

# Séries Temporais e Técnicas de Regressão: Uma Análise Comparativa a Partir de Dados de Precipitações Pluviométricas Coletados por Estação Meteorológica Automática

Ananias Caetano de Oliveira<sup>1</sup>, Adonias Caetano de Oliveira<sup>1</sup>, Rhyan Ximenes de Brito<sup>1</sup>

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)  
Av. Tabelião Luiz Nogueira de Lima S/N – Tianguá – CE – Brasil

{ananciascaetano.ifce, adonias.ifce, rxbrito}@gmail.com

**Abstract.** *Rainfall forecasts are applied in lack of rain scenarios, such as droughts and droughts, the main scenario in the Northeast. In order to contribute with possible hydrological models to support the management of water resources, this article analyzes the performance of Random Forest, Polynomial Regression and Support Vector Regression Machines in a series of flow and daily rainfall data obtained from a rainfall station. The best results presented the  $R^2$  between 49.8% and 53.2% according to the approach applied to the studied problem.*

**Resumo.** *As previsões de chuva-vazões são aplicadas em cenários de falta de chuvas, como as estiagens e as secas, principal cenário nordestino. No intuito de contribuir com possíveis modelos hidrológicos de apoio à gestão dos recursos hídricos, este artigo analisa o desempenho de Random Forest, Regressão Polinomial e Support Vector Regression Machines em uma série de dados de vazão e chuvas diárias obtidos de uma estação pluviométrica. Os melhores resultados apresentaram o  $R^2$  entre 49,8% e 53,2% de acordo com abordagem aplicada para o problema estudado.*

## 1. Introdução

A segurança hídrica está relacionada ao acesso à água para as necessidades humanas, como beber e sustentar em qualidade e quantidade aceitáveis. Para isso é preciso que a água captada e tratada atenda aos padrões básicos de saúde e bem-estar humanos, ou seja, é adequada para consumo humano promovendo a dignidade humana, a vida e a saúde [Soyapi 2017].

Contudo, vários municípios da Região Nordeste enfrentam dificuldades para garantir constantemente o acesso a água de qualidade por conta da captação desse recurso, que nem sempre é satisfatória em volume, devido as poucas chuvas durante décadas [LIMA 2018, Santana and Santos 2020].

Como ferramentas computacionais de apoio ao planejamento e gestão dos recursos hídricos, os modelos hidrológicos são instrumentos de alta importância para realizar simulações que forneçam subsídios para a regionalização de vazões em bacias hidrográficas, pois permitem a predição do regime fluvial [Schroder et al. 2020, Debastiani et al. 2019]. As previsões de chuva-vazões podem ser aplicadas em

cenários de falta de chuvas, como as estiagens e as seca e em situações de extrema precipitação, como inundações, enxurradas, alagamentos e movimentos de massa [Brubacher et al. 2020].

Nessa perspectiva, é importante salientar que os algoritmos de aprendizagem de máquinas são abordagens bastante empregadas devido aos bons resultados encontrados em distintos segmentos da hidrologia, conforme pode ser visto nos trabalhos de [Ingrisawang et al. 2008], [Sumi et al. 2012], [Santos and Celeste 2014], [de Lima and Guedes 2015], [Gorodetskaya et al. 2018] e [Schroder et al. 2020].

Diante do que foi exposto, o objetivo deste artigo é uma análise comparativa de técnicas de aprendizagem de máquina para previsão de chuva de curto prazo, incluindo (1) preparação de dados de precipitação obtidos por uma estação meteorológica automática do município de Tianguá (CE), (2) estudo e aplicação de técnicas de regressão polinomial, baseadas em árvores e otimização e (3) avaliação do desempenho dos modelos com base em Coeficiente de determinação ( $R^2$ ), *Mean Absolute Error* (MAE) e *Root Mean Squared Error* (RMSE).

Este trabalho está organizado em 6 seções: a Seção 2 apresenta os trabalhos relacionados; Seção 3 traz a fundamentação teórica sobre séries temporais de precipitação; a Seção 4 apresenta materiais e métodos; Seção 5 trata dos resultados e discussões; e a Seção 6 apresenta as considerações finais e trabalhos futuros.

## 2. Trabalhos Relacionados

Foram escolhidos três trabalhos recentes sobre modelagem chuva-vazão, via Google Scholar<sup>1</sup>, que usam técnicas do tipo Redes Neurais ou baseadas em árvore de decisão.

O artigo [Debastiani et al. 2019] analisou o desempenho da árvore modelo (M5P) e sua sensibilidade à poda e comparação com o desempenho da *MultiLayer Perceptron* (MLP) para a simulação da vazão média diária mensal da bacia hidrográfica do Alto Canoas. Além da sua simplicidade e velocidade de processamento, o M5P apresentou bom ajuste na modelagem, porém a MLP apresentou desempenho superior em todos os índice de *Nash* (NS). A MLP alcançou no teste 1,  $NS = 0,872$ ,  $R^2 = 0,578$  e  $RMSE = 23,205m^3/s$  e a M5P obteve  $NS = 0,581$ ,  $R^2 = 0,611$  e  $RMSE = 22,735m^3/s$ . No teste 2 e tratamento 4, a MLP alcançou  $NS = 0,752$ ,  $R^2 = 0,684$  e  $RMSE = 22,684m^3/s$  e o M5P obteve  $NS = 0,641$ ,  $R^2 = 0,697$  e  $RMSE = 20,288m^3/s$ .

Em [de Mendonça et al. 2021], foi simulado o fenômeno de conversão de chuvas em vazões de uma Sub-bacia Hidrográfica do Rio Guamá no Pará e Amazônia. Os modelos foram baseados na arquiteturas MLP e NARX (*Nonlinear Autoregressive with Exogenous Inputs*). O melhor dos modelos RNA-NARX foi superior ao MLP apresentando para a fase de teste,  $R^2$  igual a 99,08, RMSE igual a  $13,21m^3/s - 1$  e MAPE igual a 4,45%.

[Yan et al. 2021] propõem um modelo de previsão de chuva baseado na rede neural *Attentive Interpretable Tabular Learning* (TabNet). O experimento usou cinco anos de dados meteorológicos de 26 estações na região de Pequim-Tianjin-Hebei da China. Por apresentar clima temperado quente, há pouca chuva no inverno e é chuvoso no verão,

---

<sup>1</sup>Disponível em <https://scholar.google.com.br/?hl=pt>

foram utilizados dois modelos. O *Rainy-Model*, baseado nos dados da estação chuvosa, obteve na fase de teste o  $MAE = 0.3373$  e  $RMSE = 0.5561$ . Já o *Drier-Model*, baseado nos dados da estação não chuvosa, obteve  $MAE = 0.4825$  e  $RMSE = 0.6812$ .

### 3. Série Temporal de Precipitação

Uma série temporal consiste em um conjunto de observações de uma variável aleatória indexada no tempo, representando um conjunto de números reais positivos. As observações são dependentes entre si e o estudo consiste em analisar e modelar esta dependência [Ballini et al. 2000]. Por ser um processo estocástico, a série temporal pode ser categorizada como Estacionária ou Não Estacionária.

De acordo com [Bezerra 2006], uma série temporal é estacionária quando se desenvolve aleatoriamente, no tempo, em torno de uma média constante, refletindo alguma forma de equilíbrio estável. Porém, a maior parte das séries apresentam alguma forma de não-estacionariedade, por exemplo as séries econômicas que em geral possuem tendências lineares positivas ou negativas.

Como são consideradas estacionárias diante da maioria dos procedimentos de análise estatística, é necessário transformar os dados originais se estes não formam uma série estacionária. Tomar diferenças sucessivas da série original, até se obter uma série estacionária, é a transformação mais comum, sendo detalhada em [de Albuquerque 2019].

Outro conceito importante, é a sazonalidade, definida como um comportamento específico de uma série que tende a se repetir por períodos idênticos a medida que o tempo passa [da Costa 2019, Parmezan et al. 2016]. As variações sazonais podem ser aferidas por medidas diárias, semanais ou mensais. Tais movimentos ou flutuações com período igual ou inferior a um ano, “sistemático, mas não necessariamente regulares”, decorrem normalmente das variações climáticas relacionadas as estações do ano, ou a fatores culturais relacionados a efeitos de calendário [Figueredo 2008, de Albuquerque 2019].

Tendo em vista esses conceitos teóricos, nota-se que os objetivos deste trabalho é a análise de métodos de regressão aplicada numa série temporal de precipitação, considerada por [Almeida et al. 2011] como a que apresenta a maior variabilidade temporal e espacial dentre os parâmetros meteorológicos, portanto apresentando maior dificuldade de estimativa.

## 4. Materiais e Métodos

Essa seção objetiva descrever os materiais e métodos utilizados nesse trabalho, sendo dividida em três partes: (i) dados coletados; (ii) metodologia dos experimentos computacionais; e (iii) métricas de avaliação de desempenho.

### 4.1. Dados coletados

As séries de dados de vazão e chuva diárias estão disponíveis no banco de dados do Instituto Nacional de Meteorologia (INMET<sup>2</sup>) do Brasil. Neste caso, foi utilizada a estação pluviométrica do município de Tianguá (CE).

---

<sup>2</sup>Disponível em <https://portal.inmet.gov.br/>

O período de observação das chuvas e vazões vai de 15/03/2018 a 11/05/2021, totalizando 27655 registros por hora ao remover registro com valor nulo para último atributo “chuva (mm)” que determina a quantidade em milímetros de chuva para um período. São 19 atributos que de forma categórica representam as grandezas temperatura do ar, umidade relativa do ar, temperatura do ponto de orvalho, pressão atmosférica do ar, velocidade do vento, direção do vento, intensidade da rajada do vento e radiação solar, em que algumas dessas categorias variam entre valor instantâneo, mínimo e máximo.

Foram retirados de entrada para os modelos os atributos data, hora e radiação, por não possuírem aplicação ou conterem metade dos registros nulos como foi o caso para o terceiro excluído. Todos os registros que continham qualquer atributo vazio foram retirados, restando assim 23879 dos 27655 totais do conjunto de dados original. A Figura 1 exibe parte do *dataset* final após essa etapa de pré-processamento.

Data	Temp. Ins. (C)	Temp. Max. (C)	Temp. Min. (C)	Umi. Ins. (%)	Umi. Max. (%)	Umi. Min. (%)	Pto Orvalho Ins. (C)	Pto Orvalho Max. (C)	Pto Orvalho Min. (C)	Pressao Ins. (hPa)	Pressao Max. (hPa)	Pressao Min. (hPa)	Vel. Vento (m/s)	Dir. Vento (m/s)	Raj. Vento (m/s)	Chuva (mm)
2018-03-15 20:00:00	25.5	26.5	25.5	75.0	76.0	64.0	20.8	21.4	19.1	926.8	926.8	926.5	3.1	67.0	7.3	0.0
2018-03-15 21:00:00	23.3	25.5	23.3	83.0	85.0	75.0	20.3	21.3	20.3	927.4	927.4	926.8	2.8	41.0	5.9	0.0
2018-03-15 22:00:00	22.8	23.3	22.7	89.0	90.0	83.0	21.0	21.0	20.4	927.9	927.9	927.4	3.2	55.0	6.0	0.0

Figura 1. Conjunto de dados Tianguá (A368) de precipitações

## 4.2. Metodologia dos Experimentos Computacionais

Os experimentos computacionais foram realizados no ambiente *Google Colaboratory* utilizando a linguagem Python e os modelos de regressão do módulo *scikit-learn*<sup>3</sup>. São avaliados os seguintes modelos:

- Regressão polinomial** (*Polynomial Regression*, PR) de [Fan and Gijbels 2018]. Pelo *scikit-learn* esse modelo é implementado através do *PolynomialFeatures* junto com *LinearRegression*;
- Floresta aleatória** (*Random Forest*, RF) de [Breiman 2001]. No módulo *scikit-learn* esse método de regressão é identificado como *RandomForestRegressor*;
- Support Vector Regression Machines** (SVRM) de [Drucker et al. 1997] é implementado pelo módulo *sklearn.svm.SVR*.

Diversos experimentos combinando modelos de regressão, hiperparâmetros, métodos de pré-processamento de dados e organização do conjunto de dados foram avaliados, porém apenas três abordagens com melhores resultados são apresentadas neste artigo.

A primeira abordagem utilizou todo o conjunto de dados pré-processado, identificado neste artigo como **DF Completo**. Já a segunda abordagem dividiu o conjunto em dois subconjuntos, sendo o que **DF Chuvoso** corresponde ao período chuvoso e o **DF Seco** corresponde ao período seco assim como [Yan et al. 2021] realizaram.

<sup>3</sup>Disponível em <https://scikit-learn.org/stable/>

A terceira abordagem deu-se com a retirada da sazonalidade do conjunto de dados aplicando diferenças sucessivas da série original com base no ano. Essa operação passa a considerar a base de dados como estacionária, porém para isso foi necessário também retirar os dados do dia 15/03/2018 a 14/03/2019 por ainda ser um período não estacionário. Assim, têm-se 18649 registros do total inicial para essa abordagem. Sua denominação foi **DF Estacionário**.

Na abordagem DF Completo, o método PR teve o parâmetro *degree* atribuído com valor 2 para a transformação polinomial. Além disso, somente para esse modelo PR foram removidas as colunas do conjunto de dados com valores instantâneos. O RF foi definido com 100 árvores de decisão e a métrica MSE (*mean squared error*). O SVRM foi configurado com o kernel RBF, parâmetro *tol* = 0.0001 e dados normalizados através da técnica *z-score*. Vale salientar que foram testados diversos hiperparâmetros para cada abordagem, destacando-se aqui os que obtiveram os melhores resultados.

No caso da abordagem DF Chuvoso e DF Seco, o modelo PR foi configurado com os hiperparâmetros *degree* = 2 e *normalize* = *True*. O RF foi instanciado com 60 árvores e a métrica MSE. Por fim, o SVRM teve o kernel RBF e o *tol* = 0.0001. Para os métodos RF e SRVM, foram usados a normalização *z-score* antes da modelagem.

Para a abordagem DF Estacionário, a configuração dos hiperparâmetros do modelo PR foi *degree* = 2 e *normalize* = *True*. O RF recebeu 100 árvores e métrica MSE e o SVRM foi configurado com kernel RBF e *tol* = 0.0001. Para o RF e SRVM, a normalização Z-Score foi aplicada aos dados antes da modelagem.

### 4.3. Métricas de avaliação de desempenho

Para avaliar o desempenho dos modelos foram considerados os critérios coeficiente de determinação ( $R^2$ ), *Mean Absolute Error* (MAE) e *Root Mean Squared Error* (RMSE). Sendo que  $y_i$  é o valor observado,  $\hat{y}_i$  o valor predito pelo modelo e  $\bar{y}$  é a média da variável.

A métrica  $R^2$  apresenta valores entre 0 e 1, sendo que quanto mais próximo de 1, maior o ajuste de previsão, ou seja, esse critério diz o quanto o modelo está prevendo corretamente [de Mendonça et al. 2021, Guirado 2019].

No caso dos critérios MAE e RMSE são valores não-negativos, sendo que os valores mais baixos correspondem a melhores modelos, ou seja, melhores aproximações da predições realizadas [Faceli et al. 2011]. A diferença entre RMSE e MAE, é que o primeiro dá um peso maior para desvios grandes (pois são elevados ao quadrado), enquanto o MAE dá um peso igual a todos os desvios [Stenghele et al. 2020].

## 5. Resultados e Discussões

Conforme é apresentado pela Tabela 1, o melhor valor de  $R^2$  foi abaixo de 54%, pois os dados possuem características que dificultam as previsões via regressão, como a sazonalidade. Para avaliação dos possíveis problemas do conjunto de dados, um dos testes aplicados foi a divisão destes em um total de doze subconjuntos, um para cada mês do ano. Aplicando os mesmos métodos, no mês de janeiro com RF foi obtido  $R^2 = 79.6\%$  com apenas 20 árvores. Isso é bastante positivo dado os resultados das Tabelas 1 e 2. Ainda assim em alguns meses também aconteceram resultados ainda negativos, não gerando um grupo de modelos adequado a ser utilizado por todo o ano.

**Tabela 1. Resultados de cada modelo e abordagem**

Modelo	Abordagem	$R^2$ (%)	MAE	RMSE
PR	DF Completo	33.0	0.324	1.205
RF	DF Completo	49.8	0.172	1.043
SVRM	DF Completo	13.9	0.197	1.367
PR	DF Chuvoso	43.3	0.451	1.319
PR	DF Seco	5.7	0.207	0.448
RF	DF Chuvoso	52.2	0.265	2.211
RF	DF Seco	53.2	0.031	0.315
SVRM	DF Chuvoso	19.3	0.266	1.574
SVRM	DF Seco	2.2	0.078	0.456
PR	DF Estacionário	27.5	0.354	1.34
RF	DF Estacionário	27.4	0.247	1.341
SVRM	DF Estacionário	5.5	0.237	1.575

**Tabela 2. Média das métricas da segunda abordagem: DF Chuvoso e DF Seco**

Modelo	Abordagem	$R^2$ (%)	MAE	RMSE
PR	DF Chuvoso e DF Seco	24.5	0.329	0.884
RF	DF Chuvoso e DF Seco	52.7	0.148	1.263
SVRM	DF Chuvoso e DF Seco	10.75	0.172	1.015

O modelo que previu mais corretamente ( $R^2$ ) foi o RF em todas as abordagens (acima de 49%), com exceção do DF Estacionário. Ele também apresentou o menor valor de MAE usando a divisão em período chuvoso e seco. Com base no RMSE, o PR teve melhor desempenho. Para todos os casos de DF Chuvoso e DF Seco, foi calculada a média aritmética dos resultados de cada métrica, como mostrado na Tabela 2. O melhor método foi o RF no DF Seco para todas as métricas, conforme a Tabela 1.

O pior resultado para o problema considerando  $R^2$  ou RMSE partindo do SVRM para o MAE foi o PR, todos usando o DF Estacionário. Todavia, desconsiderando a média como mostra a Tabela 1, o pior resultado para  $R^2$  é o SVRM no DF Seco, para o MAE sendo o PR no DF Chuvoso, e para o RMSE o RF no DF Chuvoso.

Analisando os trabalhos relacionados, os resultados ainda não são satisfatórios, mesmo considerando serem outros conjuntos de dados, pois se encontram abaixo de 57,8% obtido pelos [Debastiani et al. 2019]. É preciso uma análise aprofundada dos dados da estação pluviométrica e aplicar métodos mais robustos, como o MLP usado pelos [Debastiani et al. 2019] ou mesmo NARX em [de Mendonça et al. 2021]. A metodologia aplicada por [Yan et al. 2021] pode ser um caminho a ser usado novamente, visto que trouxe o melhor desempenho.

## 6. Considerações Finais e Trabalhos Futuros

Este estudo sobre modelos hidrológicos mostraram que o método SVRM não é adequado ao conjunto de dados utilizado, pois foi o que gerou os piores modelos. O método RF

combinado a abordagem de divisão em dois períodos obteve o melhor resultado com 52.7% de média, porém ainda não é um resultado satisfatório. Este método mostrou-se como o que mais se ajustou ao problema, pois foi o que mais previu corretamente.

Uma recomendação para trabalhos futuros é avaliar junto ao RF, Redes Neurais como MLP, NARX e até *Deep Neural Networks*. Integrar isso a dados refinados e uma avaliação mais criteriosa dos atributos a serem usados, assim como buscar mais formas de atacar o problema com objetivo de obter resultados positivos.

## 7. Agradecimentos

Os autores agradecem ao IFCE pelo apoio, através do projeto PEVPI.

## Referências

- Almeida, A. Q. d., Ribeiro, A., Paiva, Y. G., Rascon Jr, N., Lima, E. P., et al. (2011). Geoestatística no estudo de modelagem temporal da precipitação. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 15:354–358.
- Ballini, R. et al. (2000). Análise e previsões de vazões utilizando modelos de séries temporais, redes neurais e redes neurais nebulosas.
- Bezerra, M. I. S. (2006). Apostila de análise de séries temporais. *UNESP: Curso de estatística. São Paulo, SP*.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.
- Brubacher, J. P., Oliveira, G. G. d., and Guasselli, L. A. (2020). Preenchimento de falhas e espacialização de dados pluviométricos: Desafios e perspectivas. *Revista Brasileira de Meteorologia*, 35:615–629.
- da Costa, E. S. (2019). Análise da série temporal de precipitação total mensal do município de Cruz das Almas-Ba. *Universidade Federal do Recôncavo da Bahia. Bacharelado em Ciências Exatas e Tecnológicas*.
- de Albuquerque, R. C. (2019). Modelagem em séries :aplicação em dados de precipitação na região do sertão de Pernambuco-Brasil. *Universidade Federal Rural de Pernambuco. Programa de Pós-graduação Em Biometria e Estatística Aplicada*.
- de Lima, P. M. and Guedes, E. B. (2015). Uma abordagem baseada em redes neurais para a previsão de chuva em Manaus, Amazonas. *Latin America Congress on Computational Intelligence (LA-CCI)*.
- de Mendonça, L. M., de Souza Gomide, I., de Sousa, J. V., and Blanco, C. J. C. (2021). Modelagem chuva-vazão via redes neurais artificiais para simulação de vazões de uma bacia hidrográfica da Amazônia. *Revista de Gestão de Água da América Latina*, 18(2021).
- Debastiani, A. B., Neto, S. L. R., and da Silva, R. D. (2019). Árvore modelo frente a uma rede neural artificial para a modelagem chuva-vazão. *Nativa*, 7(5):527–534.
- Drucker, H., Burges, C. J., Kaufman, L., Smola, A., Vapnik, V., et al. (1997). Support vector regression machines. *Advances in neural information processing systems*, 9:155–161.

- Faceli, K., Lorena, A. C., Gama, J., and Carvalho, A. (2011). *Inteligência artificial. Uma Abordagem de Aprendizado de Máquina*. Editora LTC.
- Fan, J. and Gijbels, I. (2018). *Local polynomial modelling and its applications: monographs on statistics and applied probability* 66. Routledge.
- Figueredo, C. J. (2008). Previsão de séries temporais utilizando a metodologia box & jenkins e redes neurais para inicialização de planejamento e controle da produção. *Curitiba: UFPR*.
- Gorodetskaya, Y., da Fonseca, L. G., and de Melo Ribeiro, C. B. (2018). Previsão de vazão de curto prazo utilizando redes neurais artificiais. estudo de caso: Rio paraíba do sul.
- Guirado, A. G. (2019). *Critérios robustos de seleção de modelos de regressão e identificação de pontos aberrantes*. PhD thesis, Universidade de São Paulo.
- Ingsrisawang, L., Ingsriswang, S., Somchit, S., Aungsuratana, P., and Khantiyanan, W. (2008). Machine learning techniques for short-term rain forecasting system in the northeastern part of thailand. In *Proceedings of World Academy of Science, Engineering and Technology*, volume 31. Citeseer.
- LIMA, J. R.; MAGALHÃES, A. R. (2018). Secas no nordeste: registros históricos das catástrofes econômicas e humanas do século 16 ao século 21. 23.
- Parmezan, A. R. S., Batista, G. E. d. A. P. A., et al. (2016). Descrição de modelos estatísticos e de aprendizado de máquina para predição de séries temporais.
- Santana, A. S. d. and Santos, G. R. d. (2020). Impactos da seca de 2012-2017 na região semiárida do nordeste: notas sobre a abordagem de dados quantitativos e conclusões qualitativas.
- Santos, J. R. S. and Celeste, A. B. (2014). Avaliação de estratégias de modelagem guiada por dados para previsão de vazão em rio sergipano. *Revista Ambiente & Água*, 9:488–498.
- Schroder, T. et al. (2020). Modelos chuva-vazão: Uso de técnicas de aprendizagem de máquinas para calibração de modelos em uma pequena bacia hidrográfica. *Geoambiente On-line*, (37).
- Soyapi, C. B. (2017). Water security and the right to water in southern africa: an overview. *Potchefstroom Electronic Law Journal/Potchefstroomse Elektroniese Regsblad*, 20(1).
- Stenghele, G., dos Santos, J. A. A., Miura, F. S., Pasa, L. A., and Schmidt, C. A. P. (2020). Previsão de consumo de energia elétrica na região sudeste: Um estudo de caso usando sarima/forecast of electricity consumption in the southeast region: A case study using sarima. *Revista FSA (Centro Universitário Santo Agostinho)*, 17(2):169–183.
- Sumi, S. M., Zaman, M. F., and Hirose, H. (2012). A rainfall forecasting method using machine learning models and its application to the fukuoka city case. *International Journal of Applied Mathematics and Computer Science*, 22:841–854.
- Yan, J., Xu, T., Yu, Y., and Xu, H. (2021). Rainfall forecast model based on the tabnet model. *Water*, 13(9):1272.