

Daily streamflow forecasting in Sobradinho Reservoir using machine learning models coupled with wavelet transform and bootstrapping

Samuel Vitor Saraiva, Frede de Oliveira Carvalho, Celso Augusto Guimarães Santos,
Lucas Costa Barreto, Paula Karenina de Macedo Machado Freire

Apresentação: Rodrigo Cunha

São Paulo, 7 de dezembro de 2021



Tópicos

1. Introdução
2. Objetivos
3. Transformada discreta de ondaletas (DWT)
4. Redes neurais artificiais (ANN)
5. Máquina de vetores de suporte (SVM)
6. Experimentos
7. Resultados
8. Próximos passos



1. Introdução

- *Daily streamflow forecasting in Sobradinho Reservoir using machine learning models coupled with wavelet transform and bootstrapping*, Applied Soft Computing Journal (2021)
- Pesquisadores brasileiros (UNICAMP, UFAL, UFPB)
- Problema de hidrologia
 - Previsão da vazão de água no reservatório da usina de Sobradinho, BA
- Dados de vazão obtidos da ONS (Operador Nacional do Sistema, <http://www.ons.org.br>)
 - Medidas diárias de vazão em m^3/s , compreendendo o período de 1931 a 2015.
 - Treino: 1931 a 2004
 - Teste: 2005 a 2015



2. Objetivos

- Utilizar a transformada discreta de ondaletas (DWT) para decompor a série de vazão
- As séries decompostas são utilizadas como input de modelos de *machine learning* (ML)
 - *Artificial Neural Network* (ANN) e *Support Vector Machine* (SVM)
- O objetivo é expandir o espaço de treinamento dos modelos de modo a melhorar o desempenho na previsão da vazão

3. Transformada discreta de ondaletas (DWT)

- Analogamente à transformada de Fourier, a DWT decompõe o sinal a partir de uma base gerada por uma família de funções. Porém, diferente de senos e cossenos, as ondaletas são funções localizadas no tempo. Todo o espaço $L^2(\mathbb{R})$ pode ser preenchido a partir de translações e dilatações de uma única ondaleta mãe.

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), \quad j, k \in \mathbb{Z}, \quad \Rightarrow \quad f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(t)$$

- As ondaletas podem ser geradas a partir de uma *ondaleta pai*, ou *função escala*.

$$\phi(t) = \sqrt{2} \sum_k \ell_k \phi(2t - k). \quad \Rightarrow \quad \phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k), \quad j, k \in \mathbb{Z}. \quad \Rightarrow \quad \psi(t) = \sqrt{2} \sum_k h_k \phi(2t - k),$$

- Os coeficientes ℓ_k e h_k definem respectivamente filtros *passa-baixo* e *passa-alto*

$$\ell_k = \sqrt{2} \int_{-\infty}^{\infty} \phi(t) \phi(2t - k) dt, \quad h_k = \sqrt{2} \int_{-\infty}^{\infty} \psi(t) \phi(2t - k) dt.$$



3. Transformada discreta de ondaletas (DWT)

Dessa forma, é possível representar uma função a partir da sua decomposição com a transformada de ondaletas:

$$f(t) = \sum_k c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j \geq j_0} \sum_k d_{j,k} \psi_{j,k}(t),$$

onde

$$c_{j_0,k} = \int_{-\infty}^{\infty} f(t) \phi_{j_0,k}(t) dt,$$

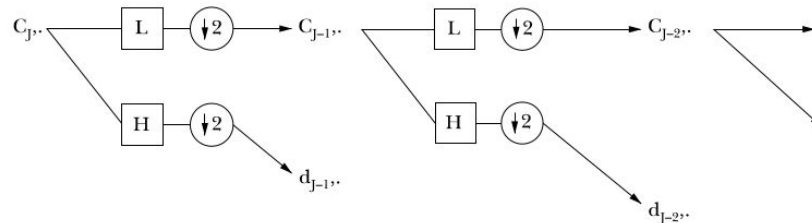
$$d_{j,k} = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt.$$

3. Transformada discreta de ondaletas (DWT)

Representando a DWT na forma da equação matricial $\mathbf{d} = \mathbf{W}^* \mathbf{x}$, onde \mathbf{x} corresponde ao sinal original e \mathbf{W} à matriz de transformação, os coeficientes da DWT representados pelo vetor \mathbf{d} podem ser obtidos a partir do algoritmo piramidal:

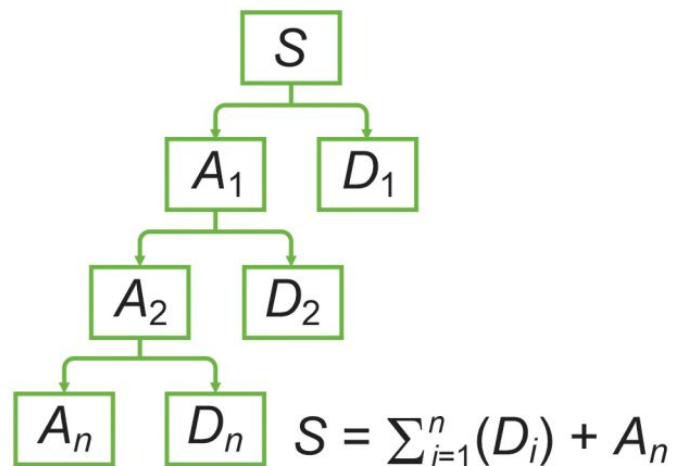
$$\mathbf{d} = (\mathbf{c}_0, \mathbf{d}_0, \mathbf{d}_1, \dots, \mathbf{d}_{J-2}, \mathbf{d}_{J-1}),$$

$$\mathbf{d} = (L^n \mathbf{x}, HL^{n-1} \mathbf{x}, \dots, HL^2 \mathbf{x}, HL \mathbf{x}, L \mathbf{x}).$$



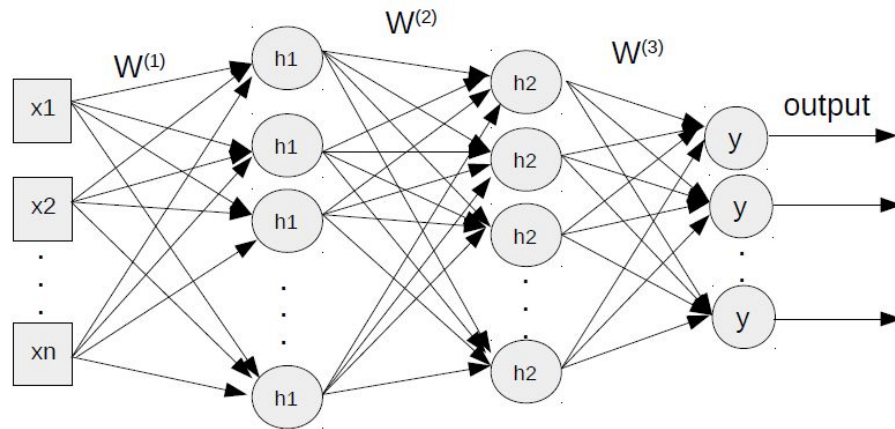
3. Transformada discreta de ondaletas (DWT)

A decomposição da série original permite que diferentes dinâmicas presentes no sinal original sejam analisadas separadamente pelos modelos de previsão.



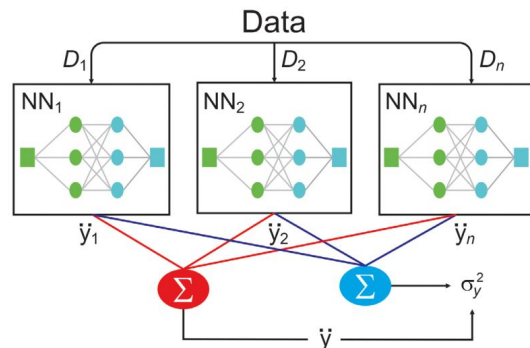
4. Redes neurais artificiais (ANN)

As redes neurais artificiais são modelos de aprendizado de máquina supervisionado inspirados no funcionamento do neurônio biológico. Em um modelo ANN com arquitetura MLP (*Multi-layer Perceptron*), cada nó corresponde a um *neurônio*, que recebe o somatório das saídas dos neurônios da camada anterior ponderados por um vetor de pesos e o submete à uma função de ativação, gerando a sua saída.



4. Redes neurais artificiais (ANN)

O autor do artigo prevê a utilização de modelos com arquitetura MLP e uma variação deste modelo que utiliza a técnica de *bootstrapping*. Nessa arquitetura, ao invés de um único modelo, **NN0**, são treinados diversos modelos (**NN1**, **NN2**, ...); cada modelo é treinado em um dataset **D_i**, que é uma subamostragem com reposição do dataset original.



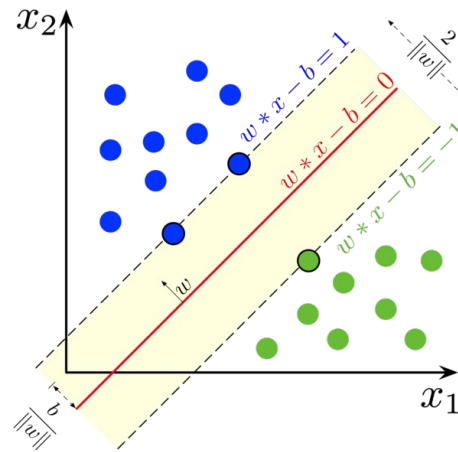
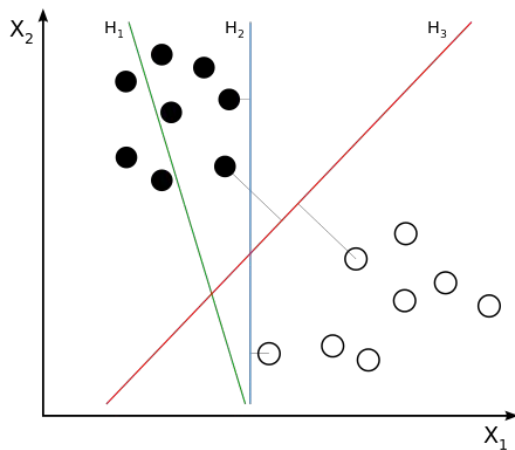


4. Redes neurais artificiais (ANN)

- Atingem bons resultados em uma ampla variedade de problemas
- Modelos flexíveis, podendo ser aplicados para problemas de classificação e regressão
- Difíceis de treinar: alta demanda por dados, tempo e capacidade computacional
- A arquitetura MLP não é a mais adequadas para problemas envolvendo séries temporais
 - Arquiteturas como as redes neurais recorrentes (RNN) obtém melhores resultados

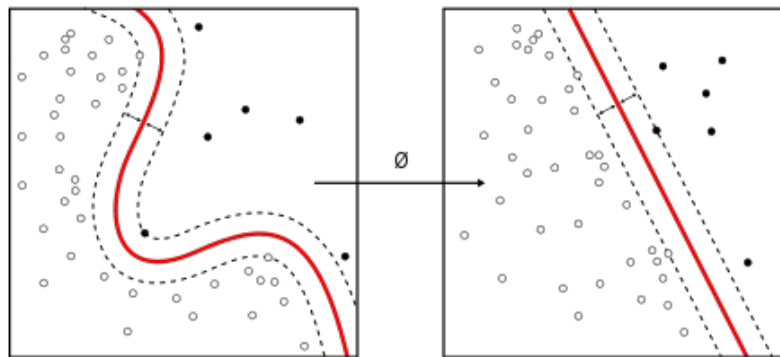
5. Máquina de Vetores de Suporte (SVM)

As SVMs (*Support Vector Machines*) são modelos de aprendizado supervisionado propostos por Vladimir Vapnik (1992, 1997). As SVM basicamente constroem hiperplanos em espaços de dimensões elevadas e podem ser utilizados em tarefas de classificação e regressão.



5. Máquina de Vetores de Suporte (SVM)

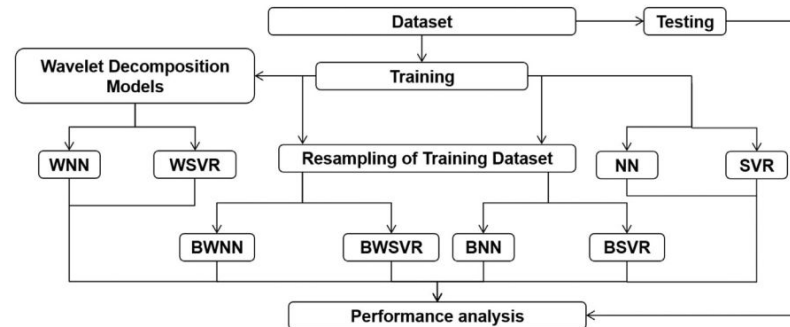
A SVM original, proposta inicialmente por Vapnik, é um classificador linear. Porém, posteriormente as SVMs foram modificadas permitindo a aplicação de uma função **kernel**, que substitui o produto escalar, permitindo uma transformação do espaço de atributos. Tal transformação permite que as SVMs resolvam problemas que não são linearmente separáveis no espaço de atributos original.



6. Experimentos

Os experimentos conduzidos pelo autor do artigo avaliam o desempenho dos modelos ANN e SVM em diferentes configurações:

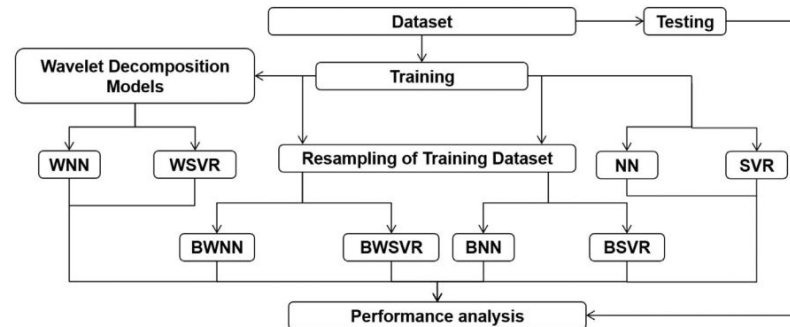
- Com e sem a decomposição utilizando ondaletas
- Com e sem a técnica de *bootstrapping*



6. Experimentos

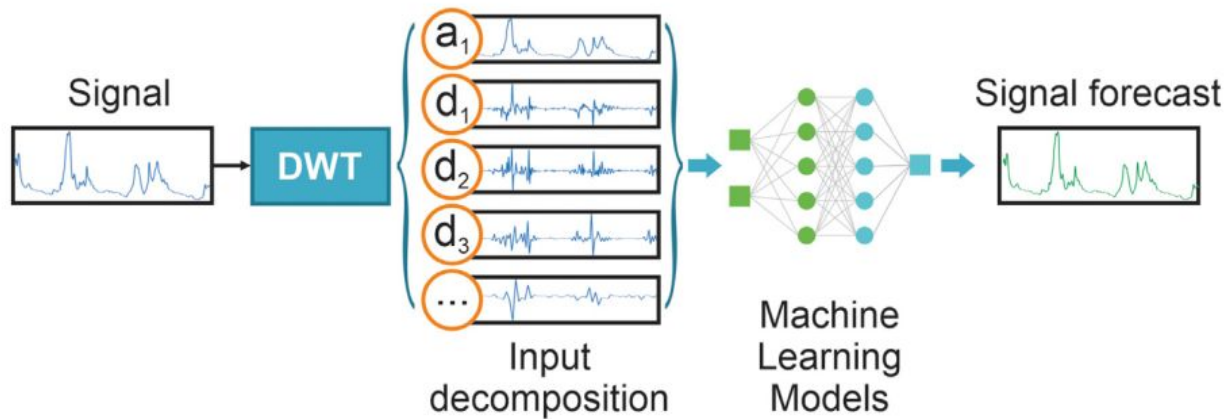
Os experimentos conduzidos pelo autor do artigo avaliam o desempenho dos modelos ANN e SVM em diferentes configurações:

- Com e sem a decomposição utilizando ondaletas
- Com e sem a técnica de *bootstrapping*



6. Experimentos

O modelo completo considerando a decomposição da série original utilizando ondaletas tem a seguinte forma:



7. Resultados

Os resultados obtidos pelo autor apontam para uma melhoria significativa no desempenho das previsões ao aplicar DWT para pré-processar os dados de entrada dos modelos.

(ii) The correlation index (R) and the Nash–Sutcliffe coefficient (E):

$$R = \sqrt{\frac{\sum_{i=1}^n [(y_i^0 - \bar{y}^0)(y_i^f - \bar{y}^f)]}{\sqrt{\sum_{i=1}^n (y_i^0 - \bar{y}^0)^2 \cdot \sum_{i=1}^n (y_i^f - \bar{y}^f)^2}}}; \quad (10)$$

$$E = 1 - \frac{\sum_{i=1}^n (y_i^0 - y_i^f)^2}{\sum_{i=1}^n (y_i^0 - \bar{y}^0)^2}$$

(iii) The Akaike information criterion (AIC):

$$AIC = n \left[\ln \left(\frac{E}{n} \right) + 2 \frac{d}{n} \right] + n [n, \ln(2\pi) + 1] \quad (11)$$

where y_i^0 represents the observed values and y_i^f represents the simulated values. Similarly, \bar{y}^0 and \bar{y}^f represent the means of the simulated and forecasted values for n observed samples, respectively. Note that R is in the range of -1 to 1 , and E is in the range of $-\infty$ to 1 , where 1 represents an almost perfect forecast.

Table 2

Performance of the models without bootstrapping for different amounts of input data.

Machine:	ANN			SVR		
	$V_t = f(V_{t-1}, V_{t-2}, \dots, V_{t-i})$					
Number of previous days:	$i = 3$	$i = 5$	$i = 7$	$i = 3$	$i = 5$	$i = 7$
R	0.99957	0.99959	0.99961	0.98924	0.98648	0.98583
E	0.99914	0.99918	0.99922	0.97786	0.9722	0.97109
MSE	23.434	23.142	22.978	80.563	89.52	85.253
MAE	1678.9	1611.7	1540.7	93905	117910	122620
AIC	67802	67846	67927	55174	55979	57371

Machine:	WNN			WSVR		
Number of previous days:	$i = 3$	$i = 5$	$i = 7$	$i = 3$	$i = 5$	$i = 7$
R	0.99997	0.99999	0.99999	0.99441	0.98909	0.98149
E	0.99995	0.99999	0.99999	0.98844	0.97753	0.96219
MAE	6.3711	2.7082	1.5698	69.131	77.52	97.819
MSE	107.11	19.381	6.4908	49023	95314	160360
AIC	67963	68190	68511	67828	67762	67551

Table 3

Calculation of the errors of the learning machines with the resampling of the validation data.

Machine:	BNN			BSVR		
	$V_t = f(V_{t-1}, V_{t-2}, \dots, V_{t-i})$					
No. of previous days:	$i = 3$	$i = 5$	$i = 7$	$i = 3$	$i = 5$	$i = 7$
R	0.99711	0.99961	0.99962	0.99281	0.98604	0.97991
E	0.99422	0.99922	0.99924	0.9852	0.97118	0.95887
MAE	22.382	22.593	22.348	70.256	81.935	99.191
MSE	2848.3	1522.3	1495.4	62780	122260	174470
AIC	68180	68620	69241	67920	68000	68101

Machines:	BWNN			BWSVR		
No. of previous days:	$i = 3$	$i = 5$	$i = 7$	$i = 3$	$i = 5$	$i = 7$
R	0.99997	0.99999	1	0.99239	0.9892	0.98423
E	0.99995	0.99999	1	0.98435	0.97789	0.96762
MAE	6.3393	2.7208	1.5219	123.2	128.79	101.29
MSE	105.87	19.081	6.0869	66388	93800	137350
AIC	69860	72059	75079	68507	68806	69040

7. Resultados

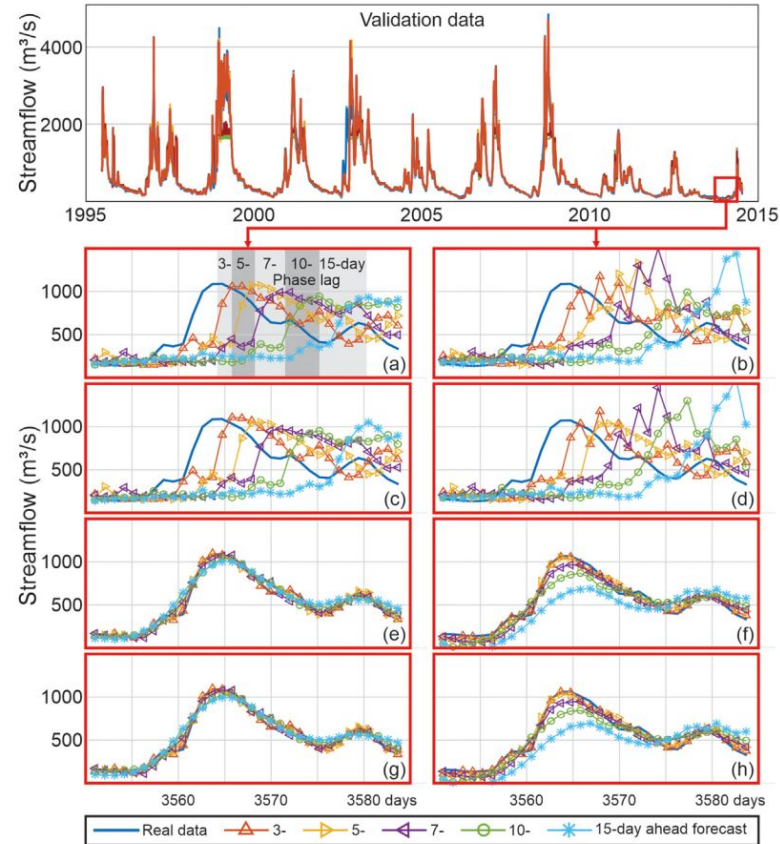


Fig. 9. Forecasts of ensemble models for several prediction horizons: (a) artificial neural network (ANN); (b) support vector regression (SVR); (c) bootstrap neural network (BNN); (d) bootstrap support vector regression (BSVR); (e) wavelet neural network (WNN); (f) wavelet support vector regression (WSVR); (g) bootstrap wavelet neural network (BWNN); (h) bootstrap wavelet support vector regression (BWSVR).

7. Resultados

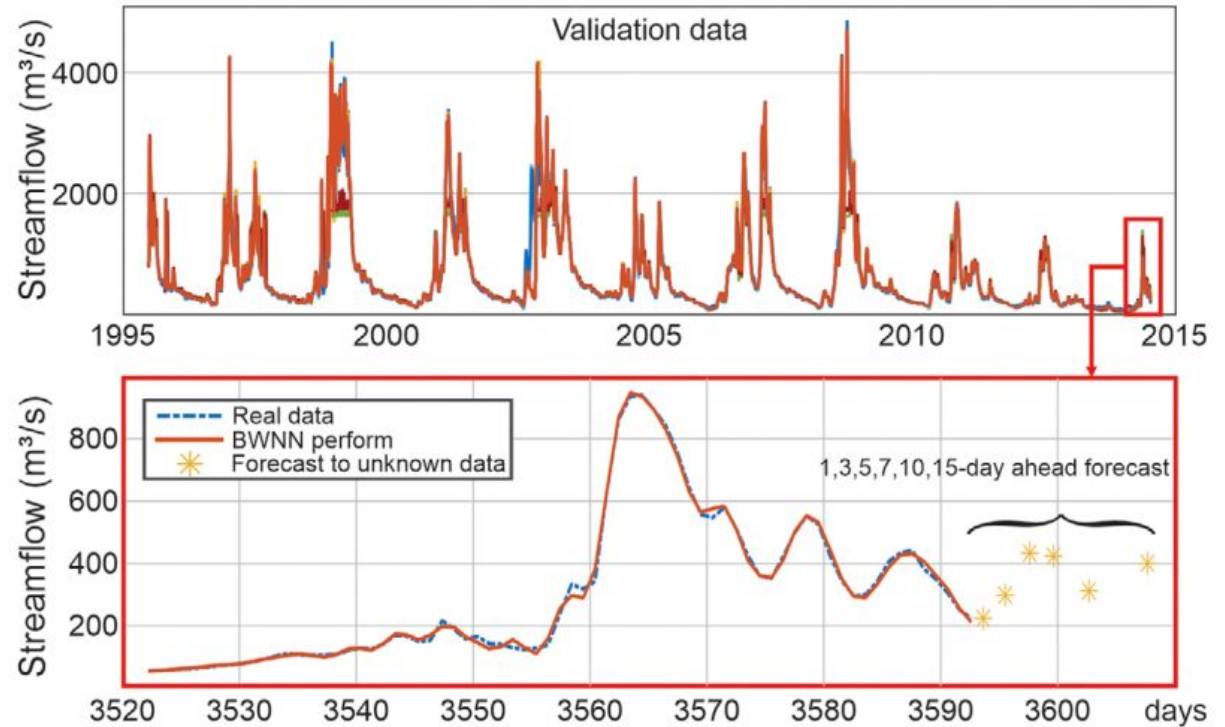


Fig. 10. Performance of the BWNN when forecasting unknown data.



5. Próximos passos

- Aplicar as ideias do artigo utilizando os dados hídricos da bacia do Rio Paraná, fornecidos pela ONS
 - Cenário atual de escassez hídrica e insegurança energética no Brasil
- Comparar os resultados obtidos com os modelos ANN e SVM com modelos ARIMA



Fim.

Obrigado pela atenção!