

UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
INSTITUTO DE CIÊNCIAS EXATAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM  
COMPUTACIONAL

Yulia Gorodetskaya

Modelos hidrológicos híbridos para a bacia do rio Paraíba do Sul: acoplando redes neurais artificiais com transformada wavelet para previsão de vazão em curto prazo com ênfase na previsão de vazões extremas

Juiz de Fora

2022

**Yulia Gorodetskaya**

**Modelos hidrológicos híbridos para a bacia do rio Paraíba do Sul: acoplando redes neurais artificiais com transformada wavelet para previsão de vazão em curto prazo com ênfase na previsão de vazões extremas**

Tese apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Instituto de Ciências Exatas da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutor em Modelagem Computacional. Área de concentração: Modelagem computacional.

Orientador: Prof. D.Sc. Leonardo Goliatt da Fonseca

Coorientador: Prof. D.Sc. Celso Bandeira de Melo Ribeiro

Juiz de Fora

2022

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF  
com os dados fornecidos pelo(a) autor(a)

Gorodetskaya, Yulia.

Modelos hidrológicos híbridos para a bacia do rio Paraíba do Sul: aco-  
plando redes neurais artificiais com transformada wavelet para previsão de  
vazão em curto prazo com ênfase na previsão de vazões extremas / Yulia  
Gorodetskaya. – 2022.

117 f.

Orientador: Leonardo Goliatt da Fonseca

Coorientador: Celso Bandeira de Melo Ribeiro

Tese (Doutorado) – Universidade Federal de Juiz de Fora, Instituto de  
Ciências Exatas . Programa de Pós-Graduação em Modelagem Computaci-  
onal, 2022.

1. Redes Neurais Artificiais. 2. Transformada Wavelet. 3. Previsão de  
vazão. 4. Rio Paraíba do Sul. I. Fonseca, Leonardo Goliatt da, orient. II.  
Ribeiro, Celso Bandeira de Melo, coorient.

**Yulia Gorodetskaya**

**Modelos hidrológicos híbridos para a bacia do rio Paraíba do Sul:** acoplando redes neurais artificiais com transformada wavelet para previsão de vazão em curto prazo com ênfase na previsão de vazões extremas

Tese apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutora em Modelagem Computacional. Área de concentração: Modelagem Computacional.

Aprovada em 05 de setembro de 2022.

**BANCA EXAMINADORA**

**Prof. Dr. Leonardo Goliatt da Fonseca** - Orientador

Universidade Federal de Juiz de Fora

**Prof. Dr. Celso Bandeira de Melo Ribeiro** - Coorientador

Universidade Federal de Juiz de Fora

**Profa. Dra. Luciana Conceição Dias Campos**

Universidade Federal de Juiz de Fora

**Prof. Dr. Heder Soares Bernardino**

Universidade Federal de Juiz de Fora

**Profa. Dra. Eliane da Silva Christo**

Universidade Federal Fluminense

**Profa. Dra. Camila Martins Saporetti**

Universidade do Estado de Minas Gerais

Juiz de Fora, 25/08/2022.



Documento assinado eletronicamente por **Luciana Conceicao Dias Campos, Professor(a)**, em 05/09/2022, às 17:30, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Camila Martins Saporetti, Usuário Externo**, em 05/09/2022, às 17:31, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo Goliatt da Fonseca, Professor(a)**, em 05/09/2022, às 17:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Heder Soares Bernardino, Professor(a)**, em 05/09/2022, às 17:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Celso Bandeira de Melo Ribeiro, Professor(a)**, em 05/09/2022, às 17:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **ELIANE DA SILVA CHRISTO, Usuário Externo**, em 05/09/2022, às 17:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf ([www2.ufjf.br/SEI](http://www2.ufjf.br/SEI)) através do ícone Conferência de Documentos, informando o código verificador **0922963** e o código CRC **C873951F**.

## AGRADECIMENTOS

Agradeço meus orientadores Leonardo Goliatt da Fonseca e Celso Bandeira de Melo Ribeiro pela confiança e ajuda em todas as etapas desta pesquisa, apoio e paciência em todos esses sete anos.

Agradeço toda minha família e amigos brasileiros e russos, em especial, a minha mãe, meu marido Giovanni e meu filho Viktor por todo apoio, intercessão e proximidade em todos esses anos.

Agradeço a todo o corpo docente e técnico da Universidade Federal de Juiz de Fora que acredita e que luta pela importância de um ensino transformador: público e de qualidade. Agradeço os membros do Programa de Pós-Graduação em Modelagem Computacional pelos ensinamentos.

Agradeço a todos que diretamente ou indiretamente contribuíram para o desenvolvimento desse trabalho.

Agradeço à Universidade Federal de Juiz de Fora pela infraestrutura que possibilitou minha formação e pelas políticas sociais que garantiram minha permanência por grande parte da minha formação. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradeço e testemunho a importância do financiamento público da pesquisa no Brasil.

Per aspera ad astra.

## RESUMO

O rio Paraíba do Sul flui através da mais importante região industrial do Brasil, entre as cidades do Rio de Janeiro e de São Paulo. De acordo com a Agência Nacional de Águas, a bacia do rio Paraíba do Sul é caracterizada por conflitos de usos múltiplos de recursos hídricos (abastecimento urbano, diluição de esgotos, irrigação e geração de energia hidrelétrica). Em função de sua importância, o desenvolvimento de modelos precisos de previsão de vazão pode assumir valor estratégico para a gestão da quantidade e da qualidade de água nesta bacia. Estes modelos podem apoiar a tomada de decisão de gestores públicos sobre alertas para condições extremas de inundação ou seca e pode se tornar um elemento-chave para proteger a sociedade e favorecer uma reação oportuna, reduzindo efetivamente danos socioeconômicos. Os modelos baseados na inteligência artificial têm sido aplicados com sucesso para resolver problemas não-lineares em hidrologia. Nas últimas duas décadas, a transformada wavelet, uma técnica de pré-processamento de dados, tem sido estudada para uso em análise de séries temporais e tem se mostrado muito eficaz em dados não estacionários. A transformada wavelet divide as séries temporais originais em subcomponentes que servem como entradas para modelos de inteligência artificial. Assim, fornecem uma representação de tempo-frequência de um sinal em diferentes períodos no domínio do tempo, além de informações sobre a estrutura física de dados. Esta pesquisa visa contribuir com o fornecimento de uma metodologia de desenvolvimento de dois tipos de modelos híbridos chamados  $WANN_{one}$  e  $WANN_{multi}$  acoplando redes neurais artificiais com uma técnica de pré-processamento de dados baseada em transformada wavelet *À Trous*. Os modelos desenvolvidos são aplicados para previsão de vazão natural média diária de 1, 3, 5 e 7 dias à frente em 8 estações fluviométricas localizadas na bacia do rio Paraíba do Sul. O desempenho dos modelos baseados em wavelets e redes neurais artificiais é comparado com o dos modelos convencionais de redes neurais artificiais (ANN). Os resultados revelam que o desempenho dos modelos  $WANN_{one}$  e  $WANN_{multi}$  são significativamente superiores aos modelos ANN em termos de medidas de eficiência da previsão MAPE, RMSE e  $R^2$  em todos os horizontes de previsão testados para todas as estações modeladas. Em termos de aplicabilidade a determinados cenários, o modelo  $WANN_{multi}$  consegue representar melhor os ciclos de vazões baixas, em cenários de seca, ao passo em que  $WANN_{one}$  consegue acompanhar melhor a magnitude das vazões extremas altas, em situações de risco de inundação.

Palavras-chave: Redes neurais artificiais. Transformada wavelet. Previsão de vazão. Rio Paraíba do Sul.

## ABSTRACT

The Paraíba do Sul River flows through the most important industrial region in Brazil, between the cities of Rio de Janeiro and São Paulo. According to the National Water Agency, the Paraíba do Sul river basin is characterized by conflicts of multiple uses of water resources (urban supply, sewage dilution, irrigation and hydroelectric power generation). Due to its importance, the development of accurate flow prediction models can assume strategic value for the management of the quantity and quality of water in this basin. These models can support decision-making by public managers on alerts for extreme flood or drought conditions and can become a key element to protect society and favor a timely reaction, effectively reducing socio-economic damage. Models based on artificial intelligence have been successfully applied to solve nonlinear problems in hydrology. In the last two decades, the wavelet transform, a data pre-processing technique, has been studied for use in time series analysis and has been shown to be very effective in non-stationary data. The wavelet transform divides the original time series into sub-components that serve as inputs to artificial intelligence models. Thus, they provide a time-frequency representation of a signal at different periods in the time domain, as well as information about the physical structure of the data. This research aims to contribute to the provision of a methodology for the development of two types of hybrid models called  $WANN_{one}$  and  $WANN_{multi}$  coupling artificial neural networks with a data pre-processing technique based on a wavelet decomposition algorithm *À Trous*. The developed models are applied to forecast the average daily natural flow of 1, 3, 5 and 7 days ahead in 8 fluvimetric stations located in the Paraíba do Sul river basin. The performance of models based on wavelets and artificial neural networks is compared with that conventional models of artificial neural networks (ANN). The results reveal that the performance of the  $WANN_{one}$  and  $WANN_{multi}$  models are significantly superior to the ANN models in terms of MAPE, RMSE and  $R^2$  forecast efficiency measures at all horizons for all modeled stations. In terms of applicability to certain scenarios, the  $WANN_{multi}$  model is able to better represent low flow cycles, in drought scenarios, while  $WANN_{one}$  can better follow the magnitude of high flows extremes, in situations of risk of flooding.

Keywords: Artificial Neural Networks. Wavelet Transform. Stream flow forecasting. Paraíba do Sul River.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Uma visão geral esquemática das resoluções de tempo e frequência das diferentes transformações em comparação com o conjunto de dados de série temporal original. O tamanho e as orientações do bloco fornecem uma indicação do tamanho da resolução. . . . .	35
Figura 2 – A diferença entre uma onda senoidal e uma Wavelet. A onda senoidal é infinitamente longa e a Wavelet está localizada no tempo. . . . .	36
Figura 3 – Dilatação e translação de uma wavelet: (a) Alongamento e compressão de uma wavelet: Dilatação ( $a_1 = a_2/2$ ; $a_3 = a_1 \times 2$ ). (b) Movendo uma wavelet: Translação. . . . .	38
Figura 4 – Os coeficientes de aproximação e detalhe da onda <i>coif5</i> aplicada em um sinal de vazão, do nível 1 ao 5. À esquerda pode-se ver uma representação esquemática dos filtros passa-alta (PA) e passa-baixa (PB) aplicados no sinal em cada nível. . . . .	41
Figura 5 – Várias famílias de Wavelets. Na primeira linha estão wavelets discretas e na segunda linha estão wavelets contínuas. . . . .	44
Figura 6 – A família de wavelet <i>Daubechies</i> para várias ordens diferentes de momentos de desaparecimento e vários níveis de refinamento. . . . .	45
Figura 7 – Uma simples Rede Neural MLP de três camadas. . . . .	48
Figura 8 – Validação Cruzada <i>Time Series Split</i> . . . . .	51
Figura 9 – Famílias Wavelets usadas. . . . .	55
Figura 10 – Decomposição wavelet AT (em 3 níveis) com wavelet mãe <i>coif4</i> de uma série temporal de vazão, do nível 1 ao 3: coeficientes de aproximação e detalhes. . . . .	57
Figura 11 – Representação esquemática do modelo $WANN_{one}$ . . . . .	58
Figura 12 – Representação esquemática do modelo $WANN_{multi}$ . . . . .	59
Figura 13 – Representação esquemática da técnica de janela de tempo movel. . . . .	60
Figura 14 – Representação do uso da técnica da janela de tempo móvel para criar um vetor de entrada $X$ e vetor de saída $Y$ no caso do modelo $WANN_{one}$ . $Q_i$ são valores de série temporal de vazão, $a_i$ são valores de sub-série de aproximação, $d_i$ são valores de sub-série de detalhe, o $N$ denota o tamanho das séries temporais e $n$ denota o número de valores antecedentes (o tamanho da janela de tempo móvel). . . . .	61
Figura 15 – Representação esquemática da validação cruzada TSS com $k = 5$ splits. . . . .	62
Figura 16 – Ajuste de hiperparâmetros e estimativa de erros em busca em grade com validação cruzada TSS. . . . .	62

Figura 17 – Bacia hidrográfica do Rio Paraíba do Sul e suas respectivas porções: paulista mineira e fluminense. . . . .	64
Figura 18 – Localização das estações fluviométricas modeladas no estudo. . .	67
Figura 19 – Serie histórica de vazão média diária da estação 58183000. . . .	68
Figura 20 – Serie histórica de vazão média diária da estação 58235100. . . .	68
Figura 21 – Serie histórica de vazão média diária da estação 58305001. . . .	69
Figura 22 – Serie histórica de vazão média diária da estação 58380001. . . .	69
Figura 23 – Serie histórica de vazão média diária da estação 58630002. . . .	70
Figura 24 – Serie histórica de vazão média diária da estação 58790002. . . .	70
Figura 25 – Serie histórica de vazão média diária da estação 58795000. . . . .	71
Figura 26 – Serie histórica de vazão média diária da estação 58880001. . . . .	71
Figura 27 – Desempenho dos modelos com as wavelets mãe investigadas na previsão de vazão de 7 dias à frente para 5 estações fluviométricas. Os valores médios das métricas de avaliação após 30 execuções. . . . .	74
Figura 28 – ANN, $WANN_{one}$ e $WANN_{multi}$ do MAPE média em todos os horizontes de previsão . . . . .	76
Figura 29 – ANN, $WANN_{one}$ e $WANN_{multi}$ do $R^2$ média em todos os horizontes de previsão . . . . .	80
Figura 30 – ANN, $WANN_{one}$ e $WANN_{multi}$ do RMSE média em todos os horizontes de previsão . . . . .	81
Figura 31 – Comparação da vazão observada e prevista por modelos WANNs e ANN para horizonte de previsão de 5 dias à frente para as estações: (a) 58880001; (b) 58790002 e (c) 58795000 . . . . .	82
Figura 32 – Comparação da vazão observada e prevista por modelos WANNs e ANN para horizonte de previsão de 7 dias à frente para as estações: (a) 58235100 e (b) 58183000. . . . .	83
Figura 33 – Comparação da vazão observada e prevista por modelos WANNs e ANN para horizonte de previsão de 7 dias à frente para as estações: (a) 58630002; (b) 58380001 e (c) 58305001 . . . . .	84
Figura 34 – Desempenho dos modelos $WANN_{one}$ e $WANN_{multi}$ em termos do MAPE médio em todos os horizontes de previsão. . . . .	85
Figura 35 – Desempenho dos modelos $WANN_{one}$ e $WANN_{multi}$ em termos do RMSE médio em todos os horizontes de previsão. . . . .	86
Figura 36 – Desempenho dos modelos $WANN_{one}$ e $WANN_{multi}$ em termos do $R^2$ médio em todos os horizontes de previsão. . . . .	87
Figura 37 – Os hidrogramas da vazão observada e prevista por modelos WANN para o horizonte de 7 dias à frente no caso da 3 estações: (a) 58630002; (b) 58795000 e (c) 58880001. . . . .	88
Figura 38 – Escolha dos extremos altos e baixos. . . . .	89

Figura 39 – O desempenho dos modelos WANN em termos do MAPE médio na previsão de valores extremos (altos e baixos) para todos os horizontes de previsão. . . . .	91
Figura 40 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58183000. . . . .	106
Figura 41 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 582351000. . . . .	107
Figura 42 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58630002. . . . .	108
Figura 43 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58380001. . . . .	109
Figura 44 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 583050001. . . . .	110
Figura 45 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58880001. . . . .	111
Figura 46 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58790002. . . . .	112
Figura 47 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58795000. . . . .	113
Figura 48 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58630002 para o horizonte de previsão de 7 dias à frente.114	
Figura 49 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58380001 para o horizonte de previsão de 7 dias à frente.114	
Figura 50 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58305001 para o horizonte de previsão de 7 dias à frente.114	
Figura 51 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58235100 para o horizonte de previsão de 7 dias à frente.115	
Figura 52 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58790002 para o horizonte de previsão de 7 dias à frente.115	
Figura 53 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58795000 para o horizonte de previsão de 7 dias à frente.115	

- Figura 54 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58880001 para o horizonte de previsão de 7 dias à frente.116
- Figura 55 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58183000 para o horizonte de previsão de 7 dias à frente.116

## LISTA DE TABELAS

- Tabela 1 – Resumo dos trabalhos pesquisados para previsão de séries hidrológicas, incluindo ano de publicação, autores, área de estudo, o métodos de IA acoplado com wavelets acoplados, o tipo de transformada wavelet, escala de tempo e variável de previsão, bem como a informação sobre a investigação das wavelets mãe e avaliação de vazões extremas. . . . . 32
- Tabela 2 – Famílias de wavelets utilizadas. . . . . 46
- Tabela 3 – Hiperparâmetros utilizados em busca em grande com validação cruzada. . . . . 63
- Tabela 4 – Informações das estações fluviométricas modeladas nesta pesquisa. 66
- Tabela 5 – Desempenho de diferentes wavelets mãe para previsão de 7 dias à frente. . . . . 73
- Tabela 6 – Tamanho de conjuntos de dados de treinamento e teste, em dias. 75
- Tabela 7 – Quantidade de *lags* para os modelos ANN, WANN<sub>one</sub> e WANN<sub>multi</sub>. 75
- Tabela 8 – Os melhores *lags* de entrada para os modelos ANN, WANN<sub>one</sub> e WANN<sub>multi</sub> após 30 execuções. . . . . 76
- Tabela 9 – Desempenho dos modelos para diferentes horizontes de previsão para as estações fluviométricas 58630002, 58380001 e 58305001. Os valores médios das métricas de avaliação após 30 execuções. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses. . . . . 77
- Tabela 10 – Desempenho dos modelos para diferentes horizontes de previsão para estações fluviométricas 58880001, 58790002 e 58795000. Os valores médios das métricas de avaliação após 30 execuções. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses. . . . . 78
- Tabela 11 – Desempenho dos modelos para diferentes horizontes de previsão para estações fluviométricas 58235100 e 58183000. Os valores médios das métricas de avaliação após 30 execuções. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses. . . . . 79
- Tabela 12 – Desempenho dos modelos WANN em extremos altos e baixos para diferentes horizontes de previsão. Os valores médios da métrica MAPE para extremos altos e baixos após 30 execuções. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses. . . . . 90

Tabela 13 – Resumo dos resultados do p-valor do teste Shapiro sobre normalidade das distribuições dos resultados das métricas avaliadas. Os itens em negrito indicam a existência da normalidade. . . . .	92
Tabela 14 – Resumo dos resultados do p-valor do teste U de Mann-Whitney <i>one – sided</i> das métricas avaliadas. Os itens em negrito indicam a existência da superioridade no desempenho do modelo $WANN_{one}$ para a previsão de 7 dias à frente. . . . .	93

## LISTA DE ABREVIATURAS E SIGLAS

ANA	Agência Nacional de Águas
ANFIS	Inferência fuzzy baseado em neural adaptativo
ANN	Redes Neurais Artificiais
AR	Modelo Auto-Regressivo
AWT	Transformada Wavelet Aditiva
AWT-ANN	Rede Neural Artificial com Transformada Wavelet Aditiva
AR	Modelo autorregressivo
ARMA	Modelo de média móvel autorregressiva
ARIMA	Modelo Autoregressivo Integrado de Média Móvel
AT	Algoritmo À Trous
CV	Validação cruzada
CWT	Transformada Wavelet Contínua
DWT	Transformada Wavelet Discreta
DWT-ANN	Rede Neural Artificial com Transformada Wavelet Discreta
IA	Inteligencia Artificial
FFNN	Rede Neural Feedforward
FT	Transformada de Fourie
LR	Regressão Linear
LSTM	Modelo de memória de longo prazo
MAPE	Erro Absoluto Médio Percentual
M-DWT	Transformada Wavelet Multi Discreta
MLP	Perceptron Multicamadas
MODWT	Transformada Wavelet Discreta de sobreposição máxima
PIB	Produto Interno Bruto
PNRH	Política Nacional de Recursos Hídricos
R <sup>2</sup>	Coefficiente de determinação
RF	Random Forest
ReLU	função de ativação Rectified Linear Unit
RH-IX	Região Baixo Paraíba do Sul
RMSE	Raiz quadrática do erro médio

RNA	Rede Neural Artificial
SIRH	Sistema de Informações sobre Recursos Hídricos
SVR	Regressão de vetores de suporte
SVM	Máquinas de Vetores de Suporte
STFT	Transformada de Fourier de curto termo
TRMM	Tropical Rainfall Measuring Mission
TSS	Time Series Split
UHE	Usina Hidrelétrica
WANN	Wavelet-Rede Neural Artificial
WT	Transformada Wavelet

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>17</b>
1.1	CARACTERIZAÇÃO DO RIO PARAÍBA DO SUL . . . . .	17
1.2	CLASSIFICAÇÃO DE PREVISÃO DE VAZÃO EM CURTO PRAZO .	19
1.3	MODELOS HIDROLÓGICOS . . . . .	19
1.3.1	<b>Modelos conceituais . . . . .</b>	<b>20</b>
1.3.2	<b>Modelos empíricos . . . . .</b>	<b>20</b>
1.3.3	<b>Seleção de variáveis de entrada . . . . .</b>	<b>21</b>
1.4	ESCOLHA DE MÉTODOS PREDITIVOS . . . . .	21
1.5	MOTIVAÇÃO . . . . .	22
1.6	CONTRIBUIÇÃO . . . . .	24
1.7	OBJETIVOS . . . . .	25
1.7.1	<b>Objetivo Geral . . . . .</b>	<b>25</b>
1.7.2	<b>Objetivos Específicos . . . . .</b>	<b>25</b>
1.8	ESTRUTURA DO TEXTO . . . . .	25
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA . . . . .</b>	<b>27</b>
2.1	APLICAÇÃO DE MODELOS HÍBRIDOS BASEADOS EM WAVELETS EM ESTUDOS HIDROLÓGICOS NA PREVISÃO DE VAZÃO FLUVIAL	27
<b>3</b>	<b>MÉTODOS . . . . .</b>	<b>34</b>
3.1	PRINCIPAIS CONCEITOS DA TRANSFORMADA WAVELET . . . . .	34
3.1.1	<b>Da Transformada de Fourier à Transformada Wavelet . . . . .</b>	<b>34</b>
3.1.2	<b>Wavelets . . . . .</b>	<b>36</b>
3.1.3	<b>Transformada Wavelet Contínua e Transformada Wavelet Discreta</b>	<b>39</b>
3.1.4	<b>A Transformada Wavelet Discreta como um banco de filtros . .</b>	<b>40</b>
3.1.5	<b>Algoritmo À Trous de decomposição wavelet . . . . .</b>	<b>42</b>
3.1.6	<b>O problema de “dados futuros” . . . . .</b>	<b>43</b>
3.1.7	<b>Famílias de Wavelet . . . . .</b>	<b>44</b>
3.2	CONCEITOS BÁSICOS DE UMA REDE NEURAL PERCEPTRON MULTICAMADAS . . . . .	47
3.3	VALIDAÇÃO CRUZADA PARA PREVISÃO DE SÉRIES TEMPORAIS	50
3.4	CRITÉRIOS DE AVALIAÇÃO DA CAPACIDADE PREDITIVA DOS MODELOS . . . . .	51
3.5	TESTES ESTATÍSTICOS . . . . .	53
<b>4</b>	<b>METODOLOGIA PROPOSTA . . . . .</b>	<b>55</b>
4.1	PRÉ-PROCESSAMENTO DE DADOS . . . . .	55
4.1.1	<b>Escolha da Wavelet mãe e nível de decomposição . . . . .</b>	<b>55</b>
4.1.2	<b>Decomposição wavelet . . . . .</b>	<b>56</b>
4.2	MODELOS HÍBRIDOS WAVELET REDES NEURAS ARTIFICIAIS .	58

4.2.1	Primeiro modelo híbrido proposto $WANN_{one}$ . . . . .	58
4.2.2	Segundo modelo híbrido proposto $WANN_{multi}$ . . . . .	59
4.2.3	Técnica de janela de tempo móvel . . . . .	60
4.2.4	Busca em grade com validação cruzada para ajuste de hiperpa- râmetros . . . . .	60
4.2.5	Configuração de modelos propostos . . . . .	63
<b>5</b>	<b>ÁREA DE ESTUDO E DADOS</b> . . . . .	<b>64</b>
5.1	ÁREA DE ESTUDO . . . . .	64
5.1.1	Caracterização da região Fluminense da bacia do rio Paraíba do Sul . . . . .	65
5.1.2	Caracterização da região Paulista da bacia do rio Paraíba do Sul	65
5.2	SELEÇÃO DE DADOS . . . . .	66
<b>6</b>	<b>RESULTADOS E DISCUSSÃO</b> . . . . .	<b>72</b>
6.1	INVESTIGAÇÃO DE FAMÍLIAS WAVELET . . . . .	72
6.2	AVALIAÇÃO GERAL DE COMPORTAMENTO DE MODELOS WANN E ANN SOBRE TODAS AS ESTAÇÕES . . . . .	73
6.3	COMPARAÇÃO DOS MODELOS $WANN_{one}$ E $WANN_{multi}$ . . . . .	85
6.3.1	Avaliação do desempenho dos modelos WANN para vazões extremas . . . . .	87
6.3.2	Análise estatística . . . . .	89
6.4	TENDÊNCIAS EM COMUM NO COMPORTAMENTO DA MODE- LAGEM DAS 8 ESTAÇÕES MODELADAS . . . . .	94
<b>7</b>	<b>CONCLUSÃO</b> . . . . .	<b>96</b>
7.1	PERSPECTIVAS DE TRABALHOS FUTUROS . . . . .	97
	<b>REFERÊNCIAS</b> . . . . .	<b>98</b>
	<b>APÊNDICE A – Resultados</b> . . . . .	<b>105</b>
	<b>APÊNDICE B – Trabalhos publicados</b> . . . . .	<b>117</b>

# 1 INTRODUÇÃO

## 1.1 CARACTERIZAÇÃO DO RIO PARAÍBA DO SUL

A bacia do rio Paraíba do Sul é de grande importância para o Brasil pois engloba os três maiores estados brasileiros em população: Rio de Janeiro, São Paulo e Minas Gerais. Desempenha, há mais de 300 anos, função estratégica nos diversos ciclos de desenvolvimento econômico e ocupação territorial brasileiros, o que vai imprimir uma de suas principais características - o conflito nos múltiplos usos de suas águas.

A bacia está localizada entre os maiores polos industriais do país, as cidades do Rio de Janeiro e São Paulo e, atualmente, a bacia comporta um complexo parque industrial com mais de 6000 fábricas e 120 usinas hidroelétricas em operação (dentre pequenas, médias e grandes geradoras) que respondem por cerca de 11% do Produto Interno Bruto nacional (PIB) (1). Ao norte, é limitada pelas bacias dos rios Grande e Doce e pelas serras da Mantiqueira, Caparaó e Santo Eduardo. A nordeste é limitada pela bacia do rio Itabapoana. Ao sul, a bacia é limitada pela Serra dos Órgãos e pela Serra do Mar e, à oeste, pela bacia do rio Tietê.

A Agência Nacional de Águas (ANA) caracteriza a bacia do rio Paraíba do Sul por seus acentuados conflitos de usos múltiplos de recursos hídricos, sendo suas águas utilizadas no abastecimento urbano, na diluição de esgotos, na indústria, agricultura e geração de energia hidrelétrica. No município de Barra do Piraí, estado do Rio de Janeiro, e através da Usina Elevatória Santa Cecília, o Paraíba do Sul tem cerca de 2/3 de suas águas transpostas para o rio Guandú, responsável pelo abastecimento de cerca 80% da população (cerca de 9 milhões de pessoas) na região metropolitana do Rio de Janeiro (2).

No Brasil, a Política Nacional de Recursos Hídricos (PNRH), instituída pela Lei n. 9.433/97, é responsável por regulamentar a utilização dos recursos hídricos estabelecendo como fundamentos a água ser um bem de domínio público; um recurso natural limitado que, em situações de escassez, deve ter seu uso priorizado para o consumo humano e a dessedentação de animais; um bem dotado de valor econômico e cujo aproveitamento deve sempre proporcionar sua múltipla utilização. Estabelece também a bacia hidrográfica como a unidade territorial através da qual deve ser implementada tanto a PNRH quanto a atuação do Sistema Nacional de Gerenciamento de Recursos Hídricos. Para isto, sua gestão deve ser descentralizada, o que significa que o poder público, os usuários, a sociedade civil e a comunidade científica devem ter participação ativa nesta tarefa (3).

A PNRH tem por objetivos: assegurar o acesso à água em padrões de qualidade adequados pelas gerações futuras; a utilização racional e integrada dos recursos hídricos, incluindo o transporte aquaviário, objetivando o desenvolvimento sustentável; a prevenção e a defesa contra eventos hidrológicos críticos de origem natural ou consequentes do uso inadequado dos recursos naturais; e o incentivo e promoção da captação, preservação e

aproveitamento de águas pluviais.

Nesse sentido, um desafio atual às pesquisas em modelagem de recursos hídricos é fazer com que o conhecimento científico seja disponibilizado de forma acessível e contribua efetivamente para a tomada de decisão em gestão ambiental. Entretanto, ainda há uma carência de ferramentas que apoiem a implementação efetiva dos dispositivos legais previstos para a adequada gestão dos recursos hídricos (4). Uma demanda permanente é a busca por tecnologias que subsidiem e fomentem, com regras de governança equilibradas e sustentáveis, a competitividade econômica, o desenvolvimento social e a proteção ambiental.

A bacia do rio Paraíba do Sul passou por diversas intervenções antrópicas que modificaram o comportamento de seus corpos hídricos em situações normais e extremas: o desmatamento, a exploração inadequada de terras e encostas (e a sedimentação e assoreamento resultantes desta ação), efluentes domésticos e industriais sem tratamento, a extração de areia, entre outras. Adicionalmente, a geração de energia elétrica por meio de força hidráulica causa impacto ambiental negativo devido à necessidade de alagamento de grandes áreas. Por outro lado, seus reservatórios são importantíssimos para o controle de inundações (5). Exercem influência sobre o regime de vazão do rio Paraíba do Sul diversas usinas e reservatórios distribuídos pela bacia, por exemplo: as usinas com reservatório de Paraibuna, Santa Branca, Jaguari, Funil; a usina de bombeamento de Santa Cecília; as usinas a fio d'água de Picada, Sobragi, Simplício, Anta, Ilha dos Pombos, Pereira Passos e Fontes e os reservatórios de Lages, Tócos, Santana e Vigário (6).

Em um contexto de significativas intervenções no curso do Rio Paraíba do Sul que somadas, impactam seu regime de vazão e aceleram sua degradação, estudos voltados para o desenvolvimento de modelos de previsão de vazão são de grande importância para uma melhor gestão tanto da disponibilidade (quantidade) quanto para a qualidade da água. A vazão é um dado de fundamental importância pois está conectado com quase todos os fatores ambientais, sociais e econômicos relacionados à água. O regime de vazão de um rio está sujeito a diversos condicionantes. Alterações no uso da terra exercem impacto sobre a frequência e intensidade de oscilações na vazão, como as decorrentes de alterações na cobertura florestal (desmatamento) e na impermeabilização do solo em áreas urbanas. A captação de água para atividade agrícola e industrial pode afetar significativamente a vazão dos rios, o prolongamento de secas e pode diminuir ou paralisar a atividade industrial, ao passo que as usinas hidrelétricas podem bloquear a vazão de seus vertedouros em horários de baixo consumo e liberarem a vazão de forma abrupta em horários de pico de consumo.

Para Tucci e Collischonn(7), a vazão de um rio é definida como “o resultado da precipitação que ocorre sobre a bacia, o balanço hídrico no solo, evapotranspiração e o escoamento pela bacia, rios, canais e reservatórios”. As previsões de vazão podem ser de longo ou curto prazo, sendo esta última aplicação o objetivo desta pesquisa.

Estudos sobre previsão de vazão podem trazer significativas contribuições científicas

para diversificadas áreas de conhecimento, a depender de sua classificação como previsão de curto ou longo prazo. No contexto do presente trabalho, a previsão de curto prazo, de acordo com (8), é assim classificada quando o horizonte de previsão é de algumas horas até cerca de três semanas. Previsões de curto prazo podem ser aplicadas na definição da vazão para um reservatório em uma bacia hidrográfica, uso de vertedouros, vazão sob pontes, gestão de geração de hidrelétricas, previsão de enchentes, emissões de alerta para população ribeirinha em cenários de inundação de cidades, riscos de colapso na operação de obras hidráulicas, irrigação e navegação fluvial. A previsão de vazões máximas atuará como precursora do planejamento, gerenciamento e mitigação do risco de inundação ou falta de água. Para um melhor uso dos recursos disponíveis, estas aplicações exigem um acompanhamento contínuo dos níveis de água acumulada. Os limites de antecedência fornecidas por metodologias e erros de estimativa limitam a previsão de vazão.

## 1.2 CLASSIFICAÇÃO DE PREVISÃO DE VAZÃO EM CURTO PRAZO

Conforme exposto por Tucci e Collischonn(7), métodos de previsão de vazão em curto prazo podem apoiar-se em dados da vazão do próprio local, à montante e na precipitação empírica.

- No caso da vazão do próprio local, tem-se como única variável a própria vazão do local em que se busca obter a previsão. A qualidade da previsão está diretamente relacionada as séries históricas de vazão e de sua auto-correlação. Não obtém bom desempenho para antecedências grandes e subestima os picos embora seja eficaz na previsão de recessão do escoamento.
- No cenário da vazão observada no próprio local e à montante, o fator de limitação é o tempo de viagem da onda de cheia entre os dois pontos. Esta abordagem de previsão obtém melhores resultados em grandes rios com baixa contribuição incremental da bacia sob modelos baseados em técnicas como redes neurais, modelos estocásticos e regressão múltipla.
- Na última abordagem, com precipitação observada e vazão do rio, há a necessidade de correta estimativa de precipitação e ajuste no modelo hidrológico com duas opções: previsão com base apenas em precipitação ou com base em precipitação e vazões à montante. Esta abordagem apoia-se fundamentalmente em monitorar a precipitação até o tempo de previsão.

## 1.3 MODELOS HIDROLÓGICOS

A modelagem hidrológica impõe diversos desafios relacionados à grande quantidade de variáveis que compõem as especificidades de cada bacia. Estas variáveis possuem

alta não-linearidade e variabilidade espacial e temporal de seus processos, como por exemplo: a condição inicial de umidade do solo, precipitação, condensação, infiltração, interceptação, escoamento superficial e subterrâneo, percolação e evapotranspiração. Todos esses fatores contribuem para a complexidade da tarefa parâmetros de modelagem incertos e a integridade dos dados de entrada (9). Atualmente existem duas abordagens na hidrologia para a previsão de vazão.

### 1.3.1 Modelos conceituais

Modelos conceituais, ou modelos distribuídos de base física (também chamados *theory-driven*) representam os sub-processos internos gerais e mecanismos físicos do ciclo hidrológico sem considerar a variabilidade espacial e as propriedades estocásticas do processo de chuva-vazão. Normalmente, eles envolvem a solução de um sistema de equações diferenciais que representam nossa melhor compreensão do processo de vazão no corpo d'água. Os parâmetros são geralmente considerados como representações agrupadas das características da bacia.

Os modelos conceituais geralmente tendem a exigir uma grande quantidade de dados de entrada (ou seja, previsões de chuva e dados de topografia) que nem sempre estão disponíveis ou podem ser difíceis de obter. Além disso, os parâmetros do modelo precisam ser testados e avaliados cuidadosamente, pois esses parâmetros são regionalmente dependentes e, às vezes, é difícil estimar ou calibrar os parâmetros que se ajustam ao modelo. Como resultado, os modelos não obtêm bom desempenho, especialmente em áreas onde os dados disponíveis são limitados.

De acordo com Le et al.(10) os modelos conceituais apresentam limitações quanto aos alertas de inundação, pois o tempo de execução desses modelos costuma ser bastante longo. Além disso, os autores Le et al.(10) sublinham que esses modelos não preveem o aumento da vazão a jusante se a vazão a montante tiver flutuações bruscas causadas pela vazão de grandes barragens e reservatórios, principalmente para grandes bacias hidrográficas.

### 1.3.2 Modelos empíricos

Modelos empíricos, ou orientados a dados (*data-driven models*, também chamados caixa preta ou *black-box*) consideram o sistema hidrológico como uma caixa-preta e buscam encontrar relações entre dados históricos de entrada (por exemplo, chuva, temperatura, etc.) e saídas (por exemplo, vazão). Um modelo empírico é baseado na análise das propriedades estatísticas da série temporal de parâmetros físicos. Tradicionalmente, modelos orientados a dados tomam emprestado técnicas desenvolvidas em áreas como estatística, inteligência computacional, aprendizado de máquina e mineração de dados. Os modelos empíricos são preferidos aos modelos conceituais quando os processos são extremamente complexos e

suas inter-relações são difíceis de conceituar (11).

### 1.3.3 Seleção de variáveis de entrada

O objetivo dos modelos empíricos é generalizar relacionamentos na forma de  $Y^m = f(X^n)$ , onde  $X^n$  é um vetor de entrada n-dimensional que representa as variáveis  $x_1, \dots, x_i, \dots, x_n$ ;  $Y^m$  é um vetor de saída m-dimensional que representa as variáveis de interesse resultantes  $y_1, \dots, y_i, \dots, y_m$ . Em hidrologia, os valores  $x_i$  podem representar variáveis causais como chuva, temperatura, vazões anteriores, evaporação, entre outros. Os valores de  $y_i$  podem ser respostas hidrológicas como vazão. Um importante elemento para a construção de uma previsão de bom desempenho é a escolha apropriada de amostras temporais que representarão a série como um todo (12, 13).

Solomatine e Dulal(13) advertem que a escolha das variáveis de entrada é normalmente baseada no conhecimento antecipado das variáveis causais em complemento com inspeções visuais na plotagem das séries temporais que potencialmente sejam utilizadas como entradas e saídas. Além disso, segundo o autor, como relevante requisito para a aplicação bem sucedida de um modelo empírico, exige-se um sólido conhecimento do sistema hidrológico em questão.

De acordo com Sudheer, Gosain e Ramasastr(14), na modelagem de vazão de um rio os parâmetros que devem ser escolhidos no vetor de entrada são o número de valores de chuva/vazão para diferentes intervalos de tempo que melhor representem o processo pelo modelo. Identificar o número de parâmetros de chuva/vazão implica determinar os atrasos de chuva/vazão que exerçam impacto significativo sobre a vazão predita. No caso da previsão de cheias Le et al.(10) notam que os dados de vazão representaram um papel mais importante que os dados de chuva. Os autores demonstram que a existência de dados de precipitação causa ruído na previsão precisa do valor da vazão máxima.

Para que modelos de aprendizado de máquina estabeleçam relacionamentos entre entradas e saídas, é fundamental a escolha de bons conjuntos de treinamento dentre as séries temporais disponíveis. Preferencialmente, é importante adicionar a maior parte possível de eventos extremos, tais como valores muito altos e muito baixos, no conjunto de treinamento (15).

## 1.4 ESCOLHA DE MÉTODOS PREDITIVOS

A escolha dos métodos preditivos depende de vários fatores, como: a disponibilidade de dados (dados de séries históricas, de radar, de previsões a partir de modelos numéricos), características físicas da bacia hidrográfica e o objetivo da previsão (previsão de curto prazo, análise de risco, cenários de mudanças climáticas, mudanças do uso da terra, entre outros) (9). De acordo com White(16), a escolha de um método adequado se apoia na busca por um equilíbrio entre três propriedades: precisão, confiabilidade e generalidade. A busca

por modelos também não deve apoiar-se exclusivamente em estatísticas. Alguns modelos representam melhor fenômenos físicos da hidrologia e devem levar em conta um balanço entre otimizar a acurácia ou otimizar a generalidade. White(16) e Solomatine e Ostfeld(17) destacam que não existe, na hidrologia, uma recomendação de aplicação universal de um modelo empírico a um dado problema. Na medida em que aplicações relacionadas à água são geralmente caracterizadas por dados de má qualidade, é recomendada a aplicação de vários tipos de técnicas e a utilização dos resultados de forma comparativa e complementar (17).

Entretanto, Shortridge, Guikema e Zaitchik(18) notam que a precisão preditiva por si só não é uma justificativa suficiente para aplicar um modelo a um determinado problema. Os modelos devem não apenas ser precisos, mas também adequados a um propósito. Por exemplo, uma representação precisa de fluxos de baixo período de retorno é mais importante em um modelo de previsão de inundações do que uma representação que visa prever quantidades médias de água disponível para retirada e consumo humano.

## 1.5 MOTIVAÇÃO

Dentre os modelos empíricos, modelos de séries temporais clássicos tais como Autoregressivo Integrado de Média Móvel (ARIMA) são amplamente utilizados na previsão de séries hidrológicas. Entretanto, tratam-se essencialmente de modelos lineares que tomam os dados como fixos, imóveis. Possuem capacidade restrita de representar relações complexas, multidimensionais e não lineares entre conjuntos de dados de entrada e saída em hidrologia (12, 19, 20). Por outro lado, progressos significativos na representação de tais relações tem sido alcançados por meio de técnicas baseadas em inteligência artificial (IA) que são consideradas a uma das mais aplicadas na área de recursos hídricos (10). No entanto, apesar da flexibilidade e utilidade de tais técnicas na modelagem de processos hidrológicos, a presença de “sazonalidade” em séries temporais, a falta de dados de entrada, pré ou pós-processamento, podem não permitir que modelos de IA lidem adequadamente com dados de fluxo natural não estacionário (21). Nesse sentido, Nourani et al.(21) e Yaseen et al.(22) apontam que modelos de IA acoplados (ou híbridos) a técnicas de pré-processamento de dados podem melhorar a precisão de modelos convencionais de IA, aproveitando a capacidade destes modelos de descrever efetivamente dados observados em termos estatísticos, bem como outras informações preliminares ocultas em registros observacionais.

De acordo com Wu, Chau e Li(23), uma série de fluxo natural pode ser vista como um sinal quase periódico, que é contaminado por vários ruídos em diferentes níveis de fluxo. Sinais mais limpos usados como entradas do modelo melhorarão o desempenho do modelo e ajudarão esses modelos a explicar a física dos processos hidrológicos (24). Nas últimas duas décadas, a transformada wavelet (WT), uma técnica de pré-processamento

de dados, tem sido estudada para uso em análise de séries temporais e tem se mostrado muito eficaz em dados não estacionários (25, 26). As revisões recentes das aplicações da transformada wavelet (21, 27, 22, 28) destacam a crescente popularidade do uso da mesma na modelagem hidrológica.

A WT como técnica matemática, inicialmente formalizada por Grossmann e Morlet(29), é útil na análise numérica e manipulação de conjuntos de sinais multidimensionais (21), fornecendo uma representação do processo e de suas relações em escala de tempo. Em uma revisão das aplicações da WT na modelagem de séries temporais hidrológicas, Sang(28) destacaram-se as seguintes informações multifacetadas que podem ser extraídas com WT: caracterização e compreensão das escalas multitemporais das séries hidrológicas, identificação de sazonalidades e tendências, e eliminação de ruídos dos dados. Portanto, a capacidade da WT de decompor sinais não estacionários em subsinais em diferentes escalas temporais (níveis) é útil para interpretar melhor os processos hidrológicos (25, 21).

A WT pode ser interpretada como uma convolução entre um filtro wavelet mãe (uma pequena função semelhante a uma onda) e uma dada série temporal, resultando em um conjunto de sub-séries temporais (coeficientes de aproximação e detalhe) que contêm informações de alta e baixa frequência da série temporal original fornecendo uma estrutura mais coerente de séries temporais complexas para um modelo orientado a dados. O conteúdo de alta e baixa frequência do conjunto de sub-séries temporais depende: 1) do nível de decomposição, que representa as diferentes escalas de mudança na série temporal e 2) do filtro wavelet mãe, cuja largura e forma determinam o tipo de características que podem ser extraídas da série temporal (por exemplo, tendências, transientes, periodicidades, comportamento polinomial, etc.) (30, 25). A seleção de uma wavelet mãe eficiente e o nível de decomposição são duas questões importantes no primeiro passo. Para determinar a wavelet mãe apropriada para um determinado problema, uma variedade de wavelets mãe deve ser testada por meio de um processo de tentativa e erro (30, 21, 28). No entanto, a semelhança na forma entre a wavelet mãe e a série temporal original é muitas vezes a melhor orientação na escolha de uma wavelet mãe robusta. Entre os algoritmos que executam a decomposição wavelet, destacam-se: transformada wavelet contínua (CWT), transformada wavelet discreta (DWT), o algoritmo Maximum Overlap DWT (MODWT) e o algoritmo *À trous* (AT) (31).

Quilty e Adamowski(31) salientam que o desenvolvimento e interpretação adequados de previsões baseadas em wavelets nem sempre são cuidadosamente estudados, o que muitas vezes resulta em previsões que não podem ser usadas em cenários da vida real. Na previsão de uma série temporal, as informações futuras não devem ser implícitas com base em valores de tempo passado. No entanto, algumas transformadas wavelet (tais como, por exemplo, DWT), devido a suas próprias estruturas, utilizam dados do futuro da série temporal (para tempos  $> t$ ) para calcular sub-séries temporais no presente (no tempo  $t$ ). Se não abordar essas condições de contorno durante a decomposição wavelet,

os modelos de previsão baseados em wavelet incorretamente desenvolvidos geralmente resultam em um desempenho muito melhor do que o que é alcançável na realidade (32, 31). Quilty e Adamowski(31) demonstram que o emprego de algoritmos de decomposição de wavelet MODWT e AT fornece o desenvolvimento correto de modelos de previsão baseados em wavelet para uso em problemas de previsão de fluxo de curto prazo no mundo real. Além disso, os autores notam que entre os trabalhos publicados no Scopus no período entre os anos de 2003 e 2018 (na modelagem hidrológica), 90% dos trabalhos utilizam a transformada DWT como um algoritmo de decomposição wavelet o que não é correto em termos do problema de uso de dados futuros no algoritmo de decomposição wavelet, conforme mencionado acima.

Durante a busca pelos métodos que auxiliam na melhora da modelagem de vazão no horizonte distante e no aumento da precisão da previsão das vazões extremas, optou-se pela utilização da transformada wavelet. Como uma técnica de pré-processamento de dados, ela alcançou ótimos resultados nas tarefas de previsão de séries temporais hidrológicas. Com objetivo de fornecer a abordagem que permita a aplicação em tarefas de previsão no mundo real foi escolhido o algoritmo de decomposição *À trous*. Previamente, durante a pesquisa, a aplicabilidade dos outros métodos de aprendizagem de máquina (33, 34, 35) na previsão de vazão em curto prazo no caso da Bacia do rio Paraíba do Sul foi investigada, porém os resultados alcançados por acoplamento wavelet e redes neurais artificiais foram muito superiores, especialmente na previsão de vazões extremas.

## 1.6 CONTRIBUIÇÃO

Nos últimos anos, registrou-se a formação de um significativo corpo de estudos sobre hidrologia na bacia do rio Paraíba do Sul. Entretanto, são raros os registros na literatura de estudos que apliquem técnicas de inteligência computacional na modelagem de vazão no contexto do Rio Paraíba do Sul. Neste sentido, dada a importância estratégica deste rio, este estudo busca fornecer contribuições científicas para a área de modelagem computacional aplicada a recursos hídricos e que também tenha desdobramentos nas esferas legal, econômica, ambiental e social.

Os resultados deste estudo podem ser úteis para engenharia hidrológica e para tomadores de decisão em previsão de vazão de rios na mitigação do impacto de períodos de seca, reação a inundações em cidades, possibilidade de colapso de obras hidráulicas e planejamento de disponibilidade hídrica em áreas urbanas, na irrigação, navegação fluvial e na distribuição sustentável de água. Os horizontes de previsão adotado neste estudo, para até 7 dias de antecedência busca um meio termo entre previsões mais precisas e disponibilidade de tempo de resposta, permitindo que, em casos extremos, o fornecimento de informações estratégicas auxilie a mobilização de recursos na esfera pública. Por eventos extremos consideramos valores compreendidos por gestores públicos como muito baixos,

que comprometam o abastecimento público de água e a atividade econômica (geração de energia elétrica, indústria, mineração e agropecuária) e valores muito altos que provoquem destruição e danos econômicos à sociedade. A previsão da vazão do rio Paraíba do Sul pode posteriormente ser utilizada por modelos de qualidade de água na previsão da disponibilidade hídrica e na redução do impacto de condições ambientais críticas.

## 1.7 OBJETIVOS

### 1.7.1 Objetivo Geral

O objetivo principal desta pesquisa é apresentar uma nova metodologia de desenvolvimento de modelos hidrológicos híbridos para o caso da bacia do rio Paraíba do Sul, combinando redes neurais artificiais com transformada wavelet para a previsão de vazão diária em curto prazo com ênfase na previsão de vazões extremas (altos e baixos).

### 1.7.2 Objetivos Específicos

Com o intuito de alcançar o objetivo geral, se propõem os seguintes objetivos específicos:

- na etapa de pré-processamento de dados, investigar a influência de diferentes famílias de wavelets na precisão da previsão de vazão do rio Paraíba do Sul e escolher a wavelet mãe apropriada para a modelagem de vazão em curto prazo;
- desenvolver robustos modelos híbridos acoplando redes neurais artificiais com uma técnica de pré-processamento de dados baseada em transformada wavelet *À Trous*;
- aplicar os modelos desenvolvidos na previsão de vazão de 1, 3, 5 e 7 dias à frente em estações fluviométricas da bacia do rio Paraíba do Sul;
- comparar a capacidade preditiva dos modelos baseados em wavelet e redes neurais artificiais com modelos de redes neurais artificiais convencionais para diferentes horizontes temporais; e
- analisar a capacidade preditiva de modelos baseados em wavelets desenvolvidos na previsão de vazões extremas baixas e altas em diferentes horizontes temporais - cenários de seca ou risco de inundação.

## 1.8 ESTRUTURA DO TEXTO

O presente trabalho está estruturado em sete capítulos mais um apêndice descritos a seguir. No Capítulo 1 encontra-se a contextualização, motivação e os objetivos do estudo. No Capítulo 2 é apresentada uma revisão bibliográfica que aborda assuntos pertinentes

ao enfoque do presente trabalho. No Capítulo 3 são explicados os conceitos básicos de métodos empregados neste estudo. A metodologia proposta é apresentada no Capítulo 4. A caracterização das regiões do estudo e os dados utilizados são apresentados no Capítulo 5. Os resultados obtidos e as discussões associadas são apresentados no Capítulo 6. As conclusões e trabalhos futuros são apresentados no Capítulo 7. Em Apêndice são encontrados os gráficos de alguns resultados e lista de trabalhos publicados durante a pesquisa.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 APLICAÇÃO DE MODELOS HÍBRIDOS BASEADOS EM WAVELETS EM ESTUDOS HIDROLÓGICOS NA PREVISÃO DE VAZÃO FLUVIAL

O número de trabalhos empregando modelos baseados em wavelets na área de recursos hídricos vem crescendo constantemente nos últimos anos e é de grande interesse devido à precisão e robustez da abordagem dada a capacidade das transformadas wavelet em análise multi-resolução, capacidade de localização tanto no domínio do tempo como da frequência e na eliminação de ruídos. A seguir, são demonstrados exemplos encontrados na literatura de estudos em modelagem hidrológica baseada em wavelets.

Adamowski e Sun(25) desenvolveram modelos que acoplam redes neurais artificiais (RNA) e transformada wavelet AT e os compararam com modelos convencionais de RNA para previsão de vazão em horizontes de 1 a 3 dias à frente para dois rios diferentes em Chipre. Determinou-se que para previsão de 1 e 3 dias à frente, os modelos wavelet-redes neurais artificiais forneceram resultados mais precisos que os modelos ANN convencionais.

No estudo de Tiwari e Chatterjee(36) a técnicas de wavelet e bootstrapping é aplicada para previsão de cheias de 1 a 10 h à frente. São usadas séries temporais de nível de água por hora para a estação das monções de cinco estações de medição na bacia do rio Mahanadi, na Índia. O método de reamostragem de bootstrap é usado para gerar diferentes realizações dos conjuntos de dados recém-construídos usando a DWT para criar um conjunto de amostras de bootstrap que são finalmente usadas como entrada para desenvolver o modelo wavelet–bootstrap–ANN (WBANN). O desempenho do modelo WBANN também é comparado com três modelos diferentes de RNAs: RNAs tradicionais, RNAs baseadas em wavelet e RNAs baseadas em bootstrap. Os resultados obtidos indicam que o modelo de previsão WBANN é superior ao tradicional ANN, bootstrap-ANN e WANN para previsão de 1 a 10 horas à frente.

Pramanik, Panda e Singh(37) propuseram modelos híbridos de wavelet e redes neurais artificiais para previsão de vazão de 1, 2 e 3 dias à frente em um local do rio Brahmani, na Índia. Os coeficientes wavelet efetivos são usados como entrada para os modelos de rede neural. Os resultados são comparados com modelos de RNA convencional e revelaram que as previsibilidades dos modelos híbridos são significativamente superiores aos modelos de RNA convencionais. Os picos altos de vazão são previstos com melhor precisão usando modelos híbridos RNA.

Santos e Silva(38) propuseram modelos híbridos de wavelets e redes neurais artificiais para previsão de vazão diária em 1, 3, 5 e 7 dias à frente no rio São Francisco, Brasil, com base nos componentes de baixa frequência do sinal original (aproximações). Demonstrou-se que os modelos híbridos propostos apresentam resultados significativamente melhores do que o modelo clássico de rede neural artificial para todas as situações testadas. Para

previsões de curto prazo (um dia à frente), informações sobre os componentes do sinal de alta frequência foram essenciais para garantir o bom desempenho do modelo. No entanto, para prever mais dias à frente, foram necessários componentes de baixa frequência como entrada para os modelos híbridos propostos.

Maslova, Ticlavilca e McKee(39) propuseram a técnica para uma previsão de vazão mensal de longo prazo com vários passos à frente. Um modelo baseado em decomposição wavelet e uma abordagem multivariada de aprendizado de máquina bayesiano é desenvolvido para previsão de vazão de 3, 6, 9 e 12 meses à frente simultaneamente. As entradas do modelo utilizam apenas os registros de vazões mensais anteriores. Eles são decompostos em componentes formulados em termos de análise multiresolução wavelet (MRA). Um estudo de simulação é realizado usando dados sintéticos e reais de vazão do rio Yellowstone na Bacia do Uinta em Utah. O modelo baseado na combinação das técnicas de wavelet e regressão Bayesiana de aprendizado de máquina (WMVRVM) é comparado com os modelos ANN. Demonstrou-se que o modelo Wavelet-MVRVM superou os modelos ANN.

O estudo de Shafaei e Kisi(40) investigou a capacidade de wavelet-redes neurais artificiais (WANN) para a previsão de vazão de curto prazo (um, dois e três dias à frente) da estação Vanyar no rio Ajichai no Irã. A série temporal de vazão foi decomposta em três níveis de decomposição usando DWT e wavelet mãe de Daubechies 5. O somatório de componentes de frequência efetiva com diferentes combinações foi usado como entrada do modelo RNA. Os resultados indicaram que os modelos WANN com capacidade de modelar relações complexas, não estacionárias e não lineares tiveram um desempenho melhor que os modelos SVM e ANN.

Londhe e Narkhede(26) desenvolveram modelos WANN para previsão de vazão de um, três, cinco e sete dias à frente em algumas estações de duas bacias hidrográficas diferentes na Índia e compararam os resultados com modelos ANN convencionais. Testou-se diferentes níveis de decomposição e concluiu-se que o desempenho de todos os níveis de decomposição (três, cinco e sete) foi excelente em termos de eliminação do atraso de fase em todos os horizontes de previsão em comparação com redes neurais convencionais. Demonstrou-se que a decomposição multinível de dados de fluxo usando a transformada wavelet ajuda a melhorar a precisão da previsão.

O estudo de Alexander, Thampi e NR(41) explora o potencial dos modelos híbridos wavelet-ANN para modelar eventos de inundação na bacia do rio Achankovil em Kerala. O nível de água por hora é previsto com horizontes de uma, três e seis horas à frente. Os três modelos híbridos wavelet-ANN desenvolvidos foram treinados considerando várias combinações de eventos de inundação. A decomposição dos sinais de entrada usando wavelet foi realizada selecionando uma wavelet mãe adequada e um nível de decomposição adequado. Os resultados demonstram a capacidade dos modelos híbridos de prever inundações satisfatoriamente com até 6 h de antecedência. A capacidade do modelo

híbrido é comprovada em comparação a um modelo convencional de RNA. Os modelos WANN desenvolvidos foram capazes de prever a magnitude do valor de pico e o tempo para o valor de pico com mais precisão quando comparado a modelos RNA.

O estudo de Santos et al.(42) propôs um modelo híbrido wavelet-rede neural artificial para previsão diária de influxo de curto prazo. Utilizou-se dados da Tropical Rainfall Measuring Mission (TRMM) juntamente com dados de influxo, que foram transformados usando wavelets para melhorar o desempenho do modelo. Os modelos foram avaliados usando os registros de afluência para um reservatório brasileiro denominado Três Marias, localizado na bacia do rio São Francisco, e estimativas diárias de precipitação do TRMM. Utilizou-se a transformada wavelet discreta como uma técnica para remover os detalhes (ou componentes de alta frequência) do sinal bruto dos dados de chuva do TRMM e afluências do reservatório de Três Marias. Várias combinações de entradas para redes neurais artificiais convencionais e híbridas foram avaliadas para prever fluxos de entrada com sete dias de antecedência e foi comprovado que o WANN teve um desempenho superior. Demonstrou-se que mesmo o modelo WANN, que utiliza apenas a aproximação no nível três dos dados pluviométricos, apresentou desempenho superior ao da RNA convencional, que utiliza os dados brutos de afluência.

O estudo de Sun, Niu e Sivakumar(19) investiga o desempenho do modelo autorregressivo de média móvel (ARMA), modelo de rede neural artificial e modelo de regressão linear para previsão de vazão de curto prazo, com ênfase na combinação de cada modelo com transformada wavelet. Os modelos convencionais e os modelos baseados em wavelets são aplicados para previsões de 1, 2 e 3 dias à frente de séries temporais diárias de vazão em duas bacias hidrográficas diferentes na China. Os resultados sugerem que os modelos WANN e ANN são mais adequados para a bacia norte, mas apresentam desempenho ruim para a bacia sul, enquanto o modelo wavelet ARMA é mais adequado para a bacia sul. Os resultados também sugerem que os modelos híbridos baseados em wavelet são melhores do que os modelos únicos, principalmente para *leadtimes* mais longos, oferecendo melhorias graduais com o aumento do *leadtime*. Ressaltou-se que os modelos baseados em wavelets são especialmente úteis para áreas que sofrem com desastres naturais frequentes, ambiente ecológico frágil e autorregulações precárias.

Freire, Santos e Silva(43) analisaram o uso de transformadas wavelet para prever vazões diárias no Reservatório de Sobradinho (Estado da Bahia, Brasil) sete dias à frente por um sistema híbrido wavelet-rede neural artificial. Determinou-se a wavelet mãe adequada para este tipo de previsão com uma RNA e comparou-se os resultados com as previsões feitas sem a aplicação de uma transformada wavelet. De acordo com os resultados, o sistema híbrido wavelet-ANN teve um desempenho melhor do que o sistema que utiliza a ANN com os dados brutos. A aproximação do nível três da wavelet mãe discreta de Meyer obteve os melhores resultados; a raiz quadrada do erro quadrático diminuiu aproximadamente 80%, enquanto o coeficiente  $R^2$  aumentou mais de 5%, em

comparação com a RNA convencional.

Em estudo de Dalkiliç e Hashimi(44) realizou-se a previsão da vazão diária usando os modelos baseados em wavelet e rede neural artificial e sistema de inferência fuzzy baseado em neural adaptativo (ANFIS). As previsões um dia à frente foram feitas em quatro estações diferentes localizadas no rio Buyuk Menderes, Turquia. O modelo WANN combinou-se pela aplicação da transformada wavelet discreta e rede neural artificial. As entradas para o modelo WANN foram obtidas através da mãe wavelet Haar. Demonstrou-se que o modelo WANN apresentou o melhor desempenho entre os modelos ANN e ANFIS.

Jain et al.(45) propuseram o modelo híbrido que combina wavelets e RNA para previsão de vazão de um dia à frente. O modelo foi aplicado para séries temporais de vazão diária do rio Malaprabha na Índia. As séries temporais observadas foram decompostas em subséries usando a transformada wavelet discreta e, em seguida, foram usadas como entradas para a rede neural. O modelo híbrido WANN foi comparado com os modelos convencionais de RNA e linear auto-regressivo (AR). Demonstrou-se que o modelo WANN é capaz de fornecer um bom ajuste com os dados observados, principalmente nos valores extremos e produzir os melhores resultados em comparação com os modelos ANN e AR.

Ni et al.(46) desenvolveram dois modelos híbridos, baseados em rede de memória de longo prazo (WLSTM), para previsão de vazão e precipitação mensal para 1, 3 e 6 passos à frente. Um modelo, WLSTM, aplicou a transformada wavelet *À trous* para fazer a decomposição das séries temporais, e o outro, CLSTM, uma rede neural convolucional acoplada com LSTM para extrair características temporais. Os modelos foram aplicados para prever vazão mensal e precipitação em estações no rio Yangtze na China. A precisão de predição de WLSTM e CLSTM foi comparada com a de multi-layer perceptron (MLP) e LSTM. Os resultados obtidos indicaram que a transformada wavelet e as camadas convolucionais melhoraram a precisão de previsão do LSTM, especialmente para previsões com passos de tempo mais longos. CLSTM e WLSTM foram alternativas superiores quando se espera uma previsão com passos de tempo mais longos.

Saraiva et al.(47) realizaram uma análise comparativa de um conjunto de modelos de aprendizado de máquina, a saber, uma RNA e uma SVM acoplada com uma transformada wavelet discreta e amostragem de dados com o método bootstrap, para uma previsão de vazão de 3 a 15 dias à frente. Utilizaram-se as séries temporais de vazão diária do reservatório de Sobradinho, localizado no nordeste do Brasil. Dentre as wavelets mãe analisadas (Daubechies, Coiflets, B-spline e Dmeyer), aquelas pertencentes à família Dmey apresentaram o melhor desempenho para a previsão de séries temporais de fluxo tanto com os modelos ANN quanto com os SVM. Além disso, a melhor combinação para previsão a vazão no reservatório de Sobradinho foi wavelet e rede neural com o método bootstrap.

Drisya, Kumar e Roshni(11) realizaram uma comparação entre o modelo de rede neural feedforward convencional (FFNN) e o modelo de rede neural artificial baseada em

wavelet para analisar sua eficácia na previsão de vazão mensal de seca para as regiões mais estressadas da bacia do rio Bharathapuzha na Índia. Os dados de séries temporais obtidos via transformada wavelet discreta são alimentados ao modelo FFNN para produzir um modelo WANN. O resultado demonstra que há uma melhora significativa na capacidade de previsão de vazão mensal por modelos WANN em comparação com modelos FFNN. A previsão de seca é realizada desenvolvendo um índice padronizado de vazão a partir da vazão prevista.

Yilmaz et al.(48) propuseram o modelo híbrido acoplado a transformada wavelet aditiva (AWT) e ANN para a previsão de vazão mensal de quatro estações na bacia de Çoruh, Turquia. Os modelos AWT-ANN foram comparados com modelos de ANN convencionais e modelos baseados em DWT e ANN. Usando AWT e DWT, as séries originais foram decompostas em subcomponentes que foram posteriormente usados como entradas para RNAs. Tanto o modelo AWT-ANN quanto o modelo DWT-ANN melhoraram o desempenho preditivo da ANN. No entanto, uma vez que os subcomponentes obtidos do AWT são mais compatíveis com a série original do que os do DWT, o modelo AWT-ANN demonstrou um desempenho superior ao modelo DWT-ANN na previsão de vazão média mensal.

Momeneh e Nourani(49) desenvolveram modelos híbridos que combinam uma rede neural artificial com transformada wavelet discreta e transformada wavelet multidiscreta (M-DWT) para prever o vazão em escala de tempo diária e mensal com um, dois, três e sete passos à frente. Os resultados indicaram que os modelos híbridos M-DWT redes neurais reduziram significativamente o erro de previsão em relação aos modelos unitários DWT-ANN e modelos usuais ANN para os próximos um, dois, três e sete dias e meses.

Xie et al.(50) desenvolveram o modelo híbrido combinando a transformada wavelet discreta com modelo de memória de longo prazo, WLSTM, para a previsão de vazão diária de um dia à frente nas estações fluviométricas do rio Yangtze na China. Os resultados demonstram que o modelo WLSTM é mais eficaz e supera o ANN e o LSTM convencionais na previsão de vazões extremas.

Um resumo do levantamento dos trabalhos relacionados à modelagem hidrológica híbrida baseada em wavelets é apresentado na Tabela 1. Destaca-se que na revisão bibliográfica feita nesse estudo, notou-se a tendência: quase todos os trabalhos revisados utilizam-se da transformada wavelet DWT em seus modelos de previsão de vazão propostos. Em contraste à maioria dos trabalhos relatados acima e com intuito de fornecer a abordagem que permita a aplicação em tarefas de previsão nos casos reais, o presente estudo adotou o uso do algoritmo de decomposição wavelet AT na modelagem de vazão média diária do rio Paraíba do Sul. Além disso, o presente estudo busca uma wavelet mãe adequada para uma modelagem de vazão do rio Paraíba do Sul em curto prazo.

Uma das arquiteturas baseadas em wavelets frequentemente encontradas na litera-

Tabela 1 – Resumo dos trabalhos pesquisados para previsão de séries hidrológicas, incluindo ano de publicação, autores, àrea de estudo, o métodos de IA acoplado com wavelets acoplados, o tipo de transformada wavelet, escala de tempo e variável de previsão, bem como a informação sobre a investigação das wavelets mãe e avaliação de vazões extremas.

<b>Autores</b>	<b>Ano</b>	<b>Área de estudo</b>	<b>Tipo de método de AI</b>	<b>Tipo de Transformada Wavelet</b>	<b>Seleção de Wavelet mãe</b>	<b>Horizonte de previsão</b>	<b>Avaliação dos picos</b>	<b>Escala de tempo, variável</b>
Adamowski e Sun(25)	2010	Rios Xeros, Kargotis/Cyprus	ANN	AT, DWT	-	1, 2, 3	-	diária, vazão
Tiwari e Chatterjee(36)	2010	Rio Mahanadi/Índia	ANN, BANN,	DWT	-	1-10	sim	por hora, nível de água
Pramanik, Panda e Singh(37)	2011	Rio Brahmani/India	ANN	DWT	-	1,2,3	sim	diária, vazão
Santos e Silva(38)	2014	Reservatório Sobradinho, Rio São Francisco	ANN	DWT	-	1, 3, 5, 7	-	diária, vazão
Maslova, Ticlavilca e McKee(39)	2016	Rio Yellowstone/EUA	MVRVM, ANN	MRA	-	3, 6, 9, 12	-	mensal, vazão
Shafaei e Kisi(40)	2017	Rio Ajichai/Iran	ANN	DWT	-	1-4	-	diária, vazão
Londhe e Narkhede(26)	2018	Bacias Krishna, Narmada/India	ANN	DWT	sim	1-7	-	diária, vazão
Alexander, Thampi e NR(41)	2018	Rio Achankovil/Kerala	ANN	DWT	sim	1, 3, 6	sim	por hora, vazão
Santos et al.(42)	2019	Três Marias reservoir/Brazil	ANN	DWT	sim	7	-	diária, vazão
Sun, Niu e Sivakumar(19)	2019	Heihe River/China	ARMA, ANN	DWT	-	1-3	-	diária, vazão
Freire, Santos e Silva(43)	2019	Sobradinho Reservoir/Brazil	ANN	DWT	sim	7	-	diária, vazão
Dalkiliç e Hashimi(44)	2020	Büyük Menderes river/Turkey	ANN	DWT	-	1	-	diária, vazão
Jain et al.(45)	2020	Bacia do Rio Malaprabha/India	ANN	DWT	-	1	-	diária, vazão
Ni et al.(46)	2020	rio Yangtze/China	LSTM	AT	-	1, 3, 6	-	mensal, vazão, precipitação
Saraiva et al.(47)	2021	Reservatório Sobradinho/Brazil	ANN, SVR, BANN, BSVR	DWT	sim	1,3,5,7,10,15	-	diária, vazão
Drisy, Kumar e Roshni(11)	2021	Rio Bharathapuzha/India	ANN	DWT	-	1,2,3,6,9,12	-	diária, mensal, vazão
Yilmaz et al.(48)	2022	Rio Çoruh/Turkey	ANN	DWT, AWT	-	1	-	mensal, vazão
Momenh e Nourani(49)	2022	Rio Gamasiab/Iran	ANN	DWT, M-DWT	-	1,2,3,7	-	diária, mensal, vazão
Xie et al.(50)	2022	Rio Yangtze/China	LSTM	DWT	-	1	sim	diária, vazão

Fonte: Elaborada pelo autor (2022).

tura (51, 52, 53, 54, 47) é a arquitetura que utiliza sub-séries de decomposição wavelet da série temporal original como dados de entrada para o método de aprendizado de máquina. Neste estudo, além dessa arquitetura, utilizou-se uma segunda arquitetura que combina redes neurais artificiais, usadas para modelar os componentes da frequência da série temporal original obtidas através da WT, as quais, em seguida, são combinadas em processo de reconstrução da série temporal original através da WT inversa. A segunda arquitetura é menos usada, mas também é encontrada na literatura, por exemplo, nos trabalhos de (26, 47). Ao contrário dos trabalhos citados acima, a nova abordagem híbrida proposta utiliza como dados de entrada para redes neurais, além das componentes da frequência da série temporal original, ainda, os dados da série temporal original. Em processo de desenvolvimento de modelos notou-se que essa abordagem aumenta significativamente a eficiência das previsões de vazão.

### 3 MÉTODOS

A seguir são apresentados os principais conceitos de métodos empregados neste estudo.

#### 3.1 PRINCIPAIS CONCEITOS DA TRANSFORMADA WAVELET

##### 3.1.1 Da Transformada de Fourier à Transformada Wavelet

Transformações matemáticas são aplicadas a sinais para obter uma informação adicional daquele sinal que não está prontamente disponível no sinal bruto. A maioria dos sinais, na prática, estão no domínio do tempo em seu formato bruto. Quando são plotados sinais no domínio do tempo, é obtida uma representação de amplitude de tempo deste sinal. Muitas vezes, as informações que não podem ser prontamente vistas no domínio do tempo podem ser vistas no domínio da frequência. Em muitos casos, a informação mais distinta está escondida no conteúdo de frequência do sinal. O espectro de frequência de um sinal mostra quais frequências existem no sinal.

Em geral, se algo (uma variável matemática ou física) muda rapidamente, pode ser dita como de alta frequência. Se essa variável não muda rapidamente, ou seja, muda suavemente, pode ser dita como de baixa frequência. Se essa variável não muda, ela tem frequência zero ou nenhuma frequência. Há muitas mudanças rápidas em séries temporais hidrológicas não estacionárias. Para obter a representação da amplitude de frequência de um sinal é utilizada a transformada de Fourier (FT).

A análise de Fourier é uma transformação do sinal bruto denotado no domínio do tempo em um sinal processado que é denotado no domínio da frequência. FT informa sobre a frequência do sinal, ou seja, explica quanto de cada frequência existe no sinal mas não informa quando no tempo esses componentes de frequência existem (55).

Se um sinal contém uma frequência de  $x$ Hz, então, esta frequência deve estar presente igualmente em qualquer lugar do sinal. Esta é uma séria desvantagem se as propriedades do sinal mudarem muito ao longo do tempo, ou seja, se contiverem características não estacionárias ou transitórias (desvio, tendências, mudanças abruptas ou início e fim de eventos). Essas características geralmente são a parte mais importante de uma série temporal e a FT não é adequada para detectá-las (56).

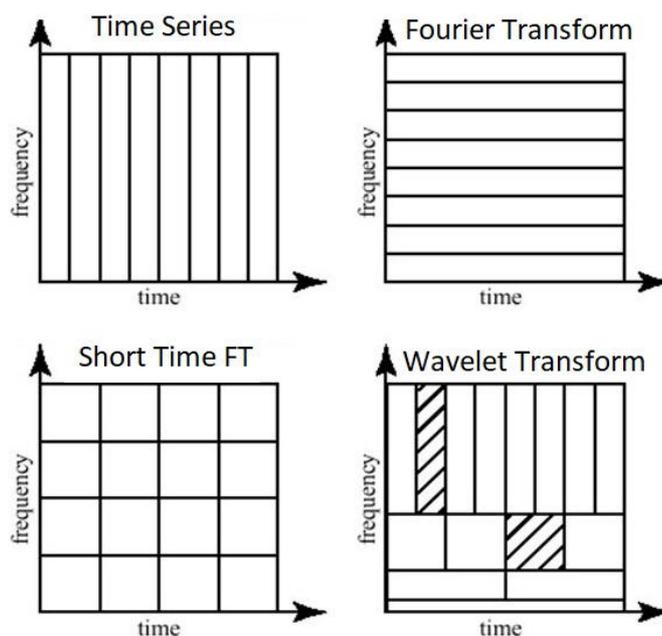
A fim de alcançar um compromisso entre frequência e tempo, a FT foi expandida em transformada de Fourier de curto termo (*Short-Time Fourier Transform*, em inglês - STFT). Nesta abordagem, o sinal é dividido em partes suficientemente pequenas (janelas) de igual comprimento usando uma janela deslizante, onde essas partes podem ser consideradas estacionárias. Como o comprimento das partes divididas é conhecido, pode-se explorar informações sobre a localização no tempo. Como um período de tempo do nosso sinal

é extraído, uma posição do período é conhecida e informações de frequência para esse tamanho de janela são obtidas. Portanto, consegue-se uma resolução constante do nosso sinal em questão de representação linear de frequência no tempo.

O principal problema com essa abordagem é atingir os limites teóricos da transformada de Fourier, conhecida como Princípio da Incerteza de Heisenberg. Quanto menor for o tamanho da janela, mais saberemos sobre onde a frequência apareceu no sinal, mas menos sobre o valor da própria frequência. Quanto maior for o tamanho da janela, mais será conhecido o valor da frequência e menos o tempo (55).

A Transformada Wavelet (WT), por outro lado, não é localizada apenas no tempo e na frequência, mas também extrapola uma divisão fixa de tempo-frequência. Ela não apenas diz quais frequências estão presentes no sinal, mas também em que momento essas frequências ocorreram. Isto é conseguido trabalhando com diferentes escalas. Primeiro, um sinal com uma grande escala/janela é observado e as características “grandes” são analisadas. Então, observa-se um sinal com uma escala menor para analisar as características menores, ou seja, wavelets têm a capacidade de decompor um sinal ou série temporal em diferentes níveis. Como resultado, esta decomposição revela a estrutura do sinal principal, bem como tendências, periodicidades, singularidades ou saltos que não podem ser observados inicialmente (57).

Figura 1 – Uma visão geral esquemática das resoluções de tempo e frequência das diferentes transformações em comparação com o conjunto de dados de série temporal original. O tamanho e as orientações do bloco fornecem uma indicação do tamanho da resolução.



Fonte: Adaptado de (58).

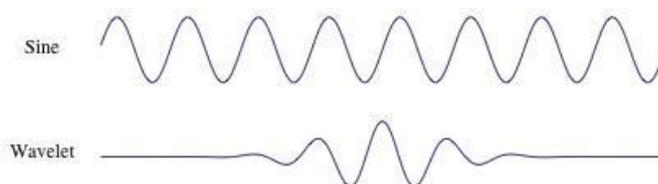
Na Figura 1 podem ser vistas as resoluções de tempo e frequência das diferentes transformações. O tamanho e a orientação dos blocos indicam quão pequenas são as características que são distinguidas no domínio do tempo e da frequência. A série temporal original tem alta resolução no domínio do tempo e resolução zero no domínio da frequência. Isso significa que pode-se distinguir características muito pequenas no domínio do tempo e nenhuma característica no domínio da frequência. Oposta a isso está a FT, que tem alta resolução no domínio da frequência e resolução zero no domínio do tempo. A STFT possui resolução de tamanho médio tanto no domínio da frequência quanto no domínio do tempo. A Transformada Wavelet tem:

- para pequenos valores de frequência - uma alta resolução no domínio da frequência, baixa resolução no domínio do tempo;
- para grandes valores de frequência - uma baixa resolução no domínio da frequência, alta resolução no domínio do tempo.

Em outras palavras, a transformada wavelet faz um *trade – off*: em escalas em que as características dependentes do tempo são interessantes, ela tem uma alta resolução no domínio do tempo; e em escalas em que as características dependentes da frequência são interessantes, tem uma alta resolução no domínio da frequência.

A abordagem wavelet é muito flexível no tratamento de séries de dados muito irregulares (57) e pode ser uma ferramenta valiosa para analisar uma ampla variedade de séries temporais e já foi usada com sucesso no processamento de imagens, eliminação de ruído de sinal, estimativa de densidade, compressão de sinal e imagem e, decomposição no tempo. As técnicas de wavelet são usadas em finanças para descobrir propriedades de valores que mudam rapidamente (56).

Figura 2 – A diferença entre uma onda senoidal e uma Wavelet. A onda senoidal é infinitamente longa e a Wavelet está localizada no tempo.



Fonte: Extraído de (59).

### 3.1.2 Wavelets

A transformada de Fourier usa uma série de ondas senoidais com diferentes frequências para analisar um sinal. Ou seja, um sinal é representado por uma combinação linear de ondas senoidais. A análise da transformada wavelet usa uma série de funções semelhantes

a ondas localizadas chamadas *wavelets*, cada uma com uma escala diferente. O termo *wavelet* significa uma pequena onda. A pequenez refere-se à condição de que esta função é de comprimento finito (suporte compactado). O *wave* refere-se à condição de que esta função é oscilatória.

Na Figura 2 pode ser vista a diferença entre uma onda senoidal e uma wavelet. A principal diferença é que a onda senoidal não está localizada no tempo (ela se estende de  $-\infty$  a  $+\infty$ ) enquanto uma wavelet está localizada no tempo. Esta localização no tempo implica que a maior parte da energia está restrita a um intervalo finito e a função é idealmente zero fora deste intervalo. Isso permite que a transformada wavelet obtenha informações de tempo além das informações de frequência. Matematicamente, uma wavelet  $\psi(t)$  é simplesmente uma função do tempo  $t$  (chamada de *wavelet mãe*) que obedece a uma regra básica, conhecida como condição de admissibilidade wavelet (55):

$$C_\psi(t) = \int_0^\infty \frac{|\psi(f)|}{f} df < \infty, \quad (3.1)$$

onde  $\psi(f)$  é a transformada de Fourier, uma função da frequência  $f$ , de  $\psi(t)$ . Essa condição garante que  $\psi(f)$  vá para zero rapidamente como  $f \rightarrow 0$ . De fato, para garantir que  $C_\psi(t) < \infty$  devemos impor  $\psi(0) = 0$ , que é equivalente a

$$\int_{-\infty}^\infty \psi(t) dt = 0. \quad (3.2)$$

Uma condição secundária imposta a uma função wavelet é a energia unitária:

$$\int_{-\infty}^\infty |\psi(t)|^2 dt = 1. \quad (3.3)$$

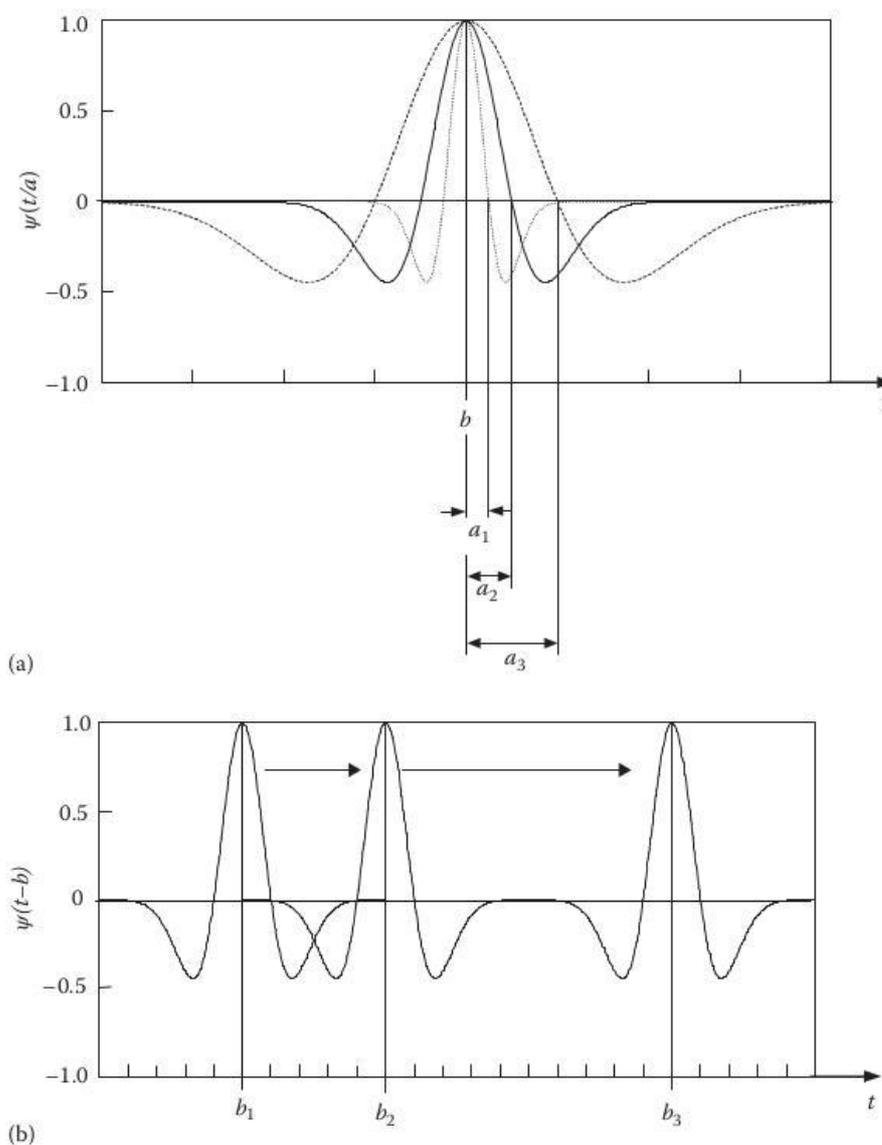
Satisfazendo ambas as Equações 3.2 e 3.3 a função wavelet deve ter entradas diferentes de zero mas todos os desvios de zero devem ser cancelados.

Dada uma wavelet mãe  $\psi(t)$ , para todo real  $a, b$  ( $a \neq 0$ ), pode ser construída uma sequência de wavelets que é manipulada através de um processo de translação (ou seja, movimentos ao longo do eixo-tempo) e dilatação (ou seja, o espalhamento da wavelet) de  $\psi(t)$  para transformar o sinal em outra forma que “desdobra” em tempo e escala:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad (3.4)$$

onde  $a$  é a escala ou parâmetro de dilatação e  $b$  é o deslocamento ou parâmetro de translação. O valor do parâmetro de escala  $a$  determina o nível de alongamento ou compressão. Versões deslocadas e dilatadas da wavelet mãe são denotadas como  $\psi\left(\frac{t-b}{a}\right)$ . O termo  $\frac{1}{\sqrt{|a|}}$  normaliza  $\|\psi(t)\| = 1$ . A análise temporal é realizada com uma versão da wavelet mãe contraída e de alta frequência, enquanto a análise de frequência é realizada com uma versão de baixa frequência dilatada da mesma wavelet mãe.

Figura 3 – Dilatação e translação de uma wavelet: (a) Alongamento e compressão de uma wavelet: Dilatação ( $a_1 = a_2/2$ ;  $a_3 = a_1 \times 2$ ). (b) Movendo uma wavelet: Translação.



Fonte: Extraído de (59).

A Figura 3(b) mostra o movimento de uma wavelet ao longo do eixo do tempo de  $b_1$  via  $b_2$  até  $b_3$ . O parâmetro *scale* na análise wavelet é semelhante à escala usada nos mapas. Como no caso dos mapas, as escalas altas correspondem a uma visão global não detalhada (do sinal) e as escalas baixas correspondem a uma visão detalhada. Da mesma forma, em termos de frequência, as baixas frequências (escalas altas) correspondem a uma informação global de um sinal (que geralmente abrange todo o sinal), enquanto as frequências altas (escalas baixas) correspondem a uma informação detalhada de um padrão oculto no sinal (que geralmente dura um tempo relativamente curto). O termo de escala é usado no denominador e, portanto, vale o oposto das declarações acima, ou seja, as escalas  $a > 1$  dilatam os sinais, enquanto as escalas  $a < 1$  comprimem o sinal.

### 3.1.3 Transformada Wavelet Contínua e Transformada Wavelet Discreta

Duas versões da WT podem ser distinguidas: a transformada wavelet contínua (CWT) e a transformada wavelet discreta (DWT). A CWT é uma correlação entre uma wavelet em diferentes escalas e o sinal com a escala (ou a frequência) sendo usada como medida de similaridade. A CWT é definida como a soma ao longo de todo o tempo do sinal multiplicado por versões escalonadas e deslocadas da função wavelet, usando um intervalo de  $a$  e  $b$  (59):

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi \left( \frac{t - b}{a} \right) dt, \quad (3.5)$$

onde  $\psi(t)$  é a wavelet mãe contínua que é escalada por um fator de  $a$  e translataada por um fator de  $b$ . Os valores dos fatores de escala e translação são contínuos, o que significa que pode haver uma quantidade infinita de wavelets.

A diferença entre CWT e Transformada Wavelet Discreta está no conjunto de escalas e posições em que cada transformação opera. A CWT pode operar em todas as escalas. No entanto, um limite superior é determinado, pois a CWT é extremamente cara computacionalmente (57). Para reduzir a carga computacional, em DWT os coeficientes wavelet são calculados apenas em um subconjunto de escalas (ou seja, valores discretos dos parâmetros de dilatação e translação,  $a$  e  $b$ , são usados). Portanto, a DWT fornece informações suficientes tanto para análise quanto para síntese do sinal original com uma redução significativa no tempo de computação (57).

Uma maneira natural de amostrar os parâmetros  $a$  e  $b$  é usar uma discretização logarítmica da escala  $a$  e vinculá-la, por sua vez, ao tamanho dos passos dados entre  $b$  localizações. Para vincular  $b$  a  $a$ , move-se em etapas discretas para cada local  $b$  que é proporcional à escala  $a$ . Este tipo de discretização da wavelet tem a forma:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} \psi \left( \frac{t - nb_0 a_0^m}{a_0^m} \right), \quad (3.6)$$

onde o inteiro  $m$  controla a dilatação da wavelet; o inteiro  $n$  controla a translação da wavelet;  $a_0$  é um parâmetro de etapa de dilatação fixo definido em um valor maior que 1;  $b_0$  é o parâmetro de localização que deve ser maior que zero.

Para eficiência computacional,  $a_0 = 2$  e  $b_0 = 1$  são comumente usados para que os resultados levem a uma dilatação binária de  $2^{-m}$  e uma translação diádica de  $n2^m$ . Substituindo  $a_0 = 2$  e  $b_0 = 1$  na Equação 3.6, vê-se que a wavelet da grade diádica pode ser escrita como

$$\psi_{m,n}(t) = 2^{-m/2} \psi \left( 2^{-m} t - n \right), \quad (3.7)$$

Wavelets discretas de uma grade diádica são geralmente escolhidas para serem ortonormais. Essas wavelets são ortogonais entre si e normalizadas por unidade de energia.

Usando a wavelet da grade diádica da Equação 3.7, a DWT pode ser escrito como

$$T(m, n) = \int_{-\infty}^{\infty} x(t) \psi_{m,n}^*(t) dt, \quad (3.8)$$

onde  $*$  denota o conjugado complexo,  $T(m, n)$  são os valores da transformada wavelet discreta dados em uma grade de localização de escala de índice  $m, n$ . Os valores  $T(m, n)$  também são conhecidos como coeficientes *wavelet* ou coeficientes de detalhe. Escolhendo uma base de wavelet ortonormal,  $\psi_{m,n}(t)$ , pode-se reconstruir o sinal original em termos dos coeficientes de wavelet,  $T(m, n)$ , usando a transformada de wavelet discreta inversa como segue:

$$x(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} T(m, n) \psi_{m,n}(t), \quad (3.9)$$

exigindo a soma sobre todos os inteiros  $m$  e  $n$ .

### 3.1.4 A Transformada Wavelet Discreta como um banco de filtros

Na prática, a DWT é implementada como um banco de filtros (uma cascata de filtros passa-alta e passa-baixa). Em outras palavras, a DWT analisa o sinal em diferentes bandas de frequência com diferentes resoluções decompondo o sinal em uma aproximação grosseira e informações detalhadas (60). Normalmente, o conteúdo de baixa frequência é a parte mais importante de um sinal. É o que dá identidade ao sinal, enquanto a componente de alta frequência geralmente contém uma grande parte do ruído no sinal (57).

Para aplicar a DWT em um sinal, começa-se com a menor escala. Como já foi explicado anteriormente, pequenas escalas correspondem a altas frequências (filtro passa-altas). Isso significa que primeiramente analisa-se o comportamento de alta frequência. No segundo nível, a escala aumenta com um fator de dois (a frequência diminui com um fator de dois) e analisa-se o comportamento em torno da metade da frequência máxima. No terceiro nível, o fator de escala é quatro e analisa-se o comportamento da frequência em torno de um quarto da frequência máxima. E isso continua até que o número de amostras em sinal se torne menor que o comprimento do filtro wavelet e seja atingido o nível máximo de decomposição. O filtro passa-baixo vai deixar passar os componentes de baixa frequência do sinal mas ele vão rejeitar os componentes de alta frequência em cada nível de decomposição e assim gerando coeficientes de aproximação. Por outro lado, o filtro passa-alto vai gerar coeficientes de detalhe. O número de coeficientes de aproximação e detalhe é reduzido pela metade a cada vez para cada nível de decomposição. Isso é chamado DWT dizimada.

Matematicamente, o filtro passa-alta leva a uma função wavelet  $\psi(t)$ , enquanto o filtro passa-baixa leva a uma função detalhe  $\phi(t)$ . A função detalhe  $\phi(t)$  pode ainda ser reduzida para o próximo nível da função wavelet e da função detalhe usando o banco de filtros recursivos do filtro passa-baixa - filtro/passa-alta e amostragem descendente. A wavelet  $\psi(t)$  é a saída da função detalhe filtrada passa-alta. O último nível de decomposição

produz uma função detalhe com um coeficiente chamado aproximação wavelet. A função detalhe está associada à suavização do sinal e tem a mesma forma da wavelet. A equação de dilatação para a função detalhe  $\phi(t)$  e a equação wavelet para a função wavelet  $\psi(t)$  em termos dos coeficientes de filtro originais são dadas por (60, 61):

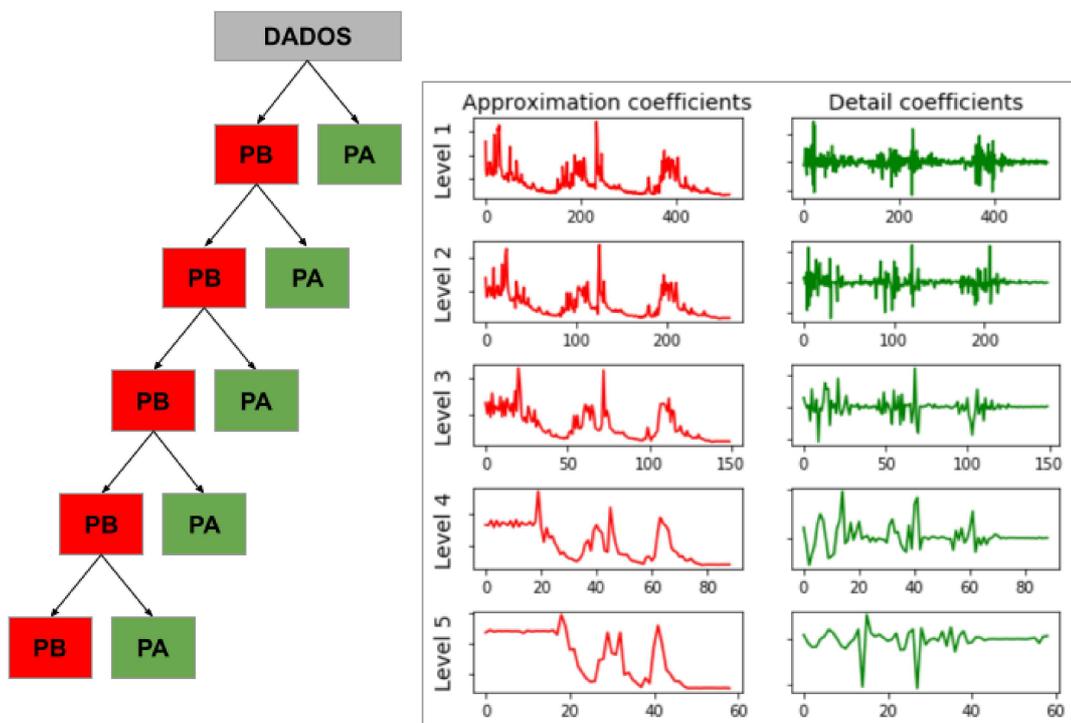
$$\phi(t) = \sqrt{2} \sum_{k=0}^N h(k) \phi(2t - k), \quad (3.10)$$

$$\psi(t) = \sqrt{2} \sum_{k=0}^N g(k) \phi(2t - k), \quad (3.11)$$

onde  $h(k)$  são os coeficientes do filtro passa-baixa e  $g(k)$  são os coeficientes do filtro passa-alta.

Pode ser notada a presença de duas escalas de tempo  $t$  e  $2t$  nas equações acima. Os coeficientes do filtro passa-alta  $g(k)$  estão relacionados aos coeficientes do filtro passa-baixa  $h(k)$  por  $g(k) = (-1)^k h(N - k)$ , onde  $N$  é o total número de coeficientes. Portanto, para uma dada wavelet, apenas os coeficientes do filtro passa-baixa (coeficientes de aproximação) precisam ser conhecidos.

Figura 4 – Os coeficientes de aproximação e detalhe da onda *coif5* aplicada em um sinal de vazão, do nível 1 ao 5. À esquerda pode-se ver uma representação esquemática dos filtros passa-alta (PA) e passa-baixa (PB) aplicados no sinal em cada nível.



Fonte: Elaborada pelo autor (2022).

Por exemplo, a Figura 4 demonstra a decomposição de um sinal de vazão em 5 níveis. PB e PA denominam filtros passa-baixa e passa-alta, respectivamente.

A aproximação wavelet de um sinal  $x(t)$  é definida por (61):

$$x(t) = \sum_k a_{J,k} \phi_{J,k}(t) + \sum_k d_{J,k} \psi_{J,k}(t) + \sum_k d_{J-1,k} \psi_{J-1,k}(t) + \cdots + \sum_k d_{1,k} \psi_{1,k}(t), \quad (3.12)$$

onde  $k$  abrange todos os valores de dados e  $J$  é o número de níveis de decomposição. Os coeficientes  $a_{J,k}, d_{J,k}, d_{J-1,k}, \dots, d_{1,k}$  são coeficientes de transformada wavelet dados pelas decomposições

$$a(J, k) = \int x(t) \phi_{J,k}(t) dt, \quad (3.13)$$

$$d(j, k) = \int x(t) \psi_{j,k}(t) dt, \quad j = 1, 2, \dots, J. \quad (3.14)$$

A magnitude desses coeficientes dá uma medida da contribuição da função wavelet correspondente para o sinal total. As funções básicas  $\psi_{j,k}$  e  $\phi_{j,k}$ ,  $j = 1, 2, \dots, J$  são as funções wavelet aproximadas geradas como versões de escala e translação de  $\phi$  e  $\psi$  com fator de escala  $2^j$  e parâmetro de translação  $2^j k$  respectivamente, definidos como

$$\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k) = 2^{-j/2} \phi\left(\frac{t - 2^j k}{2^j}\right) \quad (3.15)$$

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k) = 2^{-j/2} \psi\left(\frac{t - 2^j k}{2^j}\right), \quad j = 1, 2, \dots, J \quad (3.16)$$

O parâmetro de translação  $2^j k$  é combinado com o parâmetro de escala  $2^j$  de forma que, à medida que as funções  $\phi_{j,k}$  e  $\psi_{j,k}$  se alargam, seus passos de translação são correspondentemente maiores.

### 3.1.5 Algoritmo À Trous de decomposição wavelet

O algoritmo À Trous (AT), proposto por (62), é uma versão não dizimada da DWT. De acordo com (30, 31), o algoritmo AT é considerado o algoritmo de decomposição mais adequado para fins de previsão hidrológica. A ideia básica da transformada wavelet à *trous* é preencher as lacunas resultantes usando informações redundantes obtidas da série original com a vantagem de que as informações adicionais fornecem uma base para uma precisão de previsão aprimorada, além de promover uma compreensão abrangente das propriedades subjacentes do processo em termos de seus dados observados (30).

Ao inserir zeros para criar buracos (*trous*, em francês) no banco de filtros usado na transformada wavelet biortogonal dizimada padrão (Equação 3.12), cada banda de decomposição tem o mesmo tamanho que a imagem original, ou seja, é redundante e a dizimação não é usada. A redundância é útil para detectar características finas nos sinais de detalhes, uma vez que não surgem vieses de *aliasing* por meio da dizimação.

O AT como banco de filtros da série  $x(t) = c_0(t)$ , leva a um conjunto  $\{d_1, \dots, d_J, c_J\}$ , onde  $d_j$  são os coeficientes de detalhe em nível  $j$  e  $c_j$  são os coeficientes de aproximação

(versões mais suaves de  $x(t)$  em diferentes resoluções). A passagem de uma resolução para a seguinte é obtida por (62, 63):

$$c_{j+1}(t) = (\bar{h}_j * c_j)(t) = \sum_{k=-\infty}^{\infty} h(k)c_j(t + 2^j k). \quad (3.17)$$

O componente de detalhe de  $x(t)$  no nível  $j + 1$  é definido como

$$d_{j+1}(t) = c_j(t) - c_{j+1}(t), \quad (3.18)$$

onde  $*$  significa produto de convolução,  $k \in \mathbb{Z}$ ,  $t = 0, \dots, 2^m$ ,  $2^m$  - o tamanho do sinal original,  $j = 1, \dots, J$ ,  $j = 0$  é o nível mais fino de resolução,  $J$  é o número de níveis de decomposição,  $h(k)$  é um filtro passa-baixa com suporte compacto (62).

Para qualquer nível  $j \geq 1$ , o filtro  $h(k)$  dilatado por  $2^j$ , portanto, é definido por (64):

$$\bar{h}_j(k) = \begin{cases} h(-k/2^j), & \text{if } k/2^j \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases} \quad (3.19)$$

Pode-se ver pela Equação 3.19 que o filtro  $\bar{h}_1(k)$  é obtido inserindo-se  $2^j - 1$  zeros entre cada amostra de  $h(k)$ . A inserção de zeros no filtro cria buracos (*trous*). Por exemplo,

$$\bar{h}_1(k) = (\dots, h(-2), 0, h(-1), 0, h(1), 0, h(0), 0, h(1), 0, \dots) \quad (3.20)$$

A transformada inversa à *trous* é calculada adicionando os coeficientes wavelet de todos os níveis ao sinal final de baixa resolução, a última matriz suavizada  $c_J(t)$ , é adicionada a todas as diferenças  $d_j(t)$ , para qualquer tempo  $t$ :

$$c_0(t) = c_J(t) + \sum_{j=1}^J d_j(t) \quad (3.21)$$

A transformação 3.21 é não dizimada, ou seja, cada nível produz coeficientes de detalhe no tamanho do sinal original.

### 3.1.6 O problema de “dados futuros”

Como é mencionado na Sessão 2.5, ao lidar com modelagem de séries temporais, informações futuras não devem ser usadas em pontos de tempo passados.

O maior erro de interpretação em relação às previsões baseadas em wavelets e um “impedimento significativo” para seu uso em problemas de previsão do mundo real está relacionada à decomposição wavelets e à adição não considerada de erro (relacionada ao uso dos dados do futuro) no cálculo dos coeficientes da WT. Este problema é conhecido como “problema de dados futuros” em condição de contorno. Se for desconsiderado o

problema de contorno na decomposição wavelet, é provável que surja um problema de superestimação levando a resultados não confiáveis. (65).

A abordagem para lidar com o “problema de dados futuros” proposta por (31) é evitar o uso de coeficientes de aproximação e detalhe que contenham dados do futuro, escolhendo o algoritmo apropriado de decomposição (como, por exemplo, o algoritmo AT) e removendo os primeiros  $L_J$  de coeficientes de aproximação e detalhe determinados pela seguinte equação:

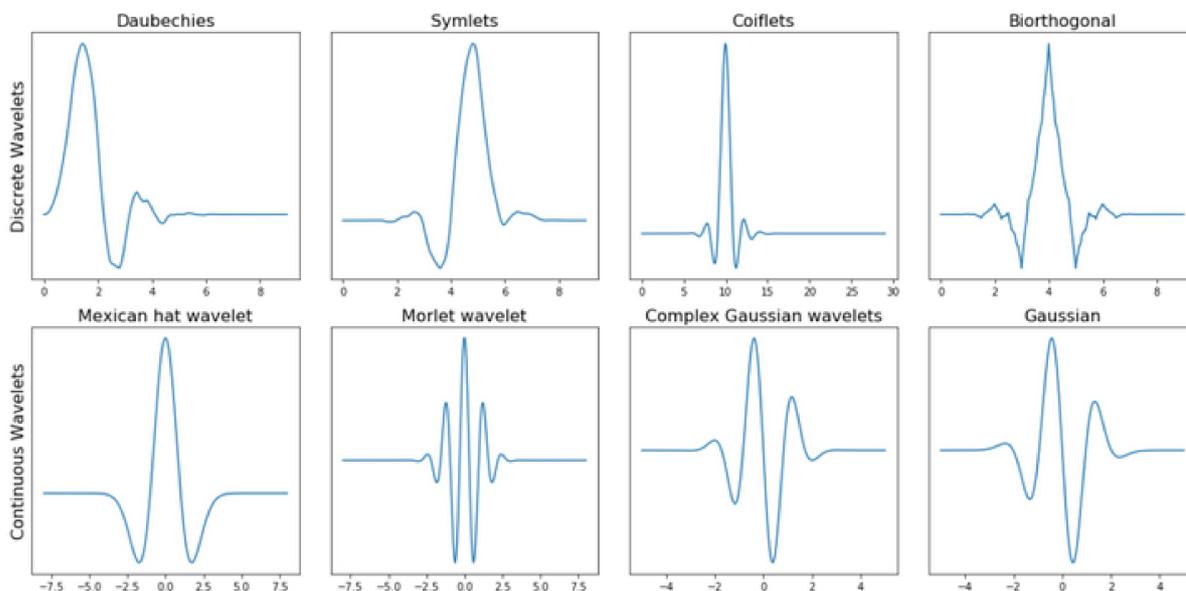
$$L_J = (2^J - 1)(L - 1) + 1, \quad (3.22)$$

onde  $L_J$  representa o número de e coeficientes de aproximação e detalhe afetados pelo “problema de dados futuros” para o nível de decomposição  $J$  e um filtro wavelet de comprimento  $L$ .

### 3.1.7 Famílias de Wavelet

Diversos wavelets tem sido apresentados pela comunidade de pesquisa em wavelet: cada wavelet mãe tendo diferentes formas de famílias e diferentes comprimentos o que leva a coeficientes de wavelet com propriedades diferentes (66).

Figura 5 – Várias famílias de Wavelets. Na primeira linha estão wavelets discretas e na segunda linha estão wavelets contínuas.

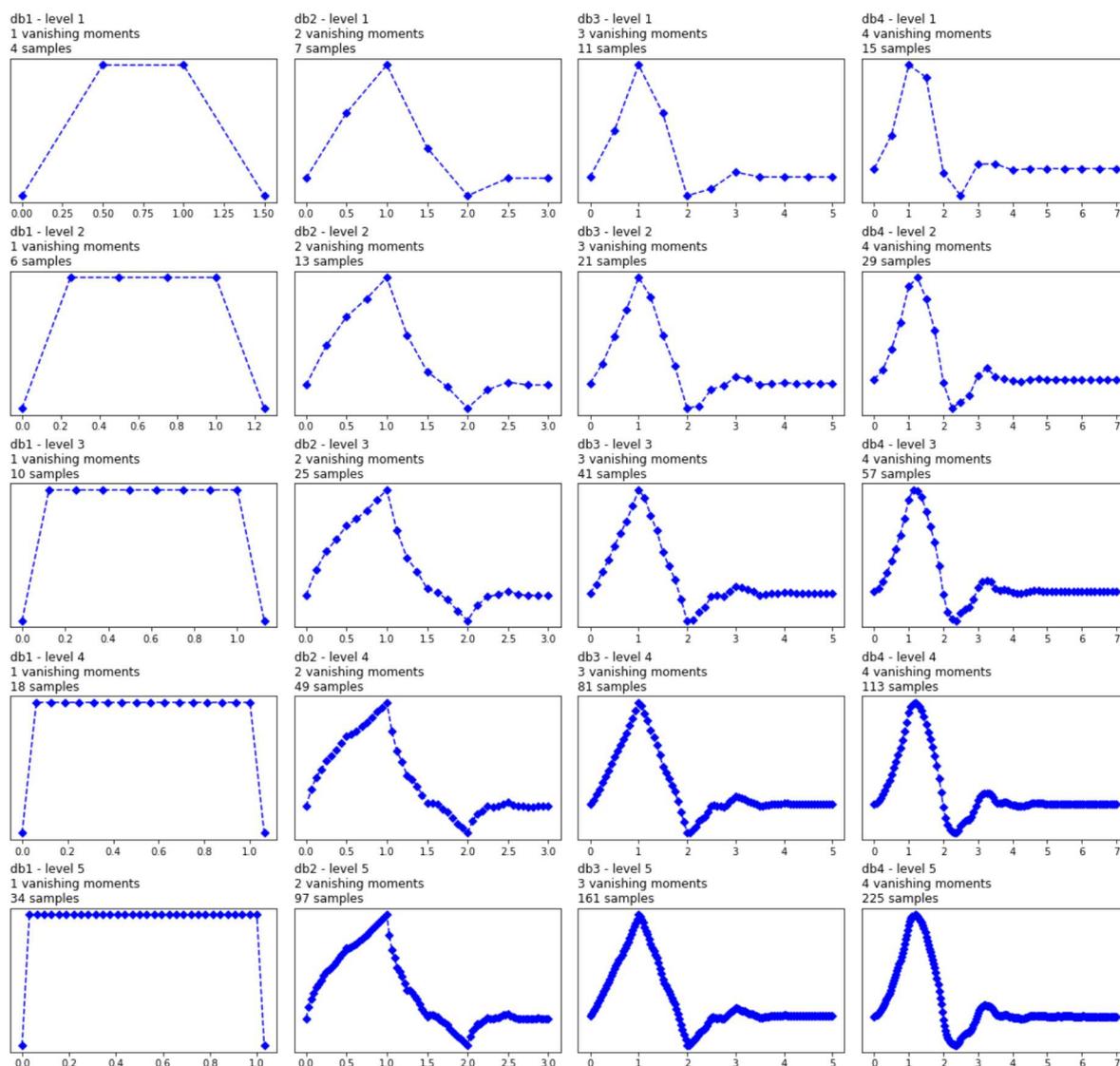


Fonte: Elaborada pelo autor (2022).

As qualidades da wavelet mãe variam de acordo com vários critérios: o comprimento do suporte da wavelet mãe, o número de momentos de desaparecimento, a simetria ou a regularidade. A região de suporte implica na extensão do comprimento de uma dada wavelet que, por sua vez, afeta suas capacidades de localização de características (67).

Uma função wavelet longa e amplamente distribuída calculará a amplitude instantânea do processo enquanto, ao mesmo tempo, abrange uma janela mais ampla do processo subjacente, resultando em um alto grau de média dos estados do processo (30). O momento de desaparecimento limita a capacidade da wavelet de representar adequadamente o comportamento ou informação polinomial em um sinal.(30) Dentro de cada

Figura 6 – A família de wavelet *Daubechies* para várias ordens diferentes de momentos de desaparecimento e vários níveis de refinamento.

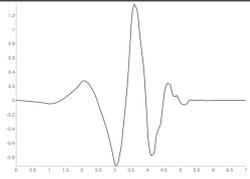
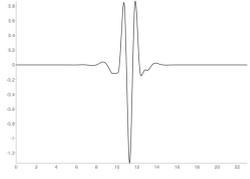
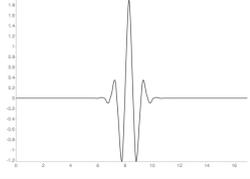
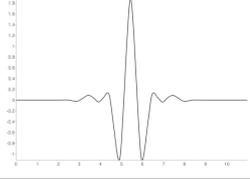
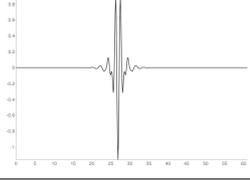


Fonte: Elaborada pelo autor (2022).

família de wavelets pode haver muitas subcategorias diferentes pertencentes a essa família. Distingui-se as diferentes subcategorias de wavelets pelo número de coeficientes (o número de momentos de desaparecimento) e o nível de decomposição. Na Figura 5 pode ser visto um gráfico com várias famílias diferentes de wavelets. A primeira linha contém quatro Wavelets discretas e a segunda linha quatro Wavelets contínuas.

As wavelets *Gaussianas*, *Morlet* e *Mexican hat* são wavelets grosseiras que só podem ser usadas em decomposição contínua. As wavelets da família de wavelets *Meyer* são wavelets infinitamente regulares que podem ser usadas em CWT e DWT. As wavelets das famílias *Daubechies*, *Symlets* e *Coiflets* são wavelets ortogonais e compactamente suportadas. Essas famílias de wavelets também podem ser usadas em CWT e DWT. As famílias de wavelets *B – splines* e *Biortogonais* são biortogonais com suporte compacto que também podem ser usados em CWT e DWT (68, 59).

Tabela 2 – Famílias de wavelets utilizadas.

Wavelet-mãe	Família	Propriedades	Função Wavelet
db4	Daubechies	asymmetric, orthogonal, biorthogonal	
sym4	Symlets	near symmetric, orthogonal, biorthogonal	
coif4	Coiflets	near symmetric orthogonal, biorthogonal	
bior2.8	Biorthogonal	symmetric, not orthogonal, biorthogonal	
rbio5.5	Reverse biorthogonal	symmetric, not orthogonal, biorthogonal	
dmey	Discrete Meyer	symmetric, orthogonal, biorthogonal	

Fonte: Elaborada pelo autor (2022).

Na Figura 6 podem ser vistas as wavelets da família *Daubechies* (db) que são uma das wavelets mais utilizadas. Na primeira coluna pode-se ver as wavelets *Daubechies* de primeira ordem (db1), na segunda coluna de segunda ordem (db2), até a quinta ordem na quinta coluna e assim sucessivamente. O número da ordem indica o número

de momentos de desaparecimento. Assim, db3 tem três momentos de desaparecimento e db4 tem 4 momentos de desaparecimento, e assim por diante. O número de momentos de desaparecimento está relacionado à ordem de aproximação e suavidade da wavelet. Se uma wavelet tiver  $p$  momentos de desaparecimento ela pode aproximar polinômios de grau  $p-1$ . Como pode ser visto, à medida que o número de momentos de desaparecimento aumenta, o grau polinomial da wavelet aumenta e ela se torna mais suave. À medida que o nível de decomposição aumenta, aumenta o número de amostras em que essa wavelet é expressa.

Na Tabela 2, são apresentadas as wavelets mães utilizadas neste estudo, especificando a família wavelet a qual pertence, suas propriedades, sua função *wavelet*, bem como sua função detalhe. Para questões adicionais e abrangentes sobre os aspectos matemáticos das wavelets, sugere-se, (68, 69, 70, 59).

### 3.2 CONCEITOS BÁSICOS DE UMA REDE NEURAL PERCEPTRON MULTICAMADAS

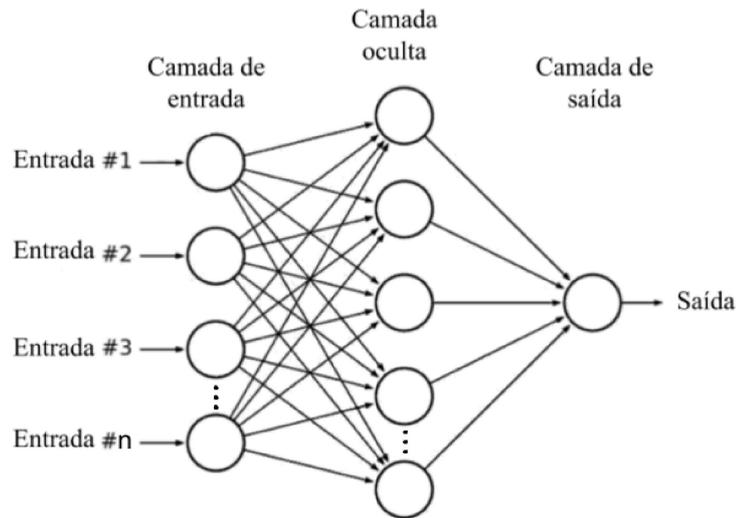
As redes neurais artificiais se inspiram nas redes neurais biológicas. São sistemas de processamento paralelo e distribuído que simulam os princípios básicos de funcionamento do cérebro biológico. Na literatura eles também são chamados de algoritmos de aprendizado de máquina.

O cérebro biológico em sua estrutura básica é uma rede de células neurais (neurônios) ligadas por meio de conexões que têm a capacidade de ajustar a potência do pulso elétrico que as percorre (sinapses). O estímulo externo na forma de pulso elétrico é transmitido como informação através de sinapses para os neurônios, onde é processado e, eventualmente, é produzida uma resposta de saída da rede. A informação é codificada como “conhecimento” através da atualização contínua das sinapses existentes entre os neurônios. Da mesma maneira, a rede neural artificial é constituída por muitos neurônios interligados (também conhecidos por unidades ocultas), cujos pesos associados determinam a força do sinal que passa por eles. As redes neurais podem aproximar qualquer função não linear desconhecida e geralmente são menos sensíveis do que as abordagens clássicas estatísticas para suposições sobre o termo de erro, portanto, eles podem ter um bom desempenho na presença de ruído e trechos caóticos (57). Graças a essas características, as RNAs possuem uma ampla gama de aplicações incluindo sua grande habilidade em modelagem e previsão de séries temporais hidrológicas não lineares. Além disso, redes neurais são adotadas por sua capacidade de atuar em problemas complexos e onde haja imprecisão, inconsistência e(ou) ausência de dados nas séries históricas. De acordo com (71), uma rede neural feed-forward com uma camada oculta fornece a complexidade adequada para recriar, em previsões, o comportamento não-linear de variáveis dos sistemas hidrológicos.

A arquitetura de rede neural mais utilizada na previsão hidrológica é a Rede Neural Perceptron Multicamadas (MLP) (72). MLP é uma rede neural artificial feed-forward

que consiste em um número de neurônios conectados por pesos de ligação, fornece um mapeamento não linear entre a entrada e a saída. Considera-se um simples MLP que consiste em uma camada de entrada, uma camada oculta e uma camada de saída (veja a Figura 7).

Figura 7 – Uma simples Rede Neural MLP de três camadas.



Fonte: Elaborada pelo autor (2022).

A camada de entrada contém  $n$  entradas e recebe os dados de entrada de fora, então os passa para a camada oculta com  $m$  neurônios ocultos, que serão encaminhados até que finalmente cheguem à camada de saída com  $k$  neurônios. A saída da  $j$ -ésima unidade oculta é então calculada por combinações lineares ponderadas de  $n$  entradas:

$$a_j = \sum_{i=0}^n w_{ij}^1 x_i, j = 1, \dots, m \quad (3.23)$$

onde  $w_{ij}^1$  é o peso que passa da entrada  $i$  para a unidade oculta  $n$  na primeira camada, que é denotada por sobrescrito (1), e onde o viés  $x_0 = 1$ .

Subsequentemente,  $a_j$  é transformado por uma função de ativação  $g()$  para obter a ativação da unidade oculta  $j$  através de  $z_j = g(a_j)$ . Posteriormente, as saídas da rede são recuperadas pela transformação desta função de ativação através de uma segunda camada oculta e assim sucessivamente (no caso de MLP de múltiplas camadas). A equação completa representa então a propagação direta através de toda a rede:

$$y_k = g\left(\sum_{j=0}^m w_{kj}^l h\left(\sum_{j=0}^n w_{ji}^{l-1} x_i\right)\right), \quad (3.24)$$

onde  $l$  descreve o número de camadas e  $k = 1, \dots, K$  o número de unidades de saída. A equação 3.24 mostra um MLP simples com uma saída.

O principal objetivo no treinamento (ensino) de uma rede MLP é buscar um conjunto ótimo de conexões (pesos) entre suas unidades de processamento para que a RNA

possa prever valores precisos de saídas para um determinado conjunto de entradas. O processo de treinamento supervisionado envolve a apresentação dos dados de treinamento em um formato padronizado. Para cada amostra de dados de entrada obtém-se a respectiva saída (resposta) desejada.

O processo de aprendizagem consiste em obter um conjunto de pesos ótimos que permitirão ao modelo produzir a melhor representação da relação entrada-saída. O aprendizado é um processo iterativo no qual os pesos são atualizados para minimizar o erro entre a rede e as saídas de destino para arquivar a generalização ótima do modelo. Esse resultado é obtido fornecendo à rede várias amostras de treinamento, cada uma consistindo em um padrão de entrada específico e uma resposta de saída correspondente (valor correto) (73). A atualização dos pesos sinápticos é proporcional à diferença entre a saída real e a saída calculada pela RNA, que é diretamente proporcional aos pesos utilizados (43).

Se for considerada matematicamente o aprendizado em RNAs como uma aproximação da função multi-variável em questão  $h(X)$  por outra função  $H(W, X)$ , onde  $X = [x_1, x_2, \dots, x_n]$  é o vetor de entrada e  $W = [w_1, w_2, \dots, w_n]$  é um vetor parâmetro (peso), a tarefa de aprendizado é encontrar o  $W$  que fornece a melhor aproximação possível de  $h(X)$  baseado no conjunto de exemplos de treinamento  $X$  (74).

O treinamento de uma rede MLP envolve duas fases: (a) feed-forward e (b) retropropagação. No processo feed-forward, os dados fluem das unidades de entrada de maneira estritamente direta para prever as saídas da rede e compará-las com os alvos medidos. No processo de retropropagação, o erro entre as saídas previstas e medidas é propagado para trás a partir da camada de saída para modificar as forças de conexão da rede a fim de melhorar seu desempenho (73). O processo de treinamento é alcançado com a obtenção de um conjunto ótimo de pesos que minimizam o erro entre os valores observados e os valores de saída determinados pela rede (75).

O otimizador (algoritmo de aprendizado) é responsável por alterar a taxa de aprendizado e o ajustar aos pesos dos neurônios na rede neural para atingir a função de perda mínima. Nesse estudo utilizaram-se os seguintes algoritmos de aprendizado:

- o algoritmo *Adam* é um algoritmo para otimização baseada em gradiente de primeira ordem de funções objetivo estocásticas, com base em estimativas adaptativas de momentos de ordem inferior (76). Este algoritmo é computacionalmente eficiente, tem poucos requisitos de memória, é adequado para problemas grandes em termos de dados e/ou parâmetros e para objetivos não estacionários e problemas com gradientes muito ruidosos e/ou esparsos (76);
- o algoritmo *Limited-Memory Broyden-Fletcher-Goldfarb-Shanno* (L-BFGS) (77). Este é um algoritmo da família de métodos Quasi-Newton adequado para encontrar extremos locais de funções, que são baseados no método de Newton para

encontrar pontos estacionários de funções. O algoritmo L-BFGS funciona bem com grandes conjuntos de dados porque precisa de menos memória do que o BFGS padrão.

A taxa de aprendizado (o hiperparâmetro no otimizador) controla o tamanho do passo para um modelo atingir a função de perda mínima. Uma taxa de aprendizado mais alta faz com que o modelo aprenda mais rápido mas pode perder a função de perda mínima e atingir apenas o seu entorno. Uma taxa de aprendizado mais baixa oferece uma chance melhor de encontrar uma função de perda mínima. Como compensação, uma taxa de aprendizado mais baixa precisa de épocas mais altas ou mais recursos de tempo e capacidade de memória. Para maiores informações a respeito das redes neurais artificiais deve-se consultar os trabalhos de (78, 79, 71).

### 3.3 VALIDAÇÃO CRUZADA PARA PREVISÃO DE SÉRIES TEMPORAIS

Um dos problemas relacionados aos modelos de previsão é o chamado sobreajuste (*overfitting*, em inglês), que se dá quando não se tem acesso completo aos dados e o modelo fica condicionado aos dados de treino, falhando, assim, na validação quando dados diferentes são utilizados. Uma alternativa para esse tipo de problema é a aplicação da técnica de validação cruzada (CV). O procedimento de validação cruzada fornece um mecanismo para avaliar o quão bem um modelo irá generalizar um conjunto de dados ainda não vistos evitando alguns problemas que podem aparecer com o uso de um único modelo em um único conjunto de dados. Um dos tipos mais comuns de validação cruzada é a validação cruzada  $k - fold$  (80).

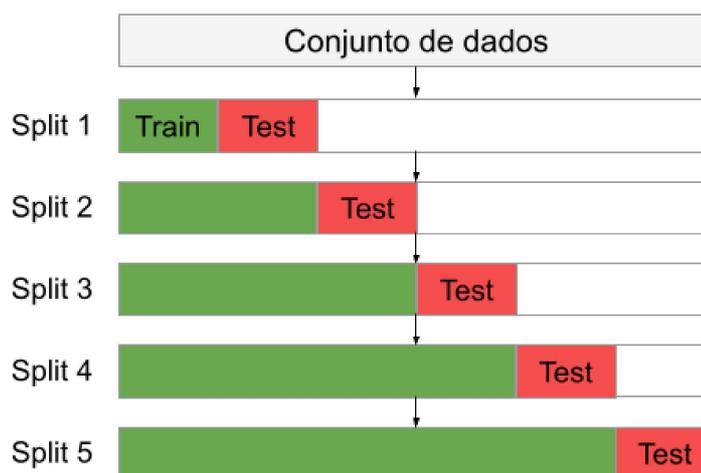
Na validação cruzada  $k - fold$ , a amostra original é dividida aleatoriamente em  $k$  subamostras de tamanhos iguais, chamadas *folds*. Das  $k$  *folds*, uma única *fold* é mantida como dados de validação para testar o modelo e as  $k - 1$  *folds* restantes são usadas como dados de treinamento. O processo de validação cruzada é então repetido  $k$  vezes, com cada uma das  $k$  *folds* usadas exatamente uma vez como dados de validação. Os  $k$  resultados podem então ser calculados para produzir uma única estimativa.

Ao lidar com dados de séries temporais, a validação cruzada  $k - fold$  não deve ser adotada porque a estrutura espacial ou temporal nos dados não pode ser ignorada.

Em uma previsão de séries temporais, uma variação da técnica  $k - fold$ , chamada *Time Series Split* (TSS), permite os mesmos ganhos de generalização como na validação cruzada  $k - fold$ , mas sem embaralhar as observações. A idéia da validação cruzada TSS (veja a Figura 8) é dividir os dados em dois subconjuntos, a saber subconjunto de treinamento e subconjunto de teste, a cada iteração, desde que o conjunto de teste sempre venha cronologicamente após o conjunto de treinamento. Esta abordagem produz  $k$  diferentes divisões de treinamento/teste. É calculada a média dos erros em cada divisão

para fornecer uma estimativa robusta do erro do modelo. Em cada divisão, estão mantidos os subconjuntos já treinados e testados como novos dados de treinamento e está agregada outra divisão totalizando em novo subconjunto de treinamento e teste. Os conjuntos de teste são do mesmo tamanho e o subconjunto de treinamento continua crescendo a cada etapa.

Figura 8 – Validação Cruzada *Time Series Split*.



Fonte: Elaborada pelo autor (2022).

### 3.4 CRITÉRIOS DE AVALIAÇÃO DA CAPACIDADE PREDITIVA DOS MODELOS

A avaliação do desempenho do modelo utiliza técnicas estatísticas que produzem uma avaliação quantitativa da capacidade preditiva do modelo. Normalmente, a forma como os dados de entrada determinam a qualidade do ajuste do modelo é dada por comparações entre as observações e o modelo previsto. Taxas quantitativas de acerto extraídas das simulações são utilizadas na medição da qualidade do modelo preditivo. Um modelo hidrológico exige, entretanto, para uma avaliação do seu desempenho, uma medida de erro relativo (tal como  $R^2$ , por exemplo) e outra de erro absoluto (tais como RMSE, MAE ou MAPE) (81, 82). Dessa forma, entre os critérios que são comumente usados para avaliação de desempenho dos modelos hidrológicos, nesta pesquisa, foram utilizadas as seguintes métricas:

- O erro percentual absoluto médio (MAPE) representa a média dos erros percentuais absolutos de cada entrada em um conjunto de dados, mostrando, em média, a precisão das quantidades previstas em comparação com as quantidades reais. Seu retorno é dado em porcentagem (%). Valores baixos de MAPE significam uma melhor adequação dos dados preditos. MAPE igual a 10% significa que o desvio médio entre o valor previsto e os valores reais foi de 10%, independentemente do desvio ser positivo ou negativo.

O MAPE é calculado de acordo com a seguinte equação:

$$\text{MAPE} = 100 \times \frac{1}{N} \sum_{i=1}^{N-1} \frac{|O_i - P_i|}{|O_i|}, \quad (3.25)$$

onde  $O$  representa a vazão observada,  $P$  a vazão predita pelo modelo e  $N$  é o tamanho da série histórica.

- O coeficiente de determinação ( $R^2$ ) descreve a proporção da variância total nos dados observados que pode ser explicada pelo modelo, que varia de  $-\infty$  a 1, sendo 1 a representação do modelo perfeito e zero significando que o modelo é tão preciso quanto os meios dos dados observados. Valores negativos significam que a média dos dados observados poderia ser um estimador melhor que o próprio modelo. De acordo com (75), um modelo pode ser considerado bom se o  $R^2$  está acima de 0,5 e muito bom se estiver acima de 0,7. O coeficiente de determinação é dado por

$$R^2 = 1 - \frac{\sum_{i=0}^{N-1} (O_i - P_i)^2}{\sum_{i=0}^{N-1} (O_i - \bar{O})^2}, \quad (3.26)$$

onde  $N$  denota o número de dados,  $O$  denota a vazão observada e  $P$  denota a vazão prevista.  $\bar{O}$  denota a média de vazões observadas para todo o período de tempo da avaliação.

Krause, Boyle e Båse(83) salientam que a principal desvantagem da  $R^2$  é o fato de que as diferenças entre os valores observados e preditos são calculados como valores quadrados. Conseqüentemente, valores maiores da série temporal são demasiadamente superestimadas e valores mais baixos são descartados. Em um contexto de previsão de vazão fluvial, isso acarreta uma superestimativa da performance do modelo nos picos de vazão e uma subestimativa em situações de baixa vazão.

- A Raiz do Erro Quadrático Médio (RMSE) é a raiz quadrada da média do quadrado da diferença entre valores observados e os valores previstos pelo modelo. O valor do RMSE possui a mesma unidade da variável e deve ser o mais baixo possível. Seu valor é dado em  $m^3/s$  e pode ser calculado como se segue:

$$\text{RMSE} = \frac{1}{N} \sqrt{\sum_{i=0}^{N-1} (O_i - P_i)^2}, \quad (3.27)$$

onde  $N$  denota o número de dados,  $O$  denota a vazão observada e  $P$  denota a vazão prevista.

Patel e Ramachandran(20) salientam que RMSE é uma métrica útil de demonstração do desvio entre um valor previsto e o real pois esta medida reúne todos os desvios ao longo dos pontos dos dados.

- Devido à crucial importância dos valores extremos na modelagem hidrológica, a Equação 3.28 pode ser utilizada na comparação da capacidade dos modelos em capturar os picos de valores na série histórica de vazão (12). O  $MAPE_{pks}$  é calculado de acordo com a seguinte equação:

$$MAPE_{pks} = 100 \times \frac{1}{N} \sum_{i=1}^{N-1} \frac{|O_{pks}^{(i)} - P_{pks}^{(i)}|}{|O_{pks}^{(i)}|}, \quad (3.28)$$

onde  $O_{pks}$  representa a vazão observada nos picos,  $P_{pks}$  a vazão predita nos picos e  $N$  é o tamanho da série histórica.

### 3.5 TESTES ESTATÍSTICOS

Ao realizar análises estatísticas usando métodos paramétricos, a verificação da suposição de normalidade é de grande importância. O teste de normalidade de Shapiro-Wilks é um dos testes de normalidade geral concebido para detectar todos os desvios da normalidade (84). O teste rejeita a hipótese de normalidade quando o p-valor for menor ou igual a 0,05. A falha no teste de normalidade permite afirmar com 95% de confiança que os dados não se ajustam à distribuição normal. Passar no teste de normalidade permite afirmar que não foi encontrado nenhum desvio significativo da normalidade. A estatística de teste é seguinte (84):

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)}\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (3.29)$$

onde  $x_i$  representa os valores ordenados da amostra,  $a_i$  as constantes geradas a partir das covariâncias, variâncias e médias da amostra a partir de uma amostra normalmente distribuída. A indicação de normalidade retornada pelo teste Shapiro-Wilks indica a possibilidade de uso de testes estatísticos paramétricos para verificar as diferenças entre os métodos. No caso de não ser possível atestar normalidade, testes não paramétricos devem ser aplicados.

O teste U de Mann-Whitney é o teste não paramétrico alternativo ao teste t de amostra independente. É usado para comparar diferenças entre dois grupos independentes. Testa a hipótese de dois grupos pertencerem a mesma população ou terem as mesmas medianas. Não assume qualquer distribuição específica (tais como distribuição normal das amostras) para o cálculo das estatísticas dos testes e dos p-valores. A classificação das amostras ou medianas (não as médias) são comparadas no teste U de Mann-Whitney com base na forma de distribuição de dois grupos independentes, o que a distingue do teste t, que compara as médias das amostras. O teste U de Mann-Whitney pode ser aplicado em amostras pequenas (5-20) e grandes ( $n > 20$ ).

As suposições do Mann-Whitney são seguintes: as observações dos dois grupos devem ser selecionadas aleatoriamente das populações-alvo; as observações são independentes uma da outra; e as observações devem ser contínuas ou ordinais.

Se houver dois grupos independentes com observações  $x_1, x_2, \dots, x_m$  e  $y_1, y_2, \dots, y_n$  amostradas das populações X e Y, então o teste U de Mann-Whitney compara cada observação  $x_i$  da amostra X com cada observação  $y_j$  da amostra Y. O maior número possível de comparações pareadas é assim:  $(n_x n_y)$ , onde  $n_x$  é o número de observações no primeiro grupo e  $(n_y)$  o número de observações no segundo. Se os dois grupos provêm da mesma população, conforme estipulado pela hipótese nula, cada dado do primeiro grupo terá igual chance de ser maior ou menor que cada dado do segundo grupo, ou seja, uma probabilidade p de 1/2.

Também pode-se afirmar a hipótese bilateral (*two – sided*, em inglês) em termos de mediana como (quando dois grupos têm a mesma forma de distribuição):

- Hipótese nula  $H_o$ : dois grupos têm mediana igual (as duas populações são iguais) versus
- Hipótese alternativa  $H_a$ : dois grupos não têm mediana igual (as duas populações não são iguais).

A Hipótese alternativa bilateral testa que há igual probabilidade de  $x_i$  ser maior ou menor que  $y_j$  (ambos os grupos vieram da mesma população). A hipótese nula é rejeitada se um grupo for significativamente maior que o outro, sem especificar a direção dessa diferença. Em uma aplicação unilateral do teste (hipótese alternativa *One – sided*, em inglês), a hipótese nula continua a mesma. No entanto, a hipótese alternativa é alterada especificando a direção da comparação (ela testa a probabilidade de  $x_i$  ser maior que  $y_j$  e vice-versa). A mediana dos testes de hipóteses alternativas unilaterais de um grupo pode ser maior ou menor do que a de outro grupo.

A estatística Mann-Whitney U é definida como o mínimo de U1 e U2 (85):

$$\begin{aligned} U_1 &= n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \\ U_2 &= n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \end{aligned} \tag{3.30}$$

onde  $n_1$  e  $n_2$  são os respectivos tamanhos de cada amostra,  $R_1$  e  $R_2$  é a soma das fileiras das observações das amostras 1 e 2, respectivamente. O p-valor é calculado com base na comparação entre o valor crítico e o valor U. Se o valor de  $U \leq$  valor crítico, rejeita-se a hipótese nula e vice-versa.

## 4 METODOLOGIA PROPOSTA

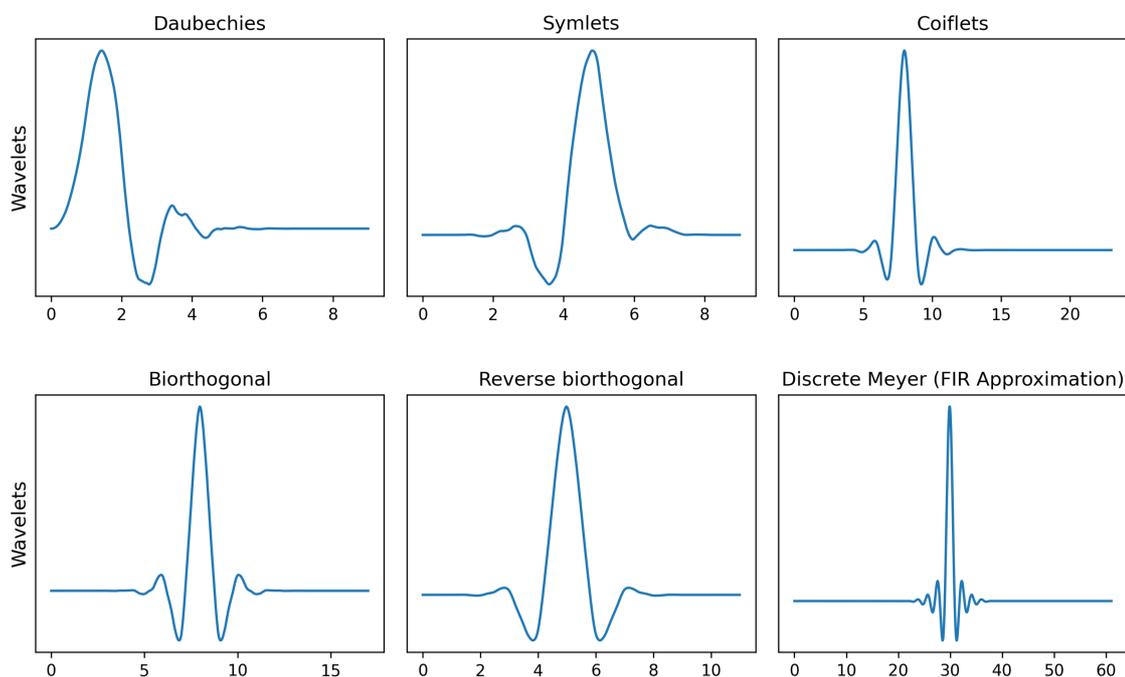
Esta seção descreve cada etapa da metodologia proposta do desenvolvimento de modelos híbridos wavelet redes neurais artificiais com apresentação dos procedimentos adotados na construção dos modelos baseados em wavelets e redes neurais artificiais para previsão de vazão diária de 1, 3, 5 e 7 dias à frente em determinados pontos da Bacia do rio Paraíba do Sul.

### 4.1 PRÉ-PROCESSAMENTO DE DADOS

#### 4.1.1 Escolha da Wavelet mãe e nível de decomposição

A escolha de uma wavelet mãe apropriada dentro da grande variedade de wavelets disponíveis é a tarefa principal em modelagem baseada em wavelets porque os resultados da modelagem dependem diretamente das wavelets usadas (24). De acordo com Alexander, Thampi e NR(41), as wavelets com propriedades ortonormais ou ortogonais são consideradas wavelets ideais para variáveis hidrológicas.

Figura 9 – Famílias Wavelets usadas.



Fonte: Elaborada pelo autor (2022).

Chourasia e Mitra(86) observem que as oscilações da wavelet mãe devem ser muito semelhantes ao sinal de entrada a ser analisado. De acordo com (30), as wavelets da família Daubechies geralmente produzem resultados confiáveis em problemas hidrológicos de séries temporais. (30) observou que a alta variabilidade na série temporal sugere a necessidade

de wavelets mãe com boa localização de tempo como Daubechies (db2) e Symlets (sym4). Além disso, Freire, Santos e Silva(43) observaram que a wavelet mãe discreta de Meyer (dmey) é um método adequado para representar a série temporal de vazão.

Na etapa de pré-processamento de dados, o objetivo foi comparar a influência de diferentes famílias de wavelets na precisão da previsão de vazão média diária. Com base nos resultados e nas inferências encontrados na literatura, este estudo fez uma avaliação comparativa das seguintes wavelets mãe: Daubechies (db4), Symlets (sym4), Coiflets (coif4), Biortogonal (bior2.8), Reverse biortogonal (rbio5.5) and Discrete Meyer (dmey). A representação visual destas wavelets são apresentadas na Figura 9.

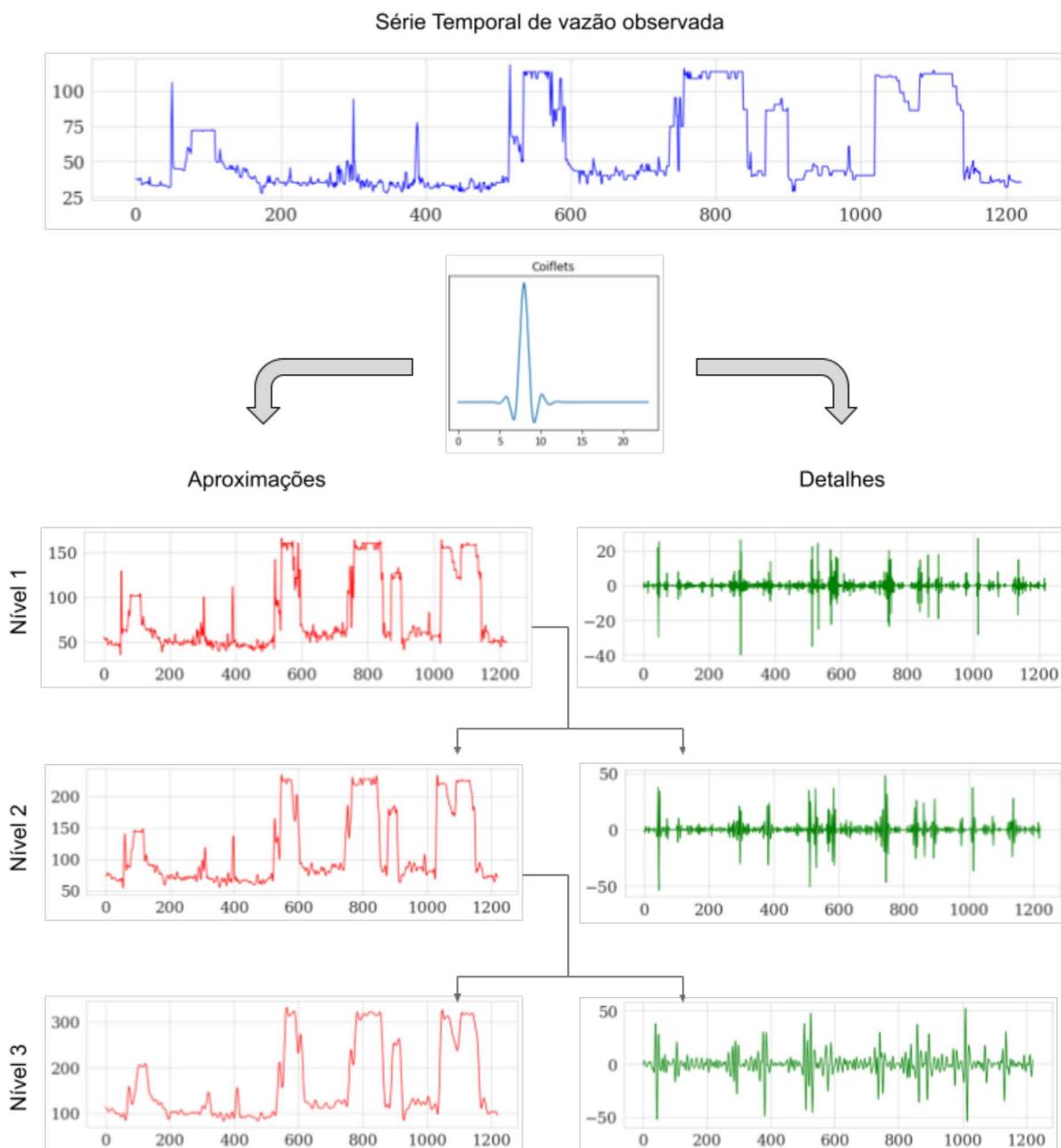
Como notam Jin e Kim(65), a escolha do nível de decomposição ótimo dos dados da série temporal tem um papel importante na conservação das informações e na minimização da distorção dos conjuntos de dados. Nourani, Komasi e Mano(87) introduziram uma fórmula para determinar o nível mínimo de decomposição  $L = \lceil \log(N) \rceil$ , onde  $N$  é o número de pontos de dados. Em alguns trabalhos (por exemplo, (47, 21, 36, 88)) pode ser encontrado o uso desta fórmula. Porém alguns autores (89, 41) argumentam que a aplicação dessa fórmula é incorreta para todas as séries temporais, motivado pelo fato de que séries temporais com o mesmo número de pontos de dados produzem tendências e características diferentes. De acordo com Sang et al.(89), a seleção da escala temporal adequada deve estar condicionada à composição e características dos dados analisados e não apenas ao comprimento dos dados. Para este estudo, a equação acima foi usada para se ter uma ideia aproximada do nível de decomposição.

Inicialmente, uma exploração preliminar de 3 níveis de decomposição ( $L=2, 3, 4$ ) foi realizada para identificar o nível de decomposição em termos de precisão do modelo. Após executar várias simulações, os modelos baseados em wavelets com nível de decomposição igual a 3 apresentaram melhores desempenhos em termos MAPE,  $R^2$  e RMSE em comparação aos modelos que utilizaram os níveis 2 e 4. Portanto, para simulações posteriores, as séries temporais originais foram transformadas usando o nível de decomposição igual a 3. Dessa forma, 3 níveis de decomposição wavelet geram 3 sub-séries temporais de detalhe e 3 sub-séries temporais de aproximação.

#### 4.1.2 Decomposição wavelet

Em diferentes algoritmos de decomposição wavelet, níveis de decomposição e filtros wavelet levam a um desempenho significativamente diferente em cada modelo (31). Para desenvolver os modelos baseados em wavelets, no presente estudo foi utilizado a transformada AT para evitar o problema dos “dados futuros”. O uso da transformada wavelet AT na modelagem hidrológica permite usar dados conhecidos apenas no presente e no passado (o que acontece em casos reais) durante o treinamento e teste dos modelos de previsão.

Figura 10 – Decomposição wavelet AT (em 3 níveis) com wavelet mãe *coif4* de uma série temporal de vazão, do nível 1 ao 3: coeficientes de aproximação e detalhes.



Fonte: Elaborada pelo autor (2022).

Utilizando a transformada AT, acontece a separação dos componentes de frequência do sinal de acordo com a magnitude, ou seja, a decomposição do sinal em seus componentes de aproximação  $cA$  (baixa frequência) e detalhes  $cD$  (alta frequência). Na primeira decomposição (Nível 1), o sinal original é decomposto em aproximações ( $cA_1$ ) e detalhes ( $cD_1$ ), atuando como filtros passa-alta e passa-baixa. Em nível 2, o componente de baixa frequência obtido ( $cA_1$ ) é novamente decomposto em  $cA_2$  e  $D_2$  e por fim, em nível 3, o componente ( $cA_2$ ) é decomposto em  $cA_3$  e  $cD_3$ . A Figura 10 demonstra a decomposição

wavelet AT (em 3 níveis) com wavelet mãe `coif4` de uma série temporal de vazão. Observe-se que 3 níveis de decomposição wavelet geram 3 sub-séries temporais de detalhe e 3 sub-séries temporais de aproximação.

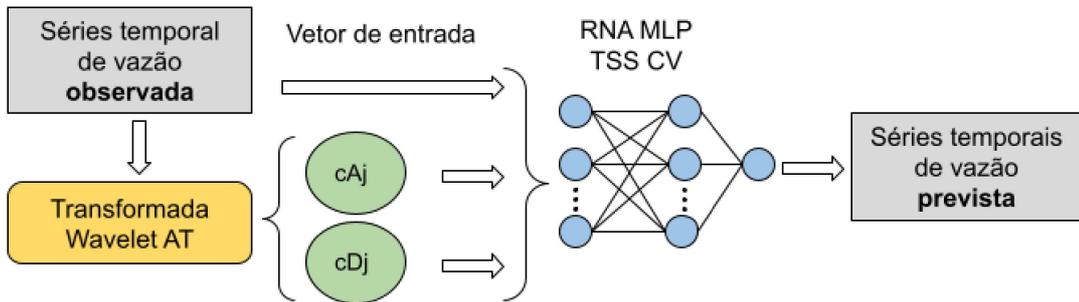
## 4.2 MODELOS HÍBRIDOS WAVELET REDES NEURAS ARTIFICIAIS

Com intuito de alcançar os objetivos impostos dentro da metodologia proposta, foram desenvolvidos dois modelos híbridos acoplando o método de pré-processamento de dados baseado em wavelets e redes neurais artificiais.

### 4.2.1 Primeiro modelo híbrido proposto $WANN_{one}$

O primeiro modelo proposto, denominado  $WANN_{one}$ , envolve o uso de método de pré-processamento baseado em wavelets (transformada AT) acoplado à rede neural MLP com uma camada oculta e uma saída. O diagrama esquemático da arquitetura desenvolvida  $WANN_{one}$  é demonstrado na Figura 11.

Figura 11 – Representação esquemática do modelo  $WANN_{one}$ .



Fonte: Elaborada pelo autor (2022).

Na primeira abordagem proposta, a série temporal de vazão é primeiramente decomposta em sub-séries temporais com diferentes escalas, a saber, sub-séries de aproximação,  $cA(t)$  (em grande escala) e sub-séries de detalhe,  $cD(t)$  (em pequena escala). Em seguida, estas sub-séries junto com a vazão observada não decomposta são usadas como entrada para uma rede neural do tipo MLP. Na saída da RNA obtém-se a vazão estimada em determinado horizonte de previsão.

A vazão estimada para um horizonte temporal  $i$  é considerada como uma função de conjuntos finitos de diferentes defasagens temporais da série temporal de vazão observada numa estação e diferentes defasagens temporais de sub-séries geradas por wavelet a partir da série temporal observada:

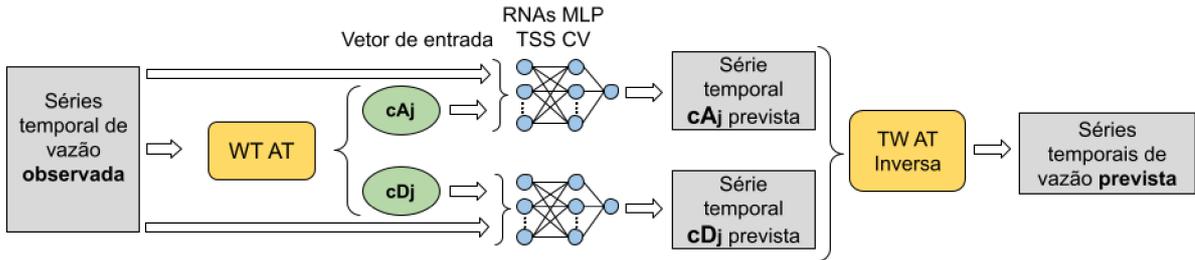
$$Q_{t+i} = F(Q_t, Q_{t-1}, \dots, Q_{t-n}, cA_t^1, cA_{t-1}^1, \dots, cA_{t-n}^1, cD_t^1, cD_{t-1}^1, \dots, cD_{t-n}^1, \dots, cA_t^l, cA_{t-1}^l, \dots, cA_{t-n}^l, cD_t^l, cD_{t-1}^l, \dots, cD_{t-n}^l), \quad (4.1)$$

onde o índice  $i$  denota o horizonte temporal,  $n$  denota o número de valores antecedentes para as variáveis de entrada, o índice superior  $l$  denota o nível de decomposição e  $F$  denota função desconhecida mapeada pelo modelo.

#### 4.2.2 Segundo modelo híbrido proposto $WANN_{multi}$

O segundo modelo desenvolvido, denominado  $WANN_{multi}$ , consiste em uma combinação de método de pre-processamento de dados baseados em wavelets e redes neurais do tipo MLP. Na segunda abordagem proposta, a séries temporal de vazão também é decomposta através da transformada wavelet AT em sub-séries temporais com diferentes escalas, a saber, sub-séries de aproximação,  $cA(t)$ , e sub-séries de detalhe,  $cD(t)$ . Em seguida, estas sub-séries são usadas como entrada para redes neurais do tipo MLP. As variáveis de vazão observada não decompostas também são usadas como entradas para redes neurais. As saídas dessas RNAs em seguida são reconstruídas usando a transformada wavelet AT inversa para produzir a vazão estimada. O diagrama esquemático do modelo  $WANN_{multi}$  é demonstrado na Figura 12.

Figura 12 – Representação esquemática do modelo  $WANN_{multi}$ .



Fonte: Elaborada pelo autor (2022).

As sub-séries de aproximação previstas (ou sub-séries de detalhe) para um horizonte temporal  $i$  foram consideradas como funções de conjuntos finitos de diferentes defasagens temporais de sub-séries de aproximação (ou sub-séries de detalhe) e diferentes defasagens temporais da série temporal de vazão observada:

$$cA_{t+i} = F(Q_t, Q_{t-1}, \dots, Q_{t-n}, cA_t^1, cA_{t-1}^1, \dots, cA_{t-n}^1, \dots, cA_t^l, cA_{t-1}^l, \dots, cA_{t-n}^l), \quad (4.2)$$

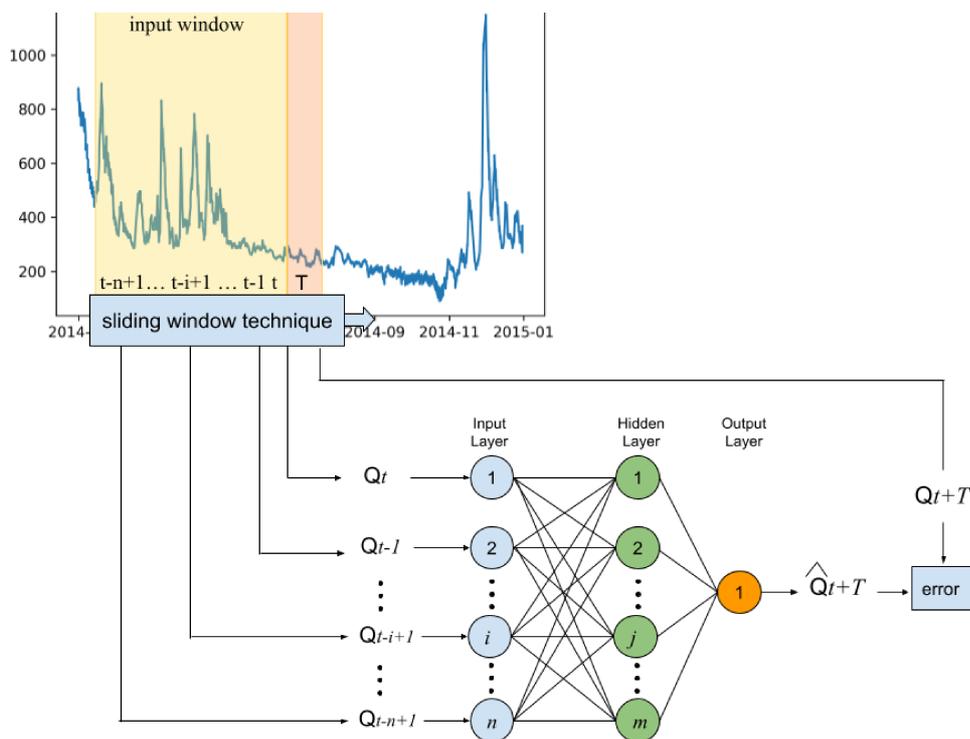
$$cD_{t+i} = F(Q_t, Q_{t-1}, \dots, Q_{t-n}, cD_t^1, cD_{t-1}^1, \dots, cD_{t-n}^1, \dots, cD_t^l, cD_{t-1}^l, \dots, cD_{t-n}^l), \quad (4.3)$$

onde  $n$  denota o número de valores antecedentes para as variáveis de entrada, o índice superior  $l$  denota o nível de decomposição e  $F$  denota função desconhecida mapeada pelo modelo.

### 4.2.3 Técnica de janela de tempo móvel

Para a geração de conjuntos de entrada e saída de dados para uma rede neural do tipo MLP utilizou-se a técnica de janela de tempo móvel (*sliding window technique*, em inglês) (Veja a Figura 13).

Figura 13 – Representação esquemática da técnica de janela de tempo movel.



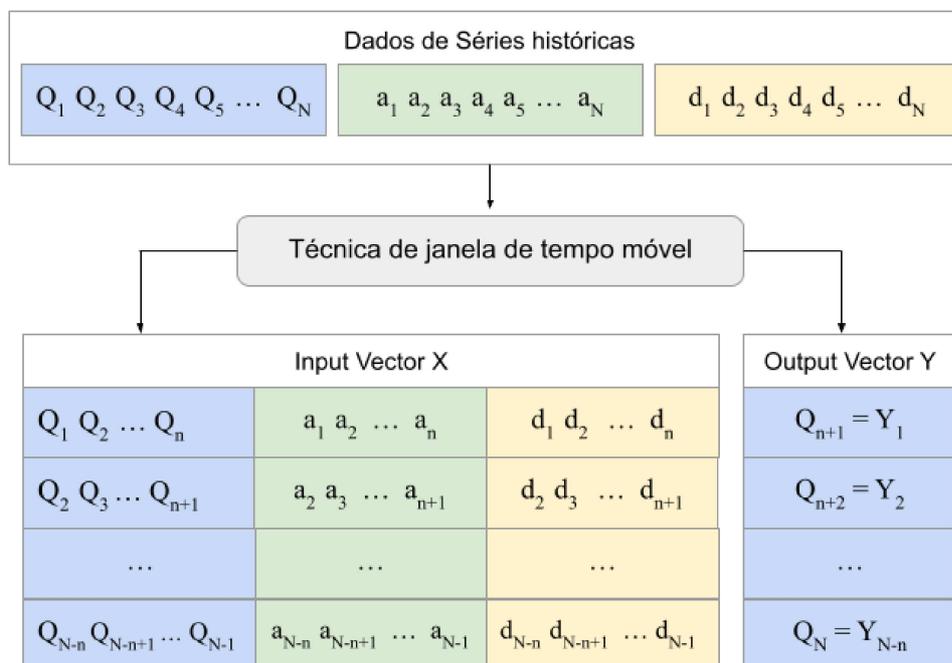
Fonte: Elaborada pelo autor (2022).

Para o caso do modelo  $WANN_{one}$ , para cada valor de vazão previsto  $Q_{t+i}$  num horizonte temporal  $i$ , que é considerado como uma saída da rede neural, um vetor de entrada correspondente contém valores de vazão  $Q(t)$ , sub-série de aproximação  $cA(t)$  e sub-série de detalhe  $cD(t)$  das janelas de tempo de  $n$  dias anteriores (Veja a Figura 14). Para o caso do modelo  $WANN_{multi}$ , para cada valor de sub-séries de aproximação previsto  $cA_{t+i}$  (ou sub-séries de detalhe previsto  $cD_{t+i}$ ) num horizonte temporal  $i$ , um vetor de entrada correspondente contém valores de vazão  $Q(t)$  e sub-séries de aproximação  $cA(t)$  (ou valores de vazão  $Q(t)$  e sub-séries de detalhe  $cD(t)$ ) das janelas de tempo de  $n$  dias anteriores.

### 4.2.4 Busca em grade com validação cruzada para ajuste de hiperparâmetros

Diferentes conjuntos de dados exigem diferentes conjuntos de hiperparâmetros do modelo para previsão mais eficiente. O ajuste de parâmetros é o processo de selecionar os valores para os parâmetros de um modelo que maximizam a precisão do modelo. Porém, não há resposta para quantas camadas são as mais adequadas, quantos neurônios são os

Figura 14 – Representação do uso da técnica da janela de tempo móvel para criar um vetor de entrada  $X$  e vetor de saída  $Y$  no caso do modelo  $WANN_{one}$ .  $Q_i$  são valores de série temporal de vazão,  $a_i$  são valores de sub-série de aproximação,  $d_i$  são valores de sub-série de detalhe, o  $N$  denota o tamanho das séries temporais e  $n$  denota o número de valores antecedentes (o tamanho da janela de tempo móvel).



Fonte: Elaborada pelo autor (2022).

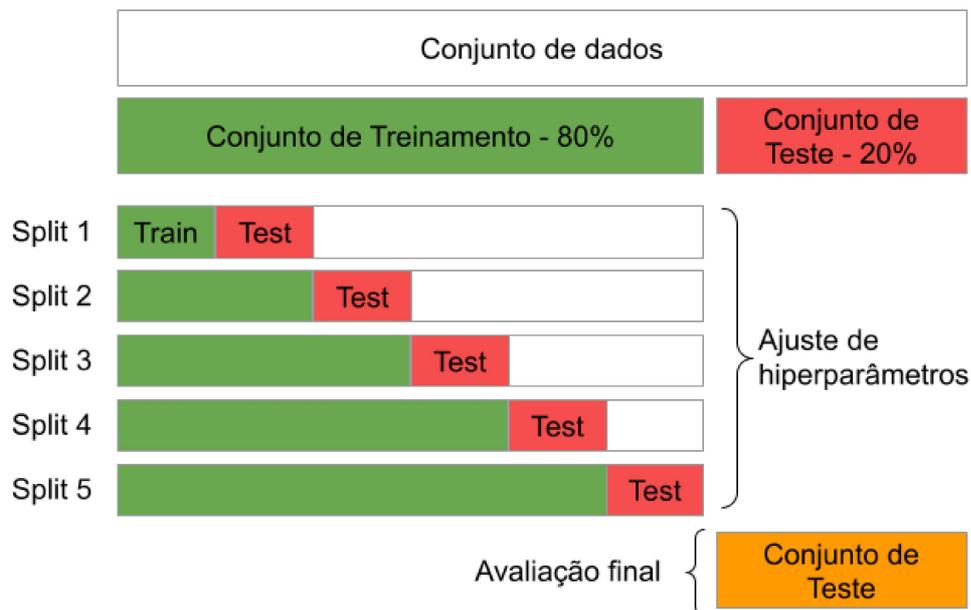
melhores ou qual otimizador se adapta melhor a todos os conjuntos de dados. Geralmente, executa-se a validação cruzada para descobrir quais configurações de parâmetros fornecem o erro mínimo antes de treinar um modelo final usando essas configurações selecionadas no conjunto de treinamento completo.

O conjunto de dados original foi dividido em duas partes: os primeiros 80% do total de dados foram usados como conjunto de treinamento e os segundos 20% (conjunto de teste) foram usados para estimar o desempenho final. Dessa maneira, os dados de treinamento são usados para ajustar hiperparâmetros e ajustar a rede neural, enquanto, os dados de teste são usados para calcular a capacidade de generalização na combinação ótima de hiperparâmetros e comparar diferentes modelos.

A arquitetura de rede que produziu os melhores resultados em termos de erro quadrático médio nas etapas de treinamento e teste determinou-se explicitamente, tentando todas as combinações possíveis dos hiperparâmetros na validação cruzada TSS de 5-*split*. Conforme a Figura 15, para encontrar os “melhores” hiperparâmetros o modelo treinou-se repetidamente do zero no subconjunto de treinamento e testou-se no subconjunto de teste em cada *split*.

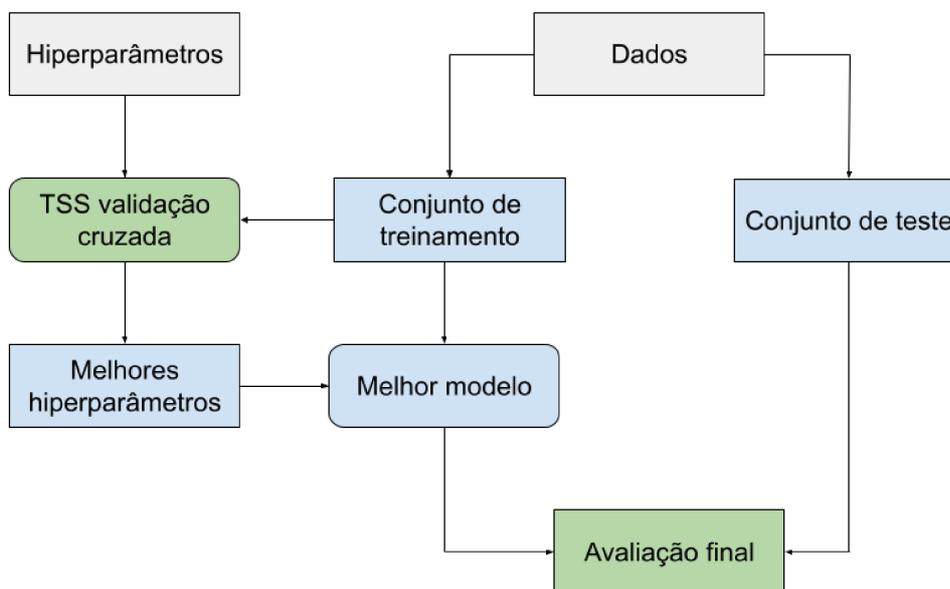
O desempenho médio do modelo em todas as *splits* de teste foi guardado e repetido

Figura 15 – Representação esquemática da validação cruzada TSS com  $k = 5$  splits.



Fonte: Elaborada pelo autor (2022).

Figura 16 – Ajuste de hiperparâmetros e estimativa de erros em busca em grade com validação cruzada TSS.



Fonte: Elaborada pelo autor (2022).

todo o procedimento para outro conjunto de hiperparâmetros. Em seguida, foi escolhido o conjunto de hiperparâmetros que corresponde ao melhor desempenho durante a validação cruzada. Ao definir os hiperparâmetros, o modelo é treinado 5 vezes, todas as vezes usando uma divisão de teste diferente. Finalmente, o modelo é treinado no conjunto de treinamento completo usando os melhores parâmetros escolhidos, e o erro no conjunto de teste é registrado na avaliação final. Na Figura 16 é demonstrado um fluxograma de fluxo

de trabalho de validação cruzada na etapa de treinamento do modelo.

#### 4.2.5 Configuração de modelos propostos

A fim de demonstrar a capacidade preditiva superior dos modelos propostos baseados em wavelets, também se optou por desenvolver os modelos de redes neurais convencionais denominados ANN. Os possíveis parâmetros dos modelos  $WANN_{one}$ ,  $WANN_{multi}$  e ANN utilizados na busca em grade com validação cruzada são apresentadas na Tabela 3.

Tabela 3 – Hiperparâmetros utilizados em busca em grande com validação cruzada.

Modelo	Método	Hyperparâmetros
$WANN_{one}$	MLPRegressor	hidden_layer_sizes: 100, 200 activation: relu solver: adam, lbfgs learning_rate_init: 0.01, 0.001 max_iter: 200
$WANN_{multi}$	MLPRegressor	hidden_layer_sizes: 10, 20, 50, 100 activation: relu solver: adam, lbfgs learning_rate_init: 0.01, 0.001 max_iter: 200
ANN	MLPRegressor	hidden_layer_sizes: 10,20,30,50 activation: relu solver: adam, lbfgs learning_rate_init: 0.01, 0.001 max_iter: 200

Fonte: Elaborada pelo autor (2022).

As redes neurais MLP foram desenvolvidas utilizando a biblioteca Scikit-Learn, empregando o algoritmo `MLPRegressor` que é treinado usando retropropagação com a função de ativação na camada oculta, *Rectified Linear Unit* (ReLU) (90), que é padrão do algoritmo e a função identidade como função de ativação na camada de saída. A função de ativação ReLU retorna

$$f(x) = \max(0, x),$$

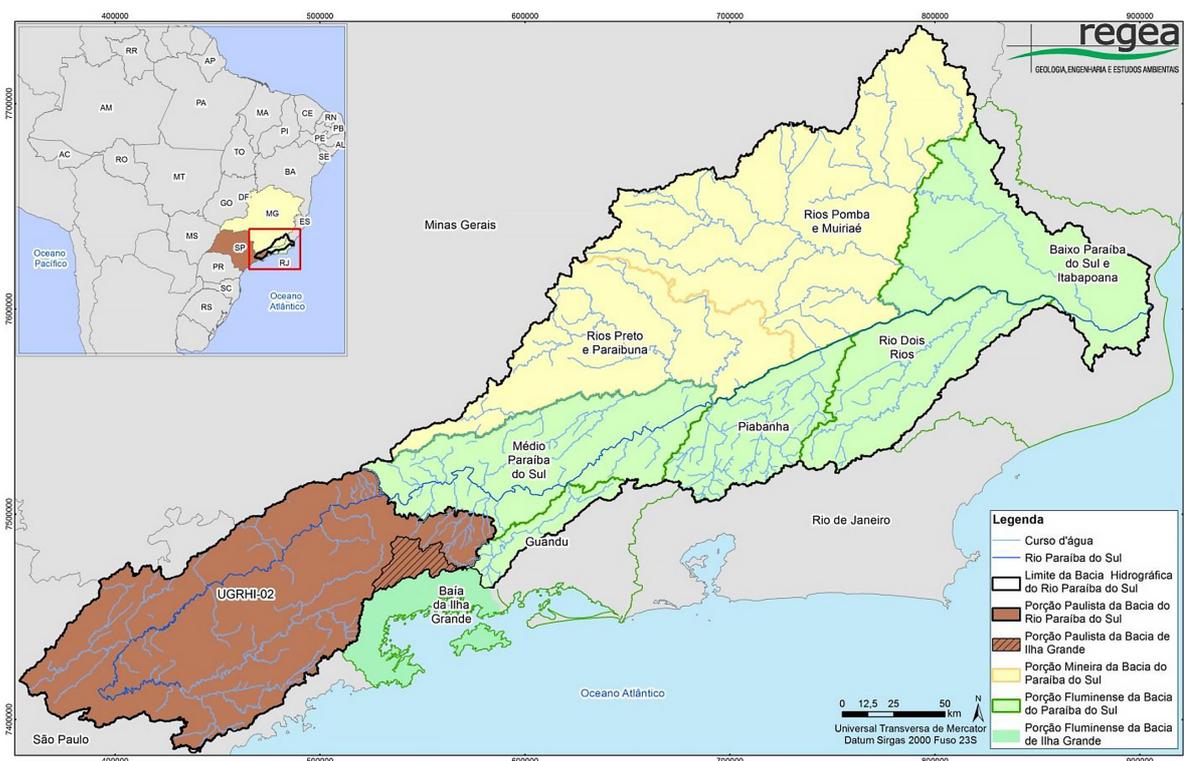
minimizando o erro médio quadrático RMSE. Como algoritmo de treinamento foram escolhidos otimizadores Adam e L-BFGS. A taxa de aprendizado inicial, `learning_rate_init`, controla o tamanho do passo na atualização dos pesos para um modelo atingir a função de perda mínima. Utilizou-se a taxa de aprendizado inicial igual 0.01 e 0.001. O critério de parada do treinamento (`max_iter`, padrão do algoritmo) é 200 iterações.

## 5 ÁREA DE ESTUDO E DADOS

Neste capítulo são apresentados a descrição das regiões de estudo e as séries históricas diárias de vazão natural das estações fluviométricas consideradas.

### 5.1 ÁREA DE ESTUDO

Figura 17 – Bacia hidrográfica do Rio Paraíba do Sul e suas respectivas porções: paulista mineira e fluminense.



Fonte: Comitê das Bacias do Rio Paraíba do Sul, 2017.

Como estudo de caso foram modeladas estações fluviométricas localizadas em duas regiões do rio Paraíba do Sul (ver a Figura 17):

- Região Fluminense da bacia do rio Paraíba do Sul caracterizada por possuir muitos rios afluentes e uma vazão no rio principal extremamente elevada, sendo estas características determinantes para uma dinâmica mais lenta;
- Região Paulista da bacia do rio Paraíba do Sul caracterizada por possuir poucos rios afluentes em comparação do região Fluminense da bacia do rio Paraíba do Sul e uma vazão no rio principal menos elevada, o que determina uma dinâmica mais rápida para esta bacia.

As breves caracterizações destas regiões são apresentadas a seguir.

### 5.1.1 Caracterização da região Fluminense da bacia do rio Paraíba do Sul

A região Fluminense do rio Paraíba do Sul caracteriza-se pela precária situação de porcentagem de cobertura florestal (10%) e pela predominância de planícies, e pastagens. Tem como principal característica a existência da agroindústria sucroalcooleira cujas plantações ocupam áreas significativas da região (12%). O cultivo de cana é feito por pequenos produtores para todo o Estado, considerado o segundo maior consumidor nacional de álcool. As plantações de café assumem importância econômica, ocorrendo de forma pontual no município de Varre-Sai. O restante destas regiões hidrográficas é ocupado por pastagens (63%), que correspondem, em sua maioria, a áreas muito pouco produtivas ou degradadas.

A população total de acordo com levantamento censitário (91) é de 913.818 habitantes, metade deles habitando a cidade de Campos dos Goytacazes. Os centros urbanos pertencentes à baixada Campista merecem atenção especial, principalmente por problemas relacionados a drenagem urbana e controle de cheias, que sofrem com as eventuais cheias (92). Geralmente, terrenos baixos e planos, situados próximos às margens dos rios, são mais propensos a inundações, provocando um processo erosivo que contribui para o assoreamento do leito dos rios.

A expansão da demanda regional pela construção civil ampliou a exploração de argila e a instalação de pedreiras. Os rios Pomba e Muriaé, principais afluentes do rio Paraíba do Sul na RH-IX, nascem na Zona da Mata mineira. A intensa exploração de recursos minerais e eventos hidrológicos críticos na parte mineira exercem grande impacto sobre a gestão de recursos hídricos na região fluminense. Além disso, nessa região pode-se notar a influência dos efeitos das usinas implantadas na bacia do rio Paraíba (Sobragi, Picada) e da UHE Simplício.

### 5.1.2 Caracterização da região Paulista da bacia do rio Paraíba do Sul

De acordo com Fundação COPPETEC(93), a região Paulista do Paraíba do Sul abrange uma área de 14.000 km<sup>2</sup> que compreende 49 municípios que totalizam cerca de 2 milhões de habitantes. Os maiores municípios desta região que possuem mais de 30.000 habitantes são Aparecida, Caçapava, Cruzeiro, Lorena, Tremembé, São José dos Campos, Taubaté, Guaratinguetá, Pindamonhangaba e Jacareí. A região está situada à montante da UHE de Funil e é formada pelos rios Paraíba e Paraitinga. Constitui o principal diferencial desta região o fato de que ela concentra o maior parque industrial de toda a bacia do rio Paraíba do Sul somando-se, ainda, a instalação de usinas hidrelétricas e reservatórios que regularizam a vazão do Paraíba do Sul, controlando o ciclo de inundações e a transposição da maior parte de suas águas para o rio Guandu, responsável pelo abastecimento da região metropolitana do Rio de Janeiro.

Esta região possui a segunda melhor distribuição percentual de florestas nos

municípios de toda a bacia do Paraíba do Sul, apenas possuindo uma menor área florestal que a região serrana, aos pés da Serra dos Órgãos. As partes paulista e a fluminense da bacia possuem uma semelhante vulnerabilidade de seus solos à erosão, se consideradas as classes mais críticas (muito alta e alta) e apresenta os melhores índices de saneamento ambiental, realizada através de estações de tratamento de esgotos domésticos instaladas em municípios de maior impacto ambiental da bacia.

O trecho paulista do Rio Paraíba do Sul possui forma estreita e alongada, e os cursos de água afluentes que drenam a bacia nessa região são de pequeno porte. A bacia é de grande importância econômica, o que torna viável os aproveitamentos de múltiplos usos instalados, Paraibuna, Santa Branca, Jaguari e Funil, destinados principalmente à geração de energia, à regularização de vazões e ao controle de cheias.

Este trecho também sofre degradação resultante de ações de natureza antrópica que agravam os processos de erosão de encostas e do solo da bacia. Agravam o quadro de elevada produção de sedimentos a extração predatória de areia, os demais usos do solo como impermeabilização das áreas urbanas, elevada poluição de resíduos sólidos, difusa, doméstica e industrial e de resíduos sólidos, que restringem a capacidade de escoamento das calhas durante as cheias. Esta degradação contribui para um aumento na frequência das enchentes urbanas, sendo urgentes ações de controle das atividades antrópicas e de recuperação das áreas já degradadas da bacia.

## 5.2 SELEÇÃO DE DADOS

As principais informações sobre as estações fluviométricas que foram selecionadas para modelagem hidrológica podem ser encontradas na Tabela 4, onde: a coluna “RG” indica a região da bacia do rio Paraíba do Sul Fluminense (FL) ou Paulista (PA); a coluna “AD” indica a área de drenagem em (km<sup>2</sup>); as colunas “Início” e “Fim” indicam início e fim da observação de vazão numa estação, respectivamente; a coluna “QTD” indica quantidade de dados observados durante o período total, em (dias); e as colunas “Q.max”, “Q.min” e “Q.med” a vazão máxima, mínima e média, respectivamente, em (m<sup>3</sup>/s).

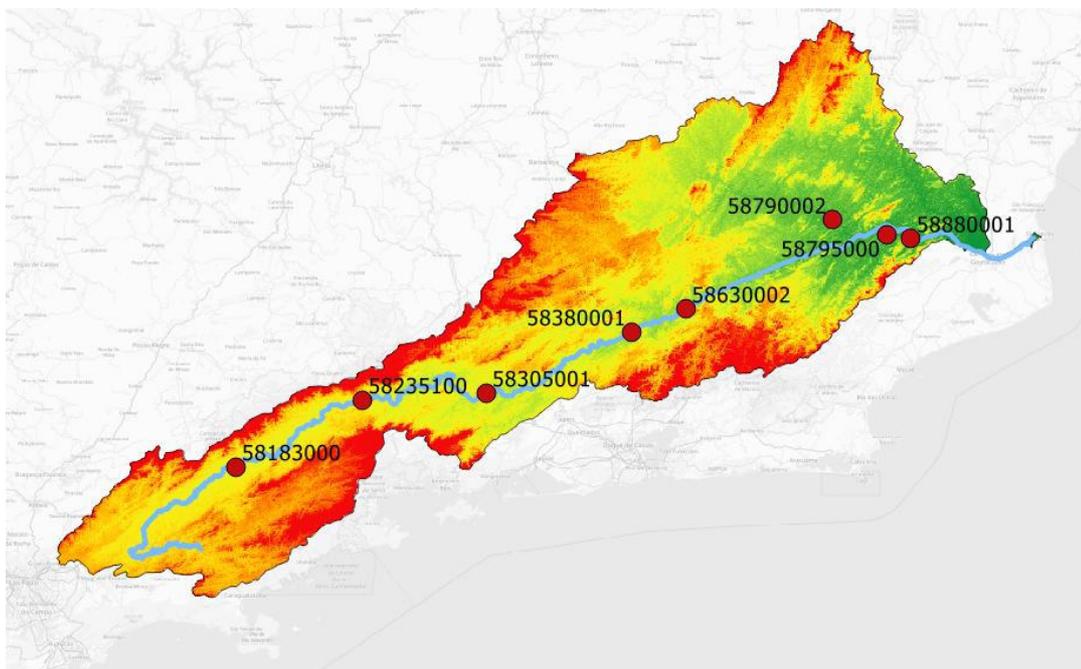
Tabela 4 – Informações das estações fluviométricas modeladas nesta pesquisa.

<b>Código</b>	<b>RG</b>	<b>AD</b>	<b>Q.max</b>	<b>Q.min</b>	<b>Q.med</b>	<b>Início</b>	<b>Fim</b>	<b>QTD</b>
58880001	FL	48900	5853.51	81.42	591.85	1973-01-01	2019-12-31	17135
58790002	FL	8210	1196.47	8.34	115.56	2001-02-01	2019-07-31	8730
58795000	FL	45300	4621.15	84.6	542.74	1973-01-01	2019-07-31	17006
58630002	FL	32700	2560.49	85.59	365.67	1973-01-01	2019-12-31	17166
58380001	FL	21400	1244.02	31.73	141.72	1973-01-01	2019-12-31	17166
58305001	PA	16000	1650.87	36.84	261.57	1973-01-01	2019-12-31	13719
58235100	PA	12800	1299.21	49.87	204.93	1973-01-01	2019-12-31	14274
58183000	PA	9600	698.03	27.63	146.42	1973-01-01	2019-10-31	17072

Fonte: Elaborada pelo autor (2022).

As séries históricas de vazão média diária destas estações foram disponibilizadas

Figura 18 – Localização das estações fluviométricas modeladas no estudo.



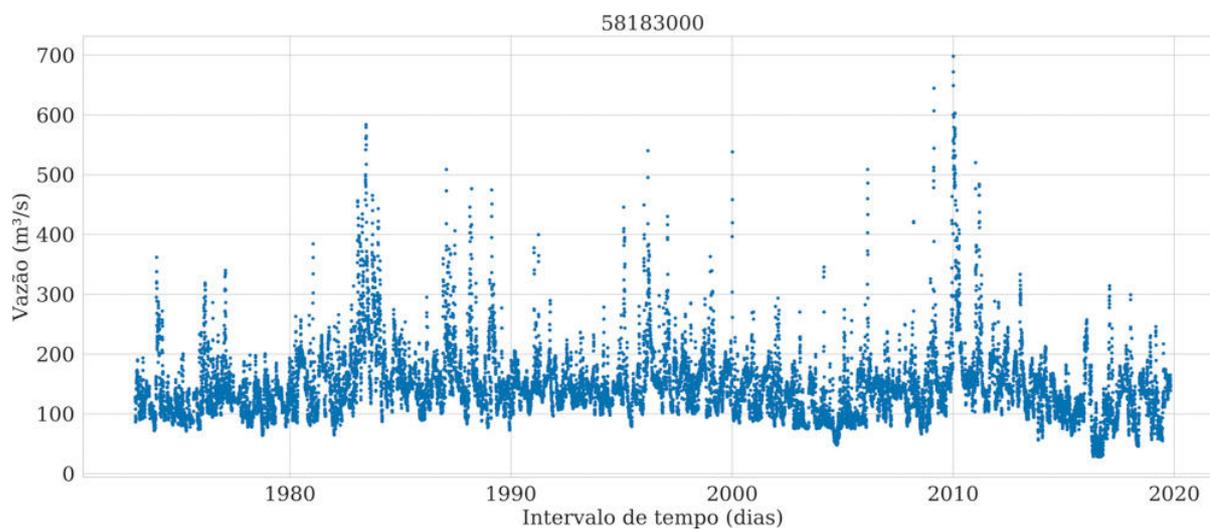
Fonte: Elaborada pelo autor (2022).

pelo Sistema Nacional de Informações Hidrológicas coordenado pela Agência Nacional de Águas, através de seu site.

Uma investigação preliminar das séries históricas de vazão (as Figuras 19, 20, 21, 22, 23, 24, 25 e 26) através de uma análise qualitativa visual indicou que os eventos mais importantes são variações sazonais e altos valores observados nas séries. Estes eventos se relacionam principalmente com o ciclos de cheias no verão e estiagens no inverno. Além disso, identificou-se muitos dados faltantes com anos de duração. Optou-se por não considerar esses dados no processo de modelagem pois, além da incerteza devido às medições, o modelo passaria a conter incertezas com tentativas de previsão desses dados, dos quais não se tem nenhuma informação.

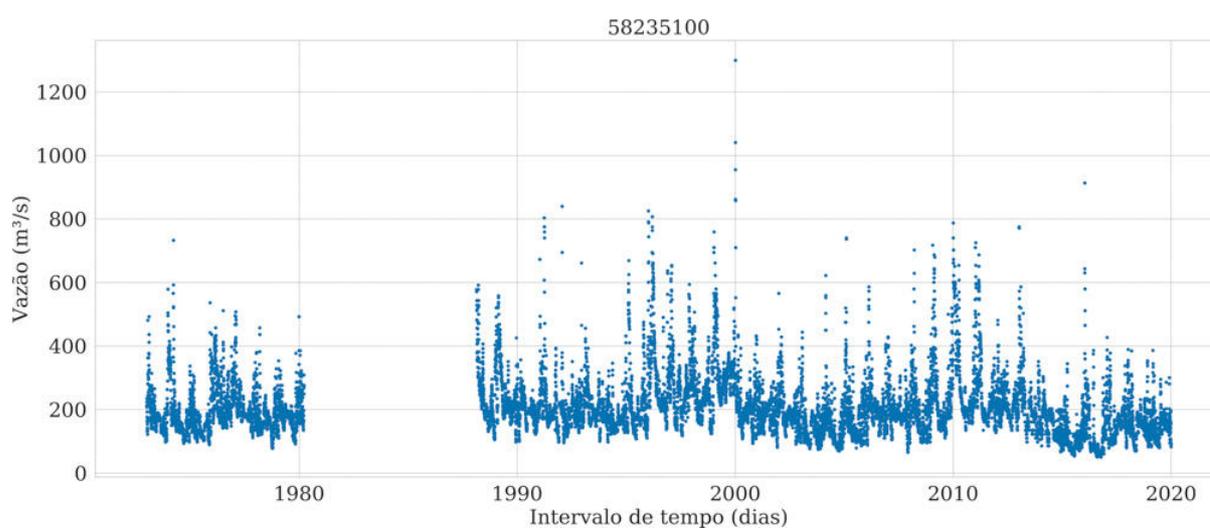
Não foi identificada na literatura uma explicação para a tendência de queda na vazão nos últimos seis anos observada nas série histórica de todas as estações. Entre as possíveis razões para esta tendência de queda cita-se: a) um prolongado período de baixos índices de Volume Útil Observado nos reservatórios à montante; b) um aumento no desvio das águas à montante para irrigação; c) alterações na regulação das bacias que incluem diminuição na defluência mínima para diluição e transporte de poluentes e maiores restrições ao uso de águas para geração e abastecimento em períodos de estiagem. Um exemplo de como interferências antropogênicas afetam o regime de vazão é a presença de 3 barragens que impactam os registros de estações à jusante.

Figura 19 – Serie histórica de vazão média diária da estação 58183000.



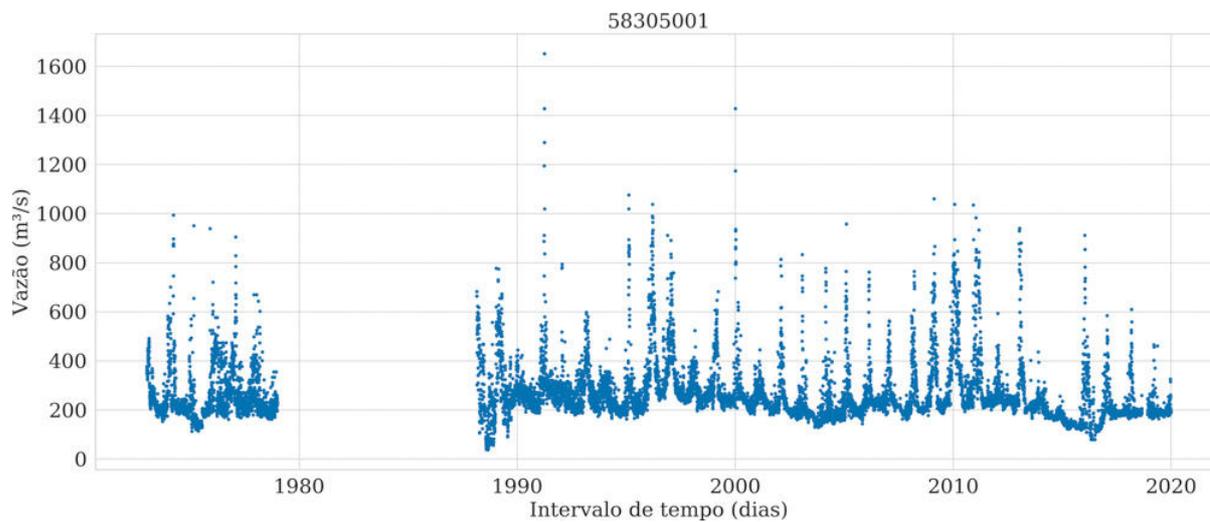
Fonte: Elaborada pelo autor (2022).

Figura 20 – Serie histórica de vazão média diária da estação 58235100.



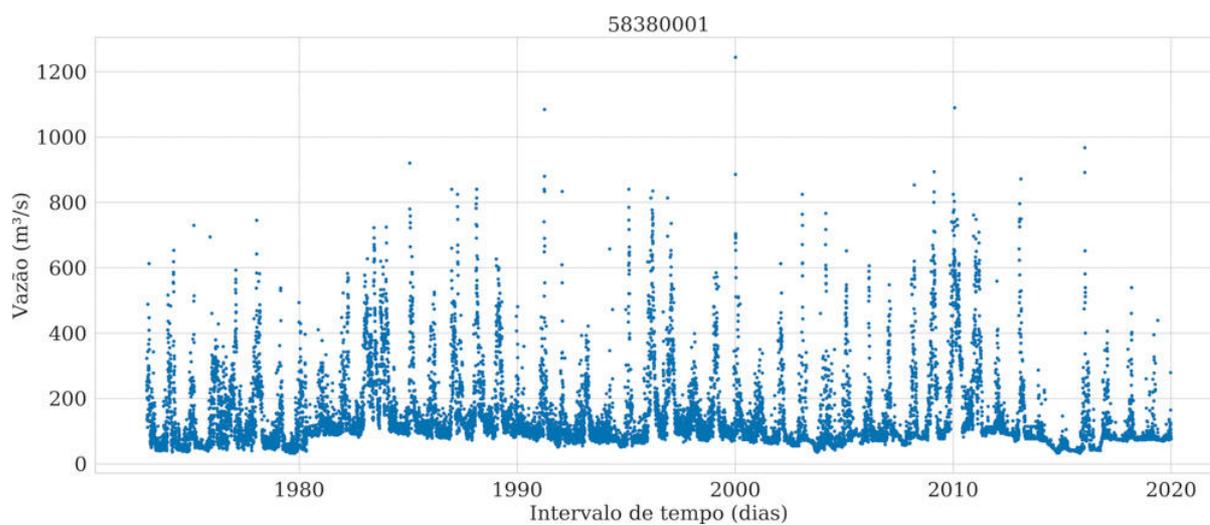
Fonte: Elaborada pelo autor (2022).

Figura 21 – Serie histórica de vazão média diária da estação 58305001.



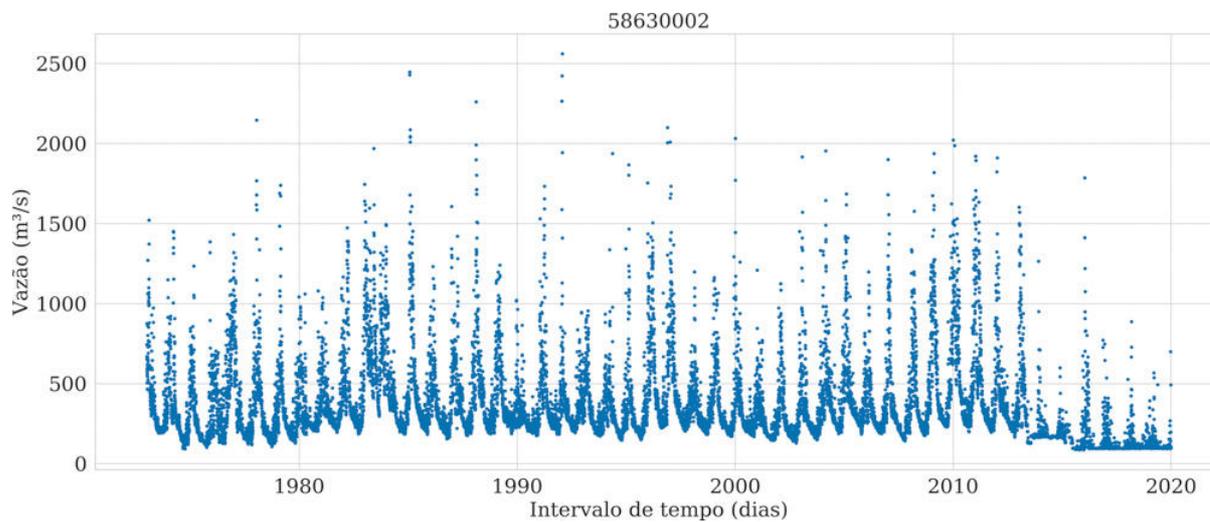
Fonte: Elaborada pelo autor (2022).

Figura 22 – Serie histórica de vazão média diária da estação 58380001.



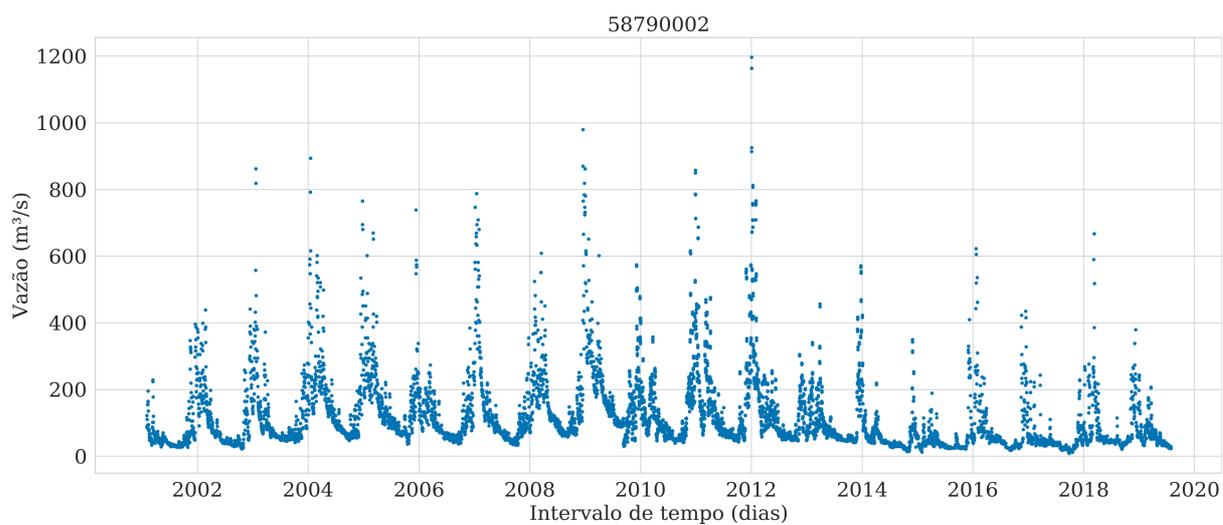
Fonte: Elaborada pelo autor (2022).

Figura 23 – Serie histórica de vazão média diária da estação 58630002.



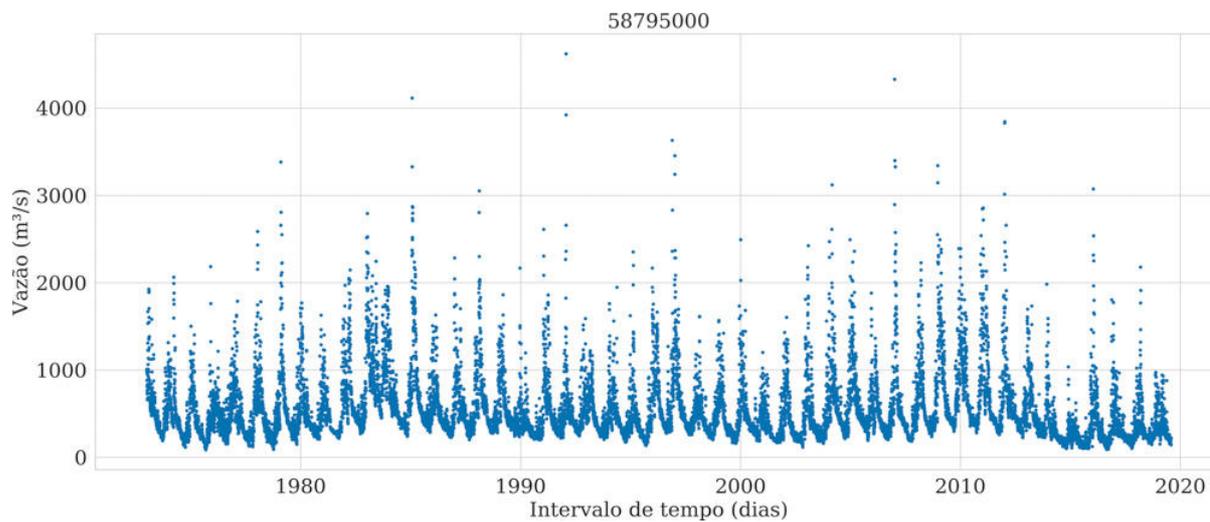
Fonte: Elaborada pelo autor (2022).

Figura 24 – Serie histórica de vazão média diária da estação 58790002.



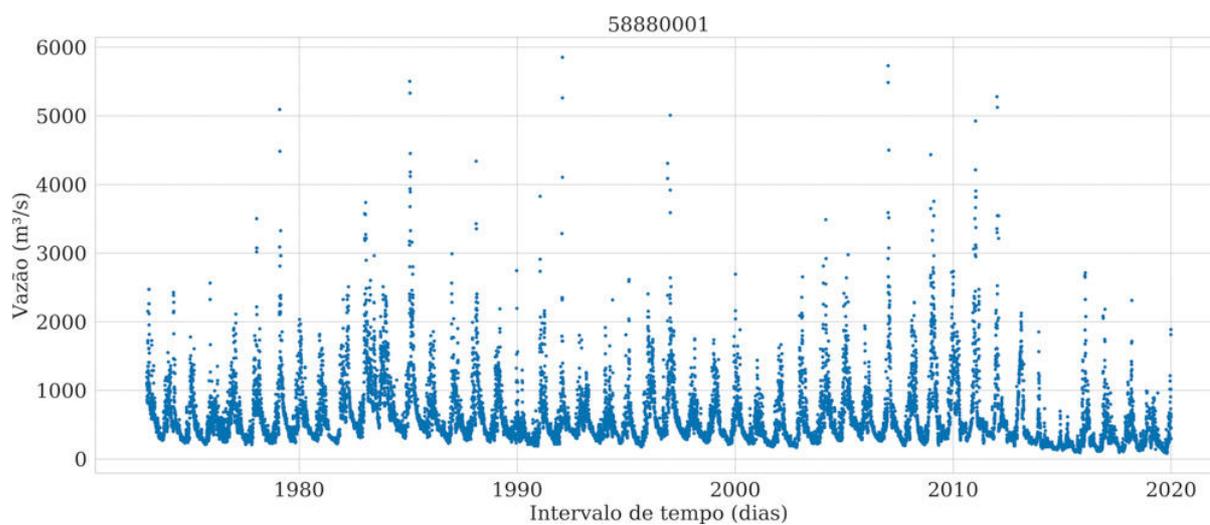
Fonte: Elaborada pelo autor (2022).

Figura 25 – Serie histórica de vazão média diária da estação 58795000.



Fonte: Elaborada pelo autor (2022).

Figura 26 – Serie histórica de vazão média diária da estação 58880001.



Fonte: Elaborada pelo autor (2022).

## 6 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados de previsão de vazão média diária de 4 horizontes temporais (1, 3, 5 e 7 dias à frente) para 8 estações fluviométricas apresentadas na tabela 4 da Bacia do rio Paraíba do Sul utilizando modelos desenvolvidos baseados em wavelets e redes neurais do tipo MLP. Com o objetivo de alcançar uma melhor generalização, e assim melhores resultados, cada um dos modelos propostos teve o processo de treinamento repetido 30 vezes. Em cada iteração, a técnica de validação cruzada 5-split TSS foi utilizada para selecionar os melhores parâmetros de cada modelo, processos de treinamento e teste. A capacidade preditiva dos modelos foi avaliada com base nos critérios listados na Seção 3.4.

Todos os processamentos de dados e desenvolvimento de modelos para este estudo foram feitos na linguagem de programação Python 3.6. As seguintes bibliotecas gerais foram utilizadas: Scikit-learn (94), PyWavelets (95), Seaborn (96) e Pandas (97).

### 6.1 INVESTIGAÇÃO DE FAMÍLIAS WAVELET

Inicialmente, uma exploração preliminar de seis diferentes famílias de wavelets comumente usadas, a saber - Haar, Daubechies (`db2`, `db3`, `db4`, `db5`, `db6`), Symlets (`sym2`, `sym3`, `sym4`), Coiflets (`coif4`, `coif5`, `coif6`), Biorthogonal(`bior1.3`, `bior2.8`) e Discrete Meyer (FIR Approximation) (`dmey`) foi realizada para identificar wavelets mãe em termos de precisão do modelo.

Após alguns testes foram escolhidas as seguintes 6 wavelets mãe para experimentos posteriores: Daubechies (`db4`), Symlets (`sym4`), Coiflets (`coif4`), Biortogonal (`bior2.8`), Reverse biortogonal (`rbio5.5`) e Discrete Meyer (FIR Approximation) (`dmey`). A sensibilidade das wavelets mãe selecionadas foi testada a partir dos dados de séries temporais de vazão diária provenientes de 5 estações fluviométricas na bacia do rio Paraíba do Sul (a saber, 58204000, 582835100, 5300001, 58880001 e 58974000).

As séries temporais da vazão foram divididas em dois conjuntos: treinamento (80%) e teste (20%), os quais posteriormente transformaram-se usando 6 wavelets mãe no nível de decomposição igual a 3. Para executar previsões de 7 dias à frente, a quantidade de dias antecedentes de sub-séries de aproximação e detalhe no vetor de entrada foi fixada igual a 10. O modelo  $WANN_{multi}$  foi adotado para executar a investigação das wavelets mãe. Dessa forma, o número total de modelos baseados em wavelets foi de 30 (6 wavelets mãe  $\times$  5 estações). O desempenho de modelos na previsão de vazão de 7 dias à frente utilizando wavelets mãe investigadas é relatado na Tabela 5.

Conforme os resultados apresentados na Tabela 5 e Figura 27 conclui-se que os modelos que empregam `coif4` e `db4` como wavelets mãe se mostraram mais eficazes em termos de valores de RMSE, MAPE e  $R^2$  na modelagem da 5 estações escolhidas. No

Tabela 5 – Desempenho de diferentes wavelets mãe para previsão de 7 dias à frente.

Estação	Metricas	Wavelets mãe					
		bior2.8	coif4	db4	dmey	rbio5.5	sym4
58204000	MAPE	18.12(2.17)	<b>3.62(0.51)</b>	5.48(1.25)	21.71(4.78)	15.62(4.58)	20.11(1.72)
	RMSE	31.10(1.33)	<b>8.19(0.91)</b>	14.50(0.76)	34.87(4.15)	26.33(4.31)	35.61(0.96)
	$R^2$	0.52(0.04)	<b>0.96(0.01)</b>	0.93(0.02)	0.34(0.16)	0.65(0.11)	0.43(0.02)
	$MAPE_{he}$	17.70(4.72)	<b>4.16(1.05)</b>	5.61(1.33)	28.18(10.37)	15.90(5.86)	26.02(3.90)
	$MAPE_{le}$	13.62(1.48)	<b>4.43(0.23)</b>	5.38(0.77)	15.6(1.04)	12.84(3.8)	17.22(1.5)
58235100	MAPE	27.09(5.91)	<b>6.6(0.65)</b>	9.13(1.65)	32.85(5.0)	24.99(6.38)	29.53(6.02)
	RMSE	50.41(3.8)	<b>18.49(1.66)</b>	24.05(3.63)	57.8(3.32)	44.35(5.39)	56.34(4.13)
	$R^2$	0.41(0.09)	<b>0.92(0.02)</b>	0.86(0.04)	0.24(0.09)	0.54(0.12)	0.26(0.12)
	$MAPE_{he}$	31.8(9.91)	<b>6.7(1.07)</b>	11.08(2.76)	47.07(9.69)	27.78(7.23)	42.59(10.19)
	$MAPE_{le}$	21.45(3.76)	<b>6.09(0.34)</b>	8.36(1.4)	21.77(2.72)	20.46(7.6)	22.73(0.72)
58300001	MAPE	10.89(1.42)	<b>5.26(0.43)</b>	6.99(0.7)	11.66(0.78)	11.37(3.13)	11.96(0.75)
	RMSE	68.57(4.54)	<b>33.01(0.56)</b>	37.91(2.43)	72.7(1.51)	55.88(4.12)	71.1(1.61)
	$R^2$	0.67(0.05)	<b>0.96(0.0)</b>	0.92(0.01)	0.66(0.01)	0.78(0.03)	0.65(0.02)
	$MAPE_{he}$	17.36(4.09)	<b>5.98(0.85)</b>	9.42(1.19)	22.16(2.45)	17.02(5.38)	23.86(5.07)
	$MAPE_{le}$	11.25(3.91)	<b>5.78(0.68)</b>	6.25(1.26)	8.82(1.15)	8.22(1.7)	9.26(2.52)
58880001	MAPE	37.55(7.0)	<b>9.09(0.22)</b>	13.32(3.05)	48.9(9.29)	33.04(9.8)	46.07(10.88)
	RMSE	224.43(12.58)	<b>88.4(10.38)</b>	101.35(15.5)	207.93(11.47)	178.13(19.12)	236.61(8.42)
	$R^2$	0.45(0.07)	<b>0.93(0.0)</b>	0.89(0.03)	0.36(0.07)	0.67(0.07)	0.42(0.04)
	$MAPE_{he}$	93.09(21.74)	<b>27.11(1.99)</b>	40.38(8.06)	89.37(19.1)	64.5(12.55)	95.24(10.95)
	$MAPE_{le}$	20.78(2.86)	<b>8.32(1.36)</b>	12.23(2.39)	35.89(9.14)	24.68(9.89)	31.4(6.42)
58974000	MAPE	34.07(5.7)	<b>8.39(0.3)</b>	14.51(4.34)	37.85(7.83)	29.2(8.01)	47.85(9.49)
	RMSE	245.7(7.18)	<b>86.01(1.32)</b>	114.54(15.99)	254.8(8.51)	194.27(30.17)	276.18(10.24)
	$R^2$	0.52(0.03)	<b>0.94(0.0)</b>	0.9(0.03)	0.43(0.04)	0.69(0.1)	0.4(0.05)
	$MAPE_{he}$	76.55(8.29)	<b>16.94(1.02)</b>	28.46(9.33)	87.86(15.06)	52.44(20.18)	103.87(15.57)
	$MAPE_{le}$	21.79(5.76)	<b>8.32(0.81)</b>	13.31(4.32)	25.84(5.61)	23.58(7.27)	32.32(3.49)

Fonte: Elaborada pelo autor (2022).

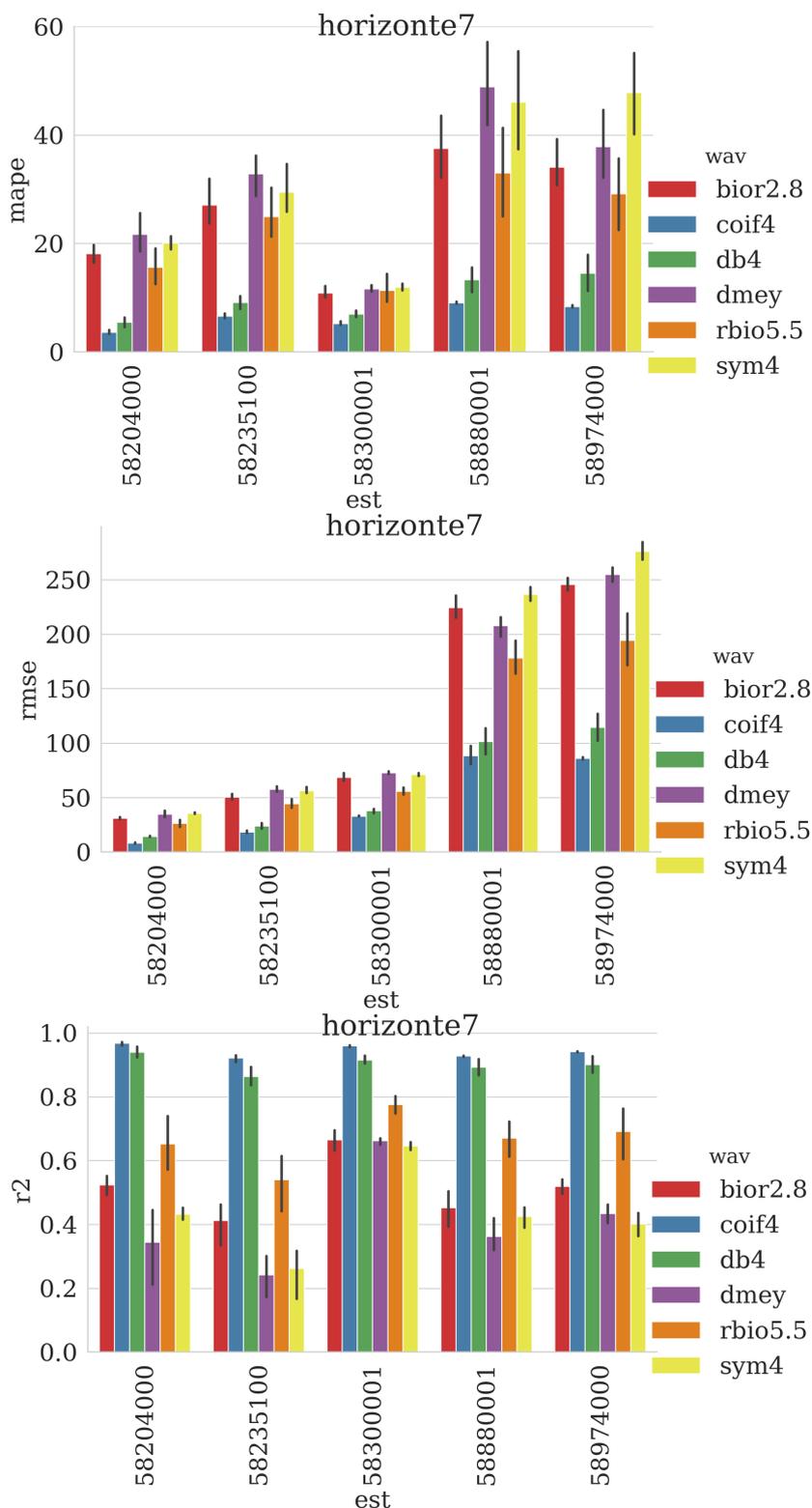
entanto, pode ser visto que os modelos que empregam `coif4` apresentaram os resultados são melhores em comparação aos modelos que empregam `db4`. Dessa maneira, a wavelet mãe `coif4` foi escolhida para experimentos posteriores.

Supõe-se que a wavelet mãe `coif4` tenha apresentado o melhor desempenho devido ao fato dela ter uma boa propriedade de localização tempo-frequência e ter sua forma quase simétrica e estreita, apropriada para a representação dos picos altos do hidrograma (98). Estas propriedades permitem que o modelo capture tanto a tendência subjacente quanto as variabilidades de curto prazo na série temporal. Além disso, existem poucos trabalhos com sua aplicação em modelos baseados em wavelets (21), embora proporcionem um alto desempenho quando acopladas com redes neurais como demonstra-se neste estudo.

## 6.2 AVALIAÇÃO GERAL DE COMPORTAMENTO DE MODELOS WANN E ANN SOBRE TODAS AS ESTAÇÕES

Para uma análise comparativa da eficiência dos modelos desenvolvidos  $WANN_{one}$ ,  $WANN_{multi}$  e ANN na previsão em curto prazo, os horizontes de previsão foram definidos para 1, 3, 5 e 7 dias à frente. Nas previsões foram utilizados dados de séries temporais de vazão diária de 8 estações fluviométricas (ver a Tabela 4) da Bacia do rio Paraíba do Sul.

Figura 27 – Desempenho dos modelos com as wavelets mãe investigadas na previsão de vazão de 7 dias à frente para 5 estações fluviométricas. Os valores médios das métricas de avaliação após 30 execuções.



Fonte: Elaborada pelo autor (2022).

Nos modelos  $WANN_{one}$  e  $WANN_{multi}$  a série temporal de vazão foi decomposta usando `coif4` wavelet mãe em 3 níveis de decomposição em sub-séries temporais de aproximação e detalhe. Em seguida, a séries temporal original e as sub-séries de aproximação e detalhe foram então divididas em conjuntos de dados de treinamento (80%) e teste (20%) conforme demonstrado na Tabela 6).

Tabela 6 – Tamanho de conjuntos de dados de treinamento e teste, em dias.

<b>Estação</b>	<b>Training 80%</b>	<b>Test 20%</b>
58880001	13708	3427
58790002	5404	1351
58795000	13604	3402
58630002	13732	3434
58380001	13732	3434
58305001	10975	2744
58235100	11419	2855
58183000	13657	3415

Fonte: Elaborada pelo autor (2022).

A Tabela 7 apresenta a quantidade de dias antecedentes testados, que explicam a vazão prevista (*lags*, em inglês), para cada horizonte temporal como as entradas para os modelos. Optou-se por testar diferentes quantidades de dias antecedentes para garantir que

Tabela 7 – Quantidade de *lags* para os modelos ANN,  $WANN_{one}$  e  $WANN_{multi}$ .

<b>Horizonte</b>	<b>Defasagens temporais</b>
1	2, 3, 5, 7
3	5, 7, 10, 14
5	7, 10, 14
7	10, 14

Fonte: Elaborada pelo autor (2022).

os modelos desenvolvidos recebam estrutura temporal suficiente para identificar preditores adequados. Dessa forma, garante-se que não sejam incluídas muitas ou poucas quantidades de variáveis explicativas que possam potencialmente diminuir o desempenho do modelo.

A técnica de validação cruzada de séries temporais TSS de *5-splits* foi empregada para ajustar, treinar e testar os modelos. Para realizar as previsões para cada um dos horizontes temporais foi desenvolvido o modelo separado. Os experimentos foram repetidos 30 vezes. Assim, nestes experimentos foram desenvolvidos os seguintes modelos:

- 416 modelos  $WANN_{one}$ : (4 horizontes temporais)  $\times$  (13 lags)  $\times$  (8 estações).
- 416 modelos  $WANN_{multi}$ : (4 horizontes temporais)  $\times$  (13 lags)  $\times$  (8 estações).
- 416 modelos ANN: (4 horizontes temporais)  $\times$  (13 lags)  $\times$  (8 estações).

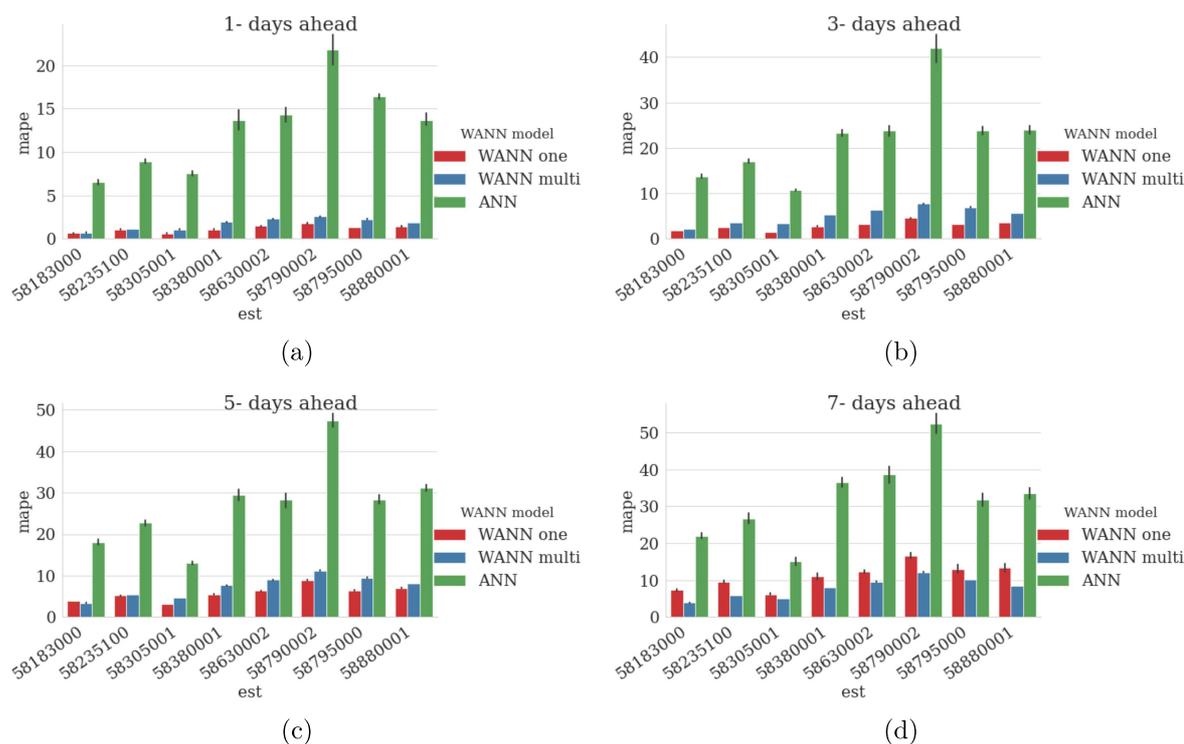
A tabela 8 demonstra defasagens de tempo com as quais os modelos exibiram, na média, melhores desempenhos. Os valores médios das métricas de desempenho definidas na Sessão 3.4 com essas defasagens temporais são representados em Tabelas 9, 11 e 10. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses.

Tabela 8 – Os melhores *lags* de entrada para os modelos ANN,  $WANN_{one}$  e  $WANN_{multi}$  após 30 execuções.

Horizonte	$WANN_{one}$	$WANN_{multi}$	ANN
1	7	7	3
3	7	7	7
5	7	7	10
7	10	10	14

Fonte: Elaborada pelo autor (2022).

Figura 28 – ANN,  $WANN_{one}$  e  $WANN_{multi}$ : variação do MAPE média em todos os horizontes de previsão: (a) 1 dia à frente; (b) 3 dias à frente; (c) 5 dias à frente; (d) 7 dias à frente.



Fonte: Elaborada pelo autor (2022).

As Figuras 28, 29 e 30 representam a variação das medidas de desempenho dos modelos em todos os horizontes de previsão para 8 estações fluviométricas.

Comparando o desempenho dos modelos, observa-se que, a partir do horizonte de previsão de 1 dia à frente, todos os modelos ANN tendem a ter dificuldade na previsão para todas as estações demonstrando erros muito altos em comparação ao modelos baseados

Tabela 9 – Desempenho dos modelos para diferentes horizontes de previsão para as estações fluviométricas 58630002, 58380001 e 58305001. Os valores médios das métricas de avaliação após 30 execuções. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses.

Hor	Modelo	RMSE	MAPE(%)	$R^2$
<b>Estação 58630002</b>				
1	WANN <sub>one</sub>	<b>8.6306(1.0198)</b>	<b>1.4656(0.14)</b>	<b>0.9984(0.0004)</b>
	WANN <sub>multi</sub>	21.3393(4.4715)	3.6208(1.1453)	0.9943(0.0035)
	ANN	89.8545(1.0727)	14.3271(1.3785)	0.8482(0.0037)
3	WANN <sub>one</sub>	<b>14.3008(0.7116)</b>	<b>3.2124(0.159)</b>	<b>0.9956(0.0004)</b>
	WANN <sub>multi</sub>	42.3667(6.8643)	7.576(1.6069)	0.9689(0.0114)
	ANN	136.8424(1.0252)	23.8723(1.9065)	0.648(0.0053)
5	WANN <sub>one</sub>	<b>25.8955(0.9433)</b>	<b>6.3451(0.2217)</b>	<b>0.9854(0.0011)</b>
	WANN <sub>multi</sub>	56.6119(17.115)	9.7972(8.2438)	0.9382(0.0757)
	ANN	154.2386(1.138)	28.2057(2.7561)	0.5529(0.0066)
7	WANN <sub>one</sub>	<b>46.9515(3.3588)</b>	12.2829(1.2845)	<b>0.9507(0.0068)</b>
	WANN <sub>multi</sub>	60.2144(2.9989)	<b>9.7153(0.288)</b>	0.9359(0.0074)
	ANN	169.2887(1.2333)	38.5442(3.59)	0.4616(0.0078)
<b>Estação 58380001</b>				
1	WANN <sub>one</sub>	<b>3.2975(0.4883)</b>	<b>1.086(0.1319)</b>	<b>0.9985(0.0004)</b>
	WANN <sub>multi</sub>	9.3445(2.5048)	3.0109(0.9786)	0.9914(0.0055)
	ANN	42.5234(0.5821)	13.629(1.8329)	0.7856(0.0059)
3	WANN <sub>one</sub>	<b>6.0173(0.459)</b>	<b>2.7273(0.1627)</b>	<b>0.9951(0.0007)</b>
	WANN <sub>multi</sub>	21.4897(4.9547)	6.8436(5.5607)	0.9446(0.0617)
	ANN	57.2111(0.2388)	23.3652(1.205)	0.6112(0.0032)
5	WANN <sub>one</sub>	<b>11.4252(0.3194)</b>	<b>5.4972(0.155)</b>	<b>0.9819(0.001)</b>
	WANN <sub>multi</sub>	26.3135(5.3314)	8.1874(2.5564)	0.9171(0.0397)
	ANN	65.5757(0.5022)	29.5117(2.3527)	0.4882(0.0078)
7	WANN <sub>one</sub>	<b>21.8058(2.0226)</b>	11.0104(2.2298)	<b>0.9312(0.0125)</b>
	WANN <sub>multi</sub>	27.6664(1.2122)	<b>8.0273(0.1703)</b>	0.9122(0.008)
	ANN	71.9832(0.4644)	36.5138(2.1179)	0.382(0.008)
<b>Estação 58305001</b>				
1	WANN <sub>one</sub>	<b>3.1227(0.4998)</b>	<b>0.6259(0.0635)</b>	<b>0.9987(0.0004)</b>
	WANN <sub>multi</sub>	9.068(2.4735)	1.6929(0.5675)	0.9912(0.0059)
	ANN	40.923(0.2096)	7.2081(0.2718)	0.7722(0.0023)
3	WANN <sub>one</sub>	<b>6.2562(0.4039)</b>	<b>1.5133(0.0717)</b>	<b>0.9949(0.0007)</b>
	WANN <sub>multi</sub>	21.0645(2.565)	3.9668(0.7553)	0.9427(0.0146)
	ANN	55.5112(0.373)	10.8057(0.3105)	0.5815(0.0056)
5	WANN <sub>one</sub>	<b>12.4647(0.4469)</b>	<b>3.0772(0.1123)</b>	<b>0.9797(0.0015)</b>
	WANN <sub>multi</sub>	25.408(7.3267)	5.0873(2.7629)	0.9096(0.0984)
	ANN	64.7149(0.3809)	13.0914(0.6499)	0.4317(0.0067)
7	WANN <sub>one</sub>	<b>26.127(1.2658)</b>	6.1729(0.7435)	0.9107(0.0081)
	WANN <sub>multi</sub>	25.923(0.31)	<b>5.0736(0.0954)</b>	<b>0.9136(0.0028)</b>
	ANN	71.2747(0.7651)	15.0224(1.8813)	0.3111(0.0149)

Fonte: Elaborada pelo autor (2022).

Tabela 10 – Desempenho dos modelos para diferentes horizontes de previsão para estações fluviométricas 58880001, 58790002 e 58795000. Os valores médios das métricas de avaliação após 30 execuções. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses.

Hor	Modelo	RMSE	MAPE(%)	$R^2$
<b>Estação 58880001</b>				
1	WANN <sub>one</sub>	<b>15.306(2.1503)</b>	<b>1.4034(0.161)</b>	<b>0.9987(0.0004)</b>
	WANN <sub>multi</sub>	36.3816(8.5079)	3.1084(1.1099)	0.9946(0.0035)
	ANN	154.4575(0.9022)	13.0286(0.8524)	0.8653(0.0016)
3	WANN <sub>one</sub>	<b>31.0308(1.6321)</b>	<b>3.529(0.125)</b>	<b>0.9948(0.0005)</b>
	WANN <sub>multi</sub>	72.2528(7.7172)	6.9166(1.8341)	0.9714(0.0156)
	ANN	262.8434(1.5547)	23.6565(2.8392)	0.6101(0.0046)
5	WANN <sub>one</sub>	<b>58.574(1.432)</b>	<b>6.9835(0.1706)</b>	<b>0.9815(0.0009)</b>
	WANN <sub>multi</sub>	96.9221(4.5961)	8.9075(7.2814)	0.9433(0.0691)
	ANN	313.1169(2.4264)	31.7331(2.8862)	0.4472(0.0086)
7	WANN <sub>one</sub>	100.3913(9.7698)	13.2605(2.7912)	0.945(0.0112)
	WANN <sub>multi</sub>	<b>101.0156(2.293)</b>	<b>8.4159(0.2667)</b>	<b>0.9451(0.0027)</b>
	ANN	354.7394(3.5433)	33.5065(2.3991)	0.2923(0.0141)
<b>Estação 58790002</b>				
1	WANN <sub>one</sub>	<b>3.2975(0.4883)</b>	<b>1.086(0.1319)</b>	<b>0.9985(0.0004)</b>
	WANN <sub>multi</sub>	9.3445(2.5048)	3.0109(0.9786)	0.9914(0.0055)
	ANN	37.3322(0.1286)	20.7213(2.5397)	0.7532(0.0017)
3	WANN <sub>one</sub>	<b>6.0173(0.459)</b>	<b>2.7273(0.1627)</b>	<b>0.9951(0.0007)</b>
	WANN <sub>multi</sub>	21.4897(4.9547)	6.8436(5.5607)	0.9446(0.0617)
	ANN	58.3333(0.3693)	41.9202(4.3041)	0.398(0.0076)
5	WANN <sub>one</sub>	<b>11.4252(0.3194)</b>	<b>5.4972(0.155)</b>	<b>0.9819(0.001)</b>
	WANN <sub>multi</sub>	26.3135(5.3314)	8.1874(2.5564)	0.9171(0.0397)
	ANN	65.5177(0.4984)	62.1989(7.4905)	0.2405(0.0115)
7	WANN <sub>one</sub>	<b>21.8058(2.0226)</b>	11.0104(2.2298)	<b>0.9312(0.0125)</b>
	WANN <sub>multi</sub>	27.6664(1.2122)	<b>8.0273(0.1703)</b>	0.9122(0.008)
	ANN	70.0271(0.4284)	56.4753(7.0217)	0.134(0.0106)
<b>Estação 58795000</b>				
1	WANN <sub>one</sub>	<b>12.3666(1.1139)</b>	<b>1.3264(0.0988)</b>	<b>0.9987(0.0002)</b>
	WANN <sub>multi</sub>	64.3323(7.629)	3.71(1.255)	0.9935(0.0043)
	ANN	145.4666(1.0502)	16.132(0.4425)	0.8426(0.0023)
3	WANN <sub>one</sub>	<b>24.3042(1.0754)</b>	<b>3.2543(0.1168)</b>	<b>0.9951(0.0004)</b>
	WANN <sub>multi</sub>	81.5091(12.9603)	8.1726(1.6094)	0.9662(0.0155)
	ANN	237.3808(1.7235)	23.7705(1.4159)	0.5761(0.0062)
5	WANN <sub>one</sub>	<b>45.3977(1.3046)</b>	<b>6.4467(0.1769)</b>	<b>0.9828(0.001)</b>
	WANN <sub>multi</sub>	99.9696(24.6135)	10.3602(4.9117)	0.9315(0.0641)
	ANN	266.9964(1.8709)	29.9947(2.4862)	0.4637(0.0075)
7	WANN <sub>one</sub>	<b>82.5145(8.6914)</b>	13.0341(2.75)	<b>0.9426(0.0126)</b>
	WANN <sub>multi</sub>	103.5653(6.1765)	<b>10.3087(0.1135)</b>	0.9287(0.0028)
	ANN	285.4033(1.846)	31.8577(2.7557)	0.3741(0.0065)

Fonte: Elaborada pelo autor (2022).

em wavelets. Por exemplo, para as previsões maiores que 3 dias á frente os valores de coeficiente  $R^2$  abaixo a 0.5 indicam a ineficiência das previsões de modelos ANN.

Tabela 11 – Desempenho dos modelos para diferentes horizontes de previsão para estações fluviométricas 58235100 e 58183000. Os valores médios das métricas de avaliação após 30 execuções. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses.

Hor	Modelo	RMSE	MAPE(%)	$R^2$
<b>Estação 58235100</b>				
1	WANN <sub>one</sub>	<b>4.3301(0.657)</b>	<b>1.0625(0.0966)</b>	<b>0.9965(0.0011)</b>
	WANN <sub>multi</sub>	9.0896(36.0261)	2.1809(3.5719)	0.7534(4.6415)
	ANN	28.3689(0.2198)	8.9115(0.2601)	0.8519(0.0023)
3	WANN <sub>one</sub>	<b>6.882(0.5937)</b>	<b>2.5234(0.1345)</b>	<b>0.9913(0.0015)</b>
	WANN <sub>multi</sub>	13.7551(3.039)	4.5894(1.3532)	0.9684(0.0155)
	ANN	48.2725(0.1985)	17.0776(0.8902)	0.5714(0.0035)
5	WANN <sub>one</sub>	<b>12.7459(1.0989)</b>	<b>5.1907(0.1725)</b>	<b>0.9701(0.0053)</b>
	WANN <sub>multi</sub>	18.1987(6.7881)	6.1338(3.7145)	0.9357(0.086)
	ANN	56.1899(0.43)	22.7029(1.1894)	0.4196(0.0089)
7	WANN <sub>one</sub>	20.5705(1.3685)	9.521(0.9592)	0.9225(0.0101)
	WANN <sub>multi</sub>	<b>18.3097(0.4295)</b>	<b>5.9324(0.0815)</b>	<b>0.943(0.0028)</b>
	ANN	61.3359(0.918)	26.6426(2.3499)	0.3088(0.0207)
<b>Estação 58183000</b>				
1	WANN <sub>one</sub>	<b>1.179(0.1493)</b>	<b>0.6438(0.0845)</b>	<b>0.9995(0.0001)</b>
	WANN <sub>multi</sub>	4.3728(0.8984)	1.2889(0.5187)	0.9974(0.002)
	ANN	13.0927(0.1577)	6.2019(0.1955)	0.946(0.0013)
3	WANN <sub>one</sub>	<b>3.2544(0.1483)</b>	<b>1.7953(0.0807)</b>	<b>0.9962(0.0003)</b>
	WANN <sub>multi</sub>	6.8686(1.9977)	3.0704(1.2044)	0.9849(0.0115)
	ANN	26.7603(0.1765)	13.7303(0.5544)	0.7773(0.0029)
5	WANN <sub>one</sub>	<b>7.0669(0.1871)</b>	<b>3.8845(0.1072)</b>	<b>0.9821(0.001)</b>
	WANN <sub>multi</sub>	8.7436(4.1114)	4.1509(2.6548)	0.9689(0.0587)
	ANN	33.9434(0.3539)	18.4905(1.005)	0.6277(0.0078)
7	WANN <sub>one</sub>	12.822(0.9091)	7.2835(0.9753)	0.941(0.0093)
	WANN <sub>multi</sub>	<b>8.0547(0.4291)</b>	<b>3.8414(0.0613)</b>	<b>0.9782(0.0008)</b>
	ANN	37.9489(0.3981)	21.7504(1.0817)	0.5371(0.0098)

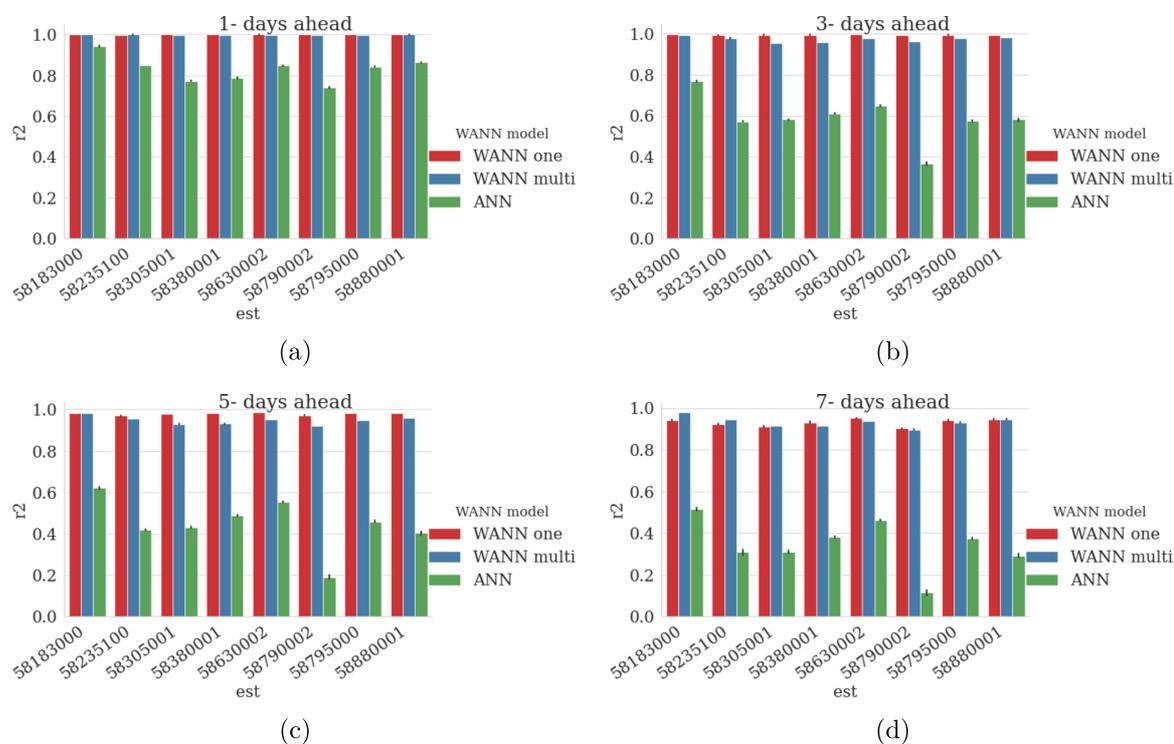
Fonte: Elaborada pelo autor (2022).

A mesma tendência pode ser observada em termos de MAPE e RMSE: na previsão de 3 dias à frente, o modelo ANN, na maioria dos casos, apresenta valores de MAPE acima de 20%, enquanto para horizonte 7, valores de MAPE acima de 30%.

Badrzadeh, Sarukkalige e Jayawardena(72) notaram que a perda de confiabilidade na previsão com o aumento de horizonte de previsão é um dos problemas inerentes a todos os métodos preditivos baseados em IA. Por exemplo, no caso da estação 58880001 obteve-se:

- Para previsões de 1 dia à frente, o modelo WANN<sub>one</sub> produziu um RMSE = 15,306 m<sup>3</sup>/s, MAPE = 1,4% e R<sup>2</sup> = 0,9987, enquanto a previsão usando o modelo ANN deu os seguintes resultados: RMSE = 154,45 m<sup>3</sup>, MAPE = 13,03% e R<sup>2</sup> = 0,8653.

Figura 29 – ANN,  $WANN_{one}$  e  $WANN_{multi}$  do  $R^2$  média em todos os horizontes de previsão: (a) 1 dia à frente; (b) 3 dias à frente; (c) 5 dias à frente; (d) 7 dias à frente.

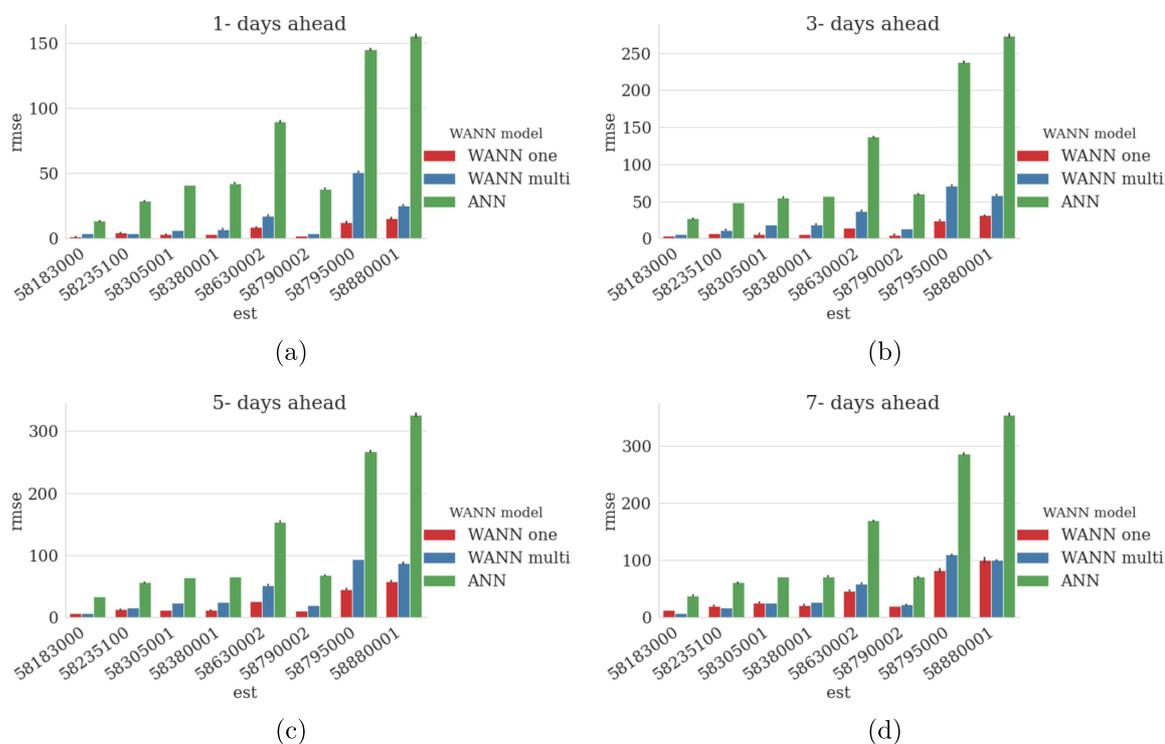


Fonte: Elaborada pelo autor (2022).

O RMSE e o MAPE diminuíram 90%, e o coeficiente  $R^2$  aumentou 13%, em relação à previsão usando o modelo ANN.

- Para previsões de 3 dias à frente, os modelos baseados em wavelets produziram um  $RMSE = 31,03 \text{ m}^3 / \text{s}$ ,  $MAPE = 3,53\%$  e  $R^2 = 0,9948$ , enquanto a previsão usando o modelo ANN deu os seguintes resultados:  $RMSE = 313,74 \text{ m}^3$ ,  $MAPE = 31,73\%$  e  $R^2 = 0,44$ . O RMSE e o MAPE diminuíram 81% e 78%, respectivamente, e o coeficiente  $R^2$  aumentou 50%, em relação à previsão usando o modelo ANN.
- Para previsões de 5 dias à frente, os modelos baseados em wavelets produziram um  $RMSE = 58,57 \text{ m}^3 / \text{s}$ ,  $MAPE = 6,98\%$  e  $R^2 = 0,9815$ , enquanto a previsão usando o modelo ANN deu os seguintes resultados:  $RMSE = 313,74 \text{ m}^3$ ,  $MAPE = 31,73\%$  e  $R^2 = 0,44$ . O RMSE e o MAPE diminuíram 81% e 78%, respectivamente, e o coeficiente  $R^2$  aumentou 50%, em relação à previsão usando o modelo ANN.
- Para previsões de 7 dias à frente, os modelos baseados em wavelets produziram um  $RMSE = 101,02 \text{ m}^3 / \text{s}$ ,  $MAPE = 8,4\%$  e  $R^2 = 0,94$ , enquanto a previsão usando o modelo ANN deu os seguintes resultados:  $RMSE = 354,7 \text{ m}^3$ ,  $MAPE = 33,5\%$  e  $R^2 = 0,29$ . O RMSE e o MAPE diminuíram 71% e 75%, respectivamente, e o coeficiente  $R^2$  aumentou 69%, em relação à previsão usando o modelo ANN.

Figura 30 – ANN,  $WANN_{one}$  e  $WANN_{multi}$ : variação do RMSE média em todos os horizontes de previsão: (a) 1 dia à frente; (b) 3 dias à frente; (c) 5 dias à frente; (d) 7 dias à frente



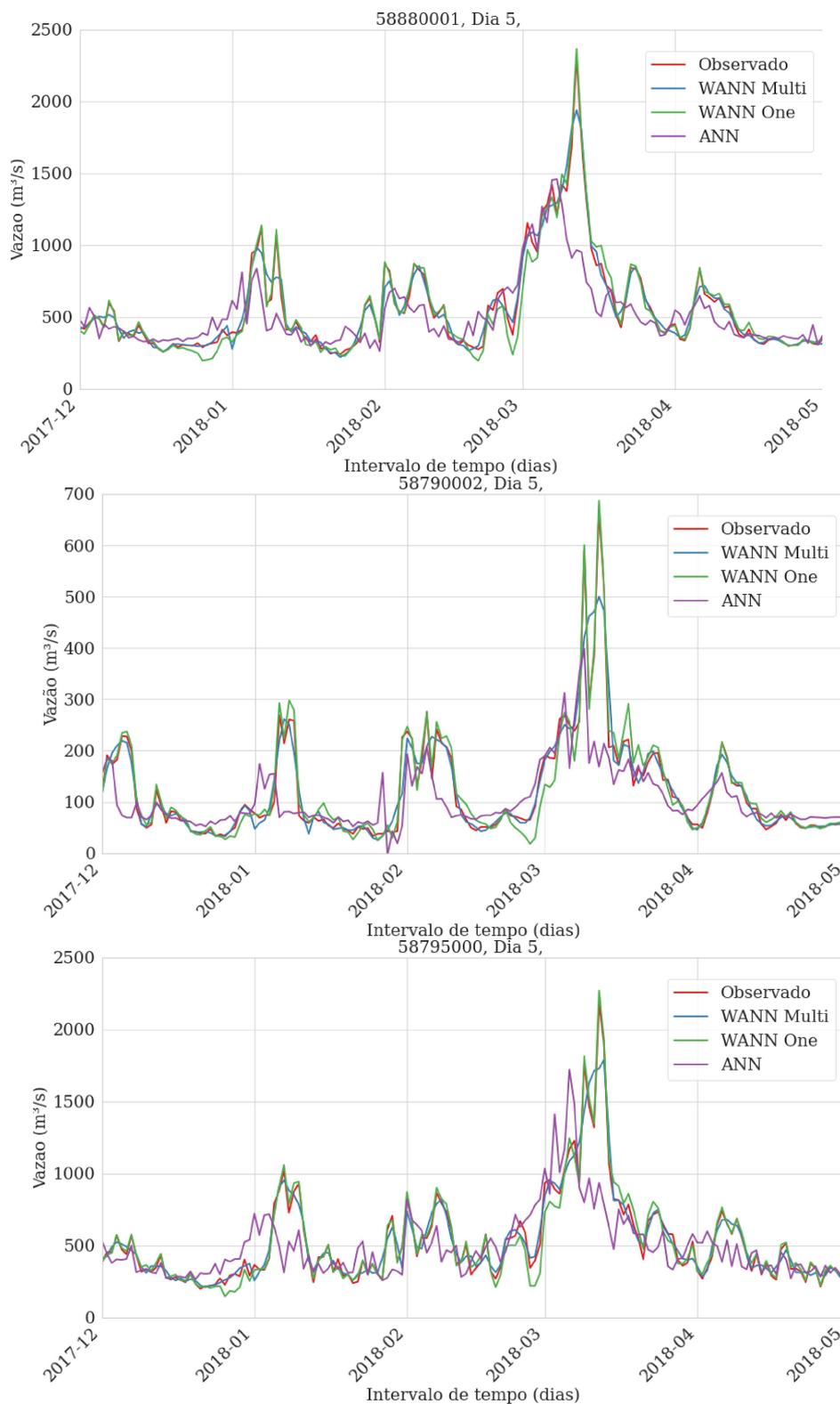
Fonte: Elaborada pelo autor (2022).

A análise das medidas de erro apresentadas suporta o comportamento bem ajustado dos modelos WANNs para todas as 8 estações modeladas. Valores consistentemente bons de  $R^2$ , bem como outras medidas de erro, descrevem a qualidade do ajuste dos modelos WANNs desenvolvidos.

Para compreender quão perto as previsões estão de suas observações correspondentes estão apresentados os gráficos de dispersão dos valores observados versus os previstos (ver as Figuras 48, 49, 50, 54, 53, 52, 51 e 55 no Apêndice A). No caso dos modelos WANN, os gráficos de dispersão descrevem a ausência de superestimativas ou subestimativas e demonstram que os modelos WANN se ajustam melhor à linha de unidade, assim provando a precisão e eficiência dos modelos. Observa-se que os modelos WANN previram valores de vazão muito mais próximos dos valores de vazão observada correspondente em comparação com os modelos ANN para horizontes temporais mais curtos e mais longos.

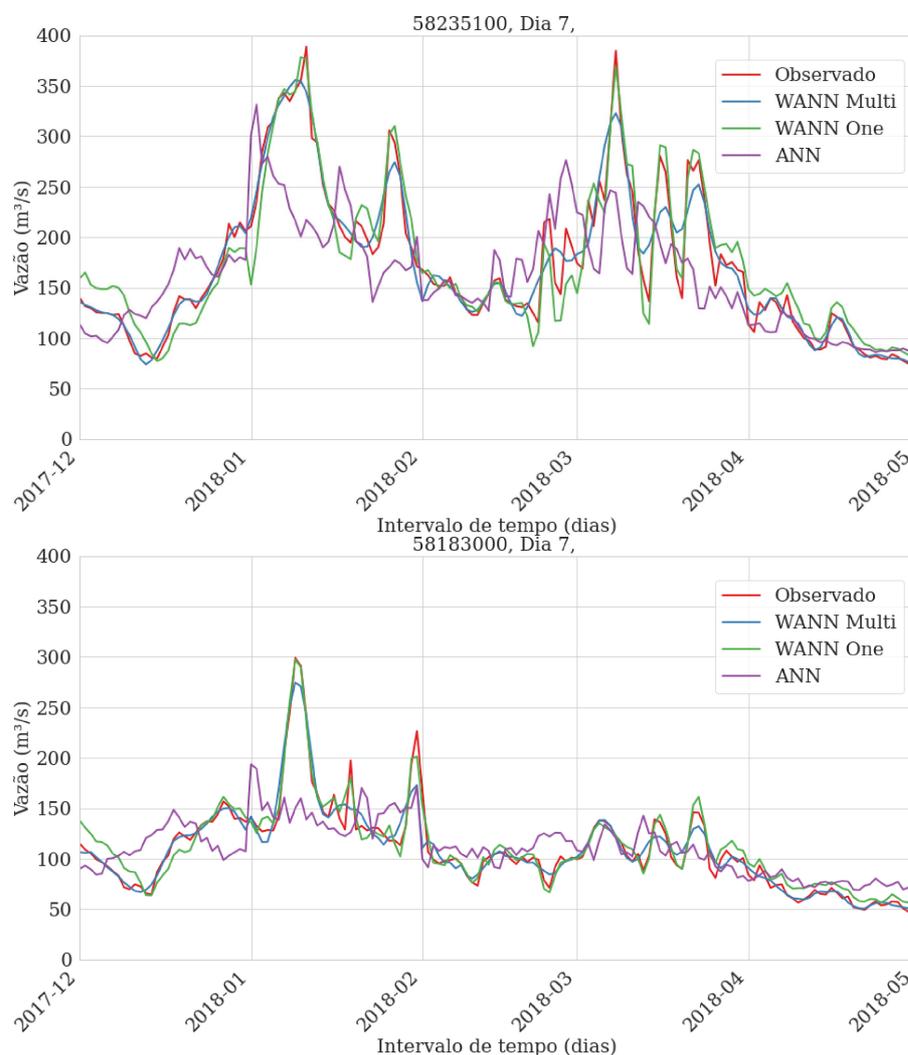
Além disso, conforme as Figuras 31, 32 e 33 e as Figuras 40, 41, 42, 43, 44, 45, 46 e 47 do Apêndice A, os modelos baseados em wavelets conseguem capturar de forma muito bem sucedida a hidrodinâmica complexa (tais como crescimento e queda de vazão) do rio durante todo o período de previsão. Em relação ao modelo ANN, notou-se que o modelo tem a tendência de sub e superestimar as vazões quando acontece a mudança no comportamento do rio, tanto com as vazões baixas, quanto com as altas. Também,

Figura 31 – Comparação da vazão observada e prevista por modelos WANNs e ANN para horizonte de previsão de 5 dias à frente para as estações: (a) 58880001; (b) 58790002 e (c) 58795000



Fonte: Elaborada pelo autor (2022).

Figura 32 – Comparação da vazão observada e prevista por modelos WANNs e ANN para horizonte de previsão de 7 dias à frente para as estações: (a) 558235100 e (b) 58183000.

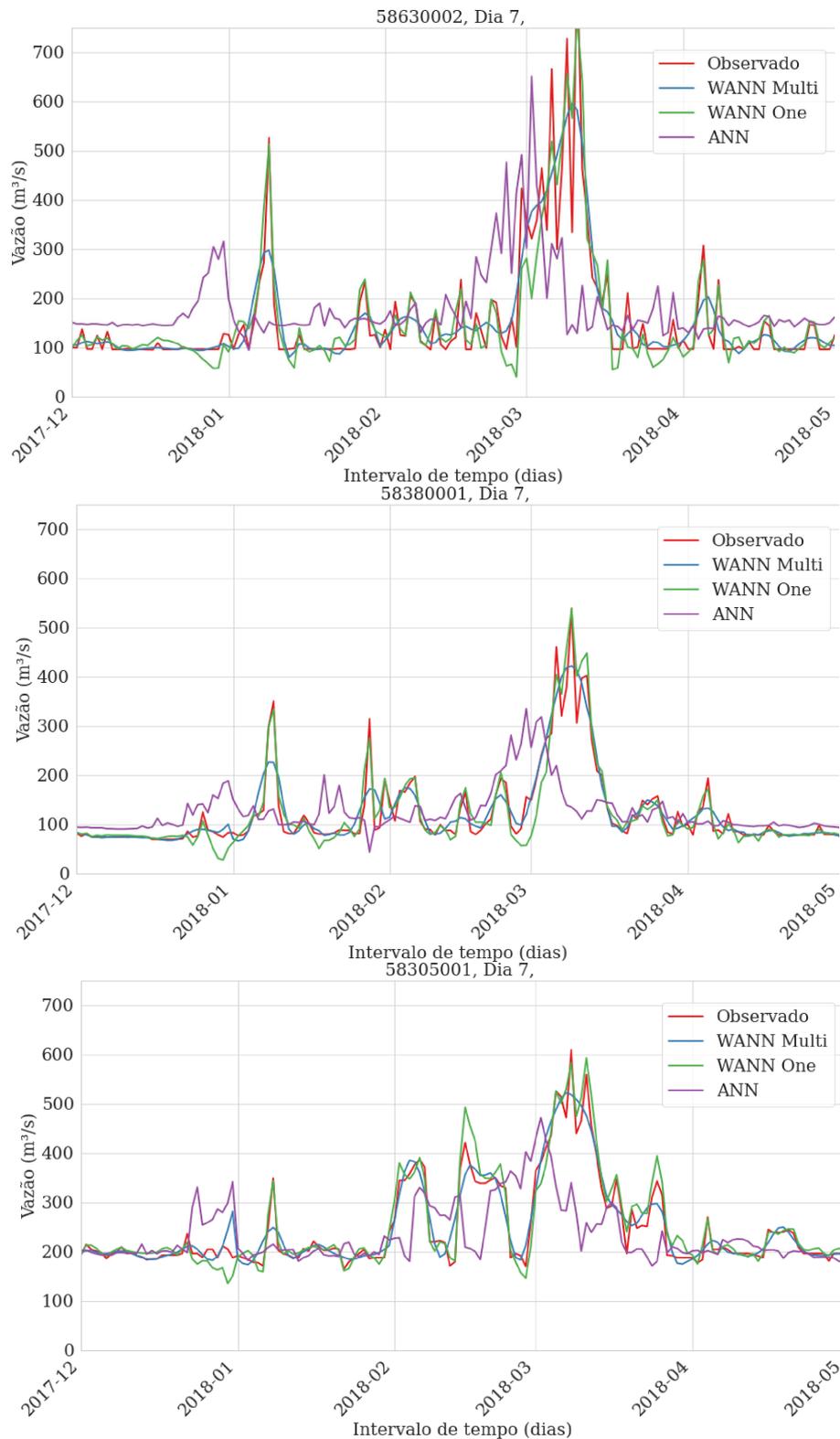


Fonte: Elaborada pelo autor (2022).

notou-se a presença de oscilações nos picos médios e altos e até nos fluxos baixos. Para horizontes mais distantes, os modelos ANN produzem subestimações significativas.

Além disso, por inspeção visual entre os valores observados e previstos com modelos de ANN, notou-se o deslocamento dos valores de vazão em alguns passos de tempo. A mudança de fase está aumentando com o aumento do horizonte da previsão. Essa mudança de fase diminui a precisão geral dos modelos de ANN resultando em valores menores de  $R^2$  e MAPE e maior RMSE. Esse fenômeno já foi destacado por alguns autores (43, 26, 24, 37) para previsões usando modelos ANN convencionais. Isso pode acontecer devido à forte autocorrelação entre o último valor observado e os valores alvo ao usar séries temporais de vazão como as entradas para a rede (26). A partir das Figuras 31, 32 e 33 pode-se notar que o problema da previsão defasada é completamente eliminado pelas wavelets, o que pode ser considerado uma importante contribuição para aumento da precisão e robustez

Figura 33 – Comparação da vazão observada e prevista por modelos WANNs e ANN para horizonte de previsão de 7 dias à frente para as estações: (a) 58630002; (b) 58380001 e (c) 58305001



Fonte: Elaborada pelo autor (2022).

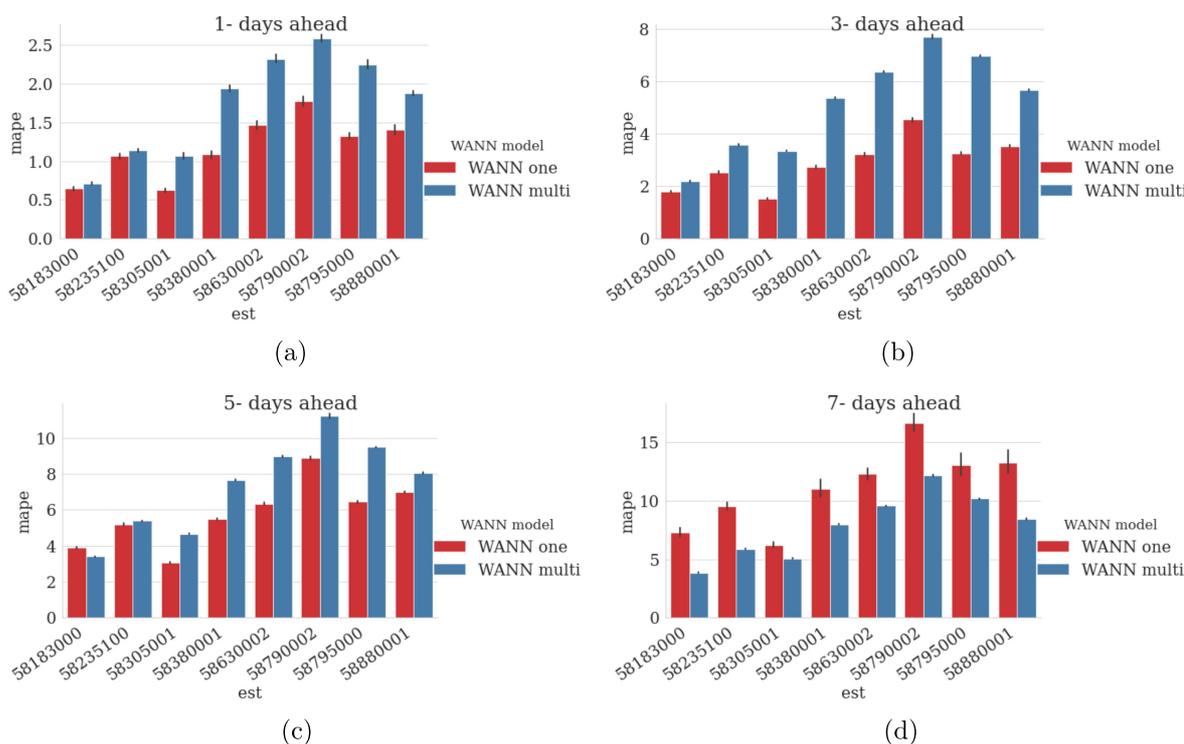
de modelos.

Conclui-se que o efeito do uso da transformada wavelets AT em combinação com RNAs foi aparente a partir da melhoria geral no resultado para todos os horizontes de previsão no caso de todas as estações modeladas. Isto pode ser devido à capacidade da análise wavelet de extrair informações úteis de tempo-frequência da série temporal de vazão original que não foi capturada pelos modelos ANN convencionais e, desta forma, simplificar a captura da não linearidade em conjuntos de dados de entrada-saída (99).

### 6.3 COMPARAÇÃO DOS MODELOS $WANN_{one}$ E $WANN_{multi}$

As Figuras 34, 35 e 36 representam a variação das medidas de desempenho no caso da modelagem por modelos WANNs em todos os horizontes de previsão para 8 estações fluviométricas.

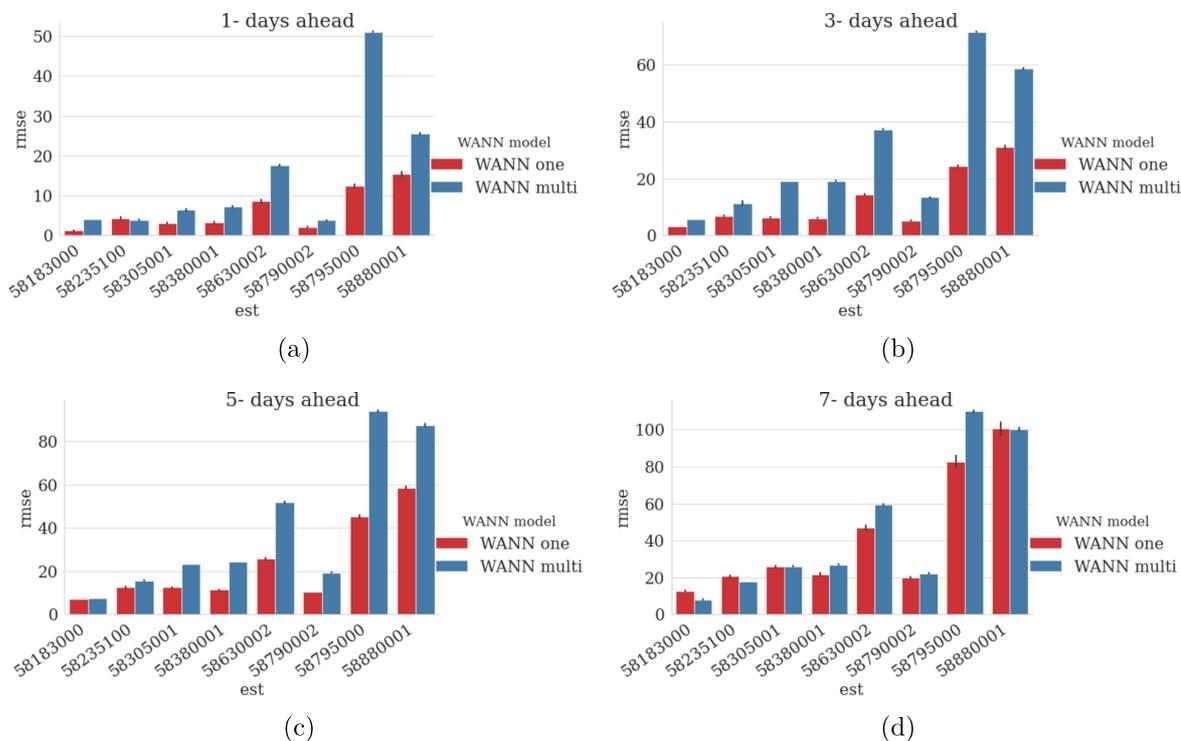
Figura 34 – Desempenho dos modelos  $WANN_{one}$  e  $WANN_{multi}$  em termos do MAPE médio em todos os horizontes de previsão: (a) 1 dia à frente; (b) 3 dias à frente; (c) 5 dias à frente; (d) 7 dias à frente.



Fonte: Elaborada pelo autor (2022).

Em geral, notou-se que os dois modelos WANN desenvolvidos obtiveram um desempenho muito satisfatório considerando a remoção do atraso de fase, bem como a obtenção dos picos. Pela Figura 37, observa-se que os modelos WANN seguem muito bem a tendência da vazão observada e são capazes de capturar o comportamento periódico e

Figura 35 – Desempenho dos modelos  $WANN_{one}$  e  $WANN_{multi}$  em termos do RMSE médio em todos os horizontes de previsão: (a) 1 dia à frente; (b) 3 dias à frente; (c) 5 dias à frente; (d) 7 dias à frente.



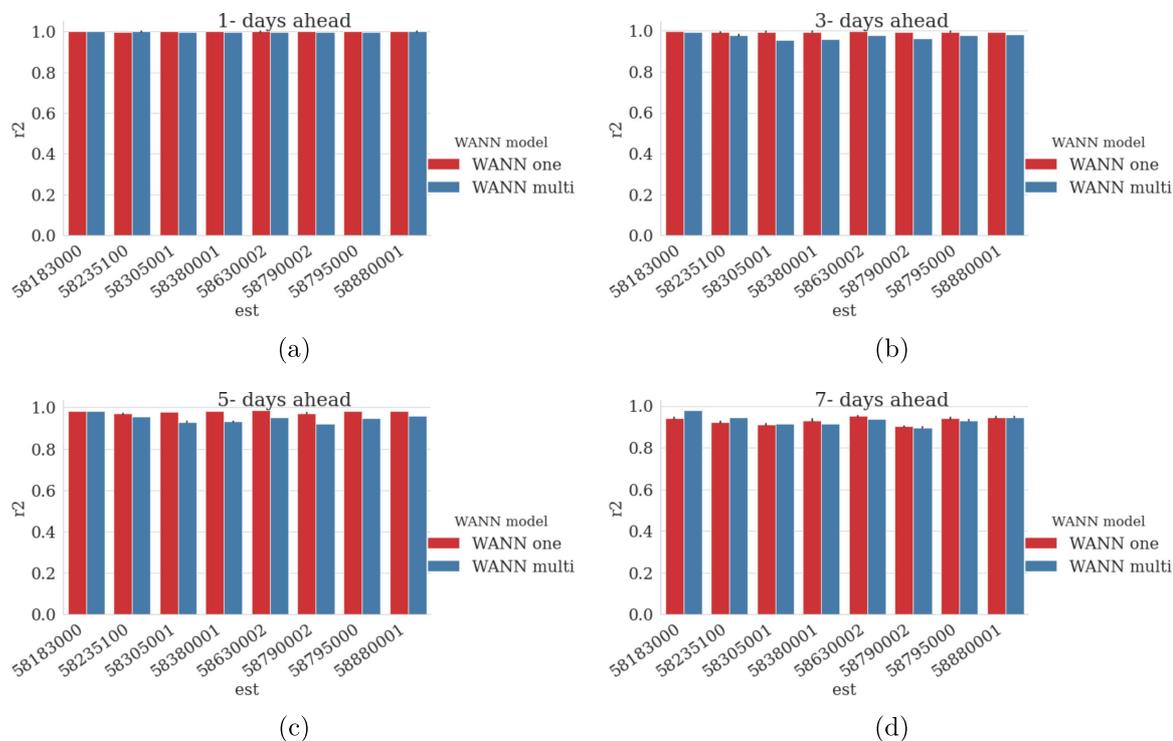
Fonte: Elaborada pelo autor (2022).

reproduzir bem os fluxos baixos e médios até para o horizonte de previsão de 7 dias à frente.

Comparando os desempenhos dos modelos baseados em wavelets observou-se que em termos das métricas MAPE, RMSE e  $R^2$ , em geral, para horizontes de previsão de 1, 3 e 5 dias à frente, os modelos  $WANN_{one}$  possuem a melhor capacidade de previsão. No entanto, à medida que o horizonte de previsão aumenta, a precisão da previsão de modelos  $WANN_{one}$  torna-se semelhante com a dos  $WANN_{multi}$  para todas as 8 estações. Observa-se que para o horizonte de previsão de 7 dias à frente para todas as estações, a métrica MAPE indicou que o modelo  $WANN_{multi}$  apresentou resultados ligeiramente melhores. Além disso, as métricas RMSE e  $R^2$  para as estações 58183000, 58235001, 58305001 e 58880001 também indicaram que os resultados para a previsão de 7 dias à frente do modelo  $WANN_{multi}$  é mais precisos.

Em relação à modelagem utilizando  $WANN_{multi}$  notou-se a tendência destes modelos em subestimar a magnitude dos picos médios, representando a previsão por uma curva suavizada (tende à valor médio entre os picos). No entanto, observou-se que o  $WANN_{multi}$ , geralmente, representa com mais precisão as vazões baixas em comparação do modelo  $WANN_{one}$ . Por outro lado, observando os resultados produzidos por modelo  $WANN_{one}$ , observa-se que este modelo tem a tendência de superestimar as vazões mais baixas, porém

Figura 36 – Desempenho dos modelos  $WANN_{one}$  e  $WANN_{multi}$  em termos do  $R^2$  médio em todos os horizontes de previsão: (a) 1 dia à frente; (b) 3 dias à frente; (c) 5 dias à frente; (d) 7 dias à frente.



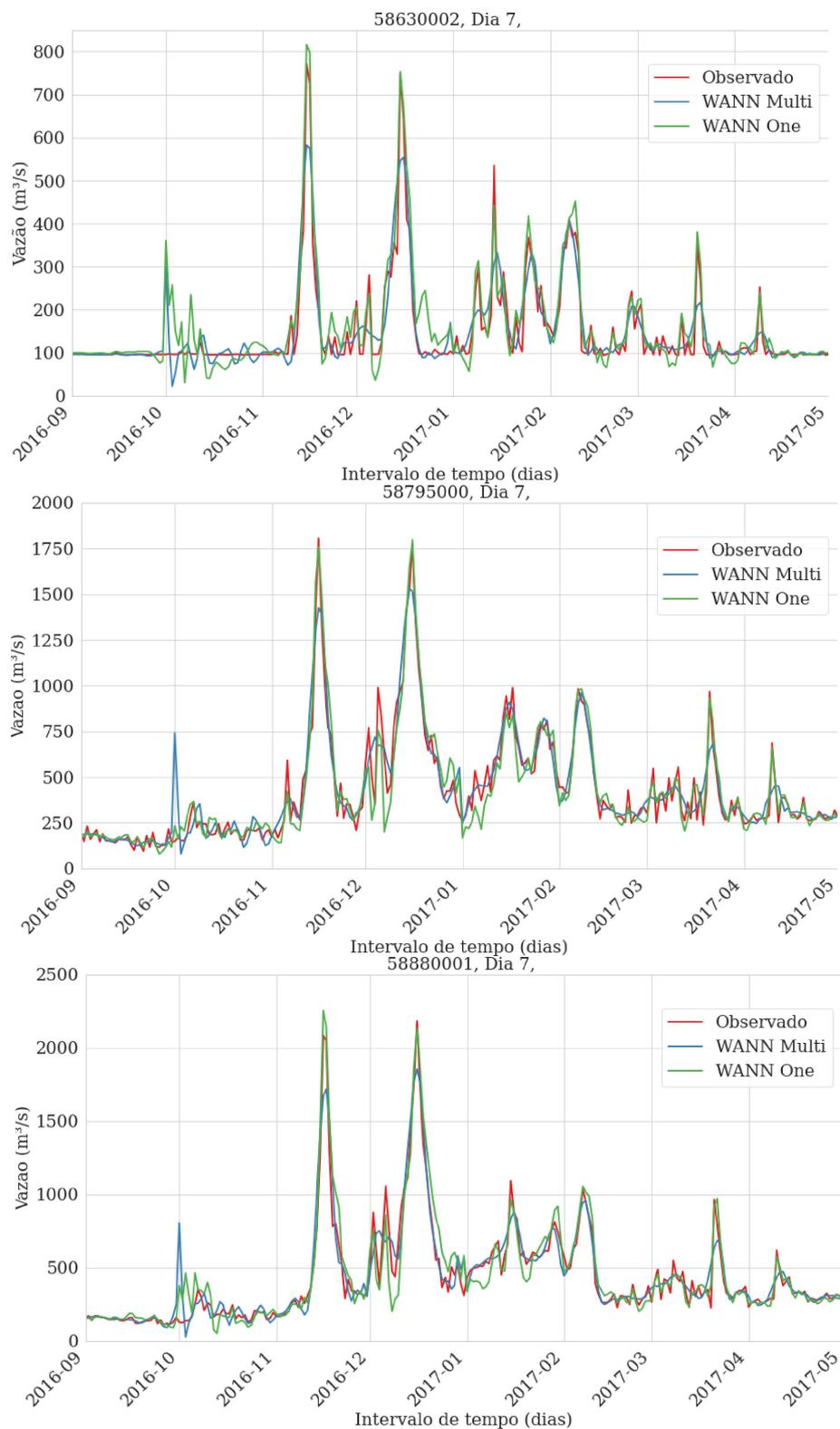
Fonte: Elaborada pelo autor (2022).

consegue acompanhar a magnitude dos picos altos de vazão melhor do que o modelos  $WANN_{multi}$  para todos os horizontes de previsão estudados.

### 6.3.1 Avaliação do desempenho dos modelos WANN para vazões extremas

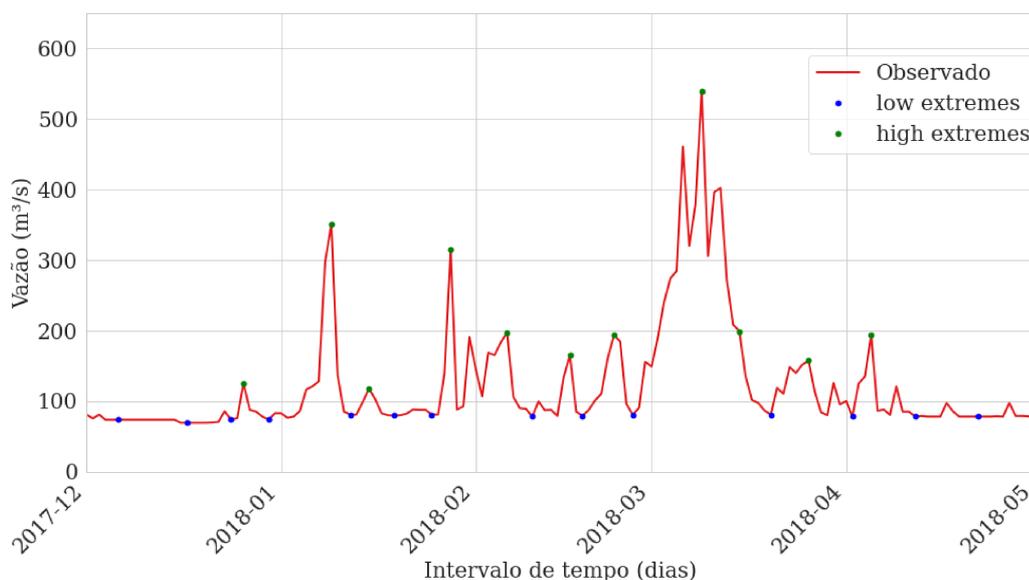
A próxima avaliação da eficácia dos modelos baseados em wavelets é comparar vazões extremas entre o valor previsto e o valor observado no conjunto de dados de teste. Para extrair valores extremos de séries temporais utilizou-se a biblioteca Pyxtremes (<https://georgebv.github.io/pyxtremes/>). Os valores extremos são extraídos de séries temporais gerando primeiramente uma série temporal de excedências selecionando valores acima (ou abaixo para valores extremos baixos) de um determinado limite e, em seguida, desagrupando a série temporal de excedência identificando clusters separados por um determinado período de tempo e, em seguida, selecionando apenas os valores mais altos (mais baixos) dentro de cada cluster. Neste estudo, a distância mínima entre clusters adjacentes foi definida como 5 dias (Veja Figura 38). O valor limite (*threshold value*, em inglês) foi definido como um valor médio (a soma de todos os valores de dados fornecidos dividido pelo número total de valores de dados fornecidos no conjunto) para cada série temporal estudada. Em seguida, a métrica MAPE (Equação 3.28) foi calculada para medir o desvio médio entre os valores previstos e os valores observados para os extremos alto e

Figura 37 – Os hidrogramas da vazão observada e prevista por modelos WANN para o horizonte de 7 dias à frente no caso da 3 estações: (a) 58630002; (b) 58795000 e (c) 58880001.



Fonte: Elaborada pelo autor (2022).

Figura 38 – Escolha dos extremos altos e baixos.



Fonte: Elaborada pelo autor (2022).

baixo. A precisão de previsão dos modelos wavelet  $WANN_{one}$  e  $WANN_{multi}$  para extremos altos e baixos para diferentes horizontes de previsão é apresentada na Tabela 12. A Figura 39 demonstra a comparação do desempenho dos modelos baseados em wavelets na previsão de valores extremos (altos e baixos) de vazão em termos da métrica MAPE.

De acordo com os resultados obtidos, comparando os desempenhos dos modelos baseados em wavelets, observa-se:

- Para previsões de 1 e 3 dias à frente os modelos  $WANN_{one}$  possuem a melhor capacidade de previsão de vazões extremas altas e baixas para todas as estações.
- Na previsão de 5 dias à frente os modelos  $WANN_{one}$  superam o desempenho do  $WANN_{multi}$  na previsão dos picos altos. Por outro lado, os picos baixos foram previstas com mais precisão por modelos  $WANN_{multi}$  para todas as estações.

Para avaliar o desempenho dos modelos baseados em wavelets para horizonte de previsão de 7 dias à frente empregou-se análise estatística.

### 6.3.2 Análise estatística

Com o objetivo de provar as conclusões feitas na sessão anterior e identificar se existe diferença estatisticamente significativa entre os desempenhos dos modelos  $WANN_{one}$  e  $WANN_{multi}$  e, também, determinar qual abordagem de previsão baseada em wavelets apresentou melhor desempenho na modelagem de vazão, empregaram-se testes estatísticos.

Os testes estatísticos de Shapiro-Wilk e o U de Mann-Whitney foram empregados nas métricas obtidas a partir de 30 execuções independentes para previsão de vazão de 7

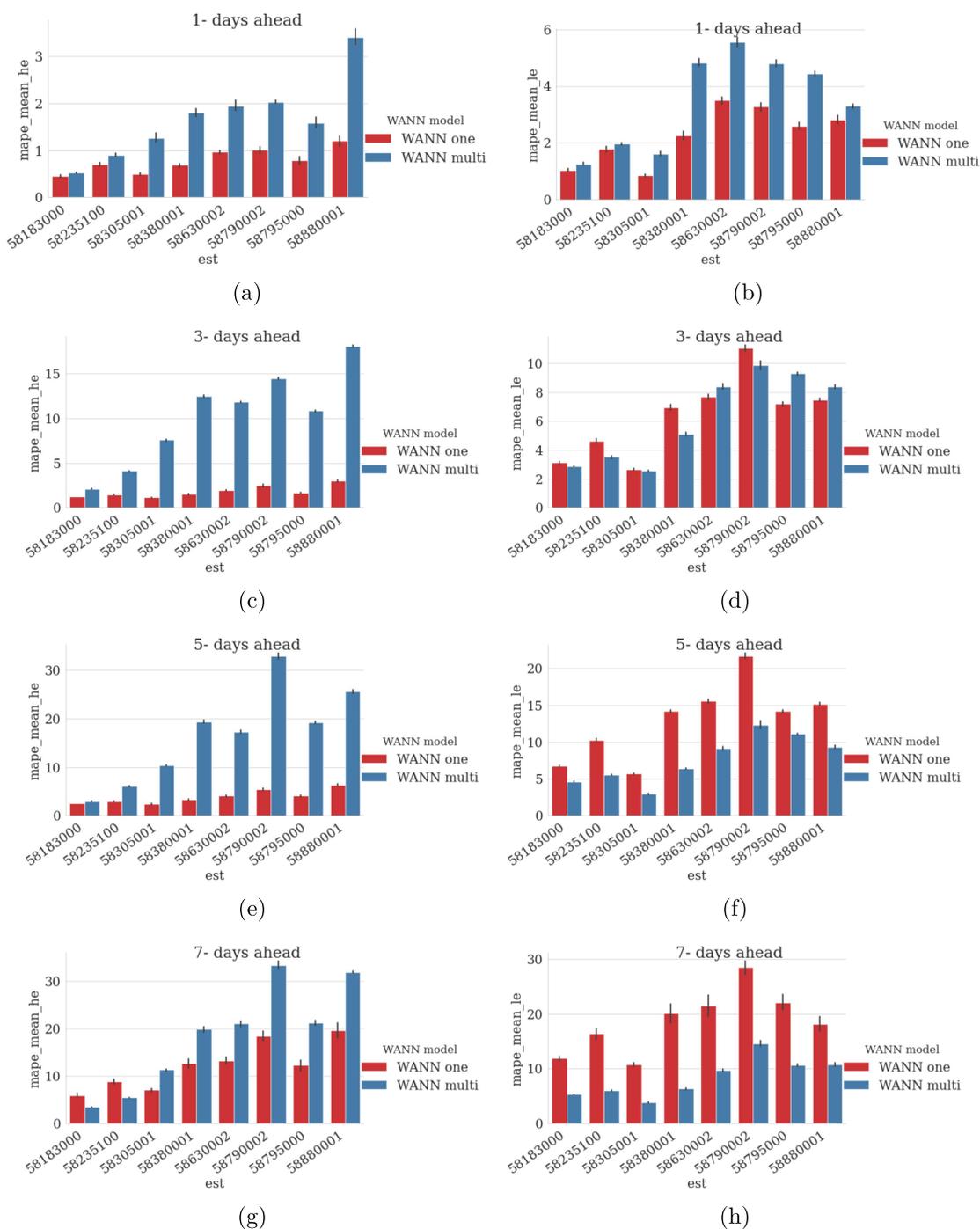
Tabela 12 – Desempenho dos modelos WANN em extremos altos e baixos para diferentes horizontes de previsão. Os valores médios da métrica MAPE para extremos altos e baixos após 30 execuções. Os melhores valores são indicados em negrito e os desvios padrão entre parênteses.

Estação	Hor	MAPE(%) Extremos altos		MAPE(%) Extremos baixos	
		WANN <sub>one</sub>	WANN <sub>multi</sub>	WANN <sub>one</sub>	WANN <sub>multi</sub>
58880001	1	<b>1.1996(0.3017)</b>	7.6107(3.8634)	<b>2.8151(0.3885)</b>	4.6858(2.5358)
	3	<b>3.0075(0.4426)</b>	20.6773(4.0318)	<b>7.4717(0.3789)</b>	9.5686(2.5744)
	5	<b>6.3393(0.765)</b>	28.3255(9.0055)	15.1722(0.5414)	<b>10.2601(6.7396)</b>
	7	<b>19.5624(4.4975)</b>	31.7751(1.362)	18.0834(3.7205)	<b>10.6513(0.9928)</b>
58790002	1	<b>0.6874(0.0975)</b>	4.176(2.0548)	<b>2.2517(0.3661)</b>	6.1373(1.1441)
	3	<b>1.5059(0.1594)</b>	14.9966(4.5617)	<b>6.9375(0.5835)</b>	8.6495(5.5912)
	5	<b>3.3713(0.3804)</b>	20.8707(4.9159)	14.2064(0.5369)	<b>6.9602(1.9627)</b>
	7	<b>12.7413(2.8114)</b>	20.5169(1.4831)	20.0361(4.9061)	<b>6.4808(1.0038)</b>
58795000	1	<b>0.7884(0.2274)</b>	3.6652(2.0132)	<b>2.5948(0.3293)</b>	6.327(1.4306)
	3	<b>1.6563(0.2257)</b>	12.6479(3.3614)	<b>7.1735(0.4611)</b>	11.3284(2.6709)
	5	<b>4.0477(0.456)</b>	19.7752(4.7095)	<b>14.1642(0.5335)</b>	11.5468(4.4176)
	7	<b>12.2249(3.1606)</b>	21.9296(1.0011)	21.9667(3.8889)	<b>10.8054(0.7096)</b>
58235100	1	<b>0.7005(0.1254)</b>	2.1969(6.0245)	<b>1.7817(0.2722)</b>	3.3665(3.9727)
	3	<b>1.4778(0.1887)</b>	4.6126(0.8201)	<b>4.6101(0.4195)</b>	5.1484(1.7895)
	5	<b>2.9719(0.2207)</b>	6.75(3.3842)	10.2765(0.6271)	<b>6.108(3.5488)</b>
	7	8.7912(1.4565)	<b>5.7474(0.5809)</b>	16.4111(2.6651)	<b>6.0491(0.4903)</b>
58183000	1	<b>0.4545(0.0684)</b>	0.9945(0.4551)	<b>1.0429(0.1619)</b>	2.0533(0.6074)
	3	<b>1.2263(0.0622)</b>	2.6199(0.7908)	<b>3.1203(0.2399)</b>	3.888(1.4063)
	5	<b>2.5241(0.0926)</b>	3.8916(2.4629)	6.7365(0.3399)	<b>5.4047(2.7212)</b>
	7	5.9067(1.3901)	<b>3.3327(0.1241)</b>	11.9005(1.1777)	<b>5.1046(0.2319)</b>
58630002	1	<b>0.9635(0.1092)</b>	4.3882(2.2284)	<b>3.4915(0.349)</b>	7.3925(1.6851)
	3	<b>1.9337(0.1629)</b>	13.1971(2.0433)	<b>7.6641(0.5368)</b>	11.5076(3.503)
	5	<b>4.0702(0.555)</b>	18.4868(7.3369)	15.5843(0.7152)	<b>10.1796(7.8014)</b>
	7	<b>13.2275(2.1984)</b>	21.3554(2.9281)	21.4323(5.7131)	<b>10.2422(1.6894)</b>
58380001	1	<b>0.6874(0.0975)</b>	4.176(2.0548)	<b>2.2517(0.3661)</b>	6.1373(1.1441)
	3	<b>1.5059(0.1594)</b>	14.9966(4.5617)	<b>6.9375(0.5835)</b>	8.6495(5.5912)
	5	<b>3.3713(0.3804)</b>	20.8707(4.9159)	14.2064(0.5369)	<b>6.9602(1.9627)</b>
	7	<b>12.7413(2.8114)</b>	20.5169(1.4831)	20.0361(4.9061)	<b>6.4808(1.0038)</b>
58305001	1	<b>0.4987(0.0761)</b>	2.5551(1.2)	<b>0.8483(0.1357)</b>	2.2089(0.5893)
	3	<b>1.1596(0.0933)</b>	8.1283(1.3296)	<b>2.6583(0.1693)</b>	3.9386(1.3591)
	5	<b>2.3958(0.2262)</b>	10.5193(2.8626)	5.6753(0.2819)	<b>3.6997(2.7332)</b>
	7	<b>7.0724(0.8559)</b>	11.3319(0.7426)	10.7522(0.9288)	<b>3.6762(0.3703)</b>

Fonte: Elaborada pelo autor (2022).

dias à frente. Para executar os testes estatísticos utilizou-se a biblioteca *Scipy* em Python. A Tabela 13 apresenta os p-valores calculados pelo método de Shapiro-Wilk, onde os itens em negrito indicam a existência da normalidade. Para o uso de teste paramétrico em dada métrica, é necessário que os métodos testados apresentem distribuição normal. Assim, conforme a Tabela 13 é possível verificar que é necessário aplicar o teste não paramétrico, pois quase nenhuma das métricas apresentou distribuição normal nos dois modelos de forma concomitante. Portanto, o teste U de Mann-Whitney é apropriado para analisar duas amostras de métricas.

Figura 39 – O desempenho dos modelos WANN em termos do MAPE médio na previsão de valores extremos (altos e baixos) para todos os horizontes de previsão: (a) Extremos altos, previsão de 1 dia à frente;(b) Extremos baixos, previsão de 1 dia à frente;(c) Extremos altos, previsão de 3 dias à frente; (d) Extremos baixos, previsão de 3 dias à frente;(e) Extremos altos, previsão de 5 dias à frente;(f) Extremos baixos, previsão de 5 dias à frente;(g) Extremos altos, previsão de 7 dias à frente; e(h) Extremos baixos, previsão de 7 dias à frente.



Fonte: Elaborada pelo autor (2022).

Tabela 13 – Resumo dos resultados do p-valor do teste Shapiro sobre normalidade das distribuições dos resultados das métricas avaliadas. Os itens em negrito indicam a existência da normalidade.

Estação	MAPE		RMSE		R <sup>2</sup>	
	WANN <sub>one</sub>	WANN <sub>multi</sub>	WANN <sub>one</sub>	WANN <sub>multi</sub>	WANN <sub>one</sub>	WANN <sub>multi</sub>
58380001	0.000	0.001	0.012	0.001	0.019	0.001
58235100	0.000	0.000	0.000	0.000	0.000	0.000
58305001	0.000	<b>0.189</b>	0.000	0.019	0.000	0.016
58630002	<b>0.116</b>	<b>0.083</b>	0.001	0.002	0.003	0.002
58183000	0.000	0.172	0.000	0.094	0.000	0.069
58790002	0.000	0.429	<b>0.394</b>	<b>0.396</b>	<b>0.239</b>	<b>0.401</b>
58880001	0.000	0.001	0.038	<b>0.236</b>	0.005	<b>0.205</b>
58795000	0.000	<b>0.093</b>	0.006	<b>0.676</b>	0.001	<b>0.925</b>

Fonte: Elaborada pelo autor (2022).

Para determinar se há diferença estatisticamente significativa entre os desempenhos dos modelos WANN<sub>one</sub> e WANN<sub>multi</sub>, empregou-se o teste U de Mann-Whitney *two-sided*. Foram comparadas as seguintes métricas: MAPE, RMSE, R<sup>2</sup>, MAPE<sub>he</sub> (o MAPE que avalia extremos altos) e MAPE<sub>le</sub> (o MAPE que avalia extremos baixos). Os resultados obtidos são os seguintes:

- Como o p-valor obtido pelo teste U de Mann-Whitney é significativo ( $p < 0,05$ ) para as métrica MAPE, MAPE<sub>he</sub> e MAPE<sub>le</sub>, conclui-se que em termos de MAPE, MAPE<sub>he</sub> e MAPE<sub>le</sub> os desempenhos de modelos WANN<sub>one</sub> e WANN<sub>multi</sub> são significativamente diferentes um do outro para todas as estações modeladas.
- Para a métrica RMSE obteve-se o p-valor significativo ( $p < 0,05$ ) para todas as estações além das 58880001 e 58305001. Neste caso, conclui-se que em termos do RMSE, os desempenhos de modelos WANN<sub>one</sub> e WANN<sub>multi</sub> são significativamente diferentes um do outro para todas as estações exceto para as 58880001 e 58305001.
- Para a métrica R<sup>2</sup> obteve-se o p-valor significativo ( $p < 0,05$ ) para todas as estações além da 58880001. Neste caso, conclui-se que em termos do R<sup>2</sup> os desempenhos de modelos WANN<sub>one</sub> e WANN<sub>multi</sub> são significativamente diferentes um do outro para todas as estações além para a 58880001.

Conclui-se que para as estações 58880001 e 58305001 os modelos apresentaram desempenho semelhante porém, para as demais estações, o desempenho demonstrado é significativamente diferente.

Para determinar qual modelo possui desempenho significativamente melhor em termos das métricas MAPE, RMSE, R<sup>2</sup>, MAPE<sub>he</sub> e MAPE<sub>le</sub> empregou-se o teste U de Mann-Whitney *one-sided*. Os resultados obtidos estão apresentados na Tabela 14.

Tabela 14 – Resumo dos resultados do p-valor do teste U de Mann-Whitney *one – sided* das métricas avaliadas. Os itens em negrito indicam a existência da superioridade no desempenho do modelo  $WANN_{one}$  para a previsão de 7 dias à frente.

Estação	MAPE	RMSE	$R^2$	$MAPE_{he}$	$MAPE_{le}$
58880001	0.000	<b>0.898</b>	<b>0.217</b>	<b>0.999</b>	0.000
58790002	0.000	<b>0.999</b>	0.000	<b>0.999</b>	0.000
58795000	0.000	<b>0.999</b>	0.000	<b>0.999</b>	0.000
58235100	0.000	0.000	<b>0.999</b>	0.000	0.000
58183000	0.000	0.000	<b>0.999</b>	0.000	0.000
58630002	0.000	<b>0.999</b>	0.000	<b>0.999</b>	0.000
58380001	0.000	<b>0.999</b>	0.000	<b>0.999</b>	0.000
58305001	0.000	0.011	<b>0.999</b>	<b>0.999</b>	0.000

Fonte: Elaborada pelo autor (2022).

São feitas as seguintes conclusões:

- Como o p-valor obtido pelo teste U de Mann-Whitney é significativo ( $p < 0.05$ ) para as métricas MAPE e  $MAPE_{le}$ , conclui-se que o modelo  $WANN_{one}$  em termos destas métricas apresentou um desempenho inferior (as métricas MAPE e  $MAPE_{le}$  do modelo  $WANN_{one}$  são significativamente maiores em comparação das do modelo  $WANN_{multi}$ ) para todas as estações modeladas.
- Para a métrica RMSE, obteve-se o significativo p-valor ( $p < 0.05$ ) somente para as estações 58235100, 58183000 e 58305001. Desta forma, conclui-se que, para estas estações, o modelo  $WANN_{one}$ , em termos do RMSE, apresentou desempenho inferior em comparação do modelo  $WANN_{multi}$ , apresentando os valores do RMSE significativamente maiores do que o modelo  $WANN_{multi}$ .
- No caso das métricas  $R^2$ , o teste indicou que o modelo  $WANN_{one}$  demonstrou o pior desempenho em termos do  $R^2$  para as estações 58880001, 58235100, 58183000 e 58305001.
- No caso da métrica  $MAPE_{he}$ , o teste indicou que o modelo  $WANN_{one}$  superou o modelo  $WANN_{multi}$  na modelagem de todas as estações além das 58235100 e 58183000.

Levando-se em consideração os resultados obtidos através da análise estatística, conclui-se que para a previsão de vazão de 7 dias à frente:

- Em geral, os modelos  $WANN_{one}$  e  $WANN_{multi}$  demonstraram os desempenhos diferentes e o ponto da diferença é a modelagem dos vazões extremas.
- $WANN_{multi}$  fornecem previsões mais precisas na modelagem das vazões extremas baixas para todas as estações. Além disso, as vazões extremas altas das esta-

ções 58183000 e 58235100 também são previstas com mais precisão por modelos  $WANN_{multi}$ .

- Em contraste, os modelos  $WANN_{one}$  fornecem a previsão mais precisa na modelagem das vazões extremas altas para as demais estações modeladas.

#### 6.4 TENDÊNCIAS EM COMUM NO COMPORTAMENTO DA MODELAGEM DAS 8 ESTAÇÕES MODELADAS

Uma possível explicação dos maiores erros relacionados às vazões extremas é fornecida: as séries históricas de vazão das estações estudadas contém as informações de valores médios e baixos em muito maior quantidade do que as informações de valores máximos. De fato, o conjunto de treinamento (estação 58880001) contém informação de vazão para o total de 4977 dias, dentre os quais só existe a informação de 152 dias de vazão acima de  $1000 \text{ m}^3/\text{s}$  (que acontecem nos períodos de chuva) o que pode ser não suficiente para o treinamento adequado do modelo para gerar os valores extremos altos. Nesse sentido, Sivakumar, Jayawardena e Fernando(15) supõe que a presença de valores muito altos no conjunto de treinamento podem levar a valores mais altos (superestimação) na previsão, ao passo em que valores muito baixos podem gerar previsões com valores mais baixos (subestimação).

Outra tendência em comum é a ocorrência dos maiores erros nos horizontes de previsão mais distantes. Possível causa para isto é a falta da informação mais recente na estrutura do modelo. Por exemplo, para a previsão no tempo  $t + 5$ , a informação mais recente é no tempo  $t$ , enquanto as informações dos tempos  $t + 4$ ,  $t + 3$ ,  $t + 2$  e  $t + 1$  estão faltantes. Como o fenômeno de vazões altas tem característica de curta duração (devido, por exemplo, à chuva repentina ou abertura de uma barragem à montante do rio), os últimos escoamentos não fornecem informações sobre o tempo correto de subida e descida do hidrograma, como no caso da previsão no tempo  $t + 1$ . Por isso, possivelmente os modelos não tenham conseguido aprender adequadamente as relações na parte de vazões altas demonstrando subestimações na previsão dos picos altos.

Os modelos híbridos baseados em wavelets  $WANN$  e convencionais  $ANN$  demonstraram seus diferentes desempenhos para a previsão de vazões de curto prazo em diferentes estações da Bacia do Rio Paraíba do Sul. Entre todas as 8 estações consideradas, a precisão do modelos de previsão para as estações localizadas na região Paulista (a saber, 58305001, 58235100 e 58183000) foi considerada a melhor, o que pode estar relacionado com a baixa variação e magnitude nas suas vazões e baixas índices de complexidades das séries históricas.

Os modelos para as séries temporais de vazões da região Fluminense, principalmente na sub-bacia Baixo Paraíba do Sul, apresentaram os desempenhos mais baixos, o que pode

estar relacionado à alta magnitude de vazões, maiores áreas de drenagem e maiores índices de complexidade das séries históricas. Além disso, através inspeção visual e uma análise estatística notou-se que a vazão do rio Paraíba do Sul nessa região possui altas flutuações de seus valores. Por exemplo, a estação 58795000 (região Fluminense) possui a área de drenagem igual 45300 km<sup>2</sup> e a vazão nela varia entre 85 m<sup>3</sup>/s e 4621 m<sup>3</sup>/s, enquanto a vazão na estação 58235100 (região Paulista) varia entre 50 m<sup>3</sup>/s e 1300 m<sup>3</sup>/s e essa estação possui a área de drenagem igual 12800 km<sup>2</sup>.

Na modelagem da estação 58790002 observa-se erros significativos e os maiores apresentados por todos os modelos, o que podem ser relacionado ao fato que a bacia hidrográfica onde está localizada a estação 58790002 possui a menor área de drenagem entre as estações modeladas (a saber, 8210 km<sup>2</sup>). Isto significa que quase não existe armazenamento de água de chuva nesta bacia, ou seja, a vazão do rio neste local possui variação muito grande. De fato, a vazão mínima na estação 58790002 tem registro de 8.34 m<sup>3</sup>/s e vazão máxima de 1196 m<sup>3</sup>/s.

Sujeita a condições geológicas, hidrológicas e ambientais, vazões podem possuir dinâmicas que variam entre relativamente simples até amplamente complexas (100). Além disso, aspectos físicos como área restrita, alta variação na elevação e alto declive afetam negativamente o desempenho do modelo. Dessa forma, conclui-se que o desempenho do modelo está relacionado direto às características hidromorfológicas locais que explica a melhoria ou deterioração da precisão das previsões (101).

## 7 CONCLUSÃO

O presente estudo foi realizado para a previsão das vazões médias diárias em 8 estações fluviométricas da Bacia do rio Paraíba do Sul. Dada a importância estratégica da região compreendida por esta bacia, este estudo fornece contribuições científicas para a área de modelagem computacional aplicada a recursos hídricos. É proposta uma abordagem eficaz para a previsão de vazão em curto prazo desenvolvendo modelos híbridos e acoplando redes neurais artificiais à uma técnica de pré-processamento de dados baseada em transformada wavelet *À Trous*.

Os modelos desenvolvidos foram aplicados para previsão de vazão média diária natural diária de 1, 3, 5 e 7 dias à frente. Na etapa de pré-processamento de dados é investigada a influência de diferentes famílias de wavelet na precisão da previsão e escolhida a wavelet mãe apropriada para a modelagem de vazão em curto prazo. Os melhores modelos híbridos WANN foram compostos pela wavelet mãe Coiflets 4 que se demonstrou um método adequado para representar as séries temporais de vazão consideradas. Os melhores resultados foram obtidos usando as aproximações e detalhes em 3 níveis de decomposição. Os modelos foram avaliados aplicando critérios como MAPE, RMSE e  $R^2$ , usando os conjuntos de dados independentes. A robustez dos modelos é avaliada e o papel de influência da transformada wavelet nos modelos híbridos é discutido.

O desempenho dos modelos baseados em wavelets e redes neurais artificiais foi comparado com o dos modelos convencionais de redes neurais inicialmente desenvolvidos para previsão de vazão. Modelos ANN falharam em atingir precisão em horizontes mais distantes e atingir vazões extremas. Além disso, a presença da previsão defasada demonstrou a inconsistência desses modelos.

Os resultados obtidos revelaram que o desempenho dos modelos WANN são significativamente superiores ao dos modelos ANN em termos de medidas de eficiência da previsão em todos os horizontes de previsão testados. Os modelos WANN se destacaram na capacidade de acompanhar a tendência dos dados de vazão observados e apresentaram desempenhos excelentes em termos de critérios de avaliação utilizados. Vale ressaltar que os eventos extremos em todas as 8 estações foram previstos com boa precisão para horizonte temporal de 7 dias à frente, juntamente com 1, 3 e 5 dias à frente. A análise estatística indicou que em termos de aplicabilidade a determinados cenários, o modelo  $WANN_{multi}$  conseguiu representar melhor os ciclos de vazões baixas ao passo em que  $WANN_{one}$  conseguiu acompanhar melhor magnitude das vazões extremas altas.

A metodologia apresentada neste estudo visa obter uma previsão de vazão precisa, configurando-a como uma ferramenta de grande importância para reduzir os riscos de inundações em áreas de desastres naturais frequentes e gestão de recursos hídricos. Os resultados deste estudo podem ser úteis para engenharia hidrológica e para tomadores de

decisão em previsão de vazão de rios, impacto econômico sobre indústrias em períodos de seca, reação a inundações em cidades, possibilidade de colapso de obras hidráulicas e planejamento de disponibilidade hídrica em áreas urbanas, na irrigação, navegação fluvial e na distribuição sustentável de água e gestão de reservatórios de energia. Os resultados deste estudo sugerem o potencial de aplicação de modelo WANN no campo da hidrologia na construção e gestão de um sistema de alerta de cheias em tempo real. Este modelo é uma opção viável para grandes bacias hidrográficas como a bacia do rio Paraíba do Sul.

## 7.1 PERSPECTIVAS DE TRABALHOS FUTUROS

- Explorar o potencial de modelos híbridos wavelet-redes neurais artificiais para horizontes de previsão mais distantes, como 10, 14 e 21 dias à frente.
- Acoplar outros algoritmos de aprendizado de máquina com a transformada wavelet;
- Os resultados apresentados no trabalho (33) indicaram que a adição de mais informações de vazão da seção à montante no conjunto de dados de entrada contribui para a redução do erro de previsão para horizontes de previsão acima de 2 dias à frente. Dessa forma, sugere-se explorar o uso das vazões à montante como variáveis de entrada para modelos WANN.
- Aplicar os modelos híbridos desenvolvidos em outros pontos da bacia do rio Paraíba do Sul.
- Explorar a complexidade das séries históricas de vazão e relacionamento com a previsibilidade.
- Criar uma metodologia que agrupe estações fluviométricas da bacia do rio Paraíba do Sul conforme a complexidade de suas séries históricas e características físicas da bacia (tais como a elevação, inclinação e área de drenagem). Determinar quais classes de modelos são apropriadas para os grupos detectados e aplicá-las na modelagem da vazão diária.

## REFERÊNCIAS

- 1 IORIS, A. A. Os limites políticos de uma reforma incompleta: a implementação da lei dos recursos hídricos na bacia do paraíba do sul. *Revista Brasileira de Estudos Urbanos e Regionais*, v. 10, n. 1, p. 61, 2008.
- 2 Agência Nacional de Águas. *Dados de Chuva e Vazão da Bacia do Rio Paraíba do Sul*. [S.l.], 2016.
- 3 BRASIL. Lei nº 9.433, de 8 de janeiro de 1997. institui a política nacional de recursos hídricos, cria o sistema nacional de gerenciamento de recursos hídricos, regulamenta o inciso xix do art. 21 da constituição federal, e altera o art. 1º da lei nº 8.001, de 13 de março de 1990, que modificou a lei nº 7.990, de 28 de dezembro de 1989. *Diário Oficial da União*, 1997.
- 4 ANDRADE, V. S.; RIBEIRO, C. B.; REIS, G. R. Desenvolvimento de sistema georreferenciado de informações hidrológicas para gestão dos recursos hídricos nas bacias hidrográficas da zona da mata mineira, tado de são paulo. *XI Simpósio de Recursos Hídricos do Nordeste*, SciELO Brasil, 2012.
- 5 MECCHI, A.; SANCHES, D. L. Impactos ambientais da mineração no estado de são paulo. *Estudos avançados*, SciELO Brasil, v. 24, n. 68, p. 209–220, 2010.
- 6 Agência Nacional de Águas. Bacia do paraíba do sul. *Boletim Diário de Monitoramento da Baía do Rio Paraíba do Sul 23 de fevereiro de 2018*, SciELO Brasil, p. 1–9, 2018.
- 7 TUCCI, C. E.; COLLISCHONN, W. Previsão de vazão em curto prazo com base na precipitação: Modelos conceituais. In: *Workshop Previsão de Vazão*. [S.l.: s.n.], 2007.
- 8 TUCCI, C. E. et al. Hidrologia: ciência e aplicação. *São Paulo: Editora da*, 2012.
- 9 CASAGRANDE, L. et al. Early flood warning in the itajaí-açu river basin using numerical weather forecasting and hydrological modeling. *Natural Hazards*, Springer, v. 88, n. 2, p. 741–757, 2017.
- 10 LE, X.-H. et al. Application of long short-term memory (lstm) neural network for flood forecasting. *Water*, MDPI, v. 11, n. 7, p. 1387, 2019.
- 11 DRISYA, J.; KUMAR, D. S.; ROSHNI, T. Hydrological drought assessment through streamflow forecasting using wavelet enabled artificial neural networks. *Environment, Development and Sustainability*, Springer, v. 23, n. 3, p. 3653–3672, 2021.
- 12 ASADI, S. et al. A new hybrid artificial neural networks for rainfall–runoff process modeling. *Neurocomputing*, Elsevier, v. 121, p. 470–480, 2013.
- 13 SOLOMATINE, D. P.; DULAL, K. N. Model trees as an alternative to neural networks in rainfall—runoff modelling. *Hydrological Sciences Journal*, Taylor & Francis, v. 48, n. 3, p. 399–411, 2003.
- 14 SUDHEER, K.; GOSAIN, A.; RAMASASTRI, K. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological processes*, Wiley Online Library, v. 16, n. 6, p. 1325–1330, 2002.

- 15 SIVAKUMAR, B.; JAYAWARDENA, A.; FERNANDO, T. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. *Journal of hydrology*, Elsevier, v. 265, n. 1-4, p. 225–245, 2002.
- 16 WHITE, E. *Predicting Unimpaired Flow in Ungauged Basins: "random Forests" Applied to California Streams*. [S.l.]: University of California, Davis, 2017.
- 17 SOLOMATINE, D. P.; OSTFELD, A. Data-driven modelling: some past experiences and new approaches. *Journal of hydroinformatics*, IWA Publishing, v. 10, n. 1, p. 3–22, 2008.
- 18 SHORTRIDGE, J. E.; GUIKEMA, S. D.; ZAITCHIK, B. F. Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. *Hydrology and Earth System Sciences*, Copernicus GmbH, v. 20, n. 7, p. 2611, 2016.
- 19 SUN, Y.; NIU, J.; SIVAKUMAR, B. A comparative study of models for short-term streamflow forecasting with emphasis on wavelet-based approach. *Stochastic Environmental Research and Risk Assessment*, Springer, v. 33, n. 10, p. 1875–1891, 2019.
- 20 PATEL, S. S.; RAMACHANDRAN, P. A comparison of machine learning techniques for modeling river flow time series: the case of upper cauvery river basin. *Water resources management*, Springer, v. 29, n. 2, p. 589–602, 2015.
- 21 NOURANI, V. et al. Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. *Journal of Hydrology*, Elsevier, v. 514, p. 358–377, 2014.
- 22 YASEEN, Z. M. et al. Application of the hybrid artificial neural network coupled with rolling mechanism and grey model algorithms for streamflow forecasting over multiple time horizons. *Water Resources Management*, Springer, v. 32, n. 5, p. 1883–1899, 2018.
- 23 WU, C.; CHAU, K.; LI, Y. Methods to improve neural network performance in daily flows prediction. *Journal of Hydrology*, Elsevier, v. 372, n. 1-4, p. 80–93, 2009.
- 24 MOUATADID, S. et al. Coupling the maximum overlap discrete wavelet transform and long short-term memory networks for irrigation flow forecasting. *Agricultural Water Management*, Elsevier, v. 219, p. 72–85, 2019.
- 25 ADAMOWSKI, J.; SUN, K. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *Journal of Hydrology*, Elsevier, v. 390, n. 1-2, p. 85–91, 2010.
- 26 LONDHE, S. N.; NARKHEDE, S. Forecasting stream flow using hybrid neuro-wavelet technique. *ISH Journal of Hydraulic Engineering*, Taylor & Francis, v. 24, n. 3, p. 275–284, 2018.
- 27 FAHIMI, F.; YASEEN, Z. M.; EL-SHAFIE, A. Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review. *Theoretical and applied climatology*, Springer, v. 128, n. 3-4, p. 875–903, 2017.
- 28 SANG, Y.-F. A review on the applications of wavelet transform in hydrology time series analysis. *Atmospheric research*, Elsevier, v. 122, p. 8–15, 2013.

- 29 GROSSMANN, A.; MORLET, J. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM journal on mathematical analysis*, SIAM, v. 15, n. 4, p. 723–736, 1984.
- 30 MAHESWARAN, R.; KHOSA, R. Comparative study of different wavelets for hydrologic forecasting. *Computers & Geosciences*, Elsevier, v. 46, p. 284–295, 2012.
- 31 QUILTY, J.; ADAMOWSKI, J. Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework. *Journal of hydrology*, Elsevier, v. 563, p. 336–353, 2018.
- 32 DU, K.; ZHAO, Y.; LEI, J. The incorrect usage of singular spectral analysis and discrete wavelet transform in hybrid models to predict hydrological time series. *Journal of Hydrology*, Elsevier, v. 552, p. 44–51, 2017.
- 33 GORODETSKAYA, Y. et al. Daily streamflow forecasting for paraíba do sul river using machine learning methods with hydrologic inputs. In: SBC. *Anais do XV Encontro Nacional de Inteligência Artificial e Computacional*. [S.l.], 2018. p. 162–173.
- 34 FONSECA, T. L. et al. A gradient boosting model optimized by a genetic algorithm for short-term riverflow forecast. *Revista Mundi Engenharia, Tecnologia e Gestão (ISSN: 2525-4782)*, v. 4, n. 3, 2019.
- 35 CAMPOS, L. C. D. et al. Short-term streamflow forecasting for paraíba do sul river using deep learning. In: SPRINGER. *EPIA Conference on Artificial Intelligence*. [S.l.], 2019. p. 507–518.
- 36 TIWARI, M. K.; CHATTERJEE, C. Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ann (wbann) hybrid approach. *Journal of Hydrology*, Elsevier, v. 394, n. 3-4, p. 458–470, 2010.
- 37 PRAMANIK, N.; PANDA, R. K.; SINGH, A. Daily river flow forecasting using wavelet ann hybrid models. *Journal of Hydroinformatics*, IWA Publishing, v. 13, n. 1, p. 49–63, 2011.
- 38 SANTOS, C. A. G.; SILVA, G. B. L. d. Daily streamflow forecasting using a wavelet transform and artificial neural network hybrid models. *Hydrological Sciences Journal*, Taylor & Francis, v. 59, n. 2, p. 312–324, 2014.
- 39 MASLOVA, I.; TICLAVILCA, A.; MCKEE, M. Adjusting wavelet-based multiresolution analysis boundary conditions for long-term streamflow forecasting. *Hydrological Processes*, Wiley Online Library, v. 30, n. 1, p. 57–74, 2016.
- 40 SHAFAEI, M.; KISI, O. Predicting river daily flow using wavelet-artificial neural networks based on regression analyses in comparison with artificial neural networks and support vector machine models. *Neural Computing and Applications*, Springer, v. 28, n. 1, p. 15–28, 2017.
- 41 ALEXANDER, A. A.; THAMPI, S. G.; NR, C. Development of hybrid wavelet-ann model for hourly flood stage forecasting. *ISH Journal of Hydraulic Engineering*, Taylor & Francis, v. 24, n. 2, p. 266–274, 2018.

- 42 SANTOS, C. A. et al. Hybrid wavelet neural network approach for daily inflow forecasting using tropical rainfall measuring mission data. *Journal of Hydrologic Engineering*, American Society of Civil Engineers, v. 24, n. 2, p. 04018062, 2019.
- 43 FREIRE, P. K. d. M. M.; SANTOS, C. A. G.; SILVA, G. B. L. da. Analysis of the use of discrete wavelet transforms coupled with ann for short-term streamflow forecasting. *Applied Soft Computing*, Elsevier, v. 80, p. 494–505, 2019.
- 44 DALKILIÇ, H. Y.; HASHIMI, S. A. Prediction of daily streamflow using artificial neural networks (anns), wavelet neural networks (wnns), and adaptive neuro-fuzzy inference system (anfis) models. *Water Supply*, IWA Publishing, v. 20, n. 4, p. 1396–1408, 2020.
- 45 JAIN, D. et al. Time series modeling of river flow using wavelet neural networks. In: *Time Series Modeling of River Flow Using Wavelet Neural Networks (December 21, 2020). e-journal-First Pan IIT International Management Conference–2018*. [S.l.: s.n.], 2020.
- 46 NI, L. et al. Streamflow and rainfall forecasting by two long short-term memory-based models. *Journal of Hydrology*, Elsevier, v. 583, p. 124296, 2020.
- 47 SARAIVA, S. V. et al. Daily streamflow forecasting in sobradinho reservoir using machine learning models coupled with wavelet transform and bootstrapping. *Applied Soft Computing*, Elsevier, v. 102, p. 107081, 2021.
- 48 YILMAZ, M. et al. Predicting monthly streamflow using artificial neural networks and wavelet neural networks models. *Modeling Earth Systems and Environment*, Springer, p. 1–17, 2022.
- 49 MOMENEH, S.; NOURANI, V. Application of a novel technique of the multi-discrete wavelet transforms in hybrid with artificial neural network to forecast the daily and monthly streamflow. *Modeling Earth Systems and Environment*, Springer, p. 1–20, 2022.
- 50 XIE, M. et al. Daily streamflow forecasting using hybrid long short-term memory model. In: IOP PUBLISHING. *Journal of Physics: Conference Series*. [S.l.], 2022. v. 2271, n. 1, p. 012019.
- 51 NOURANI, V.; KISI, Ö.; KOMASI, M. Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *Journal of Hydrology*, Elsevier, v. 402, n. 1-2, p. 41–59, 2011.
- 52 DADU, K. S.; DEKA, P. C. Wavelet-neural network conjunction model in flow forecasting of subhimalayan river brahmaputra. *International journal of civil engineering and technology (IJCIET)*, v. 3, n. 2, p. 415–425, 2012.
- 53 KRISHNA, B.; SATYAJI, R. Y.; PC, N. Time series modeling of river flow using wavelet neural networks. *Journal of Water Resource and Protection*, Scientific Research Publishing, v. 2011, 2011.
- 54 PARTAL, T. River flow forecasting using different artificial neural network algorithms and wavelet transform. *Canadian Journal of Civil Engineering*, v. 36, n. 1, p. 26–38, 2009.
- 55 GENÇAY, R.; SELÇUK, F.; WHITCHER, B. J. *An introduction to wavelets and other filtering methods in finance and economics*. [S.l.]: Elsevier, 2001.

- 56 ZAPRANIS, A.; ALEXANDRIDIS, A. Weather analysis & weather derivative pricing. 2006.
- 57 ALEXANDRIDIS, A. K.; ZAPRANIS, A. D.; ACHILLEAS, D. *Wavelet Neural Networks*. [S.l.]: Wiley Online Library, 2014.
- 58 KUMAR, P.; FOUFOULA-GEORGIU, E. Wavelet analysis for geophysical applications. *Reviews of geophysics*, Wiley Online Library, v. 35, n. 4, p. 385–412, 1997.
- 59 ADDISON, P. S. *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. [S.l.]: CRC press, 2017.
- 60 STRANG, G.; NGUYEN, T. *Wavelets and filter banks*. [S.l.]: SIAM, 1996.
- 61 DEBNATH, L.; SHAH, F. A. *Wavelet transforms and their applications*. [S.l.]: Springer, 2002.
- 62 SHENSA, M. J. et al. The discrete wavelet transform: wedding the a trous and mallat algorithms. *IEEE Transactions on signal processing*, v. 40, n. 10, p. 2464–2482, 1992.
- 63 FEIL, M.; UHL, A. Real-time image analysis using wavelets: the "a trous" algorithm on mimd architectures. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Real-Time Imaging IV*. [S.l.], 1999. v. 3645, p. 56–65.
- 64 WANG, K. C.; LI, Q.; GONG, W. Wavelet-based pavement distress image edge detection with a trous algorithm. *Transportation Research Record*, SAGE Publications Sage CA: Los Angeles, CA, v. 2024, n. 1, p. 73–81, 2007.
- 65 JIN, J.; KIM, J. Forecasting natural gas prices using wavelets, time series, and artificial neural networks. *PloS one*, Public Library of Science, v. 10, n. 11, p. e0142064, 2015.
- 66 PARAMESWARIAH, C. B. Understanding wavelet analysis and filters for engineering applications. 2003.
- 67 STOLOJESCU, C. et al. Comparison of wavelet families with application to wimax traffic forecasting. In: IEEE. *2010 12th international conference on optimization of electrical and electronic equipment*. [S.l.], 2010. p. 932–937.
- 68 MALLAT, S. *A wavelet tour of signal processing*. [S.l.]: Elsevier, 1999.
- 69 DAUBECHIES, I. *Ten lectures on wavelets*. [S.l.]: SIAM, 1992.
- 70 KAISER, G.; HUDGINS, L. H. *A friendly guide to wavelets*. [S.l.]: Springer, 1994. v. 300.
- 71 MAIER, H. R.; DANDY, G. C. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental modelling & software*, Elsevier, v. 15, n. 1, p. 101–124, 2000.
- 72 BADRZADEH, H.; SARUKKALIGE, R.; JAYAWARDENA, A. Impact of multi-resolution analysis of artificial intelligence models inputs on multi-step ahead river flow forecasting. *Journal of Hydrology*, Elsevier, v. 507, p. 75–85, 2013.

- 73 EL-CHABIB, H.; NEHDI, M. Neural network modelling of properties of cement-based materials demystified. *Advances in cement research*, Thomas Telford Ltd, v. 17, n. 3, p. 91–102, 2005.
- 74 DIBIKE, Y. B.; SOLOMATINE, D. P. River flow forecasting using artificial neural networks. *Physics and Chemistry of the Earth: B: Hydrology, Oceans and Atmosphere*, Citeseer, v. 26, n. 1, p. 1–7, 2001.
- 75 STEYN, M. et al. Short-term stream flow forecasting at australian river sites using data-driven regression techniques. 2017.
- 76 KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- 77 LIU, D. C.; NOCEDAL, J. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, Springer, v. 45, n. 1, p. 503–528, 1989.
- 78 HAYKIN, S. S. et al. *Neural networks and learning machines*. [S.l.]: Pearson Upper Saddle River, NJ, USA:, 2009. v. 3.
- 79 NISSEN, S. Neural networks made simple. *Software*, v. 2, n. 2, p. 14–19, 2005.
- 80 HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*. 2. ed. Verlag, New York: Springer, 2009. 745 p. ISBN 978-0-387-84857-0.
- 81 LEGATES, D. R.; MCCABE, G. J. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water resources research*, Wiley Online Library, v. 35, n. 1, p. 233–241, 1999.
- 82 CHAI, T.; DRAXLER, R. R. Root mean square error (rmse) or mean absolute error (mae)?—arguments against avoiding rmse in the literature. *Geoscientific model development*, Copernicus GmbH, v. 7, n. 3, p. 1247–1250, 2014.
- 83 KRAUSE, P.; BOYLE, D.; BÄSE, F. Comparison of different efficiency criteria for hydrological model assessment. *Advances in geosciences*, v. 5, p. 89–97, 2005.
- 84 YAP, B. W.; SIM, C. H. Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 81, n. 12, p. 2141–2155, 2011.
- 85 MCKNIGHT, P. E.; NAJAB, J. Mann-whitney u test. *The Corsini encyclopedia of psychology*, Wiley Online Library, p. 1–1, 2010.
- 86 CHOURASIA, V.; MITTRA, A. Selection of mother wavelet and denoising algorithm for analysis of foetal phonocardiographic signals. *Journal of medical engineering & technology*, Taylor & Francis, v. 33, n. 6, p. 442–448, 2009.
- 87 NOURANI, V.; KOMASI, M.; MANO, A. A multivariate ann-wavelet approach for rainfall–runoff modeling. *Water resources management*, Springer, v. 23, n. 14, p. 2877–2894, 2009.
- 88 SHOAI, M. et al. Hybrid wavelet neural network approach. In: *Artificial Neural Network Modelling*. [S.l.]: Springer, 2016. p. 127–143.

- 89 SANG, Y.-F. et al. Wavelet-based hydrological time series forecasting. *J. Hydrol. Eng.*, v. 21, n. 5, p. 06016001, 2016.
- 90 NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*. [S.l.: s.n.], 2010. p. 807–814.
- 91 Instituto Estadual do Ambiente. *Plano Estadual de Recursos Hídricos do Estado do Rio de Janeiro*. [S.l.]: LabHid/Coppe-UFRJ, 2014.
- 92 Fundação COPPETEC. *Elaboração do plano estadual de recursos hídricos do estado do Rio de Janeiro. R2-F-Characterização Ambiental*. 2014. <<https://goo.gl/NNdpDQ>>. Acesso em: 14 set. 2018.
- 93 Fundação COPPETEC. *Plano de Recursos Hídricos da bacia do Rio Paraíba do Sul - Resumo. Caderno de Ações Área de Atuação do CBH-PS*. 2006. <<https://bit.ly/2NhbK2w>>. Acesso em: 17 set. 2018.
- 94 PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, v. 12, n. Oct, p. 2825–2830, 2011.
- 95 LEE, G. R. et al. Pywavelets: A python package for wavelet analysis. *J. Open Source Softw*, v. 4, n. 36, p. 1237, 2019.
- 96 WASKOM, M. et al. *mwaskom/seaborn: v0.8.1 (September 2017)*. 2017. Disponível em: <<https://doi.org/10.5281/zenodo.883859>>.
- 97 MCKINNEY, W. Data structures for statistical computing in python. In: WALT, S. van der; MILLMAN, J. (Ed.). *Proceedings of the 9th Python in Science Conference*. [S.l.: s.n.], 2010. p. 51 – 56.
- 98 SEO, Y. et al. Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. *Journal of Hydrology*, Elsevier, v. 520, p. 224–243, 2015.
- 99 GORODETSKAYA, Y.; FONSECA, L. G. da; RIBEIRO, C. B. de M. Coupling the trous wavelet decomposition algorithm and artificial neural networks for short-term stream flow forecasting of the low lands of paraíba do sul river. In: *Anais do XXII ENMC – Encontro Nacional de Modelagem Computacional e X ECTM – Encontro de Ciências e Tecnologia de Materiais*. [S.l.: s.n.], 2019.
- 100 TONGAL, H.; BERNDTSSON, R. Impact of complexity on daily and multi-step forecasting of streamflow with chaotic, stochastic, and black-box models. *Stochastic environmental research and risk assessment*, Springer, v. 31, n. 3, p. 661–682, 2017.
- 101 HADI, S. J.; TOMBUL, M. Forecasting daily streamflow for basins with different physical characteristics through data-driven methods. *Water Resources Management*, Springer, p. 1–18, 2018.

## APÊNDICE A – Resultados

Nesta seção são apresentados os seguintes resultados:

- os hidrogramas da vazão observada e prevista por modelos  $WANN_{one}$ ,  $WANN_{multi}$  e ANN para o horizonte de 1, 3 e 5 dias à frente no caso das estações:
  - 58183000 (Figura 40);
  - 58235100 (Figura 41);
  - 58630002 (Figura 42);
  - 58380001 (Figura 43);
  - 58305001 (Figura 44);
  - 58880001 (Figura 45);
  - 58790002 (Figura 46);
  - 58795000 (Figura 47).
- gráficos de dispersão dos valores observados versus os previstos por modelos  $WANN_{one}$ ,  $WANN_{multi}$  e ANN para o horizonte de previsão de 7 dias à frente no caso das estações:
  - 58630002 (Figura 48);
  - 58380001 (Figura 49);
  - 58305001 (Figura 50);
  - 58235100 (Figura 51);
  - 58790002 (Figura 52);
  - 58795000 (Figura 53);
  - 58880001 (Figura 54);
  - 58183000 (Figura 55).

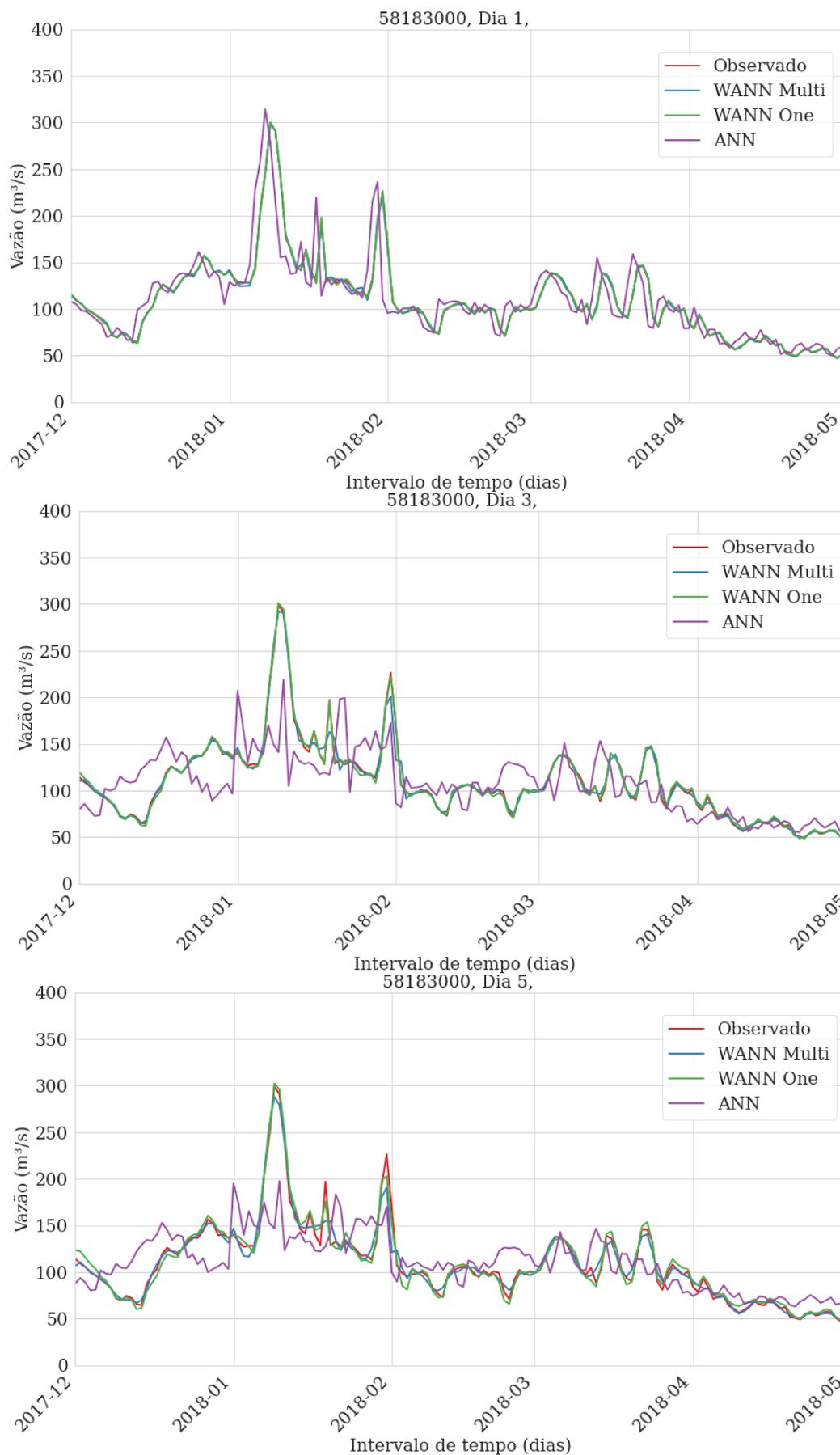


Figura 40 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58183000.

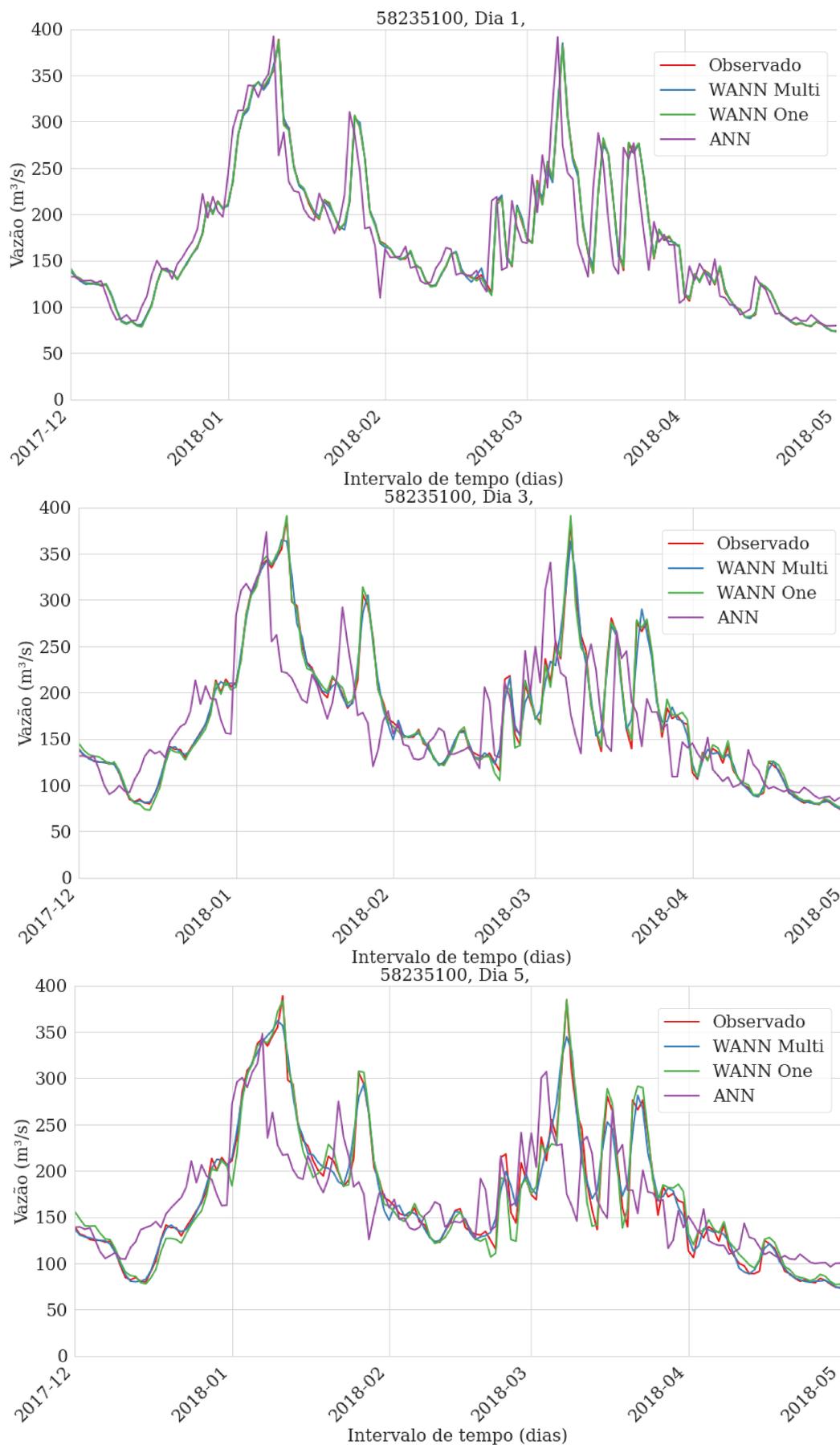


Figura 41 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 582351000.

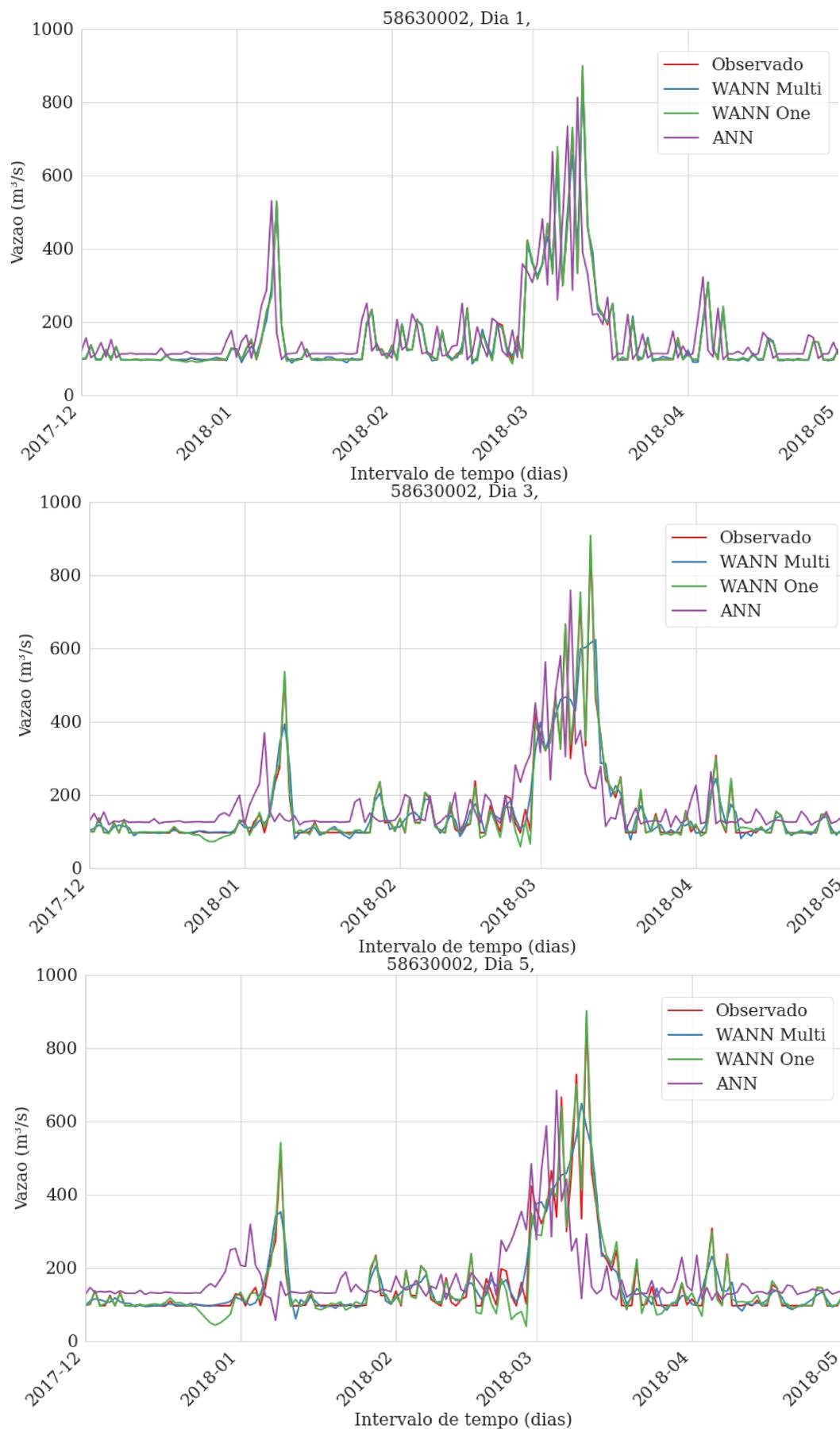


Figura 42 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58630002.

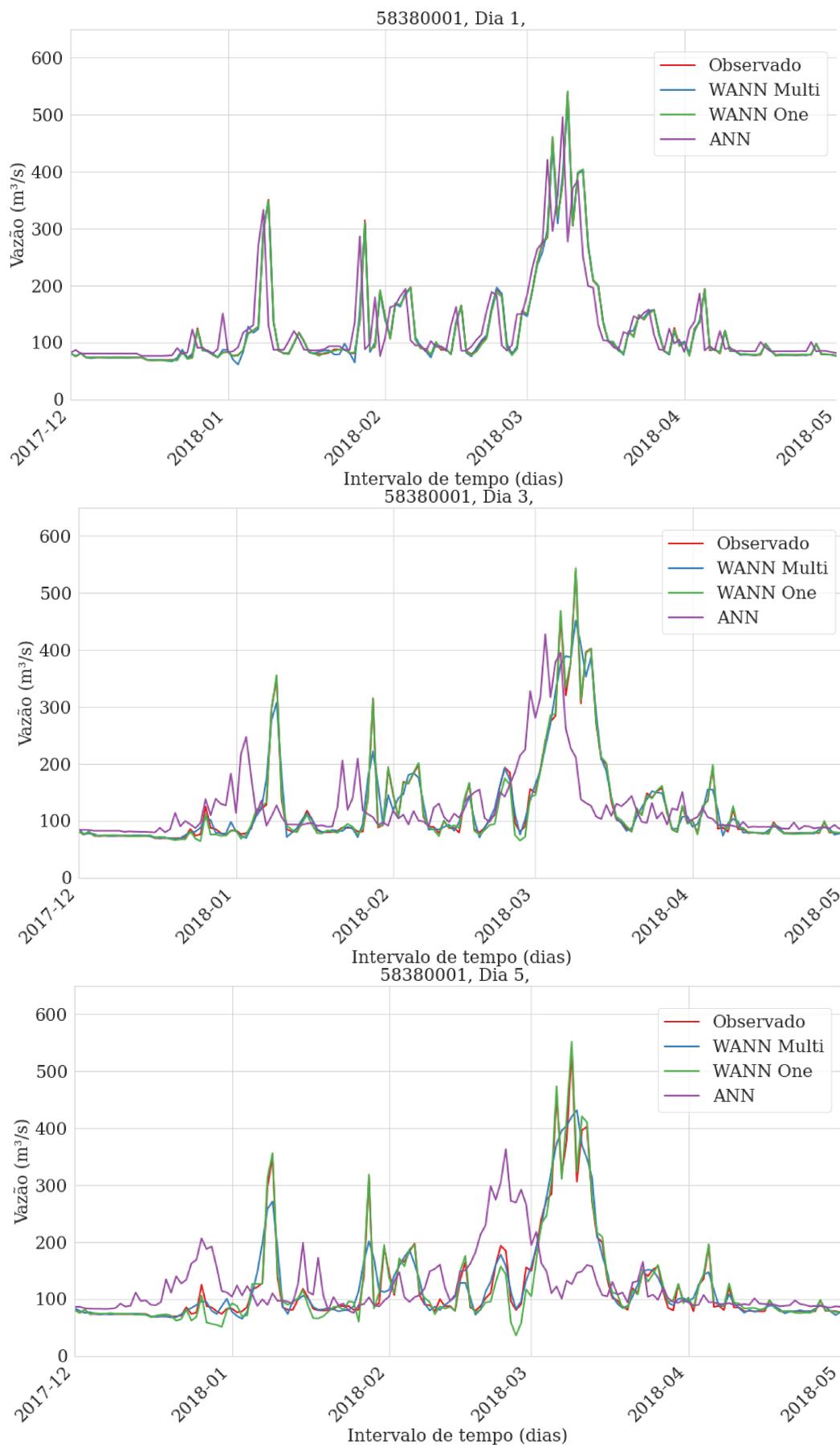


Figura 43 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58380001.

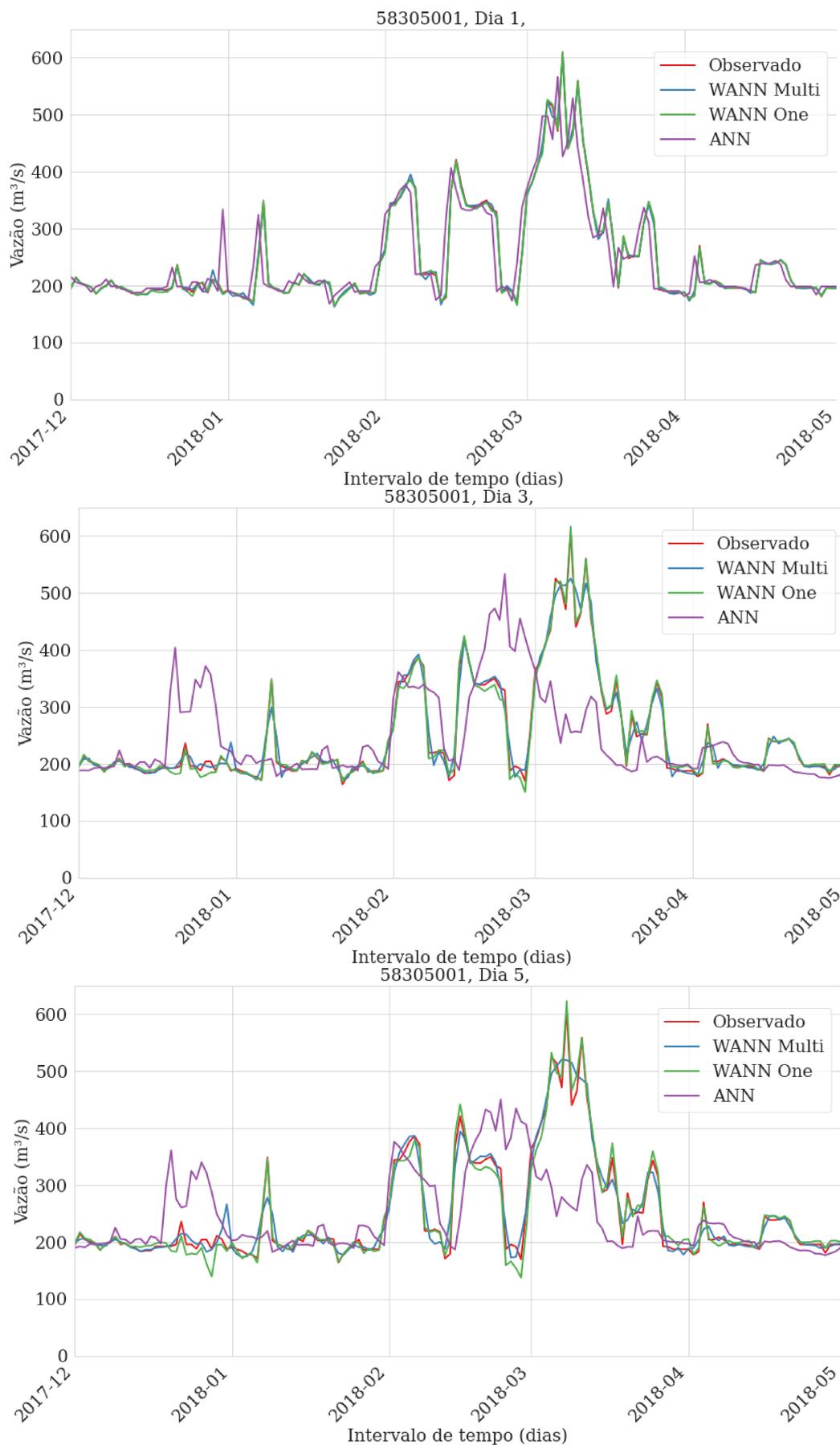


Figura 44 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 583050001.

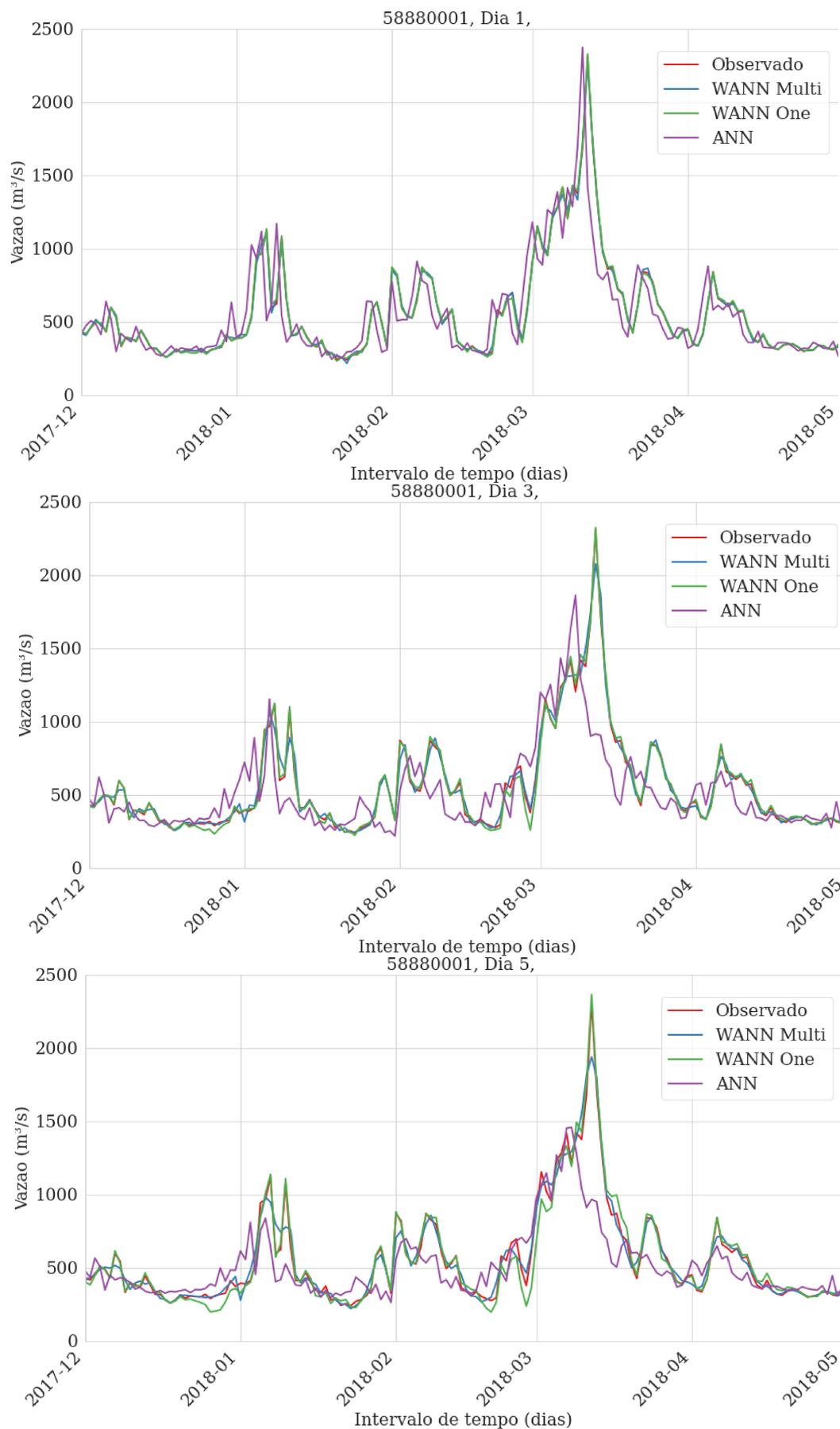


Figura 45 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58880001.

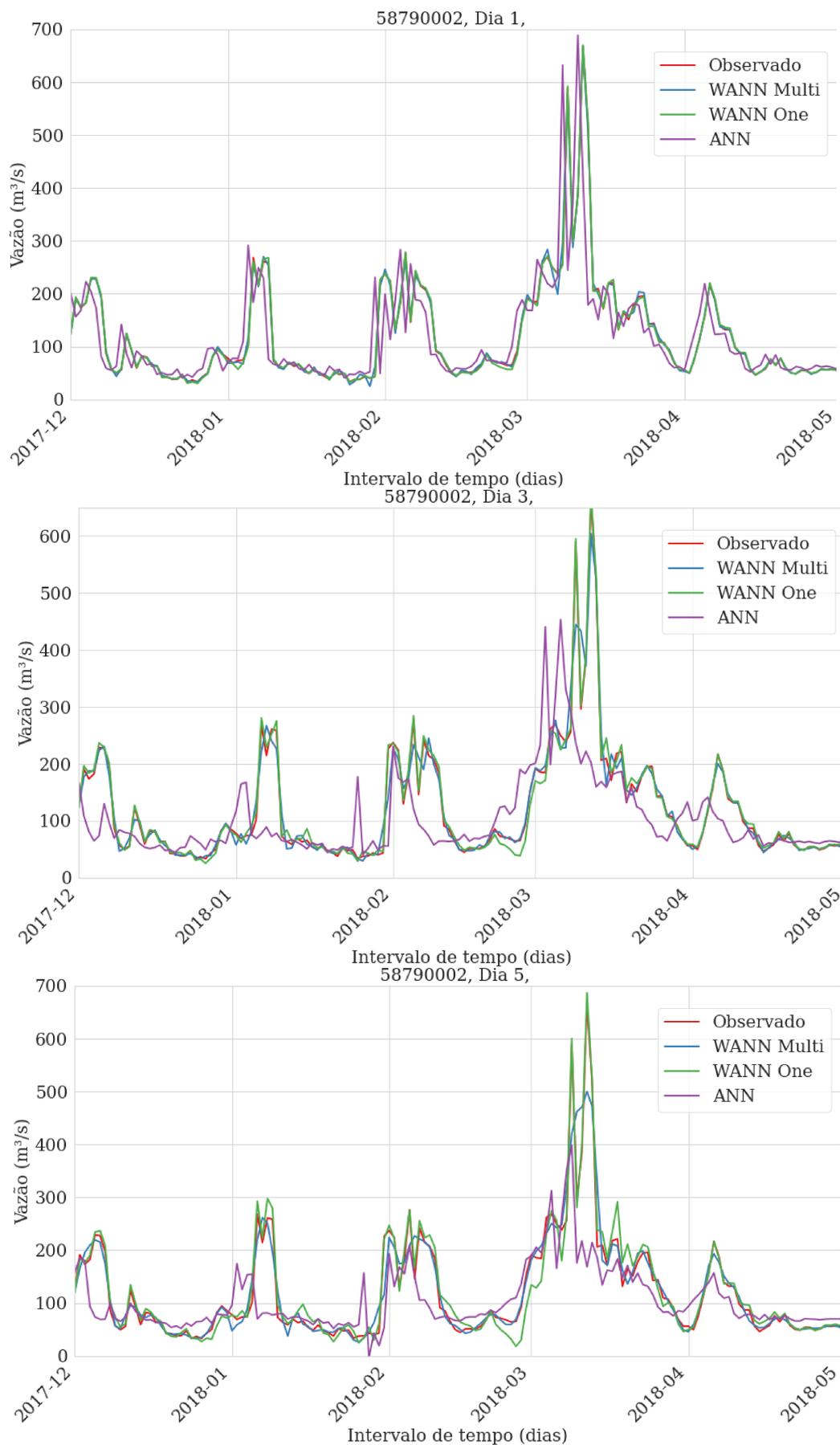


Figura 46 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58790002.

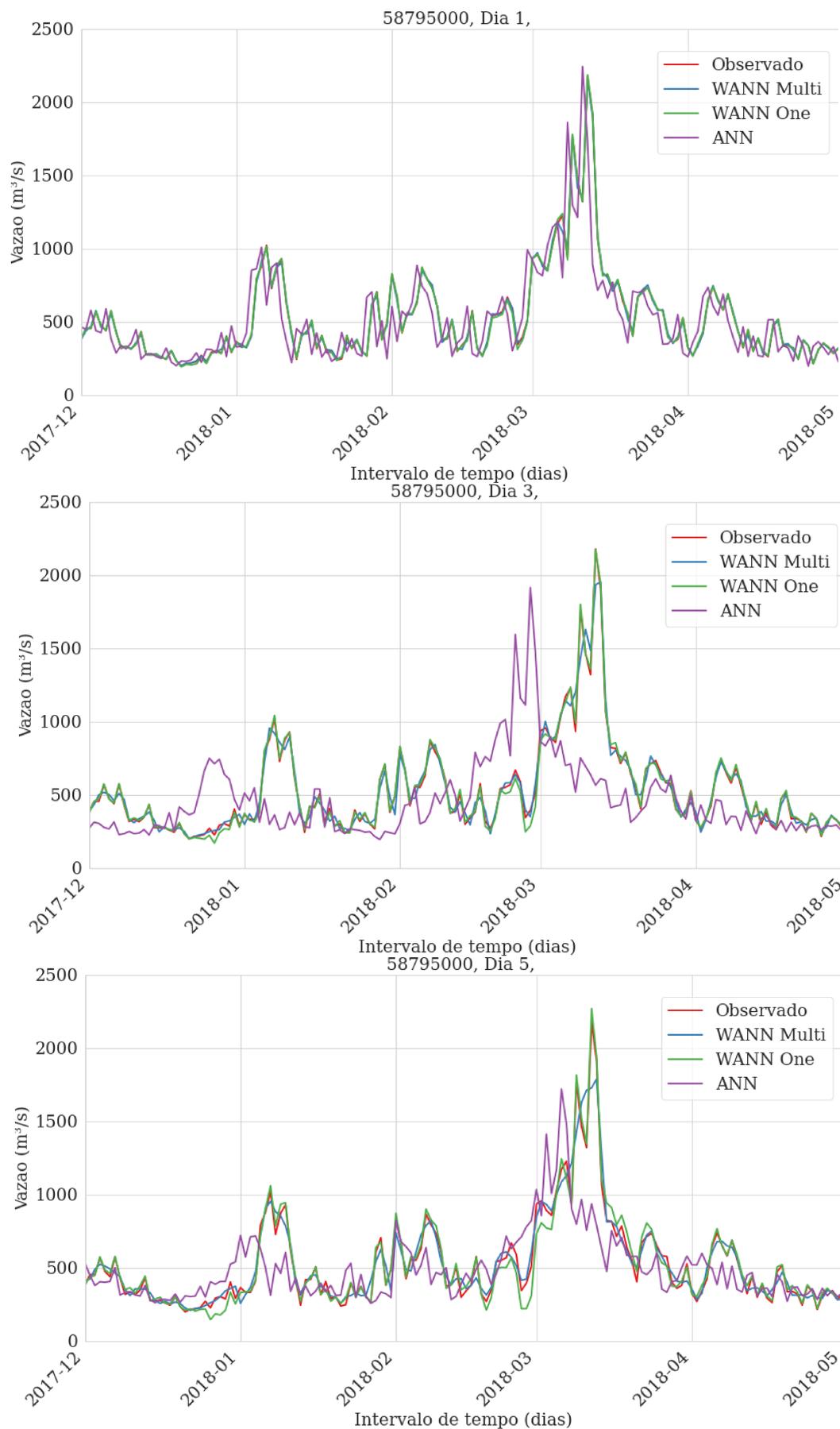


Figura 47 – Os hidrogramas da vazão observada e prevista por modelos WANN e ANN para o horizonte de 1, 3 e 5 dias à frente no caso da estação 58795000.

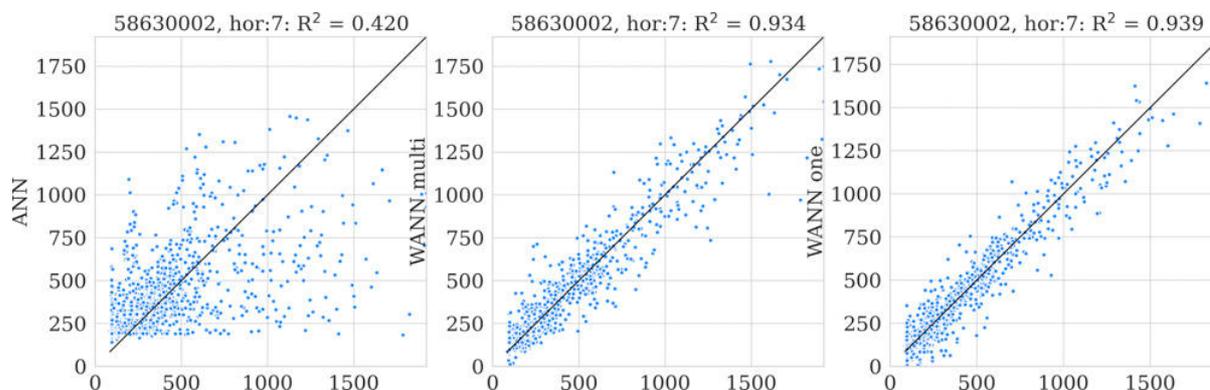


Figura 48 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58630002 para o horizonte de previsão de 7 dias à frente.

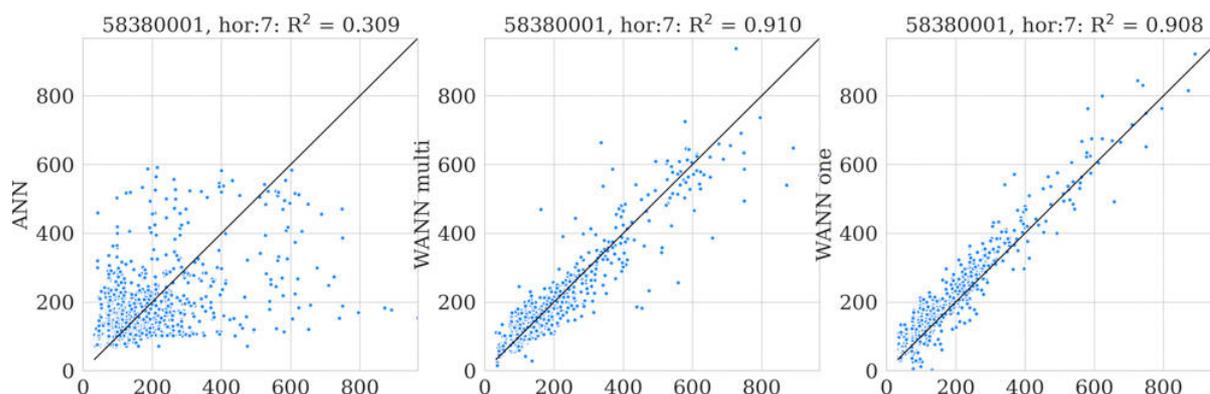


Figura 49 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58380001 para o horizonte de previsão de 7 dias à frente.

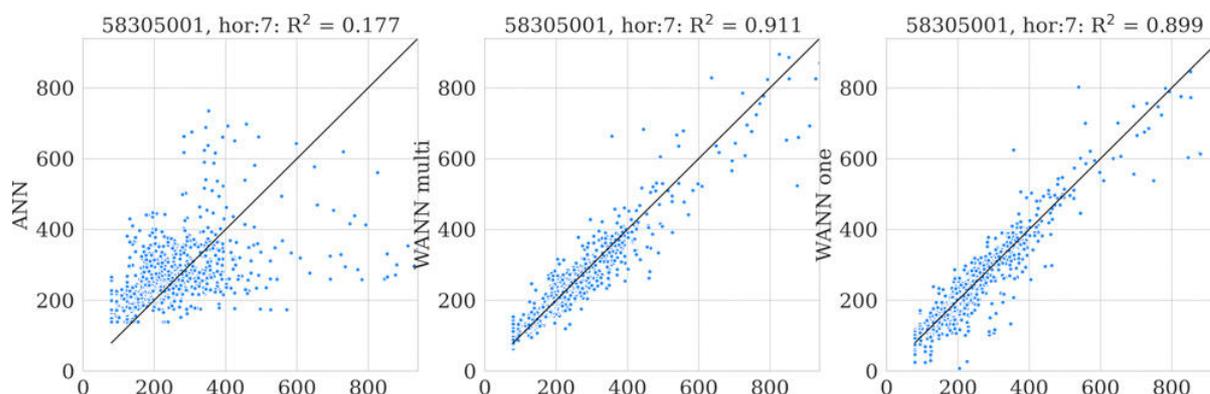


Figura 50 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58305001 para o horizonte de previsão de 7 dias à frente.

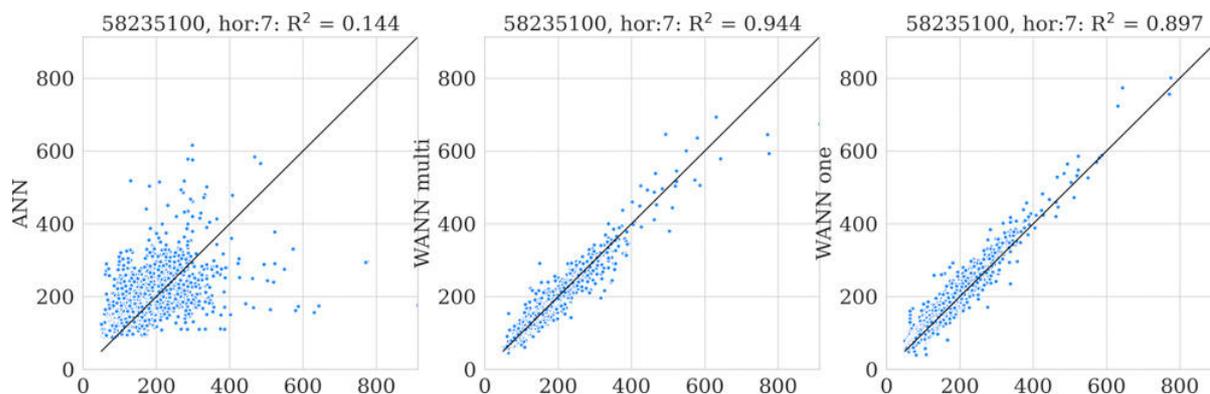


Figura 51 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58235100 para o horizonte de previsão de 7 dias à frente.

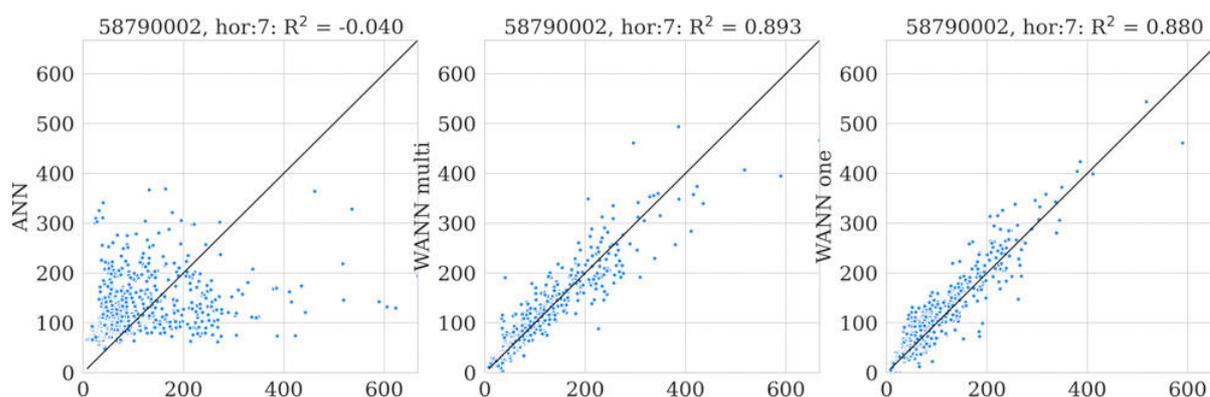


Figura 52 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58790002 para o horizonte de previsão de 7 dias à frente.

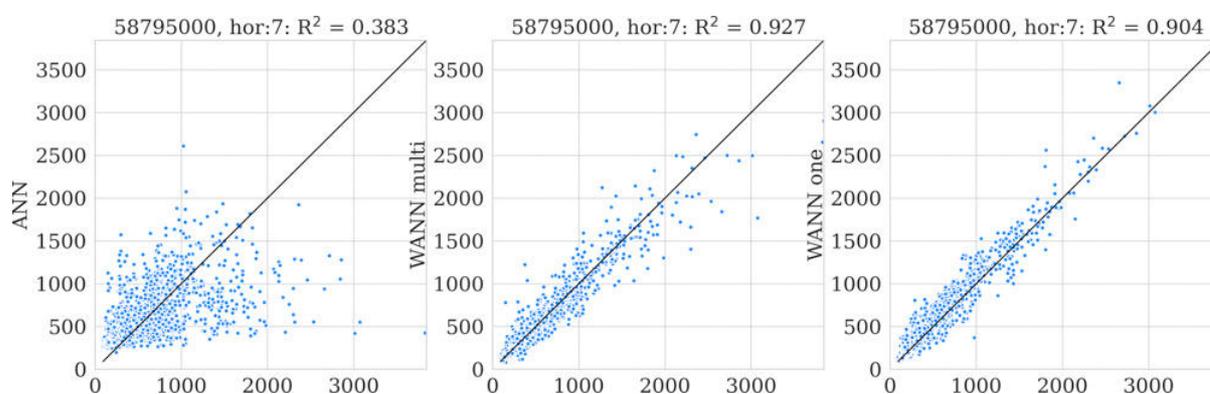


Figura 53 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58795000 para o horizonte de previsão de 7 dias à frente.

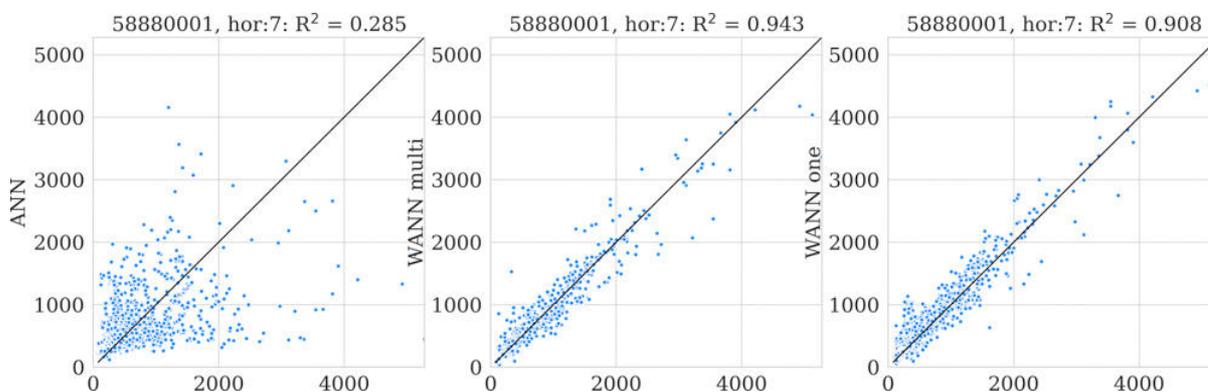


Figura 54 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58880001 para o horizonte de previsão de 7 dias à frente.

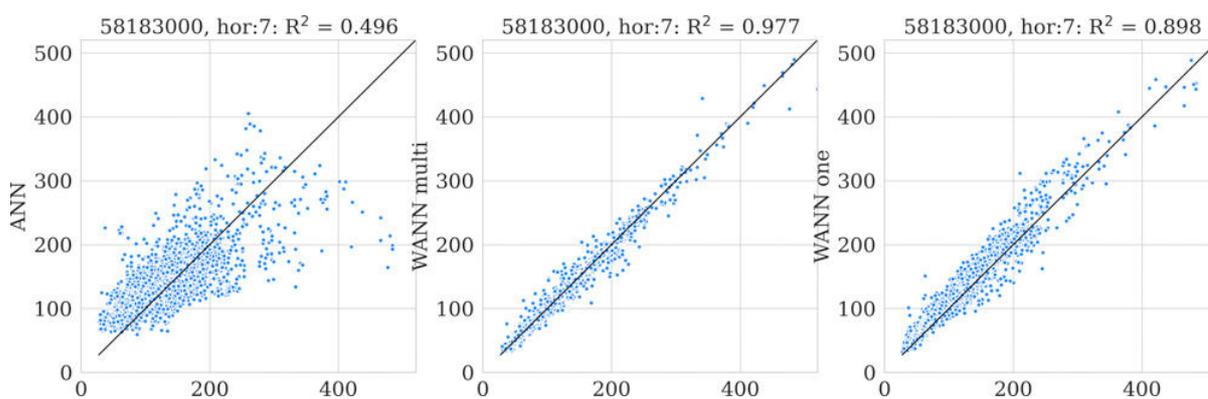


Figura 55 – Gráficos de dispersão dos valores observados versus os previstos no caso da estação 58183000 para o horizonte de previsão de 7 dias à frente.

**APÊNDICE B – Trabalhos publicados**

1. Gorodetskaya, Y.; Goliatt L. F.; Ribeiro, C. B. M. Previsão de vazão de curto prazo utilizando Redes Neurais Artificiais. Estudo De Caso: Rio Paraíba Do Sul. In: *III Simpósio de Recursos Hídricos da Bacia do Rio Paraíba do Sul*, 2018, Juiz de Fora. III SRH-PS 2018.
2. Gorodetskaya, Y.; Tavares, G. G.; Goliatt L. F.; Ribeiro, C. B. M. Comparação de métodos de aprendizado de máquina para a previsão de curto prazo de vazão do baixo curso do rio paraíba do sul. In: *XXI Encontro Nacional de Modelagem Computacional e IX Encontro de Ciência e Tecnologia de Materiais*, 2018, Búzios. ENMC & ECTM 2018.
3. Gorodetskaya, Y.; Goliatt L. F.; Tavares, G. G.; Ribeiro, C. B. M. Daily streamflow forecasting for Paraíba do Sul river using machine learning methods with hydrologic inputs. In: *XV Encontro Nacional de Inteligência Artificial e Computacional*, 2018, São Paulo. ENIAC 2018.
4. Gorodetskaya, Y.; Goliatt L. F.; Ribeiro, C. B. M. Coupling the À Trous wavelet decomposition algorithm and artificial neural networks for short-term stream flow forecasting of the low lands of Paraíba do Sul River. In: *XXII Encontro Nacional de Modelagem Computacional e X ECTM – Encontro de Ciências e Tecnologia de Materiais*, 2019, Juiz de Fora, ENMC & ECTM 2019.
5. Campos, L. C. D., Goliatt da Fonseca, L., Fonseca, T. L., Abreu, G. D. D., Pires, L. F., Gorodetskaya, Y. Short-term streamflow forecasting for Paraíba do Sul river using deep learning. In: *EPIA Conference on Artificial Intelligence*. Springer, Cham, 2019.
6. Fonseca, T. L., Gorodetskaya, Y., Tavares, G. G., de Melo Ribeiro, C. B., da Fonseca, L. G.. A Gradient Boosting Model Optimized by a Genetic Algorithm for Short-term Riverflow Forecast. In: *Revista Mundi Engenharia, Tecnologia e Gestão* (ISSN: 2525-4782), v. 4, n. 3, 2019.
7. Martinho, A. D., Ribeiro, C., Gorodetskaya, Y., Fonseca, T. L., Goliatt, L. Extreme learning machine with evolutionary parameter tuning applied to forecast the daily natural flow at Cahora Bassa dam, Mozambique. In: *International Conference on Bioinspired Methods and Their Applications* Springer, Cham, 2020. p. 255-267.