RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

ANALYSIS OF THE HISTORICAL SERIES OF THE WATER VOLUME OF THE ENGENHEIRO ÁVIDOS DAM IN PARAÍBA

ANÁLISE DA SÉRIE HISTÓRICA DO VOLUME DE ÁGUA DO AÇUDE ENGENHEIRO ÁVIDOS NA PARAÍBA

ANÁLISIS DE LA SERIE HISTÓRICA DEL VOLUMEN DE AGUA EN LA REPRESA ENGENHEIRO ÁVIDOS EN PARAÍBA

Ana Clara da Silva Morais[1], Nyedja Fialho Morais Barbosa[2], Erika Fialho Morais Xavier[3], Silvio Fernando Alves Xavier Júnior[4], Jader da Silva Jale[5]

e463438
https://doi.org/10.47820/recima21.v4i6.3438

RECEIVED: 05/07/2023 APPROVED: 06/07/2023 PUBLISHED: 06/23/2023

ABSTRACT
The Northeast region of Brazil has a climate that is different from other regions of the country, being influenced by several climatic factors that directly contribute to long periods of drought in the region. Furthermore, due to the scarcity of water resources, it is customary to build dams to guarantee water supply to the most varied sectors during critical periods. In this sense, the present work analyzed the time series of the Engenheiro Ávidos dam, located in the hinterland of Paraíba, between January 2001 and December 2022 with data available on the AESA portal. The R 4.1.2 software was used through the "forecast" package using the "auto. arima" function for data processing. The model that showed the highest accuracy was ARIMA (2,1,1), where it was possible to find possible estimates for future series values until June 2023.

KEYWORDS: Time series. ARIMA model. Prediction. Reservoir.

RESUMO
A região Nordeste do Brasil possui climatologia diferenciada das demais regiões do país, sofrendo influência de vários fatores climáticos que contribuem diretamente para longos períodos de seca na região. E virtude da escassez de recursos hídricos, costuma-se recorrer a construções de represas na tentativa de garantir o abastecimento de água para os mais variados setores durante os períodos críticos. Neste sentido, o presente trabalho analisou a série temporal do açude Engenheiro Ávidos, localizado no sertão paraibano, entre janeiro de 2001 e dezembro de 2022 com dados disponibilizados no portal da AESA, com o objetivo de encontrar o melhor modelo para descrever o comportamento da série, tendo em vista fazer predições para meses à frente. Para o processamento dos dados foi utilizado o *software* R 4.1.2 por meio do pacote *"forecast"* utilizando a função "auto.arima". O modelo que apresentou maior acurácia para modelar a série histórica do açude Engenheiro Ávidos foi o ARIMA (2,1,1), e por meio deste foi possível encontrar possíveis estimativas para valores futuros da série até junho de 2023.

PALAVRAS-CHAVE: Séries temporais. Modelo ARIMA. Predição. Reservatório.

[1] Aluna de graduação do curso de Ciências Biológicas da Universidade Estadual da Paraíba (UEPB), campus V.
[2] Universidade Estadual da Paraíba.
[3] Bacharel em Estatística pela Universidade Estadual da Paraíba, com mestrado em Biometria e Estatística Aplicada pela Universidade Federal Rural de Pernambuco e doutorado em Biometria e Estatística Aplicada pela Universidade Federal Rural de Pernambuco. Pesquisadora de pós doutorado em Estatística no Cidacs/Fiocruz Bahia.
[4] Licenciado em Matemática (UFPE). Mestrado em Biometria e Estatísitica Aplicada (UFRPE). Doutorado em Biometria e Estatística Aplicada (UFRPE). Estágio sanduíche na Texas A & M University, United States, Biological and Agricultural Engineering Department. Coordenador do curso de Estatística (CCT/UEPB), presidente do colegiado do curso de Estatística. Presidente do Núcleo Docente Estruturante. Chefe adjunto do Departamento de Estatística. Membro do PROFMAT - UEPB.
[5] Professor Adjunto A no DEINFO-UFRPE. Bacharelado em Estatística pela Universidade Federal do Ceará. Mestrado e Doutorado em Biometria e Estatística Aplicada pela Universidade Federal Rural de Pernambuco. Pós-Doutorado em Ciência da Computação pela Universidade Federal Rural de Pernambuco.

RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

RESUMEN
La región del Nordeste de Brasil tiene un clima que es diferente de otras regiones del país, siendo influenciado por varios factores climáticos que contribuyen directamente a los largos períodos de sequía en la región. Y debido a la escasez de recursos hídricos, se acostumbra a recurrir a la construcción de represas en un intento de garantizar el abastecimiento de agua a los más variados sectores en períodos críticos. En ese sentido, el presente trabajo analizó la serie temporal de la represa Engenheiro Ávidos, ubicada en el sertão de Paraíba, entre enero de 2001 y mayo de 2023 con datos disponibles en el portal AESA, con el objetivo de encontrar el mejor modelo para describir el comportamiento de la serie, con vistas a hacer predicciones para meses venideros. Para el procesamiento de datos se utilizó el software R 4.1.2 a través del paquete "forecast" utilizando la función "auto.arima". El modelo que presentó mayor precisión para modelar la serie histórica de la presa Engenheiro Ávidos fue el ARIMA (2,1,1), a través del cual fue posible encontrar posibles estimaciones para valores futuros de la serie hasta junio de 2023.

PALABRAS CLAVE: Series de tiempo. Modelo ARIMA. Predicción. Reservorio.

1 INTRODUCTION

According to Silva and Vieira (2017), Brazil is recognized for having the most remarkable diversity of water on the planet, corresponding to 12% of all freshwater. However, despite this, there are many problems related to water scarcity, especially in the Brazilian Northeast, where the population often suffers from the lack of this resource due to the climatic aspects of the region (CARDOSO et al., 2012).

The Brazilian Northeast region is affected by long periods of drought due to its climatic aspects. Given this, the solution found by the government to mitigate the problems of water scarcity in the Northeast was the creation of dams (OLIVEIRA, 2017; SANTOS, 2018). However, more than just building these reservoirs is needed to meet all the population's needs; it is also essential to improve the management of water distribution (CELESTE, 2006).

According to Agência Executiva de Gestão de Águas da Paraíba (AESA), there are currently 135 active dams in Paraíba whose maximum capacity varies between 255,744 m^3 and 744,144,694 m^3. The Engenheiro Ávidos dam, the fourth dam in Paraíba in terms of maximum water storage capacity (293,617,376.00 m^3), located in the city of Cajazeiras – Sertão Paraíba, is responsible for supplying water to the municipality of Cajazeiras, Nazarezinho, Gravatá, and the surrounding rural area, corresponding to the supply of approximately 75 thousand inhabitants (SILVA, 2018). The reservoir, also known as Boqueirão de Piranhas, receives and dams the waters coming from the Piranhas River, whose source is located in the municipality of Bonito de Santa Fé (OLIVEIRA, 2017).

This work used volumetric data from the Engenheiro Ávidos dam between January 2001 and May 2023, available on the AESA portal, to find an adequate prediction model to model the historical series and point out estimates for possible volumetric levels of the reservoir until May 2024.

2 METHODOLOGY

Time series is information on a variable ordered over a fixed time interval, where it is possible to observe its behavior and find a model that best fits the observed data and make predictions for

RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

future values. The components of a time series are the trend, seasonality, cycles, and random errors. The models used to explain the behavior of a time series are stochastic processes. Thus, we consider a family $Z = \{Z(t), t \in T\}$ where $Z(t)$ is a random variable for each $t \in T$ (MORETTIN; TOLOI, 2006).

According to Gujarati and Porter (2011), Auto Regressive Integrated Moving Average (ARIMA) modeling is one of the most used to make predictions of the behavior of time series due to the reliability of its results. Such a model is presented as ARIMA(p, d, q), where: p represents the autoregressive order of the model; d refers to the degree of differentiation, and q to the number of moving averages.

Morettin and Toloi (2006) explain that when stochastic seasonality is present in the series, it is possible to use the Seasonal Autoregressive Integrated Moving Average (SARIMA), SARIMA$(p, d, q) \times (P, D, Q)$ model, which is an extension of the ARIMA(p, d, q) model where:

- P is the number of seasonal autoregressive coefficients related to the stationarity of the series;
- D represents the number of seasonal differences to make the series stationary; and
- Q is the number of seasonal coefficients of moving averages.

If the model is well adjusted, one should expect the residuals to be randomly distributed around zero, with constant variance. According to Bueno (2011), to complement the analysis of the chosen model, it is recommended to use the Ljung-Box test. The test statistic is given by (EHLERS, 2007):

$$Q = n(n+2)\sum_{k=1}^{m} \frac{r_k^2}{n-k} \chi^2{}_{(m-p-q)} \quad (1),$$

where r_k is the autocorrelation of the series at lag k, and m is the number of tested lags. Thus, when the null hypothesis is not rejected, the model does not exhibit fit failure. From the chosen and validated model, it is possible to make predictions for future values of the series.

2.1 Data Processing

The data used in this work were collected on the website of the Agência Executivo de Gestão de Águas da Paraíba (AESA), whose observations referred to monthly measurements of water levels in the Engenheiro Ávidos reservoir between January 2001 and May 2023. The database referring to the time series of the chosen reservoir was processed in the R.4.1.2 software using the forecast package, using the functions:

1. "autoplot", to plot the series;
2. "decompose"; for the graph of the series decomposition;
3. "auto.arima" (HYNDMAN; KHANDAKAR, 2008), to find the best model for the series;
4. "checkresiduals" for analysis of residuals;
5. "forecast", to make predictions for future values of the series.

RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

3 RESULTS AND DISCUSSION

The use of the time series technique to model water data has been growing in the scientific community, such as the work of Morais and Barbosa (2022a) and Morais and Barbosa (2022b), who used the technique mentioned above to model the Coremas and Mãe D'dams water, and the Epitácio Pessoa dam, respectively, both in the State of Paraíba.

The two previously mentioned dams are the largest in terms of maximum water storage capacity in the State, followed by the Engenheiro Ávidos dam, the target of this work. The reservoir in question is located in the high hinterland of Paraíba, in the city of Cajazeiras (SILVA, 2018).

The time series of the volumetric behavior of the dam between January 2001 and May 2023 can be seen in Figure 1.

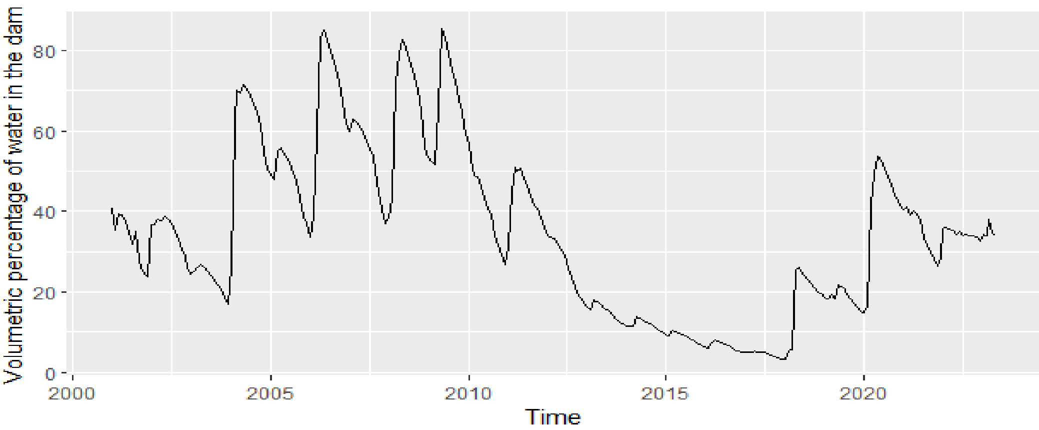

Figure 1 – Time series of the volumetric percentage of the Engenheiro Ávids dam from January 2001 to May 2023

The average value of the volumetric percentage was around 35.10%, with a standard deviation of 21.1%. The weir registered the highest volume of the series in May 2009, corresponding to 85.28% of the reservoir's total capacity. In January 2018, it reached 3.24%, the lowest value recorded during the observed period. Such results may be linked to El Nino Soutern Oscilation (ENSO) climate anomalies, which influence the rainfall regime in Northeast Brazil, as analyzed for the Coremas and Mãe D'Água dams by Morais and Barbosa (2022a). Nevertheless, it is not the only explanation for the variation in the water level of the reservoir; several factors contribute to the volumetric oscillation of water in the dams, such as climate, evaporation, silting, precipitation, as well as the withdrawal of water for population supply and irrigation of crops (FEITOSA et al.; 2021; ARAÚJO, 2017; FRANCISCO; SANTOS, 2017).

The graph of the decomposition of the time series for the Coremas dam can be seen in Figure 2, where trends, seasonality, and stochasticity are observed.

RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

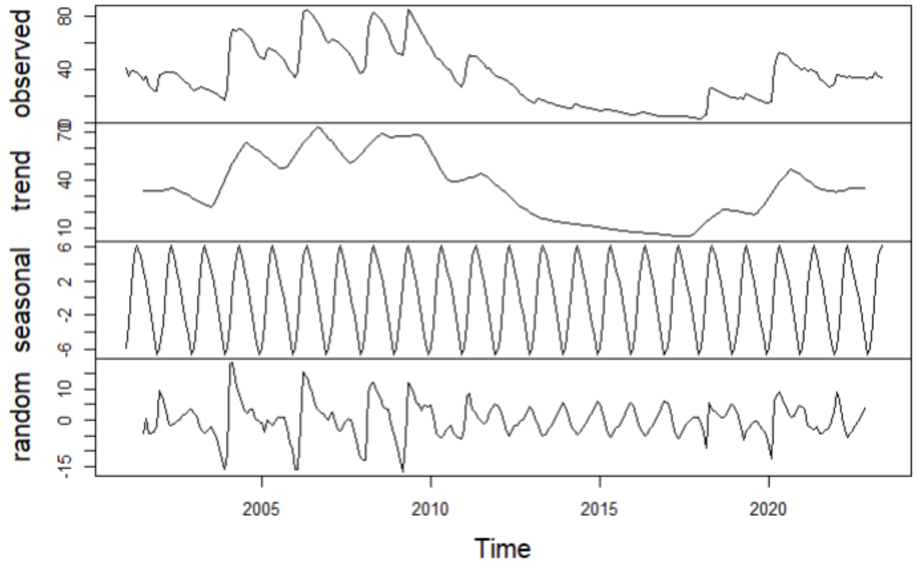

Figure 2 – Time series decomposition of the Engenheiro Ávidos dam

The 'auto.arima' function of the R forecast package automatically selected the best mathematical model for the reservoir based on the AIC information criterion calculated for each tested model (Table 1).

Table 1 - Models tested for modeling the time series of the Engenheiro Ávidos dam

Model	AIC	Model	AIC
ARIMA(0,1,0)	1622.195	**ARIMA(2,1,1)**	**1555.459**
ARIMA(2,1,1)(1,0,0)	1557.459	ARIMA(2,1,1)(0,0,1)	1557.464
ARIMA(2,1,1)(1,0,1)	1559.535	ARIMA(1,1,1)	1566.413
ARIMA(2,1,0)	1565.596	ARIMA(3,1,1)	1557.51
ARIMA(2,1,2)	1557.503	ARIMA(1,1,0)	1567.098
ARIMA(1,1,2)	1567.918	ARIMA(3,1,0)	1566.35

Of the tested models, the one with the best accuracy is the ARIMA (2,1,1) (Table 1); since it presented the lowest AIC value, it will be used to make predictions for future series values. However, before using it for this purpose, it is necessary to validate it through the diagnosis of residues (Figure 3).

RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

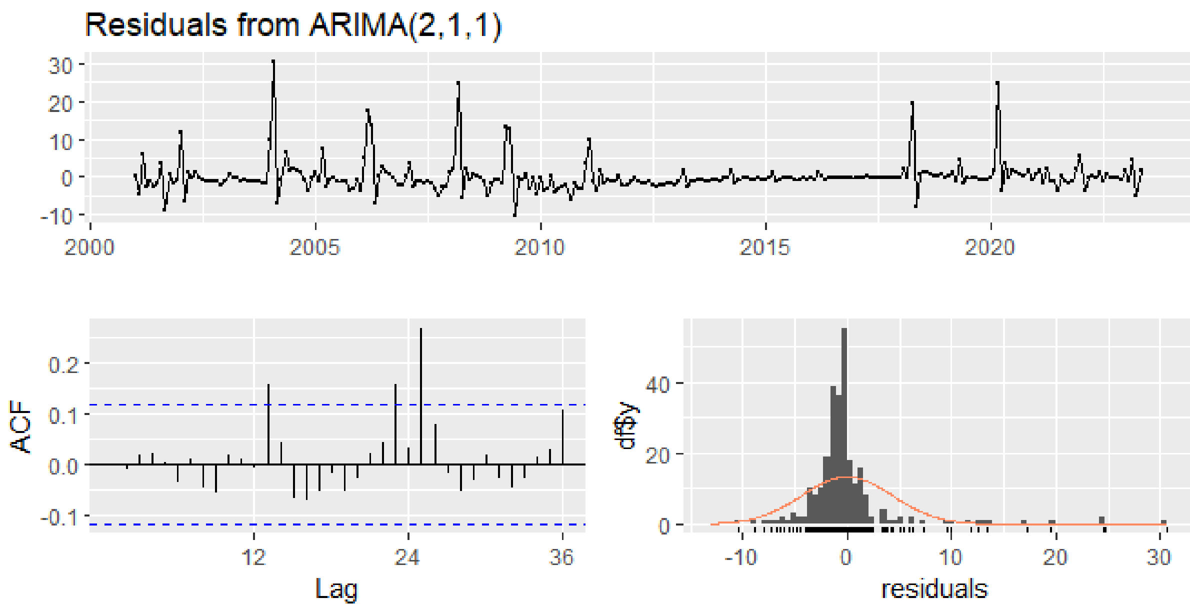

Figure 3 - Analysis of residuals from the ARIMA model (2,1,1)

Through the residual graph (Figure 4), it is possible to verify that the assumptions were satisfied, making the ARIMA model (2,1,1) valid for estimating future values of the series. Such validation can also be confirmed by the Ljung-Box test (Q=23.252; p=0.3307). Because of this, the validated model was used to predict the possible volumetric percentages of the Engenheiro Ávidos reservoir until May 2024, whose results can be seen in Table 2.

Table 2 - Forecast for future values of the volumetric percentage of the Engenheiro Ávido reservoir for 1(one), 6(six), and 12(twelve) steps ahead

Month	**Mean**	**Confidence Interval 80%**		**Confidence Interval 95%**	
June 2023	34.17	28.58	39.76	25.63	42.72
Nov 2023	34.28	16.23	52.34	6.67	61.9
May 2024	34.38	13.33	55.43	2.18	66.57

Based on the chosen model, the predicted behavior for the series until May 2024 can be seen below (Figure 4).

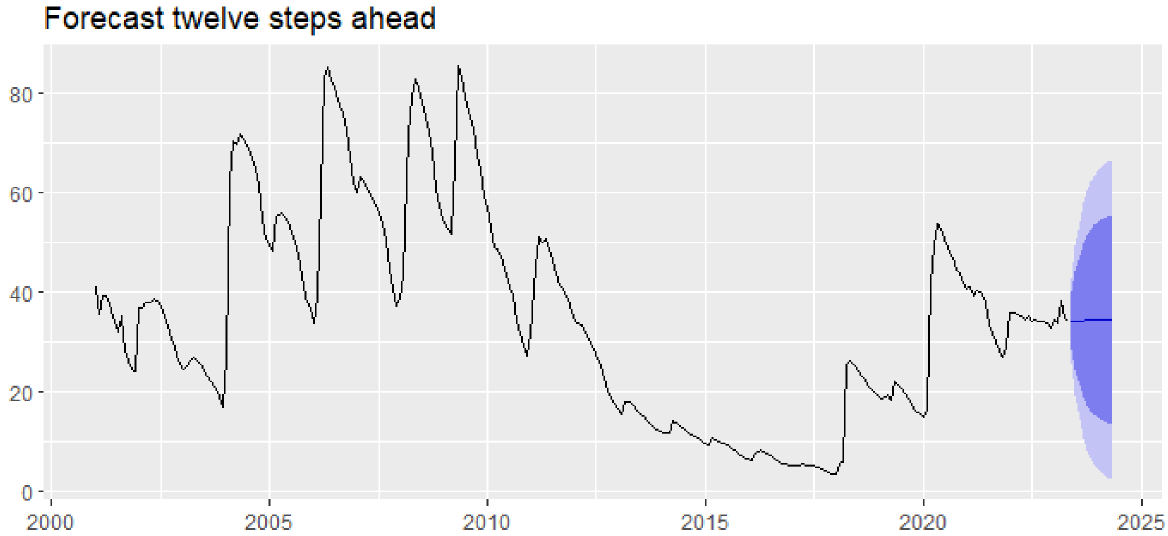

Figure 4 – Forecast for the historical series of the Engenheiro Ávidos dam until May 2024

Observing Figure 4, one can see that the volume of the Engenheiro Ávidos dam seems to remain almost constant until May 2024.

4 CONCLUSION

Through the study of the time series of volumetric data from the Engenheiro Ávidos dam, it was possible to find and validate the ARIMA model (2,1,1) to explain the behavior of the series and make predictions until May 2024, where according to the results obtained, a conservative pattern is expected for the series with few volumetric variations in the dam. For future work, it is possible to expand the study to more reservoirs, comparing prediction models for time series using artificial neural network models.

REFERENCES

ARAÚJO, J. C. de. Assoreamento em reservatórios do semi-árido: modelagem e validação. **Revista Brasileira de Recursos Hídricos**, v. 8, n. 2, p. 39-56, 2003.

BUENO, R. de L. da S. **Econometria das séries temporais**. 2. ed. São Paulo: Cengarge Learning, 2011.

CARDOSO, M. M. L. *et al*. Diversidade de peixes em poças de um rio intermitente do semiárido paraibano, Brasil. **Biotemas,** v. 25, n. 3, p.161-171, 2012.

CELESTE, A. B. **Determinação e análise de indicadores de desempenho e de sustentabilidade de seis açudes na bacia do rio Piancó - PB e de suas potenciais demandas**. 2006. Dissertação (Mestrado em Engenharia Civil e Ambiental) – Universidade Federal de Campina Grande, Campina Grande, PB. 2006.

RECIMA21 - REVISTA CIENTÍFICA MULTIDISCIPLINAR
ISSN 2675-6218

EHLERS, R. S. **Análise de séries temporais**. 4. ed. São Paulo: USP, 2007. Disponível em: http://www.each.usp.br/rvicente/AnaliseDeSeriesTemporais.pdf. Acesso em: 01 jun. 2022.

FEITOSA, G. P.; ARAÚJO, J. C. de; BARROS, M. U. G. Different methods for measuring evaporation in a tropical reservoir: the case of the Gavião reservoir in the state of Ceará. **Revista Caatinga**, v. 34, n. 2, p. 410-421, 2021.

FRANCISCO, P. R. M.; SANTOS, D. **Climatologia do Estado da Paraíba.** Campina Grande: Editora da Universidade Federal de Campina Grande - EDUFCG, 2017.

GUJARATI, D. N.; PORTER, D. C. **Econometria básica**. 5. ed. Porto Alegre: AMGH Editora Ltda., 2011.

HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for R. **Journal of Statistical Software**, v. 27, n. 1, p. 1–22, 2008.

MORAIS, A. C. da S.; Barbosa, N. F. M. Análise de dados volumétricos dos açudes Coremas e Mãe D'água da Paraíba. **Research, society and development**, v. 11, n. 10, e49111032397, 2022a.

MORAIS, A. C. da S.; Barbosa, N. F. M. Modelagem de dados volumétricos do açude Epitácio Pessoa (Boqueirão). **Research, society and development**, v. 11, n. 11, e40911113372, 2022b.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: Egard Blucher, 2006.

OLIVEIRA, I. S. de. **Construção do açude Engenheiro Ávidos: narrativas silenciadas das vítimas da seca e do progresso (São José de Piranhas/PB, 1920-1980)**. 2017. Trabalho de conclusão de curso (Graduação em história) – Universidade Federal da Paraíba, João Pessoa, PB, 2017.

SANTOS, E. P. dos. **Análise temporal e histórica da distribuição dos reservatórios superficiais no Estado da Paraíba – BR.** 2018. Trabalho de conclusão de curso (Graduação em geografia) – Universidade Federal da Paraíba, João Pessoa, PB, 2018.

SILVA, E. L. da; VIEIRA, A. S. Simulação integrada dos recursos hídricos nos reservatórios de Engenheiro Ávidos e São Gonçalo na Paraíba. **Revista verde de agroecologia e desenvolvimento sustentável**, v. 12, n. 5, p. 892-901, 2017.

SILVA, S. A. da. **Alocação da água do reservatório Engenheiro Ávidos utilizando um sistema de suporte à decisão**. 2018. Trabalho de conclusão de curso (Bacharelado em Engenharia Ambiental) – Universidade Federal de Campina Grande, Pombal, PB. 2018.