



MOHAMED LEMINE OULD SID AHMED

**COMBINAÇÃO DE PREVISÕES: UMA ABORDAGEM
USANDO WAVELETS.**

LAVRAS – MG

2018

MOHAMED LEMINE OULD SID AHMED

COMBINAÇÃO DE PREVISÕES: UMA ABORDAGEM USANDO WAVELETS.

Dissertação a ser apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Profa. Dra. Thelma Sáfadi

Orientadora

LAVRAS – MG

2018

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Sid Ahmed, Mohamed Lemine Ould.

Combinação de previsões: uma abordagem usando Wavelets /
Mohamed Lemine Ould Sid Ahmed. - 2018.

54 p. : il.

Orientador(a): Thelma Sáfadi.

.
Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2018.

Bibliografia.

1. Séries temporais. 2. Análise de wavelet. 3. Combinação de
previsões. I. Sáfadi, Thelma. . II. Título.

MOHAMED LEMINE OULD SID AHMED

COMBINAÇÃO DE PREVISÕES: UMA ABORDAGEM USANDO WAVELETS.

Dissertação a ser apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 8 de fevereiro de 2018.

Prof. Dr. Manoel Vitor de Souza Veloso UNIFAL
Prof. Dr. Paulo Henrique Sales Guimarães UFLA

Profa. Dra. Thelma Sáfadi
Orientadora

**LAVRAS – MG
2018**

*Se não puder voar, corra;
Se não puder correr, ande;
Se não puder andar, rasteje,
mas continue em frente de qualquer jeito.;*

Martin Luther King

AGRADECIMENTOS

A Deus, antes de tudo, por tudo.

À minha orientadora, Profa. Dra. Thelma Sáfydi, pela orientação, pelo apoio e pela paciência.

À minha esposa e à minha filha, Patrícia e Yasmin, pelo apoio e pelos sacrifícios.

À UFLA, pela oportunidade.

A todos os professores e os colegas que me ajudaram a concluir essa etapa.

RESUMO

A importância de previsão de alta precisão para séries de vazão natural de afluentes pode ser atribuída ao seu papel fundamental no planejamento dos recursos hídricos, devido ao impacto direto desses nos vários aspectos da atividade econômica e social. Neste trabalho foi aplicado um método híbrido de combinação de previsão de séries temporais. O método utiliza a análise de wavelet, modelagem SARIMA e programação matemática não linear multiobjetivo. Em primeiro momento, é realizada a transformação da série temporal analisada utilizando a transformada de wavelet não decimada *Maximal Overlap Discrete Wavelet Transform* (MODWT) até um certo nível de resolução. Em seguida, modelos SARIMA são ajustados às componentes wavelet, e previsões são geradas para cada componente wavelet por meio do respectivo modelo SARIMA. Na última etapa, otimização matemática é utilizada para combinar as previsões das componentes wavelet, gerando assim a previsão da série temporal original. O método foi aplicado à série mensal de vazão natural de afluentes da barragem de Samuel no estado de Rondônia, no período de março de 1973 a dezembro de 2015. Em termos de ganhos preditivos, o método proposto teve desempenho inferior ao do modelo SARIMA puro.

Palavras-chave: Análise de wavelet, Combinação de previsões, Programação matemática, Séries temporais, Modelos SARIMA, Vazão de afluentes.

ABSTRACT

The reason behind the importance of accurate forecasting of flow of tributaries lies in the fact that they are crucial to hydrological resources planning, which directly impacts many aspects of the economic and social activities. In this work, a hybrid method of combining times series forecasts is applied. Wavelet analysis is implemented, together with SARIMA models and multi-objective nonlinear mathematical programming. At first, the time series to be analyzed is decomposed via Maximal Overlap Discrete Wavelet Transform (MODWT) up to a certain level of resolution. Subsequently, SARIMA models are used to, individually, model and generate forecasts for every wavelet component. In the last stage, mathematical optimization is used to generate forecasts for the original time series in the form of the optimal linear combination of the wavelet components' forecasts, previously obtained by SARIMA models. The method was applied to the time series of the monthly flow of tributaries of Samuel's dam in the state of Rondônia. In terms of predictive gains, pure SARIMA model produced better results.

Keywords: Key words: SARIMA models, Mathematical programming, Times series, Wavelet decomposition, Flow of attributes.

LISTA DE FIGURAS

Figura 3.1 – O gráfico das funções wavelets de Haar, Morlet e Chapéu Mexicano.	17
Figura 3.2 – O gráfico da função wavelet Morlet ,a parte real, em diferentes escalas e deslocamentos.	18
Figura 3.3 – Representação multinível da função $f(x)$	20
Figura 3.4 – Os espaços da AMR	23
Figura 3.5 – Codificação em sub-bandas	25
Figura 3.6 – A arvore da TDW pelo algoritmo de Mallat em 3 níveis de resolução para uma função $f(x)$. A ceta para baixo representa o operados de decimação.	27
Figura 3.7 – MODWT em 3 níveis de resolução para um sinal X	28
Figura 5.1 – A série histórica de vazão natural da barragem de Samuel, o período de setembro de 1972 a fevereiro de 2015, media em metros cúbicos por segundo (m^3/s)	34
Figura 5.2 – A série de vazão natural(denominada "sum") e as componentes de wavelet.	36
Figura 5.3 – Amostra de treino da série de vazão (de março de 1973 a abril de 2013).	37
Figura 5.4 – As funções FAC e FACP da amostra de treino da série de vazão (de março de 1973 a abril de 2013).	37
Figura 5.5 – As funções FAC e FACP do resíduo do modelo SARIMA ajustado à amostra de treino da série original.	38
Figura 5.6 – Componente de aproximação no nível 2 de resolução, S2, e seu correlograma.	40
Figura 5.7 – Componente de detalhe no nível 1 de resolução, d1, e seu correlograma.	41
Figura 5.8 – Componente de detalhe no nível 2 de resolução, d2, e seu correlograma.	42
Figura 5.9 – Correlograma do resíduo do modelo SARIMA ajustado à componente de aproximação S2.	42
Figura 5.10 – Correlograma do resíduo do modelo SARIMA ajustado à componente de detalhe d1.	43
Figura 5.11 – Correlograma do resíduo do modelo SARIMA ajustado à componente de detalhe d2.	43
Figura 5.12 – O gráfico da amostra de treino da série de vazão natural e a sua previsão com ARIMA.	44
Figura 5.13 – O gráfico da amostra de treino da componente de aproximação S2 e a sua previsão.	44

Figura 5.14 – O gráfico da amostra de treino da componente de detalhe d1 e a sua previsão.	45
Figura 5.15 – O gráfico da amostra de treino da componente de detalhe d2 e a sua previsão.	45
Figura 5.16 – O gráfico da amostra de teste da série original e a sua previsão pelo SARIMA.	46
Figura 5.17 – O gráfico da amostra de teste da componente de aproximação S2 e a sua previsão.	46
Figura 5.18 – O gráfico da amostra de teste da componente de detalhe d1 e a sua previsão.	47
Figura 5.19 – O gráfico da amostra de teste da componente de detalhe d2 e a sua previsão.	47
Figura 5.20 – Amostra de teste e a sua previsão pelo modelo SARIMA e pela combinação das previsões das componentes wavelet.	50

LISTA DE TABELAS

Tabela 5.1 – Os coeficientes dos modelos dos modelos SARIMA ajustados às séries. . .	39
Tabela 5.2 – Os valores dos pesos ótimos e dos parâmetros adaptativos da combinação das previsões das componentes wavelet.	48
Tabela 5.3 – Tabela de comparação entre os métodos usado as estatísticas RMSE, MAPE e MAE.	48
Tabela 5.4 – As previsões da amostra de teste (40 passos a frente) da série de vazão, pelos SARIMA e Wavelet-SARIMA.	49

SUMÁRIO

1	Introdução	10
2	Revisão de literatura	11
3	Referencial teórico	13
3.1	Previsão e combinação de modelos	13
3.2	Wavelets	15
3.2.1	O conceito de wavelet	15
3.2.2	Aspectos de wavelet	18
3.2.3	Transformada contínua de wavelet	19
3.2.4	Análise de multiresolução	19
3.2.5	Transformada wavelet discreta(decimada)	23
3.2.6	Maximal overlap discrete wavelet transform (MODWT)	27
3.3	Modelos SARIMA	28
3.4	Programação matemática	30
3.4.1	Programação não linear	31
3.4.2	Programação matemática multiobjetivo	31
4	Métodos e materiais	32
4.1	Metodologia	32
4.2	Material	32
5	Resultados e discussão	34
5.1	Resultados	34
5.1.1	Decomposição wavelet	35
5.1.2	Modelagem SARIMA	36
5.1.3	Combinação de previsões	47
6	Conclusão	51
	REFERÊNCIAS	52

1 INTRODUÇÃO

A previsão de séries temporais é um dos problemas mais importantes que os analistas enfrentam sendo crucial para vários tipos de planejamento e tomada de decisão em diversas áreas como: finanças, economia, meio ambiente e outras áreas (MONTGOMERY; JENNINGS; KULAHCI, 2015). De acordo com Júnior et al. (2016), os métodos de previsão de séries temporais estocásticas podem ser divididos, de modo geral, em três grupos: estatísticos, como os modelos Box & Jenkins; os de inteligência artificial, como redes neurais e os de combinação, linear ou não linear.

Desde a obra seminal de Bates e Granger (1969), os trabalhos teóricos e empíricos vêm fornecendo evidências de que quando existe mais de um modelo plausível para fazer a previsão de uma certa variável (série temporal), a combinação das previsões desses modelos, frequentemente, fornece um desempenho mais robusto e de maior precisão em termos do erro quadrático médio de previsão comparada com a previsão de um único modelo (QIAN et al., 2015). Porém, não há consenso quanto à escolha de pesos dos modelos combinados.

Na literatura tem-se várias técnicas para escolher pesos, como em Elliott (2011), Hansen (2008), Vasnev, Claeskens e Wang (2014), e encontram-se várias combinações de modelos e métodos de previsão que incorporam, além de combinação de modelos, outros métodos auxiliares de pré-processamento e filtragem de séries temporais, como Chau e Wu (2010) e Khandelwal, Adhikari e Verma (2015). Desses últimos, destaca-se a análise de wavelets pela sua capacidade de lidar com certos problemas de séries temporais, como o problema de não estacionariedade.

Tendo isso em vista, neste trabalho foi aplicado um método híbrido de combinação de previsões, integrando à análise de wavelet, modelos SARIMA e programação matemática. Utilizou-se a transformada de wavelet “*Maximal Overlap Discrete Wavelet Transform*” (MODWT) que, de acordo com Walden (2001), possui características que a torna mais adequada para a análise de séries temporais. O método foi aplicado à série histórica de vazão natural mensal da barragem da Usina Hidroelétrica de Samuel, no estado de Rondônia, no período de março de 1973 a dezembro de 2015.

Assim, o objetivo desse trabalho é aplicar um método híbrido de combinação de previsões de séries temporais com a implementação de análise de wavelet para melhorar a qualidade de previsão.

2 REVISÃO DE LITERATURA

A literatura apresenta vários exemplos de trabalhos de previsão de séries temporais que utilizam uma combinação de técnicas e métodos de modelagem e pré-processamento com o objetivo de obter previsões com melhor precisão e mais confiabilidade. Tan et al. (2010) aplicam um método que combina a transformação wavelet com as modelagens ARIMA e GARCH para a previsão de série de preço de energia elétrica no mercado espanhol um dia para frente. Eles ajustam um modelo ARIMA-GARCH à componente wavelet de aproximação e modelos GARCH para as componentes wavelet de detalhe. A previsão da série original foi obtida na forma de composição das previsões das componentes wavelet. Eles concluem que “o método foi superior a umas das técnicas recentemente publicadas”.

Kriechbaumer et al. (2014) usam uma abordagem wavelet-ARIMA para gerar previsões para as séries de preços de metais. Eles realizam a transformação da série de preços em componentes wavelet que, em seguida, são modeladas individualmente com modelos ARIMA. A composição das previsões geradas pelos respectivos modelos ARIMA fornece a previsão da série original. Os resultados mostraram que a abordagem proposta provou ser superior em comparação com o modelo ARIMA puro e que o uso da transformação MODWT (Maximal Overlap Discrete Wavelet Transform) traz mais ganhos preditivos que o uso de DWT (Discrete Wavelet Transform).

Schlueter e Deuschle (2014) conduziram um experimento empírico com o objetivo de verificar se existem ganhos preditivos na agregação de análise de wavelet, como um método de pré-processamento, à modelagem ARIMA e ao método de Census-12¹. Eles concluíram que existem ganhos significativos somente quando existe uma estrutura (de sazonalidade) de médio a longo prazo na série.

Joo e Kim (2015) propuseram um método híbrido Wavelet-ARIMA em que eles dividem a série temporal em duas componentes, uma de tendência e outra de variação, usam a transformada wavelet na decomposição e reconstrução da série original e usam ARIMA na modelagem de cada uma das duas componentes. Eles aplicaram o método em dados simulados e concluíram que a modelagem ARIMA gera resultados de maior precisão quando usada em conjunto com a filtragem via wavelet.

¹ X-12-ARIMA, é um programa usado pelo *Census Bureau* dos Estados Unidos para ajustar séries sazonais.

Júnior et al. (2015) usam uma metodologia que integra a modelagem de Box & Jenkins e decomposição wavelet e programação matemática na previsão de série de vazão mensal do sul do Brasil. Eles usam a programação matemática para combinar as previsões das componentes wavelets da série original para obter sua previsão. Concluem que o pré-processamento com wavelet traz ganhos preditivos em comparação com um modelo ARIMA puro.

Júnior et al. (2016) usam os métodos Box & Jenkins, Redes Neurais Artificiais (RNA), Wavelet-RNA e Wavelet-ARIMA para obter previsões da série de vazão mensal no posto 266 (Itaipu). Eles mostram que os métodos compostos Wavelet-RNA e Wavelet-ARIMA forneceram previsões melhores em relação aos modelos individuais ARIMA e RNA.

3 REFERENCIAL TEÓRICO

3.1 Previsão e combinação de modelos

De acordo com Montgomery, Jennings e Kulahci (2015), a análise de dados temporais e previsão de valores futuros de séries temporais estão entre os problemas mais importantes que os analistas enfrentam em varias áreas como: finanças, economia, meio ambiente e outras. Previsão é a predição de um ou mais eventos futuros. A predição de eventos futuros é crucial para vários tipos de planejamento e tomada de decisão em diversas áreas. As técnicas de previsão se dividem em qualitativas, de natureza subjetiva, e quantitativas, que fazem uso de dados históricos e de modelos de previsão. Os mais importantes tipos de previsão são modelos de regressão, modelos de suavização e modelos gerais de séries temporais. Esses últimos usam as características estatísticas dos dados históricos para identificar um modelo formal e, em seguida, estimam seus parâmetros desconhecidos, normalmente usando o método dos mínimos quadrados que, de fato, fornece a base para quase todos os estudos teóricos. Segundo Morettin e Tolo (2006) e Brockwell e Davis (2016), após construir o modelo é necessário especificar uma função-perda, como o erro quadrático médio, para verificar a precisão das previsões geradas pelo modelo. Neste caso, se os dados históricos forem realmente gerados pelo modelo ajustado, esse procedimento resultará em um erro quadrático mínimo.

Em 1963, Barnard introduziu um método simples de combinação de previsões usando a média aritmética de dois conjuntos de previsões obtidas a partir de dois modelos diferentes de previsão (BATES; GRANGER, 1969).

Bates e Granger (1969) apresentaram um método mais geral para calcular as combinações de previsões geradas por modelos diferentes. Eles observaram que seria mais razoável considerar pesos maiores para as previsões que parecem ter EQM (erro quadrático médio) menor, ao invés de considerar pesos iguais para todas as previsões. Assim, em uma generalização do método de Barnard, eles consideraram uma combinação de pesos relativos usando otimização (CLEMENTS, 2017).

Clemen (1989) fez uma revisão de literatura sobre combinação de modelos para previsão. Ele verificou que durante os vinte anos desde a publicação de Bates e Granger (1969), foram desenvolvidos modelos para encontrar combinação ótima de previsões, foram feitas simulações e estudos empíricos para testar os modelos e foram apresentadas interpretações Bayesianas. O

resultado foi praticamente unânime: combinação de múltiplas previsões aumenta a precisão da previsão.

Em uma revisão mais recente de literatura sobre as abordagens de combinação de previsões, Mancuso e Werner (2013) mostraram que o método continuou a sua trajetória de evolução, ganhando mais contribuições em ambos os campos teórico e empírico. Os autores apresentaram novos métodos além dos métodos clássicos. Vários outros trabalhos foram publicados desde esta última revisão.

Para ilustrar o método de Bates e Granger (1969), suponha que se tenha duas previsões distintas f_{1t} e f_{2t} para h passos a frente da série y_t , assumindo que os erros das previsões têm média zero. Os autores sugerem a seguinte fórmula para calcular a combinação f_{ct} das duas previsões:

$$f_{ct} = (1 - \lambda)f_{1t} + \lambda f_{2t}. \quad (3.1)$$

Quando $0 \leq \lambda \leq 1$, tem-se uma média ponderada simples das duas previsões, sendo que λ pode ser escolhido com base na precisão relativa das previsões individuais f_{1t} e f_{2t} . Se o erro de previsão associado a f_{ct} é $\varepsilon_t = y_t - f_{ct}$, então

$$E(\varepsilon_t^2) = \lambda^2 \sigma_1^2 + (1 - \lambda)^2 \sigma_2^2 + 2\rho\lambda\sigma_1(1 - \lambda)\sigma_2, \quad (3.2)$$

em que σ_1^2 e σ_2^2 são respectivamente, o erro quadrático esperado de f_{1t} e f_{2t} e ρ é a correlação entre os erros de previsão ε_{1t} e ε_{2t} . Assim, se as duas previsões, f_{1t} e f_{2t} , são estacionárias e não viesadas, o λ que minimiza $E(\varepsilon_t^2)$ pode ser encontrado resolvendo o seguinte problema de otimização

$$\lambda_{opt} = \underset{\lambda}{\operatorname{argmin}} E(\varepsilon_t^2) = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - \rho\sigma_1\sigma_2}. \quad (3.3)$$

Quando $\rho = 0$,

$$\lambda_{opt} = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}. \quad (3.4)$$

Pode-se verificar que a variância da combinação f_{ct} , quando o peso ótimo λ_{opt} é usado, é dada por

$$(\sigma_c^*)^2 = \frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}, \quad (3.5)$$

que é menor que σ_1^2 e σ_2^2 , ou seja, $(\sigma_c^*)^2 \leq \min(\sigma_1^2, \sigma_2^2)$. Quando $\sigma_1^2 = \sigma_2^2 = \sigma^2$ tem-se que:

$$(\sigma_c^*)^2 = \frac{1+\rho}{2}\sigma^2. \quad (3.6)$$

Dado que $|\rho| \leq 1$, o erro quadrático esperado associado à previsão combinada f_{ct} , usando o peso ótimo λ_{opt} é menor que o erro quadrático esperado de qualquer uma das duas previsões individuais f_{1t} e f_{2t} para todo $\rho \neq 1$. A equação (3.5) mostra que, se as previsões não forem perfeitamente correlacionadas, a previsão combinada supera qualquer uma das previsões individuais, e em particular quando essas têm precisões muito próximas ou iguais.

De fato, com este método de combinação ótima, tem-se sempre uma melhoria, a previsão continua não viesada e a variância diminui. Bates e Granger notaram que o valor ótimo de λ não é conhecido no início do processo da combinação das previsões. Sendo que o valor muda com cada passo a frente à medida que mais evidências sobre o desempenho relativo das previsões individuais f_{1t} e f_{2t} vão se acumulando. Logo, a previsão combinada para um período T é dada por

$$f_{CT} = (1 - \lambda_T)f_{T1} + \lambda_T f_{T2}. \quad (3.7)$$

(CLEMENTS, 2017; VASNEV; CLAESKENS; WANG, 2014; BATES; GRANGER, 1969).

3.2 Wavelets

3.2.1 O conceito de wavelet

De acordo com Percival e Walden (2006), wavelet, como o nome sugere, é uma "onda pequena" que basicamente cresce e decai em tempo limitado. Em outras palavras, wavelet é uma pequena onda localizada. Hoje, a palavra é frequentemente usada para descrever uma base ortonormal do espaço das funções de quadrado integrável, $L^2(\mathbb{R})$ (ECKLEY, 2001). A base de wavelet é construída a partir de aplicação de dilatações e translações de uma função básica ψ , chamada wavelet mãe. Segundo Meyer (1993), uma wavelet mãe pode ser definida da seguinte maneira:

Seja $m \in \mathbb{N}$. Então para $x \in \mathbb{R}$ a função $\psi(x)$ é chamada wavelet mãe de ordem m se as seguintes propriedades se aplicam:

¹ Demonstração: $\frac{\sigma^2 \sigma^2 (1-\rho^2)}{\sigma^2 + \sigma^2 - \rho \sigma^2} = \frac{1-\rho^2}{2-\rho} \sigma^2 = \frac{1-\rho^2}{1+(1-\rho)} \sigma^2 \frac{1-(1-\rho)}{1-(1-\rho)} = \frac{1-\rho^2}{1-(1-\rho)^2} \rho \sigma^2 = \frac{(1-\rho)(1+\rho)}{2(1-\rho)} \sigma^2 = \frac{1+\rho}{2} \sigma^2$

² Se $f(x) \in L^2(\mathbb{R})$, então $\|f\|^2 = \int_{-\infty}^{\infty} |f(t)|^2 dt < \infty$.

- i. Se $m = 0$, $\psi(x) \in L^\infty(\mathbb{R})^3$, então $\psi(x)$ e todas as suas derivadas até a ordem m pertencem a $L^\infty(\mathbb{R})$.
- ii. $\psi(x)$ e todas as suas derivadas até a ordem m decrescem rapidamente quando $x \rightarrow \pm \infty$.
- iii. Para todo $k \in \{0, \dots, m\}$, tem-se que

$$\int_{-\infty}^{+\infty} x^k \psi(x) dx = 0. \quad (3.8)$$

- iv. A coleção $\{\psi_{j,k}(x)\}_{j,k \in \mathbb{Z}}$ constitui uma base ortogonal de $L^2(\mathbb{R})$, sendo que a $\psi_{j,k}$ é construída a partir de $\psi(x)$ aplicando dilatações e translações pela fórmula:

$$\psi_{j,k}(x) = 2^{j/2} \psi(2x - k), \quad (3.9)$$

em que j e k são parâmetros de dilatação (escala) e translação (localização), respectivamente (ECKLEY, 2001).

Como pode ser inferido da definição, existe, potencialmente, um número infinito de wavelets, e essas são construídas de acordo com o objetivo que elas podem servir. De fato, existe uma gama de funções wavelets utilizadas na construção de bases para transformadas wavelets, que serão definidas nas próximas sessões. Exemplos de wavelets são Haar, a wavelet de Morlet e Chapéu Mexicano, Figura 3.1. A figura 3.2 é uma representação de versões dilatadas e transladadas da wavelet de Morlet.

O primeiro e o mais simples sistema de wavelet com base ortogonal é o sistema Haar, que tem como wavelet mãe a função matemática

$$\psi(x) = \begin{cases} 1, & \text{se } x \in [0, 1/2) \\ -1, & \text{se } x \in [1/2, 1) \\ 0, & \text{caso contrario.} \end{cases} \quad (3.10)$$

³ O conjunto de funções- L^p ($p \geq 1$) é uma generalização do espaço L^2 , em que as funções são mensuráveis ao em vez de serem necessariamente de quadrado integrável. Em um espaço mensurável X , a norma L^p de uma função f é dada por $\|f\|_{L^p} = (\int_X |f|^p)^{\frac{1}{p}}$. O espaço L^∞ é uma generalização do espaço L^p para $p = \infty$.

Essa função, apesar de ser muito simples, apresenta vários aspectos característicos de wavelets (NASON, 2010). Duas características importantes dessa função são: a oscilação, que pode ser representada matematicamente pela condição $\int_{-\infty}^{+\infty} \psi(x)dx = 0$, uma propriedade compartilhada por todas as wavelets de suporte compacto, sendo que nem todas as wavelets tem suporte compacto mas elas têm que decair rapidamente para zero.

As representações matemáticas das funções chapéu mexicano e Morlet wavelet são, respectivamente

$$\psi(x) = (1 - x^2)e^{-x^2/2} \quad (3.11)$$

e

$$\psi(x) = e^{-2\pi(x-x_0)^2}, \quad (3.12)$$

em que x_0 é uma constante.

Figura 3.1 – O gráfico das funções wavelets de Haar, Morlet e Chapéu Mexicano.

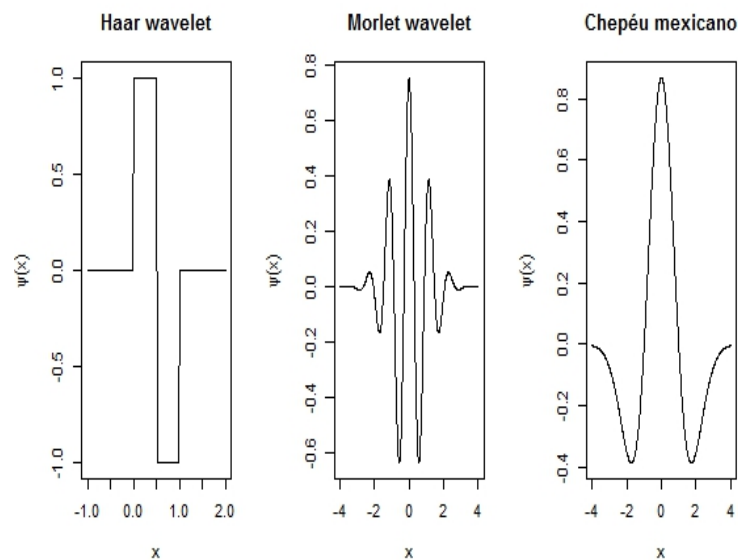
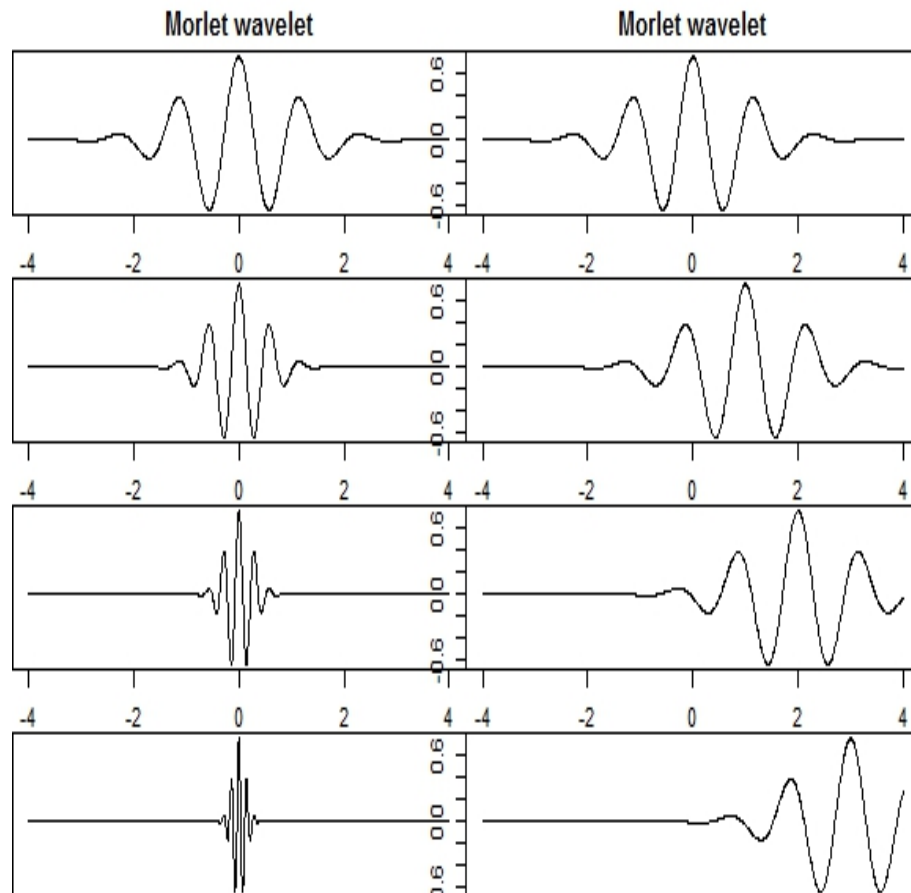


Figura 3.2 – O gráfico da função wavelet Morlet ,a parte real, em diferentes escalas e deslocamentos.



3.2.2 Aspectos de wavelet

A análise de wavelet foi desenvolvida como ferramenta capaz de analisar localmente sinais complexos com setores não estacionários em várias escalas. Seu desenvolvimento acelerado levou a um número enorme de aplicações em vários campos. O seu sucesso é intimamente ligado à sua capacidade de ser adaptada para resolver uma vasta gama de questões teóricas e práticas (MISITI et al., 2013).

De acordo com Misiti et al. (2013), a análise de wavelets mostrou-se como ferramenta poderosa para resolver três tipos de problemas. O primeiro é a análise para investigar dados e identificar suas regularidades em alta resolução. Devido à sua localização nos dois domínios, de frequência e de tempo, wavelet é capaz de descrever as propriedades de um sinal em várias escalas de tempo. O segundo problema é apuração ou estimação de sinais. Os métodos de apuração de sinais baseados em wavelets criam algoritmos normalmente simples, poderosos e fáceis de adaptar. O princípio é transformar o sinal com wavelets, eliminar o ruído e em seguida aplicar a inversa da transformação de wavelets para reconstruir o sinal. O resultado é o sinal

puro. O terceiro problema é a compressão, de forma particular compressão de imagens, no qual wavelet é um método muito competitivo. A compressão padrão de imagens JPEG2000 é baseada em wavelet. A razão principal para a eficiência de wavelet vem da sua capacidade de, geralmente, concentrar a energia do sinal em poucos coeficientes significativamente não nulos, o que gera uma estrutura de decomposição esparsa e capaz de ser codificada com poucas informações.

3.2.3 Transformada contínua de wavelet

De acordo com Starck, Murtagh e Bijaoui (1998), pela definição de Morlet-Grossman, a transformada de wavelets contínua para um sinal unidimensional $f(x) \in L^2(\mathbb{R})$ é dada por:

$$W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x) \psi\left(\frac{x-b}{a}\right) dx \quad (3.13)$$

em que

$W(a, b)$ é o coeficiente de wavelet para a função $f(x)$,

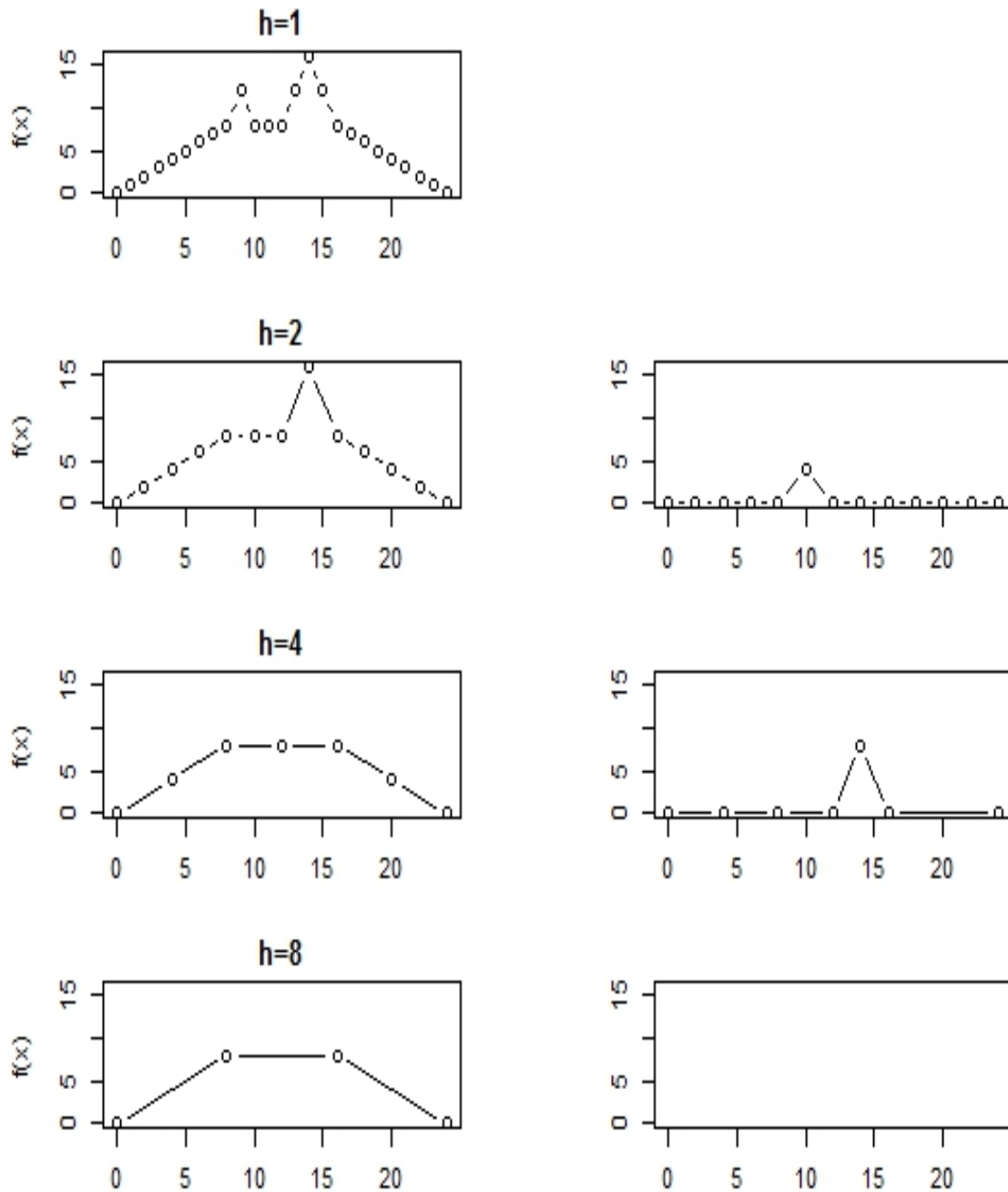
$\psi(x)$ é a wavelet da análise,

$a > 0$ é o parâmetro de escala,

b é o parâmetro de posição.

3.2.4 Análise de multiresolução

Para representar uma função $f(x)$ contínua em um certo nível de aproximação, tem-se que discretizá-la usando um passo h , determinado pelo seu seguimento de mais rápida variação. Isso gera um grande número de pontos. A representação da função por meio de discretizações com passos de vários tamanhos, permite uma redução significativa do número de pontos necessários para uma representação precisa. A aproximação mais suave, determinada pelo tamanho do passo h , junto com os detalhes em cada nível de resolução representam a função original de forma completa. Na Figura 3.3, o tamanho do passo h é dobrado em cada nível de resolução. As aproximações estão representadas nos gráficos do lado esquerdo da figura e os detalhes estão representados nos gráficos do lado direito (GOSWAMI; CHAN, 2011).

Figura 3.3 – Representação multinível da função $f(x)$.

Este exemplo pode ser considerado uma apresentação não formal da análise de multiresolução.

Goswami e Chan (2011) afirmam que a análise de multiresolução forma os mais importantes elementos base para a construção das funções de escala e wavelets (a base da análise de wavelets) e para desenvolver algoritmos.

Segundo Mallat (1989), a análise de multiresolução é resultado de subconjuntos embebidos gerados pela interpolação em escalas diferentes.

Definição: Uma sequência $\{V_j : j \in \mathbb{Z}\}$ de subespaços fechados de $L^2(\mathbb{R})$ é uma aproximação de multiresolução se as seguintes propriedades são satisfeitas

- Os espaços tem uma intercessão trivial

$$\bigcap_{j \in \mathbb{Z}} V_j = 0; \quad (3.14)$$

- A união dos espaços é densa em $L^2(\mathbb{R})$

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}); \quad (3.15)$$

-

$$f(x) \in V_j \Rightarrow f(2x) \in V_{j-1} \quad \forall j \in \mathbb{Z}; \quad (3.16)$$

-

$$f(x) \in V_0 \Rightarrow f(x-k) \in V_0 \quad \forall k \in \mathbb{Z}; \quad (3.17)$$

- Existe uma função escala $\phi \in V_0$ de modo que $\{\phi(x-n)\}$ é uma base ortogonal de V_0 .

Pode-se notar que as condições 3.16 e 3.17 implicam que $\{\phi_{j,k} : k \in \mathbb{Z}\}$ constitui uma base ortogonal para V_j . Além disso, como $V_0 \subset V_1$, a função $\phi(x) \in V_0$ pode ser representada como combinação linear de V_1 . Isso significa que, para alguns coeficientes $\{h_k\}_{k \in \mathbb{Z}}$, pode-se expressar $\phi(x)$ como

$$\phi(x) = \sum_{k \in \mathbb{Z}} h_k \phi_{1,k}(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k \phi(2x-k) \quad \text{com} \quad \sum |h|^2 < +\infty. \quad (3.18)$$

Esta é a equação principal da análise de multiresolução (AMR) e tem solução única, de forma a permitir que os coeficientes $\{h_k\}$ sejam usadas para determinar de forma única a função de escala $\phi(x)$. Os coeficientes $\{h_k\}$ são chamados de *coeficientes do filtro passa-baixa* ou *filtro de escala*.

A AMR também define subespaços $W_{j-1} \subset L^2(\mathbb{R})$ chamados de espaços wavelets. Esses subespaços são os complementares ortogonais de V_{j-1} em V_j . Assim

$$V_j = V_{j-1} \oplus W_{j-1}, \quad (3.19)$$

sendo que $W_j \perp V_j$ e $W_j \perp W_{j'}$ caso $j' \neq j$. Portanto, $L^2(\mathbb{R})$ pode ser decomposto em subespaços mutuamente ortogonais. Além disso, o $\{W_j\}$ herda a propriedade de escalonamento em 4.16 de $\{V_j\}$, ou seja

$$f(x) \in W_j \Rightarrow f(2x) \in W_{j-1}. \quad (3.20)$$

Portanto, se existe uma função ψ de forma que as suas translações formam uma base ortogonal para W_0 , então as relações 4.20 e 4.21 garantem que

$$\{\psi_{j,k} : \psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^{\frac{j}{2}}x - k)\}_{j,k \in \mathbb{Z}}$$

forma uma base ortonormal para $L^2(\mathbb{R})$.

Para derivar uma função wavelet ψ a partir de uma função de escala ϕ , percebendo que $W_0 \subset V_1 \Rightarrow \psi(x) \in V_1$, $\psi(x)$ pode ser representada da seguinte forma

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k \phi(2x - k). \quad (3.21)$$

para alguns coeficientes $\{g_k\}_{k \in \mathbb{Z}}$, chamados de "filtro de passa alta",⁴ tal que $\sum_{n \in \mathbb{Z}} |g_k|^2 < +\infty$.

Os coeficientes de escala, h , e de detalhe, g , são ligados pela seguinte relação

$$g_n = (-1)^n h_{(n-1)}. \quad (3.22)$$

Essa relação é chamada "*quadrature mirror filter*". Os filtros g_k e h_k podem ser calculados a partir de um sistema wavelet ortogonal por meio das relações

$$h_k = \sqrt{2} \sum_{k \in \mathbb{Z}} \phi(x) \phi(2x - k) \quad (3.23)$$

e

$$g_k = \sqrt{2} \sum_{k \in \mathbb{Z}} \psi(x) \phi(2x - k). \quad (3.24)$$

O exemplo mais simples da AMR pode ser obtido usando a função escala Haar

$$\phi(x) = \begin{cases} 1, & \text{se } 0 \leq x \leq 1 \\ 0, & \text{caso contrario} \end{cases} \quad (3.25)$$

⁴ Considerando uma sinal passando por um filtro, os filtros "passa baixa" e "passa alta" tem como saída as componentes, respectivamente, de baixa frequência e de alta frequência do sinal.

Os coeficientes h_k e g_k neste caso são

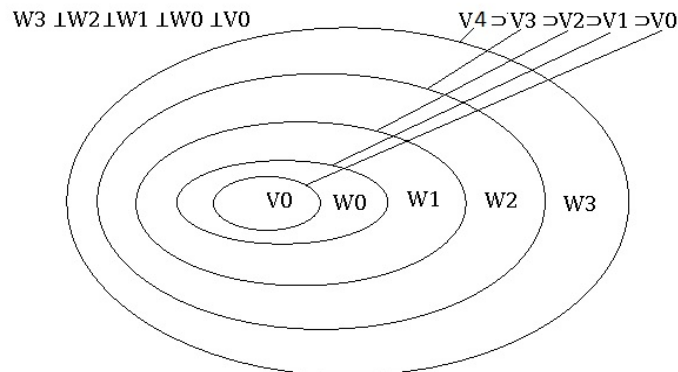
$$h_k = \begin{cases} \frac{1}{\sqrt{2}}, & \text{se } k = 0 \\ \frac{1}{\sqrt{2}}, & \text{se } k = 1 \\ 0, & \text{se } k \neq 0, 1 \end{cases} \quad (3.26)$$

e

$$g_k = \begin{cases} \frac{1}{\sqrt{2}}, & \text{se } k = 0 \\ -\frac{1}{\sqrt{2}}, & \text{se } k = 1 \\ 0, & \text{se } k \neq 0, 1 \end{cases} \quad (3.27)$$

A Figura 3.4 é uma ilustração das relações $\dots V_j \subset V_{j+1} \dots$ e $V_{j+1} = V_j \oplus W_j$ ⁵ sobre o espaço da análise de multiresolução para $j = 0, 1, 2, 3$. (IN; KIM, 2013; STARCK; MURTAGH; BIJAQUI, 1998; OLIVEIRA, 2007; ECKLEY, 2001)

Figura 3.4 – Os espaços da AMR



3.2.5 Transformada wavelet discreta(decimada)

A transformada de wavelet contínua é calculada fazendo translações e escalonamentos contínuos de uma função sobre um sinal. Esse processo demanda muito tempo e muitos recursos computacionais e ainda gera um grande volume de informações. Portanto, esta transformada não seria muito versátil. As wavelets discretas foram introduzidas para superar este obstáculo. Elas não são transladas nem escalonadas continuamente, mas sim em intervalos discretos. Isto

⁵ \oplus é o símbolo da soma direta, ou a soma ortogonal.

pode ser feito com uma pequena modificação na wavelet contínua da seguinte forma

$$\Psi_{a,b}(x) = \frac{1}{\sqrt{a}} \Psi\left(\frac{x-b}{a}\right) \implies \Psi_{j,k} = \frac{1}{a_0^j} \Psi\left(\frac{x - kb_0 a_0^j}{a_0^j}\right), \quad (3.28)$$

em que j e k são inteiros; $a_0 > 1$ é um parâmetro de dilatação fixo; b_0 é o fator de translação o qual depende do fator de dilatação.

Geralmente escolhe-se $a_0 = 2$ para que se tenha uma amostragem da frequência chamada de amostragem diádica e $b_0 = 1$ é escolhido para que a amostragem temporal também seja diádica. Isto resulta em:

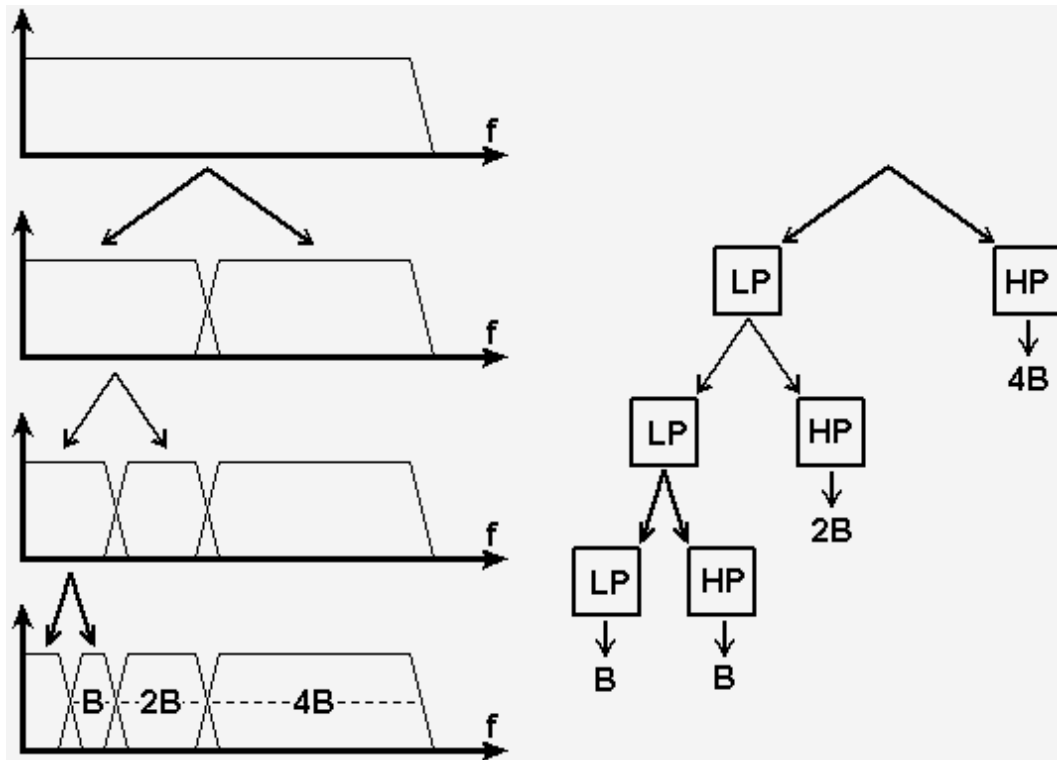
$$\Psi_{j,k}(x) = \sqrt{2^j} \Psi(2^j x - k). \quad (3.29)$$

O princípio usado no cálculo da Transformada de Wavelet Discreta é chamado "codificação em sub-bandas" (Subband Coding) e consiste em projetar filtros passa-alta (HPF) e passa-baixa (LPF) de tal modo que "quebre" o espectro do sinal (função) exatamente ao meio. As componentes resultantes da filtragem passa-alta, contêm as altas frequências, que já nos fornecem as informações com os mínimos detalhes, e os componentes de baixas frequências, resultantes da filtragem passa-baixa, que fornecem uma visão global destas frequências. Passa-se novamente este trecho do espectro, as baixas frequências, por outro par de filtros, um LPF e um HPF. Da mesma maneira supracitada, a parte do HPF fornece detalhes. Divide-se novamente a porção LPF, até considerar satisfatório o número de bandas criadas.

O sinal, ao passar por um banco de filtros digitais, resulta na TWD. O procedimento inicia-se com a passagem do sinal através de um filtro passa-baixa digital de meia banda com resposta ao impulso, h_n . Após o sinal passar pelo filtro passa-baixa de meia banda metade das amostras, a informação de alta frequência do sinal é eliminada. Assim, o sinal fica subamostrado por dois. A escala do sinal fica então multiplicada por dois. Nota-se que o LPF remove a informação de alta frequência do sinal, porém mantém sua escala inalterada. A escala é alterada apenas pelo processo de subamostragem. A resolução, por outro lado, é afetada pela filtragem, pois metade das frequências do sinal é eliminada. Como foi mencionado, a TDW emprega dois tipos de funções, chamadas de funções *escala* e funções *wavelet*, que estão, por sua vez, associadas com os LPFs e HPFs, respectivamente. A decomposição do sinal em bandas de frequências diferentes é simplesmente obtida pela filtragem sucessiva do sinal (OLIVEIRA, 2007).

A Figura 3.5 é uma ilustração de como funciona a codificação em sub-bandas.

Figura 3.5 – Codificação em sub-bandas



Fonte: Oliveira (2007)

Em uma AMR, dada uma função $f(x) \in L^2(\mathbb{R})$, sua discretização em um nível de resolução j pode ser representada como sendo a sua projeção nos espaços V_j , cuja expansão em série é dada por:

$$P_{V_j}f(x) = \sum_{k=-\infty}^{+\infty} c_{j,k}\phi_{j,k}(x), \quad (3.30)$$

sendo $\{c_{j,k}\}_{k \in \mathbb{Z}}$ denominados coeficientes de escala.

A projeção de f no espaço V_j pode ser dada pela soma das projeções de f em V_{j-1} e W_{j-1} , $P_{V_j}f = P_{V_{j-1}}f + P_{W_{j-1}}f$, produzindo a nova expansão:

$$P_{V_j}f(x) = \sum_{k=-\infty}^{+\infty} c_{j-,k}\phi_{j-,k}(x) + \sum_{k=-\infty}^{+\infty} d_{j-,k}\psi_{j-,k}(x). \quad (3.31)$$

Os coeficientes $\{d_{j-,k}\}_{k \in \mathbb{Z}}$ são denominados coeficientes wavelets (ou detalhes) e são responsáveis pela representação de informações complementares existentes entre as resoluções j e $j-1$. Dessa forma, os coeficientes de escala, $\{c_{j-,k}\}_{k \in \mathbb{Z}}$ e $\{c_{j-1,k}\}_{k \in \mathbb{Z}}$ representam as informações médias presentes em ambas resoluções.

Assim, para representar f em vários níveis de resolução partindo de um nível mais refinado J até um nível mais grosseiro j_0 , a decomposição de f é dada pela seguinte forma:

$$P_{V_j} f(x) = \sum_{k=-\infty}^{+\infty} c_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^{J-1} \sum_{k=-\infty}^{+\infty} d_{j-1,k} \psi_{j-1,k}(x). \quad (3.32)$$

Na prática, todos os espaços V_j e W_j tem dimensão finita $T_j = 2^j$ e, analogamente, os coeficientes de escala e wavelet podem ser representados como vetores de tamanho $T_j = 2^j$ conforme o nível j .

Muitas famílias de wavelets não possuem uma fórmula analítica, e são definidas apenas por meio de sistemas obtidos através de suas propriedades como a relação de escala, que associa cada família de funções $\{\phi_{j,k}(x)\}_{k \in \mathbb{Z}}$ e $\{\psi_{j,k}(x)\}_{k \in \mathbb{Z}}$ aos filtros h_k e g_k com largura finita $2N$, permitindo a obtenção dos valores dessas funções de forma recursiva

$$\phi(x) = \sqrt{2} \sum_{k=0}^{2N-1} h_k \phi(2x - k), \quad (3.33)$$

$$\psi(x) = \sqrt{2} \sum_{k=0}^{2N-1} g_k \phi(2x - k). \quad (3.34)$$

A determinação desses filtros também depende da capacidade das funções escala de representarem exatamente funções polinomiais até um certo grau N tornando nulos todos os seus coeficientes wavelet. Essa propriedade é denominada *propriedade dos momentos nulos*.

Considerando a relação $g_l = (-1)^l h_{2N-1-l}$, com $l = 0, 1, 2, \dots, 2N-1$, e considerando a ortogonalidade da base, as equações (4.31) e (4.32) induzem uma relação entre os coeficientes de escala $\{c_{j,k}\}_{k \in \mathbb{Z}}$ de um nível mais fino j com os coeficientes $\{c_{j-1,k}\}_{k \in \mathbb{Z}}$ e $\{d_{j-1,k}\}_{k \in \mathbb{Z}}$ no nível imediatamente mais grosseiro, envolvendo apenas os filtros h_k e g_k . Isso produz um algoritmo rápido e acurado, denominado algoritmo de Mallat ou algoritmo piramidal que define a transformada discreta de wavelet (TDW) rápida dada por

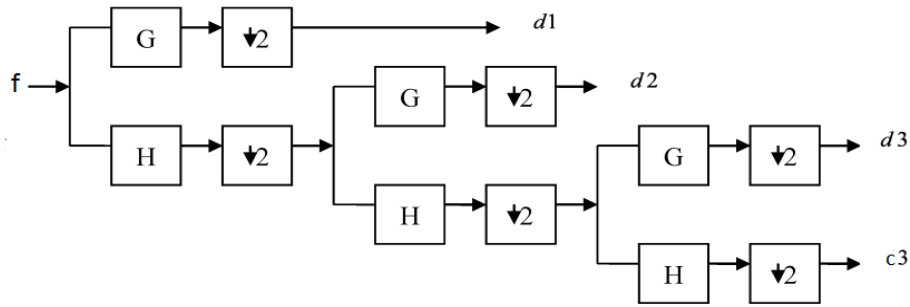
$$c_{j-1,k} = \sum_{l=1}^{2N-1} h_l c_{j,2k+l}, \quad (3.35)$$

$$d_{j-1,k} = \sum_{l=1}^{2N-1} g_l c_{j,2k+l}. \quad (3.36)$$

(KOZAKEVICIUS; BAYER, 2014)

A Figura 3.6 é uma representação de decomposição de uma função $f(x)$ usando a TWD.

Figura 3.6 – A árvore da TDW pelo algoritmo de Mallat em 3 níveis de resolução para uma função $f(x)$. A seta para baixo representa o operador de decimação.



Fonte: adaptado de (DAQROUQ et al., 2017).

3.2.6 Maximal overlap discrete wavelet transform (MODWT)

Maximal overlap discrete wavelet transform (MODWT) é uma transformada semelhante à TDW que transforma um sinal, da mesma forma que a TDW, em coeficientes de detalhes e de aproximação. Porém, a MODWT foi desenhada para superar algumas deficiências da TDW. Por exemplo, a MODWT pode ser definida para qualquer comprimento do sinal, enquanto a TDW é limitada aos comprimentos de sinais que são múltiplos de potências de 2. Ao contrário da TDW, MODWT não é sensível à escolha da origem, ou ponto de partida da transformação, do sinal. Finalmente, o número dos coeficientes de wavelet e de escala da TDW, N_j , diminui por um fator de 2 a cada aumento de nível de resolução, enquanto MODWT tem sempre o mesmo número de coeficientes em todos os níveis, isso, porque a MODWT mantém os valores subamostrados enquanto a TDW os descarta. Mais detalhes sobre as diferenças entre TDW e MODWT, e sobre a MODWT, podem ser encontrados em (PERCIVAL; WALDEN, 2006).

Para um sinal $X = \{X_t, t = 0, 1, \dots, n-1\}$, os elementos do j -ésimo nível dos coeficientes de wavelet, W_j , e de escala, V_j , de MODWT, podem ser determinados pelas seguintes equações:

$$V_{j,t} = \sum_{l=0}^{n-1} \tilde{h}_{j,l}^{\circ} X_{t-l} \text{ mod } n \quad j = 1, 2, \dots, L \quad (3.37)$$

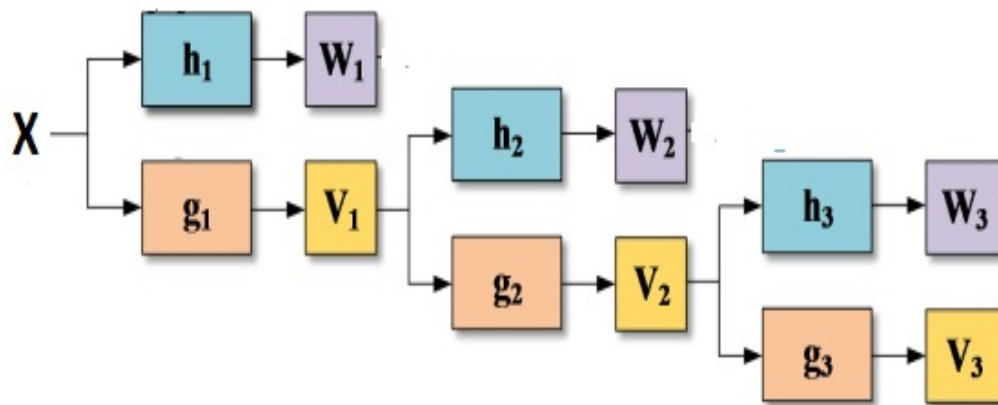
$$W_{j,t} = \sum_{l=0}^{n-1} \tilde{g}_{j,l}^{\circ} X_{t-l} \text{ mod } n, \quad (3.38)$$

em que, $W_{j,t}$ é o t -ésimo elemento dos coeficientes wavelet do j -ésimo nível da MODWT; $V_{j,t}$ é o t -ésimo elemento dos coeficientes de escala do j -ésimo nível da MODWT; $\{\tilde{h}_{j,l}^{\circ}\}$ e $\{\tilde{g}_{j,l}^{\circ}\}$ são, respectivamente, os filtros HP e LP de "periodizing" de $\{\tilde{h}_{j,l}\}$ e $\{\tilde{g}_{j,l}\}$ a um comprimento n ,

respectivamente; $\{\tilde{h}_{j,l} : \tilde{h}_{j,l} = h_{j,l}/2^{j/2}\}$ e $\{\tilde{g}_{j,l} : \tilde{g}_{j,l} = g_{j,l}/2^{j/2}\}$ são os filtros HP e LP do j -ésimo nível da MODWT; $\{h_{j,l}\}$ e $\{g_{j,l}\}$ são os filtros HP e LP do j -ésimo nível da TDW; L é o nível mais alto da decomposição. Os filtros são determinados dependendo da wavelet mãe da TDW. A MODWT, ao contrário da TDW, não é ortogonal (SEO; CHOI; CHOI, 2017).

A Figura 3.7 é uma ilustração dessa apresentação teórica da MODWT.

Figura 3.7 – MODWT em 3 níveis de resolução para um sinal X .



Fonte: (SEO; CHOI; CHOI, 2017).

3.3 Modelos SARIMA

A introdução de autocorrelação como um fenômeno que pode ser gerado por relações defasadas levou à proposição dos modelos auto-regressivos AR e os modelos auto-regressivos e médias móveis $ARMA$ para modelar séries estacionárias⁶, e o modelo auto-regressivo integrável de médias móveis $ARIMA$ que é uma generalização do $ARMA$ que incorpora uma classe ampla de séries não estacionárias. Esses modelos tornaram-se populares depois dos trabalhos de (BOX, 1970).

Um modelo autorregressivo de ordem p representado por $AR(p)$ é dado por

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t, \quad (3.39)$$

⁶ Uma série temporal é dita *estritamente estacionária* se as suas propriedades não são afetados pela escolha do ponto de origem no tempo. Isso significa que a distribuição de probabilidade conjunta das observações $y_t, y_{t+1}, \dots, y_{t+n}$ é a mesma distribuição de probabilidade conjunta das observações $y_{t+k}, y_{t+k+1}, \dots, y_{t+k+n}$. (MONTGOMERY; JENNINGS; KULAHCI, 2015)

em que Z_t é uma série estacionária e $\phi_1, \phi_2, \dots, \phi_p$ são constantes, sendo $\phi_p \neq 0$, e a_t é suposto um ruído branco, i.e, $cov(a_t, a_s) = 0, t \neq s$. Usando o operador translação para o passado (B), pode-se escrever o AR(p) como

$$\phi(B)Z_t = a_t, \quad (3.40)$$

em que

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p. \quad (3.41)$$

Um modelo de médias móveis de ordem q , descrito por MA(q), pode ser representado por

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (3.42)$$

em que q é o número de defasagens das médias móveis e $\theta_1, \theta_2, \dots, \theta_q$, com $\theta_q \neq 0$, são parâmetros e o a_t é suposto um ruído branco.

O modelo (??) pode ser representado na forma

$$Z_t = \theta(B)a_t, \quad (3.43)$$

em que

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \quad (3.44)$$

Z_t é um processo ARIMA(p,q) se

$$Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}. \quad (3.45)$$

Se d é um número inteiro e $Z_t = (1 - B)^d X_t$ é um processo ARMA(p, q), então X_t é um processo ARIMA(p, d, q). Isso significa que X_t satisfaz a equação de diferença da forma:

$$\phi(B)(1 - B)^d X_t = \theta(B)a_t, \quad (3.46)$$

em que a_t é um ruído branco⁷ com média zero e variância constante⁷.

Se d e D são números inteiros não negativos, então X_t é um processo ARIMA sazonal, SARIMA(p, d, q)x(P, D, Q)_s, com período s se a série diferenciada $Z_t = (1 - B)^d (1 - B^s)^D X_t$ é

⁷ Se uma série temporal é composta de observações não correlacionadas e tem variância constante, então ela é dita Ruído Branco (RB).

um processo *ARMA* definido por

$$\phi(B)\Phi(B^s)Z_t = \theta(B)\Theta(B^s)a_t, \quad a_t \sim RB(0, \sigma^2), \quad (3.47)$$

em que $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, $\Phi(B) = 1 - \Phi_1 B - \dots - \Phi_p B^p$, $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ e $\Theta(B) = 1 + \Theta_1 B + \dots + \Theta_Q B^Q$ (BROCKWELL; DAVIS, 2016).

3.4 Programação matemática

Programação matemática refere-se a um conjunto de procedimentos estabelecidos para lidar com a análise de problemas de otimização. Os problemas de otimização são, geralmente, aqueles nos quais o tomador de decisão deseja otimizar (minimizar ou maximizar) uma medida de satisfação (ou mais de uma), chamada *função objetivo*, selecionando valores para um conjunto de variáveis, chamadas de *variáveis de decisão*. Essas variáveis podem ser restritas, ou não, por condições externas ao problema. Essas condições são chamadas de *restrições* (MCCARL; SPREEN, 1997).

Um problema geral de Programação Matemática (PM) pode ser representado da seguinte forma:

$$\text{minimize} \quad f(x) \quad (3.48)$$

$$\text{sujeito a} \quad h_i(X) = 0, \quad i = 1, 2, \dots, m \quad (3.49)$$

$$g_j \leq 0, \quad j = 1, 2, \dots, r \quad (3.50)$$

$$x \in S, \quad (3.51)$$

em que X é um vetor n -dimensional de variáveis de decisão desconhecidas, $X = (x_1, x_2, \dots, x_n)$, as funções $f, g_1, g_2, \dots, g_m, h_1, h_2, \dots, h_r$ são funções definidas sobre R^n e S é um subespaço de R^n . A função f é, normalmente, chamada a *função objetivo*. Cada uma das restrições $h_i(X) = 0$ para $i = 1, 2, \dots, m$ é chamada *restrição de igualdade*, e cada uma das restrições $g_j \leq 0$ para $j = 1, 2, \dots, r$ é chamada uma *restrição de desigualdade*. Escusado será dizer que o problema de otimização pode ser escrito como um problema de maximização (LUENBERGER; YE et al., 1984).

Em geral, os problemas de otimização são divididos em classes de acordo com as características da função objetivo e das restrições.

3.4.1 Programação não linear

A Programação não linear é o termo usado para descrever um problema de otimização no qual a função objetivo ou as funções restrição são não lineares, e não sendo identificadas como convexas. De fato, não há um método que garante a melhor solução para um problema de programação não linear. Portanto, os métodos para solução deste tipo de problemas adotam várias abordagens das quais cada uma envolve algum tipo de comprometimento (BOYD; VANDENBERGHE, 2004).

3.4.2 Programação matemática multiobjetivo

A Programação matemática multiobjetivo se refere a qualquer programação matemática envolvendo mais de uma função objetivo. Ela implica que o tomador de decisão está respondendo a múltiplos objetivos entre os quais, geralmente, existe um conflito de forma a não permitir que todas as objetivas cheguem, simultaneamente, aos seus respectivos níveis ótimos (MCCARL; SPREEN, 1997).

4 MÉTODOS E MATERIAIS

4.1 Metodologia

O método a ser aplicado consiste nos seguintes passos: (1) Determinar a série a ser analisada e definir o período da análise; (2) Realizar a decomposição da série utilizando a transformada *Maximum Overlapping Discrete Wavelet Transform* (MODWT) com momentos nulos e nível de resolução adequados; (3) Modelar cada uma das componentes wavelet, obtidas no passo anterior, por meio de um modelo SARIMA; (4) Obter previsões em h passos a frente para cada componente wavelet utilizando o respectivo modelo SARIMA; (5) Obter a previsão da série de vazão na forma de combinação linear ótima das previsões das componentes de wavelet, obtidas no passo anterior, de modo a minimizar o erro de previsão ¹ utilizando o método dos mínimos quadrados para determinar os pesos ρ_i , $i = 1, 2, \dots, p$ das previsões das componentes wavelet e o método de programação matemática multiobjetivo para obter os parâmetros adaptativos α e β na seguinte equação:

$$\hat{Y}_{CL} = [\hat{Y}_{Sp}\rho_1 + \sum_{i=1}^p \hat{Y}_{di}\rho_{i+1}]\alpha + \beta, \quad (4.1)$$

em que

\hat{Y}_{CL} : A previsão linearmente combinada da série de vazão analisada.

\hat{Y}_{Sp} : A previsão da componente wavelet de aproximação no nível p de resolução.

$\hat{Y}_{d,i}$: A previsão da componente wavelet de detalhe no nível i , $i = 1, 2, \dots, p$ de resolução.

ρ_1 : Peso da previsão da componente de aproximação no nível p de resolução.

ρ_{i+1} : Peso da previsão da componente wavelet de detalhe no nível i de resolução.

α : Parâmetro adaptativo multiplicativo.

β : Parâmetro adaptativo aditivo.

4.2 Material

A metodologia proposta foi aplicada à série de vazão natural da usina hidroelétrica de Samuel no rio Jamari, no município de Candeias do Jamari no estado de Rondônia.

Vazões naturais de afluentes a um determinado aproveitamento são aquelas que ocorriam em uma seção do rio caso não houvesse a operação de reservatórios a montante, nem a

¹ Erro de previsão é a diferença entre o observado Y e o previsto \hat{Y} .

vazão evaporada pelos lagos artificiais, nem mesmo retiradas de água consumidas com abastecimento e irrigação a montante. Em síntese, são as vazões providas pela própria natureza (ONS).

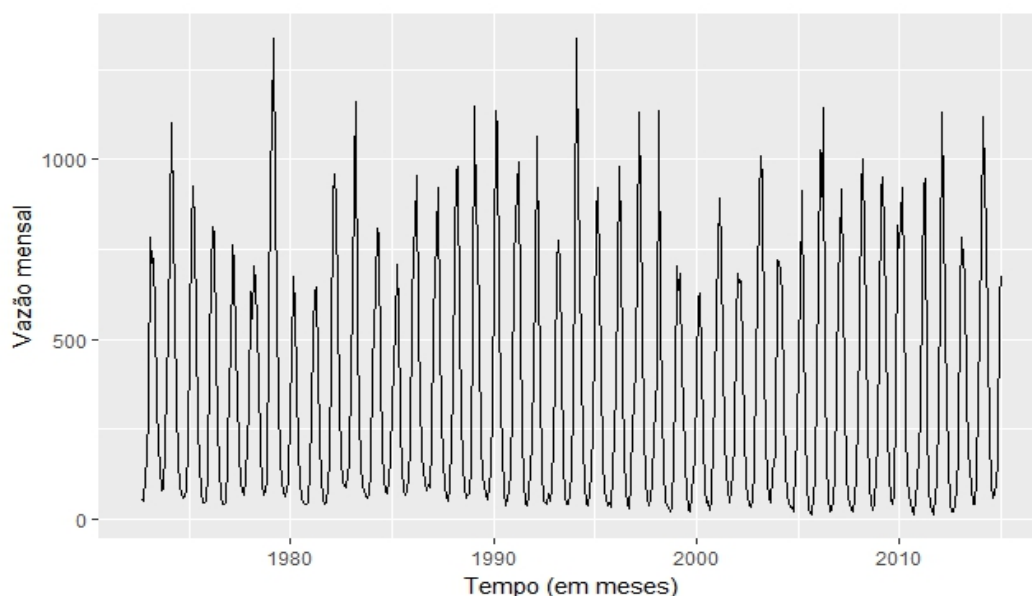
A relevância das séries de vazão natural de afluentes consiste, em um lado, no fato dessas estarem diretamente ligadas aos níveis de água nas barragens e nos rios que formam recursos hídricos para abastecimento urbano e agrícola, além de outros aspectos da atividade econômica e social, e em outro lado por serem caracterizadas pela alta complexidade e por serem difíceis de modelar.

5 RESULTADOS E DISCUSSÃO

5.1 Resultados

Com o objetivo de aplicar a metodologia proposta, foi realizado um experimento computacional incluindo análise de wavelet, modelagem SARIMA e combinação de previsões. Toda análise foi feita no programa R (R Core Team, 2017). O experimento foi aplicado à série de vazão natural mensal da usina hidroelétrica de Samuel, medida em metros cúbicos por segundo (m^3/s), que foi obtida do website do Operador Nacional do Sistema Elétrico (ONS, 2017) no dia primeiro de maio de 2017. A série utilizada contém 510 observações que representam o período do mês de setembro de 1972 ao mês de fevereiro de 2015 e está representada na Figura 5.1. O gráfico mostra que a série tem média constante e sazonalidade anual.

Figura 5.1 – A série histórica de vazão natural da barragem de Samuel, o período de setembro de 1972 a fevereiro de 2015, medida em metros cúbicos por segundo (m^3/s)



Em primeiro passo, foi realizada a transformação de toda a série em componentes de wavelet e, em seguida, a série original e cada uma das componentes wavelet (série de dados) foi dividida em duas amostras: amostra de treino (i.e., a amostra utilizada para ajustar os modelos) constituída pelos 470 primeiros dados que representem o período de mês de março de 1973 ao mês de abril de 2013, e amostra de teste (i.e., amostra usada para verificar o desempenho do modelo) constituída pelos 40 dados restantes que representam o período de mês de maio de 2013 ao mês de dezembro de 2015. No segundo passo, cada uma das amostras de treino foi ajustada a um modelo SARIMA, utilizando o Critério de Informação de Akaike (AIC) para

selecionar o melhor modelo, e por cada um desses foram geradas previsões 40 passos a frente para a respectiva série. No terceiro passo foi realizada a combinação das previsões geradas pelos modelos ajustados às componentes wavelet, e para efeito de comparação, foram calculadas as estatísticas RMSE ¹: $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$, em que: y_i é o valor observado, \hat{y}_i é o valor previsto, MAPE ² para as previsões combinadas e para as previsões da série original obtidas pelo modelos SARIMA.

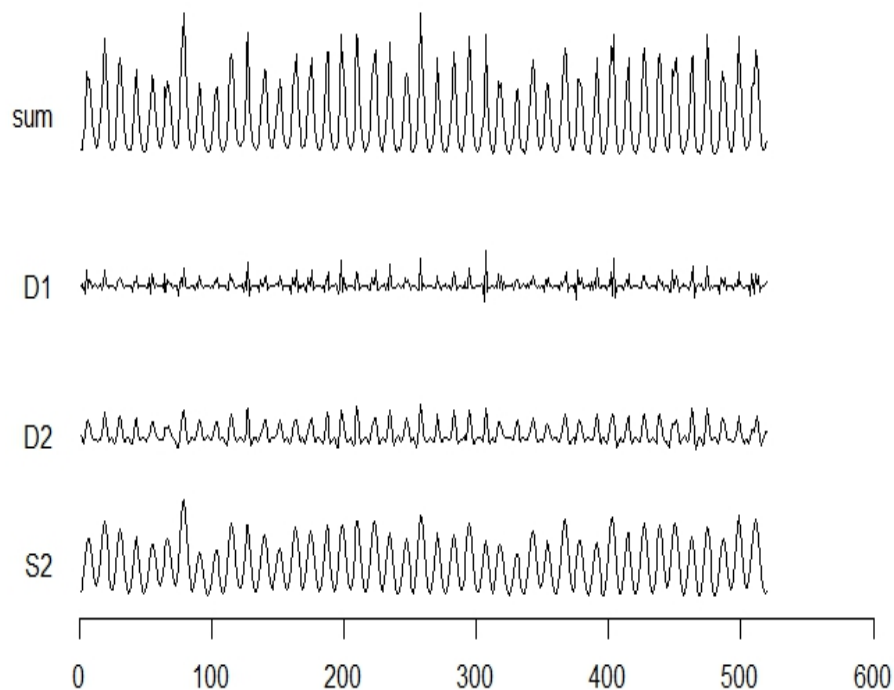
5.1.1 Decomposição wavelet

Neste estágio foi realizada a transformação de toda a série em componentes wavelet utilizando a transformada não decimada de *Maximum Overlap Discrete Wavelet Transform* (MODWT) usando o pacote wmtsa. Foram testadas várias bases de Daubechies: haar, d2, d4, d6, d8, d10 e d20; e de symmlet: s2, s4, s8, s10 e s20. Os melhores resultados foram obtidos com a implementação da base haar e a realização da transformação em dois níveis de resolução, gerando assim uma componente de aproximação (S2) e uma de detalhe (d2), ambas no nível 2 de resolução, e uma componente de detalhe (d1) no nível 1 de resolução. As três componentes estão representadas, junto com a série original, na figura 5.2. Os resíduos dos modelos ajustados às componentes wavelets obtidas com o uso das outras bases wavelets apresentaram altas e persistentes correlações comprometendo a validade dos modelos.

¹ *Root Mean Square Error*

² *Mean Absolute Percentage Error*: $MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100$ e MAE ³: $MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$.

Figura 5.2 – A série de vazão natural(denominada "sum") e as componentes de wavelet.



5.1.2 Modelagem SARIMA

As amostras de treino da série original e das componentes wavelet foram todas modeladas com modelos SARIMA. A Figura 5.3 e a Figura 5.4 apresentam, respectivamente, o gráfico da amostra de treino da série original e o seu correlograma (Os gráficos das FAC⁴ e FACP⁵). Todos indicam uma sazonalidade anual.

⁴ Função de Autocorrelação

⁵ Função de Autocorrelação Parcial

Figura 5.3 – Amostra de treino da série de vazão (de março de 1973 a abril de 2013).

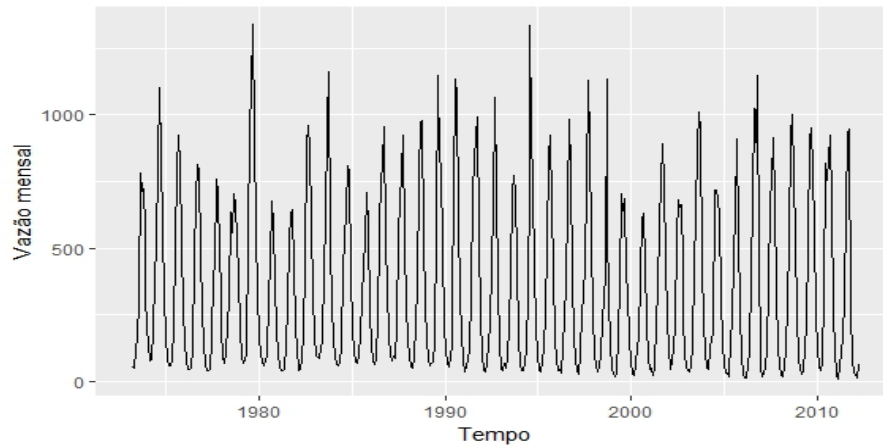
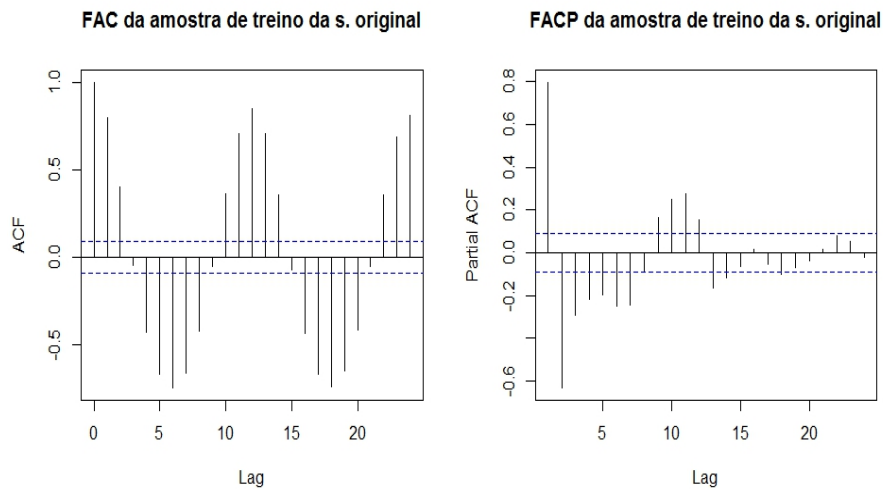
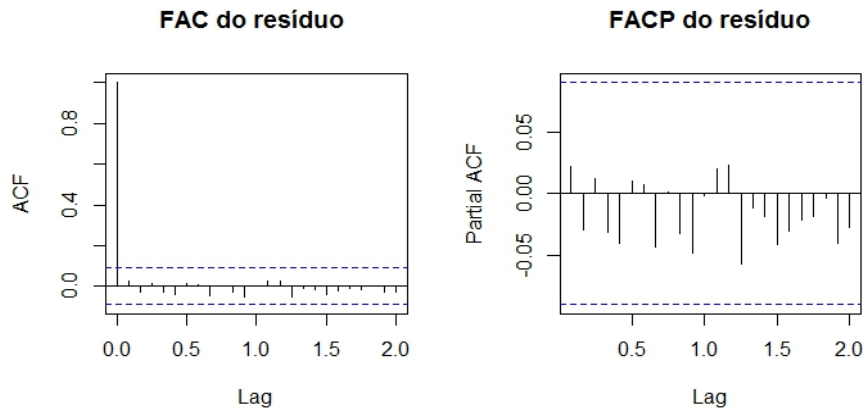


Figura 5.4 – As funções FAC e FACP da amostra de treino da série de vazão (de março de 1973 a abril de 2013).



A essa série foi ajustado um modelo $SARIMA(1,0,0) \times (1,1,1)_{12}$. Os coeficientes do modelo podem ser encontrados na Tabela 5.1 junto com os desvios padrão respectivos. A Figura 5.5 representa o correlograma do resíduo do modelo que mostra ser, ao nível de confiança de 90%, um ruído branco (i.e., com média zero e variância constante). Isso significa que, a este nível de confiança, toda a estrutura de autocorrelação dentro da amostra foi modelada.

Figura 5.5 – As funções FAC e FACP do resíduo do modelo SARIMA ajustado à amostra de treino da série original.



Às amostras de treino das componentes de wavelet, representadas graficamente junto com seus respectivos correlogramas nas Figuras 5.6, 5.7 e 5.8 foram ajustados os modelos $\text{SARIMA}(1, 0, 1) \times (0, 1, 1)_{12}$ para d1, $\text{SARIMA}(1, 0, 3) \times (1, 1, 1)_{12}$ para d2 e $\text{SARIMA}(1, 0, 3) \times (1, 0, 1)_{12}$ para s2. Os coeficientes dos modelos podem ser encontrados na Tabela 5.1 juntamente com os desvios padrão respectivos. As Figuras 5.9, 5.10 e 5.11 apresentam os correlogramas dos resíduos desses modelos. Os correlogramas mostram que todos os coeficientes de correlação dos resíduos dos modelos ajustados, até 24 lags, estão contidos no intervalo de confiança para a hipótese de não significância desses. Isso significa que ao nível de confiança de, aproximadamente 90%, os resíduos são ruídos brancos.

Tabela 5.1 – Os coeficientes dos modelos dos modelos SARIMA ajustados às séries.

Série	Modelo	Coeficiente	Desvio padrão
Vazão	SARIMA(1,0,0)x(1,1,1) ₁₂	$\phi_1 = 0,5533$	0,0391
		$\Phi_1 = 0,1150$	0,0543
		$\Theta_1 = -0,9267$	0,0317
d1	SARIMA(1,0,1)x(1,1,1) ₁₂	$\phi_1 = 0,554085$	0,040660
		$\theta_1 = -1,0000$	0,010221
		$\Phi_1 = 0,115227$	0,056658
		$\Theta_1 = -0,912718$	0,033468
d2	SARIMA(1,0,3)x(0,1,1) ₁₂	$\phi_1 = 0,565036$	0,039383
		$\theta_1 = 0,971260$	0,019532
		$\theta_2 = -0,999140$	0,023026
		$\theta_3 = -0,970857$	0,019198
		$\Theta_1 = -0,867674$	0,037549
S2	SARIMA(1,0,3)x(1,0,2) ₁₂	$\phi_1 = 0,552480$	0,039774
		$\theta_1 = 1,006342$	0,027091
		$\theta_2 = 1,011478$	0,018283
		$\theta_3 = 0,994803$	0,031104
		$\Phi_1 = -0,844906$	0,104064
		$\Theta_1 = 0,049054$	0,094633
		$\Theta_2 = -0,812333$	0,074761

Figura 5.6 – Componente de aproximação no nível 2 de resolução, S2, e seu correlograma.

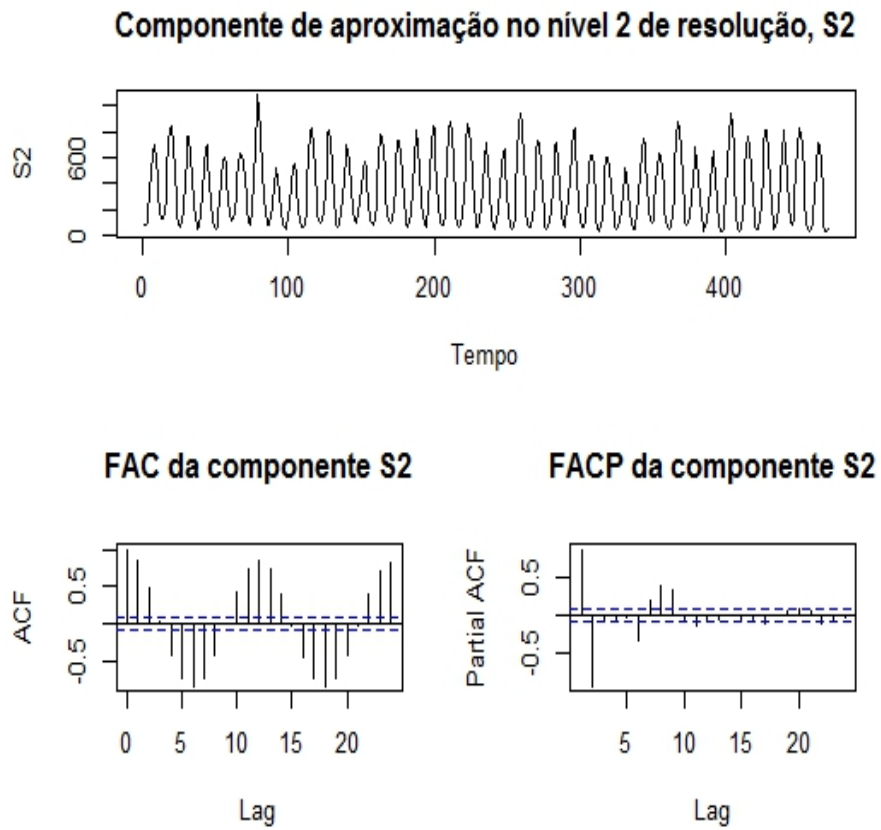


Figura 5.7 – Componente de detalhe no nível 1 de resolução, d1, e seu correlograma.

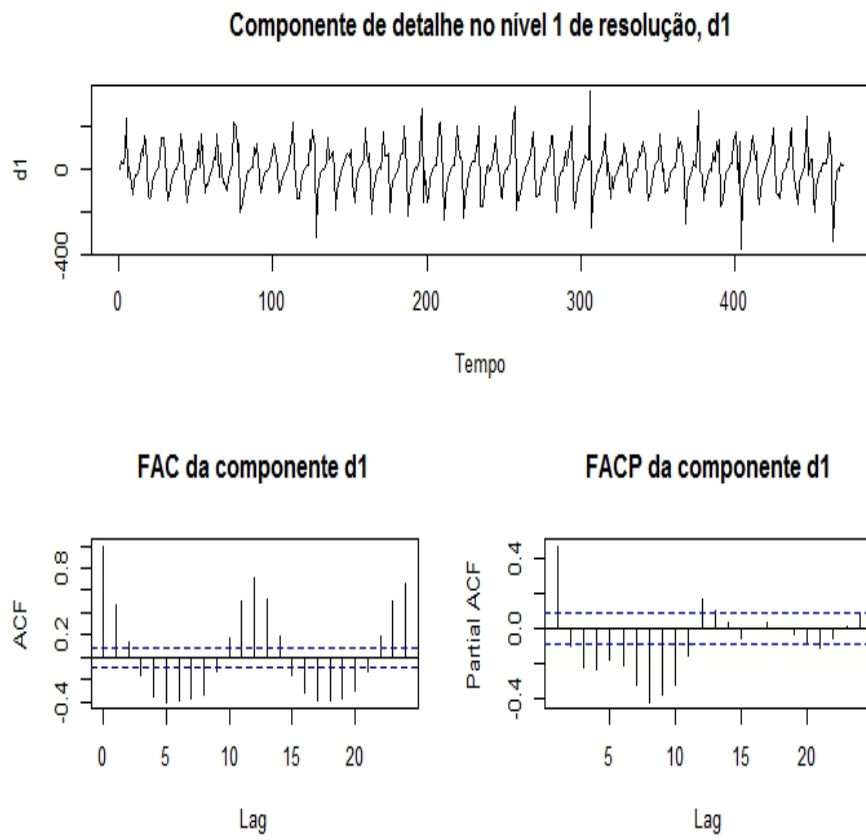


Figura 5.8 – Componente de detalhe no nível 2 de resolução, d2, e seu correlograma.

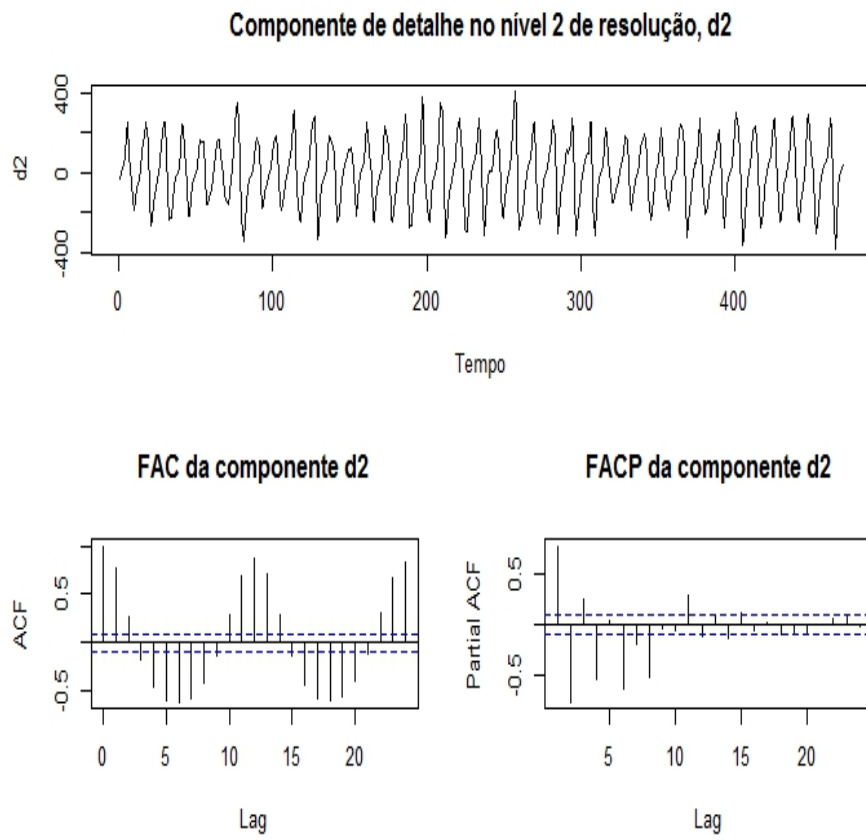


Figura 5.9 – Correlograma do resíduo do modelo SARIMA ajustado à componente de aproximação S2.

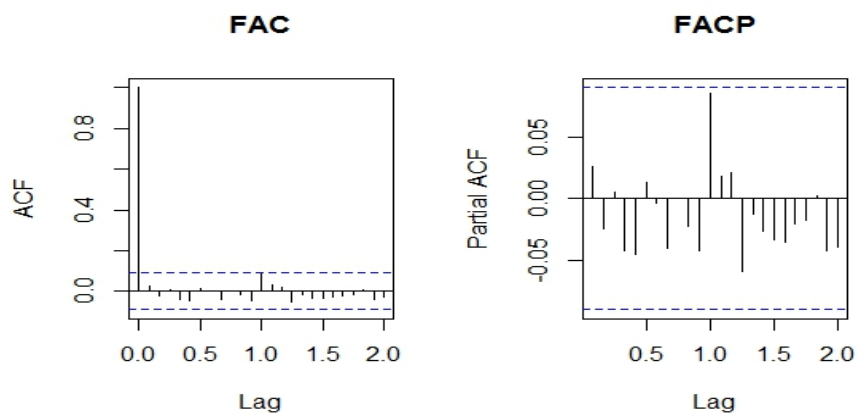


Figura 5.10 – Correlograma do resíduo do modelo SARIMA ajustado à componente de detalhe d1.

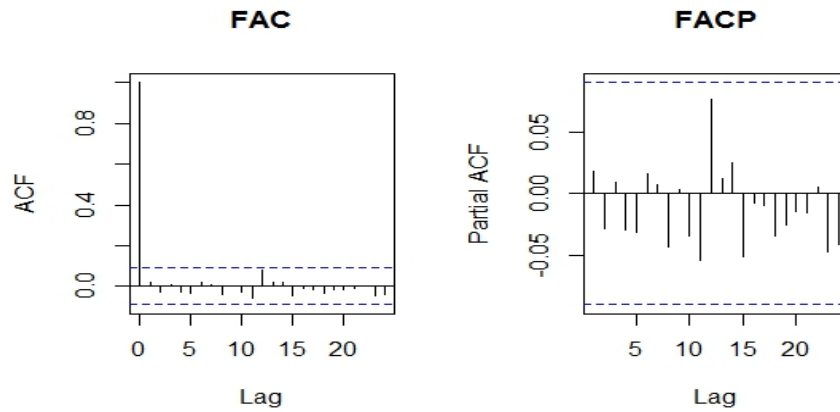
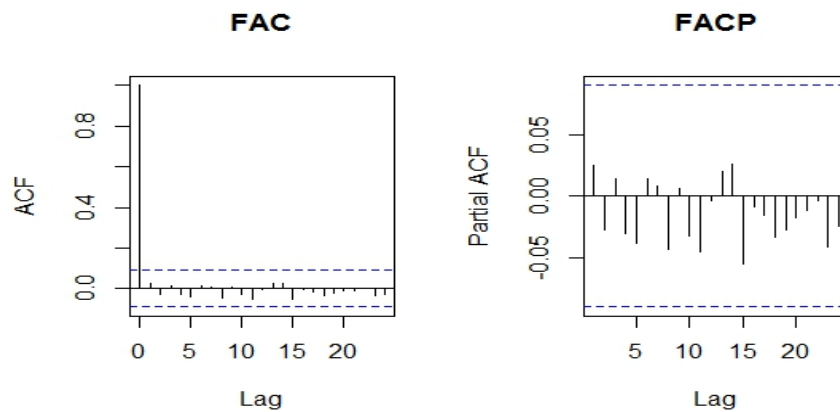


Figura 5.11 – Correlograma do resíduo do modelo SARIMA ajustado à componente de detalhe d2.



As Figuras 5.12, 5.13, 5.14 e 5.15 apresentam os gráficos das previsões dentro da amostra para a série original e para as componentes de wavelet. Os gráficos mostram que os modelos desempenharam melhor para as séries com mais regularidade. O modelo ajustado à componente d1 não conseguiu desempenhar uma boa previsão dos picos devido à alta irregularidade desses.

Figura 5.12 – O gráfico da amostra de treino da série de vazão natural e a sua previsão com ARIMA.

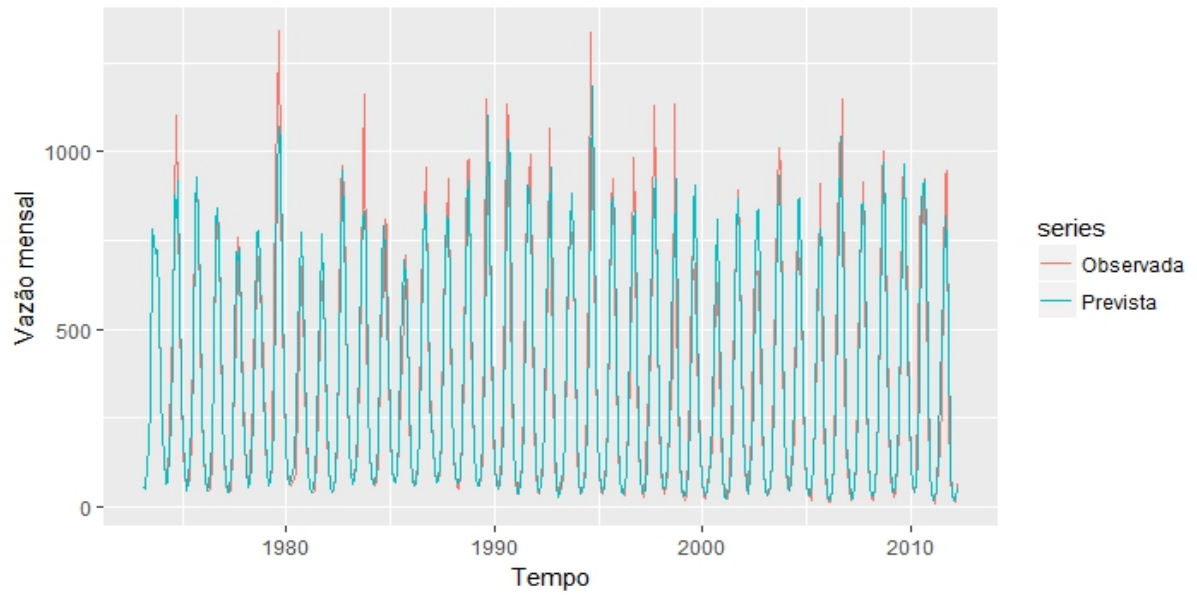


Figura 5.13 – O gráfico da amostra de treino da componente de aproximação S2 e a sua previsão.

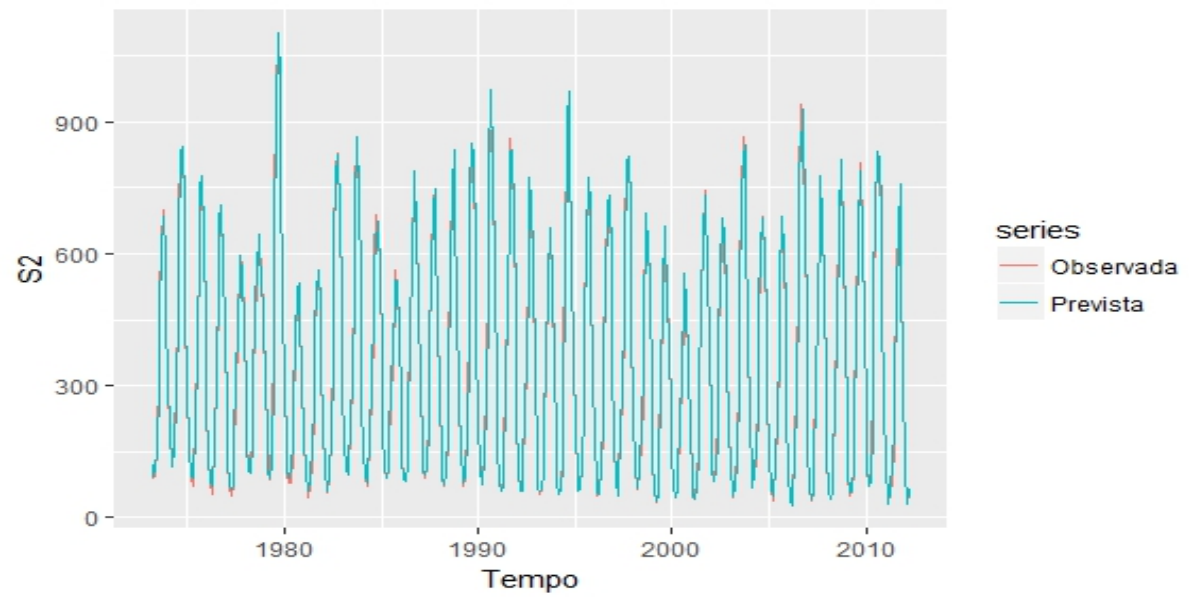


Figura 5.14 – O gráfico da amostra de treino da componente de detalhe d1 e a sua previsão.

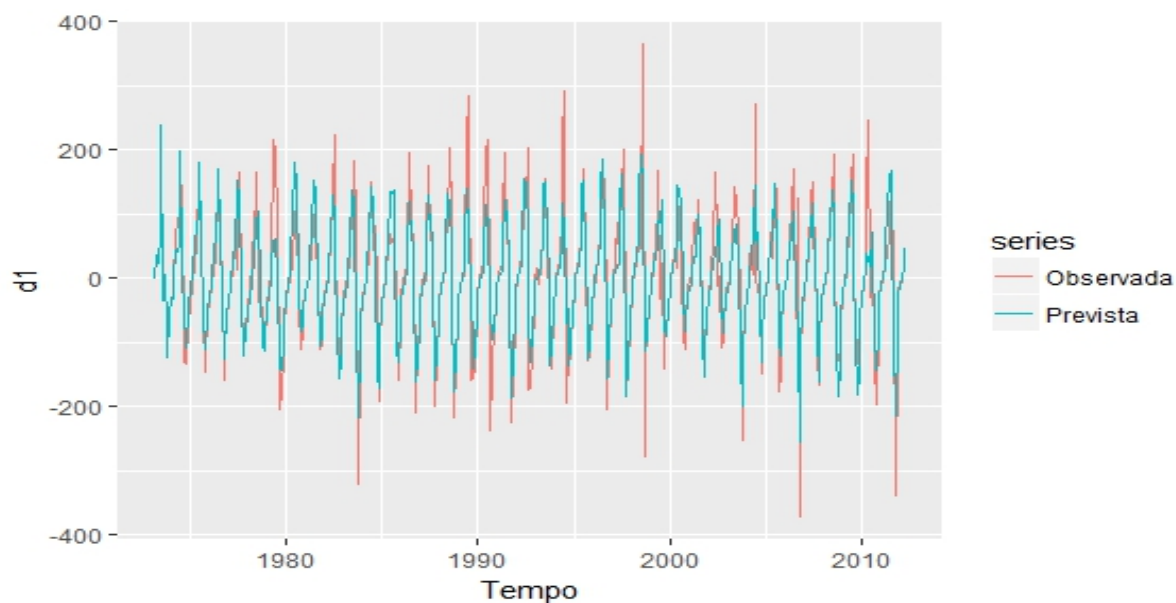
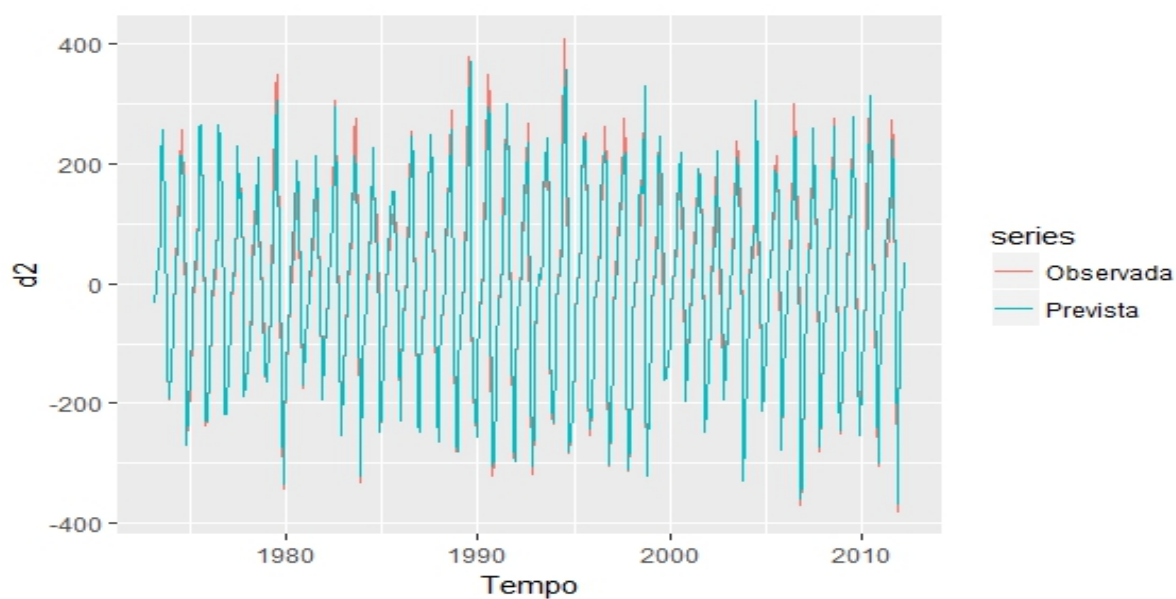


Figura 5.15 – O gráfico da amostra de treino da componente de detalhe d2 e a sua previsão.



Para cada um dos modelos SARIMA ajustados à série original e as componentes wavelet, foram geradas previsões 40 passos a frente para a respectiva série de dados, com o objetivo de combinar as previsões das componentes wavelet e, por fim, comparar o desempenho dos dois métodos de previsão (SARIMA e wavelet-SARIMA combinado). Os gráficos das amostras de teste e suas respectivas previsões podem ser encontradas nas Figuras 5.16, 5.17, 5.18 e 5.19.

Figura 5.16 – O gráfico da amostra de teste da série original e a sua previsão pelo SARIMA.

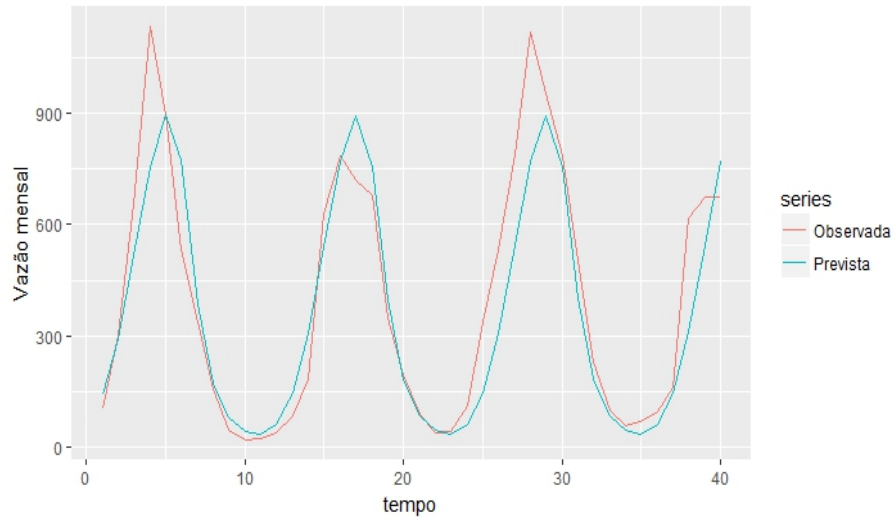


Figura 5.17 – O gráfico da amostra de teste da componente de aproximação S2 e a sua previsão.

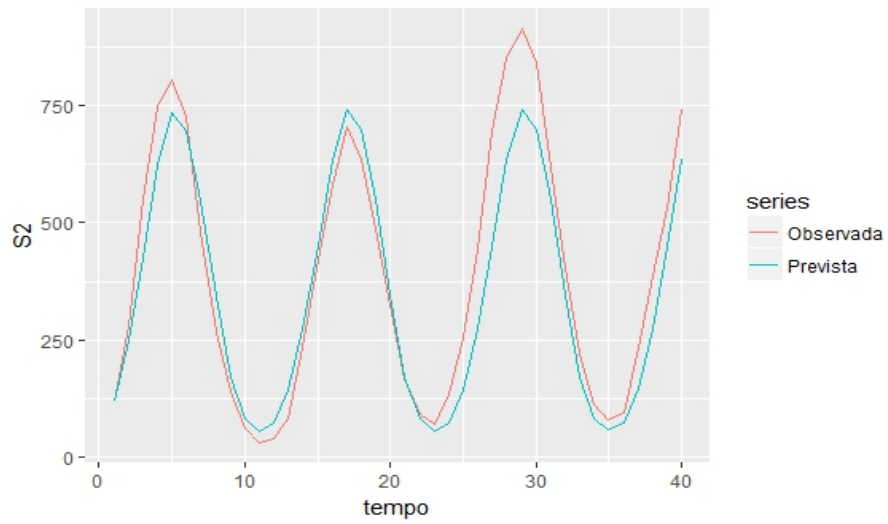


Figura 5.18 – O gráfico da amostra de teste da componente de detalhe d1 e a sua previsão.

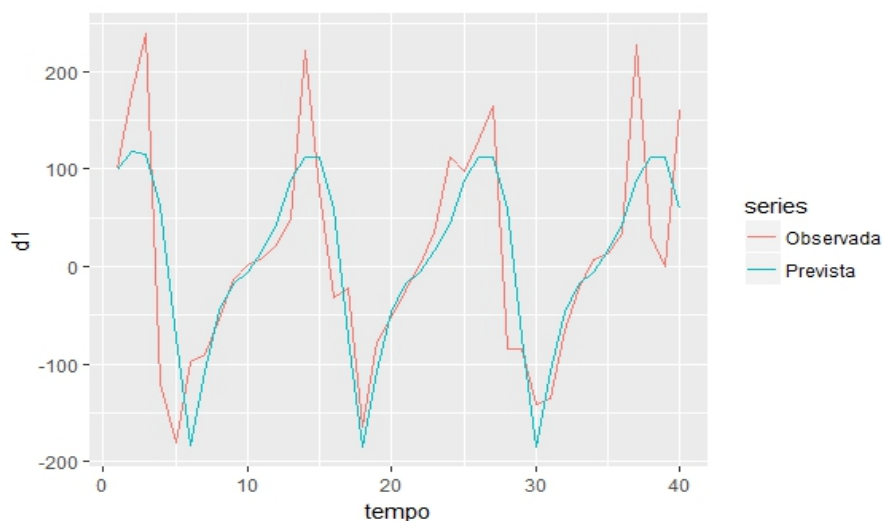
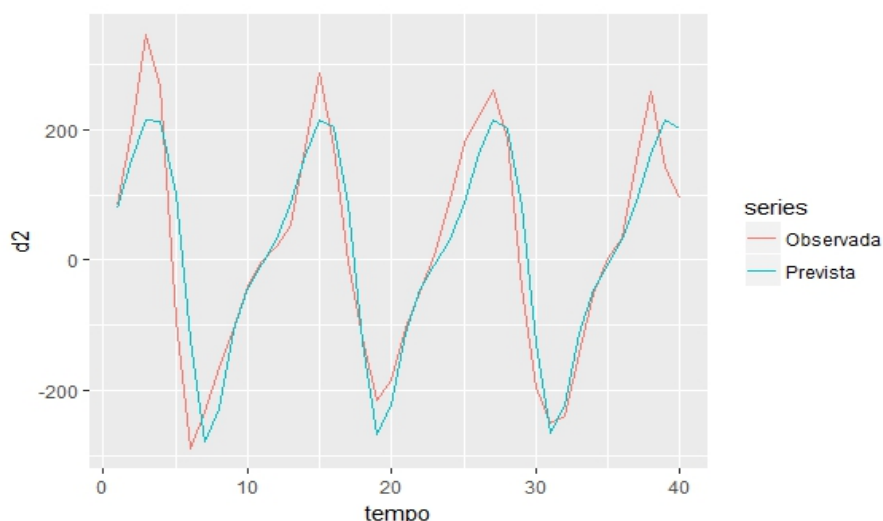


Figura 5.19 – O gráfico da amostra de teste da componente de detalhe d2 e a sua previsão.



5.1.3 Combinação de previsões

Apos a modelagem da série original e das componentes wavelet usando modelos SARIMA, foram realizados os cálculos para determinar os pesos ótimos ρ_i , $i = 1, 2, 3$ e os parâmetros adaptativos α e β da fórmula 4.1 para combinar linearmente as previsões das componentes wavelet. Os cálculos foram realizados em duas etapas. Na primeira etapa foram determinados os pesos ótimos da combinação linear empregando o método dos mínimos quadrados, e na segunda etapa foram determinados os valores dos parâmetros adaptativos, α e β . Nessa etapa foi formulado um problema de programação matemática não linear multiobjetivo que foi resolvido implementando o pacote (mopsocd) do programa R Core Team (2017), no qual foram utilizadas

as estatísticas RMSE , MAPE e MAE da amostra do treino como funções objetivo. Os valores dos pesos ótimos e dos parâmetros adaptativos podem ser encontrados na tabela 5.2. Os valores das previsões fora da amostra de 40 passos a frente para a série de vazão, utilizando o modelo ARIMA e a previsão combinada, podem ser encontrados na tabela 5.2 . A Figura 5.20 é uma representação desses valores e da amostra de teste da série de vazão.

Tabela 5.2 – Os valores dos pesos ótimos e dos parâmetros adaptativos da combinação das previsões das componentes wavelet.

Pesos de combinação	
Componente wavelet	Peso
ρ_1	0,9900822
ρ_2	-0,9893648
ρ_3	1,0548765
Parâmetros adaptativos	
α	0,9758786
β	-9,0670230

Para efeito de comparação, a tabela 5.3 apresenta os valores das estatísticas RMSE, MAPE e MAE para a previsão dentro e fora da amostra para os dois métodos de previsão.

Tabela 5.3 – Tabela de comparação entre os métodos usado as estatísticas RMSE, MAPE e MAE.

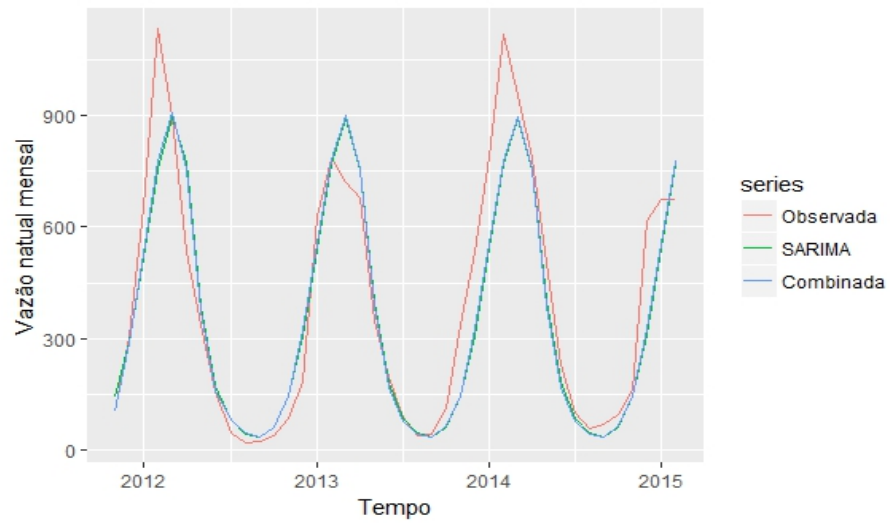
	Amostra de treino			Amostra de teste		
	RMSE	MAPE	MAE	RMSE	MAPE	MAE
SARIMA	100,58	20,95%	61,06	132,78	29,53%	88,16
Método proposto	104,38	29,41%	67,63	137,34	32,86%	93,76

Apesar de todos os trabalhos citados no capítulo 2 que analisaram séries de vazão de afluentes terem obtido melhores resultados empregando wavelet, os resultados deste trabalho mostraram superioridade do modelo SARIMA puro. Perceba-se que todos esses trabalhos utilizaram a transformada de wavelet decimada.

Tabela 5.4 – As previsões da amostra de teste (40 passos a frente) da série de vazão, pelos SARIMA e Wavelet-SARIMA.

Previsões de 1 a 20		Previsões de 21 a 40	
SARIMA	Wavelet-SARIM	SARIMA	Wavelet-SARIM
142,9151	130,5929	84,84396	36,26794
292,6062	295,0791	45,90885	17,56063
517,8198	523,0241	34,85673	13,59459
755,9718	740,48	62,88768	45,87149
896,7762	864,5414	146,9573	136,2451
775,7205	735,7161	308,2823	302,716
386,9039	350,745	543,9944	525,6035
171,6381	149,5237	769,6954	737,4737
79,7879	63,87114	892,7938	859,4889
43,88256	31,97719	753,7771	726,5088
32,56594	18,49435	401,7625	364,6507
62,86164	54,23831	182,9576	143,0322
146,5408	146,1986	85,42515	57,15646
306,6662	331,3849	46,14178	28,36525
541,2959	566,5537	35,12006	17,65679
768,2805	784,0478	62,89067	51,93752
893,2044	895,8954	147,0051	145,0242
756,0395	734,4087	308,468	327,6156
400,2306	349,4749	544,3046	560,731
181,7906	119,4648	769,858	776,2578

Figura 5.20 – Amostra de teste e a sua previsão pelo modelo SARIMA e pela combinação das previsões das componentes wavelet.



6 CONCLUSÃO

O método proposto, WAVLET-SARIMA, teve desempenho inferior ao desempenho do modelo SARIMA puro. O melhor resultado foi obtido com o uso da base Haar e a decomposição em dois níveis de resolução. Será feito um trabalho futuro de comparação, utilizando as transformadas DWT e MODWT, aplicado na mesma série analisada neste trabalho.

REFERÊNCIAS

- BATES, J. M.; GRANGER, C. W. The combination of forecasts. **Or**, JSTOR, p. 451–468, 1969.
- BOX, G. Ep, and jenkins, g. M. **"Time Series Analysis, forecasting, and Control,"** San Francisco: **Holden-Day**, 1970.
- BOYD, S.; VANDENBERGHE, L. **Convex optimization**. [S.l.]: Cambridge university press, 2004.
- BROCKWELL, P. J.; DAVIS, R. A. **Introduction to time series and forecasting**. [S.l.]: springer, 2016.
- CHAU, K.; WU, C. A hybrid model coupled with singular spectrum analysis for daily rainfall prediction. **Journal of Hydroinformatics**, IWA Publishing, v. 12, n. 4, p. 458–473, 2010.
- CLEMEN, R. T. Combining forecasts: A review and annotated bibliography. **International journal of forecasting**, Elsevier, v. 5, n. 4, p. 559–583, 1989.
- CLEMENTS, M. P. Sir clive wj granger's contributions to forecasting. **European Journal of Pure and Applied Mathematics**, v. 10, n. 1, p. 30–57, 2017.
- DAQROUQ, K. et al. Off-line handwritten signature recognition by wavelet entropy and neural network. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 19, n. 6, p. 252, 2017.
- ECKLEY, I. A. **Wavelet methods for time series and spatial data**. Tese (Doutorado) — University of Bristol, 2001.
- ELLIOTT, G. Averaging and the optimal combination of forecasts. **University of California, San Diego**, 2011.
- GOSWAMI, J. C.; CHAN, A. K. **Fundamentals of wavelets: theory, algorithms, and applications**. [S.l.]: John Wiley & Sons, 2011. v. 233.
- HANSEN, B. E. Least-squares forecast averaging. **Journal of Econometrics**, Elsevier, v. 146, n. 2, p. 342–350, 2008.
- IN, F.; KIM, S. **An introduction to wavelet theory in finance: a wavelet multiscale approach**. [S.l.]: World scientific, 2013.
- JOO, T. W.; KIM, S. B. Time series forecasting based on wavelet filtering. **Expert Systems with Applications**, Elsevier, v. 42, n. 8, p. 3868–3874, 2015.
- JÚNIOR, L. A. T. et al. Combinação linear wavelet sarima-rna com estágios múltiplos na previsão de séries temporais. **Blucher Marine Engineering Proceedings**, Blucher, v. 2, n. 1, p. 295–307, 2016.
- JÚNIOR, L. A. T. et al. Box-jenkins modeling, wavelet decomposition and mathematical programming in time series forecasting. **Revista Pesquisa Naval, Brasília-DF**, n. 27, p. 74–81, 2015.
- KHANDELWAL, I.; ADHIKARI, R.; VERMA, G. Time series forecasting using hybrid arima and ann models based on dwt decomposition. **Procedia Computer Science**, Elsevier, v. 48, p. 173–179, 2015.

- KOZAKEVICIUS, A. d. J.; BAYER, F. M. Filtragem de sinais via limiarização de coeficientes wavelet. **Ciência e Natura**, Universidade Federal de Santa Maria, v. 36, 2014.
- KRIECHBAUMER, T. et al. An improved wavelet–arima approach for forecasting metal prices. **Resources Policy**, Elsevier, v. 39, p. 32–41, 2014.
- LUENBERGER, D. G.; YE, Y. et al. **Linear and nonlinear programming**. [S.l.]: Springer, 1984. v. 2.
- MALLAT, S. G. A theory for multiresolution signal decomposition: the wavelet representation. **IEEE transactions on pattern analysis and machine intelligence**, Ieee, v. 11, n. 7, p. 674–693, 1989.
- MANCUSO, A. C. B.; WERNER, L. Review of combining forecasts approaches. **Independent Journal of Management & Production**, Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, v. 4, n. 1, 2013.
- MCCARL, B. A.; SPREEN, T. H. Applied mathematical programming using algebraic systems. **Cambridge, MA**, 1997.
- MEYER, Y. Algorithms and applications. **SIAM, philadelphia**, JSTOR, 1993.
- MISITI, M. et al. **Wavelets and their Applications**. [S.l.]: John Wiley & Sons, 2013.
- MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to time series analysis and forecasting**. [S.l.]: John Wiley & Sons, 2015.
- MORETTIN, P. A.; TOLOI, C. **Análise de séries temporais**. [S.l.]: Blucher, 2006.
- NASON, G. **Wavelet methods in statistics with R**. [S.l.]: Springer Science & Business Media, 2010.
- OLIVEIRA, H. de. **Análise de Fourier e Wavelets**. [S.l.]: Hélio Magalhaes de Oliveira, 2007.
- ONS. **Séries Históricas de Vazões**. 2017. Disponível em: <<http://www.ons.org.br>>.
- PERCIVAL, D. B.; WALDEN, A. T. **Wavelet methods for time series analysis**. [S.l.]: Cambridge university press, 2006. v. 4.
- QIAN, W. et al. On the forecast combination puzzle. **arXiv preprint arXiv:1505.00475**, 2015.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>.
- SCHLUETER, S.; DEUSCHLE, C. Wavelet-based forecasting of arima time series-an empirical comparison of different methods. **Managerial Economics**, v. 15, n. 1, p. 107, 2014.
- SEO, Y.; CHOI, Y.; CHOI, J. River stage modeling by combining maximal overlap discrete wavelet transform, support vector machines and genetic algorithm. **Water**, Multidisciplinary Digital Publishing Institute, v. 9, n. 7, p. 525, 2017.
- STARCK, J.-L.; MURTAGH, F. D.; BIJAOU, A. **Image processing and data analysis: the multiscale approach**. [S.l.]: Cambridge University Press, 1998.

TAN, Z. et al. Day-ahead electricity price forecasting using wavelet transform combined with arima and garch models. **Applied Energy**, Elsevier, v. 87, n. 11, p. 3606–3610, 2010.

VASNEV, A. L.; CLAESKENS, G.; WANG, W. A simple theoretical explanation of the forecast combination puzzle. **SSRN: [http://dx. doi. org/10.2139/ssrn](http://dx.doi.org/10.2139/ssrn)**, v. 2342841, 2014.

WALDEN, A. T. Wavelet analysis of discrete time series. **PROGRESS IN MATHEMATICS-BOSTON-**, BIRKHAEUSER, v. 202, p. 627–641, 2001.