Kevin Fernandez

Changes in the Xingu River flooded vegetation caused by Belo Monte Dam: an approach using cointegration analysis and remote sensing images

Thesis/Dissertation submitted to the Department of Geosciences of the University of São Paulo as a partial requirement to obtain the Master of Science degree.

Area of Concentration: Geochemistry and Geotectonics
Orientador(a): Prof. Dr. Fabiano do Nascimento Pupim

São Paulo

2023

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

Changes in the Xingu River flooded vegetation caused by Belo Monte Dam: an approach using cointegration analysis and remote sensing images

KEVIN ANTHONY FERNANDEZ MOLINA

Orientador: Prof. Dr. Fabiano do Nascimento Pupim

Dissertação de Mestrado

Nº 922

COMISSÃO JULGADORA

Dr. Fabiano do Nascimento Pupim

Dra.Breylla Campos Carvalho

Dr. Carlos Henrique Grohmann

SÃO PAULO
2023

Serviço de Biblioteca e Documentação do IGc/USP
Ficha catalográfica gerada automaticamente com dados fornecidos pelo(a) autor(a) via programa desenvolvido pela Seção Técnica de Informática do ICMC/USP

Bibliotecários responsáveis pela estrutura de catalogação da publicação:
Sonia Regina Yole Guerra - CRB-8/4208 | Anderson de Santana - CRB-8/6658

Fernandez Molina, Kevin Anthony
Changes in the Xingu River flooded vegetation caused by Belo Monte Dam: an approach using cointegration analysis and remote sensing images Sao Paulo 2023 / Kevin Anthony Fernandez Molina; orientador Fabiano Nascimento Pupim. -- São Paulo, 2023.
39 p.

Tese (Doutorado - Programa de Pós-Graduação em Geoquímica e Geotectônica) -- Instituto de Geociências, Universidade de São Paulo, 2023.

1. Time series. 2. Ecosystem. 3. Cointegration. 4. Dam impact. 5. Normalized difference vegetation index. I. Nascimento Pupim, Fabiano, orient. II. Título.

RESUMO

A operação da usina hidrelétrica de Belo Monte causou impactos significativos no ciclo hidrológico natural do rio Xingu. A hipótese deste trabalho sugere que a construção da usina de Belo Monte causou um impacto significativo no ecossistema, principalmente por meio de mudanças na dinâmica hidrológica, que, por sua vez, induziram alterações na vegetação da planície de inundação da região. Para testar essa hipótese, o estudo teve como objetivo determinar a existência de uma relação causal entre os dados da série temporal do Índice de Vegetação por Diferença Normalizada (NDVI) e o nível de água do rio (NCA), usando o modelo autorregressivo de defasagens distribuídas (ARDL). Como resultado, foi obtida uma relação causal estrutural entre essas séries, o que indica uma relação dinâmica entre o NDVI e o NCA que pode gerar efeitos de longo prazo, como variações na biodiversidade, mudanças nas estruturas florestais e até mesmo alterações no ecossistema ao redor do rio Xingu.

Palavras-chave: série temporal; ecossistema; cointegração; barragem; NDVI

Abstract

The operation of the Belo Monte hydroelectric plant has had significant impacts on the natural hydrological cycle of the Xingu River. The hypothesis of this work suggests that the construction of the Belo Monte plant has caused a significant impact on the ecosystem, mainly through changes in hydrological dynamics, which in turn have induced alterations in the region's floodplain vegetation. To test this hypothesis, the study aimed to determine the existence of a causal relationship between the time series data of the Normalized Difference Vegetation Index (NDVI) and the river water level (NCA), using the autoregressive model of Distributed Lags (ARDL). Where a structural causal relationship between these series was obtained as a result, which indicates a dynamic relationship between the NDVI and the NCA which can generate long-term effects such as variations in biodiversity, changes in forest structures and even alterations in the ecosystem around the Xingu river.

Keywords: time series; ecosystem ;cointegration; dam impact; NDVI

List of Figures

List of Tables

Summary

Chapter 1

Introduction

The implementation of the Belo Monte dam on the Xingu River has been the subject of intense debate in recent decades, due to its great impact on biodiversity and on the indigenous population that inhabit the affected area. It is one of the largest hydroeletric plant in the world, which has been built in one of the most important ecosystems on the planet, the Amazon forest[1]. Regardless of the contribution to national economic development, the ecological and environmental influence has attracted significant attention from researchers, due to the damage it has generated in the environment, in the fauna and flora of the areaJiang, Xiandie and Lu, Dengsheng and Moran, Emilio and Calvi, Miquéias Freitas and Dutra, Luciano Vieira and Li, Guiying (2018) and Mayer, Adam and Lopez, Maria Claudia and Moran, Emilio F (2022).

In addition, another study published by Cunha, Denise de Andrade and Ferreira, Leandro Valle (2012) found that the construction of hydroelectric dams can have a far-reaching negative impact on the biodiversity of rivers and floodplains. This study found that the construction of dams can alter the habitat of fish and other aquatic organisms, which can have a cascading effect on the food chain and biodiversity in general [2]. In addition to this, the construction of dams can alter the flow of water and consequently the sedimentation, which can affect the quality of the water and the health of aquatic ecosystems.

[1] https://es.mongabay.com/2018/10/el-legado-de-belo-monte-el-dano-de-la-represa-del-amazonas-no-acabo-con-su-construccion/

[2] https://aida-americas.org/es/haciendo-que-brasil-se-responsabilice-por-los-da-os-de-la-represa-belo-monte

Consequently, it is important to note that the construction and operation of said dam has negative impact on the environment and local communities see Castro-Diaz, Laura and Lopez, Maria Claudia and Moran, Emilio (2018). Therefore, the hypothesis of this study states that the construction of the Belo Monte power plant has had a significant impact on the ecosystem, which is caused by the variation of the hydrological dynamics, which in turn has induced changes in the flooded vegetation in the region of the Vuelta Grande of the Xingu River. That is why to test this hypothesis, the existence of a causal relationship between data from time series of the Normalized Difference Vegetation Index (NDVI) and the river water level (NCA) will be determined, in a period that covers the before and after the construction of the plant.

In summary, the hypothesis of this work intends to study the variation in nature after the alteration of the hydrological dynamics of the Xingu River that was caused by the implantation of the Belo Monte hydroelectric plant.Therefore, to evaluate this hypothesis, it is necessary to analyze data from the NCA in the Xingu River and data from the NDVI in a time series that occupies periods before and after the implementation of hydroelectric power. To do this, it is intended to evaluate whether there is a short- and long-term relationship between these series, using the autoregressive Distributed Lags (ARDL) model, through which we seek to estimate a causal relationship between these variables. In addition, we will work statistically with geolocated data, which involves analyzing and visualizing geographic information to identify patterns and frequencies through data collection, cleaning, and preparation, exploratory analysis, and selection of statistical methods.

1.1 Related literature

There are several studies that have investigated the relationship between NDVI and NCA. Some of these studies have used remote sensing data, such as satellite imagery, to analyze the spatial and temporal relationship between NDVI and NCA Xiao, Xiangming and Zhang, Qingyuan and Braswell, Bobby and Urbanski, Shawn and Boles, Stephen and Wofsy, Steven and Moore III, Berrien and Ojima, Dennis (2004a). Other studies have used in situ data, water level measurements, and vegetation sampling, to examine the relationship between NDVI and river water level at different spatial and temporal scales Jiang et al. (2015). On the other hand, regarding the short-term relationship, some studies have found a positive short-term relationship between NDVI and the water level in rivers. For example, during flood events, NDVI can increase due to increased soil moisture and water availability for vegetation Xiao, Xiangming and Zhang, Qingyuan and Braswell, Bobby and Urbanski, Shawn and Boles, Stephen and Wofsy, Steven and Moore III, Berrien and Ojima, Dennis (2004b). However, this relationship can vary depending on the type of vegetation and local environmental conditions. In addition to this, the long-term relationship between NDVI and NCA in rivers may be more complex, for example, changes in land use and vegetation cover can affect the flow and quality of water in rivers, which in turn can influence NDVIJiang et al. (2015). Likewise, factors such as climate change and interannual variability in precipitation can affect both NDVI and NCA over time Xiao, Xiangming and Zhang, Qingyuan and Braswell, Bobby and Urbanski, Shawn and Boles, Stephen and Wofsy, Steven and Moore III, Berrien and Ojima, Dennis (2004a).

In all the cited references above, various statistical models and methods were used, including linear and non-linear regression models, spatial and temporal correlation analysis. Unlike these models, however, the model we use in this paper, the ARDL model, leads to an analytically manageable framework where parameters can be easily estimated from empirical data, and in turn models time series with different integration orders which is not possible with conventional cointegration methods.

The utility of this type of model, ARDL, occurs when the variables of interest have different integration orders, that is, when some are stationary and others are not, which is mentioned in Pesaran, M Hashem and Shin, Yongcheol and Smith, Richard J (2001). Because of this, the ARDL model differs from other time series models in several aspects. First, it allows greater flexibility in modeling the dynamics of adjustment of the variables. In other words, you can capture how variables respond to changes in other variables over time. Second, it can be applied even when the time series variables are I(0), I(1) or a mixture of both, which is not possible in other models such as the vector error correction model (VECM) which requires all variables to be I(1). Third, it uses a special approach to estimate the long-run relationship using the cointegration technique and the short-run relationships using the residuals from the first step. Finally, the ARDL model is capable of providing robust long-term parameter estimates, even in small samples, which can be a significant advantage in many research contexts. Next, for a better understanding of the model, a flowchart was generated with the name of the figure 1.1, which explains the dynamics of the ARDL and its main components.

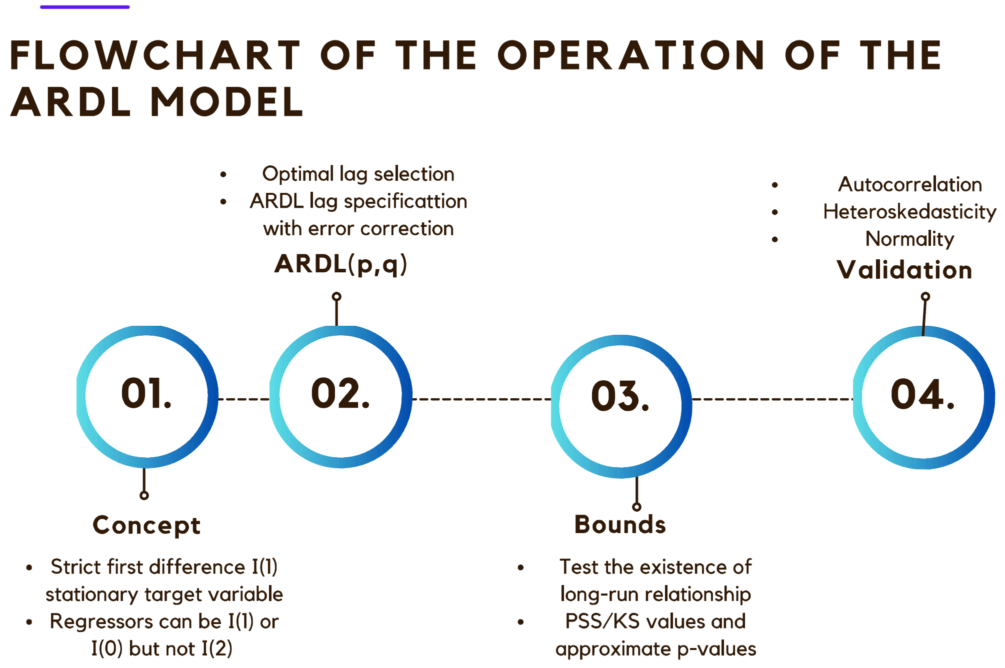

Figure 1.1: ARDL model simulation flowchart

1.2 *General objective and specific objectives*

Our general objective is to identify the causal relationship, conintegration analysis, between the time series of the water level in the Xingu River (NCA) and the time series of the normalized difference vegetation index (NDVI) using the ARDL model. In addition, using this structural relationship to understand the effects of the variation in the water level of the Xingu River on the biodiversity and ecosystem of the areas involved.

Specific objectives: **i)** Organize and compile the spatial and hydrological information available, building a georeferenced data bank, with free distribution products. **ii)** Apply various image processing techniques to data from the Landsat satellite, to obtain NDVI information. **iii)** Analyze the temporal characteristics of the NDVI and NCA series; such as stationarity, order of integration and autocorrelogram. **iv)** Modeling the short- and long-term relationships between both series using the ARDL model.

Outline

Our article is organized as follows. In section 2 we describe the study area as well as the environments in drought and floods. In section 3 we present the materials and methods, in this chapter you can see the collection and preprocessing of data as well as the analysis of the ARDL model. In section 4 resents the results of our data after the established analysis, as well as its discussion and final conclusions.

Chapter 2

Study Area

The study was carried out in the Xingú River, one of the largest tributaries of the Amazon River, which runs through the states of Mato Grosso and Pará. It lies between the southern latitude 11° 56' 19" S and the west longitude 53° 32' 48" O, see Figure 2.1. This basin covers an area of 531.25 km^2, with an elongated shape 1450 km long and about 350 km in average width.

The Xingu River is home to a diversity of vegetation types that play a crucial role in the river ecosystem. On its shores, you can find different types of vegetation that adapt to the unique conditions of this aquatic environment. One of the most outstanding plant formations is the tropical forest, known for its exuberant biodiversity and its vital role in regulating the global climate according to Acuña, Isaías Tobasura (1996). The Xingu River Rainforest is home to a wide variety of tree, plant and animal species, creating a vibrant and complex ecosystem.

In addition to the tropical forest, there are also mangroves, which is known to be home to trees adapted to fluctuations in water level. These Amazonian ecosystems vary considerably with respect to hydrology, water and soil fertility, vegetation cover, diversity of plant species, and primary and secondary productivity, creating a fascinating landscape full of life along the river Xingu Junk, Wolfgang J and Piedade, Maria Teresa Fernandez and Schöngart, Jochen and Cohn-Haft, Mario and Adeney, J Marion and Wittmann, Florian (2011). Another type of vegetation found in the Xingu River is water grasslands. These meadows, composed mainly of aquatic plants such as duckweed and water lilies, are essential for the health of the river ecosystem. Because they not only provide shelter and food for many aquatic species, but also help maintain water quality by filtering nutrients and reducing shoreline erosion.

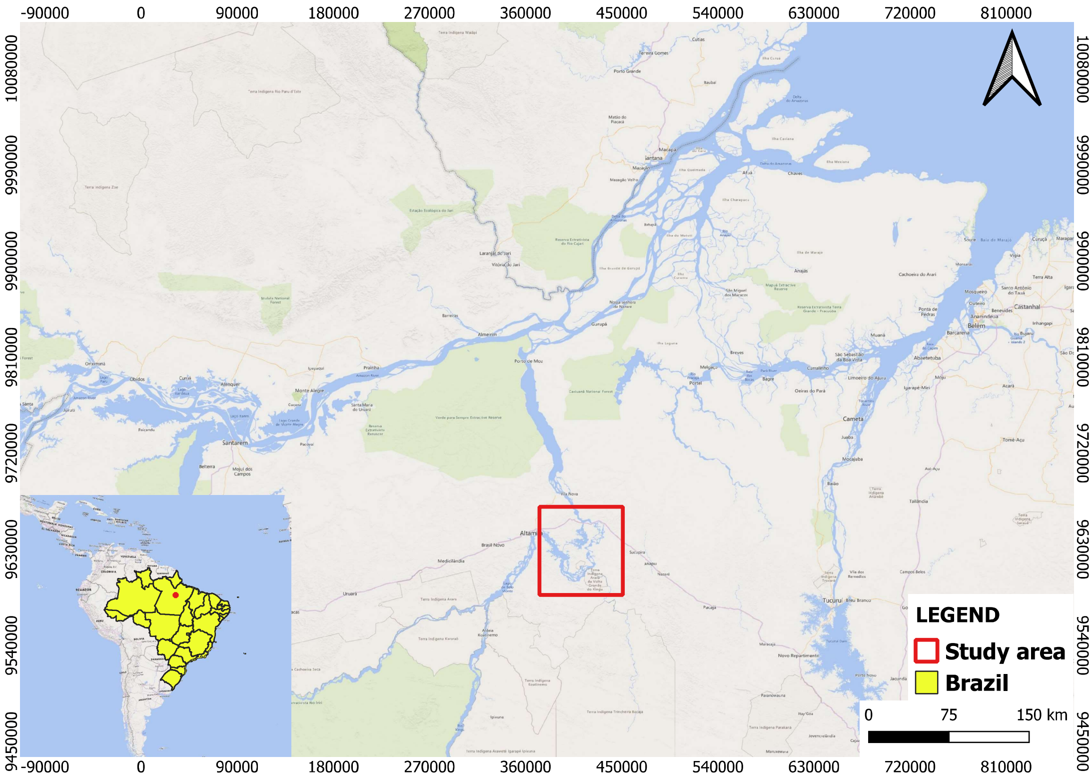

Figure 2.1: Location map of the study area

Another factor that plays an important role is the natural hydrological cycle, which is a fundamental process in the dynamics of rivers and their environment. In the case of the Xingu River, its flow, water level and seasonality play a crucial role in the dynamics of the surrounding vegetation.

The hydrological cycle begins with the evaporation of water at the surface. In the case of the Xingu River, evaporation occurs mainly in its basin and in the water bodies that feed it. This vapor condenses to form clouds that are driven by the winds that move over the Xingu River basin and discharge their contents as precipitation. The amount of precipitation can vary throughout the year, resulting in different water levels and flow in the river.

These seasonal variations in the flow and water level of the Xingu River could have a significant impact on the dynamics of the vegetation, as can be seen in the figure 2.2. According to Baker, Jessica CA and Garcia-Carreras, Luis and Gloor, Manuel and Marsham, John H and Buermann, Wolfgang and da Rocha, Humberto R and Nobre, Antonio D and de Araujo, Alessandro Carioca and Spracklen, Dominick V (2021), the variations in the flow of the Xingu river are related to the seasonality of rainfall in the basin. For this reason, it is important to highlight that the natural hydrological cycle can be affected by external factors, such as climate change and human intervention, which can alter precipitation patterns and, consequently, affect the dynamics of vegetation in the Xingu River and its surroundings[1].

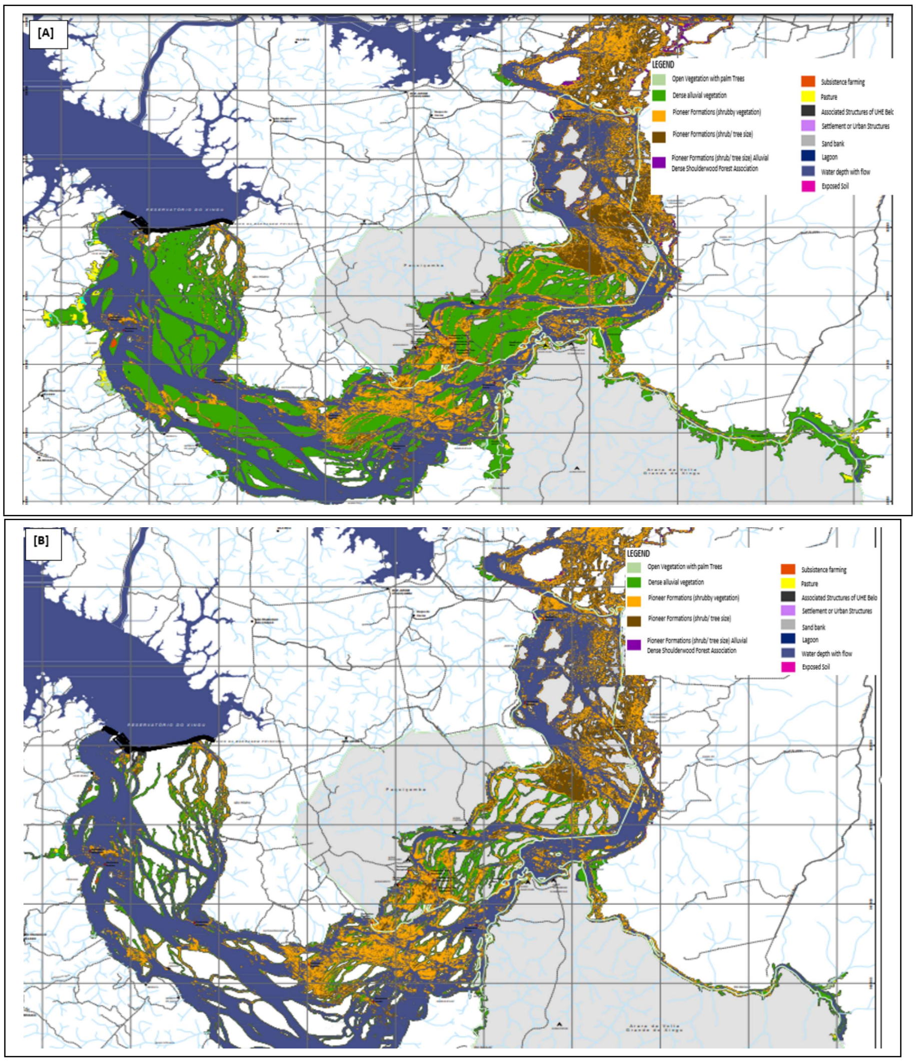

Figure 2.2: vegetation zones in the study area when the water level rises (A) and when the water level falls (B)

[1] https://www.socioambiental.org/noticias-socioambientais/solucao-para-o-caos-ambiental-de-belo-monte-esta-na-mesa-do-ibama

Chapter 3

Material and Methods

3.1 Data collection and Preparation

Using digital analysis techniques on satellite images and the information extracted from the study area, multiple vegetation units were acquired. for which the normalized difference vegetation index (NDVI) in wetlands and the river water level (NCA) was concluded. To determine the NDVI, Landsat images with a resolution of 30 meters were used using the cit2 platform of Google Earth Engine as a tool[1], for which calculation algorithms were applied to obtain satellite images considering the dates and cloudiness, which is expressed by the formulation

$$NDVI = \frac{\rho NIR - \rho RED}{\rho NIR + \rho RED}, \tag{3.1}$$

where ρRED is the fraction of the electromagnetic spectrum in the range of 0.6-0.7 μm and ρNIR is the near-infrared portion of the electromagnetic spectrum range from 0.75-1.5 μm, for more detail on the equation (3.1) consult Price, John C (1994); Tucker, Compton J (1979). This index is based on the absorption of vegetation in the red and infrared bands of the Landsat satellite.

Once the image has been generated with the NDVI filter in the study area, the process of extraction, analysis and compilation of georeferenced images begins, which after a data purification and cleaning process is collected in a data bank. This database generates a time series of the NDVI where the state and health of the vegetation can be determined according to its index at a given time, see Fig.3.1.

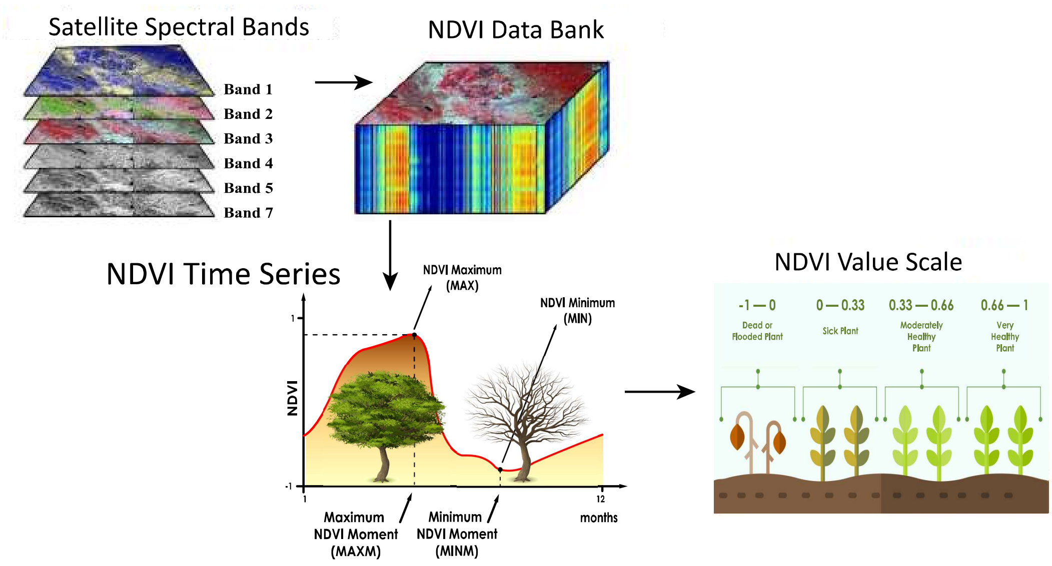

Figure 3.1: It is possible to observe the process of interpreting the various bands of a study area and then deposit all the information in a database which is used to generate, based on time, a time series of NDVI with its scale of vegetation values

Although the data obtained generally does not have a regular measurement frequency set, in addition to the fact that they tend to be noisy, they require optimal pre-processing where not too much information is lost or too much noise is retained, see Li, Shuang and Xu, Liang and Jing, Yinghong and Yin, Hang and Li, Xinghua and Guan, Xiaobin (2021). Therefore, we apply different smoothing methods, such as moving averages which can be seen in the figure 3.2 and exponential smoothing, see the Appendix. After smoothing out short-term fluctuations, long-term patterns were observed and factored into the modeling process. In addition, imputation methods were used for empty observations, generated by omissions in the Landsat platform, which consist of temporal statistical interpolation, see Figure 3.3. With the objective of obtaining time series with a monthly frequency of observations without the presence of omissions.

To finish this section, two monthly time series graphs will be presented, which are the NDVI, see in figure 3.4 and the NCA see in figure 3.5. These data were subjected to the statistical processes mentioned above to obtain a more accurate and continuous representation of trends over time. These data cover a period of twenty years, where uniform monthly data free of random fluctuations was collected for each year, which allows us to identify long-term patterns.

[1] https://earthengine.google.com/platform/

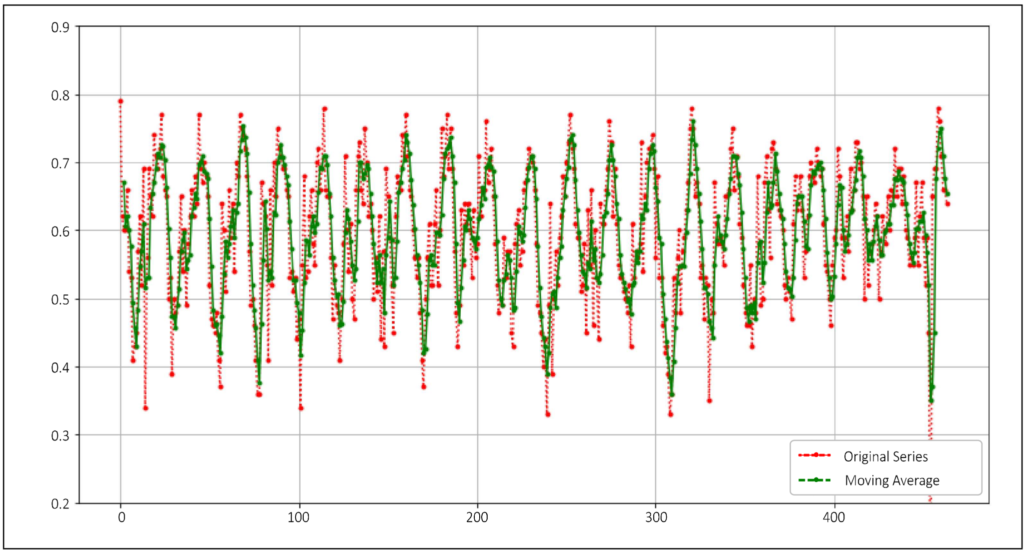

Figure 3.2: Data Moving average smoothing

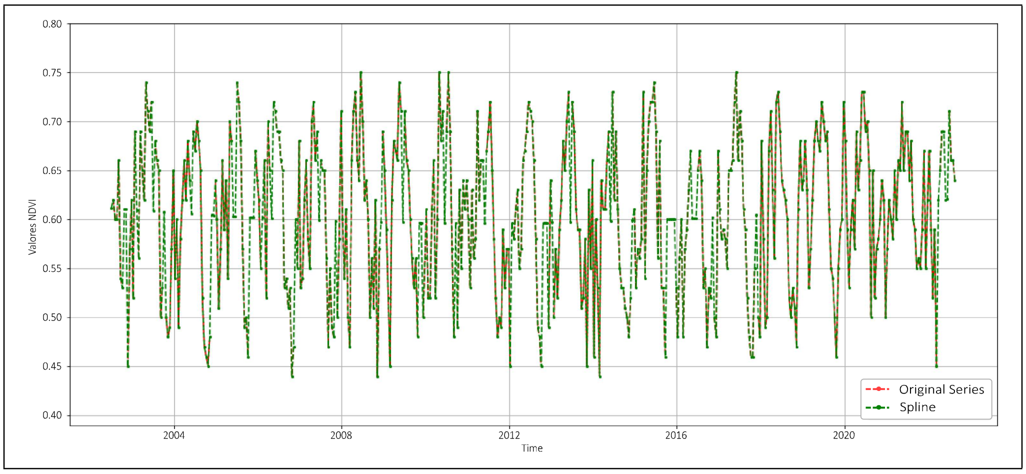

Figure 3.3: spline interpolation

3.2 Data Analysis

In this study, the normalized difference vegetation index (NDVI) is the dependent variable and the water level in the Xingu River (NCA) is the independent variable. Therefore, the functional relationship between these two variables can be established as follows

$$NDVI = \mathbf{f}(NCA), \tag{3.2}$$

where the objective is, based on historical observations of both series, to estimate the functional relationship $\mathbf{f}$ that links both quantities. Since both variables are time series, the causal relationship has to be of a dynamic nature, otherwise there is a risk of obtaining a spurious relationship, see Granger, Clive WJ and Newbold, Paul (1974). Therefore, the verification of the stationary nature of both series is mandatory prior to applying a time series model Ewing, Bradley T and Sari, Ramazan and Soytas, Ugur (2007). This paper, the stationarity condition of the variables was verified by using the augmented Dickey-Fuller tests Dickey, David A and Fuller, Wayne A (1979). As will be detailed later, Section 4.1, the NDVI variable is of a stationary nature, usually denoted by I(0), while the NCA variable is non-stationary of first order, usually denoted by I(1). This mixed nature of the integration order of the NDVI and NCA variables leads to selecting the ARDL model Pesaran, M Hashem and Shin, Yongcheol and Smith, Richard J (2001), which is determined as follows

$$y_t = c_0 + c_1 t + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=0}^{q} \beta_i' x_{t-i} + u_t, \tag{3.3}$$

where y_t are the dependent variables, in this case the NDVI series, x_t are the independent variables, in this case the NCA series, for which p and q are lags of each variable. Furthermore, the coefficients c_0 and c_1 are the intercept and the trend coefficient respectively, and in turn, u_t is a gaussian white noise error that models the random fluctuations. The ARDL model has advantages over the traditional cointegration test method, one of them is that you can mix variables of I(0) and I(1), which is the case with our data. Another advantage is that you can estimate the short-run and long-run relationship between variables in parallel using the ARDL limit test technique. For the purposes of this work, since our variable of interest is the NDVI and our exogenous variable is the NCA, the equation (3.3) takes the representation

$$NDVI_t = \alpha + \sum_{i=1}^{n_1} \gamma NDVI_{t-i} + \sum_{i=0}^{n_2} \beta_1 NCA_{t-i} + \varepsilon_t, \tag{3.4}$$

where n_1 and n_2 are optimally selected based on the data. The Bound Test is developed by converting the equation (3.3) and including the short-term and long-term dynamics. This approach guarantees us to carry out the F-test on the selected ARDL bound test equation with lag lengths. In this case, a maximum of 3 lags was selected at the variable level and then an optimal maximum lag length was selected based on the Akaike Information Criterion (AIC).

If The F-test confirms the presence of cointegration, we can meaningfully estimate the long-run equilibrium relationship between the variables. the order of integration of the variables I(0) and I(1) are assumed by the critical values of Dickey-Fuller test. If the F-statistic generated by the bound testing equation is above the upper bound, the null hypothesis of no cointegration is rejected; if it is below the lower limit, the test does not reject the null hypothesis, for a more detailed consultation see Pesaran, M Hashem and Shin, Yongcheol and Smith, Richard J (2001). After confirmation of the long-run relationship of the variables, we can obtain the short-run dynamics by converting the equation (3.4) in an error correction model (ECM) which is specified as

$$\Delta NDVI_t = \alpha + \sum_{i=1}^{n_1} \gamma \Delta NDVI_{t-i} + \sum_{i=0}^{n_2} \beta_1 \Delta NCA_{t-i} + \delta EC_{t-1} + \varepsilon_t, \tag{3.5}$$

where δ is the speed of adjustment, EC_{t-1} represents the unbalance or error correction term and Δ interprets the first difference the variables. The coefficient of EC_{t-1} shows the rate of readjustment to equilibrium in the long run when short-term disturbances produce disequilibrium. Causality is represented with a negative sign and significant in the error correction coefficient δ, see Shahbaz, Muhammad and Hye, Qazi Muhammad Adnan and Tiwari, Aviral Kumar and Leitão, Nuno Carlos (2013). It is important to diagnose the model specification, as well as the assumptions of independence, heteroscedasticity and normality of the errors. Therefore, serial independence the errors is verified using the Breush-Godfrey serial correlation Lagrange multiplier (LM) test, while the ARCH test will be used to verify the heteroscedasticity problem in the model. The CUSUM of squares Brown, Robert L and Durbin, James and Evans, James M (1975) and Pesaran e Pesaran (1999) they are used to detect errors of specification in the model of autoregressive structures. It is important to mention that these tests are also used to ensure the stability of certain model parameters.

As was previously mentioned, the ARDL model is not restricted by the integration order of the variables, this model reveals the existence of a linear relationship between non-stationary or mixed variables. However, this test has an important disadvantage since it does not reveal the direction of the relationship between variables. Therefore, to understand the direction between cause and effect of the variables, i.e., the relationship between the vegetation index and the water level of the Xingu River, the Bound Test ARDL must be applied for which one of the five cases must be selected in equation 3.6.

Once the case is selected, which for our study is type I, a candidate model is obtained, for which, if the calculated F statistic falls beyond the upper critical value, the null hypothesis of no cointegration is rejected. If the value is below the lowest value, the null hypothesis of no cointegration cannot be rejected. If the value falls within the band of critical values, the inference is not conclusive and other cointegration procedures such as Johansen, Soren and Juselius, Katarina and others (1990) must be verified.

$$
\begin{aligned}
&\textbf{CASE I} \\
&\qquad c_0 = c_1 = 0 \\
&\qquad ECT_t = y_{t-1} - (\sum_{j=1}^{k} \theta_j xj, t-1), \\
&\textbf{CASE II} \\
&\qquad c_0 = c_1 = 0 \\
&\qquad ECT_t = y_{t-1} - (\mu + \sum_{j=1}^{k} \theta_j xj, t-1), \\
&\textbf{CASE III} \\
&\qquad c_1 = 0 \\
&\qquad ECT_t = y_{t-1} - (\sum_{j=1}^{k} \theta_j xj, t-1), \\
&\textbf{CASE IV} \\
&\qquad c_1 = 0 \\
&\qquad ECT_t = y_{t-1} - (\Delta_{t-1} + \sum_{j=1}^{k} \theta_j xj, t-1), \\
&\textbf{CASE V} \\
&\qquad ECT_t = y_{t-1} - (\Delta_{t-1} + \sum_{j=1}^{k} \theta_j xj, t-1).
\end{aligned}
\tag{3.6}
$$

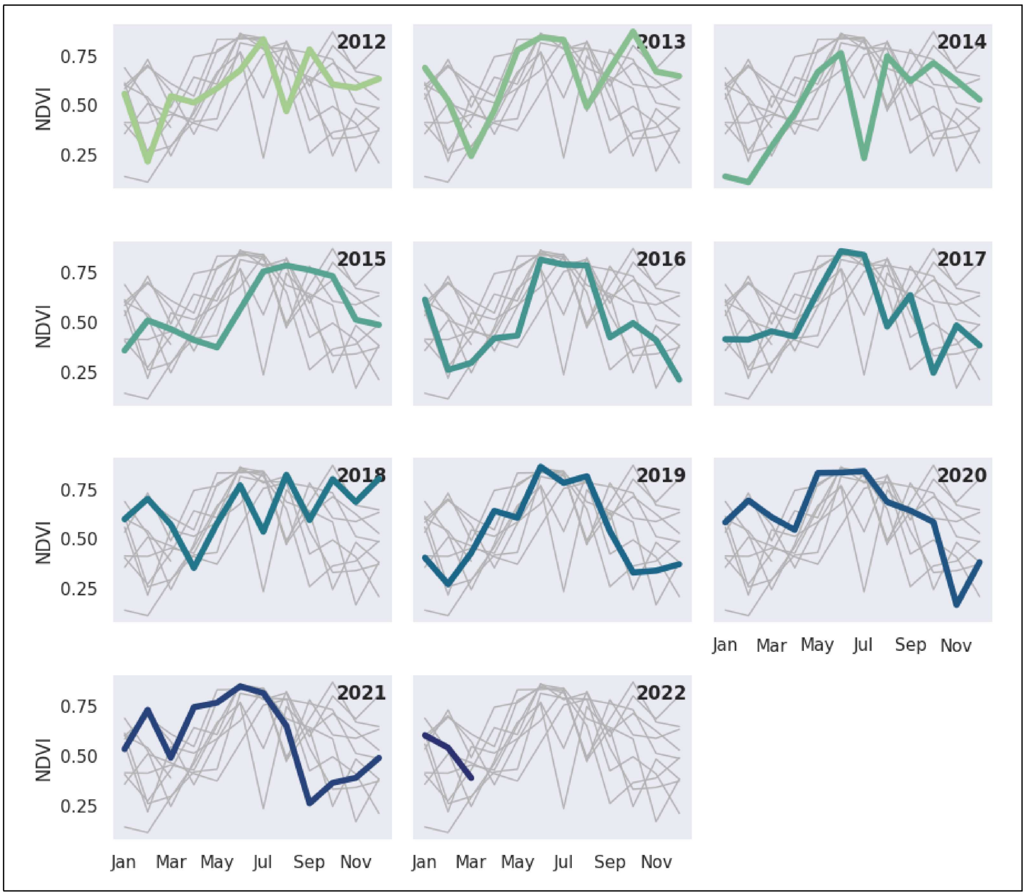

Figure 3.4: NDVI time series monthly for each year

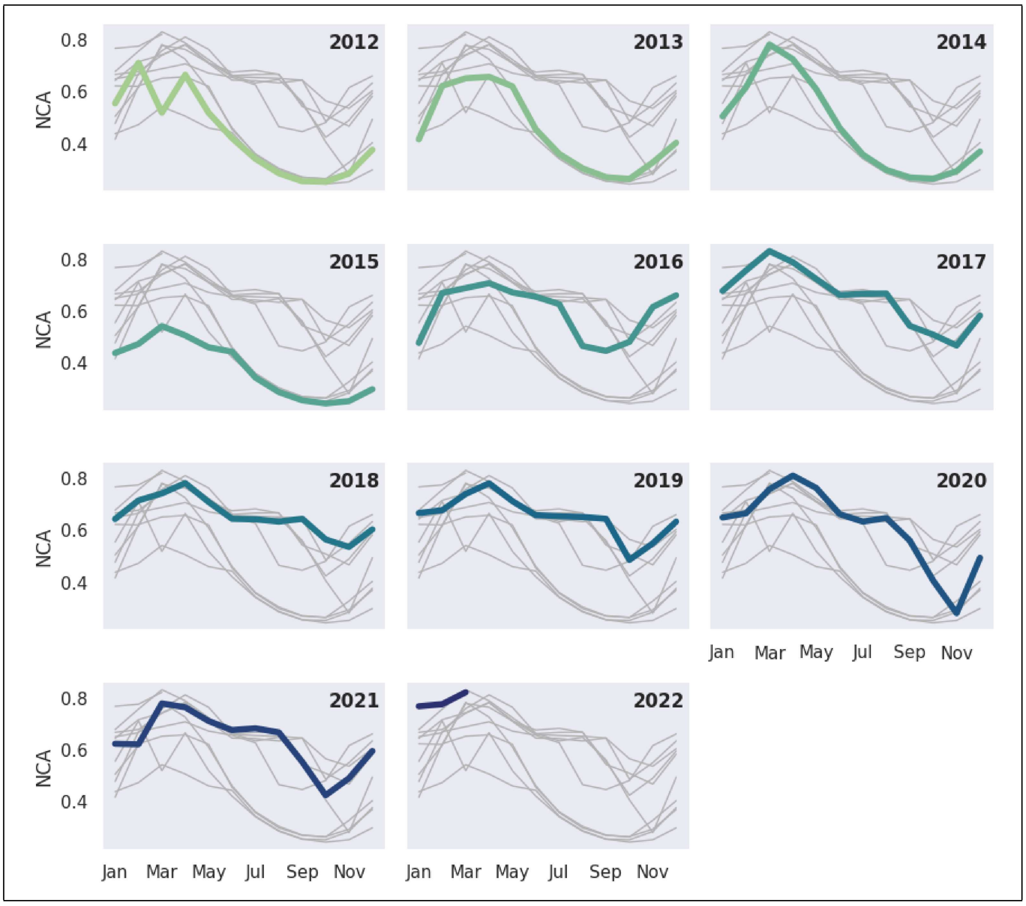

Figure 3.5: NCA time series monthly for each year

Chapter 4

Results and Discussion

4.1 Results

Once the data was processed, the time series analysis of the two variables was carried out, where the temporal evolution's in the NDVI and NCA measured monthly are observed. The figure 4.1 shows the graph of the two time series of both variables in the same time interval, with the aim of exploratory visualization of the dynamic relationship between them. In addition, a regime change can be visually observed in the dynamic relationships between both variables after the inauguration of the Belo Monte hydroelectric plant that began operating in 2016, which diverts a large part of the flow of the Xingu River.

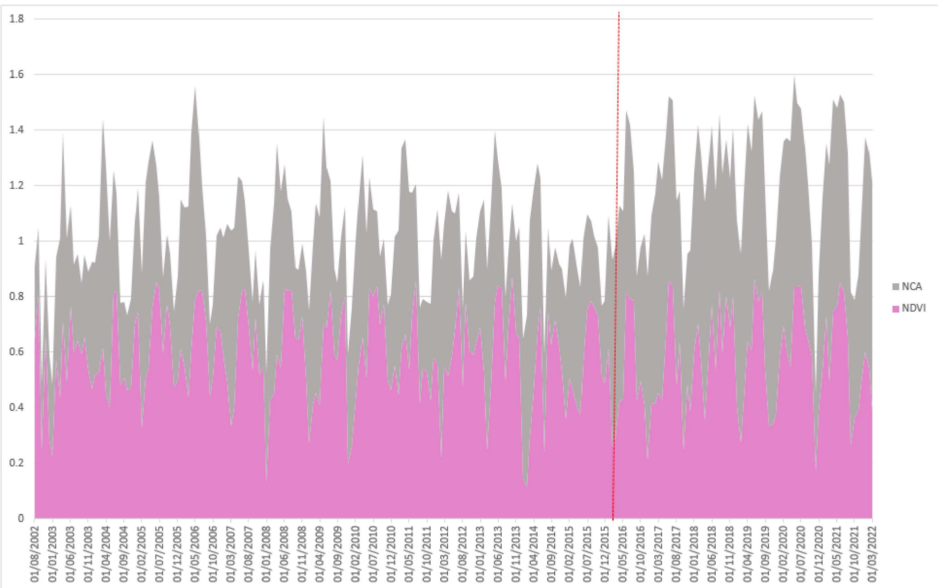

Figure 4.1: You can see the NDVI time series in pink and the NCA time series in lead during the period from 2002 to 2022. In addition, you can see its variation before and after the operation of the Belo Monte Plant, which is marked with a red line (April 20016)

The central interest of this work is to estimate the impact of the NCA in the Xingu river on the NDVI, this impact can be observed through satellite images as can be seen in figure 4.2, figure 4.3 and figure 4.4 where the NDVI images vary in the study area for the same hydrological moments over time. For this reason, our objective in particular is to obtain a dynamic relationship, both in the short and long term, between both magnitudes.

Figure 4.2: satellite image of the study area in 2014

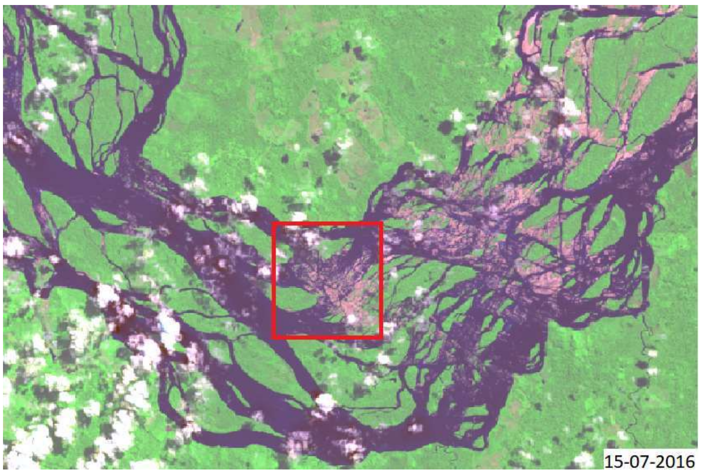

Figure 4.3: satellite image of the study area in 2016

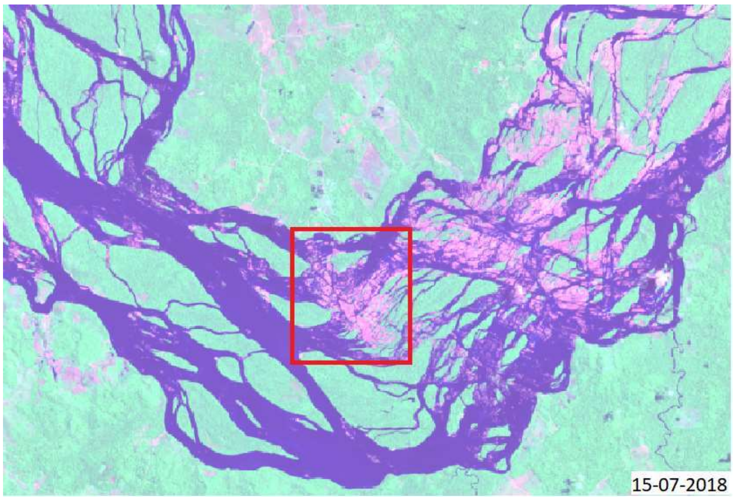

Figure 4.4: satellite image of the study area in 2018

For this, the descriptive statistics show that in Table 4.2, the p-value is less than 0.5, which leads to rejecting the hypothesis of non-stationarity for the Dickey-Fuller test (ADF) for the NDVI variables, which means say of order of integration is I(0), this is equivalent to saying that the NDVI variable is not a function of time and does not present a trend. While in table 4.1, which contains values of the stationarity test for the water level (NCA), it leads to not rejecting the hypothesis of non-stationarity, concluding that the order of integration of the NCA series is I(1), this means that the NCA systematically tends to increase or decrease over time. Therefore, the series are of order I(0) and I(1), which prevents us from working with conventional cointegration methodologies such as those described in Granger, Clive WJ (1981) and Johansen, Søren (1991). Therefore, a suitable methodology for our objectives is the ARDL-ECM model described in Pesaran e Pesaran (1999) and Pesaran, M Hashem and Shin, Yongcheol and Smith, Richard J (2001).

Once the Dickey Fuller test is finished, the ARDL model is shown, with which the lag number is determined, which was conveniently chosen to later formulate an error correction

Dickey-Fuller Test Results for Column: River Level	
statistical test	-1.371771
p-value	0.595732
No Lags Used	15.000000
Number of observations used	220.000000
Critical Value (1%)	-3.460428
Critical Value (5%)	-2.874769
Critical Value (10%)	-2.573821
Conclusions:	The null hypothesis cannot be rejected
	The data is not stationary

Table 4.1 -

Dickey-Fuller test results for spine: NDVI	
statistical test	-4.075307
p-value	0.001064
No Lags Used	11.000000
Number of observations used	229.000000
Critical Value (1%)	-3.459233
Critical Value (5%)	-2.874245
Critical Value (10%)	-2.573541
Conclusions:	Reject the null hypothesis
	The data is stationary

Table 4.2 -

model (ECM), which will be a particular type of model ARDL that removes spurious regressions. This can be seen in table 4.3 and table 4.4 where the dependent variable will be the NDVI, in addition it can be seen that the autoregressive order of the dependent variable is three and the distributed lag of the independent variable is two.

ECM Model Results

Dep. Variable:	NDVI	No. Observations:	236
Model:	ECM(3, 2)	Log Likelihood	117.081
Method:	Conditional MLE	S.D. of innovations	0.622
Date:	Mon, 27 Feb 2023	AIC	-218.161
Time:	20:59:37	BIC	-190.553

Table 4.3 -

	coef	std err	z	$P > \|z\|$	[0.025	0.975]
const	0.5201	0.072	7.204	0.000	0.378	0.662
NDVI L1	-0.9141	0.088	-10.341	0.000	-1.088	-0.740
NIVEL AGUA L1	0.0218	0.071	0.308	0.758	-0.118	0.162
D. NDVI L1	0.206	0.076	2.711	0.007	0.056	0.357
D. NDVI L2	0.1901	0.063	2.996	0.003	0.065	0.315
D. NIVEL AGUA L0	-0.3760	0.138	-2.731	0.007	-0.647	-0.105
D. NIVEL AGUA L1	-0.4332	0.138	3.134	0.002	0.161	0.706

Table 4.4 -

Once the appropriate delay structure for the ARDL model has been determined, it is important to ensure that the model errors are serially independent. Because the focus is on the long-run relationship, the ECM corrects for the imbalance between the variables, allowing the short-run relationship to be evaluated as well as the long-run cointegration relationship. After the correction, the cointegration vector is determined, which is presented in table 4.5 with the objective of obtaining the normalized estimates of the cointegration relationship in addition to the estimated values associated with the relationship between the NCA and the NDVI.

Cointegrating Vector

	coef	std err	t	$P > \|t\|$	[0.025	0.975]
const	-0.5689	0.041	-14.021	0.000	-0.649	-0.489
NDVI L1	1.0000	0	nan	nan	1.000	1.000
Nivel agua	-0.0239	0.078 -	0.305	0.760	-0.178	0.131

Table 4.5 -

After making sure that the model is dynamically stable, the Bound Test was carried out, with the objective of observing if there is evidence of a long-term relationship between variables. This test provides two sets of critical values p. If the test statistic is below the critical value for the lower bound, then there appears to be no relationship of levels regardless of order or integration on the variables. If it is above the upper bound, then again there appears to be a relationship of levels, regardless of the order of integration of the variables. For this test, case three was used, which is when the constant is included in the model but not in the test.

Bounds Test Result	
Stat:	29.20969
Upper P value:	3.54e-14
Lower P value:	1.17e-14
Null:	No Cointegration
Alternative:	Possible Cointegration

Table 4.6-

Percentile	Lower	Upper
90.0	2.156	3.168
95.0	2.684	3.800
99.0	3.877	5.177
99.9	5.542	7.045

Table 4.7-

The value of our F statistic in Table 4.6 indicates that it is 29.20 and we have $(k+1) = 2$ variables (NDVI and NCA) in our model. So when we go to the critical value limit test tables, we have $k = 1$. Since the value of our F statistic exceeds the upper limit at the 5% significance level seen in Table 4.7, we can conclude that there is evidence of a cointegration or also called a period of long-term relationship between the two time series. This means that they share a common long-term trend, although they may deviate from this trend in the short term.

This period of relationship could be due to the fact that the flooded vegetation depends on the river water for its survival, and therefore, changes in the river water level can have a long-term impact on the vegetation. Furthermore, it is important to note that although the ARDL model provides evidence of a long-term relationship, there may also be short term factors that affect the relationship between NDVI and NCA in the Xingu River. These short-term factors may include seasonal variations in river water level and vegetation, as well as extreme events such as droughts or floods.

Although analysis of variance (ANOVA) is a statistical technique that allows comparing the means of different groups or treatments. In this case, we used ANOVA to evaluate whether there is a significant difference in the normalized difference vegetation index (NDVI) before and after the diversion of water in the Xingu River, which was generated by the implementation of the Belo Monte dam. To do this, we separated into two groups, where before the implementation was determined from 08/31/2002 to 04/30/2016 and after the implementation was defined from 05/31/2016 to 08/31/2022. Obtaining graph 4.5 in response to said analysis and seeing the box plot that is obtained after this analysis.

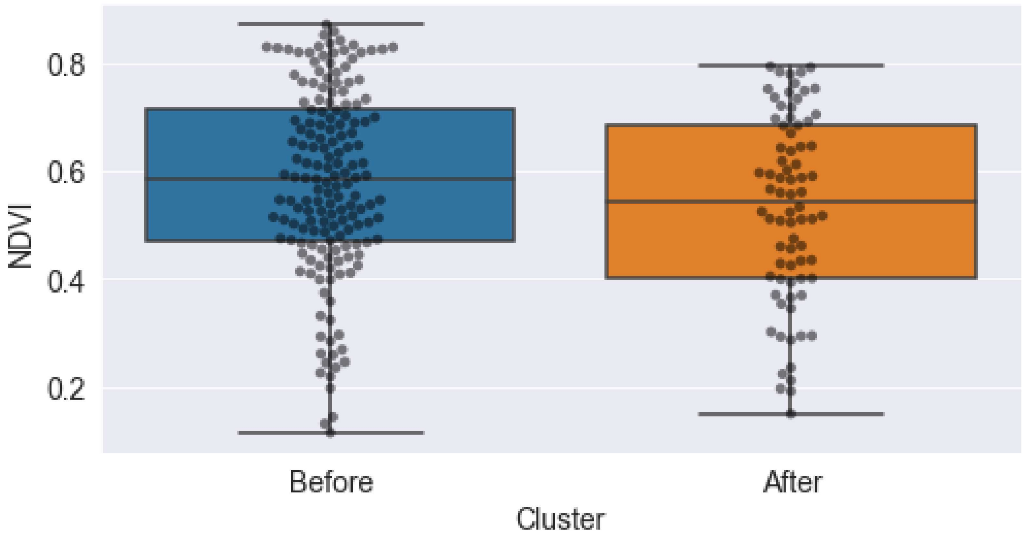

Figure 4.5: ANOVA analysis before and after the operation of the Belo Monte dam, Brazil

In addition to this, to obtain better answers from the ANOVA analysis, it was developed and separated into two more specific groups, with the aim of obtaining better answers about the study. Therefore, it was decided to generate this analysis for rainy seasons (February, March and April) and also for drought seasons (August, September and October). In this sense, you can see the response of this analysis in the rainy season in graph 4.6 and also in the dry season in graph 4.7.

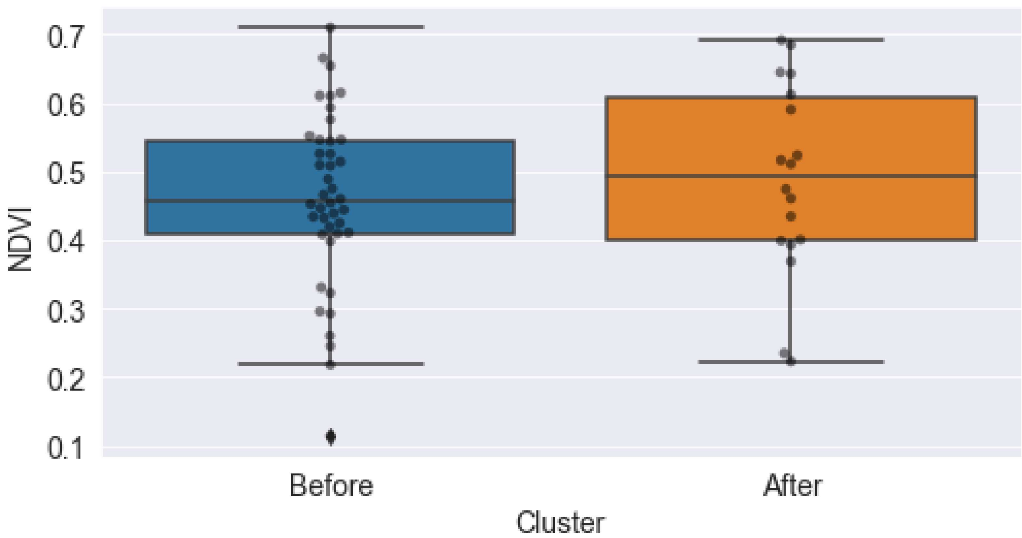

Figure 4.6: ANOVA analysis before and after the operation of the Belo Monte dam in the rainy season, Brazil

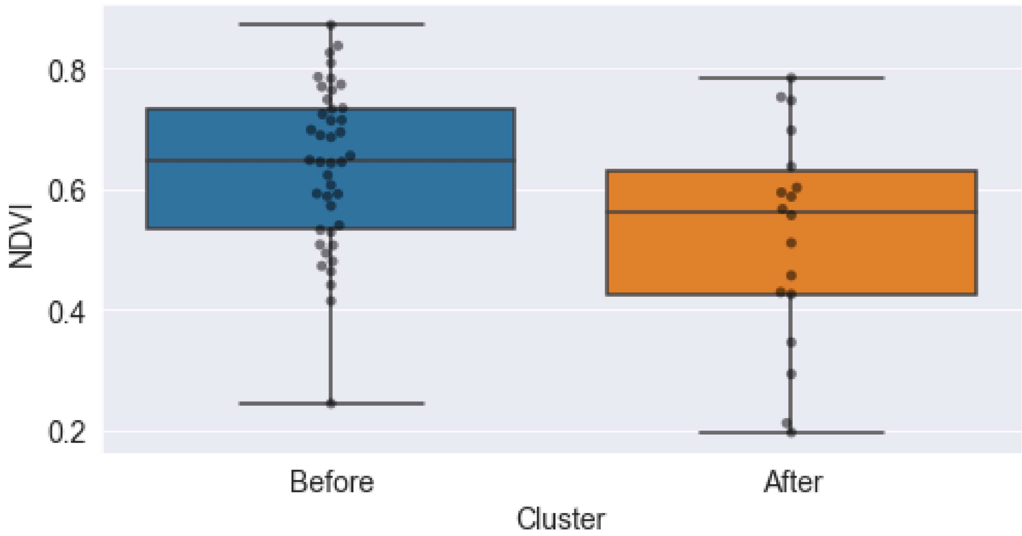

Figure 4.7: ANOVA analysis before and after the operation of the Belo Monte dam during the drought season, Brazil

It can be seen in figure 4.5, which studies the general variations before and after the operation of the Belo Monte dam, where the P value is 0.045715, which suggests that it is less than the significance level (0.05). Indicating that there is evidence to reject the null hypothesis, which suggests that there are significant differences between groups. In addition, the result of the ”F” statistic value is 4.034, which is high and means that the difference between the groups is large. Additionally, in Figure 4.6 that studies the changes before and after the implementation of the dam for the rainy season, it is observed that the value of P is 0.4, which is greater than 0.05. Which implies that there are no significant differences between groups, likewise, the “F” statistical value is 0.5, which is low and means that the difference between groups is minimal. Finally, in figure 4.7 where the changes in the implementation of the dam in times of drought are studied, it is observed that the value of P is 0.0076, which is less than 0.05, which implies that there are significant differences between groups. Furthermore, the “F” statistic value is 7.638, which is high and means that the difference between groups is large.

4.2 Discussion

The commissioning of the Belo Monte plant has generated various environmental changes that have implications for both biodiversity and society in the region. In terms of biodiversity, the construction of the plant has caused the flooding of large areas of tropical forest, which has resulted in the loss of natural habitats and the fragmentation of ecosystems. This can have negative consequences for the plant and animal species that depend on these habitats to survive. In addition, the construction of the plant has altered the natural patterns of water flow, which can affect water quality and the availability of resources for aquatic life, since according to Nagendra, Harini (2001), there is a relationship between the indices of various types of biodiversity and various NDVI values obtained from satellite images of different ecosystems. Likewise, according to Freitas, Simone R and Mello, Marcia CS and Cruz, Carla BM (2005) who evaluated the variations of the forest structures in the vegetation index by fragments of the Atlantic Forest, to determine the existence of a correlation between the NDVI and the structure of the humid forests. Furthermore, Chen, T and De Jeu, RAM and Liu, YY and Van der Werf, GR and Dolman, AJ (2014) who used soil moisture, vegetation productivity and CO_2 concentration to understand ecosystem dynamics and subsequently found a strong positive relationship between these variables with NDVI.

In terms of society, the construction of the plant has had a significant impact on the indigenous and local communities that depend on the Xingu River for their livelihood. Many people have been displaced from their land and homes, leading to social and economic tensions. In addition, the construction of the plant has led to changes in the traditional livelihoods of the communities, such as fishing and agriculture, which has affected the food security and livelihood of these communities, which may have lasting negative effects on their well-being and quality of life.

Therefore, in the future it is important to consider how these environmental changes will affect society in the region and implement mitigation and compensation measures to minimize negative impacts on biodiversity and local communities. In addition, sustainable practices and alternative energy should be promoted to reduce dependence on projects that generate major environmental changes.

Chapter 5

Conclusions

In conclusion, the use of the ARDL (AutoRegressive Distributed Lag) cointegration method to analyze and model NDVI and NCA time series has proven to be a valuable tool in satellite detection and understanding of changes in vegetation and water resources. Since, the advantage of this cointegration approach lies in its ability to capture long-term relationships and provide accurate and consistent estimates. This makes it possible to analyze the influence of multiple environmental factors, such as weather conditions and water availability, on the dynamics of the vegetation.

Likewise, satellite detection based on NDVI time series provides invaluable information in the management of natural resources, as well as in the evaluation of the impact of environmental changes on terrestrial ecosystems. This technique allows us to better understand fluctuations in vegetation cover over time and to identify patterns of ecosystem change and performance. Furthermore, the interaction between vegetation and water bodies is complex and depends on several factors. For this reason, the use of time series allows us to discern the direct and indirect effects of environmental variables and understand how they influence the dynamics of ecosystems.

Therefore, the use of the ARDL method in satellite detection through the analysis of time series of NDVI and NCA is important for the study of water resources and vegetation. Since this methodology gives us a deeper understanding of environmental changes and helps us make informed decisions in the management and conservation of our ecosystems. For this reason, in this study we have presented compelling evidence to affirm that the construction of the Bello Monte power plant, commissioned in April 2016, has had a negative impact on the ecosystem of the Volta Grande do Xingu, which is caused by the alteration of the natural hydrological dynamics, which, in turn, has induced changes in the flooded vegetation. These conclusions were inferred from the identification of a causal, structural, and negative short-term relationship between the NCA and the NDVI, series. This implies that this dynamic relationship between series can influence a chain of effects over time.

In conclusion, further study conducted using the ANOVA method has revealed significant changes in the Normalized Difference Vegetation Index (NDVI) in the fish feeding areas before and after the operation of the Belo Monte dam. These results indicate that there have been notable variations in the ecosystem of the Belo Monte reserve.

Additionally, additional ANOVA studies were conducted to determine if these changes had a differential impact during the wet and dry seasons. The results showed that the changes before and after the implementation of the dam significantly affected during times of drought. These findings underscore the importance of considering the ecological impacts of such infrastructure projects. It is crucial to continue monitoring and studying these effects to mitigate any negative impacts on the ecosystem and ensure the long-term sustainability of the Belo Monte reserve.

Bibliography

Acuña, Isaías Tobasura Una visón integral de la biodiversidad en colombia, Luna Azul, 1996, pp 2–de

Baker, Jessica CA and Garcia-Carreras, Luis and Gloor, Manuel and Marsham, John H and Buermann, Wolfgang and da Rocha, Humberto R and Nobre, Antonio D and de Araujo, Alessandro Carioca and Spracklen, Dominick V Evapotranspiration in the Amazon: spatial patterns, seasonality, and recent trends in observations, reanalysis, and climate models, Hydrology and Earth System Sciences, 2021, vol. 25, p. 2279

Brown, Robert L and Durbin, James and Evans, James M Techniques for testing the constancy of regression relationships over time, Journal of the Royal Statistical Society: Series B (Methodological), 1975, vol. 37, p. 149

Castro-Diaz, Laura and Lopez, Maria Claudia and Moran, Emilio Gender-differentiated impacts of the Belo Monte hydroelectric dam on downstream fishers in the Brazilian Amazon, Human Ecology, 2018, vol. 46, p. 411

Chen, T and De Jeu, RAM and Liu, YY and Van der Werf, GR and Dolman, AJ Using satellite based soil moisture to quantify the water driven variability in NDVI: A case study over mainland Australia, Remote Sensing of Environment, 2014, vol. 140, p. 330

Cunha, Denise de Andrade and Ferreira, Leandro Valle Impacts of the Belo Monte hydroelectric dam construction on pioneer vegetation formations along the Xingu River, Pará State, Brazil, Brazilian Journal of Botany, 2012, vol. 35, p. 159

Dickey, David A and Fuller, Wayne A Distribution of the estimators for autoregressive time series with a unit root, Journal of the American statistical association, 1979, vol. 74, p. 427

Ewing, Bradley T and Sari, Ramazan and Soytas, Ugur Disaggregate energy consumption and industrial output in the United States, Energy policy, 2007, vol. 35, p. 1274

Freitas, Simone R and Mello, Marcia CS and Cruz, Carla BM Relationships between forest structure and vegetation indices in Atlantic Rainforest, Forest ecology and management, 2005, vol. 218, p. 353

Granger, Clive WJ Some properties of time series data and their use in econometric model specification, Journal of econometrics, 1981, vol. 16, p. 121

Granger, Clive WJ and Newbold, Paul Spurious regressions in econometrics, Journal of econometrics, 1974, vol. 2, p. 111

Jiang W., Yuan L., Wang W., Cao R., Zhang Y., Shen W., Spatio-temporal analysis of vegetation variation in the Yellow River Basin, Ecological Indicators, 2015, vol. 51, p. 117

Jiang, Xiandie and Lu, Dengsheng and Moran, Emilio and Calvi, Miquéias Freitas and Dutra, Luciano Vieira and Li, Guiying Examining impacts of the Belo Monte hydroelectric dam construction on land-cover changes using multitemporal Landsat imagery, Applied Geography, 2018, vol. 97, p. 35

Johansen, Søren Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica: journal of the Econometric Society, 1991, pp 1551–1580

Johansen, Soren and Juselius, Katarina and others Maximum likelihood estimation and inference on cointegration–with applications to the demand for money, Oxford Bulletin of Economics and statistics, 1990, vol. 52, p. 169

Junk, Wolfgang J and Piedade, Maria Teresa Fernandez and Schöngart, Jochen and Cohn-Haft, Mario and Adeney, J Marion and Wittmann, Florian A classification of major naturally-occurring Amazonian lowland wetlands, Wetlands, 2011, vol. 31, p. 623

Li, Shuang and Xu, Liang and Jing, Yinghong and Yin, Hang and Li, Xinghua and Guan, Xiaobin High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques, International Journal of Applied Earth Observation and Geoinformation, 2021, vol. 105, p. 102640

Mayer, Adam and Lopez, Maria Claudia and Moran, Emilio F Uncompensated losses and damaged livelihoods: Restorative and distributional injustices in Brazilian hydropower, Energy Policy, 2022, vol. 167, p. 113048

Nagendra, Harini Using remote sensing to assess biodiversity, International journal of remote sensing, 2001, vol. 22, p. 2377

Pesaran B., Pesaran H., , 1999 Microfit 4.1 interactive econometric analysis

Pesaran, M Hashem and Shin, Yongcheol and Smith, Richard J Bounds testing approaches to the analysis of level relationships, Journal of applied econometrics, 2001, vol. 16, p. 289

Price, John C How unique are spectral signatures?, Remote sensing of environment, 1994, vol. 49, p. 181

Shahbaz, Muhammad and Hye, Qazi Muhammad Adnan and Tiwari, Aviral Kumar and Leitão, Nuno Carlos Economic growth, energy consumption, financial development, international trade and CO2 emissions in Indonesia, Renewable and Sustainable Energy Reviews, 2013, vol. 25, p. 109

Tucker, Compton J Red and photographic infrared linear combinations for monitoring vegetation, Remote sensing of Environment, 1979, vol. 8, p. 127

Xiao, Xiangming and Zhang, Qingyuan and Braswell, Bobby and Urbanski, Shawn and Boles, Stephen and Wofsy, Steven and Moore III, Berrien and Ojima, Dennis Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data, Remote sensing of environment, 2004a, vol. 91, p. 256

Xiao, Xiangming and Zhang, Qingyuan and Braswell, Bobby and Urbanski, Shawn and Boles, Stephen and Wofsy, Steven and Moore III, Berrien and Ojima, Dennis Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data, Remote sensing of environment, 2004b, vol. 91, p. 256

Chapter 6

Appendix

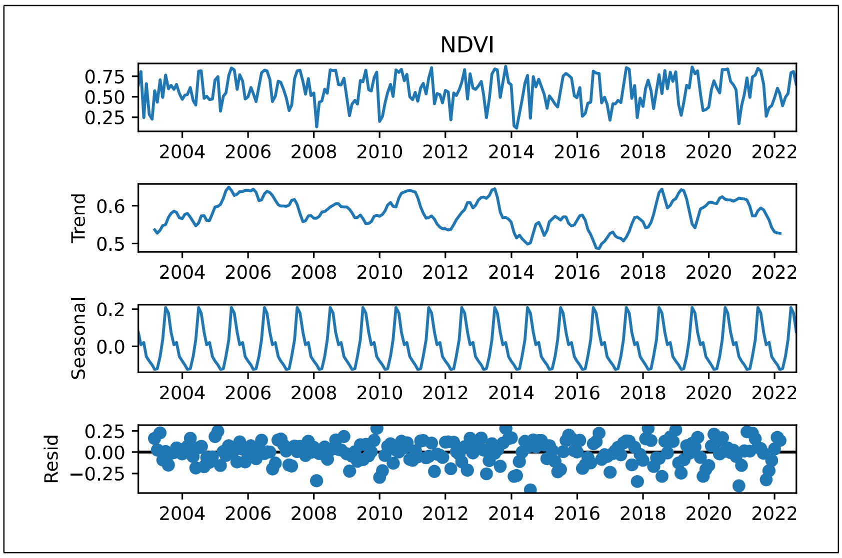

Figure 6.1: NDVI time series decomposition in the time period from 2002 to 2022

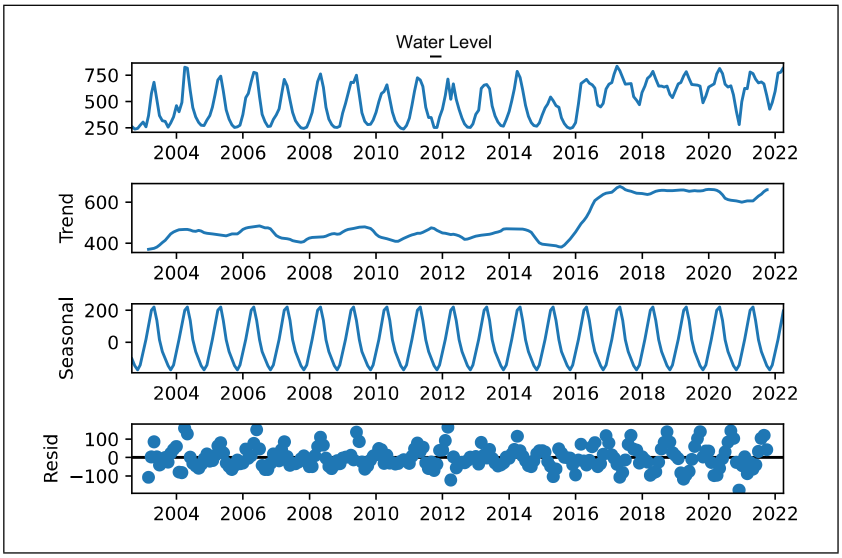

Figure 6.2: NCA time series decomposition in the time period from 2002 to 2022

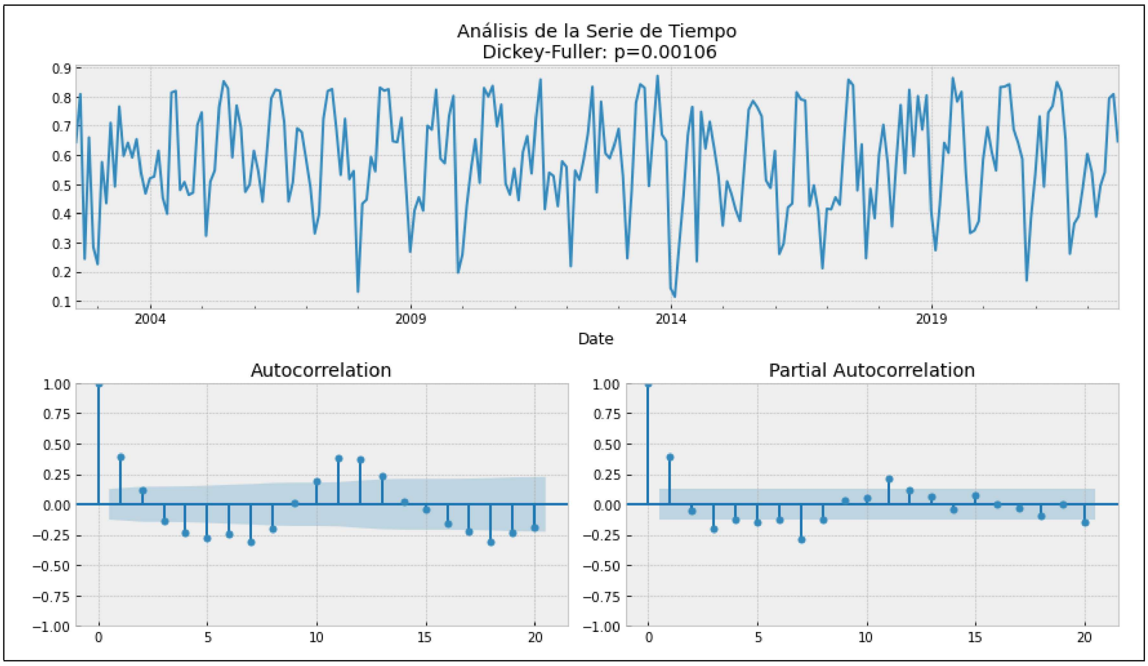

Figure 6.3: Autocorrelation and partial correlation of NDVI, it can help to check the stationarity of the variable

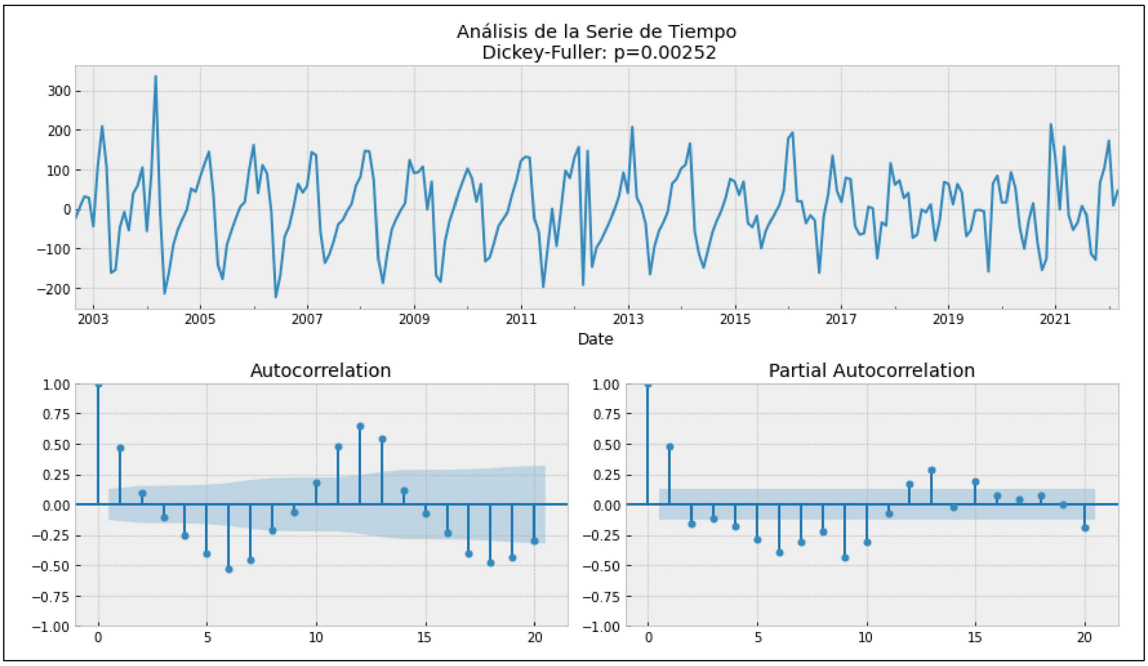

Figure 6.4: Autocorrelation and partial correlation of NCA, it can help to check the stationarity of the variable