza o mapeamento entre os pontos do passado  $(X_{t-1}, X_{t-2}, X_{t-3}, \dots, X_{t-d})$  e do futuro  $(X_t)$ . O termo  $rt$  representa um termo de ruído. O ruído decresce com o incremento de  $d$  e tende a zero quando  $d$  excede a dimensão mínima necessária para a correta descrição do sistema gerador da série temporal (ARAÚJO, 2012b) (FERREIRA, 2006) (NETO, 2012).

### 2.3 DILEMA DO PASSEIO ALEATÓRIO

De acordo com (MALKIEL, 2007), em seu livro “A random walk down Wall Street”, um passeio aleatório é a teoria que os preços têm a mesma distribuição de probabilidade e são

<sup>1</sup> Dois espaços topológicos dizem-se homeomorfos se existir uma aplicação entre esses espaços que seja contínua, invertível e a sua inversa seja contínua.

independentes um do outro, de modo que o movimento passado não pode ser utilizado para prever o movimento futuro. Quando o termo é aplicado no mercado de ações, significa que as mudanças nos preços das ações de curto prazo não podem ser previstas.

Entretanto, estudos atuais (ARAÚJO, 2012b) (FERREIRA, 2006) (NETO, 2012) (DOUCOURE; AGBOSSOU; CARDENAS, 2016) mostram que pode-se modelar um passeio aleatório através da seguinte formalização:

$$P_t = P_{t-1} + a_t \quad (2.7)$$

em que  $P_t$  é o preço observado no instante  $t$  da série temporal,  $P_{t-1}$  é o preço no instante anterior a  $P_t$ , e  $a_t$  é um erro com distribuição gaussiana com média zero e desvio padrão dado por  $\sigma(N(0, \sigma))$ . O ruído da série temporal é definido em termos de uma componente aleatória, ou seja, um ruído branco (*white noise*). Desta forma, pode-se perceber que os valores futuros de uma série temporal não podem ser previstos, e o valor que mais se aproxima do valor futuro é o último valor analisado, ou seja,  $P_t$  é o melhor valor para  $P_{t-1}$ , uma vez que  $P_{t-1}$  representa a previsão com menor custo e desempenho aceitável. Araújo (ARAÚJO, 2016) afirma que este comportamento é característico em mercados financeiros e na economia mundial, sendo formalmente conhecido como Dilema do Passeio Aleatório (*Random Walk Dilemma*, RWD).

A fim de comprovar tal afirmação, vamos assumir a hipótese apresentada por Araújo (ARAÚJO, 2012b) em que um modelo de previsão será desenvolvido para construir o valor estimado de  $X_t$ , dado por  $X'_t$ . Então o valor esperado  $E[\cdot]$  da diferença entre  $X'_t$  e  $X_t$  deve tender a zero.

$$E[X'_t - X_t] \rightarrow 0. \quad (2.8)$$

De acordo com Araújo (ARAÚJO, 2012b), se um fenômeno gerador da série temporal é composto por componentes lineares dominantes e componentes não lineares subdominantes (definida por  $g(t)$ ) e assumindo que  $E[r_t] = 0$ , o valor esperado da diferença entre  $X'_t$  e  $X_t$  (assumindo que  $X_t = X_{t-1} + g(t) + rt$ , ou seja, que a série não é gerada num processo puramente aleatório) tem-se:

$$E[X'_t - (X_{t-1} + g(t) + rt)] \rightarrow 0, \quad (2.9)$$

$$E[X'_t] - E[X_{t-1}] - E[g(t)] - E[rt] \rightarrow 0, \quad (2.10)$$

$$E[X'_t] - E[X_{t-1}] - E[g(t)] \rightarrow 0, \quad (2.11)$$

$$E[X'_t] - E[X_{t-1}] + E[g(t)]. \quad (2.12)$$

Como  $E[g(t)] \rightarrow 0$ , devido à magnitude da componente não linear subdominante em relação a componente linear dominante, tem-se que:

$$E[X_{t-1}] + E[g(t)] \rightarrow E[X_{t-1}] \quad (2.13)$$

Nesse contexto:

$$E[X'_t] \rightarrow E[X_{t-1}] \quad (2.14)$$

Neste mesmo diapasão, pode-se perceber que a equação 2.14 garante que qualquer modelo de previsão arbitrário aplicado ao problema de previsão de séries temporais financeiras terá em suas previsões um característico atraso de um passo a frente em relação aos valores reais, tornando o dilema do passeio aleatório um problema complexo de ser resolvido.

### 2.3.1 O conceito do Ajuste de Fase Temporal

De acordo com Malkiel (MALKIEL, 2007), superar o dilema do passeio aleatório (*random walk*, RW) é uma tarefa bastante complexa, pois como o fenômeno gerador de séries temporais financeiras em F-frequência é bastante similar ao fenômeno gerador de passeio aleatório, qualquer modelo de previsão tenderá a ser guiado por esse comportamento (SITTE; SITTE, 2002) (MALKIEL, 2007).

Entretanto, Ferreira (FERREIRA; VASCONCELOS; ADEODATO, 2008) apresentou um procedimento, chamado de procedimento de ajuste de fase (*phase adjustment procedure*, PAP), utilizando testes estatísticos de hipótese para determinar se há ou não necessidade de ajustar a fase da série temporal. Caso a série tenha um comportamento *Random Walk*, um procedimento manual é feito para ajustar a defasagem temporal entre os dados reais da série temporal e sua respectiva previsão.

Tal técnica foi utilizada num método de previsão de série temporal denominado TAEF (*Time-lag Added Evolucionary Forecasting Method*), modelo de RNA (similar a *multilayer perceptron*) construído por meio da junção entre um algoritmo evolucionário (LEUNG et al., 2003b) e uma técnica baseada em gradiente descendente (WIDROW; HOFF et al., 1960) (BATTITI, 1992) (HAGAN; MENHAJ, 1994) (MØLLER, 1993), (RIEDMILLER; BRAUN, 1993). Essa técnica mostra de forma prática a resolução do problema do dilema do passeio aleatório no caso particular do problema de previsão de séries temporais financeiras, em frequência diária. Posteriormente Araújo (ARAÚJO, 2012b) desenvolveu um procedimento de ajuste de fase automático, que durante o processo de treinamento da rede neural é verificado se a série temporal tende para um modelo *Random Walk* (RW). Se a série temporal não se comportar como um RW, não é necessário o procedimento de ajuste de fase. Porém se o comportamento da série temporal for um RW direto, o algoritmo já executa o procedimento de ajuste de fase automático.

O processo de ajuste de fase automático desenvolvido por Araújo possui dois passos: primeiramente, um conjunto de retardos temporais  $x_1, x_2, \dots, x_n$  é apresentado ao modelo de previsão para geração de uma saída ( $y_1$ ); no segundo passo, o padrão de entradas é rearranjado, incluindo a saída gerada do primeiro passo, formando o conjunto de retardos temporais  $y_1, x_1, x_2, \dots, x_n - 1$ . A partir deste momento, o novo conjunto será a nova

entrada para o modelo de previsão, gerando uma nova saída ( $y_2$ ). Tal saída  $y_2$  representa a previsão do modelo com a fase ajustada, ou seja, sem o atraso de um passo a frente em relação aos valores reais da série temporal financeira.

Até o momento, esse procedimento só foi estudado em séries temporais financeiras em frequências diárias (FERREIRA, 2006) e alta-frequência (ARAÚJO; OLIVEIRA; MEIRA, 2015) (ARAÚJO, 2016). Entretanto, de acordo com (KIM; EOM et al., 2008), a modificação na frequência de uma série temporal implica na mudança das características do seu fenômeno gerador.

Dessa forma, surge a necessidade de realização de um estudo adicional sobre as características do fenômeno gerador de séries temporais financeiras, em outras frequências, visando o desenvolvimento de modelos de previsão, mediante a inclusão do conceito de ajuste de fase em seus processos construtivos, para solucionar o dilema.

Neste estudo, foi utilizado o procedimento automático de Araújo para ajustar a fase das séries temporais financeiras utilizadas. A funcionalidade desse procedimento foi pesquisada em outras duas frequências menores que as frequências diárias, quais sejam, semanal e quinzenal, pois de acordo com (COVEL, 2004), investidores de ações trabalham com mais segurança nestas frequências. O argumento para essa escolha, é o prazo de liquidação de uma ordem que pode durar mais de três dias.

## 2.4 RESUMO DO CAPÍTULO

O presente capítulo mostrou os principais conceitos que compõem o estudo de Séries Temporais, desde a sua definição, passando pelo processo de análise de uma série temporal, do problema do passeio aleatório, e finalizando com o conceito do Ajuste de Fase Temporal utilizado em Séries temporais Financeiras.

### 3 O ESTADO DA ARTE

Este capítulo detalha o estado da arte no âmbito das metodologias utilizadas para previsão de séries temporais financeiras. Foram pesquisadas metodologias usando modelos Estatísticos, modelos de Redes Neurais Artificiais, modelos de Regressor de Vetor de Suporte e Modelos Híbridos. Além dos modelos para previsão de séries temporais, foram analisadas algumas metodologias para otimização de modelos de previsão de séries temporais, mediante uso de Programação Genética. Desse modo, o objetivo dessa seção é discutir alguns trabalhos relacionados à matéria, com a finalidade de fornecer uma visão geral dos vários aspectos abordados atualmente por outros pesquisadores da área.

#### 3.1 INTRODUÇÃO

Previsão de séries temporais financeiras é um tema bastante estudado na literatura. São encontrados estudos com foco em Modelos Estatísticos, Redes Neurais Artificiais e Metodologias Híbridas que utilizam sistemas híbridos para definir a composição de um portfólio baseado em análise fundamental e indicadores técnicos, utilizando Algoritmos Evolucionários (SILVA; NEVES; HORTA, 2015). Outros trabalhos fazem uso apenas de indicadores técnicos e Algoritmos Genéticos para obter os valores ótimos das análises dos gráficos (MATSUI; SATO, 2010), havendo também aqueles que utilizam conjuntos apenas valores de uma série temporal para encontrar o valor futuro de uma ação (FERREIRA, 2006; ARAÚJO, 2012b; SALGADO et al., 2010; ARAÚJO; SUSSNER, 2010; ARAÚJO; FERREIRA, 2013; ARAÚJO; OLIVEIRA; MEIRA, 2015).

Porém, para esta pesquisa, os principais trabalhos pesquisados no Estado da Arte foram os que tinham como principal foco o estudo de previsão de séries temporais financeiras. O processo de previsão de séries temporais financeiras consiste em representar as características (Constantes, Ciclos, Sazonalidades, Tendências e Flutuações Aleatórias) da série temporal em um modelo que possa generalizar tais características temporais (FERREIRA, 2006). Para o modelo ser válido, este deve, em qualquer instante de tempo (passado, futuro e presente), representar valores de previsão muito próximos aos valores reais da série. Por exemplo: se o valor futuro de uma ação é cinco dólares, o modelo deve prever esse valor ou chegar muito próximo dele, baseado nas observações passadas.

Segundo Araújo (ARAÚJO, 2012b), apesar da maioria dos modelos de previsão considerarem apenas a análise das observações de uma dada série temporal, existem alguns modelos que consideram a inclusão de outras séries temporais para análise de uma série de interesse. A literatura divide os modelos de séries temporais em três classes (ARAÚJO, 2012b; FERREIRA, 2006):

- Univariados – mais utilizados nas soluções de problema de séries temporais. Esse

modelo considera apenas o histórico de uma série e a relação entre esses dados. Em alguns casos pode existir a adição de outras informações, como por exemplo um modelo bayesiano (ARAÚJO, 2012b; FERREIRA, 2006).

- Multivariados – utiliza mais de uma série para gerar a previsão. As séries temporais utilizadas não podem haver correlações entre si.
- Função de transferência – tipo de modelo multivariado que utiliza mais de uma série para a previsão, sendo que estas devem ter alguma correlação entre si.

Independente da classificação, existem diversos modelos de previsão para descrever uma série temporal. Neste trabalho, falaremos daqueles encontrados na literatura, relacionados com o tema em análise.

Estudos relacionados com esse tema mostram que podem ser utilizados Modelos Estatísticos para descrição do comportamento de uma série temporal.

## 3.2 MODELOS ESTATÍSTICOS

Os modelos estatísticos podem ser definidos como modelos lineares e modelos não lineares.

### 3.2.1 Modelos Lineares

Os modelos lineares têm sua análise de regressão definida por uma função linear simples, enquanto um modelo de regressão linear tem como objetivo principal expressar uma variável como função de outra variável. Todavia, quando se trata de única variável e esta se relacionando com ela mesma, é definido como um modelo de autorregressão linear.

Os modelos de autorregressão linear podem ser utilizados com o intuito de resolver o problema de previsão de séries temporais. Um modelo autorregressivo é uma representação de um tipo de processo aleatório utilizado para descrever processos em variações no tempo no sistema financeiro, nas indústrias e na natureza.

De acordo com (MILLS, 1991), pode-se definir um modelo autorregressivo AR(p) através da função:

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = a_t \quad (3.1)$$

ou

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = \Phi(B) X_t = a_t \quad (3.2)$$

Segundo (FERREIRA, 2006), a função autocorrelação do modelo AR pode ser estimada por:

$$r_j = \frac{C_j}{C_0}, \quad j = 1, \dots, p - 1 \quad (3.3)$$

em que

$$C_k = \frac{1}{N} \sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z}), \quad k = 0, 1, \dots, N-1 \quad (3.4)$$

Em que  $N$  é o tamanho da série temporal.

Além do modelo autorregressivo, também se utiliza o Modelo de Médias Móveis ( $MA$ ) para analisar a combinação de choques aleatórios ocorrida no passado. O modelo  $MA(q)$ , sendo  $q$  a ordem do modelo, pode ser definido por:

$$Z_t = \mu + r_t - \phi r_{t-1} - \dots - \phi r_{t-q}. \quad (3.5)$$

Utilizando o fato de que  $\tilde{Z}_t = Z_t - \mu$ , pode-se escrever,

$$\tilde{Z}_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) r_t = \Theta(B) a_t \quad (3.6)$$

Em que

$$\Theta(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (3.7)$$

é operador de médias móveis.

O Modelo de Médias Móveis tem como função autocorrelação (FERREIRA, 2006):

$$\gamma_j = E[\tilde{Z}_t \tilde{Z}_{t-j}] \quad (3.8)$$

$$= E \left( \left[ r_t - \sum_{k=1}^q \theta_k r_{t-k} \right] \left[ r_{t-j} - \sum_{l=1}^q \theta_l r_{t-j-l} \right] \right) \quad (3.9)$$

$$= E(r_t r_{t-j}) - \sum_{k=1}^q \phi_k E(r_{t-j} r_{t-k}) - \sum_{l=1}^q \phi_l E(r_t r_{t-j-l}) + \sum_{k=1}^q \sum_{l=1}^q \phi_k \phi_l E(r_{t-k} r_{t-j-l}) \quad (3.10)$$

Lembrando que,

$$\gamma_r(j) = E(r_t r_{t-j}) = \begin{cases} \sigma_r^2, & j = 0 \\ 0, & j \neq 0 \end{cases} \quad (3.11)$$

e

$$\gamma_0 = Var(Z_t) = \sigma_Z^2 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_a^2 \quad (3.12)$$

Pode-se organizar os índices da equação 3.10 reescrevendo-a em termos de  $\gamma_j$ , resultando em:

$$\gamma_j = \gamma_r(j) - \sum_{k=1}^q \phi_k \gamma_r(k-j) - \sum_{l=1}^q \phi_l \gamma_r(j+l) + \sum_{k=1}^q \sum_{l=1}^q \phi_k \phi_l \gamma_r(j+l-k) \quad (3.13)$$

Levando a:

$$\gamma_j = \begin{cases} (-\theta_j + \sum_{l=1}^q \theta_l \theta_{l+j}) \sigma_r^2 = (-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \cdots + \theta_q \theta_{q-j}) \sigma_r^2, & j = 1, \dots, q \\ 0, & j > q \end{cases} \quad (3.14)$$

Ou simplesmente:

$$\rho_j = \begin{cases} \frac{-\theta_j + \theta_1 \theta_{j+1} + \theta_2 \theta_{j+2} + \cdots + \theta_q \theta_{q-j}}{1 + \theta_1^2 + \cdots + \theta_q^2}, & j = 1, \dots, q \\ 0, & j > q \end{cases} \quad (3.15)$$

Observe que a Função de Autocorrelação - Autocorrelation Function (FAC) de um processo  $MA(q)$  é igual a zero retardos temporais de ordem superior a  $q$ .

Um modelo misto foi desenvolvido a partir dos modelos Médias Móveis - Moving Average (MA) e Modelo Autorregressivo - Autoregressive (AR), denominado de Modelo Modelo Autorregressivo de Médias Móveis - Autoregressive Moving Average (ARMA). O objetivo principal do ARMA é gerar um modelo de previsão compacto com o menor número de parâmetro possível. O *ARMA* só conta com dois parâmetros: o  $p$  representa o número de termos autorregressivo e o  $q$  representa o número de termos da média móvel (ordem).

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \cdots + \phi_p \tilde{Z}_{t-p} + r_t - \theta_1 r_{t-1} - \cdots - \theta_q r_{t-q}. \quad (3.16)$$

Sendo  $\Phi(B)$  e  $\Theta(B)$  os operadores auto-regressivos e de médias móveis, introduzidos anteriormente, pode-se escrever a equação 3.16 em uma forma compacta:

$$\Phi(B) \tilde{Z}_t = \Theta(B) r_t \quad (3.17)$$

Multiplicando-se a equação 3.16 por  $\tilde{Z}_{t-j}$  e tomando-se o valor esperado tem-se:

$$\gamma_j = E(\tilde{Z}_t \tilde{Z}_{t-j}) = E\{(\phi_1 \tilde{Z}_{t-1} + \cdots + \phi_p \tilde{Z}_{t-p} + r_t - \theta_1 r_{t-1} - \cdots - \theta_q r_{t-q}) \tilde{Z}_{t-j}\} \quad (3.18)$$

ou

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \cdots + \phi_p \gamma_{j-p} - \theta_1 \gamma_{Z_r}(j-1) - \cdots - \theta_q \gamma_{Z_r}(j-q) \quad (3.19)$$

Em que  $\gamma_{j-r}(j)$  é a covariância cruzada entre  $Z_t$  e  $r_t$ , definida por:

$$\gamma_{j-r}(j) = E(r_t \tilde{Z}_{t-j}) \quad (3.20)$$

Como  $\tilde{Z}_{t-j}$  só depende de choques  $r_t$  ocorridos até o instante  $t - j$ :

$$\gamma_{j-r}(j) \longrightarrow \begin{cases} = 0, & j > 0 \\ \neq 0, & j \leq 0 \end{cases} \quad (3.21)$$

de modo que

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \cdots + \phi_p \gamma_{j-p}, \quad j > q \quad (3.22)$$

Desta forma, obtem-se

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p}, \quad j > q \quad (3.23)$$

Deduzindo-se que as auto-correlações de retardos temporais  $1, 2, \dots, q$  serão afetados pelos parâmetros de médias móveis. Por sua vez, para  $j > q$ , as mesmas comportam-se como nos modelos auto-regressivos. Pode-se verificar que se  $q < p$ , a função de auto-correlação consiste de uma mistura de exponenciais e/ou senóides amortecidas; entretanto, se  $q \geq p$ , os primeiros  $q-p+1$  valores  $\rho_0, \rho_1, \dots, \rho_{q-p}$ , não seguirão este padrão (FERREIRA, 2006).

De acordo com (ARAÚJO, 2012b) e (FERREIRA, 2006), a grande maioria dos modelos utilizados para previsão de séries temporais é linear, e o mais utilizado na literatura é o método de Box e Jenkins (Modelo Autorregressivo Integrado de Médias Móveis, Modelo Autorregressivo Integrado de Médias Móveis - Autoregressive Integrated Moving Average (ARIMA)) (BOX et al., 2015), pois este modelo é o mais apropriado para a descrição de séries estacionárias (também chamada de série convergente), isto é, séries que flutuam em torno de uma mesma média ao longo do tempo.

O modelo ARIMA baseia-se na inserção de um filtro passa-alta para permitir a passagem das frequências altas com facilidade e reduzir (atenuar) a amplitude das frequências abaixo da frequência de corte, sendo estes aplicados no Modelo *ARMA*. O *ARIMA*( $p, d, q$ ) (ARAÚJO, 2012b) é um estudo algébrico que mostra que a variável de uma série temporal ( $X$ ) é relacionada com seus próprios valores passados ( $X_{t-1}, X_{t-2}, \dots, X_{t-n-1}, X_{t-n}$ ) e os valores passados dos termos de ruído ( $r_{t-1}, r_{t-2}, \dots, r_{t-n-1}, r_{t-n}$ ), diferenciando  $d$  vezes, uma vez que algumas séries temporais quando não estacionárias se tornam estacionárias quando aplicado o processo de diferenciação.

Pode-se definir o Modelo *ARIMA*( $p, d, q$ ) da seguinte forma (BOX et al., 2015):

$$w_t = \varphi_1 w_{t-1} + \cdots + \varphi_p w_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \quad (3.24)$$

Em que  $w_t = \nabla^d z_t$ .

Contudo, outras áreas também estão pesquisando o tema com a finalidade de diminuir a complexidade das técnicas utilizadas e aumentar o desempenho dos algoritmos automáticos. A principal área de pesquisa relacionada a este estudo é a Inteligência Artificial

(Inteligência Artificial - Artificial Intelligence (AI)), que é vista como uma alternativa para suprir as dificuldades técnicas encontradas na utilização de modelos estatísticos. Dentro da Inteligência Artificial, as Redes Neurais Artificiais são as estruturas mais utilizadas para a predição de séries temporais.

### 3.3 MODELOS DE REDES NEURAS ARTIFICIAIS

O primeiro modelo artificial de um neurônio biológico foi desenvolvido por (MCCULLOCH; PITTS, 1943). Neste trabalho, o autor apresenta uma discussão de redes lógicas de neurônios artificiais, além de ideias sobre a máquina de estados finitos, elementos de decisão de limiar lineares e representações lógicas de várias formas de comportamento e memória (BRAGA; CARVALHO; LUDERMIR, 2000).

A década de 1980 marcou o ressurgimento da área de Redes Neurais Artificiais (Redes Neurais Artificiais (RNA)s), também conhecidas como conexionismo ou sistemas de processamento paralelo e distribuídos. Essa forma de computação não algorítmica é caracterizada por sistemas que, em algum nível, relembram a estrutura do cérebro humano. Por não ser baseada em regras, a computação neural se constitui uma alternativa à computação algorítmica convencional (BRAGA; CARVALHO; LUDERMIR, 2000).

A primeira aplicação que utilizou RNA como estimador, em vez de modelos matemáticos como ARIMA, foi o de (WHITE, 1988). Apesar de não desenvolver resultados experimentais comparativos com os modelos matemáticos, o artigo teve uma forte aceitação na comunidade científica, pois seus experimentos mostraram a possibilidade do uso de RNA para predição dos valores de uma série temporal financeira.

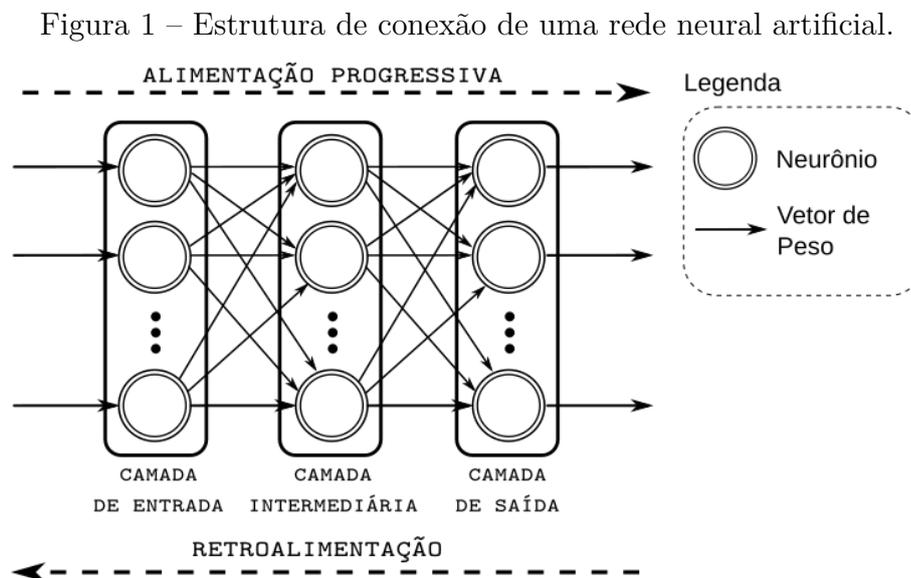
Segundo (SOARES, 2008), em 1991 a comunidade científica conheceu dois trabalhos notoriamente reconhecidos até os dias atuais: o de (JANG, 1991) e o de (MATSUBA, 1991). O primeiro conseguiu prever, com pequenos erros, algumas séries econômicas utilizando duas RNAs do tipo MLP (*MultiLayer Perceptron*) com algoritmo *backpropagation*. Uma rede foi treinada para prever a tendência de preço em longo prazo e outra rede foi treinada para prever a tendência de preço do mesmo ativo no curto prazo. Os valores utilizados foram valores de picos e valores de baixa, conhecidos no mercado financeiro como resistência e suporte, respectivamente, e os valores de fechamento diário e o volume operado cada dia. O segundo trabalho empregou uma rede neural *feedforward* com os  $n$  últimos valores de índice como entrada na rede, e os  $p$  preços futuros como saída da rede neural. O pesquisador também utilizou duas RNAs, porém, uma para aprender as características globais (janelas longas) e outra para treinar com características locais (janelas curtas).

(BRAGA; CARVALHO; LUDERMIR, 2000) afirmam que a solução de problemas por meio de RNAs é bastante atrativa, já que a forma como estes são representados internamente pela rede e o paralelismo natural inerente à arquitetura das RNAs criam a possibilidade de um desempenho superior ao dos modelos convencionais. Em RNAs, o procedimento usual na solução de problemas passa inicialmente por um processo chamado aprendizagem.

Nesse processo, busca-se utilizar um conjunto de exemplos para informar a rede e esta extrair as características necessárias para representar a informação fornecida.

Desta forma, pode-se perceber que uma rede neural necessita de dados organizados inicialmente, para que possa “aprender” de forma correta todas as características de uma rede. Após ter “aprendido” todas as características do conjunto de informações, deve-se utilizar outras informações do evento para validar a rede e outro grupo para desenvolver os testes na rede neural artificial. As redes são capazes de extrair características não apresentadas de forma explícita através dos dados de aprendizagem, assim como atuar como mapeadores universais de funções multivariáveis, com custo computacional que cresce apenas linearmente conforme o número de variáveis.

Portanto, as RNAs são estruturas adequadas para problemas nos quais as soluções requerem conhecimento que não pode ser adquirido de forma simples, mas que está contido nos dados de aprendizagem da rede. A figura 1



Fonte: O autor.

Já visto neste trabalho que o problema de previsão de séries temporais tem o objetivo de gerar valores futuros, com base em informações do passado e do presente. Partindo do pressuposto que as RNAs desenvolvem um processo de aprendizagem para retirar as características de um conjunto de valores passados, e assim gerar valores futuros baseados nas características dos conjuntos, pode-se perceber que as RNAs podem realmente ser utilizadas para o estudo de previsão de séries temporais, inclusive as financeiras.

Outro fator importante é que as RNAs são mapeadores universais de funções multivariáveis. Afirmando que as séries temporais podem ser representadas por funções com uma ou mais de uma variável, podemos perceber que as RNAs também podem ser utilizadas para expressá-las.

Outra vantagem do uso de RNAs para previsão de séries temporais é que, por esta

ter seu crescimento linear com relação ao número de variáveis, pode ser utilizada para auxiliar os modelos estatísticos não lineares a solucionar problemas complexos, visto que, de acordo com (FERREIRA, 2006) e (NETO, 2012), o desempenho de previsão de tais modelos não lineares está bem abaixo do esperado e o problema permanece aberto nos dias atuais.

(FERREIRA, 2006) afirma que uma desvantagem do modelo unicamente dependente de dados vem do fato que regras a respeito do sistema em estudo são, de forma geral, não evidentes e as observações geralmente marcadas por ruídos, podendo o modelo fracassar devido a este fato.

Nas seções seguintes serão analisados alguns tipos de redes neurais artificiais que foram encontrados na literatura, sendo utilizadas para o tratamento do problema de previsão de séries temporais financeiras.

### 3.3.1 Processo de treinamento de uma Rede Neural Artificial utilizada para Previsão de Séries Temporais Financeiras

Para o correto estudo da utilização de uma rede neural, é necessário dividir o conjunto de dados históricos de uma série temporal em três partes (FERREIRA, 2006) (ARAÚJO, 2012b) (PRECHELT et al., 1994):

- Aprendizagem – a maior parte dos dados históricos da série temporal será utilizada para a aprendizagem da rede neural, pois assim ela poderá definir uma maior quantidade de características daquela série. Geralmente é utilizado de 70% a 90% dos valores para esta etapa, quando não há a etapa de validação, e 50% quando se faz a validação;
- Validação – o processo de validação nem sempre é visto na literatura, mas é de extrema importância para determinar o melhor momento para o término do treinamento da rede neural. Nesta etapa, a literatura mostra que os pesquisadores utilizam 25% do conjunto de dados históricos da série temporal;
- Teste – esta etapa tem como finalidade testar o desempenho da rede neural e verificar se a rede conseguiu modelar o sistema de previsão aceitável. Utilizam-se os outros 25% dos dados para fazer esta análise, de acordo com estudos na literatura.

### 3.3.2 Rede de *Perceptron* Multicamadas

As não linearidades são incorporadas a modelos de redes neurais através das funções de ativação (não lineares) de cada neurônio da rede e da composição da sua estrutura em camadas sucessivas. Assim, a resposta da camada mais externa da rede corresponde à composição das respostas dos neurônios das camadas anteriores (BRAGA; CARVALHO; LUDERMIR, 2000). Portanto, uma rede composta de mais de uma camada composta por

neurônios com funções de ativação sigmoidais nas camadas intermediárias é denominada de Perceptron de Múltiplas Camadas (MLPs- *Multilayer Perceptron*). Em 1969 foi publicado o livro intitulado de Perceptron (MINSKY; PAPERT, 1969), propondo pela primeira vez uma rede Perceptron com Múltiplas Camadas, na tentativa de superar limitações de uma rede Perceptron simples.

Diferente de redes de única camada, nas quais se pode obter o erro através da diferença entre o valor de saída corrente e o valor desejado de previsão, nas redes MLPs devemos atentar para as valores intermediários, pois elas também produzem saídas com erros e estes devem ser estimados e calculados de alguma forma.

A fim de solucionar esse problema, foi desenvolvida em 1980 a descrição do algoritmo de retropropagação de erros, ou *back-propagation* (WILLIAMS; HINTON, 1986). O algoritmo de *back-propagation* utiliza gradientes descendentes para estimar o erro das camadas intermediárias e, assim, estimar o efeito que estas camadas causam no erro da camada de saída. Desse modo, o erro calculado na camada de saída serve como entrada para o ajuste dos pesos de cada neurônio das camadas intermediárias de forma proporcional aos valores das conexões entre as camadas.

As redes MPL são compostas por uma camada de entrada (*input layer*), camadas intermediárias ou escondidas (*hidden layers*) e a camada de saída (*output layer*). A camada de entrada é responsável por receber as informações e distribuí-las sem modificá-las. As camadas intermediárias são compostas por neurônios conectados às camadas adjacentes. As camadas recebem os valores da camada de entrada e ajustam os pesos para o processamento, gerando valores de saída que serão entradas para a camada de saída. Esta tem a função de receber os valores da camada anterior e apresentar a resposta final *target*. Segundo Ferreira, caso todos os neurônios de uma determinada camada se conectem a todos os neurônio da camada subsequente, a rede MPL é chamada de rede MLP totalmente conectada.

Redes MLP apresentam um poder computacional expressivamente maior quando comparadas a redes de uma única camada. Segundo (CYBENKO, 1989), em problemas de classificação, por exemplo, MLPs podem lidar com conjuntos de dados que não sejam linearmente separáveis. Teoricamente, redes com duas camadas intermediárias podem implementar qualquer função, seja ela linearmente separável ou não. A qualidade da previsão obtida dependerá da complexidade da rede, ou seja, do número de neurônios que serão utilizados nas camadas intermediárias.

Baseado em (FERREIRA, 2006), será definida a rede MLP utilizada para problemas de previsão. Em um problema de previsão, as entradas de uma RNA são usualmente as variáveis de predição. Portanto, a função estimada pode ser escrita por:

$$y = f(x_1, x_2, x_3, \dots, x_{p-1}, x_p) \quad (3.25)$$

Em que  $x_1, x_2, x_3, \dots, x_{p-1}, x_p$ , são  $p$  variáveis independentes e  $y$  é uma variável dependente

dessas.

Assim, pode-se afirmar que a RNA funciona de forma equivalente a um modelo de regressão não linear.

Caso o problema seja encarado como uma extrapolação, (FERREIRA, 2006) afirma que as entradas da rede são tipicamente as observações passadas da série, e a saída é o respectivo valor futuro. Então, a rede dependerá de uma função de mapeamento, em que o valor dependente será o  $y_{t-1}$  e dependerá dos valores do presente  $y_t$  e do passado, como mostra a equação 3.26:

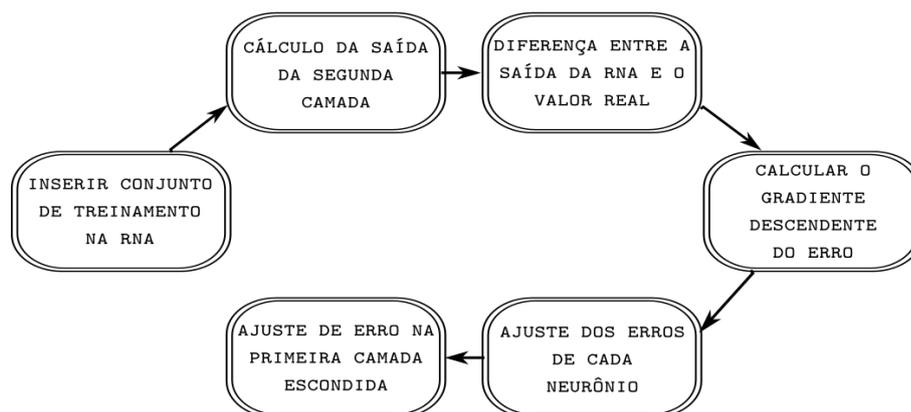
$$y_{t+1} = f(y_t, y_{t-1}, \dots, y_{t-p}) \quad (3.26)$$

Portanto, podemos definir que a RNA pode ser equivalente a um modelo autorregressivo não linear.

Pode-se inserir numa RNA as variáveis necessárias para o processo de previsão e os retardos temporais que são necessários para a análise de uma série temporal. Assim, a RNA poderá definir características na aprendizagem. Para a aprendizagem de uma RNA ser válida, é necessário que ela seja supervisionada. Para cada padrão de entrada da rede, deverá existir um padrão de saída, forçando a rede a ajustar seus pesos (aprendizado) de forma a realizar o mapeamento entre o padrão de entrada e o padrão de saída, através da minimização do erro entre a saída da rede e o valor desejado.

O treinamento de uma RNA ocorre de acordo com o fluxo mostrado na figura 2:

Figura 2 – Fluxo de treinamento de uma rede neural artificial.



**Fonte:** O autor.

O fluxo de treinamento de uma rede MLP inicia-se com a inserção dos conjuntos de treinamento da RNA. O vetor de entrada  $x$  é apresentado às entradas da rede e as saídas da primeira camada são calculadas. As saídas das camadas escondidas são calculadas a partir das saídas da primeira camada. Esse fluxo de cálculo continuará até a camada de saída. Quando as saídas da última camada escondida são recebidas pela camada de saída,

os neurônios dessa camada comparam às saídas desejadas  $y$ , para o vetor de entrada  $x$ , e o erro correspondente à diferença entre o valor real e o valor previsto é calculado.

A fase de *forward* tem como objetivo obter o erro de saída. A etapa de *backward* segue o seguinte fluxo: o erro calculado na camada de saída é utilizado para ajustar diretamente os pesos de cada neurônio, utilizando para isso o valor do gradiente descendente do erro; os erros são propagados para a camada anterior, utilizando-se os pesos das conexões entre camadas, que serão multiplicados pelos erros correspondentes, definindo, assim, um valor estimado para o erro de cada neurônio da camada escondida; os erros são utilizados para ajustar os pesos de cada neurônio pelo gradiente descendente; repete-se o processo até chegar na primeira camada escondida da RNA MLP. O método do gradiente descendente busca otimizar o erro baseado no mínimo local de uma função, utilizando a informação local do gradiente e fazendo sempre com que a busca vá no sentido negativo do gradiente.

A rede MLP foi a rede utilizada por (FERREIRA, 2006) para o desenvolvimento de uma metodologia híbrida inteligente para previsão de séries temporais. Todavia, estudos recentes mostram que redes MLPs não são os melhores tipos de RNAs para se trabalhar com previsões de séries temporais financeiras, pois estas séries têm uma particularidade importante: memórias longas. Portanto, as redes MLPs não são bons aproximadores quando se trata de séries com memórias longas (FERREIRA, 2006) (MENEZES JR.; BARRETO, 2008).

### 3.4 MODELO AUTORREGRESSIVO NÃO-LINEAR COM ENTRADA EXÓGENA (*NONLINEAR AUTOREGRESSIVE EXOGENOUS MODEL*) – NARX

Conforme vimos na literatura, a equação do modelo preditor é dada por:

$$\hat{y}_t = f(\phi_t, \omega) \quad (3.27)$$

na qual  $f$  é uma função não linear com entrada exógena<sup>1</sup> usada pela rede neural.  $\phi_t$  é o vetor de regressores, que contém entradas passadas  $\phi_t$  e saídas passadas observadas  $y_t$  do modelo no tempo  $t$ ,

$$\phi_t = (u_{(t-D_u)}, \dots, u_{(t-1)}, u_t, y_{(t-D_y)}, \dots, y_{(t-1)}) \quad (3.28)$$

$D_u$  e  $D_y$  são os atrasos do conjunto de entrada e do conjunto de saída.  $\omega$  é o vetor dos pesos,

$$\omega = [a_1, \dots, a_m, b_0, \dots, b_m] \quad (3.29)$$

O trabalho de (OLIVEIRA, 2001) utiliza a rede NARX (*Nonlinear Autoregressive Exogenous Model*) na tentativa de prever os valores de alguns ativos da Bolsa de Valores de São Paulo, porém os resultados não se mostram satisfatórios, permanecendo com o problema

<sup>1</sup> informações externas que influenciam na série a ser prevista

de RWD. Para melhorar o desempenho da NARX na previsão de séries temporais, tal rede terá suas entradas exógenas direcionadas e seguindo o teorema da imersão de Floris Takens, matemático que proveu importantes contribuições científicas nas áreas de sistemas dinâmicos e Teoria do Caos. Na estatística, a incorporação de atrasos dá as condições em que um sistema caótico dinâmico pode ser reconstruído a partir de uma sequência de observações do estado do sistema. Esta reconstrução preserva as propriedades do sistema dinâmico que não sofre alterações sob mudanças coordenadas, mas não mantém a forma geométrica de estrutura no espaço de fase.

O Teorema de Takens também é conhecido como método de atraso e pode ser aplicado a qualquer série temporal univariada  $X_t$ ,  $t = 1, \dots, T$  para construir um vetor  $V_n$ ,  $n = 1, \dots, N$ .

O processo de construção inicia com a escolha de alguns valores consecutivos de amplitudes de  $X_t$ , que representarão coordenadas dos vetores. Dada uma série suficientemente longa, todo o espaço de estado será bem amostrado e podemos ter informações sobre a construção do espaço a partir de informações do sistema dinâmico.

Para a construção do espaço de estados, assume-se que  $X_t$  é uma projeção dinâmica que opera num espaço de alta-dimensão. O valor de  $\tau$  e  $d_e$  são utilizados para transformar a série temporal em vetores de Takens, também chamados de espaço de fase ou vetores de atrasos.

O vetor  $V_n$  é descrito por:

$$V_n = [V_1, V_2, \dots, V_N] = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(d_E-1)\tau} \\ x_2 & x_{2+\tau} & \cdots & x_{2+(d_E-1)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_N & x_{N+\tau} & \cdots & x_{N+(d_E-1)\tau} \end{bmatrix} \quad (3.30)$$

A dimensão incorporada,  $d_e$ , representa a dimensão do espaço de estados utilizados para gerar a dinâmica do sistema subjacente separados por  $\tau$  intervalos entre as sucessivas amostras observadas, utilizadas para construir o vetor de incorporação com dimensão  $d_e$ .

Simplificando o Teorema de Takens, podemos reconstruir um espaço de fase como valores passados  $d_e$ , espaçados por uma constante  $\tau$ , como por exemplo:  $\tau : y_{t-\tau}, y_{t-2\tau}$ . Em que  $\tau$  indica o espaçamento entre os valores passados que serão utilizados na função que fará o mapeamento. A utilização deste teorema provê à rede NARX a absorção de todas as características de memória de longo prazo, ou seja, o Teorema de Takens provê a organização das entradas exógenas (entradas com maior atraso) da Rede Neural e, assim, pode-se trabalhar com valores de curto e longo prazo. O valor de  $\tau$  ótimo é encontrado de forma empírica, tendo que analisar um conjunto de experimentos para identificar o melhor resultado para cada série temporal.

O trabalho intitulado “*Long-term time series prediction with the NARX network: An empirical evaluation*” de (MENEZES JR.; BARRETO, 2008) define uma metodologia para

analisar os ganhos da rede NARX na predição de séries temporais de longo prazo. O autor desenvolve duas propostas para redes NARX-Takens: paralela, que utiliza os valores do passado para previsão de mais de um valor futuro, e a serial-paralela, que utiliza os valores passados para prever apenas um valor futuro. Na comparação entre a rede neural NARX-Takens proposta, a rede de Elman e a Time Delay Neural Network (TDNN), a rede proposta se mostrou superior às rede commparadas.

Objetivando a previsão do próximo valor futuro, o presente trabalho analisará o comportamento de uma rede similar a Rede NARX-Takens serial paralela, de Menezes e Barreto, para a previsão de séries temporais financeiras. A saída regressiva da rede NARX-Takens serial paralela é dada pela função:

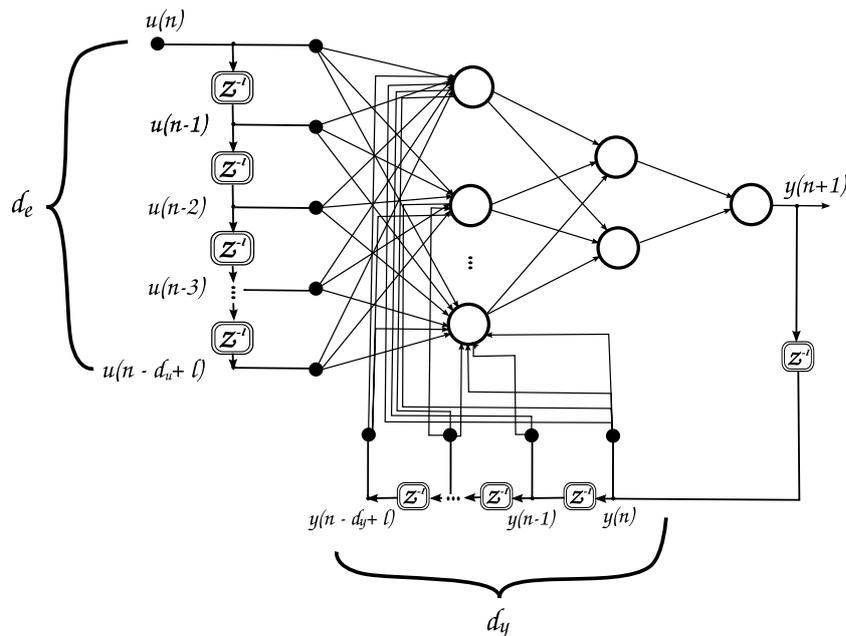
$$\hat{y}(n + 1) = \hat{f}[\mathbf{y}_{sp}(n), \mathbf{u}(n)] \tag{3.31}$$

$$= \hat{f}[y(n), \dots, y(n - d_y + 1); u(n), u(n - 1), \dots, u(n - d_u + 1)] \tag{3.32}$$

Em que ( $\hat{\cdot}$ ) é utilizado para denotar o valor estimado pela função,  $d_y$  representa o atraso das saídas sequenciais (componente linear) e  $d_u$  os atrasos para as entradas da função.

A utilização da rede NARX para previsão de sistemas não lineares pode ser observada literatura, desde a previsão de trocadores de calor e desperdício de água tratada (LIN et al., 1997), até equações caóticas como a série temporal de (DIACONESCU, 2008).

Figura 3 – Arquitetura da rede NARXTAKENS.



Fonte: O autor.

A figura 3 mostra a rede neural utilizando tanto as entradas exógenas ( $u(n)$ ) quanto as entradas sequenciais ( $y(n)$ ).

A função também segue o método do gradiente descendente, buscando o mínimo local da função a partir do uso de informação local do gradiente que fará com que a busca vá sempre em direção negativa do gradiente. Durante todo esse estudo, a menção do termo "rede NARX" refere-se à rede NARX-Takens serial paralela.

### 3.5 REGRESSÃO LOGÍSTICA

A técnica de Regressão Logística foi modelada como uma técnica estatística, com objetivo de produzir, a partir de observações, um modelo que permita a predição de valores tomados por uma variável categórica. Este modelo é utilizado em diversas áreas: na medicina, detectando doenças e analisando fatores que caracterizam um grupo de enfermos em relação a indivíduos sãos (LIU et al., 2012) (WAGNER et al., 2002); na econometria (ENDERS, 2004), para explicar variáveis discretas; e na economia, em instituições financeiras para vários tipos de análises, detectando os grupos de risco para subscrição de crédito e previsão de valores futuros de um ativo (WANG et al., 2012).

A regressão logística pode ser binomial, ordinal ou multinomial (JR; LEMESHOW; STURDIVANT, 2013). O modelo binomial trata situações em que os resultados observados para uma variável dependente pode ter apenas dois tipos de saídas, 0 ou 1 (Exemplo: morto VS. Vivo, doente VS. Sadio, fracasso VS. Sucesso). O modelo multinomial trata de situações com mais de duas variáveis. e a regressão logística ordinal lida com variáveis dependentes, que são ordenadas.

A forma de análise de dados da Regressão Logística - Logistic Regression (LOGREG) segue dados distribuídos binomialmente da forma:

$$Y_i \sim B(p_i, n_i), \quad \text{para } i = 1, \dots, m, \quad (3.33)$$

Em que os números de ensaios de Bernoulli  $n_j$  são conhecidos e as probabilidades de êxito  $p_i$  são desconhecidas (AGRESTI; KATERI, 2011).

Uma função logística ou uma curva logística é estruturada no formato de curva sigmoide padrão com equação (JR; LEMESHOW; STURDIVANT, 2013):

$$f(x) = \frac{L}{1 + e^{-k(x-x_0)}} \quad (3.34)$$

Em que  $x_0$  o valor de  $x$  no ponto médio da curva sigmoide,  $L$  o valor máximo da curva e  $k$  a declividade da curva. Para valores de  $x \in \mathbb{R}$  tal que  $-\infty < x < +\infty$ , a curva sigmoide à direita é obtida com o gráfico de  $f$  se aproximando a  $L$  conforme o  $x$  se aproxima de  $+\infty$ , e se aproximando a zero conforme o  $x$  se aproxima de  $-\infty$  (JR; LEMESHOW; STURDIVANT, 2013).

Enquanto método de predição, a regressão logística (LOGREG) é comparável às técnicas supervisionadas propostas em aprendizagem automática (Árvores de Decisão, Redes

Neurais, Máquinas de Vetor de Suporte para Regressão, entre outras), ou até outros métodos estatísticos, como ARIMA (AGRESTI; KATERI, 2011) (GREENE, 2003).

O cálculo da derivada função logística padrão ( $k = 1$ ,  $x_0 = 0$  e  $L = 1$ ) é dado por:

$$\frac{d}{dx}f(x) = f(x) \cdot (1 - f(x)) \quad (3.35)$$

A função logística é a solução de equações diferenciais não lineares de primeira ordem. Funções logísticas são frequentemente usadas em redes neurais para introduzir não linearidade no modelo e/ou para captar sinais dentro de um espaço temporal específico. Um elemento neural popular computa uma combinação linear de seus sinais e aplica uma função logística relacionada com o resultado. Esse modelo pode ser visto com uma variante suavizada do clássico neurônio limiar. Uma escolha comum para a função de ativação, usada para recortar grandes magnitudes, a fim de manter a resposta da rede neural delimitada, é:

$$g(h) = \frac{1}{1 + e^{-2\beta h}} \quad (3.36)$$

Em que  $\beta$  é o parâmetro de inclinação da suavização da função sigmoide. Devido à sua natureza sigmoide (antissimétrica), busca-se a utilização do algoritmo *backpropagation* para levar à convergência mais rapidamente.

### 3.6 REDE DE FUNÇÃO DE BASE RADIAL

Uma Rede Neural de Função de Base Radial, ou rede RBF (BRAGA; CARVALHO; LUDERMIR, 2000), é um modelo de RNA que pode ser definido dinamicamente durante o treinamento da rede, retirando ou inserindo neurônios das camadas intermediárias, e retirando ou inserindo conexões à rede já existente.

(BRAGA; CARVALHO; LUDERMIR, 2000) define RBF como uma rede neural multicamadas em que a ativação de um neurônio pode ser função da distância entre seus vetores de entrada e de peso, e que utiliza funções base radiais na camada intermediária. Vimos anteriormente que as funções utilizadas pela rede neural MLP são de bases sigmoidais. Uma rede MLP também constrói aproximadores globais com maior capacidade de generalização para regiões onde não há dados de treinamento. Contudo, isso permite que neurônios completamente diferentes daqueles utilizados no treinamento sejam classificados em uma das classes existentes, fato este conhecido como problema dos “falsos padrões” (BRAGA; CARVALHO; LUDERMIR, 2000).

Funções de base radial são funções não lineares que podem ser utilizadas em qualquer tipo de modelo de regressão não linear, bem como função de ativação de qualquer rede multicamada. Trata-se de uma função sobre os números reais cujos valores dependem apenas da distância a partir da origem. Funções radiais representam uma classe especial de funções, cujo valor diminui ou aumenta em relação à distância de um ponto central

(BRAGA; CARVALHO; LUDERMIR, 2000). Existem diferentes funções de bases radiais que podem ser utilizadas em RBF:

Função Gausseana:

$$\phi(r) = e^{(\varepsilon r)^2} \quad (3.37)$$

Função multiquadrática:

$$\phi(r) = \sqrt{1 + (\varepsilon r)^2} \quad (3.38)$$

Função *thin-plate-spline*:

$$\phi(r) = r^2 \ln(r) \quad (3.39)$$

Função Quadrática Inversa:

$$\phi(r) = \frac{1}{1 + (\varepsilon r)^2} \quad (3.40)$$

Função Multiquadrática Inversa:

$$\phi(r) = \frac{1}{\sqrt{1 + (\varepsilon r)^2}} \quad (3.41)$$

Em que  $v = \|x - \mu\|$ , que é dado geralmente pela distância euclidiana,  $x$  é o vetor de entrada e  $\mu$  e  $\sigma$  representam o centro e a largura da função radial, respectivamente. Geralmente, as redes de base radiais são construídas com apenas uma camada intermediária. A resposta gerada por um dado neurônio de saída é definida pela equação (BRAGA; CARVALHO; LUDERMIR, 2000):

$$Y_j = \sum_{i=1}^K w_{ji} \phi(\|x - \mu_i\|) + w_{j0} \quad (3.42)$$

Em que  $w_{j0}$  representa o termo de polarização.

Assim como as redes MLP, as redes RBF são aproximadores universais de funções e podem ser ditas como funções equivalentes. Porém, algumas diferenças são encontradas entre as duas redes:

Tabela 1 – Diferença entre redes MLP e RBF.

Redes MLP <i>versus</i> RBF	
MLP	RBF
É vista na literatura com uma ou mais camadas intermediárias	Geralmente só é construída com uma camada intermediária
Neurônios das camadas escondidas e de saída têm funções semelhantes	Neurônios da camada intermediária têm funções diferentes dos neurônios da camada de saída
Neurônios de saída nem sempre são lineares	Neurônios de saída sempre são lineares
Neurônios das camadas escondidas têm funções sigmoidais como função de ativação	Neurônios das camadas escondidas têm funções de base radiais como função de ativação
Separa padrões de entradas com hiperplanos	Separa padrões de entradas com hiperelipsoides
Constrói aproximadores globais para o mapeamento entrada-saída	Constrói aproximadores locais para o mapeamento entrada-saída
Geralmente utilizada para problemas complexos	Geralmente utilizada para problemas bem definidos

**Fonte:** (BRAGA; CARVALHO; LUDERMIR, 2000).

### 3.7 MODELO DE REGRESSÃO DE VETOR SUPORTE

Pode-se definir (CORTES; VAPNIK, 1995) (HEARST et al., 1998) (CRISTIANINI; SHAW-TAYLOR, 2000) (LORENA; CARVALHO, 2007) uma máquina de vetor de suporte (SVMs - do inglês *Support Vector Machine*) como um método de aprendizagem de máquina que recebe dados de entrada e classifica-os em uma entre duas categorias. Semelhantes a RNAs, as SVMs devem ser desenvolvidas a partir de um conjunto de dados de entrada e de saída de treinamento para construir o modelo de máquina de vetores de suporte que pode ser utilizado para classificação de novos dados, ou seja, deve-se dividir o conjunto de dados históricos em três subconjuntos: treinamento, validação e testes.

O subconjunto de treinamento é utilizado para construir a Máquina de Vetor de Suporte. A partir dos dados de entrada e saída, o treinamento mapeia os dados de entrada no espaço multidimensional e, utilizando regressão, encontra-se o hiperplano<sup>2</sup> que melhor separa as duas classes. Após o treinamento, a máquina é validada e a capacidade de avaliação das novas entradas em relação ao hiperplano divisor é testada, sendo essas novas entradas classificadas entre as duas categorias criadas no treinamento.

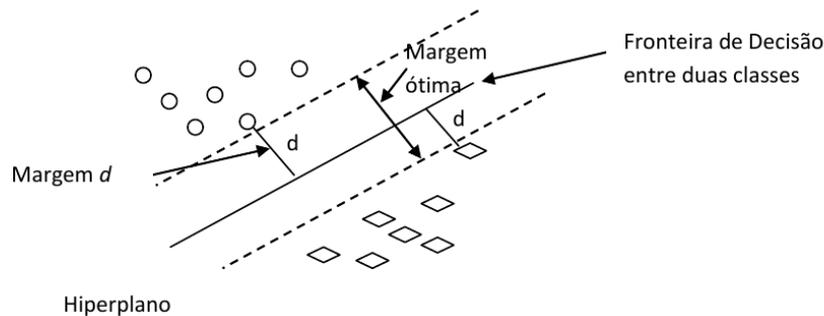
Segundo (BRAGA; CARVALHO; LUDERMIR, 2000), esses resultados são alcançados pelo emprego dos conceitos da Teoria de Aprendizado Estatístico, que apresentam diversos limites na capacidade de generalização de um classificador linear. Dessa forma, dado um conjunto de treinamento  $X$  com  $i$  pares  $(x_i, y_i)$ , em que  $x_i \in \mathbb{R}$  e  $y_i \in -1, +1$ , as SVMs buscam o classificador linear  $g(x) = \text{sgn}(w \cdot x + b)$ , capaz de separar os dados pertencentes a  $X$  com erro mínimo nas duas classes criadas pelo hiperplano, maximizando a margem  $d$  de separação entre as classes.

Dada uma função linear  $f(x) = w \cdot x + b$ , a margem  $d(x_i, y_i)$  utilizada para classificar

<sup>2</sup> um hiperplano é uma superfície em espaço de  $n$  dimensões que o separa em duas metades de espaço

um padrão  $x_i$  é fornecida por  $y_i f(x_i)$ . Tal função mede a distância entre a margem  $d$  e o valor  $x_i$ . A margem  $d$  é a margem mínima observada em todo o conjunto de treinamento. Pode-se observar na figura 4.

Figura 4 – Exemplo de uma máquina de vetor de suporte.



**Fonte:** O autor.

Todavia, dois problemas surgem na busca do desenvolvimento de uma SVM. O primeiro deles refere-se a encontrar um hiperplano de separação que generalize bem a separação das classes, pois a dimensionalidade do espaço pode ser enorme e nem todos os hiperplanos conseguem generalizar o hiperplano ótimo. O segundo problema é a ordem de grandeza das funções, pois podem existir funções polinomiais de grau 2, 3, 4, 200 ou mais, o que resultaria na construção de um hiperplano com mais de um bilhão de dimensões (BRAGA; CARVALHO; LUDERMIR, 2000).

O primeiro problema foi solucionado por (VAPNIK; KOTZ, 1982), em 1965, para o caso de hiperplanos que tivessem classes aptas a separá-los. Um hiperplano ótimo é definido como a função de decisão linear com margem máxima entre os vetores das duas categorias. Observou-se que para construir tais hiperplanos ótimos é necessário uma quantidade de informações para treinar e determinar os locais das margens  $d$ , os chamados vetores de apoio. No trabalho supracitado foi mostrado que se os vetores de treinamento são separados sem erros por um hiperplano ótimo, a esperança da probabilidade do erro  $E$  é limitada pela razão entre a esperança do número de vetores de suporte e o número de vetores de treinamento.

O segundo problema foi superado e demonstrado em 1992 por (BOSER; GUYON; VAPNIK, 1992). Foi desenvolvido um algoritmo de treinamento de margem máxima, que possibilita a obtenção da margem ótima.

A separação das classes é desenvolvida através da técnica de mapeamento de qualquer função, linear ou não, num espaço linear. O objetivo é mapear cada padrão do conjunto de treinamento  $E$  para um novo espaço, denominado espaço característica. Um aspecto desse espaço é que a escolha de uma função de mapeamento  $\varphi$  apropriada torna o conjunto de treinamento mapeado linearmente separável. As funções utilizadas para esses mapeamentos são chamadas de funções *Kernel*, sendo as mais utilizadas as polinomiais, RBF

(Gaussiana) e sigmoidais. A escolha da função Kernel e dos valores para seus parâmetros livres, assim como constantes, exercem influência direta no desempenho do classificador.

Pode-se observar vários estudos na literatura utilizando SVM (KAZEM et al., 2013) (ZBIKOWSKI, 2015) (XIONG; BAO; HU, 2014) (CHEN; CHEN, 2015) (JIANG et al., 2015). Sua expansão foi motivada principalmente pela observância de sua capacidade de generalização, semelhante às Redes Neurais Artificiais (RNAs), e aos resultados obtidos quando se trabalha com dados de grande dimensionalidade.

(RANCO et al., 2015) mostra, em seu trabalho, o desenvolvimento de uma metodologia para analisar os efeitos da rede social Twitter, analisando os sentimentos e efeitos no mercado financeiro dessas informações (análise fundamentalista). Para a classificação, o citado autor utilizou uma Máquina de Suporte de Vetores e mostrou uma ligação entre os eventos ocorridos no Twitter e a flutuação de preço dos ativos no mercado financeiro.

A regressão de uma SVM é denominada de SVR (*Support Vector Regression*) e foi proposta em 1996 por Vapnik (CORTES; VAPNIK, 1995)(HEARST et al., 1998). O modelo de regressão depende apenas de um subconjunto de dados de entrada para o treinamento, pois a função custo só busca os pontos que estão na margem  $d$ . Assim como em outras regressões, o objetivo da regressão de vetores de suporte é aproximar as previsões de saída (valores futuros) baseadas nos valores de entrada, utilizando uma função não linear entre as entradas dadas e os correspondentes dados de saída (target) na etapa de treinamento (MENG; GAO; SHI, 2007).

Segundo (SMOLA et al., 1996), SVMs podem ser utilizadas como alternativa para se trabalhar com problemas de regressão. Para isso, faz-se necessário a modificação da função de perda para incluir uma medida de distância. Alguns exemplos de funções de perdas são: Quadrática, Laplace, Huber e  $\epsilon$ -insensitive.

A figura 5 exemplifica a forma de cada função de perda. As funções Quadrática, Laplace e Huber não produzirão novos vetores de suportes. Por esse motivo, Vapnik propôs a função  $\epsilon$ -insensitive como uma aproximação da função de perda de Huber, que permite a obtenção de um conjunto de vetores de suporte.

### 3.7.1 Regressão de Vetor de Suporte Linear

A Regressão de Vetor de Suporte Linear é uma das formas de regressão para vetores de suporte. O trabalho de (GUNN et al., 1998) descreve a utilização da regressão de vetor de suporte linear.

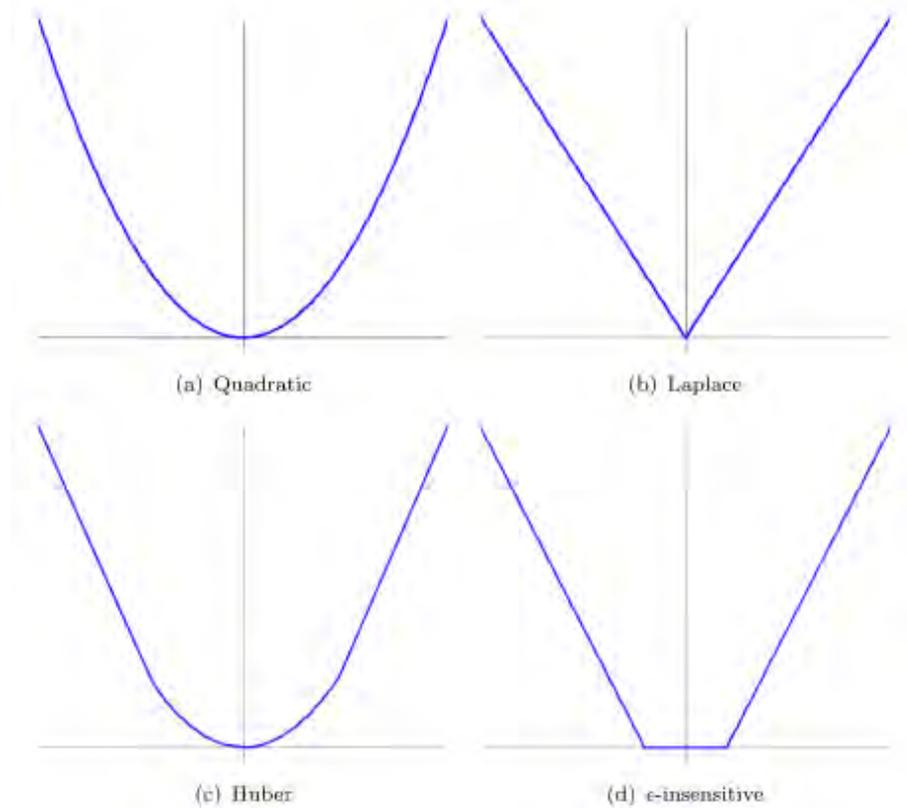
Considere-se um problema de aproximação para um conjunto dado por:

$$D = \{(x^1, y^1), \dots, (x^l, y^l), \quad x \in \mathbb{R}^n, y \in \mathbb{R}\} \quad (3.43)$$

Com a função linear

$$f(x) = \langle \omega, x \rangle + b \quad (3.44)$$

Figura 5 – Funções de perda utilizadas em regressão de vetor de suporte.



**Fonte:** Fonte:(MENG; GAO; SHI, 2007).

A função de regressão ótima é dada pelo mínimo da função:

$$\Phi(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_i (\xi_i^- + \xi_i^+) \quad (3.45)$$

no qual  $C$  é o valor pré-especificado e  $\xi_-, \xi_+$  são variáveis de entrada (mínimo e máximo) que restringem o valor da saída do sistema.

Pode-se utilizar quaisquer uma das quatro funções de perda citadas acima. Neste trabalho, exemplificaremos a regressão de vetor de suporte linear utilizando apenas a função  $\epsilon$ -insensitive e a função Quadrática, pois estas são as mais utilizadas na literatura.

### 3.7.1.1 Regressão linear com função de erro $\epsilon$ -insensitive

Utilizando a função  $\epsilon$ -insensitive como função de perda:

$$L_\epsilon(y) = \begin{cases} 0 & \text{para } |f(x) - y| < \epsilon \\ |f(x) - y| - \epsilon & \text{caso contrário} \end{cases} \quad (3.46)$$

Tem-se a solução dada por:

$$\max_{\alpha, \alpha^*} W(\alpha, \alpha^*) = \max_{\alpha, \alpha^*} -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) \langle x_i, x_j \rangle + \sum_{i=1}^l \alpha_i (y_i - \epsilon) - \alpha_i^* (y_i + \epsilon). \quad (3.47)$$

ou

$$\alpha, \alpha^* = \arg \min_{\alpha, \alpha^*} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) \langle x_i, x_j \rangle - \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i + \sum_{i=1}^l (\alpha_i - \alpha_i^*) \epsilon. \quad (3.48)$$

Com as seguintes restrições:

$$\begin{aligned} 0 \leq \alpha_i, \alpha_i^* \leq C, \quad i = 1, \dots, l \\ \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \end{aligned} \quad (3.49)$$

Após a resolução da equação 3.47 com as restrições equação 3.50, será determinado o multiplicador de (CLARKE, 1976) (LAGRANGE, 1853), de modo que a função de regressão linear da equação 3.44 terá como variáveis:

$$\bar{\omega} = \sum_{i=1}^l (\bar{\alpha}_i - \bar{\alpha}_i^*) x_i$$

e

$$\bar{b} = -\frac{1}{2} \langle \bar{\omega}, (x_r, x_s) \rangle.$$

O próximo passo é verificar as condições de Karush-Kuhn-Tucker (KKT) (MORGAN, 1999) que satisfazem a equação, obtendo como resultado:

$$\bar{\alpha}_i \bar{\alpha}_j = 0, \quad i = 1, \dots, l. \quad (3.50)$$

Portanto, o vetor de suporte será os pontos onde exatamente um dos multiplicadores de Lagrange é maior que 0. Quando  $\epsilon = 0$ , a função de perda  $L_1$  é utilizada e a otimização do problema é simplificada,

$$\min_{\beta} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \beta_i \beta_j \langle x_i, x_j \rangle - \sum_{i=1}^l \beta_i y_i \quad (3.51)$$

Com a restrição,

$$\begin{aligned} -C \leq \beta_i \leq C, \quad i = 1, \dots, l \\ \sum_{i=1}^l \beta_i = 0 \end{aligned} \quad (3.52)$$

E a função de regressão linear tendo como variáveis,

$$\begin{aligned}\bar{\omega} &= \sum_{i=1}^l \beta_i x_i \\ \bar{b} &= -\frac{1}{2} \langle \bar{\omega}, (x_r + x_s) \rangle.\end{aligned}\quad (3.53)$$

### 3.8 REGRESSÃO LINEAR COM FUNÇÃO DE ERRO QUADRÁTICA

Utilizando a função de erro quadrática na regressão de suporte de vetores linear tem-se:

$$L_{quad}(f(x) - y) = (fx(x) - y)^2 \quad (3.54)$$

A solução para esta função é encontrada por:

$$\begin{aligned}\max_{\alpha, \alpha^*} W(\alpha, \alpha^*) &= \max_{\alpha, \alpha^*} -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) \langle x_i, x_j \rangle \\ &\quad + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i - \frac{1}{2C} \sum_{i=1}^l (\alpha_i^2 + (\alpha_i^*)^2).\end{aligned}$$

Semelhante à utilização da função de erro  $\epsilon$ -insensitive, a otimização pode ser encontrada através do uso de KKT equação 3.50 e  $\beta_i^8 = |\beta_i|$ . Logo, o resultado da otimização é dado por:

$$\min_{\beta} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \beta_i \beta_j \langle x_i, x_j \rangle - \sum_{i=1}^l \beta_i y_i + \frac{1}{2C} \sum_{i=1}^l \beta_i^2 \quad (3.55)$$

Restritos por:

$$\sum_{i=1}^l \beta_i = 0.$$

Dessa forma, a função de regressão é encontrada através da equação 3.44 e da equação 3.54 (GUNN et al., 1998).

#### 3.8.1 Regressão de Vetor de Suporte Não Linear

Pode-se utilizar regressão de vetor de suporte não linear para mapear funções não lineares. O comportamento do mapeamento será semelhante a um classificador de vetor de suporte não linear, com a finalidade de mapear os dados em um espaço de grande dimensão, onde a regressão linear é definida. A solução SRV não linear, utilizando como função de perda  $\epsilon$ -insensitive é dada por:

$$\max_{\alpha, \alpha^*} W(\alpha, \alpha^*) = \max_{\alpha, \alpha^*} \sum_{i=1}^l \alpha_i^* (y_i - \epsilon) - \alpha_i (y_i + \epsilon) - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(x_i, x_j) \quad (3.56)$$

restrito por:

$$0 \leq \alpha_i, \alpha_i^* \leq C, \quad i = 1, \dots, l \quad (3.57)$$

$$\sum_{I=1}^l (\alpha_i - \alpha_i^*) = 0$$

A partir da equação equação 3.56 com as restrições da equação equação 3.58 é determinado o multiplicador de Lagrange,  $\alpha_i, \alpha_i^*$ , e a função de regressão é dada por,

$$f(x) = \sum_{SV_s} (\bar{\alpha}_i - \bar{\alpha}_i^*) K(x_i, x) + \bar{b} \quad (3.58)$$

No qual

$$f(x) = \sum_{SV_s} (\bar{\alpha}_i - \bar{\alpha}_i^*) K(x_i, x) + \bar{b}$$

e

$$\bar{b} = -\frac{1}{2} \sum_{i=1}^l (\bar{\alpha}_i - \bar{\alpha}_i^*) (K(x_i, x_r) + K(x_i, x_s).)$$

A restrição de igualdade pode ser descartada se o *kernel* contiver o *bias* dentro da função. Portanto, a função de regressão é dada por:

$$f(x) = \sum_{i=1}^l (\bar{\alpha}_i - \bar{\alpha}_i^*) K(x_i, x) \quad (3.59)$$

De forma similar, os critérios de otimização utilizando outras funções de perdas citadas nesta seção podem ser obtidos sempre substituindo o produto dos pontos por uma função *Kernel*.

### 3.8.1.1 Kernel Polinomial

Com o *Kernel* polinomial, para polinômios de grau- $d$ , a função é dada por:

$$K(x, y) = (x^\top y + c)^d \quad (3.60)$$

em que  $x$  e  $y$  são vetores no espaço de entrada, ou seja, vetores calculados a partir da série temporal;  $c$  é uma constante positiva ou neutra, que influencia a função polinomial com valores tanto para alto quanto para baixo do valor de origem. Quando  $c$  é neutra ( $c = 0$ ) têm-se uma função *Kernel* chamada de função homogênea. A função *Kernel* corresponde a um produto interno com um espaço de mapeamento  $\varphi$  definido por:

$$K(x, y) = \langle \varphi(x), \varphi(y) \rangle \quad (3.61)$$

Para encontrar o espaço de mapeamento  $\varphi$  faz-se necessário a resolução do polinômio da função *Kernel*. Tendo  $d$  da equação 3.60 igual a 2, teremos um caso especial denominado de *Kernel* quadrático, no qual a função é definida por:

$$K(x, y) = \left( \sum_{i=1}^n x_i y_i + c \right)^2 = \sum_{i=1}^n (x_i^2)(y_i^2) + \sum_{i=2}^n \sum_{j=1}^{i-1} (\sqrt{2}x_i x_j)(\sqrt{2}y_i y_j) + \sum_{i=1}^n (\sqrt{2c}x_i)(\sqrt{2c}y_i) + c^2 \quad (3.62)$$

Sendo, portanto, o espaço de mapeamento  $\varphi$  dado por:

$$\varphi(x) = \langle x_n^2, \dots, x_1^2, \sqrt{2}x_n x_{n-1}, \dots, \sqrt{2}x_n x_1, \sqrt{2}x_{n-1} x_{n-2}, \dots, \sqrt{2}x_{n-1} x_1, \dots, \sqrt{2}x_2 x_1, \sqrt{2c}x_n, \dots, \sqrt{2c}x_1, c \rangle$$

### 3.8.1.2 Kernel Função de Base Radial

Pode-se utilizar como função *Kernel* uma função de base radial (RBF), bastante utilizada em Classificação de Vetores de Suporte (SVC). Uma função RBF *kernel* de duas amostras  $x$  e  $x'$  pode ser representada por:

$$K(x, x') = \exp \left( -\frac{\|x - x'\|^2}{2\sigma^2} \right) \quad (3.63)$$

$\|x - x'\|^2$  é reconhecida como o quadrado da distância euclidiana entre os dois vetores.  $\sigma$  é um parâmetro livre denominado de medida de similaridade. Quando este parâmetro é igual a 1, a função *kernel* é dada por:

$$\begin{aligned} \exp \left( -\frac{1}{2}\|x - x'\|^2 \right) &= \sum_{j=0}^{\infty} \frac{(x^\top x')^j}{j!} \exp \left( -\frac{1}{2}\|x\|^2 \right) \exp \left( -\frac{1}{2}\|x'\|^2 \right) \\ &= \sum_{j=0}^{\infty} \sum_{\sum n_i=j} \exp \left( -\frac{1}{2}\|x\|^2 \right) \frac{x_1^{n_1} \cdots x_k^{n_k}}{\sqrt{n_1! \cdots n_k!}} \exp \left( -\frac{1}{2}\|x'\|^2 \right) \frac{x_1'^{n_1} \cdots x_k'^{n_k}}{\sqrt{n_1! \cdots n_k!}} \end{aligned} \quad (3.64)$$

A função original de treinamento de um modelo de regressão de máquina de vetores de suporte é definida pela fórmula:

$$\begin{aligned} &\text{minimizar} && \frac{1}{2}\|\omega\|^2 \\ &\text{sujeito para} && \begin{cases} y_i - \langle \omega, x_i \rangle - b \leq \varepsilon \\ \langle \omega, x_i \rangle + b - y_i \leq \varepsilon \end{cases} \end{aligned}$$

em que  $x_i$  é a amostra de treinamento e o  $y_i$  o *target* (saída real) da série temporal. A expressão é o valor de previsão da amostra, e  $\varepsilon$  é o limiar da diferença entre o valor real e o valor previsto no treinamento, ou seja, o valor limite de ruído (FALAT; PANCIKOVA; HLINKOVA, 2015).

(FALAT; PANCIKOVA; HLINKOVA, 2015) utiliza SVM para desenvolver um modelo de predição para séries temporais de alta volatilidade no mercado financeiro. Em seus experimentos, o autor compara o modelo desenvolvido com o método ARIMA, obtendo uma significativa melhora de desempenho quando comparado o valor MAPE (*Mean Absolute Percentage Error*). Essa é uma das métricas que pode ser utilizada na comparação entre os modelos de previsões desenvolvidos.

Tabela 2 – Comparação entre Modelagens para Séries.

<b>Característica</b>	<b>ARIMA</b>	<b>MLP</b>	<b>NARX</b>	<b>LOGREG</b>	<b>RBF</b>	<b>SVR</b>
Tipo de modelo	Estatístico	Rede Neural	Rede Neural	Estatístico	Rede Neural	SVM
Aplicado a séries financeiras	Sim	Sim	Sim	Sim	Sim	Sim
Modelo Linear	Sim	Sim	Sim	Sim	Sim	As vezes
Análise em séries financeiras de frequências distintas	Não	Não	Não	Não	Não	Não

**Fonte:** O autor.

### 3.9 RESUMO DO CAPÍTULO

Neste capítulo foram apresentados os modelos de previsão de séries temporais mais utilizados na literatura atual. Foi detalhado o estado da arte no âmbito das metodologias utilizadas para previsão de séries temporais financeiras, e descritos conceitos importantes para o bom andamento desta pesquisa, que darão suporte ao estudo realizado com redes neurais e algoritmos genéticos.

## 4 ANÁLISE DAS SÉRIES TEMPORAIS FINANCEIRAS

Neste capítulo serão descritas as séries temporais financeiras utilizadas no presente estudo. Serão apresentadas a descrição e a motivação de cada série, além de um estudo minucioso mostrando o comportamento destas. Após o detalhamento de cada uma das séries será analisada quanto ao seu comportamento no final do capítulo.

### 4.1 ANÁLISE DAS SÉRIES TEMPORAIS FINANCEIRAS

Com a finalidade de testar o modelo proposto perante aos existentes na literatura sobre predição de séries temporais financeiras de baixa frequência, foram escolhidos quatro dos principais ativos da Bolsa de Valores de São Paulo (IBovespa): Vale S/A (VALE5), Petrobras PN (PETR4), Bradesco PN (BBDC4) e Companhia Siderúrgica Nacional (CSNA3). Esses ativos foram escolhidos por fazerem parte do grupo dos principais papeis da Ibovespa, conhecidos como Bluechips, possuindo uma quantidade de informações que se adequa à criação das séries temporais em diferentes frequências que serão analisadas. A escolha das Bluechips foi baseada no volume de negociação de cada ativo e nos trabalhos estudados na literatura.

Com o objetivo de mostrar que o modelo proposto pode ser utilizado em diferentes frequências de tempo, os ativos serão analisados nas frequências diária, semanal e quinzenal. Para tanto, deve-se analisar cada série, a fim de verificar suas semelhanças comportamentais, dentre elas: se suas funções são crescentes ou decrescentes, a relação dos retardos temporais com os valores da série, e se esse comportamento também pode ser visto nos outros ativos do Ibovespa.

Toda a análise foi desenvolvida com base nos valores de fechamento do preço, utilizando a técnica do *lagplot* (PERCIVAL; WALDEN, 1993; KANTZ; SCHREIBER, 2004). O *lagplot* é uma técnica de análise de gráficos que mostra o comportamento de um valor com os valores passados para se determinar o fenômeno gerador de uma série temporal. Esse método foi utilizado também por (ARAÚJO, 2016), e mostra uma forte dependência da interpretação humana do gráficos, sendo que em alguns casos as relações contidas nos gráficos podem não refletir claramente as características do fenômeno gerador da série. Para auxiliar no processo de análise dos gráficos, foram utilizados também o *variance ratio test* (VRT) (LO; MACKINLAY, 1988; LO; MACKINLAY, 1989; LO; MACKINLAY, 2002; CHARLES; DARNÉ, 2009), *autocorrelation function* (Função de Autocorrelação - autocorrelation function (ACF)) (BOX et al., 2015), *partial autocorrelation function* (PACF)] (BOX et al., 2015), *mean mutual information* (MMI) (FRASER; SWINNEY, 1986; KRASKOV; STGBAUER; GRASSBERGER, 2004) e *Hurst parameter* (HP) (HURST, 1951; CLEGG, 2006). Para verificar se a série temporal financeira poderá ser utilizada, foram feitos estudos de

autocorrelação, autocorrelação parcial e o *lagplot* (PERCIVAL; WALDEN, 1993) (KANTZ; SCHREIBER, 2004).

Na estatística, autocorrelação é uma medida que informa o quanto o valor de uma variável aleatória pode influenciar seus vizinhos, ou seja, identificar o nível de influência do valor atual com valores do passado. No mercado financeiro, podemos exemplificar da seguinte forma: quanto o valor de uma ação subiu em quinze minutos influenciou numa alta da mesma ação no dia, semana ou quinzena seguintes?

A autocorrelação está entre os valores 1, que seria a correlação perfeita, e  $-1$ , que seria a anticorrelação perfeita. O valor 0 significa a total ausência de correlação.

Supondo uma variável aleatória  $X_t$  discreta estacionária, dependente do tempo, com média  $\mu$ , sua autocorelação  $R(k)$  é definida como:

$$R(k) = \frac{E[(X_t - \mu)(X_{t+k} - \mu)]}{\sigma^2} \quad (4.1)$$

em que  $E[\cdot]$  é o valor médio, esperança matemática ou expectativa da expressão,  $k$  é o deslocamento no tempo e  $\sigma^2$  é a variância da variável  $X_t$ .

Com os gráficos *lagplots*, pode-se analisar o comportamento de cada série temporal financeira com relação aos atrasos (pontos do passado). Dessa forma, consegue-se observar se há uma componente linear e não linear que influencie a série temporal financeira, identificando, caso exista, qual seria a componente dominante e qual seria a subdominante (PERCIVAL; WALDEN, 1993) (KANTZ; SCHREIBER, 2004) (ARAÚJO, 2016). O *lagplot* é um gráfico de dispersão que mede a relação entre os retardos temporais de uma série. Caso exista alguma relação entre os retardos, aparecerá uma estrutura bem definida identificando sua forma.

O gráfico de autocorrelação parcial é útil na identificação da ordem de um modelo autorregressivo. Essa análise utiliza modelos AR(p). Se o gráfico da autocorrelação amostral indicar que um modelo AR pode ser apropriado, então o gráfico de autocorrelação parcial amostral é examinado para ajudar identificar a ordem. Busca-se o ponto no gráfico onde as autocorrelações parciais tornam-se essencialmente zero. Para isso, foi estipulado um intervalo de confiança de 95% para a significância estatística. O cálculo da VRT é um procedimento utilizado para verificar se o fenômeno gerador da série tende a um processo aleatório ou não. As funções ACF e PACF foram utilizadas para analisar o comportamento da componente linear dominante do mapeamento. A função *Average Mutual Information* Informação Médua Média - Average Mutual Information (AMI) mede o quanto uma variável aleatória nos diz sobre outra. No contexto da análise de séries temporais, a AMI ajuda a quantificar a quantidade de conhecimento adquirido do valor de  $x(t + \tau)$  ao observar  $x(t)$ . A MMI foi utilizada para reforçar a análise da componente não linear sub dominante, verificando se existe dependência não linear no fenômeno gerador da série, uma vez que ela representa uma medida de dependência não linear. Para auxiliar no estudo da componente sub dominante não linear, também foi utilizada a *Hurst Parameter*, parâmetro

que verifica se tal dependência é caracterizada como de longo ( $0,5 < HP \leq 1,0$ ) ou de curto prazo ( $0,0 < HP \leq 0,5$ ), quando utilizado o método R/S. (CLEGG, 2006).

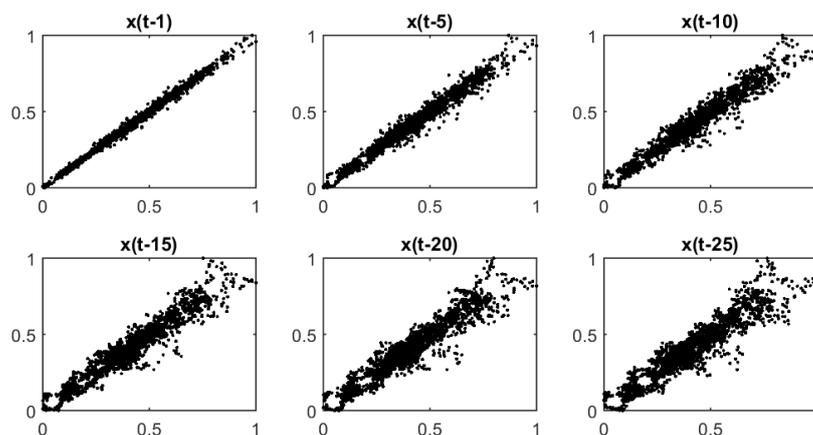
Destarte, serão analisados quatro dos principais ativos mais negociados na BM&FBovespa, baseado na carteira teórica do Ibovespa válida para 14/03/2016 (BM&FBOVESPA, 2016). Os quatro ativos escolhidos serão analisados nas frequências diária, semanal e quinzenal, com a finalidade de observar as características dos fenômenos geradores. Foram escolhidos os papéis de acordo com três regras: participação em porcentagem, quantidade teórica e diversificação da carteira. Por exemplo, as ações da Petrobras ON e PNA têm uma participação em percentual e a quantidade teórica de negociações maiores do que as ações da Vale do Rio Doce Vale5 PNA. Porém, a Vale5 PNA foi inserida e o ativo PETR3 ON foi retirada para uma análise mais diversificada da carteira.

Os ativos escolhidos para o estudo foram: Petrobras PN (PETR4), VALE PNA (VALE5), Bradesco (BBDC4), Companhia Siderúrgica Nacional (CSNA3).

#### 4.1.1 Análise do ativo PETR4 na frequência diária

Para o ativo Ações Preferenciais da Empresa Petrobras (PETR4) com dados de fechamento na frequência diária temos o *lagplot* com atrasos de baixa ordem (1 a 25) conforme mostra a figura 6.

Figura 6 – Lagplot de baixa ordem da série temporal PETR4 na frequência diária.



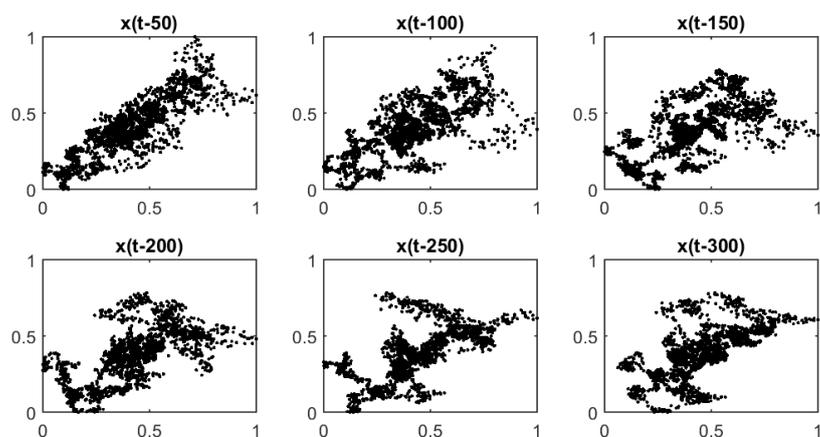
Fonte: O autor.

A figura 7 expõe os gráficos dos retardos de alta ordem, entre (50 a 300).

Os gráficos de ACF e Função de Autocorrelação Parcial - Partial Autocorrelation Function (PACF) são ilustradas nas figura 8, figura 9, respectivamente.

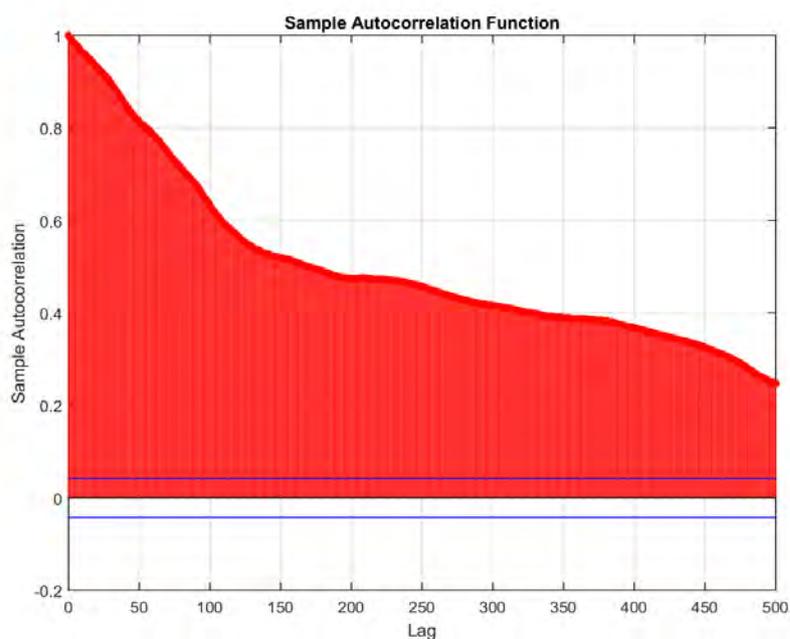
Os valores estatísticos calculados para a série temporal PETR4 na frequência diária pode ser vista na tabela 3.

Figura 7 – Lagplot de alta ordem da série temporal PETR4 na frequência diária.



Fonte: O autor.

Figura 8 – Gráfico de autocorrelação dos dados da PETR4 na frequência diária.



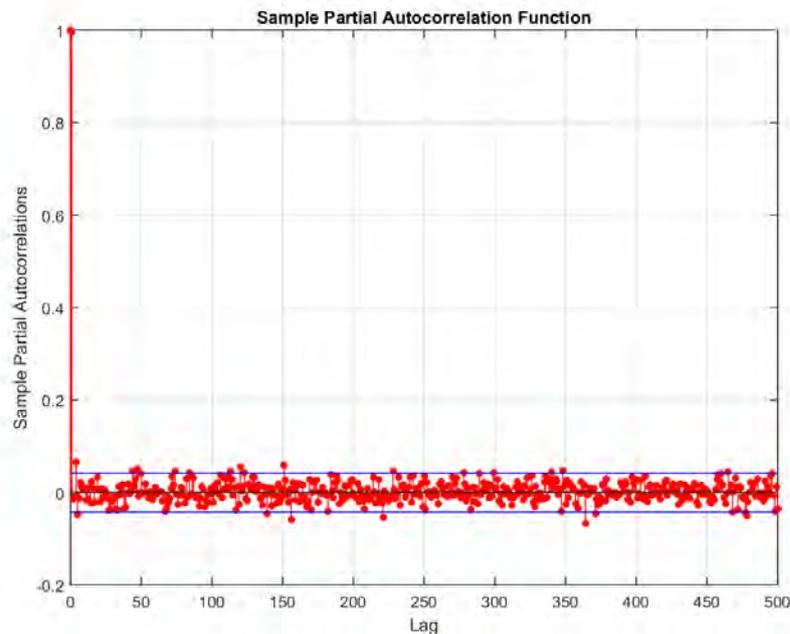
Fonte: O autor.

Tabela 3 – Estatísticas PETR4 na frequência diária.

PETR4	VRT	HP	MAI
D1	0,00	1,00	0,700501

Fonte: O autor.

Figura 9 – Gráfico de autocorrelação parcial dos dados da PETR4 na frequência diária.

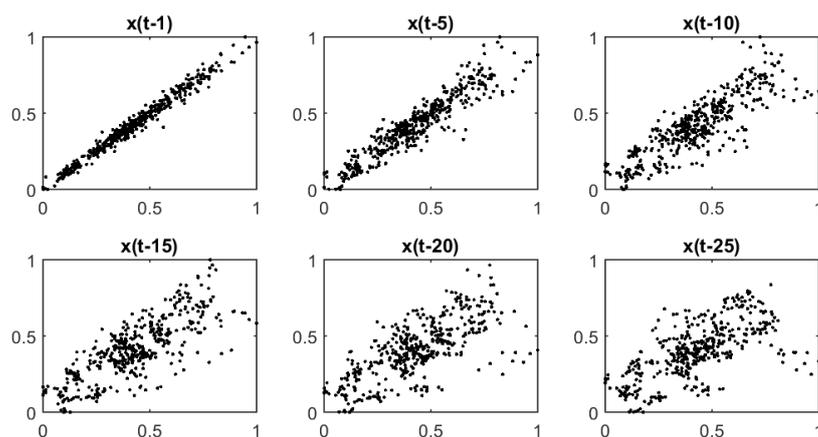


Fonte: O autor.

#### 4.1.2 Análise do ativo PETR4 na frequência semanal

Para o ativo PETR4 com dados de fechamento na frequência semanal o *lagplot* com atrasos de baixa ordem (1 a 25) é mostrado na figura 10. A figura 11 expõe os gráficos

Figura 10 – Lagplot de baixa ordem da série temporal PETR4 na frequência semanal.



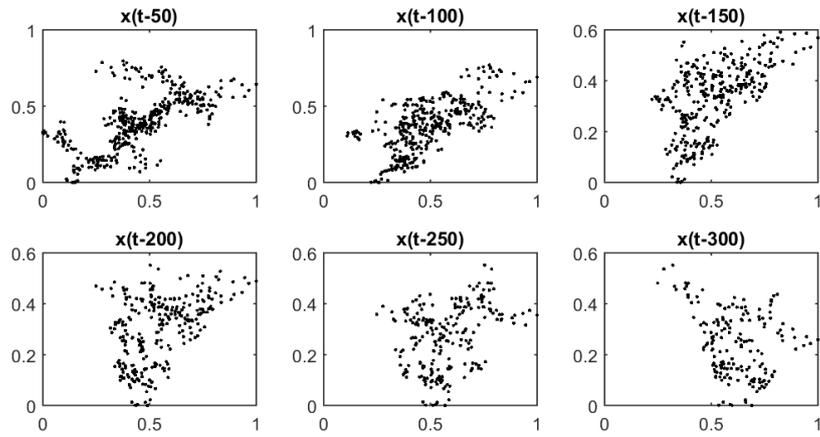
Fonte: O autor.

dos retardos de alta ordem, entre (50 a 300).

Os gráficos de ACF e PACF são ilustradas nas figura 12, figura 13, respectivamente.

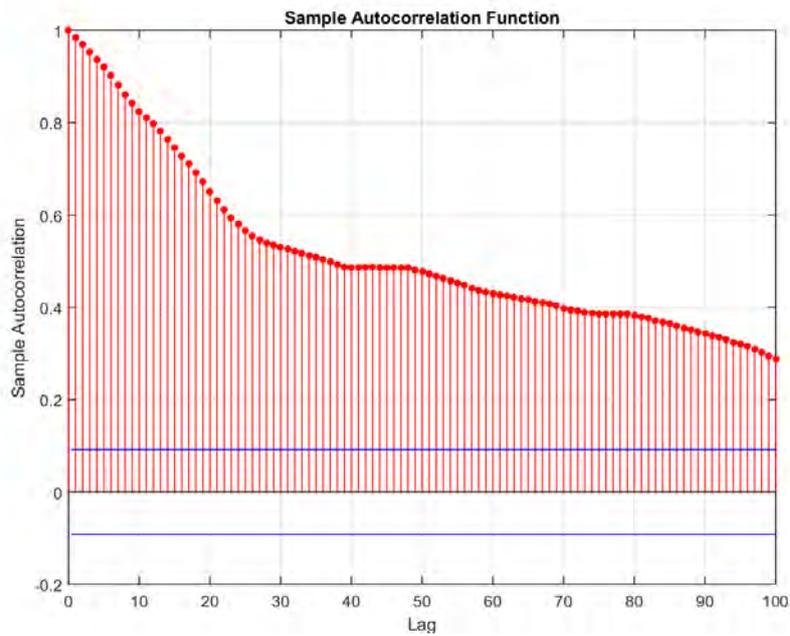
Os valores estatísticos calculados para a série temporal PETR4 na frequência semanal pode ser vista na tabela 4.

Figura 11 – Lagplot de alta ordem da série temporal PETR4 na frequência semanal.



Fonte: O autor.

Figura 12 – Gráfico de autocorrelação dos dados da PETR4 na frequência semanal.



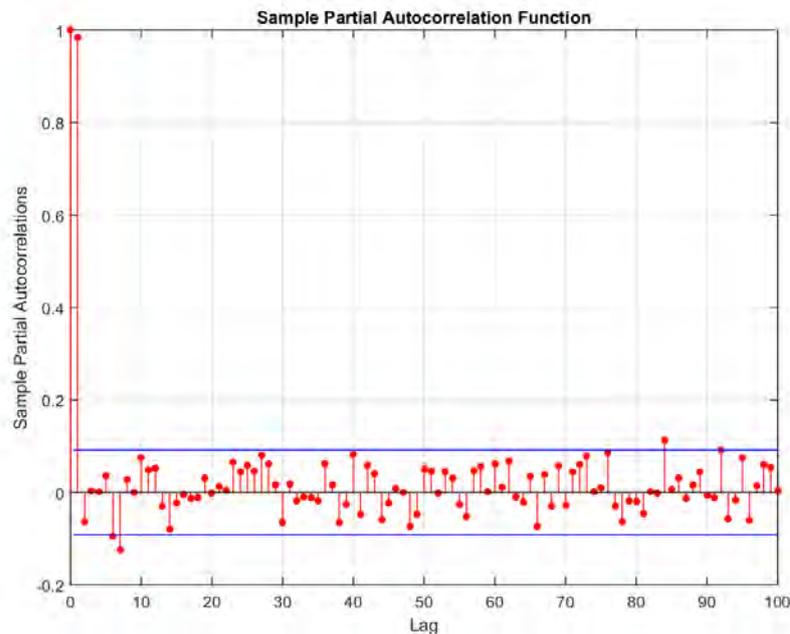
Fonte: O autor.

Tabela 4 – Estatísticas PETR4 na frequência semanal.

PETR4	VRT	HP	MAI
W1	0,00	0,906	0,661712

Fonte: O autor.

Figura 13 – Gráfico de autocorrelação parcial dos dados da PETR4 na frequência semanal.

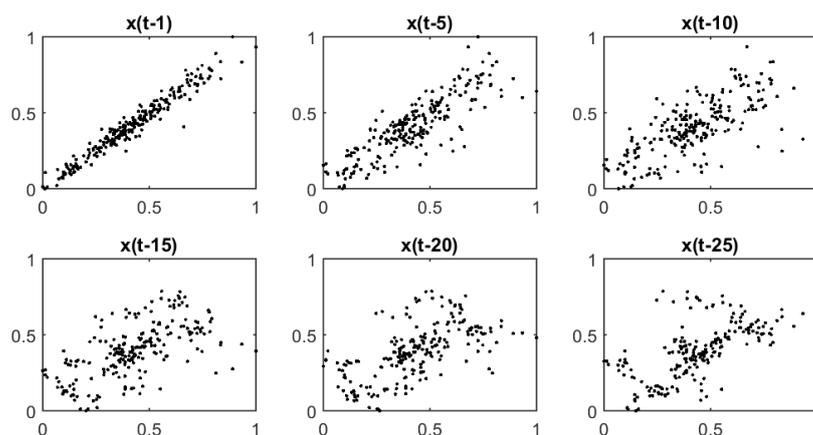


Fonte: O autor.

#### 4.1.3 Análise do ativo PETR4 na frequências quinzenal

Para o ativo PETR4 com dados de fechamento na frequência quinzenal temos o *lagplot* com atrasos de baixa ordem (1 a 25) conforme mostra a figura 14.

Figura 14 – Lagplot de baixa ordem da série temporal PETR4 na frequência quinzenal.



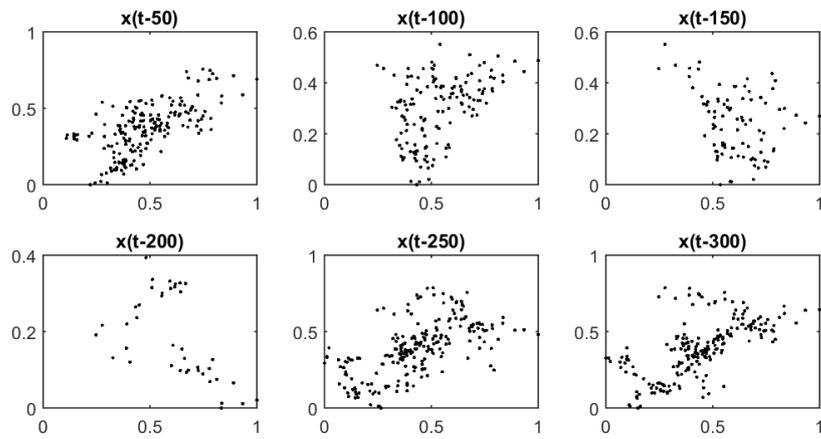
Fonte: O autor.

A figura 15 expõe os gráficos dos retardos de alta ordem, entre (50 a 300).

Os gráficos de ACF e PACF são ilustradas nas figura 16, figura 17, respectivamente.

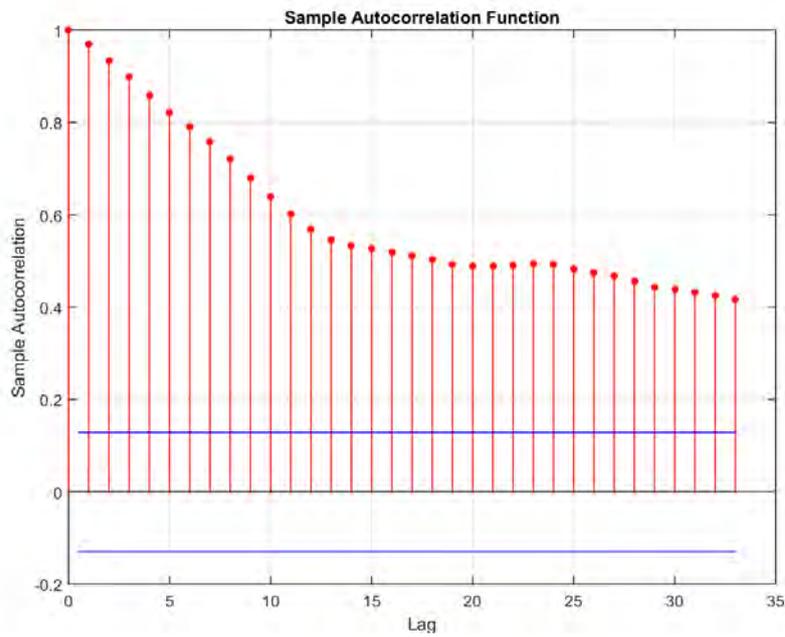
Os valores estatísticos calculados para a série temporal PETR4 na frequência quinzenal pode ser vista na tabela 5.

Figura 15 – Lagplot de alta ordem da série temporal PETR4 na frequência quinzenal.



Fonte: O autor.

Figura 16 – Gráfico de autocorrelação dos dados da PETR4 na frequência quinzenal.



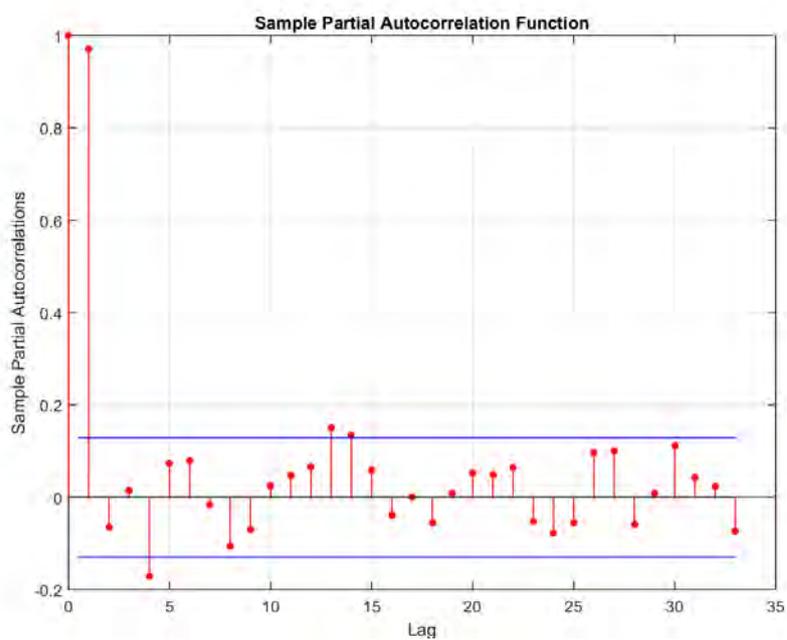
Fonte: O autor.

Tabela 5 – Estatísticas PETR4 na frequência quinzenal.

PETR4	VRT	HP	MAI
D15	0,00	0,949	0,829943

Fonte: O autor.

Figura 17 – Gráfico de autocorrelação parcial dos dados da PETR4 na frequência quinzenal.

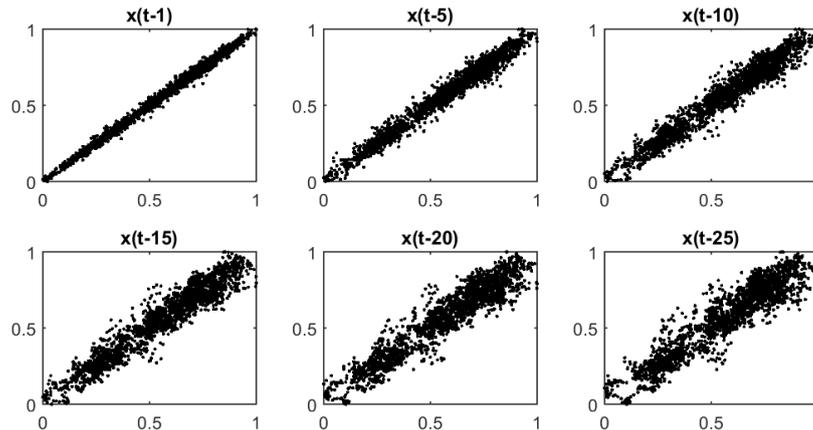


Fonte: O autor.

#### 4.1.4 Análise do ativo VALE5 na frequência diária

Para o ativo Ações Preferenciais da Empresa Vale SA (VALE5) com dados de fechamento na frequência diária temos o *lagplot* com atrasos de baixa ordem (1 a 25) conforme mostra a figura 18.

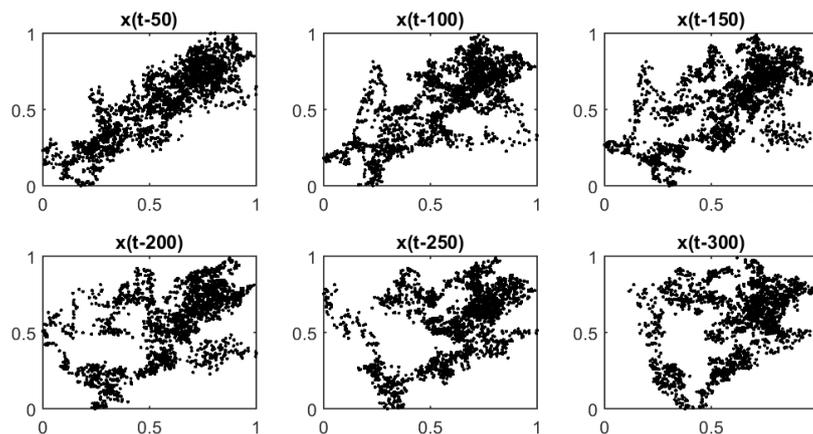
Figura 18 – Lagplot de baixa ordem da série temporal VALE5 na frequência diária.



Fonte: O autor.

A figura 19 expõe os gráficos dos retardos de alta ordem, entre (50 a 300).

Figura 19 – Lagplot de alta ordem da série temporal VALE5 na frequência diária.

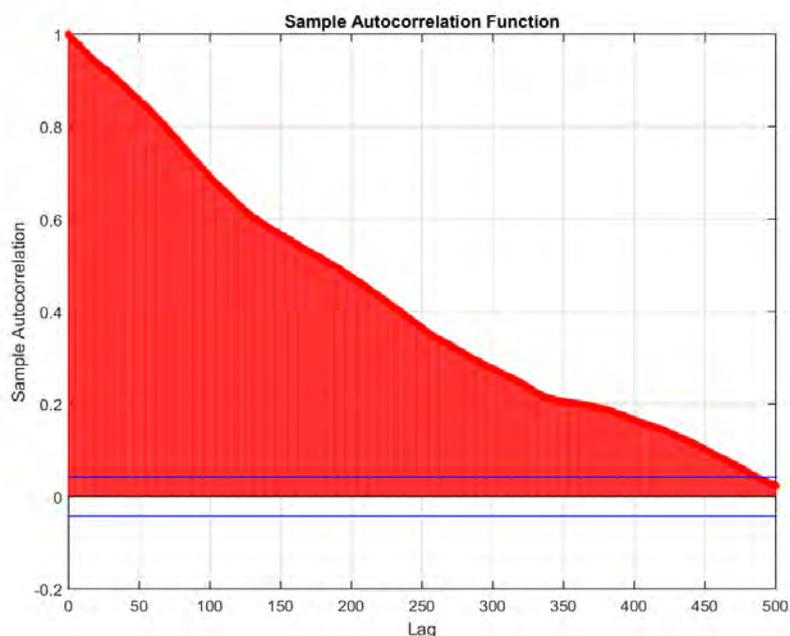


Fonte: O autor.

Os gráficos de ACF e PACF são ilustradas nas figura 20, figura 21, respectivamente.

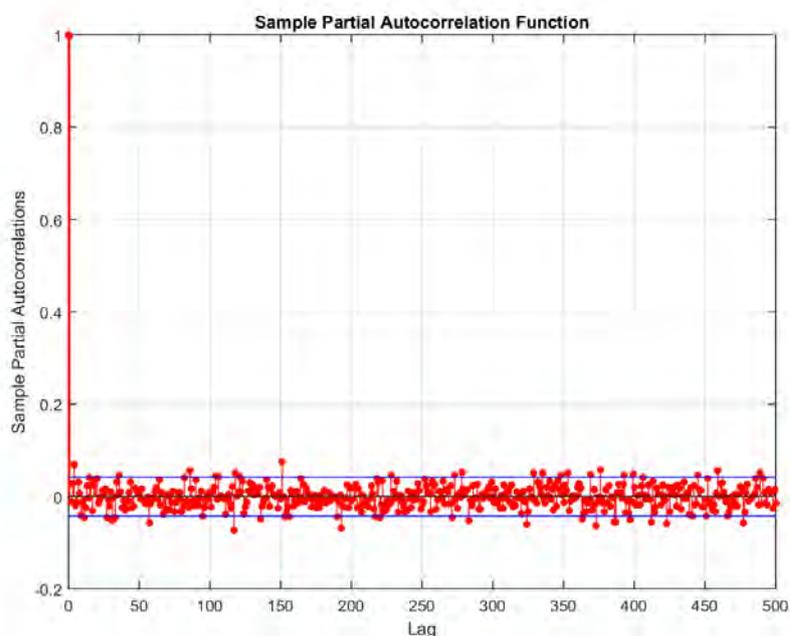
Os valores estatísticos calculados para a série temporal VALE5 na frequência diária pode ser vista na tabela 6.

Figura 20 – Gráfico de autocorrelação dos dados da VALE5 na frequência diária.



Fonte: O autor.

Figura 21 – Gráfico de autocorrelação parcial dos dados da VALE5 na frequência diária.



Fonte: O autor.

#### 4.1.5 Análise do ativo VALE5 na frequência semanal.

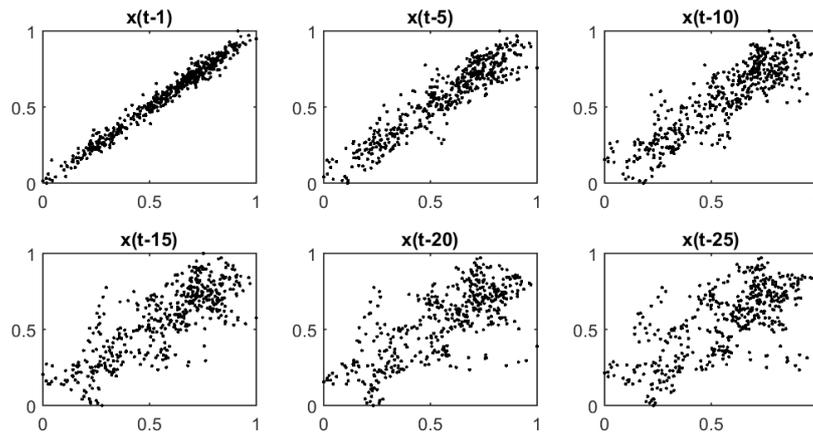
Para o ativo VALE5 com dados de fechamento na frequência semanal o *lagplot* com atrasos de baixa ordem (1 a 25) é mostrado na figura 22. A figura 23 expõe os gráficos

Tabela 6 – Estatísticas VALE5 na frequência diária.

VALE5	VRT	HP	MAI
D1	0,00	0,997	0,665224

Fonte: O autor.

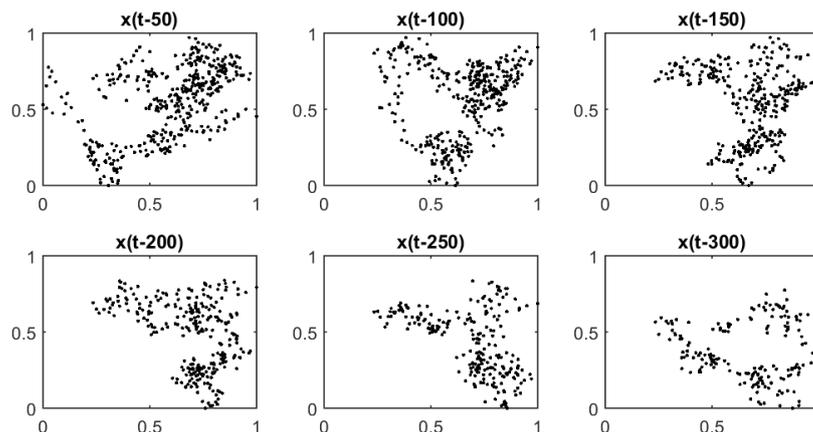
Figura 22 – Lagplot de baixa ordem da série temporal VALE5 na frequência semanal.



Fonte: O autor.

dos retardos de alta ordem, entre (50 a 300).

Figura 23 – Lagplot de alta ordem da série temporal VALE5 na frequência semanal.

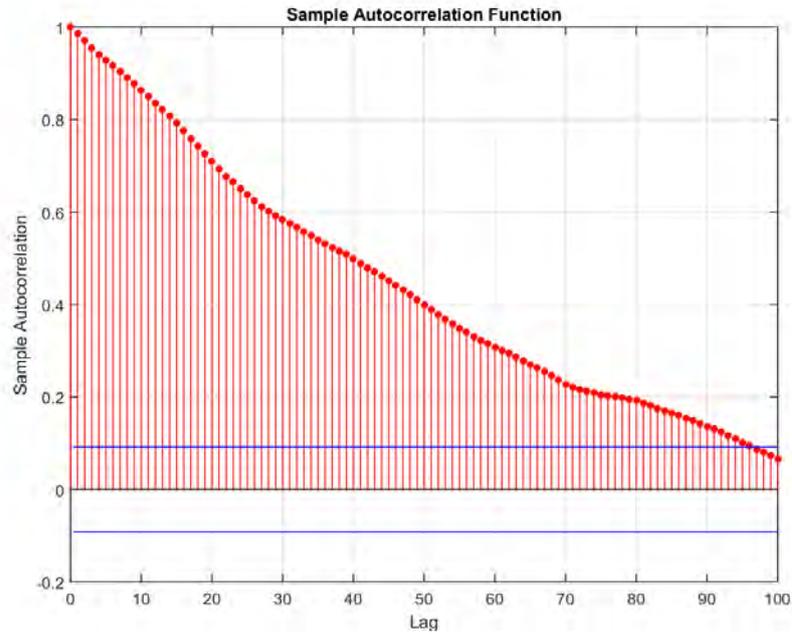


Fonte: O autor.

Os gráficos de ACF e PACF são ilustradas nas figura 24, figura 25, respectivamente.

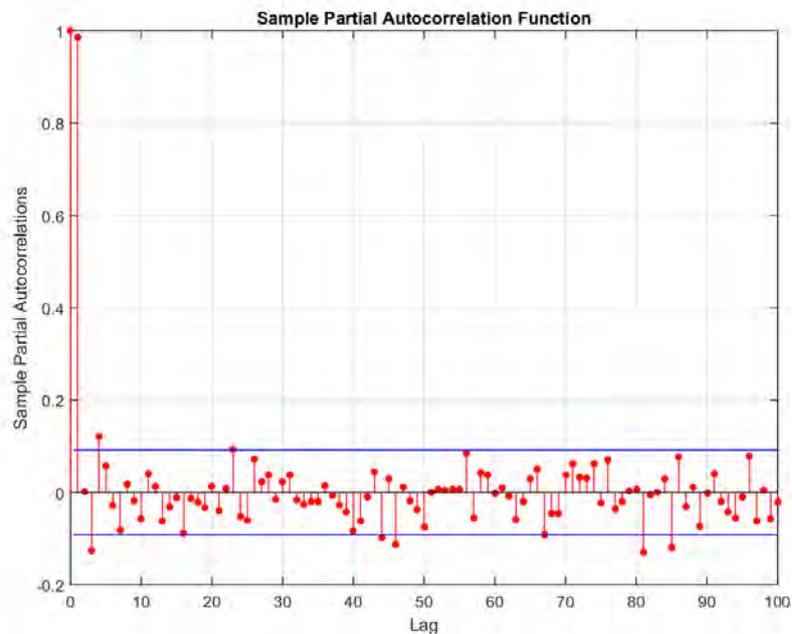
Os valores estatísticos calculados para a série temporal VALE5 na frequência semanal pode ser vista na tabela 7.

Figura 24 – Gráfico de autocorrelação dos dados da VALE5 na frequência semanal.



Fonte: O autor.

Figura 25 – Gráfico de autocorrelação parcial dos dados da VALE5 na frequência semanal.



Fonte: O autor.

#### 4.1.6 Análise do ativo VALE5 na frequências quinzenal

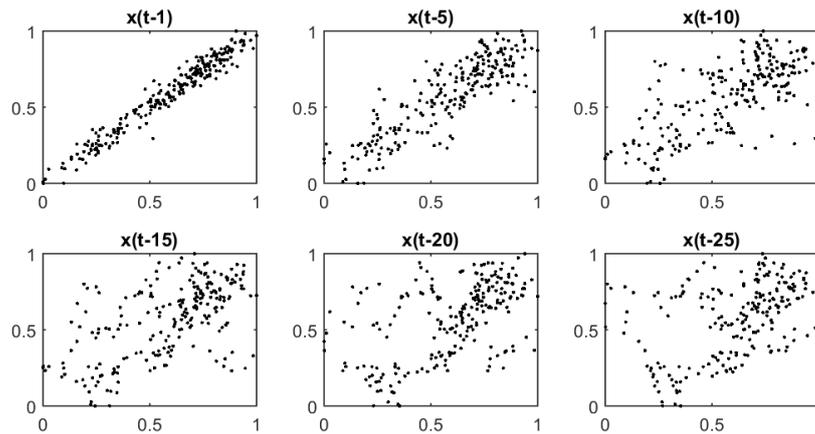
Para o ativo VALE5 com dados de fechamento na frequência quinzenal temos o *lagplot* com atrasos de baixa ordem (1 a 25) conforme mostra a figura 26.

Tabela 7 – Estatísticas VALE5 na frequência semanal.

VALE5	VRT	HP	MAI
W1	0,00	1,0	0,620915

Fonte: O autor.

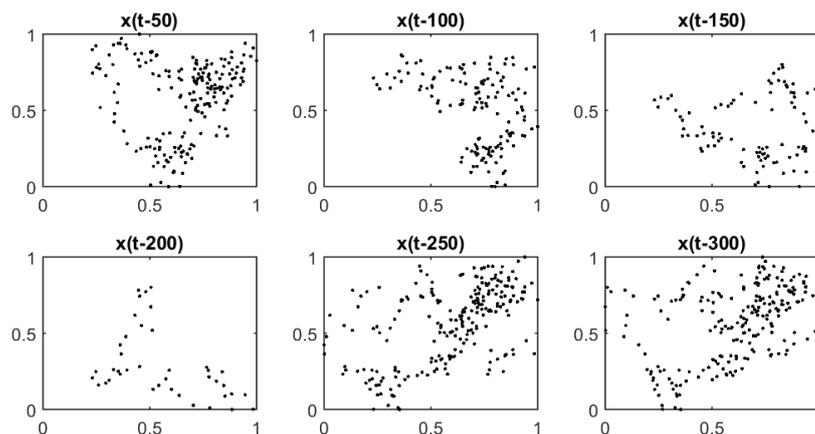
Figura 26 – Lagplot de baixa ordem da série temporal VALE5 na frequência quinzenal.



Fonte: O autor.

A figura 27 expões os gráficos dos retardos de alta ordem, entre (50 a 300).

Figura 27 – Lagplot de alta ordem da série temporal VALE5 na frequência quinzenal.

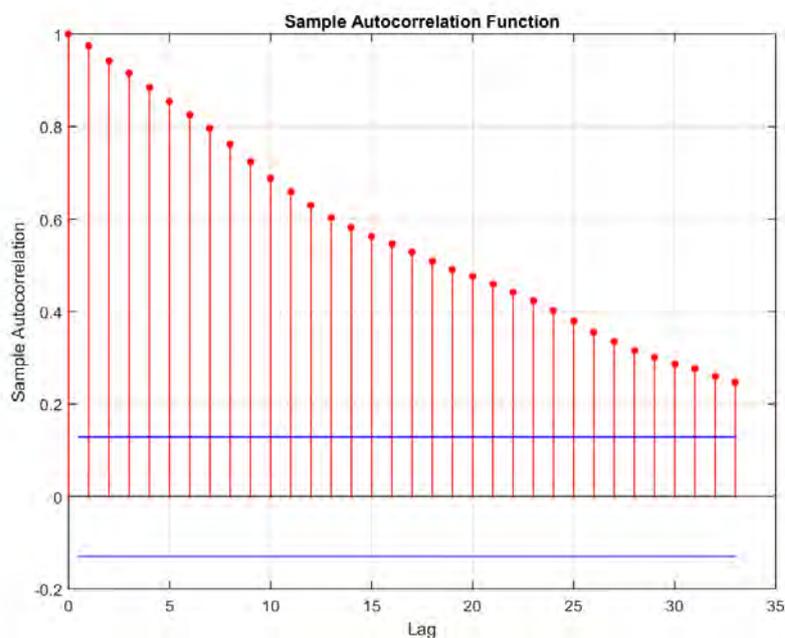


Fonte: O autor.

Os gráficos de ACF e PACF são ilustradas nas figura 28, figura 29, respectivamente.

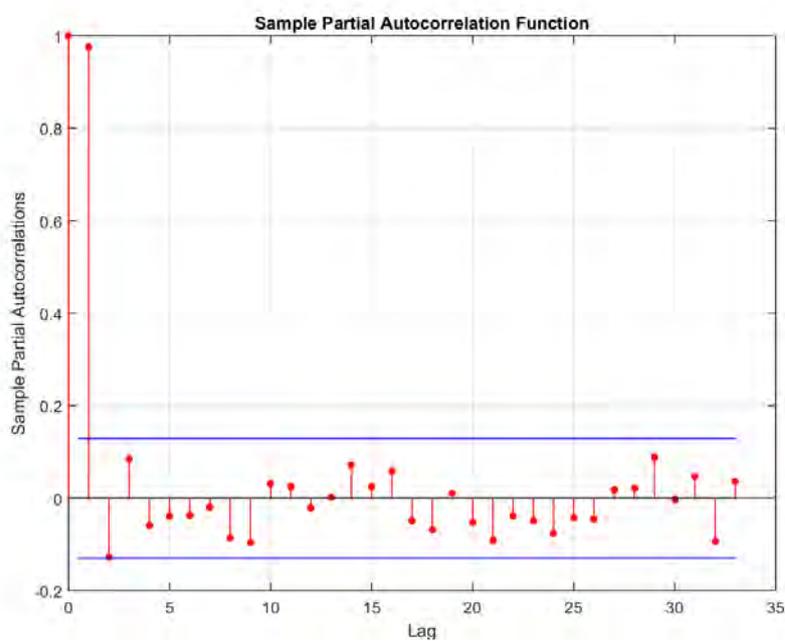
Os valores estatísticos calculados para a série temporal VALE5 na frequência quinzenal pode ser vista na tabela 8.

Figura 28 – Gráfico de autocorrelação dos dados da VALE5 na frequência quinzenal.



Fonte: O autor.

Figura 29 – Gráfico de autocorrelação parcial dos dados da VALE5 na frequência quinzenal.



Fonte: O autor.

Tabela 8 – Estatísticas VALE5 na frequência quinzenal.

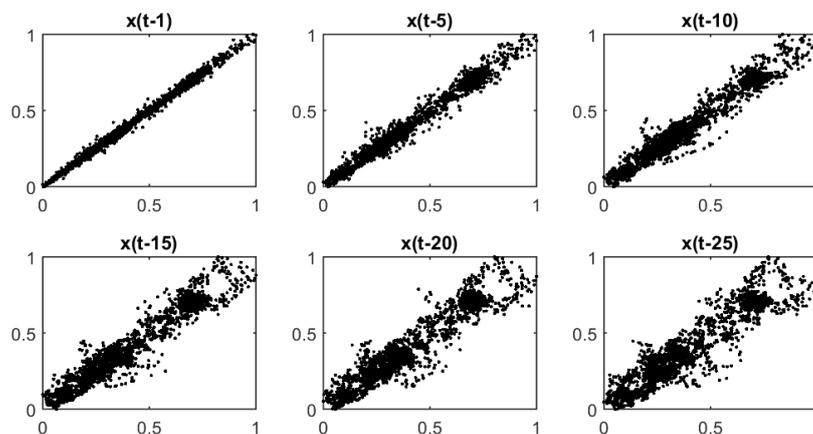
VALE5	VRT	HP	MAI
D15	0,00	0,930	0,569510

Fonte: O autor.

#### 4.1.7 Análise do ativo CSNA3 na frequência diária

Para o ativo Ações Ordinárias da Empresa Companhia Siderúrgica Nacional (CSNA3) com dados de fechamento na frequência diária temos o *lagplot* com atrasos de baixa ordem (1 a 25) conforme mostra a figura 30.

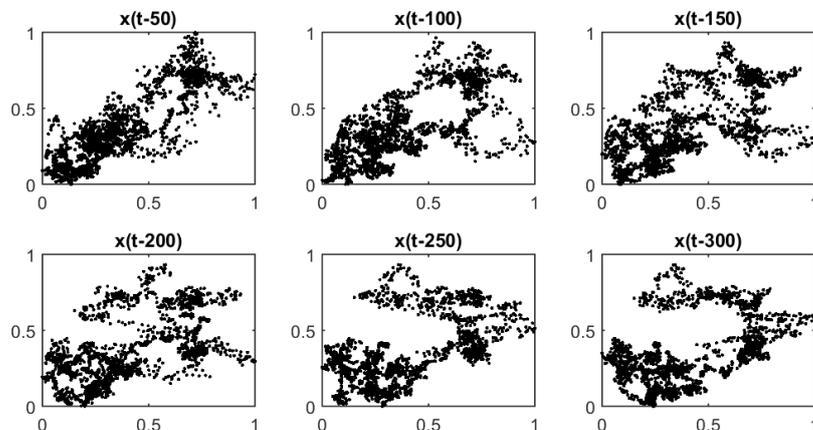
Figura 30 – Lagplot de baixa ordem da série temporal CSNA3 na frequência diária.



Fonte: O autor.

A figura 31 expõe os gráficos dos retardos de alta ordem, entre (50 a 300).

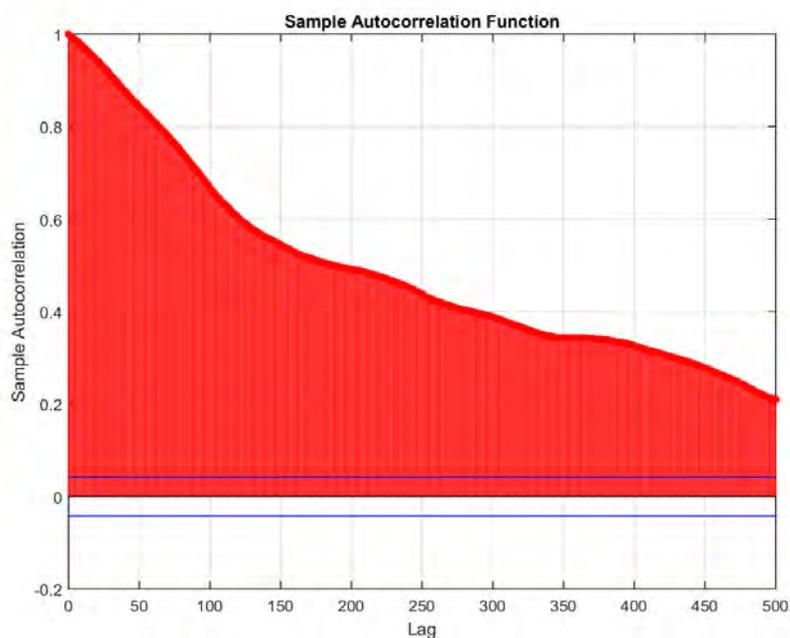
Figura 31 – Lagplot de alta ordem da série temporal CSNA3 na frequência diária.



Fonte: O autor.

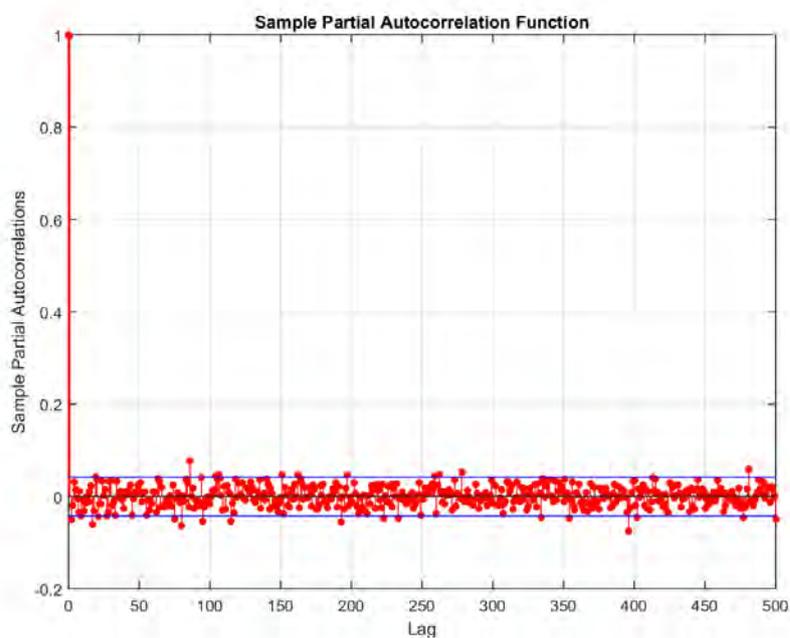
Os gráficos de ACF e PACF são ilustradas nas figura 32, figura 33, respectivamente.

Figura 32 – Gráfico de autocorrelação dos dados da CSNA3 na frequência diária.



Fonte: O autor.

Figura 33 – Gráfico de autocorrelação parcial dos dados da CSNA3 na frequência diária.



Fonte: O autor.

Os valores estatísticos calculados para a série temporal CSNA3 na frequência diária pode ser vista na tabela 9.

Tabela 9 – Estatísticas CSNA3 na frequência diária.

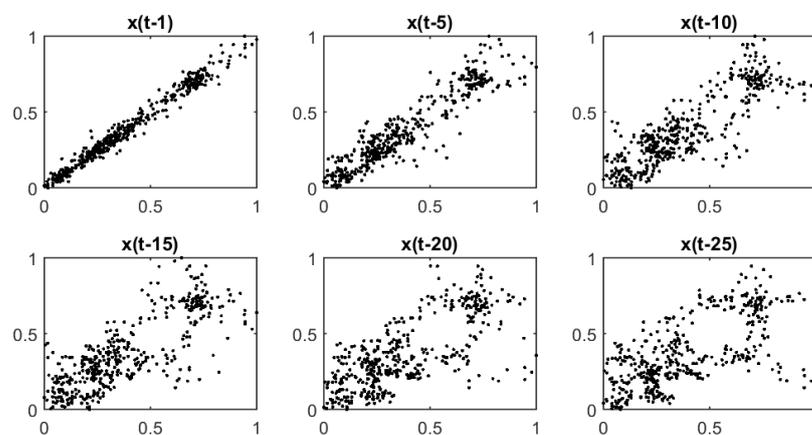
CSNA3	VRT	HP	MAI
D1	0,00	0,985	0,467339

Fonte: O autor.

#### 4.1.8 Análise do ativo CSNA3 na frequência semanal

Para o ativo CSNA3 com dados de fechamento na frequência semanal o *lagplot* com atrasos de baixa ordem (1 a 25) é mostrado na figura 34.

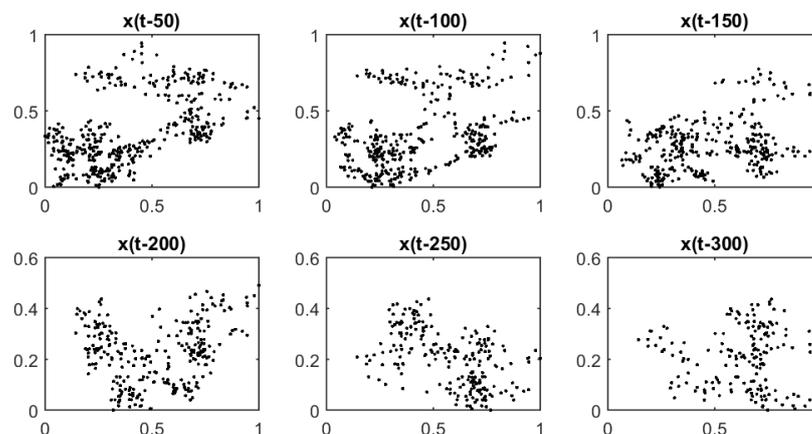
Figura 34 – Lagplot de baixa ordem da série temporal CSNA3 na frequência semanal.



Fonte: O autor.

A figura 35 expõe os gráficos dos retardos de alta ordem, entre (50 a 300).

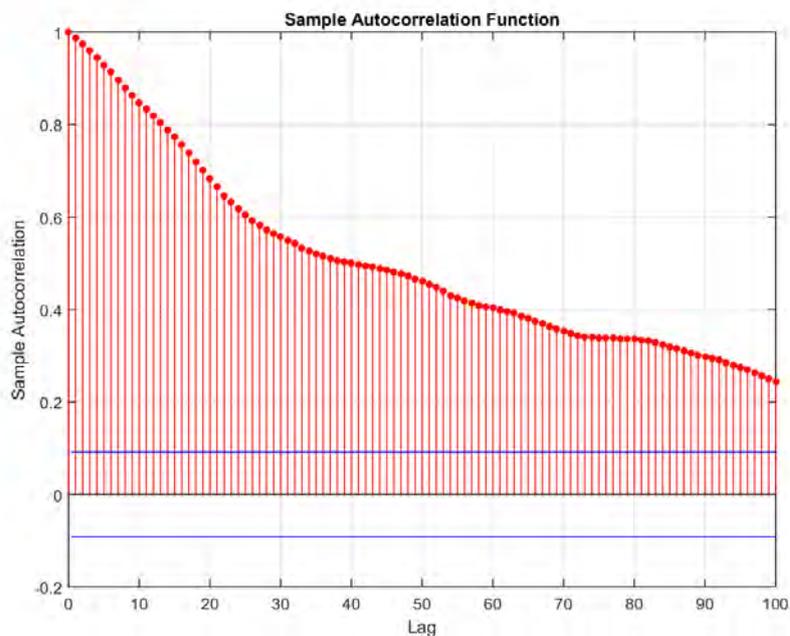
Figura 35 – Lagplot de alta ordem da série temporal CSNA3 na frequência semanal.



Fonte: O autor.

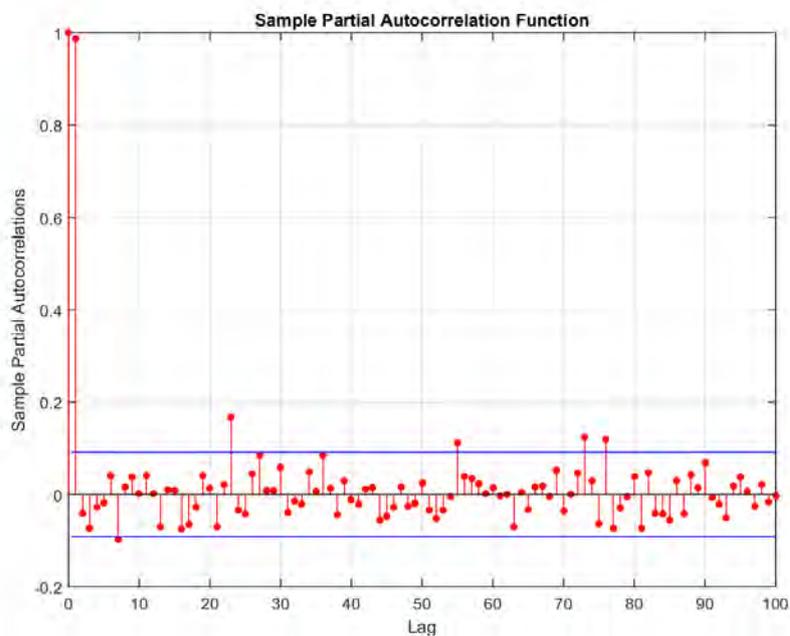
Os gráficos de ACF e PACF são ilustradas nas figura 36, figura 37, respectivamente.

Figura 36 – Gráfico de autocorrelação dos dados da CSNA3 na frequência semanal.



Fonte: O autor.

Figura 37 – Gráfico de autocorrelação parcial dos dados da CSNA3 na frequência semanal.



Fonte: O autor.

Os valores estatísticos calculados para a série temporal CSNA3 na frequência semanal pode ser vista na tabela 10.

Tabela 10 – Estatísticas CSNA3 na frequência semanal.

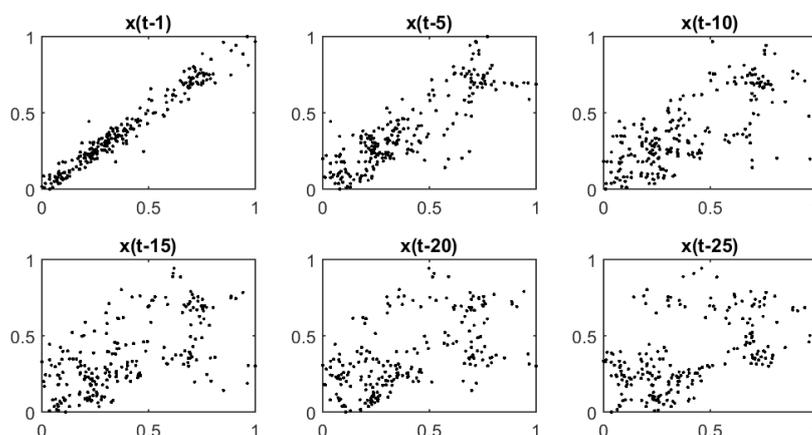
CSNA3	VRT	HP	MAI
W1	0,00	1,0	0,566508

Fonte: O autor.

#### 4.1.9 Análise do ativo CSNA3 na frequências quinzenal

Para o ativo CSNA3 com dados de fechamento na frequência quinzenal temos o *lagplot* com atrasos de baixa ordem (1 a 25) conforme mostra a figura 38.

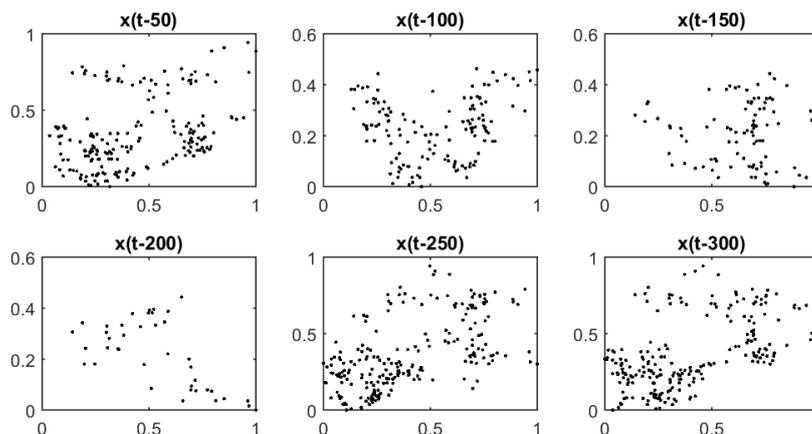
Figura 38 – Lagplot de baixa ordem da série temporal CSNA3 na frequência quinzenal.



Fonte: O autor.

A figura 39 expões os gráficos dos retardos de alta ordem, entre (50 a 300).

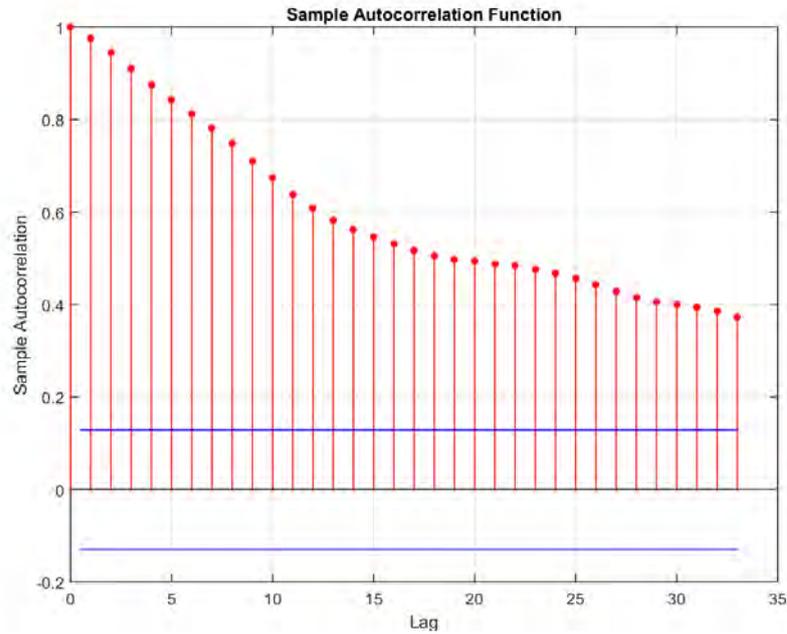
Figura 39 – Lagplot de alta ordem da série temporal CSNA3 na frequência quinzenal.



Fonte: O autor.

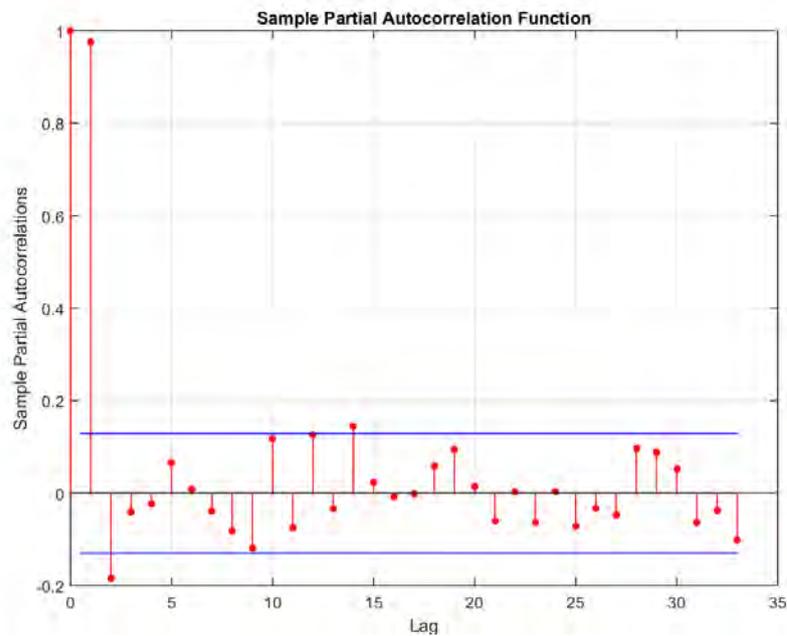
Os gráficos de ACF e PACF são ilustradas nas figura 40, figura 41, respectivamente.

Figura 40 – Gráfico de autocorrelação dos dados da CSNA3 na frequência quinzenal.



**Fonte:** O autor.

Figura 41 – Gráfico de autocorrelação parcial dos dados da CSNA3 na frequência quinzenal.



**Fonte:** O autor.

Os valores estatísticos calculados para a série temporal CSNA3 na frequência quinzenal pode ser vista na tabela 11.

Tabela 11 – Estatísticas CSNA3 na frequência quinzenal.

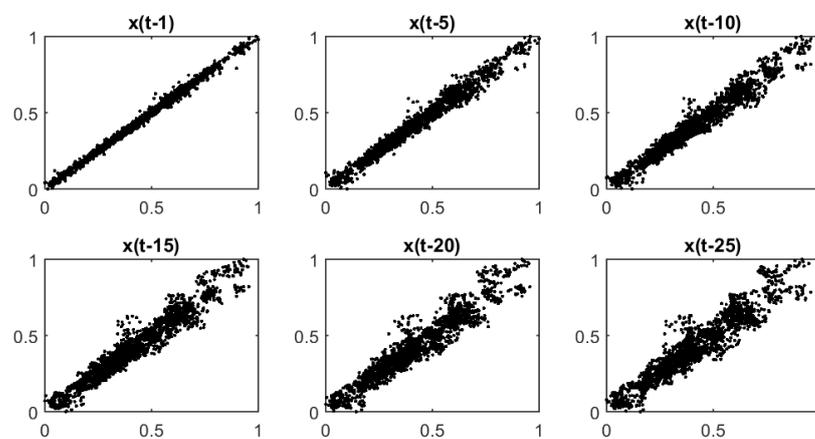
CSNA3	VRT	HP	MAI
D15	0,00	0,988	0,715684

Fonte: O autor.

#### 4.1.10 Análise do ativo BBDC4 na frequência diária

Para o ativo Banco Bradesco SA (BBDC4) com dados de fechamento na frequência diária temos o *lagplot* com atrasos de baixa ordem (1 a 25) conforme mostra a figura 42.

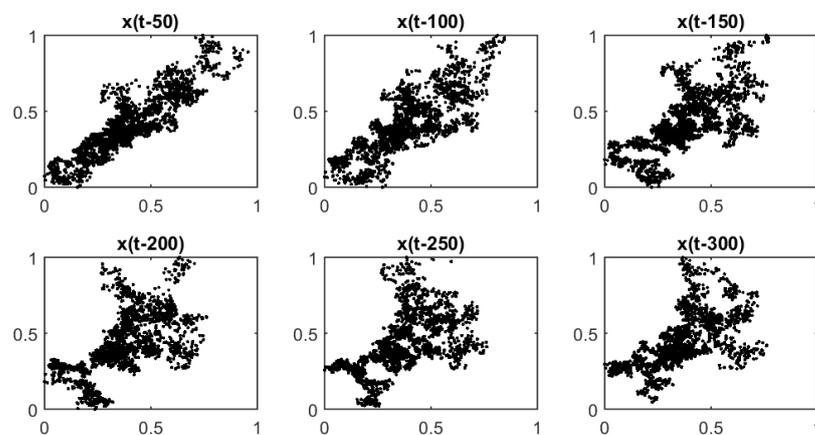
Figura 42 – Lagplot de baixa ordem da série temporal BBDC4 na frequência diária.



Fonte: O autor.

A figura 43 expõe os gráficos dos retardos de alta ordem, entre (50 a 300).

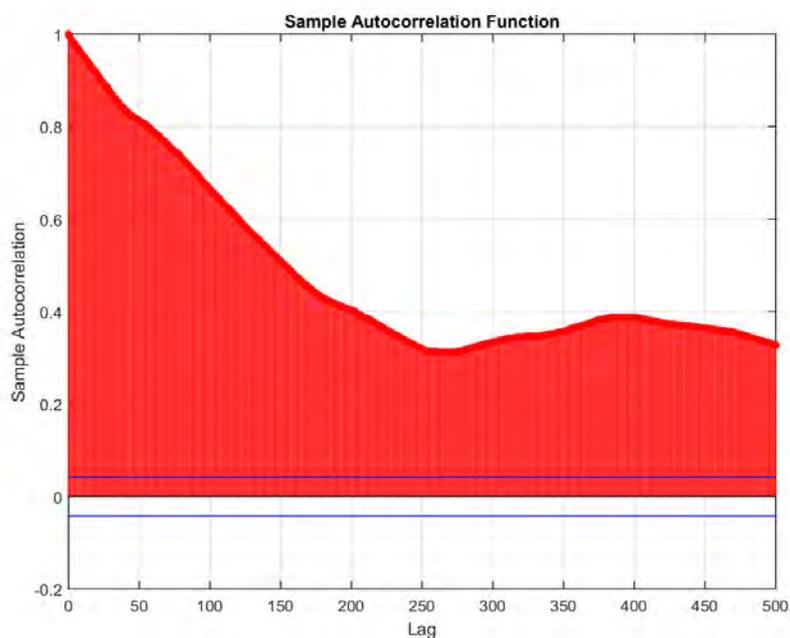
Figura 43 – Lagplot de alta ordem da série temporal BBDC4 na frequência diária.



Fonte: O autor.

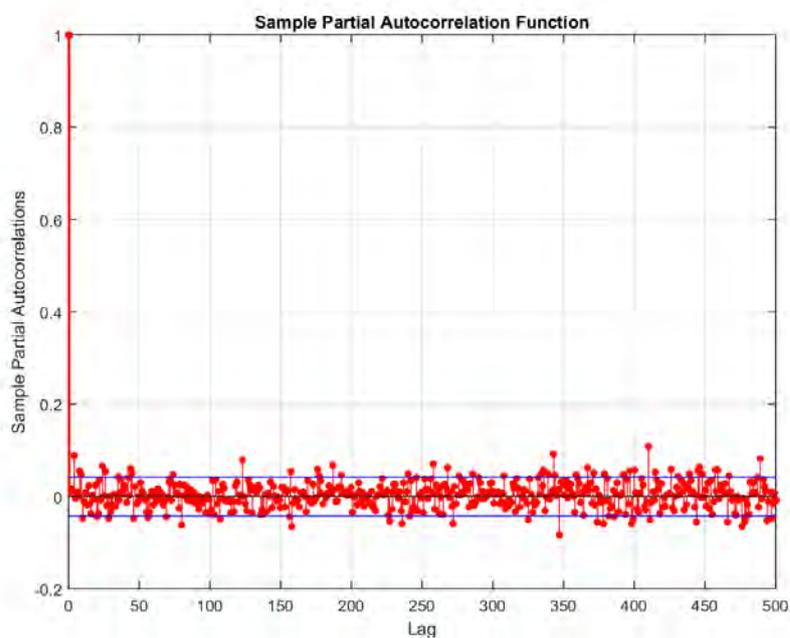
Os gráficos de ACF e PACF são ilustradas nas figura 44, figura 45, respectivamente.

Figura 44 – Gráfico de autocorrelação dos dados da BBDC4 na frequência diária.



Fonte: O autor.

Figura 45 – Gráfico de autocorrelação parcial dos dados da BBDC4 na frequência diária.



Fonte: O autor.

Os valores estatísticos calculados para a série temporal BBDC4 na frequência diária pode ser vista na tabela 12.

Tabela 12 – Estatísticas BBDC4 na frequência diária.

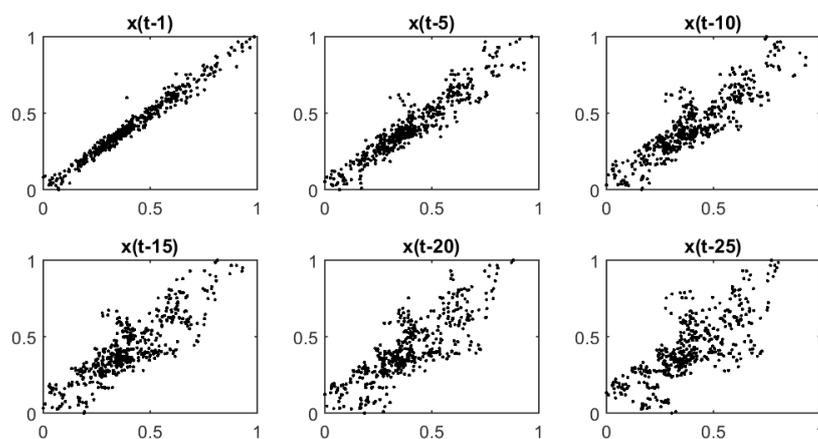
BBDC4	VRT	HP	MAI
D1	0,00	1,0	0,958024

Fonte: O autor.

#### 4.1.11 Análise do ativo BBDC4 na frequência semanal

Para o ativo BBDC4 com dados de fechamento na frequência semanal o *lagplot* com atrasos de baixa ordem (1 a 25) é mostrado na figura 46.

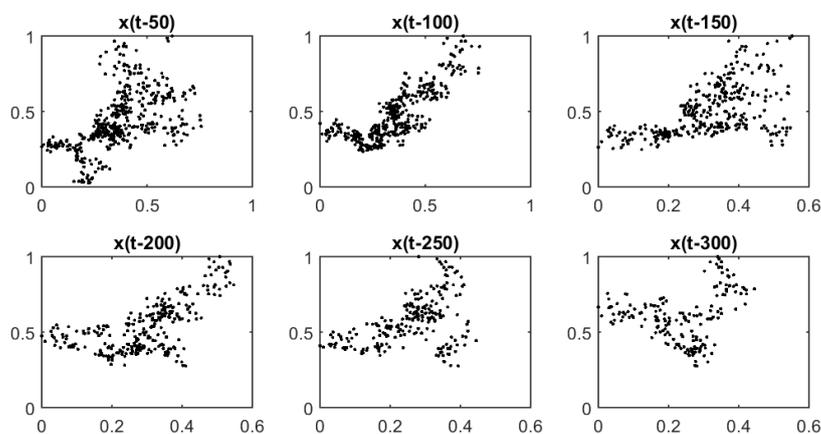
Figura 46 – Lagplot de baixa ordem da série temporal BBDC4 na frequência semanal.



Fonte: O autor.

A figura 47 expõe os gráficos dos retardos de alta ordem, entre (50 a 300).

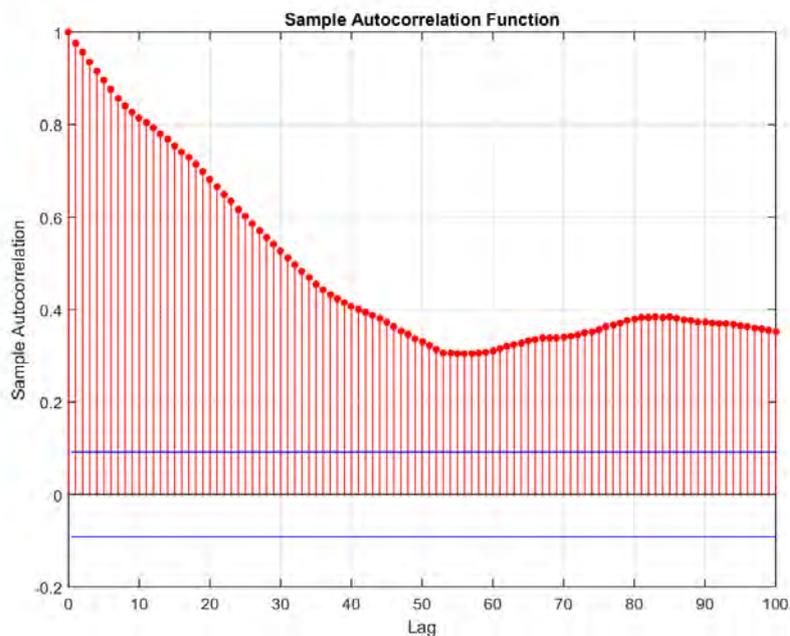
Figura 47 – Lagplot de alta ordem da série temporal BBDC4 na frequência semanal.



Fonte: O autor.

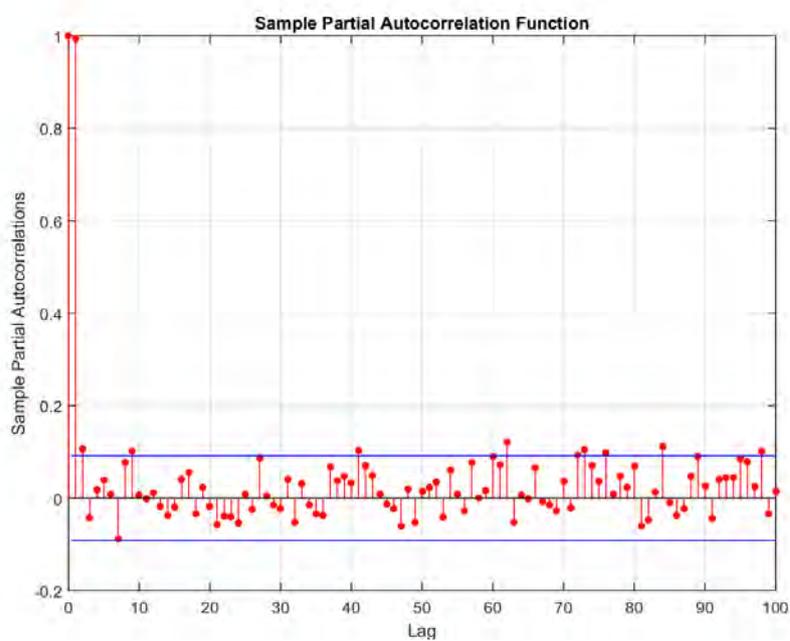
Os gráficos de ACF e PACF são ilustradas nas figura 48, figura 49, respectivamente.

Figura 48 – Gráfico de autocorrelação dos dados da BBDC4 na frequência semanal.



Fonte: O autor.

Figura 49 – Gráfico de autocorrelação parcial dos dados da BBDC4 na frequência semanal.



Fonte: O autor.

Os valores estatísticos calculados para a série temporal BBDC4 na frequência semanal pode ser vista na tabela 13.

Tabela 13 – Estatísticas BBDC4 na frequência semanal.

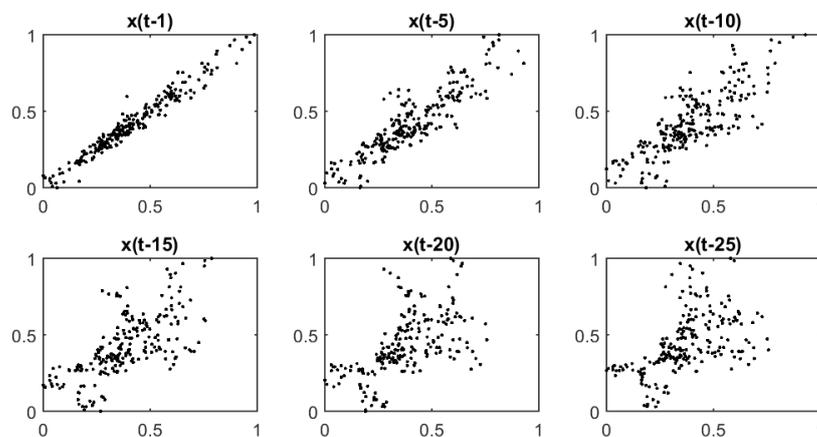
BBDC4	VRT	HP	MAI
W1	0,00	1,0	0,928732

Fonte: O autor.

#### 4.1.12 Análise do ativo BBDC4 na frequências quinzenal

Para o ativo BBDC4 com dados de fechamento na frequência quinzenal temos o *lagplot* com atrasos de baixa ordem (1 a 25) conforme mostra a figura 50. A figura 51 expõe os

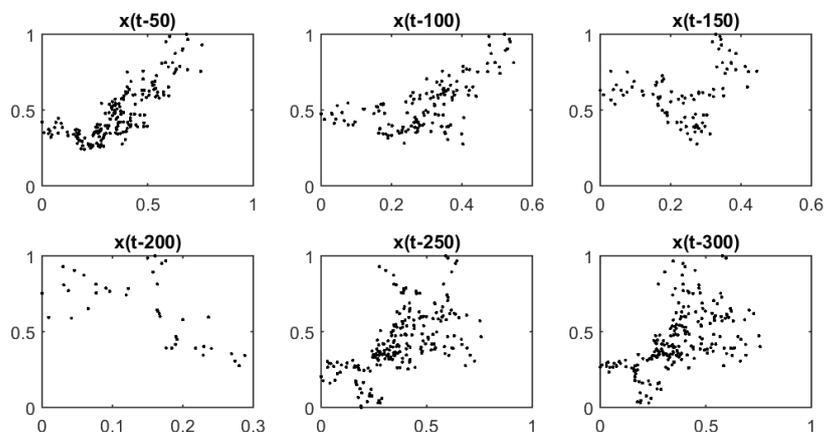
Figura 50 – Lagplot de baixa ordem da série temporal BBDC4 na frequência quinzenal.



Fonte: O autor.

gráficos dos retardos de alta ordem, entre (50 a 300).

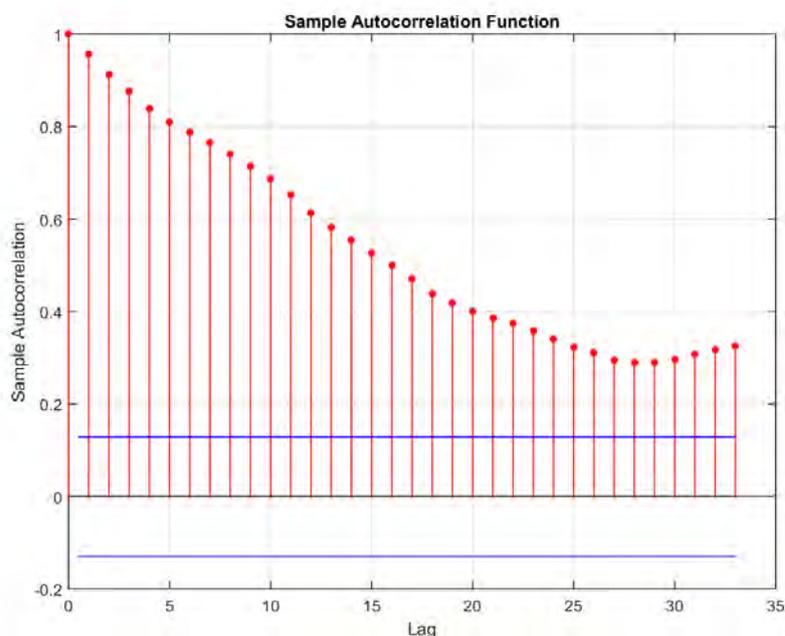
Figura 51 – Lagplot de alta ordem da série temporal BBDC4 na frequência quinzenal.



Fonte: O autor.

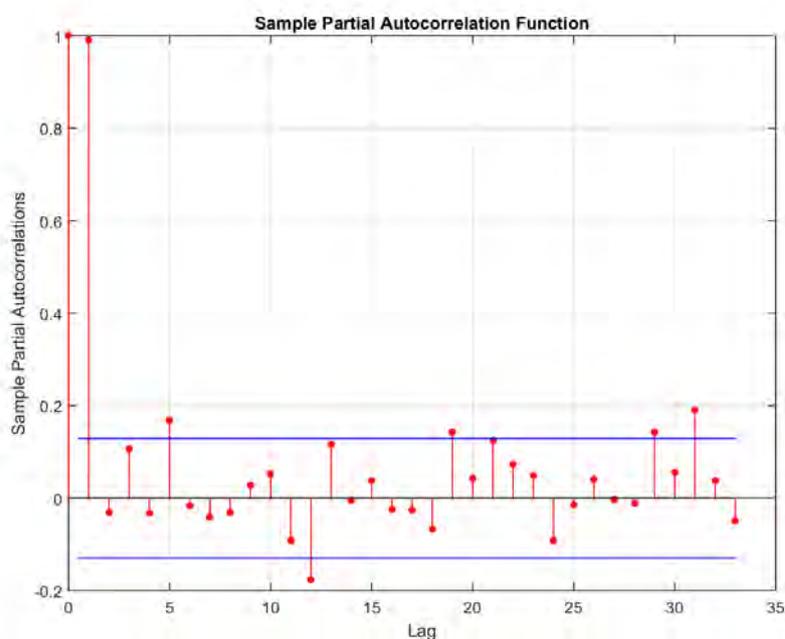
Os gráficos de ACF e PACF são ilustradas nas figura 52, figura 53, respectivamente.

Figura 52 – Gráfico de autocorrelação dos dados da BBDC4 na frequência quinzenal.



Fonte: O autor.

Figura 53 – Gráfico de autocorrelação parcial dos dados da BBDC4 na frequência quinzenal.



Fonte: O autor.

Os valores estatísticos calculados para a série temporal BBDC4 na frequência quinzenal pode ser vista na tabela 14.

Tabela 14 – Estatísticas BBDC4 na frequência quinzenal.

BBDC4	VRT	HP	MAI
D15	0,00	1,0	0,627652

Fonte: O autor.

#### 4.1.13 Considerações analíticas sobre as séries estudadas

Considerando que o principal problema na caracterização do fenômeno gerador de uma série temporal é a escolha correta de seus retardos temporais, pode-se observar que a análise dos *lagplots* de todas as séries temporais financeiras deste estudo evidencia um comportamento crescente, com uma componente linear dominante de curto prazo, menos visualizada quando os retardos são muito antigos. A presença de uma componente não linear sub dominante de longo prazo, sendo dificilmente identificada com os retardos muito próximos do valor atual, mas facilmente vista com retardos de alta ordem. Portanto, o relacionamento dominante linear foi caracterizado em retardos temporais de baixa ordem (1 a 25).

Entretanto, com o aumento da dimensionalidade dos retardos temporais (50 a 300) podemos observar o aparecimento de uma estrutura não linear complexa, o que caracteriza o relacionamento sub dominante não linear. Com base em (ARAÚJO, 2016), pode-se afirmar que as séries temporais financeiras estudadas não são construídas por um processo aleatório, mas por um processo gerado a partir de uma combinação balanceada entre componentes dominantes lineares e componentes subdominantes não lineares.

Quando analisado as diferentes frequências é possível observar que, quanto maior a frequência de uma série temporal financeira, menor é o retardo temporal para identificar a componente não linear, sugerindo que esta componente ganha mais força no processo de previsão. Porém os gráficos de ACF e PACF confirmam a dominância da componente linear em todos os casos.

Através dos gráficos de ACF e PACF mostra-se o comportamento da componente linear para todos os ativos analisados. Observa-se um decaimento hiperbólico em todos os *lagplot*, mostrando que há uma forte dependência linear de curto prazo nas séries temporais financeiras, bem como uma baixa correlação entre valores de alta ordem (retardos temporais acima de 50). O *variance ratio test* (Teste de Razão de Variância - Variance Ratio Test (VRT)) encontrado em todas as séries temporais analisadas, mostra que o teste de decisão rejeitou a hipótese nula de um processo aleatório, a 5% de nível de significância. Esses valores reforçam a análise do *lagplot* e confirmam que o fenômeno gerador não tende a um processo de passeio aleatório (Dilema do Passeio Aleatório - Random Walk Dilemma (RWD)).

Com o intuito de provar a existência de uma componente não linear subdominante, foi utilizado primeiramente o AMI com um retardo temporal de 150. Os valores encontrado

para todas as séries mostram que as informações com este atraso ainda são relevâtes para o mapeamento do fenômeno gerador da série. Para especificar melhor tal dependência, foi utilizado o parâmetro de hurst(HP) e, com ele, pôde ser observado que a existência de uma componente não linear de longo prazo, pois todos os valores deram acima de 0,5 e próximos ou igual a 1,0.

Fica evidenciado também nos gráficos *lagplots* analisados a existência de componentes de tendência com comportamento crescente. Isso indica que o fenômeno gerador das séries temporais financeiras em baixa frequência podem ser mapeadas por uma função crescente que consiga absorver os comportamentos das componentes linear e não-linear.

A Componente não-linear crescente pode ser matematicamente representada por  $g_t$  e definida por:

$$g_t = \theta u(t) \quad (4.2)$$

em que  $u(t)$  representa um mapeamento não-linear crescente.

Dessa forma, uma série temporal financeira nas frequências analisadas, pode ser representada por uma combinação balanceada entre uma componente dominante linear, uma componente sub-dominante não linear, e um termo de ruído, que pode ser definida por:

$$x_t = \lambda(x_{t-1}) + (1 - \lambda)g_t + r_t \quad (4.3)$$

a qual  $g_t$  representa o mapeamento não-linear e  $r_t$  representa o termo de ruído. O termo  $x_{t-1}$  representa o valor anterior observado na série temporal.

Após a análise técnica executada em todas as séries temporais foi investigada as características das series temporais em baixa frequência (diária, semanal e quinzenal) e das séries temporais em alta frequência estudadas por Araújo (ARAÚJO, 2016). Foi identificado que o comportamento crescente das séries são semelhantes.

## 4.2 RESUMO DO CAPÍTULO

Este capítulo apresentou uma análise técnica e fundamentada sobre o fenômeno gerador de séries temporais financeiras em baixa frequência (diária, semanal e quinzenal). Foram identificadas algumas evidências que uma série temporal financeira pode ser mapeada através de uma combinação balanceada entre um maior percentual de componente linear identificada nas janelas temporais de curto prazo e uma menor porcentagem para a componente subdominante não-linear identificada nas janelas temporais de longo prazo, ambas com comportamento crescente.

## 5 O MODELO PROPOSTO

Este capítulo descreve o modelo de neurônio híbrido proposto para previsão de séries temporais financeiras em baixas frequências (diária, semanal e quinzenal). Em seguida, um método baseado em gradiente descendente (FLETCHER; POWELL, 1963) (BRAGA; CARVALHO; LUDERMIR, 2000) é apresentado para treinamento do modelo proposto, com base nas ideias do algoritmo de retropropagação do erro. Por fim, é mostrado o método de ajuste de fase utilizado no treinamento da rede neural.

### 5.1 INTRODUÇÃO

A partir do estudo desenvolvido no capítulo 4, pode-se concluir que para um modelo prever uma série temporal financeira de baixa frequência deve ser capaz de estimar, de forma balanceada (calculando seus pesos através de experimentos), as componentes linear e não-linear (com comportamento crescente) da série. Após a conclusão das análises desenvolvidas no capítulo 4, foi observado que as séries temporais financeiras estudadas têm um comportamento linear dominante, levando a escolha de um tipo de Rede Neural que tenha como resposta um operador linear de resposta finita. Dessa forma, é proposto a utilização de um modelo de Rede Neural Artificial com neurônios Perceptron. Porém, após a constatação do comportamento não linear subdominante ser crescente, fez-se necessário uma modificação nesta Rede Neural para buscar a melhor forma de mapear este comportamento.

Analisando os modelos da literatura, foi identificado que os operadores morfológico são fortemente utilizados para mapear o comportamento crescente e decrescente de um operador não-linear.

Partindo desta constatação, é proposto um modelo de previsão para séries temporais financeiras, em baixa-frequência, que atenda a tal requisito. O modelo é definido inicialmente pela criação de um neurônio híbrido, que consistirá de uma combinação balanceada entre operadores não-lineares crescentes e um operador linear, tendo a habilidade de determinar, automaticamente, os percentuais de utilização das suas componentes através da decomposição de erosões e dilatações. Araújo (ARAÚJO; OLIVEIRA; MEIRA, 2015) afirma que a inclusão de operadores morfológicos para otimização do desempenho de uma rede neural artificial para previsão de séries temporais financeira é uma boa escolha. Devido ao fato de os operadores morfológicos de dilatação e erosão possuírem características crescentes (ARAÚJO; OLIVEIRA; MEIRA, 2015), estes operadores não-lineares podem mapear o comportamento crescente das séries. Dessa forma, tais operadores serão utilizados como base para construção do modelo proposto.

Para melhor entender o modelo em questão, faz-se necessário uma breve fundamenta-

ção teórica sobre conceitos basilares.

## 5.2 FUNDAMENTAÇÃO TEÓRICA RELEVANTE

### 5.2.1 Teoria do Reticulado Completo e a Morfologia Matemática

A teoria de reticulados (*lattice theory*, LT) surgiu a partir do trabalho de (BIRKHOFF, 1940), e posteriormente foi melhorada por (BIRKHOFF, 1993) e (BIRKHOFF et al., 1948).

**Definição 5.1** *Um reticulado completo é um conjunto parcialmente ordenado  $(L, \leq)$  se todo subconjunto  $A$  de  $L$  tem um limite inferior (ínfimo,  $\wedge$ ) e um limite superior (supremo,  $\vee$ ).*

De acordo com (BIRKHOFF, 1993), o conjunto dos números reais estendidos  $\mathbb{R}_{\pm\infty}$  e o intervalo unitário  $[0, 1]$  representam exemplos específicos de reticulados completos.

**Definição 5.2** *Seja um conjunto  $A$  um conjunto não vazio, e seja  $\leq$  uma relação binária em  $A$ . Então  $\leq$  é uma ordem parcial se, e somente se, as seguintes propriedades forem válidas ( $\forall a, b, c \in A$ ) (SUSSNER; ESMI, 2011):*

- 1) *Reflexividade:*  $a \leq a, \forall a \in A$ ;
- 2) *Anti-Simetria:*  $a \leq b, b \leq a \Rightarrow a = b \forall a, b \in X$
- 3) *Transitividade:*  $a \leq b, b \leq c \Rightarrow a \leq c \forall a, b, c \in X$ ;

Seja  $A$  um conjunto parcialmente ordenado e não vazio, e seja  $B \subseteq A$ . Define-se que um elemento  $a \in A$  é um limite inferior de  $A$  se  $a \leq b, \forall b \in B$ . É importante ressaltar que  $a$ , é limitante inferior de  $B$ . O maior limitante inferior de  $B$ , quando existir, é definido como ínfimo de  $B$  e denotado por  $\wedge B$ . Semelhante a definição de um limite inferior, pode-se definir o limite superior como sendo  $a_0 \in A$ . Pode-se dizer que  $a_0 \in A$  é o ínfimo de  $B$  se  $l$  satisfaz as seguintes propriedades:

- i)  $l_0$  representa o *limite inferior* de  $b$
- ii)  $l \leq l_0$  para todos os limites inferiores  $l$  de  $B$ .

Similarmente pode ser definida a noção de supremo. É denotado o ínfimo de  $B$ , por  $\wedge B$ , e o supremo por  $\vee B$ . Escreve-se  $a \wedge b$  em vez de  $\wedge\{a, b\}$  e  $a \vee b$  em vez de  $\vee\{a, b\}$ . Um *conjunto parcialmente ordenado*  $A$  é chamado um *reticulado* se, e somente se, existe um *ínfimo* e um *supremo* em  $A$  para todo subconjunto *finito* de  $A$ . Um *reticulado* é considerado completo se todo subconjunto de  $A$  tem um *ínfimo* e um *supremo* em  $A$ . Neste trabalho, denota-se um *reticulado completo* pela letra  $\mathbb{L}$ .

Suponha que  $X$  e  $Y$  são reticulados. Uma função  $f : X \rightarrow Y$  que satisfaz as seguintes equações para todo  $x \in X$  e para todo  $y \in Y$  recebe o nome de *homomorfismo de reticulados*.

$$f(a \wedge b) = f(a) \wedge f(b), \quad (5.1)$$

$$f(a \vee b) = f(a) \vee f(b). \quad (5.2)$$

Por definição, um *isomorfismo de reticulados* é um homomorfismo bijetivo dos reticulados.

Portanto, será aplicada a morfologia matemática sobre os reticulados completos  $\wp(X)$ , em que  $X$  denota o espaço euclidiano  $\mathbb{R}^d$  ou o espaço  $\mathbb{Z}^d$ . De acordo com (SILVA, 2007), as operações elementares da morfologia matemática são: *erosão*, *dilatação*, *anti-erosão*, *anti-dilatação*. Sejam  $\varepsilon$ ,  $\delta$ ,  $\bar{\varepsilon}$ ,  $\bar{\delta}$  operadores do reticulado completo  $\mathbb{L}$  para o reticulado completo  $\mathbb{M}$ , e seja  $Y \subseteq \mathbb{L}$ .

$$\varepsilon \text{ é a erosão} \iff \varepsilon(\bigwedge Y) = \bigwedge_{y \in Y} \varepsilon(y); \quad (5.3)$$

$$\delta \text{ é a dilatação} \iff \delta(\bigvee Y) = \bigvee_{y \in Y} \delta(y); \quad (5.4)$$

$$\bar{\varepsilon} \text{ é a anti-erosão} \iff \bar{\varepsilon}(\bigwedge Y) = \bigvee_{y \in Y} \bar{\varepsilon}(y); \quad (5.5)$$

$$\bar{\delta} \text{ é a anti-dilatação} \iff \bar{\delta}(\bigvee Y) = \bigwedge_{y \in Y} \bar{\delta}(y); \quad (5.6)$$

Sejam  $A$  e  $B$  reticulados. Um operador  $\psi : A \rightarrow B$  é chamado de crescente (*increasing*) se, e somente se, a seguinte condição for verdadeira ( $\forall a, b \in A$ ) (MATHERON, 1975):

$$a \leq b \Rightarrow \Psi(a) \leq \Psi(b), \quad (5.7)$$

A morfologia matemática é uma teoria baseada na álgebra, topologia e geometria integral, tratando a decomposição de transformações entre reticulados completos em termos das operações elementares como um ponto central.

Neste cenário, é importante mostrar o teorema de Banon e Barrera (MATHERON, 1975), que define um mapeamento crescente  $\psi : \mathbb{L} \rightarrow \mathbb{M}$ , onde  $\mathbb{L}$  e  $\mathbb{M}$ , para qualquer mapeamento entre reticulados completos, um conjunto de mapeamentos sup-geradores derivados são menores ou iguais ao seu núcleo e pode ser decomposto como supremo de erosões ou um ínfimo de dilatações.

Considerando que todo reticulado completo é limitado (BIRKHOFF, 1993), bem como as definições 5.1 e 5.2, tem-se o seguinte teorema:

Seja  $\psi : A \rightarrow B$  um mapeamento crescente entre reticulados completos  $A$  e  $B$  (MATHERON, 1975), então:

1) existem dilatações  $\delta^i$  para um conjunto de índices  $I$ , de forma que

$$\Psi = \bigwedge_{i \in I} \delta^i, \quad (5.8)$$

2) existem erosões  $\epsilon^i$  para um conjunto de índices  $I$ , de forma que

$$\Psi = \bigvee_{i \in I} \epsilon^i, \quad (5.9)$$

Este teorema será utilizado como base para a construção do modelo proposto, pois segundo (ARAÚJO, 2011b) o relacionamento não-linear das séries temporais financeiras pode ser estimado a partir de operadores crescentes. A partir do teorema, pode-se analisar os reticulados completos como vetores  $a^i$  e  $b^j \in \mathbb{R}^n$  para um conjunto finito de índices  $i$  e  $j$  diferentes de 0, e então assumir que um operador crescente  $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$  pode ser aproximado por:

$$\Psi \approx \bigwedge_{i \in I} \delta^i, \quad (5.10)$$

e

$$\Psi \approx \bigvee_{i \in I} \epsilon^i, \quad (5.11)$$

Para auxiliar no processo de aproximação faz-se o uso da Álgebra Minimax (CUNINGHAME-GREEN, 2012), modelada a partir de 1950 por diferentes pesquisadores, que definiram a estrutura algébrica formada pelo conjunto  $\mathbb{R}$  munido das operações binárias de máximo, mínimo e adição (BIRKHOFF, 1993). Tal conceito foi utilizado para a criação do novo modelo de neurônio, onde se utiliza a Álgebra Minimax para calcular os produtos máximo e mínimos de matrizes por vetores que assumem valores no BLOG (*bounded lattice ordered group*). Define-se um BLOG como um reticulado limitado  $A$  cujo conjunto de elementos finitos  $B$  forma um grupo definido por  $A \cup \{-\infty, +\infty\}$  cujo os termos  $-\infty, +\infty$  denotam  $\wedge A$  (o menor elemento de  $A$ ) e  $\vee A$  (o maior elemento de  $A$ ) respectivamente. O BLOG tem a seguinte tupla  $(\mathbb{G}, \vee, \wedge, +, +')$ , onde  $\mathbb{G} = \bar{\mathbb{R}}$ , no qual  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$  ou  $\mathbb{G} = \bar{\mathbb{Z}}$ , que  $\bar{\mathbb{Z}} = \mathbb{Z} \cup \{-\infty, +\infty\}$ . Dadas duas matrizes  $A \in \bar{\mathbb{G}}^{m \times n}$  e  $B \in \bar{\mathbb{G}}_{\pm\infty}^{p \times n}$ , o *max-product*  $C = A \vee B$  resultante do produto entre elas é dado por:

$$c_{ij} = \bigvee_{k=1}^p (a_{ik} + b_{kj}), \quad (5.12)$$

E o *min-product*  $D = A \wedge B$  resultante do produto entre elas é dado pela função:

$$d_{ij} = \bigwedge_{k=1}^p (a_{ik} + b_{kj}), \quad (5.13)$$

Partindo do princípio da existência de uma série temporal  $X$  contendo  $n$  valores espaçados por uma mesma constante temporal, com a série tendo como mapeamento uma

função crescente, então pode-se mostrar que  $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ , em que  $n$  representa o número de atrasos no tempo anteriores:

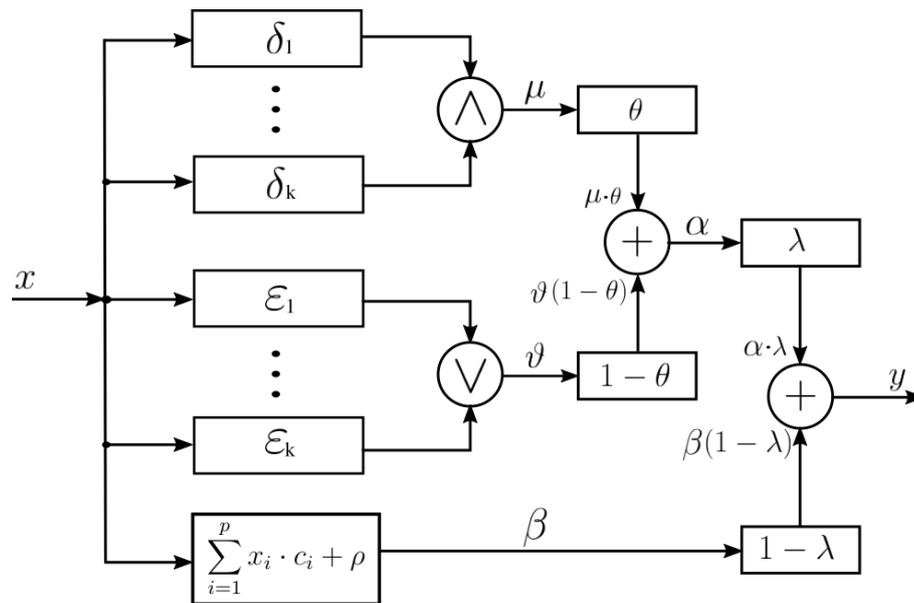
$$(x_{s-n}, \dots, x_{s-1}) \leq (x_{t-n}, \dots, x_{t-1}) \Rightarrow x_s \leq x_t \quad . \quad (5.14)$$

### 5.3 DEFINIÇÃO DO MODELO GIMP

Após observar as proposições de (MATHERON, 1975), e a proposição de um teorema mostrando que qualquer transformação  $\psi : \mathbb{L} \rightarrow \mathbb{M}$  pode ser decomposta como um supremo de ínfimos (e também como um ínfimo de supremos), nas teorias da decomposição do reticulados completos e da Álgebra Minimax, uma função crescente  $\psi : \mathbb{L} \rightarrow \mathbb{M}$  pode ser aproximada utilizando o modelo morfológico GIMP figura 54.

Portanto, o modelo proposto utiliza a morfologia matemática para calcular a componente subdominante não-linear, componente identificada no estudo desenvolvido no capítulo 4.

Figura 54 – Arquitetura do modelo GIMP.



Fonte: o autor.

De acordo com a figura 54, o modelo escolhe a melhor aproximação entre um conjunto de erosões morfológicas ou conjunto de dilatações morfológicas (saída não linear), e fará uma combinação com um modelo morfológico linear (saída linear) e, posteriormente, essas saídas serão combinadas para criar a função de aproximação do modelo proposto.

O modelo GIMP é descrito da seguinte forma:

$$y = \lambda\alpha + (1 - \alpha)\beta, \quad \lambda \in [0, 1], \quad (5.15)$$

onde

$$\alpha = \theta\mu + v(1 - \theta), \quad \theta \in [0, 1] \quad (5.16)$$

e

$$\beta = \sum_{i=1}^n x_i \cdot c_i + \rho \quad (5.17)$$

com

$$\mu = \bigwedge \{\mathbf{u}_1, \dots, \mathbf{u}_k\} \quad , \quad (5.18)$$

$$v = \bigvee \{\mathbf{v}_1, \dots, \mathbf{v}_k\} \quad , \quad (5.19)$$

e

$$\mathbf{u}_i = \delta_i(\mathbf{x}) = \bigvee_{j=1}^n (a_j^i + \mathbf{x}_j) \quad (5.20)$$

$$\mathbf{v}_i = \varepsilon_i(\mathbf{x}) = \bigwedge_{j=1}^n (b_j^i + \mathbf{x}_j) \quad (5.21)$$

em que o termo  $n$  denota a dimensionalidade do sinal de entrada  $\mathbf{x}$ , os termos  $\lambda, \theta \in \mathbb{R}$ , e os termos  $a, b$  e  $c \in \mathbb{R}^n$ . O vetor  $c$  são os coeficientes (pesos) relacionado ao cálculo do operador linear.

$\rho$  é o limiar de ativação que especifica qual será o patamar apropriado para que o resultado produzido pelo combinador linear possa gerar um disparo em direção à saída do neurônio. O termo  $\beta$  representa a saída do operador linear de resposta finita ao impulso, resultado da combinação linear e o limiar de ativação, no qual o termo de de combinação é dado por  $\theta$ . Por outro lado, o termo  $\alpha$  representa a combinação linear por aproximação, resultante da decomposição por dilatação e erosão entre os operadores morfológicos. Os termos  $\mu$  e  $v$  representam a resultante da decomposição morfológica por dilatação e a resultante da decomposição morfológica por erosão, respectivamente. Pode-se perceber que a saída  $y$  do modelo *General Increasing Morphological Perceptron* é dada pela combinação linear entre o operador linear e uma combinação linear de operadores não lineares crescentes. O termo de combinação da saída é  $\lambda$ .

É importante ressaltar que este modelo diferencia do modelo de Araújo por dois fatores: a generalização definida por (ARAÚJO, 2016) utiliza apenas um valor de dilatação e um valor para erosão, enquanto o GIMP faz a decomposição de  $n$  valores de dilatação e  $n$  Valores para erosão; e o modelo de Araújo já tenta generalizar operadores crescentes e decrescentes, enquanto o GIMP foca apenas em melhorar o desempenho preditivo de funções crescentes.

### 5.3.1 Fundamentos do Processo de Aprendizagem

Algumas definições devem ser levadas em conta para fundamentar o processo de aprendizagem do modelo proposto GIMP.

**Definição 5.3** A  $r$ -ésima função de posto (função de rank) de um vetor  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$  é dada pelo  $r$ -ésimo elemento do vetor  $\mathbf{x}$ , ordenado de forma crescente ( $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}$ ). Portanto, a  $r$ -ésima função de posto do vetor  $\mathbf{x}$  é definida por (PESSOA; MARAGOS, 2000), (ARAÚJO, 2012b) e (ARAÚJO; SUSSNER, 2010).

$$\mathfrak{R}_r(\mathbf{x}) = x_{(r)}, \quad r = 1, 2, \dots, n. \quad (5.22)$$

**Definição 5.4** A função de impulso unitário,  $q(x)$ , é definida por (PESSOA; MARAGOS, 2000) (ARAÚJO; OLIVEIRA; MEIRA, 2015)

$$q(x) = \begin{cases} 1, & \text{se } x = 0, \\ 0, & \text{caso contrário.} \end{cases} \quad (5.23)$$

na qual  $x \in \mathbb{R}$

Ao se aplicar uma função de impulso unitário a um vetor  $x$ , esta retorna um vetor de funções de impulso unitário ( $Q(x)$ ), definido em (ARAÚJO; OLIVEIRA; MEIRA, 2015) como

$$Q(\mathbf{x}) = [q(x_1), q(x_2), \dots, q(x_n)] \quad \forall \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n, \quad (5.24)$$

**Definição 5.5** O  $r$ -ésimo índice indicador de função  $\mathbf{c}$  de um vetor  $\mathbf{x}$  é definido por (ARAÚJO; OLIVEIRA; MEIRA, 2015)

$$\mathbf{c} = \frac{Q(z \cdot \mathbf{1} - \mathbf{x})}{Q(z \cdot \mathbf{1} - \mathbf{x}) \cdot \mathbf{1}^T}, \quad (5.25)$$

no qual  $z = \mathfrak{R}_r(x)$  e  $\mathbf{1} = (1, 1, \dots, 1)$

**Proposição 5.1** De acordo com (ARAÚJO; OLIVEIRA; MEIRA, 2015) e (ARAÚJO, 2016), o vetor indicador de posto marca as posições dos elementos do vetor  $\mathbf{x}$ , onde ocorre o valor  $z$  e, obrigatoriamente, deve ter uma área unitária. Em outras palavras, o vetor indicador de posto tem o papel de marcar as localizações em  $\mathbf{x}$ , onde  $\mathbf{z} = \mathfrak{R}_r(x)$  ocorrem.

Dessa forma, foi proposto em (ARAÚJO, 2012b) que dado um vetor  $\mathbf{x} \in \mathbb{R}^n$ ,  $r \in 1, 2, \dots, n$  e  $\mathbf{c} = \mathbf{c}(X, r)$ , então:

1.  $\mathbf{c} \cdot \mathbf{1}^T = 1$  (área unitária)
2.  $\mathbf{c} \cdot \mathbf{x}^T = \mathfrak{R}_r(\mathbf{x})$  (representação via produto interno)
3.  $\mathbf{c} \cdot (\mathbf{z} \mathbf{1} - \mathbf{x})^T = 0$  em que  $z = \mathfrak{R}_r(\mathbf{x})$

4. Se  $r$  é fixado, então  $\mathbf{c}$  é uma função constante por partes do vetor  $\mathbf{x}$ , com exatamente  $2^n - 1$  valores distintos possíveis e para todos os pontos  $\mathbf{x}_0 \in \mathbb{R}^n$  dotados com componentes desiguais,  $x_{0,i} \neq x_{0,j} \forall i \neq j$ , é possível verificar uma vizinhança em torno delas, de modo que

$$\|\mathbf{x} - \mathbf{x}_0\|_\infty < \frac{1}{2} \min_{i \neq j} |x_{0,i} - x_{0,j}|, \quad (5.26)$$

em que  $\mathbf{c}$  é constante e dado por  $\mathbf{c}(\mathbf{x}, r) = \mathbf{c}(\mathbf{x}_0, r)$ .

Outra preposição de (PESSOA; MARAGOS, 2000) é que considera-se para  $r$  fixo,  $c = c(\mathbf{x}, r)$  seja constante em uma certa vizinhança de  $\mathbf{x}_0$ , então a função posto  $\mathfrak{R}_r(\mathbf{x})$  é diferenciável em  $\mathbf{x}_0$  e sua derivada é dada por:

$$\left. \frac{\partial \mathfrak{R}_r(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_0} = \left. \frac{\partial}{\partial \mathbf{x}} (\mathbf{c}(\mathbf{x}_0, r) \cdot \mathbf{x}^T) \right|_{\mathbf{x}=\mathbf{x}_0} = \mathbf{c}(\mathbf{x}_0, r) \quad (5.27)$$

Nos pontos da vizinhança nos quais  $\mathbf{c}$  não é constante, a função posto não é diferenciável. Portanto, se  $\mathbf{x}$  satisfaz as hipóteses da proposição 1(4), tem-se que  $\mathbf{x}$  e  $\mathbf{x}_0$  apresentam a mesma ordem decrescente de suas componentes, e, conseqüentemente, como o posto  $r$  é fixo, o vetor indicador posto  $\mathbf{c}$  de  $\mathbf{x}$  e  $\mathbf{x}_0$  serão os mesmos, pois o vetor indicador posto simplesmente marca as localizações em  $\mathbf{x}$  onde  $z = \mathfrak{R}_r(\mathbf{x})$  ocorrem. Em outras palavras, obtém-se que  $\mathbf{c}(\mathbf{x}, r) = \mathbf{c}(\mathbf{x}_0, r)$ . Neste sentido, considerando  $\mathbf{c}(\mathbf{x}, r)$  constante em uma vizinhança de  $\mathbf{x}_0$ , e usando a representação de produto interno para a função posto  $\mathfrak{R}_r(\mathbf{x})$ , então a  $r$ -ézima função de posto  $\mathfrak{R}_r(\mathbf{x})$  é diferenciável em  $\mathbf{x}_0$  (ARAÚJO, 2012b):

$$\begin{aligned} q_\sigma(x) &= q_\sigma(-x) \text{ (simetria)} \\ q_\sigma(x) &\rightarrow q(x) \quad \forall x \text{ com } \sigma \rightarrow 0, \\ q_\sigma(x) &\rightarrow 1 \quad \forall x \text{ com } \sigma \rightarrow \infty. \end{aligned} \quad (5.28)$$

Caso a função posto não seja diferenciável, é possível assim torná-la, atribuindo ao vetor  $\mathbf{c}$  um valor unilateral da descontinuidade  $\frac{\partial z}{\partial \mathbf{x}}$ . Utilizando este artifício, o vetor indicador de posto pode ser utilizado para estimar o gradiente baseado na descontinuidade da função de impulso unitário. A prova para esta afirmação pode ser encontrada em (ARAÚJO, 2012b). Como a proposta desse modelo foca na utilização dos operadores morfológicos de erosão e dilatação, o principal objetivo é estimar seus gradientes nos pontos de não-diferenciabilidade.

A fim de resolver esta questão, é utilizada uma extensão do vetor indicador de posto para os operadores de dilatação e erosão. Dessa forma, pode-se observar que  $(\mathfrak{R}_r(\mathbf{x} + \mathbf{a}) = \vee \mathbf{x} + \mathbf{a})$  quando  $r = 1$  (dilatação) e  $(\mathfrak{R}_r(\mathbf{x} + \mathbf{b}) = \wedge \mathbf{x} + \mathbf{b})$  quando  $r = n$  (erosão), com  $\mathbf{x}, \mathbf{a}, \mathbf{b} \in \mathfrak{R}^n$ . Entretanto, (ARAÚJO; OLIVEIRA; MEIRA, 2015) e (ARAÚJO, 2016) afirmam que utilizar apenas a extensão do vetor indicador de posto para estimar os gradientes dos

operadores morfológicos de dilatação e erosão podem levar a mudanças abruptas na estimativa do gradiente e, conseqüentemente, comprometer a robustez numérica. (ARAÚJO; OLIVEIRA; MEIRA, 2015)(ARAÚJO, 2016) superou este problema substituindo a função impulso unitário por uma função de impulso suave (*smoothed impulse function*),  $q_\sigma(x)$ , que depende de um fator de escala  $\sigma \geq 0$ . Assim, os operadores de dilatação e erosão podem ser descritos em termos diferenciáveis para o cálculo de seus gradientes. A função de rank suavizada é definida por (PESSOA; MARAGOS, 1998):

**Definição 5.6** *Se  $q_\sigma(x)$  é um impulso suavizado satisfatório, então o  $r$ -ésimo elemento da função rank suavizada é definido por:*

$$R_{r,\sigma}(x) \equiv \mathbf{c}_\sigma \cdot \mathbf{1}^T. \quad (5.29)$$

Araújo afirma, ainda, que a função de impulso suave deve ter as seguintes propriedades:

$$Q_\sigma(x) = [q_\sigma(x_1), q_\sigma(x_2), \dots, q_\sigma(x_n)]. \quad (5.30)$$

Duas funções podem ser utilizadas como funções de impulso suave: a função  $sech^2\left(\frac{x}{\sigma}\right)$  e a função  $exp\left[-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right]$ . Neste trabalho, será utilizada como função de impulso unitário suave a função  $exp\left[-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right]$ , que, de acordo com (ARAÚJO; OLIVEIRA; MEIRA, 2015)(ARAÚJO, 2016), é uma boa escolha para  $q_\sigma(\mathbf{x})$ , e seu comportamento será semelhante à função de impulso unitário de vetor  $Q(x)$ , sendo definida como:

$$Q_\sigma(\mathbf{x}) = (q_\sigma(x_1), q_\sigma(x_2), \dots, q_\sigma(x_i)), \quad (5.31)$$

onde

$$q_\sigma(x_i) = exp\left[\frac{1}{2}\left(\frac{x_i}{\sigma}\right)^2\right], \quad \forall i = 1, \dots, n. \quad (5.32)$$

Baseado na proposição 5.1.3 (2), temos que  $\mathbf{c}_\sigma \cdot x^T = R_{r,\sigma}(x)$  onde:

$$\mathbf{c}_\sigma \cdot x^T = \frac{Q_\sigma(z \cdot \mathbf{1} - x)}{Q_\sigma(z \cdot \mathbf{1} - x) \cdot \mathbf{1}^T}, \quad z = \mathfrak{R}_r(x), \quad (5.33)$$

no qual  $\mathbf{c}_\sigma$  representa o vetor indicador de posto suave. Como a finalidade da utilização da função  $\mathbf{c}_\sigma$  é a aproximação da função  $c$ , Araújo utiliza outra proposição para auxiliar nos cálculos dos gradientes. Proposição 5.1.4. Seja  $x \in \mathbb{R}^n$ ,  $r \in 1, 2, \dots, n$ ,  $\mathbf{c}_\sigma = \mathbf{c}_\sigma(x, r)$ ,  $z = R_{r,\sigma}(x)$ . Temos que:

$$(1) \quad \mathbf{c}_\sigma \cdot \mathbf{1}^T = 1 \text{ (área unitária)}$$

$$(2) \quad \lim_{\sigma \rightarrow 0} \mathbf{c}_\sigma = \mathbf{c}$$

$$(3) \quad \lim_{\sigma \rightarrow 0} \mathbf{z}_\sigma = \mathbf{z}$$

- (4)  $\lim_{\sigma \rightarrow \infty} \mathbf{c}_\sigma = \frac{1}{n}$
- (5)  $\lim_{\sigma \rightarrow \infty} \mathbf{z}_\sigma = \frac{1}{n} \sum_{j=1}^n t_j$ .

### 5.3.2 Proposta do Algoritmo de Treinamento com o Ajuste de Fase Temporal

A utilização do gradiente descendente foi mostrado por (ARAÚJO; OLIVEIRA; MEIRA, 2015)(ARAÚJO, 2016) uma boa escolha para o desenvolvimento do algoritmo de treinamento. Dessa forma, o processo de treinamento definido para o GIMP também foi o de gradiente descendente. A partir do modelo definido para o GIMP, é possível notar que o modelo requer uma configuração de parâmetros de entrada  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^d$  e  $\lambda, \theta \in \mathbb{R}$ . O vetor peso  $\mathbf{w}$  onde  $\mathbf{w} \in \mathbb{R}^{3d+2}$  é definido por:

$$\mathbf{w}^T = (\lambda, \theta, \mathbf{a}, \mathbf{b}, \mathbf{c}), \quad (5.34)$$

Os valores dos pesos são valores que servirão para ponderar cada uma das variáveis de entrada da rede, permitindo-se quantificar as suas relevâncias em relação à funcionalidade do modelo do neurônio. Durante o processo de treinamento proposto, os pesos do modelo GIMP são ajustados de forma iterativa, e de acordo com um critério de erro ( $e$ ) até que a convergência seja alcançada. Para que isto ocorra, faz-se necessário definir uma função objetivo em termo dos pesos  $\mathbf{J}(\mathbf{w})$ , a ser minimizada durante o processo de treinamento. Seguindo o mesmo conceito definido por (ARAÚJO; OLIVEIRA; MEIRA, 2015)(ARAÚJO, 2016), a função peso pode ser definida por:

$$J(\mathbf{w}) = \sum_{m=1}^M e^2(m). \quad (5.35)$$

em que  $M$  representa a quantidade de padrões apresentados ao modelo no processo de treinamento, e  $e(m)$  representa o erro instantâneo do modelo para o  $m$ -ésimo padrão de treinamento, sendo definido por:

$$e(m) = d(m) - y(m), \quad (5.36)$$

sendo  $d(m)$  o valor da saída desejada e  $y(m)$  o valor da saída do modelo.

Pode-se observar que a função  $J(w)$  gera uma superfície de erro que reside no espaço  $\mathbb{R}^{3d+2}$ . O objetivo é a busca do ponto ótimo neste espaço que minimize o erro entre a saída obtida pelo modelo e a saída desejada, ou seja, encontrar um  $J(w)$  mínimo possível. No presente estudo também é proposto um algoritmo de aprendizagem supervisionado, empregando um método de gradiente descendente com um método de ajuste de fase automático, baseado no algoritmo de retropropagação de erro (*back propagation*, BP) (HAYKIN et al., 2009) (BRAGA; CARVALHO; LUDERMIR, 2000). Nesse algoritmo, o peso  $w$  para o  $m$ -ésimo padrão de treinamento é dado pela função iterativa

$$\mathbf{w}(i+1) = \mathbf{w}(i) - \mu \nabla J(\mathbf{w}), \quad (5.37)$$

onde  $i \in \{1, 2, \dots, n\}$  e o termo  $\mu$  representa o tamanho do passo ou taxa de aprendizagem, devendo ser maior que 0, e  $\mathbf{w}(i)$  é a condição inicial para o treinamento do modelo GIMP. O termo  $\nabla J(\mathbf{w})$  é dado pelo gradiente de  $J$  em relação ao  $\mathbf{w}$  nos pontos em que existe gradiente. Dessa forma, tem-se que:

$$\nabla J(\mathbf{w}) = \frac{\partial J}{\partial \mathbf{w}} = \left( \frac{\partial J}{\partial \lambda}, \frac{\partial J}{\partial \theta}, \frac{\partial J}{\partial \mathbf{a}}, \frac{\partial J}{\partial \mathbf{b}}, \frac{\partial J}{\partial \mathbf{c}} \right) \quad (5.38)$$

onde

$$\frac{\partial J}{\partial \lambda} = -2e(m) \frac{\partial y}{\partial \lambda}, \quad (5.39)$$

$$\frac{\partial J}{\partial \theta} = -2e(m) \frac{\partial y}{\partial \theta}, \quad (5.40)$$

$$\frac{\partial J}{\partial \mathbf{c}} = -2e(m) \frac{\partial y}{\partial \mathbf{c}}, \quad (5.41)$$

$$\frac{\partial J}{\partial \mathbf{b}} = -2e(m) \left( \frac{\partial y}{\partial \mathbf{b}^1}, \frac{\partial y}{\partial \mathbf{b}^2}, \dots, \frac{\partial y}{\partial \mathbf{b}^k} \right) \quad (5.42)$$

$$\frac{\partial J}{\partial \mathbf{a}} = -2e(m) \left( \frac{\partial y}{\partial \mathbf{a}^1}, \frac{\partial y}{\partial \mathbf{a}^2}, \dots, \frac{\partial y}{\partial \mathbf{a}^k} \right) \quad (5.43)$$

A existência do gradiente de  $J$  em relação a  $w$  depende apenas da existência do gradiente  $\frac{\partial y}{\partial \mathbf{a}^i}$  para  $i = 1, 2, \dots, k$  e  $\frac{\partial y}{\partial \mathbf{b}^i}$  para  $i = 1, 2, \dots, k$ , onde os pontos de não diferenciabilidade  $\frac{\partial y}{\partial \mathbf{a}^i}$  e  $\frac{\partial y}{\partial \mathbf{b}^i}$  serão gerados baseados no conceito do vetor indicador de posto (*rank indicator vector*), serão utilizadas as equações 5.31 e 5.33 (PESSOA; MARAGOS, 2000). Neste processo de treinamento deve-se ser utilizado o parâmetro de escala da função rank  $\sigma \rightarrow 0$ , pois este afeta diretamente a estimação e interpolação de  $\frac{\partial y}{\partial \mathbf{a}^i}$  e  $\frac{\partial y}{\partial \mathbf{b}^i}$ . A partir deste momento serão calculados todos os termos presentes na fórmula que define o gradiente descendente da função de erro do algoritmo de treinamento. Cálculo de  $\frac{\partial y}{\partial \lambda}$

$$\frac{\partial y}{\partial \lambda} = \frac{\partial}{\partial \lambda} \left[ \underbrace{\lambda \alpha}_{\alpha} + \underbrace{(\beta - \lambda \beta)}_{-\beta} \right] = \alpha - \beta \quad , \quad (5.44)$$

Cálculo do termo  $\frac{\partial y}{\partial \theta}$

$$\frac{\partial y}{\partial \theta} = \frac{\partial y}{\partial \alpha} \frac{\partial \alpha}{\partial \theta} \quad , \quad (5.45)$$

onde

$$\frac{\partial y}{\partial \alpha} = \frac{\partial}{\partial \alpha} \left[ \underbrace{\lambda \alpha}_{\lambda} + \underbrace{(1 - \lambda) \beta}_0 \right] = \lambda \quad , \quad (5.46)$$

e

$$\frac{\partial \alpha}{\partial \theta} = \frac{\partial}{\partial \theta} \left[ \underbrace{\theta \mu}_{\mu} + \underbrace{(1 - \theta)v}_{-v} \right] = \mu - v \quad , \quad (5.47)$$

Portanto,

$$\frac{\partial y}{\partial \theta} = \lambda(\mu - v) \quad . \quad (5.48)$$

O termo  $\frac{\partial y}{\partial \mathbf{c}}$  é dado por:

$$\frac{\partial y}{\partial \mathbf{c}} = \frac{\partial y}{\partial \beta} \frac{\partial \beta}{\partial \mathbf{c}} \quad . \quad (5.49)$$

onde

$$\frac{\partial y}{\partial \beta} = (1 - \lambda) \quad , \quad (5.50)$$

e

$$\frac{\partial \beta}{\partial \mathbf{c}} = \mathbf{x} \quad . \quad (5.51)$$

O termo  $\frac{\partial y}{\partial \mathbf{b}^i}$  é definido por:

$$\frac{\partial y}{\partial \mathbf{b}^i} = \frac{\partial y}{\partial \alpha} \frac{\partial \alpha}{\partial v} \frac{\partial v}{\partial \mathbf{b}^i} = \lambda \left( \frac{\partial \alpha}{\partial v} \frac{\partial v}{\partial \mathbf{b}^i} \right) \quad , \quad (5.52)$$

sendo

$$\frac{\partial \alpha}{\partial v} = \frac{Q_{\sigma}(\alpha \cdot \mathbf{1} - v)}{Q_{\sigma}(\alpha \cdot \mathbf{1} - v) \cdot \mathbf{1}} \quad (5.53)$$

e

$$\frac{\partial v}{\partial \mathbf{b}^i} = \frac{Q_{\sigma}(\mathbf{v}_i \cdot \mathbf{1} - \mathbf{x} + \mathbf{b}^i)}{Q_{\sigma}(\mathbf{v}_i \cdot \mathbf{1} - \mathbf{x} + \mathbf{b}^i) \cdot \mathbf{1}} \quad (5.54)$$

e o termo  $\frac{\partial y}{\partial \mathbf{a}^i}$  é obtido por:

$$\frac{\partial y}{\partial \mathbf{a}^i} = \frac{\partial y}{\partial \alpha} \frac{\partial \alpha}{\partial \mu} \frac{\partial \mu}{\partial \mathbf{a}^i} = \lambda \left( \frac{\partial \alpha}{\partial \mu} \frac{\partial \mu}{\partial \mathbf{a}^i} \right) \quad , \quad (5.55)$$

Sendo

$$\frac{\partial \alpha}{\partial \mu} = \frac{Q_{\sigma}(\alpha \cdot \mathbf{1} - \mu)}{Q_{\sigma}(\alpha \cdot \mathbf{1} - \mu) \cdot \mathbf{1}} \quad (5.56)$$

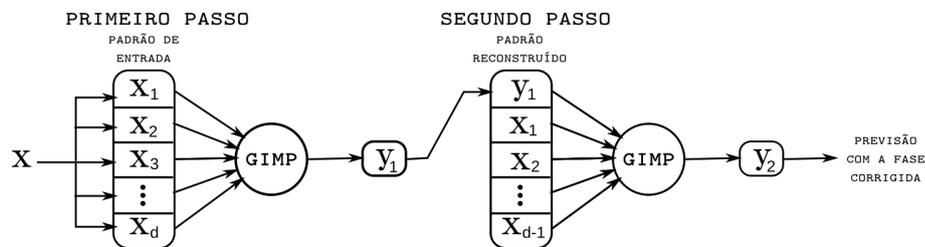
and

$$\frac{\partial \mu}{\partial \mathbf{a}^i} = \frac{Q_{\sigma}(\mathbf{u}_i \cdot \mathbf{1} - \mathbf{x} + \mathbf{a}^i)}{Q_{\sigma}(\mathbf{u}_i \cdot \mathbf{1} - \mathbf{x} + \mathbf{a}^i) \cdot \mathbf{1}} \quad (5.57)$$

### 5.3.3 Utilização do Ajuste de Fase Temporal no modelo proposto

Seguindo o processo de treinamento, o algoritmo faz o ajuste das distorções da fase temporal que ocorrem nas séries temporais diárias, semanais e quinzenais. O procedimento de ajuste segue o mesmo processo de Araújo (ARAÚJO, 2012b) (ARAÚJO; OLIVEIRA; MEIRA, 2015) (ARAÚJO, 2016) e do criador do procedimento, (FERREIRA, 2006) (FERREIRA; VASCONCELOS; ADEODATO, 2008), no qual se faz uma primeira previsão, sendo posteriormente o resultado utilizado numa outra iteração para calcular o resultado final da previsão do próximo valor, conforme pode-se observar na figura 55. Para a execução do processos de treinamento, faz-se necessário a definição de três condições de paradas (PRECHELT et al., 1994): perda de generalização, processos de treinamento e o estabelecimento do número máximo de épocas de treinamento que o algoritmo pode executar.

Figura 55 – Procedimento de ajuste de fase automático.



Fonte: O autor.

## 5.4 RESUMO DO CAPÍTULO

Neste Capítulo foi abordado o modelo proposto para previsão de séries temporais financeiras em F-frequências (diária, semanal, quinzenal) de baixa frequência. Foram mostrados os principais fundamentos que embasam teoricamente a criação do novo neurônio, bem como a arquitetura do modelo proposto e o detalhamento do cálculo de cada uma das componentes do modelo. Posteriormente, foi apresentado um método de gradiente descendente para treinamento do modelo proposto, baseado em ideias do algoritmo de retropropagação do erro, e apresentada sua forma de treinamento, finalizando com a definição do processo de ajuste de fase utilizado para superar o problema do Dilema do Passeio Aleatório. Após observar toda formulação matemática para o cálculo do erro é possível identificar que cada uma das entradas do neurônio tem um papel importante para o seu treinamento. Cada operador será responsável pelo mapeamento dos valores que farão parte do cálculo do valor que pretende-se prever.

## 6 SIMULAÇÕES E RESULTADOS EXPERIMENTAIS

Neste capítulo serão apresentados os experimentos executados com o modelo proposto, bem como as medidas utilizadas para a avaliação dos resultados preditivos alcançados. Esses resultados serão analisados e discutidos, utilizando um conjunto relevante de medidas de desempenho, e comparados aos modelos comumente empregados na literatura de previsão de séries temporais financeiras. A metodologia adotada no procedimento experimental foi principalmente baseada nos trabalhos do estado da arte (FERREIRA, 2006; ARAÚJO, 2012b; ARAÚJO, 2016; CAVALCANTE et al., 2016).

### 6.1 METODOLOGIA

O processo para a análise dos resultados seguiu os passos: Exposição detalhada do processo de obtenção dos dados estatísticos das séries temporais utilizadas; Identificação dos trabalhos do estado da arte que tenham parâmetros para comparação; Análise, na literatura, dos estudos que conseguiram superar o problema do passeio aleatório; Exposição de análise comparativa com os modelos utilizados para o estudo de séries temporais financeira e outras técnicas de aprendizado de máquina. Para que as séries temporais possam ser utilizadas no modelo proposto, é necessário que haja um processo de normalização ou escalonamento. Esta etapa é formalmente definida como pré-processamento (ZHANG et al., 2015; ARAÚJO, 2012b; ARAÚJO, 2016). Além da normalização, a etapa de pré-processamento também pode ser utilizada para organizar os dados da série sem gerar nenhuma modificação série/tempo nos dados estudados. Por exemplo, se uma ação sofreu um processo de *split*, dividindo seu valor na metade e criando o dobro de ações existentes, esse dado deve ser reorganizado para não influenciar no processo de aprendizagem da rede neural e para que esta rede não absorva uma característica errada, como se as ações tivessem tido uma queda de cinquenta por cento em único dia.

O processo de normalização pode ser desenvolvido de diversas formas, dependendo do tipo de problema em questão. Nos experimentos desenvolvidos neste trabalho, foi definido o uso da normalização linear para o intervalo  $[0, 1]$ , dando a possibilidade de utilização de todo o domínio de atuação do GIMP.

Após a normalização, todas as séries utilizadas no estudo foram divididas em três conjuntos, seguindo (PRECHELT et al., 1994), que em seu trabalho busca padronizar a divisão e o percentual do conjunto de dados em problemas de classificação e previsão (ARAÚJO, 2012b) (ARAÚJO, 2016):

- i) Conjunto de treinamento – utilizado no processo de aprendizagem da rede neural do modelo de previsão;

- ii) Conjunto de validação – utilizado para a validação da aprendizagem do modelo de previsão;
- iii) Conjunto de teste – utilizado para confirmar o desempenho prático do modelo de previsão;

Seguindo Prechelt, os conjuntos foram separados nos percentuais de 50% dos dados para o processo de aprendizagem, 25% para o processo de validação e 25% para o processo de testes.

Com o intuito de comparar o desempenho preditivo do modelo proposto, foram escolhidos, no estado da arte, modelos já utilizados na predição de séries temporais com diversas abordagens que obtiveram resultados relevantes, utilizando diferentes formas de mapeamento da função de previsão de séries temporais de baixa frequência. Os modelos escolhidos foram: ARIMA (BOX et al., 2015), Regressão Logística, Rede MLP, Rede de Base Radial, Modelo Regressor de Vetor de Suporte Linear, Modelo Regressor de Vetor de Suporte Polinomial, Modelo Regressor de Vetor de Suporte com função de Base Radial e a rede NARX-Takens (MENEZES JR.; BARRETO, 2008). Tal escolha se baseou na literatura em um conjunto de formas existentes para o desenvolvimento de modelos matemáticos e modelos que utilizam a inteligência artificial para prever valores futuros. Após os experimentos e a análise comparativa dos resultados, será feita outra análise comparativa com os dados diretos dos trabalhos encontrados na literatura. Essa última comparação foi diferenciada, utilizando apenas o melhor modelo encontrado em cada trabalho, pois estes não utilizavam as médias de  $n$  iterações desenvolvida no modelo, podendo-se observar, assim, apenas o melhor caso de cada modelo.

A Regressão Logística foi analisada com os parâmetros série temporal financeira  $X$  e saídas esperadas  $Y$  como entrada, para obter os valores  $B$  da previsão. A função sigmóide utilizada segue a equação padrão equação 3.34, assim como a função de ativação sendo equação 3.36.

Os experimentos executados com a Regressão de Vetores de Suporte foram desenvolvidos utilizando três funções *kernel*: linear, polinomial e função base radial.

Os experimentos foram realizados com o modelo ARIMA(p;q;d), no qual o termo de diferenciação  $d$  sendo igual a 1, como sugerido por Box et al. (2015), e Araújo (2016). Utilizou-se uma arquitetura de três camadas para os experimentos com a rede MLP, sendo elas: camada de entrada, camada intermediária e camada de saída. A rede MLP pode ser formalmente descrita por MLP(I;H;O), em que I(input), H(hidden), O(output) representam a quantidade de unidades de processamento das camadas de entrada, intermediária e saída, respectivamente. I é definido pela quantidade de retardos temporais utilizados para a descrição da série temporal.

Conforme o trabalho de (FERREIRA; VASCONCELOS; ADEODATO, 2008), foram usados os valores dos parâmetros 1, 3, 5 e 10 para os retardos temporais utilizados na rede MLP.

A quantidade de unidades de processamento na camada escondida foi determinada através do estudo de (ARAÚJO, 2016) os valores 1, 5, 10, 25 e 50. A quantidade de unidades de processamento na camada de saída foi fixada em 1, pois o objetivo é fazer a previsão apenas de um-passo-adiante  $O = 1$ . A função de ativação utilizada na rede MLP foi a sigmóide logística, para todas as unidades de processamento, enquanto para a unidade de saída, utilizou-se uma função de ativação linear. Esta parametrização mostra nos trabalhos de (FERREIRA; VASCONCELOS; ADEODATO, 2008), (SALGADO et al., 2010) e (ARAÚJO, 2016) o melhor desempenho para previsão de séries temporais financeiras.

Para o treinamento da rede, optou-se pelo algoritmo de levenberg-marquardt (HAGAN; MENHAJ, 1994), como empregado nos trabalhos de (ARAÚJO, 2016) (ARAÚJO; OLIVEIRA; MEIRA, 2015)(ARAÚJO; SOARES; OLIVEIRA, 2012) de (FERREIRA; VASCONCELOS; ADEODATO, 2008) e de (SALGADO et al., 2010), mediante os critérios de parada sugeridos por Prechelt (PRECHELT et al., 1994): *i*) A quantidade máxima de épocas de treinamento  $10^4$  *ii*) O aumento no erro de validação (generalization loss) maior que 5%, e *iii*) A queda no erro de treinamento ou process training ser menor que  $10^{-6}$ .

Na Rede de Função Base Radial, o parâmetro *Goal* (condição de parada) foi ajustado para 0,001; o *Spread* utilizado foi 0,2; no processo, foram variados os parâmetros de *Delay* (quantidade de valores passados) entre [1-15] e DF (Número de neurônios adicionados à camada escondida) entre [20,40]. Para os experimentos com Modelos de Regressor de Vetor Suporte, foram utilizados os seguintes dados: comprimento da margem  $C = 100$ ; função *Kernel* = linear, polinomial de ordem 3 e função de base radial com margem de tolerância  $\epsilon = 25$  e argumento de variação 0,5. Todos os parâmetros utilizados foram definidos de acordo com uma metodologia empírica, criada através de uma sequência de experimentos, onde foram analisados os melhores valores que deveriam fazer parte de cada modelo de previsão de série temporal financeira de baixa frequência.

A escolha da rede NARX se baseou nas evidências de (MENEZES JR.; BARRETO, 2008), que mostram que os sistemas estatísticos lineares não conseguem suportar a não linearidade. Os métodos estatísticos não lineares têm baixo desempenho, alta complexidade e, geralmente, necessitam de intervenção humana para o funcionamento. Além desses fatos, as Redes MLP são geralmente formadas por meio de variantes baseadas em gradiente temporal do algoritmo de retropropagação (PEARLMUTTER, 1995), onde as novas entradas serão apenas retroalimentadas pelas saídas anteriores. Porém, são necessários alguns cuidados para que uma rede NARX consiga realmente absorver as características de longo prazo. Diferentes das redes MLP, que aprendem a executar tarefas em que as dependências temporais presentes nos sinais de entrada-saída estendem por longos intervalos de tempo, os algoritmos de aprendizagem baseados em gradientes podem ter dificuldades nesse tipo de tarefa. A NARX se mostra a melhor Rede da literatura, pelo fato de sua aprendizagem ter dependências temporais de longo prazo (LIN et al., 1996). Isso ocorre porque a NARX tem o vetor de entrada do modelo construído através de duas linhas de

atraso: um deslizamento sobre o sinal de entrada e o outro passando pelo conjunto de saídas da rede. Para a análise experimental com esta rede, foi utilizada uma arquitetura que consiste de uma rede NARX formalmente definida  $NARX(D_y, \tau, D_e, H)$  no modo Serial-Paralelo, e seus parâmetros foram determinados com base no estudo dos lagplots do capítulo 3. O parâmetro  $D_y$  foi parametrizado com o valor 50, devido ao relacionamento dominante linear encontrando nos primeiros cinquenta retardos temporais via *lagplot*. O  $\tau = 5$  fará com que suas entradas exógenas sejam direcionadas e que, seguindo o teorema da imersão de Floris Takens, haja um espaçamento entre cada observação da série. Esse parâmetro foi definido com base no trabalho de (ARAÚJO, 2016), que mostrou um bom comportamento na previsão de séries temporais financeiras para a NARX. O  $D_e = 60$  foi calculado baseado na observação da componente não linear, via estudo do *lagplot*, onde a dimensionalidade máxima utilizada no estudo foi  $D = 300$ . Observa-se que o valor de  $D_e$  pode ser calculado pela divisão da dimensionalidade pelo valor de  $\tau \frac{D}{\tau}$ . O cálculo da quantidade de processamento na camada escondida ( $H$ ) foi realizado através de uma metodologia empírica, seguindo (ARAÚJO, 2016) com valores de 1, 5, 10, 25 e 50.

O processo de treinamento da rede NARX foi desenvolvido utilizando o algoritmo de levenberg-marquardt (HAGAN; MENHAJ, 1994), como utilizado nos trabalhos de (ARAÚJO, 2016)(ARAÚJO; OLIVEIRA; MEIRA, 2015) (ARAÚJO; SOARES; OLIVEIRA, 2012), com os critérios de parada sugeridos por (PRECHELT et al., 1994): *i*) A quantidade máxima de épocas de treinamento  $10^4$  *ii*) O aumento no erro de validação (*generalization loss*) maior que 5%, e *iii*) A queda no erro de treinamento ou *process training* ser menor que  $10^{-6}$ .

O GIMP foi executado seguindo a quantidade de módulos das decomposições morfológicas determinada pelo estudo dos *lagplots* e no trabalho de (ARAÚJO; SOARES; OLIVEIRA, 2012). Uma arquitetura básica foi definida para o modelo, sendo formalmente descrita pela notação  $GIMP(lags;k;)$ , onde *lags* representa os retardos temporais utilizados para a descrição da série temporal, e *k* representa a quantidade de módulos das decomposições morfológicas. Os valores escolhidos para os atrasos foram fixados entre 1e200, intervalo no qual se verifica uma forte presença da componente linear e não linear das séries estudadas e os valores de  $k = 1, 5, 10, 25, 50$  e 100. A quantidade de módulos utilizada no experimento foi 100. O valor de  $\mu$  é determinado empiricamente através de uma série de experimentos, a partir dos quais foi escolhida a quantidade 0,001. O valor de  $\sigma$  também é determinado empiricamente através de uma série de experimentos, a partir dos quais foi eleita a quantidade 0,05. Os valores iniciais do vetor de pesos do modelo GIMP são:  $\mathbf{a}, \mathbf{b}, \mathbf{p} \in [-1,1]$  e  $\theta, \lambda \in [0,1]$ .

Assim como na rede MLP, o processo de treinamento da GIMP foi baseado no algoritmo de retro-propagação do erro *back-propagation*, como utilizado nos trabalhos de (ARAÚJO, 2016)(ARAÚJO; OLIVEIRA; MEIRA, 2015)(ARAÚJO; SOARES; OLIVEIRA, 2012), de (FERREIRA; VASCONCELOS; ADEODATO, 2008) e de (SALGADO et al., 2010), com os critérios de parada sugeridos por (PRECHELT et al., 1994): *i*) A quantidade máxima de

épocas de treinamento  $10^4$  ii) O aumento no erro de validação (generalization loss) maior que 5%, e iii) A queda no erro de treinamento ou *process training* ser menor que  $10^{-6}$ . O objetivo desta análise é verificar o comportamento de diferentes modelos de previsão utilizados em diferentes área na literatura, com foco no estudo de previsões de séries temporais financeiras de baixa frequência, e a contribuição do ajuste de fase no mapeamento gerado pelo modelo proposto.

Para cada modelo investigado foram realizados cinquenta experimentos, analisando as estatísticas básicas de cada resultado obtido: média (*MEAN*), raiz do erro quadrático médio (RMS), mediana (*MEDIAN*) e os percentis 2, 5% (PERC25) e 97, 5% (PERC975). Além desta análise, também foi desenvolvido o teste de Wilcoxon, uma alternativa não paramétrica para o teste de experimentos em blocos ao acaso, verificando se existe uma sazonalidade determinística e considerando a série temporal livre de tendência. Algumas medidas de desempenho de previsão também foram calculadas para analisar o ganho do modelo GIMP com relação aos outros encontrados na literatura, relatadas a seguir.

## 6.2 MEDIDAS PARA DESEMPENHO DE PREVISÃO

O desempenho de um modelo de previsão de uma série temporal é frequentemente definido em termos do erro instantâneo de previsão, que é calculado pela subtração entre o valor real e o valor previsto pelo modelo.

$$Erro_j = valor_{real} - valor_{previsto}; \quad (6.1)$$

em que  $j$  é o valor do erro com relação ao valor correto de previsão. A partir desse conceito, algumas medidas são utilizadas pelos pesquisadores da área (ARAÚJO, 2012b) e (FERREIRA, 2006) foram adotados para o cálculo da função de desempenho, denominado neste trabalho de função de avaliação (*evaluation function*, EF). Esta função foi proposta por (ARAÚJO, 2007), e vem sendo utilizada até o momento por pesquisadores da área. Formalmente, a função Função de Avaliação - Evaluation Function (EF) é definida por

$$EF = \frac{POCID}{1 + MSE + MAPE + THEIL + ARV} \quad (6.2)$$

Onde o Previsão de Mudança na Direção - Prediction of Change in Direction (POCID) é a previsão de mudança de direção, Erro Médio Quadrático - Mean Squared Error (MSE) é o valor do erro médio quadrático, o Erro Médio Percentual Absoluto - Mean Absolute Percentage Error (MAPE) corresponde ao valor de Erro Médio Percentual Absoluto, Estatística U de Theil - U of Theil statistics (THEIL) corresponde ao valor U de Theil e Variância Relativa Média - Average Relative Variance (ARV) corresponde à Variância Relativa Média.

POCID é uma métrica utilizada que mapeia o acerto de previsão na direção futura da série temporal, indicando se o valor futuro da série aumentará ou diminuirá em relação ao passado. A equação que define esta métrica é:

$$POCID = \frac{100}{N} \sum_{j=1}^N DP_j \quad (6.3)$$

na qual

$$DP_j = \begin{cases} 1, & \text{se } (x_j - x_{j-1})(\hat{x}_j - x_{j-1}) \geq 0 \\ 0, & \text{caso contrário} \end{cases} \quad (6.4)$$

MSE é a principal métrica utilizada na literatura (ARAÚJO, 2012b), (FERREIRA, 2006) (OLIVEIRA, 2001) (FREITAS; SOUZA; ALMEIDA, 2009) (DOUCOURE; AGBOSSOU; CARDENAS, 2016), e objetiva guiar o modelo de previsão no processo de aprendizagem. A equação que define esta métrica é dada por:

$$MSE = \frac{1}{N} \sum_{j=1}^N e_j^2, \quad (6.5)$$

em que  $N$  é a quantidade de padrões e  $e_j$  é o erro instantâneo para o padrão  $j$ .

MAPE é a métrica utilizada para identificar os desvios do modelo de previsão. Pode-se calcular o MAPE através da fórmula (6.6).

$$MAPE = \frac{1}{N} \sum_{j=1}^N \left| \frac{e_j}{x_j} \right|, \quad (6.6)$$

no qual  $e_j$  é o erro instantâneo para o padrão  $j$ ,  $N$  a quantidade de padrões e  $x_j$  é o valor real instantâneo para o padrão  $j$ .

A métrica  $U$  de Theil revela-se de suma importância para a previsão de séries temporais, pois determina o relacionamento entre um modelo de previsão e um modelo de passeio aleatório, comparando seu desempenho.  $U$  de Theil é descrita pela equação:

$$THEIL = \frac{\sum_{j=1}^{N-1} (e_j)^2}{\sum_{j=1}^{N-1} (x_j - x_{j+1})^2}, \quad (6.7)$$

Em que, se  $THEIL = 1$ , implica que o modelo de previsão tem a mesma performance de um RW. Se  $THEIL > 1$  implica que o modelo de previsão tem seu desempenho pior que um RW e se  $THEIL < 1$ , então o modelo de previsão é melhor que um RW. Importante frisar que o modelo de previsão perfeito  $THEIL \rightarrow 0$ .

A métrica Variância Relativa Média (ARV) associa o desempenho do modelo de previsão em relação à previsão da média da série temporal. Podemos calcular o ARV utilizando a equação (6.8),

$$ARV = \frac{1}{N} \frac{\sum_{t=1}^N (saida_t - target_t)^2}{\sum_{t=1}^N (saida_t - target_t)^2}, \quad (6.8)$$

No qual  $N$  é a quantidade de pontos preditos,  $target_t$  é a saída desejada para o modelo preditor no tempo  $t$ , saída  $t$  é a saída do modelo preditor e  $\overline{target}$  é a média da série temporal.  $ARV = 1$  implica que poderemos calcular a média do modelo de previsão e utilizar este valor como previsão; se  $ARV > 1$ , temos que o modelo de previsão tem seu desempenho pior que a média da série temporal; e se  $ARV < 1$ , então o modelo de previsão é melhor quando comparado à previsão através da média da série. Importante frisar que o modelo de previsão perfeito  $ARV \rightarrow 0$ .

Após o cálculo das medidas, foi executado, em todas as séries temporais o teste de decisão ( $h$ ), denominado teste não-paramétrico de Wilcoxon. O teste de Wilcoxon ou teste dos postos sinalizados de Wilcoxon é um teste de hipóteses não paramétrico utilizado quando se deseja comparar duas amostras relacionadas, amostras emparelhadas ou medidas repetidas em uma única amostra para avaliar se os postos médios populacionais diferem (WILCOXON, 1945). O objetivo deste teste é revelar se, estatisticamente, o modelo proposto obtém desempenho preditivo superior aos modelos investigados neste trabalho. Por fim, é desenvolvida a análise dos ruídos gerados pelo modelo GIMP, buscando verificar se este tem um comportamento de ruído branco, e assim validar o modelo.

### 6.3 RESULTADOS

Nesta seção serão apresentados os resultados de todas as séries temporais analisadas no capítulo 4. Serão mostrados os resultados das medidas analisadas, através de gráficos e tabelas, comentando-se o comportamento de cada uma das séries temporais e modelo estudados.

#### 6.3.1 PETR4 - Frequência Diária

Os experimentos com a série temporal PETR4 na frequência diária foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 15.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal PETR4 na frequência diária. Os resultados podem ser observados na tabela tabela 16.

Através da figura 56 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal PETR4 diário.

O ganho do modelo GIMP em relação aos outros modelos para esta série pode ser vista na tabela 17

Tabela 15 – Desempenho preditivo para a série PETR4 - Frequência Diária (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	6.798749e-06	7.395614e-03	2.256309e-03	9.870610e+01	1.805990e-01	8.520867e+01
	RMS	2.607441e-03	8.599776e-02	4.750062e-02	9.935094e+00	4.249694e-01	9.230854e+00
	MEDIAN	2.823030e-06	5.629362e-03	9.409003e-04	9.870610e+01	7.498975e-02	9.126284e+01
	PERC25	1.426659e-07	1.265618e-03	4.759232e-05	9.870610e+01	3.789716e-03	5.850585e+01
	PERC975	2.492566e-05	1.672884e-02	8.246839e-03	9.870610e+01	6.621145e-01	9.820495e+01
ARIMA	MEAN	3.782365e-05	1.573913e-02	1.268272e-02	4.658041e+01	1.005877e+00	2.289709e+01
	RMS	6.150093e-03	1.254557e-01	1.126176e-01	6.824984e+00	1.002934e+00	4.785091e+00
	MEDIAN	3.782365e-05	1.573913e-02	1.268272e-02	4.658041e+01	1.005877e+00	2.289709e+01
	PERC25	3.782365e-05	1.573913e-02	1.268272e-02	4.658041e+01	1.005877e+00	2.289709e+01
	PERC975	3.782365e-05	1.573913e-02	1.268272e-02	4.658041e+01	1.005877e+00	2.289709e+01
LOGREG	MEAN	1.109353e-04	3.062969e-02	4.501482e-02	4.658041e+01	2.951776e+00	1.156550e+01
	RMS	1.053258e-02	1.750134e-01	2.121670e-01	6.824984e+00	1.718073e+00	3.400808e+00
	MEDIAN	1.109353e-04	3.062969e-02	4.501482e-02	4.658041e+01	2.951776e+00	1.156550e+01
	PERC25	1.109353e-04	3.062969e-02	4.501482e-02	4.658041e+01	2.951776e+00	1.156550e+01
	PERC975	1.109353e-04	3.062969e-02	4.501482e-02	4.658041e+01	2.951776e+00	1.156550e+01
MLP	MEAN	3.767513e-05	1.571382e-02	1.269283e-02	4.658041e+01	1.002095e+00	2.293991e+01
	RMS	6.138007e-03	1.253548e-01	1.126625e-01	6.824984e+00	1.001047e+00	4.789563e+00
	MEDIAN	3.767714e-05	1.571459e-02	1.269375e-02	4.658041e+01	1.002150e+00	2.293927e+01
	PERC25	3.766331e-05	1.570931e-02	1.268752e-02	4.658041e+01	1.001774e+00	2.293927e+01
	PERC975	3.767716e-05	1.571460e-02	1.269376e-02	4.658041e+01	1.002151e+00	2.294366e+01
NARXTAKENS	MEAN	7.552359e-05	2.410442e-02	3.414197e-02	9.870610e+01	2.008786e+00	3.476985e+01
	RMS	8.690431e-03	1.552560e-01	1.847755e-01	9.935094e+00	1.417317e+00	5.896597e+00
	MEDIAN	5.946506e-05	2.191916e-02	2.529596e-02	9.870610e+01	1.581766e+00	3.760439e+01
	PERC25	4.068828e-05	1.746480e-02	1.645830e-02	9.870610e+01	1.082445e+00	2.055302e+01
	PERC975	1.391250e-04	3.385339e-02	6.811697e-02	9.870610e+01	3.700402e+00	4.663849e+01
RBF	MEAN	6.677792e-05	2.323870e-02	2.525438e-02	4.658041e+01	1.776657e+00	1.648738e+01
	RMS	8.171776e-03	1.524424e-01	1.589163e-01	6.824984e+00	1.332913e+00	4.060465e+00
	MEDIAN	6.677792e-05	2.323870e-02	2.525438e-02	4.658041e+01	1.776657e+00	1.648738e+01
	PERC25	6.677792e-05	2.323870e-02	2.525438e-02	4.658041e+01	1.776657e+00	1.648738e+01
	PERC975	6.677792e-05	2.323870e-02	2.525438e-02	4.658041e+01	1.776657e+00	1.648738e+01
SVR-LIN	MEAN	4.138811e-05	1.686238e-02	1.419492e-02	4.658041e+01	1.101173e+00	2.184544e+01
	RMS	6.433359e-03	1.298552e-01	1.191424e-01	6.824984e+00	1.049368e+00	4.673910e+00
	MEDIAN	4.138811e-05	1.686238e-02	1.419492e-02	4.658041e+01	1.101173e+00	2.184544e+01
	PERC25	4.138811e-05	1.686238e-02	1.419492e-02	4.658041e+01	1.101173e+00	2.184544e+01
	PERC975	4.138811e-05	1.686238e-02	1.419492e-02	4.658041e+01	1.101173e+00	2.184544e+01
SVR-POL	MEAN	4.794006e-05	1.895278e-02	1.703248e-02	4.658041e+01	1.275486e+00	2.015142e+01
	RMS	6.923876e-03	1.376691e-01	1.305085e-01	6.824984e+00	1.129374e+00	4.489034e+00
	MEDIAN	4.794006e-05	1.895278e-02	1.703248e-02	4.658041e+01	1.275486e+00	2.015142e+01
	PERC25	4.794006e-05	1.895278e-02	1.703248e-02	4.658041e+01	1.275486e+00	2.015142e+01
	PERC975	4.794006e-05	1.895278e-02	1.703248e-02	4.658041e+01	1.275486e+00	2.015142e+01
SVR-RBF	MEAN	1.493810e-04	3.638807e-02	6.118604e-02	4.658041e+01	3.975257e+00	9.182059e+00
	RMS	1.222215e-02	1.907566e-01	2.473581e-01	6.824984e+00	1.993805e+00	3.030191e+00
	MEDIAN	1.493810e-04	3.638807e-02	6.118604e-02	4.658041e+01	3.975257e+00	9.182059e+00
	PERC25	1.493810e-04	3.638807e-02	6.118604e-02	4.658041e+01	3.975257e+00	9.182059e+00
	PERC975	1.493810e-04	3.638807e-02	6.118604e-02	4.658041e+01	3.975257e+00	9.182059e+00

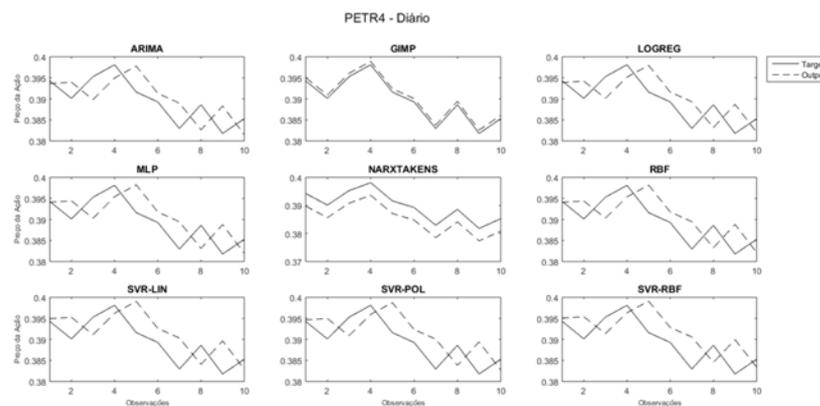
Fonte: O autor.

Tabela 16 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série PETR4 diário (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Figura 56 – Gráfico de previsão da série temporal PETR4 diária (conjunto de teste).



Fonte: O autor.

Tabela 17 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal PETR4 Diário

Modelo Proposto/Modelo Investigado	Ganho GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	82,0251377	53,011291	82,2095812	111,904747	82,0456179	272,137551
GIMP/LOGREG	93,8714287	75,854754	94,9876307	111,904747	93,8816834	636,748692
GIMP/MLP	81,9542786	52,935607	82,2237515	111,904747	81,9778564	271,442913
GIMP/NARXTAKENS	90,9978472	69,3184321	93,3913919	0	91,0095451	145,064819
GIMP/RBF	89,8188668	68,1754401	91,0656726	111,904747	89,8348978	416,811464
GIMP/SVR-LIN	83,5731832	56,1413395	84,1048136	111,904747	83,5993981	290,052432
GIMP/SVR-LOG	85,8182301	60,9787377	86,7529039	111,904747	85,8407697	322,842013
GIMP/SVR-RBF	95,448719	79,6757179	96,3123794	111,904747	95,4569227	827,990879

Fonte: O autor.

### 6.3.2 PETR4 - Frequência Semanal

Os experimentos com a série temporal PETR4 na frequência semanal foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 18. Foram utilizados os valores de fechamento para a criação de cada série.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos

Tabela 18 – Desempenho preditivo para a série PETR4 - Frequência Semanal (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	1.882565e-05	1.219583e-02	4.955887e-03	1.000000e+02	1.138669e-01	8.950258e+01
	RMS	4.338854e-03	1.104347e-01	7.039806e-02	1.000000e+01	3.374417e-01	9.460580e+00
	MEDIAN	8.294389e-06	9.647297e-03	2.200799e-03	1.000000e+02	5.016858e-02	9.416409e+01
	PERC25	1.781145e-07	1.416546e-03	4.736380e-05	1.000000e+02	1.077325e-03	7.450591e+01
	PERC975	5.043464e-05	2.383667e-02	1.323460e-02	1.000000e+02	3.050537e-01	9.974650e+01
ARIMA	MEAN	1.657644e-04	3.152387e-02	4.481545e-02	5.591398e+01	1.010868e+00	2.678677e+01
	RMS	1.287495e-02	1.775496e-01	2.116966e-01	7.477565e+00	1.005419e+00	5.175594e+00
	MEDIAN	1.657644e-04	3.152387e-02	4.481545e-02	5.591398e+01	1.010868e+00	2.678677e+01
	PERC25	1.657644e-04	3.152387e-02	4.481545e-02	5.591398e+01	1.010868e+00	2.678677e+01
	PERC975	1.657644e-04	3.152387e-02	4.481545e-02	5.591398e+01	1.010868e+00	2.678677e+01
LOGREG	MEAN	2.547247e-04	4.605283e-02	8.532619e-02	5.591398e+01	1.554107e+00	2.081883e+01
	RMS	1.596010e-02	2.145992e-01	2.921065e-01	7.477565e+00	1.246638e+00	4.562766e+00
	MEDIAN	2.547247e-04	4.605283e-02	8.532619e-02	5.591398e+01	1.554107e+00	2.081883e+01
	PERC25	2.547247e-04	4.605283e-02	8.532619e-02	5.591398e+01	1.554107e+00	2.081883e+01
	PERC975	2.547247e-04	4.605283e-02	8.532619e-02	5.591398e+01	1.554107e+00	2.081883e+01
MLP	MEAN	1.843933e-04	3.609586e-02	5.274423e-02	5.591398e+01	1.121971e+00	2.543667e+01
	RMS	1.357915e-02	1.899891e-01	2.296611e-01	7.477565e+00	1.059231e+00	5.043478e+00
	MEDIAN	1.736501e-04	3.470426e-02	4.963809e-02	5.591398e+01	1.057266e+00	2.610627e+01
	PERC25	1.700246e-04	3.399770e-02	4.807434e-02	5.591398e+01	1.035167e+00	2.024883e+01
	PERC975	2.693358e-04	4.705475e-02	7.865395e-02	5.591398e+01	1.635365e+00	2.640679e+01
NARXTAKENS	MEAN	9.867087e-05	2.728074e-02	3.920081e-02	1.000000e+02	6.009583e-01	6.470277e+01
	RMS	9.933321e-03	1.651688e-01	1.979919e-01	1.000000e+01	7.752150e-01	8.043803e+00
	MEDIAN	6.488040e-05	2.340673e-02	2.175089e-02	1.000000e+02	3.955274e-01	6.941671e+01
	PERC25	4.382212e-05	1.879415e-02	1.399890e-02	1.000000e+02	2.673626e-01	3.053086e+01
	PERC975	3.376595e-04	5.641353e-02	1.658985e-01	1.000000e+02	2.052725e+00	7.691127e+01
RBF	MEAN	2.619717e-03	1.131406e-01	1.000924e+00	5.053763e+01	1.601035e+01	2.787971e+00
	RMS	5.118317e-02	3.363638e-01	1.000462e+00	7.108983e+00	4.001293e+00	1.669722e+00
	MEDIAN	2.619717e-03	1.131406e-01	1.000924e+00	5.053763e+01	1.601035e+01	2.787971e+00
	PERC25	2.619717e-03	1.131406e-01	1.000924e+00	5.053763e+01	1.601035e+01	2.787971e+00
	PERC975	2.619717e-03	1.131406e-01	1.000924e+00	5.053763e+01	1.601035e+01	2.787971e+00
SVR-LIN	MEAN	1.836823e-04	3.646886e-02	5.305144e-02	5.591398e+01	1.117818e+00	2.532885e+01
	RMS	1.355294e-02	1.909682e-01	2.303290e-01	7.477565e+00	1.057269e+00	5.032777e+00
	MEDIAN	1.836823e-04	3.646886e-02	5.305144e-02	5.591398e+01	1.117818e+00	2.532885e+01
	PERC25	1.836823e-04	3.646886e-02	5.305144e-02	5.591398e+01	1.117818e+00	2.532885e+01
	PERC975	1.836823e-04	3.646886e-02	5.305144e-02	5.591398e+01	1.117818e+00	2.532885e+01
SVR-POL	MEAN	2.159186e-04	4.150588e-02	6.662072e-02	5.591398e+01	1.315431e+00	2.306898e+01
	RMS	1.469417e-02	2.037299e-01	2.581099e-01	7.477565e+00	1.146922e+00	4.803018e+00
	MEDIAN	2.159186e-04	4.150588e-02	6.662072e-02	5.591398e+01	1.315431e+00	2.306898e+01
	PERC25	2.159186e-04	4.150588e-02	6.662072e-02	5.591398e+01	1.315431e+00	2.306898e+01
	PERC975	2.159186e-04	4.150588e-02	6.662072e-02	5.591398e+01	1.315431e+00	2.306898e+01
SVR-RBF	MEAN	5.874546e-04	7.425605e-02	2.018331e-01	5.591398e+01	3.580742e+00	1.151105e+01
	RMS	2.423746e-02	2.724996e-01	4.492584e-01	7.477565e+00	1.892285e+00	3.392794e+00
	MEDIAN	5.874546e-04	7.425605e-02	2.018331e-01	5.591398e+01	3.580742e+00	1.151105e+01
	PERC25	5.874546e-04	7.425605e-02	2.018331e-01	5.591398e+01	3.580742e+00	1.151105e+01
	PERC975	5.874546e-04	7.425605e-02	2.018331e-01	5.591398e+01	3.580742e+00	1.151105e+01

Fonte: O autor.

modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal PETR4 na frequência semanal. Os resultados podem ser observados na tabela 19.

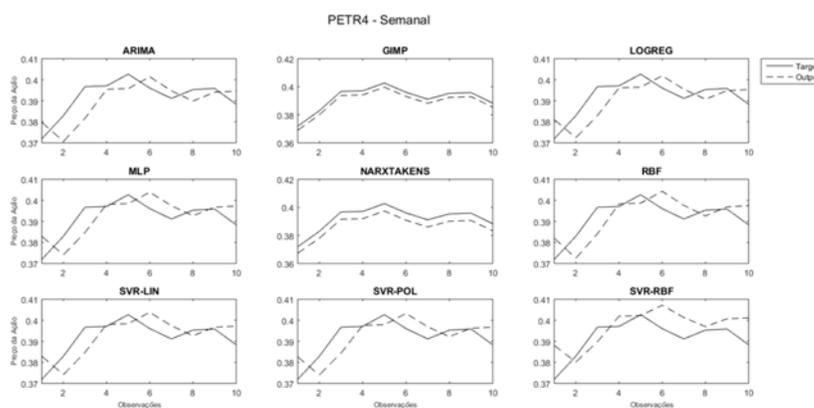
Tabela 19 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série PETR4 semanal (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Através da figura 57 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal PETR4 semanal. O ganho do modelo GIMP em relação aos outros modelos para esta série pode ser

Figura 57 – Gráfico de previsão da série temporal PETR4 semanal (conjunto de teste).



Fonte: O autor.

vista na tabela 20

### 6.3.3 PETR4 - Frequência Quinzenal

Os experimentos com a série temporal PETR4 na frequência quinzenal foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 21. Foram utilizados os valores de fechamento para a criação de cada série.

Tabela 20 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal PETR4 Semanal (conjunto de teste)

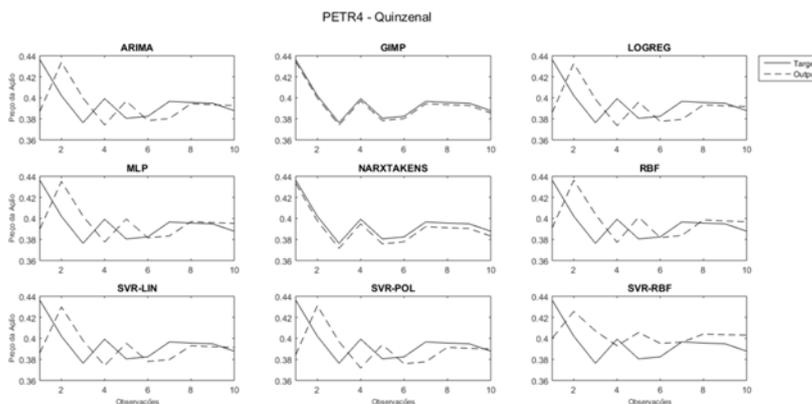
Modelo Proposto/Modelo Investigado	Ganho Gimp					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	88,6431284	61,312396	88,9415659	78,846149	88,7357301	234,129796
GIMP/LOGREG	92,6094132	73,5177404	94,1918337	78,846149	92,6731621	329,911671
GIMP/MLP	89,7904913	66,2126626	90,6039258	78,846149	89,8511726	251,864375
GIMP/NARXTAKENS	80,9207621	55,295091	87,3576924	0	81,0524457	38,3288227
GIMP/RBF	99,2813861	89,2206423	99,5048688	97,8723577	99,2887919	3110,31245
GIMP/SVR-LIN	89,7509722	66,5582363	90,6583365	78,846149	89,813467	253,362194
GIMP/SVR-LOG	91,2811356	70,6166211	92,5610426	78,846149	91,3437573	287,978055
GIMP/SVR-RBF	96,7953864	83,575978	97,5445618	78,846149	96,8200194	677,536193

Fonte: O autor.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal PETR4 na frequência quinzenal. Os resultados podem ser observados na tabela tabela 22.

Através da figura 58 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal PETR4 quinzenal.

Figura 58 – Gráfico de previsão da série temporal PETR4 quinzenal (conjunto de teste).



Fonte: O autor.

O ganho do modelo GIMP em relação aos outros modelos para esta série pode ser vista na tabela 23

### 6.3.4 VALE5 - Frequência Diária

Os experimentos com a série temporal VALE5 na frequência diária foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 24. Foram utilizados os valores de fechamento para a criação de cada série.

Tabela 21 – Desempenho preditivo para a série PETR4 - Frequência Quinzenal (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	2.189636e-05	1.331458e-02	4.771461e-03	1.000000e+02	5.859639e-02	9.328237e+01
	RMS	4.679355e-03	1.153888e-01	6.907576e-02	1.000000e+01	2.420669e-01	9.658280e+00
	MEDIAN	1.575687e-05	1.330098e-02	3.455885e-03	1.000000e+02	4.216663e-02	9.443471e+01
	PERC25	5.503845e-07	2.486682e-03	1.211184e-04	1.000000e+02	1.472873e-03	7.988888e+01
	PERC975	7.684638e-05	2.938318e-02	1.663168e-02	1.000000e+02	2.056470e-01	9.959354e+01
ARIMA	MEAN	3.774662e-04	4.530904e-02	8.528012e-02	6.470588e+01	1.038032e+00	2.983214e+01
	RMS	1.942849e-02	2.128592e-01	2.920276e-01	8.043997e+00	1.018839e+00	5.461881e+00
	MEDIAN	3.774662e-04	4.530904e-02	8.528012e-02	6.470588e+01	1.038032e+00	2.983214e+01
	PERC25	3.774662e-04	4.530904e-02	8.528012e-02	6.470588e+01	1.038032e+00	2.983214e+01
	PERC975	3.774662e-04	4.530904e-02	8.528012e-02	6.470588e+01	1.038032e+00	2.983214e+01
LOGREG	MEAN	3.862837e-04	5.371062e-02	1.071607e-01	6.470588e+01	1.063028e+00	2.909064e+01
	RMS	1.965410e-02	2.317555e-01	3.273541e-01	8.043997e+00	1.031032e+00	5.393574e+00
	MEDIAN	3.862837e-04	5.371062e-02	1.071607e-01	6.470588e+01	1.063028e+00	2.909064e+01
	PERC25	3.862837e-04	5.371062e-02	1.071607e-01	6.470588e+01	1.063028e+00	2.909064e+01
	PERC975	3.862837e-04	5.371062e-02	1.071607e-01	6.470588e+01	1.063028e+00	2.909064e+01
MLP	MEAN	3.557622e-04	4.771948e-02	8.667972e-02	6.470588e+01	9.746442e-01	3.067777e+01
	RMS	1.886166e-02	2.184479e-01	2.944142e-01	8.043997e+00	9.872407e-01	5.538752e+00
	MEDIAN	3.512763e-04	4.671697e-02	8.768271e-02	6.470588e+01	9.659361e-01	3.080212e+01
	PERC25	3.496506e-04	4.525096e-02	8.171293e-02	6.470588e+01	9.609299e-01	3.001355e+01
	PERC975	3.743036e-04	5.181540e-02	9.009750e-02	6.470588e+01	1.016025e+00	3.093331e+01
NARXTAKENS	MEAN	6.345041e-05	2.259774e-02	1.758380e-02	1.000000e+02	1.729888e-01	8.281740e+01
	RMS	7.965577e-03	1.503254e-01	1.326039e-01	1.000000e+01	4.159192e-01	9.100407e+00
	MEDIAN	5.620078e-05	2.157595e-02	1.507445e-02	1.000000e+02	1.533668e-01	8.403482e+01
	PERC25	3.194980e-05	1.627101e-02	8.166919e-03	1.000000e+02	8.731141e-02	6.976534e+01
	PERC975	1.321401e-04	3.412789e-02	4.028821e-02	1.000000e+02	3.588283e-01	8.994575e+01
RBF	MEAN	8.080290e-04	7.992501e-02	2.650792e-01	6.176471e+01	2.219614e+00	1.732323e+01
	RMS	2.842585e-02	2.827101e-01	5.148584e-01	7.859052e+00	1.489837e+00	4.162119e+00
	MEDIAN	8.080290e-04	7.992501e-02	2.650792e-01	6.176471e+01	2.219614e+00	1.732323e+01
	PERC25	8.080290e-04	7.992501e-02	2.650792e-01	6.176471e+01	2.219614e+00	1.732323e+01
	PERC975	8.080290e-04	7.992501e-02	2.650792e-01	6.176471e+01	2.219614e+00	1.732323e+01
SVR-LIN	MEAN	3.558304e-04	4.839409e-02	9.273107e-02	6.470588e+01	9.792632e-01	3.051093e+01
	RMS	1.886347e-02	2.199866e-01	3.045178e-01	8.043997e+00	9.895773e-01	5.523670e+00
	MEDIAN	3.558304e-04	4.839409e-02	9.273107e-02	6.470588e+01	9.792632e-01	3.051093e+01
	PERC25	3.558304e-04	4.839409e-02	9.273107e-02	6.470588e+01	9.792632e-01	3.051093e+01
	PERC975	3.558304e-04	4.839409e-02	9.273107e-02	6.470588e+01	9.792632e-01	3.051093e+01
SVR-POL	MEAN	3.760508e-04	5.031189e-02	9.876657e-02	6.470588e+01	1.035637e+00	2.961244e+01
	RMS	1.939203e-02	2.243031e-01	3.142715e-01	8.043997e+00	1.017662e+00	5.441732e+00
	MEDIAN	3.760508e-04	5.031189e-02	9.876657e-02	6.470588e+01	1.035637e+00	2.961244e+01
	PERC25	3.760508e-04	5.031189e-02	9.876657e-02	6.470588e+01	1.035637e+00	2.961244e+01
	PERC975	3.760508e-04	5.031189e-02	9.876657e-02	6.470588e+01	1.035637e+00	2.961244e+01
SVR-RBF	MEAN	3.144504e-03	1.791007e-01	9.802292e-01	6.470588e+01	8.644126e+00	5.987626e+00
	RMS	5.607588e-02	4.232029e-01	9.900653e-01	8.043997e+00	2.940089e+00	2.446963e+00
	MEDIAN	3.144504e-03	1.791007e-01	9.802292e-01	6.470588e+01	8.644126e+00	5.987626e+00
	PERC25	3.144504e-03	1.791007e-01	9.802292e-01	6.470588e+01	8.644126e+00	5.987626e+00
	PERC975	3.144504e-03	1.791007e-01	9.802292e-01	6.470588e+01	8.644126e+00	5.987626e+00

Fonte: O autor.

Tabela 22 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série PETR4 quinzenal (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

**Fonte:** O autor.

Tabela 23 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal PETR4 quinzenal(conjunto de teste)

Modelo Proposto/Modelo Investigado	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	94,1991203	70,6138554	94,4049551	54,5454602	94,3550497	212,690843
GIMP/LOGREG	94,3315341	75,2105263	95,5473779	54,5454602	94,4877849	220,661113
GIMP/MLP	93,8452258	72,0982291	94,4952972	54,5454602	93,9879199	204,071548
GIMP/NARXTAKENS	65,4905934	41,0800372	72,8644491	0	66,1270614	12,6361972
GIMP/RBF	97,2901517	83,3411594	98,1999866	61,9047511	97,360064	438,481392
GIMP/SVR-LIN	93,8464055	72,4871777	94,8545175	54,5454602	94,0162777	205,734273
GIMP/SVR-LOG	94,1772867	73,5359177	95,1689514	54,5454602	94,3419953	215,010752
GIMP/SVR-RBF	99,3036625	92,5658694	99,5132301	54,5454602	99,3221248	1457,91912

**Fonte:** O autor.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal VALE5 na frequência diária. Os resultados podem ser observados na tabela tabela 25.

Através da figura 59 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal VALE5 diária.

A tabela tabela 26 mostra o ganho do modelo GIMP em relação aos outros modelos estudados.

### 6.3.5 VALE5 - Frequência Semanal

Os experimentos com a série temporal VALE5 na frequência semanal foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 27. Foram utilizados os valores de fechamento para a criação de cada série.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal VALE5

Tabela 24 – Desempenho preditivo para a série VALE5 - Frequência Diária (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	3.286745e-06	4.189648e-03	3.204570e-04	9.907579e+01	2.924165e-02	9.597781e+01
	RMS	1.812938e-03	6.472749e-02	1.790131e-02	9.953682e+00	1.710019e-01	9.796827e+00
	MEDIAN	1.456631e-06	3.462505e-03	1.421187e-04	9.907579e+01	1.295941e-02	9.746133e+01
	PERC25	6.837798e-08	7.503771e-04	6.672319e-06	9.907579e+01	6.083481e-04	8.720891e+01
	PERC975	1.393634e-05	1.071262e-02	1.358070e-03	9.907579e+01	1.239894e-01	9.894069e+01
ARIMA	MEAN	1.123958e-04	2.168310e-02	1.111285e-02	4.879852e+01	1.001815e+00	2.398288e+01
	RMS	1.060169e-02	1.472518e-01	1.054175e-01	6.985594e+00	1.000907e+00	4.897232e+00
	MEDIAN	1.123958e-04	2.168310e-02	1.111285e-02	4.879852e+01	1.001815e+00	2.398288e+01
	PERC25	1.123958e-04	2.168310e-02	1.111285e-02	4.879852e+01	1.001815e+00	2.398288e+01
	PERC975	1.123958e-04	2.168310e-02	1.111285e-02	4.879852e+01	1.001815e+00	2.398288e+01
LOGREG	MEAN	1.263943e-04	2.455762e-02	1.290508e-02	4.879852e+01	1.126536e+00	2.254885e+01
	RMS	1.124252e-02	1.567087e-01	1.136005e-01	6.985594e+00	1.061384e+00	4.748563e+00
	MEDIAN	1.263943e-04	2.455762e-02	1.290508e-02	4.879852e+01	1.126536e+00	2.254885e+01
	PERC25	1.263943e-04	2.455762e-02	1.290508e-02	4.879852e+01	1.126536e+00	2.254885e+01
	PERC975	1.263943e-04	2.455762e-02	1.290508e-02	4.879852e+01	1.126536e+00	2.254885e+01
MLP	MEAN	1.130534e-04	2.197392e-02	1.129214e-02	4.879852e+01	1.007676e+00	2.390854e+01
	RMS	1.063266e-02	1.482360e-01	1.062645e-01	6.985594e+00	1.003831e+00	4.889636e+00
	MEDIAN	1.132460e-04	2.201560e-02	1.131712e-02	4.879852e+01	1.009394e+00	2.388759e+01
	PERC25	1.123597e-04	2.182427e-02	1.119194e-02	4.879852e+01	1.001492e+00	2.384730e+01
	PERC975	1.136194e-04	2.208123e-02	1.137636e-02	4.879852e+01	1.012720e+00	2.398409e+01
NARXTAKENS	MEAN	1.732508e-04	2.871916e-02	1.958943e-02	9.907579e+01	1.544123e+00	4.108007e+01
	RMS	1.316248e-02	1.694673e-01	1.399622e-01	9.953682e+00	1.242627e+00	6.409374e+00
	MEDIAN	1.482643e-04	2.725416e-02	1.634014e-02	9.907579e+01	1.321402e+00	4.191155e+01
	PERC25	8.211337e-05	1.915968e-02	8.713072e-03	9.907579e+01	7.318288e-01	2.122378e+01
	PERC975	4.007350e-04	4.731874e-02	4.860599e-02	9.907579e+01	3.571825e+00	5.629998e+01
RBF	MEAN	1.823728e-04	3.215399e-02	1.978859e-02	4.879852e+01	1.625541e+00	1.822428e+01
	RMS	1.350455e-02	1.793153e-01	1.406719e-01	6.985594e+00	1.274967e+00	4.268990e+00
	MEDIAN	1.823728e-04	3.215399e-02	1.978859e-02	4.879852e+01	1.625541e+00	1.822428e+01
	PERC25	1.823728e-04	3.215399e-02	1.978859e-02	4.879852e+01	1.625541e+00	1.822428e+01
	PERC975	1.823728e-04	3.215399e-02	1.978859e-02	4.879852e+01	1.625541e+00	1.822428e+01
SVR-LIN	MEAN	1.126734e-04	2.185160e-02	1.128674e-02	4.879852e+01	1.004270e+00	2.394995e+01
	RMS	1.061477e-02	1.478229e-01	1.062391e-01	6.985594e+00	1.002133e+00	4.893869e+00
	MEDIAN	1.126734e-04	2.185160e-02	1.128674e-02	4.879852e+01	1.004270e+00	2.394995e+01
	PERC25	1.126734e-04	2.185160e-02	1.128674e-02	4.879852e+01	1.004270e+00	2.394995e+01
	PERC975	1.126734e-04	2.185160e-02	1.128674e-02	4.879852e+01	1.004270e+00	2.394995e+01
SVR-POL	MEAN	1.323701e-04	2.532089e-02	1.381817e-02	4.879852e+01	1.179837e+00	2.199014e+01
	RMS	1.150522e-02	1.591254e-01	1.175507e-01	6.985594e+00	1.086203e+00	4.689365e+00
	MEDIAN	1.323701e-04	2.532089e-02	1.381817e-02	4.879852e+01	1.179837e+00	2.199014e+01
	PERC25	1.323701e-04	2.532089e-02	1.381817e-02	4.879852e+01	1.179837e+00	2.199014e+01
	PERC975	1.323701e-04	2.532089e-02	1.381817e-02	4.879852e+01	1.179837e+00	2.199014e+01
SVR-RBF	MEAN	3.416068e-04	4.699615e-02	3.963044e-02	4.879852e+01	3.044835e+00	1.181046e+01
	RMS	1.848261e-02	2.167860e-01	1.990740e-01	6.985594e+00	1.744946e+00	3.436636e+00
	MEDIAN	3.416068e-04	4.699615e-02	3.963044e-02	4.879852e+01	3.044835e+00	1.181046e+01
	PERC25	3.416068e-04	4.699615e-02	3.963044e-02	4.879852e+01	3.044835e+00	1.181046e+01
	PERC975	3.416068e-04	4.699615e-02	3.963044e-02	4.879852e+01	3.044835e+00	1.181046e+01

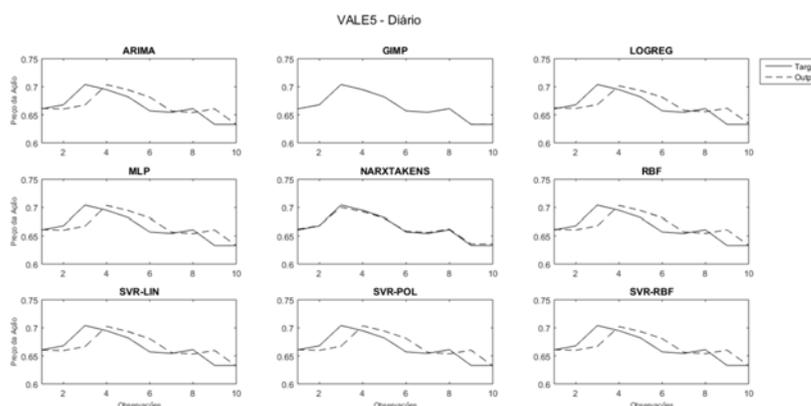
Fonte: O autor.

Tabela 25 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série VALE5 diário (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Figura 59 – Gráfico de previsão da série temporal VALE5 diária (conjunto de teste).



Fonte: O autor.

Tabela 26 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal VALE5 Diário

Modelo Proposto/Modelo Investigado	Ganho GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	97,0757404	80,6778182	97,1163383	103,030317	97,0811327	300,193013
GIMP/LOGREG	97,3996098	82,9395194	97,5168151	103,030317	97,4042862	325,643924
GIMP/MLP	97,09275	80,933543	97,1621234	103,030317	97,0981099	301,437353
GIMP/NARXTAKENS	98,1028976	85,4116625	98,3641331	0	98,1062616	133,635946
GIMP/RBF	98,1977877	86,9700526	98,3805971	103,030317	98,2011127	426,648021
GIMP/SVR-LIN	97,082945	80,8268136	97,1607656	103,030317	97,0882681	300,743258
GIMP/SVR-LOG	97,5170035	83,4537886	97,6809013	103,030317	97,5215517	336,458385
GIMP/SVR-RBF	99,0378573	91,0851251	99,1913867	103,030317	99,039631	712,650904

Fonte: O autor.

na frequência semanal. Os resultados podem ser observados na tabela tabela 28.

Através da figura 60 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal VALE5 semanal.

Tabela 27 – Desempenho preditivo para a série VALE5 - Frequência Semanal (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	1.338059e-05	9.256775e-03	1.118884e-03	1.000000e+02	2.411201e-02	9.674890e+01
	RMS	3.657949e-03	9.621214e-02	3.344972e-02	1.000000e+01	1.552804e-01	9.836102e+00
	MEDIAN	7.071141e-06	7.723734e-03	5.922257e-04	1.000000e+02	1.274230e-02	9.794076e+01
	PERC25	6.359900e-08	7.457232e-04	5.330132e-06	1.000000e+02	1.146063e-04	9.071301e+01
	PERC975	4.388456e-05	1.958883e-02	3.664438e-03	1.000000e+02	7.908062e-02	9.991350e+01
ARIMA	MEAN	5.546681e-04	4.842100e-02	5.030651e-02	5.483871e+01	1.007238e+00	2.603284e+01
	RMS	2.355139e-02	2.200477e-01	2.242911e-01	7.405316e+00	1.003613e+00	5.102239e+00
	MEDIAN	5.546681e-04	4.842100e-02	5.030651e-02	5.483871e+01	1.007238e+00	2.603284e+01
	PERC25	5.546681e-04	4.842100e-02	5.030651e-02	5.483871e+01	1.007238e+00	2.603284e+01
	PERC975	5.546681e-04	4.842100e-02	5.030651e-02	5.483871e+01	1.007238e+00	2.603284e+01
LOGREG	MEAN	5.556132e-04	5.048684e-02	5.293770e-02	5.483871e+01	1.008768e+00	2.595610e+01
	RMS	2.357145e-02	2.246928e-01	2.300819e-01	7.405316e+00	1.004375e+00	5.094713e+00
	MEDIAN	5.556132e-04	5.048684e-02	5.293770e-02	5.483871e+01	1.008768e+00	2.595610e+01
	PERC25	5.556132e-04	5.048684e-02	5.293770e-02	5.483871e+01	1.008768e+00	2.595610e+01
	PERC975	5.556132e-04	5.048684e-02	5.293770e-02	5.483871e+01	1.008768e+00	2.595610e+01
MLP	MEAN	5.499402e-04	4.876722e-02	5.102083e-02	5.483871e+01	9.988931e-01	2.612324e+01
	RMS	2.345080e-02	2.208330e-01	2.258779e-01	7.405316e+00	9.994464e-01	5.111090e+00
	MEDIAN	5.501524e-04	4.878950e-02	5.109385e-02	5.483871e+01	9.993214e-01	2.611662e+01
	PERC25	5.491424e-04	4.866085e-02	5.076403e-02	5.483871e+01	9.973040e-01	2.611165e+01
	PERC975	5.503258e-04	4.888912e-02	5.113233e-02	5.483871e+01	9.996895e-01	2.614756e+01
NARXTAKENS	MEAN	1.903779e-04	3.151307e-02	1.835786e-02	1.000000e+02	3.465388e-01	7.351935e+01
	RMS	1.379775e-02	1.775192e-01	1.354912e-01	1.000000e+01	5.886754e-01	8.574343e+00
	MEDIAN	1.829425e-04	3.238467e-02	1.719137e-02	1.000000e+02	3.330843e-01	7.232368e+01
	PERC25	3.874600e-05	1.318028e-02	3.257225e-03	1.000000e+02	6.999246e-02	5.053011e+01
	PERC975	4.792392e-04	5.584329e-02	4.984879e-02	1.000000e+02	8.728466e-01	9.204131e+01
RBF	MEAN	1.781848e-02	2.996974e-01	1.334162e+00	5.053763e+01	3.245204e+01	1.439666e+00
	RMS	1.334859e-01	5.474462e-01	1.155059e+00	7.108983e+00	5.696670e+00	1.199861e+00
	MEDIAN	1.781848e-02	2.996974e-01	1.334162e+00	5.053763e+01	3.245204e+01	1.439666e+00
	PERC25	1.781848e-02	2.996974e-01	1.334162e+00	5.053763e+01	3.245204e+01	1.439666e+00
	PERC975	1.781848e-02	2.996974e-01	1.334162e+00	5.053763e+01	3.245204e+01	1.439666e+00
SVR-LIN	MEAN	5.541666e-04	4.899455e-02	5.209886e-02	5.483871e+01	1.007337e+00	2.600243e+01
	RMS	2.354074e-02	2.213471e-01	2.282517e-01	7.405316e+00	1.003662e+00	5.099258e+00
	MEDIAN	5.541666e-04	4.899455e-02	5.209886e-02	5.483871e+01	1.007337e+00	2.600243e+01
	PERC25	5.541666e-04	4.899455e-02	5.209886e-02	5.483871e+01	1.007337e+00	2.600243e+01
	PERC975	5.541666e-04	4.899455e-02	5.209886e-02	5.483871e+01	1.007337e+00	2.600243e+01
SVR-POL	MEAN	5.753130e-04	5.094014e-02	5.614666e-02	5.483871e+01	1.045401e+00	2.547009e+01
	RMS	2.398568e-02	2.256992e-01	2.369529e-01	7.405316e+00	1.022448e+00	5.046790e+00
	MEDIAN	5.753130e-04	5.094014e-02	5.614666e-02	5.483871e+01	1.045401e+00	2.547009e+01
	PERC25	5.753130e-04	5.094014e-02	5.614666e-02	5.483871e+01	1.045401e+00	2.547009e+01
	PERC975	5.753130e-04	5.094014e-02	5.614666e-02	5.483871e+01	1.045401e+00	2.547009e+01
SVR-RBF	MEAN	1.157059e-03	9.099254e-02	1.305268e-01	5.483871e+01	2.105724e+00	1.647599e+01
	RMS	3.401556e-02	3.016497e-01	3.612849e-01	7.405316e+00	1.451111e+00	4.059063e+00
	MEDIAN	1.157059e-03	9.099254e-02	1.305268e-01	5.483871e+01	2.105724e+00	1.647599e+01
	PERC25	1.157059e-03	9.099254e-02	1.305268e-01	5.483871e+01	2.105724e+00	1.647599e+01
	PERC975	1.157059e-03	9.099254e-02	1.305268e-01	5.483871e+01	2.105724e+00	1.647599e+01

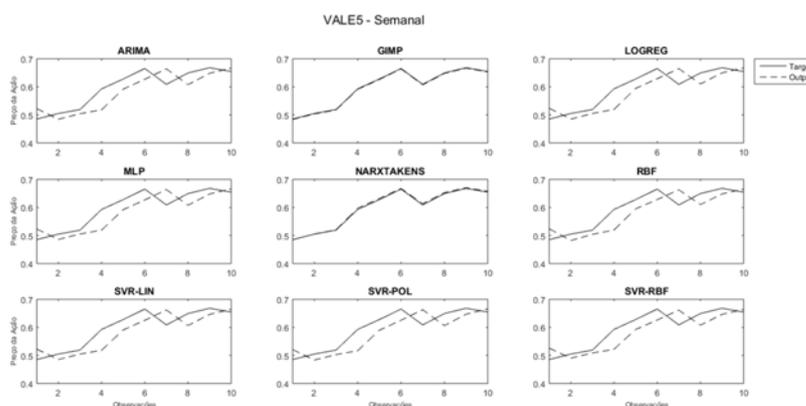
Fonte: O autor.

Tabela 28 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série VALE5 semanal (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Figura 60 – Gráfico de previsão da série temporal VALE5 semanal (conjunto de teste).



Fonte: O autor.

Tabela 29 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal VALE5 Semanal (conjunto de teste)

Modelo Proposto/Modelo Investigado	Ganho GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	97,5876402	80,8827265	97,7758664	82,3529401	97,6061259	271,641742
GIMP/LOGREG	97,5917437	81,6649745	97,8864137	82,3529401	97,6097567	272,740512
GIMP/MLP	97,5669009	81,0184485	97,8070055	82,3529401	97,5861271	270,355668
GIMP/NARXTAKENS	92,9715634	70,6256007	93,9051502	0	93,0420461	31,5965117
GIMP/RBF	99,9249061	96,9112929	99,9161358	97,8723577	99,9256996	6620,23233
GIMP/SVR-LIN	97,5854572	81,106521	97,852383	82,3529401	97,6063611	272,076379
GIMP/SVR-LOG	97,6742069	81,828132	98,0072118	82,3529401	97,6935157	279,852996
GIMP/SVR-RBF	98,8435689	89,8268858	99,1427937	82,3529401	98,8549302	487,211451

Fonte: O autor.

A tabela 29 mostra o ganho do modelo GIMP em relação aos outros modelos estudados.

### 6.3.6 VALE5 - Frequência Quinzenal

Os experimentos com a série temporal VALE5 na frequência quinzenal foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 30.

Foram utilizados os valores de fechamento para a criação de cada série.

Tabela 30 – Desempenho preditivo para a série VALE5 - Frequência Quinzenal (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	1.985700e-05	1.243216e-02	1.071837e-03	1.000000e+02	1.495320e-02	9.726888e+01
	RMS	4.456120e-03	1.114996e-01	3.273892e-02	1.000000e+01	1.222833e-01	9.862499e+00
	MEDIAN	1.014975e-05	9.684836e-03	5.486778e-04	1.000000e+02	7.643208e-03	9.824276e+01
	PERC25	5.850229e-06	7.353419e-03	3.163268e-04	1.000000e+02	4.405480e-03	9.254145e+01
	PERC975	6.859208e-05	2.517908e-02	3.696292e-03	1.000000e+02	5.165286e-02	9.880631e+01
ARIMA	MEAN	1.390395e-03	7.823712e-02	8.816805e-02	6.176471e+01	1.075633e+00	2.753139e+01
	RMS	3.728801e-02	2.797090e-01	2.969310e-01	7.859052e+00	1.037127e+00	5.247036e+00
	MEDIAN	1.390395e-03	7.823712e-02	8.816805e-02	6.176471e+01	1.075633e+00	2.753139e+01
	PERC25	1.390395e-03	7.823712e-02	8.816805e-02	6.176471e+01	1.075633e+00	2.753139e+01
	PERC975	1.390395e-03	7.823712e-02	8.816805e-02	6.176471e+01	1.075633e+00	2.753139e+01
LOGREG	MEAN	1.380925e-03	8.062791e-02	9.693022e-02	6.176471e+01	1.067406e+00	2.749564e+01
	RMS	3.716080e-02	2.839505e-01	3.113362e-01	7.859052e+00	1.033154e+00	5.243629e+00
	MEDIAN	1.380925e-03	8.062791e-02	9.693022e-02	6.176471e+01	1.067406e+00	2.749564e+01
	PERC25	1.380925e-03	8.062791e-02	9.693022e-02	6.176471e+01	1.067406e+00	2.749564e+01
	PERC975	1.380925e-03	8.062791e-02	9.693022e-02	6.176471e+01	1.067406e+00	2.749564e+01
MLP	MEAN	1.356400e-03	7.633476e-02	8.892350e-02	6.176471e+01	1.048615e+00	2.788197e+01
	RMS	3.682933e-02	2.762875e-01	2.982004e-01	7.859052e+00	1.024019e+00	5.280338e+00
	MEDIAN	1.356678e-03	7.619965e-02	8.913353e-02	6.176471e+01	1.048655e+00	2.787858e+01
	PERC25	1.350071e-03	7.617171e-02	8.782342e-02	6.176471e+01	1.044218e+00	2.773147e+01
	PERC975	1.370939e-03	7.714547e-02	8.927859e-02	6.176471e+01	1.060073e+00	2.794920e+01
NARXTAKENS	MEAN	1.730202e-04	3.207645e-02	1.054175e-02	1.000000e+02	1.338855e-01	8.554634e+01
	RMS	1.315371e-02	1.790990e-01	1.026730e-01	1.000000e+01	3.659036e-01	9.249126e+00
	MEDIAN	1.924807e-04	3.544583e-02	1.151671e-02	1.000000e+02	1.492029e-01	8.358860e+01
	PERC25	3.200735e-05	1.308265e-02	1.797788e-03	1.000000e+02	2.375673e-02	7.563102e+01
	PERC975	3.269280e-04	4.793672e-02	2.052113e-02	1.000000e+02	2.534240e-01	9.627705e+01
RBF	MEAN	4.435706e-02	4.991776e-01	1.467716e+00	5.000000e+01	3.438289e+01	1.337108e+00
	RMS	2.106112e-01	7.065250e-01	1.211493e+00	7.071068e+00	5.863693e+00	1.156334e+00
	MEDIAN	4.435706e-02	4.991776e-01	1.467716e+00	5.000000e+01	3.438289e+01	1.337108e+00
	PERC25	4.435706e-02	4.991776e-01	1.467716e+00	5.000000e+01	3.438289e+01	1.337108e+00
	PERC975	4.435706e-02	4.991776e-01	1.467716e+00	5.000000e+01	3.438289e+01	1.337108e+00
SVR-LIN	MEAN	1.362946e-03	7.601084e-02	9.042365e-02	6.176471e+01	1.052945e+00	2.781263e+01
	RMS	3.691809e-02	2.757006e-01	3.007053e-01	7.859052e+00	1.026131e+00	5.273768e+00
	MEDIAN	1.362946e-03	7.601084e-02	9.042365e-02	6.176471e+01	1.052945e+00	2.781263e+01
	PERC25	1.362946e-03	7.601084e-02	9.042365e-02	6.176471e+01	1.052945e+00	2.781263e+01
	PERC975	1.362946e-03	7.601084e-02	9.042365e-02	6.176471e+01	1.052945e+00	2.781263e+01
SVR-POL	MEAN	1.468649e-03	8.120304e-02	1.046839e-01	6.176471e+01	1.136511e+00	2.657842e+01
	RMS	3.832295e-02	2.849615e-01	3.235490e-01	7.859052e+00	1.066073e+00	5.155426e+00
	MEDIAN	1.468649e-03	8.120304e-02	1.046839e-01	6.176471e+01	1.136511e+00	2.657842e+01
	PERC25	1.468649e-03	8.120304e-02	1.046839e-01	6.176471e+01	1.136511e+00	2.657842e+01
	PERC975	1.468649e-03	8.120304e-02	1.046839e-01	6.176471e+01	1.136511e+00	2.657842e+01
SVR-RBF	MEAN	1.938141e-03	1.108124e-01	1.658932e-01	6.176471e+01	1.490068e+00	2.230810e+01
	RMS	4.402433e-02	3.328849e-01	4.072999e-01	7.859052e+00	1.220683e+00	4.723146e+00
	MEDIAN	1.938141e-03	1.108124e-01	1.658932e-01	6.176471e+01	1.490068e+00	2.230810e+01
	PERC25	1.938141e-03	1.108124e-01	1.658932e-01	6.176471e+01	1.490068e+00	2.230810e+01
	PERC975	1.938141e-03	1.108124e-01	1.658932e-01	6.176471e+01	1.490068e+00	2.230810e+01

Fonte: O autor.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal VALE5 na frequência quinzenal. Os resultados podem ser observados na tabela 31.

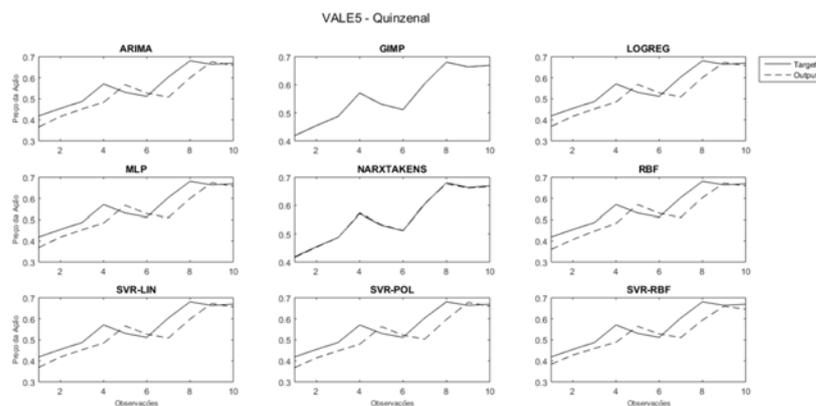
Através da figura 61 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal VALE5 diária.

Tabela 31 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série VALE5 quinzenal (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão(h)					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Figura 61 – Gráfico de previsão da série temporal VALE5 quinzenal (conjunto de teste).



Fonte: O autor.

Tabela 32 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal VALE5 quinzenal (conjunto de teste)

Modelo Proposto/Modelo Investigado	Ganho GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	98,5718447	84,10964	98,7843249	61,9047511	98,6098232	253,30174
GIMP/LOGREG	98,5620508	84,5808232	98,8942179	61,9047511	98,5991085	253,761105
GIMP/MLP	98,5360513	83,7136319	98,7946527	61,9047511	98,5740048	248,859424
GIMP/NARXTAKENS	88,5233054	61,2420951	89,8324567	0	88,8313522	13,7031462
GIMP/RBF	99,9552337	97,5094716	99,9269725	100	99,9565098	7174,57169
GIMP/SVR-LIN	98,5430824	83,6442276	98,8146497	61,9047511	98,5798688	249,72917
GIMP/SVR-LOG	98,6479411	84,6900313	98,9761205	61,9047511	98,684289	265,969384
GIMP/SVR-RBF	98,9754615	88,7808946	99,3538994	61,9047511	98,9964753	336,024942

Fonte: O autor.

A tabela 32 mostra o ganho do modelo GIMP em relação aos outros modelos estudados.

### 6.3.7 CSNA3 - Frequência Diária

Os experimentos com a série temporal CSNA3 na frequência diária foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 33. Foram

utilizados os valores de fechamento para a criação de cada série.

Tabela 33 – Desempenho preditivo para a série CSNA3 - Frequência Diária (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	8.663509e-06	8.684991e-03	1.674615e-03	9.833641e+01	1.089667e-01	8.851747e+01
	RMS	2.943384e-03	9.319330e-02	4.092207e-02	9.916472e+00	3.301010e-01	9.408372e+00
	MEDIAN	7.792176e-06	9.366886e-03	1.508139e-03	9.833641e+01	9.800734e-02	8.872066e+01
	PERC25	2.203313e-07	1.585522e-03	4.271001e-05	9.833641e+01	2.771251e-03	7.670127e+01
	PERC975	2.087708e-05	1.543367e-02	4.030769e-03	9.833641e+01	2.625849e-01	9.790566e+01
ARIMA	MEAN	7.958836e-05	1.952257e-02	1.551500e-02	5.304991e+01	1.001963e+00	2.604213e+01
	RMS	8.921231e-03	1.397232e-01	1.245592e-01	7.283537e+00	1.000981e+00	5.103149e+00
	MEDIAN	7.958836e-05	1.952257e-02	1.551500e-02	5.304991e+01	1.001963e+00	2.604213e+01
	PERC25	7.958836e-05	1.952257e-02	1.551500e-02	5.304991e+01	1.001963e+00	2.604213e+01
	PERC975	7.958836e-05	1.952257e-02	1.551500e-02	5.304991e+01	1.001963e+00	2.604213e+01
LOGREG	MEAN	1.260350e-04	3.161379e-02	2.859294e-02	5.304991e+01	1.586579e+00	2.004219e+01
	RMS	1.122653e-02	1.778027e-01	1.690945e-01	7.283537e+00	1.259595e+00	4.476851e+00
	MEDIAN	1.260350e-04	3.161379e-02	2.859294e-02	5.304991e+01	1.586579e+00	2.004219e+01
	PERC25	1.260350e-04	3.161379e-02	2.859294e-02	5.304991e+01	1.586579e+00	2.004219e+01
	PERC975	1.260350e-04	3.161379e-02	2.859294e-02	5.304991e+01	1.586579e+00	2.004219e+01
MLP	MEAN	7.967602e-05	1.956261e-02	1.560915e-02	5.304991e+01	1.003020e+00	2.602804e+01
	RMS	8.926142e-03	1.398664e-01	1.249366e-01	7.283537e+00	1.001509e+00	5.101768e+00
	MEDIAN	7.931609e-05	1.952434e-02	1.546618e-02	5.304991e+01	9.986185e-01	2.608556e+01
	PERC25	7.931607e-05	1.952069e-02	1.546618e-02	5.304991e+01	9.986183e-01	2.551884e+01
	PERC975	8.286124e-05	1.991295e-02	1.688914e-02	5.304991e+01	1.041967e+00	2.608557e+01
NARXTAKENS	MEAN	7.110575e-05	2.283282e-02	1.727066e-02	9.833641e+01	8.941298e-01	5.616659e+01
	RMS	8.432423e-03	1.511053e-01	1.314179e-01	9.916472e+00	9.455844e-01	7.494437e+00
	MEDIAN	4.871132e-05	2.026203e-02	1.099883e-02	9.833641e+01	6.127445e-01	6.004806e+01
	PERC25	1.056657e-05	9.223075e-03	2.078903e-03	9.833641e+01	1.329203e-01	2.936923e+01
	PERC975	1.799363e-04	3.851725e-02	4.744283e-02	9.833641e+01	2.262140e+00	8.594091e+01
RBF	MEAN	8.792705e-05	2.282829e-02	1.830758e-02	5.304991e+01	1.106881e+00	2.469615e+01
	RMS	9.376942e-03	1.510903e-01	1.353055e-01	7.283537e+00	1.052084e+00	4.969522e+00
	MEDIAN	8.792705e-05	2.282829e-02	1.830758e-02	5.304991e+01	1.106881e+00	2.469615e+01
	PERC25	8.792705e-05	2.282829e-02	1.830758e-02	5.304991e+01	1.106881e+00	2.469615e+01
	PERC975	8.792705e-05	2.282829e-02	1.830758e-02	5.304991e+01	1.106881e+00	2.469615e+01
SVR-LIN	MEAN	8.266133e-05	2.124747e-02	1.651733e-02	5.304991e+01	1.040955e+00	2.551946e+01
	RMS	9.091827e-03	1.457651e-01	1.285198e-01	7.283537e+00	1.020272e+00	5.051679e+00
	MEDIAN	8.266133e-05	2.124747e-02	1.651733e-02	5.304991e+01	1.040955e+00	2.551946e+01
	PERC25	8.266133e-05	2.124747e-02	1.651733e-02	5.304991e+01	1.040955e+00	2.551946e+01
	PERC975	8.266133e-05	2.124747e-02	1.651733e-02	5.304991e+01	1.040955e+00	2.551946e+01
SVR-POL	MEAN	1.250107e-04	3.198963e-02	2.746480e-02	5.304991e+01	1.574269e+00	2.014159e+01
	RMS	1.118082e-02	1.788564e-01	1.657251e-01	7.283537e+00	1.254699e+00	4.487938e+00
	MEDIAN	1.250107e-04	3.198963e-02	2.746480e-02	5.304991e+01	1.574269e+00	2.014159e+01
	PERC25	1.250107e-04	3.198963e-02	2.746480e-02	5.304991e+01	1.574269e+00	2.014159e+01
	PERC975	1.250107e-04	3.198963e-02	2.746480e-02	5.304991e+01	1.574269e+00	2.014159e+01
SVR-RBF	MEAN	1.726870e-04	3.964567e-02	4.026439e-02	5.304991e+01	2.174888e+00	1.629812e+01
	RMS	1.314104e-02	1.991122e-01	2.006599e-01	7.283537e+00	1.474750e+00	4.037093e+00
	MEDIAN	1.726870e-04	3.964567e-02	4.026439e-02	5.304991e+01	2.174888e+00	1.629812e+01
	PERC25	1.726870e-04	3.964567e-02	4.026439e-02	5.304991e+01	2.174888e+00	1.629812e+01
	PERC975	1.726870e-04	3.964567e-02	4.026439e-02	5.304991e+01	2.174888e+00	1.629812e+01

Fonte: O autor.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal CSNA3 na frequência diária. Os resultados podem ser observados na tabela 34.

Através da figura 62 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente,

Tabela 34 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série CSNA3 diária (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

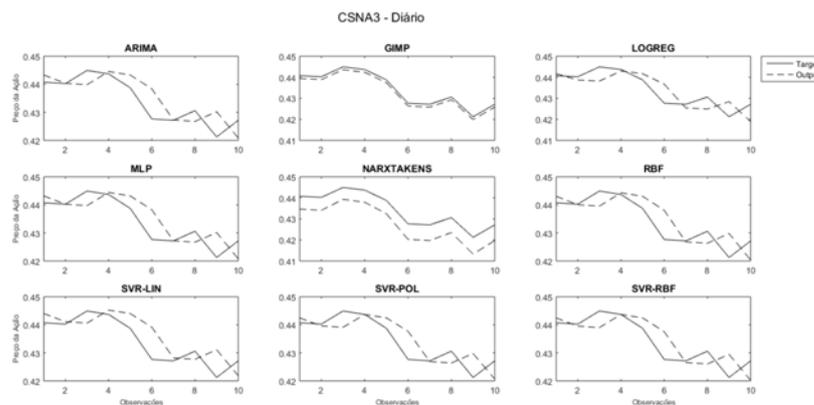
Tabela 35 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal CNSA3 Diário (conjunto de teste)

Modelo Proposto/Modelo Investigado	Ganho do GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	89,1146029	55,5130754	89,2064776	85,3658376	89,1246783	239,901037
GIMP/LOGREG	93,1261086	72,5278399	94,143257	85,3658376	93,1319714	341,655677
GIMP/MLP	89,1265791	55,6041295	89,2715811	85,3658376	89,1361389	240,085039
GIMP/NARXTAKENS	87,8160219	61,9626879	90,3037	0	87,8131005	57,5980846
GIMP/RBF	90,1469354	61,9551399	90,8528872	85,3658376	90,1555181	258,426192
GIMP/SVR-LIN	89,5192722	59,1245875	89,8614667	85,3658376	89,5320451	246,86263
GIMP/SVR-LOG	93,069786	72,850605	93,9026863	85,3658376	93,0782668	339,476079
GIMP/SVR-RBF	94,9831145	78,093469	95,8409528	85,3658376	94,9897788	443,114605

Fonte: O autor.

considerando os últimos dez pontos do conjunto de testes da série temporal CSNA3 diária.

Figura 62 – Gráfico de previsão da série temporal CSNA3 diária (conjunto de teste).



Fonte: O autor.

A tabela 35 mostra o ganho do modelo GIMP em relação aos outros modelos estudados.

### 6.3.8 CSNA3 - Frequência Semanal

Os experimentos com a série temporal CSNA3 na frequência semanal foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 36. Foram utilizados os valores de fechamento para a criação de cada série.

Tabela 36 – Desempenho preditivo para a série CSNA3 - Frequência Semanal (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	1.586117e-05	1.023668e-02	2.682001e-03	9.892473e+01	2.986344e-02	9.509096e+01
	RMS	3.982608e-03	1.011765e-01	5.178804e-02	9.946091e+00	1.728104e-01	9.751459e+00
	MEDIAN	4.640359e-06	7.100628e-03	7.899419e-04	9.892473e+01	8.736877e-03	9.730894e+01
	PERC25	3.333567e-08	6.146647e-04	5.679993e-06	9.892473e+01	6.276446e-05	8.437760e+01
	PERC975	7.026028e-05	2.821880e-02	1.182994e-02	9.892473e+01	1.322862e-01	9.885720e+01
ARIMA	MEAN	5.279752e-04	5.360635e-02	9.254160e-02	5.698925e+01	1.001866e+00	2.652461e+01
	RMS	2.297771e-02	2.315305e-01	3.042065e-01	7.549122e+00	1.000933e+00	5.150205e+00
	MEDIAN	5.279752e-04	5.360635e-02	9.254160e-02	5.698925e+01	1.001866e+00	2.652461e+01
	PERC25	5.279752e-04	5.360635e-02	9.254160e-02	5.698925e+01	1.001866e+00	2.652461e+01
	PERC975	5.279752e-04	5.360635e-02	9.254160e-02	5.698925e+01	1.001866e+00	2.652461e+01
LOGREG	MEAN	5.513896e-04	5.964469e-02	1.140482e-01	5.698925e+01	1.047403e+00	2.565180e+01
	RMS	2.348169e-02	2.442226e-01	3.377102e-01	7.549122e+00	1.023427e+00	5.064761e+00
	MEDIAN	5.513896e-04	5.964469e-02	1.140482e-01	5.698925e+01	1.047403e+00	2.565180e+01
	PERC25	5.513896e-04	5.964469e-02	1.140482e-01	5.698925e+01	1.047403e+00	2.565180e+01
	PERC975	5.513896e-04	5.964469e-02	1.140482e-01	5.698925e+01	1.047403e+00	2.565180e+01
MLP	MEAN	5.214691e-04	5.404142e-02	9.326948e-02	5.698925e+01	9.894930e-01	2.666381e+01
	RMS	2.283570e-02	2.324681e-01	3.054005e-01	7.549122e+00	9.947326e-01	5.163702e+00
	MEDIAN	5.214023e-04	5.403591e-02	9.332373e-02	5.698925e+01	9.893975e-01	2.666435e+01
	PERC25	5.212632e-04	5.403268e-02	9.280789e-02	5.698925e+01	9.890907e-01	2.665435e+01
	PERC975	5.222251e-04	5.407463e-02	9.332804e-02	5.698925e+01	9.907161e-01	2.666804e+01
NARXTAKENS	MEAN	8.685018e-05	2.825875e-02	1.883293e-02	9.892473e+01	1.638969e-01	8.215213e+01
	RMS	9.319344e-03	1.681034e-01	1.372331e-01	9.946091e+00	4.048418e-01	9.063781e+00
	MEDIAN	1.048443e-04	3.236306e-02	2.261914e-02	9.892473e+01	1.982421e-01	7.893036e+01
	PERC25	3.219350e-05	1.752202e-02	6.129463e-03	9.892473e+01	6.119800e-02	7.398823e+01
	PERC975	1.421269e-04	3.709729e-02	3.279175e-02	9.892473e+01	2.670022e-01	9.118481e+01
RBF	MEAN	5.368007e-04	5.807789e-02	1.081689e-01	5.698925e+01	1.020027e+00	2.606044e+01
	RMS	2.316896e-02	2.409935e-01	3.288905e-01	7.549122e+00	1.009964e+00	5.104942e+00
	MEDIAN	5.368007e-04	5.807789e-02	1.081689e-01	5.698925e+01	1.020027e+00	2.606044e+01
	PERC25	5.368007e-04	5.807789e-02	1.081689e-01	5.698925e+01	1.020027e+00	2.606044e+01
	PERC975	5.368007e-04	5.807789e-02	1.081689e-01	5.698925e+01	1.020027e+00	2.606044e+01
SVR-LIN	MEAN	5.193061e-04	5.439431e-02	9.424885e-02	5.698925e+01	9.855272e-01	2.669674e+01
	RMS	2.278829e-02	2.332259e-01	3.069998e-01	7.549122e+00	9.927372e-01	5.166888e+00
	MEDIAN	5.193061e-04	5.439431e-02	9.424885e-02	5.698925e+01	9.855272e-01	2.669674e+01
	PERC25	5.193061e-04	5.439431e-02	9.424885e-02	5.698925e+01	9.855272e-01	2.669674e+01
	PERC975	5.193061e-04	5.439431e-02	9.424885e-02	5.698925e+01	9.855272e-01	2.669674e+01
SVR-POL	MEAN	5.308175e-04	5.724688e-02	1.066984e-01	5.698925e+01	1.008846e+00	2.622218e+01
	RMS	2.303948e-02	2.392632e-01	3.266471e-01	7.549122e+00	1.004413e+00	5.120760e+00
	MEDIAN	5.308175e-04	5.724688e-02	1.066984e-01	5.698925e+01	1.008846e+00	2.622218e+01
	PERC25	5.308175e-04	5.724688e-02	1.066984e-01	5.698925e+01	1.008846e+00	2.622218e+01
	PERC975	5.308175e-04	5.724688e-02	1.066984e-01	5.698925e+01	1.008846e+00	2.622218e+01
SVR-RBF	MEAN	5.493430e-04	6.038447e-02	1.119764e-01	5.698925e+01	1.042995e+00	2.571827e+01
	RMS	2.343807e-02	2.457325e-01	3.346288e-01	7.549122e+00	1.021271e+00	5.071318e+00
	MEDIAN	5.493430e-04	6.038447e-02	1.119764e-01	5.698925e+01	1.042995e+00	2.571827e+01
	PERC25	5.493430e-04	6.038447e-02	1.119764e-01	5.698925e+01	1.042995e+00	2.571827e+01
	PERC975	5.493430e-04	6.038447e-02	1.119764e-01	5.698925e+01	1.042995e+00	2.571827e+01

Fonte: O autor.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal CSNA3

na frequência semanal. Os resultados podem ser observados na tabela 37.

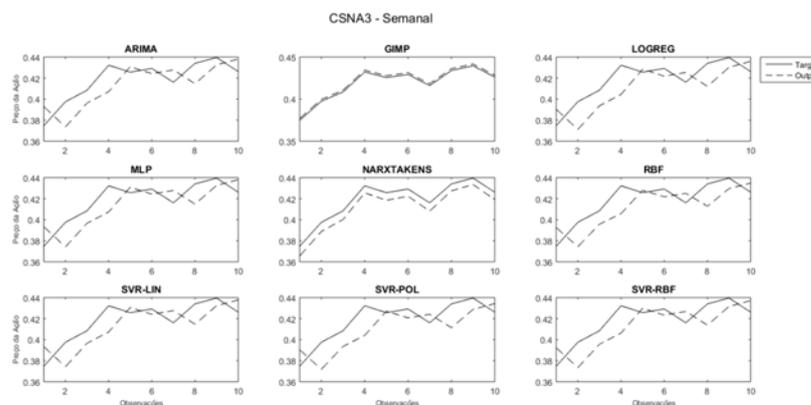
Tabela 37 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série CSNA3 semanal (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Através da figura 63 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal CSNA3 semanal.

Figura 63 – Gráfico de previsão da série temporal CSNA3 semanal (conjunto de teste).



Fonte: O autor.

A tabela 38 mostra o ganho do modelo GIMP em relação aos outros modelos estudados.

### 6.3.9 CSNA3 - Frequência Quinzenal

Os experimentos com a série temporal CSNA3 na frequência quinzenal foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 39. Foram utilizados os valores de fechamento para a criação de cada série.

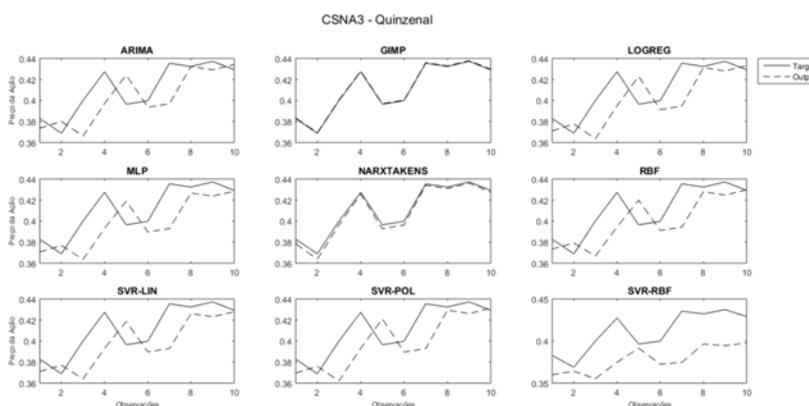
Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal CSNA3 na frequência quinzenal. Os resultados podem ser observados na tabela 40.

Tabela 38 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal CNSA3 Semanal

Modelo Proposto/Modelo Investigado	Ganho GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	96,9958494	80,9039787	97,1018428	73,5848954	97,0192181	258,500879
GIMP/LOGREG	97,1234187	82,8372316	97,6483618	73,5848954	97,1488109	270,698976
GIMP/MLP	96,9583682	81,0577146	97,1244602	73,5848954	96,9819453	256,629304
GIMP/NARXTAKENS	81,7373205	63,7751847	85,7589817	0	81,7791307	15,7498412
GIMP/RBF	97,0452404	82,3742219	97,5205433	73,5848954	97,0722893	264,886241
GIMP/SVR-LIN	96,9456993	81,1806051	97,1543409	73,5848954	96,9698005	256,189407
GIMP/SVR-LOG	97,0119354	82,1183617	97,4863719	73,5848954	97,0398416	262,635601
GIMP/SVR-RBF	97,1127019	83,0474955	97,6048516	73,5848954	97,136761	269,740888

Através da figura 64 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal CSNA3 quinzenal.

Figura 64 – Gráfico de previsão da série temporal CSNA3 quinzenal (conjunto de teste).



Fonte: O autor.

A tabela 41 mostra o ganho do modelo GIMP em relação aos outros modelos estudados

### 6.3.10 BBDC4 - Frequência Diária

Os experimentos com a série temporal BBDC4 na frequência diária foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 42. Foram utilizados os valores de fechamento para a criação de cada série.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal BBDC4 na frequência diária. Os resultados podem ser observados na tabela 43.

Através da figura 65 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal BBDC4 diária.

Tabela 39 – Desempenho preditivo para a série CSNA3 - Frequência Quinzenal (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	2.651117e-05	1.423170e-02	4.585950e-03	1.000000e+02	1.833621e-02	9.647713e+01
	RMS	5.148900e-03	1.192967e-01	6.771964e-02	1.000000e+01	1.354113e-01	9.822277e+00
	MEDIAN	1.772584e-05	1.303456e-02	3.080087e-03	1.000000e+02	1.225992e-02	9.724159e+01
	PERC25	1.401746e-06	3.691238e-03	2.443052e-04	1.000000e+02	9.695051e-04	9.279786e+01
	PERC975	6.145893e-05	2.444161e-02	1.060049e-02	1.000000e+02	4.250750e-02	9.951175e+01
ARIMA	MEAN	1.426365e-03	7.505325e-02	2.497942e-01	5.588235e+01	1.015061e+00	2.386773e+01
	RMS	3.776725e-02	2.739585e-01	4.997941e-01	7.475450e+00	1.007503e+00	4.885461e+00
	MEDIAN	1.426365e-03	7.505325e-02	2.497942e-01	5.588235e+01	1.015061e+00	2.386773e+01
	PERC25	1.426365e-03	7.505325e-02	2.497942e-01	5.588235e+01	1.015061e+00	2.386773e+01
	PERC975	1.426365e-03	7.505325e-02	2.497942e-01	5.588235e+01	1.015061e+00	2.386773e+01
LOGREG	MEAN	1.437289e-03	7.480467e-02	2.828406e-01	5.588235e+01	1.023006e+00	2.345940e+01
	RMS	3.791159e-02	2.735044e-01	5.318276e-01	7.475450e+00	1.011437e+00	4.843490e+00
	MEDIAN	1.437289e-03	7.480467e-02	2.828406e-01	5.588235e+01	1.023006e+00	2.345940e+01
	PERC25	1.437289e-03	7.480467e-02	2.828406e-01	5.588235e+01	1.023006e+00	2.345940e+01
	PERC975	1.437289e-03	7.480467e-02	2.828406e-01	5.588235e+01	1.023006e+00	2.345940e+01
MLP	MEAN	1.531175e-03	8.175826e-02	2.876122e-01	5.588235e+01	1.090061e+00	2.296702e+01
	RMS	3.913023e-02	2.859340e-01	5.362949e-01	7.475450e+00	1.044060e+00	4.792392e+00
	MEDIAN	1.374063e-03	7.323233e-02	2.743667e-01	5.588235e+01	9.782142e-01	2.401294e+01
	PERC25	1.365286e-03	7.280509e-02	2.665218e-01	5.588235e+01	9.719840e-01	1.681161e+01
	PERC975	2.661370e-03	1.295956e-01	3.990234e-01	5.588235e+01	1.894438e+00	2.408132e+01
NARXTAKENS	MEAN	2.509070e-05	1.371316e-02	4.711623e-03	1.000000e+02	1.775661e-02	9.653707e+01
	RMS	5.009062e-03	1.171032e-01	6.864126e-02	1.000000e+01	1.332539e-01	9.825328e+00
	MEDIAN	1.787681e-05	1.176088e-02	3.262010e-03	1.000000e+02	1.272412e-02	9.729855e+01
	PERC25	1.318808e-05	1.028046e-02	2.399302e-03	1.000000e+02	9.380995e-03	9.247962e+01
	PERC975	6.386734e-05	2.359591e-02	1.241146e-02	1.000000e+02	4.524803e-02	9.784028e+01
RBF	MEAN	1.381365e-03	7.616942e-02	2.824565e-01	5.588235e+01	9.835038e-01	2.384557e+01
	RMS	3.716672e-02	2.759881e-01	5.314663e-01	7.475450e+00	9.917176e-01	4.883192e+00
	MEDIAN	1.381365e-03	7.616942e-02	2.824565e-01	5.588235e+01	9.835038e-01	2.384557e+01
	PERC25	1.381365e-03	7.616942e-02	2.824565e-01	5.588235e+01	9.835038e-01	2.384557e+01
	PERC975	1.381365e-03	7.616942e-02	2.824565e-01	5.588235e+01	9.835038e-01	2.384557e+01
SVR-LIN	MEAN	1.369851e-03	7.291830e-02	2.719848e-01	5.588235e+01	9.752677e-01	2.407124e+01
	RMS	3.701150e-02	2.700339e-01	5.215216e-01	7.475450e+00	9.875564e-01	4.906245e+00
	MEDIAN	1.369851e-03	7.291830e-02	2.719848e-01	5.588235e+01	9.752677e-01	2.407124e+01
	PERC25	1.369851e-03	7.291830e-02	2.719848e-01	5.588235e+01	9.752677e-01	2.407124e+01
	PERC975	1.369851e-03	7.291830e-02	2.719848e-01	5.588235e+01	9.752677e-01	2.407124e+01
SVR-POL	MEAN	1.452049e-03	7.511463e-02	2.814192e-01	5.588235e+01	1.033746e+00	2.336481e+01
	RMS	3.810576e-02	2.740705e-01	5.304896e-01	7.475450e+00	1.016733e+00	4.833716e+00
	MEDIAN	1.452049e-03	7.511463e-02	2.814192e-01	5.588235e+01	1.033746e+00	2.336481e+01
	PERC25	1.452049e-03	7.511463e-02	2.814192e-01	5.588235e+01	1.033746e+00	2.336481e+01
	PERC975	1.452049e-03	7.511463e-02	2.814192e-01	5.588235e+01	1.033746e+00	2.336481e+01
SVR-RBF	MEAN	1.716314e-03	1.019242e-01	7.654423e-01	5.588235e+01	1.201645e+00	1.819841e+01
	RMS	4.142842e-02	3.192557e-01	8.748956e-01	7.475450e+00	1.096195e+00	4.265959e+00
	MEDIAN	1.716314e-03	1.019242e-01	7.654423e-01	5.588235e+01	1.201645e+00	1.819841e+01
	PERC25	1.716314e-03	1.019242e-01	7.654423e-01	5.588235e+01	1.201645e+00	1.819841e+01
	PERC975	1.716314e-03	1.019242e-01	7.654423e-01	5.588235e+01	1.201645e+00	1.819841e+01

Fonte: O autor.

Tabela 40 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série CSNA3 quinzenal (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	0	0	1	0	0	0
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

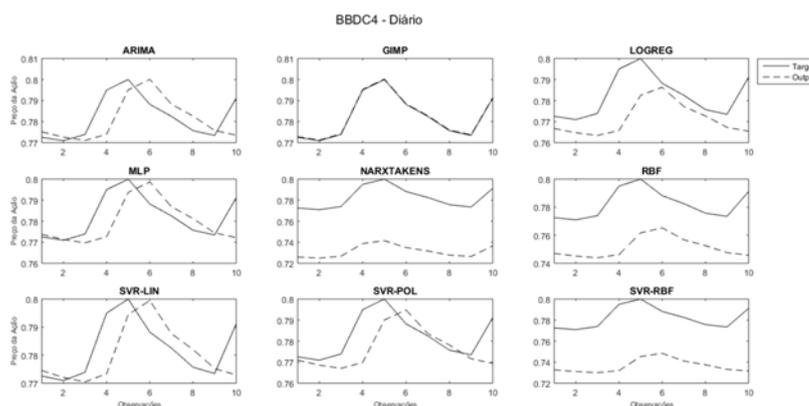
Fonte: O autor.

Tabela 41 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal CNSA3 quinzenal (conjunto de teste)

Modelo Proposto/Modelo Investigado	Ganho do GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	98,1413474	81,0378631	98,1641087	78,9473778	98,1935854	304,215776
GIMP/LOGREG	98,155474	80,9748509	98,3786097	78,9473778	98,2076146	311,251481
GIMP/MLP	98,2685735	82,5929515	98,4055092	78,9473778	98,317873	320,068124
GIMP/NARXTAKENS	-5,66134066	-3,78133122	2,66729745	0	-3,26413657	-0,06209014
GIMP/RBF	98,0807991	81,3157301	98,3764049	78,9473778	98,1356239	304,591419
GIMP/SVR-LIN	98,0646676	80,4826772	98,3138947	78,9473778	98,1198793	300,798339
GIMP/SVR-LOG	98,1742235	81,053358	98,3704204	78,9473778	98,2262364	312,91639
GIMP/SVR-RBF	98,4553427	86,0369765	99,4008758	78,9473778	98,4740743	430,140435

Fonte: O autor.

Figura 65 – Gráfico de previsão da série temporal BBDC4 diária (conjunto de teste).



Fonte: O autor.

O ganho do modelo GIMP em relação aos outros modelos para esta série pode ser vista na tabela 44

Tabela 42 – Desempenho preditivo para a série BBDC4 - Frequência Diária (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	9.614725e-06	4.789472e-03	8.927018e-04	9.889094e+01	6.611487e-02	9.265891e+01
	RMS	3.100762e-03	6.920601e-02	2.987812e-02	9.944393e+00	2.571281e-01	9.625950e+00
	MEDIAN	6.215077e-06	4.447411e-03	5.777558e-04	9.889094e+01	4.273747e-02	9.439882e+01
	PERC25	3.270743e-07	1.032456e-03	3.042321e-05	9.889094e+01	2.249100e-03	8.097660e+01
	PERC975	3.031331e-05	9.939513e-03	2.811786e-03	9.889094e+01	2.084470e-01	9.856447e+01
ARIMA	MEAN	1.451449e-04	1.634050e-02	1.360803e-02	4.898336e+01	9.959854e-01	2.417643e+01
	RMS	1.204761e-02	1.278300e-01	1.166535e-01	6.998812e+00	9.979907e-01	4.916954e+00
	MEDIAN	1.451449e-04	1.634050e-02	1.360803e-02	4.898336e+01	9.959854e-01	2.417643e+01
	PERC25	1.451449e-04	1.634050e-02	1.360803e-02	4.898336e+01	9.959854e-01	2.417643e+01
	PERC975	1.451449e-04	1.634050e-02	1.360803e-02	4.898336e+01	9.959854e-01	2.417643e+01
LOGREG	MEAN	1.597926e-04	1.712307e-02	1.483941e-02	4.898336e+01	1.092428e+00	2.305588e+01
	RMS	1.264091e-02	1.308552e-01	1.218171e-01	6.998812e+00	1.045193e+00	4.801653e+00
	MEDIAN	1.597926e-04	1.712307e-02	1.483941e-02	4.898336e+01	1.092428e+00	2.305588e+01
	PERC25	1.597926e-04	1.712307e-02	1.483941e-02	4.898336e+01	1.092428e+00	2.305588e+01
	PERC975	1.597926e-04	1.712307e-02	1.483941e-02	4.898336e+01	1.092428e+00	2.305588e+01
MLP	MEAN	1.462354e-04	1.642965e-02	1.380293e-02	4.898336e+01	1.002858e+00	2.409146e+01
	RMS	1.209279e-02	1.281782e-01	1.174859e-01	6.998812e+00	1.001428e+00	4.908305e+00
	MEDIAN	1.460182e-04	1.643330e-02	1.384131e-02	4.898336e+01	1.001172e+00	2.411082e+01
	PERC25	1.457160e-04	1.634889e-02	1.355590e-02	4.898336e+01	1.000174e+00	2.392482e+01
	PERC975	1.483119e-04	1.654049e-02	1.402267e-02	4.898336e+01	1.016676e+00	2.412707e+01
NARXTAKENS	MEAN	2.188958e-04	1.119662e-02	2.513056e-02	9.889094e+01	1.468278e+00	4.300841e+01
	RMS	1.479513e-02	1.058141e-01	1.585262e-01	9.944393e+00	1.211725e+00	6.558079e+00
	MEDIAN	1.884096e-04	1.067427e-02	2.068675e-02	9.889094e+01	1.262319e+00	4.317060e+01
	PERC25	8.507755e-05	7.742506e-03	8.667317e-03	9.889094e+01	5.679817e-01	2.371662e+01
	PERC975	4.603469e-04	1.744560e-02	5.720844e-02	9.889094e+01	3.094575e+00	6.241237e+01
RBF	MEAN	2.158463e-04	1.943525e-02	2.265963e-02	4.898336e+01	1.460884e+00	1.956834e+01
	RMS	1.469171e-02	1.394104e-01	1.505312e-01	6.998812e+00	1.208670e+00	4.423612e+00
	MEDIAN	2.158463e-04	1.943525e-02	2.265963e-02	4.898336e+01	1.460884e+00	1.956834e+01
	PERC25	2.158463e-04	1.943525e-02	2.265963e-02	4.898336e+01	1.460884e+00	1.956834e+01
	PERC975	2.158463e-04	1.943525e-02	2.265963e-02	4.898336e+01	1.460884e+00	1.956834e+01
SVR-LIN	MEAN	1.453922e-04	1.636690e-02	1.367025e-02	4.898336e+01	9.974142e-01	2.415834e+01
	RMS	1.205787e-02	1.279332e-01	1.169198e-01	6.998812e+00	9.987063e-01	4.915113e+00
	MEDIAN	1.453922e-04	1.636690e-02	1.367025e-02	4.898336e+01	9.974142e-01	2.415834e+01
	PERC25	1.453922e-04	1.636690e-02	1.367025e-02	4.898336e+01	9.974142e-01	2.415834e+01
	PERC975	1.453922e-04	1.636690e-02	1.367025e-02	4.898336e+01	9.974142e-01	2.415834e+01
SVR-POL	MEAN	1.463732e-04	1.644601e-02	1.394264e-02	4.898336e+01	1.002397e+00	2.409494e+01
	RMS	1.209848e-02	1.282420e-01	1.180790e-01	6.998812e+00	1.001198e+00	4.908659e+00
	MEDIAN	1.463732e-04	1.644601e-02	1.394264e-02	4.898336e+01	1.002397e+00	2.409494e+01
	PERC25	1.463732e-04	1.644601e-02	1.394264e-02	4.898336e+01	1.002397e+00	2.409494e+01
	PERC975	1.463732e-04	1.644601e-02	1.394264e-02	4.898336e+01	1.002397e+00	2.409494e+01
SVR-RBF	MEAN	3.208908e-04	2.244307e-02	3.580554e-02	4.898336e+01	2.165718e+00	1.519200e+01
	RMS	1.791343e-02	1.498101e-01	1.892235e-01	6.998812e+00	1.471638e+00	3.897691e+00
	MEDIAN	3.208908e-04	2.244307e-02	3.580554e-02	4.898336e+01	2.165718e+00	1.519200e+01
	PERC25	3.208908e-04	2.244307e-02	3.580554e-02	4.898336e+01	2.165718e+00	1.519200e+01
	PERC975	3.208908e-04	2.244307e-02	3.580554e-02	4.898336e+01	2.165718e+00	1.519200e+01

Fonte: O autor.

Tabela 43 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série BBDC4 diária (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Tabela 44 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal BBDC4 diária (conjunto de teste)

Modelo Proposto/Modelo Investigado	Ganho GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	93,3757748	70,6895627	93,4398895	101,886804	93,3618635	283,261342
GIMP/LOGREG	93,9829973	72,029128	93,98425	101,886804	93,9478968	301,888412
GIMP/MLP	93,4251727	70,848606	93,5325195	101,886804	93,4073548	284,613095
GIMP/NARXTAKENS	95,6076247	57,2239479	96,4477441	0	95,497115	115,4437
GIMP/RBF	95,5455688	75,356777	96,0603867	101,886804	95,4743245	373,514412
GIMP/SVR-LIN	93,3870421	70,7368408	93,4697478	101,886804	93,3713727	283,548332
GIMP/SVR-LOG	93,4313624	70,877605	93,5973259	101,886804	93,4043228	284,557546
GIMP/SVR-RBF	97,0037393	78,6594615	97,5068054	101,886804	96,9472078	509,919102

Fonte: O autor.

### 6.3.11 BBDC4 - Frequência Semanal

Os experimentos com a série temporal BBDC4 na frequência semanal foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 45. Foram utilizados os valores de fechamento para a criação de cada série.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal BBDC4 na frequência semanal. Os resultados podem ser observados na tabela 46.

Através da figura 66 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal BBDC4 semanal.

O ganho do modelo GIMP em relação aos outros modelos para esta série pode ser vista na tabela 47

Tabela 45 – Desempenho preditivo para a série BBDC4 - Frequência Semanal (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	1.816176e-05	6.668184e-03	1.352053e-03	1.000000e+02	2.824140e-02	9.658464e+01
	RMS	4.261661e-03	8.165895e-02	3.677028e-02	1.000000e+01	1.680518e-01	9.827748e+00
	MEDIAN	9.484609e-06	5.611472e-03	7.073369e-04	1.000000e+02	1.474850e-02	9.793584e+01
	PERC25	7.256154e-07	1.552181e-03	5.414985e-05	1.000000e+02	1.128327e-03	9.302382e+01
	PERC975	3.901584e-05	1.138176e-02	2.903304e-03	1.000000e+02	6.066934e-02	9.972721e+01
ARIMA	MEAN	6.355307e-04	3.319758e-02	4.907670e-02	4.408602e+01	9.969102e-01	2.119704e+01
	RMS	2.520973e-02	1.822020e-01	2.215326e-01	6.639731e+00	9.984539e-01	4.604024e+00
	MEDIAN	6.355307e-04	3.319758e-02	4.907670e-02	4.408602e+01	9.969102e-01	2.119704e+01
	PERC25	6.355307e-04	3.319758e-02	4.907670e-02	4.408602e+01	9.969102e-01	2.119704e+01
	PERC975	6.355307e-04	3.319758e-02	4.907670e-02	4.408602e+01	9.969102e-01	2.119704e+01
LOGREG	MEAN	7.004295e-04	3.527699e-02	5.966559e-02	4.408602e+01	1.080130e+00	2.026224e+01
	RMS	2.646563e-02	1.878217e-01	2.442654e-01	6.639731e+00	1.039293e+00	4.501360e+00
	MEDIAN	7.004295e-04	3.527699e-02	5.966559e-02	4.408602e+01	1.080130e+00	2.026224e+01
	PERC25	7.004295e-04	3.527699e-02	5.966559e-02	4.408602e+01	1.080130e+00	2.026224e+01
	PERC975	7.004295e-04	3.527699e-02	5.966559e-02	4.408602e+01	1.080130e+00	2.026224e+01
MLP	MEAN	3.873818e-03	5.720362e-02	1.761908e-01	4.408602e+01	5.908982e+00	1.819707e+01
	RMS	6.224000e-02	2.391728e-01	4.197509e-01	6.639731e+00	2.430840e+00	4.265803e+00
	MEDIAN	6.628489e-04	3.457236e-02	5.407208e-02	4.408602e+01	1.034598e+00	2.075412e+01
	PERC25	6.563831e-04	3.343688e-02	4.790294e-02	4.408602e+01	1.024862e+00	8.535414e-01
	PERC975	3.234265e-02	2.524063e-01	1.273415e+00	4.408602e+01	4.909255e+01	2.085415e+01
NARXTAKENS	MEAN	1.519333e-04	1.099678e-02	1.354134e-02	1.000000e+02	2.076140e-01	8.180901e+01
	RMS	1.232612e-02	1.048655e-01	1.163673e-01	1.000000e+01	4.556468e-01	9.044833e+00
	MEDIAN	1.328335e-04	1.079914e-02	1.160424e-02	1.000000e+02	1.805931e-01	8.320826e+01
	PERC25	4.019404e-05	5.366864e-03	3.263806e-03	1.000000e+02	5.287627e-02	7.035391e+01
	PERC975	2.761352e-04	1.512108e-02	2.552231e-02	1.000000e+02	3.804656e-01	9.420213e+01
RBF	MEAN	2.373920e-01	2.328113e-01	7.744101e-01	4.408602e+01	2.806180e+02	1.558566e-01
	RMS	4.872289e-01	4.825053e-01	8.800058e-01	6.639731e+00	1.675166e+01	3.947868e-01
	MEDIAN	2.373920e-01	2.328113e-01	7.744101e-01	4.408602e+01	2.806180e+02	1.558566e-01
	PERC25	2.373920e-01	2.328113e-01	7.744101e-01	4.408602e+01	2.806180e+02	1.558566e-01
	PERC975	2.373920e-01	2.328113e-01	7.744101e-01	4.408602e+01	2.806180e+02	1.558566e-01
SVR-LIN	MEAN	6.740087e-04	3.495494e-02	5.593432e-02	4.408602e+01	1.050692e+00	2.057925e+01
	RMS	2.596168e-02	1.869624e-01	2.365044e-01	6.639731e+00	1.025033e+00	4.536436e+00
	MEDIAN	6.740087e-04	3.495494e-02	5.593432e-02	4.408602e+01	1.050692e+00	2.057925e+01
	PERC25	6.740087e-04	3.495494e-02	5.593432e-02	4.408602e+01	1.050692e+00	2.057925e+01
	PERC975	6.740087e-04	3.495494e-02	5.593432e-02	4.408602e+01	1.050692e+00	2.057925e+01
SVR-POL	MEAN	6.361191e-04	3.294542e-02	4.802884e-02	4.408602e+01	9.996938e-01	2.118192e+01
	RMS	2.522140e-02	1.815087e-01	2.191548e-01	6.639731e+00	9.998469e-01	4.602382e+00
	MEDIAN	6.361191e-04	3.294542e-02	4.802884e-02	4.408602e+01	9.996938e-01	2.118192e+01
	PERC25	6.361191e-04	3.294542e-02	4.802884e-02	4.408602e+01	9.996938e-01	2.118192e+01
	PERC975	6.361191e-04	3.294542e-02	4.802884e-02	4.408602e+01	9.996938e-01	2.118192e+01
SVR-RBF	MEAN	1.363567e-03	4.816184e-02	1.432131e-01	4.408602e+01	2.051014e+00	1.359106e+01
	RMS	3.692651e-02	2.194581e-01	3.784351e-01	6.639731e+00	1.432136e+00	3.686605e+00
	MEDIAN	1.363567e-03	4.816184e-02	1.432131e-01	4.408602e+01	2.051014e+00	1.359106e+01
	PERC25	1.363567e-03	4.816184e-02	1.432131e-01	4.408602e+01	2.051014e+00	1.359106e+01
	PERC975	1.363567e-03	4.816184e-02	1.432131e-01	4.408602e+01	2.051014e+00	1.359106e+01

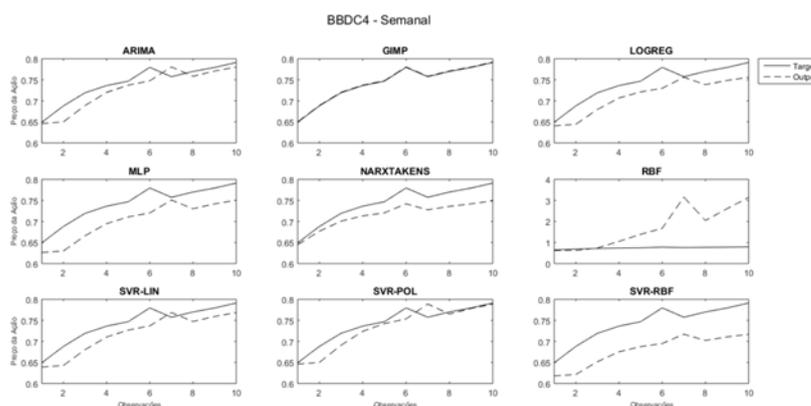
Fonte: O autor.

Tabela 46 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série BBDC4 semanal (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Figura 66 – Gráfico de previsão da série temporal BBDC4 semanal (conjunto de teste).



Fonte: O autor.

Tabela 47 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal BBDC4 semanal (conjunto de teste)

Modelo Proposto/Modelo Investigado	Ganho GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	97,1422687	79,9136443	97,2450205	126,829276	97,1671069	355,651544
GIMP/LOGREG	97,4070538	81,097639	97,7339485	126,829276	97,3853703	376,673063
GIMP/MLP	99,5311664	88,3430734	99,23262	126,829276	99,5220598	430,770283
GIMP/NARXTAKENS	88,0462282	39,3623952	90,0153678	0	86,3971601	18,061128
GIMP/RBF	99,9923495	97,1357988	99,8254087	126,829276	99,989936	61870,1957
GIMP/SVR-LIN	97,3054116	80,9234861	97,5827846	126,829276	97,3121143	369,330223
GIMP/SVR-LOG	97,144912	79,7599059	97,1849143	126,829276	97,174995	355,976795
GIMP/SVR-RBF	98,6680698	86,154632	99,0559153	126,829276	98,6230518	610,648323

Fonte: O autor.

### 6.3.12 BBDC4 - Frequência Quinzenal

Os experimentos com a série temporal BBDC4 na frequência quinzenal foram desenvolvidos, e os resultados do desempenho de todos os modelos podem ser vistos na tabela 48. Foram utilizados os valores de fechamento para a criação de cada série.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos

Tabela 48 – Desempenho preditivo para a série BBDC4 - Frequência Quinzenal (conjunto de teste).

Modelo	Estatística	Medida					
		MSE	MAPE	ARV	POCID	THEIL	EF
GIMP	MEAN	2.245224e-05	7.089379e-03	1.439633e-03	1.000000e+02	1.830985e-02	9.743330e+01
	RMS	4.738380e-03	8.419845e-02	3.794249e-02	1.000000e+01	1.353139e-01	9.870831e+00
	MEDIAN	1.249822e-05	6.241607e-03	8.029492e-04	1.000000e+02	1.019233e-02	9.830432e+01
	PERC25	2.745045e-08	2.925241e-04	1.764971e-06	1.000000e+02	2.238590e-05	9.365419e+01
	PERC975	6.125660e-05	1.381858e-02	3.923147e-03	1.000000e+02	4.995488e-02	9.996834e+01
ARIMA	MEAN	1.206785e-03	4.588378e-02	7.946731e-02	6.470588e+01	1.001269e+00	3.040938e+01
	RMS	3.473881e-02	2.142050e-01	2.818995e-01	8.043997e+00	1.000634e+00	5.514470e+00
	MEDIAN	1.206785e-03	4.588378e-02	7.946731e-02	6.470588e+01	1.001269e+00	3.040938e+01
	PERC25	1.206785e-03	4.588378e-02	7.946731e-02	6.470588e+01	1.001269e+00	3.040938e+01
	PERC975	1.206785e-03	4.588378e-02	7.946731e-02	6.470588e+01	1.001269e+00	3.040938e+01
LOGREG	MEAN	1.643707e-03	5.137819e-02	1.344090e-01	6.470588e+01	1.295762e+00	2.605754e+01
	RMS	4.054266e-02	2.266676e-01	3.666183e-01	8.043997e+00	1.138315e+00	5.104658e+00
	MEDIAN	1.643707e-03	5.137819e-02	1.344090e-01	6.470588e+01	1.295762e+00	2.605754e+01
	PERC25	1.643707e-03	5.137819e-02	1.344090e-01	6.470588e+01	1.295762e+00	2.605754e+01
	PERC975	1.643707e-03	5.137819e-02	1.344090e-01	6.470588e+01	1.295762e+00	2.605754e+01
MLP	MEAN	2.121923e-03	5.588535e-02	1.795870e-01	6.470588e+01	1.680714e+00	2.538123e+01
	RMS	4.606434e-02	2.364008e-01	4.237770e-01	8.043997e+00	1.296423e+00	5.037979e+00
	MEDIAN	1.510433e-03	4.929108e-02	1.194170e-01	6.470588e+01	1.213866e+00	2.714076e+01
	PERC25	1.375414e-03	4.730740e-02	1.024327e-01	6.470588e+01	1.117159e+00	8.264714e+00
	PERC975	7.746790e-03	1.163749e-01	7.441419e-01	6.470588e+01	5.960910e+00	2.852647e+01
NARXTAKENS	MEAN	4.093738e-04	1.912884e-02	3.353579e-02	1.000000e+02	2.759586e-01	7.617643e+01
	RMS	2.023299e-02	1.383070e-01	1.831278e-01	1.000000e+01	5.253176e-01	8.727911e+00
	MEDIAN	4.374087e-04	1.997182e-02	3.561898e-02	1.000000e+02	2.959176e-01	7.399321e+01
	PERC25	9.419782e-05	1.187994e-02	6.922935e-03	1.000000e+02	6.322420e-02	6.320325e+01
	PERC975	7.283756e-04	2.432568e-02	6.156088e-02	1.000000e+02	4.955821e-01	9.241108e+01
RBF	MEAN	9.332556e-01	4.971578e-01	8.806701e-01	6.176471e+01	5.642597e+02	1.088229e-01
	RMS	9.660516e-01	7.050941e-01	9.384403e-01	7.859052e+00	2.375415e+01	3.298832e-01
	MEDIAN	9.332556e-01	4.971578e-01	8.806701e-01	6.176471e+01	5.642597e+02	1.088229e-01
	PERC25	9.332556e-01	4.971578e-01	8.806701e-01	6.176471e+01	5.642597e+02	1.088229e-01
	PERC975	9.332556e-01	4.971578e-01	8.806701e-01	6.176471e+01	5.642597e+02	1.088229e-01
SVR-LIN	MEAN	2.334944e-03	6.188848e-02	2.165811e-01	6.470588e+01	1.826379e+00	2.082461e+01
	RMS	4.832126e-02	2.487740e-01	4.653827e-01	8.043997e+00	1.351436e+00	4.563399e+00
	MEDIAN	2.334944e-03	6.188848e-02	2.165811e-01	6.470588e+01	1.826379e+00	2.082461e+01
	PERC25	2.334944e-03	6.188848e-02	2.165811e-01	6.470588e+01	1.826379e+00	2.082461e+01
	PERC975	2.334944e-03	6.188848e-02	2.165811e-01	6.470588e+01	1.826379e+00	2.082461e+01
SVR-POL	MEAN	1.367649e-03	4.678336e-02	1.017975e-01	6.470588e+01	1.129137e+00	2.839116e+01
	RMS	3.698174e-02	2.162946e-01	3.190572e-01	8.043997e+00	1.062609e+00	5.328336e+00
	MEDIAN	1.367649e-03	4.678336e-02	1.017975e-01	6.470588e+01	1.129137e+00	2.839116e+01
	PERC25	1.367649e-03	4.678336e-02	1.017975e-01	6.470588e+01	1.129137e+00	2.839116e+01
	PERC975	1.367649e-03	4.678336e-02	1.017975e-01	6.470588e+01	1.129137e+00	2.839116e+01
SVR-RBF	MEAN	6.146771e-03	1.051865e-01	7.812895e-01	6.470588e+01	4.658226e+00	9.877480e+00
	RMS	7.840134e-02	3.243246e-01	8.839058e-01	8.043997e+00	2.158292e+00	3.142846e+00
	MEDIAN	6.146771e-03	1.051865e-01	7.812895e-01	6.470588e+01	4.658226e+00	9.877480e+00
	PERC25	6.146771e-03	1.051865e-01	7.812895e-01	6.470588e+01	4.658226e+00	9.877480e+00
	PERC975	6.146771e-03	1.051865e-01	7.812895e-01	6.470588e+01	4.658226e+00	9.877480e+00

Fonte: O autor.

modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas da série temporal BBDC4 na frequência quinzenal. Os resultados podem ser observados na tabela 49.

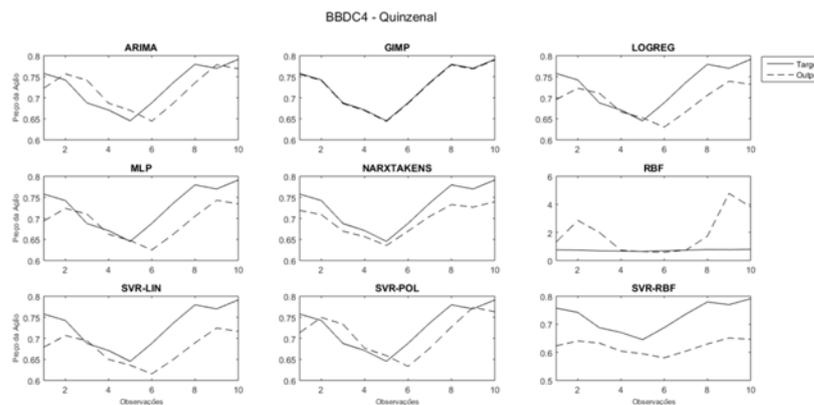
Tabela 49 – Teste de decisão ( $h$ ) não-paramétrico de Wilcoxon para a série BBDC4 quinzenal (conjunto teste).

Modelo Proposto/Modelo Investigado	Teste de Decisão( $h$ )					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	1	1	1	1	1	1
GIMP/LOGREG	1	1	1	1	1	1
GIMP/MLP	1	1	1	1	1	1
GIMP/NARXSTAKENS	1	1	1	0	1	1
GIMP/RBF	1	1	1	1	1	1
GIMP/SRV-LIN	1	1	1	1	1	1
GIMP/SRV-LOG	1	1	1	1	1	1
GIMP/SRV-RBF	1	1	1	1	1	1

Fonte: O autor.

Através da figura 67 pode-se observar o gráfico comparativo entre os valores reais da série e os valores preditos médio de cada um dos modelos analisado experimentalmente, considerando os últimos dez pontos do conjunto de testes da série temporal BBDC4 quinzenal.

Figura 67 – Gráfico de previsão da série temporal BBDC4 quinzenal (conjunto de teste).



Fonte: O autor.

O ganho do modelo GIMP em relação aos outros modelos para esta série pode ser vista na tabela 50

### 6.3.13 Considerações

É importante salientar que dentre os diversos modelos utilizados neste estudo, a medida EF média das 50 execuções do conjunto de teste são utilizados para a comparação dos resultados. Pode-se observar que o modelo GIMP obteve um desempenho preditivo superior

Tabela 50 – Ganho do modelo GIMP em Relação aos Simulados para Série Temporal BBDC4 quinzenal (conjunto de teste)

Modelo Proposto/Modelo Investigado	Ganho GIMP					
	MSE	MAPE	ARV	POCID	THEIL	EF
GIMP/ARIMA	98,1394996	84,5492699	98,188396	54,5454602	98,1713356	220,405414
GIMP/LOGREG	98,6340485	86,2015789	98,9289162	54,5454602	98,5869434	273,915957
GIMP/MLP	98,9418919	87,3144232	99,1983646	54,5454602	98,910591	283,879347
GIMP/NARXTAKENS	94,5154673	62,9387929	95,7071743	0	93,3650011	27,9047863
GIMP/RBF	99,9975942	98,5740183	99,8365298	61,9047511	99,9967551	89433,8205
GIMP/SVR-LIN	99,0384249	88,5449134	99,3352915	54,5454602	98,9974781	367,875749
GIMP/SVR-LOG	98,3583332	84,8463663	98,5857875	54,5454602	98,3784209	243,181821
GIMP/SVR-RBF	99,6347311	93,2601817	99,8157363	54,5454602	99,6069351	886,418601

**Fonte:** O autor.

aos experimentos dos modelos encontrados na literatura sobre sistemas preditivos. Para esse conjunto de séries, também observa-se ganhos consideráveis para todas as medidas analisadas. Em termos globais, um ganho de mais de 200%. Assim como no conjunto do estudo com as séries temporais do ativo PETR4, foi observado que alguns modelos obtiveram desempenho muito aquém do esperado e não mostraram um desempenho satisfatório na predição destas séries temporais. Os modelos SVR-POL, LOGREG, SVR-RBF e RBF obtiveram valores muito abaixo em relação ao modelo GIMP, que obteve ganhos acima de 300% na comparação entre os modelos para todas as séries temporais analisadas.

Outra característica observada foi a semelhança entre os resultados para diferentes séries temporais de baixa frequência, mostrando que o comportamento das séries temporais financeiras em baixa frequência tendem a ter o mesmo comportamento preditivo. Pode-se constatar que para as séries temporais financeiras, quanto mais baixa a sua frequência, melhor é o desempenho preditivo dos modelos, ficando o GIMP com um ganho em desempenho preditivo menor. Os únicos modelos que não seguem essa tendência são o RBF e o SVR-RBF, aumentando o ganho do GIMP nas frequências semanal e quinzenal. Esse fato provavelmente está atrelado ao fracasso da utilização de Função de Base Radial para previsão de séries temporais financeiras em baixa frequência.

Para comprovar o desempenho preditivo superior do modelo GIMP, em relação aos modelos que não utilizam a morfologia matemática, foi utilizado o teste de decisão ( $h$ ) não-paramétrico de Wilcoxon, para todas as medidas investigadas e em todas as séries temporais e os resultados mostram que o modelo proposto obteve resultados superiores na maioria dos casos. Nota-se na CSNA3 quinzenal tabela 40 o modelo NARKSTAKENS alguns valores superiores ao GIMP, obtendo um desempenho melhor para esta série temporal estudada.

A partir do gráfico de previsão de cada série temporal, constata-se o dilema do passeio aleatório na maioria dos modelos. Porém o GIMP e em alguns casos figura 64, a NARXTAKENS também consegue superar esse problema. O fato de utilizar o Teorema de Takens auxilia o modelo a criar uma memória de longo prazo, podendo calcular a componente não linear sub-dominante identificada nos estudos dos *lagplots*. Segundo (ARAÚJO, 2016),

o motivo da incapacidade dos modelo ARIMA e MLP superarem o dilema do passeio aleatório para este tipo particular de série temporal pode estar associado a uma falha no processo de construção do mapeamento utilizado para prevêê-las, o que, na prática, leva o modelo a ter um comportamento similar ao modelo Random Walk, ou seja, a previsão gerada apresenta um característico atraso de um passo a frente em relação aos valores reais da série temporal. Tal fato também pode ser atribuído aos modelos LOGREG, RBF, SVR-LIN, SVR-LOG e SVR-RBF.

Vale mencionar que o modelo NARX foi capaz de superar o dilema do passeio aleatório e no caso da série temporal CSNA3 quinzenal obteve um melhor resultado, quando comparado ao GIMP. Tal fato sugere que o sucesso para superar o dilema é condicionado ao tipo do mapeamento gerado pelo modelo em análise, sendo este capaz de estimar as dependências linear e não-linear características do fenômeno gerador deste tipo de série temporal. Esse fato pode ser atribuído à utilização do teorema de Takens que prover ao modelo uma memória de longo prazo, podendo assim calcular a componente não-linear subdominante.

A análise do erro preditivo gerados pelo modelo proposto, através dos histogramas, gráficos de autocorrelação e autocorrelação parcial mostrou que tais resíduos não se comportam como uma distribuição normal padrão e estes também não são autocorrelacionados.

Portanto, analisando os resultados experimentais alcançados, é possível encontrar evidências de que o modelo proposto confirma a análise apresentada no capítulo 4, mostrando que através da combinação dos termos lineares e não-lineares com comportamento crescente, adicionando o procedimento de ajuste de fase temporal e o uso da morfologia matemática com a decomposição de Matharon, pode-se gerar um modelo de previsão de séries temporais financeiras com desempenho superior àqueles analisados no Estado da Arte.

#### 6.3.13.1 Comparação com outros modelos do Estado da Arte PETR4 diário e Vale5 diário

Dentre outros estudos do Estado da Arte, foi analisado o trabalho de (ARAÚJO, 2012b), que comparou o modelo DELP (*Dilatation-Erosion-Linear Perceptron*) com outros trabalhos do Estado da Arte em 2012. Foi realizada uma comparação entre o GIMP e os resultados obtidos no trabalho de Araújo, cabendo algumas ressalvas: *i*) a série temporal PETR4 e a VALE5 na frequência diária conta com um conjunto diferente de pontos (o GIMP utilizou dados atualizados entre 02 de janeiro de 2008 até 06 de março de 2017); *ii*) o DELP mostrou resultados melhores que os métodos SHIF, TAEF e os modelos analisados por Araújo; *iii*) o GIMP utiliza apenas o módulo linear para o cálculo do operador crescente; *iv*) na comparação feita no trabalho, Araújo utiliza sempre o melhor modelo proposto, diferente do GIMP, que utiliza a média de 50 experimentos. Dessa forma, os dados mostram uma função de avaliação EF do DELP de 99,9524 para o melhor modelo,

contra 85,2087 médio do GIMP na série temporal da Petrobras e uma função de avaliação EF do DELP de 99,2497 para o melhor modelo, contra 95,9778 médio do GIMP para a série temporal VALE5. Apesar do desempenho do GIMP, nestes casos, se mostrar abaixo do DELP, o modelo *General Increasing Morphological Perceptron* foi superior a todos os estudados por Araújo.

### 6.3.13.2 Análise dos valores de $\lambda$ e $\theta$

Pôde-se observar que os valores obtidos pelo GIMP provêm de uma combinação balanceada entre as componentes lineares dominantes e componentes não-lineares sub-dominantes,  $\lambda$  e  $\theta$ . Os valores analisados são do conjunto de experimentos executados para cada ativo financeiro onde foi calculado a média e o desvio padrão dos valores de  $\lambda$  e  $\theta$ . Os resultados podem ser observados na tabela 51.

Tabela 51 – Média dos valores de  $\lambda$  e  $\theta$  para os experimento analisados.

Ações	Frequências	$\lambda$		$\theta$	
		Média	Desvio Padrão	Média	Desvio Padrão
BBDC4	D	0.9739	0.0027	0.3754	0.0178
	S	0.9508	0.0005	0.3650	0.0497
	Q	0.9899	0.0010	0.4875	0.0608
CSNA3	D	0.9326	0.0002	0.4043	0.0269
	S	0.9694	0.0012	0.5133	0.0147
	Q	0.9295	0.0001	0.4793	0.0241
PETR4	D	0.9791	0.0001	0.5343	0.0146
	S	0.9739	0.0001	0.5263	0.0287
	Q	0.9899	0.0010	0.4875	0.0608
VALE5	D	0.9422	0.0003	0.5386	0.0156
	S	0.9499	0.0001	0.5592	0.0238
	Q	0.9592	0.0001	0.4663	0.0169

**Fonte:** O autor.

Observa-se que, para o modelo proposto, são utilizados aproximadamente 0,0261% da componente não-linear e 0,9739% da componente linear, com um desvio padrão  $DP$  de 0,0027%, variando os valores entre  $0,9712\% < V < 0,9766\%$  para o ativo BBDC4 na frequência diária. O valor de  $\theta$  médio para este mesmo ativo foi de 0,3754% variando os valores dos experimentos entre  $0,3576\% < V < 0,3932\%$ . Esse resultado mostra que foi utilizado na média 0,3754% de decomposição por dilatações e 0,6246% de decomposição por erosões. É visto para o ativo BBDC4 na frequência semanal, a utilização de aproximadamente 0,0492% da componente não-linear e 0,9508% da componente linear, com um desvio padrão  $DP$  de 0,0005%, variando os valores entre  $0,9503\% < V < 0,9513\%$ . O valor de  $\theta$  médio dos experimentos para este ativo foi de 0,3650% variando os valores entre  $0,3156\% < V < 0,4147\%$ . Esse resultado mostra que foi utilizado na média 0,3650% de decomposição por dilatações e 0,6246% de decomposição por erosões. Da mesma forma podemos fazer a análise de todos os outros valores da tabela 51. É importante observar

o baixo desvio padrão do valor de  $\lambda$ , mostrando a homogeneidade dos experimentos executados, e o percentual de cada uma das decomposições para o cálculo da componente não-linear existente em cada série estudada. Pode-ser constatar também a superioridade da componente linear da função, tendo seus valores sempre acima de 90% do valor de saída da função do GIMP. Apesar do baixo percentual da componente não-linear, foi identificado que esta tem um papel fundamental na precisão do valor de saída da função, mapeando os valores muito mais próximos aos valores reais de previsão.

#### 6.4 RESUMO

Neste capítulo foi apresentada uma análise experimental comparativa entre os resultados obtidos pelos modelos investigados. Inicialmente, foram apresentados todos os modelos que fariam parte da análise experimental, investigando os diversos modelos existentes no Estado da Arte. Em seguida, foi apresentada a metodologia empregada para a realização dos experimentos, conceituadas todas as medidas encontradas na literatura para análise dos resultados das simulações. Para cada série temporal analisada, foram mostrados os resultados obtidos através de tabelas e gráficos dos dez últimos valores previstos, identificando inclusive os modelos que não conseguiram superar o dilema do passeio aleatório. Foi constatado que as séries temporais em baixa frequência têm um comportamento muito parecido, quanto às características do fenômeno gerador. Por fim, foi identificado, através da análise dos termos de combinação do modelo proposto, que as séries estudadas podem ser geradas através de uma combinação entre uma componente linear dominante e uma componente não-linear subdominante com comportamento crescente, viabilizando a previsão.

## 7 CONCLUSÃO

Com base em informações do passado de algumas ações da Bolsa de Mercados e Futuro Bovespa de São Paulo, o presente trabalho analisou o fenômeno gerador de séries temporais em baixa frequência de ativos, em diferentes frequências. Conforme observado na literatura, mesmo com modelos de previsão mais sofisticados surge uma limitação para o caso de previsões de séries financeiras, o que denomina-se de dilema do passeio aleatório (RWD). Trata-se de um problema de complexa solução, chegando alguns autores a afirmarem que tais séries são formadas por um processo puramente aleatório.

Foi identificada a importância de ferramentas para o auxílio de investidores na tomada de decisão de compra e venda de um ativo, e que a utilização de modelos preditivos é uma escolha interessante, tendo em vista seu desempenho com relação a outros modelos encontrados no Estado da Arte. Este trabalho estudou detalhadamente quatro ativos da BOVESPA (PETR4, CSNA3, VALE5 e BBDC4), em três frequências diferentes (diária, semanal e quinzenal), identificando a semelhança de comportamento dessas séries.

Através do estudo da literatura, foram observados alguns modelos preditivos que são utilizados em séries temporais financeiras. Também foi identificado o processo de análise de cada série temporal financeira, mostrando todas as variáveis que caracterizaram cada ativo. Modelos lineares de previsão também foram estudados, verificando a possibilidade de utilização dos mesmos no mercado financeiro.

Seguindo com o trabalho, utilizou-se a análise de *lagplot*, que permitiu a identificação de estruturas que caracterizaram a presença de um relacionamento dominante linear (baixo atraso) e subdominante não-linear (alto atraso) nos retardos temporais. Desta forma, a análise do *lagplot* levantou a hipótese de que o fenômeno gerador de séries temporais financeiras em baixa frequência é construído a partir de uma combinação entre componentes dominantes lineares e componentes subdominantes não-lineares. Esse procedimento foi desenvolvido seguindo o método de estudos das séries temporais idealizado por (ARAÚJO, 2012b). Devido tal análise ser fortemente subjetiva, uma vez que as relações contidas nestes gráficos podem não refletir claramente as características do fenômeno gerador da série, utilizou-se o teste da variância, que rejeitou, a 5% de nível de significância, a possibilidades dessas séries serem geradas de forma aleatória, confirmando estatisticamente a análise subjetiva do *lagplot*.

Para confirmar, de forma estatística, a presença da componente linear, foram utilizadas algumas análises de métodos estatísticos (autocorrelação, autocorrelação parcial) que mostraram uma forte dependência das séries em todos os casos. Para a confirmação da componente não-linear foi utilizado o método de informação mútua média. Embora confirmada a hipótese da existência de ambas as componentes, observou-se a necessidade de identificação da natureza da dependência não-linear das séries temporais. Para isso, foi

utilizado o parâmetro de Hurst, que comprovou a existência de dependência não-linear de longo prazo com comportamento auto-similar (anti-persistente) neste tipo particular de fenômeno gerador.

Baseado nas evidências encontradas na análise do fenômeno gerador destas séries temporais, foi proposto um modelo de neurônio híbrido de previsão, sendo composto por uma combinação de um operador não-linear crescente, calculado através da decomposição de erosões e dilatações, operações advindas da morfologia matemática, combinado com um operador linear. O modelo criado se baseia em gradiente descendente e utiliza ideias do algoritmo de retropropagação do erro das redes MLP e de uma abordagem sistemática, para superar o problema da não-diferenciabilidade das operações morfológicas, através do uso de uma função de impulso suave, uma vez que estes não são diferenciáveis de forma usual. Para finalizar a definição do modelo de previsão, foi empregado um procedimento de ajuste de fase que auxiliou no desempenho preditivo do modelo.

Após a proposta do modelo, este trabalho definiu uma metodologia experimental para analisar seu desempenho preditivo do modelo proposto, comparando-o com os modelos estudados no Estado da Arte, criando, assim, um nível de referência para análise dos resultados obtidos. Para cada modelo investigado neste trabalho, foram realizados 50 experimentos e, para cada medida de desempenho, foram calculadas cinco estatísticas relevantes, a média dos experimentos (MEAN), a raiz do erro quadrático médio (RMS), a mediana (MEDIAN), e os percentis 2,5% (PERC25) e 97,5% (PERC975).

A análise dos resultados obtidos com o modelo proposto, no conjunto de treinamento, revelou que o processo de aprendizagem teve comportamento estável, e no conjunto de teste revelou que não houve perda significativa na ordem de grandeza dos resultados obtidos para as medidas de desempenho em relação ao conjunto de treinamento.

Os resultados experimentais mostraram um ganho em cada uma das medidas analisadas, chegando até a 510% de ganho quando compara-se o GIMP aos outros modelos no ativo BBDC4. Identificou-se que alguns modelos utilizados na comparação não são modelos ideais para o uso na predição de séries temporais financeiras. Viu-se que os métodos estatísticos utilizados no trabalho conseguiu identificar os reais ganhos do modelo GIMP quando comparado aos outros modelos analisados. Em termos globais para esse conjunto de séries, também observa-se ganhos consideráveis para todas as medidas analisadas.

Observar-se, através dos resultados, que apenas o modelo GIMP e o Modelo NARX-TAKENS conseguiram superar o dilema do passeio aleatório mostrando qualitativamente que não foi identificado nenhum atraso entre os valores de predição e os valores reais e quantitativamente através do uso de U de THEIL indicando os valores sempre tendendo a 0. Foi identificado que em um caso o modelo NARXTAKENS obteve um melhor desempenho em relação ao GIMP, porém seu ganho foi mínimo, podendo ser considerado um desempenho similar ao GIMP. Tal fato ocorreu devido a componente crescente não ser tão importante naquela série temporal financeira, mostrando assim que o GIMP deve

ser utilizado apenas em séries temporais financeiras crescentes. Em todos os outros experimentos o Modelo GIMP mostrou um melhor desempenho de previsão em relação a NARXTAKENS, obtendo valores muito próximos dos reais.

Acredita-se que o fato de um modelo de previsão superar o problema do passeio aleatório está condicionado ao tipo do mapeamento utilizado para estimar a componente dominante linear e a componente sub-dominante não-linear presentes no fenômeno gerador deste tipo particular de série temporal, associadas à inclusão do procedimento de ajuste de fase no processo construtivo do modelo, como comprovado no GIMP. No caso da Rede NARX, foi a utilização do teorema de Takens que possibilitou a superação do problema.

Por fim, foi identificado que as séries temporais de baixa frequência (diária, semanal e quinzenal) têm comportamentos semelhantes, podendo o modelo GIMP ser utilizado para previsão de séries temporais financeiras nestas frequências. Em todos os ativos estudados, foi observado que os modelos ARIMA, LOGREG, MLP, RBF, SVR linear, SVR polinomial e SVR-RBF não superaram o problema do passeio aleatório, fato estatisticamente comprovado através do teste de Theil.

## 7.1 TRABALHOS FUTUROS

Apesar do desempenho preditivo do modelo GIMP ter alcançado ótimos resultados experimentais, foi visto através do Estado da Arte que há formas de melhorá-lo, em relação ao tempo de simulação. Além disso, o trabalho em questão abre um conjunto de possibilidades de estudos futuros, visando o estudo mais aprofundado do modelo a nível de software e de hardware. Quando observado o tempo de simulação do GIMP e de outros modelos de previsão de séries, fica nítida a importância do desenvolvimento de hardwares dedicados e o uso da computação paralela para acelerar o processo de localização do melhor modelo de previsão de cada série analisada, pois, de acordo com (ARAÚJO; OLIVEIRA; MEIRA, 2015), o crescimento das negociações no mercado financeiro em alta-frequência requer cada vez modelos mais ágeis para tomada de decisão.

O estudo do comportamento do GIMP nas séries temporais financeiras em alta frequência deve ser desenvolvido para verificar a possibilidade da inserção do cálculo da componente decrescente, elemento constatado por Araújo em séries financeiras de alta frequência.

Deve-se observar a formalização e uma investigação mais detalhada sobre as propriedades do modelo proposto, visando determinar as suas limitações práticas e teóricas em outras séries temporais financeiras e em outras frequências (BOVESPA, DOW JONES, NASDAQ, entre outras).

Considerando o fato que operadores morfológicos não são diferenciáveis de maneira usual, deve-se investigar outras técnicas para estimação do gradiente destes operadores.

Devido aos resultados obtidos com o modelo NARXTAKENS, é importante a realização de um estudo adicional visando a otimização dos parâmetros, uma vez que tal

modelo demonstrou resultados promissores mesmo sem a necessidade de utilização do procedimento de ajuste de fase. Dessa forma, é importante verificar se o uso do teorema de Takens também pode melhorar o desempenho preditivo dos outros modelos estudados.

A investigação de outros modelos com foco num comportamento crescentes existentes no Estado da Arte deve ser desenvolvido para analisar o comportamento do GIMP perante a estes. Alguns modelos do Estado da Arte estão descritos no capítulo A.

É importante investigar se o comportamento do modelo GIMP em ativos da bolsa de valores BOVESPA denominados *smallcaps* é similar as *bluechips* analisadas.

Por fim, é importante estudar, futuramente, a criação de um sistema automático para utilização em bolsas de valores que sirva como ferramenta de auxílio para investidores, limitando perdas e aumentando os ganhos nos ativos negociados.

## REFERÊNCIAS

- ADEODATO, P. J.; ARNAUD, A. L.; VASCONCELOS, G. C.; CUNHA, R. C.; MONTEIRO, D. S. Mlp ensembles improve long term prediction accuracy over single networks. *International Journal of Forecasting*, Elsevier, v. 27, n. 3, p. 661–671, 2011.
- ADHIKARI, R.; AGRAWAL, R. An introductory study on time series modeling and forecasting. *arXiv preprint arXiv:1302.6613*, 2013.
- AGRESTI, A.; KATERI, M. *Categorical data analysis*. [S.l.]: Springer, 2011.
- ANAGNOSTOPOULOS, K.; MAMANIS, G. The mean–variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multiobjective evolutionary algorithms. *Expert Systems with Applications*, Elsevier, v. 38, n. 11, p. 14208–14217, 2011.
- ARAÚJO, R. d. A. Swarm-based hybrid intelligent forecasting method for financial time series prediction. *Learning and Nonlinear Models*, v. 5, n. 2, p. 137–154, 2007.
- ARAÚJO, R. d. A. Adjusting time phase distortions in financial forecasting via morphological-rank-linear evolutionary approach. *Learning and Nonlinear Models*, v. 8, p. 157–162, 2010.
- ARAÚJO, R. d. A. Hybrid intelligent methodology to design translation invariant morphological operators for brazilian stock market prediction. *Neural Networks*, Elsevier, v. 23, n. 10, p. 1238–1251, 2010.
- ARAÚJO, R. d. A. A hybrid intelligent morphological approach for stock market forecasting. *Neural Processing Letters*, Springer, v. 31, n. 3, p. 195–217, 2010.
- ARAÚJO, R. d. A. Swarm-based translation-invariant morphological prediction method for financial time series forecasting. *Information Sciences*, Elsevier, v. 180, n. 24, p. 4784–4805, 2010.
- ARAÚJO, R. d. A. A class of hybrid morphological perceptrons with application in time series forecasting. *Knowledge-Based Systems*, Elsevier, v. 24, n. 4, p. 513–529, 2011.
- ARAÚJO, R. d. A. Translation invariant morphological time-lag added evolutionary forecasting method for stock market prediction. *Expert Systems with Applications*, Elsevier, v. 38, n. 3, p. 2835–2848, 2011.
- ARAÚJO, R. d. A. A morphological perceptron with gradient-based learning for brazilian stock market forecasting. *Neural Networks*, Elsevier, v. 28, p. 61–81, 2012.
- ARAÚJO, R. d. A. *Proposta de uma Classe de Perceptrons Híbridos com Aprendizagem Baseada em Gradiente Descendente*. Tese (Doutorado) — Universidade Federal De Pernambuco, 2012.
- ARAÚJO, R. d. A. *Mercado de Ações Brasileiro em Alta-Frequência: Evidências de Sua Previsibilidade com Modelagem Morfológica-Linear*. Tese (Doutorado) — Universidade Federal De Pernambuco, 2016.

- ARAÚJO, R. d. A.; OLIVEIRA, A. L.; MEIRA, S. A hybrid model for high-frequency stock market forecasting. *Expert Systems with Applications*, Elsevier, v. 42, n. 8, p. 4081–4096, 2015.
- ARAÚJO, R. d. A.; SOARES, S.; OLIVEIRA, A. L. Hybrid morphological methodology for software development cost estimation. *Expert Systems with Applications*, Elsevier, v. 39, n. 6, p. 6129–6139, 2012.
- ARAÚJO, R. d. A.; SUSSNER, P. An increasing hybrid morphological-linear perceptron with pseudo-gradient-based learning and phase adjustment for financial time series prediction. *Neural Networks (IJCNN), The 2010 International Joint Conference on*, p. 1–8, 2010.
- ARAÚJO, R. de A.; FERREIRA, T. A. E. A morphological-rank-linear evolutionary method for stock market prediction. *Information Sciences*, v. 237, p. 3–17, 2013. ISSN 0020-0255. Prediction, Control and Diagnosis using Advanced Neural Computations. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0020025509003053>>.
- ASADI, S.; HADAVANDI, E.; MEHMANPAZIR, F.; NAKHOSTIN, M. M. Hybridization of evolutionary levenbergmarquardt neural networks and data preprocessing for stock market prediction. *Knowledge-Based Systems*, Elsevier, v. 35, p. 245–258, 2012.
- ASADI, S.; TAVAKOLI, A.; HEJAZI, S. R. A new hybrid for improvement of autoregressive integrated moving average models applying particle swarm optimization. *Expert Systems with Applications*, Elsevier, v. 39, n. 5, p. 5332–5337, 2012.
- BATTITI, R. One step secant conjugate gradient. *Neural Computation*, v. 4, p. 141–166, 1992.
- BILLARD, L. Exponential autoregressive models. In: \_\_\_\_\_. *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd, 2014. ISBN 9781118445112. Disponível em: <<http://dx.doi.org/10.1002/9781118445112.stat03477.pub2>>.
- BIRKHOFF, G. *Lattice theory*. [S.l.]: American Mathematical Soc., 1940. v. 25.
- BIRKHOFF, G. *Lattice theory* 3rd ed. *Amer. Math. Soc. Colloq. Publ*, American Mathematical Society Providence, 1993.
- BIRKHOFF, G.; BIRKHOFF, G.; BIRKHOFF, G.; MATHÉMATICIEN, E.-U.; BIRKHOFF, G. *Lattice theory*. [S.l.]: American Mathematical Society New York, 1948. v. 25.
- BM&FBOVESPA. *Carteira Teórica do Ibovespa Válida para 14/03/16*. [http://www.bmfbovespa.com.br/pt\\_br/produtos/indices/indices-amplos/indice-ibovespa-ibovespa-composicao-da-carteira.htm](http://www.bmfbovespa.com.br/pt_br/produtos/indices/indices-amplos/indice-ibovespa-ibovespa-composicao-da-carteira.htm): [s.n.], 2016. Acesso em: 05/03/2016.
- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory*, p. 144–152, 1992.
- BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. *Time series analysis: forecasting and control*. [S.l.]: John Wiley & Sons, 2015.

- BRAGA, A. d. P.; CARVALHO, A.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. [S.l.]: Livros Técnicos e Científicos, 2000.
- CAVALCANTE, R. C.; BRASILEIRO, R. C.; SOUZA, V. L.; NOBREGA, J. P.; OLIVEIRA, A. L. Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, Elsevier, v. 55, p. 194–211, 2016.
- CECEN, A.; UGUR, A. On testing for nonlinear dependence and chaos in financial time series data. *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, v. 1, p. 203–208 Vol. 1, Oct 2005.
- CHARLES, A.; DARNÉ, O. Variance-ratio tests of random walk: an overview. *Journal of Economic Surveys*, Wiley Online Library, v. 23, n. 3, p. 503–527, 2009.
- CHEN, M.-Y.; CHEN, B.-T. A hybrid fuzzy time series model based on granular computing for stock price forecasting. *Information Sciences*, Elsevier, v. 294, p. 227–241, 2015.
- CHONG, T. T.-L.; LAM, T.-H. Predictability of nonlinear trading rules in the us stock market. *Quantitative Finance*, Taylor & Francis, v. 10, n. 9, p. 1067–1076, 2010.
- CLARKE, F. H. A new approach to lagrange multipliers. *Mathematics of Operations Research*, v. 1, n. 2, p. 165–174, 1976.
- CLEGG, R. G. A practical guide to measuring the hurst parameter. *arXiv preprint math/0610756*, 2006.
- CLEMENTS, M. P.; FRANSES, P. H.; SWANSON, N. R. Forecasting economic and financial time-series with non-linear models. *International Journal of Forecasting*, Elsevier, v. 20, n. 2, p. 169–183, 2004.
- CORTES, C.; VAPNIK, V. Support-vector networks. *Machine learning*, Springer, v. 20, n. 3, p. 273–297, 1995.
- COVEL, M. *Trend following: how great traders make millions in up or down markets*. [S.l.]: FT Press, 2004.
- CRISTIANINI, N.; SHAWE-TAYLOR, J. *An introduction to support vector machines and other kernel-based learning methods*. [S.l.]: Cambridge university press, 2000.
- CUMMING, D.; JOHAN, S.; SCHWEIZER, D. Information systems, agency problems, and fraud. *Information Systems Frontiers*, v. 19, n. 3, p. 421–424, Jun 2017. ISSN 1572-9419. Disponível em: <<https://doi.org/10.1007/s10796-017-9761-3>>.
- CUNINGHAME-GREEN, R. A. *Minimax algebra*. [S.l.]: Springer Science & Business Media, 2012. v. 166.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, Springer, v. 2, n. 4, p. 303–314, 1989.
- DIACONESCU, E. The use of narx neural networks to predict chaotic time series. *Wseas Transactions on computer research*, Citeseer, v. 3, n. 3, p. 182–191, 2008.

- DOUCOURE, B.; AGBOSSOU, K.; CARDENAS, A. Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data. *Renewable Energy*, v. 92, p. 202 – 211, 2016. ISSN 0960-1481. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0960148116301045>>.
- EBERHART, R.; KENNEDY, J. A new optimizer using particle swarm theory. *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*, p. 39–43, 1995.
- ENDERS, W. Applied econometric time series, by walter. *Technometrics*, v. 46, n. 2, p. 264, 2004.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 987–1007, 1982.
- FALAT, L.; PANCIKOVA, L.; HLINKOVA, M. Prediction model for high-volatile time series based on svm regression approach. *Information and Digital Technologies (IDT), 2015 International Conference on*, p. 77–83, 2015.
- FERREIRA, T. A.; VASCONCELOS, G. C.; ADEODATO, P. J. A new intelligent system methodology for time series forecasting with artificial neural networks. *Neural Processing Letters*, Springer, v. 28, n. 2, p. 113–129, 2008.
- FERREIRA, T. A. E. *Uma Nova Metodologia Híbrida Inteligente para a Previsão de Séries Temporais*. Tese (Doutorado) — Universidade Federal De Pernambuco, 2006.
- FLETCHER, R.; POWELL, M. J. A rapidly convergent descent method for minimization. *The computer journal*, Br Computer Soc, v. 6, n. 2, p. 163–168, 1963.
- FOGEL, L.; OWENS, A.; WALSH, M. *Artificial Intelligence Through Simulated Evolution*. [S.l.]: John Wiley & Sons, 1966.
- FRASER, A. M.; SWINNEY, H. L. Independent coordinates for strange attractors from mutual information. *Physical review A*, APS, v. 33, n. 2, p. 1134, 1986.
- FREITAS, F. D.; SOUZA, A. F. D.; ALMEIDA, A. R. de. Prediction-based portfolio optimization model using neural networks. *Neurocomputing*, Elsevier, v. 72, n. 10, p. 2155–2170, 2009.
- FRIEDBERG, R. M. A learning machine: Part i. *IBM Journal of Research and Development*, IBM, v. 2, n. 1, p. 2–13, 1958.
- GREENE, W. H. *Limdep*. [S.l.]: Econometric Software Bellport, 1995.
- GREENE, W. H. *Econometric analysis*. [S.l.]: Pearson Education India, 2003.
- GUNN, S. R. et al. Support vector machines for classification and regression. *ISIS technical report*, v. 14, p. 85–86, 1998.
- HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the marquardt algorithm. *IEEE transactions on Neural Networks*, IEEE, v. 5, n. 6, p. 989–993, 1994.

- HANSEN, N.; MÜLLER, S. D.; KOUMOUTSAKOS, P. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es). *Evolutionary computation*, MIT Press, v. 11, n. 1, p. 1–18, 2003.
- HAYKIN, S. S.; HAYKIN, S. S.; HAYKIN, S. S.; HAYKIN, S. S. *Neural networks and learning machines*. [S.l.]: Pearson Upper Saddle River, NJ, USA:, 2009. v. 3.
- HEARST, M. A.; DUMAIS, S. T.; OSMAN, E.; PLATT, J.; SCHOLKOPF, B. Support vector machines. *Intelligent Systems and their Applications, IEEE*, IEEE, v. 13, n. 4, p. 18–28, 1998.
- HOLLAND, J. H. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. [S.l.]: U Michigan Press, 1975.
- HURST, H. E. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, v. 116, p. 770–808, 1951.
- JANG, G. An intelligent stock portfolio management system based short-term trend prediction using dual-module neural network. *International Conference on Artificial Neural Network*, v. 1, p. 447–452, 1991.
- JIANG, P.; QIN, S.; WU, J.; SUN, B. Time series analysis and forecasting for wind speeds using support vector regression coupled with artificial intelligent algorithms. *Mathematical Problems in Engineering*, Hindawi Publishing Corporation, 2015.
- JOHANSON, J.; VAHLNE, J. The internationalization process of the firm-a model of knowledge development and increasing foreign market commitments. *Journal of International Business Studies*, Palgrave Macmillan Journals, v. 8, n. 1, p. 23–32, 1977. ISSN 00472506, 14786990. Disponível em: <<http://www.jstor.org/stable/254397>>.
- JR, D. W. H.; LEMESHOW, S.; STURDIVANT, R. X. *Applied logistic regression*. [S.l.]: John Wiley & Sons, 2013. v. 398.
- JÚNIOR, J. M. P. D. M.; BARRETO, G. A. Redes baseadas em projeções aleatórias para predição não-linear de séries temporais caóticas. *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, v. 1, n. 1, 2013.
- KANTZ, H.; SCHREIBER, T. *Nonlinear time series analysis*. [S.l.]: Cambridge university press, 2004. v. 7.
- KAR, B. P.; NAYAK, S. K.; NAYAK, S. C. Opposition-based ga learning of artificial neural networks for financial time series forecasting. In: *Computational Intelligence in Data Mining–Volume 2*. [S.l.]: Springer, 2016. p. 405–414.
- KAZEM, A.; SHARIFI, E.; HUSSAIN, F. K.; SABERI, M.; HUSSAIN, O. K. Support vector regression with chaos-based firefly algorithm for stock market price forecasting. *Applied soft computing*, Elsevier, v. 13, n. 2, p. 947–958, 2013.
- KENNEDY, J. *Particle swarm optimization*. [S.l.]: Springer, 2011. 760–766 p.
- KIM, S.; EOM, C. et al. Long-term memory and volatility clustering in high-frequency price changes. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 387, n. 5, p. 1247–1254, 2008.

- KÖMM, H.; KÜSTERS, U. Forecasting zero-inflated price changes with a markov switching mixture model for autoregressive and heteroscedastic time series. *International Journal of Forecasting*, Elsevier, v. 31, n. 3, p. 598–608, 2015.
- KRASKOV, A.; STGBAUER, H.; GRASSBERGER, P. A new auto-associative memory based on lattice algebra. *Phys. Rev. E*, v. 69, n. 6, 2004.
- LAGRANGE, J. L. *Mécanique analytique*. [S.l.]: MalletBachelier, 1853. v. 1.
- LEUNG, F. H.; LAM, H.-K.; LING, S.-H.; TAM, P. K. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *Neural Networks, IEEE Transactions on*, IEEE, v. 14, n. 1, p. 79–88, 2003.
- LEUNG, F. H. F.; LAM, H. K.; LING, S. H.; TAM, P. K. S. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Transactions on Neural Networks*, v. 14, n. 1, p. 79–88, Jan 2003. ISSN 1045-9227.
- LIN, T.; HORNE, B. G.; TIÑO, P.; GILES, C. L. Learning long-term dependencies in narx recurrent neural networks. *Neural Networks, IEEE Transactions on*, IEEE, v. 7, n. 6, p. 1329–1338, 1996.
- LIN, T.-N.; GILES, C. L.; HORNE, B. G.; KUNG, S.-Y. A delay damage model selection algorithm for narx neural networks. *Signal Processing, IEEE Transactions on*, IEEE, v. 45, n. 11, p. 2719–2730, 1997.
- LIU, L.; JOHNSON, H. L.; COUSENS, S.; PERIN, J.; SCOTT, S.; LAWN, J. E.; RUDAN, I.; CAMPBELL, H.; CIBULSKIS, R.; LI, M. et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *The Lancet*, Elsevier, v. 379, n. 9832, p. 2151–2161, 2012.
- LO, A. W.; MACKINLAY, A. C. Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of financial studies*, Soc Financial Studies, v. 1, n. 1, p. 41–66, 1988.
- LO, A. W.; MACKINLAY, A. C. The size and power of the variance ratio test in finite samples: A monte carlo investigation. *Journal of econometrics*, Elsevier, v. 40, n. 2, p. 203–238, 1989.
- LO, A. W.; MACKINLAY, A. C. *A non-random walk down Wall Street*. [S.l.]: Princeton University Press, 2002.
- LORENA, A. C.; CARVALHO, A. C. de. Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, v. 14, n. 2, p. 43–67, 2007.
- LOUKERIS, N.; MATSATSINIS, N. Corporate financial evaluation and bankruptcy prediction implementing artificial intelligence methods. *WSEAS Transactions on Business and Economics*, WSEAS Press, v. 3, n. 4, p. 343, 2006.
- MALKIEL, B. G. *A random walk down Wall Street: the time-tested strategy for successful investing*. [S.l.]: WW Norton & Company, 2007.
- MANAGEMENT, F. money. *Forex Money Management. Trade safe building stable gains*. <http://www.forex-money-management.com>: [s.n.], 2016. Acesso em: 22/02/2016.

- MATHERON, G. Random sets and integral geometry. Wiley New York, 1975.
- MATSUBA, I. Neural sequential associator and its application to stock price prediction. *Industrial Electronics, Control and Instrumentation, 1991. Proceedings. IECON '91., 1991 International Conference on*, v. 2, p. 1476–1479, Oct 1991.
- MATSUI, K.; SATO, H. Neighborhood evaluation in acquiring stock trading strategy using genetic algorithms. *Soft Computing and Pattern Recognition (SoCPaR), 2010 International Conference of*, p. 369–372, 2010.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, n. 4, p. 115–133, 1943.
- MENEZES JR., J. M. P.; BARRETO, G. A. Long-term time series prediction with the narx network: An empirical evaluation. *Neurocomput.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 71, n. 16-18, p. 3335–3343, oct 2008. ISSN 0925-2312. Disponível em: <<http://dx.doi.org/10.1016/j.neucom.2008.01.030>>.
- MENG, J.; GAO, Y.; SHI, Y. Support vector regression model for measuring the permittivity of asphalt concrete. *Microwave and Wireless Components Letters, IEEE, IEEE*, v. 17, n. 12, p. 819–821, 2007.
- MILLS, T. C. *Time series techniques for economists*. [S.l.]: Cambridge University Press, 1991.
- MINSKY, M.; PAPER, S. *Perceptrons*. [S.l.]: MIT press, 1969.
- MØLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks*, Elsevier, v. 6, n. 4, p. 525–533, 1993.
- MORGAN, A. H. Invexity and the kuhn-tucker theorem. *Journal of Mathematical Analysis and Applications*, v. 236, n. 2, p. 594-604, 1999. ISSN 0022247X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022247X99964843>>.
- NEKOUKAR, V.; BEHESHTI, M. T. H. A local linear radial basis function neural network for financial time-series forecasting. *Applied Intelligence*, Springer, v. 33, n. 3, p. 352-356, 2010.
- NETO, P. S. G. M. *Um Método para Análise de Mercado de Ações Utilizando Séries Temporais de Índices Financeiros*. Tese (Doutorado) — Universidade Federal De Pernambuco, 2012.
- NIU, H.; WANG, J. Financial time series prediction by a random data-time effective rbf neural network. *Soft Computing*, Springer, v. 18, n. 3, p. 497-508, 2014.
- OLIVEIRA, E. M. d. J. *Previsão da série temporal do Índice da Bolsa de Valores de São Paulo usando redes neurais e estatística*. 122 p. Tese (Doutorado) — Universidade Federal De Pernambuco, Recife, 2001.
- ÖSTERMARK, R. Incorporating asset growth potential and bear market safety switches in international portfolio decisions. *Applied Soft Computing*, Elsevier, v. 12, n. 8, p. 2538–2549, 2012.

- PEARLMUTTER, B. A. Gradient calculations for dynamic recurrent neural networks: A survey. *Neural Networks, IEEE Transactions on*, IEEE, v. 6, n. 5, p. 1212–1228, 1995.
- PERCIVAL, D.; WALDEN, A. Spectral analysis for physical applications: multitaper and conventional univariate techniques, 583 pp. *Cambridge Univ. Press, New York*, 1993.
- PESSOA, L. F.; MARAGOS, P. Mrl-filters: A general class of nonlinear systems and their optimal design for image processing. *IEEE Transactions on Image Processing*, IEEE, v. 7, n. 7, p. 966–978, 1998.
- PESSOA, L. F.; MARAGOS, P. Neural networks with hybrid morphological/rank/linear nodes: a unifying framework with applications to handwritten character recognition. *Pattern Recognition*, Elsevier, v. 33, n. 6, p. 945–960, 2000.
- PRECHELT, L. et al. Proben1: A set of neural network benchmark problems and benchmarking rules. *journal*, Citeseer, 1994.
- RANCO, G.; ALEKSOVSKI, D.; CALDARELLI, G.; GRČAR, M.; MOZETIČ, I. The effects of twitter sentiment on stock price returns. *PloS one*, Public Library of Science, v. 10, n. 9, p. e0138441, 2015.
- RAO, T.; SRIVASTAVA, S. Modeling movements in oil, gold, forex and market indices using search volume index and twitter sentiments. *Proceedings of the 5th Annual ACM Web Science Conference*, ACM, New York, NY, USA, p. 336–345, 2013. Disponível em: <<http://doi.acm.org/10.1145/2464464.2464521>>.
- RAO, T. S.; GABR, M. *An introduction to bispectral analysis and bilinear time series models*. [S.l.]: Springer Science & Business Media, 2012. v. 24.
- RATHER, A. M.; AGARWAL, A.; SASTRY, V. Recurrent neural network and a hybrid model for prediction of stock returns. *Expert Systems with Applications*, Elsevier, v. 42, n. 6, p. 3234–3241, 2015.
- RIEDMILLER, M.; BRAUN, H. A direct adaptive method for faster backpropagation learning: The rprop algorithm. *Neural Networks, 1993., IEEE International Conference on*, p. 586–591, 1993.
- RODRIGUEZ, J.; BLAZQUEZ, J.; SANTISO, J. Angel or devil: Chinas trade impact on latin american emerging markets. *CEPAL*, papers.ssrn.com,, v. 90, p. 15–41, 2008.
- RUSSELL, S.; NORVIG, P.; INTELLIGENCE, A. A modern approach. *Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs*, Citeseer, v. 25, p. 27, 1995.
- SALGADO, P.; LIMA, A.; FERREIRA, T.; CAVALCANTI, G. An intelligent perturbative approach for the time series forecasting problem. *International joint conference on neural networks*, 2010.
- SHAHRABI, J.; HADAVANDI, E.; ASADI, S. Developing a hybrid intelligent model for forecasting problems: Case study of tourism demand time series. *KnowledgeBased Systems*, Elsevier, v. 43, p. 112–122, 2013.
- SILVA, A.; NEVES, R.; HORTA, N. A hybrid approach to portfolio composition based on fundamental and technical indicators. *Expert Systems with Applications*, Elsevier, v. 42, n. 4, p. 2036–2048, 2015.

- SILVA, A. M. *Redes Neurais Morfológicas: Alguns aspectos teóricos e resultados experimentais em problemas de classificação*. Tese (Doutorado) — Universidade Estadual de Campinas, 2007.
- SILVA, A. M.; DAVIS, R. I.; PASHA, S. A.; LEONG, P. H. Forecasting financial time series with grammar-guided feature generation. *Computational Intelligence*, Wiley Online Library, 2016.
- SINGH, P. *Applications of Soft Computing in Time Series Forecasting: Simulation and modeling techniques*. [S.l.]: Springer International Publishing, 1992. v. 1. ISSN 1434-9922. ISBN 978-3-319-26293-2.
- SINGH, P. Fuzzy time series modeling approaches: A review. In: *Applications of Soft Computing in Time Series Forecasting*. [S.l.]: Springer, 2016. p. 11–39.
- SITTE, R.; SITTE, J. Neural networks approach to the random walk dilemma of financial time series. *Applied Intelligence*, v. 16, n. 3, p. 163–171, 2002. ISSN 1573-7497.
- SMOLA, A. J. et al. *Regression estimation with support vector learning machines*. Tese (Doutorado) — Master's thesis, Technische Universität München, 1996.
- SOARES, A. S. *Predição de Séries Temporais Econômicas por meio de Redes Neurais Artificiais e Transformadas de Wavelet: Combinando Modelo Técnico e Fundamentalista*. 70 p. Tese (Doutorado) — Universidade de São Paulo, 2008.
- SUSSNER, P.; ESMI, E. L. Morphological perceptrons with competitive learning: Lattice-theoretical framework and constructive learning algorithm. *Information Sciences*, Elsevier, v. 181, n. 10, p. 1929–1950, 2011.
- SUSSNER, P.; VALLE, M. E. Morphological and certain fuzzy morphological associative memories for classification and prediction. In: *Computational intelligence based on lattice theory*. [S.l.]: Springer, 2007. p. 149–171.
- TAKENS, F. Dynamical systems and turbulence, warwick 1980: Proceedings of a symposium held at the university of warwick 1979/80. In: \_\_\_\_\_. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981. cap. Detecting strange attractors in turbulence, p. 366–381. ISBN 978-3-540-38945-3. Disponível em: <<http://dx.doi.org/10.1007/BFb0091924>>.
- VAPNIK, V. N.; KOTZ, S. *Estimation of dependences based on empirical data*. [S.l.]: Springer-Verlag New York, 1982. v. 40.
- WAGNER, A. K.; SOUMERAI, S. B.; ZHANG, F.; ROSS-DEGNAN, D. Segmented regression analysis of interrupted time series studies in medication use research. *Journal of clinical pharmacy and therapeutics*, Wiley Online Library, v. 27, n. 4, p. 299–309, 2002.
- WANG, J.-J.; WANG, J.-Z.; ZHANG, Z.-G.; GUO, S.-P. Stock index forecasting based on a hybrid model. *Omega*, Elsevier, v. 40, n. 6, p. 758–766, 2012.
- WHITE, H. Economic prediction using neural networks: The case of ibm daily stock returns. *Neural Networks, 1988., IEEE International Conference on*, p. 451–458, 1988.
- WIDROW, B.; HOFF, M. E. et al. Adaptive switching circuits. *IRE WESCON convention record*, v. 4, n. 1, p. 96–104, 1960.

WILCOXON, F. Individual comparisons by ranking methods. *Biometrics bulletin*, JSTOR, v. 1, n. 6, p. 80–83, 1945.

WILLIAMS, D. R. G. H. R.; HINTON, G. Learning representations by back-propagating errors. *Nature*, v. 323, p. 533–536, 1986.

XIONG, T.; BAO, Y.; HU, Z. Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting. *Knowledge-Based Systems*, Elsevier, v. 55, p. 87–100, 2014.

ŻBIKOWSKI, K. Using volume weighted support vector machines with walk forward testing and feature selection for the purpose of creating stock trading strategy. *Expert Systems with Applications*, Elsevier, v. 42, n. 4, p. 1797–1805, 2015.

ZHANG, X.; HU, Y.; XIE, K.; ZHANG, W.; SU, L.; LIU, M. An evolutionary trend reversion model for stock trading rule discovery. *Knowledge-Based Systems*, Elsevier, v. 79, p. 27–35, 2015.

## Apêndices

## APÊNDICE A – OUTROS MODELOS PARA PREVISÃO DE SÉRIES TEMPORAIS

Este apêndice descreve outros modelos de previsão lineares e não-lineares comumente encontrados na literatura de previsão de séries temporais financeiras.

### A.1 INTRODUÇÃO

A pesquisa desenvolvida neste trabalho encontrou outros modelos que utilizam Rede Neural Artificial para previsão de séries temporais financeiras e modelos estatísticos não-lineares. A seguir serão relatados aqueles que tem importância para este trabalho e alguns conceitos importante para cada um deles.

#### A.1.1 Modelos Não-Lineares

O modelo de previsão de séries temporais financeiras mais utilizado até o momento é o modelo ARIMA. Isso se deve ao fato do seu nível de complexidade ser baixo e, principalmente, os pesquisadores não levarem em consideração as componentes lineares subdominantes de uma série temporal financeira. Outro fator que limita os modelos Lineares é a introdução da inferência da estacionariedade da série, ou a série pode ser transformada numa série estacionária. No entanto, a literatura mostra que as séries temporais financeiras não são séries puramente lineares, mas há uma componente subdominante não linear que influencia nos valores de previsão (ARAÚJO, 2012b; FERREIRA, 2006; NETO, 2012).

Dessa forma, uma série temporal não linear pode ser definida pela seguinte equação (ARAÚJO, 2012b; FERREIRA, 2006; NETO, 2012):

$$Z_t = h(Z_{t-1}, \dots, Z_{t-p}, r_{t-1}, \dots, r_{t-q}) + r_t \quad (\text{A.1})$$

Onde  $r_t$  representa uma componente aleatória, ou termo de ruído. As variáveis  $p$  e  $q$  são índices  $\in \mathbb{N}^+$  que determinam as janelas temporais e os termos de ruídos, que serão utilizadas na função não linear de transferência  $h(\cdot)$ . A função  $h(\cdot)$  tem como objetivo o mapeamento entre os valores futuros e os valores do passado.

Com o intuito de superar as limitações dos modelos lineares, a literatura mostra alguns modelos não lineares estatísticos propostos: modelo heterocedásticos condicionais autorregressivo, modelos bilineares, modelos de limiar autorregressivo, modelos de estados gerais dependentes, modelos autorregressivo exponenciais e modelo de limiar autorregressivo autoexcitado.

Os modelos heterocedásticos condicionais autorregressivos, também conhecidos como *ARCH*, são modelos que pressupõem que a variância da estrutura de erro é constante ao longo do tempo. Engle (ENGLE, 1982), autor do modelo *ARCH*, afirma que se a variância

do erro  $E1$  não é constante ao longo do tempo, é possível estimar qualquer tendência de movimentos, tendo como base uma função *ARMA*. A representação do modelo *ARCH* de ordem  $q$  é definido como um modelo *AR*( $q$ ) onde  $r_t$  é condicionado à informação avaliada no tempo  $t - 1$ , com variância,

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2 \quad (\text{A.2})$$

como  $\alpha_0 > 0$  e  $\alpha_i \geq 0$  Desta forma, o modelo ARCH é dado pela equação,

$$Z_t = r_t \left[ \alpha_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2 \right]^{\frac{1}{2}} \quad (\text{A.3})$$

Neste modelo a não-linearidade vem da variância condicional de  $r_t$  (FERREIRA, 2006). Seja  $\{\hat{\varepsilon}_t\}$ , os resíduos estimados. Para calcular a variância de  $y_{t+1}$  teremos:

$$\text{var}(y_{t+1}/y_t) = E_t[(y_{t+1} - a_0 - a_1 y_t)^2] = E_t \varepsilon_{t+1}^2 \quad (\text{A.4})$$

Caso a variância não seja constante no tempo, utiliza-se a função *ARMA* de ordem  $q$  (*AR*( $q$ )) com o quadrado dos resíduos:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 + v_t \quad (\text{A.5})$$

Onde  $v_t$  é um processo ruído branco. Um processo ruído branco se caracteriza pelo fato de ser identicamente e independentemente distribuído, com média zero. A sua função de autocorrelação é igual a zero para todas as defasagens (lags), exceto para a defasagem de ordem zero. Outro aspecto relevante de um ruído branco é sua aleatoriedade, tornando impossível prever seu comportamento em qualquer processo de estimação.

Engle (ENGLE, 1982) mostrou que a melhor forma de tratar  $v_t$  é na forma multiplicativa e propôs o modelo multiplicativo condicional heterocedástico:

$$\varepsilon_t^2 = v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \quad (\text{A.6})$$

Representando o modelo autorregressivo de ordem um.

Modelo autorregressivo de ordem maior que um:

$$\varepsilon_t^2 = v_t \sqrt{\alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2} \quad (\text{A.7})$$

Onde  $v_t$  é o processo ruído branco, tal que  $\sigma_v^2 = 1$ ,  $e_{t-1}$  e  $v_t$  são independentes um do outro,  $E1 \in | N$ ,  $\alpha_0 > 0$  e  $0 < \alpha_1 < 0$ .

## A.2 MODELO DE REDE NEURAL COM A UTILIZAÇÃO DO ALGORITMO GENÉTICO MODIFICADO (MGA)

O Algoritmo Genético inserido no sistema tem a finalidade de melhorar a busca da sub-rede ótima, reduzindo o tempo de processamento e encontrando o mínimo global da função de previsão da série temporal.

A etapa de Avaliação dos indivíduos é responsável por gerar uma lista de indivíduos da população, indicando aqueles que poderão ser retirados caso o processo de melhoramento genético gere indivíduos melhores do que eles. Nesse processo não é modificada a sequência em que a população foi criada, apenas elenca os piores indivíduos para, caso necessário, serem retirados da população, de modo que outros indivíduos melhores ocupem seus lugares.

Escolher os pais é a primeira ação que deve ser desenvolvida na etapa de Gerar Cruzamento. O sistema irá escolher aleatoriamente dois indivíduos da população para fazer o cruzamento e gerar os novos filhos. A escolha dos indivíduos da população deverá ser aleatória para seguir um dos princípios da evolução utilizados em algoritmos genéticos, afirmando que não são os mais inteligentes e nem os mais fortes que sobrevivem, mas aqueles que mais se adaptam ao ambiente. Essa forma de escolha é conhecida na literatura como *Spinning Roulette Wheel* e foi mostrado em Holland (HOLLAND, 1975) que este método tem um alto potencial de geração de descendentes melhorados. Dessa forma, não podemos utilizar apenas os melhores indivíduos, mas conjuntos diferentes de indivíduos existentes na população. Com esta ação, o sistema suportará a criação de uma população panmítica de sub-redes ótimas, ou seja, todos os cruzamentos podem ocorrer com igual probabilidade, casualmente, permitindo uma perfeita distribuição dos seus genes entre todos os seus indivíduos.

Para cada cromossomo da população, teremos a evolução definida por uma função denominada função de aptidão (*fitness*) ou valor adaptativo. O valor de aptidão é uma medida de sucesso reprodutivo, onde os cromossomos que sobrevivem e produzem maior número de prole capaz de se reproduzir, garantem a manutenção do conjunto de pais da população. Em um dado momento do fluxo do sistema, esses cromossomos ocupam uma determinada área do programa (vetor) e trocarão informações livremente entre si para gerar seus filhos (novos cromossomos), que formarão a base para o fundo genético da geração seguinte.

O cruzamento entre os dois pais escolhidos dará início à criação de alguns outros indivíduos com probabilidade de participar da população já existente. O processo de cruzamento gerará quatro novos indivíduos filhos, seguindo a metodologia criada por Leung (LEUNG et al., 2003b), que faz o cruzamento através do seguinte processo:

$$P_{max} = [para_{max}^1 \quad para_{max}^2 \quad \cdots \quad para_{max}^{no\_vars}] \quad (A.8)$$

$$P_{min} = [para_{min}^1 \quad para_{min}^2 \quad \cdots \quad para_{min}^{no\_vars}] \quad (A.9)$$

$$os_c^1 = [os_1^1 \quad os_2^1 \quad \cdots \quad os_{no\_vars}^1] = \frac{P_1 + P_2}{2} \quad (A.10)$$

$$\begin{aligned} os_c^2 &= [os_1^2 \quad os_2^2 \quad \cdots \quad os_{no\_vars}^2] \\ &= p_{max}(1 - w) + max(p1, p2)w \end{aligned} \quad (A.11)$$

$$\begin{aligned} os_c^3 &= [os_1^3 \quad os_2^3 \quad \cdots \quad os_{no\_vars}^3] \\ &= p_{min}(1 - w) + min(p1, p2)w \end{aligned} \quad (A.12)$$

$$\begin{aligned} os_c^4 &= [os_1^4 \quad os_2^4 \quad \cdots \quad os_{no\_vars}^4] \\ &= \frac{(p_{max} + p_{min})(1 - w) + (p_1 + p_2)w}{2} \end{aligned} \quad (A.13)$$

Onde  $w \in [0, 1]$  denota o peso determinado pelo usuário na etapa um de ajuste de parâmetros.

$max(p1, p2)$  é o vetor obtido com os valores máximos de cada gene dos cromossomos  $p_1$  e  $p_2$ . Como exemplo, dada função Max  $([2, 5, 8], [4, 3, 6])$ ; Para essa instância teremos como resposta o vetor  $[4, 3, 8]$ . Para o cálculo de mínimo, o sistema procederá de forma semelhante, porém criando o vetor de saída com os valores mínimos de cada gene dos cromossomos utilizados para o cruzamento.

O valor de *fitness* é obtido da função de evolução da população:

$$fitness = f(p_i), \quad (A.14)$$

A forma da função de aptidão dependerá diretamente da aplicação que será desenvolvida. Em problemas de previsão de série temporal, as métricas mais utilizadas são frequentemente definidas em termos do erro instantâneo de previsão, que é definido pela subtração entre o valor real e o valor previsto.

$$Erro_j = valor_{real} - valor_{previsto}; \quad (A.15)$$

### A.3 MODELOS HÍBRIDOS

Os seres humanos são exemplos de máquinas que processam informações de forma híbrida (ARAÚJO, 2012b; FERREIRA, 2006). Isso ocorre pelo fato de recebermos informações genéticas do nosso corpo e adquirirmos informações por meios de aprendizagem. Quando nascemos, já temos informações necessárias para a sobrevivência, como respirar, dormir, entre outras. Com o passar do tempo, vamos adquirindo informações e “aprendendo” por meio dos nossos sentidos: visão, paladar, olfato, audição e tato. A combinação destes diferentes tipos de processamento é indispensável para a vida do ser humano, pois a dinâmica

da vida é complexa e o ser humano tem que se adaptar às modificações que ocorre no ambiente. Isso ocorre não só com os seres humanos, mas com qualquer organismo. De acordo com a teoria da seleção natural de Charles Darwin, os organismos mais bem adaptados ao meio têm maiores chances de sobrevivência do que os menos adaptados, deixando um número maior de descendentes.

Um modelo híbrido na Inteligência Artificial, portanto, é um sistema que pode trabalhar com informações através de um sistema inteligente em seu núcleo, e se adaptando as mudanças que ocorrem quando estão sendo treinados. Segundo Araújo (ARAÚJO, 2012b), os sistemas híbridos se tornam necessários quando os problemas são complexos e com um tipo particular de processamento.

Na literatura podemos encontrar diversos sistemas híbridos utilizados para o estudo do mercado financeiro. O trabalho de Loukeris (LOUKERIS; MATSATSINIS, 2006) descreve um sistema híbrido que utiliza as redes de Elman e Jordan e Programação Genética para selecionar o melhor portfólio de investimento dentre as ações disponíveis.

Anagnostopoulos (ANAGNOSTOPOULOS; MAMANIS, 2011) aplica o teorema de Markowitz (modelo da média-variância) para definir os portfólios e utiliza Algoritmos Evolucionários para melhorar a estratégia de investimento com diferentes níveis de riscos. Os resultados mostrados no trabalho confirmam que o uso de um modelo estatístico (Markowitz) e de um algoritmo evolucionário em conjunto obteve resultados melhores que o uso apenas de um Algoritmo Evolucionário.

No trabalho intitulado *A hybrid model for high-frequency stock market forecasting*, de Araújo (ARAÚJO; OLIVEIRA; MEIRA, 2015), o autor desenvolve um modelo híbrido de neurônio para previsão de séries financeiras em alta frequência. O trabalho mostra resultados melhores que outros modelos de previsão, tais como MLP e ARIMA.

Os sistemas híbridos inteligentes podem não somente representar a combinação de diferentes técnicas da Inteligência Artificial, mas pode também integrar essas técnicas com sistemas computacionais. Segundo Ferreira (FERREIRA, 2006), podemos citar três motivos para a criação de modelos híbridos inteligentes:

- Técnica de Intensificação – Integrar mais de uma das diferentes técnicas da Inteligência Artificial com o objetivo de uma auxiliar a outra;
- Multiplicidade de Aplicações em Tarefas – criação de um sistema híbrido para tratamento de uma tarefa que não pode ser executada por um modelo singular;
- Realização de Múltiplas Funcionalidades – criação de sistemas híbridos que emulam diferentes técnicas de processamento.

#### A.4 MODELOS DE PROGRAMAÇÃO GENÉTICA

O campo de Programação Genética é um campo muito similar ao de Algoritmos Genéticos. Todavia, há uma grande diferença entre elas: as representações que sofrem mutações e combinações são programas, em vez de cadeias de bits. Os programas são representados sob a forma de árvore de expressões, e as expressões podem estar em uma linguagem de programação ou podem ser projetadas especificamente para representar circuitos numa linguagem de descrição de hardware ou linguagens ESL (*Electronic System Level*). O cruzamento envolve a união de sub-árvores e não de subcadeias. Nesse contexto, a programação genética garante que as mutações e combinações gerarão expressões descendentes bem formadas, o que não ocorreria se fossem utilizados Algoritmos Genéticos com cadeia de bits.

De acordo com Russel (RUSSELL; NORVIG; INTELLIGENCE, 1995), a programação genética teve início por volta 1958 com os experimentos com código de máquina desenvolvidos por Friedberg (FRIEDBERG, 1958) e com os autômatos de estado finito, desenvolvidos por (FOGEL; OWENS; WALSH, 1966).

O objetivo da programação genética é criar um sistema onde cada indivíduo é um programa de computador capaz de executar uma tarefa predefinida. Cada programa é codificado como um conjunto de genes que, quando executado, são modificados (evoluídos), através de combinações e mutações nos genes.

Quando se trata de previsões de séries temporais financeiras, estudos na literatura mostram que a programação genética é utilizada para buscar os indivíduos ótimos de um conjunto de indivíduos, ou seja, otimização dos parâmetros de entrada de alguns modelos de previsão de séries temporais financeiras.

O trabalho de Araújo (ARAÚJO, 2011b) utiliza um algoritmo genético para otimizar uma rede morfológica, buscando os melhores valores subótimos de um modelo de RNA morfológico (*Modular Morphological Neural Network*, MMNN).

Os resultados mostram que a programação genética realmente melhorou o desempenho da RNA. No trabalho de Lam (LEUNG et al., 2003a), o autor modifica um algoritmo genético, melhorando a estrutura e os parâmetros de uma RNA. Os resultados mostram que as modificações no processo de combinação e mutação dos indivíduos geram uma melhora significativa quando comparado ao algoritmo genético padrão.

Já Ostermark (ÖSTERMARK, 2012) utilizou um modelo híbrido de algoritmo genético para otimizar o modelo desenvolvido que analisa um conjunto de portfólios para prever a melhor carteira que deve ser utilizada por investidores. Os resultados mostram que o uso do algoritmo genético otimizou a organização das carteiras e pode servir para qualquer investidor utilizar no processo de gestão de sua carteira.

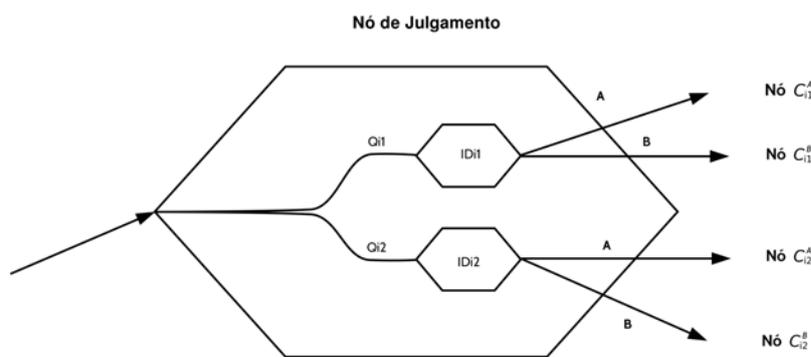
### A.4.1 Genetic Network Programming (GNP) and Mean Conditional Value-at-Risk Model (GNP-CVaR)

O trabalho desenvolvido por (CHEN; CHEN, 2015) descreve um sistema de negociações de ações híbrido baseado em Redes Genéticas Programáveis (GNP) e um Modelo Condicional Valor-Risco (GNP-CVaR). De acordo com o autor o sistema combina a vantagem de algoritmos evolutivos e modelos estatísticos, gerando estratégias para investidores com diferentes visões de Risco. A carteira é construída pelo modelo de média-variância de Markowitz similar ao modelo de média-CVaR.

O autor afirma que esta é a primeira vez que se mistura um modelo de análise de risco em combinação com uma Rede Genética Programável. Outra contribuição informada foi fixar a distribuição do capital inicial na GNP sem considerar o risco enquanto a variância média e o GNP-CVaR calculam a medida do risco e constrói o portfólio de cada carteira.

A Rede Gerérica Programável (GNP) consiste em um nó de início, nós de julgamento (figura 68) e nós de processamento figura (??), e as conexões entre estas. O nó de partida é utilizado para marcar o primeiro nó a ser executado. Nós julgamento tem ramificações condicionais que possam tornar-se decisões do caminho que seguirá. Os nós de processamento não têm ramos e são usados para tomar decisões de compra e venda de ações. A estrutura do modelo GNP e do gene pode ser visto na figura 70.

Figura 68 – Estrutura do nó de julgamento do GNP-CVaR.

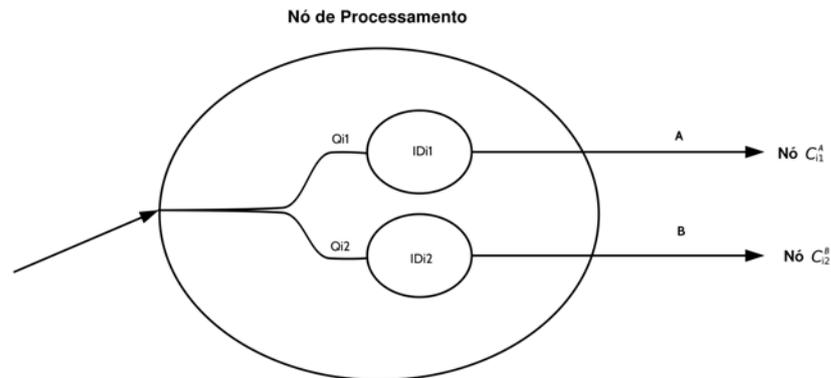


**Fonte:** (CHEN; CHEN, 2015) com ajustes pelo autor.

Dentro do nó de processamento, há o algoritmo que é executado para julgar o sub-nó que irá alcançar. O algoritmo é desenvolvido baseado em análise técnica existentes no mercado financeiro: Rate of Derivation (ROD), Relative Strength Index (RSI), Rate of Change (ROC), Volume Ratio (VR), Rank Correlation Index (RCI), Stochastics, Gold/-Dead Cross and Moving Average Convergence and Divergence (MACD). Para cada análise técnica o autor define as constantes utilizadas na análise.No ROC pode variar entre  $-6.5$  e  $6.5$  e no RSI pode variar entre  $0$  e  $100$ .

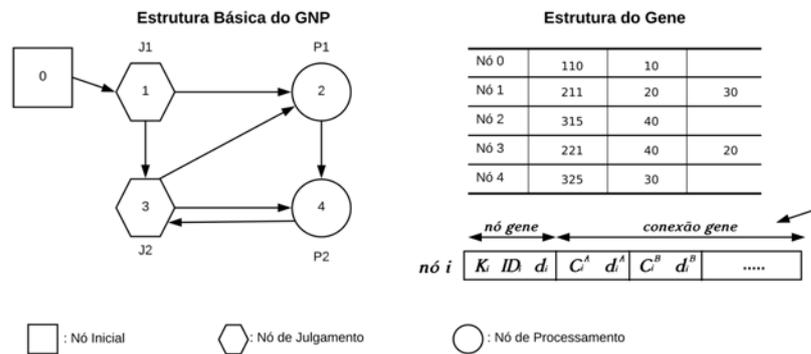
Para o sistema de Otimização de Portfólio, Chen e Wang escolhe um conjunto de portfólio  $x$  de um vetor  $X$  que pertence a  $R_n$  inserindo pesos a cada ativo da carteira.

Figura 69 – Estrutura do nó de julgamento do *Genetic Network Programming (GNP) and Mean Conditional Value-at-Risk Model (GNP-CVaR)*.



Fonte: (CHEN; CHEN, 2015) com ajustes pelo autor.

Figura 70 – Estrutura do modelo *Genetic Network Programming (GNP) and Mean Conditional Value-at-Risk Model (GNP-CVaR)*.



Fonte: adaptado de (CHEN; CHEN, 2015).

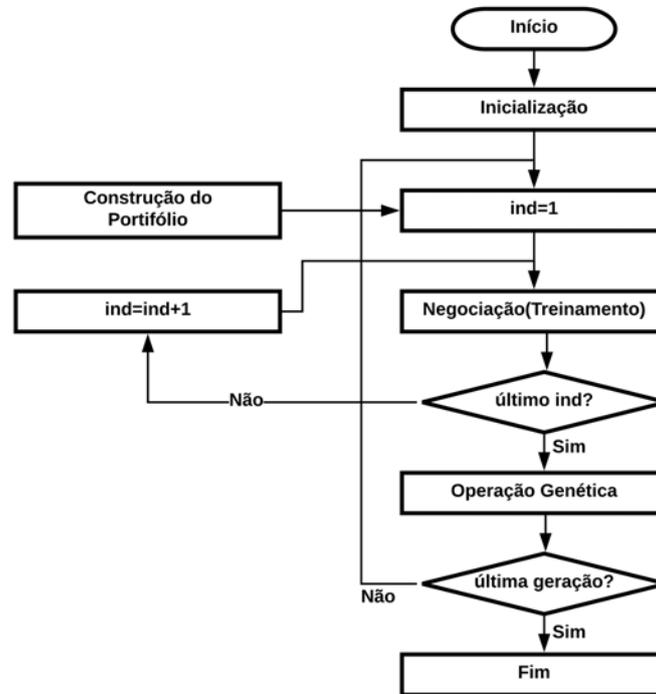
Um conjunto aleatório  $y$  em  $R_m$ , com distribuição de probabilidade subjacente é assumido por conveniência para calcular a densidade  $p(y)$ . Após isso, calcula-se a Variância de cada ativo, medida de ficou popular no mercado financeiro para análise de risco e foi descrito até nos regulamentos do setor financeiro. Após a nova forma de cálculo do portfólio o autor concluiu com os resultados experimentais que as carteiras destinadas a maximizar o risco-recompensa de volatilidade pode obter mais lucro no mercado utilizando seu sistema, e eles superar outros métodos de definição de carteiras.

A arquitetura do sistema híbrido para desenvolver o cálculo da carteira pode ser visto na figura 71:

#### A.4.2 Multi-Objective Evolutionary Algorithms-MOEA

O trabalho de Silva (SILVA; NEVES; HORTA, 2015) mostra um sistema híbrido com o uso de indicadores fundamentalistas e algoritmos evolucionários multiobjetos (MOEA,

Figura 71 – Estrutura do modelo *Genetic Network Programming (GNP) and Mean Conditional Value-at-Risk Model (GNP-CVaR)*.



**Fonte:** adaptado de (CHEN; CHEN, 2015).

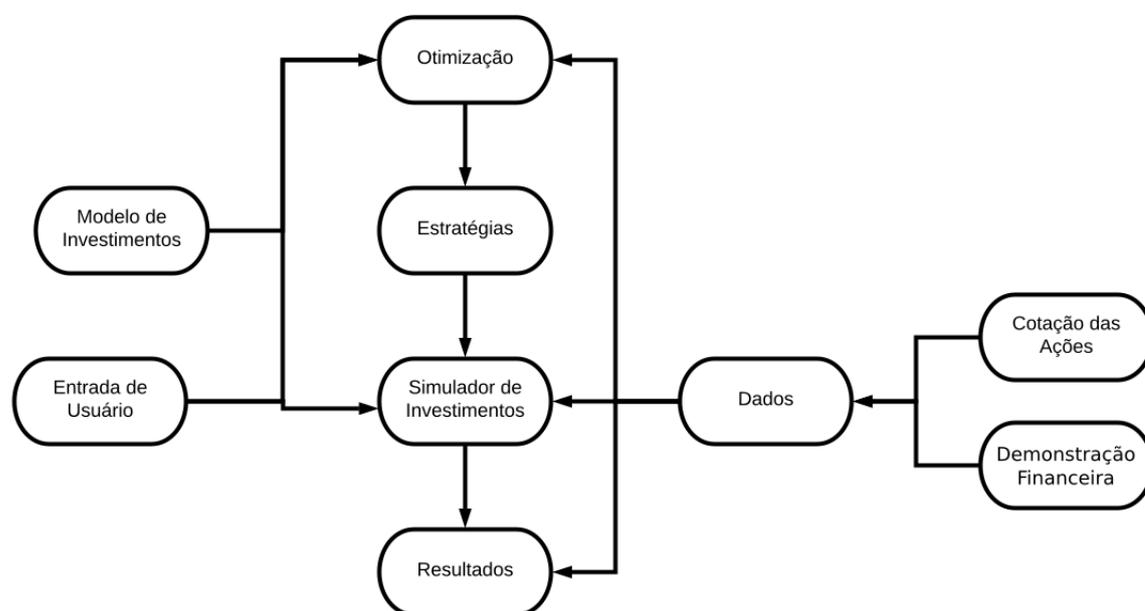
*Multi-Objective Evolutionary Algorithms*). Com o uso do sistema híbrido, Silva obtém resultados melhores que o uso apenas da análise fundamentalista. Este trabalho descreve uma nova abordagem para a gestão de carteiras através de existências. Os modelos de investimento testado incorporar uma abordagem fundamental e técnica usando índices financeiros e indicadores técnicos. *Multi-Objective Evolutionary Algorithms* (MOEA) tem duas saídas: o retorno de ganho capital e o risco. São utilizados para otimizar os modelos três cromossomos diferentes para representar o investimento em diferentes visões com restrições reais equivalentes às enfrentadas pelos gestores de carteira.

Para validar a presente solução foram utilizados de três estudos de caso na bolsa de valores *S&P500* no período de junho de 2010 até o final de 2014. Simulações demonstram que a seleção de ações com base em índices financeiros pode ser usado para escolher as melhores empresas em termos operacionais, a obtenção de retornos acima da média do mercado com baixas variações em seus retornos. O aumento dos indicadores fundamentais melhora a qualidade dos cromossomos encontrados pelo MOEA, e o resultados de simulações reais tornam-se mais preciso. Alguns dos melhores cromossomos encontrados pelos algoritmos são utilizados em investimento em ações com alto retorno sobre o patrimônio líquido (ROE). Seu uso em conjunto com a elevada taxa de crescimento da rede

rende uma margem de lucro alta. Para obter ações com alto potencial de valorização, é necessário escolher empresas com uma capitalização menor que a média de capitalização de mercado, baixo (PER), altas taxas de crescimento da receita e de alta alavancagem operacional.

O sistema proposto otimiza um modelo de investimento usando indicadores fundamentais e indicadores técnicos. O objetivo do algoritmo é escolher quais ações devem ser incluídos em uma carteira de ações. O fundamento indicadores são baseados nas demonstrações financeiras informações de cada empresa. A pré-seleção dos indicadores fundamentais utilizados no trabalho permite a descoberta de empresas estáveis com bom crescimento, um preço razoável, e que têm o potencial para um maior retorno. No entanto, é o algoritmo que é responsável por dizer se indicador deve ter um maior ou um menor peso na decisão final. As ações com um valor mais alto na soma de todos os indicadores fundamentais multiplicado por cada peso mais a soma de todos os indicadores técnicos são escolhidos para serem incluídos na carteira. Portanto, é uma melhor seleção de empresas na *S&P500* que vai permitir o algoritmo atingir rendimentos mais elevados do que o próprio índice *S&P500*. A estrutura do sistema pode ser visto na figura 72

Figura 72 – Estrutura do modelo *Multi-Objective Evolutionary Algorithms*.



**Fonte:** o autor.

Pode-se observar na figura 72 uma estrutura constituída por três módulos principais: o simulador de investimento, o otimizador e os dados. O módulo de dados MOEA é acessado

pelo simulador de investimentos para testar as estratégias e os instrumentos do módulo de otimização. O otimizador usa os resultados da rentabilidade acumulada e a variância, obtidos no bloco investidor para calcular a aptidão funcional e avaliar as estratégias. Com base na avaliação, este seleciona os cromossomos para a reprodução e aplica os métodos de cruzamento e mutação para criar novos elementos. O bloco de dados usa as demonstrações financeiras e cotações de ações para calcular as proporções utilizadas pelo simulador investidor. O Investidor Simulador testa as estratégias obtidos na otimização, dependendo do modelo de investimento e usando as entradas dadas pelo utilizador, para obter os dados para o teste. Cada dia em que avalia as ações e calcula a rentabilidade acumulada da carteira de o início da simulação e a variância da mensal Retorna.

Cada estratégia é testada no simulador, utilizando uma série temporal diária, onde haverá o incremento de informações diariamente. Os índices financeiros selecionados são avaliados e os cálculos são realizada para fazer as decisões de negociação do dia. Depois de terminar o período de tempo de cada período de formação, os resultados são registrados, e a reprodução da população é realizada. Este processo repete-se para a obtenção do número de interação que indica o final do treinamento. Após a conclusão do treinamento da população, o ficheiro externo com as melhores soluções é usado em um teste real. O sistema utiliza os seguintes parâmetros: o período de formação, e o período do teste real. Além disso, os custos de transações Também estão incluídos a 2% do valor do estoque.

#### **A.4.3 Modelo MGA-MLP**

O modelo de rede neural MGA-MLP é um sistema híbrido que junta conceitos de redes neurais e algoritmos genéticos. Um algoritmo genético faz parte de uma classe de algoritmos que realizam buscas evolucionárias em espaço de estados com o objetivo de buscar por uma configuração de indivíduo subótima de uma população inicial. Neste caso, foi criado um algoritmo genético modificado que se mostrou mais eficiente do que o algoritmo genético padrão (LEUNG et al., 2003b). Leung fez um modificação nas etapas de cruzamento e mutação que geraram melhorias na busca de um indivíduo com as mesmas características da população inicial. Para o trabalho de Ferreira (FERREIRA, 2006; LEUNG et al., 2003b), o MGA foi utilizado para, a partir de uma população inicial de configurações de redes MLP, o algoritmo buscará através do algoritmo modificado de Leung, uma rede mais evolucionária quando comparada a população inicial, com um mapeamento mais próximo da função que representa a série temporal financeira.

Para tanto, Ferreira (FERREIRA, 2006; LEUNG et al., 2003b) criou um conjunto de parâmetros que necessitam de modificações para encontrar a melhor rede MLP que absorva as características da função de previsão do modelo. A partir desses parâmetros é criado um conjunto de redes MLP iniciais (População) e depois o algoritmo genético modificado é executado para encontrar a melhor rede MLP subótima que se aproxima do modelo de previsão da série temporal financeira.

Desta forma o sistema híbrido MGA-MLP utiliza três modelos distintos para descrever a arquitetura da rede MLP:

1. Quantidade de retardos temporais, parâmetro relevante para representar a série temporal. Este valor é escolhido num intervalo  $[1, Maxlags]$ , onde  $Maxlags$  é a quantidade máxima de retardos temporais utilizada pela rede (Janela Temporal).
2. Quantidade de unidades de processamento na camada intermediária da rede MLP: este parâmetro é utilizado para expandir e diminuir a quantidade de neurônios que serão utilizados na camada escondida da rede MLP, podendo variar num intervalo  $[1, MaxHidden]$  onde  $MaxHidden$  é o número máximo de neurônios existentes na camada intermediária.
3. Configuração do peso e *bias* da rede MLP.

O sistema híbrido MGA-MLP funciona de acordo com a estrutura da figura 73.

Na figura 73 pode-se observar o ciclo de funcionamento do algoritmo MGA-MLP. Após a configuração dos parâmetros iniciais é criada a população inicial de redes e cada uma das redes é avaliada como solução do problema. Após esse processo, é escolhido um par de indivíduos pais e aplicado os operadores de cruzamento e mutação neles. No processo de cruzamento é criado três novos indivíduos que sofrerão mutação e inseridos na população novamente, no lugar dos 3 piores indivíduos existentes anteriormente. A condição de parada então é verificada e o processo se repetirá até que uma condição de parada seja encontrada.

A codificação de cada cromossomo de feita através dos seguintes parâmetros (FERREIRA, 2006):

$W_{ij}$ : matriz de pesos das conexões da camada de entrada  $i$  para a camada intermediária  $j$  da rede MLP;

$W_{ij}$ : matriz de pesos das conexões da camada intermediária  $j$  para a camada de saída  $k$  da rede MLP;

$b_j$ : vetor com os bias das unidades de processamento da camada intermediária  $j$  da rede MLP;

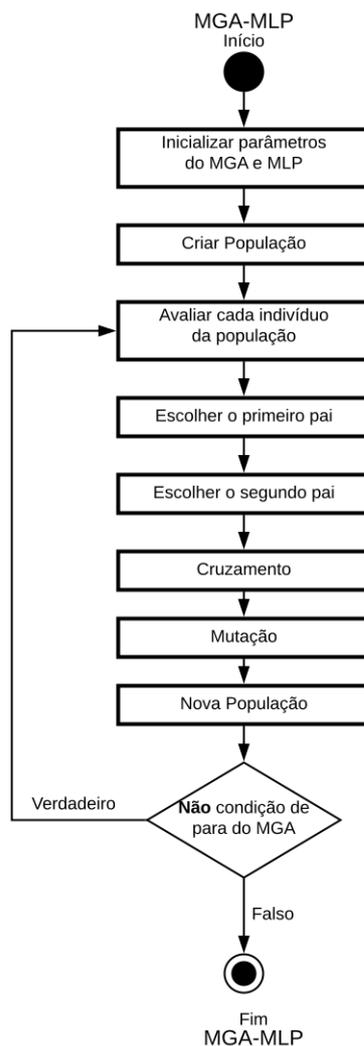
$b_k$ : vetor com os bias das unidades de processamento da camada de saída  $k$  da rede MLP;

**MlpMod**: modelo da rede MLP;

**MlpHiddenUnit**: quantidade de unidades de processamento na camada escondida da rede MLP;

**Lag**: vetor contendo o máximo de elementos  $Maxlags$ , onde cada um destes é utilizado para determinar se um retardo temporal específico será utilizado.

Figura 73 – Estrutura do modelo MGA-MLP.



**Fonte:** adaptado de (ARAÚJO, 2012b).

Para cada um dos indivíduos é avaliado o *fitness* de cada indivíduo. No trabalho de Ferreira o *fitness* é calculado por:

$$fitness = \frac{1}{1 + MSE}, \quad (A.16)$$

Como critério de parada são utilizadas três possibilidades:

1. Quantidade Máxima de gerações, definido no parâmetro (*MgaMaxGer*);
2. Aumento no erro de validação (*generalization loss*, GL) definido  $> 5\%$ ;
3. A queda no erro de treinamento (*Process training*, Pt) definido  $\leq 10^{-6}$ .

Ferreira (FERREIRA, 2006) utiliza três formas distintas de modelagem da rede MLP apenas modificando o uso da função de ativação sigmóide logística. O modo 1 ele utiliza a

equação A.17 onde a função de ativação sigmóide logística é utilizada para todas unidades de processamento escondidas e uma função de ativação linear é utilizada na unidade de processamento de saída com seu bias passando por uma função sigmóide logística. A função de ativação sigmóide logística utilizada é a equação A.18.

$$y_k(t) = \sum_{j=1}^{n_k} W_{jk} \text{Sig} \left[ \sum_{i=1}^{n_{in}} W_{ij} x_i(t) + b_j^1 \right] + \text{Sig}(b_k^2) \quad (\text{A.17})$$

Onde  $x_i(t)$  ( $i = 1, 2, \dots, n_{in}$ ) são os retardos temporais de entrada da rede MLP,  $n_{in}$  e  $n_h$  são parâmetros da rede MLP que indicam a quantidade de unidade de processamento de cada camada escondida, respectivamente.

$$\text{Sig}(x) = \frac{1}{1 + \exp(-x)} \quad (\text{A.18})$$

Como o objetivo de Ferreira é prever um passo a frente, utiliza-se apenas uma unidade de processamento na camada de saída ( $k=1$ ).

O segundo modelo de rede MLP definida  $MLPMod=2$  é a equação A.19 que consiste em utilizar uma função de ativação sigmoide logística na camada escondida e uma função de ativação linear na unidade de processamento da camada de saída, diferenciando do modo 1 por não utilizar a função de ativação sigmóide no bias.

$$y_k(t) = \sum_{j=1}^{n_h} W_{jk} \text{Sig} \left[ \sum_{i=1}^{n_{in}} W_{ij} x_i(t) + b_j^1 \right] + b_k^2 \quad (\text{A.19})$$

Para o modelo de rede MLP  $MLPMod=3$  a função de ativação sigmoide logística é utilizada em todas as unidades de processamento. Dessa forma a saída da rede MLP é dada pela equação A.20:

$$y_k(t) = \left\{ \sum_{j=1}^{n_k} W_{jk} \text{Sig} \left[ \sum_{iu=1}^{n_{in}} W_{ij} x_i(t) + b_j^1 \right] + b_k^2 \right\} \quad (\text{A.20})$$

#### A.4.4 Otimizador por Enxame de Partícula e Redes Neurais Multicamadas

O trabalho de Kennedy e Erberhart (KENNEDY, 2011; EBERHART; KENNEDY, 1995) introduz um conceito para a otimização de funções não lineares usando metodologia de enxames de partículas. O método foi descoberto através da simulação de um modelo social simplificado baseado na teori do enxame das aves (*bird flock*) e dos cardumes de peixes (*fish school*). O Otimizador por Enxame de Partícula (*PSO*) é um mecanismo de busca mais eficiente que o Algoritmo Genético e o Algoritmo Genético modificado. O objetivo de fazer a junção entre o PSO e as Redes Neurais MLP é desenvolver um sistema híbrido semelhante a MGA-MLP, porém mais otimizada na busca da configuração subótima (ou até mesmo ótima) da rede MLP para o problema em questão, modificando os parâmetros que compõe a rede MLP.

No sistema híbrido PSO-MLP cada partícula do enxame do PSO representa uma rede MLP candidata a solução do problema, onde os parâmetros de cada rede MLP da população são evoluídos a cada iteração do PSO. De forma semelhante ao MGA-MLP, o sistema híbrido PSO-MLP segue o paradigma de intensificação, onde o PSO melhora o processo de busca da rede MLP subótima. O PSO-MLP foi utilizado por Araújo (ARAÚJO, 2007) no estudo *Swarm-Based Hybrid Intelligent Forecasting Method For Financial Time Series Prediction - SHIF* para previsão de séries temporais financeiras. e se mostrou melhor que a rede MGA-MLP.

Assim como o MGA-MLP, o PSO-MLP consiste em três modelos distintos para descrever a arquitetura da rede MLP onde cada modelo é treinado e ajustado pelo Otimizador por Enxame de Partícula para determinar os melhores parâmetros:

1. Quantidade de retardos temporais, parâmetro relevante para representar a série temporal. Este valor é escolhido num intervalo  $[1, Maxlags]$ , onde *Maxlags* é a quantidade máxima de retardos temporais utilizada pela rede (Janela Temporal).
2. Quantidade de unidades de processamento na camada intermediária da rede MLP: este parâmetro é utilizado para expandir e diminuir a quantidade de neurônios que serão utilizados na camada escondida da rede MLP, podendo variar num intervalo  $[1, MaxHidden]$  onde *MaxHidden* é o número máximo de neurônios existentes na camada intermediária.
3. Configuração do peso e *bias* da rede MLP.

O sistema híbrido PSO-MLP funciona de acordo com a estrutura da figura 74.

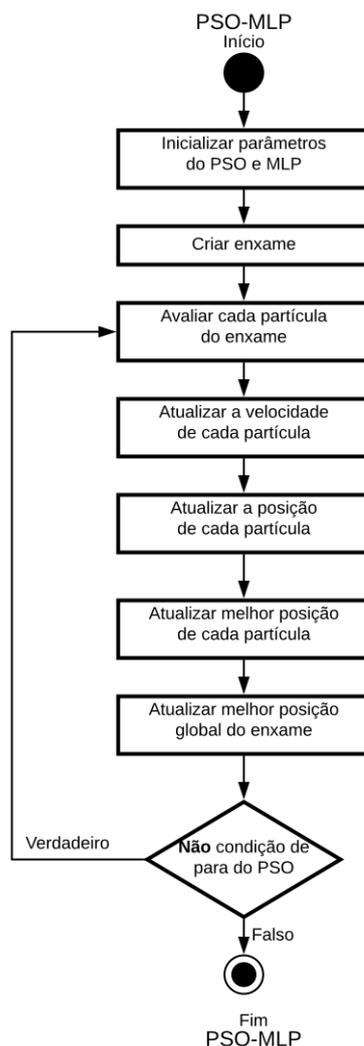
Na figura 74 pode-se observar o ciclo de funcionamento do algoritmo PSO-MLP. Após a configuração dos parâmetros iniciais é criado a população inicial de redes do Otimizador por Enxame de Partícula e cada uma das redes é avaliada como solução do problema. Após esse processo, é atualizada a velocidade, a posição, e a melhor posição individual de cada partícula, e a melhor posição global de cada partícula encontrada pelo enxame. A condição de parada é verificada e o processo se repetirá iterativamente até que uma das condições de paradas seja alcançada.

Os critérios de paradas do sistema híbrido PSO-MLP são (ARAÚJO, 2007):

1. Quantidade Máxima de gerações, definido no parâmetro (*MgaMaxGer*);
2. Aumento no erro de validação (*generalization loss*, GL) definido  $> 5\%$ ;
3. A queda no erro de treinamento (*process training*, Pt) definido  $\leq 10^{-6}$ .

Segundo Araújo, os sistemas MGA-MLP e PSO-MLP tem limitações. Esses sistemas híbridos não abordam o meio convencional de aprendizagem de redes MLP: aprendizagem supervisionada baseada no gradiente descendente (HAYKIN et al., 2009). Esse tipo de

Figura 74 – Estrutura do modelo PSO-MLP.



**Fonte:** adaptado de (ARAÚJO, 2010d).

aprendizagem é importante para se realizar buscas locais pela melhor rede subótima. O algoritmo baseado em gradiente descendente é uma opção para se trabalhar as buscas locais e suprir as necessidades dos sistemas híbridos MGA-MLP e PSO-MLP. Dessa forma surgiram dois outros sistemas híbridos para previsão de séries temporais financeiras: TAEF e SHIF.

#### **A.4.5 TAEF – *Time-delay Added Evolucionary Forecasting***

O Método TAEF foi desenvolvido por Ferreira (FERREIRA, 2006). A partir de sua definição pode-se verificar que o sistema híbrido incorpora duas técnicas de busca (global e local), no qual uma (global) é responsável pela definição dos retardos temporais, modelo MLP e algoritmo de treinamento da rede MLP, e a outra (local) é responsável pelo treinamento da rede MLP.

O objetivo desse método é tentar reconstruir o espaço de fase de uma série temporal realizando uma busca pela mínima dimensionalidade necessária para reproduzir o fenômeno da série temporal.

Para este método os parâmetros importantes são (FERREIRA, 2006; FERREIRA; VASCONCELOS; ADEODATO, 2008):

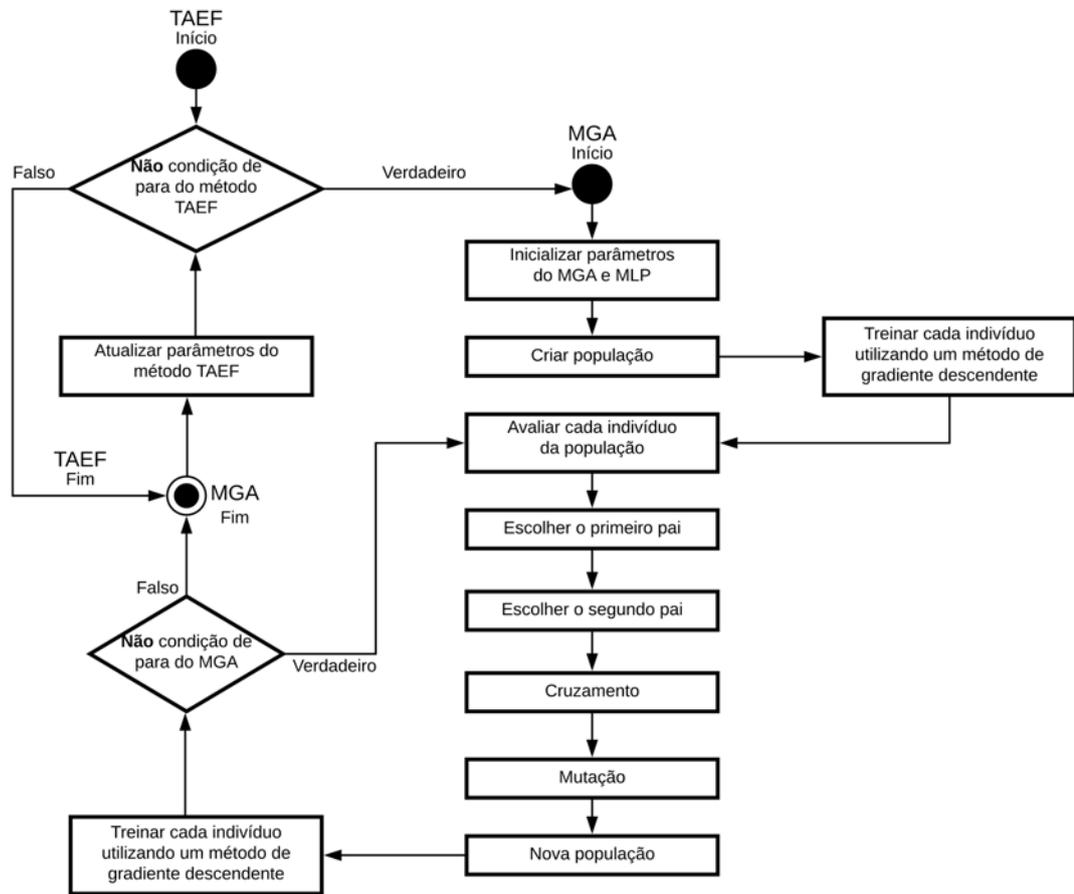
1. Quantidade de retardos temporais, parâmetro relevante para representar a série temporal. Este valor é escolhido num intervalo  $[1, Maxlags]$ , onde *Maxlags* é a quantidade máxima de retardos temporais utilizada pela rede (Janela Temporal).
2. Quantidade de unidades de processamento na camada intermediária da rede MLP: este parâmetro é utilizado para expandir e diminuir a quantidade de neurônios que serão utilizados na camada escondida da rede MLP, podendo variar num intervalo  $[1, MaxHidden]$  onde *MaxHidden* é o número máximo de neurônios existentes na camada intermediária.
3. O modelo e a arquitetura da rede MLP.
4. Algoritmo de treinamento da rede MLP para busca local (*resilient back-propagation* (RIEDMILLER; BRAUN, 1993), *Levenberg Marquardt* (HAGAN; MENHAJ, 1994), *Scaled Conjugate Gradient* (MØLLER, 1993), e *One Step Secant Conjugate Gradient* (BATTITI, 1992)).

A figura 75 mostra o esquemático do método TAEF:

Pode-se observar que inicialmente o método TAEF verifica se a condição de parada é satisfeita. Caso negativo, o algoritmo passa para a próxima etapa; inicializar os parâmetros do MGA-MLP e, posteriormente, cria uma população de indivíduos. Neste ponto é que se adiciona o algoritmo de busca local para o treinamento de cada rede MLP; cada rede neural da população é treinada utilizando um método de gradiente descendente e após o treinamento todo indivíduo da população é avaliado para identificar aqueles que serão escolhidos para o cruzamento. Nesta etapa são escolhidos dois pais para iniciar o processamento do algoritmo genético melhorado e estes pais passarão pelo processo de cruzamento e mutação. Após essas duas etapas, a busca local volta a acontecer novamente com a utilização do método de gradiente descendente para cada novo indivíduo (rede MLP) criado no processo. Baseado no MGA, alguns novos indivíduos são inseridos na população anterior e outros membros que faziam partes da população serão excluídos, finalizando a busca global pela melhor sub-rede ótima. Após este momento, os parâmetros do método TAEF são atualizados e este procedimento se repetirá até que uma condição de parada do método TAEF seja encontrada.

O método TAEF tem a codificação dos seus cromossomos através dos seguintes parâmetros:

Figura 75 – Estrutura do modelo TAEF.



Fonte: adaptado de (FERREIRA, 2006).

**MlpMod:** modelo da rede MLP;

**MlpHiddenUnit:** quantidade de unidades de processamento na camada escondida da rede MLP;

**MlpTrain:** algoritmo de treinamento MLP;

**Lag:** vetor contendo o máximo de elementos  $Maxlags$ , onde cada um destes é utilizado para determinar se um retardo temporal específico será utilizado.

As três formas de redes MLP (equação A.17, equação A.19, equação A.20) utilizadas nos sistemas híbridos MGA-MLP e PSO-MLP também são utilizadas no método TAEF. Assim como, para cada um dos indivíduos é avaliado o *fitness* de acordo com a equação A.16.

Os critério de parada para a MGA (FERREIRA, 2006) são:

1. Quantidade Máxima de gerações, definido no parâmetro (*MgaMaxGer*);
2. Aumento no erro de validação (*generalization loss*, GL) definido  $> 5\%$ ;
3. A queda no erro de treinamento (*Process training*, Pt) definido  $\leq 10^{-6}$ .

Os critérios de parada de treinamento de cada rede MLP (busca local) são:

1. Quantidade máxima de épocas (*MlpEpoch*);
2. O aumento no erro de validação ou *generalization loss* (*GL*) (PRECHELT et al., 1994)  $GL > 5\%$ ;

Já os critérios de parada do método TAEF foram definidos sendo:

1. A quantidade Máxima de iterações (*textititerMax*);
2. Valor mínimo da função de avaliação (*minFit*);
3. A queda no erro de treinamento (*Process training*, Pt) definido  $\leq 10^{-6}$ .

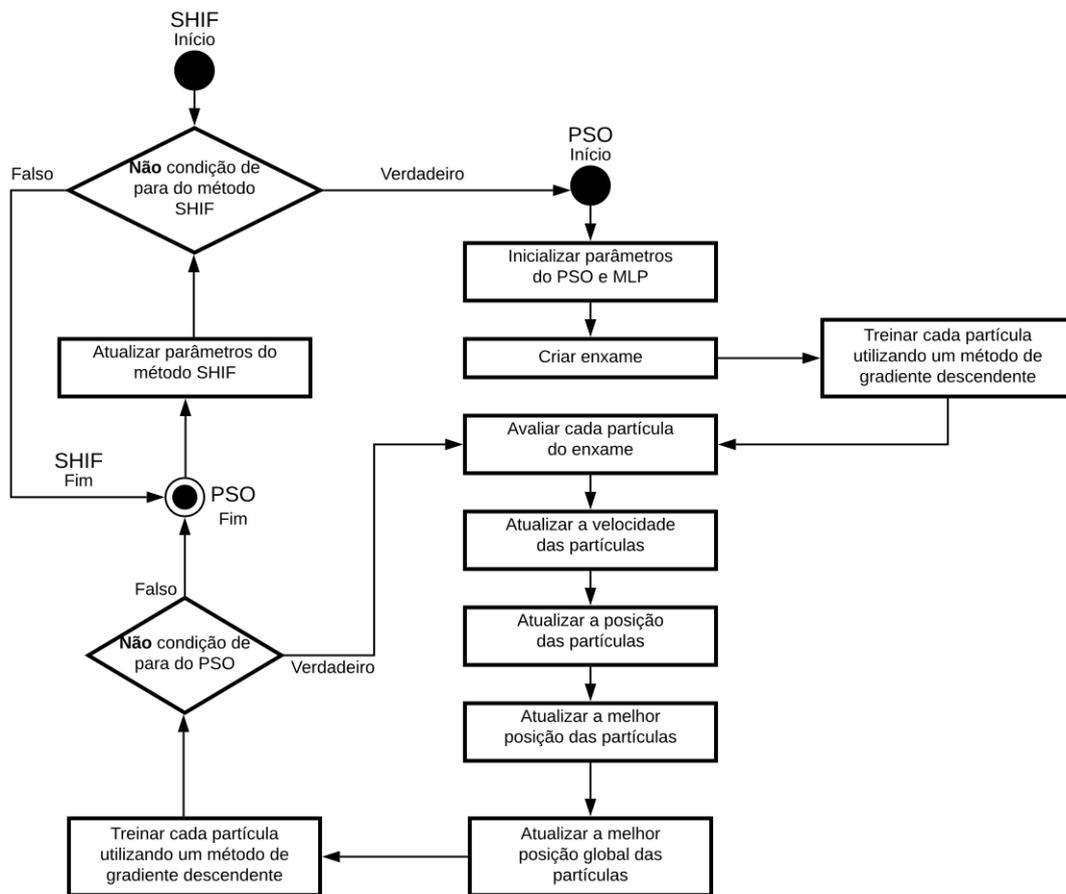
Por fim, o indivíduo de maior valor de avaliação do conjunto de validação dentre todos os gerados para cada execução do MGA é escolhido para ser o modelo de previsão gerado pelo método (FERREIRA, 2006; FERREIRA; VASCONCELOS; ADEODATO, 2008).

#### **A.4.6 SHIF – *Swarm – based Hybrid Intelligent Forecasting***

Assim como no método TAEF, o método *Swarm-based Hybrid Intelligent Forecasting* (SHIF) (ARAÚJO, 2007; ARAÚJO, 2010d) utiliza duas técnicas de buscas: a busca local e a busca global. A busca global definirá os retardos temporais, modelo, e algoritmo da rede MLP e a busca local é responsável pelo treinamento da rede MLP. Uma grande limitação no processo evolucionário do método TAEF é encontrado devido aos pesos e *bias* das redes MLPs serem inicializados aleatoriamente para cada indivíduo criado. A figura 76 explicita a forma que o método híbrido funciona.

De acordo com a figura 76, o modelo SHIF executa iterativamente o processo do PSO para determinar os parâmetros ótimos do modelo de previsão. Inicialmente, em cada iteração do modelo, os parâmetros do PSO são inicializados. Após este procedimento, o enxame do PSO é criado, isto é, um conjunto de redes MLP candidatas a solução do problema. Cada partícula (rede MLP) é então treinada por um método baseado em gradiente descendente (definido distintamente para cada partícula do enxame do PSO). Posteriormente, cada partícula do enxame do PSO é avaliada. A posição e a velocidade de cada partícula são atualizados. A melhor posição individual e global são então atualizadas. Tem-se então um novo enxame, que é treinado por um método baseado em gradiente descendente (definido individualmente para cada partícula do enxame do PSO). A condição de parada do

Figura 76 – Estrutura do modelo SHIF.



Fonte: adaptado de (ARAÚJO, 2007).

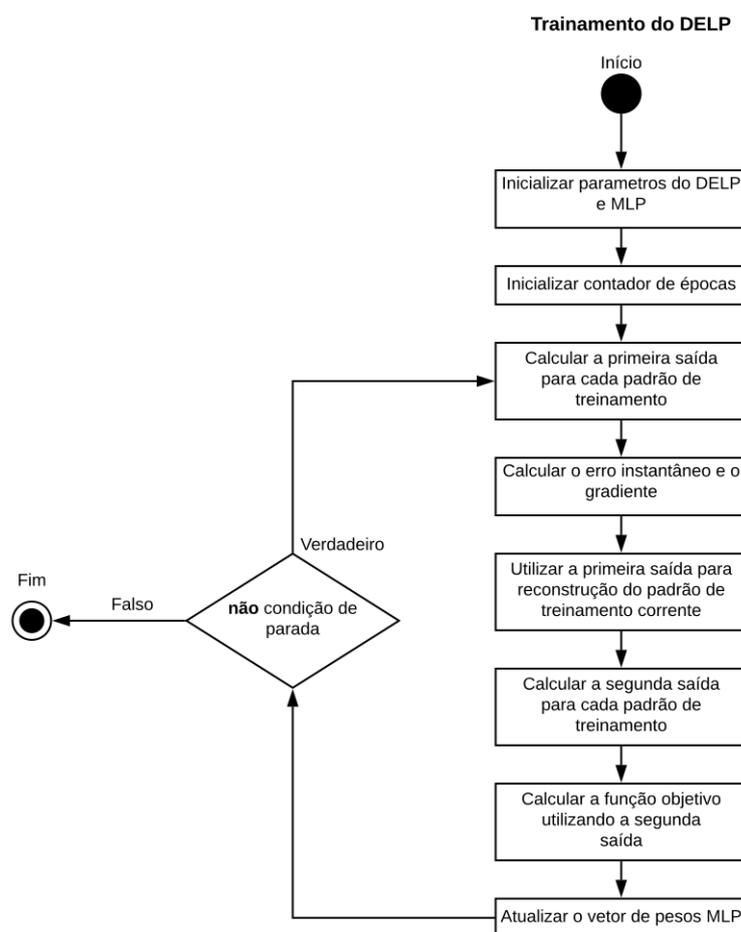
PSO é então verificada. Todo este procedimento é iterativamente repetido até que uma condição de parada do PSO seja encontrada. Após o término do processo do PSO, os parâmetros do modelo SHIF são atualizados e este procedimento é iterativamente repetido até que uma condição de parada do modelo SHIF seja encontrada (ARAÚJO, 2016).

#### A.4.7 DELP – *Dilatation Erosion Linear Perceptron*

O modelo de previsão DELP foi desenvolvido por araújo (ARAÚJO, 2012b) com o objetivo de criar um moledo de previsão para séries temporais financeiras com maior desempenho em relação aos existentes na literatura até aquele ano. O autor utilizou a morfologia matemática para calcular a componente não-linear identificada no estudos das séries analisadas em seu trabalho. A figura 77 mostra a arquitetura DELP. O modelo consiste de uma combinação linear entre operadores não-lineares (do tipo morfológicos via dilatação e erosão) e um operador linear (do tipo resposta finita ao impulso). A partir das modificações dos parâmetros de balanceamento de cada uma das componentes do modelo e do processo de

ajuste de fase o modelo consegue um desempenho preditivo estatisticamente superior a todos os modelos de previsão investigados em seu trabalho, sob as mesmas condições de experimentação, para todas as medidas de desempenho analisadas.

Figura 77 – Estrutura do modelo DELP.



**Fonte:** O autor.

Os resultados do trabalho de Araújo identificou que o modelo DELP obteve desempenho superior aos modelos ARIMA, MLP, MGA-MLP, PSO-MLP, TAEF e SHIF.