

UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ  
CAMPUS FOZ DO IGUAÇU  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
ENGENHARIA ELÉTRICA E COMPUTAÇÃO

DISSERTAÇÃO DE MESTRADO

**MÉTODO AUTOMÁTICO HÍBRIDO SSA-ARIMA-NEURAL PARA PREVISÃO  
MULTI-STEP DE SÉRIES TEMPORAIS ESTOCÁSTICAS**

GIONEI GAIO

FOZ DO IGUAÇU

2018



Gionei Gaio

**Método Automático Híbrido SSA-ARIMA-Neural  
para Previsão Multi-Step de Séries Temporais Estocásticas**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica e Computação; Área de concentração: Sistemas Dinâmicos e Energéticos.

Orientador: Prof. Dr. Edgar Manuel Carreño Franco

Coorientador: Prof. Dr. Luiz Albino Teixeira Jr.

Foz do Iguaçu

2018

Ficha de identificação da obra elaborada através do Formulário de Geração Automática do Sistema de Bibliotecas da Unioeste.

Gaio, Gionei

Método automático híbrido SSA-ARIMA-Neural para previsão multi-step de séries temporais estocásticas / Gionei Gaio; orientador(a), Edgar Manuel Carreño Franco; coorientador(a), Luiz Albino Teixeira Jr., 2018.  
89 f.

Dissertação (mestrado), Universidade Estadual do Oeste do Paraná, Centro de Engenharias e Ciências Exatas, Programa de Pós-Graduação em Engenharia Elétrica e Computação, 2018.

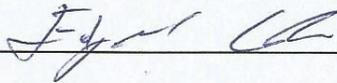
1. Forecasting. 2. Modelos Híbridos. 3. Decomposição SSA. 4. Redes Neurais Artificiais. I. Carreño Franco, Edgar Manuel. II. Teixeira Jr., Luiz Albino. III. Título.

**Método Automático Híbrido SSA-ARIMA-Neural  
para Previsão de Séries Temporais Estocásticas**

Gionei Gaio

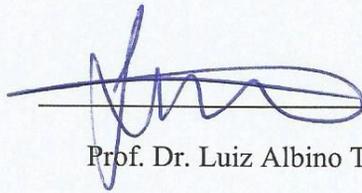
Esta Dissertação de Mestrado foi apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação e aprovada pela Banca Examinadora:

Data da defesa pública: 03/09/2018



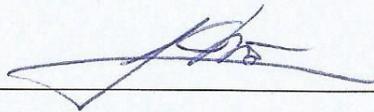
Prof. Dr. Edgar Manuel Carreño Franco – (Orientador)

Universidade Estadual do Oeste do Paraná – UNIOESTE



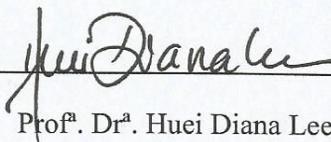
Prof. Dr. Luiz Albino Teixeira Jr. – (Coorientador)

Universidade Federal da Integração Latino-Americana – UNILA



Prof. Dr. Júlio César Royer

Instituto Federal do Paraná - IFPR



Profª. Drª. Hwei Diana Lee

Universidade Estadual do Oeste do Paraná – UNIOESTE



# Resumo

O desenvolvimento de metodologias para modelagem e previsão de séries temporais tem como objetivo fundamental diminuir as incertezas inerentes à maioria dos eventos (que não sejam determinísticos) futuros. A informação sobre o comportamento futuro de determinadas variáveis possibilita um melhor planejamento para as situações que hão de vir, independentemente de quais sejam. Para tal, aproveitando-se da ideia de que a maioria das séries temporais reais não possuem estruturas de interdependência puramente lineares ou não lineares, mas sim combinações delas, propõem-se uma nova metodologia híbrida automática designada SSA-ARIMA-neural.

Essa metodologia consiste em decompor a série original por meio de *Singular Spectrum Analysis* e, após isso, fazer o *forecast* separadamente para cada uma das componentes geradas que não tenham sido classificadas como ruído. A componente de tendência, por ser entendida como mais puramente linear, é modelada por meio da metodologia Box-Jenkins e, as componentes oscilatórias, por sua vez, tendo características de não linearidade, são aproximadas por Redes Neurais Artificiais. Ao final, então, todos os *forecasts* independentes são somados gerando a predição final.

Em busca de validar tal metodologia, foi realizado experimento computacional com um conjunto de dados advindo da instrumentação de uma barragem de concreto de grande porte de gravidade aliviada, onde os *forecasts* provenientes deste método foram comparados com os provenientes dos métodos consagrados na literatura. Os resultados mostraram uma melhoria relevante.

**Palavras-chave:** *Forecasting*, Modelos Híbridos, Decomposição SSA, Redes Neurais Artificiais, ARIMA.

# Abstract

The development of forecasting methods fundamentally aims to reduce the uncertainty inherent in predicting non-deterministic future events. It is necessary because the information about the future behavior of variables allows a better planning of situations to come, independent of what they are. For that, using the idea that real world time series are neither pure linear nor non-linear but instead a combination of those, a new hybrid automatic method called "SSA-ARIMA-Neural" is proposed.

This method consists in decomposing the original time series by means of Singular Spectrum Analysis and forecasting each component not classified as noise independently. The trend component, which is understood as more purely linear, is modeled by the Box-Jenkins methodology and the oscillatory components, in turn, having non-linear behavior, are approximated by Artificial Neural Networks. In the end, all independent forecasts are summed, generating the final prediction.

Aiming at validate this method, a computational experiment was performed using a data set obtained from instrumentation of a large concrete gravitational dam, where the resulting forecasts where compared with the ones generated using consecrated methods from the literature. Those comparisons showed a relevant improvement.

**Key-words:** Forecasting, Hybrid Models, SSA Decomposition, Artificial Neural Networks, ARIMA.

A meus pais, irmão, amigos, professores e amada esposa.  
Que me apoiaram nesta jornada.



# Agradecimentos

Gostaria de, primeiramente, agradecer a meus familiares que, em muitos momentos durante a jornada me apoiaram, me fazendo sempre seguir em frente. Em especial, agradeço a meu pai, a minha mãe e a meu irmão, sempre presentes e tão importantes em minha vida.

Estendo este agradecimento a meus amigos, que sempre me possibilitaram momentos de descontração e discussões divertidíssimas sobre os mais variados assuntos. Cito aqui os mais próximos, já me desculpando por eventuais esquecimentos: Samuel Barreto Siqueira, Ademir Wille da Cruz, Marcos Müller, Thaionara Ferreira, Gustavo Bodaneze, Maria Luísa Teixeira, Géssyca Fernanda, Yakine Alves, Jean Cassio, Vera Regina e Guilherme Portilho.

Não seria justo não agradecer aqui a empresa em que tenho trabalhado, Intelltech Tecnologias Inteligentes, por me disponibilizar a possibilidade de ocupar alguns de meus dias nesta dissertação. Em especial, agradeço a Jefferson Royer Chaurais e a Arthur Endlein Correia pela revisão do documento e pelas ideias concedidas.

Agradeço enormemente a todos os professores que, em algum momento, passaram em minha trajetória acadêmica. Porém, destaco aqui Edgar Manuel Carreño Franco e Luiz Albino Teixeira Jr. Por, respectivamente, orientar e coorientar esta dissertação com tanta paciência e disponibilidade.

Finalmente, agradeço a minha esposa, por sempre estar a meu lado me auxiliando em tudo que preciso e sempre me apoiar em todos meus projetos. Te amo.



# Sumário

<b>Lista de Figuras</b>	<b>xiii</b>
<b>Lista de Tabelas</b>	<b>xv</b>
<b>Lista de Símbolos</b>	<b>xvii</b>
<b>Capítulo 1 Problema</b>	<b>1</b>
1.1 Objetivos.....	4
1.1.1 Objetivo Geral.....	4
1.1.2 Objetivos Específicos.....	4
1.2 Pressupostos.....	5
1.3 Hipóteses.....	5
1.4 Estrutura do Trabalho.....	5
<b>Capítulo 2 Referencial Teórico</b>	<b>7</b>
2.1 Características de uma série temporal.....	7
2.2 Análise de Espectro Singular.....	9
2.3 Modelos ARIMA.....	10
2.3.1 SARIMA.....	14
2.3.2 <i>Forecasting</i> a partir dos modelos ARIMA.....	14
2.4 Redes Neurais Artificiais.....	15
2.4.1 Neurônios artificiais.....	15
2.4.2 Classificações de uma RNA.....	18
2.4.3 Transformação.....	18
2.4.4 RNA <i>Multi-Layer Perceptron Feedforward</i> .....	19
2.4.5 Previsão a partir de RNAs.....	20
2.5 Modelagem Híbrida.....	21
<b>Capítulo 3 Materiais e métodos</b>	<b>27</b>
3.1 Método SSA-ARIMA-Neural.....	27
3.2 Plataforma de Programação R.....	30
3.2.1 Pacote Forecast.....	30
3.2.2 Pacote Rssa.....	31
3.2.3 Pacote RSNNS.....	33
3.2.4 Pseudocódigo do Método SSA-ARIMA-Neural Automatizado.....	35

3.3	Dados Utilizados .....	36
3.3.1	Pêndulos.....	36
3.3.2	Piezômetros.....	37
3.3.3	Bases de Alongâmetro .....	37
<b>Capítulo 4 Resultados e Discussões</b>		<b>39</b>
<b>Capítulo 5 Conclusão</b>		<b>47</b>
<b>Referências Bibliográficas</b>		<b>49</b>
<b>Apêndice A Modelagem de Séries Temporais utilizando modelos de alisamento</b>		<b>55</b>
A.1	Médias Móveis .....	55
A.2	Suavização Exponencial De Primeira Ordem .....	56
A.3	Suavização Exponencial de Holt e Winters .....	57
<b>Apêndice B Gráficos das Séries Temporais e dos Forecasts</b>		<b>61</b>
<b>Apêndice C Parâmetros Finais</b>		<b>69</b>

# Lista de Figuras

2.1 - Esquema simplificado de um neurônio biológico.....	16
2.2 - Esquema de um neurônio artificial. ....	16
2.3 - Estrutura básica de uma RNA <i>multi-layer perceptron</i> . ....	19
2.4 - previsão utilizando RNA MLP. ....	20
3.1 - Fluxograma do método SSA-ARIMA-Neural.....	28
4.1 - Número de vezes que cada um dos métodos conseguiu o menor MAPE se comparado aos demais. ....	41
4.2 - Número de vezes que cada um dos métodos conseguiu o menor ou segundo menor MAPE se comparado aos demais.....	42
4.3 – Quantidade de vezes que uma metodologia obteve o pior resultado se comparada as demais. ....	42
4.4 – Predições baseadas na série temporal das leituras da base de alongâmetro 8.....	43
4.5 – Comparação do desempenho do método SSA-ARIMA-Neural automatizado comparado a cada um dos benchmarks.....	44
4.6 – Gráfico dos resultados obtidos com a série histórica da Base de alongâmetro 7...	44
B.1 – Dados do Instrumento Pêndulo Direto. ....	61
B.2 – Dados do Instrumento Pêndulo Invertido.....	61
B.3 – Dados do Instrumento Base de Alongâmetro 1.....	62
B.4 – Dados do Instrumento Base de Alongâmetro 2.....	62
B.5 – Dados do Instrumento Base de Alongâmetro 3.....	62
B.6 – Dados do Instrumento Base de Alongâmetro 4.....	63
B.7 – Dados do Instrumento Base de Alongâmetro 5.....	63
B.8 – Dados do Instrumento Base de Alongâmetro 6.....	63
B.9 – Dados do Instrumento Base de Alongâmetro 7.....	64
B.10 – Dados do Instrumento Base de Alongâmetro 8.....	64
B.11 – Dados do Instrumento Base de Alongâmetro 9.....	64
B.12 – Dados do Instrumento Base de Alongâmetro 10.....	65
B.13 – Dados do Instrumento Base de Alongâmetro 11.....	65
B.14 – Dados do Instrumento Base de Alongâmetro 12.....	65
B.15 – Dados do Instrumento Piezômetro 1. ....	66
B.16 – Dados do Instrumento Piezômetro 2. ....	66

B.17 – Dados do Instrumento Piezômetro 3.....	66
B.18 – Dados do Instrumento Piezômetro 4.....	67
B.19 – Dados do Instrumento Piezômetro 5.....	67
B.20 – Dados do Instrumento Piezômetro 6.....	67
B.21 – Dados do Instrumento Piezômetro 7.....	68

# Lista de Tabelas

Tabela 4.1 - Ajuste das previsões de acordo com o MAPE obtido.....	40
Tabela 4.2 – Propriedades estatísticas dos MAPEs .....	40
Tabela C.1 – Parâmetros finais de cada modelo. ....	69



# Lista de Símbolos

$\{.\}_{i=1}^N$	Conjunto de variáveis de tamanho $N$
$L_T$	Componente linear do processo estocástico não observado
$N_T$	Componente não linear do processo estocástico não observado
$X_I$	Matriz proveniente do I-ésimo grupo de autotriplas
$X_i$	Matriz proveniente da i-ésima autotripla da matriz trajetória
$Y_T$	Série temporal diferenciada
$Z_T$	Processo estocástico gerador de série temporal não observado
$b_k$	<i>Bias</i> do neurônio $k$
$v_k$	Campo local induzido
$w_{kj}$	Peso sináptico que liga a entrada $j$ ao neurônio $k$
$x_{I(i,j)}$	elemento da linha $i$ e coluna $j$ da matriz $X_I$
$y_t$	Elemento observado de série temporal diferenciada
$z_{I_t}$	Observação da I-ésima componente SSA no instante $t$
$z_t$	Observação de uma série temporal
$z_t^*$	Observação de série temporal transformado
$\mathbf{z}_I$	Vetor representante da I-ésima componente SSA no espaço $\mathbb{R}^N$
$\nabla$	Operador diferença
AR	Autoregressivo
ARIMA	<i>Autoregressive Integrated Moving Average</i>
ARMA	Autoregressivo e médias móveis
MA	<i>Moving Average</i>
MAPE	<i>Mean Absolute Percentual Error</i>
mlp	<i>Multi-Layer Perceptron</i>
RNA	Rede Neural Artificial
SARIMA	<i>Seasonal Autoregressive Integrated Moving Average</i>

SSA	<i>Singular Spectrum Analysis</i>
$\Theta(B)$	Polinômio de atraso da parte médias móveis Sazonal de um modelo SARIMA.
$\Phi(B)$	Polinômio de atraso da parte autorregressiva Sazonal de um modelo SARIMA.
$\Omega$	Espaço amostral
$B$	Operador Atraso
$E(\cdot)$	Média de uma variável aleatória.
$L$	Tamanho de Janela SSA
$X$	Matriz trajetória
$Z(t)$	Variável aleatória que compõem o processo estocástico $Z_T$ no instante $t$
$\gamma(\cdot)$	Função autocovariância
$\theta(B)$	Polinômio de atraso da parte médias móveis de um modelo ARIMA.
$\mu(\cdot)$	Média populacional.
$\varphi(\cdot)$	Função de ativação
$\phi(B)$	Polinômio de atraso da parte autorregressiva de um modelo ARIMA.

# Capítulo 1

## Problema

Este trabalho tem como principal motivação a apresentação de uma metodologia para previsão de series temporais, que consiga combinar as qualidades de previsores lineares, sendo o mais notável deles o ARIMA – *Autoregressive Integrated Moving Average*, qual será utilizado, com previsores não lineares, dos quais se destacam as redes neurais artificiais, também presente na metodologia a ser abordada, para isso fazendo uso de decomposição pela metodologia *Singular Spectrum Analysis* (SSA).

O desenvolvimento de metodologias para modelagem e previsão de séries temporais tem como objetivo fundamental diminuir as incertezas inerentes à maioria dos eventos, que não sejam determinísticos, futuros. A informação sobre o comportamento futuro de determinadas variáveis possibilita um melhor planejamento para as situações que hão de vir, independentemente de quais sejam.

Segundo Morettin & Tolo (2006), de maneira informal, uma série temporal é qualquer conjunto de observações coletadas em sequência no decorrer do tempo. Essa é composta por quatro elementos: ciclos, que representam ondulações nas séries, ou seja, padrões de crescimento e decrescimento que se repetem no tempo, nem sempre em intervalos regulares; sazonalidades, que são padrões que se repetem em intervalos regulares de um ano; tendência, que diz respeito ao deslocamento da série no longo prazo e; ruído aleatório, também chamado erro, que não pode ser modelado, por ser fruto de aleatoriedade e imprecisões. Assim sendo, todo o histórico de medições tomado no decorrer do tempo pode ser tratado como uma série temporal, cabendo assim modelagem matemática sobre ela.

Ainda de acordo com Morettin & Tolo (2006), a análise de séries temporais pode estar interessada em: investigar o mecanismo gerador das séries; descrever seu comportamento; procurar periodicidades e; fazer previsões dos valores porvir. Neste trabalho o foco principal é neste último ponto, já que o campo de aplicação de uma técnica que tenha resultados satisfatórios é muito abrangente. Por exemplo, muitas vezes é interessante antecipar índices econômicos, demanda de energia, comportamento de estruturas, refletidos nas medidas de instrumentos instalados, etc. principalmente com a finalidade de prever eventos, de forma a evitar, ou se precaver contra, comportamentos indesejados, ou seja, o que se visa é fornecer mais informação para melhores tomadas de decisão em vários níveis.

A projeção das séries temporais, também chamado *forecast*, é possível desde que: os dados apresentem estruturas de interdependência, também chamadas autocorrelação; que a

modelagem efetuada sobre os dados seja adequada e; a série respeite condições de regularidade, ou seja, não haja mudanças no processo gerador no período que se deseja estimar.

Uma metodologia bastante conhecida para modelagem e previsão de séries temporais é a proposta por Box & Jenkins (1970), que consiste no ajuste de um modelo ARIMA a um conjunto de dados. Segundo Moretin & Toloí (2006), este método é baseado em um processo iterativo que começa com a identificação do modelo, após isso, se estimam os parâmetros e, feito isso, há a verificação e diagnóstico do modelo por meio de análise de resíduos.

Contudo, conforme Zhang (2003), uma importante limitação dos modelos ARIMA reside no fato de este ser linear, pois muitas séries reais apresentam características não lineares que não podem ser satisfatoriamente aproximadas por relações lineares. Neste sentido Redes Neurais Artificiais (RNAs) têm sido amplamente estudadas para fins de previsão, devido, principalmente à sua flexibilidade com relação à modelagem não linear. Segundo Zhang, Patuwo & Hu (1998) e Yan (2012), as Redes Neurais mais utilizadas no contexto da previsão são as *feedforward multi-layer perceptron*, treinadas através do algoritmo *backpropagation* ou derivações do mesmo.

Zhang, Patuwo & Hu (1998) e Royer (2016) argumentam que, um problema inerente à utilização de Redes Neurais para a modelagem de séries temporais é o fato de não existirem valores predefinidos ou meios algébricos para determinação, por exemplo, do número de neurônios na camada de entrada e na camada escondida, necessitando experimentação para uma escolha adequada.

Apesar da popularidade, Khashei & Bijari (2011) argumentam, baseado em uma série de trabalhos, que em algumas situações específicas seu desempenho pode ser inconsistente, podendo até mesmo ser pior do que os modelos lineares, principalmente quando utilizadas para modelar problemas lineares.

Ainda segundo Khashei & Bijari (2011), cada vez mais tem se tornado uma prática comum a utilização de modelagens híbridas, sempre em busca de superar as limitações de um ou outro método, principalmente os que buscam decompor as séries em parte linear e não linear.

Pioneiro neste aspecto, Zhang (2003), partindo da ideia que a maioria das séries reais complexas não é puramente linear ou não linear, ou seja, dada uma série temporal  $Z_T$ , ela é tal que:  $Z_T = L_T + N_T$ , onde  $L_T$  representa a parte linear e  $N_T$  a parte não linear. A fim de tentar captar tanto os padrões lineares quanto não lineares Zhang propõem ajustar um modelo ARIMA (captando a estrutura linear) e depois, nos resíduos, uma RNA *feedforward* (capturando a estrutura não linear que esses possam conter). Desde então, muitos autores apresentam métodos similares, dentre eles Khashei & Bijari (2011), Gairaa *et al.* (2016), Han *et al.* (2017), etc. Contudo, entende-se que esta abordagem, ao ajustar a série original a partir de um modelo, este fará aproximações de acordo com suas limitações, ou seja, aproximações lineares das não linearidades, fazendo restar pouco da dinâmica não linear para ser analisada pelo modelo subsequente.

Pensando nisso, neste trabalho, propõe-se que a série original seja decomposta antes da modelagem. Alguns estudos como os de Chau & Wu (2010), Wu, Chau & Fan (2010) e Wu &

Chau (2011) para predição de chuvas, e Wang *et al.* (2015) para predição de escoamentos, têm feito desta forma, utilizando principalmente a análise de espectro singular (SSA) ou *ensemble empirical mode decomposition* (EEMD) como métodos de decomposição. Contudo, são poucos os trabalhos que utilizam técnicas de decomposição para combinar modelagens lineares e não lineares, deixando assim de aproveitar os pontos fortes que poderiam vir de suas características.

Assim, é proposta uma técnica que consiste, primeiro, na decomposição da série original via SSA, em vários componentes, um de tendência, um ou mais oscilatórios e um de ruído. Para a componente tendencial, é ajustado um modelo ARIMA, dado que esta, por ter oscilações lentas é satisfatoriamente aproximada pela modelagem linear deste modelo. Os demais componentes não classificados como ruído são modelados por RNAs *feedforward multi-layer perceptron*, buscando capturar a estrutura de dependência destes componentes utilizando as propriedades não lineares que neles contenham. Feitas as modelagens de todas as componentes e suas respectivas previsões elas são somadas (já que a técnica de decomposição SSA gera componentes que, se somadas, são iguais à série original) para gerar a previsão final. Esta técnica será referida como SSA-ARIMA-Neural no decorrer do trabalho. Técnicas consagradas na literatura são utilizadas como comparativo, a saber: ARIMA; Redes Neurais Artificiais; médias móveis e amortecimento exponencial. A técnica proposta por Zhang (2003), qual traz o pressuposto que inspira este trabalho, também será incluída no comparativo.

Dois pontos valem ser ressaltados. Primeiro, mesmo havendo na literatura algumas propostas de modelagem que utilizem técnicas de decomposição para hibridização de modelos lineares e não lineares, estas são escassas e não há, até a publicação deste trabalho, metodologia que combine modelos ARIMA, RNAs *feedforward* e SSA da maneira apresentada. Segundo, a metodologia proposta engloba também métodos de escolha automático da maioria dos parâmetros necessários para os preditores envolvidos, sendo que, os demais parâmetros são todos escolhidos *a priori*. Ou seja, o método foi automatizado de tal forma que independe da ação do analista em passos intermediários.

A nova metodologia foi validada fazendo-se a modelagem e previsão de séries de medidas de instrumentos instalados em uma barragem de concreto de gravidade aliviada de grande porte. Neste sentido, conseguiu-se, com a metodologia a ser exposta a seguir, estimativas confiáveis do comportamento estrutural da barragem na ausência de mudanças bruscas de regime (já que, neste caso, fica inviável a previsão por qualquer método que seja), mesmo que as séries estimadas sejam bastante heterogêneas, pois, com a utilização de diferentes tipos de modelagem, é razoável que o método aqui proposto tenha ganho versatilidade. Os instrumentos usados para *benchmark* englobam:

- Piezômetros, que medem a subpressão causada pela água na barragem, responsável pela diminuição do coeficiente de atrito entre a estrutura e a base sobre a qual está assentada, podendo levar até a escorregamentos;
- Pêndulos diretos e invertidos, que medem a inclinação da represa, na direção montante-jusante e na direção paralela a barragem, monitorados visando evitar tombamentos e;

- Bases de alongômetros, que medem abertura e fechamento entre juntas de monólitos, que são utilizados para identificar surgimento de rachaduras na estrutura.

Séries temporais extraídas destes instrumentos apresentam, normalmente, comportamento bastante heterogêneo, podendo apresentar tendências de longo prazo e componentes cíclicas, algumas vezes complexas, o que dificulta o trabalho de modelagem e previsão por técnicas convencionais, assim sendo, os efeitos de decomposição devem ser sentidos, justificando o uso de uma metodologia distinta das convencionais. Assim sendo, o método aqui apresentado é flexível e o conjunto de dados selecionados visa explicitar esta característica. Ressalta-se que se pode aplicar este método também a outros dados, dada sua flexibilidade esperada, desde que sejam igualmente espaçados no tempo.

## 1.1 Objetivos

### 1.1.1 Objetivo Geral

Este trabalho tem por objetivo geral propor um novo método para modelagem e previsão de séries temporais, mais assertivo e flexível que os métodos convencionais, capaz de aproveitar as estruturas de auto dependência lineares e não lineares, utilizando, para isso, a decomposição via *Singular Spectrum Analysis* e a modelagem por ARIMA ou via RNAs de cada um dos componentes não classificados como ruído aleatório, de acordo com suas características.

### 1.1.2 Objetivos Específicos

- Implementar um script que utilize a biblioteca RSSA para performar a decomposição das séries temporais por meio de SSA.
- Utilizar a biblioteca forecast para, dada uma série temporal, selecionar o modelo ARIMA que melhor se ajuste aos dados de tendência e calcule as previsões desta componente.
- Desenvolver uma metodologia que seja capaz de selecionar os melhores parâmetros a serem utilizados na modelagem via RNAs, a saber, quantidade de neurônios na camada de entrada e quantidade de neurônios na camada escondida e, além disto, utilizando os demais parâmetros estabelecidos previamente (tipo de normalização, tipo de função de ativação e método de aprendizagem), treinar a rede neural artificial e calcular as previsões.
- Unir as implementações anteriores no método SSA-ARIMA-Neural e utilizá-lo para prever o comportamento das medidas de sensores de uma barragem de concreto de gravidade aliviada de grande porte de forma a validar os resultados, comparando o desempenho do novo método com outros já consagrados na literatura, a saber: RNA; ARIMA; SARIMA; Holt-Winters; além do método proposta por Zhang (2003).

## 1.2 Pressupostos

São pressupostos que norteiam este trabalho:

- As séries temporais possuem parte linear e parte não linear, de acordo com o que argumenta Zhang (2003).
- A técnica SSA é capaz de separar as séries em componentes com estruturas de autodependência mais simples, mais puramente lineares (principalmente a primeira componente) e outras mais puramente não lineares.

## 1.3 Hipóteses

As hipóteses consideradas são:

- As RNAs se beneficiarão das estruturas de autodependência mais simples e, com isso, produzirão melhores aproximações das funções geradoras e, conseqüentemente, melhores previsões.
- A combinação das técnicas SSA, ARIMA e RNAs pode produzir previsões de séries temporais mais acuradas do que o uso de outros métodos já consagradas na literatura.

## 1.4 Estrutura do Trabalho

A partir daqui, este trabalho apresentará: No capítulo 2, o referencial teórico por trás dos métodos utilizados na modelagem e previsão das séries temporais assim como uma revisão da literatura abordando algumas das diversas formas propostas para *forecasting* em artigos recentes; No capítulo 3 uma apresentação dos métodos utilizados neste trabalho; No capítulo 4 a apresentação dos resultados e, por fim, as conclusões.



# Capítulo 2

## Referencial Teórico

### 2.1 Características de uma série temporal

Uma série temporal, segundo Yafee & McGee (2000), é uma sequência de observações de uma variável de interesse ordenadas no tempo. Elas podem ser encontradas abundantemente em vários ramos da ciência, pode-se citar, como exemplo, sem a pretensão de ser exaustivo, as sociais (como a taxa de aprovação de um presidente e proporção da população com nível superior), as biológicas (tamanho de uma população de bactérias, pressão arterial de um indivíduo) e também as exatas (Energia elétrica gerada por uma usina, consumo de energia).

Conceito bastante íntimo ao de série temporal é o de processo estocástico. Um processo estocástico, segundo Moretini & Toloi (2006) é uma família de variáveis aleatórias tal que, se denotarmos o processo estocástico por  $Z_T$ , então  $Z_T = \{Z(t), t \in T\}$ , onde, para cada  $t \in T$ , sendo  $T$  um conjunto arbitrário, normalmente tomado nos inteiros,  $Z(t)$  é uma variável aleatória. É comum supor que tal família vem de um mesmo espaço de probabilidades  $(\Omega, F, P)$ <sup>1</sup>, assim, na verdade,  $Z_T$  é função de duas variáveis,  $t$  e  $\omega$ , sendo que  $\omega \in \Omega$ . Sob esta ótica, mantendo-se constante  $\omega$ , tem-se  $Z_T$  como função do tempo, ou seja, uma realização do processo estocástico, que nada mais é que uma série temporal. O processo estocástico, por sua vez, pode ser interpretado como uma família de trajetórias.

Vale destacar que as variáveis aleatórias pertencentes ao processo estocástico nem sempre possuem a mesma função de densidade de probabilidade (f.d.p.), para o caso contínuo, ou função de massa de probabilidade (f.m.p.), para o caso discreto, em diferentes instantes de tempo e, em termos práticos, estas distribuições não são conhecidas<sup>2</sup>. Assim sendo, o que se faz comumente é limitar o estudo aos momentos de baixa ordem. Para isso, normalmente se utilizam as funções média e autocovariância, definidas na Equação 2.1 e 2.2, respectivamente (Moretini & Toloi, 2006):

$$\mu(t) = E\{Z(t)\} = \int_{-\infty}^{\infty} zf(z; t)dz \quad (2.1)$$

---

<sup>1</sup> Composto por espaço amostral ( $\Omega$ ), eventos ( $F$ ) e medida de probabilidade ( $P$ ).

<sup>2</sup> Daqui para frente, por simplicidade, somente o caso contínuo será considerado.

$$\gamma(t_1, t_2) = E\{Z(t_1)Z(t_2)\} - E\{Z(t_1)\}E\{Z(t_2)\}, t_1, t_2 \in T \quad (2.2)$$

Segundo Sousa & Camargo (2004), em muitos casos, somente uma série temporal é observada e, a partir dela, é estimado um processo estocástico que possivelmente a gerou, obtendo-se estimativas que o caracterizem. Yaffe & McGee (2000) dizem que, nestes casos, cada observação deve ser tomada como uma média amostral da distribuição de probabilidades de uma população subjacente. Além disso, sob a assunção de ergodicidade<sup>3</sup>, também se estimam os outros momentos de interesse do processo estocástico gerador, normalmente variância e autocovariância.

Algumas classificações são úteis na definição dos processos estocásticos geradores das séries temporais. Morettin & Tolo (2006) apresentam três delas:

- Processos estacionários ou não estacionários. Existem dois tipos principais de estacionariedade: em sentido estrito, caracterizada pela constância das distribuições de probabilidade associadas a cada uma das variáveis aleatórias pertencentes ao processo estocástico. Assim, a média e a variância que a caracterizam são constantes, logo  $\mu(t) = \mu$  e  $V(t) = V \forall t \in T$ , onde  $V$  é a variância do processo; estacionário de segunda ordem ou fracamente estacionário, qual devem respeitar três condições mostradas na equação 2.3. São elas média constante e variância finita em todo o espaço de estados, assim como autocovariância como função somente da diferença de tempo entre os instantes e não dos instantes em si.

$$\begin{aligned} (i) E\{Z(t)\} &= \mu(t) = \mu, \forall t \in T \\ (ii) E\{Z^2(t)\} &< \infty, \forall t \in T \\ (iii) \gamma(t_1, t_2) &= COV\{Z(t_1), Z(t_2)\} = f(|t_1 - t_2|) \end{aligned} \quad (2.3)$$

- Processos Gaussianos ou não-Gaussianos. Diz respeito à normalidade ou não das f.d.p. associadas às variáveis aleatórias do processo estocástico, assim, diz-se gaussiano um processo estocástico em que as variáveis aleatórias que o compõem sejam também gaussianas.
- Processos Markovianos ou não-Markovianos. Neste a discriminação é feita à independência, em dado instante, dos valores do processo, com relação aos instantes precedentes.

Existe ainda uma outra classe de processos estocásticos bastante importante, chamado não estacionário homogêneo. Montgomery, Jennings & Kulahci (2008) o definem como um processo estocástico não estacionário, porém com primeira diferença ou diferença de ordem mais alta estacionária.

Como este trabalho pretende a modelagem de séries temporais e, além disso, tendo como um dos pilares o modelo ARIMA, faz-se necessário então, supor que os dados utilizados, ou

---

<sup>3</sup> Em um processo estocástico ergódico, quanto mais a população de realizações se aproxima do infinito, mais os momentos amostrais se aproximam dos momentos populacionais.

pelo menos as componentes que serão modeladas estatisticamente, venham de processos ergódicos, gaussianos e homogêneos de primeira ou segunda ordem, ou não estacionário homogêneo. Não só os dados utilizados neste trabalho suprem estes pressupostos como também uma vasta gama de outras séries reais também o faz.

## 2.2 Análise de Espectro Singular

Neste trabalho, a análise de espectro singular será utilizada a fim de separar séries temporais complexas em componentes com estruturas de interdependência mais simples, mais puramente lineares ou puramente não lineares e de mais fácil predição.

Esta técnica é uma poderosa ferramenta para análise de séries temporais. Ela incorpora elementos de estatística multivariada, análise clássica de séries temporais, geometria multivariada, sistemas dinâmicos e processamento de sinais. Através da SSA, é possível decompor uma série original em um pequeno número de componentes interpretáveis, a saber: tendência de baixa frequência de variação, componentes oscilatórios e ruído branco (Alonso, Castillo & Pintado, 2005) (Hassain, 2007) (Golyadina & Zhiglavsky, 2013).

Desta forma, sua utilidade é bastante variada e abrange, entre outros, a extração de tendências complexas e não lineares, detecção de mudanças de regime, remoção de componentes sazonais de amplitude e períodos variáveis, além da suavização das séries. A SSA conta com vantagens para a realização destas tarefas, dentre as quais se destacam a necessidade de poucos dados para a obtenção de resultados satisfatórios e o fato de esta depender de poucos parâmetros, demandando pouco conhecimento prévio sobre as séries (Elsner & Tsonis, 1996) (Golyadina, Nekrutkin & Zhiglavsky, 2001).

Golyadina, Nekrutkin & Zhiglavsky (2001), Alonso, Castillo & Pintado (2005), Marques *et al.* (2006) e Wu & Chau (2011) e descrevem o algoritmo SSA como segue:

O procedimento pode ser dividido em duas etapas: decomposição e reconstrução. Cada uma destas duas etapas, por sua vez, abrange mais duas subetapas. Assim, a decomposição é formada pela incorporação e pela decomposição em valores singulares, enquanto a reconstrução abrange o agrupamento e a média diagonal.

Durante a subetapa incorporação, seja  $\{z_t\}_{t=1}^N$  o conjunto de observações de uma série temporal de tamanho  $N$  unidimensional, a qual deseja-se aplicar a SSA,  $\{z_t\}_{t=1}^N$  é transformado em uma sequência de vetores  $L$ -dimensionais (com  $L$  sendo chamado tamanho da janela), que são usados para construir uma matriz trajetória  $X$ , com formato:

$$X = [\mathbf{x}_1, \dots, \mathbf{x}_k] = \begin{pmatrix} z_1 & z_2 & \dots & z_K \\ z_2 & z_3 & \dots & z_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ z_L & z_{L+1} & \dots & z_N \end{pmatrix} \quad (2.4)$$

Esta matriz, como pode ser visto, tem  $L$  linhas e  $K$  colunas, sendo  $K = N - L + 1$ . Não existe uma forma fechada para a determinação de  $L$ , porém, este deve pertencer ao intervalo  $2 \leq L \leq N - 1$ . Maneiras para encontrar o seu valor ótimo ainda são tópicos de pesquisa, uma sugestão proposta por Khan & Poskitt (2013) é utilizar  $L = (\ln N)^c$ , com  $1.5 \leq c \leq 3$ , e  $L \leq N/2$ .

A decomposição por valores singulares consiste em transformar a matriz trajetória em uma soma de matrizes de rank 1. Sendo  $S = XX^T$ ,  $\lambda_1, \lambda_2, \dots, \lambda_L$  seus respectivos autovalores em ordem decrescente, assim como  $U_1, U_2, \dots, U_L$  os auto vetores associados a  $\lambda_1, \lambda_2, \dots, \lambda_L$ , além de  $V_i = X^T U_i / \sqrt{\lambda_i}$ , a matriz trajetória  $X$  pode ser escrita como  $X = X_1 + X_2 + \dots + X_L$ , onde  $X_i = \lambda_i U_i V_i^T$ . A coleção  $(\lambda_i, U_i, V_i)$  é chamada  $i$ -ésima autotripla da matriz  $X$  (Menezes, Souza & Pessanha, 2015).

A subetapa agrupamento, por sua vez, tem como objetivo a junção das autotriplas, de forma que os grupos remanescentes sejam interpretáveis como tendência, um ou mais componentes oscilatórios com períodos variáveis e ruído aleatório. Assim sendo, nesta subetapa as  $L$  autotriplas são separadas em  $m$  conjuntos disjuntos, e cada um deles gerará uma matriz  $X_I$ . Seja  $I = \{i_1, i_2, \dots, i_p\}$  os índices das  $p$  autotriplas agrupadas em um dos  $m$  subconjuntos desejados, a matriz  $X_I$  se define, pela soma das  $p$  matrizes  $X_i$  com  $i \in I$ . Ou seja  $X_I = X_{i_1} + X_{i_2} + \dots + X_{i_p}$ .

Por fim, a subetapa média diagonal é responsável por transformar cada uma das  $m$  matrizes  $X_I$  em uma subsérie temporal univariada de tamanho  $N$ , chamada “componente SSA”. Esta etapa consiste em calcular a média dos elementos de  $X_I$ , denotados  $x_{I(i,j)}$ , onde  $i + j = k$ , sendo  $k$  uma constante qualquer. Ou seja, sendo  $L^* = \min(L, K)$  e  $K^* = \max(L, K)$ , e sendo  $\{z_{I_t}\}_{t=1}^N$  o conjunto de observações de uma componente SSA, cada elemento de  $z_{I_t}$  será dado por:

$$z_{I_t} \begin{cases} \frac{\sum_{i=1}^t x_{I(i,t-i+1)}}{t}, \forall 1 \leq t \leq L^* \\ \frac{\sum_{j=1}^{L^*} x_{I(i,t-i+1)}}{L^*}, \forall L^* \leq t \leq K^* \\ \frac{\sum_{j=t-K^*}^{N-K^*-1} x_{I(i,t-i+1)}}{N-K^*-1}, \forall K^* \leq t \leq N \end{cases} \quad (2.5)$$

## 2.3 Modelos ARIMA

Modelos ARIMA e suas variações (por exemplo ARIMA sazonal e ARIMA com variável exógena) são, ainda hoje, as metodologias mais usadas para a previsão de séries temporais dentro da família dos métodos estatísticos lineares. Pode-se ver, como exemplo, os trabalhos de Gao, Lo & Lu (2017), Reddy, Jadhav & Gaddi (2017) e Darmawan, Kurniasih & Ahmar (2018).

Neste trabalho, a modelagem ARIMA assumirá um duplo papel. Primeiro, será utilizada para modelar a primeira componente SSA das séries temporais, considerada com o comportamento mais linear e, segundo, será utilizada como *benchmark*, dado a sua aceitação na literatura tanto para o caso sazonal (utilizando sua variante sazonal) quanto para o caso não sazonal.

Um modelo ARIMA é um modelo com parte auto regressiva (AR), integrado (I) e com parte médias móveis (do inglês *moving average*, MA), a metodologia mais utilizada para o uso de tais modelos foi apresentada por Box e Jenkins no ano de 1970.

Segundo Morettin & Toloí (2006), a estratégia para se construir o modelo baseia-se em um ciclo iterativo, sendo que a escolha do modelo é tomada de acordo com a estrutura dos próprios dados. Os passos a serem seguidos no processo são:

- Primeiramente, se considera uma classe geral de modelos para análise, a isso se dá o nome especificação;
- Após isso, segue-se a etapa de identificação, na qual as ordens  $(p, d, q)$  do modelo são escolhidas, segundo alguma estratégia, que pode envolver análise das funções de autocorrelação e autocorrelação parcial, métodos baseados em funções penalizadoras ou outros;
- Seguindo tem-se a etapa de estimação, onde os coeficientes do modelo identificado são estimados;
- Finalmente, verifica-se se o modelo é adequado por meio de uma análise de resíduos, caso o modelo não seja considerado adequado, volta-se à fase de identificação.

Estes modelos somente são aplicáveis a funções estacionárias homogêneas, ou seja, processos estacionários ou que podem tornar-se estacionários por meio de diferenciações (Hamilton, 1994). Assim sendo, se  $Z_T$  for uma série temporal não estacionária na média, mas que pode tornar-se estacionária aplicando-se o operador diferença, neste caso, a série diferenciada  $Y_T$  pode ser representada por um modelo autorregressivo com médias móveis ARMA  $(p, q)$ , sendo a relação entre as observações  $z_t$  e  $y_t$  é tal que:

$$\nabla^d z_t = y_t = (1 - B)^d z_t \quad (2.6)$$

onde  $\nabla$  é o operador diferença,  $B$  é o operador de atraso,  $d$  é um parâmetro inteiro que indica a quantidade de diferenciações necessárias para que se atinja a estacionariedade e  $y_t$  as observações da série  $Y_T$  resultante das diferenciações.

Isto ocorre, pois, os modelos ARIMA estão baseados na teoria geral dos sistemas lineares, a qual supõe que a passagem de um ruído branco por um filtro de memória infinita resulta em um processo estacionário de segunda ordem. (Souza & Camargo, 2004).

Desta forma cada observação  $y_t$  pode ser escrita:

$$y_t = \mu + a_t + \sum_{k=1}^{\infty} \psi_k a_{t-k} \quad (2.7)$$

Adotando-se, por conveniência,  $\tilde{y}_t = y_t - \mu$  e escrevendo os vários coeficientes  $\psi_k$  utilizando um polinômio do operador atraso  $B$ ,  $\psi(B) = (1 + \psi_1 B^1 + \psi_2 B^2 + \dots)$ , tem-se  $\tilde{y}_t = \psi(B)a_t$ , ou seja, um filtro linear com entrada  $a_t$ , saída  $\tilde{y}_t$ , e função transferência  $\psi(B)$ , sendo  $a_t$  um processo estocástico i.i.d. (independente e identicamente distribuído).

Contudo, um filtro de memória infinita possui um forte inconveniente para o caso prático, pois, afinal de contas, precisa da estimação de infinitos parâmetros. Montgomery, Jennings & Kulahci (2008) apresentam 3 alternativas para contornar este problema.

A primeira seria simplesmente considerar que os termos passados do processo estocástico influenciam no valor presente da série temporal  $y_t$  de maneira desprezível, podendo ser desconsiderados. Assim sendo, o filtro, ao invés de ter uma memória infinita passa a ser truncado em um modelo médias móveis de ordem finita e, conseqüentemente, se elimina o problema da estimação de infinitos parâmetros.

A segunda opção, ainda assumindo que o efeito do termo de médias móveis tem efeito reduzido conforme passa o tempo, porém, facilitando a modelagem de efeitos mais “persistentes”, consiste em utilizar termos autorregressivos. De volta à equação 2.7,  $y_t = \mu + a_t + \sum_{k=1}^{\infty} \psi_k a_{t-k}$ , pode-se assumir que os pesos possuem magnitudes descendentes, uma forma simples de se fazer isso é atribuir aos pesos  $\psi_k$  um decaimento exponencial. Por exemplo, utilizando  $\psi_k = \phi^k$  (obviamente, considerando  $|\psi_1| < 1$ ), a Equação 2.7 se torna:

$$y_t = \mu + a_t + \sum_{k=1}^{\infty} \phi^k a_{t-k} \quad (2.8)$$

E, por consequência:

$$y_{t-1} = \mu + a_{t-1} + \sum_{k=2}^{\infty} \phi^k a_{t-k} \quad (2.9)$$

Como  $\sum_{k=1}^{\infty} \phi^k a_{t-k} = \phi y_{t-1} - \phi \mu$ , fazendo  $\delta = \mu - \phi \mu$  e, considerando o caso de se utilizar somente um termo autorregressivo por simplicidade<sup>4</sup>, é possível reduzir a Equação 2.8 a:

$$y_t = \delta + a_t + \phi y_{t-1} \quad (2.10)$$

Um modelo autorregressivo (AR) de ordem finita genérico pode ser representado por:

$$\phi(B)y_t = \delta + a_t \quad (2.11)$$

Sendo  $\phi(B)$  o polinômio de atraso, de maneira que  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ .

A terceira via, por sua vez, é útil quando o padrão de decaimento não pode ser aproximado de maneira satisfatória por modelos autoregressivos (AR) de ordens baixas e tampouco se podem desconsiderar as ordens superiores de um modelo médias móveis (MA) por truncamento. Nestes casos, a solução passa a ser o uso misto dos dois cenários em um, já citado, modelo ARMA, do qual os modelos AR e MA descritos acima são casos especiais. Um modelo ARMA de ordens finitas genérico é escrito da seguinte forma:

$$y_t = \delta + \sum_{k=1}^p \phi_k y_{t-k} + a_t + \sum_{k=1}^q \theta_k a_{t-k} \quad (2.12)$$

Ou, utilizando os polinômios de atraso:

$$\phi(B)y_t = \delta + \theta(B)a_t \quad (2.13)$$

A identificação das ordens de um modelo apropriado para representar uma série temporal específica pode ser feita de várias formas. A mais tradicional delas passa pela análise das funções de autocorrelação e autocorrelação parcial. Contudo, esta abordagem inclui fatores subjetivos, que fazem o resultado variar de analista para analista. Com o desenvolvimento das tecnologias da informação, que aumentaram o poder de cálculo automatizado, as abordagens

---

<sup>4</sup> Raciocínio análogo pode ser seguido para ordens superiores, como pode ser visto em Montgomery, Jennings & Kulahci (2008) ou Morettin & Toloí (2006). Utilizar modelos com ordem superior a 1 significa seguir padrões mais complexos de decaimento de  $\psi_k$  do que o exponencial.

baseadas em critérios de informação<sup>5</sup> vêm se popularizando. Este método consiste em ajustar vários modelos à série em análise e comparar o desempenho destes modelos por meio de funções que premiam de acordo com a verossimilhança do modelo à série e punem pela quantidade de parâmetros, dando preferência a modelos parcimoniosos. Para mais detalhes consultar Morettin & Toloi (2006).

A estimação dos parâmetros, por sua vez, pode ser feita de diversas formas distintas. Podem ser utilizados Mínimos Quadrados Ordinários, Máxima Verossimilhança, etc. Um método muitas vezes utilizado é o de Mínimos Quadrados Ordinários iterativos (Box, Jenkins & Reinsel, 1994).

### 2.3.1 SARIMA

Um modelo SARIMA, ou ARIMA multiplicativo, é uma extensão dos modelos ARIMA para séries que tenham componentes sazonais bastante fortes. Uma vez identificado período da sazonalidade, é preciso seguir um procedimento similar aos descritos anteriormente, porém, sempre levando em conta a parte sazonal. Ou seja, precisa-se testar a necessidade ou não de diferenciações sazonais e, após isso, identificar as ordens da parte autoregressiva e de médias móveis sazonais do modelo, depois disso, estimar os parâmetros sazonais e não sazonais e calcular os *forecasts* (Montgomery, Jennings & Kulahci, 2008).

Um modelo SARIMA genérico pode ser representado por SARIMA  $(p, d, q)x(P, D, Q)s$ , ou, fazendo uso dos polinômios de atraso, por:

$$\phi(B)\Phi(B)(1 - B)^d(1 - B^S)^D z_t = \delta + \theta(B)\Theta(B)a_t \quad (2.14)$$

Nesta equação,  $\phi(B)$  é o polinômio de atraso da parte autoregressiva,  $\Phi(B)$  o polinômio de atraso da parte autoregressiva sazonal,  $(1 - B)^d$  as  $d$  diferenciações necessárias para se atingir a estacionariedade,  $(1 - B^S)^D$  as  $D$  diferenciações sazonais,  $\theta(B)$  o polinômio de atraso da parte médias móveis e  $\Theta(B)$  o polinômio de atraso da parte médias móveis sazonais. Não é comum que os parâmetros  $P, D$  ou  $Q$  sejam maiores que 1.

### 2.3.2 Forecasting a partir dos modelos ARIMA

O *forecasting*  $h$  passos a frente por meio da metodologia Box-Jenkins é feito simplesmente pelo cálculo do valor esperado da variável aleatória  $Z(t + h)$  condicionada à amostra  $(z_t, \dots, z_1)$ , ou seja, a estimação é dada por:

---

<sup>5</sup> Também chamados critérios baseados em função penalizadora.

$$E[Z(t+h)|z_t, \dots, z_1] = \hat{Z}_t(h) \quad (2.15)$$

Tal esperança pode ser calculada mais facilmente por meio das equações diferença do modelo, utilizando processos iterativos para  $h > 1$ .

## 2.4 Redes Neurais Artificiais

Da mesma forma que os modelos ARIMA, neste trabalho, também as Redes Neurais Artificiais terão duplo papel. Como *benchmark* para comparação dos resultados e como parte do método proposto. Isso se justifica pelo fato de as RNAs serem amplamente utilizadas no contexto de previsão de séries temporais. Exemplos podem ser vistos em Chang *et al.* (2014), Li *et al.* (2015) e Kanchimalay *et al.* (2017).

Redes Neurais Artificiais são modelos computacionais não lineares constituídos por unidades de processamento simples, comumente chamados neurônios artificiais ou processadores. Normalmente, os neurônios artificiais são organizados em camadas que são interligadas por uma vasta gama de pesos sinápticos. Seu funcionamento é baseado nos neurônios biológicos e em sua estrutura de processamento altamente paralelizada e tem como fortaleza a capacidade de adquirir, armazenar e utilizar conhecimento experimental, o que torna possível aplicá-la em tarefas como reconhecimento de padrões, clusterização e previsão (Haykin, 2001).

### 2.4.1 Neurônios artificiais

Um neurônio biológico, segundo Haykin (2001), é constituído basicamente de um corpo celular, onde está o núcleo do neurônio, uma grande gama de dendritos, responsáveis por receber os impulsos elétricos e um axônio, cuja função é enviar pulsos elétricos. Tais pulsos passam de um neurônio a outros por meio do que se chama sinapses, que são contatos entre neurônios ativados por pulsos elétricos e reações químicas. O conjunto de neurônios próximos interagindo por meio de sinapses, mediados por meio de neurotransmissores compõem uma

rede neural biológica de transmissão e processamento de informações. Uma esquematização simplificada de um neurônio biológico pode ser vista na Figura 2.1.

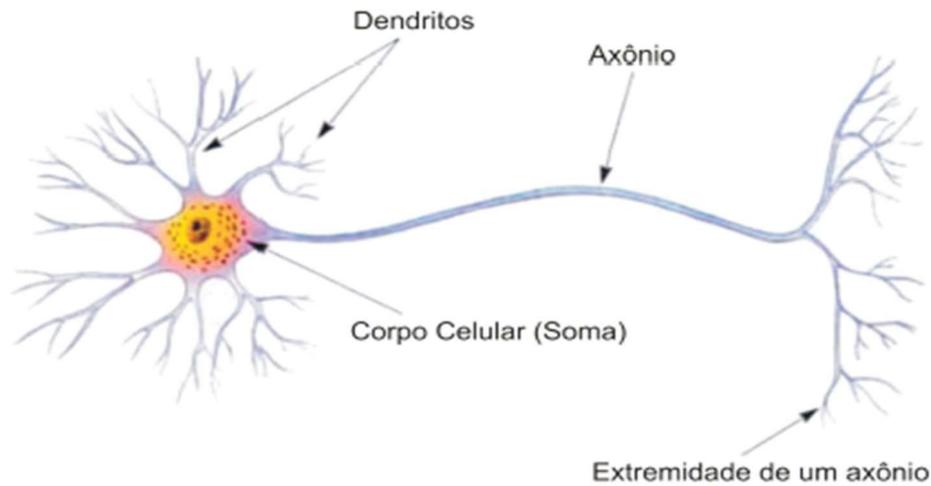


Figura 2.1 - Esquema simplificado de um neurônio biológico. Fonte: Paiva (2015)

Um neurônio artificial, por sua vez, segundo Haykin (2009), é composto por uma série de sinapses ou conexões, um somador e uma função de ativação. As sinapses são constituídas de pesos que ponderam as entradas de um neurônio. Normalmente, o peso que liga a entrada  $j$  ao neurônio  $k$  são denotados por  $w_{kj}$ , podendo, inclusive, assumir valores negativos. O somador, a seu turno, fará a combinação linear das entradas ponderadas pelos pesos sinápticos. Finalmente, a função de ativação será responsável por limitar a amplitude dos valores de saída. Um modelo de neurônio artificial é mostrado a seguir.

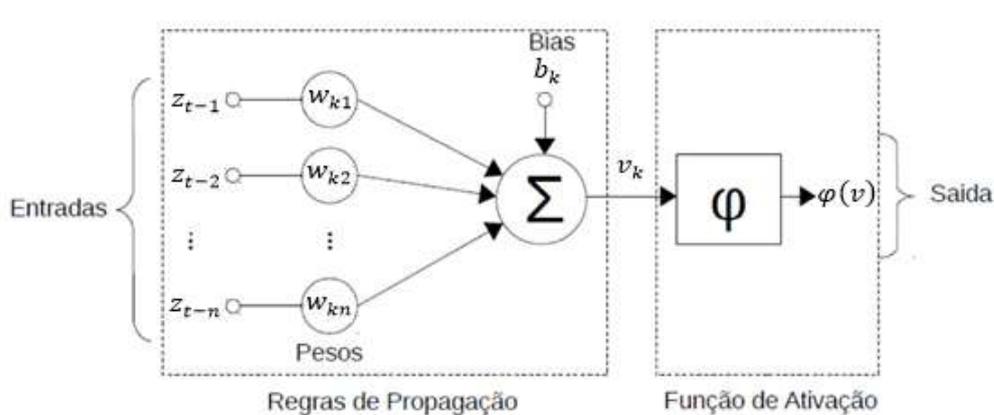


Figura 2.2 - Esquema de um neurônio artificial. Fonte: Adaptado de Haykin (2009).

O exemplo mostrado na Figura 2.2 traz as entradas denotadas por  $z_t$ , pois, neste trabalho os padrões de entrada foram as próprias observações da série temporal. A Figura 2.2 contém ainda um bias ( $b_k$ ), qual tem a função de aumentar ou diminuir arbitrariamente o valor das entradas da função de ativação.

Matematicamente, então, um neurônio pode ser descrito pelo par de equações.

$$u_k = \sum_{j=1}^m z_j w_{kj} \quad (2.16)$$

$$\varphi(v_k) = \varphi(u_k + b_k) \quad (2.17)$$

Nestas equações,  $z_j$  são as  $m$  entradas do  $k$  –ésimo neurônio,  $w_{kj}$  seus pesos respectivos,  $u_k$  a combinação linear das entradas dados os pesos,  $\varphi(\cdot)$  a função de ativação, a soma ( $u_k + b_k$ ) pode ser denotada por  $v_k$  e é chamada campo local induzido e  $\varphi(v_k)$  é a saída do neurônio.

Estrutura similar a esta foi proposta por McCulloch & Pitts (1943). Já nos anos 50 foram utilizadas para resolução de problemas de classificação de padrões linearmente separáveis (Rosenblatt, 1958; 1962).

As funções de ativação de neurônios artificiais são utilizadas para limitar a amplitude das saídas de um neurônio, assim como, em alguns casos, adicionar não linearidade ao modelo (Campos, 2010). Segundo Haykin (2009) elas podem ser:

- Funções degrau, ou função de Heaviside:

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad (2.18)$$

- Função Sinal:

$$\varphi(v) = \begin{cases} 1, & \text{se } v > 0 \\ 0, & \text{se } v = 0 \\ -1, & \text{se } v < 0 \end{cases} \quad (2.19)$$

- Função sigmoide:

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (2.20)$$

- Funções tangente hiperbólica:

$$\varphi(v) = \tanh(v) \quad (2.21)$$

Haykin (2001), cita também a função linear, onde  $\varphi(v) = v$ , que, segundo Zhang, Patuwo & Hu (1998) é comumente empregada nos neurônios da camada de saída de Redes Neurais Artificiais cujo o propósito seja o *forecasting*.

## 2.4.2 Classificações de uma RNA

De acordo com Teixeira Jr. (2013), é possível classificar as RNAs de acordo com: funções de ativação, como apresentado previamente; a arquitetura utilizada, que, dentre outros, pode ser *feedforward* ou recorrente, e; o treinamento adotado, que pode ser supervisionado ou não;

No que diz respeito à arquitetura, a classificação diz respeito a como os neurônios estão ligados. Pode-se dizer que, uma RNA *feedforward* é aquela que não possui neurônios com conexões com outros que estejam na mesma camada ou em camada anterior. As RNAs recorrentes, ao contrário, possuem tais conexões, ou seja, as entradas dos neurônios das camadas iniciais podem ser compostas de padrões de entrada e saídas dos neurônios de camadas mais avançadas.

Em se tratando do treinamento, classifica-se de acordo com o critério utilizado para o aprendizado de uma RNA, ou seja, ajuste de seus pesos sinápticos de acordo com algum algoritmo de otimização plausível. Um treinamento supervisionado é aquele em que são apresentados à RNA os padrões de entrada em conjunto com a saída desejada respectiva a estes padrões, ajustando os pesos sinápticos iterativamente de forma a reduzir uma função do erro entre as saídas obtidas e as saídas desejadas. No treinamento não supervisionado, não existem padrões de saída desejados nem tampouco sinais de erro, então, em geral, neste tipo de treinamento são extraídas propriedades estatísticas do conjunto de entrada, gerando agrupamentos disjuntos de padrões similares.

Para mais detalhes é possível consultar Haykin (2001), Teixeira Jr. (2013) ou Campos (2010)

## 2.4.3 Transformação

Haykin (2001) recomenda que os padrões apresentados à Rede Neural para treinamento sejam transformados para evitar um problema conhecido como paralisia neural. Teixeira Jr. (2013) coloca que, as formas mais comuns em aplicações envolvendo projeção de séries temporais são as que levam os pontos transformados para intervalos fechados  $[0, 1]$  ou  $[-1, 1]$  ou os padronizam, ou seja, transformam para um conjunto com média 0 e desvio padrão 1. Obviamente, para fins de previsão é preciso fazer o procedimento inverso para que os dados retornem à escala original.

Neste trabalho, foi utilizado a normalização *premnmx*, que transforma os dados utilizando como base os valores máximo e mínimo do conjunto. Sendo  $\{z_t\}_{t=1}^N$  o conjunto dos padrões, e definindo  $z_{min} = \text{Min}(\{z_t\}_{t=1}^N)$  e  $z_{max} = \text{Max}(\{z_t\}_{t=1}^N)$  a normalização pode ser escrita como:

$$z_t^* = 2 \times \left[ \frac{z_t - z_{min}}{z_{max} - z_{min}} \right] - 1 \quad (2.22)$$

#### 2.4.4 RNA *Multi-Layer Perceptron Feedforward*

A RNA *multi-layer perceptron feedforward* é a classe mais comum de Redes Neurais aplicada no contexto da previsão de séries temporais (Moreno, Pol & Gracia, 2011). Isso talvez se dê por ela ser um aproximador universal. Ou seja, é possível aproximar qualquer função arbitrária real, desde que contínua e definida em um intervalo fechado (Cybenko, 1989). Esses são os principais motivos de ela ter sido utilizada neste estudo.

A Figura 2.3 abaixo exemplifica a estrutura de uma RNA *multi-layer perceptron feedforward*.

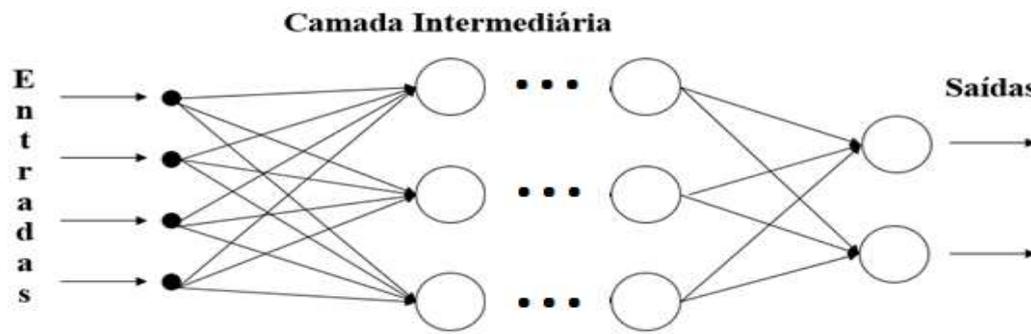


Figura 2.3 - Estrutura básica de uma RNA *multi-layer perceptron*. Fonte: Adaptado de Haykin (2009)

Como pode ser visto na Figura 2.3, estas RNAs possuem uma camada de entrada, uma ou mais camadas de neurônios ocultas (não recebem padrões de entrada diretamente e não são a última camada) e uma camada de saída. São estes neurônios ocultos que permitem extrair progressivamente as características mais significativas dos padrões de entrada. (Haykin, 2009)

Normalmente, no contexto das séries temporais, o treinamento adotado é o supervisionado. O mais comum dos algoritmos de aprendizagem é o *backpropagation* (Campos, 2010), porém, este além de possuir uma grande gama de variações (aprendizado *on-line* ou por lotes, tipo de inicialização dos pesos sinápticos, etc.) (ver Haykin, 2001) é somente um dos possíveis. Por apresentar um desempenho mais satisfatório, neste estudo foi utilizado o *Scaled Conjugate Gradient* (ver Møller, 1993).

Estes dois algoritmos se baseiam em duas fases básicas, a *forward* e a *backward*. Na primeira, calcula-se a resposta da rede com base nos padrões de entrada. Na segunda, o erro entre a resposta real e a esperada é propagado para trás, atualizando os pesos sinápticos.

Além do método de treinamento, alguns outros parâmetros precisam ser definidos para a aplicação da RNA a um problema específico. O mais relevante deles é a arquitetura a ser utilizada, ou seja, quantos neurônios serão utilizados na camada de entrada, na camada escondida e na camada de saída. Zhang (2005) aponta que não há uma maneira fácil de fazer

isso e que, para cada série em específico é necessária a realização de vários testes. Outros autores como Moreno, Pol & Gracia (2011) também pontuam que não há um procedimento bem estabelecido para a aplicação de Redes Neurais para a previsão.

### 2.4.5 Previsão a partir de RNAs

Como as observações de uma série temporal podem ser interpretadas como pontos de uma função, as RNAs *feedforward multi-layer perceptron* (mlp), por serem aproximadores universais, podem ser utilizadas para aproximar este tipo de mapeamento. Assim sendo, havendo uma estrutura de dependência (linear ou não), pode-se interpretar os valores passados da série como sendo os padrões de entrada e o valor da série no instante como o padrão de saída desejado.

Segundo Haykin (2001), Redes Neurais em que a estrutura de memória está na camada de entrada podem ser chamadas TLFN (*Time Lagged Forward Network*). No contexto de séries temporais, este tipo de RNA é considerado um modelo autorregressivo não linear, onde a camada de entrada recebe o nome de janela de tempo e o tamanho da janela é dito ordem do modelo (Campos, 2010).

A figura abaixo exemplifica o uso de uma RNA MLP para previsão de séries temporais.

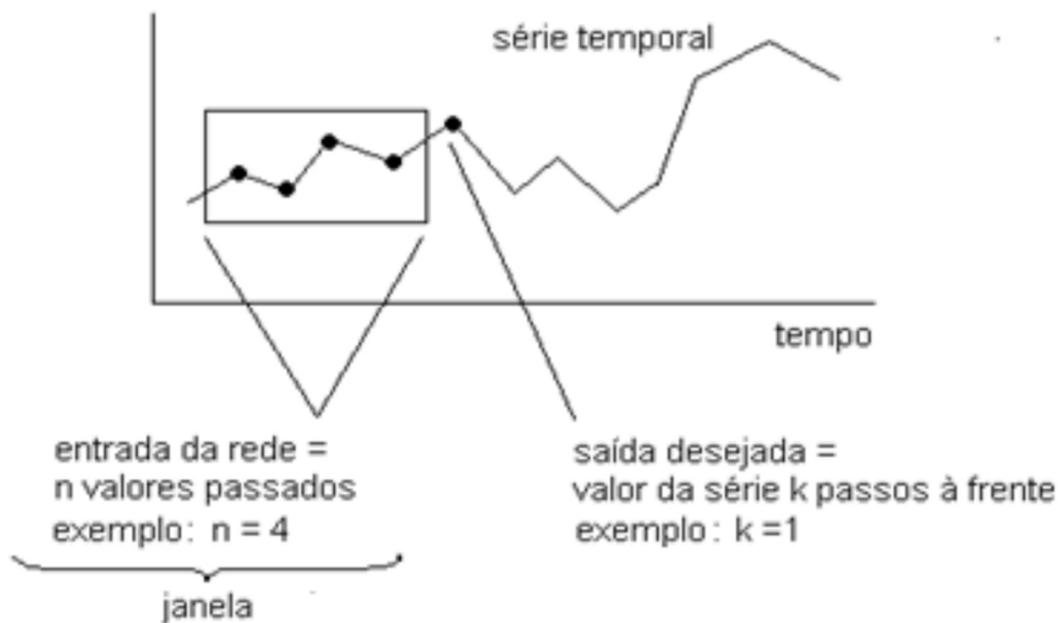


Figura 2.4 - Previsão utilizando RNA MLP. Fonte: Teixeira Jr. (2013)

No exemplo acima, a janela considerada é de 4 dados e o horizonte ( $k$ ) é de apenas 1 dado. Previsões deste tipo são chamadas *one-step*, enquanto que horizontes maiores que um são tratados como *multi-step*, qual são o foco deste trabalho. Neste contexto, caso se queira fazer previsões *multi-step* é preciso executar um processo iterativo onde, após o primeiro *forecast*, é

preciso incorporar o valor estimado aos padrões de entrada e refazer a previsão, repetindo o procedimento até que se atinja o horizonte desejado.

## 2.5 Modelagem Híbrida

Neste trabalho, optou-se pela utilização de uma modelagem híbrida. É preciso, agora, explorar as vantagens de se utilizar tais metodologias ao invés de preditores individuais, justificando então essa decisão. Este será o tópico desta seção, juntamente a uma revisão da literatura recente voltada à previsão, que, em grande medida, está concentrada neste tipo de modelagem.

Convencionalmente, a predição de séries temporais foi feita baseando-se em métodos estatísticos lineares, como, por exemplo, o supracitado ARIMA. Contudo, como previamente colocado, as redes neurais artificiais têm ganhado muita atenção da comunidade científica por conta de características que são as suas (não linear, dirigida pelos dados, sem necessidade de modelo explícito, flexível, etc.). Por consequência, milhares de artigos sobre previsão baseada em RNAs foram publicados nas últimas décadas. Nestes artigos, vários mostram uma performance superior das RNAs em comparação a métodos estatísticos, contudo, em tantos outros, os modelos lineares conseguiram melhores resultados (Yan, 2012).

Pode-se citar como exemplo de estudo que mostra melhor performance das RNAs: Valipour, Banihabib & Behbahani (2013), que comparam Redes Neurais Artificiais a modelos ARMA e ARIMA para predição de ingresso mensal em reservatório de barragem; Bozkurt, Biricik, & Tayşi (2017), que comparou RNAs e modelos SARIMA no contexto da previsão de curvas de carga e; Marín, Orozco & Velilla (2018), estudando o preço da energia elétrica na Colômbia.

Exemplos contrários podem ser encontrados em: Baquero, Santana & Chiaravalloti-Neto (2018), que, buscando a antecipação do número de casos de dengue em São Paulo – SP, obtiveram desempenho superior de GAMs (*Generalized Additive Models*) sobre RNAs e Maleki *et al.* (2018) que buscando o *forecasting* de características da água para tratamento obteve melhores resultados com modelos ARIMA do que com RNAs autorregressivas.

Por sua vez, existem também trabalhos que encontram resultados ambíguos, por exemplo: Co *et al.* (2017), ao estudar a propagação de dengue comparando SARIMA e SVM constatou que a resposta varia cidade a cidade; já Navares *et al.* (2018), comparando modelos ARIMA e RNAs para previsão de entradas em hospital por problemas circulatórios e respiratórios identificou respostas concorrentes entre os dois casos estudados.

Este tipo de conclusões ambíguas mostradas por vários estudos levou vários autores a buscarem, como alternativa, as modelagens híbridas, pretendendo, com isso, melhorar o desempenho preditivo aproveitando a fortaleza de diversos previsores para cada série de tempo. Segundo Teixeira Jr. *et al.* (2016) modelos híbridos, ou combinação de métodos preditivos individuais, pode ser entendido em sentido amplo, tanto para combinação de previsões como para combinação de densidades preditivas bayesianas.

Esta visão não é propriamente nova. Já em 1969 surgiram ideias de se usar combinações lineares de preditores distintos para uma previsão mais eficiente (Bates & Granger, 1969). Com a ascensão da popularidade dos modelos baseados em inteligência artificial não lineares, os modelos híbridos passaram a incorporar também estas técnicas.

Ainda em 2003, Zhang (2003) propôs uma modelagem híbrida, englobando modelos ARIMA e Redes Neurais Artificiais *feedforward*. O que motivou o uso de uma modelagem híbrida, segundo o próprio autor, é o fato de ser bastante difícil identificar se o processo gerador de uma série temporal é linear ou não linear, assim como é difícil encontrar o método preditivo capaz de gerar as melhores previsões fora da amostra. Além disso, as séries de dados reais raramente são puramente lineares ou não lineares, contendo, muito frequentemente, os dois tipos de padrões e, sendo esse o caso, tanto o modelo ARIMA como RNAs não seriam adequados para a previsão, pois o primeiro não conseguiria lidar com as não linearidades, enquanto o segundo mostrou resultados ambíguos, principalmente quando lidando com problemas lineares. Estas ideias são motivadoras também da presente dissertação.

Assim sendo, considerando que uma série temporal pode ser descrita como  $Z_T = L_T + N_T$ , onde  $L_T$  é a componente linear e  $N_T$  é a componente não linear, a metodologia por Zhang (2003) proposta consiste em: primeiro, ajustar um modelo ARIMA ao conjunto de dados, pretendendo assim ajustar a parte linear do problema, e; num segundo procedimento, utilizar uma RNA *feedforward* para captar as estruturas de auto dependência não lineares restantes nos resíduos do modelo linear previamente ajustado.

A modelagem híbrida continua em voga nos dias atuais. De acordo com Rahimi e Khashei (2018) existem mais de 400 mil artigos relacionados às palavras chave “*hybrid model*” na base de dados das publicadoras Springer e Elsevier, sendo que boa parte dos autores converge para a conclusão de que a hibridização de modelos pode melhorar o desempenho dos previsores individuais.

Muitos autores, utilizando-se da ideia proposta por Zhang (2003), propuseram aplicações ou variações de seu método. Por exemplo, Gairaa *et al.* (2016) utiliza uma metodologia bastante similar, aplicando-a para a previsão da radiação solar diária. Han *et al.* (2017) propõem, para a predição de velocidade do vento, duas abordagens: A primeira delas consiste em modelar os dados brutos por técnicas lineares (ARMA) e, em seguida, os resíduos por técnicas não paramétricas; a segunda segue ordem oposta, aplicar técnicas não paramétricas aos dados brutos e, em seguida, modelos lineares aos resíduos. Yan & Ma (2016) sugere, para predição de nível freático, modelar os dados com ARIMA extraíndo a parte linear e, em seguida aplicar aos seus resíduos uma RNA RBF (*Radial Basis Function*) para captar a interdependência não linear. Moeeni, Bonakdarj & Ebtehaj (2017), para predição de influxo mensal de uma bacia hidrográfica, sugere que: primeiramente, a série temporal fosse pré-processada, executando as transformações necessárias; após isso, ajustar um modelo SARIMA, pretendendo identificar os componentes linear e não linear dos dados; em seguida, testar a adequação e a presença de não linearidades nos resíduos do modelo SARIMA; finalmente, a componente não linear é modelada ou por uma RNA ou por uma ANFIS (*Adaptative Neuro-Fuzzy Inference System*). Outro exemplo é o modelo proposto por Kashei & Bijari (2011), qual prevê: em um primeiro

passo, ajustar aos dados um modelo ARIMA; na sequência, treinar uma RNA utilizando, como padrões de entrada os resíduos do modelo linear, as estimativas do modelo linear, assim como os dados originais e; por fim, executar o *forecasting* a partir desta Rede Neural Artificial conforme o procedimento padrão. Chaâbane (2014 a), por sua vez, sugere que, primeiramente, seja ajustado aos dados um modelo ARFIMA (*Auto Regressive Fractionally Integrated Moving Average*), considerando que esta seja suficiente para capturar as estruturas de dependência linear dos dados, após isso, modelar os resíduos deste modelo por meio de um modelo auto regressivo não linear determinado por uma RNA. Já Chaâbane (2014 b), de maneira similar, usa também o modelo ARFIMA para a parte linear, porém usa uma SVM (*Support Vector Machine*) para modelar a parte não linear do problema (que seria os resíduos do modelo ARFIMA). Com uma abordagem um pouco mais diferenciada, Babazadeh (2017) utiliza, para a predição do consumo de gasolina, a previsão de um modelo ARIMA como uma entrada de uma RNA *feedforward*, juntamente aos dados passados.

Modelos deste tipo podem ser chamados de sequenciais, pois, aplicam primeiro um previsor e, subsequentemente, utilizam-se de seus resultados para etapas posteriores (Rahimi & Khashei, 2018).

Outra forma que tem se utilizado para a hibridização de modelos são os modelos paralelos. Ou seja, modelos que fazem a previsão por mais de um método separadamente e, ao final, combinam as estimativas ou densidades preditivas de acordo com algum critério, normalmente linear (Rahimi & Khashei, 2018). Rahimi & Khashei (2018), por exemplo, propõem que sejam ajustados ao conjunto de dados, separadamente, um modelo ARIMA e uma RNA *feedforward*, tomando como predição final uma média ponderada das duas, onde os pesos são ajustados de acordo com a minimização dos erros das previsões dentro da amostra. No caso de Wang *et al.* (2017), no contexto da predição de índices de qualidade do ar, propõem fazer previsões dos dados utilizando a metodologia Box-Jenkins e uma SVM (*Support Vector Machine*), utilizando GARCH (*Generalized Auto Regressive Conditional Heterocedastic*) para estimar os pesos para a combinação linear dos valores previstos. Cavaleri & Ribeiro (2011), para a previsão de séries financeiras utilizaram vários modelos estatísticos assim como várias formas de combinações lineares dos resultados para obtenção da previsão final. Martins (2011), para a modelagem de séries industriais, utilizou a combinação linear de previsões advindas de modelos ARIMA e de RNAs *feedforward*, chegando à conclusão que a combinação pelo método da variância mínima obteve melhores resultados que pela média aritmética. Já Machado *et al.* (2016), utilizou, para realizar a predição de captação de água para uso urbano, a combinação linear das previsões de modelos Holt-Winters e de modelos SARIMA, utilizando, para isso, médias aritméticas, harmônicas e geométricas.

Uma outra forma de modelos híbridos que se tornou popular é a que conta com técnicas de decomposição da série original, para a predição das componentes em separado e a posterior junção dos *forecastings*. Por exemplo, Babu & Reddy (2014) propõem que a série temporal seja decomposta em uma componente de alta volatilidade e uma de baixa volatilidade por meio de médias móveis e, após isso, que a primeira seja modelada utilizando uma RNA *feedforward* e a segunda utilizando um modelo ARIMA. Wang *et al.* (2015), por sua vez, no contexto da predição de séries hidrológicas, indicam a decomposição dos dados originais por EEMD

(*Ensemble Empirical Mode Decomposition*) para que cada uma das sub-séries seja ajustada por um modelo ARIMA distinto e, ao final, as estimativas sejam somadas. Já Du *et al.* (2017), estudando a velocidade dos ventos, propõem a decomposição das séries temporais por CEEMD (*Complete Ensemble Empirical Mode Decomposition*) e, em seguida, ajustar a cada uma das componentes uma RNA Elman, utilizando a meta heurística MOALO (*Multi Objective Ant Lion Optimization*) em seu treinamento para, enfim, agregar cada uma das previsões. Também no contexto da previsão da velocidade dos ventos, Xiao, Qian & Shao (2017) propõem 4 métodos distintos que consistem na combinação das decomposições por SSA ou FEEMD (*Fast Ensemble Empirical Mode Decomposition*) ajustando aos componentes gerados RNAs RBF (*Radial Basis Function*) ou GR (*General Regression*) utilizando a meta heurística *bat algorithm* em seu treinamento, somando as previsões para obter um *forecasting* final. Latifoglu, Kişi & Latifoglu (2015), pensando na modelagem de variáveis hidrológicas, propõem que os dados sejam decompostos por meio de SSA e, depois disso, cada uma das componentes SSA sejam preditas por uma RNA *feedforward* distinta e, ao final, os *forecastings* sejam somados. Para previsão de curvas de carga, Lee & Ko (2011) utilizaram *lifting scheme* para a decomposição das séries temporais, em seguida, modelaram cada uma das sub-séries por ARIMA e agregaram a previsão utilizando o *lifting scheme* inverso. Teixeira Jr., Franco & Souza (2016), para a previsão da série temporal *canadian lynx*, aplicam, preliminarmente, decomposição *Wavelet* aos dados e, em seguida, prevê cada uma das componentes por RNAs *feedforward* distintas, agregando os resultados para obter o *forecasting* final. Wang *et al.* (2018), propõem uma metodologia pouco parcimoniosa para o *forecasting* do preço de petróleo bruto. O autor sugere realizar 4 decomposições distintas nos dados originais (SSA, *Wavelet*, EMD e *Variational Mode Decomposition*) e, em seguida aplicar modelos AR, ARIMA, RNAs e SVRs (*Support Vector Regression*) em cada um dos componentes de cada uma das decomposições, gerando 16 *forecastings* a serem combinados para a obtenção de uma predição final. Teixeira Jr. (2013), por fim, propôs uma modelagem envolvendo a remoção de ruído da série temporal por SSA, após isso, a decomposição *Wavelet* da série filtrada e, em seguida, a modelagem de cada uma das componentes *Wavelet* por uma RNA distinta para, finalmente, executar a combinação linear das previsões por meio de um problema de programação matemática multiobjetivo MINIMAX.

Os modelos híbridos têm sido também utilizados no contexto aqui estudado. Ou seja, para predição do comportamento estrutural de barragens, identificado por meio das medições de sua instrumentação. Corrêa (2015), utiliza o modelo WARIMAX-GARCH para a predição de medidas de pêndulo. Este modelo consiste em realizar a decomposição *Wavelet* da série original para, depois, “completar” cada componente *Wavelet* por previsões vindas de modelos ARIMA-GARCH, após isso, essas componentes são utilizadas como variáveis exógenas de um modelo SARIMAX que gera as previsões finais. Royer (2016), utiliza, para a predição do comportamento de piezômetros, uma metodologia que consiste em combinar linearmente por média aritmética as previsões de um modelo ARIMAX e de uma Combinação Neural SSA-*Wavelet*, qual consiste em filtrar os dados por SSA, realizar a decomposição *Wavelet* da série filtrada, “completar” as componentes *Wavelet* com o uso de 5 RNAs distintas, combinadas não linearmente por outra RNA e, por fim, somar os *forecasts* de cada componente para obter as previsões finais. Wei *et al.* (2017) utiliza uma modelagem híbrida que engloba decomposição *Wavelet*, Modelos ARIMA, RNAs *feedforward* e modelos estatísticos para a previsão do

deslocamento de barragens. Rodrigues *et al.* (2015), para modelar o deslocamento de um bloco da Usina Hidrelétrica de Itaipu, primeiramente ajustam um modelo SARIMA aos dados e, após isso, caso haja estrutura de dependência não linear nos resíduos esses passam por decomposição *Wavelet* e são modelados por distintas SVRs, ao fim as previsões são todas somadas para se chegar ao *forecast* final.

Neste trabalho, optou-se pelo uso de decomposição por SSA, alinhando-se com o terceiro grupo de artigos aqui abordados, por entender-se que a decomposição de uma série pode simplificar a estrutura de interdependência presente nos dados. Isto pode ser interpretado como tornar os componentes mais puramente lineares ou não lineares, alinhando-se com a ideia proposta por Zhang (2003), assim possibilitando o melhor desempenho de cada previsor individual.



# Capítulo 3

## Materiais e métodos

Nesta seção, são apresentados com mais detalhes as séries temporais utilizadas neste estudo, os instrumentos que as geraram, assim como o método utilizado para a previsão destas séries temporais.

Para tal, primeiramente a abordagem é explicada como um todo e, após isso, os pacotes utilizados para implementação são abordados, sempre trazendo os detalhes relevantes, tanto no tocante a implementação em si, quanto ao que tange às ferramentas necessárias.

### 3.1 Método SSA-ARIMA-Neural

Este trabalho tem como foco principal a modelagem e previsão de séries temporais. Visando conseguir previsões mais acuradas, buscou-se utilizar um método híbrido, que consiste em, primeiramente, decompor a série de tempo original utilizando, para isso, o algoritmo SSA, conforme descrito na Seção 2.2 deste trabalho. As componentes resultantes serão modeladas separadamente: A primeira por meio da metodologia Box-Jenkins (abordada no Tópico 2.3); a última, interpretada como ruído, é descartada e; as componentes oscilatórias, utilizando RNAs *feedforward*, também já referidas neste trabalho.

Interpretando-se as componentes SSA como vetores no espaço  $\mathbb{R}^N$ , seja  $\mathbf{z}$  o vetor que contém, em cada posição, um elemento do conjunto de observações  $\{z_t\}_{t=1}^N$  de uma série temporal original  $Z_T$  e  $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$ , os vetores que trazem as observações das  $n$  componentes SSA após a decomposição, desde todas as auto triplas estejam contempladas na reconstrução temos que:

$$\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2 + \dots + \mathbf{z}_n \quad (3.1)$$

Com efeito, para um horizonte preditivo  $h$  e, mais uma vez interpretando-se as componentes e suas respectivas previsões como vetores, agora no espaço  $\mathbb{R}^{N+h}$ , espera-se que, a estimativa dos valores de  $Z(N+h)$  possam ser dadas por

$$E(Z(N+h)) = E(Z_1(N+h) + Z_2(N+h) + \dots + Z_n(N+h)) \quad (3.2)$$

Isso por que, as estimativas  $Z(N+h)$  são a média da distribuição de probabilidade estimada de uma variável aleatória a qual supõe-se fazer parte do processo estocástico gerador da série temporal, assim como as estimativas dos componentes  $Z_1(N+h)$ ,  $Z_2(N+h)$ , ...,  $Z_n(N+h)$ , são médias de distribuições de probabilidades de variáveis aleatórias pertencentes aos processos estocásticos geradores de cada uma das componentes SSA. Logo, como a média da soma de variáveis aleatórias é igual a soma das médias, isso permite que as estimativas sejam feitas separadamente. Para maiores informações sobre soma de variáveis aleatórias ver, por exemplo, Becker (2015).

Logo, isso equivale a dizer, então, que as previsões finais serão dadas pela soma das previsões individuais de cada componente SSA.

Assim sendo, de forma resumida, o método pode ser descrito conforme o fluxograma apresentado na Figura 3.1.

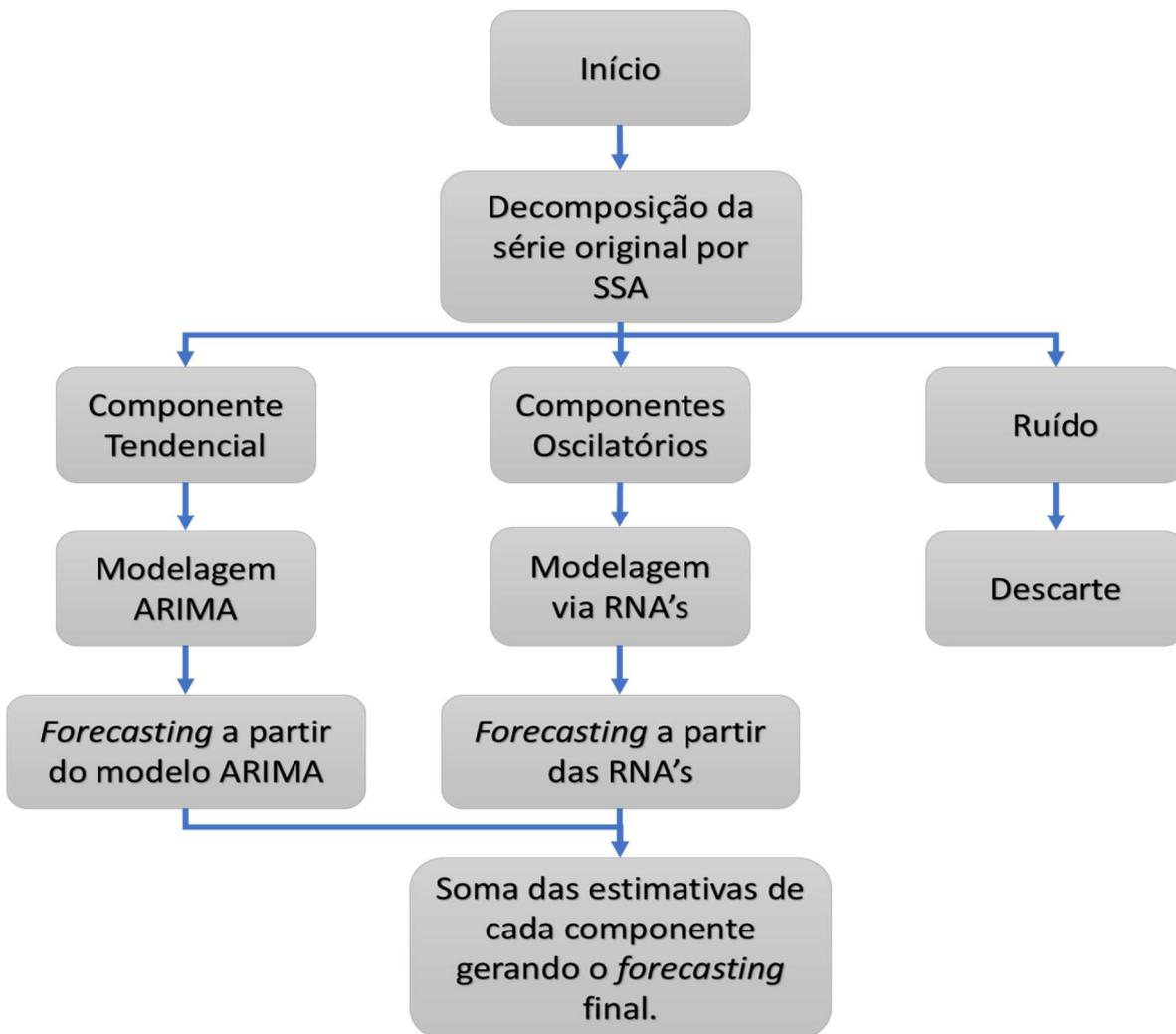


Figura 3.1 - Fluxograma do método SSA-ARIMA-Neural. Fonte: elaboração própria.

Optou-se por decompor as séries temporais em detrimento de outras formas de hibridização por considerar-se que utilizar os ruídos de modelos lineares como parte não linear de uma série temporal pode não ser adequado, já que modelos lineares podem fazer aproximações, ainda que sub ótimas, de não linearidades, fazendo com que reste pouco da interdependência não linear para ser modelada.

A escolha do SSA para a decomposição das séries advém de vários fatores. Pode-se destacar o fato de esta ser uma técnica relativamente nova no campo das séries temporais, de o SSA ter uma complexidade matemática e computacional baixa e depender de pouquíssimos parâmetros e, principalmente, haver uma escassez de trabalhos que façam uso desta técnica para decompor séries temporais visando a previsão.

O uso de um previsor linear e um não linear é proposital, pois, busca-se captar as interdependências destes dois tipos presentes nas séries, baseando-se na ideia proposta por Zhang (2003). A decisão de se utilizar modelagens ARIMA e RNAs do tipo *feedforward* advém do fato de elas serem as mais utilizadas na literatura no campo dos previsores estatísticos e baseados em inteligência artificial, respectivamente, além de possuírem características desejáveis. O ARIMA, por exemplo, é teoricamente capaz de aproximar qualquer função estacionária ou estacionária homogênea. Além disso, possui implementações eficientes e de fonte aberta para serem utilizadas e, o mais importante, apresenta resultados sólidos, por isso sendo amplamente utilizada. RNAs *feedforward* também são aproximadores universais, como garante Cybenko (1989), e graças aos bons resultados que normalmente demonstra para a previsão de séries temporais ganhou muita popularidade, o que fez com que surgissem também boas implementações de uso livre. Além disso, os problemas que se apresentam para o uso de uma RNA *feedforward* são similares aos enfrentados na arquitetura de outras RNAs, por exemplo, quantos neurônios usar, qual função de ativação, etc. Outros previsores lineares e não lineares poderiam ser utilizados no lugar destes, contudo, verificar a adequação do uso de outros previsores em conjunto a decomposição foi uma tarefa deixada para estudos posteriores.

A metodologia aqui apresentada foi utilizada para realizar previsões *multi-step*. O horizonte preditivo utilizado foi de 60 dados e a eficácia das predições fora da amostra foi medida por meio dos MAPEs (*Mean Absolute Percentual Error*)<sup>6</sup> obtidos entre os *forecasts* e os dados observados. Ou seja, os últimos 60 dados disponíveis foram omitidos dos previsores (tanto para o método SSA-ARIMA-Neural como para os métodos utilizados como *benchmark*) e comparados com os *forecasts* obtidos.

Toda a implementação para este estudo foi realizada por meio da linguagem voltada para programação estatística R (Ihaka & Gentleman, 1996), utilizando tanto seu pacote básico quanto alguns outros, a saber: *forecast*, *Rssa*, e *RSNNS*. Assim sendo, convém explorar estas

---

<sup>6</sup> MAPE ou Mean Percentual Absolute Error é uma medida de adequação descrita pela seguinte equação:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{z_t - \hat{z}_t}{z_t} \right|.$$

ferramentas mais detidamente, trazendo também pormenores importantes da implementação utilizada.

## 3.2 Plataforma de Programação R

De acordo com Chambers (2009), o R foi apresentado ao mundo por meio de um artigo de Ihaka & Gentleman (1996), onde fica descrito um pouco das origens e implementações próprios da linguagem.

Ainda segundo Chambers (2009), são 6 as características da linguagem que se destacam: versátil, quer dizer, é uma interface para procedimentos computacionais de vários tipos; interativa, ou seja, pode ser executado em tempo real; funcional na forma de programação; orientado a objetos; modular e; por fim, colaborativo, grátis e livre. Estas características, principalmente as duas últimas, fazem com que a comunidade envolvida com o R tenha crescido sobremaneira, inclusive dentro da comunidade científica.

A escolha desta linguagem de programação se deu principalmente por: ela ser livre; não ter custo para a aquisição de licenças; ter uma grande disponibilidade de pacotes estatísticos com várias funcionalidades bastante úteis já implementadas e; tais pacotes possuem uma documentação bastante vasta e estarem validados pela comunidade científica.

### 3.2.1 Pacote Forecast

Neste trabalho, o pacote forecast foi utilizado para fazer o ajuste dos modelos ARIMA e SARIMA utilizados. Tanto para a *forecasting* da componente SSA tendencial, na metodologia aqui descrita, como para a modelagem das séries a ser utilizada como *benchmark*, a metodologia foi bastante parecida, conforme descrito a seguir.

Deste pacote, a principal função utilizada foi a `auto.arima`. Esta faz uma busca no espaço de modelos possíveis, escolhendo dentre eles o que for considerado mais adequado.

Segundo Hyndman & Khandakar (2008), o maior desafio para o uso de um modelo ARIMA  $(p, d, q)(P, D, Q)_s$  é a escolha dos valores dos parâmetros, pois muitas recomendações levam a fazê-la de forma subjetiva, partindo-se de análises qualitativas das funções de autocorrelação e autocorrelação parcial, por exemplo. Contudo, estes autores apresentam uma metodologia objetiva para a escolha do modelo, qual é utilizada na já citada função `auto.arima`.

Tal metodologia consiste em:

Primeiramente, encontrar os valores adequados para os parâmetros  $d$  e  $D$ , relativos a quantidade necessária de diferenciações, sazonais ou não, para que a série resultante se torne estacionária. Caso a série tenha padrão de comportamento cíclico, no máximo uma diferenciação sazonal é tomada ( $D \leq 1$ ), a necessidade ou não desta diferenciação é dada por um teste de Canova-Hansen (Canova & Hansen, 1995) onde a hipótese nula é a de uma série

com padrão sazonal estável. Após isso, a série diferenciada sazonalmente (caso  $D = 1$ ) ou original (caso  $D = 0$ ), é submetida a testes KPSS (Kwiatkowski *et al.* 1992) sucessivos, ou seja, é feito um teste unitário sobre a série, caso o resultado seja não significativo, uma diferença não sazonal é tomada e o teste é novamente aplicado sobre a série diferenciada. O procedimento é repetido até que o resultado seja o aceite da hipótese de estacionariedade ou até que o valor máximo de diferenciações definido por parâmetro seja atingido, esse valor por default é 2 e assim foi mantido neste trabalho.

Em seguida, usa-se uma heurística para percorrer o espaço dos modelos, já que uma busca por força bruta seria bastante dispendiosa em termos de processamento. Tal heurística consiste em, inicialmente, ajustar quatro possíveis modelos. Caso seja um modelo não sazonal os modelos são: ARIMA(2,  $d$ , 2); ARIMA(0,  $d$ , 0); ARIMA(1,  $d$ , 0) e ARIMA(0,  $d$ , 1). Do contrário, ou seja, para uma série com sazonalidade, o ajuste é feito para: SARIMA(2,  $d$ , 2)(1,  $D$ , 1); SARIMA(0,  $d$ , 0)(0,  $D$ , 0); SARIMA(1,  $d$ , 0)(1,  $D$ , 0) e SARIMA(0,  $d$ , 1)(0,  $D$ , 1); Desses, o modelo que resulte em menor BIC<sup>7</sup> (*Bayesian Information Criterion*) é tomado como modelo corrente.

A partir do modelo corrente são ajustados até treze modelos adicionais. Parte deles variando, alternadamente os parâmetros  $p$ ,  $q$ ,  $P$  e  $Q$  em uma unidade para cima e para baixo do modelo corrente, outra parcela mudando os parâmetros  $p$  e  $q$  em conjunto em +1 ou -1. Alguns mais somando aos parâmetros  $P$  e  $Q$  conjuntamente +1 e -1 e, finalmente, caso o modelo corrente não tenha termo constante, um é atribuído. Dentre estes, o com menor BIC é tomado como novo modelo corrente, após isso, este passo é repetido até que não seja possível encontrar, por este meio, um modelo com menor valor para a função de avaliação que o modelo corrente.

Esta função traz, como default, valores máximos limitantes para os parâmetros, é dizer, os valores tomados como limites no espaço dos modelos é: 2 para  $P$  e  $Q$  e 5 para  $p$  e  $q$ . Estes limites foram mantidos no presente estudo, visando evitar problemas de performance e manter a parcimônia dos modelos. Vale ressaltar que, quaisquer conjuntos de parâmetros que resultem em um modelo não estacionário ou não invertível são rejeitados.

O cálculo das estimativas é feito com o uso da função `forecast`. Esta é uma função genérica, capaz de realizar o cálculo a partir de objetos retornados de várias funções de modelagens pertencentes a este pacote e não vai além de aplicar ao modelo encontrado o conjunto de dados de interesse.

### 3.2.2 Pacote Rssa

---

<sup>7</sup> Por default, a função utiliza o critério AIC (*Akaike Information Criterion*), contudo não há diferença significativa entre o AIC e o BIC, uma vez que  $AIC = -2 \log(L_p) + 2[(p + 1) + 1]$  e  $BIC = -2 \log(L_p) + 2[(p + 1) + 1] \log(n)$ , onde  $L_p$  é a função de máxima verossimilhança,  $p$  é a quantidade de parâmetros do modelo, e  $n$  é o tamanho da amostra. Todavia, o critério BIC tende a trazer modelos mais parcimoniosos, uma vez que a parte penalizadora da função é ligeiramente maior.

O pacote Rssa foi a ferramenta utilizada para proceder a decomposição por SSA. Este pacote foi desenvolvido tendo como base os procedimentos descritos em Golyadina, Nekrutkin & Zhiglavsky (2001) e Golyadina & Zhiglavsky (2013) (conforme mostrado no Tópico 2.2). Este pacote, segundo Golyadina & Korobeynikov (2014) é uma implementação eficiente, possivelmente, a mais rápida disponível. Para maiores detalhes sobre a implementação, no que tange a eficiência computacional em termos de memória e processamento, para a realização do algoritmo, pode ser consultado Korobeynikov (2010).

Como visto na Seção 2.2, o algoritmo para análise de espectro singular tem duas etapas (decomposição e reconstrução), com duas subetapas cada, totalizando quatro procedimentos, a saber: incorporação, decomposição em valores singulares, agrupamento e média diagonal. Destas, o agrupamento precisa de um maior detalhamento, por não ter uma metodologia fechada para execução, Golyadina, Nekrutkin & Zhiglavsky (2001) inclusive tratam a forma de agrupar as auto triplas como um parâmetro do modelo.

Existem várias formas de se fazer tal agrupamento, mais ou menos indicadas a depender da aplicação desejada. Contudo, muitas destas maneiras envolvem análises subjetivas. Estas abordagens foram evitadas por trazerem dificuldades para a automatização dos processos. Assim buscando-se um método que traga componentes separáveis e, como a medida de separabilidade entre duas séries é a falta de *w-correlation* entre elas, se buscou uma metodologia baseada neste critério<sup>8</sup>.

Além do critério de agrupamento, é preciso definir também quantas componentes gerar, ou seja, quantos grupos utilizar. O uso de demasiados componentes não se justifica por dois motivos: primeiro, pelo fato de muitas delas trazerem padrões de comportamento similares e, por isso, não serem devidamente separáveis umas das outras; segundo, é desejável que modelos preditivos sejam parcimoniosos, reduzindo assim o custo computacional e, conseqüentemente, o tempo necessário para o cálculo das previsões. Assim sendo, foram utilizadas decomposições em 3 e 4 componentes, sendo, uma de tendência, uma ou duas componentes harmônicas e uma componente de ruído.

Desta forma, o agrupamento das auto-triplas em cada uma das componentes foi feito pela clusterização hierárquica da matriz de *w-correlations* das componentes. Ou seja, se tomou como medida de distância para a classificação a matriz correlações ponderadas, de forma que auto vetores similares sejam colocados no mesmo grupo.

Outro tema em aberto em termos de SSA é o como definir o parâmetro  $L$ , ou, o tamanho da janela SSA, a ser utilizado, não havendo consenso sobre um valor adequado. Contudo, dado que, como mostra Golyadina, Nekrutkin & Zhiglavsky, (2001) quando existem harmônicos em uma série temporal, se obtém uma melhor separabilidade nos dados usando-se um valor de  $L$

---

<sup>8</sup> Se uma série temporal é a soma de duas outras, temos que elas são fracamente separáveis se todas as linhas da matriz trajetória de uma forem ortogonais a todas as linhas da matriz trajetória da outra. Para uma separabilidade forte, além de elas serem fracamente separáveis, o conjunto dos valores singulares de suas matrizes trajetórias devem ser disjuntos. Estas condições contudo, raramente se apresentam em casos reais, por isso, consideram-se separáveis séries que tenham correlações ponderadas próximas de zero.

múltiplo do período do harmônico, escolheu-se para este parâmetro o valor 24, já que a maior parte dos dados apresentam uma ciclicidade com período igual a 12.

### 3.2.3 Pacote RSNNS

Tanto o ajuste das redes neurais utilizadas como benchmark quanto o treinamento das RNAs utilizadas na metodologia SSA-ARIMA-Neural foram feitas utilizando o pacote RSNNS, mais especificamente, a função `mlp`, qual é específica para redes neurais do tipo *multi-layer perceptron*.

Segundo Bergmeir & Benítez (2012), este pacote é uma interface para usuários R das funcionalidades do SNNS (*Stuttgart Neural Network Simulator*) que, ainda segundo o autor, é uma das mais completas, confiáveis e rápidas implementações dos procedimentos padrões relacionados a RNAs.

Este pacote é bastante abrangente, tanto em termos de arquiteturas disponíveis como das variações de parâmetros para cada arquitetura. Estão disponíveis as implementações para, dentre outras, Redes Neurais Artificiais MLP (Rosenblatt, 1958), Elman (Elman, 1990), RBF (Poggio & Girosi, 1989), ART1 (Carpenter & Grossberg, 1987b), ART2 (Carpenter & Grossberg, 1987a) e ARTMAP (Carpenter & Grossberg, 1991). No que tange a MLP, empregada neste trabalho, é possível utilizar:

- Padrões de entrada de tamanho flexível, o que significa poder utilizar quantos neurônios se queira na camada de entrada;
- Padrões de saída também de diferentes tamanhos, o que no contexto da previsão é equivalente a disponibilizar *forecasts one-step* e *multi-step* com diferentes horizontes;
- Uma ou mais camadas escondidas, com quantidade de nós variável;
- Número de iterações para treinamento variável;
- Diferentes formas de inicialização dos pesos neurais;
- Diferentes métodos de aprendizagem;
- Funções de ativação distintas camada a camada;

Diante de tamanha flexibilidade da ferramenta utilizada, é preciso detalhar a aplicação dela, realizada na metodologia proposta. Primeiramente, vale destacar que foi adotado procedimento semelhante para o ajuste das RNAs utilizadas para o *forecast* da série completa ou para o *forecast* de componentes SSA. Assim sendo, salvo quando especificado o contrário, a forma aqui apresentada foi utilizada nos dois casos.

Em termos de arquitetura, foram utilizadas RNAs *feedforward multi-layer perceptron* com somente uma camada escondida, já que, para a grande maioria dos problemas isso é suficiente (Moreno, Pol & Gracia, 2011). Em termos de função de ativação, na camada escondida foi utilizada a arco tangente hiperbólica, com isso, fornecendo não linearidades ao sistema (Haykin, 2001), já para a camada de saída uma função de ativação linear foi utilizada, como é desejável para previsão (Zhang & Qi, 2005). A inicialização dos pesos neurais foi aleatória, a partir de uma distribuição uniforme, conforme recomenda Haykin (2001). Também

segundo suas orientações os padrões de entrada foram padronizados, utilizando, para isto, a padronização *premnmx*. Já no que concerne a função de aprendizagem, utilizou-se o *Scaled Conjugate Gradient* (Møller, 1993), por ter apresentado, durante testes preliminares, convergência mais rápida e melhor desempenho em previsões dentro e fora da amostra.

Definir a quantidade ótima de neurônios na camada de entrada e na camada intermediária não é uma tarefa fácil e não há uma maneira consolidada de fazê-lo. Neste trabalho, primando-se pelo automatismo e na ausência de alguma implementação que buscasse a arquitetura ótima de uma RNA *multi-layer perceptron* optou-se pelo desenvolvimento de um algoritmo para cumprir tal função. Para tal, o caminho escolhido foi tentar várias arquiteturas. Assim sendo, como para séries sazonais (como é o caso da maioria dos dados testados) é recomendável que algumas observações sazonais estejam incluídas no modelo (Zhang & Qi, 2005) e como uma busca que passe por todos os valores possíveis é inviável computacionalmente, avaliaram-se as arquiteturas com quantidades de neurônios na camada de entrada que cobrissem os ciclos dos dados. Considerou-se, durante a busca, RNAs com quantidade de neurônios na camada de entrada que respeitasse uma das seguintes regras:

$$n = kS \quad (3.3)$$

$$n = kS + 2 \quad (3.4)$$

$$n = kS - 2 \quad (3.5)$$

onde,  $n$  é o número de neurônios,  $S$  a sazonalidade considerada para os dados, no caso,  $S = 12$ , e  $k$  um inteiro entre 1 e 5.

Sobre a quantidade de neurônios na camada oculta, Zhang, Patuwo & Hu (1998) apontam que RNAs com uma quantidade menor desses neurônios possuem, geralmente, melhor capacidade de generalização e menos problemas como *overfitting*. Contudo, um número muito pequeno dos nós na camada escondida pode fazer com que a RNA não tenha força suficiente para compreender os padrões dos dados. Neste trabalho optou-se por utilizar uma aproximação do sugerido por Kang (1991), testando os valores próximos a  $n/2$ , onde  $n$ , mais uma vez, é o número de neurônios na camada de entrada, a saber,  $\frac{n}{2} - 1$ ,  $\frac{n}{2} - 2$ ,  $\frac{n}{2}$ ,  $\frac{n}{2} + 1$ ,  $\frac{n}{2} + 2$  e  $\frac{n}{2} + 3$ .

Tendo em mente estas regras, buscou-se, de maneira exaustiva, dentre as arquiteturas consideradas, a que apresentasse menor erro de previsões dentro da amostra. Para ranquear estes *forecastings*, foi utilizado o MAE (*mean absolute error*) de tais previsões. Desta forma, para o *forecasting* fora da amostra somente a RNA com o desempenho considerado mais satisfatório foi utilizada.

### 3.2.4 Pseudocódigo do Método SSA-ARIMA-Neural Automatizado

Para a melhor compreensão do até aqui exposto, pode ser de grande valia a exposição de um pseudocódigo do método utilizado. Para tal, esse será apresentado abaixo.

#### Algoritmo SSA-ARIMA-Neural

```

1  Function PrevisaoHibrida (Serie, horizonte)
    Input: array com a série a ser modelada e horizonte preditivo.
    Output: array com a previsão realizada para cada instante.

2  For  $nc$  in 3: 4 do
    2      Inicializa matriz de componentes SSA de tamanho  $T$  (tamanho da série) x  $nc$ 
        (número de componentes)
    3      Preenche a matriz de componentes SSA pela decomposição da série original
        em  $c$  componentes utilizando SSA conforme o Tópico 3.2.2.
    4      Inicializa a matriz de forecasts de tamanho 60 (horizonte) x  $(nc - 1)$  (número
        de componentes menos 1)
    5      For  $i$  in 1:  $(nc - 1)$  do
    6          If  $i == 1$  then
                Testa vários ajustes de modelos ARIMA, de acordo com o Tópico 3.2.1, na
                primeira componente
    7              A partir do melhor ajuste, realiza os forecasts
    8              Armazena os forecasts na primeira linha da matriz de forecasts
    9          Else
    10             Testa várias arquiteturas de RNA, de acordo com o Tópico 3.2.3, na  $i$ -ésima
                componente
    11             A partir da arquitetura com melhor desempenho dentro da amostra,
                realiza os forecasts
    12             Armazena os forecasts na  $i$ -ésima coluna da matriz de forecasts
    13          End If else
    14      End For
    15 End For

```

- 16 Inicializa vetor do *forecast* final com 60 elementos (horizonte)
- 17 Armazena no vetor de *forecast* final a soma das colunas da matriz de *forecasts*
- 18 **Return** (*forecast* final)

### 3.3 Dados Utilizados

Os dados utilizados são provenientes do processo de auscultação e inspeção de uma barragem de concreto de gravidade aliviada de grande porte. Auscultação, por sua vez, é o conjunto de processos que têm por objetivo identificar, observar e caracterizar possíveis deteriorações que representem risco a segurança da barragem. A auscultação é feita por meio de inspeções visuais e da instrumentação, onde, as inspeções visuais se preocupam em descrever as eventuais anomalias de maneira qualitativa enquanto que a instrumentação é responsável por adquirir, registrar e processar dados advindos de instrumentos instalados nas estruturas de interesse (Fonseca, 2003).

Os dados disponíveis foram coletados entre outubro de 1982 e agosto de 2012. A coleta foi feita manualmente, ou seja, sem uma comunicação direta entre os instrumentos e um banco de dados. O intervalo de tempo entre as leituras é de 1 mês, de acordo com o que recomenda a ICOLD (*International Commission on Large Dams*), assim sendo, estas séries podem ser classificadas como mensais. Mais detalhes sobre estas recomendações estão disponíveis em Buzzi (2007).

Vale ressaltar que a metodologia aqui apresentada não se limita a este tipo de dados, podendo ser aplicada a qualquer conjunto de dados igualmente espaçados no tempo. Ademais disso, o método SSA-ARIMA-Neural tem como característica a flexibilidade e a escolha destes dados visa refletir isso. Estão presentes instrumentos com: leituras com sazonalidade bastante marcada ou mais difícil de ser percebida; com amplitudes do componente sazonal constantes ou variáveis; sem tendência ou com tendências lineares ou não lineares e; com valores bastante variáveis entre si.

Uma breve explicação sobre a função e algumas características dos instrumentos responsáveis pelas medições é apresentada a seguir.

#### 3.3.1 Pêndulos

Pêndulos são instrumentos utilizados para aferir o deslocamento horizontal da crista da barragem com relação a um ponto da fundação considerado estático (Garcia, 2016). Os pêndulos são categorizados de duas formas: diretos e invertidos. Os pêndulos diretos são fixos na crista da barragem e vão até o contato concreto rocha, já os pêndulos invertidos são fixados na fundação, indo também até a crista (Buzzi, 2007).

As medidas são realizadas em conjunto a um coördinometro ótico, que permite escrever as medições em termos das direções montante - jusante e margem direita – margem esquerda (as vezes lidas como eixo X e eixo Y). Sua medida normalmente é tomada em milímetros, dado que espera-se que a magnitude do deslocamento aferido seja pequena.

É normal que barragens apresentem determinados níveis de deslocamento. Até por isso, são determinados limites já em seu projeto. Contudo, deslocamentos excessivos são indicativos de problemas estruturais que podem levar a rupturas ou acidentes.

### 3.3.2 Piezômetros

Piezômetros são instrumentos convencionalmente utilizados para medir poro-pressões<sup>9</sup> em estruturas geotécnicas. Estes instrumentos podem ter diferentes princípios de funcionamento e diferentes naturezas, medindo a pressão exercida pela água de forma direta, por exemplo, por meio da altura da coluna d'água no interior de um tubo (nos chamados piezômetros Casagrande ou de tubo aberto), ou de forma indireta, medindo, entre outros, a deformação de uma membrana (nos piezômetros de membrana) (Fonseca, 2003).

Mesmo sendo tão distintos entre si e utilizando técnicas diferentes para a realização das medições, o objetivo dos piezômetros é sempre aferir a subpressão atuante nas estruturas monitoradas. Assim sendo, é usual que, a leitura, seja ela qual for, seja convertida para unidades de pressão, normalmente, kPa (kiloPascal) ou m.c.a. (metros de coluna d'água) para que possa ser analisada. Os dados utilizados neste estudo são provenientes de piezômetros Casagrande e foram medidos em metros de coluna d'água.

Caso a subpressão seja demasiadamente elevada, isso pode trazer problemas muito graves às estruturas monitoradas. Por exemplo, a água reduz o coeficiente de atrito entre as superfícies de discontinuidades presentes na fundação, além disso, caso a pressão da água sob a fundação da barragem for suficiente, esta pode provocar deslizamentos da estrutura (CBDB, 1999).

### 3.3.3 Bases de Alongâmetro

Bases de alongâmetro são medidores da distância relativa entre pontos da barragem que podem apresentar deslocamento relativo (Mendonça, 2017). Ou seja, são instrumentos colocados entre os blocos ou entre fissuras das estruturas para medir aberturas ou fechamentos, recalques e deslizamentos. Visto de outra forma, as bases de alongâmetro são instrumentos mecânicos utilizados para medir movimentos de abertura e fecho e deslizamento de discontinuidades estruturais. Observação destas grandezas permite acompanhar a evolução

---

<sup>9</sup> Poro-pressões, pressão de poro e subpressão, são todos sinônimos e são utilizados para referir-se a pressão de baixo para cima exercida pela água nas estruturas de interesse.

dessas discontinuidades contribuindo assim para a garantia da segurança estrutural (Martins & Ribeiro, 2016).

A unidade de medida geralmente utilizada nestes instrumentos é o milímetro, o que é conveniente para captar variações não muito expressivas nas fendas sob monitoramento. Valores que estejam crescendo sobremaneira podem indicar, por exemplo, surgimentos de rachaduras na estrutura monitorada, o que, por sua vez, pode representar problemas consideráveis na segurança estrutural.

# Capítulo 4

## Resultados e Discussões

A fim de verificar a eficiência do método proposto, realizou-se experimento computacional utilizando as séries descritas. Como *benchmark* foram utilizadas metodologias consagradas, a saber: ARIMA, SARIMA, Holt-Winters, RNAs *feedforward* além do método proposto por Zhang (2003). Os ajustes dos modelos ARIMA e SARIMA foram feitos automaticamente conforme metodologia descrita no Seção 2.4, assim como a definição da arquitetura da RNA. Os modelos Holt-Winters foram ajustados com o auxílio da função `HoltWinters`, disponível no pacote `stats`, para o software de programação R, esta função automatiza a busca dos parâmetros  $\alpha$ ,  $\beta$  e  $\gamma$ . Para mais detalhes é possível consultar Hyndman e Khandakar (2008). Para uma explicação da teoria que embasa este método pode-se consultar o Apêndice A. O *forecasting* pelo método proposto por Zhang foi feito ajustando-se um modelo SARIMA, também seguindo a metodologia apresentada na Seção 2.4 e, em seguida, os resíduos deste modelo foram utilizados para treinar uma RNA *feedforward*, com arquitetura definida automaticamente, também seguindo o escrito no Capítulo 3. Todos os *forecastings* realizados consideraram um horizonte de 60 dados para previsão *multi-step*.

Os resultados do experimento computacional serão mostrados neste tópico. Contudo, por conveniência, os gráficos das séries temporais e os parâmetros utilizados em cada modelo para cada série serão deixados nos anexos B e C respectivamente.

O critério utilizado para medida de adequação dos *forecasts* dos distintos métodos foi o MAPE, não só por ser extremamente popular, mas também por dar ao analista uma medida de proporção do erro de previsão em vista dos dados aferidos.

A Tabela 4.1 traz todos os MAPEs obtidos por cada método para a previsão das séries estudadas.

Tabela 4.1 - Ajuste das previsões de acordo com o MAPE obtido.

Instrumento \ Método	RNA	ARIMA	SARIMA	Holt-Winters	Zhang	SSA-ARIMA-NEURAL
Pêndulo Direto	0,0276	0,0726	0,0726	0,0445	0,0746	0,0244
Pêndulo Invertido	0,7782	1,7002	1,0603	0,7973	1,1804	0,7006
Base de alongâmetro 1	0,3051	0,2593	0,2593	0,1525	0,2765	0,1662
Base de alongâmetro 2	0,0771	0,0803	0,073	0,0456	0,0775	0,048
Base de alongâmetro 3	1,1109	0,8194	0,6067	0,6898	1,1485	0,9008
Base de alongâmetro 4	0,1006	0,094	0,0751	0,0615	0,0812	0,0564
Base de alongâmetro 5	0,0576	0,1111	0,1111	0,0846	0,1074	0,0559
Base de alongâmetro 6	0,1411	0,0442	0,0956	0,1243	0,0921	0,0332
Base de alongâmetro 7	0,0344	0,0488	0,041	0,0901	0,0429	0,0306
Base de alongâmetro 8	0,1715	0,1141	0,0836	0,0822	0,0938	0,0762
Base de alongâmetro 9	0,085	0,0481	0,0398	0,1076	0,0398	0,0346
Base de alongâmetro 10	0,0762	0,0173	0,0643	0,0477	0,0643	0,0418
Base de alongâmetro 11	0,1042	0,0935	0,087	0,0415	0,0874	0,0453
Base de alongâmetro 12	0,0537	0,0177	0,0242	0,0133	0,0252	0,0109
Piêzometro 1	0,0427	0,0167	0,016	0,0131	0,0168	0,0105
Piêzometro 2	0,0042	0,0048	0,0035	0,0033	0,0037	0,0039
Piêzometro 3	0,0206	0,0235	0,0314	0,0516	0,0323	0,0172
Piêzometro 4	0,0081	0,003	0,003	0,0035	0,0029	0,0047
Piêzometro 5	0,0014	0,0014	0,0014	0,0014	0,0012	0,0024
Piêzometro 6	0,0031	0,0024	0,0025	0,0036	0,0030	0,0020
Piêzometro 7	0,0177	0,0127	0,0077	0,0136	0,0091	0,0062

Tabela 4.2 – Propriedades estatísticas dos MAPEs

	RNA	ARIMA	SARIMA	Holt-Winters	Zhang	SSA-ARIMA-NEURAL
Média	0,1534	0,1707	0,1314	0,1177	0,1648	0,1082
Desvio Padrão	0,2707	0,3830	0,2448	0,2079	0,3297	0,2296

Na Tabela 4.1, as células contendo os dados foram coloridas conforme a classificação do *forecasting* de cada método, aspirando, com isso, melhorar a visualização. Em vermelho estão os piores MAPEs de cada série, em azul os segundos melhores e, por fim, em verde os melhores desempenhos para cada série temporal.

Pode-se perceber que o método SSA-ARIMA-Neural automatizado aqui proposto performou melhor do que os demais métodos utilizados como *benchmark*. As estimativas geradas por tal metodologia conseguiram os menores MAPEs em 13 das 21 séries testadas, além do mais, obteve o segundo melhor resultado em mais 4 séries, tendo o pior ajuste em apenas 1 dos conjuntos de dados estudados.

A Tabela 4.2 traz algumas propriedades estatísticas dos resultados de MAPE obtido. Estas informações também corroboram para mostrar o melhor desempenho do método SSA-ARIMA-Neural, já que, a média dos MAPEs obtidos por esse método foi de 0.1082, a menor dentre todas as observadas. Além disso, o desvio padrão foi de 0.2296, mostrando a consistência do método.

Mais detalhes sobre as classificações podem ser vistos nas Figuras Figura 4.1 até Figura 4.3. Na Figura 4.1 as RNAs foram omitidas por não terem conseguido o menor MAPE nenhuma vez.

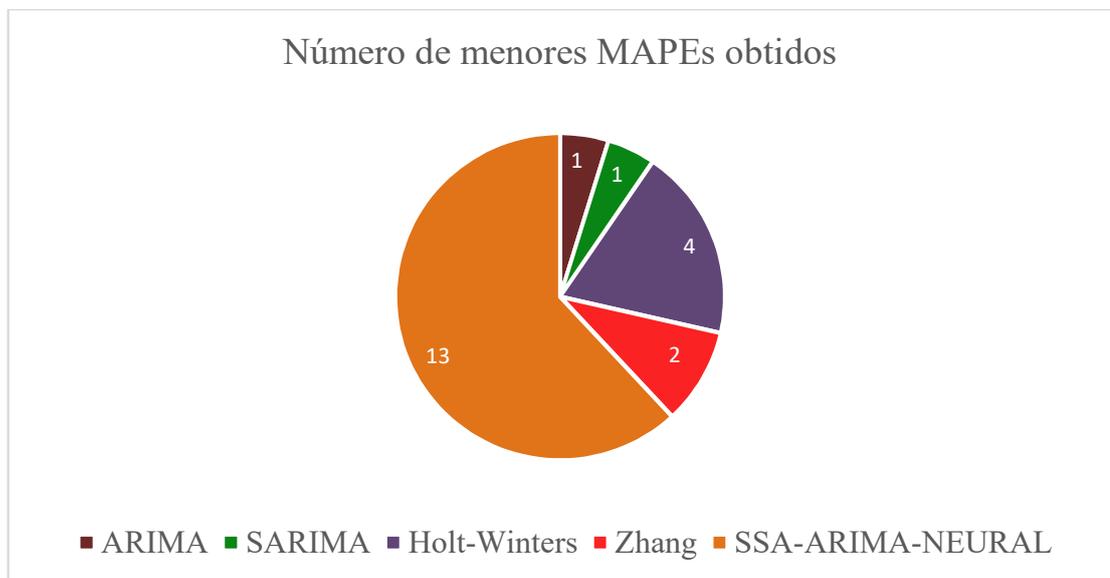


Figura 4.1 - Número de vezes que cada um dos métodos conseguiu o menor MAPE se comparado aos demais. Fonte: elaboração própria.

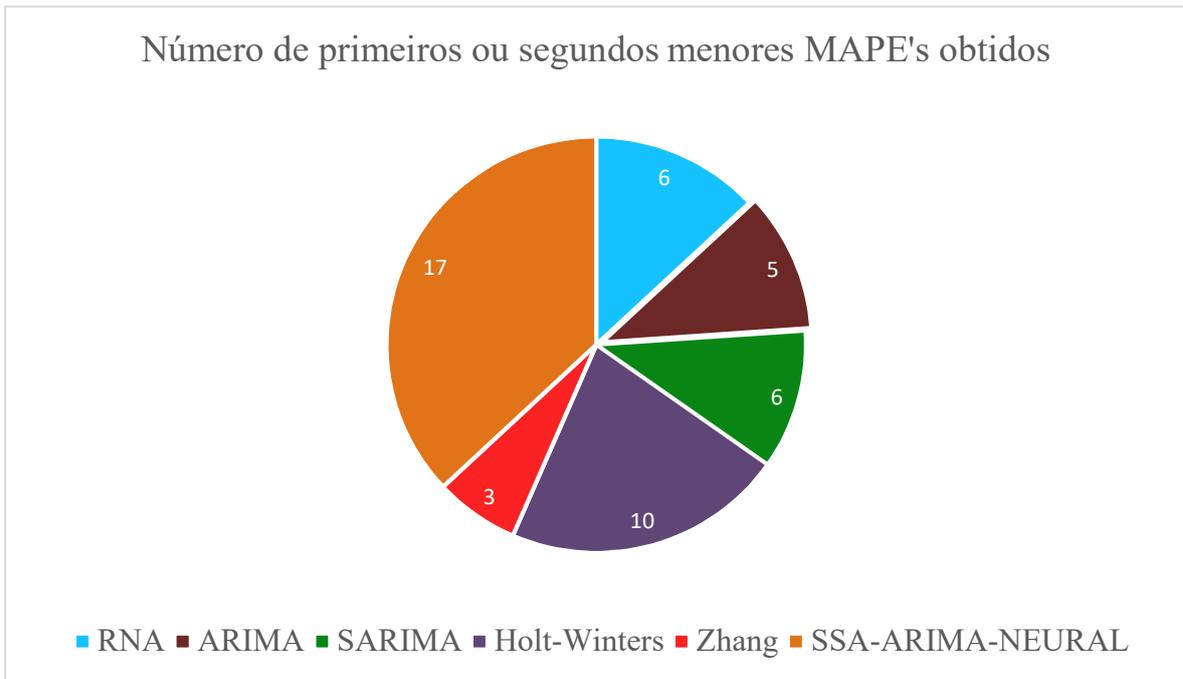


Figura 4.2 - Número de vezes que cada um dos métodos conseguiu o menor ou segundo menor MAPE se comparado aos demais. Fonte: elaboração própria.

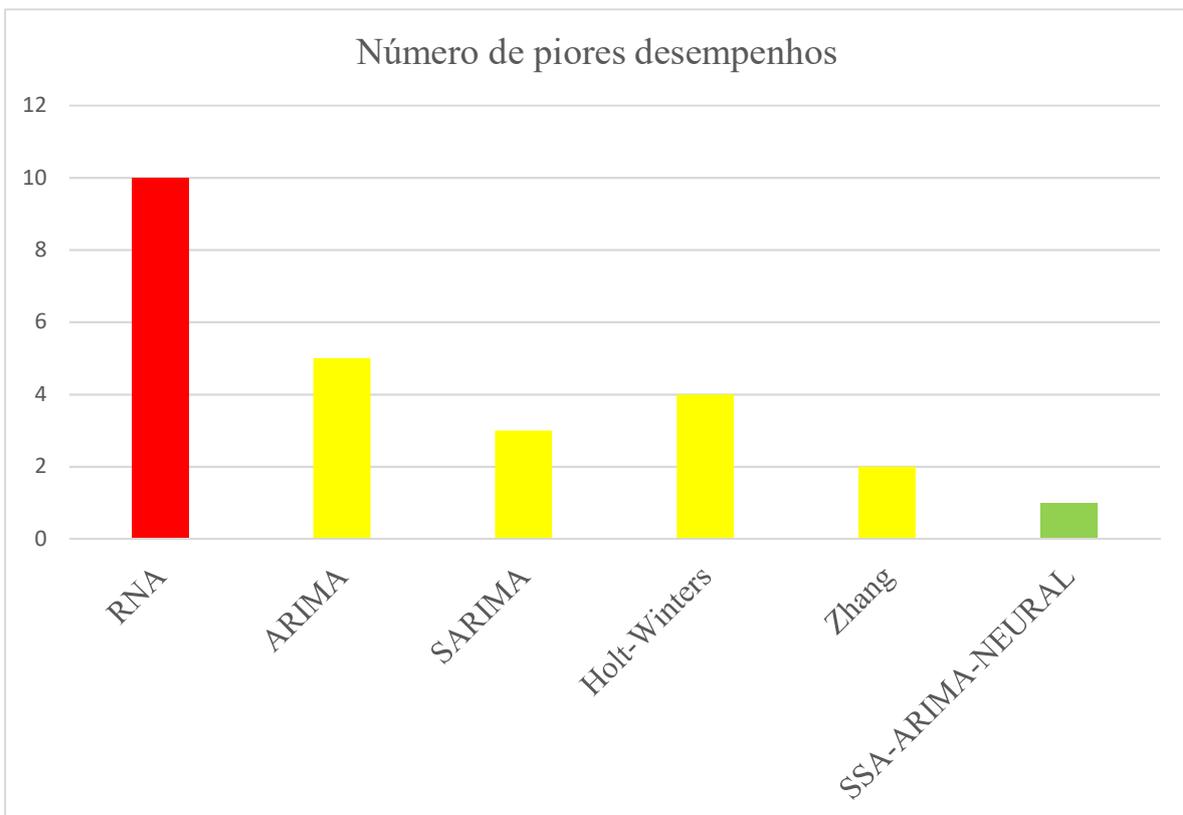


Figura 4.3 – Quantidade de vezes que uma metodologia obteve o pior resultado se comparada as demais. Fonte: elaboração própria.

Observando os resultados apresentados na Figura 4.3 é possível constatar que a maioria dos piores resultados foram obtidos utilizando a RNA isoladamente. Não é trivial identificar o motivo que levou a tal desempenho, contudo, pode-se conjecturar que em grande parte isso se deu pela dificuldade que a RNA *feedforward* apresenta em ultrapassar máximos e mínimos históricos. A modelagem ARIMA, por sua vez, foi a que apresentou o segundo maior número de piores resultados, isso se deve ao fato de esta metodologia não conter ferramentas para lidar com sazonalidades presentes em várias das séries. O SARIMA, por ser mais abrangente que o ARIMA, levando em conta também componentes sazonais, naturalmente teve melhor desempenho. A modelagem Holt-Winters obteve quatro dos piores resultados. Averiguou-se que em todos estes casos as séries históricas apresentaram mudança de tendência bem definida, o que pode compor a causa dos maus resultados observados. O método proposto por Zhang se destaca por raras vezes ter sido responsável por trazer os piores resultados, o que pode advir do fato de ser uma metodologia híbrida, que como tal se vale das fortalezas de diferentes metodologias. Contudo, poucas vezes este predictor permitiu obter os melhores resultados, indicando que talvez ele não traga melhoras tão significativas em termos de acurácia preditiva. Em se tratando da metodologia SSA-ARIMA-Neural automatizada, não só foi única a vez em que o método trouxe a pior estimativa em termos de MAPE, mas também foram 13 as vezes em que o método trouxe as previsões mais acuradas, mostrando que esta hibridização tem potencial para sobrepujar algumas das dificuldades dos métodos considerados em separado, exemplificado pelos resultados obtidos utilizando os dados históricos da base de alongâmetro 8 na Figura 4.4.

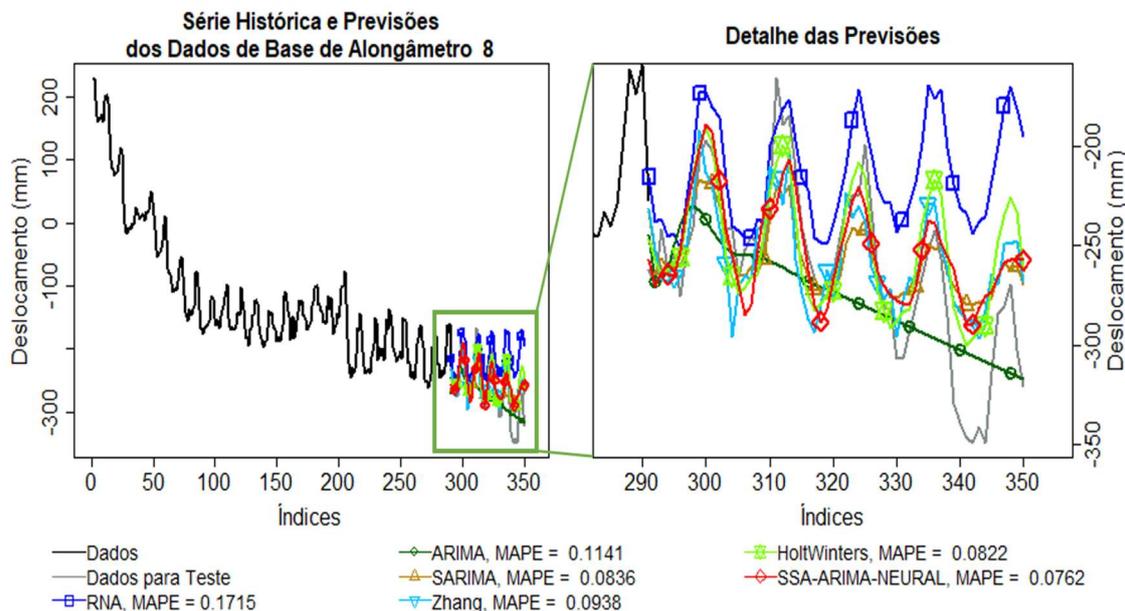


Figura 4.4 – Previsões baseadas na série temporal das leituras da base de alongâmetro 8.

Fonte: elaboração própria.

A Figura 4.5 traz a comparação do método proposto com cada um dos demais previsores utilizados como *benchmarks*, mostrando em números e em porcentagem a quantidade de vezes que o método proposto foi superior ao método usado como comparativo. Por exemplo,

comparando individualmente o método SSA-ARIMA-Neural e as RNAs, vê-se que o método proposto neste trabalho foi melhor em 20 das 21 séries. Pode-se verificar que o SSA-ARIMA-Neural automatizado superou os métodos de referência na grande maioria dos casos. O predictor que em mais séries temporais foi superior ao método híbrido proposto foi o Holt-Winters (7/21), contudo, esse não apresentou regularidade na qualidade das previsões, como ilustra o fato de 4 vezes ter trazido os piores resultados e em algumas vezes sendo bastante inferiores aos demais métodos, como demonstra a Figura 4.6.

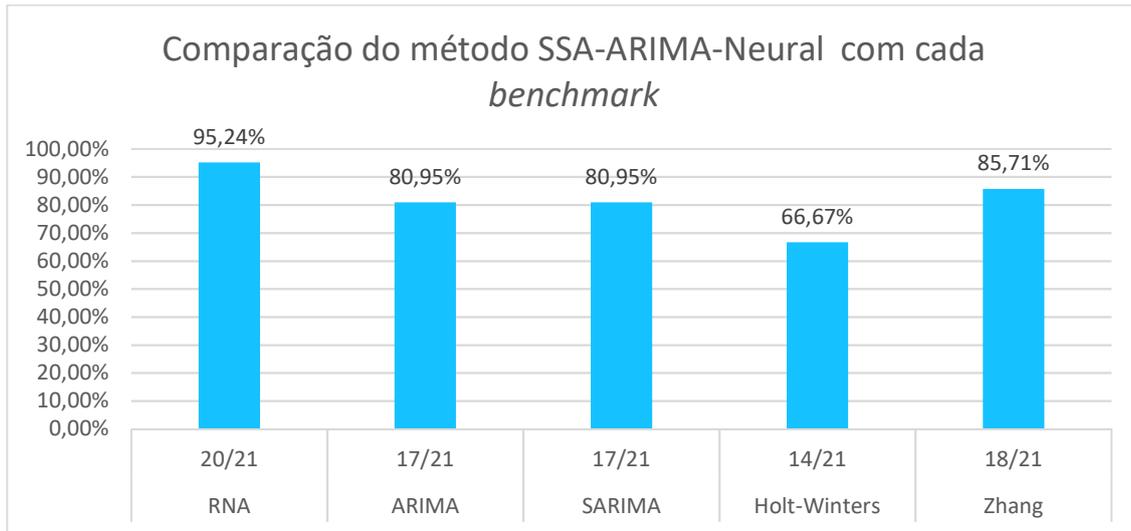


Figura 4.5 – Comparação do desempenho do método SSA-ARIMA-Neural automatizado comparado a cada um dos benchmarks. Fonte: elaboração própria.

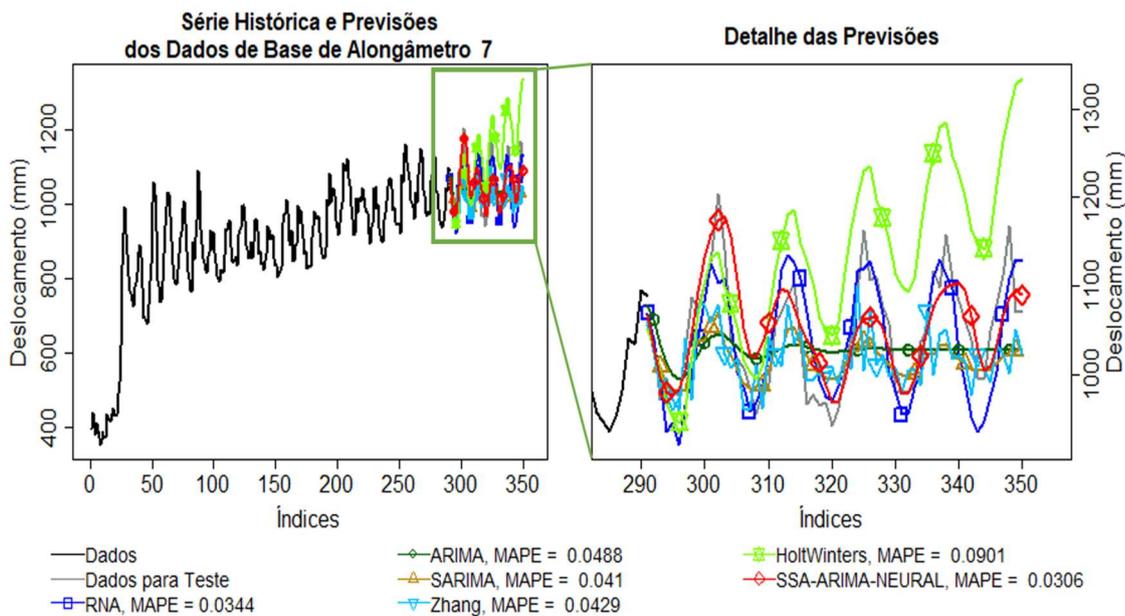


Figura 4.6 – Gráfico dos resultados obtidos com a série histórica da Base de alongômetro 7. Fonte: elaboração própria.

Estes resultados demonstram a versatilidade do método SSA-ARIMA-Neural automatizado, pois, mesmo quando submetido a um conjunto de séries temporais bastante diversas entre si, pode performar de maneira adequada, obtendo resultados efetivos mesmo quando não chegou ao MAPE mínimo entre os previsores testados.



# Capítulo 5

## Conclusão

Neste trabalho foi apresentada uma metodologia híbrida para o *forecasting* de séries temporais. Tal metodologia, baseada na ideia de Zhang (2003), considera que as séries temporais são geradas por processos não puramente lineares ou não lineares e, sendo assim, combina métodos lineares com métodos não lineares.

Na literatura existem várias formas de modelos preditivos híbridos, das quais, podem ser destacados 3. Os modelos sequenciais, que utilizam os ruídos e/ou resultados de uma predição como *input* para outro modelo preditivo; os modelos paralelos, que consistem em modelar as séries de tempo separadamente com previsores distintos, combinando (normalmente linearmente) os *forecasts* ao fim e; finalmente, modelos que se utilizam de decomposições do sinal original, predizendo as subséries em separado e agrupando as séries posteriormente.

Decidiu-se, neste trabalho, utilizar a terceira forma. Isso por que considerou-se que ajustar os ruídos de um modelo linear como parte não linear de uma série temporal poderia ser inadequado, pois os modelos lineares podem fazer aproximações lineares das não linearidades, fazendo restar pouco a ser modelado.

Desta forma, o assim chamado SSA-ARIMA-Neural automatizado consiste em: decompor a série original utilizando *Singular Spectrum Analysis* em 3 ou 4 componentes, uma tendencial, uma ou duas cíclicas e uma considerada ruído aleatório; a componente considerada como tendência, vista como mais puramente linear e com menos volatilidade, é ajustada e prevista por um modelo ARIMA com seus parâmetros escolhidos automaticamente; na sequência, a(s) componente(s) harmônicas são utilizadas para o treino de RNAs *feedforward* com a arquitetura definida por heurística, qual proverá as predições para esta(s) componente(s); a componente de ruído, por sua vez, é descartada; por fim, os *forecastings* de cada componente são somados para se chegar à predição final.

Esse previsor foi posto à prova em experimento computacional envolvendo 21 séries temporais distintas. Tais séries são provenientes de instrumentos de monitoramento de uma barragem de concreto de gravidade aliviada de grande porte. Essas são bastante heterogêneas, possuindo ou não tendências lineares ou não lineares, sazonalidades constantes e de amplitude variável, etc.

Os resultados e tal experimento computacional mostraram a adequação da metodologia proposta. O método SSA-ARIMA-Neural automatizado performou melhor que os métodos utilizados para *benchmark* na maioria dos casos, mostrando que tal metodologia possui

versatilidade, robustez e contribui para o ganho de acurácia preditiva em comparação a métodos consagrados na literatura.

Assim sendo, a principal contribuição deste trabalho é apresentar um método híbrido e automatizado para previsões *multi-step* que consiga tirar proveito das estruturas de interdependência lineares e não lineares presentes nas séries temporais. O fato de ser automatizado dá a este método muitas possibilidades de aplicação prática, por não exigir a interferência do usuário em passos intermediários.

Fica a cargo de trabalhos futuros a experimentação da adequação de outros métodos lineares e não-lineares em substituição ao ARIMA e às RNA MLP e a verificação dos resultados destas mudanças a fim de encontrar um método híbrido ainda melhor.

# Referências Bibliográficas

Alonso, F. J., Castillo, J. M. & Pintado, P. (2005). Application of singular spectrum analysis to the smoothing of raw kinematic signals, *Journal of biomechanics* **38**(5): 1085–1092.

Babazadeh, R. (2017). A Hybrid ARIMA-ANN approach for optimum estimation and forecasting of gasoline consumption, *RAIRO-Oper. Res.* **51**(3): 719–728.

Babu, C. N. & Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data, *Applied Soft Computing* **23**: 27-38.

Baquero, O. S., Santana, L. M. R. & Chiaravalloti-Neto, F. (2018). Dengue forecasting in São Paulo city with generalized additive models, artificial neural network and seasonal autoregressive integrated moving average models, *Plos one* **13**(4): 1-12.

Bates, J. M. & Granger, C. W. J. (1969). The Combination of Forecasts, *Journal of the operational research society* **20**(4): 451 – 468.

Becker, J. L. (ed 1) (2015). *Estatística Básica: Transformando Dados em Informação*, Bookman, Porto Alegre.

Bergmeier, C. & Benítez, J. M. (2012). Neural Networks in R Using the Stuttgart Neural Network Simulator: RSNNS, *Journal of Statistical Software*, **46**(7): 1 – 26.

Box, G. E. P. & Jenkins, G. M. (ed 1) (1970). *Time Series Analysis: Forecasting and control*, Holden-Day, San Francisco.

Box, G.E.P., Jenkins, G.M. & Reinsel, G.C. (ed 3) (1994). *Time Series Analysis: Forecasting and Control*, Prentice Hall, Englewood Cliff, New Jersey.

Bozkurt, Ö. Ö., Biricik, G. & Tayşi, Z. C. (2017). Artificial neural network and SARIMA based models for power load forecasting in Turkish electricity market, *Plos one* **12**(4): 1-24. Acesso em: 15/02/2017. Disponível em: <https://doi.org/10.1371/journal.pone.0175915>.

Buzzi, M. F. (2007). *Avaliação das correlações de séries temporais de leituras de instrumentos de monitoração geotécnico-estrutural e variáveis ambientais em barragens: Estudo de caso Itaipu*, Dissertação de Mestrado, Universidade Federal do Paraná, Curitiba, Paraná, Brasil.

Campos, L. C. D. (2010). *Modelo estocástico periódico baseado em redes neurais*, Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brasil.

Canova, F. & Hansen, B. E. (1995). Are Seasonal Patterns Constant Over Time? A Test for Seasonal Stability, *Journal of Business and Economic Statics* **13**: 237-252.

Carpenter. G. A. & Grossberg. S. (1987a). ART 2: Self Organization of Stable Category Recognition Codes for Analog Input Patterns, *Applied Optics* **26**(23): 4919-4930.

Carpenter. G. A. & Grossberg. S. (1987b). A Massively Parallel Architecture for a Self-Organizing Neural Network, *Computer Vision, Graphics and Image Processing* **37**: 54-115.

Carpenter. G. A. & Grossberg. S. (1991). ARTMAP: Supervised Real-Time Learning and Classification of Nonstationary Data by a Self-Organizing Neural Network, *Neural Networks* **4**(5): 565 - 588.

Cavaleri, R. & Ribeiro, E. P. (2011). Combinação de Previsões de Volatilidade: Um Estudo, *Revista Economia* **12**(2): 239 – 261.

CBDB. (1999). *Guia Básico de Segurança de Barragens – Comitê Brasileiro de Barragens*, XXIII Seminário Nacional de Grandes Barragens, Belo Horizonte, Minas Gerais, Brasil.

Chaâbane, N. (2014 a). A hybrid ARFIMA and network model for electricity price prediction, *Electrical Power and Energy System* **55**: 187 – 194.

Chaâbane, N. (2014 b). A novel auto-regressive fractionally integrated moving average-least-squares support vector machine model for electricity spot prices prediction, *Journal of Applied Statistics* **41**(3): 635 – 651.

Chambers, J. M. (2009). Facets of R, *The R Journal* **1**(1): 5 - 8.

Chang, F.-J., Chen, P.-A., Lu, Y.-R., Huang, E. & Chang, K.-Y. (2014). Real-time-multi-step-ahead water level forecasting by recurrent neural networks for urban flood control, *Journal of Hydrology* **517**: 836 – 846.

Chau, K. W. & Wu, C. L. (2010). A hybrid model coupled with singular spectrum analysis for daily rainfall prediction, *Journal of Hydroinformatics* **12**(4): 458-473.

Co, J., Tan, J. A., Estuar. M. R. J. & Espina. K. (2017). Dengue Spread Modeling in the Absence of Sufficient Epidemiological Parameters: Comparison of SARIMA and SVM Time Series Models, *Proceedings of the RAIS Conference: The Future of Ethics, Education and Research*, Batimore, Maryland, EUA, pp. 232-242. Acesso em: 15/02/2018. Disponível em: <https://ssrn.com/abstract=3086161>.

Corrêa, J. M., Neto, A. C., Teixeira Jr., L. A., Franco, E. M. C. & Faria Jr., A. E. (2015). Time series forecasting with WARIMAX-GARCH method, *Neurocomputing* **216**(5): 805 – 815.

Cybenko, G. (1989). Approximation by superposition of a sigmoidal function, *Mathematics of Control, Signals and Systems* **2**(4): 303-314.

Dantas, T. M., Oliveira, F. L. C. & Repolho, H. M. V. (2017). Air transportation demand forecast through Bagging Holt Winters methods, *Journal Of Air Transport Management* **59**: 116 – 123.

Darmawan, G., Kurniasih, N. & Ahmar, A. S. (2018). Accuracy of Periodogram Analysis for Identification of Multiplicative SARIMA Models, *IOP Conf. Series: Journal of Physics: Conf. Series* **1028**: 1-9.

Du, P., Wang, J., Guo, Z. & Yang, W. (2017). Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting, *Energy Conversion and Management* **15**: 90 – 107.

Elman, J. L. (1990). Finding Structure in Time, *Cognitive Science* **14**: 179 - 211.

Elsner, J. & Tsonis, A. (ed 1) (1996). *Singular Spectrum Analysis. A New Tool in Time Series Analysis*, Plenum Press, Nova York, EUA.

Fonseca, A. R. (2003). *Auscultação por instrumentação de barragens de terra e enronçamento para geração de energia elétrica – Estudo de caso das barragens da UHE São Simão*, Dissertação de Mestrado, Universidade Federal de Ouro Preto, Ouro Preto, Minas Gerais, Brasil.

- Gairaa, K., Khellaf, A., Messlem, Y. & Chellali, F. (2016). Estimation of the daily global solar radiation based on Box-Jenkins and ANN models: A combined approach, *Renewable and Sustainable Energy Reviews* **57**: 238 – 249.
- Gao, G., Lo, K. & Lu, J. (2017). Risk Assessment due to Electricity Price Forecast Uncertainty in UK Electricity Market, *52<sup>nd</sup> International Universities Power Engineering Conference*, Creta, Grécia, pp. 1-6.
- Garcia, S. R. P. (2016). *Modelagem e previsão de deslocamentos em barragens de concreto: aplicação a dados de instrumentação da usina hidrelétrica de Itaipu*. Tese de Doutorado, Universidade Federal do Paraná, Curitiba, Paraná, Brasil.
- Golyadina, N. & Korobeynikov, A. (2014). Basic Singular Spectrum Analysis and Forecasting with R. *Computational Statistics and Data Analysis* **71**: 934-954.
- Golyadina, N. & Zhigljavsky, A. A. (ed 1) (2013). *Singular Spectrum Analysis for Time Series*. Springer, Nova York, EUA.
- Golyadina, N., Nekrutkin V. & Zhigliavsky A. A. (ed 1) (2001). *Analysis of Time Series Structure: SSA and Related Techniques*, CRC Press, Boca Raton, Florida.
- Haikin, S. (ed 2) (2001). *Redes Neurais: Princípios e Prática*. Bookman, Porto Alegre, Rio Grande do Sul.
- Haikin, S. (ed 3) (2009). *Neural Networks and Learning Machines*, Pearson, Hamilton, Canada.
- Hamilton, J. (1994). *TimeSeries Analysis*, Princeton University Press, Princeton, Nova Jersey, EUA.
- Han, Q., Meng, F., Hu, T. & Chu, F. (2017). Non-parametric hybrid models for wind speed forecasting, *Energy Conversion and Management* **148**: 554 - 568.
- Hassain, H. (2007). Singular Spectrum Analysis: Methodology and Comparison, *Journal of Data Science* **5**: 239 – 257.
- Holt, C.C. (1957). *Forecasting trends and seasonals by exponentially weighted averages*, Pittsburgh ONR memorandum no. 52, Carnegie Institute of Technology, Pittsburgh, Pensilvânia, EUA.
- Hyndman, R. J. & Khandakar, Y. (2008). Automatic Time Series Forecasting: The forecast Package for R, *Journal of Statistical Software* **27**(3): 1-22.
- Ihaka, R. & Gentleman, R. (1996). R: A language for data analysis and graphics, *Journal of Computational and Graphical Statistics* **5**(3): 299-314.
- Kanchimalay, K., Salim, N., Sukprasert, A., Krishnan, R. & Hashim, U. R. (2017). Multivariate Time Series Forecasting of Crude Palm Oil Using Machine Learning Techniques, *IOP Conference Series: Materials Science and Engineering*, **226**: 1-9.
- Kang, S. (1991). *An Investigation of the Use of FeedForward Neural Networks for forecasting*. Tese de Doutorado, Kent State University, Ohio, EUA.
- Kashei, M. & Bijari, M. (2011). A Novel hybridization of artificial neural networks and ARIMA models for time series forecasting, *Applied Soft Computing* **11**(2): 2664-2675.
- Khan, M. A. R. & Poskitt, D. S. (2013). A Note on Window Length Selection in Singular Spectrum Analysis, *Australian & New Zealand Journal of Statistic*, **55**(2): 87-108.
- Korobeynikov, A. (2010). Computation and Space Efficient Implementation of SSA, *Statistics and Its Interface*, **3**: 357-368.

Kwiatkowski, D., Phillips, P. C., Schmidt, P. & Shin, Y. (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of Unit Root, *Journal of Econometrics*, **54**(1-3): 159-178.

Latifoglu, L., Kişi, Ö. & Latifoglu, F. (2015). Importance of hybrid models for forecasting hydrological variable, *Neural Computing and Applications*, **26**(7): 1669 – 1680.

Lee, C.-M. & Ko, C.-N. (2011). Short-term load forecasting using lifting scheme and ARIMA models, *Expert Systems With Applications*, **38**(5): 5902 – 5911.

Li, W., Wang, W., Wang, X., Liu, S., Pei, L. & Guo, F. (2014). A dynamic relearning neural network model for time series analysis of online marine data, *Computers & Geoscience*, **73**: 99 – 107.

Machado, R. H. S., Lima, J. D., Batistus, D. R., Adamczuk, G. O. & Trentin, M. G. (2016). Estudo comparativo de modelos clássicos e técnicas de combinação para a previsão do volume de captação de água para uso urbano – um estudo de caso, *Enciclopédia Biosfera*, **13**(23): 1778 – 1793.

Maleki, A., Nasser, S., Aminabad, M. S. & Hadi, M. (2018). Comparisson of ARIMA and NNAR Models for Forecasting Water Treatment Plant's Influent Characteristics, *KSCE Journal of Civil Engineering*. Acesso em 10/05/2018. Disponível em: <https://doi.org/10.1007/s12205-018-1195-z>.

Marín, J. B., Orozco, E. T. & Velilla, E. (2018). Forecasting Electricity Price in Colombia: A Comparison Between Neural Network, ARMA Process and Hybrid Models, *Inernational Journal of Energy Economics and Policy*, **8**(3): 97-106.

Marques, C. A. F., Ferreira, J. A., Rocha, A., Castanheira, J. M., Melo-Gonçalves, P., Vaz, N. & Dias, J. M. (2006). Singular spectrum analysis and forecasting of hydrological time series, *Physics and Chemistry of the Earth*, **31**(18): 1172- 1179.

Martins, V. L. M. (2011). *Comparação de combinação de previsões correlacionadas e não correlacionadas com as suas previsões individuais: um estudo com series industriais*, Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brasil.

Martins, L. F. L. & Ribeiro, A. S. (2016). Caracterização metrológica de um alôngmetro, *Relatório técnico*, Centro de Instrumentação Científica, Lisboa, Portugal.

McCulloch, W. S. & Pitts, W. H. (1943). A logical calculus of the ideas immanent in nervous activity, *Mathematical Biophysics*, **52**(1-2): 115 – 133.

Mendonça, G. G. (2017). *Comportamento hidromecânico de fundações de barragens gravidade*, Dissertação de Mestrado, Universidade Nova de Lisboa, Lisboa, Portugal.

Menezes, L. D., Souza, R. C. & Pessanha, J. F. M. (2015) Electricity consumption forecasting using singular spectrum analysis, *Dyna(Medellín)*, **82**(190): 138-146.

Moenni, H., Bonakdari, H. & Ebtehaj, I. (2017) Integrated SARIMA with Neuro-Fuzzy Systems and Neural Network for Monthly Inflow Prediction, *Water Resources Management*, **31**(7): 2141 – 2156.

Møller, M. F. (1993). A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning, *Neural Networks*, **6**: 525 – 533.

Montgomery, D. C., Jennings, C. L. & Kulahci, M. (ed 2) (2008). *Introduction to Time Series Analysis and Forecasting*, John Wiley & Sons, Nova York.

Montgomery, D. C., Johnson, A. L., & Gardiner, J. S. (ed 2) (1990). *Forecasting & Time Series Analysis*, McGraw-Hill, Nova York.

Moreno, J. J. M., Pol, A. P. & Gracia, P. M. (2011). Artificial neural networks applied to forecasting time series, *Psicothema*, **23**(2): 322 – 329.

Morettin, P. A. & Toloí, C. M. C. (ed 1) (1981). *Modelos para Previsão de Séries Temporais*, IMPA, Rio de Janeiro.

Morettin, P. A. e Toloí, L. M. C. (ed 2) (2006). *Análise Séries Temporais*, Blucher, São Paulo.

Navares, R., Díaz, J., Linares, C. & Aznarte, J. L. (2018). Comparing ARIMA and computational intelligence methods to forecast daily hospital admissions due to circulatory and respiratory causes in Madrid, *Stochastic Environmental Research and Risk Assessment*. Acessado em: 11/04/2018. Disponível em: <https://doi.org/10.1007/s00477-018-1519-z>.

Paiva, A. B. D. (2015) Dia a Dia Educação, *Secretaria de Educação – Governo do Estado do Paraná*. Acessado em: 11/06/2017. Disponível em: <http://www.biologia.seed.pr.gov.br/modules/galeria/detalhe.php?foto=256&evento=3>.

Poggio, T. & Girosi, F. (1989). A Theory of Network Approximation and Learning, *Relatório técnico*, A. I. Memo No. 1140, C.B.I.P. paper No. 31. MIT Artificial Intelligence Laboratory, Cambridge, Massachusetts.

Rahimi, H. Z. & Khashei, M. (2018). A least squares-based parallel hybridization of statistical and intelligent models for time series forecasting, *Computers & Industrial Engineering* **118**: 44-53.

Reddy, B. V. C., Jadhav, V. & Gaddi, G. M. (2017). Application of ARIMA for Forecasting Agricultural Prices, *Journal of Agricultural Science & Technology*, **19**(5): 981 – 992.

Rodrigues, S. B., Silva, A. C. L., Teixeira Jr., L. A., Teixeira, L. L., Siqueira, P. H. & Franco, E. M. C. (2015). Forecasting of the displacements in block III of the Itaipu hydroelectric plant dam through a ARIMA-SVR hybrid approach, *Espaço Energia*, **22**(22): 28–37.

Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, *Psychological Review*, **65**(6): 386–408.

Rosenblatt, F. (1962). *Principles of Neurodynamics*, Spartan Books, Washington, DC.

Royer, J. C. (2016). *Metodologia para definição de alertas dinâmicos para subpressão em fundações de barragens de concreto: o caso de Itaipu*, Tese de Doutorado, Universidade Federal do Paraná, Curitiba, Paraná, Brasil.

Souza, R. C. & Camargo, M. E. (ed 2) (2004). *Análise e Previsão de Séries Temporais: Os Modelos Arima*, Gráfica e Editora Reginal, Rio de Janeiro.

Teixeira Jr., L. A. (2013) *Combinação SSA-Wavelet de Métodos Preditivos com Ajuste Numérico MINIMAX, na Geração de Previsões e de Cenário*, Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, Brasil.

Teixeira Jr., L. A., Faria Jr., A. E., Pereira, R. V. B., Sousa, R. C. & Franco, E. M. C. F. (2016). Combinação Linear Wavelet SARIMA-RNA com Estágios Múltiplos na Previsão de Séries Temporais, *Blucher Marine Engineering Proceedings*, **2**(1): 295 – 397.

- Teixeira Jr., L. A., Franco, E. M. C. & Souza, R. M. (2016). Previsão de curto prazo da série temporal Canadian lynx via redes neurais artificiais e decomposição wavelet, *XVIII Simpósio de Pesquisa Operacional & Logística da Marinha*, **2**(1): 1-12.
- Valipour, M., Banihabib, M. E. & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir, *Journal of Hydrology*, **476**: 433 - 441.
- Wang, J., Li, X., Hong, T., Wang, S. (2018). A semi-heterogeneous approach to combining crude oil price forecasts, *Information Sciences*, **460 – 461**: 279 – 292.
- Wang, P., Zhang, H., Qin, Z. & Zhang, G. (2017). A novel hybrid-Gach model based on ARIMA and SVM for PM<sub>2.5</sub> concentration forecasts, *Atmospheric Pollution Research*, **8**(5): 850-860.
- Wang, W., Chau, K., Xu, D. & Chen, X. (2015). Improving Forecasting Accuracy of Annual Runoff Time Series Using ARIMA on EEMD Decomposition, *Water Resources Management*, **29**(8): 2655 – 2675.
- Wei, B., Yuan, D., Li, H. & Xu, Z. (2017). Combination forecast model for concrete dam displacement considering residual correction, *Structural Health Monitoring*. Acesso em: 15/05/2018. Disponível em: <https://doi.org/10.1177/1475921717748608>.
- Winters, P. R. (1960). Forecasting Sales by Exponentially Weighted Moving Average, *Management Science*, **6**(3): 324-342.
- Wu, C. L., Chau, K. W. & Fan, C. (2010). Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques, *Journal of Hydrology*, **389**(1-2): 146-167.
- Wu, C.L. & Chau, K.W. (2011). Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis, *Journal of Hydrology*, **399**(3-4): 394-409.
- Xiao, L., Qian, F. & Shao, W. (2017). Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm, *Energy Conversion and Management*. **143**: 410 – 430.
- Xiong, T. Li, C. & Bao, Y. (2017). Interval-valued time series forecasting using a novel hybrid Holt and MSVR model, *Economic Modelling*, **60**: 11 – 23.
- Yaffe, R. A. & McGee, M. (ed 1) (2000). *Introduction to Time Series Forecasting with Applications of SAS and SPSS*, Academic Press, San Diego, California.
- Yan, Q. & Ma, C. (2016). Application of integrated ARIMA and RBF network for ground water level forecasting, *Environ. Earth Sci.*, **75**(5): 1 – 13.
- Yan, W. (2012). Toward Automatic Time-Series Forecasting Using Neural Networks, *IEEE Transactions on Neural Network and Learning Systems*, **23**(7): 1028-1039.
- Zhang, G. B., Patuwo, E. & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of art, *International Journal of Forecasting*, **14**: 35-62.
- Zhang, G. P. & Qi, M. (2005). Neural network forecast for seasonal and trend time series, *Computing, Artificial Intelligence and Information Technology*, **160**: 501-514.
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing*, **50**: 159-175.

# Apêndice A

## Modelagem de Séries Temporais utilizando modelos de alisamento

O suposto de que o passado contém informações sobre os padrões de uma série temporal é obrigatório para qualquer método preditivo de séries de tempo, o objetivo principal destes métodos seria então, distinguir tais padrões de ruído que, eventualmente, possa estar presente nas observações e utilizá-los para estimar os valores futuros da série. Os modelos de alisamento assumem que os valores extremos das observações são fruto de aleatoriedade, então, alisando estes extremos seria possível identificar o padrão básico por trás do processo gerador da série temporal. (Morettin & Toloï, 1981).

Esses métodos têm como atrativos sua simplicidade, facilidade para implementação, baixo custo computacional e resultados razoáveis, o que fez com que fossem bastante populares no passado. Hoje em dia suas versões mais simples não são amplamente utilizadas, porém, novos desenvolvimentos e aplicações desta família e desenvolvimentos ainda são frequentes na literatura, como podem exemplificar, por exemplo, Dantas *et al.* (2017) e Xiong *et al.* (2017). Neste trabalho, eles são utilizados como *benchmark*, a fim de validar a acurácia da metodologia SSA-ARIMA-Neural proposta.

### A.1 Médias Móveis

Este é o mais simples dos modelos de alisamento e não é muito utilizado para fins de previsão, contudo, como serve de base para os demais, é conveniente dedicar alguns parágrafos a ele.

Segundo a explicação de Morettin & Toloï (1981), esta técnica pode ser usada em séries temporais compostas por seu nível e ruído aleatório, isto é, para uma série temporal  $Z_t = \mu_t + a_t$ , onde  $\mu_t$  é um parâmetro desconhecido de lenta variação e  $a_t$  tem média zero e variância constante.

Esta técnica consiste em simplesmente calcular a média aritmética das  $r$  observações mais recentes da série, ou seja:

$$m_t = \frac{z_t + z_{t-1} + \dots + z_{t-r+1}}{r} \quad (\text{A.1})$$

Ou ainda

$$m_t = m_{t-1} + \frac{z_t - z_{t-r}}{r} \quad (\text{A.2})$$

Assim sendo, a previsão de um valor futuro é dada simplesmente pela última média móvel calculada, no caso de um horizonte preditivo maior que um (previsão *multistep*) as previsões  $h$  passos à frente incorporam nas médias móveis as previsões anteriores.

O único parâmetro a ser escolhido para esta técnica é a quantidade de observações a serem utilizadas ( $r$ ). Uma forma objetiva de se escolher este parâmetro é testar os vários possíveis tomando como adequado aquele que produzir as melhores previsões dentro da amostra (*backforecasting*).

Ainda sobre o parâmetro  $r$ , existem dois casos especiais merecedores de destaque. Se  $r = T$ , com  $T$  sendo a quantidade de dados observados, a estimativa será igual à média amostral da série. Já se  $r = 1$ , a última observação será tomada como estimativa dos valores futuros. Este segundo caso particular é chamado modelo ingênuo, Naive ou *Random walk*. Sua utilidade esta principalmente atrelada à verificação de adequação de modelos mais complexos (Montgomery, Johnson & Gardiner, 1990).

Os modelos de médias móveis ganharam ainda outros desenvolvimentos que visavam, principalmente, incorporar tendências lineares ou quadráticas, a saber, médias móveis duplas e triplas. Mais detalhes podem ser encontrados em Montgomery, Johnson & Gardiner (1990).

## A.2 Suavização Exponencial De Primeira Ordem

Este modelo busca conseguir um suavizador que reaja mais rápido as mudanças do processo, para isso atribui mais peso às observações recentes do que às antigas. Para isso é utilizado um fator de amortecimento  $\theta$ <sup>10</sup>, que deve ser tal que  $|\theta| < 1$  a fim de que os pesos diminuam de uma maneira exponencial. Além disto, um fator com valor  $(1 - \theta)$  é multiplicado ao modelo para que o somatório dos pesos de cada observação tenda a 1 quando o tamanho da amostra cresça indefinidamente. Assim, de acordo com Montgomery, Jennings & Kulahci (2008) a expressão que caracteriza o estimador por suavização exponencial simples é:

---

<sup>10</sup> O parâmetro  $\theta$  é utilizado em outras metodologias com significados distintos do apresentado nessa seção, contudo, por se tratarem de notações consagradas, optou-se por mantê-las conforme se encontram na literatura, exigindo certo cuidado do leitor.

$$\tilde{z}_t = (1 - \theta)(z_T + \theta z_{T-1} + \theta^2 z_{T-2} + \dots + \theta^{T-1} z_1) \quad (A.3)$$

Que também pode ser escrita em sua forma recursiva:

$$\tilde{z}_t = (1 - \theta)z_T + \theta \tilde{z}_t \quad (A.4)$$

Ou ainda, substituindo-se  $(1 - \theta)$  por  $\lambda$  obtêm-se outra forma bastante usual:

$$\tilde{z}_t = \lambda z_T + (1 - \lambda) \tilde{z}_t \quad (A.5)$$

Percebe-se assim, que o amortecimento exponencial de primeira ordem de uma série temporal nada mais é que a combinação linear da observação no tempo  $t$  com o amortecimento exponencial obtido no tempo  $t - 1$ .

O parâmetro  $\lambda$  tem papel fundamental nesta modelagem. Quanto menor seu valor, mais peso se dará as observações antigas e vice-versa, quanto maior seu valor mais peso será dado às observações presentes, o que implica que, quanto menor o valor de  $\lambda$  mais estáveis serão as estimativas, já que aleatoriedades ocorridas no presente terão sua capacidade de influência no modelo reduzida, contudo, analogamente ao parâmetro  $r$  das médias móveis simples, isso faz com que o modelo reaja mais lentamente a alterações. Uma forma objetiva de obter tal parâmetro é estimar o valor que minimize os resíduos para o *backforecasting* (Morettin & Toloï, 2006).

As previsões neste modelo são calculadas pelo amortecimento exponencial obtido na observação anterior ao instante a ser previsto. No caso de uma previsão  $h$  instantes a frente, utiliza-se, para o cálculo do amortecimento exponencial, ao invés do valor observado no instante  $(h - 1)$ , sua estimativa.

Tal qual ao modelo de médias móveis simples, a metodologia de amortecimento exponencial também possui desenvolvimentos que permitem melhor adequação a séries com tendências lineares, a saber, o amortecimento exponencial de segunda ordem, e também para séries com tendências quadráticas ou superiores, por meio de amortecimentos exponenciais de maiores ordens, contudo, nestes casos, modelos ARIMA ou transformações nos dados devem ser considerados (Montgomery, Jennings & Kulahci, 2008).

### A.3 Suavização Exponencial de Holt e Winters

---

<sup>11</sup> A notação  $\tilde{Z}_t$  diz respeito ao amortecimento exponencial calculado no instante  $t$ , e não como a estimativa de  $Z$  para o instante  $t$ , assim sendo,  $\tilde{Z}_t$  é, na verdade, um estimador para  $Z_{t+1}$ . Apesar de isto poder gerar alguma confusão, optou-se por manter a notação desta forma por ser a mais encontrada na literatura.

Este método foi apresentado por C. C. Holt (Holt, 1957) e seu estudante Peter Winters (Winters, 1960) e trata-se de um desenvolvimento do método de suavização exponencial anteriormente apresentado que visa alcançar melhores resultados em séries de dados que apresentam comportamento cíclico ou sazonal.

Basicamente, adiciona-se uma componente aditiva ou multiplicativa à componente de tendência de um amortecimento exponencial de segunda ordem.

No caso aditivo, supõe-se que o processo pode ser explicado por uma equação do tipo  $z_t = L_t + S_t + e_t$ , onde  $L_t$  é a tendência linear do modelo no instante  $t$ ,  $S_t$  é a componente sazonal, de modo que  $S_t = S_{(t+s)} = S_{(t+2s)} \dots \forall t = 1, 2, \dots, s-1$ , sendo o subscrito  $s$  a quantidade de dados em um período do ciclo e,  $e_t$  é o termo do erro aleatório, com média 0 e variância constante. Assim sendo, segundo Montgomery, Jennings & Kulahci (2008) para se ajustar um modelo Holt-Winters aditivo é preciso seguir os seguintes passos:

Primeiro, atualizar as estimativas de  $L_T$ , por meio da equação:

$$\hat{L}_T = \alpha(z_T - \hat{S}_{T-s}) + (1 - \alpha)(L_{T-1} + \hat{\beta}_{1,T-1}) \quad (\text{A.6})$$

Onde  $\alpha$  é o primeiro parâmetro de amortecimento e  $\hat{\beta}_{1,T-1}$  é a estimativa do coeficiente angular da tendência linear no instante  $T - 1$ , assim, nesta equação, a primeira parte pode ser vista como o “valor atual” de  $L$  e a segunda parte como a sua predição para o instante futuro.

Segundo, atualizar a estimativa de  $\beta_1$ , para isso utiliza-se:

$$\hat{\beta}_{1,T} = \beta(\hat{L}_T - \hat{L}_{T-1}) + (1 - \beta)\hat{\beta}_{1,T-1} \quad (\text{A.7})$$

Com  $\beta$  sendo o segundo parâmetro de amortecimento.

Em seguida, faz-se a estimativa da parte cíclica do modelo, de acordo com:

$$\hat{S}_T = \gamma(z_T - \hat{L}_T) + (1 - \gamma)\hat{S}_{T-s} \quad (\text{A.8})$$

Sendo que  $\gamma$  é a terceira e última constante de amortecimento.

Finalmente, O *forecasting*  $\tau$  instantes a frente é dado simplesmente por  $\hat{z}_{T+\tau}(T) = \hat{L}_t + \hat{\beta}_{1,T}\tau + \hat{S}_t(\tau - s)$ .

Para o caso multiplicativo, parte-se da premissa de que a série temporal pode ser explicada por uma equação do tipo  $z_t = L_t S_t + e_t$ , mais uma vez com  $L_t$  representando o componente da tendência linear,  $S_t$  a componente sazonal, também respeitando a igualdade descrita acima dos efeitos com relação à posição no ciclo e  $e_t$  como parte aleatória do modelo.

De modo similar ao ocorrido em relação ao caso aditivo, Montgomery, Jennings & Kulahci (2008) trazem também um passo a passo para esta situação, que, envolvem os mesmos passos anteriormente descritos, porém, as atualizações neste caso respeitam as seguintes equações:

$$\hat{L}_T = \alpha \frac{z_T}{\hat{S}_{T-s}} + (1 - \alpha)(L_{T-1} + \hat{\beta}_{1,T-1}) \quad (\text{A.9})$$

$$\hat{\beta}_{1,T} = \beta(\hat{L}_T - \hat{L}_{T-1}) + (1 - \beta)\hat{\beta}_{1,T-1} \quad (\text{A.10})$$

$$\hat{S}_T = \gamma \frac{z_T}{\hat{L}_T} + (1 - \gamma)\hat{S}_{T-s} \quad (\text{A.11})$$

O *forecasting*  $\tau$  instantes a frente, por sua vez, é feito através da equação  $\hat{Z}_{T+\tau}(T) = (\hat{L}_t + \hat{\beta}_{1,T}\tau) \hat{S}_t(\tau - s)$ .

Para ambos os casos, os valores iniciais podem ser obtidos por meio de regressões dos dados amostrais, para mais detalhes ver Morretin & Toloï (2006) e Montgomery, Jennings & Kulahci (2008).



# Apêndice B

## Gráficos das Séries Temporais e dos Forecasts

Seguem os gráficos das séries históricas e dos *forecasts* das séries testadas.

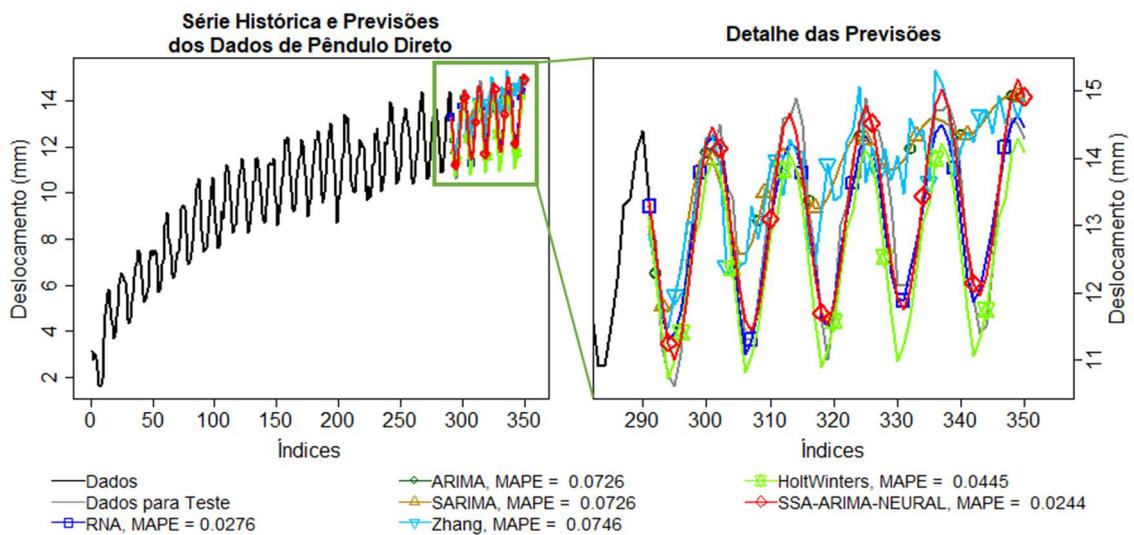


Figura B.1 – Dados do Instrumento Pêndulo Direto. Fonte: elaboração própria.

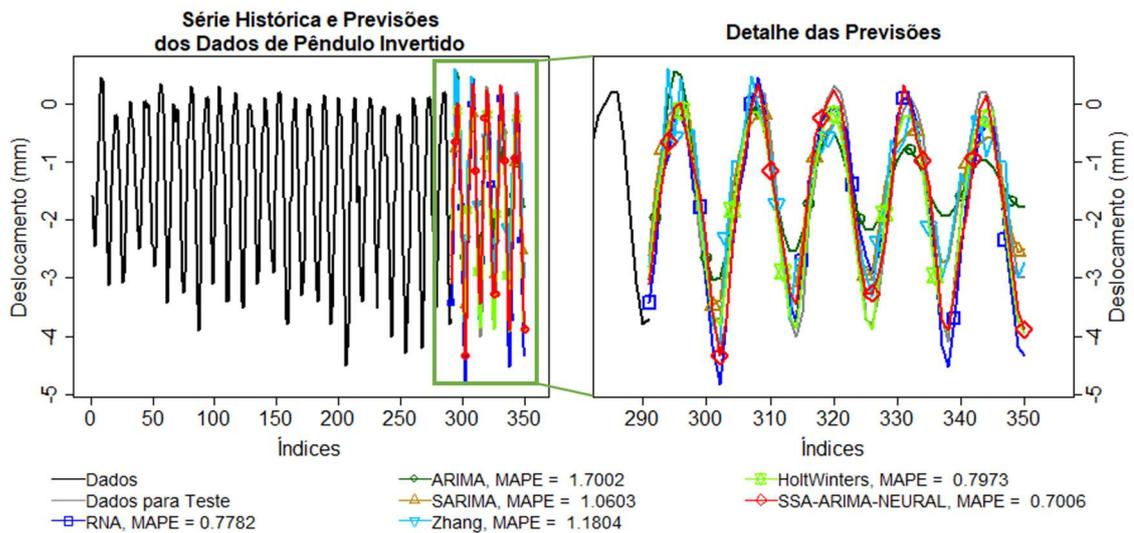


Figura B.2 – Dados do Instrumento Pêndulo Invertido. Fonte: elaboração própria.

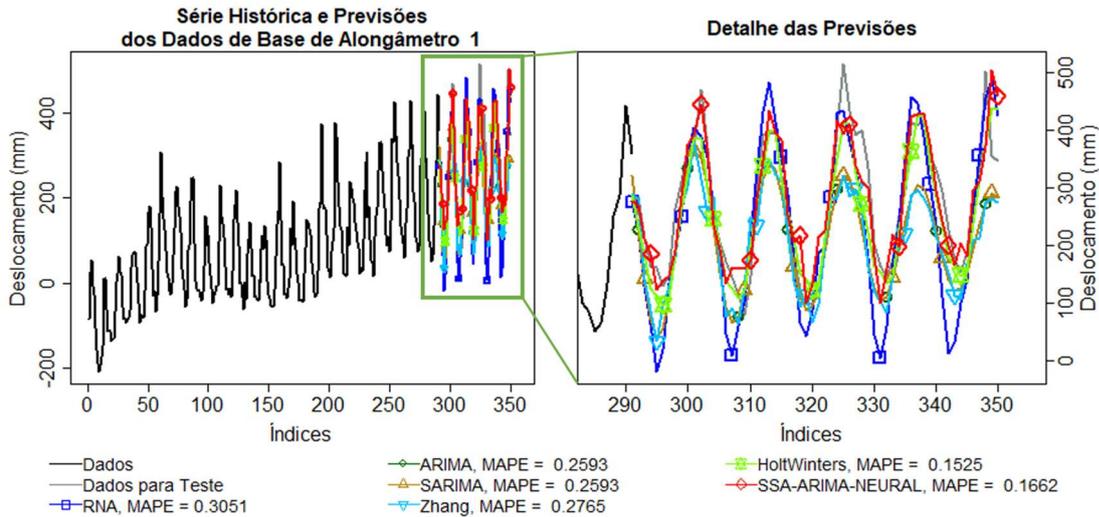


Figura B.3 – Dados do Instrumento Base de Alongâmetro 1. Fonte: elaboração própria.

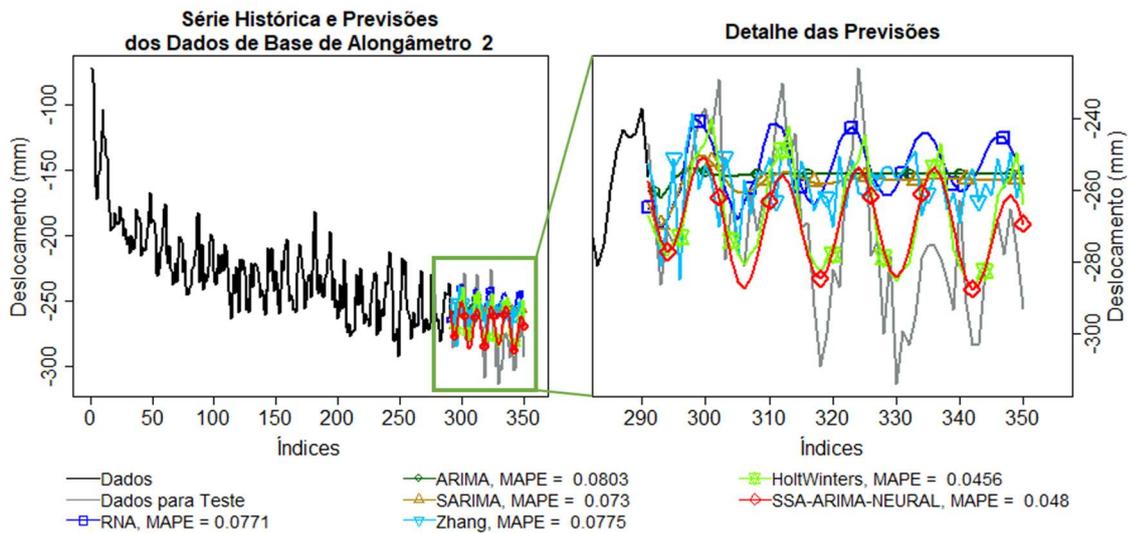


Figura B.4 – Dados do Instrumento Base de Alongâmetro 2. Fonte: elaboração própria.

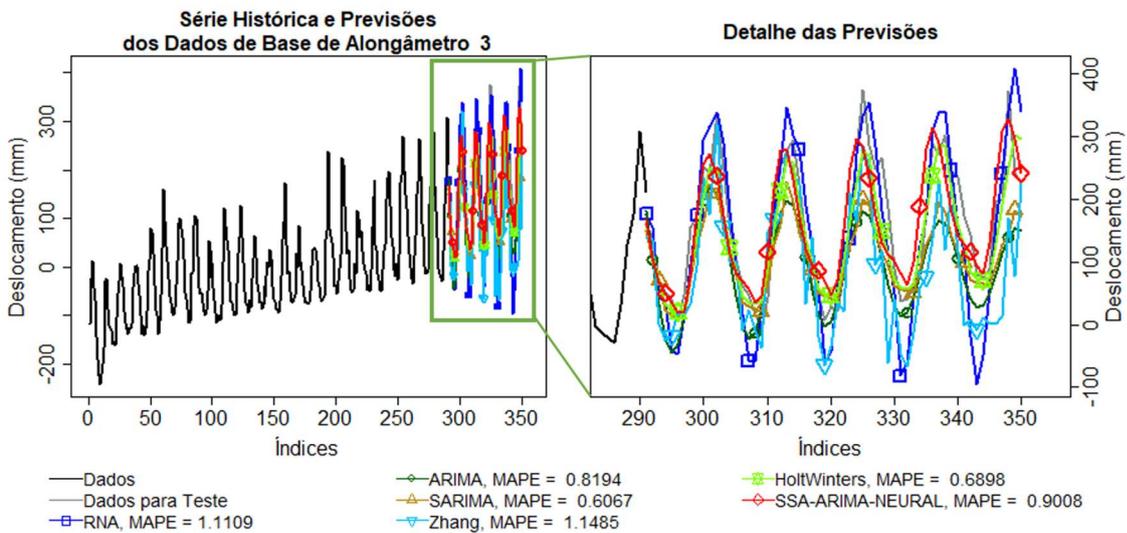


Figura B.5 – Dados do Instrumento Base de Alongâmetro 3. Fonte: elaboração própria.

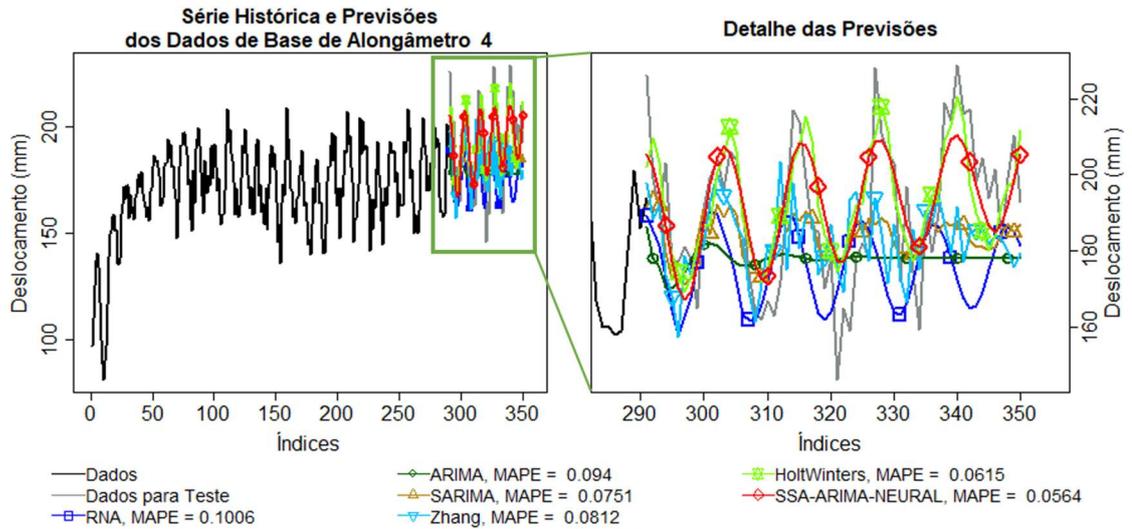


Figura B.6 – Dados do Instrumento Base de Alongamento 4. Fonte: elaboração própria.

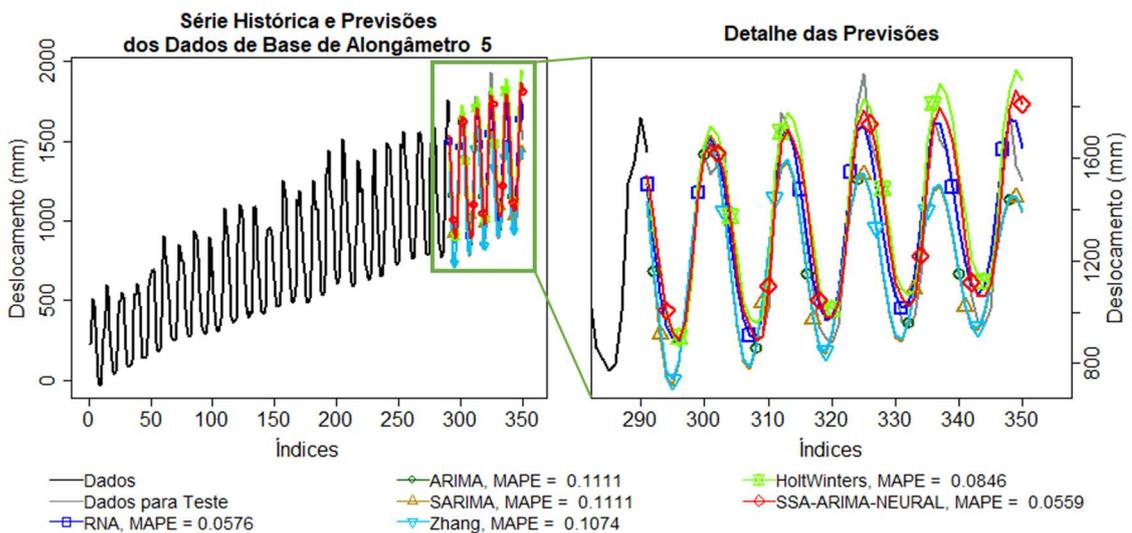


Figura B.7 – Dados do Instrumento Base de Alongamento 5. Fonte: elaboração própria.

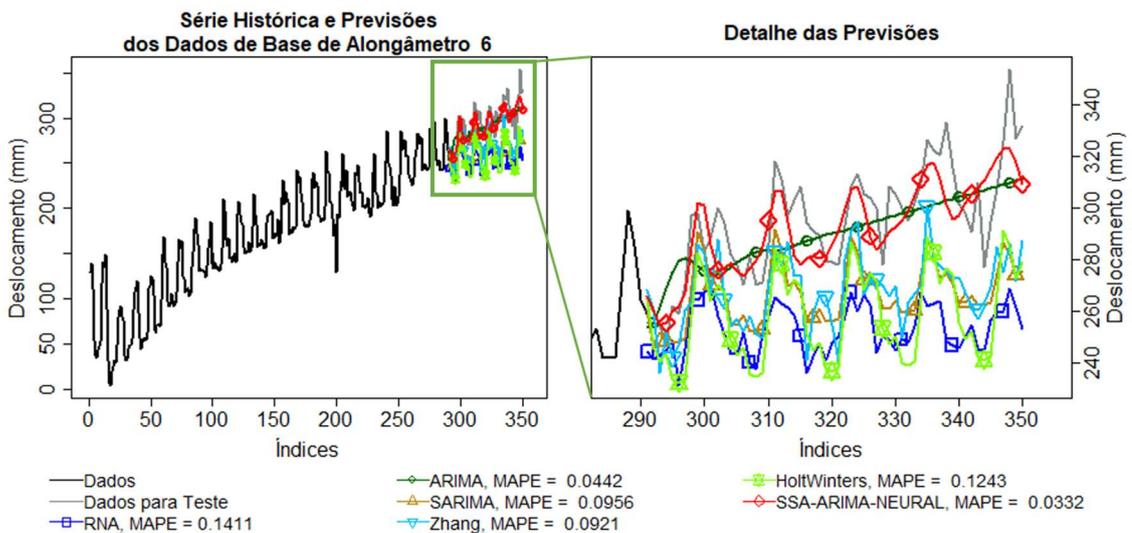


Figura B.8 – Dados do Instrumento Base de Alongamento 6. Fonte: elaboração própria.

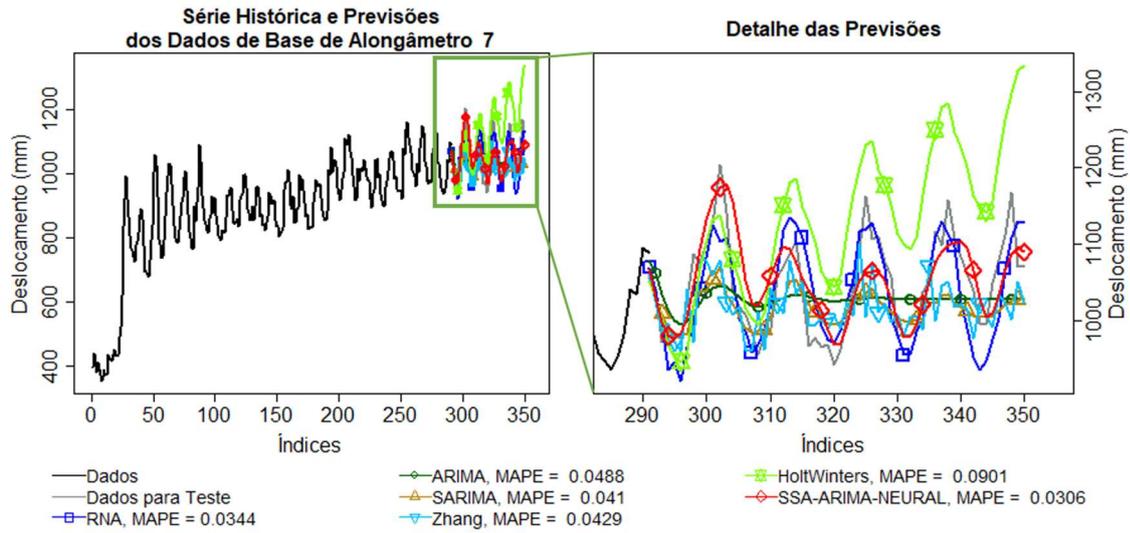


Figura B.9 – Dados do Instrumento Base de Alongâmetro 7. Fonte: elaboração própria.

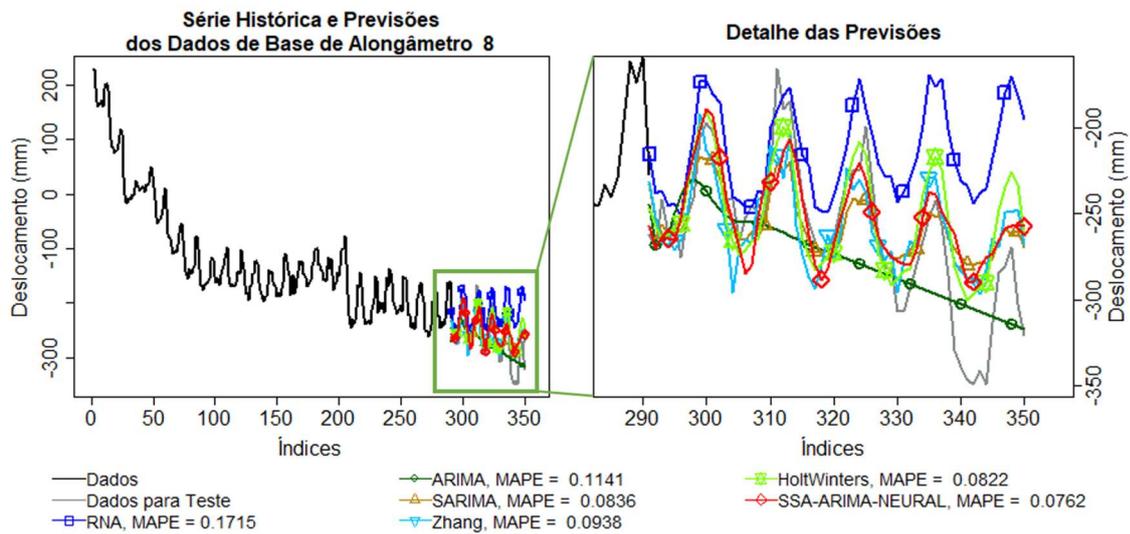


Figura B.10 – Dados do Instrumento Base de Alongâmetro 8. Fonte: elaboração própria.

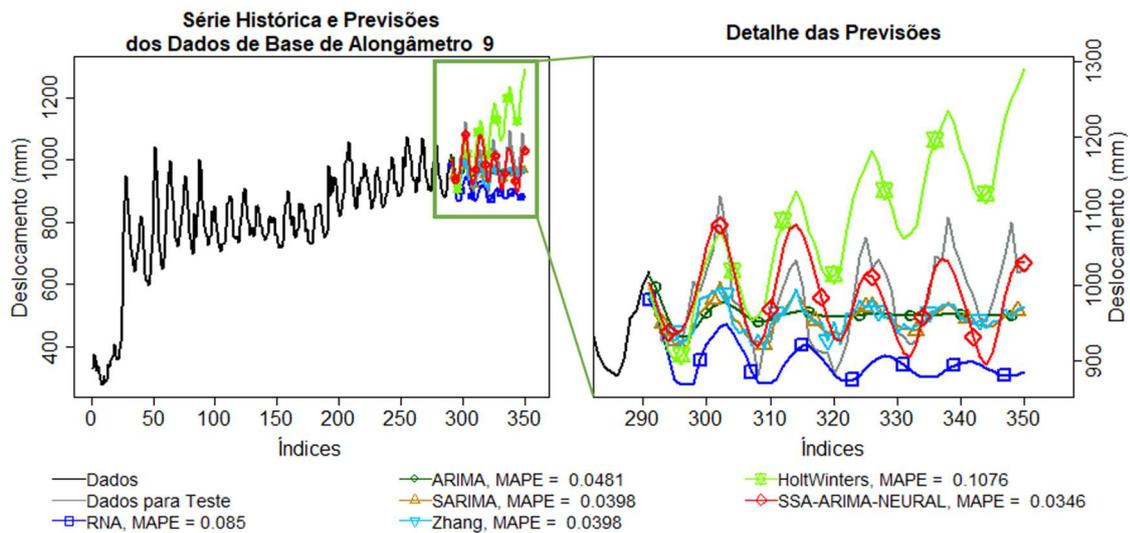


Figura B.11 – Dados do Instrumento Base de Alongâmetro 9. Fonte: elaboração própria.

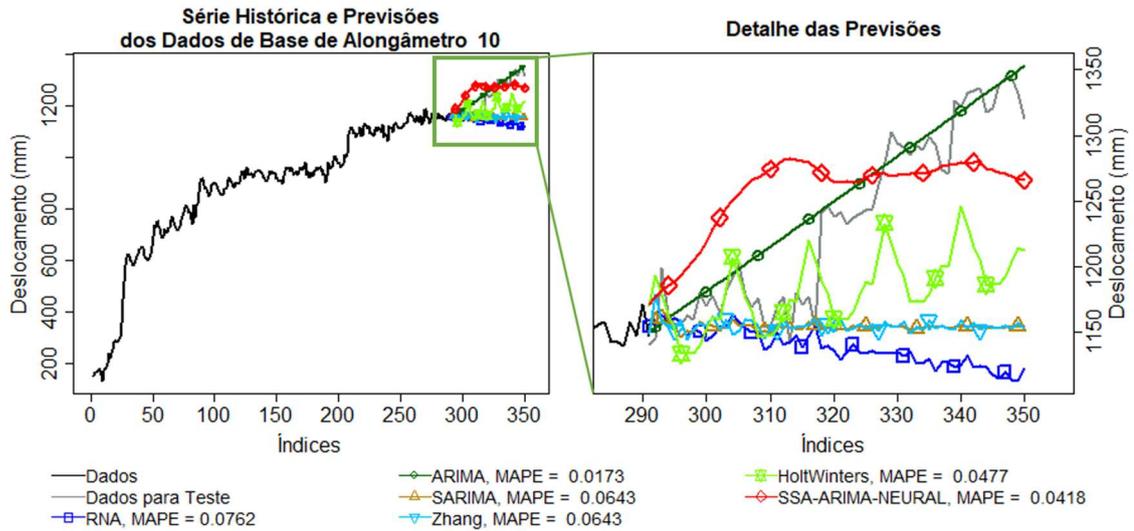


Figura B.12 – Dados do Instrumento Base de Alongâmetro 10. Fonte: elaboração própria.

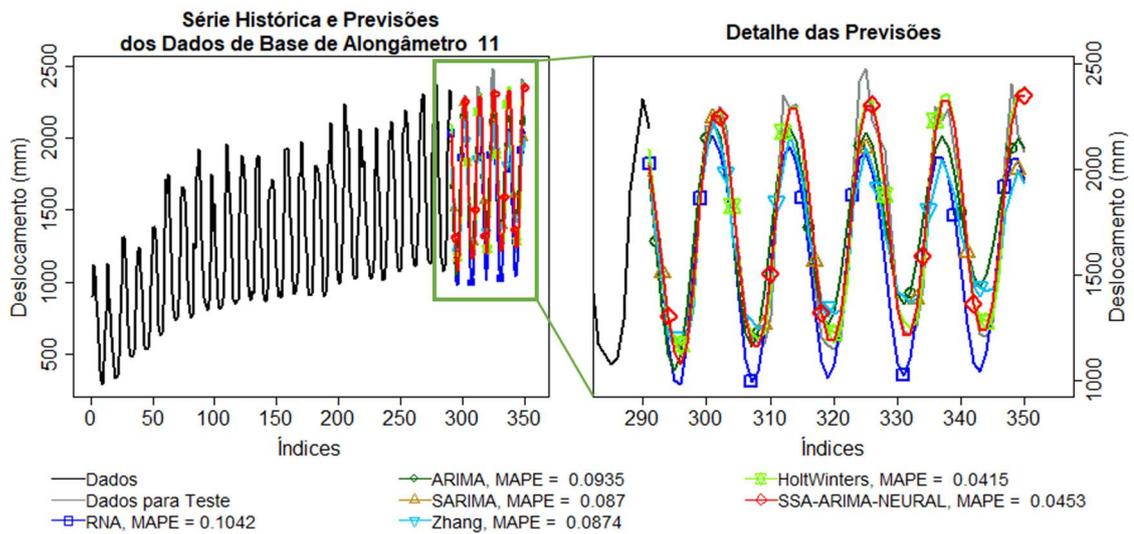


Figura B.13 – Dados do Instrumento Base de Alongâmetro 11. Fonte: elaboração própria.

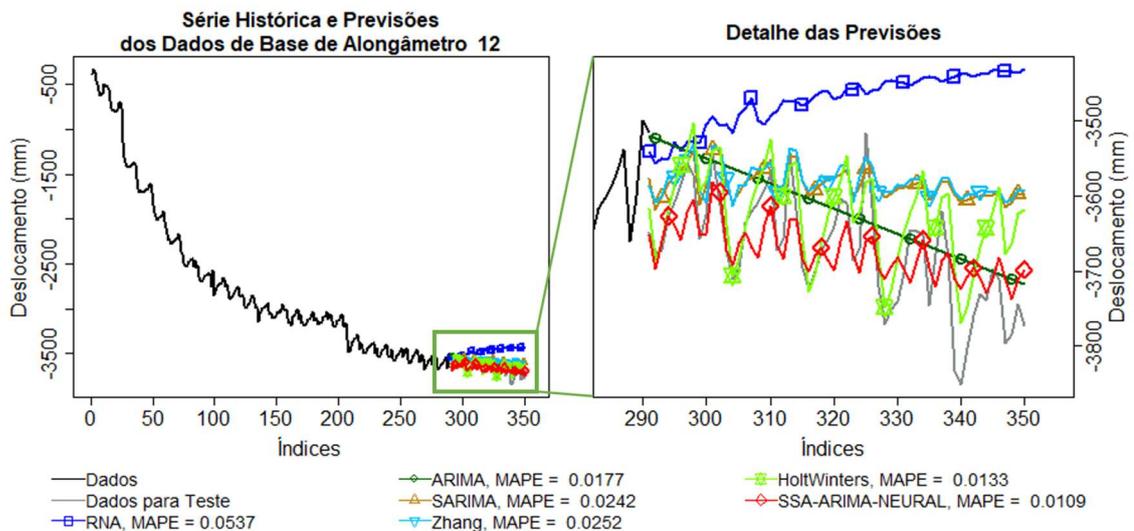


Figura B.14 – Dados do Instrumento Base de Alongâmetro 12. Fonte: elaboração própria.

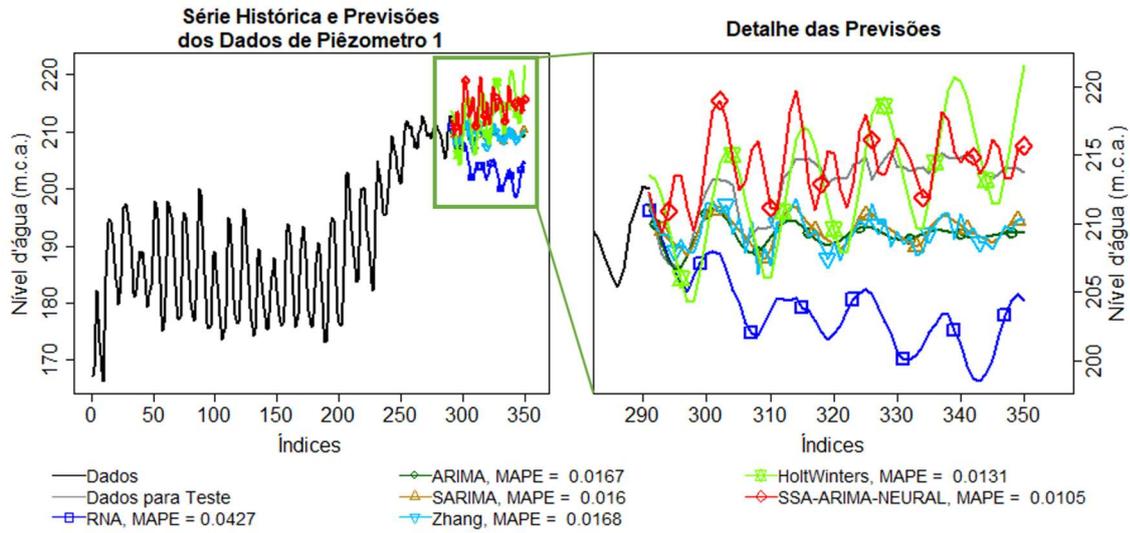


Figura B.15 – Dados do Instrumento Piezômetro 1. Fonte: elaboração própria.

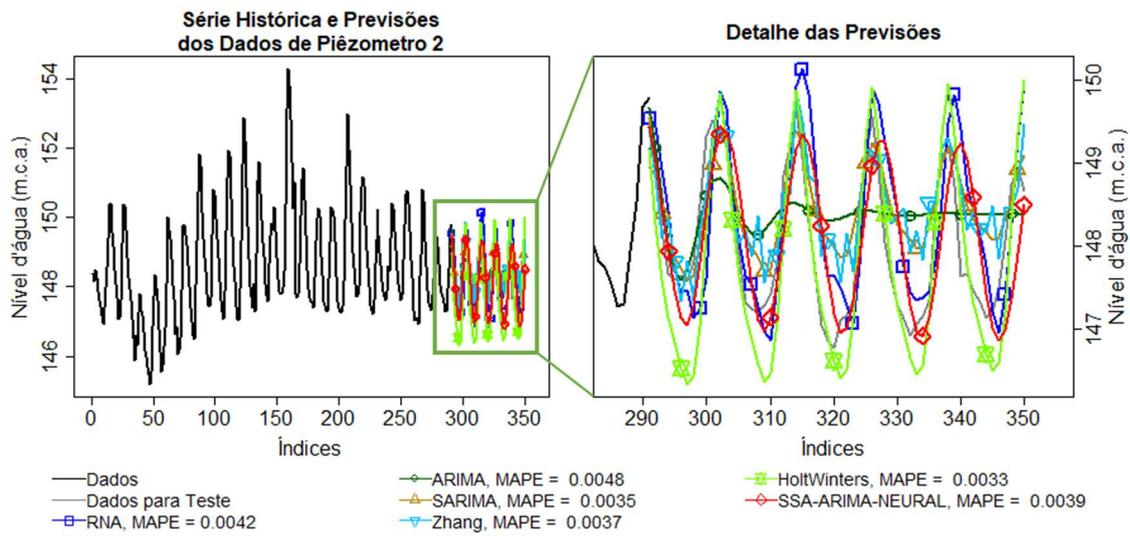


Figura B.16 – Dados do Instrumento Piezômetro 2. Fonte: elaboração própria.

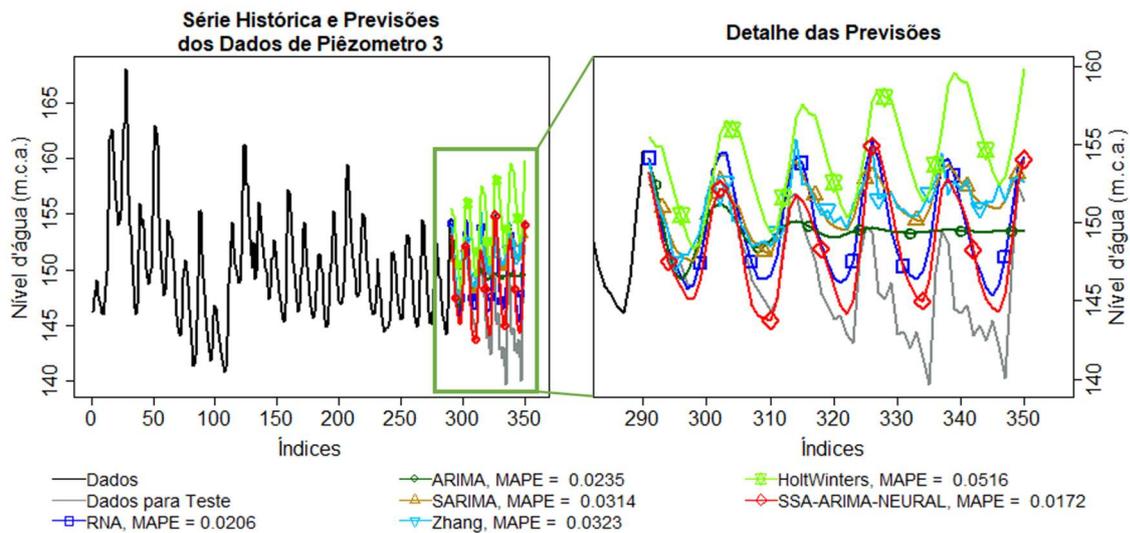


Figura B.17 – Dados do Instrumento Piezômetro 3. Fonte: elaboração própria.

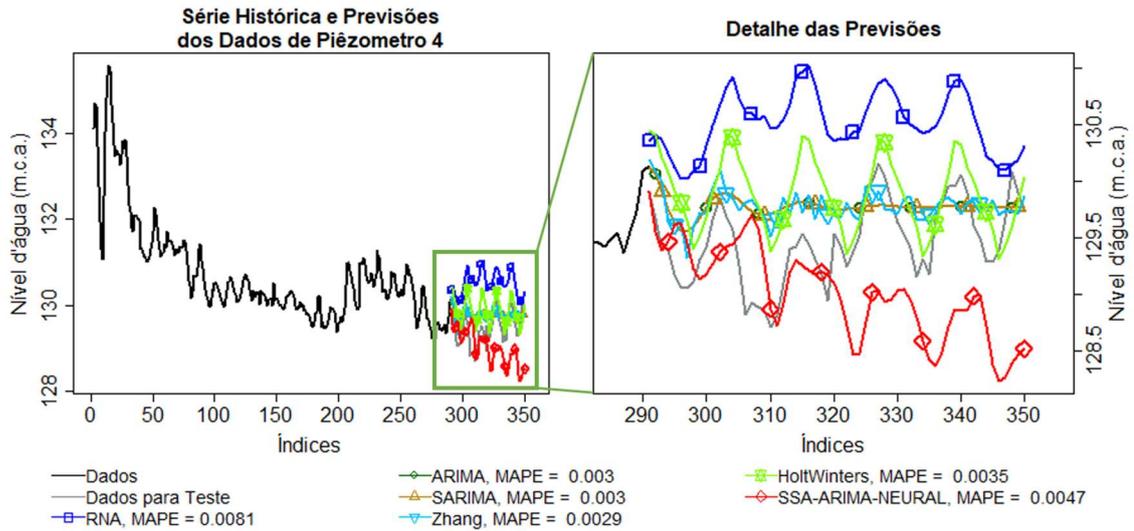


Figura B.18 – Dados do Instrumento Piezômetro 4. Fonte: elaboração própria.

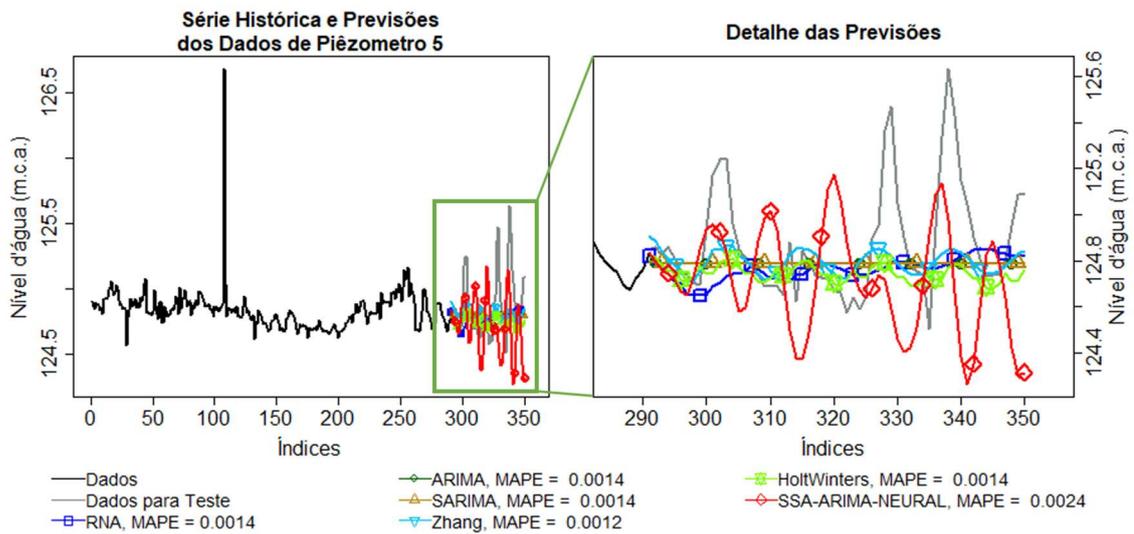


Figura B.19 – Dados do Instrumento Piezômetro 5. Fonte: elaboração própria.

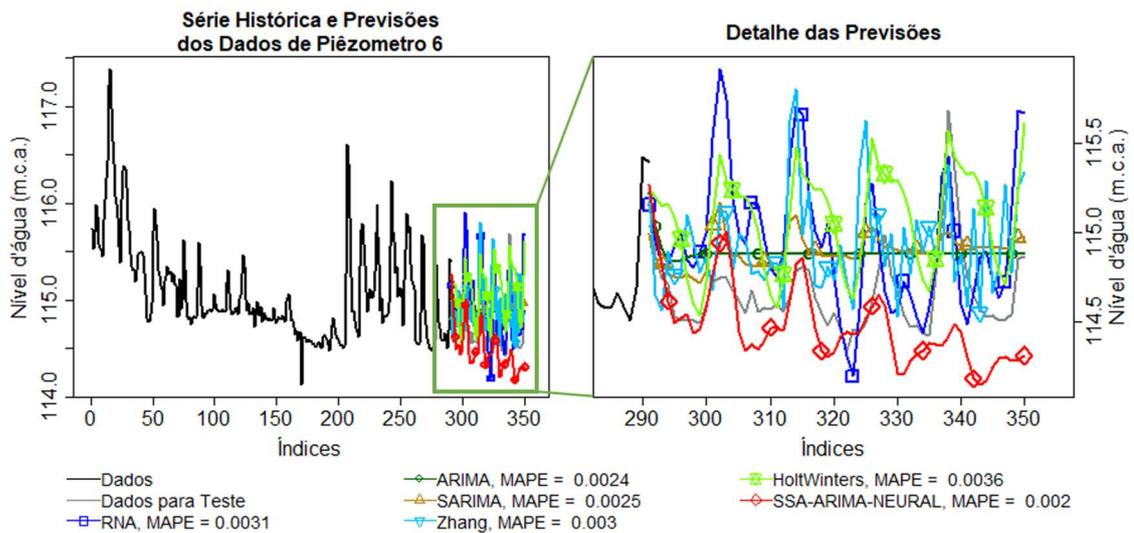


Figura B.20 – Dados do Instrumento Piezômetro 6. Fonte: elaboração própria.

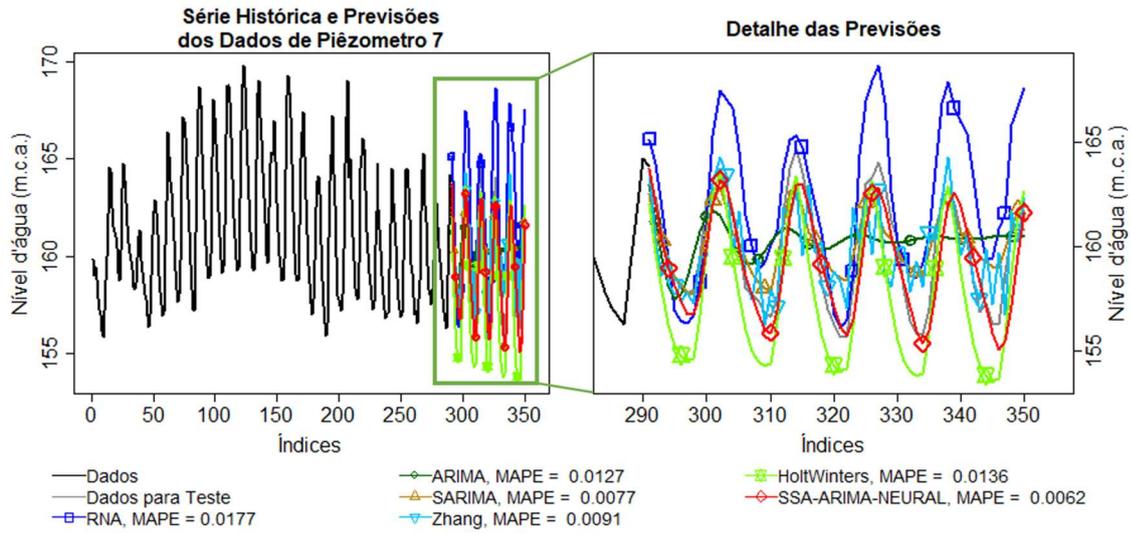


Figura B.21 – Dados do Instrumento Piezômetro 7. Fonte: elaboração própria.

# Apêndice C

## Parâmetros Finais

Tabela C.5.1 – Parâmetros finais de cada modelo.

	RNA		ARIMA			SARIMA					Holt-Winters			Zhang						SSA-ARIMA-NEURAL									
	nós na camada de entrada	nós na camada de saída	p	d	q	p	d	q	P	D	Q	alpha	beta	gamma	p	d	q	P	D	Q	nós na camada de entrada	nós na camada de saída	p	d	q	nós na camada de entrada rede 1	nós na camada escondida rede 1	nós na camada de entrada rede 2	nós na camada escondida rede 2
Pêndulo Direto	62	30	4	1	1	4	1	1	0	0	0	0.5758	0.0326	0.2863	4	1	1	0	0	0	42	21	1	2	0	36	20	12	5
Pêndulo Invertido	60	31	2	0	1	0	0	2	1	0	0	0.341	0.0135	0.2231	0	0	2	1	0	0	42	23	3	2	1	12	7	-	-
Base de alongâmetro 1	62	34	3	1	2	3	1	2	0	0	0	0.1954	0	0.2719	3	1	2	0	0	0	34	18	1	2	0	26	15	-	-
Base de alongâmetro 2	24	9	4	1	1	2	1	2	0	0	0	0.2789	0.0396	0.3591	2	1	2	0	0	0	42	22	3	2	0	34	16	-	-
Base de alongâmetro 3	62	34	3	1	2	1	1	1	2	0	0	0.2112	0.0073	0.2677	1	1	1	2	0	0	40	22	2	2	0	14	5	10	3
Base de alongâmetro 4	22	9	4	1	1	2	1	1	2	0	0	0.339	0.0241	0.4734	2	1	1	2	0	0	42	23	3	2	0	10	5	-	-
Base de alongâmetro 5	26	15	3	1	2	3	1	2	0	0	0	0.1715	0.0029	0.3856	3	1	2	0	0	0	8	3	2	1	1	26	13	22	11
Base de alongâmetro 6	62	34	4	1	1	2	1	1	2	0	0	0.262	0.0328	0.651	2	1	1	2	0	0	40	22	4	2	0	22	13	-	-
Base de alongâmetro 7	62	30	4	1	1	0	1	0	2	0	0	0.8196	0	0.8755	0	1	0	2	0	0	40	23	0	2	0	24	15	62	31
Base de alongâmetro 8	62	30	3	1	1	1	1	0	2	0	0	0.6866	0.024	0.3378	1	1	0	2	0	0	26	11	0	2	0	12	7	22	9
Base de alongâmetro 9	10	3	4	1	1	3	1	0	2	0	0	0.8885	0	0	1	3	1	0	2	0	42	23	0	2	0	60	30	62	30
Base de alongâmetro 10	58	30	0	1	0	3	1	0	2	0	0	0.8467	0.0206	0.9779	3	1	0	2	0	0	32	18	0	2	0	60	30	-	-
Base de alongâmetro 11	10	8	2	1	2	0	1	3	1	0	0	0.2611	0.0086	0.3363	0	1	3	1	0	0	6	1	2	2	0	10	4	12	9
Base de alongâmetro 12	22	13	1	2	1	1	2	2	2	0	0	0.5475	0.0309	0.8095	1	2	2	2	0	0	38	19	0	2	0	10	3	12	5
Piêzometro 1	60	33	4	1	1	2	1	1	2	0	0	0.7384	0.0169	0	1	2	1	1	2	0	42	24	4	2	1	14	8	62	34
Piêzometro 2	62	30	3	1	1	0	1	0	2	0	0	0.8914	0.0079	0.891	0	1	0	2	0	0	42	23	3	2	0	10	8	-	-
Piêzometro 3	60	32	4	1	1	0	1	1	2	0	0	0.878	0.0309	0	1	0	1	2	0	0	42	23	4	1	0	62	33	-	-
Piêzometro 4	60	32	4	1	1	4	1	1	0	0	0	0.4858	0.0162	0	1	4	1	1	0	0	42	21	4	2	1	24	14	60	31
Piêzometro 5	46	24	1	1	1	1	1	1	0	0	0	0.271	0.0139	0.0427	1	1	1	0	0	0	34	20	1	2	0	58	28	58	32
Piêzometro 6	60	28	2	1	1	0	1	3	2	0	0	0.8836	0.01	0	1	0	1	3	2	0	40	20	3	2	0	14	6	22	12
Piêzometro 7	60	28	4	0	1	0	0	3	2	0	0	0.8392	0.0128	0	1	0	0	3	2	0	42	24	3	2	0	10	6	-	-