

Identificação de índices de precipitação para previsão mensal de vazão no posto de Tucuruí – bacia do Tocantins

Identification of rainfall indexes for monthly streamflow forecast in the Tucuruí station – Tocantins basin

DOI:10.34117/bjdv8n12-118

Recebimento dos originais: 04/11/2022

Aceitação para publicação: 12/12/2022

Amanda Rodrigues de Souza

Mestranda em Ciência e Tecnologia Ambiental pela Universidade Federal do ABC (UFABC)

Instituição: Universidade Federal do ABC (UFABC)

Endereço: Avenida dos Estados, 5001, Bangu, Santo André - SP, CEP: 08210-580

E-mail: amanda.souza@ufabc.edu.br

Andrea de Oliveira Cardoso

Doutora em Meteorologia pela Universidade de São Paulo (USP)

Instituição: Universidade Federal do ABC (UFABC)

Endereço: Avenida dos Estados, 5001, Bangu, Santo André - SP, CEP: 08210-580

E-mail: andrea.cardoso@ufabc.edu.br

Willians Bini Silva

Mestre em Meteorologia pela Universidade de São Paulo (USP)

Instituição: Climatempo

Endereço: Av. Queiroz Filho, 1700, Vila Hamburguesa, São Paulo - SP,

CEP: 05319-000

E-mail: willians.bini@climatempo.com.br

RESUMO

Devido a relevância da Bacia Hidrográfica do Rio Tocantins para o abastecimento, geração de energia e desenvolvimento no entorno da bacia é importante o aprimoramento da predição de vazão. Neste sentido, este estudo explora a relações chuva-vazão, visando identificar índices de precipitação úteis na previsão de vazão mensal no posto de Tucuruí. Para tanto, foram aplicadas diversas análises aos valores mensais de vazão natural e precipitação de 1980 a 2020, que permitiram: caracterizar seis regiões pluviométricas homogêneas, através de análise de *cluster*; comparar os padrões sazonais de chuva e vazão, por médias de longo prazo; classificar os tipos de eventos de vazão, destacando os extremos, através de quantis; relacionar índices de precipitação anual móvel de cada região com a vazão mensal, por correlação simples; e estimar vazão mensal, por modelos de regressão linear múltipla, usando índices de precipitação como preditores. Os resultados destacaram os índices de acumulado de chuva e de limiares quantílicos secos e chuvosos como os melhores preditores. Os modelos indicam a potencialidade dos índices de precipitação para a previsibilidade de vazão mensal, destacando a precipitação ocorrida no extremo leste e próximo ao posto de Tucuruí para a predição da vazão.

Palavras-chave: índices climáticos, relação chuva-vazão, previsão, UHE Tucuruí, R de Pearson.

ABSTRACT

Due to the relevance of the Tocantins River Basin for supply, energy generation and development around the basin, it is important to improve the streamflow prediction. In this sense, this study explores the rainfall-streamflow relationships, aiming to identify precipitation indexes that are useful in predicting monthly streamflow at the Tucuruí station. Therefore, several analyzes were applied to the monthly values of natural streamflow and precipitation from 1980 to 2020, which allowed: to characterize six homogeneous rainfall regions, through cluster analysis; to compare seasonal patterns of rainfall and streamflow, by long-term averages; classify the types of streamflow events, highlighting the extremes, through quantiles; to relate indexes of movable annual precipitation of each region with the monthly streamflow, by simple correlation; and estimate monthly streamflow, by multiple linear regression models, using precipitation indexes as predictors. The results highlighted the indexes of accumulated rainfall and dry and wet quantile thresholds as the best predictors. The models indicate the potential of precipitation indices for the predictability of monthly streamflow, highlighting the precipitation occurring in the extreme east and near to the Tucuruí station for the prediction of streamflow.

Keywords: climate indices, rain-streamflow relation, forecast, Tucuruí HPP, Pearson correlation.

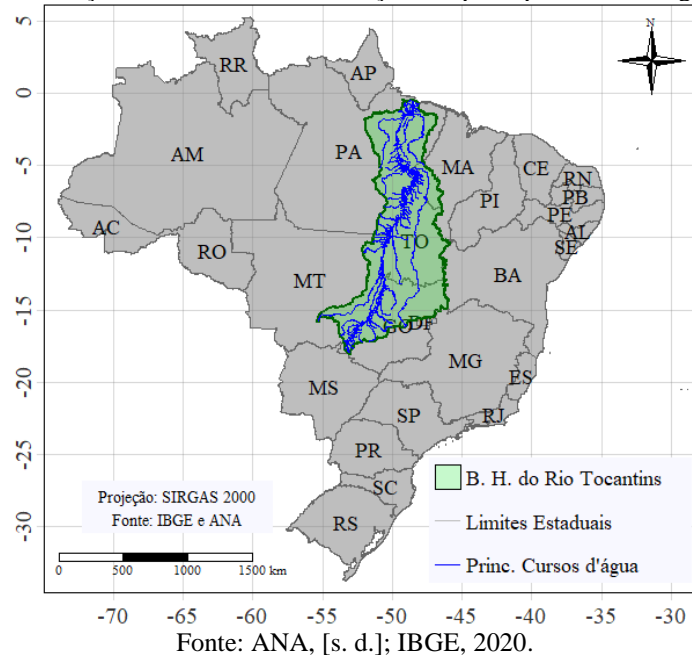
1 INTRODUÇÃO

A variabilidade da vazão de um rio depende das características climáticas e físicas da região de sua bacia. A precipitação contribui para a ocorrência de escoamento superficial, infiltração, recarga dos aquíferos e conseqüentemente contribui para vazão (TUCCI, 2001). Assim, dependendo da região, a precipitação pode afetar a vazão de um rio tanto em curto, como em longo prazo. Os regimes de chuva estão associados aos sistemas atmosféricos atuantes, que variam espaço e temporalmente (REBOITA et al., 2010).

O Brasil apresenta vazão média anual de 260 m³/s, 12% da disponibilidade mundial (ANA, 2018). Porém observa-se grande variabilidade espaço-temporal, sendo que tais variabilidades apresentam fortes relações com alguns fenômenos climáticos (CAPOZZOLI et al., 2017).

A Bacia Hidrográfica do Rio Tocantins (BHRT) corresponde a cerca de 10% do território nacional, com 920 mil km² de área e 1.960 km de extensão, abrangendo os estados do Goiás, Distrito Federal, Mato Grosso, Tocantins, Maranhão e Pará (foz) (Figura 1). Em parâmetros nacionais, a bacia destaca-se por expansão da fronteira agrícola e o potencial hidroenergético (ANA, 2015).

Figura 1. Localização da BHRT com a indicação dos principais cursos d'água da bacia.



Estão instaladas sete usinas hidroelétricas (UHE) de grande porte na BHRT. Além da UHE Tucuruí, UHE Serra da Mesa também possui capacidade de regularização das vazões (GHISLENI et al., 2022). As UHE atendem tanto o Mercado Cativo quanto o Mercado Livre, aumentando o interesse industrial na região (PENHA; BARROS, 2021), bem como o desenvolvimento socioeconômico, que apresentou aumento no Índice FIRJAN de Desenvolvimento Municipal na parte da bacia pertencente ao MATOPIBA (Maranhão, Tocantins, Piauí e Bahia) entre o período de 2005 a 2016 (CRUZ et al., 2019).

A retirada de água desassociada aos regimes de vazão podem provocar crise de abastecimento, como ocorrido na sub-bacia do Formoso, na BHRT, no final da década de 2010 (VOLKEN et al., 2022).

A precipitação sobre a BHRT apresenta grande variabilidade anual. Os maiores acumulados ocorrem próximo a foz, no estado do Pará e os menores no estado de Goiás e na porção oeste do estado de Tocantins (ALVARES et al., 2013). De acordo com Loureiro e Fernandes (2013), o volume de precipitação sobre a BHRT se manteve aproximadamente constante para o período de 1983 a 2003 porém, identificou-se variação na distribuição espacial, o que afetou a vazão no rio Tocantins. Neste contexto, o conhecimento dos padrões de chuva, suas variações e influência sobre a vazão, podem dar subsídios para um melhor planejamento e uso do recurso hídrico da região.

A utilização de índices de precipitação pode fornecer uma descrição abrangente da variação de extremos de precipitação (KARL et al., 1999), podendo ser útil para

avaliação da influência da precipitação na vazão do rio em diferentes escalas, bem como parâmetros para a predição da vazão, como destacado por Cardoso *et al.* (2022).

Este estudo tem como objetivo caracterizar os padrões e os extremos de vazão em Tucuruí, associá-los com precipitação e identificar índices de precipitação que possam ser preditores de vazão mensal em Tucuruí, sendo esse tipo de predição útil para o bom manejo dos recursos hídricos, principalmente para a geração de energia hidrelétrica.

2 METODOLOGIA

Para a realização do estudo foram utilizados dados de vazão natural mensal do posto de Tucuruí e precipitação mensal e diária de estações pluviométricas presentes na BHRT. A vazão natural é definida como a vazão que ocorreria em um corpo hídrico caso não houvesse ação antrópica a montante da bacia.

Os dados de vazão de 1980 a 2019 foram coletados do Operador Nacional de Sistema Elétrico (ONS, 2020) e, para completar a série até dezembro de 2020, foram utilizados os dados do Sistema de Acompanhamento de Reservatórios da Agência Nacional das Águas (ANA) (ANA, [s. d.]), totalizando 41 anos de dados, de 1980 a 2020.

Os dados históricos de precipitação diárias foram coletados da ANA, a partir do portal HIDROWEB. Utilizando as coordenadas geográficas das estações, os limites da BHRT (ANA, 2016), cursos d'água da bacia (ANA, 2018) e o *software* QGIS (QGIS, [s. d.]), foi realizada a visualização espacial dos elementos citados. Desse modo, foram selecionadas 149 estações pluviométricas localizadas a montante do ponto de vazão, i. e., estações que contribuem para a vazão no posto de Tucuruí.

Com o intuito de analisar e resumir os dados de precipitação, foi realizado o agrupamento dos dados. A regionalização foi feita através da análise de *cluster* (WILKS, 2011). Como informações de entrada, foram utilizados os valores das médias e quantis mensais de 15%, 35%, 65% e 85% dos dados de precipitação de cada estação.

Dada uma determinada série de dados P_i , com i variando de 1 a n observações, é necessário determinar o número de ordem $w(P_i)$ de cada elemento i desta série, com observações em ordem crescente (BUSSAD; MORETTIN, 2013). Então, calcula-se os quantis a partir da equação 1

$$q(P_i) = \frac{1}{2 \cdot n} + \frac{[w(P_i) - 1]}{n} \quad (1)$$

Na análise de *cluster*, para definir a quantidade de grupos, foi avaliada a distância euclidiana entre eles, identificando o ponto de parada do agrupamento pelo salto na distância (diferença) entre a formação de um grupo e o anterior. Para o refinamento da regionalização foi considerada a disposição geográfica das estações de cada grupo.

Para garantir a qualidade dos dados de precipitação, foram removidos dados duvidosos das séries históricas, os quais foram identificados como *outliers* no diagrama *boxplot*, com separação mensal dos dados, e comparados com a precipitação observada em estações vizinhas, pertencentes a mesma região homogênea. Neste processo, cada *outlier* foi avaliado e verificado se é maior (menor) do que três desvios padrões mais (menos) a média da série, sendo classificado como um evento duvidoso. Caso não tenha ocorrido eventos semelhantes em pontos vizinhos da mesma região homogênea de precipitação, o dado foi retirado da série.

O *boxplot* é constituído por seis elementos: três quartis, os limites inferiores e superiores e os *outliers*. O quartil inferior corresponde ao limite de 25% dos dados, o superior corresponde ao limite de 75% dos dados e a mediana (quartil de 50%). Os limites inferiores e superiores são definidos por 1,5 vezes a diferença entre o terceiro quartil e o primeiro. Valores que excedem essa marca, são classificados como um *outliers*, ou seja, um valor discrepante (DEVORE, 2006).

Para a caracterização dos padrões e variabilidades da precipitação e vazão, foram calculadas as médias mensais dos dados. Para a vazão também foi realizada a classificação dos eventos de vazão de acordo com os limiares quantílicos, seguindo metodologia realizada por outros autores, como Dionisio *et al.*, (2021, 2022), Tori *et al.*, (2021) e Souza e Cardoso (2021).

A separação das categorias assumiram o aplicado por Xavier, Xavier *et al.* (2003), adaptado ao caso de vazão, conforme a Tabela 1:

Tabela 1. Classificação dos intervalos quantílicos para a vazão.

Classificação	Intervalos Quantílicos
Muito Baixa	$q(X_i) \leq 0,15$
Baixa	$0,15 < q(X_i) \leq 0,35$
Normal	$0,35 < q(X_i) \leq 0,65$
Alta	$0,65 < q(X_i) \leq 0,85$
Muito Alta	$q(X_i) > 0,85$

Fonte: Adaptado de Xavier, Xavier *et al.* (2003).

Com o objetivo de relacionar a vazão mensal com a precipitação anual, foram calculados índices de precipitação anuais móveis baseados nas séries diárias das regiões homogêneas de precipitação, visando enfatizar padrões secos e chuvosos.

Para a obtenção de cada valor da série dos índices considerou-se doze meses móveis de dados, com passos mensais (i. e. dados de janeiro a dezembro, de fevereiro a janeiro e assim por diante). Na construção dos índices anuais foi admitido até 10% de dados faltantes. Na Tabela 2 estão expostos os índices de precipitação utilizados, a sigla e sua fonte. Os índices ND15q-, ND35q-, ND65q+ e ND85q+ foram construídos para o estudo. Para os índices que utilizam os limiares quantílicos foram utilizados apenas os dias com precipitação.

Tabela 2. Índices de precipitação, siglas e fonte

Sigla	Nome	Fonte
CDD-MAX e CDD-MED	Máximo e Média de dias Consecutivos Secos	Peterson, <i>et al.</i> (2001)
CWD-MAX e CWD-MED	Máximo e Média de dias Consecutivos Úmidos	Peterson, <i>et al.</i> (2001)
PRCTOT	Acumulado da Precipitação em dias Úmidos	Zhang; ETCCDI (2009)
SDII	Índice Simples de Intensidade Diária	Peterson, <i>et al.</i> (2001)
ND15q- e ND35q-	Número de dias com precipitação abaixo do limiar de 15% e 35%	
ND65q+ e ND85q+	Número de dias com precipitação acima do limiar de 65% e 85%	

Fonte: Elaborado pelos autores.

Para avaliar a relação chuva-vazão, foram calculadas correlações mensal, “sem” e “com” defasagem (precipitação precedendo a vazão), entre cada um dos índices de precipitação e a série de vazão mensal de Tucuruí, através de Análise de Correlação Simples (WILKS, 2011). Nesta análise, o grau de relacionamento entre as variáveis contínuas é sintetizado por um coeficiente de correlação conhecido como “r de Pearson”, conforme:

$$r_{xy} = \frac{\left[\frac{1}{(n-1)} \right] \sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\left\{ \left[\frac{1}{(n-1)} \right] \sum_{i=1}^n (x_i - \bar{x})^2 \right\}^{1/2} \left\{ \left[\frac{1}{(n-1)} \right] \sum_{i=1}^n (y_i - \bar{y})^2 \right\}^{1/2}} \quad (2)$$

Os resultados da correlação variam de -1 a +1, onde +1 demonstra correlação direta perfeita, -1 mostra uma correlação oposta perfeita (linear) e zero indica ausência de correlação (WILKS, 2011). Visando minimizar possíveis erros nas correlações

causados pela presença de dados faltantes, foram realizadas as correlações apenas de séries com mais de trinta anos. Para tanto, utilizou-se a linguagem de programação R (ANACONDA INC., 2021) e o pacote ‘*tidyverse*’ (WICKHAM, 2017).

Para avaliar a significância da correlação aplicou-se o Teste t de *Student* (WILKS, 2011). Para rejeitar ou aceitar uma hipótese nula, foi avaliado se o nível de significância dos resultados era maior que 0,05 (ou seja, confiança de 95%).

Foram avaliados diversos índices. Porém, entre eles há índices semelhantes por definição, sendo eles: (a) CWD – MAX e CWD – MED; (b) CDD – MAX e CDD – MED; (c) PRCPTOT e SDII; (d) ND35q- e ND15q-; (e) ND65q+ e ND85q+.

Para selecionar os melhores índices por região, foram utilizados critérios de eliminação, como a amplitude da correlação e a quantidade de defasagens significativas de cada índice. Para os índices de contagem de dias por limiares (itens *d*, e *e*), em caso de empate nos critérios, foi selecionado o que representa o cenário mais crítico.

A Regressão Linear Múltipla (RLM) (WILKS, 2011) foi utilizada neste estudo para identificar o principal conjunto de índices de precipitação que influenciam a vazão e que podem contribuir na previsibilidade da vazão padronizada. A própria vazão defasada foi utilizada como um dos parâmetros de entrada no modelo, devido a sua importância como condição inicial na previsão da vazão. Além da vazão, foram utilizados como preditores os índices de precipitação selecionados pelo critério de seleção, que apresentaram correlação significativa com a vazão na defasagem em que está sendo construída o modelo. A equação da RLM é definida como:

$$y = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3 + \dots + b_k \cdot X_k, \quad (3)$$

sendo *y* a variável dependente; X_n as variáveis independentes; b_0 a intersecção da reta com o eixo de *y*; b_n os coeficientes angulares obtidos da regressão; *k* o número de variáveis independentes (preditores do modelo).

Para tanto, o conjunto de dados foi separado em dois períodos, sendo um para calibração e outro para a validação. Na calibração foi utilizado o período de janeiro de 1980 a dezembro de 2010, para o ajuste das equações de regressão e obtenção do coeficiente de determinação. O coeficiente de determinação (R^2) é definido pela equação (4) (WILKS, 2011), onde *n* é o número de observações, \hat{y} são os valores estimados de *y* e \bar{y} a média da variável *y*.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (\hat{y} - \bar{y})^2} \quad (4)$$

Foi construído um modelo de RLM para cada defasagem analisada. Primeiramente, foi realizada a filtragem dos preditores (os índices de precipitação e a vazão defasada) que apresentaram multicolinearidade com os demais selecionados para a defasagem considerada e posteriormente, a seleção do conjunto de preditores para compor o modelo.

Um modo de análise da multicolinearidade de preditores é através do Fator de Inflação da Variação (VIF – *Variance Inflation Factor*). Na literatura, o VIF (equação 5) tem sido um método comum de identificação de multicolinearidade.

$$VIF_i = \frac{1}{1 - R_i^2} \quad (5)$$

O termo R_i^2 representa o coeficiente de determinação (Equação 4) de cada um dos preditores com as demais variáveis explicativas (DRAPER; SMITH, 1998). De acordo com Draper e Smith (1998), os valores de VIF maiores que 10 podem causar problemas na estimação dos coeficientes de regressão devido a problemas de multicolinearidades. Portanto, foram retirados os índices que apresentaram $VIF \geq 10$. No presente estudo, foi utilizada a função ‘*vif()*’ (NAIMI et al., 2014).

Para a seleção dos melhores índices de precipitação para a previsão de vazão, foi utilizada a função *regsubsets()* presente no pacote ‘*leaps*’ (LUMLEY, 2020) na linguagem de programação R (R CORE TEAM, 2019), que realiza uma busca pelos melhores subconjuntos de preditores.

Para a seleção do melhor modelo entre aqueles indicados pela função, utilizou-se o Critério de Informação Bayesiano (BIC – *Bayesian Information Criterion*), o menor valor de BIC indica o melhor modelo. Sendo BIC definido pela equação a seguir:

$$BIC = n \log(RSS_{p_c} / n) + \log(n)p_c \quad (6)$$

Onde RSS é a soma do quadrado dos resíduos, p_c é o número de regressões e n o tamanho da amostra (WEISBERG, 2014).

Nesta modelagem, como parâmetros de entrada, tem-se a série histórica de vazão (y), os preditores de índices de precipitação e a própria vazão (x). Os valores são inseridos em defasagem, com os preditores precedendo a vazão a ser prevista.

Além da calibração, é necessário avaliar as estimativas para os dados não considerados na calibração, visando assim reproduzir a realidade e avaliar a qualidade do modelo. Para tanto, foi necessário usar parte do conjunto dos dados para a realização da validação, que corresponde ao período de janeiro de 2011 a dezembro de 2020. Para tanto, foi calculado o Erro Viés (Equação 7), a Raiz do Erro Quadrático Médio (Equação 8), o Erro Absoluto (Equação 9), o Erro Relativo (Equação 10) (WILKS, 2011).

$$VIÉS = \frac{1}{N} \sum (F - O) \quad (7)$$

$$REQM = \frac{1}{n} \sqrt{\sum (F - O)^2} \quad (8)$$

$$EA = \frac{1}{n} \sum |F - O| \quad (9)$$

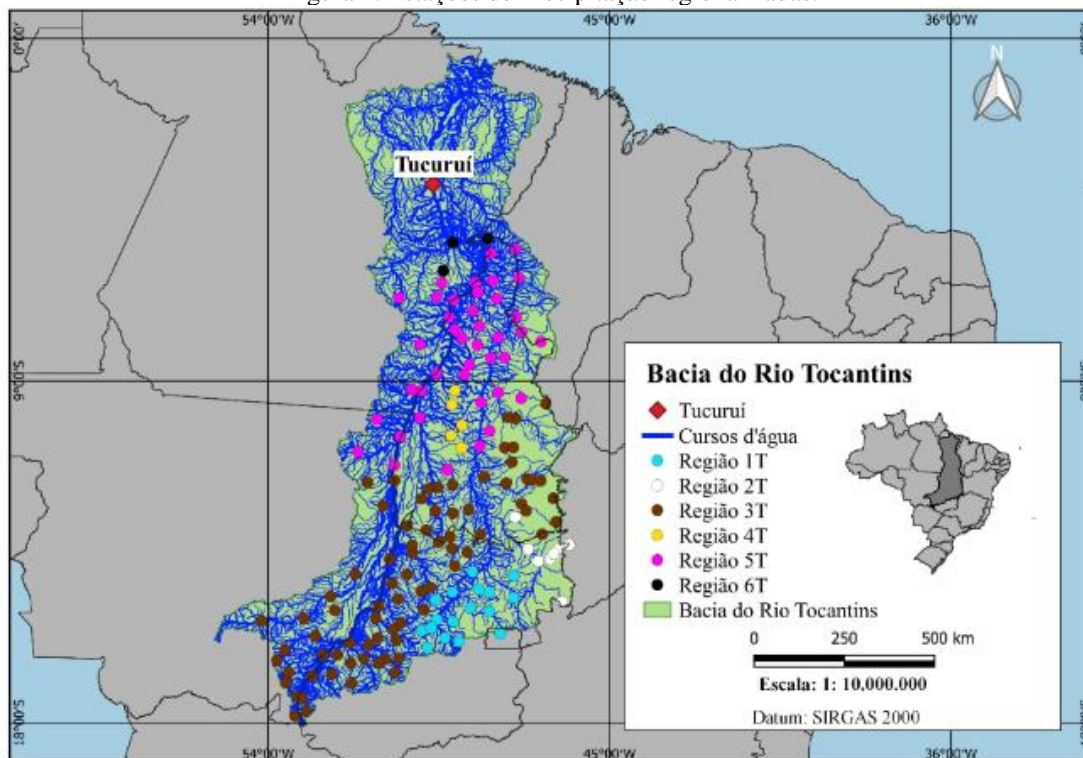
$$ER = \frac{1}{n} \sum \frac{|F - O|}{O} \quad (10)$$

Para as equações acima, F é definido como o valor estimado, O é o valor observado e n é a quantidade de casos avaliados.

3 RESULTADOS

A regionalização da precipitação separou o conjunto de estações em seis regiões, como mostrado na Figura 2. A região 1T (17 estações) se encontra no Centro-Oeste brasileiro, na porção leste da cabeceira da bacia. A região 2T (8 estações) está situada à direita da região 1T, com a maioria das estações em Goiás. A região 3T (76 estações) ocupa o lado oeste da cabeceira, contornando as regiões 1T e 2T, tendo estações na porção oeste de Goiás, leste do Mato Grosso e sul do estado do Tocantins. A região 4T está localizada na área central da bacia, no estado do Tocantins. A região 5T (5 estações) ocupa o extremo nordeste do Mato Grosso, sudeste do Pará, centro-norte do Tocantins e oeste do Maranhão. A região 6T (3 estações) encontra-se próximo a Tucuruí, localizada na porção leste do Pará e oeste da Maranhão.

Figura 2. Estações de Precipitação regionalizadas.



Fonte: Elaborado pelos autores.

Na Figura 3 estão expostas médias mensais da precipitação em cada uma das seis regiões de precipitação homogêneas, dispostas em linhas, e as médias mensais de vazão no posto de Tucuruí em barras verticais. No geral, as regiões de precipitação apresentam valores mínimos no inverno. Os valores máximos ocorrem no verão para as regiões 1T a 4T e no outono nas regiões 5T e 6T. Os valores mais altos de precipitação ocorrem majoritariamente para as regiões 4T e 6T e os valores mais baixos para a região 2T.

A região 6T apresentou o maior valor médio mensal, com precipitação de 339 mm em março. O valor mais baixo encontrado na bacia ocorreu na região 2T, com 0,5 mm em julho. Tais valores corroboram com o encontrado por Marcuzzo e Goularte (2013), que identificaram os períodos secos da bacia de junho a agosto, sendo mais severos na porção sul e mais brandos na porção norte. Comparando a regionalização realizada com o estudo de Loureiro e Fernandes (2013), tem-se que as regiões 4T e 5T concentram aproximadamente 35% da precipitação anual.

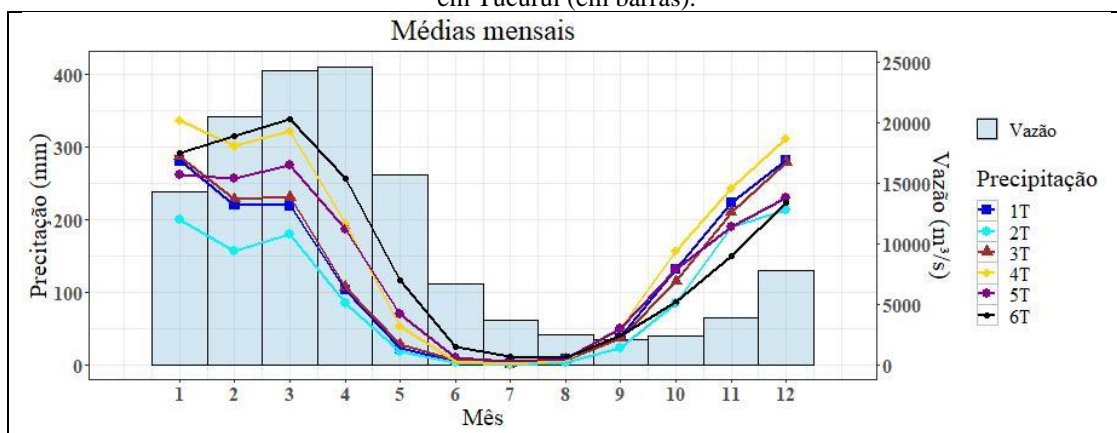
Para as regiões 1T e 2T, o pico de precipitação ocorre em dezembro, para as regiões 3T e 4T em janeiro, e nas regiões 5T e 6T, em março. Com isso, tem-se defasagens de quatro, três e um mês em relação ao pico de vazão em Tucuruí, que ocorre em abril (24631 m³/s).

As menores médias de precipitação ocorreram em julho para as regiões 1T, 2T, 3T, 4T e 5T e em agosto para 6T. Como o mínimo de vazão é atingido em setembro (2128 m^3/s), tem-se defasagem de dois e um mês entre os mínimos de precipitação e vazão.

Em Tucuruí o máximo de vazão ocorre no outono e o mínimo no inverno (Figura 3). Os meses de setembro e outubro apresentam médias abaixo de 2550 m^3/s , que é o valor de vazão de estiagem para a bacia (MATOS et al., 2007).

O padrão sazonal da vazão mostra um aumento gradativo e uma queda abrupta. Para Lima *et al.* (2022), que estudou a sub-bacia Itacaiúnas (Pará), o regime fluviométrico varia de acordo com dois períodos hidrológicos definidos: de chuvas, nos meses de outubro a abril quando a vazão tende a aumentar significativamente, e de estiagem, de maio a setembro, quando a vazão da bacia diminui. Porém, para a sub-bacia, o máximo de vazão ocorre em março, indicando uma defasagem de um mês com a vazão em Tucuruí.

Figura 3. Médias mensais de precipitação das regiões homogêneas de precipitação (em linha) e da vazão em Tucuruí (em barras).



Fonte: Elaborado pelos autores.

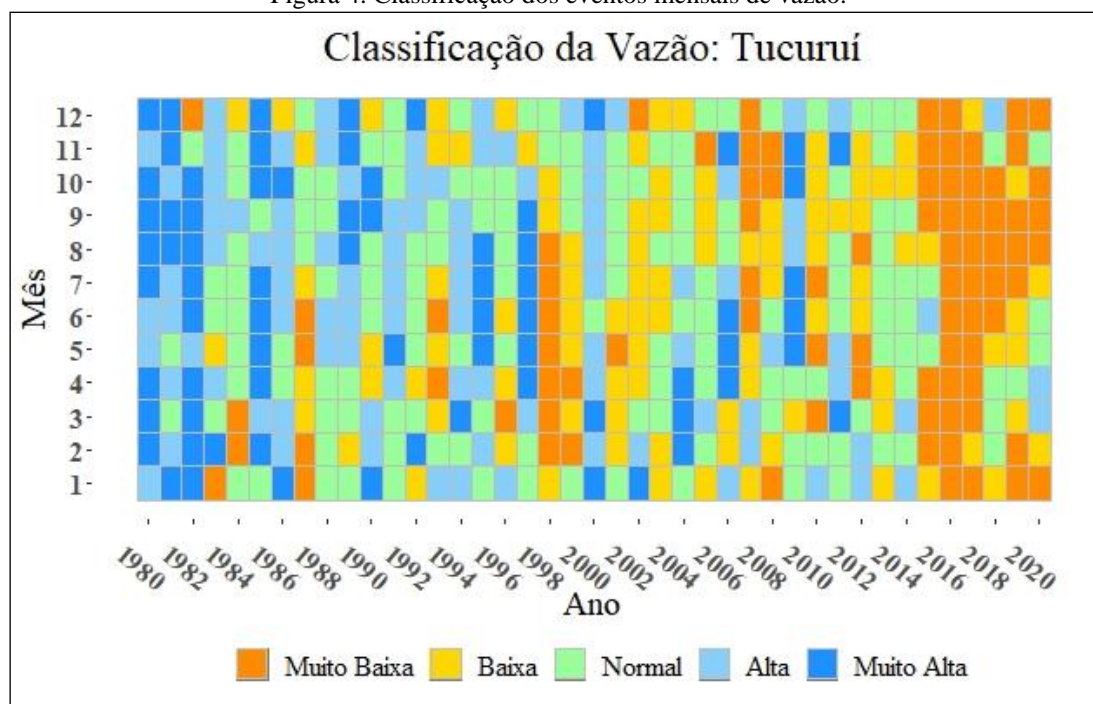
Na Figura 4 está exposto a classificação da vazão, os tons quentes indicam eventos abaixo do normal, os tons azulados acima do normal e verde a normalidade. Os tons escuros sinalizam os eventos extremos. Nota-se que os três primeiros anos apresentam maior quantidade de eventos acima do normal, como o ano de 1983, com nove eventos de vazão muito alta, sendo justificada pela ocorrência de um *El Niño* que influenciou a precipitação na América do Sul (KAYANO; MOURA, 1986). Feischmann *et al.*, (2020) identificaram o *El Niño* de 1983 como o responsável pelas inundações ocorridas na BHRT.

Destaca-se também o período maio de 1989 a março de 1990, uma sequência de 11 meses de vazão acima do normal. Os cinco últimos anos do gráfico (de 2015 a 2020) apresenta comportamento oposto, com muitos eventos abaixo do normal, onde o período de 2015 a 2017 apresentou 17 meses seguidos de vazão muito baixa (Figura 4).

O acúmulo de eventos acima do normal no início da série e abaixo do normal no final de série indica uma tendência de decréscimo, como encontrado por Sousa e Moura (2020) para o rio Caiapó (cabeceira da BHRT). Os autores também analisaram os valores de vazão média mensal no intervalo de 1972 a 2009 e de 2010 a 2020, e encontraram um decréscimo na vazão média mensal da última década para o rio Caiapó e Paranã (cabeceira do BHRT) (SOUSA; MOURA, 2020), principalmente no período de cheia.

Estes baixos valores de vazão nos últimos anos resultaram em problemas de irrigação às margens do rio Urubu, presente na sub-bacia do rio Formoso, na porção sudoeste da BHRT (VOLKEN et al., 2022).

Figura 4. Classificação dos eventos mensais de vazão.



Fonte: Elaborado pelos autores.

Analisando os resultados das correlações simples mostradas na Figura 5, nota-se que apenas a região 2T apresentou algum índice com resultados significativos em todas as defasagens e a região 3T apresentou as correlações mais fortes.

Para a região 1T (Figura 5-a), apenas o índice ND35q- apresentou correlações significativas entre os índices que indicam seca. Destaca-se também o índice ND65q+, com resultados significativos até seis meses para a região.

De forma geral, os índices da região 2T (Figura 5-b) possuem correlações significativas de sete a doze meses. Os índices PRCTOT, ND85q+ e ND65q+ apresentaram correlações significativas com a vazão em todas as defasagens.

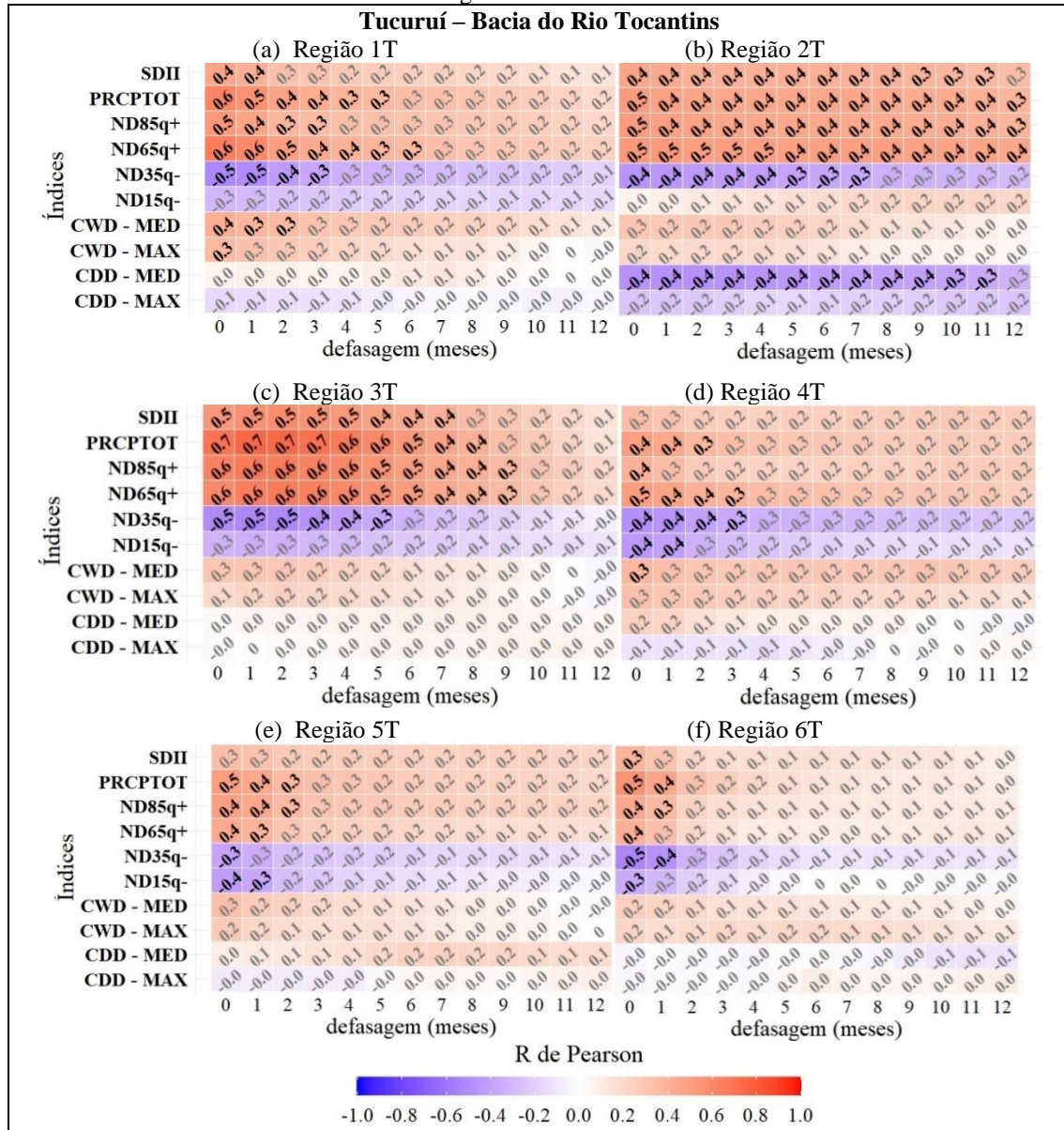
A região 3T (Figura 5-c) apresentou as correlações mais intensas, para o índice PRCTOT. Assim como na região 1T, apenas o índice ND35q- apresentou correlação significativa entre os índices que representam seca.

Para a região 4T (Figura 5-d), os índices a partir dos limiares quantílicos menos extremos (ND65q+ e ND35q-) apresentaram a maior quantidade de correlações significativas (até três meses).

Para a região 5T (Figura 5-e), os índices PRCTOT e ND85q+ apresentaram 12 meses de correlações significativas. O índice ND15q- apresentou mais defasagens significativas que o ND35q-, indicando a relevância de extremos na região para a variação da vazão.

A região 6T (Figura 5-f) apresentou correlações significativas apenas para casos sem defasagem dos índices SDII, ND65q+ e ND15q- e até um mês de defasagem para os índices PRCTOT, ND85q+ e ND35q-.

Figura 5. Correlação anual dos índices de precipitação das regiões de precipitação homogênea com a vazão no posto de Tucuuruí, na Bacia do Rio Tocantins. Resultados significativos escritos em preto e os não-significativos em cinza.



Fonte: Elaborado pelos autores.

Utilizando os critérios de pré-seleção, os índices de precipitação selecionados foram reduzidos para três ou quatro por região. Na Tabela 3 estão expostos os índices selecionados, indicando por 'X' as defasagens com correlações significativas com a vazão.

Tabela 3. Tabela dos índices pré-selecionados pela correlação simples anual móvel.

Região	Índice	Defasagens (mês)											
		1	2	3	4	5	6	7	8	9	10	11	12
1T	PRCPTOT	X	X	X	X	X							
	CWD – MED		X										
	ND35q-	X	X	X									
	ND65q+	X	X	X	X	X	X						
2T	PRCPTOT	X	X	X	X	X	X	X	X	X	X	X	X
	CDD – MED	X	X	X	X	X	X	X	X	X	X	X	
	ND35q-	X	X	X	X	X	X	X					
	ND65q+	X	X	X	X	X	X	X	X	X	X	X	X
3T	PRCPTOT	X	X										
	ND35q-	X											
	ND85q+	X	X										
4T	PRCPTOT	X	X										
	ND35q-	X	X	X									
	ND65q+	X	X	X									
5T	PRCPTOT	X	X										
	ND15q-	X											
	ND85q+	X	X										
6T	PRCTOT	X											
	ND35q-	X							X				
	ND85q+	X											

Fonte: Elaborado pelos autores.

Foi avaliada a multicolinearidade através do VIF para o conjunto de preditores (índices de precipitação e a vazão defasada) de cada defasagem. Os preditores que apresentaram valores de $VIF \geq 10$ foram eliminados da construção do modelo de regressão simples.

Na Tabela 4 estão expostos os resultados da calibração e validação dos modelos construídos. Nas equações, o Vaz_{def} representa a vazão defasada e os subscrito dos índices de precipitação indica a região de precipitação no qual aquele índice pertence.

O modelo gerado para a defasagem de um mês apresentou o maior valor do coeficiente de determinação ($R^2 = 0,58$). Nota-se uma queda acentuada dos valores de R^2 com o aumento da defasagem, atingindo o mínimo no modelo de 12 meses, com $R^2 = 0,12$, indicando a queda no poder explicativo dos modelos.

Como podemos observar nas equações da Tabela 4, a vazão defasada não foi selecionada para os modelos de defasagem de 4 a 7 meses. O índice ND65q+ da região 2T foi o preditor selecionado para a maioria dos modelos, de 3 a 12 meses, indicando que a precipitação alta na região tem influência na vazão a partir do terceiro mês de defasagem.

Em relação ao erro sistemático no período de validação, há um comportamento distinto dos modelos de defasagem de um e dois meses, pois superestimaram os valores observados e os demais subestimaram os valores observados para a vazão em Tucuruí,

como observado pela diferença de sinal dos valores encontrados para o viés. Além disso, pode-se observar que os erros mais baixos foram encontrados para o modelo com defasagem de um mês.

Tabela 4. Equação, coeficiente de determinação e os erros dos modelos.

lag	Calibração - Período: 1980 a 2010	R ²	Validação - Período: 2011 a 2020			
	Equação		VIÉS	REQM	EA	ER(%)
1	$y = 0,02 + 0,7 * Vaz_{def} + 0,15 * PRCTOT_{2T} - 0,11 * ND35q_{6T} - 0,2 * ND85q_{6T}$	0,58	187,14	2224,83	1399,28	17,72
2	$y = 0,05 + 0,42 * Vaz_{def} - 0,26 * CDD - MED_{2T}$	0,27	107,80	2953,75	1889,00	27,47
3	$y = -0,05 + 0,2 * Vaz_{def} + 0,43 * ND65q_{2T} - 0,2 * CDD - MED_{2T}$	0,19	-1230,33	3996,36	2408,93	25,20
4	$y = -0,07 + 0,6 * ND65q_{2T} - 0,19 * CDD - MED_{2T}$	0,14	-1349,80	4262,64	2615,69	28,83
5	$y = -0,08 + 0,75 * ND65q_{2T}$	0,13	-1521,68	4454,62	2746,65	31,54
6	$y = -0,09 + 0,76 * ND65q_{2T}$	0,13	-1545,65	4528,23	2789,77	32,14
7	$y = -0,05 + 0,56 * ND65q_{2T} + 0,23 * PRCTOT_{2T} + 0,22 * ND35q_{6T}$	0,17	-1006,80	3849,22	2281,12	25,35
8	$y = -0,09 - 0,15 * Vaz_{def} + 0,84 * ND65q_{2T}$	0,14	-1442,32	4630,06	2852,84	34,02
9	$y = -0,09 - 0,17 * Vaz_{def} + 0,84 * ND65q_{2T}$	0,14	-1415,32	4651,70	2883,36	35,03
10	$y = -0,08 - 0,17 * Vaz_{def} + 0,82 * ND65q_{2T}$	0,13	-1283,75	4540,33	2831,37	34,63
11	$y = -0,06 - 0,21 * Vaz_{def} + 0,79 * ND65q_{2T}$	0,13	-1014,08	4380,80	2744,39	33,41
12	$y = -0,04 - 0,22 * Vaz_{def} + 0,73 * ND65q_{2T}$	0,12	-643,38	4185,37	2629,02	31,64

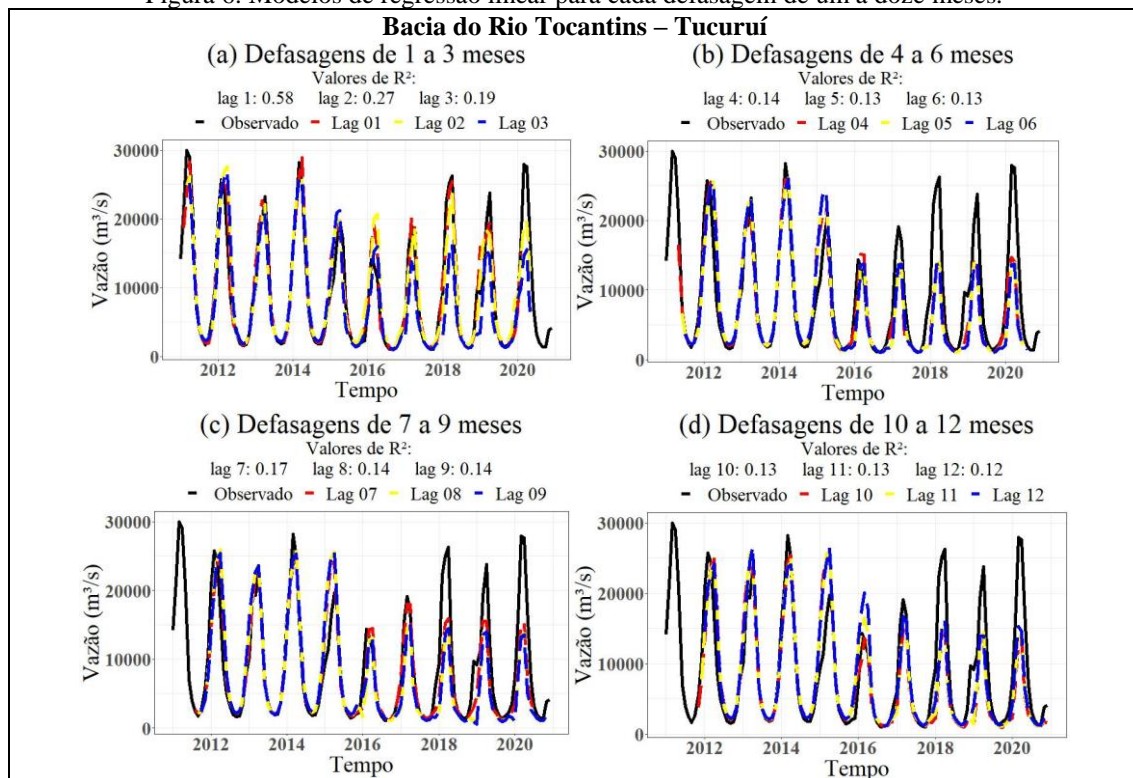
Fonte: Elaborado pelos autores.

Na Figura 6 estão expostos os valores obtidos nos modelos construídos e comparados com os valores observados de vazão no posto de Tucuruí para o período de validação. De modo geral, os modelos construídos apresentaram valores próximos aos observados em seus valores mínimos de vazão. Porém os picos de vazão são subestimados, principalmente no período mais recente, a partir de 2017. Como já indicado na tabela anterior, os modelos de defasagens mais curtas possuem os melhores desempenhos, com destaque para o modelo de um mês de defasagem. No entanto, este modelo também não foi capaz de prever os picos de vazão de 2019 e 2020. Todos os modelos desenvolvidos apresentaram valores próximos ao observado para os primeiros anos testados (até 2014), onde tanto os valores previstos mínimos quanto os máximos foram próximos ao observado, porém, aumentando a diferença entre os valores

modelados e os observados com a defasagem. Percebe-se que a partir da defasagem de três meses, os últimos quatro picos de vazão não foram bem previstos, ou seja, há grande diferença entre os valores modelados e observado para os eventos, onde os valores previstos foram próximos a metade dos valores observados para o mesmo período.

Contudo, pode-se inferir que os índices de precipitação escolhidos são bons preditores de vazão mensal. Sugere-se a necessidade de recalibração recorrente dos modelos, considerando um maior número de casos passados, para melhorar o ajuste e o acerto de casos futuros. Ademais, estes modelos incluem como preditores somente precipitação e vazão defasada, que explicam apenas uma parte da variância da vazão em Tucuruí. A inclusão de outros índices climáticos poderia trazer benefícios para a previsibilidade de extremos de vazão, com prazos mais longos, conforme sugerido por Cardoso *et al.* (2022).

Figura 6. Modelos de regressão linear para cada defasagem de um a doze meses.



Fonte: Elaborado pelos autores.

4 CONCLUSÃO

A separação das estações de precipitação em regiões homogêneas, além de reduzir a quantidade de dados utilizados, foi útil para a redução de dados faltantes. Todas as regiões apresentaram valores mínimos no inverno, sendo os máximos no verão para as regiões 1T, 2T, 3T e 4T (cabecera e centro da bacia) e no outono para as regiões 5T e 6T

(próximo a Tucuruí). A vazão em Tucuruí apresenta valores máximos em abril e mínimos em setembro, indicando defasagem com a precipitação.

As correlações entre a vazão e os índices de precipitação indicaram relevância dos valores acumulados de precipitação em dias úmidos (PRCTOT) para a vazão em todas as regiões de precipitação. Os valores médios de dias úmidos e secos consecutivos (CWD – MED e CDD – MED) foram selecionados para as regiões 1T e 2T, respectivamente. Os índices com limiares quantílicos foram selecionados em todas as regiões, porém com intensidades diferentes. Para as regiões 1T, 2T e 4T, foi selecionado o índice de limiar quantílico chuvoso ND65q+, enquanto para as demais, foi selecionada o mais extremo (ND85q+). As regiões 1T, 2T, 3T, 4T e 6T indicaram o índice ND35q- como o mais bem correlacionado com a vazão, enquanto a região 5T indicou o índice ND15q-.

Os modelos de RLM indicaram que os índices de precipitação anual móvel apresentam potencialidade para a predição da vazão, podendo ser usado conjuntamente com outros índices climáticos. O modelo de defasagem de um mês apresentou os menores erros e o maior coeficiente de determinação ($R^2 = 0,58$).

Pelos resultados dos modelos, podemos destacar a relevância da precipitação nas regiões 2T e 6T para o prognóstico de vazão em Tucuruí, principalmente através do índice CDD – MED da região 2T, que indica a influência de dias secos consecutivos na vazão, do índice ND85q+ na região 6T, que contabiliza os eventos de precipitação muito alta que ocorre na região 6T e o índice ND65q+ na região 2T, que foi selecionado para vários dos modelos construídos. Nota-se também que as regiões 2T e 6T são regiões fisicamente distantes e geograficamente distintas, com comportamentos de precipitação mensal diferentes.

AGRADECIMENTOS

Os autores agradecem à Universidade Federal do ABC (UFABC), ao Conselho Nacional de Desenvolvimento Científico e Tecnológica (CNPq) pelo fomento à pesquisa por meio de Bolsa de Mestrado da primeira autora e a empresa *Climatepo* pela parceria.

REFERÊNCIAS

ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; DE MORAES GONÇALVES, J. L.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, [s. l.], v. 22, n. 6, p. 711–728, 2013.

ANA. **Base Hidrográfica**. [S. l.: s. n.], 2018. Disponível em: <https://metadados.snirh.gov.br/geonetwork/srv/por/catalog.search#/metadata/4fd91f0d-f34f-4fca-a961-c2dcb3e0446e>. Acesso em: 19 jul. 2021.

ANA. **Conjuntura de Recursos Hídricos 2018**. Brasília: BRASIL, 2018.

ANA. **Conjuntura dos Recursos Hídricos no Brasil: Regiões Hidrográficas Brasileiras**. Brasília: BRASIL, 2015.

ANA. **Divisão de Bacias**. [S. l.: s. n.], 2016. Disponível em: <https://metadados.snirh.gov.br/geonetwork/srv/por/catalog.search#/metadata/fe192ba0-45a9-4215-90a5-3fba6abea174>. Acesso em: 19 jul. 2021.

ANA. **Divisões hidrográficas do Brasil**. [S. l.: s. n.], [s. d.]. Disponível em: https://www.ana.gov.br/aguas-no-brasil/panorama-das-aguas/copy_of_divisoes-hidrograficas. Acesso em: 28 jun. 2021.

ANA. **SAR - Sistema de Acompanhamento de Reservatórios**. [S. l.: s. n.], [s. d.]. Disponível em: <https://www.ana.gov.br/sar0/MedicaoSin#>. Acesso em: 20 abr. 2021.

ANACONDA INC. **Anaconda | Individual Edition**. [S. l.: s. n.], 2021. Disponível em: <https://www.anaconda.com/products/individual#windows>. Acesso em: 19 jul. 2021.

BUSSAD, W. O.; MORETTIN, P. A. **Estatística Básica**. 8. ed. São Paulo, SP: [s. n.], 2013.v. 51.

CAPOZZOLI, C. R.; CARDOSO, A. O.; FERRAZ, S. E. T. Padrões de Variabilidade de Vazão de Rios nas Principais Bacias Brasileiras e Associação com Índices Climáticos. **Revista Brasileira de Meteorologia**, [s. l.], v. 32, n. 2, p. 243–254, 2017.

CARDOSO, A. O.; MALFATTI, M. G. L.; DIAS, P. L. S. Predictability of extreme streamflow in Brazilian rivers through climate indices. **International Journal of Hydrology**, [s. l.], v. 6, n. 5, p. 213–222, 2022.

CRUZ, L. N.; HERREROS, M. M. A. G.; VILARINHO, C. C.; DEMETRIO NETO, E. N.; MARTINS, G. C. Desenvolvimento socioeconômico na região de Matopiba, Brasil. **Brazilian Journal of Development**, [s. l.], v. 5, n. 8, p. 12538–12556, 2019.

DEVORE, J. L. **Probabilidade e Estatística para Engenharia e Ciências**. 6. ed. São Paulo, SP: Cengage Learning, 2006. 2006.

DIONISIO, C. B.; CARDOSO, A. O.; SOUZA, A. R. Eventos Extremos de Precipitação e Relação chuva-vazão na bacia do rio São Francisco. In: III CONGRESSO BRASILEIRO INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA, 2022, Diamantina-MG. **Anais [...]**. Diamantina-MG: Even3, 2022. p. 10.

DIONISIO, C. B.; CARDOSO, A. O.; SOUZA, A. R. Padrão e extremos de vazões na bacia do rio São Francisco. *In: II CONGRESSO BRASILEIRO INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA*, 2021, Diamantina-MG. **Anais [...]**. Diamantina-MG: Even3, 2021. p. 10.

DRAPER, N. R.; SMITH, H. **Applied Regression Analysis**. 3rd.ed. [S. l.: s. n.], 1998. 1998.

FLEISCHMANN, A. S.; SIQUEIRA, V. A.; WONGCHUIG, S.; COLLISCHONN, W.; DE PAIVA, R. C. D. The great 1983 floods in South American large rivers: a continental hydrological modelling approach. **Hydrological Sciences Journal**, [s. l.], p. 35, 2020.

GHISLENI, J.; MARTINEZ, C. B.; GUIMARÃES, A. G. REGRA OPERATIVA DA BACIA DO RIO TOCANTINS. **I Simpósio Nacional de Mecânica dos Fluidos e Hidráulica**, [s. l.], v. 1, n. 3, p. 1–10, 2022.

IBGE. **Geociência: Organização do Território**. [S. l.: s. n.], 2020. Disponível em: <https://www.ibge.gov.br/geociencias/downloads-geociencias.html>. Acesso em: 19 jul. 2021.

KARL, T. R.; NICHOLLS, N.; GHAZI, A. **Weather and Climate Extremes: Changes, Variations and a Perspective from the Insurance Industrial**. [S. l.]: Springer Netherlands, 1999. 1999.v. 42.

KAYANO, M. T.; MOURA, A. D. O El Niño de 1982-1983 e a precipitação da América do Sul. **Revista Brasileira de Geofísica**, [s. l.], v. 4, p. 201–214, 1986.

LIMA, S. S.; LACERDA, I. O.; LOUREIRO, G. E. Regimes Mensais e Anuais das Vazões Médias, Máximas e Mínimas na Sub-Região Hidrográfica Itacaiúnas (SRHI). **Research, Society and Development**, [s. l.], v. 11, n. 8, p. 1–11, 2022.

LOUREIRO, G. E.; FERNANDES, L. L. Variação da precipitação por método de interpolação geoestatística. **Revista Ambiente & Água - An Interdisciplinary Journal of Applied Science**, [s. l.], v. 8, n. 2, 2013.

LUMLEY, T. **Leaps: Regression Subset Selection**. [S. l.]: R package version 3.1, 2020.

MARCUZZO, F. F. N.; GOULARTE, E. R. P. Caracterização do Ano Hidrológico e Mapeamento Espacial das Chuvas nos Períodos Úmido e Seco do Estado do Tocantins. **Revista Brasileira de Geofísica**, [s. l.], v. 6, n. 1, p. 91–99, 2013.

MATOS, B. A.; TEIXEIRA, A. L. F.; BURNETT, J. A.; ZOBY, J. L. G.; FREITAS, M. A. S. Disponibilidade e Demandas de Recursos Hídricos nas 12 regiões hidrográficas do Brasil. *In: XVII SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS*, 2007, São Paulo, SP. **Anais [...]**. São Paulo, SP: [s. n.], 2007. p. 20.

NAIMI, B.; HAMM, N. A. S.; GROEN, T. A.; SKIDMORE, A. K.; TOXOPEUS, A. G. Where is positional uncertainty a problem for species distribution modelling?. **Ecography**, [s. l.], v. 37, n. 2, p. 191–203, 2014.

ONS. **Atualização de Séries Históricas de Vazões - Período de 1931 a 2019**. Rio de Janeiro: ONS DOP-REL 0142/2020, 2020.

PENHA, L. R.; BARROS, L. L. G. Energia E a Indústria Da Celulose Na Amazônia: O Território Das Localidades Centrais Ao Entorno Da Usina Hidrelétrica De Estreito No Maranhão. **Brazilian Journal of Development**, [s. l.], v. 7, n. 3, p. 23927–23943, 2021.

PETERSON, T. C.; FOLLAND, C.; GRUZA, G.; HOGG, W.; MOKSSIT, A.; PLUMMER, N. Report on the Activities of the Working Group on Climate Change Detection and Related Rapporteurs Report on the Activities of the Working Group on Climate Change Detection and Related Rapporteurs Methods with emphasis on analyses of extreme events. **Wmo**, [s. l.], n. 48, p. 144, 2001.

QGIS. **Aplicativo para GIS**. [S. l.: s. n.], [s. d.]. Disponível em: https://qgis.org/pt_BR/site/about/index.html. Acesso em: 19 jul. 2021.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2019.

REBOITA, M. S.; GAN, M. A.; ROCHA, R. P.; AMBRIZZI, T. Regimes de precipitação na América do Sul: uma revisão bibliográfica. **Revista Brasileira de Meteorologia**, [s. l.], v. 25, n. 2, p. 185–204, 2010.

SOUSA, F. A.; MOURA, D. M. B. Evapotranspiração potencial (ETp) e sua influência na vazão de rios do Cerrado Brasileiro. **Elisée, Rev. Geo. UEG**, [s. l.], v. 11, n. 1, p. 1–23, 2020.

SOUZA, A. R.; CARDOSO, A. O. Padrão de vazão da bacia do rio Paraíba do Sul: Estação Paraibuna. In: II CONGRESSO BRASILEIRO INTERDISCIPLINAR EM CIÊNCIA E TECNOLOGIA, 2021, Diamantina-MG. **Anais [...]**. Diamantina-MG: Even3, 2021. p. 6.

TORI, S. A.; CARDOSO, A. O.; SOUZA, A. R. Padrão de vazão e extremos de na parte baixa da bacia do rio Paraná. In: SIMPÓSIO EM CLIMA, ÁGUA, ENERGIA E ALIMENTOS, 2021, Campinas, SP. **Anais [...]**. Campinas, SP: Galoá, 2021. p. 2.

TUCCI, C. E. M. **Hidrologia: Ciência e Aplicação**. 2. ed. Porto Alegre: Editora da Universidade Federal do Rio Grande do Sul, 2001. 2001.

VOLKEN, N. J.; MINOTI, R. T.; ALVES, C. M.; VERGARA, F. E. Analyzing the impact of agricultural water-demand management on water availability in the Urubu River basin – Tocantins, Brazil. **Revista Ambiente e Água**, [s. l.], v. 17, n. 4, 2022.

WEISBERG, S. **Applied Linear Regression**. fourth edied. Minneapolis, MN: [s. n.], 2014. 2014.

WICKHAM, H. **Tidyverse: Easily Install and Load the “Tidyverse”**. Versão R package version 1.2.1. [S. l.: s. n.], 2017.

WILKS, D. S. **Statistical Methods in the Atmospheric Sciences**. 3. ed. Oxford: Elsevier Inc., 2011. 2011.v. 100.

XAVIER, T. M. B.; XAVIER, A. F. S.; DIAS, M. A. F. S.; DIAS, P. L. S. Interrelações Entre Eventos ENOS (ENSO), a ZCIT (ITCZ) no Atlântico e a Chuva nas Bacias Hidrográficas do Ceará. **Revista Brasileira de Recursos Hídricos**, [s. l.], v. 8, n. 2, p.

111–126, 2003.

ZHANG, X.; ETCCDI. **Climate Changes Indices**. [S. l.: s. n.], 2009. Disponível em: http://etccdi.pacificclimate.org/list_27_indices.shtml. Acesso em: 22 mar. 2022.