

Rede Neural Convolucional 1D aplicada à previsão da vazão no Rio Madeira

Felipe Oliveira Barino e Alexandre Bessa dos Santos

Resumo— Este trabalho apresenta um esquema de previsão multi-dia para vazão no Rio Madeira. O modelo preditivo é baseado em uma rede neural convolucional unidimensional. Para o projeto do modelo nós apresentamos uma análise da série temporal de vazão e turbidez no rio estudado. Esta análise indica que as séries têm período de 362 dias e estão 38 dias fora de fase (vazão atrasada). Além do mais, concluímos que ao usar a vazão e turbidez como entrada do modelo preditivo, este se torna ~ 5 vezes mais simples que o modelo utilizando apenas vazão.

Palavras-Chave— aprendizado de máquinas, redes neurais convolucionais, séries temporais, previsão

Abstract— This work presents a multi-day ahead streamflow forecasting scheme for the Madeira River. The forecasting model is based on unidimensional convolutional neural networks. To design the model we present an analysis of the streamflow and turbidity time series at the studied river. This analysis indicated that they have a period of 362 days and are 38 days out of phase (lagged streamflow). Furthermore, we conclude that the proposed forecasting model is ~ 5 times smaller when using the river's turbidity alongside its streamflow.

Keywords— machine learning, convolutional neural networks, time series, forecasting, streamflow

I. INTRODUÇÃO

Segundo o Balanço Energético Nacional de 2019, em 2018 388.971 GW h dos 601.396 GW h gerados no Brasil foram provenientes de hidroelétricas [1]. Ou seja, 64,68% da geração elétrica brasileira depende dos recursos hídricos disponíveis no país. Neste cenário, podemos destacar a importância do monitoramento dos recursos hídricos do país no âmbito do suprimento energético nacional. Consequentemente, o monitoramento da vazão e sua previsão são atividades que impactam diretamente no planejamento energético nacional e na diminuição dos custos da energia elétrica.

De fato, o planejamento energético pode ser feito com base em simulações que utilizam vazão do rio para determinar quando a utilização de usinas térmicas é necessária e assim traçar melhores estratégias para redução do custo energético. Neste contexto, a previsão de vazão é ainda mais importante em usinas que não dispõem de represas, as chamadas usinas

Felipe Oliveira Barino, Programa de Pós-Graduação em Engenharia Elétrica (PPEE-UFJF), Universidade Federal de Juiz de Fora, Juiz de Fora-MG, e-mail: felipe.barino@engenharia.ufjf.br; Alexandre Bessa dos Santos, Departamento de Circuitos Elétricos e PPEE-UFJF, Universidade Federal de Juiz de Fora, Juiz de Fora-MG, e-mail: alexandre.bessa@engenharia.ufjf.br. Este trabalho foi parcialmente financiado pelo Programa de Bolsas de Pós-Graduação da Universidade Federal de Juiz de Fora (PBPQ-UFJF) através do Programa de Pós-Graduação em Engenharia Elétrica (PPEE-UFJF), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), O Instituto Nacional de Energia Elétrica (Inerge-UFJF) e Santo Antônio Energia.

a fio d'água. Porém o problema de estimar dados futuros de vazão em um rio é estocástico, não-linear e não-estacionário [2]. Além do mais, um modelo matemático para a vazão de um rio depende de um grande número de variáveis, como precipitação, evapotranspiração, temperatura ambiente, características do solo e urbanização [3], portanto abordagens *data-driven* têm ganhado atenção nos últimos anos. Essas abordagens utilizam apenas dados passados, do parâmetro que se deseja modelar, para estimar os valores futuros. Dentre estas técnicas, as que utilizam aprendizado de máquina têm se destacado nos últimos tempos. Redes Neurais Artificiais (RNAs), por exemplo, oferecem fácil implementação e escalabilidade. Além do mais, modelos preditivos baseados em RNAs se mostram equivalentes a técnicas tradicionais para modelagem de séries temporais, como: modelos auto-regressivos de médias móveis (ARMA) e modelo auto-regressivo de médias móveis não lineares (NARMA) [4].

Dentre os trabalhos que fazem uso de RNAs para previsão de vazão em rios, pode-se citar [5], em que os autores apresentam um estudo de previsão da vazão no rio *Pisuena* utilizando dados da precipitação e época do ano. Já em [6] os autores utilizam dados de precipitação em diferentes estações para prever a vazão do rio para até seis horas futuras. Por outro lado, [7] combina precipitação e vazão nas duas últimas horas para estimar vazão em até cinco horas seguintes, além de comparar duas topologias diferentes de RNAs (*Feed Forward Neural Networks* e *Recurrent Neural Networks*).

Com o crescente interesse e aplicação do aprendizado de máquinas e o desenvolvimento do *Deep Learning*, Redes Neurais Convolucionais (CNN, do inglês *Convolutional Neural Network*) também começaram a ser aplicadas na análise e previsão de séries temporais. CNNs são utilizadas na classificação de atividade humana com dados de diversos sensores [8] e também na identificação de tendências do mercado financeiro [9], por exemplo. Além disso, são utilizadas como ferramenta de previsão, como em [10], onde os autores apresentam um esquema utilizando camadas convolucionais juntamente com camadas recorrentes para estimar a geração futura de energia fotovoltaica. Também no contexto de energia elétrica, [11] desenvolve um modelo preditivo para estimativa da energia elétrica consumida em uma residência, comparando a proposta com outros modelos preditivos e mostrando que as CNNs são uma alternativa viável para previsão de séries temporais.

Este trabalho apresenta o desenvolvimento de um sistema para prever a vazão do Rio Madeira vários dias à frente, nas proximidades da Usina Hidroelétrica Santo Antônio, em Porto Velho, RO. Por se tratar de uma usina a fio d'água, a previsão da vazão é diretamente relacionada à previsão da energia

gerada e é, portanto, de grande valor para o planejamento energético. Ao longo do trabalho nós propomos dois modelos preditivos: CNNI que utiliza os dados passados da vazão para estimar o valor futuro, e CNNII que utiliza também dados passados da turbidez do rio, aproveitando a correlação entre esses sinais. Para o desenvolvimento destes modelos, nós apresentamos uma análise das séries temporais de turbidez e vazão, análise esta feita por meio da transformada de Fourier dos sinais e da correlação entre eles.

II. MÉTODOS

A. Análise das séries temporais

As séries temporais de vazão e turbidez, cobrindo o intervalo de 01/01/2011 a 04/12/2018, com amostragem diária, tiveram os valores faltantes preenchidos por meio de interpolação linear e os dados foram separados no tempo em três grupos: treino (45%), teste (25%) e validação (30%), isto é, os primeiros 45% pontos da série temporal foram destinados ao grupo de treino, os 30% seguintes ao grupo de validação e os 25% restantes ao grupo de teste. Note que a separação no tempo foi realizada para simular uma aplicação real, ou seja, nem um ponto futuro é conhecido pelo modelo.

A componente fundamental foi extraída por meio da análise de Fourier, com o objetivo de se obter a periodicidade dos sinais. Sabe-se que durante eventos hidrológicos, vazão e concentração de sedimentos possuem certa relação, [12], portanto é de se esperar que a componente fundamental observada na transformada discreta de Fourier dos sinais seja igual.

De fato, a análise qualitativa entre essas grandezas sugere que as séries possuem comportamento semelhante, porém com picos de turbidez ocorrendo antes dos picos de vazão, conforme mostra a Fig. 1. Nesta figura, $t[n]$ representa o sinal de vazão e $d[n]$ o sinal de turbidez. Analogamente, a Fig. 2 mostra a transformada discreta de Fourier dos sinais, representados por $T[f]$ e $D[f]$ no domínio da frequência.

Note que, se os sinais possuem mesma frequência fundamental deve existir uma diferença de fase nestes sinais, já que os picos de turbidez ocorrem antes. Portanto, existe um intervalo de atraso k no sinal $t[n]$ tal que a correlação entre $d[n]$ e $t[n - k]$ é máxima. Portanto, calcula-se a correlação

entre $d[n]$ e $t[n - k]$ para diferentes valores de k , variando entre 0 e 90 dias.

A partir dos resultados da análise de correlação destas séries, pode-se determinar qual o número de amostras passadas que será utilizado como entrada do modelo preditivo multiparâmetro para vazão. Tal modelo trata-se de uma Rede Neural Convolucional unidimensional (1D-CNN), que consiste em uma rede neural *multilayer perceptron* precedida por uma camada convolucional para extração de parâmetros.

B. Redes Neurais

RNAs são modelos matemáticos baseados no sistema nervoso de animais que são capazes de mapear dados de entrada em dados de saída com base em um conhecimento adquirido durante uma fase de treinamento desse modelo. Consiste em uma combinação de pequenas estruturas, chamadas neurônios, interconectadas em camadas. Essas camadas são, também, conectadas entre si, formando uma rede. Os neurônios artificiais de uma rede neural recebem valores de entrada x_i , processam essa informação e geram uma saída. Esse processamento é feito por meio da soma ponderada das entradas por coeficientes w_i (pesos sinápticos) e um termo independente b (*bias*). O resultado dessa soma é aplicado a uma função $\Phi(\cdot)$ (função de ativação) e o resultado é, por sua vez, a saída do neurônio.

A capacidade de mapear dados de entrada em dados de saída faz com que as RNAs possam realizar a regressão de qualquer função contínua, sendo consideradas aproximações universais [13], [14]. Essas estruturas apresentam algumas características importantes para análise de séries temporais, como: robustez à ruído, não-linearidade e a capacidade de lidar com múltiplas entradas e saídas, por exemplo. RNAs vêm sido utilizadas como ferramentas para análise e previsão de sinais no tempo desde a década de noventa [15].

C. Deep Learning

Deep learning, aprendizado profundo, é o nome dado a uma classe de algoritmos de aprendizado de máquinas que utilizam várias camadas de processamento. Geralmente as primeiras camadas são utilizadas para extração de parâmetros,

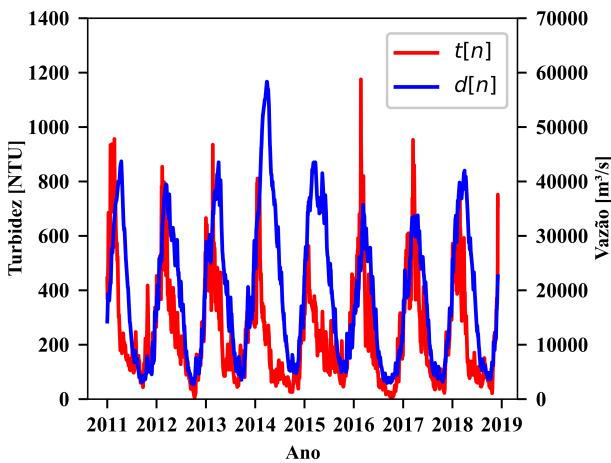


Fig. 1. Séries temporais de turbidez e vazão no período de estudo.

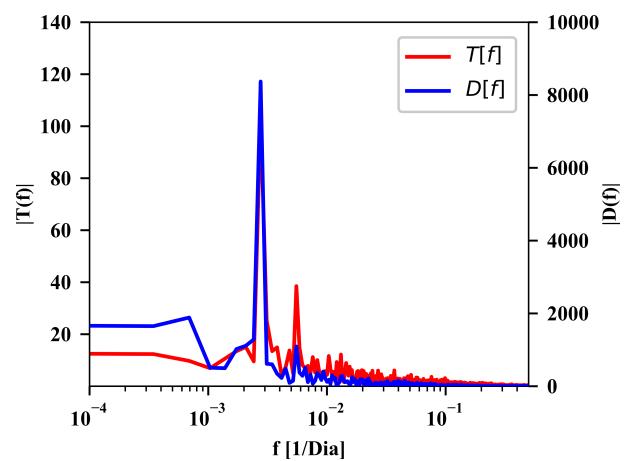


Fig. 2. Amplitude da DFT dos sinais de vazão e turbidez.

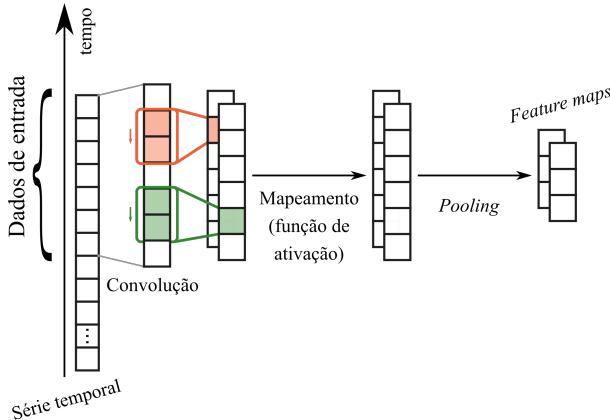


Fig. 3. Esquema de uma camada convolucional unidimensional.

processo que é normalmente feito pelo projetista em algoritmos de aprendizado rasos. Dentre esses algoritmos, pode-se destacar as CNNs. Elas foram originalmente criadas para tarefas envolvendo imagens [16], dados em duas dimensões, e são basicamente CNNs alimentadas por dados pré-processados através de convoluções na imagem de entrada por certos filtros (comumente chamados de *kernels*).

Em suma, uma camada convolucional funciona da seguinte maneira: dada uma matriz de entrada, faz-se a convolução desta por matrizes de menor ordem (os *kernels*). O resultado dessa operação tem alguns valores selecionados em matrizes de menor ordem, por um processo conhecido como *pooling*. Dessa forma, a informação contida na matriz de entrada é filtrada e concentrada em uma matriz de menor ordem, chamadas de *feature maps* (mapas de parâmetros). Os *feature maps* podem então ser entrada de outra camada convolucional, se existirem mais camadas convolucionais, ou de uma rede neural.

Na análise de séries temporais, os dados são em uma dimensão, portanto pode-se substituir a convolução bidimensional realizada na camada convolucional por uma convolução unidimensional, operando com vetores ao invés de matrizes. O processo de convolução unidimensional para séries temporais está ilustrado na Fig. 3, note que a entrada da CNN é somente a série temporal e sem pré-processamento. Dessa forma, uma camada convolucional unidimensional é utilizada para extrair informações da série temporal por meio dos filtros e utilizar os dados filtrados como entrada de um *Multilayer Perceptron*, que tem em sua saída a previsão desejada. Como o operador convolução é aplicado em subsequências da série de entrada, um padrão aprendido em uma localização pode também ser identificado em outras posições [17].

Para o desenvolvimento do modelo preditivo deste trabalho, a série temporal foi partida em sub-séries, separando dados de entrada e saída por meio de janelas deslizantes. Esse processo consistiu em criar uma janela com $N_{hist} + N_{prev}$ dados da série temporal, separando as primeiras N_{hist} amostras desta janela como dados de entrada e o último dado N_{prev} como saída. Essa janela, então, é deslizada por todas as amostras disponíveis, criando conjuntos entrada e saída em toda a extensão da série temporal. O número de dias a frente a ser

previsto, N_{prev} , e o total de dados de entrada, N_{hist} , neste trabalho foram determinados por meio da análise discutida na Seção II-A.

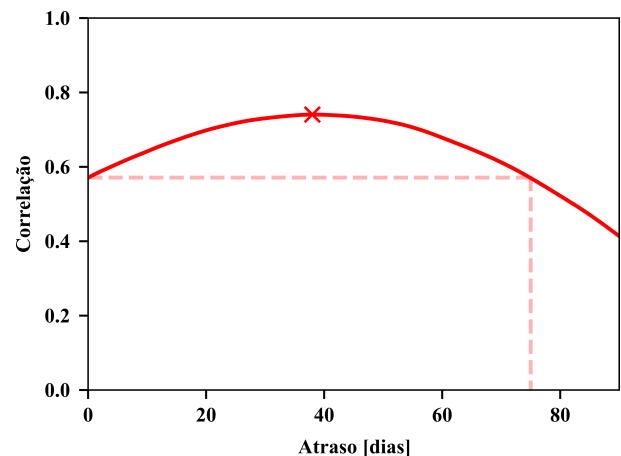
D. Modelos propostos

Com relação à topologia dos modelos desenvolvidos, ambos são 1D-CCNs de uma única camada convolucional, porém com diferentes dados de entrada. Um deles é composto apenas pela série temporal de vazão, enquanto o outro também tem a série temporal de turbidez como dado de entrada. Damos ao primeiro modelo o nome de CNNI (apenas vazão) e ao segundo de CNNII (vazão e turbidez). Para ambos, nós realizamos uma seleção de parâmetros com busca exaustiva no seguinte sub-espaco de hiper-parâmetros: $M \in \{6, 8, 12\}$ filtros de tamanho $P \in \{2, 4, 8\}$, com *pooling* de tamanho $Q \in \{1, 2, 4\}$ e camada escondida com $N_h \in \{10, 30, 50\}$ neurônios. A combinação destes parâmetros totaliza 324 modelo, que foram treinados cinco vezes com o método ADAM [18] para minimizar o erro quadrático médio da estimativa. Por fim, selecionamos os melhores conjuntos de parâmetro para cada modelo, considerando o erro absoluto médio da estimativa e sua variação ao longo dos cinco testes.

III. RESULTADOS

Analizando os sinais na frequência, pelos gráficos mostrados na Fig. 2, obtém-se que ambos os sinais possuem a mesma componente fundamental, com frequência $f_0 = 0,0028 \text{ dia}^{-1}$. Logo o período dos sinais é de $T \sim 362$ dias. Realizando o procedimento de correlacionar os sinais deslocados no tempo, obtém-se que a correlação aumenta até um pico em $k = 38$ dias de atraso e cai para o valor inicial quando $k = 75$ dias. Este resultado pode ser visto na Fig. 4, com o ponto de maior correlação marcado e uma reta indicando o período em que a correlação volta ao patamar inicial.

Logo, escolhe-se 75 dias como intervalo de entrada em ambos os modelos de previsão, sendo esse o tamanho da camada de entrada da CNN e 38 dias como intervalo de previsão futuro.

Fig. 4. Coeficiente de correlação entre os sinais $d[n]$ e $t[n - k]$ para k no intervalo de 0 a 90 dias

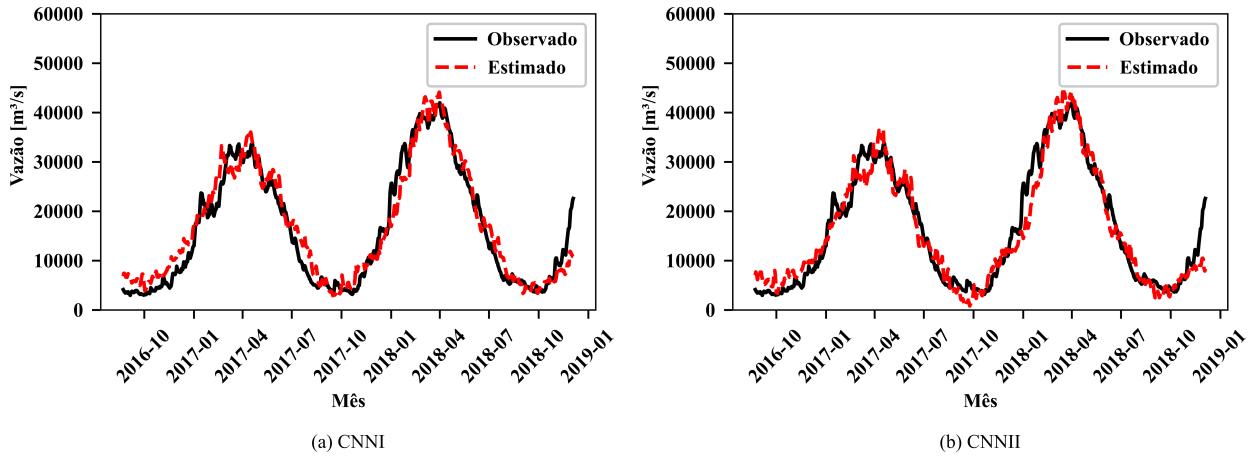


Fig. 5. Comparação entre os dados observados e a previsão do 38º dia a frente para o modelo (a) CNNI e (b) CNNII.

TABELA I
PARÂMETROS DOS MODELOS.

| Modelo | M | P | Q | N_h |
|--------|-----|-----|-----|-------|
| CNNI | 12 | 2 | 1 | 50 |
| CNNII | 8 | 4 | 2 | 30 |

Já em relação à seleção do modelo, os valores obtidos para os parâmetros M , P , Q e N_h são mostrados na Tabela I. Nós escolhemos a configuração com o melhor erro absoluto médio percentual (MAPE) que obteve a menor variância dentre os 5 testes realizados, para CNNI o MAPE médio para os parâmetros escolhidos é de 22,6%, enquanto para o CNNII ele é igual a 25,6%, no conjunto de validação.

Note que o modelo CNNI tem um maior número de parâmetros, com mais filtros e neurônios escondidos, além de ter *pooling* unitário, que é equivalente a nenhuma camada de *pooling*, aumentando o número de parâmetros dos *feature maps*. Sendo assim, o modelo CNNI tem um número de parâmetros treináveis muito maior que CNNII, de forma que sua variância é maior e, portanto, o resultado obtido no treinamento pode não ser ótimo ou até mesmo confiável. Ou seja, pode-se dizer que, além de ter um custo computacional menor, o modelo CNNII é mais confiável, pois utiliza o mesmo número de dados para estimar um menor número de parâmetros.

A performance dos modelos mostrados na Tabela I, na última partição do banco de dados (teste) é: 25,2% de erro absoluto médio percentual para CNNI e 22,8% para CNNII. Nota-se que ambos os valores são extremamente próximos e, considerando que se está prevendo um valor para 38 dias a frente, são pequenos. E, de fato, ambos os modelos explicam bem os dados observados, já que o coeficiente de determinação para ambos os modelos foi estimado em 93%. Além do mais, pode-se notar na Fig. 5 que os resultados para previsão da vazão por ambos os modelos são visivelmente semelhantes.

Apesar do desempenho dos modelos ser similar, o modelo que só utiliza a vazão (CNNI) apresenta um número muito maior de parâmetros e, conforme discutido, pode ser menos confiável. Na verdade a relação entre o número de parâmetros

é 4,94 vezes mais parâmetros treináveis em CNNI do que em CNNII. Sendo assim, apesar de os valores obtidos neste trabalho serem muito próximos, podemos dizer que o uso da série temporal de turbidez juntamente com a vazão facilita a previsão da vazão no Rio Madeira, pois necessita de um número muito menor de parâmetros para obter a mesma qualidade na estimativa. E, de fato, é um resultado esperado, já que as séries temporais são extremamente correlacionadas, especialmente quando a vazão está atrasada. Ou seja, a turbidez no Rio Madeira tem grande parcela na previsão da vazão neste rio.

IV. CONCLUSÃO

Nós apresentamos um breve estudo sobre as propriedades das séries temporais de turbidez e vazão no Rio Madeira, o maior tributário do Amazonas, e concluímos que essas séries possuem o mesmo período, $T \sim 362$ dias, e estão defasadas, em média 38 dias (vazão atrasada). Com essas informações nós projetamos dois modelos preditivos para estimar a vazão 38 dias no futuro, ambos utilizando CNN unidimensional.

Analisando os modelos propostos e desenvolvidos, nós concluímos que é possível estimar a vazão do Rio Madeira com mais de um mês e meio de antecedência (38 dias) com erro absoluto de $\sim 23\%$, além do mais, mostramos que ao usar a série temporal de turbidez juntamente com a de vazão, neste cenário, o modelo preditivo se torna ~ 5 vezes menor e, portanto, mais confiável.

REFERÊNCIAS

- [1] EPE, Empresa Brasileira de Pesquisa Energética (2019), Balanço Energético Nacional 2019: Ano base 2018. Rio de Janeiro: EPE.
 - [2] R. R. Aquino, M. M. Lira, M. H. Marinho, I. A. Tavares, & L. F. Cordeiro (2010). "Inflow forecasting models based on artificial intelligence". *The 2010 International Joint Conference on Neural Networks (IJCNN)*, Barcelona, 2010, 1-6.
 - [3] C. Sudheer, R. Maheswaran, B. Panigrahi & S. Mathur, "A hybrid SVM-PSO model for forecasting monthly streamflow," *Neural Computing and Applications*, vol. 24, no. 6, pp. 1381-1389, 2013
 - [4] H. Maier & G. Dandy, "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications," *Environmental Modelling & Software*, vol. 15, no. 1, pp. 101-124, 2000
 - [5] J. L. Crespo, & E. Mora (1993). "Drought estimation with neural networks." *Advances in Engineering Software*, 18(3), 167-170.

- [6] R. Golob, T. Štokelj, & D. Grgić (1998). “Neural-network-based water inflow forecasting.” *Control engineering practice*, 6(5), 593-600.
- [7] M. Aqil, I. Kita, A. Yano, & S. Nishiyama (2007). “Neural networks for real time catchment flow modeling and prediction.” *Water Resources Management*, 21(10), 1781-1796.
- [8] J. Yang, M. N. Nguyen, P. P. San, X. L. Li, & S. Krishnaswamy (2015). “Deep convolutional neural networks on multichannel time series for human activity recognition.” em *Twenty-Fourth International Joint Conference on Artificial Intelligence*.
- [9] J. F. Chen, W. L. Chen, C. P. Huang, S. H. Huang, & A. P. Chen (2016). “Financial time-series data analysis using deep convolutional neural networks.” *2016 7th International Conference on Cloud Computing and Big Data (CCBD)*, 87-92.
- [10] W. Lee, K. Kim, J. Park, J. Kim, & Y. Kim (2018). “Forecasting Solar Power Using Long-Short Term Memory and Convolutional Neural Networks.” *IEEE Access*, 6, 73068-73080.
- [11] K. Amarasinghe, D. L. Marino, & M. Manic (2017). “Deep neural networks for energy load forecasting.” *2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*, 1483-1488.
- [12] G. P. Williams (1989). “Sediment concentration versus water discharge during single hydrologic events in rivers.” *Journal of Hydrology*, 111(1-4), 89-106.
- [13] J.-G. Cordeiro and G. Pagès (1997). “Approximations of functions by a multilayer perceptron: a new approach”, *Neural networks* 10(6): 1069-1081.
- [14] J. L. Castro, C. J. Mantas, & J. M. Benítez (2000). “Neural networks with a continuous squashing function in the output are universal approximators.” *Neural Networks*, 13(6), 561-563.
- [15] G. Dorffner (1996). “Neural networks for time series processing.” em *Neural network world*, 6, 447-468
- [16] Y. LeCun, et. al. (1990). “Handwritten digit recognition with a back-propagation network.” em *Advances in neural information processing systems* (pp. 396-404).
- [17] Q. Tao, F. Liu, Y. Li, & D. Sidorov (2019). “Air Pollution Forecasting Using a Deep Learning Model Based on 1D Convnets and Bidirectional GRU.” *IEEE Access*, 7, 76690-76698.
- [18] D. P. Kingma, & J. Ba (2014). “Adam: A method for stochastic optimization”. *3rd International Conference for Learning Representations*, (arXiv:1412.6980).