UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA HIDRÁULICA E AMBIENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS HÍDRICOS

RENAN VIEIRA ROCHA

BAYESIAN NETWORKS AND NETWORK SCIENCE APPLIED TO WATER RESOURCES: STREAMFLOW ANALYSIS AND FORECAST INCORPORATING THE NON-STATIONARITY

FORTALEZA

2021

RENAN VIEIRA ROCHA

BAYESIAN NETWORKS AND NETWORK SCIENCE APPLIED TO WATER RESOURCES: STREAMFLOW ANALYSIS AND FORECAST INCORPORATING THE NON-STATIONARITY

Thesis defended at the Water Resources and Environmental Sanitation Post-Graduate Program of the Technology Center at the Federal University of Ceará, as partial requirement to obtain the doctor degree in Civil Engineering. Concentration Area: Water Resources.

Advisor: Prof. Dr. Francisco de Assis de Souza Filho.

FORTALEZA

2021

RENAN VIEIRA ROCHA

REDES BAYESIANAS E CIÊNCIA DE REDES APLICADA AOS RECURSOS HÍDRICOS: ANÁLISE E PREVISÃO DAS VAZÕES INCORPORANDO A NÃO-ESTACIONARIEDADE

Tese apresentada ao Programa de Pós-Graduação em Recursos Hídricos e Saneamento Ambiental da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Doutor em Engenharia Civil. Área de concentração: Recursos Hídricos.

Orientador: Prof. Dr. Francisco de Assis de Souza Filho.

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R576b Rocha, Renan Vieira.
Bayesian Networks and Network Science applied to Water Resources : Streamflow analysis and forecasting incorporating the non-stationarity / Renan Vieira Rocha. – 2021.
174 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Civil: Recursos Hídricos, Fortaleza, 2021.
Orientação: Prof. Dr. Francisco de Assis de Souza Filho.

1. Streamflow Forecasting. 2. Change point detection. 3. Dynamic Bayesian Networks. 4. Network Science. 5. Low Frequency. I. Título.

CDD 627

RENAN VIEIRA ROCHA

BAYESIAN NETWORKS AND NETWORK SCIENCE APPLIED TO WATER RESOURCES: STREAMFLOW ANALYSIS AND FORECAST INCORPORATING THE NON-STATIONARITY

Thesis defended at the Water Resources and Environmental Sanitation Post-Graduate Program of the Technology Center at the Federal University of Ceará, as partial requirement to obtain the doctor degree in Civil Engineering. Concentration Area: Water Resources.

Approved on: ___/___/_____.

EXAMINING COMMITTEE

_____

Prof. Dr. Francisco de Assis de Souza Filho (Advisor)
Universidade Federal do Ceará (UFC)

_____

Prof. Dr. Eduardo Sávio Passos Rodrigues Martins
Universidade Federal do Ceará (UFC)

_____

Prof. Dr. Samiria Maria Oliveira da Silva
Universidade Federal do Ceará (UFC)

_____

Prof. Dr. Carlos de Oliveira Galvão
Universidade Federal de Campina Grande (UFCG)

_____

Prof. Dr. Nilo de Oliveira Nascimento
Universidade Federal de Minas Gerais (UFMG)

AGRADECIMENTOS

A elaboração dessa tese é parte de um processo de orientação que se iniciou na graduação e perpassou o mestrado. Todos esses trabalhos marcam um aprendizado e a construção de um conhecimento que só foi possível devido a brilhante orientação do Prof. Francisco de Assis, da qual sou profundamente grato pela mentoria, paciência e confiança.

É fruto também de um aprendizado junto aos colegas do laboratório de Gerenciamento do Risco Climático (GRC) nos diversos projetos e trabalhos que realizamos juntos, além da participação nos congressos que tornaram esse tempo de pesquisa mais interessante. Deixo aqui o meu agradecimento aos colegas por esse período de convívio e amizade. Também agradeço aos professores do Departamento de Engenharia Hidráulica e Ambiental pelo conhecimento adquirido e os membros da banca pelas sugestões e avaliação desse trabalho.

Esse longo período de pesquisa e estudos não seriam possíveis sem o apoio integral da minha família, principalmente dos meus pais Claudenilson e Socorro e meu irmão Thales, que foram meu porto seguro durante esses tempos difíceis de pandemia (COVID-19). Agradeço também a meus amigos mais próximos que acompanharam essa trajetória completa, sem vocês esse percurso seria mais sofrido.

“A ciência nunca resolve um problema sem criar pelo menos outros dez”

George Bernard Shaw

RESUMO

A violação da premissa de estacionariedade em séries temporais de vazão impõe o desenvolvimento de metodologias para (1) identificar a existência de mudanças nas séries e sua localização, (2) incorporar esse aspecto na modelagem e previsão de vazões e (3) analisar a extensão do impacto da não-estacionariedade. Esses aspectos foram analisados utilizando como estudo de caso as vazões naturalizados do setor elétrico brasileiro. O problema de detecção de mudanças nas propriedades estatísticas das séries de vazão é uma lacuna da literatura atual com preocupações acerca da confiabilidade dos resultados das diferentes metodologias disponíveis. Três metodologias foram utilizadas para detectar mudanças na média, utilizando a convergência das metodologias para analisar a confiabilidade dos resultados. Essa abordagem demonstrou um grande potencial, visto que as diferentes metodologias apresentaram uma alta taxa de convergência para o ponto de mudança correto e uma taxa menor para os pontos incorretos. As mudanças detectadas coincidiram com mudanças de fase de oscilações de baixa frequência dos oceanos Atlântico e Pacífico, podendo ser associadas a seus impactos na Zona de Convergência do Atlântico Sul. Uma primeira tentativa de incorporar essa não-estacionariedade foi realizada utilizando Redes Bayesianas Gaussianas (GBN), incluindo nas redes variáveis discretas representando as diferentes fases de oscilações de baixa frequência, permitindo diferentes parâmetros de rede de acordo com as fases. O foco em Redes Bayesianas derivou de artigos recentes que apontaram as Redes Bayesianas como uma ferramenta promissora em estudos hidroclimáticos, simultaneamente fornecendo bons resultados para modelagem e possibilitando uma descoberta causal através da análise da estrutura da rede. Os resultados demonstraram um grande potencial da GBN para prever vazões com um a oito meses de antecedência. Os resultados também revelaram uma boa performance da previsão de vazão via Inferência Bayesiana (*Likelihood Weighting simulations*). O uso das fases resultou na melhoria da performance para algumas estações, porém, não resultou em uma melhora para as estações que apresentaram mudança nas séries, indicando modificações significativas entre as estruturas de rede de cada período homogêneo. Esse aspecto foi analisado obtendo uma estrutura de rede para cada período homogêneo via diferentes metodologias. Os resultados corroboraram a suposição inicial, indicando profundas diferenças entre as estruturas de rede de cada período homogêneo, com alterações nas relações entre as variáveis e nas suas funções de autocorrelação. Portanto, a utilização do mesmo conjunto de *parents* para a série completa pode não compreender a extensão das alterações observadas. Finalmente, o impacto da não-estacionariedade na relação entre as séries de vazão foi analisado, este aspecto é importante na geração de previsões espacialmente correlacionadas. Uma metodologia para obtenção de redes

complexas ponderadas entre as estações foi proposta, utilizando a teoria da ciência de redes para detectar e analisar mudanças nos resultados de agrupamento. Foram observadas mudanças nos agrupamentos ao longo do tempo demonstrando a necessidade de uma abordagem mais complexa para correlacionar corretamente as previsões. A utilização de uma matriz de correlação para cada fase homogênea pode ser uma solução viável visto que foram encontradas semelhanças entre as mudanças no valor médio e na relação entre as séries.

Palavras-chave: Previsão de Vazões. Detecção de ponto de mudança. Redes Gaussianas. Redes Bayesianas Dinâmicas. Redes Complexas. Ciência de Redes. Baixa frequência.

ABSTRACT

The violation of the stationarity assumption in streamflow timeseries requires the development of methodologies to (1) identify the existence of changes in the series and its location, (2) incorporate this aspect in the streamflow modelling and forecasting framework and (3) analyze the full extension regarding its impact. Naturalized streamflow of the Brazilian electricity sector was used as a case study to analyze these aspects. The problem of detecting changes in the statistical properties of streamflow series is currently an open question with concerns regarding the reliability of the results of the different methodologies available. Three methodologies were used to detect changes in the mean value and the change point reliability was assessed evaluating the convergence among them. This approach showed great potential since the different methodologies presented a high convergence rate for the correct change point and a lower convergence rate for the incorrect points. The changes detected coincided to phase shift of low frequency oscillations of the Atlantic and Pacific oceans and its impact in the South Atlantic Convergence Zone. A first attempt to incorporate this non-stationarity was made using Gaussian Bayesian Networks (GBN). Discrete variables representing the different phases of low frequency oscillations were included in the networks, allowing different network parameters according to the phases. The focus on Bayesian Networks relies in recent articles that indicated Bayesian Networks as a promising tool in hydroclimate studies, simultaneously providing good modelling results and allowing causal discovery through the analysis of the network structure. The results demonstrated a great potential of the GBN to forecast streamflow with lead times from one to eight months. The results also unveiled a good streamflow forecasting potential via Bayesian Inference based on Likelihood Weighting simulations. The use of the phases resulted in the performance improvement for some stations, however, it did not improve the results of the stations that presented changes in the timeseries, suggesting significant changes between the network structures of each homogeneous periods. Network structures were obtained through different methodologies for each homogeneous period to analyze this aspect. The results confirmed the initial hypothesis, showing significant differences between the network structures of each homogeneous periods, with alterations in the relationship between the variables and in its autocorrelation function. Therefore, the use of the same set of parents for the complete series may not comprise the extension of the changes observed. Finally, an analysis was made to evaluate the non-stationarity impact in the relationship between the streamflow series, this aspect is important in the generation of spatially coherent streamflow forecasts. A framework was proposed to obtain weighted complex networks between the stations, using network science theory to detect and analyze changes in

the clustering results. The results showed significant changes in the clustering results across time, demonstrating the necessity of a more complex approach to correct correlate the streamflow forecasts. The use of a correlation matrix for each homogeneous phase could be a viable solution since similarities were found between the changes in the mean value and in the relationship between the stations.

Keywords: Streamflow Forecasting. Change point detection. Gaussian Networks. Dynamic Bayesian Networks. Complex Networks. Network Science. Low Frequency.

LIST OF FIGURES

Figure 1 – Location of streamflow stations. Key stations are highlighted (triangles) and labelled. The stations color represents its region according to ONS standards.....23

Figure 2 – Penalty values of the synthetic series obtained from different penalty functions and CROPS method. Maximum, minimum and median values of each penalty function are shown at the bottom left, bottom right and top, respectively. CROPS maximum, minimum and median values were 6.53, 0.05 and 0.65. 31

Figure 3 – Change point results for the mean value (bold lines) of climate indices and of the streamflow series of key stations. 33

Figure 4 – Percentage difference between the complete series mean value and the latest change point mean value, along with the convergence among the three change-point methodologies. 35

Figure 5 – Comparison between mean value change point results of climate indices and annual streamflow. Numbers on the right are station numbers; 168, 6 and 266 refer to the key stations Sobradinho, Furnas and Itaipu, respectively.....37

Figure 6 – Wavelet coherence results between the key stations and climate indices, AMO (left) and PDO (right). The thick black contour line indicates the 5% significance level. The cone of influence is represented in white and indicates a border effect in the highlighted results. 38

Figure 7 – Geographical location of the higher frequency climate SST indices used.....49

Figure 8 – Results for the calibration and simulation periods for streamflow non-normalized (Original) and normalized through log 10 and SRI. Boxplot without outliers for the results of (a) monthly series (b) NRMSE for each month. (c) NASH spatial results for the original streamflow, dark red dots represent negative values of NASH.....59

Figure 9 – Spatial dispersion of the network typologies found clustered by similarity (left) and the network typologies (right). This result refers to the networks obtained using the original non-normalized streamflow and with lead time of one month. The arrow pointing in and out from the same node refers to an autoregressive component. The networks typologies refer to the monthly series, thus, an arc between the index and the station is produced if at least one month has the index as a parent. 61

Figure 10 – (a) Median of the good-of-fit results of the stations, for the calibration and simulation period, through direct use of the parameters of the Gaussian Bayesian Networks (GBN), obtained independently for each lead time, and through Bayesian inference via likelihood weighting simulations (LW), based in the networks obtained with lead time of one month. The results refer to the networks obtained for the original non-normalized streamflow and after the removal of the stations that performed poorly for the lead time of one month. (b) Spatial dispersion of the NASH results (LW) for the lead times of two and eight months. Stations shown in dark red presented negative values of NASH. (c) Spatial dispersion of the percentage difference in the R2 coefficient between the two methods (GBN and LW), calculated in relation to the GBN results. 63
Figure 11 – Results of the different networks through LW method and for the lead times from two to eight months. The PH suffix indicates the use of the AMO and PDO phases as discrete nodes, the final suffix corresponds to the networks obtained with the use of a normalization procedure for the streamflow series (Log10, SRI) or the direct use of the original streamflow (Original). (a) Presents the median of the good-of-fit results of the different models for the calibration and simulation period. (b) Presents the percentage difference of the R2 coefficient between the different networks and the LW-PH-Original networks for the simulation period. 66
Figure 12 – Streamflow sub-network for the second period of station 99 (left) and first period of station 281 (right), through HC and PCMCI, respectively. Q nodes stands for the streamflow series in the different time-steps. 86
Figure 13 – HC network structure comparison between the two periods of station 188. Black, red and blue-dotted arcs represent true positives, false positives and false negatives arcs, respectively. Q nodes stands for the streamflow series in the different time-steps. 88
Figure 14 – PCMCI network structure comparison between the two periods of station 188. Black, red and blue-dotted arcs represent true positives, false positives and false negatives arcs, respectively. Q nodes stands for the streamflow series in the different time-steps. 89
Figure 15 – Comparison metrics and its relative percentual of the total number of arcs……...90
Figure 16 – Comparison through NMI metric between the network obtained for the complete series and the networks obtained using sliding windows of 20 and 30 years (top).

Modularity result of the networks obtained using the sliding windows (bottom). 104
Figure 17 – Results of the NMI metric applied to intercompare the networks obtained using sliding windows. Continuous results (left) and clustered (right). The axis value corresponds to the initial year of the window. 105
Figure 18 – Histogram plot of the network's degree distribution. Solid lines represent the degree distribution of the correspondent random network (same number of nodes and links). 107
Figure 19 – Spatial dispersion of the communities (clusters) for different time windows: (a) 1934-1963, (b)1969-1998, (c) 1980-2009 and (d) 1931-2015. The top 10 stations with highest values of bridge closeness are highlighted. Stations that were classified as a standalone community are referred as isolated nodes. 108
Figure 20 – Sankey diagram of the community's differences across the selected time-windows. The colors follow the same standard of the spatial plot with the black rectangle standing for the stations that were classified as a standalone community. The grey lines represent the station's trajectory through time. 109

LIST OF TABLES

Table 1 – Detection and convergence rate of Pettitt test, Bai and Perron's algorithm and Pruned Exact Linear Time method for the synthetic series. 31

Table 2 – Statistical properties of each segment: mean (μ), standard deviation (σ) and coefficient of variation (CV)..... 33

Table 3 – NASH and R^2 results (via LW) for four key stations for the lead time of one to eight months. The results corresponds to the best obtained among the different network configurations tested in this work. 67

Table 4 – ARTIVA change points detected..... 83

Table 5 – Goodness-of-fit results for the different methodologies and periods. The period with the largest differences is highlighted in bold. 83

Table 6 – Parents and ancestors of the networks from different methodologies. 85

Table 7 – Different configurations of the sliding windows communities. 105

Table 8 – Clustering coefficient results for the different time-windows networks (C) and correspondent random networks ($Crandom$). Random networks generated with the same number of nodes and links. 106

LIST OF ABBREVIATIONS AND ACRONYMS

AIC	Akaike Information Criterion
AMO	Atlantic Multidecadal Oscillation
ANA	National Water Management Agency
AR	Autoregressive model
ARMA	Autoregressive Moving Average model
ARTIVA	Auto Regressive TIme VArying models
BIC	Bayesian Information Criterion
BP	Bai and Perron change point algorithm
CI	Conditional Independence
CROPS	Changepoints for a Range of PenaltieS
CSN	Complete Series Network
CV	Coefficient of Variation
DBN	Deep Belief Networks
EEMD	Ensemble Empirical Mode Decomposition
ENR	Elastic Net Regression
ENSO	El Niño Southern Oscillation
ERSST V4	Extended Reconstructed Sea Surface Temperature Version 4
FN	False Negatives arcs
FP	False Positives arcs
GBN	Gaussian Bayesian Network
GEV	Generalized Extreme Value
GOF	Goodness-of-Fit measures
Grad	Atlantic Gradient index
GUM	Gumbel
HC	Hill Climbing algorithm
HQ	Hannah-Quinn penalty function
IMF	Intrinsic Mode Function
ITCZ	Intertropical Convergence Zone
LASSO	Least Absolute Shrinkage and Selection Operator
LGS	Logistic
LNO	Lognormal

LW	Likelihood Weighting simulations
MBIC	Modified Bayes Criterion
MCI	Momentary Conditional Independence
MCMC	Monte-Carlo Markov Chain
MI	Mutual Information
NHDBN	Non-Homogeneous Dynamic Bayesian Networks
NMI	Normalized Mutual Information
NOAA	US National Oceanic and Atmospheric Administration
NRMSE %	Normalized Root Mean Squared Error
NSE	Nash-Sutcliffe Efficiency coefficient
ONS	Electric System National Operator
PAR	Periodic Autoregressive
PARX	Periodic Autoregressive with exogenous variables
PBIAS %	Percentual Bias
PDO	Pacific Decadal Oscillation
PE3	Pearson type III
PELT	Pruned Exact Linear Time
R2	Coefficient of determination
RF	Random Forest
RGELMAN	R statistic of Gelman and Rubin
SACZ	South Atlantic Convergence Zone
SRI	Standardized Runoff Index
SST	Sea Surface Temperature
SVR	Support Vector Regression
SW20	Sliding Window of 20 years
SW30	Sliding Window of 30 years
TNA	Tropical Northern Atlantic index
TP	Ture Positives arcs
TSA	Tropical Northern Atlantic index
VE	Volumetric Efficiency
VI	Variation of Information
WEB	Weibull
WTC	Wavelet Coherency Analysis

XBR	eXtreme Gradient Boosting
XTC	Cross Wavelet Analysis

SUMMARY

1 INTRODUCTION 14

2 OBJECTIVES 16

2.1 Main objective 16

2.2 Specific objectives 16

3 METHODOLOGICAL STRUCTURE AND ARTICLES INTERRELATION 17

4 MAPPING ABRUPT STREAMFLOW SHIFT IN AN ABRUPT CLIMATE SHIFT THROUGH MULTIPLE CHANGE POINT METHODOLOGIES: BRAZIL CASE STUDY 19

4.1 Abstract 19

4.2 Introduction 19

4.3 Case Study Background 21

4.4 Data and Methodology 22

4.4.1 Case Study Data and Teleconnections 22

4.4.2 Preliminary Detection Skill Assessment 25

4.4.3 Pruned Exact Linear Time and Penalty Selection 26

4.4.4 Bai and Perron's Algorithm 27

4.4.5 Pettitt test 28

4.4.6 Wavelet Transform and Wavelet Coherence Analysis 28

4.5 Results 29

4.5.1 Preliminary Detection Skill Assessment 29

4.5.2 Abrupt shift in climate indices 32

4.5.3 Abrupt shift in streamflow series 34

4.6 Discussion 39

4.6.1 Multi-change-point discovery performance 39

4.6.2 Case Study 40

4.7 Conclusions 42

5 STREAMFLOW FORECAST THROUGH DYNAMIC HYBRID GAUSSIAN NETWORKS INCORPORATING THE INFLUENCE OF LOW-FREQUENCY SST PHASES 44

5.1 Abstract 44

5.2 Introduction 44

5.3 Data and Methodology 47

5.3.1 Hydroclimatic data and teleconnections 48

5.3.2 Dynamic Hybrid Bayesian Networks 51

5.3.3 Bayesian Networks Structure Learning in face of periodicity 52

5.3.4 Analysis of the impact of normalization procedures 55

5.3.5 Forecast performance analysis 56

5.4 Results **58**
5.4.1 Forecast performance analysis for 1-month lead time ***58***
5.4.2 Network Structures Coherency analysis ***61***
5.4.3 Forecast performance analysis for longer lead times ***62***
5.5 Discussion **67**
5.6 Conclusions **70**
6 STREAMFLOW MODELLING THROUGH DYNAMIC NON-HOMEGENEOUS BAYESIAN NETWORKS **73**
6.1 Abstract **73**
6.2 Introduction **73**
6.3 Data and Methodology **75**
6.3.1 Study Area Background and exogeneous predictors ***75***
6.3.2 Data preprocessing ***76***
6.3.3 ARTIVA framework and change point detection ***77***
6.3.4 Network Structure learning through PCMCI and HC ***78***
6.3.5 Streamflow modelling through Gaussian Bayesian Networks ***81***
6.3.6 Network performance and structure changes analysis ***82***
6.4 Results **82**
6.4.1 ARTIVA change point detection ***82***
6.4.2 Streamflow modelling performance and predictors analysis ***83***
6.5 Discussion **90**
6.6 Conclusions **92**
7 STREAM GAUGE CLUSTERING AND ANALYSIS FOR NON-STATIONARY TIME SERIES THROUGH COMPLEX NETWORKS **94**
7.1 Abstract **94**
7.2 Introduction **94**
7.3 Data and Methodology **97**
7.3.1 Study Area Background and Data ***97***
7.3.2 Complex Network through Mutual Information and MCMC threshold selection ***98***
7.3.3 Community Detection for Weighted Networks ***99***
7.3.4 Community changes detection across time ***101***
7.3.5 Complex Network Analysis ***101***
7.4 Results **103**
7.4.1 Community changes detection across time ***103***
7.4.2 Complex Network Analysis and Spatial dispersion ***106***
7.5 Discussion **110**
7.6 Conclusions **112**
8 FINAL CONCLUSIONS AND REMARKS **114**

REFERENCES **120**
Appendix A – Supplemental files of chapter 5 ..... **135**
Appendix B – Supplemental files of chapter 6 ..... **143**
Appendix C – Supplemental files of chapter 7 ..... **154**
Appendix D – Further information of the streamflow stations and analysis of the streamflow elasticity to climate **156**
Appendix E – Synthesis of the key packages used **166**

1 INTRODUCTION

In a strong hydropower dependent country such as Brazil, information regarding forth coming streamflow events translates into better water management and a reduction of the operational costs, since it diminishes the necessity of complementary thermal power that has an inherent higher production cost. Also, the knowledge of the underlying natural periodic cycles that modulates the streamflow behavior reduces the uncertainties existent, allowing the development of more fitted policies that diminishes the opportunity cost generated from a conservative policy and the losses derived by an erroneous policy.

However, the development of streamflow forecasting models and the comprehension regarding the underlying causes of its strong fluctuations and changes across time is not a simple task. One reason is that the streamflow is known to respond to large-scale atmospheric teleconnections patterns of both low and high frequency (RAVINDRANATH et al., 2019). Another one is the presence of non-stationarity in hydrologic time series (MILLY et al., 2008), which requires methodologies to correct detect if and when a change of the statistical properties of the time series occurs, and to incorporate these changes in streamflow modelling.

The detection of changes in streamflow time series is still an open-question with recent articles comparing diverse methodologies to verify its reliability (RYBERG; HODGKINS; DUDLEY, 2019). Thus, requiring further studies to correct assess its existence, true location and underlying causes, also analyzing if the changes can be associated to large scale climate patterns that drives the streamflow behavior.

When dealing with time series that present changes across time, is important to assess how this particularity can be included in the development of streamflow forecasting models. Different methodologies were proposed in the literature to cope with this problem (DUTTA; MAITY, 2020; LUO et al., 2012), including methodologies developed to both model the time series and detect the occurrence and location of changes (LÈBRE et al., 2010).

In this doctoral thesis the focus was on modelling streamflow through Bayesian networks, including detected changes into the modelling framework. The choice of this methodology relies in recent articles that showed Bayesian Networks as a promising tool in hydroclimate studies, where two interesting aspects emerges for streamflow modelling: (1) the possibility to obtain good results for streamflow modelling, outperforming standard methodologies (DUTTA; MAITY, 2020), and (2) the possibility to be used to discover causality

between variables through the analysis of the graph structure obtained through structure learning algorithms (EBERT-UPHOFF; DENG, 2012, 2015; RUNGE et al., 2019a). Causal discovery aims to extract causality from the structure of the Bayesian network, enabling the analysis of the physical aspect of the model, which is important to verify assumptions about the influence of large-scale climate phenomena in streamflow and its long-term projections. Also, the use of Bayesian Network unveils different possibilities such as streamflow forecasting through Bayesian Inference and performing conditional queries to analyze the direct or indirect impact among the variables of the network.

Finally, it is also important to evaluate the extension of the consequences of the non-stationarity, regarding for example, if it generates changes in the relationship between different streamflow series. This latter information is important to the operation of an interconnected system of reservoirs such as the Brazilian National Interconnected System (SIN). The stationary hypothesis regarding the relationship of different reservoirs is a core aspect of methodologies developed to produce streamflow forecast and projections. Historically, the use of a single correlation matrix, where the residue is considered spatially correlated, is adopted as standard into statistical autoregressive models used for local streamflow forecast and projections (CEPEL, 2006).

The Network Science theory provides an interestingly solution to analyze the relationship among streamflow series. Network Science comprises studies of complex networks in its different types (such as small-world, scale-free and random networks), analyzing its underlying properties and allowing conclusions to be drawn from the network, regarding its stability, efficiency and robustness (AGARWAL et al., 2018; HALVERSON; FLEMING, 2015; SIVAKUMAR; WOLDEMESKEL, 2014). Another aspect is the detection and analysis of the communities existent in a complex network. The community denomination arises from Social Network Analysis and for hydroclimatic related studies a simply parallel can be made to clusters obtained from clustering techniques. Similarly to clusters, a community comprises nodes of the network that shares similarities, thus, allowing analysis regarding the relationship between the variables represented by the nodes that composes the network. Therefore, the study of a set of streamflow gauges as a network allows not only the use of a community detection algorithm as a clustering technique but also unveils different possibilities of analysis from the Network Science perspective and through network properties such as centrality, distance, and degree.

2 OBJECTIVES

2.1 Main objective

To develop a methodology to detect changes in streamflow time series and incorporate the non-stationarity existent using Bayesian networks in a streamflow modelling and forecasting framework, also developing a methodology to assess the non-stationarity impact in the relationship between streamflow time series across time using Network Science.

2.2 Specific objectives

1. Detect change points in streamflow time series through different methodologies and develop a framework to analyze the reliability of the results, also assessing the underlying probable causes behind the changes detected;
2. Analyze strategies to incorporate the existent non-stationarity in streamflow modelling and forecasting through Bayesian Networks;
3. Analyze the use of Bayesian networks to forecast streamflow with lead times of one to eight months and the impact of normalization procedures into the streamflow forecasting results;
4. Develop methodology to detect and analyze changes in the relationship between streamflow time series across time through Complex Networks;
5. Analyze the low-frequency impact in streamflow behavior and stream gauge clustering.

3 METHODOLOGICAL STRUCTURE AND ARTICLES INTERRELATION

To contemplate the particularities mentioned in the introduction and the proposed objectives, this doctoral thesis is organized into four articles and a final section with the summary of the key conclusions obtained throughout this work and remarks for future works. The first article contemplates the problem of detecting changes in streamflow series, evaluating the performance of different change point methodologies and the use of multiple methods simultaneously as a solution to assess the reliability of the results. Also, this article uses Wavelet Coherency Analysis to comprehend the possible underlying causes of the changes detected.

The second and third articles aimed to investigate frameworks to incorporate the non-stationarity in streamflow forecasting and modelling. The results from the first article indicated the phases of low-frequency sea surface temperature oscillations as the possible cause behind the changes detected. In this way, the second article explicitly incorporated these phases as discrete variables in the networks structure, investigating if this network configuration was able to account for the non-stationarity existent and increase the streamflow modelling and forecasting performance. Through this approach the entire period of analysis was represented by a single network structure. This article also analysis the forecasting performance for a lead time from one to eight months by two different approaches: (1) with the direct use of a Bayesian Network obtained for each lead time and (2) through Bayesian Inference from a base network obtained with a one-month lead time. Also, the impact of normalization procedures in the performance was assessed in this article.

The third article approaches the non-stationarity incorporation with the use of Non-Homogeneous Bayesian Networks, thus, obtaining one network structure for each homogeneous phase delimited by the change points found in the first article. This methodologic choice derives from the results of the second article, in which the use of the discrete variables did not result in an improvement for some of the stations with a change point, indicating possible significant changes across time in the network structures. This aspect was analyzed by the comparison of the network structures obtained. Also, this article revisits the change point detection problem with the use of a different methodology that can simultaneously discover the change points and the network structure.

The second and third article obtained the networks structures and parameters independently for each station, modelling the streamflow without accounting for spatial

correlation. The final article analyzes the non-stationarity impact in the relationship between the streamflow stations across time, thus, providing important information regarding the spatial correlation inclusion in face of the non-stationarity. To perform this task the article used complex networks and a community detection algorithm, relying in the mutual information metric combined with an automated threshold selection to construct weighted complex networks. Also, network science theory was used to analyze the typologies of the networks obtained.

In this context the articles comprise significant steps when dealing with non-stationary series: its detection, inclusion in the modelling framework and analysis of its derived impact in the generation of forecasts and projections with spatial coherence. Therefore, this methodological structure is not limited to streamflow related studies and can also be used in different fields of science that deal with non-stationary series. Furthermore, the framework proposed in the final article for the construction of the complex networks and community detection can be used as a data-driven clustering approach to any set of time series.

This thesis also presents a set of Appendix (A, B and C) that further illustrates the articles results and discussion. The Appendix D provides further context regarding the different streamflow characteristics across Brazil, presenting streamflow summary statistics and drainage basin area for each streamflow station used throughout this work. Also, a complementary analysis of the streamflow elasticity to precipitation and evapotranspiration is presented. This analysis expands the discussion regarding the limitations of the model used in the second article. Finally, the development of all articles was made with the use of codes written in the programming languages R and Python and the Appendix E synthesizes the key packages used.

4 MAPPING ABRUPT STREAMFLOW SHIFT IN AN ABRUPT CLIMATE SHIFT THROUGH MULTIPLE CHANGE POINT METHODOLOGIES: BRAZIL CASE STUDY[1]

4.1 Abstract

Three change point methodologies were used to detect changes the mean value of annual streamflow time series and analyze simultaneous changes in large scale global SST oscillations. To verify the relationship between the variables we used Wavelet Coherence Analysis. A preliminary detection skill test was performed using synthetic series and PELT presented the best results among the methods used (Pettitt test, Bai and Perron algorithm) when combined to a penalty selection through CROPS method. However, the use of classical penalty functions resulted in a poor performance of PELT. The three methods showed an extremely high convergence rate (> 90%) for the correct change points and a smaller rate for false positives (< 24%). Changes in the streamflow mean value coincided to phase shift of the low-frequency indices AMO and PDO, also corroborated by the wavelet results. Mostly of the changes can be associated to phase shift impacts in South Atlantic Convergence Zone.

4.2 Introduction

The variability of sea surface temperature (SST) at different oscillation scales, such as interannual, decadal and multi-decadal, is considered one of the drivers behind changes in hydrologic variables (ANDREOLI; KAYANO, 2007; KAYANO; CAPISTRANO, 2014; TANG et al., 2014; WANG et al., 2018). Understanding in detail the different periodic SST cycles and its effect on hydrologic variables can improve climate-based forecast models and water granting policies, and also provides better predictability to the water resources system management, enhancing its resilience.

---

[1] The Version of Record of this manuscript has been published and is available in the Hydrological Sciences Journal (Volume 65, 2020 – Issue 16, 18 Nov 2020)
https://www.tandfonline.com/doi/full/10.1080/02626667.2020.1843657

Approaches ranging from the use of artificial intelligence to periodic regression models incorporate SST indices in the streamflow forecasting (BLOCK et al., 2009; BRADLEY; HABIB; SCHWARTZ, 2015; FRITIER et al., 2012; LEHNER et al., 2017; LIMA; LALL, 2010b; SANKARASUBRAMANIAN et al., 2009; SOUZA FILHO; LALL, 2003; YANG et al., 2017). Selecting the correct indices in climate-based forecasting is based on the knowledge of either the physical or the statistical relationship of the phenomenon and the streamflow.

One classical approach is to base the selection in correlation coefficients and develop a single equation to represent the streamflow behavior (LIMA; LALL, 2010b). However, different phases of periodic cycles of the Atlantic and Pacific oceans can possibly modify the behavior of the streamflow and affect the series' statistical properties, increasing the complexity of the model required to correctly incorporate these changes.

The non-stationarity present in streamflow series can pose a challenge not just to the development of climate-based forecast models but, more importantly, to the water systems management. Periods with diverging mean values impose either changes in water granting policies, to adjust the values to the reality faced, or a high opportunity cost generated from a conservative policy that uses a fixed low streamflow value to base the granting policies. Also, different variance periods can either increase or decrease uncertainties, which in turn affect the risk associated with water allocation. Thus, the statistical properties mean, and variance are the two fundamental properties which define the possible water management policies.

Many recent articles uses change point analysis to detect changes in the streamflow series statistical properties through diverse methodologies (IVANCIC; SHAW, 2017; RYBERG; HODGKINS; DUDLEY, 2019; ZHU et al., 2019). To explore the relationship between the changes detected in streamflow series and climate indices, Tamaddun et al. (2019) used Cross Wavelet Analysis (XTC) and Wavelet Coherency Analysis (WTC), identifying that significant shifts occurred during the coupled phases of the climate signals. The use of XTC and WTC to investigate the influence of climate indices in hydroclimatic variables can be found in diverse other articles with solid results (ROCHA; SOUZA FILHO; SILVA, 2019; TAMADDUN et al., 2017; TAMADDUN; KALRA; AHMAD, 2017; TANG et al., 2014). However, we have not found the use of change point analysis to identify phases shifts of climate indices and matching change points in streamflow time series.

Despite the numerous change point methodologies found in the literature, the discovery if a change occurs in a time series and its true location is still an open question, since different methodologies imposes diverging results. This problem is amplified with the use of methods based either in likelihood cost-functions that are strongly dependent on the penalty selection (where a high value of penalty inhibits the change detection and low values imposes a great number of false positives) or on methods with the predefinition of the number of change points. These particularities result in a more subjective analysis. Therefore, more studies are needed to provide reliability to the change point results and to explore the usability of classical penalty functions or other predefinitions strategies, specifically for change point detection in hydroclimatic variables.

This paper detected and analyzed changes in the mean value of 88 Brazilian naturalized streamflow series and search for associations to changes detected in the mean values of climate SST indices. To accomplish this task a preliminary complementary study was performed to assess the detection skill and resulting convergence of three different change point methodologies with the use of synthetic series already used in a previous study (RYBERG; HODGKINS; DUDLEY, 2019). The methods used here were the Pettitt test (non-parametric), the Pruned Exact Linear Time (parametric version, likelihood cost-function based), and the dynamic algorithm proposed by Bai & Perron (2003) to estimate change points in time series regression models. WTC was also used to verify the associations between the changes observed and the SST indices.

4.3 Case Study Background

The streamflow series is of major importance to the Brazilian hydropower system, since it is the country's main power source, accounting for 66.6% of its total energy supply (636,4 TWh) (MME/EPE, 2019). Brazil's continental size also enables a spatial analysis that encompasses regions with diverse climate behavior where the atmosphere-ocean dynamic imposes different local impacts.

In recent years, a reduction in the inflow of hydroelectric power basins located in the country's Northeast region was reported by the Electric System National Operator (ONS). A technical report of Brazil’s National Water Management Agency (ANA) also identified a reduction in 80% of 125 rainfall monitoring stations (ANA, 2013), which can justify the

observed reduction in inflow to corresponding reduction of precipitation. However, a different behavior was reported for the South and Southeast regions, with a precipitation increase. Due to the magnitude of the changes reported, this study focused on evaluating large scale global phenomenon.

Castro, Souza Filho, & Silveira (2013) indicate a possible relationship between low-frequency SST oscillation and variation patterns observed in ONS streamflow time series. A possible association between the inflow changes mentioned and a low-frequency periodic phenomenon can suggest that this occurrence is not a punctual event due to abnormal years but instead a periodic behavior with a long-lasting effect, with significant impacts to energy production. Also, these characteristics would demand a water management capable of comprising to the dynamic climate risk associated.

Thus, AMO (Atlantic Multidecadal Oscillation) and PDO (Pacific Decadal Oscillation) indices were selected to represent the low-frequency SST oscillations, due to studies which impose a relationship between those indices and rainfall in Brazil and South America (ANDREOLI; KAYANO, 2007; KAYANO; CAPISTRANO, 2014; ROCHA; SOUZA FILHO; SILVA, 2019).The El Niño index was used due to its phenomenon impact on Brazilian rainfall was already extensively reported, and it is the main index used in Brazilian streamflow forecast models.

4.4 Data and Methodology

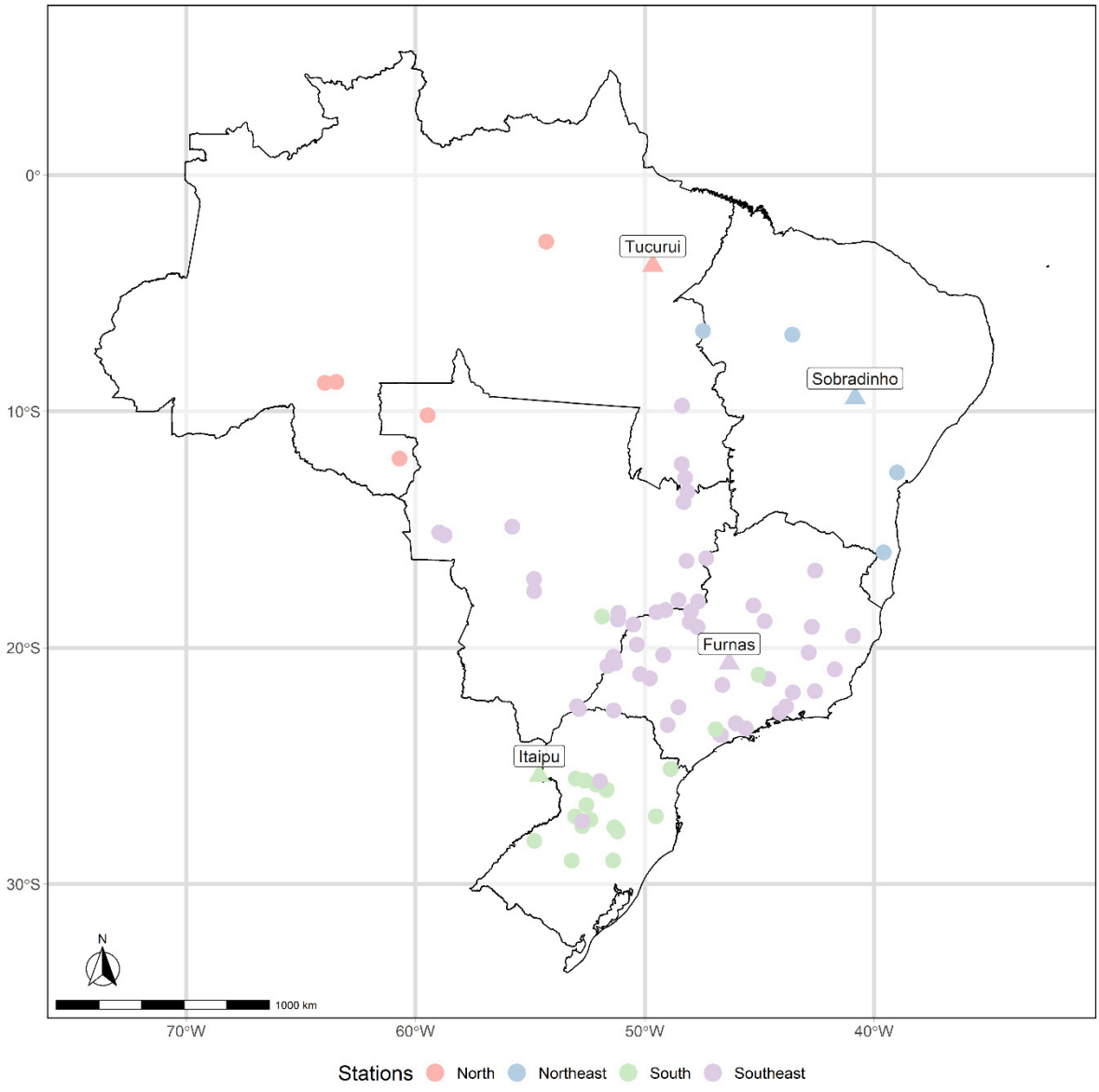
4.4.1 Case Study Data and Teleconnections

Monthly data of naturalized streamflow was used. We had access to data from 1931 to 2016, taken from 88 stations grouped in 4 regions distributed throughout Brazil, following the ONS standard (Figure 1)[2]. These regions diverge of the geographical division due to the operational interrelation between hydropower dams. The naturalized flows are acquired after a process of consolidation and consistency of the measurements from streamflow gauges and the account of anthropogenic derived factors: reservoir operations upstream of the streamflow,

[2] Appendix A provides complete information of the stations used (Table A1). Appendix D presents further information.

evaporation of the reservoir, water withdraw along the basins and river flow alterations due to infrastructures built along a natural course and operational pumping. Details of these processes can be found in ONS (2017). Annual streamflow series were obtained by averaging monthly values.

Figure 1 – Location of streamflow stations. Key stations are highlighted (triangles) and labelled. The stations color represents its region according to ONS standards.



Source: Prepared by the author

Four key stations highlighted in Figure 1, were selected for a more extensive analysis due to their relevance: Furnas, Sobradinho, Tucuruí and Itaipu. Itaipu is the world's

second largest hydropower dam in electricity production with an installed capacity of 14,000 MW accounting for 15% of the energy used in Brazil and 90% of the energy used in Paraguay. Tucuruí is considered the largest Brazilian exclusive hydropower plant with a capacity of 8,370 MW being the main source of electricity for the North subsystem. Furnas is one of the most important dams for the southeast region due to its location close to three large capitals with intensive energy demand (Rio de Janeiro, São Paulo, and Belo Horizonte) and the high potential for energy generation (1,216 MW). Sobradinho is a dam with one of the world largest flooded area having an installed capacity of 1,050.3 MW also having a local regularization purpose for the São Francisco River.

The indices AMO, PDO, and Niño 3.4 were acquired from the US National Oceanic and Atmospheric Administration (NOAA) database through the Working Group on Surface Pressure website (https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/, data retrieved in July of 2018) (ENFIELD; MESTAS-NUÑEZ; TRIMBLE, 2001; MANTUA et al., 1997; RAYNER et al., 2003).

El Niño Southern Oscillation (ENSO) is considered the most relevant oceanic-atmospheric mode to interannual climate variability on a global scale (KAYANO et al., 2016). El Niño's influence in South America rainfall is the result of alterations in the Walker and Hadley circulation cells. The El Niño 3.4 index is calculated by averaging SST values over the region delimited by 5N-5S, 170W-120W, considered a key region to analyze the ENSO related coupled ocean-interactions. This index is provided with or without removing the mean value from the 1981-2000 period; in this work, the anomaly version was used.

PDO is the principal component of the SST variability of the Pacific Ocean. The homonymous index measures the anomaly of North Pacific SST intensity by the number of standard deviations from the historical mean values (MANTUA et al., 1997). PDO variability is appointed as behaving symmetrically with interdecadal climatic fluctuations in North and South hemispheres, having a particular influence in both South and North America (CASTRO; FILHO; SILVEIRA, 2013). According to Andreoli & Kayano (2007), PDO modulates ENSO influence in South American rainfall, with ENSO showing stronger (or weaker) teleconnection with the rainfall according to phase convergence (or divergence) between the two oscillation modes. Castro, Souza Filho & Silveira (2013) also verified a positive correlation between the index and step changes observed in a series of maximum daily streamflow. Therefore, PDO influence in streamflow series is expected in the region.

AMO is considered the primary low-frequency variability mode of the Atlantic Ocean. The index is based on mean SST anomaly of the North Atlantic (0°), calculated by removing climate change derived interference through series detrending (ENFIELD; MESTAS-NUÑEZ; TRIMBLE, 2001). The index is available in two forms with or without the use of a ten years moving average smoothing. In this work, we used the index without smoothing. According to Kayano & Capistrano (2014), ENSO extremes are related to AMO, with stronger El Niño (La Niña) events in the cold (warm) phase of AMO. Thus, ENSO related precipitation anomalies are amplified with phase divergence between the two phenomena (PDO and AMO).

4.4.2 Preliminary Detection Skill Assessment

Ryberg, Hodgkins, & Dudlley (2019) generated sythetic series with random change points to assess the detection skill of eight detection methods. The series were based in peak-streamflow series of six U.S. streamgages and consists of 600 series with one change point and 600 series with two change points that can be reproduced using R code supplemental file of the article.

In this same work Ryberg, Hodgkins, & Dudlley (2019) found that for the mean value detection the parametric methods used did not perform well and other non-parametric methods were discarded due to the unacceptable number of false positives, concluding that the Pettitt's test is the most appropriate method for change-point analysis of peak streamflows.

However, we suspected that the use of the parametric method Pruned Exact Linear Time (PELT) (KILLICK; FEARNHEAD; ECKLEY, 2012) combined with the classical penalty function Modified Bayes Criterion (MBIC) could produce exceedingly high penalty values, inhinbiting the detection of the change points.

Thus, we decided to revisit and extend the results obtained by Ryberg, Hodgkins, & Dudlley (2019) for the parametric PELT methodology using a penalty value selected after an extensive penalty analysis. Also, we explored the convergence of results with the Pettitt's test, since it was considered the most appropriate method and was used in other recent articles for change point detection in streamflow time series (TAMADDUN; KALRA; AHMAD, 2019, 2016).

Since the Pettitt's test is a single change point detection method we also included in the analysis the change point algorithm detection proposed by Bai and Perron (2003), henceforth BP, which allows for multiple points discovery and was not used in the work of Ryberg, Hodgkins, & Dudlley (2019).

The convergence rate of the three methods was assessed in order to exclude eventual false positives. Following the previous work it was considered a correct detection of the change point if it is located in a distance of 5 years to the correct year and a base 10 logarithm was used when required to approximate normality (PELT and BP).

4.4.3 Pruned Exact Linear Time and Penalty Selection

The PELT change point detection methodology is an approach based in the asymptotic distribution of the likelihood ration test to detect change points in the mean value of a sequence of normally distributed observations and derives from Hinkley's (1970) original paper. More recently, this methodology was extended to diverse distributions (i.e. gamma, exponential, binomial) and also for the detection of variance changes (KILLICK; ECKLEY, 2013). The likelihood ratio test compares the fit quality of two models, detecting a change point when the null hypothesis of no change is rejected for the desired level of significance. Upon detecting a shift, the series is divided into two segments.

This change point methodology resumes in the exact minimization of an cost-based equation (KILLICK; ECKLEY, 2013):

$$\sum_{i=1}^{m+1}\left[C\left(y_{(\tau_{i-1}):\tau_i}\right)\right] + \beta f(m) \qquad \text{(Eq. 1)}$$

where the first term is the sum of each segment likelihood cost, the second term is the penalty factor and *m* is the number of change points.

The cost function depends on the assumptions imposed on the statistical distribution of the observations and also on the statistical property from which changes are desired to be detected, given that each segment's cost is based on its likelihood to the statistical property.

The addition of a change point usually imposes a reduction of equation (Eq. D1), resulting in a natural tendency of overfitting, thus, demanding the penalty factor. The literature

provides various functions to select the appropriate penalty value, e. g: Akaike information criterion (AIC) and Bayesian information criterion (BIC).

A more recently proposed approach is the use of a range of penalty values (Changepoints for a Range of PenaltieS - CROPS), allowing the analysis of the resulting segmentation from different penalty values (HAYNES; ECKLEY; FEARNHEAD, 2017). The CROPS algorithm finds the penalty thresholds that produce different numbers of change points from a given penalty interval. The penalty value can then be chosen by analyzing graphically its value versus the number of change points, searching an optimal or parsimonious value, e.g. a point where small variations of penalty do not generate a significant increase in the number of change points.

For the preliminary test the PELT method was used with all classical penalty functions provided in the R package *changepoint* (KILLICK; ECKLEY, 2013): BIC, MBIC, AIC and Hannah-Quinn (HQ).

The penalty values were then compared to the obtained through CROPS in which it was selected the value corresponding of the known change points or the closest higher number. The penalty behavior was then analyzed numerically to check if the value used could be selected straight forward by a graphical analysis. Finally, the results obtained with the "optimal" penalty value was compared to the other methodologies.

This framework provides the best detection skill of the combined PELT and CROPS method, since it removes the uncertainty regarding the number of the change points existent. Also, the previous knowledge of the correct number of change points would impose a probable strong bias to the penalty selection through a graph analysis.

For the present case study, CROPS was used to select the adequate penalty value where standard penalty functions did not provide satisfactory results.

4.4.4 Bai and Perron's Algorithm

The BP algorithm (also referred as Segment Neighborhood) selects the change points by identifying the global minimizers of the sum of squared residuals, using dynamic programming to efficient identify the optimal partitions with varying number of segments. The

algorithm provides discovery options for the minimum segment length and the maximum number of change points (BAI; PERRON, 2003; ERDMAN; EMERSON, 2007).

Through the maximum number approach, each number of change points is associated with a residual squared sum and can be combined with an information criteria to select the number of change points in the time series from either a graphical analysis or by choosing the number which minimizes the information criteria.

The BP maximum detection performance was also assessed by imposing the correct number of change points. We also evaluated the correspondence between the number of change points and the minimum value of BIC. This algorithm was used through the R package *strucchange* (ZEILEIS et al., 2002, 2003).

4.4.5 Pettitt test

The Pettitt test (1979) is a classical non-parametric single change point test for the median value, based on the Mann-Whitney U test. The change point location is defined by the maximum value of its test statistic and its corresponding p-value is derived by a two-sided test. The test was performed through the R package *trend*.

4.4.6 Wavelet Transform and Wavelet Coherence Analysis

The Wavelet transform is used to decompose a time series into a set of high and low frequency functions, enabling the analysis of the multi-frequency patterns that composes the original time series. The set of functions is obtained by the dilation and translation of a mother wavelet, in the scale and time domain, respectively (SIVAKUMAR, 2017). For geophysical variables Torrence and Compo (1998) suggests the use of Morlet Wavelet, also used in this study.

Wavelet transform of two time series can be analyzed through Cross Wavelet Analysis (XTC), to observe the common shared power in the time-frequency space, and Wavelet Coherence Analysis (WTC), to identify the relationship of spectrum bands along the time domain. XTC and WTC has been used successfully to assess the relationship between

climate indices and hydroclimatic variables (ROCHA; SOUZA FILHO; SILVA, 2019; TAMADDUN; KALRA; AHMAD, 2017, 2019; TANG et al., 2014). In this work, we focused only on the results of WTC.

WTC results refers to the wavelet squared coherence and can be analyzed similarly to classical correlation. The synchronization of the wavelet phases can be assessed to analyze the incurrence of a lagged relationship between the spectrum bands of the variables. The phase difference of individual phases can be converted in an angle in the interval of $[-\pi, \pi]$ and represented by arrows in the time-period domain plot. An absolute angle value larger than $\pi/2$ indicates that the two series move out-phase whereas the opposite indicates an in-phase relationship. The sign of the phase difference (arrows pointing up or down) indicates which series shows the leading pattern. In this work WTC was performed through the R package *WaveletComp*, more information about the methodology can be found in the work of Torrence & Webster (1999).

4.5 Results

4.5.1 Preliminary Detection Skill Assessment

The penalty values for PELT from each penalty function and the upper threshold value for CROPS, which is the highest value required to obtain the number of change points closest to the correct one, is shown in Figure 2 for the synthetic series. It is noticeable that CROPS obtained smaller values compared to the ones calculated by the penalty functions, where in just a few series the values ranged between AIC and HQ penalty functions and in three series it was close to BIC minimum value. The use of HQ and AIC resulted in the detection of change points just in a few of the time series and the use of MBIC and BIC penalty functions did not detect any change points. MBIC resulted in the highest values, its minimum value (11.74) is almost 40% higher than the maximum obtained through CROPS (8.42), confirming the hypothesis that the use of this penalty value inhibited the change point discovery in these series.

An “instantaneous” overfit pattern was noticed for the cases when the first found penalty threshold was remarkably close to zero, such as the minimum value obtained by CROPS (0.05), imposing the discovery of several change points.

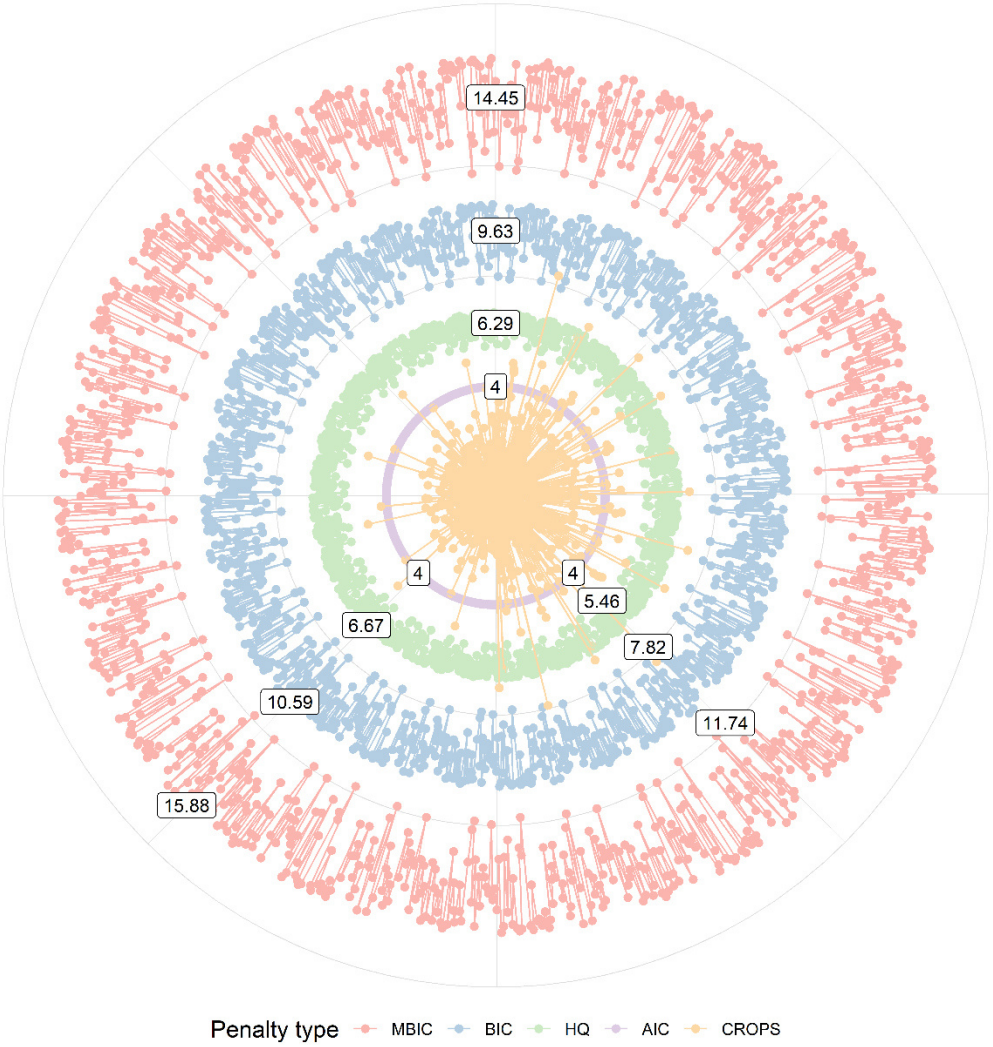
The analysis of the penalty values found by CROPS for each synthetic series implies that the correct penalty could be selected through a graphical analysis, especially when the PELT algorithm successfully found the change points. Most single and double change point series required a reduction ranging from 60% to 98% into the upper threshold value for the PELT algorithm to discover an extra change point. Further reductions in this decreased penalty, ranging from 20 percent to numbers close to 0, imposes further detections, of up to five change points. This behavior could be easily observed in a graphical analysis where the optimal or parsimonious value could be selected by avoiding an overfit point where small reductions in the penalty would lead to the discovery of extra change points.

However, the CROPS method could not find a penalty to produce the exact number of known change points, incurring in the presence of extra points in 27% and 24% of the single and double change point series, respectively. When the methodology correctly discovers the change points the presence of extra points is reduced, i.e. 9% and 14%. The combined PELT and CROPS successfully found 66% of the single change points, having an 48% discovery rate of the two change points simultaneously and 87% of at least one (Table 1).

Similarly, BP showed a remarkably high discovery rate of at least one of the two change points (88%) and a smaller rate for detecting the two change points (37%) simultaneously, having a 57% chance of success in the single change point series.

The use of BIC to select the correct number of change points successfully worked in 60% and 44% of the single and double change point series, respectively. This methodology was more likely to underestimate the number of change points, since an extra change point occurred for only 2 and 3% of the series.

Figure 2 – Penalty values of the synthetic series obtained from different penalty functions and CROPS method. Maximum, minimum and median values of each penalty function are shown at the bottom left, bottom right and top, respectively. CROPS maximum, minimum and median values were 6.53, 0.05 and 0.65.



Source: Prepared by the author.

Table 1 – Detection and convergence rate of Pettitt test, Bai and Perron's algorithm and Pruned Exact Linear Time method for the synthetic series.

Detection rate (%)				Convergence rate with PELT (%)					
Method	Single Changepoint	Double change points		Method	Correct change points			Incorrect change points	
		All	Any		Single	Double (all)	Double (any)	Single	Double
PELT	66,17	47,83	89,00	BP	95,10	90,31	91,27	52,52	57,92
BP	57,83	37,83	88,17	Pettitt	97,79	-	97,88	20,59	25,97
Pettitt	45,33	-	39,33	BP and Pettitt	95,59	-	91,27	12,61	23,12

Source: Prepared by the author.

We obtained a success rate of 45% with the Pettitt test for the single change point series. This value was 3% higher than the obtained by Ryberg et al. (2019). The reason for this small difference is probably related to small differences in the random generation of the change point locations due to the use of different R versions. The R version used herein was the 3.6.2.

Even though the Pettitt test is intended to single change point detection it could successfully find at least one change point in 39% of the double change point series.

The BP and Pettitt test presented an extremely high convergence rate with PELT for the correct change points. In most cases (> 91%) when any of these methodologies successfully found the change points, PELT was also successful in finding them (Table 1). For the incorrect change points the convergence rate was smaller. This especially true for the three methodologies simultaneously, where an erroneous change point provided by PELT only matched the other two in 12% of the single change point time series. However, the convergence rate with BP for incorrect change point was higher than 55%.

These results demonstrate the superiority of PELT and CROPS method and indicate that the concomitant use of the three methodologies herein analyzed provides more reliability to the change point results, since the convergence is more likely for the correct points. For the case study, PELT method was chosen as reference for the streamflow time series, classifying the change points according to the convergence of the methodologies. For the climate indices only PELT method was used, and the results were compared to those in the literature.

4.5.2 Abrupt shift in climate indices

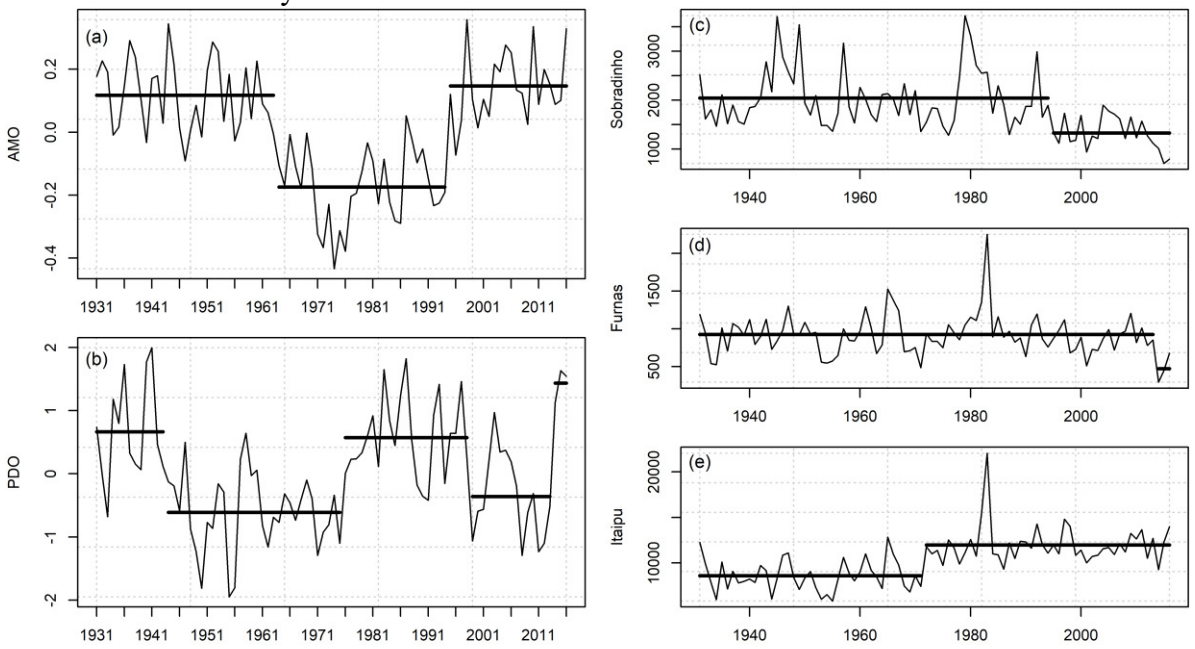
Preliminary tests with the penalty functions did not provide any change points for AMO and Nino3.4. Using CROPS, it was possible to select a parsimonious value for AMO, however, for Nino3.4 no value could be selected since the first found penalty resulted in five change points and small reductions further amplifies its number, showing a overfitting pattern. This penalty behavior imposes that there were no change points for Nino3.4.

The change point analysis segmented the AMO index into three distinct phases (Figure 3a): warm (from 1931 to 1963), cold (from 1964 to 1994) and warm (from 1995 to 2016). Similar results were obtained with the use of the smoothed index (ENFIELD; MESTAS-NUÑEZ; TRIMBLE, 2001), although with a small divergence between the phase change year.

The three periods presented similar values of standard deviation and long durations, over 20 years long (Table 2), where the shorter duration of the last period was affected by the series time window.

For PDO, MBIC did not result in any change points, BIC resulted in two and both AIC and HQ four. Through CROPS, it was noticed that AIC and HQ values were located between the optimal value interval.

Figure 3 – Change point results for the mean value (bold lines) of climate indices and of the streamflow series of key stations.



Source: Prepared by the author.

Table 2 – Statistical properties of each segment: mean (μ), standard deviation (σ) and coefficient of variation (CV)

Index	Period	μ	σ	CV	Duration
AMO	1931 – 1963	0,12	0,11	0,97	33
	1964 – 1994	-0,17	0,12	-0,69	31
	1995 – 2016	0,15	0,11	0,77	22
PDO	1931 -1943	0,66	0,81	1,21	13
	1944 - 1975	-0,61	0,62	-1,02	32
	1976 - 1998	0,57	0,64	1,11	23
	1999 - 2013	-0,36	0,67	-1,88	15
	2014 - 2016	1,44	0,27	0,19	3

Source: Prepared by the author.

Annual PDO series (Figure 3b) shows five distinct phases, warm (1931 – 1943), cold (1944 – 1975), warm (1976 – 1998), cold (1999 – 2013) and warm (2014 – 2016). The change point years are close related to values found in literature, Mantua et al. (1997) mark the years of 1947 and 1977 as PDO polarity reversal times, 3 years after the years obtained. Mary Toshie Kayano, Andreoli, & Souza (2019) considers the period from 1999 to 2011 as a cold PDO phase whereas Qin et al. (2018) defines the cold phase period between 2003-2014, although, the authors of this latter paper states that the filtering method used influenced in the period obtained. For the time window analyzed (1931 – 2016) the indices phase duration varied between 13 to 32 years with similar absolute values of the coefficient of variation (CV) (Table 2), except for the latest period, affected by the time window used.

It is noticeable a warm phase convergence between the PDO and AMO indices for the initial years of the series (1931-1942) and during two periods of three years (1994-1997 and 2013-2016). A cold phase convergence occurred between 1964 and 1975.

4.5.3 Abrupt shift in streamflow series

The Sobradinho station (Figure 3c) shows an accentuated reduction in the mean value for the latest years of the streamflow series, from 2044.71 to 1327.84, comparing to the mean value of the complete series (1861.33) the first period value represented an increase of 9.85% whereas the latest period a decrease of 28.66%. Also, the last three years presented the series' lowest values.

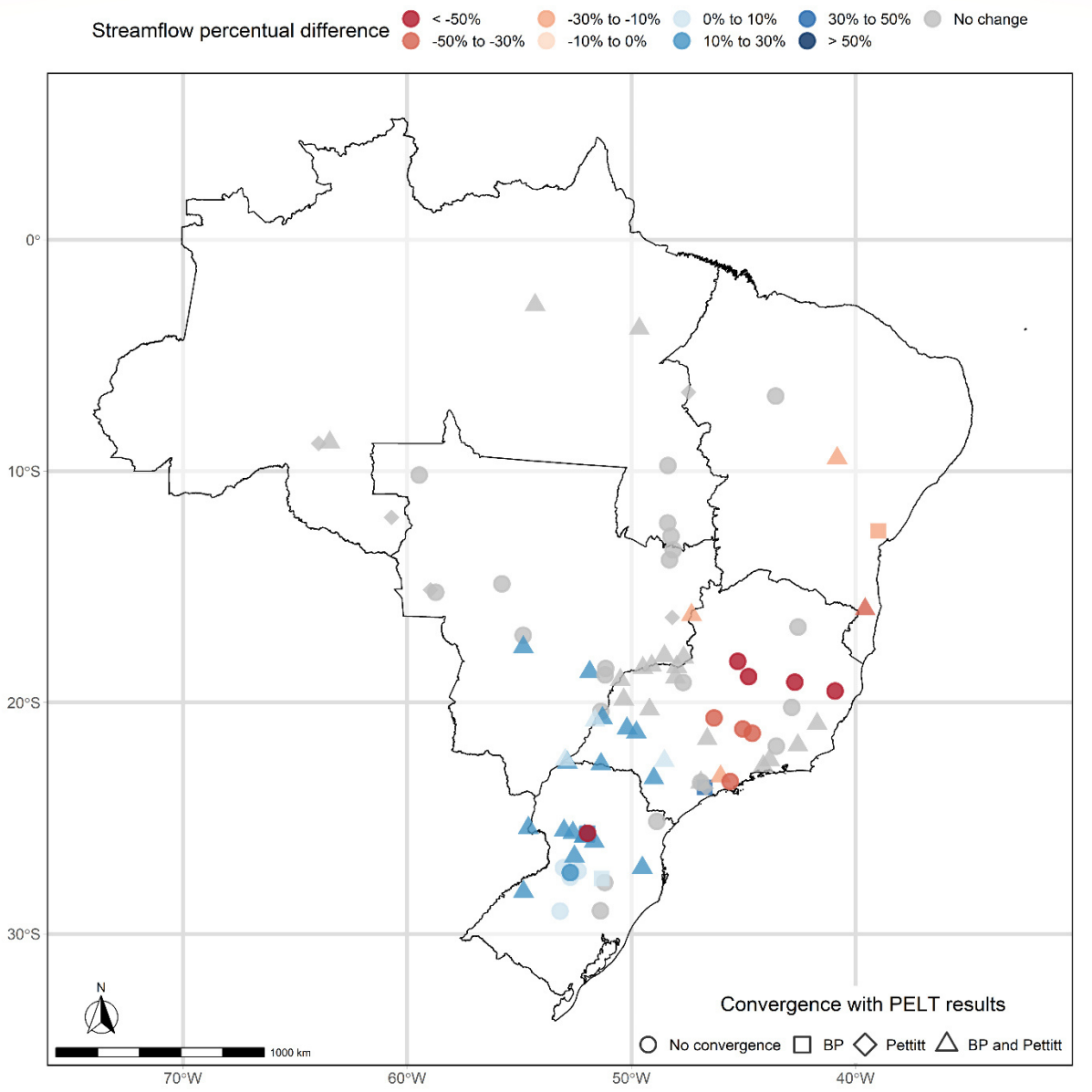
Similarly, Furnas station (Figure 3d) also exhibited a significant reduction in its streamflow mean value, from 925.15 to 474, a percentual difference of 1.73% and -47,87% compared to the mean value of the series (909.41).

The Itaipu station (Figure 3e), however, showed an increase in the streamflow after the change point, from 8597.09 to 11980.08, a difference of -17% and 15% to the complete series mean value (10367.26).

Figure 4 shows spatially the percentual difference between the complete series mean and the mean value of its latest change point period, also showing the convergence between the three change point methodologies. For Sobradinho and Itaipu stations all

methodologies indicated the same change point location, whereas for Furnas, the PELT results did not converge with the other methodologies.

Figure 4 – Percentage difference between the complete series mean value and the latest change point mean value, along with the convergence among the three change-point methodologies.



Source: Prepared by the author.

None of the methodologies used found a change point for the Tucuruí station. For the remaining northern stations, the Pettitt test indicated a change point for three western stations, where for one of these, BP also found a change point. Stations closer to the North ones,

also did not present any change point with the use of PELT, however, for mostly of these stations Pettitt test and BP indicated a change point.

It is noticeable that the stations with streamflow reduction were localized in the eastern region of Brazil, whereas stations with opposite behavior were mostly concentrated close to the south region. A transition zone can be seen between the stations with opposite pattern, where mostly stations showed no change point for all the three methodologies used. The change point location obtained for the southern stations matched between the three methodologies. PELT result for mostly of the stations with a decrease pattern did not match with the other methodologies.

With regard to the penalty functions used, the classical functions did not discover any change point for mostly of the stations, but also imposed a discovery of a large number of change points for a few number of stations. Thus, the CROPS method was used to analyze and select the penalty value.

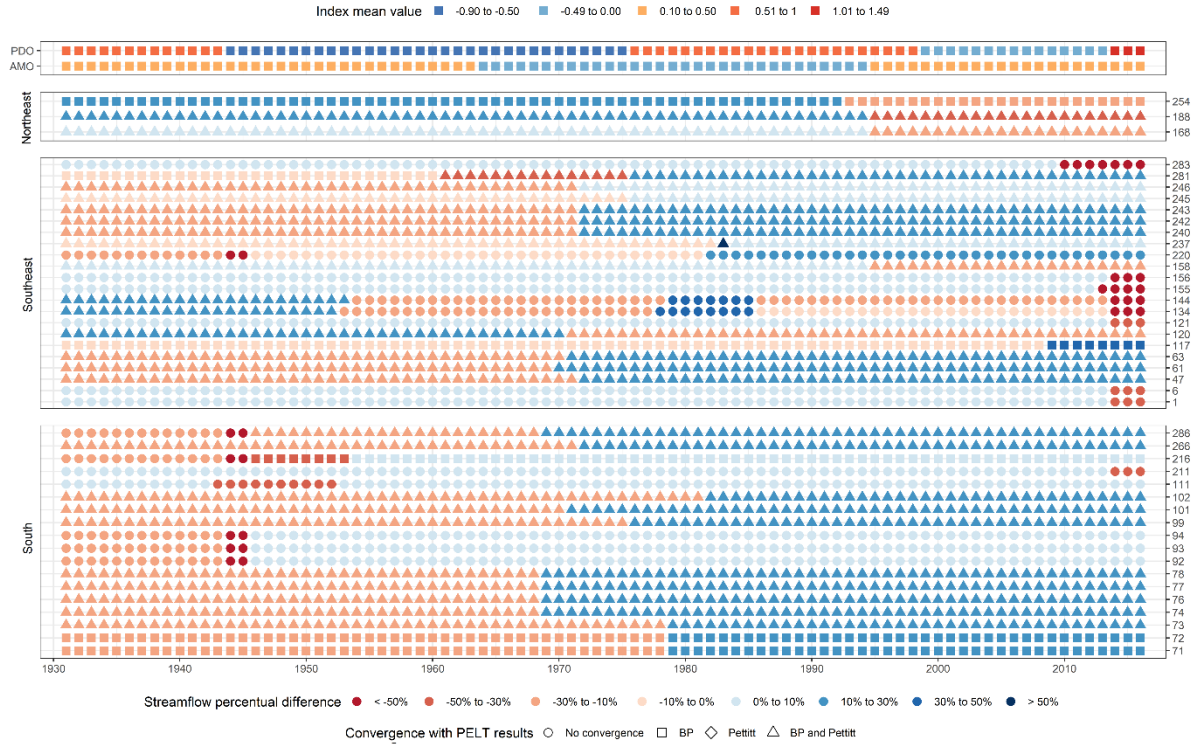
Figure 5 illustrates the indices' mean value of each period and the percentual difference between the period mean value and the streamflow series mean value. The results are shown for each station that presented changes, grouped by ONS standards, along with the convergence between PELT and the other change point methodologies.

The changes detected for the northeast stations occurred in 1994 and 1992, close to the cold-warm phase shift of the AMO index, around 1994 and up to six years before the warm-cold phase shift of the PDO (1998). This also occurred for southeast station 158.

It is noticeable, for the Southeast region, the presence of stations with similar streamflow reduction behavior to Furnas (6), with an accentuated decrease in the latest years, mostly higher than 50%. This also occurred for one south station (211). The reduction observed matches the PDO cold-warm phase shift of 2013. Two of these stations (134 and 144) presented another change point around the year of 1980, close to the cold-warm phase shift of 1975. Also, another change was found in 1952 and 1953 where this change point location matched between the three methodologies but is not too close to any phase shift of the indices.

Stations 281, 246 and 99 showed an increase in the change point matching the 1980 cold-warm phase shit of the PDO, where the station 281 also showed a change close to the warm-cold phase shift of the AMO, this change point location was matched between BP and PELT.

Figure 5 – Comparison between mean value change point results of climate indices and annual streamflow. Numbers on the right are station numbers; 168, 6 and 266 refer to the key stations Sobradinho, Furnas and Itaipu, respectively.



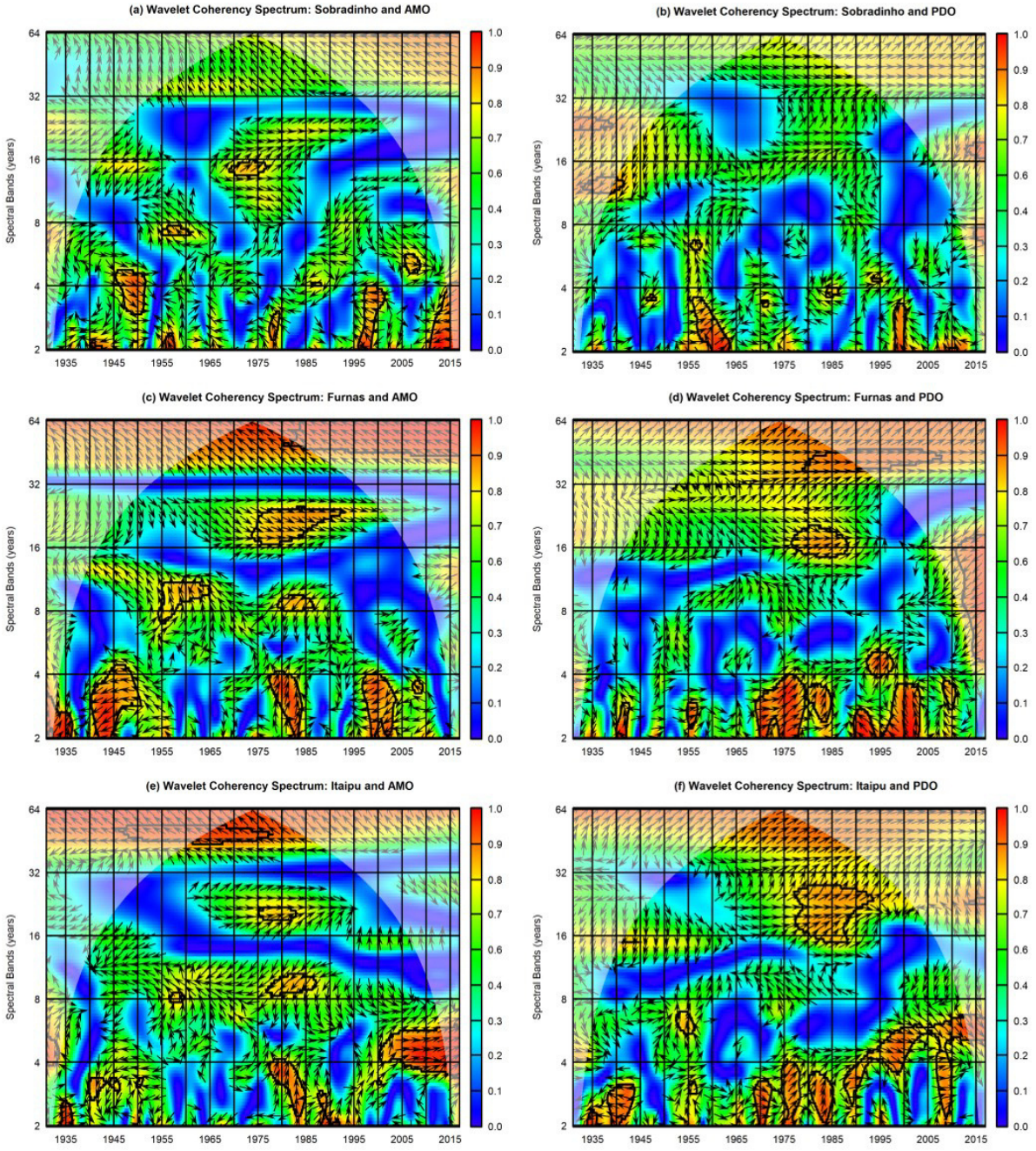
Source: Prepared by the author.

Several stations presented a streamflow increase around the year of 1970, including Itaipu (266). This change point location stands between phase shifts of the two indices and within the cold phase convergence of the indices. One station (120) presented opposite pattern around the same year. Three south stations presented a change in 1978, close to the cold-warm phase shift of the PDO.

PELT detected a change point associated to low streamflow years in five south and one southeast stations, these points, along with the periods found for stations 111 and 216, were close to the warm-cold PDO phase shift.

Figure 6 shows the results of the WTC between the key stations with change points and the AMO and PDO indices. It its noticeable that the indices had correlation (wavelet squared coherency) with the streamflow series for the low-frequency bands (32 to 64 years), especially for Furnas and Itaipu (> 0.8). However, the arrows indicated an opposite lag relationship between the indices low-frequency bands and the streamflow, with AMO out of phase and leading (arrows pointing up and left) and PDO almost in phase, showing a slight lag pattern (arrow point right and slight up).

Figure 6 – Wavelet coherence results between the key stations and climate indices, AMO (left) and PDO (right). The thick black contour line indicates the 5% significance level. The cone of influence is represented in white and indicates a border effect in the highlighted results.



Source: Prepared by the author.

WTC between Sobradinho and PDO (Figure 6b) shows a correlation (~0.6) between the index and the streamflow for almost the complete series for the spectral bands from approximately 12 to 32 years, whereas no correlation is found from 1995 to 2010, for bands from 8 to 32 years. This pattern combined with the cold-warm phase shift of AMO could be the reason behind the change point found in the year of 1995, also low-frequency bands showed

the highest correlation in the latest years, however, outside the cone of influence. AMO (Figure 6a) presented a more continuous influence along the complete series, except within 1950 to 1970 and 2000 to 2016 for the spectral bands of 16 to 32 years.

Furnas (Figure 6c and Figure 6d) presented a consistent correlation with the spectral bands of 16 to 32 for both indices. PDO low-frequency bands presented higher values starting from 1975 where the low-frequency influence of AMO remained constant along the whole streamflow series. The latest cold-warm PDO phase shift combined with the higher influence of PDO spectral bands from 4 to 16 years, could be the reason behind the change point observed in the latest years, however these values stand outside the cone of influence.

Itaipu (Figure 6d and Figure 6e) presented a consistent high correlation with the low-frequency bands of both indices, also with PDO presenting a clearer lag pattern (arrow pointing up and right) from 1935 to 1995. Since 1965 the spectral bands from 16 to 32 years of both indices presented higher correlation values and in-phase relationship. This pattern lasted until 1995 with AMO and for the remaining years with PDO. The leading and lagging pattern of the low-frequency bands of the indices could be one of the reasons behind the discovery of change points in the period after the AMO phase shift of 1963 and before PDO phase shift of 1975.

4.6 Discussion

4.6.1 Multi-change-point discovery performance

PELT showed good results when combined to CROPS method to select the adequate penalty value, whereas for both the synthetic climate indices and annual streamflow time series the penalty functions AIC, BIC, MBIC and Hannah Quinn mostly inhibited the change point discovery. The overfitting pattern found to generate extra change points can be distinguished in a graphical analysis reducing the subjectivity of the penalty selection. However, false positives were produced even when the number of change point was known due to the impossibility of obtaining the exact number of change points and were more common when PELT failed.

The presence of change points obtained for the period of one or two years in the case study indicates that extreme short-period events can influence in the change point

discovery of PELT. This observation is corroborated by Dorcas Wambui et al.(2015) who concluded that the power of PELT method is increased with the size of change. This particularity implies that PELT failure in some of the synthetic series could be a result of small changes imposed to the synthetic time series, a hypothesis that could be confirmed by future research, since it was not analyzed in this work.

Compared to the BP and the Pettitt tests, PELT presented the best performance. The convergence results showed that the use of the three methodologies combined has a small chance of converging to an erroneous change point, and a high chance to converge for the true change point location. The combined use of just BP and PELT had a higher convergence rate for false positives. In other words, there is a clear trade-off between the overall discovery rate and the reliability of the results since the use of multiple methodologies limits the overall performance to the worst discovery rate between the methods. The use of several methodologies to select the correct change point can be found in some of the literature (JIMÉNEZ-RUANO; RODRIGUES MIMBRERO; DE LA RIVA FERNÁNDEZ, 2017; TONGAL, 2019), but it is not a common practice. We find that our multi-test approach is currently the best option since an optimum method is far from reality. Further studies are required to improve the overall discovery rate, verifying the convergence rate of different change point methodologies and selecting the best complementary group of methods. The necessity of further studies is magnified for hydrological time series, since related articles that analyses the convergence of the methods were not found in the literature.

4.6.2 Case Study

The spatial pattern obtained in our case study shows some similarities to the cluster results of Lima and Lall (2010b) through principal component analysis, where the inflow of the reservoirs was subdivided mainly in two clusters with an transition zone cluster in between. The physical explanation given by Lima and Lall regarding to the grouping of Southern Northeast and Southeast stations in the same cluster is the influence of the South Atlantic Convergence Zone (SACZ) in the rainy season. In comparison to the results obtained here, the Southeast stations show two opposite patterns, and the transition zone suffered an upward right displacement, resembling more closely the spatial pattern of SACZ. This opposite pattern was

already reported by Muza et al. (2009) in their analysis of extreme precipitation in the South America.

AMO influence in the SACZ was suggested by Chiessi et al. (2009) when analyzing proxy records of the discharge of La Plata River Drainage Basin. They concluded that during cold phases of AMO, anomalous warming of the South Atlantic would increase the activity of SACZ and displace the main belt of South Atlantic Summer Monsoon to the south, with the opposite occurring in the warm phase of AMO. The displacement of the SACZ southward is associated to increased rainfall in the southern Brazil (BARROS et al., 2000) and justifies the change point location and the streamflow increase pattern found for several stations within the cold phase of AMO. The WTC results for Itaipu corroborate this conclusion since they also show a consistent strong low-frequency AMO influence.

The direct influence of PDO on the SACZ is not clear, but the correlation shown in the WTC of Itaipu, especially for the period starting in 1975, could be the reason why the cold–warm change of AMO did not produce another change point. The influence of PDO could be related indirectly via ENSO's influence on SACZ. According to Cavalcanti (2016) warm phases of ENSO apparently reinforce the persistence of oceanic SACZ. Ferreira, Sanches, & Silva Dias (2004 *apud* CAVALCANTI, 2016) also found that a more intense oceanic activity is favored during warm El Niño phases.

Although no change point was found for Niño 3.4, the literature associates the 1975 cold-warm phase shift of the PDO both to changes in atmospheric and oceanic circulation over the North Pacific and to the occurrence of more intense and frequent El Niño events (KELLER et al., 2009). This justifies the corresponding increased correlation of PDO found in the WTC of Furnas and Itaipu, and the change point location of Furnas occurring within a large influence of PDO in the spectral bands 4 to 16 and 32 to 64. Although the change point and WTC results show strong coherence, it is important to emphasize that only PELT found Furnas change point and its WTC results stands outside the cone of influence.

The reason behind the reduced correlation in the WTC among the spectral bands of 4 to 32 years of Sobradinho and PDO for the period between 1995 to 2010 is not entirely understood but finds similarities with the findings of Rocha et al. (2019), where the PDO influence in the rainfall of an northeast basin did not occur along the complete series, but mainly until 1975, and AMO correlation was consistent along the whole period. The similarity with a result obtained in a region not much affected by the SACZ indicates the concomitant influence

of another atmospheric pattern. The highest correlation of low-frequency bands of PDO for the latest years can explain the series lowest values occurring simultaneously with the index's latest warm phase and shows that despite the localized reduction noticed for higher frequency bands the low-frequency phase change of PDO had a strong influence in the latest values of the streamflow series. However, this result should be treated with caution since it occurred outside of the WTC cone of influence. The complex multi climate pattern influence in Sobradinho was expected since it is a dam of São Francisco river which has a large interstate hydrologic basin with significant climate differences between each state.

4.7 Conclusions

In this paper we used multiple change point detection methodologies to find and analyze changes in the mean value of Brazilian annual streamflow time series and its possible association to large scale global SST oscillations. We also used wavelet coherence to verify coincidences found between abrupt changes observed in the streamflow and climate indices. A preliminary test with synthetic series was performed to investigate the detection skill of the three change point methods used (PELT, Bai and Perron's algorithm and Pettitt test) and the particularities regarding PELT penalty selection specifically for hydroclimatic time series.

From the results a few methodologic and local conclusions arise:

- PELT, with the use of CROPS to select the penalty value, had the best performance of all three methods. On the other hand, the use of standard penalty functions produced poor results.
- PELT has a tendency of producing false positives that cannot be overlooked.
- The concomitant use of the three methods provides reliability to the results since they were more likely to agree (> 90% chance) for the correct change points, with a smaller convergence for erroneous results (< 24%).
- Brazilian streamflow shows a clear opposite spatial pattern with Southern stations showing an increase in the streamflow mean value after the change point and upper Southeast and Northeast stations with a decrease.

- The changes found show strong coherence with AMO and PDO phase shift and a complex concomitant influence of both SST low-frequency oscillation pattern.
- Mostly, the changes observed seems to be related to alterations in the SACZ imposed by the low-frequency oscillations and possibly associated to PDO related changes in ENSO.

5 STREAMFLOW FORECAST THROUGH DYNAMIC HYBRID GAUSSIAN NETWORKS INCORPORATING THE INFLUENCE OF LOW-FREQUENCY SST PHASES

5.1 Abstract

The streamflow forecasting performance of a data-driven Dynamic Hybrid Gaussian Network was analyzed for the lead time from one to eight months. Different network configurations were used to investigate the influence of low-frequency SST oscillations phases and the impact of two normalization procedures. The streamflow was also predicted through two different methods: (1) direct use of the Gaussian Bayesian Networks parameters and (2) via Bayesian inference, i.e. Likelihood Weighting simulations (LW). The Bayesian networks, in general, presented a good performance for streamflow forecasting, outperforming previous local studies, also having a great flexibility: using complex and simpler networks when required. The normalization of the streamflow resulted in the improvement of the forecast for some stations, also the use of the phases of low-frequency sea surface temperature indices as discrete variables resulted in a significant improvement in the performance for some stations. When the 1-month lead time network shows a good performance, it can be used to predict streamflow of higher lead times through LW. However, different set of node parents were obtained for the different network configurations tested. This aspect needs to be thoughtfully addressed and the network structure obtained profound analyzed before being used for causal inference and for the analysis of phenomena impact in the streamflow. For the case study the Bayesian networks presented good results for the Northeast, North and Center-west regions of Brazil. The southeast region presented relative inferior results and the methodology used did not present good results for mostly stations in the South region.

5.2 Introduction

Streamflow forecasting plays an important role in water resources management and planning, being the main objective of diverse articles through numerous methodologies (BEATRIZ et al., 2020; CHU; WEI; QIU, 2018; LI et al., 2019; LIMA; LALL, 2010c; SILVEIRA et al., 2017). One key aspect in the development of forecasting models is the

selection of the exogenous predictors, for this instance, climate indexes have already been widely used in the literature (CHU; WEI; QIU, 2018; LI et al., 2019; LIMA; LALL, 2010c; SILVEIRA et al., 2017). The reason for this is that streamflow is known to respond to large-scale atmospheric teleconnections patterns of both low and high frequency (RAVINDRANATH et al., 2019). Also, the use of climate indexes can provide lead time to the predictions, an important factor in water management, especially in countries with strong dependency of water resources for hydroelectric energy generation.

However, the use of multi-scale climate indexes as predictors in streamflow forecasting models is not a trivial task, especially due to the existence of an intrinsic inter-relationship between the large-scale climate phenomena synthesized by the indexes and its simultaneous complex time-lagged influence in the precipitation patterns that drives streamflow. To cope with this problem, authors have recently used sophisticated methodologies. For instance, Chu et al. (2018) have combined three different methodologies: Ensemble Empirical Mode Decomposition (EEMD); Least Absolute Shrinkage and Selection Operator (LASSO); and Deep Belief Networks (DBN) to screen a significant number of climate predictors, select the adequate ones for each Intrinsic Mode Function (IMF) obtained and predict monthly streamflow. Another example is Li et al. (2019) multi-model integration combining the results of four different methods, Elastic Net Regression (ENR), Support Vector Regression (SVR), Random Forest (RF) and eXtreme Gradient Boosting (XBR) to predict monthly streamflow using 130 climate indices, selecting the adequate ones for each model through regression mechanisms.

Although these approaches present a high performance, it is hard to confirm if the resulting models correctly represent the physical relationship between the climate phenomena they are based on. One reason for this is that these approaches were not initially developed to achieve such goals. Furthermore, extracting the causes from a deep learning black box is commonly not possible, even though interpreting deep learning models is an active area of research (MONTAVON; SAMEK; MÜLLER, 2018; RUNGE et al., 2019a).

The physical aspect of the model predictors is important to verify assumptions about the influence of large-scale climate phenomena in streamflow and its long-term projections, for instance, regarding the impacts of phase change of low frequency patterns, e.g. Atlantic Multidecadal Oscillation (AMO) and Pacific Decadal Oscillation (PDO). The phases of these low frequency sea surface temperature (SST) oscillations have been associated to changes in the frequency and intensity of El Nino Southern Oscillation (ENSO) events and also

in the strength of its teleconnections with rainfall patterns of South America (KAYANO; ANDREOLI, 2007; KAYANO; CAPISTRANO, 2014). Recent studies show that the phase shifts of these patterns can also be related to changes detected in the mean values of streamflow series (ROCHA; DE SOUZA FILHO, 2020).

From these studies, a model considering explicitly the phases of low frequency oscillations presents an interesting hypothesis. Such a model was not found in the literature and, if proven to be a good approach, can be a solution to improve the forecast and diminish a dynamic risk imposed by the different phases of low frequency oscillations.

A Bayesian network is a causal graphical model which consist in two parts: the graph structure, determining dependencies between variables, and the conditional probabilities associated, simultaneously allowing for a probabilistic forecast and a direct analysis of the graph structure for physical causalities. Bayesian networks have been recently used to discover causality in climate science as a data-driven approach, obtaining strong results and showing that physical meaning could be extracted from the graph structure obtained through causal network learning algorithms (EBERT-UPHOFF; DENG, 2012, 2015; RUNGE et al., 2019a). Another aspect is that this methodology has the flexibility to be combined with classical autoregressive methods (AR, ARMA) producing good streamflow forecasting results (MEHDIZADEH; KOZEKALANI SALES, 2018) and be used with mixed data (continuous and discrete).

Causal discovery algorithms for Bayesian networks are either based on a series of conditional independence tests (CI), to detect independence relationship and learn the graph structure as a constraint-based optimization, or based on score functions minimization, e.g. Bayesian Information Criteria (BIC). Although only recently used in climate science, Glymour et al. (2005) already found that this structure learning approach could possibly retrieve relationship between climate indices, their results showed a partly agreement with the literature expected behavior when testing for improvements in the classical PC algorithm (GLYMOUR; SPIRTES, 1991). Furthermore, Ebert-Uphoff & Deng (2012) also obtained consistent results when analyzing four climate indexes in a daily step, thus, emphasizing the usefulness of this approach for climate indices relationship extraction.

In the Water Resources field, Bayesian networks are mostly commonly found in the literature related to water management, especially those focusing on water quality and strategic decision-making (PHAN et al., 2016). In streamflow forecasting, for example, Ramadas et al.

(2015) used a gaussian graphical model for predictor selection from a set of eight physical variables, concluding that this methodology was more robust than a predictor selection through partial mutual information method. In a more recent article, Dutta & Maity (2020) used time varying Bayesian networks to account for nonstationary and forecast streamflow using data of the previous month, comparing the results with existing models and obtaining significant results. This latter study, however, did not analyze the performance Bayesian networks to forecast streamflow for higher lead times or used climate indices in the networks. Thus, the possibilities of a data-driven Bayesian network to improve streamflow forecasting and aid in answering state-of-the-art questions such as how streamflow is affected by low-frequency oscillations is not yet fully explored.

This paper investigates the streamflow forecasting performance of a data-driven Dynamic Hybrid Gaussian Network, testing diverse network configurations with the use of mixed data (continuous and discrete variables), incorporating the influence of low-frequency SST oscillations phases, analyzing the impact of two normalization procedures and the forecast performance with lead times from one to eight months. This methodology seeks to determine the best network structure among the cases tested and verify the necessity of a normalization procedures, obtaining a framework that could be used for both causal discovery of streamflow forcing and streamflow forecast.

We used Brazil as case study due to its continental size, enabling a relevant spatial analysis and comprising regions with different climate behavior caused by the dynamic iteration of the atmosphere with both Pacific and Atlantic Ocean, imposing a challenge for the structure learning algorithms. The country also has a strong dependence of its water resources for hydroelectric energy generation since it is the main country's power source, 66.6% of its total energy supply (636,4 TWh) (Ministry of Mines and Energy, 2019).

5.3 Data and Methodology

This section is organized in five topics. The first, presents the climate indices used as exogenous predictors and a brief revision of their influence in Brazilian climatology. The second and third topics present the particularities regarding the use of hybrid Bayesian networks for modelling streamflow series with a strong seasonal component and briefly revises the literature on previous local research, which were the basis for building the methodologic

approach in this work. The fourth topic presents the analyzed normalization procedures. Finally, the good-of-fit measures and details about the data separation periods (calibration and simulation) is presented, both used to evaluate the Bayesian network streamflow forecasting performance through two different methods: direct use of the Gaussian Bayesian Networks parameters, and through Bayesian inference.

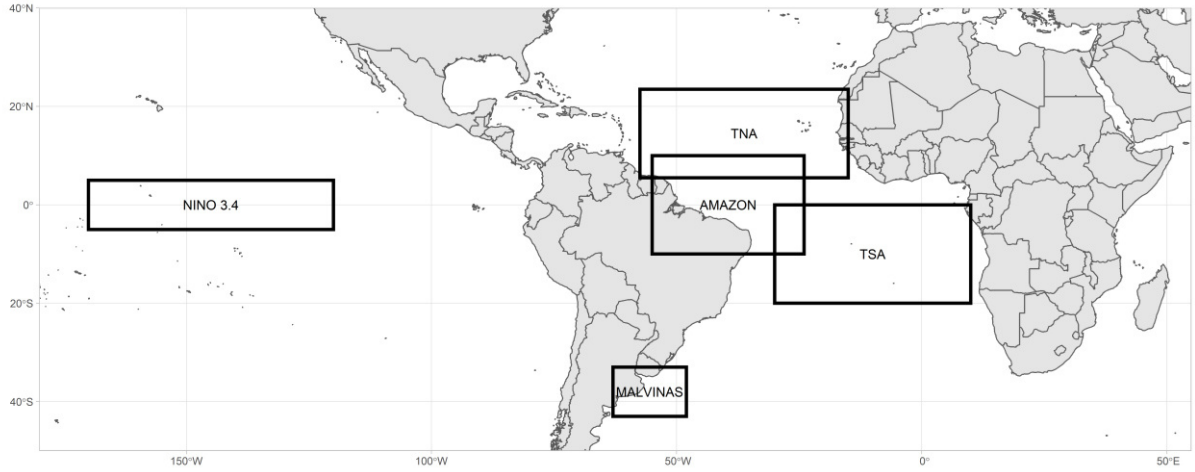
5.3.1 Hydroclimatic data and teleconnections

We used monthly naturalized streamflow data for 88 stations distributed throughout Brazil provided by the Electric System National Operator (ONS). The naturalized flows were obtained after a process detailed in ONS (2017). Briefly, this process starts with the consolidation and consistency of the measurements from streamflow gauges, followed by the account of anthropogenic factors such as reservoir operations and water withdraw along the basins.[3]

The climate indices used were Niño 3.4, TNA and TSA (Tropical Northern and Southern Atlantic Indices), AMO, PDO, Malvinas (CATALDI et al., 2010) and Amazon (Figure 7). Data for the first four indices were acquired from the National Oceanic and Atmospheric Administration (NOAA) timeseries database. The TNA and TSA indices start in 1948 and were extended to 1931 using NOAA Extended Reconstructed Sea Surface Temperature Version 4 (ERSST V4). These indices were merged into one, here named Atlantic Gradient (Grad), obtained by the difference between TNA and TSA. Malvinas index is calculated averaging the sea surface temperature (SST) of the area between 43°S to 33°S and 63°W to 48°W. The index here named Amazon refers to the region comprised by 10°N-10°S; 55ºW-24ºW, both indices were calculated using data from the NOAA ERSST V4.

[3] Further information about the stations is presented in the Appendix A and D.

Figure 7 – Geographical location of the higher frequency climate SST indices used.



Source: Prepared by the author.

The index selection was based in previous regional studies. El Niño Southern Oscillation (ENSO) is recognized as the most significant oceanic-atmospheric mode to interannual climate variability on a global scale (KAYANO et al., 2016). Its influence in South America rainfall is derived of associated alterations in the Walker and Hadley circulation cells. Its impact on Brazilian rainfall has already been broadly reported, and the index was also used as a streamflow predictor. Niño 3.4 index is defined by the region 5°N-5°S, 170°W-120°W, which is considered the main region to analyze ENSO related coupled ocean-interactions.

AMO and PDO are low-frequency SST oscillations which studies impose a relationship between those indices and rainfall in Brazil and South America (ANDREOLI; KAYANO, 2007; KAYANO; CAPISTRANO, 2014; ROCHA; SOUZA FILHO; SILVA, 2019). Those low-frequency oscillations modulate ENSO influence in South America Rainfall, where phase convergence between PDO and ENSO seems to amplify the expected effects of the phenomenon and similar behavior is expected with phase divergence between ENSO and AMO (KAYANO; ANDREOLI, 2007; KAYANO; CAPISTRANO, 2014). AMO index is obtained detrending the mean SST anomaly of the North Atlantic (0°), to remove climate change interferences (ENFIELD; MESTAS-NUÑEZ; TRIMBLE, 2001). It is considered the dominant low-frequency variability mode of the Atlantic Ocean.

PDO is the main component of the SST variability of the Pacific Ocean. The index measures the anomaly of North Pacific SST intensity by the total of standard deviations from the mean values of the climatology period (MANTUA et al., 1997).

Rocha and Souza Filho (2020) detected change points in a significant number of the streamflow series herein used and associated the changes to phase shifts of AMO and PDO. This result combined with the reported influence of the indices phases in ENSO influence in South America rainfall led to the discretization of the series:

- AMO: Warm from 1931 to 1963 and 1995 to 2016; Cold from 1964 to 1994.
- PDO: Warm from 1931 to 1943, 1976 to 1998 and 2014 to 2016, Cold from 1944 to 1975 and 1999 to 2013

These indices were used in the discretized form, included as variables into the networks, thus, the henceforth mention for the climate indices AMO and PDO in this work refers to the discretized series.

Anomalous SST meridional gradient in the off-equatorial tropical Atlantic is known to be associated with the displacement of the Intertropical Convergence Zone (ITCZ) affecting the rainfall distribution in the northern of the Northeast region of Brazil and in the eastern Amazon (NOBRE; SHUKLA, 1996). This convergence zone is the main system behind the rainfall of the northeastern region of Brazil and its anomalous displacement is the primary source of the common droughts faced in the region. ENSO also influences the formation of the meridional gradients of anomalous SST due to atypical east-west circulation cell between the eastern Pacific and the equatorial South Atlantic (NOBRE; SHUKLA, 1996).

The Brazil-Malvinas Confluence is considered one of the most important features of the South Atlantic Ocean. The confluence between the cold waters of Malvinas Current and the warm waters of Brazil Current forms a thermal front that is associated to alterations in vertical circulation patterns. Years with a positive anomaly are related to a reduction in the rainfall of the South region of Brazil and an increase in the Southeast region. The physical mechanism behind these alterations is the link between the Brazil-Malvinas Confluence and the South Atlantic Convergence Zone (SACZ), where positive anomaly years are related to better conditions to the zone formation (CATALDI et al., 2010). Although the main influence of this phenomena is mainly observed in the South and Southeast region of Brazil, its influence in other regions of Brazil is also mentioned in the work of Cataldi (2010).

The Amazon index area was selected based in previous studies that suggests low-frequency oceanic responses to Amazon runoff (HUANG; MEHTA, 2010; JAHFER et al., 2020; JAHFER; VINAYACHANDRAN; NANJUNDIAH, 2017; VIZY; COOK, 2010). The

results obtained by Jahfer et al. (2020) simulating the impacts of the removal of the amazon runoff and doubling its value shows an significant impact in the areas comprised by the TNA and TSA indices, thus its effect is indirectly accounted in the derived Grad index. The area selected in this study stands between the two indices in a SST region not much affected by the Amazon runoff and can possibly work as a control area for the mentioned effect. Jahfer et al. (2020) also associate the Amazon runoff in the displacement of ITCZ, in concordance to the mentioned effect of an anomalous SST meridional gradient.

According to Nnamchi et al., (2016), this area also do not comprise regions statistically significant to ATL3 index (ZEBIAK, 1993) during its peak phase in the boreal summer (JJA – June through August). ATL3 is used to account to the Atlantic Niño phenomena that dominates the variability of tropical Atlantic and was not included in this work due to its region be partially comprised by TSA.

5.3.2 Dynamic Hybrid Bayesian Networks

Dynamic Bayesian networks differ from static networks by representing a time series into several different nodes according to the time lag of interest. Dynamic networks follows some premises: the stochastic process is a first-order Markovian; the random variable X(t) is conditionally independent given the observations at a previous time t-1; and the process is homogenous through time (NAGARAJAN; SCUTARI; LÈBRE, 2013).

However, the presence of change points in the streamflow values, violates, for these cases, the homogeneity premise. A textbook solution to this problem is the use of Non-Homogeneous Dynamic Bayesian Networks (NHDBN), where one network is obtained for each time series segment, before and after the detected change. For the case study, however, diverse change points are located close to the beginning or the end of the series making it impossible to use NHDBN due to the lack of sufficiently long series.

The use of the AMO and PDO as discrete variables encompasses the possible modulating effect of the phases in the rainfall of the region, and indirectly in the streamflow, and pose as an alternative solution to the use of NHDBN, since several of the change point reported by Rocha and Souza Filho (2020) occurs close to phase-shift of these indices.

The inclusion of discrete variables in a Bayesian network defines a hybrid Bayesian network and results in different network parameters for each discrete value of the discrete parent node. Therefore, a node with AMO and/or PDO as a parent would have a model for each phase or phase combination (in the case of a joint influence of both AMO and PDO). This methodologic decision can result in models fitted for longer periods, where shorter periods only occurs if the joint influence of AMO and PDO is detected by the network learning algorithm.

The implicit premise of this methodology is that the indices influence occurs for the entire series, since the node parents remain constant through time, but the nodes influence can modify in scale and in signal according to the phase. This characteristic can aid in the analysis of different influences according to the phases of these low-frequency SST modes, however, has as a disadvantage of a NHDBN the impossibility to the detect a temporary phase-limited influence.

5.3.3 Bayesian Networks Structure Learning in face of periodicity

Historically, several studies model the streamflow of the Brazilian electric sector considering explicitly the seasonality and obtaining good results. The good performance of this practice was first demonstrated by Maceira *et al*. (1987 *apud* Alexandre, 2012) who compared the performance of a Periodic Autoregressive (PAR) to stochastic models used at the time. Latter studies of Alexandre (2012), Lima & Lall (2010) and Silveira et al. (2017) used PAR and PARX (Periodic Autoregressive with exogenous variables) to represent the local streamflow and reaffirmed the good results derived by considering explicitly the seasonality. Alexandre (2012) concluded that a PAR model can obtain similar results to a PARX for some stations, even with the use of many climate indices (27) as exogeneous predictors. This result reinforces the idea that the use of a separate equation for each month is the main reason of the good results of these studies. The performance of the periodic autoregressive family models is also recognized by ONS which use these models as standard to forecast streamflow and generate future scenarios.

The PARX studies mentioned all have in common the use of a fixed lag between the streamflow series and the predictors indices. This method has an implicit premise that the influence of different indices occurs in the same temporal lag. However, the influence of the SST of different regions of the ocean or of different oceans in the local rainfall is likely to occur

with variable lags and with possible significant spatial variations. Therefore, the inclusion of variable lags to the exogeneous predictors has the potential to improve the good performance mentioned.

The dynamic Bayesian networks seems to be an promising solution to the inclusion of this factor in the streamflow modelling, since network structure learning algorithms have been proved efficient into finding complex lagged causal relationships in climatic variables (NOWACK et al., 2020; RUNGE et al., 2019a).

Several preliminary tests were performed with a reduced number of stations to evaluate the structure configuration of the networks. First it was used one node for each continuous variable for each lag from one to up to 24 months. Second, for the initial configuration it was added an extra discrete variable corresponding to the month, which results in the approximation of a model to each month, explicitly considering the seasonality. However, the overall best result was obtained for the latest tests, which considered each month as a node and another set is added to contemplate a one-year lag, resulting in a set of 24 nodes for each continuous variable. This latter methodology was adopted for this study.

This result derives from the fact that the use of lagged nodes for the entire series imposes the use of the same predictors for all months since the parents are constant. Also, this configuration is consistent with the local results since it more closely resembles the methodology of the aforementioned studies. In any case, preliminary tests were performed to a reduced number of stations and may not apply for all 88 stations.

Score function minimization was selected as the network structure learning methodology. This choice was based in the knowledge that both methodologies can successfully obtain good results and also in the recent findings of Scutari *et al*. (2019) which concluded that score-based algorithms can often obtain network structures with better accuracy than constraint-based algorithms and that both of those methodologies outperformed hybrid algorithms that combine score-based functions and conditional independence tests. All algorithms were employed through the R package *bnlearn* (SCUTARI, 2010).

We used the *Hill-Climbing* (HC) algorithm, which performs a *greedy search* exploring the search space from an initial network structure, usually an empty graph, with the addition, removal or reversal of one arc at a time until the score function value cannot be further improved. This algorithm presented good results in the test made by Marco Scutari et al. (2019)

using real world geospatial data. Briefly, it follows the following steps (NAGARAJAN; SCUTARI; LÈBRE, 2013):

1. *Choose a network structure **G**;*
2. *Compute the score of **G** (**ScoreG**);*
3. *Set the maximum score equal to **ScoreG;***
4. *Repeat the further steps until the maximum score stops increasing:*
    a. *For every possible arc manipulation (addition, removal or reversal) that do not result in a cyclic network:*
        i. *Compute the score of the new modified network **G*** (**ScoreG***)*
        ii. *If **ScoreG*** > **ScoreG**, set **G = G*** and **ScoreG = ScoreG****
    b. *Update the maximum score with its new value*
5. *Return the directed acyclic graph G.*

For this work we used the BIC score. To prevent the stop of the search process in points of local minima the structure learning process was run several times through bootstrapping, averaging the networks obtained with a significance threshold selected through the methodology proposed by Scutari and Nagarajan (2012 *apud* NAGARAJAN; SCUTARI; LÈBRE, 2013). This latter step produced an increase in the overall performance of the network in the preliminary tests, up to 10% in good-of-fit measures such as R^2 and NASH.

One common additional step is to prevent the network structure learning algorithm from finding unwanted arcs that are either considered impossible due to physical reasons or expert knowledge. The use of the network learning algorithm with monthly variables and a lead time can incur in violations of the physical and temporal meaning of the network if used without restrictions. Thus, restrictions were imposed to the learning algorithm, certifying:

- Temporal coherence: it is not allowed directed nodes from a previous time step, i.e. a directed arc from March streamflow to February.
- Lead time coherence: it is not allowed arcs from variables in a time step incompatible with the lead time, i.e. for a network with a three months lead time and for the node of march streamflow, it is not allowed arcs coming from February and January of the same year.
- Physical coherence: it is not allowed directed arcs from streamflow nodes to indices.

Finally, the networks were created independently for each streamflow gauge and for each of the different lead times analyzed.

5.3.4 Analysis of the impact of normalization procedures

Gaussian networks are a particular case of Bayesian networks, being widely used due to the simplicity granted by the use a multivariate normal and linear additive terms to condition the parents influence to the node.

The normality assumption is commonly not valid to monthly streamflow series. However, the mentioned studies used PAR and PARX linear models to model the streamflow series without a normalization procedure, as shown in the works of Alexandre (2012) and Silveira et al. (2017) who used only a scale adjustment between the different stations through an standardization procedure. This fact indicates that despite the violation of normality assumption, a gaussian network applied directly to the flows can successfully represent the flows from the study area.

To verify the normalization procedure impact in the network forecasting results, two different methodologies were tested:

a) The use of a base 10 logarithm: the literature of hydrological studies commonly presents this approach to approximate normality to streamflow series, mostly in annual series.

b) The use of Standardized Runoff Index (SRI): not commonly used in studies with a streamflow normalization objective, the calculation process of this index in the monthly time step (SRI-1) explicitly imposes normality for each month streamflow series.

The SRI-1 calculation process consists of two-steps: the streamflow series of each month is approximated to a probability distribution that is then transformed to a standard normal with zero mean and standard deviation of one. To allow the comparison between different locations, it is usually used the same probability distribution for all months of the series and for all stations analyzed. Studies shows that the choice of the probability distribution impacts in the correct representation of the streamflow and in the index results. In this study, we used the methodology proposed by Vicente-Serrano *et al.* (2012) which stems from the Kolmogorov-

Smirnov test to select the probability distribution accordingly to the test statistics. In our study one distribution for all months was selected for each of the stations analyzed. The distributions used were *Generalized Extreme Value* (GEV)*, Pearson type III* (PE3)*, Gumbel* (GUM)*, lognormal* (LNO)*, logistic* (LGS) *and Weibull* (WEB).

5.3.5 Forecast performance analysis

The network performance was evaluated using the NASH coefficient, R2, volumetric efficiency (VE), percentual bias (PBIAS %) and Normalized Root Mean Squared Error (NRMSE %). For the normalized cases the coefficients were calculated after returning the streamflow to the original scale.

The network performance was analyzed for two distinct periods: calibration (1937-1997), period where the data was used to obtain the network structure and parameters; and simulation period (1932-1936 and 1998-2016). The data separation aims to verify overfitting problems, which could happen due to the high degree of freedom resulting from the use of variable lags.

The period selection aimed a continuous period for calibration, corresponding for 70% of the data (~61 years), and the remaining years for simulation (~24 years), also aiming to obtain a similar number of years of the AMO and PDO phases for both periods. The evolution of the number of years for each AMO and PDO phase combination for the calibration and simulation periods according to the initial year of the calibration period is presented in the appendix A.

5.3.5.1 Streamflow forecast through Gaussian Bayesian Networks

The global probability distribution of a set of random variables $(X_1, X_2, \dots, X_N)$ within a known network structure is dependent of the local probability distribution of each node (DUTTA; MAITY, 2020):

$$p(X_1, X_2, \dots, X_n) = \prod_{j=1}^{N} p\left(X_j | \, pa_G(X_j)\right) \quad \text{(Eq. 2)}$$

where $\left(X_j | \, pa_G(X_j)\right)$ is the local conditional probability associated to the node *j*. The Markov premise mentioned previously imposes that the state of a random variable is dependent only on its parents. This fact reduces the number of the parameters required to identify the global distribution to the sum of the parameters of the local distributions, which is further reduced due to the conditional independence relationships encoded in the network structure, alleviating the "curse" of dimensionality (NAGARAJAN; SCUTARI; LÈBRE, 2013). The global probability distribution is now given by:

$$p(X_1, X_2, \dots, X_n | \alpha) = \prod_{j=1}^{N} p\left(X_j | \, pa_G(X_j), \alpha_j\right) \quad \text{(Eq. 3)}$$

where α_j is the parameter vector for the conditional distribution of X_j, estimated using the data of the calibration period. Two different methods can be used to estimate the network parameters, either the Bayesian or the maximum likelihood estimation. In this study, the latter was adopted.

The use of Gaussian Networks translates the streamflow forecast to the use of a simple set of linear equations, one for each month, where for this study adopt the following form:

$$Q_{i,m} = \beta_{k|\omega_i} + \sum_{j}^{n} \alpha_{j,k|\omega_i} \theta_{j,i} \quad \text{(Eq. 4)}$$

where: $Q_{i,m}$ is the streamflow of the year *i* and month *m;* $\beta_{m,k}$ is the intercept for the *k* period, according to ω_i; $\alpha_{j,k|\omega_i}$ is the linear coefficient that modulates the influence of the *j* parent in the streamflow node for the *k* period; $\theta_{j,i}$ is the value of the parent *j* in the year *i*; *n* is the number of continuous parents of the node; ω_i is the joint phase of the discrete parents nodes.

These linear equations were used to forecast streamflow for the networks obtained with lead time from one to eight months and its results are referred in this work as GBN.

5.3.5.2 Streamflow forecast through Bayesian Inference

Bayesian Inference can be used to forecast streamflow with a longer lead time using as reference the network obtained with lead time of one month. For this purpose, Likelihood Weighting (LW) was used in this work. In this method the values available for the lead time desired were considered evidence and the streamflow values were computed by averaging likelihood weighting simulations (FUNG; CHANG, 1990) and selecting the expected streamflow value of the resulting conditional distribution. Since this method is based in stochastic simulation small differences can be obtained for each simulation. In this study, 10,000 random samples were used for each new observation, the large number of samples seeks to diminish these small differences.

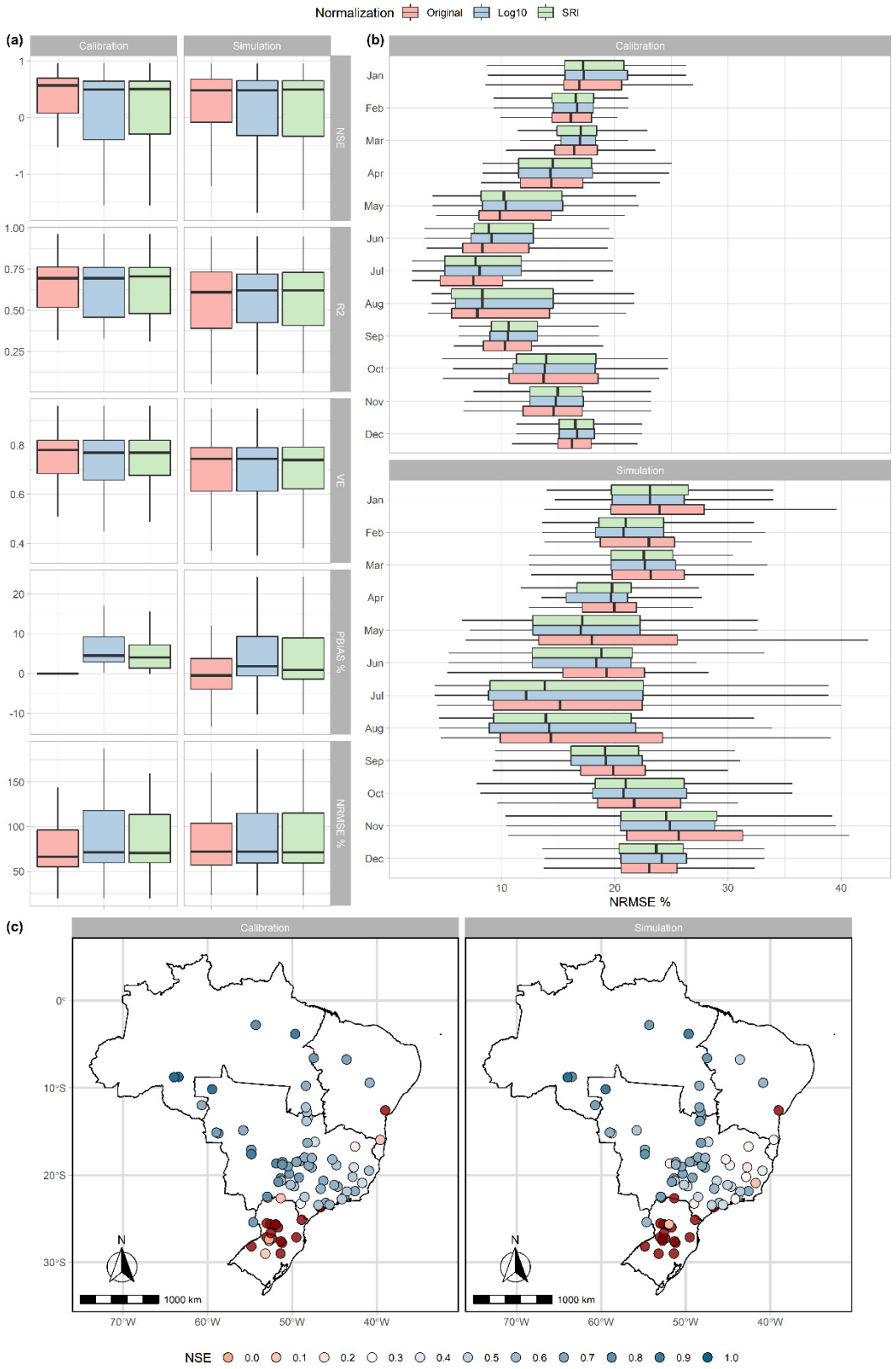
The network obtained with 1-month lead time can be used for causal inference, extracting causality from the links of the nodes of the resulting graph. The direct use of networks obtained independently for each lead time can possibly yield good results, since the network structure learning algorithm optimize the streamflow parameters for the information available. Nonetheless, the analysis of the physical coherence between the streamflow and the indices is harder in these longer lead times networks. Therefore, if the network obtained with a 1-month lead time shows satisfactory results to forecast streamflow through Bayesian inference, this method could be used also to better understand the impacts of climate phenomena in streamflow and the causal chain involved.

5.4 Results

5.4.1 Forecast performance analysis for 1-month lead time

Figure 8a presents the results obtained for the monthly streamflow with and without a normalization procedure for the lead time of one month. The outliers were removed to ease visualization. For both periods analyzed the normalization procedures resulted in lower NASH values for the lower whisker and first quartile, showing a higher frequency of negative values and worst performance for some stations. In a similar way, NRMSE % shows higher values for the upper quartile and higher maximum values, in agreement with the NASH results.

Figure 8 – Results for the calibration and simulation periods for streamflow non-normalized (Original) and normalized through log 10 and SRI. Boxplot without outliers for the results of (a) monthly series (b) NRMSE for each month. (c) NASH spatial results for the original streamflow, dark red dots represent negative values of NASH.



Source: Prepared by the author.

The PBIAS % results show that the normalization procedures incurred in a positive bias, even for the calibration period, demonstrating an intrinsic bias generated by the normalization procedures. For the simulation period, the networks obtained using non-normalized streamflow presented a negative median, close to zero and between -5 to 5% for the lower and upper quartile. The bias presented for the simulation period for the remaining networks showed a reduction in the median value in comparison to the calibration period, but an extension of the upper quartile, especially for the results derived from a normalization through SRI.

The coefficient of determination shows similar median values and boxplot configurations of the non-normalized streamflow and SRI networks. The networks with the use of log10 presented the lower values for first quartile of the calibration period but higher values for the simulation period. The results for the VE shows similar medians for both periods, the log10 presented the worst performance for the calibration period whereas the SRI presented slight better results for the simulation period.

Figure 8b presents the results for the NRMSE % of each month streamflow series. The analysis of this index alone derives from the fact that the networks have the flexibility to not impose a parent to a node, which results in the direct use of the mean value of the month. In such cases, metrics as NASH and R2 cannot be used. It is noticeable for the calibration period lower median values of the original streamflow and similar boxplots for the three cases analyzed. This fact demonstrates that the general performance of the normalized networks did not suffer from a worsening due to the normalization of a specific month. For the simulation period the original streamflow presented the highest values for the median and upper quartile in several months. This result showed that normalization procedures produced smaller errors for the isolated months when the calibrated networks are used for simulation, however, as shown, it did not produce an increase in the overall performance of the monthly continuous series. In general, the best performance occurred for the months from May to August.

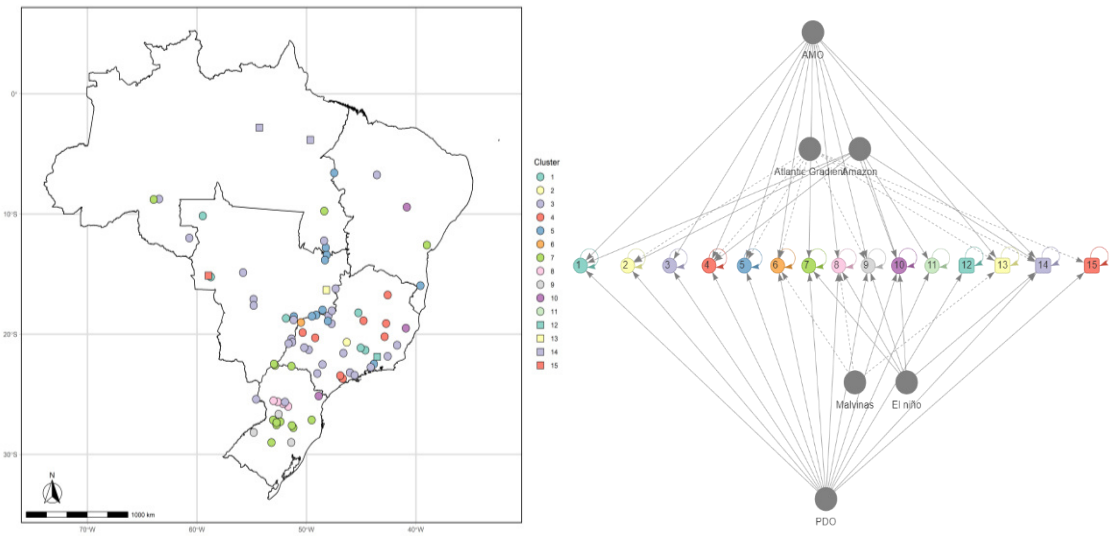
The spatial NASH results for the networks without normalization is shown in Figure 8c, no significant spatial differences were noticed from the normalized networks (not shown). The proposed network methodology did not present a good performance for the representation streamflow of mostly south stations and for some stations located in the east region of the country, close to the division between the Southeast and Northeast region. A deterioration from the performance obtained for the calibration period can be seen in some stations close to the regions border mentioned and for a few stations in the Southeast region.

For the remaining regions it is noticeable a good performance, with stations presenting values up to 0.9 and a small deterioration of the performance obtained for the calibration period.

5.4.2 Network Structures Coherency analysis

Figure 9 presents the network typologies found for Brazil. The typologies are clustered by similarity and refer to the monthly series. It requires only one of the months to have an index as parents to produce an arc between the station and the predictor index in the network typologies. It is noticeable that all stations used an autoregressive component to represent the streamflow series and the PDO phases, except the typology 13.

Figure 9 – Spatial dispersion of the network typologies found clustered by similarity (left) and the network typologies (right). This result refers to the networks obtained using the original non-normalized streamflow and with lead time of one month. The arrow pointing in and out from the same node refers to an autoregressive component. The networks typologies refer to the monthly series, thus, an arc between the index and the station is produced if at least one month has the index as a parent.



Source: Prepared by the author.

The spatial dispersion of the network typologies resembles the NASH results. All stations from typology 8 and 9 presented a negative or close to zero value of NASH. Similar pattern is observed for most of the stations in typology 7, which represented the second largest number of stations (14), showing great similarities for the stations where the networks

performed poorly. Typology 3 represents the largest number of stations, this typology used only the autoregressive component estimated for each phase of AMO and/or PDO, but presented a good performance for the NASH coefficient (> 0.7), with the exception of a few stations.

The typology five represented the third largest number of stations (11) and was located mostly in the center of the country, with good results for both calibration and simulation period and using as exogeneous predictor only the Atlantic gradient, besides the phases of AMO and PDO. Typology four represented the fourth largest number of stations (9) with a joint influence of the indices Grad and Amazon in addition to the low-frequency phases. These stations showed good results in general for the NASH coefficient, but this typology encompassed few stations that showed a decline in the performance for the simulation period. Typology one represented six stations and used as exogenous predictor only the Amazon index, its stations were located along the Center-west and Southeast region. These typologies presented spatial coherence regarding the predictors.

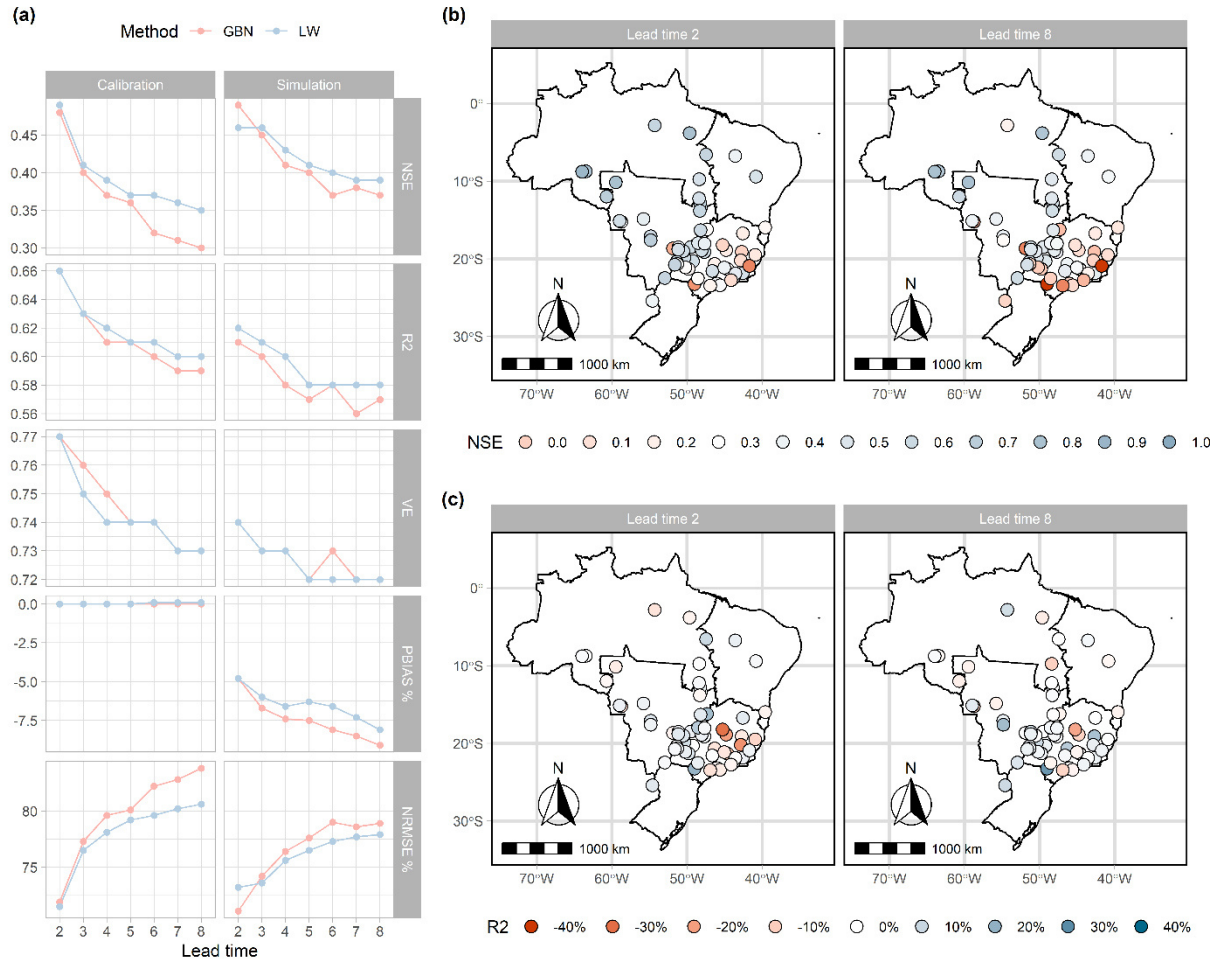
The remaining typologies refers to a low number of stations, where typologies 2, 6, 11, 12, 13 and 15 represent only one station each, typology14 two, and typologies 9 and 10 three stations. It is important to emphasize that the Malvinas index presented a discrete influence, being used as exogeneous predictor to only three typologies. The supplementary files present interactive plots of Figure 2 that eases the visualization of the network typologies.

5.4.3 Forecast performance analysis for longer lead times

Figure 10a presents the median of the results of the good-of-fit measures for the two streamflow forecasting methodologies used: GBN and LW. The median was calculated after the removal of the stations that presented NASH values lower or equal to zero in the previous section.

In general, the forecast through LW showed a slight superiority, especially for lead times higher than three months. For the NASH, NRMSE % and R^2, the superiority of this methodology is clear for both periods, but no significant difference can be distinguished from the VE results. The two methodologies did not produce bias for the calibration period, however, a sub estimation bias is shown for the simulation period, especially for the GBN method. These results refer to the networks obtained with the use of the original, non-normalized streamflow.

Figure 10 – (a) Median of the good-of-fit results of the stations, for the calibration and simulation period, through direct use of the parameters of the Gaussian Bayesian Networks (GBN), obtained independently for each lead time, and through Bayesian inference via likelihood weighting simulations (LW), based in the networks obtained with lead time of one month. The results refer to the networks obtained for the original non-normalized streamflow and after the removal of the stations that performed poorly for the lead time of one month. (b) Spatial dispersion of the NASH results (LW) for the lead times of two and eight months. Stations shown in dark red presented negative values of NASH. (c) Spatial dispersion of the percentage difference in the R2 coefficient between the two methods (GBN and LW), calculated in relation to the GBN results.



Source: Prepared by the author.

Figure 10b presents the spatial dispersion of the NASH results (LW) for the same group of stations used for the median calculation, for the lead times of two and eight months. It can be seen that the median value is a result of two distinct spatial patterns, the stations localized in the Southeast region showed the worst performance and a decrease of the performance with the increase of the lead time, whereas the stations in the Northeast, Center-west and North regions presented the best performance and a slower worsening with the increase of the lead time.

The worst performance for this region can be associated with the reduced use of exogeneous predictors since many stations in this region presented the network typology 3. This behavior is in contrast with the good results obtained for the Tucuruí station, localized in the North region, which presented a value of NASH of 0.87 for the lead time of one month, 0.8 for the lead time of two months and 0.76 when forecasting with a lead time of eight months. The good results obtained for this station derives from the use of exogeneous predictors with a significant time lag for the network obtained with lead time of one month and the existence of a strong autocorrelation component. Also, this station presented a more pronounced periodic behavior, thus, the streamflow forecast for this station greatly benefited from the use of month nodes. This station presented as parent of the January streamflow node the amazon index with a 6 month lag (July of the previous year) and used the Grad index with a lag higher than 12 months as a parent of the May streamflow node (February of the previous year). This particular result demonstrates how the use of a variable lag can result in the improvement of the forecasts, since the influence of the indices occurred with different time lags. Plot of Tucuruí results can be found in the appendix A.

The percentual difference between the two methodologies, GBN and LW, is shown in Figure 10c. It is noticeable that the GBN method presented better results for some stations, especially in the Southeast region, although, this region concentrated most stations with relative bad performance (Figure 8). In contrast, the superiority of LW method is shown for diverse stations across the country. The biggest differences occurred with lead times up to five months, where higher lead times shows a stabilization tendency and a higher convergence to similar results.

This result demonstrates that, for the case study, LW method presented a tendency to sustain the good performance of the networks obtained with lead time of one month, but produced worse results when compared to a network specifically obtained for each lead time for stations that did not show a great performance of the base networks.

In comparison to the normalized streamflow networks, no significant overall differences were observed for the GBN method regarding the two normalization procedures. However, through LW method it is noticeable from Figure 4a that the median of the results shows an overall worst performance of the networks obtained from the use of normalized streamflow (LW-PH-SRI and LW-PH-SRI), especially for the calibration period. Comparing the results of the two normalization procedures, the use of SRI produced better results than the

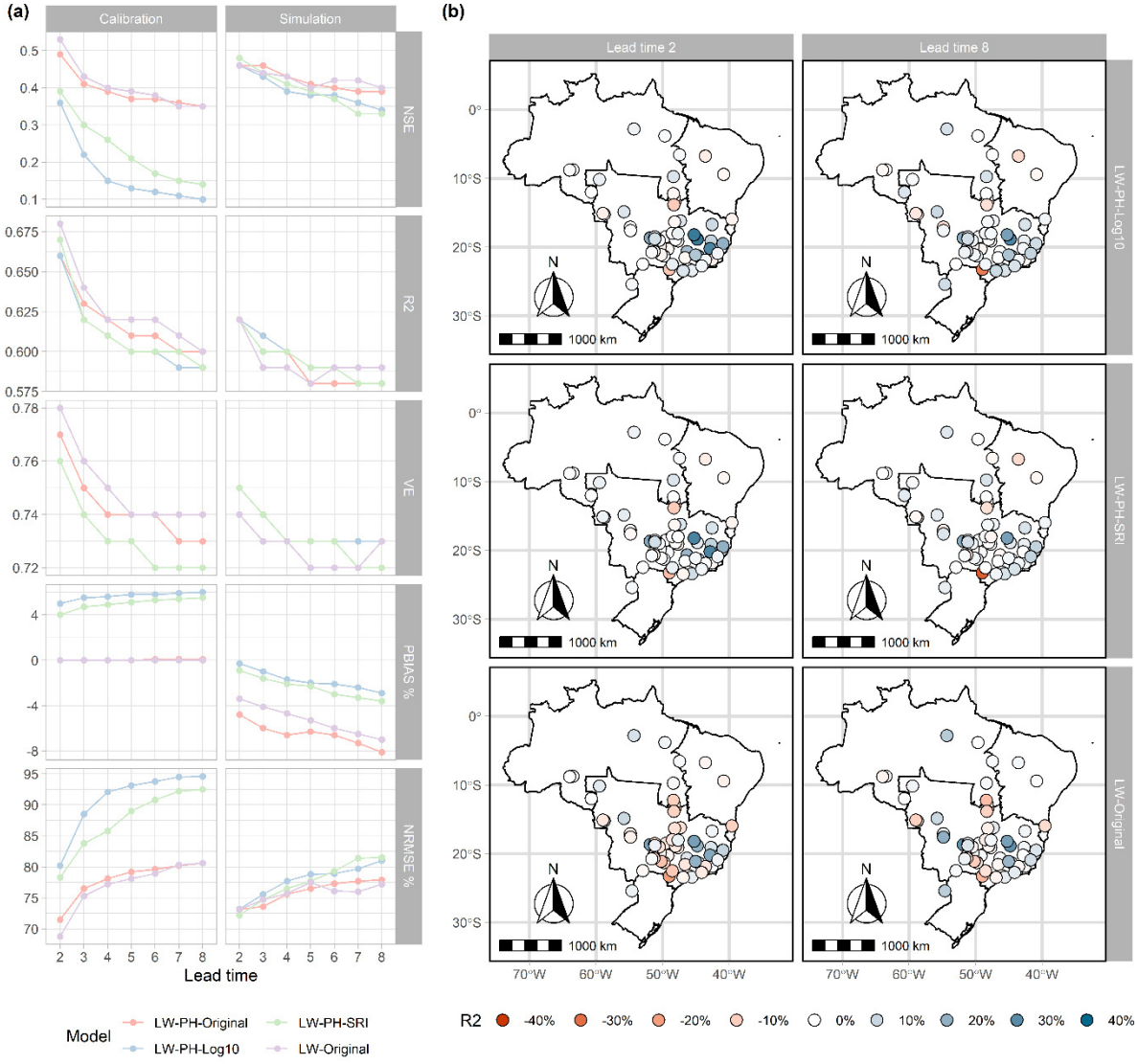
use of log 10 for this period. For the simulation period, the normalized results are close to the non-normalized (LW-PH-Original).

Analyzing the spatial dispersion of the percentage difference of the R2 between the normalized networks and the non-normalized (Figure 11b) it is noticeable that for some stations the use of a normalization procedure resulted in an increase in the overall performance of the forecast. With the increase of the lead time, however, the difference between the networks fade, converging for similar results (Figure 11b lead time of eight months).

According to the median of the results, a slight overall superiority was obtained for the networks without the use of the low-frequency phases (LW-Original). However, in a similar way to the normalized networks, the use of the low-frequency phases as a discrete node resulted in the performance increase of some stations, mainly located in the central region of the country and close to the borders between the Center-west and Southeast region.

Table 3 presents the best results via LW among the network structure analyzed, for four key stations that have been the subject of recent studies. The good results obtained for Tucuruí (LW-PH-Original) were further improved with the removal of the discrete nodes. This improvement can be explained by changes in the network structure with an increase in the use of both endogenous and exogeneous variables as parents (especially endogenous), where the predictors mentioned early remained unchanged. Structure changes were also observed for the other three stations, sometimes in in opposite directions: Itaipu presented as parent only endogenous variables for the LW-PH-Original case, but the structure that provided the best results (LW-Original) had also exogeneous variables as parents, whereas Furnas had only endogenous variables as parents for the LW-PH-SRI in contrary to the LW-PH-Original that had both. Sobradinho had changes in just the endogenous variables with an increase in the parents for the normalized cases, with both SRI and Log10 normalization procedures providing equal results. Table 3 results is further analyzed in the discussion section, also the supplementary files present the plot of the series for all lead time for the simulation period of these stations.

Figure 11 – Results of the different networks through LW method and for the lead times from two to eight months. The PH suffix indicates the use of the AMO and PDO phases as discrete nodes, the final suffix corresponds to the networks obtained with the use of a normalization procedure for the streamflow series (Log10, SRI) or the direct use of the original streamflow (Original). (a) Presents the median of the good-of-fit results of the different models for the calibration and simulation period. (b) Presents the percentage difference of the R2 coefficient between the different networks and the LW-PH-Original networks for the simulation period.



Source: Prepared by the author.

Table 3 – NASH and R² results (via LW) for four key stations for the lead time of one to eight months. The results correspond to the best obtained among the different network configurations tested in this work.

Station	Furnas		Sobradinho		Itaipu		Tucuruí	
Methodology	LW-PH-SRI		LW-PH-(SRI or Log10)		LW-Original		LW-Original	
	NSE	R2	NSE	R2	NSE	R2	NSE	R2
Lead time 1	0,62	0,71	0,67	0,70	0,62	0,70	0,89	0,90
Lead time 2	0,51	0,65	0,55	0,61	0,55	0,62	0,82	0,84
Lead time 3	0,46	0,63	0,51	0,6	0,5	0,58	0,81	0,83
Lead time 4	0,43	0,62	0,47	0,59	0,47	0,58	0,80	0,82
Lead time 5	0,41	0,61	0,43	0,58	0,47	0,58	0,79	0,82
Lead time 6	0,39	0,61	0,4	0,58	0,45	0,57	0,78	0,81
Lead time 7	0,37	0,61	0,37	0,58	0,42	0,55	0,77	0,81
Lead time 8	0,36	0,61	0,35	0,58	0,4	0,55	0,76	0,8

Source: Prepared by the author.

5.5 Discussion

In comparison to the local studies found in the literature, the proposed methodology resulted in an increase of the forecasting performance. Comparing the results with Lima & Lall (2010), for the key stations, the proposed methodology outperformed the R² results for all the lead times. Noteworthy is the significant improvement in the performance for Tucuruí station, whereas Lima and Lall results start with 0.4 for the lead time of one month and has a fast decrease for higher lead times. For Furnas station, Oliveira and Lima (2016) using an PARX model obtained a NASH value of 0.35 and 0.30 for the forecasts with a lead times of one and two months. The present work obtained values of 0.62 and 0.51, respectively, where the result for a lead time of eight months matched the results for two month of lead time. However, it is important to emphasize that the aforementioned studies used different set of indices as exogeneous predictors and different periods of calibration and simulation.

The superiority of the results can be explained by the use of a specific lag according to the climate indices and for each monthly streamflow node, resulting, for instance, in the use of the Amazon index with a lag of six months for the forecasting of the streamflow of January observed for Tucuruí, but also for Sobradinho. This latter station also used Grad index with one month lag to forecast the streamflow of November. The use of variables with a significant lag for the network with lead time of one month combined with the use of Bayesian inference resulted in a slow decrease of the performance with the increase of the lead times.

The use of Bayesian networks to predict streamflow showed a great flexibility, allowing the use of complex networks with lagged exogeneous predictors for some stations and simple networks, strictly endogenous, that produced good results for a significant number of stations. Also allowing the direct use of the mean value for months where no parent was found for the node. Therefore, this methodology encompassed the qualities observed in PAR and PARX models, besides the flexibility to predict streamflow through Bayesian inference.

The strictly endogenous networks obtained agree with the results from Alexandre (2012) which concluded that for some stations in the study area the PAR model presented a good performance. The spatial dispersion of some of these networks resembles the SACZ, with a straight-line layout linking the North region to the Southeast region. Also resembles the transition zone mentioned by Rocha and Souza Filho (2020) which splits the country in two regions with opposite change point pattern: streamflow increase for the southern stations and decrease for upper Southeast and Northeast stations. The absence of exogeneous predictors for these stations can be derived by two hypotheses: 1) the indices used alone cannot represent the complex behavior of this convergence zone, even with flexibility of the networks that also allows a co-influence between the indices; 2) the impact of this indices in the streamflow could not be verified due to the absence of some variable. A clear occult variable, that was not used in the network construction, was the precipitation series, which can present a complex spatial pattern of occurrence inside significant large basins (the case of mostly of the stations used) that cannot be observed by the direct analysis of the streamflow series. Other important variable is the wind patterns that links the ocean TSM influence on precipitation pattern, however, wind data for the long period analyzed in this work was not found.

The flexibility observed for the use of continuous variable was not observed for discrete variables. The networks showed a tendency to use the phases of AMO and PDO, even for stations where the use did not result in the increase of streamflow representation. Analyzing the spatial pattern of the typologies obtained in this study, it was noticed that several stations which benefit from the use of these phases did not present a change point in the results of Rocha and Souza Filho (2020), in opposite to the initial hypothesis that the use of these phases could be an alternative to the use of a more complex methodology that explicitly considers the heterogeneity of these streamflow series. This conclusion was also corroborated by the plot of the Sobradinho station results, which can be found in the appendix A.

The two normalization procedures adopted in this work did not result in a generalized improvement of the network performance for streamflow forecasting, however, it

was observed that some stations presented an increase in the performance with both normalization process. The use of SRI resulted in a slight superiority in comparison to the use of a base ten logarithm, especially for the forecast of streamflow through LW for the calibration period.

The results obtained through LW shows that a one-month lead time network can be used to forecast streamflow with several months in advance, even for the stations that did not used exogeneous predictors. However, this method presented good results when the base network was able to successfully represent the streamflow, where the use of a specific network for each lead time is superior when a good performance is not verified for the base network.

Thus, the results show that each station needed to be analyzed regarding the use of the AMO and PDO phases, the normalization procedures, and the forecasting method (BGN or LW). The present case study shows that the networks presented best results for the North, Northeast and Center-west regions of Brazil. Conversely, the South and Southeast regions presented the worst results, which means that the applicability of Bayesian networks for those regions would require further investigation. It should be noted, however, that the Bayesian networks outperformed the results obtained in other studies found in the literature for the Itaipu station localized in the South region.

The networks found in this work have a wide margin to be improved, with the addition of other climate indices and other pertinent variables such as the precipitation in the hydrological basin of each station. The inclusion of occult variables in the construction of similar networks will allow a more comprehensive analysis of possible spurious relations and the detection of influences that were not found with the methodology used.

The networks were constructed analyzing general aspects of the study area. We find that networks constructed specifically for each station with a more profound analysis of the predictors for coherence with regional physical phenomenon can produce better results. These refined networks could be also used to better comprehend the impacts in the streamflow in short and long term of phenomena such as El Niño and the low-frequency phases of AMO and PDO. However, differences in the set of parents imposed by different network configurations shows that this aspect needs to be thoughtfully addressed before using a network for causal inference and phenomena impact analysis.

The use of the Bayesian networks also allows for an analysis of indirect effects due to the interrelationship between the TSM indices, since the network between the indices is part of the network obtained. This aspect was not analyzed directly in this work.

Another aspect of the use of the one-month lead time network to forecast streamflow with several months in advance through Bayesian inference is possibly a better estimation of the uncertainty involved, since it considers the multiple time-steps involved in the streamflow estimation for the desired lead time. This uncertainty analysis allows for the estimation of a streamflow interval instead of only the expected value and can potentially result in the improvement of intervals in comparison to standard methodologies based on linear model errors.

5.6 Conclusions

In this work we used Bayesian networks to analyze the potential of different network configurations to forecast monthly streamflow with a lead time of one to eight months. We used networks with monthly nodes (one for each month) and lagged variables of the previous year, analyzing the impact in the performance of two normalization procedures (using base ten logarithm and SRI) and of the inclusion of discrete nodes in the network representing the phases of low-frequency SST oscillations of the Atlantic Ocean (AMO) and Pacific Ocean (PDO). Moreover, we analyzed the forecast through two different methodologies: (1) obtaining a specific Bayesian Gaussian network for each lead time and forecasting the streamflow through the direct use of the network parameters; and (2) using the network constructed with a lead time of one month to predict streamflow with longer lead times through Bayesian inference (Likelihood Weighting simulations).

From the results obtained a few methodologic and local conclusions arise:

- The Bayesian networks, in general, presented a good performance to the streamflow forecast and a great flexibility, resulting in the use of complex networks with lagged exogeneous variables as parents but also simpler, strictly endogenous networks with good forecasting results. Despite the high degree of freedom provided by variable lags between the streamflow node and the parents and also the great number of nodes, the networks did not

show overfitting patterns, i.e. a significant decrease in performance for the simulation period.

- In some stations the normalization of the streamflow resulted in the improvement of the forecast for both methods used (SRI and Log10). Given that normalization of the streamflow did not improve for all stations, this requirement needs to be evaluated for each case studied. The use of normalization also modified the network structure obtained by the network structure learning algorithm and, in general, imposed a bias to the forecasted streamflow.

- The incorporation of the phases of the indices AMO and PDO as discrete variables resulted in a significant improvement in the performance for some stations. However, the networks obtained with the use of the original non-normalized streamflow presented a higher tendency of using the discrete variables, even in networks where its use did not result in the improvement of the predicted streamflow.

- A network obtained with a lead time of one month can be used to predict streamflow of higher lead times through Bayesian inference. This method also showed better results than the use of a network obtained specifically for the lead time desired in some of the stations analyzed. However, this performance enhancement only occurred for the stations that presented a good performance for the initial network.

- The different set of parents obtained, for the different network configurations tested, show that this aspect needs to be thoughtfully addressed and the network structure obtained profound analyzed before using it for causal inference and for the analysis of phenomena impact in the streamflow.

- The Bayesian networks presented good results for the Northeast, North and Center-west regions of Brazil. The southeast region presented relative inferior results and the methodology did not present good results for most stations in the South region. The exception in the South were the Itaipu results, which outperformed the results found in the literature.

- The use of the phases of AMO and PDO did not result in the improvement of the performance for stations that presented a change point in the mean value of the series in similar periods of the phase-change of these indices.

6 STREAMFLOW MODELLING THROUGH DYNAMIC NON-HOMEGENEOUS BAYESIAN NETWORKS

6.1 Abstract

The non-stationarity impact in the development of Dynamic Bayesian Networks to streamflow modelling was evaluated using three methodologies, also discussing its possible impacts in the streamflow forecasting. The AutoRegressive Time VArying (ARTIVA) framework was used to detect and model streamflow series accounting for existent non-stationarity, evaluating its change point detection skill and modelling performance. Hill Climbing (HC) and PCMCI algorithms were used to discover a more complex lagged relationships between the variables, evaluating the non-stationarity impact in the network structures obtained before and after the change point existent in the streamflow series. ARTIVA presented a change point detection skill of 15%. This result was affected by the single change point constrain imposed and the possible existence of change points in the time series of the climate indices used as exogeneous predictors. The simpler network obtained by ARTIVA produced similar results to HC and PCMCI for streamflow modelling with one month lead time. The complex lagged relationship obtained between the streamflow and the indices indicates that for streamflow forecasting with higher lead times the use of climate indices may have a larger impact. The networks obtained before and after the change point presented significant differences, not only in the relationship of the variables but also in its autocorrelation function. This aspect needs to be included in the development of the Dynamic Non-Homogeneous Bayesian Networks since it may influence in the streamflow forecasting performance through inference methodologies such as Likelihood Weighting Simulations.

6.2 Introduction

The non-stationarity is a common feature of hydrological time series and emerges as a challenge for the development of models and water policies to successfully incorporate its complex implications. One natural reason of this characteristic derives from the hydrological cycle and the complex sea-atmosphere interaction, where the variability of the sea surface

temperature (SST) is considered by some studies the driver behind the non-stationarity observed (DUTTA; MAITY, 2020; ROCHA; DE SOUZA FILHO, 2020).

An example of implications for water policies is the requirement of dynamic water granting policies aiming to diminish the opportunity cost imposed from using a single policy developed assuming stationarity. The non-stationarity gains a particular relevance for the streamflow forecasting models, since this feature can also be a result of anthropogenic factors such as the constructions of dams, water withdraw along the course of the river and land-use changes. Therefore, this issue needs to be correctly addressed to prevent the model to incur in structural errors.

To cope with this problem studies have analyzed different methodologies, for instance, Luo et al. (2012) used different schemes to calibrate a hydrological model, such as using different lengths of records (only the years with similar climate conditions and individual monthly records) concluding that the best schematic is dependent of the catchment. More recently, Dutta & Maity (2020) used the concept of temporal networks to model streamflow with one-month lead time through Bayesian networks, for this method the non-stationarity was assessed by using a moving window of 30 years to obtain the network parameters and updating the structure after a fixed time interval, optimized for each season. The results obtained through the temporal networks-based model outperformed the other methodologies used and showed that Bayesian networks can perform well for streamflow modelling.

Bayesian networks is currently a promising tool in hydroclimate studies, with recent articles exploring its use to discover causality between variables through the analysis of the graph structure obtained through structure learning algorithms (EBERT-UPHOFF; DENG, 2012, 2015; RUNGE et al., 2019a), although its usefulness for obtaining relationship within climate indices was already mentioned by Glymour, Chu and Danks (2005). Whereas for this field of study the use of Bayesian Networks is not fully explored, this methodology is more stablished for the Systems Biology field, with the development of mechanisms to incorporate non-stationarity with time-varying networks (LÈBRE et al., 2010).

Static Bayesian Networks represent each variable as a single node in the network whereas Dynamic Bayesian Networks unfolds the series in time, using several nodes to represent the lagged information for the maximum lag desired (NAGARAJAN; SCUTARI; LÈBRE, 2013). Bayesian Networks, as mostly methodologies, has the stationarity of the data as a core premise. A solution to account for non-stationarity is the use of Non-Homogeneous

Bayesian Networks, which obtain a set of networks, one for each of the homogeneous phases existent in the time series.

The ARTIVA (Auto Regressive TIme VArying models) statistical modeling framework proposed by Lèbre et al. (2010) aims to both detect when a change occurs within a time series and discover the network structure for each shift detected. This methodology, if proven adequate for streamflow related studies, can aid in the answer to currently state-of-the-art questions in hydrologic field.

Change point detection in streamflow series is still an open question with recent articles comparing diverse methodologies to verify its reliability (RYBERG; HODGKINS; DUDLEY, 2019) and stablish a framework that discover only the true change points, disregarding eventual false positives (ROCHA; DE SOUZA FILHO, 2020).

This paper analyzed the performance of the ARTIVA methodology for both change point detection and structure learning of data-driven Dynamic Non-Homogeneous Bayesian Networks constructed for streamflow modelling. The results were compared with previous change point studies and with different structure learning methodologies: a standard score-based algorithm (Hill Climbing) and the recent PCMCI algorithm framework proposed by Runge et al. (2019b) based in conditional independence tests. These latter methodologies allow the discovery of more complex lagged networks and were also used to investigate changes in the network structures before and after the change-points and its implications for streamflow modelling and forecast.

6.3 Data and Methodology

6.3.1 Study Area Background and exogeneous predictors

In this work we used monthly naturalized streamflow data for 26 stations located in Brazil country, for the period from 1931 to 2016. The presence of change point in the mean value of these streamflow series was observed by Rocha & Souza Filho (2020), qualifying these series as an interesting application for the ARTIVA framework. The naturalization procedure

is detailed in a technical report of the maintainer, the Electric System National Operator (ONS, 2017).[4]

Based in a previous study, this work included the climate indices Niño 3.4; Grad (climate index resultant of the difference between the Tropical Northern and Southern Atlantic Indices), Malvinas (CATALDI et al., 2010) and Amazon as potential exogeneous predictors for the streamflow series. Malvinas index was calculated using data from the NOAA (National Oceanic and Atmospheric Administration) ERSST V4 (Extended Reconstructed Sea Surface Temperature Version 4) averaging the sea surface temperature (SST) of the area between 43°S to 33°S and 63°W to 48°W. Similarly, it was calculated the Amazon index averaging the area between 10°N-10°S; 55ºW-24ºW. The remaining indices were also obtained from the NOAA database, extending the data of the climate indices used to calculate the derived Grad index with data from the NOAA ERSST V4, since the time series of those indices starts in 1948.

The use of these indices intended to represent its associated large-scale phenomena: Niño 3.4 for the El Niño Southern Oscillation (ENSO), Grad for the meridional gradient in the off-equatorial tropical Atlantic, Malvinas for the Brazil-Malvinas Confluence and Amazon as control area to account to low-frequency impact of Amazon river runoff in Atlantic Ocean. The influence of these phenomena in Brazilian climatology was already reported in the literature (CATALDI et al., 2010; JAHFER et al., 2020; JAHFER; VINAYACHANDRAN; NANJUNDIAH, 2017; KAYANO et al., 2016; NOBRE; SHUKLA, 1996).

6.3.2 Data preprocessing

To assess the normality assumption of the ARTIVA framework and Hill-Climbing algorithm the streamflow series were normalized through the Tukey Ladder of Powers transformation. This step was performed through the function *transformTukey* from the R package *rcompanion.* This function performs the transformation by interactively changing an lambda (λ) parameter until is reached a power distribution that most closely fits the W statistic of the Shapiro-Wilk normality test (SHAPIRO; WILK, 1965), following the set of equations:

[4] Further information about the stations is presented in the Appendix A and D.

$$\begin{cases} \lambda > 0 \rightarrow Q' = Q^{\lambda} \\ \lambda = 0 \rightarrow Q' = \log(Q) \\ \lambda < 0 \rightarrow Q' = -1 \times Q^{\lambda} \end{cases} \quad \text{(Eq. 5)}$$

where Q stands for the streamflow series and Q' the normalized series.

The choice of this normalization procedure was based on preliminary test where this framework presented better overall results than the Box-cox transformation and a direct use of a base 10 logarithm.

After the normalization procedure the seasonality was removed from the streamflow time series. This step was intended to remove the non-stationarity derived from a strong seasonal component, also referred as cyclostationarity. The seasonal component was removed through the function *decompose* from the R package *stats* which performs a classical seasonal decomposition with the use of moving averages, in this work it was used an additive model:

$$\mathrm{Q} = T(t) + S(t) + e(t) \quad \text{(Eq. 6)}$$

where T(t), S(t) and e(t) stands, respectively, for the trend, seasonal and random component of the streamflow time series.

6.3.3 ARTIVA framework and change point detection

The ARTIVA framework aims to simultaneously discover the (1) occurrence and location of change points in the time series and (2) the graph structure of each period delimited by the change points. In order to accomplish this task, each node of the network is modelled in a regression framework (LÈBRE et al., 2010):

$$X^i(t) = \sum_{j \in Pa_h^i} a_h^{ij} X^j(t-1) + b_h^i + e^i(t), \text{ with } e^i(t) \sim \mathcal{N}(0, (\sigma_h^i)^2) \quad \text{(Eq. 7)}$$

where $X^i(t)$ is the value of the node *i* at time *t*, Pa_h^i is the set of parents of the node *i* for the phase h, $X^j(t-1)$ is the value of the node *j* at time $t-1$, a_h^{ij} is the coefficient that condition the influence of the node *j* in the node *i* for the phase h, b_h^i is a constant for the node *i* and phase h and $\mathcal{N}(0, (\sigma_h^i)^2)$ is a multivariate normal distribution centered at zero with diagonal covariance matrix $(\sigma_h^i)^2$.

The phases *h* refers to the homogeneous periods delimited by the change points, which result in phase-specific parameters and network topologies. From this framework the addition and removal of a change point changes the system's state-space dimension, this characteristic imposes a high degree of freedom to the model. To surpass this problem Lèbre et al. (2010) used a Reversible Jump Markov Chain Monte Carlo procedure (RJ-MCMC) to infer the posterior distribution from the observed data over all of the system's parameters. This procedure allows the generation of an ergodic Markov chain since it uses a reversible Markov chain sampler that is able to jump between parameter sub-spaces of different dimensions. The posterior distribution is given by the equilibrium distribution, where the equilibrium is assumed according to the R statistic of Gelman and Rubin (1992) (herein mentioned as RGELMAN) for the convergence of multiple makov chains or after an extremely high number of iterations. In this study ARTIVA framework was used through its homonymous R package with a maximum of 200,000 interations of the sampler evaluating the convergence of the chains with a RGELMAN value of 1.1. More information about the prior distributions assumed for the RJ-MCMC procedure and the methodology can be found in the original article.

The ARTIVA framework uses a fixed lag between the multiple nodes of the network. This characteristic comes from its conceptualization, that models the influence of the nodes by a vector-autoregressive process. The use of a fixed lag does not allow to discover complex variable lags between the different climate indices and the streamflow series, which can possibly occur since the indices represent different phenomena. In this work a first order autoregressive was used, as stated in (Eq. 7).

To compare the change point results to a previous local study performed by Rocha and Souza Filho (2020), the number of change points allowed was set to one. The mentioned article assessed the reliability of the change points found by analyzing the convergence between the multiple methodologies used, which limited the number of change points that could be considered reliable since one of the methods was a single change point discovery test.

6.3.4 Network Structure learning through PCMCI and HC

The PCMCI is a recent method developed by Runge et al. (2019b) that aims to obtain a reliable network structure that would allow to infer causality between the variables

through the analysis of the resulting graph. The proposed methodology is a framework based on the classical PC algorithm (GLYMOUR; SPIRTES, 1991) that relies on conditional independence tests to obtain the network structure. Briefly, this method is a two-step process: (1) first the PC algorithm is applied with a liberal significance level to obtain the possible relevant conditions, i.e. the set of potential parents of the nodes, possibly including some false positives; (2) for this set of parents is applied a second test called Momentary Conditional Independence (MCI) that aims to remove eventual spurious connections that may occur in highly interdependent time series, selecting only the strongest parents. Both stages rely on conditional independence tests. PCMCI can be combined with a false discovery rate control, to adjust the resulting p-values of the MCI stage for the complete graph.

The strength of this framework relies in the reduction of the effect of dimensionality that affects the detection power of the conditional independence tests (RUNGE et al., 2019b). The MCI test is applied only to verify connections between relevant variables, since the irrelevant conditions are removed in the first step that applied PC algorithm for all possible connections of the network. The curse of dimensionality is also a common problem faced by the PCMCI since its focus on time-lagged dependencies, unfolding each time series across several nodes through time and allowing the discovery of complex lagged relationship. This process results in large networks even for few variables. In this work, the algorithm was applied to search for time-lagged dependencies from one to twelve months.

The causal interpretation of the network obtained through the PCMCI leans in the Causal Sufficiency assumption of the method. This assumption imposes that all common drivers are among the observed variables included in the network. For this study, there are clear hidden variables that links the influence of the large-scale climate phenomena represented by the indices and its effect on the streamflow. One example is the precipitation of the hydrologic basins that generate the local streamflow and wind patterns that links the sea surface temperature patterns to precipitation patterns. Runge et al. (2019a) states that for such cases the PCMCI method requires further studies to better understand its performance.

Another assumption of this method is the stationarity of the data. In this work, this method was used to obtain a network before and after the change point present in the time series, i.e. for each homogenous phase, accounting for the initial assumption. The algorithm was used through the python code referred in the original article (RUNGE et al., 2019b), using the false discovery rate control implemented, the Hochberg-Benjamini approach (BENJAMINI; YOSEF, 1995).

Network Structure can also be learnt from data through the minimization of score functions. Both the use of conditional independence tests and score functions can yield good results and a recent study state that these methods outperforms hybrid algorithms that combine both score-based functions and conditional independence tests (SCUTARI; GRAAFLAND; GUTIÉRREZ, 2019).

A classic method to obtain the network structure through score minimization is the Hill-Climbing (HC) algorithm. HC consist in the addition, removal or reversal of a single arc of the initial network structure (commonly an empty graph), computation of the score results and the repetition of these steps until the score value can no longer be improved (NAGARAJAN; SCUTARI; LÈBRE, 2013). This algorithm presented a good performance for structure learning from real world geospatial data, as stated in the study of Scutari; Graafland & Gutiérrez (2019).

The step-by-step process of the HC method can result in the stop of the algorithm in points of local minima. This problem can be diminished with the use of bootstrapping, running the algorithm several times, and averaging the networks obtained using a significance threshold (NAGARAJAN; SCUTARI; LÈBRE, 2013). In this method, for example, assuming a significance threshold of 85%, only the arcs that occurs in 85% of the networks found are considered significant and averaged to obtain the final network. The bootstrapping process produced an increment in the streamflow modelling performance of the networks in the tests made[5], with an increase of up to 10% in good-of-fit measures such as R^2 and NASH in comparison to networks obtained without the use of bootstrap.

The selection of the threshold is an important step and can be assessed with the use of statistically motivated algorithms to reduce the subjectivity of relying on ad-hoc values (NAGARAJAN; SCUTARI; LÈBRE, 2013). In this work, the networks were obtained via HC with the use of the R package bnlearn (SCUTARI, 2010) that implements the threshold selection algorithm proposed by Scutari and Nagarajan (2013). The score function used was the Bayesian Information Criteria (BIC). Analogously to the PCMCI, the HC algorithm was applied to obtain two network structures, before and after the change point. For both the PCMCI and HC algorithms the changepoints obtained by Rocha and Souza Filho (2020) were considered the correct change points, thus, the algorithms were applied for each homogeneous

[5] The preliminary tests of the previous article (chapter 5).

phase delimited by the changepoint adopted as reference, resulting in two network structures for each station.

To obtain networks through HC that can retrieve time-lagged dependencies, the time series of the variables were unfolded in time with the inclusion of one node for each of the lags used, from one to twelve months. To prevent the algorithm from finding unwanted arcs that violate physical and temporal meaning of the network a set of restrictions were imposed to certify the temporal and physical coherence. In this study it was considered a physical violation the existence of an arc from a streamflow node to a climate index node. This latter restriction was also imposed to the PCMCI algorithm. For the PCMCI the restrictions for temporal coherence are already embedded.

6.3.5 Streamflow modelling through Gaussian Bayesian Networks

The network structure obtained through the different algorithms was used to model streamflow through Gaussian Bayesian Networks. Bayesian networks represents each random variable $(X_1, X_2, \dots, X_N)$ as a node and the dependency between the nodes is determined by links. The Markov assumption of the Bayesian Networks impose that the state of a node is only dependent on its parents. This characteristic modulates the local distribution of each node and determine the global probability distribution of the network:

$$p(X_1, X_2, \dots, X_N|\alpha) = \prod_{j=1}^{N} p\left(X_j|\, pa_G(X_j), \alpha_j\right) \qquad \text{(Eq. 8)}$$

where $p(X_1, X_2, \dots, X_N|\alpha)$ is the global probability distribution of the network given the set of parameters α, $\left(X_j|\, pa_G(X_j)\right)$ is the local conditional probability associated to the node *j* and α_j is the parameter vector that modulates the parents influence in the conditional distribution of X_j. In this work, the parameters were obtained through maximum likelihood estimation. The adoption of Gaussian networks results in a set of linear equations to represent the parents influence.

6.3.6 Network performance and structure changes analysis

The streamflow modelling performance of the networks was evaluated using the NASH coefficient (NSE), R2, volumetric efficiency (VE), percentual bias (PBIAS %) and Normalized Root Mean Squared Error (NRMSE %). The performance was analyzed for each network structure obtained, before and after the streamflow change point, comparing the different methodologies.

The network structure obtained before and after the change point for each station was analyzed graphically and according to the number of similar and different arcs, using the first network as reference. Therefore, three comparison metrics were obtained:

- True Positives (TP): number of arcs in the first period that were also presented in the second period;
- False Positives (FP): number of arcs present in the second period and not present in the first period;
- False Negatives (FN): number of arcs present in the first period but not in the second period.

6.4 Results

6.4.1 ARTIVA change point detection

The ARTIVA algorithm converged for mostly stations around 51,000 iterations, in concordance to the default maximum iteration number proposed by the R function (50,000). However, few stations did not reach the convergence of the chains even with the maximum iterations (200,000). A further increase of the iterations number did not seem feasible due to the already long-running time of the algorithm for the maximum iteration cases, higher than 3 hours for each station (computer specifications can be found in Appendix B), combined with the absence of an inherent parallel computing alternative and problems found in the R package. Although, the total running time was diminished through parallel computing of each station. We also find that it was hard to evaluate if an increase in the iteration number would result in the convergence of the chains, since the evolution of the RGELMAN statistics is not an output of the ARTIVA standard function and critical bugs leads to abnormous stop of the algorithm when using the R function searching for the convergence of the chains. Therefore, the ARTIVA

results should be taken with caution. Further information regarding the bugs found along with runtime details can be found in the Appendix B.

From the 26 stations analyzed, ARTIVA framework with a single change point constrain detected changes in 7 stations. Two of these stations presented a change remarkably close to the end of the time series (~2012 and 2013) and were discarded. The change point location of four stations closely matches Rocha & Souza Filho (2020) results, where for the remaining station the detected change point presented a distance around 26 years (Table 4). Thus, the following sections focused in these five stations (99, 120, 134, 188, 281).

Table 4 – ARTIVA change points detected

Station Number	Change point year	Distance (years) from previous results*
99	1974	-1
120	1968	-2
134	1951	-1
188	1993	+1
281	1950	-26

*Rocha & Souza Filho (2020)

Source: Prepared by the author.

6.4.2 Streamflow modelling performance and predictors analysis

The three methodologies presented remarkably similar results of the goodness-of-fit measures used for four of the five stations analyzed (Table 5). The largest difference was observed for station 281 which the discovered change point presented a significant distance from previous studies. The change point provided by ARTIVA resulted in more than a 10% improvement of the goodness-of-fit (GOF) measures for the period after the change point in comparison to the two other methodologies, while also maintaining similar results for the period before the change point.

Table 5 – Goodness-of-fit results for the different methodologies and periods. The period with the largest differences is highlighted in bold.

ARTIVA										
Station	99		120		134		188		281	
Period	1973-	1974+	1968-	1969+	1950-	1951+	1992-	1993+	1949-	**1950+**
NSE	0,8	0,7	0,75	0,58	0,67	0,53	0,46	0,39	0,9	**0,87**
R2	0,81	0,7	0,75	0,59	0,68	0,54	0,47	0,42	0,91	**0,87**
VE	0,92	0,9	0,83	0,77	0,79	0,71	0,6	0,54	0,96	**0,93**

PBIAS %	-0,6	-1,2	-2,8	-4,4	-3,5	-6,9	-10,1	-19,3	-0,2	**-0,2**
NRMSE %	44,1	55,1	50	64,5	57,1	68,3	73,2	78	31,8	**36,5**
PCMCI										
Station	99		120		134		188		281	
Period	1974-	1975+	1969-	1970+	1951-	1952+	1991-	1992+	1975-	**1976+**
NSE	0,8	0,7	0,74	0,6	0,66	0,53	0,46	0,45	0,9	**0,78**
R2	0,8	0,7	0,75	0,61	0,66	0,54	0,47	0,49	0,9	**0,78**
VE	0,92	0,9	0,83	0,78	0,78	0,71	0,61	0,55	0,94	**0,93**
PBIAS %	-0,6	-1,1	-2,9	-4,3	-4,2	-6,9	-9,9	-19,3	-0,2	**-0,2**
NRMSE %	44,2	55,1	50,6	63,3	58,4	68,2	73,2	73,8	31,3	**47,2**
HC										
Station	99		120		134		188		281	
Period	1974-	1975+	1969-	1970+	1951-	1952+	1991-	1992+	1975-	**1976+**
NSE	0,8	0,7	0,75	0,6	0,66	0,54	0,47	0,45	0,9	**0,78**
R2	0,8	0,7	0,75	0,61	0,66	0,55	0,48	0,49	0,9	**0,78**
VE	0,92	0,9	0,84	0,78	0,78	0,71	0,62	0,55	0,94	**0,93**
PBIAS %	-0,6	-1,1	-3	-4,3	-4,2	-7,1	-9,8	-19,3	-0,2	**-0,2**
NRMSE %	44,2	55,1	50,3	63,3	58,4	67,8	72,6	73,8	31,3	**47,2**

Source: Prepared by the author.

The station 281 presented the best performance in the GOF measures for both periods, followed by station 99 that also presented a good performance before and after the change point. Stations 120 and 134 presented a good performance for the first period, yielding NSE and R2 values close or higher than 0,7, for the second period these stations presented values lower than 0,6 for both measures and a NRMSE higher than 60%. Station 188 presented the worst results by all measures used and a strong negative bias for the second period. It was noticeable that all stations presented a negative bias.

The similar performance of all three methodologies can be explained by the strong presence of autocorrelation which resulted in an almost strictly endogenous modelling of the streamflow (Table 6). The modelling performance seems to be mostly related to the lag-1 component, where the additional lags imposed by HC for the first period of the stations 99 and 188 resulted in a slight increase in the performance (around the second decimal), which explains the similar performance of the simpler model imposed by ARTIVA, although small differences can also be related to the distance between the change points. This observation is also corroborated by the PCMCI results of the station 281, in which the use of climate indices did not improve the modelling performance.

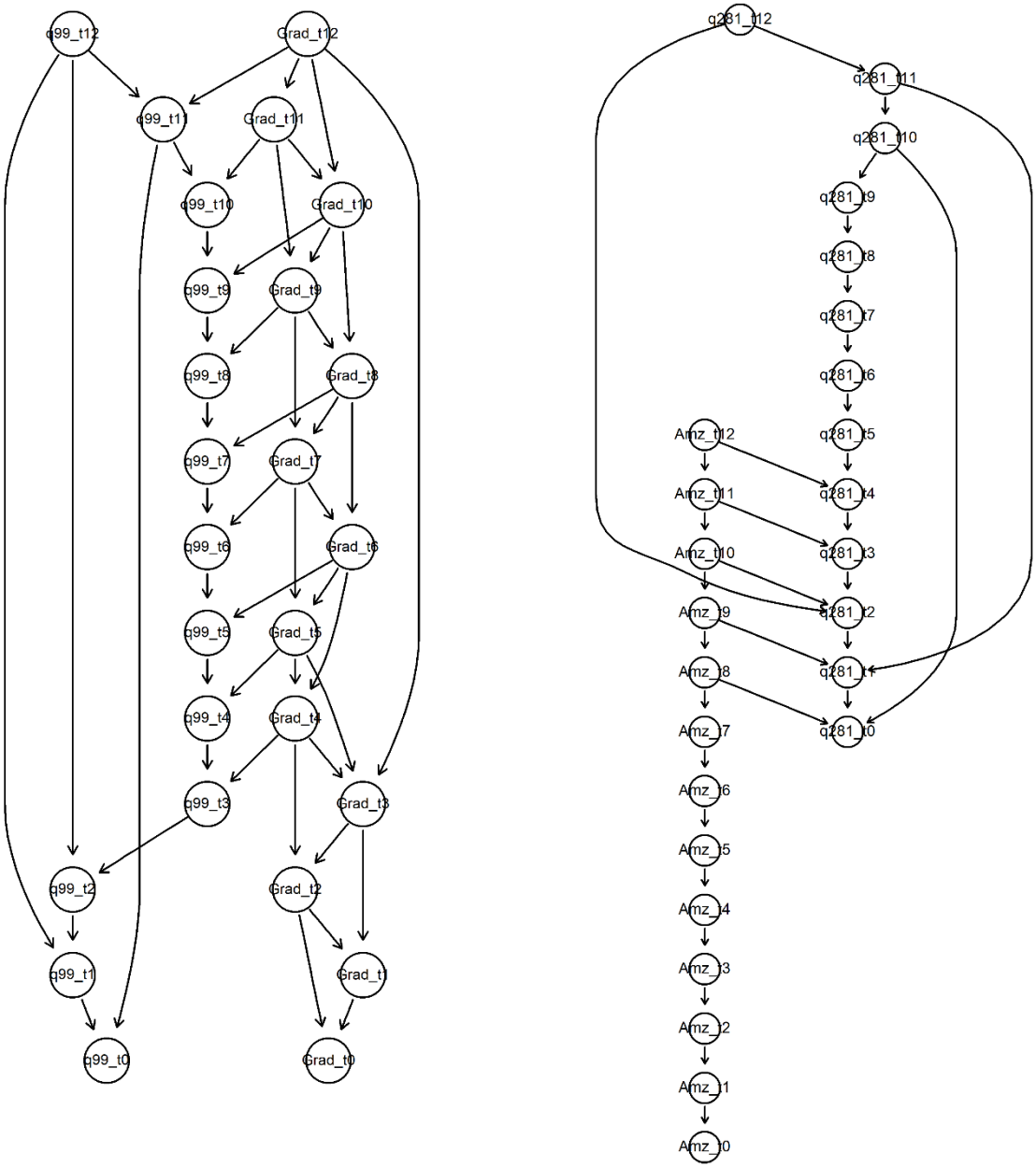
Table 6 – Parents and ancestors of the networks from different methodologies.

Station	99		120		134		188		281	
Period	1	2	1	2	1	2	1	2	1	2
Exogeneous Parents (Streamflow t0 only)										
ARTIVA							Malvinas			
PCMCI									Amazon	Malvinas
HC										
Exogeneous Ancestors (Parents of Streamflow t-1, t-2, ... t-12)										
PCMCI										
HC		Grad					Malvinas / Nino			
Lag of the endogenous component (Streamflow t0 only)										
ARTIVA	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
PCMCI	-1; -2	-1; -11	-1; -2	-1; -3	-1	-1	-1; -12	-1	-1; -10	-1
HC	-1; -2	-1; -11	-1; -2; -5	-1; -3	-1	-1; -4	-1; -4; -12	-1	-1; -10	-1

Source: Prepared by the author.

The presence of the indices as ancestors and parents to streamflow indicates that although with no influence in the streamflow modelling with a one-month lead time, the use of the indices can possibly aid in the streamflow forecast with higher lead-times. This hypothesis was corroborated by the streamflow sub-networks of the stations 99 (second period – HC results) and 281 (first period – PCMCI results), shown in Figure 12. The complete networks are presented in Appendix B. From this figure, it is noticeable that the Grad influence in station 99 occurs for streamflow events in the absence of longer lagged observations (11-month lag). Similar pattern was observed for station 281, where the Amazon influence occurred with a lag of eight months and may improve the long-lead time forecasting performance.

Figure 12 – Streamflow sub-network for the second period of station 99 (left) and first period of station 281 (right), through HC and PCMCI, respectively. Q nodes stands for the streamflow series in the different time-steps.



Source: Prepared by the author.

Malvinas presented a 12-months lag influence in the streamflow series of the second period of station 281 (network shown in Appendix B), and for the first period of station 188 with also longer lags as seem from the arcs from Malvinas at the time steps t-12 and t-11 to streamflow at t-11 and t-10, respectively (Figure 13).

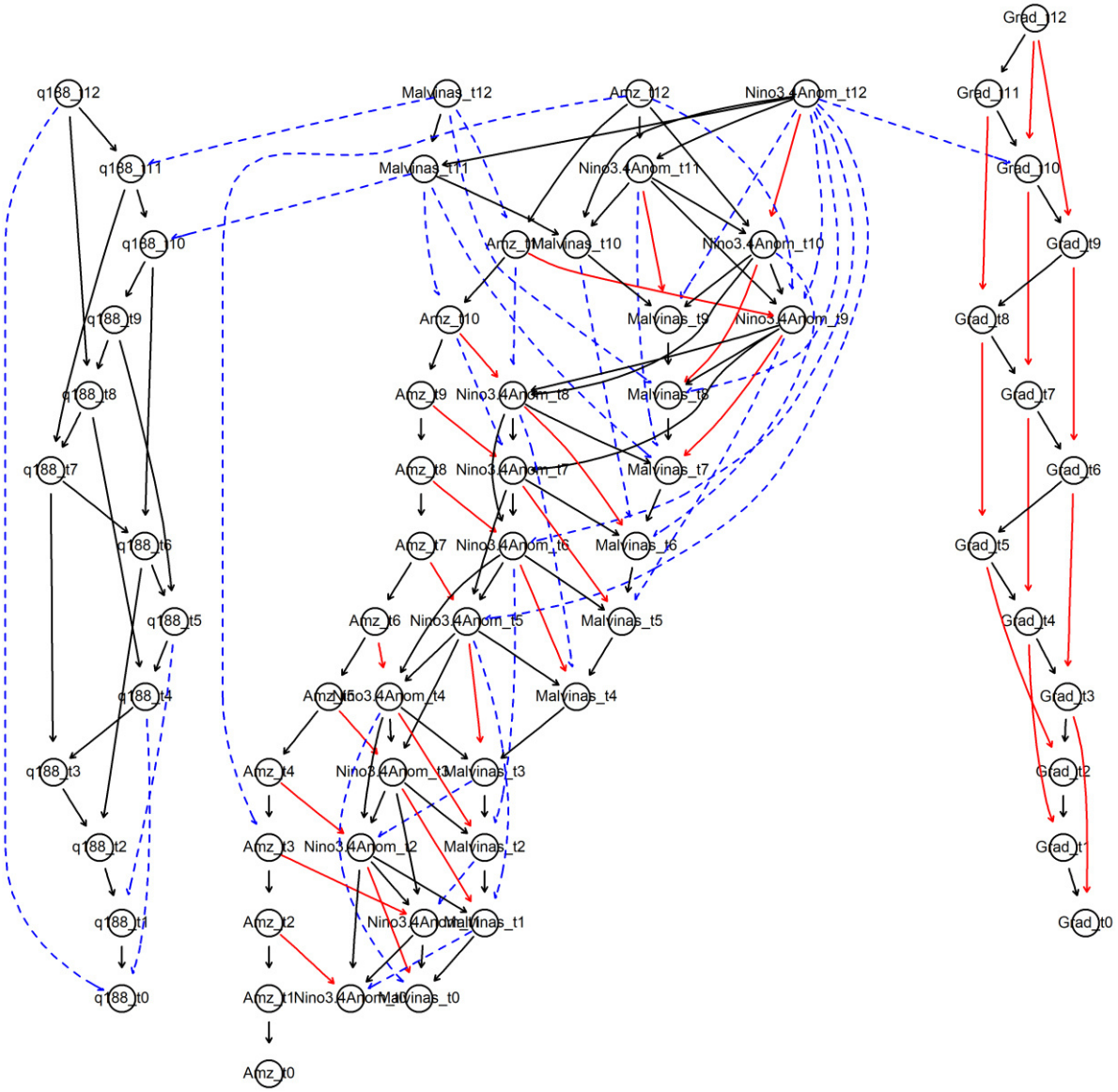
It was noticeable that all exogeneous parents and ancestors did not remain constant for both periods, indicating significant changes in the underlying network structure. The only convergence between the ancestors and parents was observed between ARTIVA and HC for the first period of station 188 for the Malvinas index.

Changes in the network structure were noticeable from the graphical comparison between the results for both periods of station 188 (Figure 13 and Figure 14). The changes observed were not limited to modifications of the relationship between the variables, but also affected the autocorrelation extension of the variables for the network structure obtained through HC and PCMCI. The two methodologies obtained networks with similarities, with significant changes in the autocorrelation of the Grad index between the two periods and a strong correlation between the Nino3.4 and Malvinas. HC also found correlation between these indices and the Amazon index that was not present for any period for PCMCI, whereas, on the other hand, PCMCI indicates a relationship between Grad index and Malvinas that was not obtained by HC. PCMCI showed fewer alterations in the network structure between the two periods in comparison to the HC results and PCMCI also presented fewer arcs.

The autocorrelation present in the Grad index for the first period was further extended for the second period for both methodologies. From HC results, this period also presented a higher correlation between the Amazon, Nino3.4 and Malvinas indices within lower time-steps, although with a reduction of the autocorrelation of Nino3.4 and Malvinas indices and the correlation between them for higher time-steps. This autocorrelation reduction for higher time-steps of Nino3.4 and Malvinas were also obtained by PCMCI. The relationship between Grad and Malvinas occurred only for the second period.

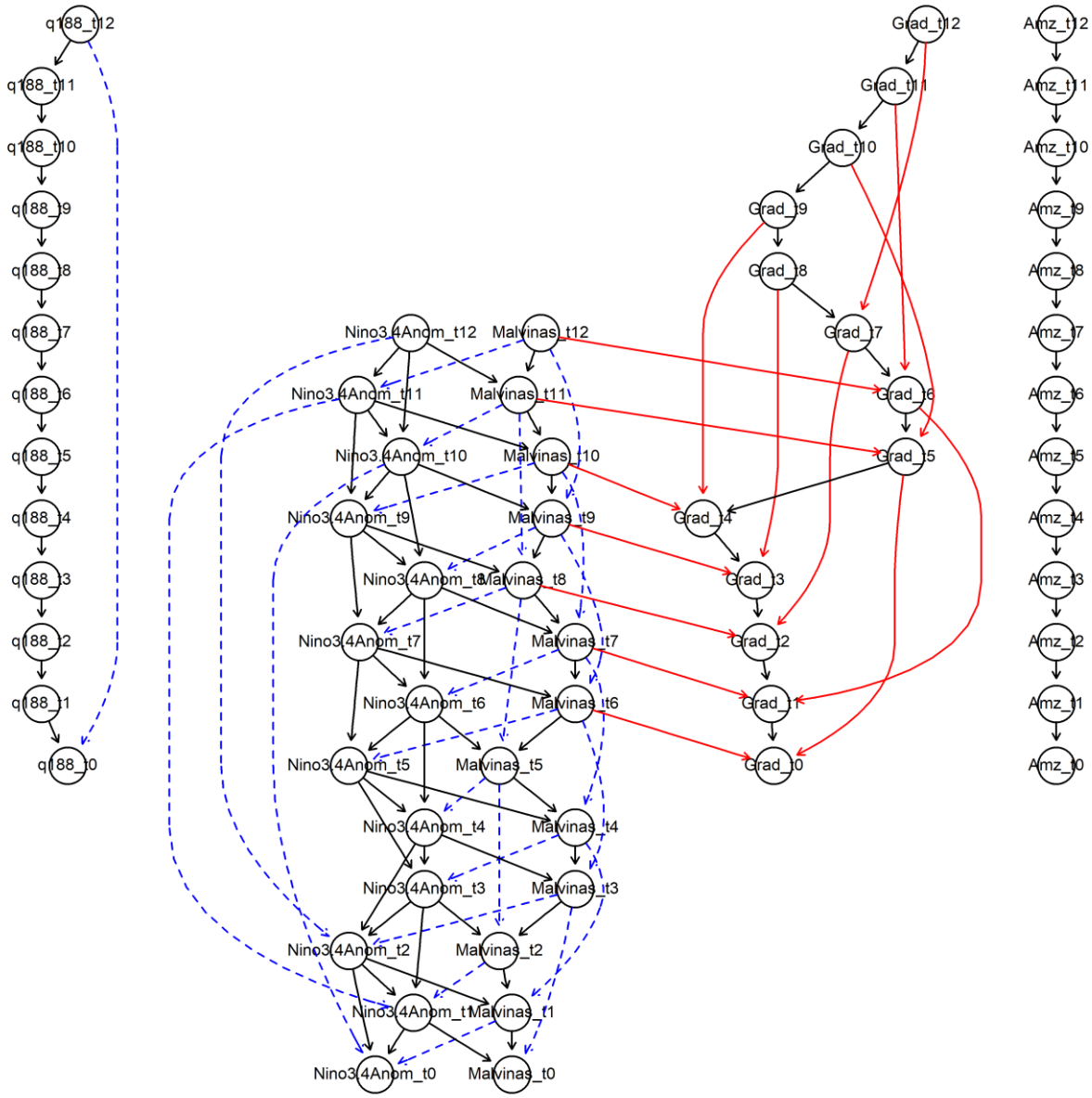
The streamflow also presented changes in its autocorrelation extension, with differences between the HC and PCMCI results. HC networks presented an autocorrelation within various time-steps for both periods, with a reduction for the second period, removing the t-12 influence in the t0 streamflow and also an influence with a 4-months lag for the t0 and t-1. PCMCI results presented a single lag-1 relationship for both periods, with a similar removal of the t-12 influence, resembling the results for this period to the obtained through HC. The plots of the networks structure from HC and PCMCI of each period is also presented in the Appendix B.

Figure 13 – HC network structure comparison between the two periods of station 188. Black, red and blue-dotted arcs represent true positives, false positives and false negatives arcs, respectively. Q nodes stands for the streamflow series in the different time-steps.



Source: Prepared by the author.

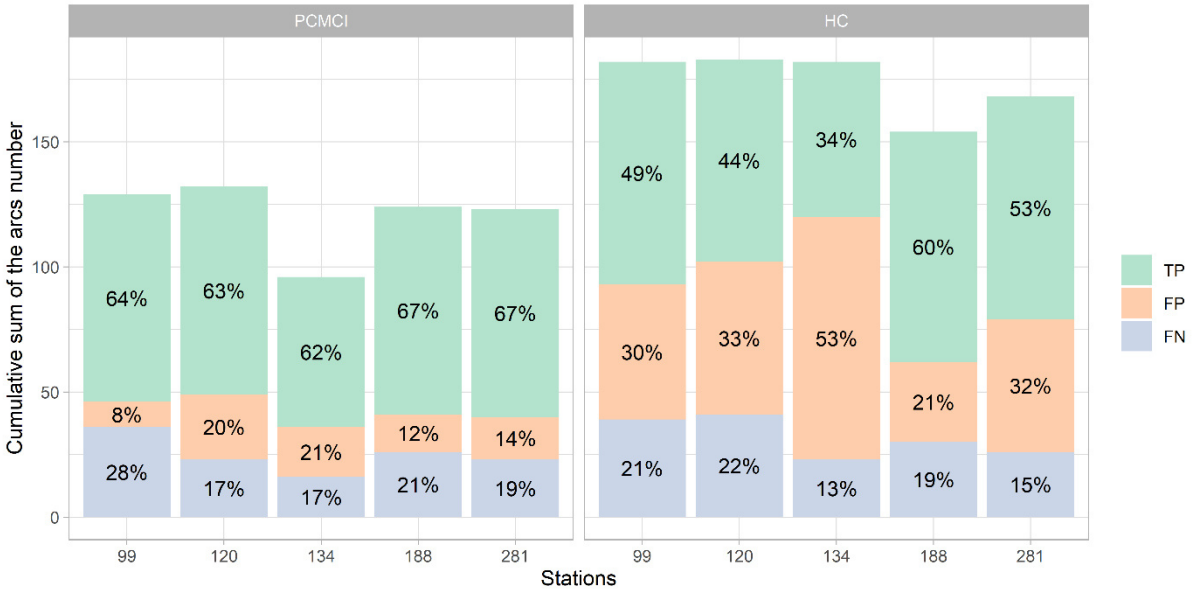
Figure 14 – PCMCI network structure comparison between the two periods of station 188. Black, red and blue-dotted arcs represent true positives, false positives and false negatives arcs, respectively. Q nodes stands for the streamflow series in the different time-steps.



Source: Prepared by the author.

The differences between the networks can be observed through the number of TP, FP and FN arcs (Figure 15). From this figure it was noticeable that some patterns observed for station 188 occurred for the remaining cases analyzed, HC presented a larger number of arcs and alterations between the first and the second period. Whereas the number of TP arcs in PCMCI results comprised more than 60% of total arcs for all stations, for HC this number reached as low as 34% for station 134.

Figure 15 – Comparison metrics and its relative percentual of the total number of arcs.



Source: Prepared by the author.

PCMCI changes from the first to the second period were more related to the arcs loss between the periods, where HC structure changes were mostly related to an increase in the number of arcs. These results indicate changes in the network structure according to the methodology, which was also observed from station 188 results.

6.5 Discussion

ARTIVA framework presented a small change point detection skill, discovering the correct change points for only 15% of the stations analyzed. However, the results obtained for station 281 show that, although not detecting the correct change point, a change imposed by ARTIVA can result in a better performance of the model. This particularity can be explained by the conceptual differences between the two change point approaches, since ARTIVA relies in a linear model and its performance to model the variable certainly influences in the change point results. Also, the single change point constrain imposed to ARTIVA affected the results. A visual inspection of the streamflow time-series of this series (presented in Appendix B) indicates that a single segmentation of the time-series may not correct comprise the step changes existent. This observation is also corroborated by Rocha and Souza Filho (2020) results in

which this particular station presented more than one change point through Pruned Exact Linear Time methodology.

The similar performance obtained for streamflow modelling shows that a simple autoregressive model can produce good results for the study area, especially after the normalization and seasonal remotion procedure. These latter steps resulted in an improvement of the performance for preliminary tests made without the use of single or both procedures. This result was expected due to the strong autocorrelation existent due to the groundwater component.

The use of climate indices and complex lagged influences did not result in the improvement of the streamflow modelling for the time-step t0 for the stations analyzed. However, the influence for longer lead-times of the indices was observed in a couple of stations, indicating it usefulness to forecast streamflow with longer lead-times, in concordance to the results obtained in chapter 5. Although the interrelationship between the indices do not direct influence the modelling performance through the modelling approach used in this work due to the Markov premise of the Bayesian Networks, it may certainly influence the forecast through another approach such as Likelihood Weighting simulations (LW) for longer lead times. This hypothesis derives from the fact that in the absence of the information in determined time-step, LW can be used to infer its value, being thus affected by the existent causal chain of the network. The results obtained in the preliminary chapter already demonstrated the usefulness to forecast streamflow through this approach.

In this way, the diverse networks changes observed in the relationship between the indices and in its autocorrelation needs to be evaluated for the development of the set of Bayesian Networks that comprise all the different structures existent. The temporal networks proposed by Dutta & Maity (2020) can be an interestingly solution to comprise the changes mentioned, however, the structure update framework and the sliding-window size needs to be assessed accounting for the mentioned indirect effect in longer-lead time streamflow forecast. Also, we find that the absence of the bootstrapping process for the HC structure retrieval of the sliding-window framework may result in inferior results and should be incorporated in the methodology. This additional step would greatly increase the run time but would impose more robust results.

The physical reasons behind the changes observed were not analyzed in this work and remains unknown, they can be related to low-frequency oscillations since the time series

used in this work were sufficiently long to capture the influence of these natural oscillations. Also, to correctly assess the causality of the obtained links, a robust analysis of the arcs strength needs to be performed.

The absence of this latter analysis prejudices the direct comparison between HC and PCMCI results and its use could diminish or increase the differences between the networks obtained. The PCMCI networks with fewer arcs than HC was expected since this methodology is intended to obtain networks that can be used to extract causality, thus requiring a more robust analysis. Also, the reduction of the arcs was derived to the use of the false discovery control. The results obtained also showed that PCMCI performed well even though the requirement of causal sufficiency was not fulfilled.

6.6 Conclusions

In this work we used ARTIVA framework to detect and model streamflow series accounting for existent non-stationarity, evaluating its change point detection skill, and modelling performance. Also, the differences in the network structure obtained before and after the change point were analyzed, comparing ARTIVA results to complex network structures obtained through two methodologies (HC and PCMCI) unfolding the streamflow series across time within a given maximum lag of interest, allowing the discovery of lagged relationships between the variables. The streamflow modelling performance obtained through HC and PCMCI was also compared to ARTIVA.

ARTIVA change point detection performance was close to 15%, although the single change point constrain imposed to match previous studies possibly influenced in the results obtained. Also, ARTIVA was used considering only a change point in the streamflow series, whereas the network comparison results indicate the possible existence of change points in the indices series used. The bugs observed for the ARTIVA version used significant reduces its potential and useability.

The networks obtained through HC and PCMCI indicate influences of the climate indices for streamflow forecasting with longer lead times, although with no significant impact in the streamflow modelling with one month lead time, yielding similar performance results to ARTIVA framework that produces simpler networks. The differences between the networks

before and after the change point were not limited to the relationship between the variables but also in its autocorrelation. From these differences emerges the possibility of change points in the climate indices. It was also noticeable that the climate indices found as ancestors to streamflow for some of the stations did not remain constant through time, changing with the change point.

The changes observed in the relationship between the indices had no direct effect through the streamflow modelling framework used, that obtains streamflow values through the use of its direct parents value. However, it may certainly affect the streamflow forecast through another methodologies, such as through LW that infer the streamflow value considering the causal chain of the network. Therefore, this aspect needs to be accounted in the development of streamflow forecasting frameworks.

7 STREAM GAUGE CLUSTERING AND ANALYSIS FOR NON-STATIONARY TIME SERIES THROUGH COMPLEX NETWORKS

7.1 Abstract

This article presents an application of a framework developed to cluster stream gauge stations and analyze changes in the clustering results across time through complex networks and community detection algorithms. Weighted networks were obtained through Mutual Information combined with an automated threshold selection based on Monte-Carlo Markov Chain bootstrapping simulations. The complex networks were obtained for the entire series and for sliding windows of 30 and 20 years, comparing the results through Complex Network Analysis and using the bridge centrality metric to identify transition zones. The weighted network construction procedure presented consistent results, obtaining networks with coherent spatial dispersion of its communities and degree distributions that do not fit a random network typology. The 30 years sliding windows networks results demonstrated that significant changes occurred across-time, obtaining three different configurations of the communities, thus, the use a single network can result in a miss representation of local characteristic and induce wrong conclusions. The communities evolution across time showed spatial-time coherence with the physical phenomenology and previous studies. The changes were associated to phase-shifts of low-frequency sea surface temperature oscillations of both the Atlantic and the Pacific oceans through its direct and indirect influence in the South Atlantic Convergence Zone. The bridge closeness metric presented good results, successfully discovering stations that presented changes in its community across-time and characterizing a coherent transition zone between different communities. The 20 years sliding window did not produce good results.

7.2 Introduction

The clustering of stream gauges is important to discover similar stations regarding the variability of interest for the purpose of the research, such as seasonal and interannual. The analysis of the grouping results enables the investigation of the physical reasons of the found similarities, such as common teleconnections and climate forcing. Also, the grouping results enables the analysis of spatial patterns and extension of stations that show simultaneous or

opposite behavior, an important information when dealing with the operation of an interconnected system of reservoirs. Finally, clustering methodologies can reduce the dimensionality of the streamflow modelling process, allowing a more in-depth analysis of fewer stations that were selected to encompass the different regional behaviors, an important aspect since streamflow modelling often requires complex methodologies to contemplate the non-linearity present in streamflow time-series.

The clustering problem, however, is a not simple straightforward one and involves different aspects, such as the selection of the clustering technique, from diverse methodologies present in the literature, and the analysis of the clustering results, to select the number of clusters adequate, which can involve the use of dendrograms, quantitative indices and plot analysis (CHARRAD et al., 2014). The diversity of materials presented in the literature shows that this is a still on-going problem that has room to be improved.

Recent field-related articles uses complex network and community detection to cluster rainfall and stream gauges, a methodology derived from social network analysis that has been surging as an strong alternative to classic methods (HAN et al., 2020; JOO et al., 2020). This methodology involves two steps: (1) the creation of a network with the aid of a similarity metric to determine the links between the nodes that represent the stations, and (2) the use of a community detection algorithm to encompass similar stations in a community (cluster).

The results obtained in these articles emphasizes the usefulness of this approach for stream gauges clustering. Comparing to the K-Means technique, Joo et al. (2020) concluded that community detection was a superior alternative to this classic method, providing a more effective clustering in terms of hydrologic similarity, persistence and connectivity. Halverson & Fleming (2015) observed that the communities obtained presented a befitting regional spatial dispersion, with stream gauges of the same community presenting similar seasonality patterns. Also, the communities detected were able to capture physical characteristics such as drainage area and elevation, where stations with similar characteristics belonged to the same community.

The use of complex networks also unveils different aspects that emerges from complex network analysis, which can be used to characterize the type of the network obtained, allowing the analysis of its stability, efficiency and robustness, and also to classify the nodes in terms of its role in the community, enabling the analysis of transition zones between the clusters and other possibilities (AGARWAL et al., 2018; HALVERSON; FLEMING, 2015; SIVAKUMAR; WOLDEMESKEL, 2014).

The two-step methodology mentioned, although, involves several particularities that can change the outcome results. Starting with the network creation, correlation is commonly used as a similarity metric (FANG; SIVAKUMAR; WOLDEMESKEL, 2017; HALVERSON; FLEMING, 2015; JOO et al., 2020; SCARSOGLIO; LAIO; RIDOLFI, 2013; SIVAKUMAR; WOLDEMESKEL, 2014). However, when comparing networks obtained using mutual information and correlation, Donges et al. (2009) observed high degree of similarity in two of the scales analyzed but a larger differences for the third one, concluding that the use of mutual information (MI) can yield promising new perspectives when dealing with nonlinear physical processes. The benefits of using MI for nonlinear data is also stated by Kim et al. (2019), that obtained superior results when dealing with nonlinear rainfall data, and by Donges et al. (2009b), which concluded that the results obtained could not be captured by classical linear methods.

The selection of the threshold of the similarity metric emerges as another particularity for the network creation. Diverse approaches can be found in the articles mentioned, ranging for the direct use of different thresholds values, to the selection of a threshold that produces a fitting value of a complex network analysis metric such as edge density and modularity, this latter obtained after the application of the community detection algorithm. The significant impact of the threshold value in the results has been stated by Fang, Sivakumar & Woldemeskel (2017); and Sivakumar & Woldemeskel (2014). The variety of threshold selection methods and the fact that this value can significantly impact in the network obtained reveals the necessity of further studies aimed to diminish this uncertainty.

Another aspect is the use of unweighted and weighted networks. Although mentioned by Halverson & Fleming (2015) as an enhancement possibility of the results obtained, the use of weighted networks is not commonly found when dealing with streamflow data. We find that a combination of weighted networks and an automated approach for the selection of the threshold could be a viable promising solution for the uncertainty mentioned, where the weight of the network would reduce the impact of eventual links created between nodes that do not present a high degree of similarity.

Finally, the presence of non-stationarity in streamflow data can alter the conclusions drawn on the network obtained, stating that they may not be valid for the entire time series, but only for a determined period. Although the non-stationarity is commonly mentioned in the hydrology field, it was not yet contemplated in articles that used community detection for clustering, to the best of our knowledge.

Therefore, in this article we aimed to present an application of a framework developed to contemplate all the particularities mentioned, obtaining weighted networks through MI with a threshold selection based on Monte-Carlo Markov Chain (MCMC) bootstrapping simulations and comparing the network obtained for the entire series to networks obtained for sliding windows of 30 and 20 years. Also presenting an application of a novel complex network analysis metrics called bridge centrality (JONES; MA; MCNALLY, 2019) to analyze the community results.

7.3 Data and Methodology

7.3.1 Study Area Background and Data

Brazil country was selected as study area. This country has strong dependence of its hydropower electricity production, elevating the importance of further studies of its streamflow data. The non-stationarity present in some stations was observed by at least two recent articles using different methodologies (ROCHA; DE SOUZA FILHO, 2020; ROLIM; FILHO, 2020), qualifying the series as a good application of the framework. Finally, the analysis and identification of eventual changes in the relationship of the stations has a local relevance since a stationarity relationship between the stations is assumed and used for the development of projections, scenarios and also for streamflow forecasting (CEPEL, 2006).

This study used monthly naturalized streamflow data of 88 stations scattered along Brazil, with a higher concentration in the Southeast and South region of the country, regions that encompasses mostly of the hydropower production. The naturalization process can be found in detail in ONS (2017) and aims to remove anthropogenically derived factors that modifies the natural variability of the streamflow process. Annual naturalized streamflow, from 1931 to 2015, was obtained averaging monthly values according to the hydrologic year. The use of annual streamflow data aims to cluster stations with similar interannual variability, diminishing the seasonal oscillations. We find that interannual clustering can provide relevant information for long-term energy production projections.

7.3.2 Complex Network through Mutual Information and MCMC threshold selection

A network represents the relationship between the variables (nodes or vertices) through directed or undirected connections (links or edges). Therefore, a network can be expressed as $G = \{V, E\}$, with V representing a set of N nodes $(V_1, V_2, \dots, V_N)$ and E a set of n edges. The link creation between two nodes relies on theorical assumptions based on the expertise of the researchers or in similarity metrics combined with a threshold criteria.

Mutual Information measures the mutual dependency of two variables, being also sensible to non-linear dependencies which cannot be detected by simpler metrics, e.g. linear correlation coefficient, thus, acting as a measure of true dependence (KRASKOV; STÖGBAUER; GRASSBERGER, 2004). The construction of a network based on this metric results in an undirected network.

MI is a function of entropy and exact computation is considered feasible for just discrete variables and special cases with known probabilities distribution (BELGHAZI et al., 2018). In this paper, a discretization approach was used through the R package *muti* (SCHEUERELL, 2017), with the default five symbolic representation of the streamflow annual series:

$$\begin{cases} Q_{t-1} < Q_t > Q_{t+1} \rightarrow Peak \\ Q_{t-1} < Q_t < Q_{t+1} \rightarrow Increasing \\ Q_{t-1} = Q_t = Q_{t+1} \rightarrow Same \\ Q_{t-1} > Q_t > Q_{t+1} \rightarrow Decreasing \\ Q_{t-1} > Q_t > Q_{t+1} \rightarrow Trough \end{cases} \quad \text{(Eq. 9)}$$

where Q_t, Q_{t-1} and Q_{t+1} stands for the streamflow event in the timestep t, immediately before t-1 and after t+1.

The entropy $H(X)$ of a series X can be calculated through the following equations:

$$H(X) = E\big[-\log\big(p(X)\big)\big] = -\sum_{i=1}^{L} p(x_i) log_b p(x_i) \quad \text{(Eq. 10)}$$

where $p(X)$ is a probability mass function, L is the length of the time series and b is the logarithm base. *muti* package uses a logarithm of base two, measuring the information in units of *bits*.

The joint entropy between two series (X and Y) is given by the following equation:

$$H(X,Y) = -\sum_{i=1}^{L} p(x_i, y_i) log_b p(x_i, y_i) \quad \text{(Eq. 11)}$$

where $p(x_i, y_i)$ is the probability that $X = x_i$, and $Y = y_i$.

Finally, MI can be calculated and normalized (MI^*) through the equations below:

$$MI(X;Y) = H(X) + H(Y) - H(X,Y) \quad \text{(Eq. 12)}$$

$$MI^*(X;Y) = \frac{MI(X;Y)}{\sqrt{H(X)H(Y)}} \quad \text{(Eq. 13)}$$

Normalized MI ranges between 0 and 1, with 1 standing for a perfect simultaneous behavior of series X and Y. The results were used to construct a weighted network representing the interrelation between the stations, where each node represents a station and an edge linking one node to another is created if the result is higher than a 95% significance threshold obtained through MCMC bootstrapping. The weight of the edge is the MI* value and stands as a similarity measure.

The MCMC bootstrapping is performed by obtaining the transition matrix between discrete categories of the X and Y time series, randomly choosing the initial value, and then pooling values according to the transition matrix probabilities until a series with same length is obtained.

7.3.3 Community Detection for Weighted Networks

The detection of communities in a complex network aims similar objectives as clustering techniques and in a broad sense the communities can also be referred as clusters. The community detection algorithms aim to group nodes that are closely connected and, therefore, share similar attributes and presents a certain degree of dissimilarity or independence from the remaining nodes outside the belonged community.

Although presenting similar goals, classical clustering techniques and community detection have a significant methodologic difference: the first one relies only in the direct similarity between the time series data, where the second is based in both the similarity of the data and the network theory (JOO et al., 2020).

To optimize the resulting communities, mostly algorithms are based on a measure of the division quality called modularity. Modularity was developed to compare the links within the communities obtained to random theoretical connections. The concept derives from the idea that if a natural random expected division of the network exists, the connection within a community need to surpass the expected random theoretical value while presenting opposing behavior for the connections between the communities (CHANG; LEAHY; PANTAZIS, 2012). Newman (2004) states that the probability of an edge existing between nodes i and j is given by $k_i k_j / 2m$, where k_i is the degree of node i and m is the total sum of the weights of all edges in the network. In weighted networks the degree of a given node is the sum of the weights of its links, thus:

$$k_i = \sum_j A_{ij} \therefore \; m = \frac{1}{2}\sum\nolimits_{i,j} A_{ij} = \frac{1}{2}\sum\nolimits_i k_i \qquad \text{(Eq. 14)}$$

where A_{ij} is the weight of the link between nodes i and j, an element of the weighted adjacency matrix **A**. The Modularity is then given by (NEWMAN, 2004):

$$Q = \frac{1}{2m}\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2m}\right)\delta\left(c_i, c_j\right) \qquad \text{(Eq. 15)}$$

where $\delta(c_i, c_j)$ is a function that returns one if the community of the node i is equal to node j $(c_i = c_j)$ and zero otherwise.

Modularity metric ranges between zero and one, with values close to one indicating a strong community structure within the network and values close to zero express a close to random number of within-community edges. Newman & Girvan (2004) states that typical values of not theoretical networks usually range from 0.3 to 0.7. Newman (2004) also states that values around 0.3 or higher indicates good divisions.

Modularity optimization, however, is not a simple task, which lead to the development of diverse algorithms to overcome the inherent complexity. Although no significant differences were detected by Halverson & Fleming (2015) when testing different methodologies, Sousa & Zhao (2014) found that the *walktrap* algorithm (PONS; LATAPY, 2005) leads to better results on average. Both articles used the same R package *igraph* (CSARDI; NEPUSZ, 2006), also used in this work. This algorithm is based in a simple assumption that short random walks along the network have the tendency to stay within the same community and was used in this work.

7.3.4 Community changes detection across time

To investigate if significant differences occur across time, the communities detected for the network obtained for the entire series were compared to the results obtained using sliding windows of 20 and 30 years through five different metrics: the variation of information (vi) (MEILA, 2003), the normalized mutual information (NMI) (DANON et al., 2005), split-join distance (split.join) (VAN DONGEN, 2000), the Rand index (rand) (RAND, 1971) and the adjusted Rand index (adjusted.rand) (HUBERT; ARABIE, 1985).

The results within the sliding windows were also compared to analyze the different configurations that occurred across time, and to check the stability of the communities. For this procedure, only the metric that produced the best results in the previous step was used. The different configurations were obtained with the aid of a hierarchical clustering. For each different community configuration, a representative time-window was selected based in the metric results within the cluster, therefore, the time-window that presented the highest mean value of the metric was selected as representative for the time-windows that belonged to the same cluster.

7.3.5 Complex Network Analysis

A more in-depth analysis was made for the network of the entire series and for the networks of the selected time-windows. Firstly, a simple analysis was made to verify if the topology of the networks obtained classify as a random network. The detection of the networks communities does not require assumptions regarding the network topology (HALVERSON; FLEMING, 2015), however, a random network typology states that the links between the nodes could occur by chance which is not in concordance for a network constructed based on streamflow variability. The streamflow variability is a product of climate variability and physical characteristics which impose a spatial consistent behavior. Therefore, a random network typology would raise concerns regarding the reliability of the network construction methodology. The network typology was evaluated using the clustering coefficient and analyzing the degree distribution.

The clustering coefficient (C) evaluates the tendency of the nodes to cluster together and is expected to produce much larger values in real networks when compared to a correspondent random network with the same number of nodes and links ($C \gg C_{random}$). Weighted networks has a slightly different calculation of this metric to correctly include the weights of the network, in this work it was followed the definition of Barrat et al. (2004).

The degree of a node corresponds to the number of connections with other nodes. The complex network theory established that the degree distribution of a random network follows a binomial distribution. Therefore, the histogram comparison between the network obtained and the correspondent random network allows the analysis of the network typology.

Secondly, the bridge closeness metric was used to analyze zones of transition between the different communities. Bridge closeness evaluates the average distance between a node and other nodes outside its community (JONES; MA; MCNALLY, 2019). The distance between nodes of an unweighted network corresponds to the shortest path between two nodes, i.e. the number of edges. In weighted networks, the weights are interpreted according to theorical concepts. For example, the classic Dijkstra algorithm interpret the weights as costs when finding the shortest paths among nodes, and thus, the distance needs to be based in the inverse of the edge weights if they act as a similarity measure.

Bridge closeness is calculated using the inverse of the average distance and nodes presenting higher values of this metric are more likely to share common attributes with different communities, acting as a bridge (transition) between the communities. To decide which node act as a bridge Jones, Ma and McNally (2019) suggests the evaluation of the metrics distribution. Through a preliminary analysis, we decided to consider the stations with top ten highest values as bridge, which result in a more restrictive threshold (~ 11%) than used in the original article after theorical simulations (20% highest values). Also, no significative changes were observed when considering the other bridge centrality measures (strength, betweenness and expected influence). The results were obtained through the R package *networktools*.

7.4 Results

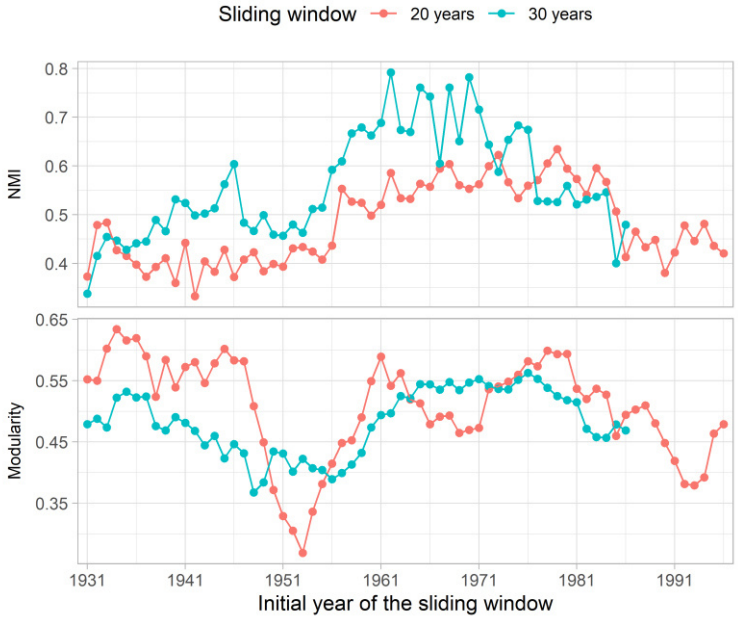
7.4.1 Community changes detection across time

All five metrics suggested changes in the communities across time when used to compare the results obtained for the complete series network (CSN) to both the sliding windows networks. We decided to adopt the NMI metric as standard for this work, since it ranges within a defined interval (0 to 1) and the values obtained showed significant variation along it scale. All five metrics presented similar patterns and are presented in the Appendix C.

Figure 16 shows that the communities obtained for the CSN shared more similarities with the communities obtained using the sliding window of 30 years (SW30), where the highest values were obtained for the sliding windows starting within 1961 to 1976. The sliding window of 20 years (SW20) showed increasing similarity since 1950 until close to 1980, with low values for the latest ones, in concordance to the results obtained for the SW30.

Figure 16 also presents the modularity results of the networks obtained. The SW20 showed a high variation across-time of its modularity values, from 0.26 to 0.63, with a strong V-shaped pattern between 1946 to 1961. The SW30 presented a steady oscillation between 0.36 to 0.56 without a significant shift. The modularity of the CSN was 0.44. These results indicates that the communities obtained for the SW20 presented strong variations within small changes of the sliding windows, which can result in a fragile stability of the communities.

Figure 16 – Comparison through NMI metric between the network obtained for the complete series and the networks obtained using sliding windows of 20 and 30 years (top). Modularity result of the networks obtained using the sliding windows (bottom).

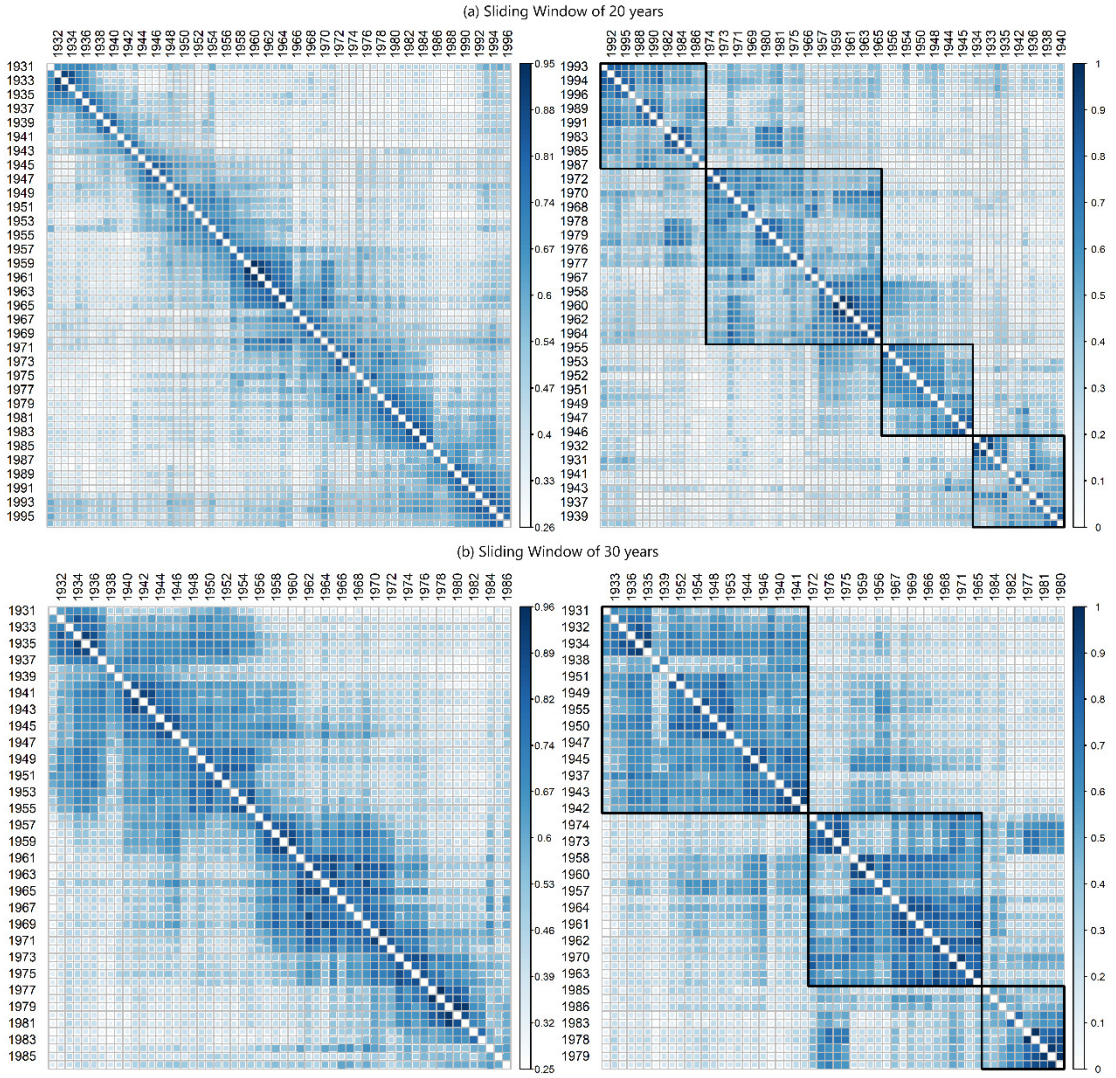


Source: Prepared by the author.

The results presented in the Figure 17 (left) also corroborates this observation, the similarities between close windows shows a fast decrease within a small number of years for the SW20. On the other hand, the SW30 shows a marked cohesion for several years indicating a stronger stability of the communities obtained, without large oscillations.

The hierarchical clustering analysis of the NMI results between the sliding windows, indicates four clusters for the SW20 and three for the SW30 (Figure 17 – right), these clusters were selected by the analysis of different cluster numbers and detail regarding each different configuration can be found in Table 7.

Figure 17 – Results of the NMI metric applied to intercompare the networks obtained using sliding windows. Continuous results (left) and clustered (right). The axis value corresponds to the initial year of the window.



Source: Prepared by the author.

Table 7 – Different configurations of the sliding windows communities.

	SW20				SW30		
Cluster	1	2	3	4	1	2	3
Initial year	1931	1944	1957	1982	1931	1956	1977
Final year	1943	1956	1981	1997	1955	1976	1986
Representative period	1936-1955	1950-1969	1970-1989	1994-2013	1934-1963	1969-1998	1980-2009

Source: Prepared by the author.

The clusters found for SW30 also shows a stronger cohesion than SW20 clusters, where diverse periods grouped together do not show high NMI values. However, SW30 also

presented time windows with abnormal low NMI values within the same cluster, which can be clearly noticed for the windows starting in the years of 1938 and 1939 for the first cluster, 1973 to 1976 in the second and 1984 to 1986 in the third. For the second cluster, the years mentioned are close to the start of the third cluster and delimitates a transition between them. The years mentioned for the first and third cluster are close to the limits of the series that suffers from border effects that suggests a previous and forthcoming configuration before the first cluster and after the third one, respectively. The years of change in the clusters of both sliding windows shows some resemblance, occurring close to the year of 1956 and 1980.

The fragile stability observed for the communities detected for the SW20 raised concerns regarding its reliability and were discarded, considering only SW30 results for the remaining analysis.

7.4.2 Complex Network Analysis and Spatial dispersion

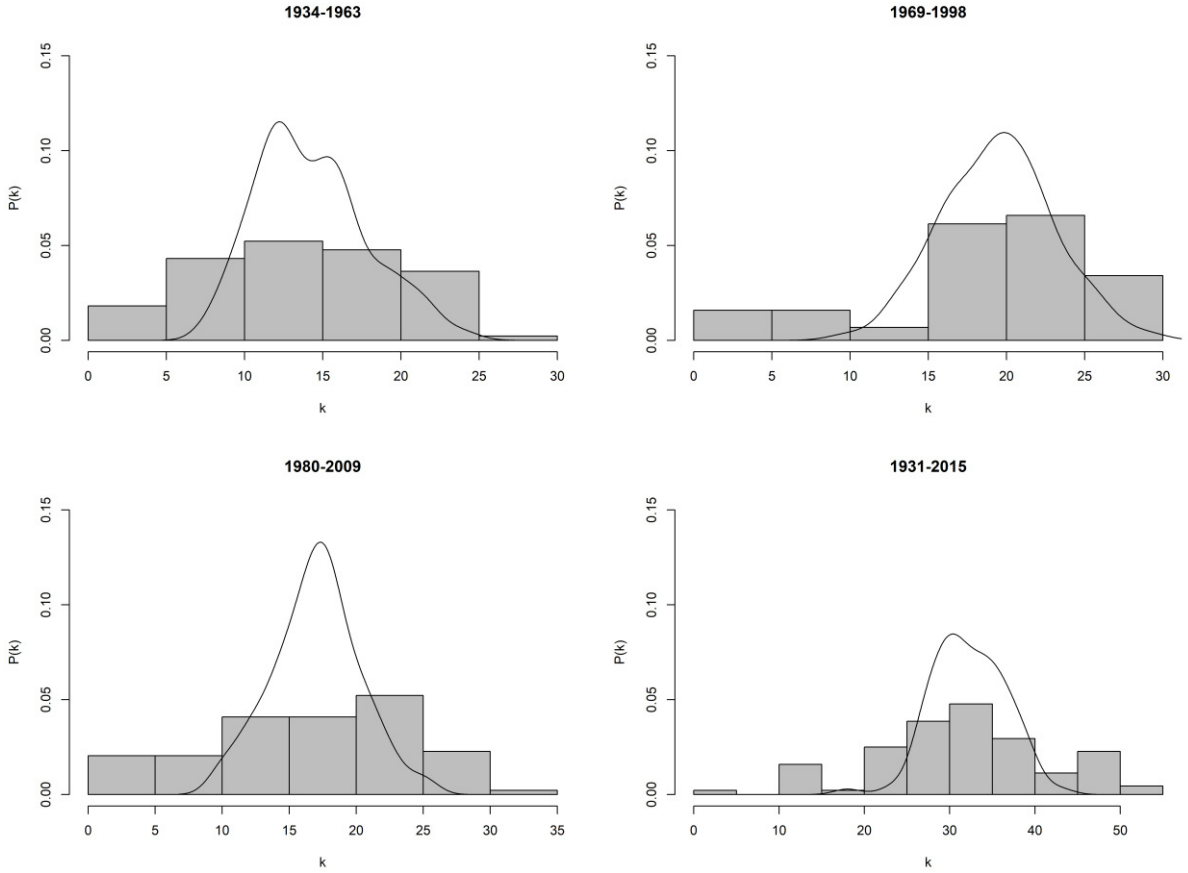
The clustering coefficient results (Table 8) indicated that none of the selected networks presented a random network typology. The CSN was the network that presented more similarities, since its clustering coefficient showed the smaller difference in comparison to the correspondent random network value, only 1.75 times the C_{random} value where the remaining time-windows were close to 3 times the C_{random}. The histogram plots also corroborate these conclusions (Figure 18), none of them follow the equivalent random distribution, with the CSN showing the closest results.

Table 8 – Clustering coefficient results for the different time-windows networks (C) and correspondent random networks (C_{random}). Random networks generated with the same number of nodes and links.

Time-windows	C	C_{random}
1934-1963	0.56	0.17
1969-1998	0.65	0.22
1980-2009	0.63	0.19
1931-2015	0.63	0.36

Source: Prepared by the author.

Figure 18 – Histogram plot of the network's degree distribution. Solid lines represent the degree distribution of the correspondent random network (same number of nodes and links).



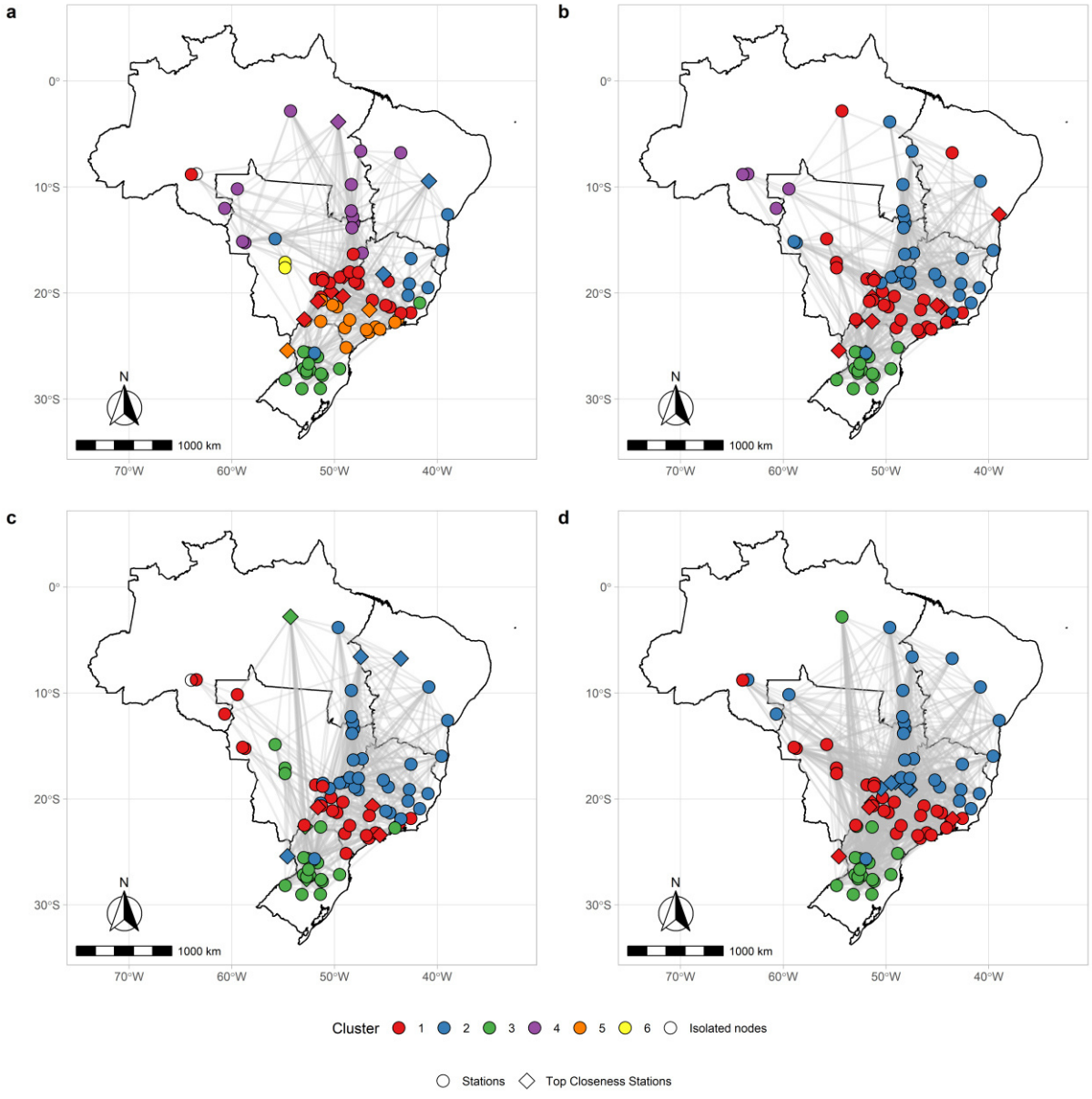
Source: Prepared by the author.

The spatial dispersion of the communities is shown in Figure 19 where the evolution of the communities is presented in Figure 20. It was notable that mostly southern stations remained in the same cluster (3) for all the periods analyzed. One station (Itaipu) was considered a bridge station through the bridge closeness criteria. This classification was in concordance with the communities' results since this station belonged to different communities across the time-windows which reinforces the similarities shared with diverse stations. Another station belonged to the cluster 2 throughout the different periods. The cluster 3 itself presented the largest changes in the third period (1980-2009, Figure 19c) with the presence of additional stations in the Middle-west and North regions, showing smaller changes across the other periods.

The remaining clusters showed more significant changes, especially when compared to the first time-window (1934-1963, Figure 19a). The major differences were delimited to the clusters 4 and 5. The significant relevance of the cluster 4 in the first period

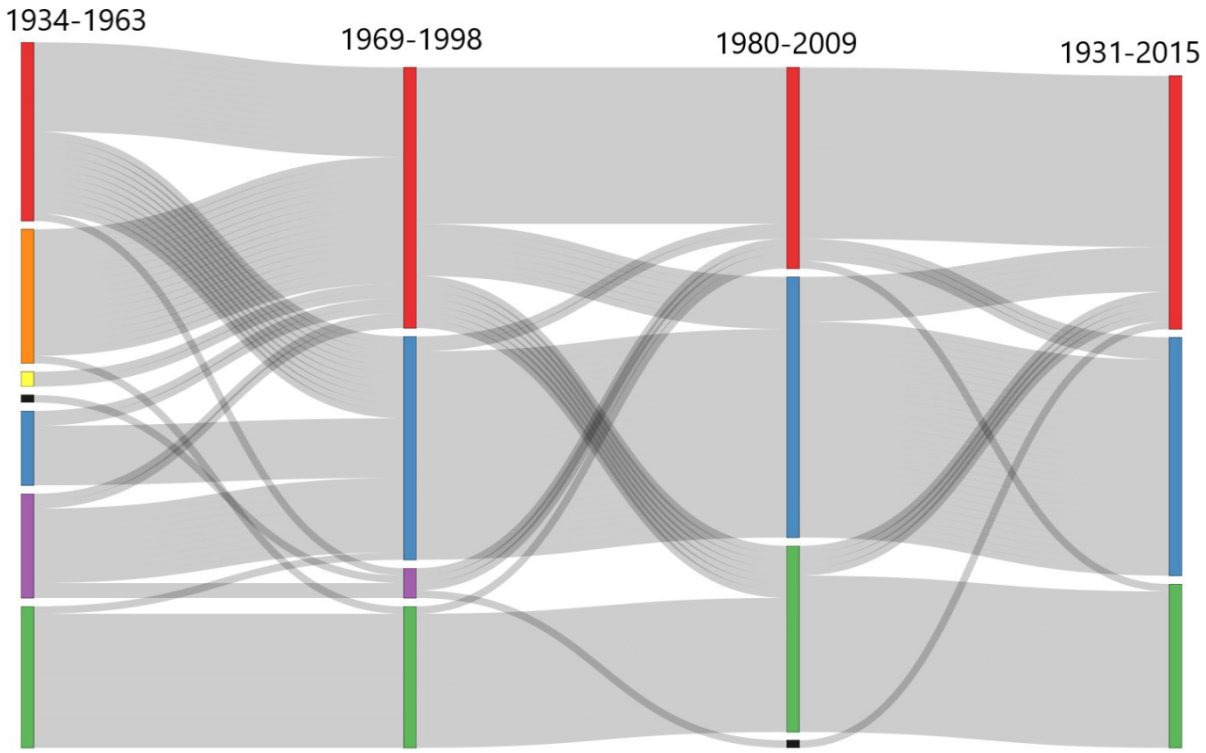
translates to a small presence concentrated in the North region for the second period (1969-1998, Figure 19b), mostly due to the western expansion of the cluster 2 combined with a south-west displacement of the cluster 1. This displacement comprised the stations that belonged to the cluster 5 and 6, also resulted in a migration of cluster 1 stations to cluster 2.

Figure 19 – Spatial dispersion of the communities (clusters) for different time windows: (a) 1934-1963, (b)1969-1998, (c) 1980-2009 and (d) 1931-2015. The top 10 stations with highest values of bridge closeness are highlighted. Stations that were classified as a standalone community are referred as isolated nodes.



Source: Prepared by the author.

Figure 20 – Sankey diagram of the community's differences across the selected time-windows. The colors follow the same standard of the spatial plot with the black rectangle standing for the stations that were classified as a standalone community. The grey lines represent the station's trajectory through time.



Source: Prepared by the author.

The third period show a continuity of the previous observed movements, with the cluster 1 showing a western expansion, comprising the remaining stations of cluster 4, and cluster 2 absorbing diverse cluster 1 stations, where the mentioned cluster 3 expansion comprised some stations of the cluster 1. The cluster evolution across these three periods indicates a reduction of the different variability patterns across time, resulting in three large spatial consistent patterns. Both the third and the second period presented a station that did not belong to any community.

The CSN (1931-2015, Figure 19d) also shows similar spatial concentration of the bridge stations. Therefore, the Southeast area close to the lower limits between the Middle-west and the Southeast region seems to act as a transition zone, since it concentrated most of the bridge stations. The bridge classification could also detect stations that presented cluster changes across time.

It was clear that the CSN shared strong similarities with the second period but diverged from the spatial pattern obtained for the first and third. These results were in

concordance to the NMI results and indicates that it should not be used as representative for the complete series. The presence of more links between the stations indicates that the methodology resulted in a more densely connected graph for the CSN, which was corroborated by the network plot found in the Appendix C.

7.5 Discussion

The proposed network construction methodology yielded solid results for the SW30, presenting a stability of the communities obtained within close time-windows, smaller fluctuations of the modularity values when compared to the SW20 and network typologies that do not follow a random distribution. However, the histogram plot of the degree distribution of the CSN shared small resemblance to a random network. This result is similar to the obtained by Halverson and Fleming (2015).

The higher similarity observed for the CSN can be explained by its more densely connected graph, which was produced by time-window delimited correspondent behavior between the stations, the same one that produced different communities across time. A time-window delimited correspondence, that do not stand for the complete series, produces a link with a comparable smaller value of MI. The consistent spatial dispersion obtained for the CSN indicates that the use of the weighted networks can successfully diminish the impact of these comparable weaker connections and result in a network that represents the local overall behavior. However, the NMI results and the spatial dispersion of the communities both show that significant variations of the communities can occur across time and, therefore, the use a single network can result in a miss representation of local characteristic and induce wrong conclusions. Therefore, through the proposed methodology a densely connected graph can be an indicative that the time-window used do not correct represent the local characteristic.

The selection of the correct time-window is an aspect that needs further studies, here we used two fixed windows and obtained good results with the SW30. The stability of the SW20 community results were certainly affected by the smaller time-window and the discretization procedure, which further reduces the statistical significance due to the comparison of the streamflow event (t) with the previous (t-1) and next (t+1) events, ending with only 18 observations for each time-window. Therefore, smaller time-windows than 20 years do not seems feasible, but a middle range (~25 years) can possibly yield good results.

The spatial pattern obtained in this work is in coherence with the reported influence of the South Atlantic Convergence Zone (SACZ) in the local streamflow (LIMA; LALL, 2010a; ROCHA; DE SOUZA FILHO, 2020). Rocha & Souza Filho (2020) associated step-changes observed in Brazilian streamflow series to the impacts of the phase-shift of low-frequency SST oscillations in the SACZ. SACZ is known to extend from the Southeast to the North region of Brazil in a straight-line pattern, with some variations in its extension and north-south displacements. The cluster 1 pattern obtained for the CSN shows remarkable similarities with the SACZ spatial extension.

The cluster evolution corroborates the low-frequency influence in the communities detected. Chiessi et al. (2009) observed southward displacements of the SACZ occurring during cold phases of the Atlantic Multidecadal Oscillation (AMO), similar displacement was observed in the cluster evolution. A cold phase of the AMO reported occurred around 1964 to 1994 (ENFIELD; MESTAS-NUÑEZ; TRIMBLE, 2001; ROCHA; DE SOUZA FILHO, 2020). The second configuration found comprised time-windows with initial years between 1956 and 1976, mostly of them belonged to a cold phase of AMO. Similarly, half of the years of the third representative time-window (1980-2009) occurred within the same cold phase.

The initial year of the third configuration is also consistent to the changes reported in the streamflow series and to a cold-warm phase shift of the Pacific Decadal Oscillation (PDO) that occurred around the same year (1976). The 1975 phase-shift of the PDO was associated by Keller et al. (2009) to more frequent and intense events of El Niño. The El Niño Southern Oscillation (ENSO) influence in SACZ was reported by Cavalcanti (2016) as a possible influence in its persistence. Therefore, the years that delimits the similar time-windows seems to be influenced by both the PDO and AMO phases, indicating a strong low-frequency impact in the relationship between the stations, in total concordance to the results of Rocha & Souza Filho (2020).

Finally, the bridge closeness metric presented good results in discovering the stations that shared similarities outside its community. Diverse stations classified as bridges presented changes in its belonging community across time, in concordance to the results obtained through the metric. Also, the area characterized as a transition zone matches the works of Lima & Lall (2010a) and Rocha & Souza Filho (2020). The threshold selection of this metric relies on the analysis of its value distribution and, although presenting good results in this work for the threshold selected, we find that further studies specifically analyzing streamflow

networks can be useful to stablish a standard selection procedure, which was not the goal of this article.

7.6 Conclusions

In this paper we presented an application of a framework developed to construct weighted complex networks, detect and analyze communities changes across time. The network construction was based on mutual information combined with a threshold selection through MCMC bootstrapping simulations. The communities detected for the network obtained for the entire time series were compared to the results of networks constructed using sliding windows of 30 and 20 years. We also used the bridge closeness metric to detect the nodes that act as a bridge between the communities, i.e. that share similarities with nodes outside its own community.

The weighted network construction procedure presented consistent results, obtaining networks that showed coherent spatial dispersion of its communities and degree distributions that do not fit a random network typology. The use of a weighted network could effectively diminish the impact of weaker connections that were a result of time-delimited similarities between the stations, successfully obtaining a network for the entire series that could represent the local overall behavior. However, the comparison of the results with the 30 years sliding windows networks demonstrated that significant changes occurred across-time and the use a single network can result in a miss representation of local characteristic and induce wrong conclusions. The sliding window of 20 years did not present consistent results due to its reduced statistical significance combined with the discretization procedure used for the mutual information calculation.

The normalized mutual information community comparison metric was successful in detecting the different communities configurations existent. The results obtained showed three different configurations and the communities evolution across time showed spatial-time coherence with the physical phenomenology and previous studies. The changes were associated to phase-shifts of low-frequency sea surface temperature oscillations of both the Atlantic (AMO) and the Pacific (PDO) oceans. Both these low-frequency patterns affected the stations relationship through its direct and indirect influence in the South Atlantic convergence zone.

Finally, the bridge closeness metric presented good results, successfully discovering stations that presented changes in its community across-time and characterizing a coherent transition zone between different communities. However, we find that further related studies are required to develop a standard selection procedure for a threshold value of this metric, which was not the goal of this article.

8 FINAL CONCLUSIONS AND REMARKS

In this study it was developed a methodology to detect changes in the streamflow time series and incorporate the detected non-stationarity in streamflow modelling and forecasting, also assessing the non-stationarity impact in the relationship structure between the different stations used.

The change point detection framework showed a good performance, with PELT outperforming the different methodologies used (Pettitt test and Bai and Perron dynamic algorithm). As mentioned in the first article, the combined use of different methodologies with a similar reliability assessment goal was not found in the literature and the framework here proposed unveils a promising new approach to assess the reliability of the results. However, the significant convergence chance for erroneous results shows that the choice of complementary methods still has a wide margin to be improved. Also, although PELT presented the best results among the methodologies used, its tendency of discovering false positives need to be assessed with caution. Further, the penalty analysis performed demonstrated that the choice of penalty has strong influence in the results and a thoughtfully penalty value selection is required to obtain the correct results through PELT.

The change point detection problem was revisited in the third article, showing that the ARTIVA framework could successfully detect the correct change points, although, with a remarkably low detection rate. This article also showed that the use of a single change point methodology may lead to erroneous segmentation of the data, since at least one of the stations analyzed may have more than one change point. In this way, the use of Pettitt test to assess change point reliability imposed a limitation to ARTIVA results. The use of Pettitt test in the first article was required to prove PELT performance, since a previous work that used this methodology obtained poor results, being substantially outperformed by the Pettitt test. We find that further studies aimed to find the best complementary change point methods should focus on multiple change point discovery methodologies to comprise similar cases.

The combined use of change point analysis and Wavelet Transform successfully identified potential causes for the changes detected. The changes detected were associated to the phase-shift of low-frequency oscillations of both the Atlantic and Pacific (AMO and PDO) and its impacts in the SACZ. This conclusion was corroborated by the results obtained in the fourth article, the spatial clustering patterns obtained presented a parallel to displacements of

the SACZ induced directly or indirectly by low frequency oscillations. However, the influence in the change point results of anthropogenic factors that were not contemplated by the streamflow naturalization process, such as land cover changes, were not analyzed and stands as a limitation of this thesis.

The observed possible low-frequency impact led to the development of a modelling framework that directly incorporate this characteristic, accounting for the non-stationarity existent. In this way, it was included in the Gaussian Dynamic Networks discrete variables representing the different phase of the AMO or/and PDO climate indices, resulting in different network parameters according to the different phases, if they were a streamflow parent node. This characteristic of the network resulted in the improvement of the results for some stations, although, not improving the results of the stations that presented a change point. This result indicated that significant underlying changes could occur within the network structure for the periods before and after the change point, thus, leading to the development of the third article.

The third article used a different approach to account for the non-stationarity observed, obtaining a network for each homogeneous period. The network structures obtained through the methodologies HC and PCMCI showed significant changes before and after the change point, in the relationship between the variables and in its autocorrelation function. This result confirmed the hypothesis that significant changes occurred within the network structure, imposing changes to the streamflow predictors and, thus, the use of the same set of parents for the complete series may not comprise the full extension of the changes observed. Therefore, the use of different networks for each homogeneous period is recommended.

The complex lagged relationship obtained in the third article showed that these particularities need to be assessed to forecast streamflow with longer lead times. In this way, the use of ARTIVA cannot capture this complexity. Although, ARTIVA framework successfully obtained a simpler model that presented similar streamflow modelling performance with a one-month lead time, indicating its usefulness for streamflow modelling only. ARTIVA framework shows conceptual potentiality since it simultaneously addresses two complex problems, however, the version used presented critical bugs that strongly affected its useability. The development of a similar framework that could also comprise the influence of lagged relationships may present good results. However, the inclusion of this possibility would impose a significant increase in the problem dimension and certainly strong impact the algorithm run time, requiring further studies regarding its feasibility.

The observed complex lagged relationship incorporated in the network structure was one of the reasons behind the good forecasting results with a lead time of up to eight months. Another aspect that influenced in the good result was the strong autocorrelation observed within the streamflow time series. The Bayesian networks showed high flexibility to deal with these characteristics, obtaining not only simpler strictly endogenous network but also complex networks with lagged relationship, in this way, the resulting linear model can resemble different Autoregressive models such as AR, PAR and PARX. Also, the use of separate nodes for each month results in a streamflow forecast through the combination of different approaches since each node can present exogeneous and/or endogenous variable as parents. It also allows nodes without any parents, resulting in the direct use of the mean value of the month. Although presenting this extreme flexibility, the networks obtained did not show overfitting patterns, even for the simulation period, demonstrating its usefulness to streamflow forecast.

The good performance obtained for streamflow forecasting through LW simulations showed another potential use of Bayesian Networks obtained with a one-month lead time. Through this approach it was not required to obtain an independent network structure for each forecasting lead time and still obtain similar or enhanced forecasting results fixing the information available at the desired lead time and inferring the streamflow value through LW. However, the good performance was limited to networks that presented a good streamflow modelling for the base network. In this way, a robust network obtained through structure learning algorithms combined with expertise analysis presents two potential uses: (1) to forecast streamflow for higher lead times and (2) analyze the structure to extract causality.

The normalization of the streamflow series did not strong affected the results obtained, resulting in the forecasting improvement for only some stations. However, its use modified the set of parents and needs to be thoughtfully addressed before using the network to extract causality.

The results obtained through the Bayesian networks in its different configurations demonstrated its potential in many forms, however, this work presents some limitations since it did not fully analyze diverse underlying details. We find that the network structure obtained in this work have a wide margin to be improved. Firstly, the climate indices choice can be refined for each station, thoughtfully considering its climate forcing. The indices used here were chosen to contemplate general aspects and it was avoided the use of a large number of indices, which would incur in overlapping SST areas and in multiple indices to represent the same climate phenomena. Some stations can certainly yield better results with a more detailed climate

index selection. Secondly, the inclusion of occult variables would allow a more comprehensive analysis of possible spurious relations and the detection of influences that were not found within the methodologies used. The inclusion of precipitation and potential evapotranspiration may specially improve the results for the stations in the South region, since the elasticity analysis performed in Appendix D showed strong streamflow elasticity for these climate variables. Also, this work did not perform a causality analysis of the links obtained, analyzing the arc strength and other particularities. This step combined with the inclusion of the occult variables may strong affect the final networks. In this work the causality analysis was not performed since it was not its main goal and would require further studies in face of the Causal Sufficiency assumption violation.

Another limitation of this work was the lack of an uncertainty analysis regarding the streamflow forecast, we find that the inclusion of this step could enhance the results and further investigate the potential of the forecasting through the LW simulations. Another limitation is that the performance of the results obtained in this work were not directly compared to other methodologies. This latter aspect needs to be analyzed in future works to fully comprehend how the performance of the Bayesian networks as used in this work stands within current methodologies.

The final article aimed to contemplate other limitation of the Bayesian networks used. The stations were analyzed independently with its own network structure and parameters, therefore producing independent forecasts. In this article, a framework was developed to obtain the weighted complex networks between the stations, detect and analyze changes in the clustering (communities) results across time.

The combined use of MI with a threshold selection through MCMC bootstrapping presented consistent results, obtaining networks that presented coherent spatial dispersion of its communities. The use of the MI value as the weight of the links diminished the impact of weaker connections imposed by time-delimited similarities between the stations, successfully obtaining a network for the entire series that could represent the local overall behavior. The comparison of the results with the 30 years sliding windows networks discovered that significant changes occurred across-time and the use a single network can result in a miss representation of local characteristic and induce wrong conclusions.

This result demonstrated that the non-stationarity may impose a more complex approach to correct correlate the streamflow forecasts than a simpler solution adopted in

previous studies, that correlated the forecast incorporating a single correlation matrix obtained for the entire series in the forecasting model. How to incorporate the relationship changes between the stations in the forecasting models was not assessed in this work and requires further studies. However, the similarity between the temporal location of the changes detected in the streamflow series and in stations relationship indicates that the use of a model combined with a correlation matrix for each homogeneous phase could be a viable solution.

The use of a sliding window combined with comparison metrics, such as the normalized mutual information and bridge closeness, can successfully detected the different clustering configurations across time and evaluate transition zones between the different groups of stations. Although, we find that the use of the bridge closeness metric for this purpose requires further related studies to asssess its sensibility.

In summary, the main objective of this work was contemplated through all articles combined, in which each one presents a different methodologic step to detect changes in the streamflow series and incorporate the non-stationarity detected in a modelling and forecasting framework, also assessing its impact into the relationship structure between the stations.

The first specific objective was contemplated by the first article, in which three different change point methodologies were used and the reliability of the results was assessed through convergence analysis. The possible reasons behind the changes detected were also investigated through Wavelet Coherence Analysis, indicating the low-frequency influence in the changes.

The second specific objective was contemplated through the second and third article, in which two different approaches were employed to incorporate the non-stationarity in an Bayesian Network modelling framework. A first attempt was made (second article) with the direct incorporation of the low-frequency oscillations phases as discrete variables in the network. This objective was further assessed by obtaining a network for each homogeneous phase in the third article.

The third objective was also assessed in the second article with the use of two different approaches to forecast streamflow with lead times from one to eight months: obtaining one network structure for each lead time and forecasting streamflow using a one-month lead time network through LW simulations. The impact of the normalization procedures in the streamflow forecasting results was also contemplated by the second article.

The fourth objective was contemplated in the fourth article, in which a methodology based on Complex Network theory was proposed contemplating all steps necessary to detect and analyze changes in the relationship between streamflow time series. Finnally, the low-frequency impact in streamflow behavior was assessed, directly or indirectly in all four articles.

The remarks for future studies may also be synthesized into:

- Further change point studies are required to find the best complementary methods to be used simultaneously to assess the reliability of the results, especially with the use of methodologies that are able to discover multiple change points;
- The forecasting potential of a one-month lead network thoughfully developed with the inclusion of occult variables and arc strength analysis needs to be further investigated and compared to different methodologies;
- The uncertainty regarding the streamflow forecasts through LW simulations needs further investigation;
- Further studies are required to incorporate changes in the relationship structure between the stations into a streamflow forecasting framework;
- Further studies are required regarding the development of a framework that could simultaneously detect change points and the network structures, since ARTIVA alternative presents several problems in the version used in this work and cannot account for the presence of complex lagged relationships;

REFERENCES

AGARWAL, A.; MARWAN, N.; MAHESWARAN, R.; MERZ, B.; KURTHS, J. Quantifying the roles of single stations within homogeneous regions using complex network analysis. **Journal of Hydrology**, *[S. l.]*, v. 563, n. May, p. 802–810, 2018. DOI: 10.1016/j.jhydrol.2018.06.050. Disponível em: https://doi.org/10.1016/j.jhydrol.2018.06.050.

ALEXANDRE, Alan Michell Barros. **Previsão de vazões mensais para o Sistema Interligado Nacional utilizando informações climáticas**. 2012. Universidade Federal do Ceará, *[S. l.]*, 2012.

ANA. **Análise de estacionaridade de séries hidrológicas na bacia do rio São Francisco e usos consuntivos a montante da UHE Sobradinho**. *Technical Note nº 006/2013/SPR*. Brasília: National Water Agency, 2013.

ANDREOLI, Rita Valéria; KAYANO, Mary Toshie. A importância relativa do atlântico tropical sul e pacífico leste na variabilidade de precipitação do Nordeste do Brasil. **Revista Brasileira de Meteorologia**, *[S. l.]*, v. 22, n. 1, p. 63–74, 2007. DOI: 10.1590/S0102-77862007000100007. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-77862007000100007&lng=pt&tlng=pt.

BAI, Jushan; PERRON, Pierre. Computation and analysis of multiple structural change models. **Journal of Applied Econometrics**, *[S. l.]*, v. 18, n. 1, p. 1–22, 2003. DOI: 10.1002/jae.659.

BARRAT, A.; BARTHÉLEMY, M.; PASTOR-SATORRAS, R.; VESPIGNANI, A. The architecture of complex weighted networks. **Proceedings of the National Academy of Sciences of the United States of America**, *[S. l.]*, v. 101, n. 11, p. 3747–3752, 2004. DOI: 10.1073/pnas.0400087101.

BARROS, V.; GONZALEZ, M.; LIEBMANN, B.; CAMILLONI, I. Influence of the South Atlantic convergence zone and South Atlantic Sea surface temperature on interannual summer rainfall variability in Southeastern South America. **Theoretical and Applied Climatology**, *[S. l.]*, v. 67, n. 3–4, p. 123–133, 2000. DOI: 10.1007/s007040070002.

BEATRIZ, Carla; ARAÚJO, Costa De; ASSIS, Francisco De; FILHO, De Souza; MARTINS, Luiz; JÚNIOR, De Araújo; SILVEIRA, Silva. Previsão Sazonal de Vazões para a Bacia do Orós (Ceará , Brasil) Utilizando Redes Neurais e a Técnica De Reamostragem dos K-vizinhos

Seasonal Flow Forecast for the Orós Dam (Ceará , Brazil) Using Neural Networks and the Resampling Technique of K-neigh. *[S. l.]*, n. 2001, p. 197–207, 2020.

BELGHAZI, Mohamed Ishmael; BARATIN, Aristide; RAJESWAR, Sai; OZAIR, Sherjil; BENGIO, Yoshua; COURVILLE, Aaron; HJELM, R. Devon. Mutual information neural estimation. **35th International Conference on Machine Learning, ICML 2018**, *[S. l.]*, v. 2, p. 864–873, 2018.

BENJAMINI, Yoav; YOSEF, Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. **Journal of the Royal Statistical Society. Series B (Methodological)**, *[S. l.]*, v. 57, n. 1, p. 289–300, 1995. Disponível em: www.jstor.org/stable/2346101.

BLOCK, Paul J.; SOUZA FILHO, Francisco Assis; SUN, Liqiang; KWON, Hyun-Han. A Streamflow Forecasting Framework using Multiple Climate and Hydrological Models. **JAWRA Journal of the American Water Resources Association**, *[S. l.]*, v. 45, n. 4, p. 828–843, 2009. DOI: 10.1111/j.1752-1688.2009.00327.x. Disponível em: http://doi.wiley.com/10.1111/j.1752-1688.2009.00327.x.

BRADLEY, A. Allen; HABIB, Mohamed; SCHWARTZ, Stuart S. Climate index weighting of ensemble streamflow forecasts using a simple Bayesian approach. **Water Resources Research**, *[S. l.]*, v. 51, n. 9, p. 7382–7400, 2015. DOI: 10.1002/2014WR016811. Disponível em: http://doi.wiley.com/10.1002/2014WR016811.

CASTRO, Bruno Costa Alves; FILHO, Francisco de Assis de Souza; SILVEIRA, Cleiton da Silva. Análise de Tendências e Padrões de Variação das Séries Históricas de Vazões do Operador Nacional do Sistema (ONS). **Revista Brasileira de Recursos Hídricos**, *[S. l.]*, v. 18, n. 4, p. 19–34, 2013.

CATALDI, Marcio; ASSAD, Luiz Paulo de Freitas; TORRES JUNIOR, Audalio Rebelo; ALVES, José Luis Drummond. Estudo da influência das anomalias da TSM do Atlântico Sul extratropical na região da Confluência Brasil-Malvinas no regime hidrometeorológico de verão do Sul e Sudeste do Brasil. **Revista Brasileira de Meteorologia**, *[S. l.]*, v. 25, n. 4, p. 513–524, 2010. DOI: 10.1590/s0102-77862010000400010.

CAVALCANTI, Iracema Fonseca de Albuquerque et al. **Tempo e Clima no Brasil**. São Paulo: Oficina de textos, 2009.

CEPEL. Manual de Referência do Modelo NEWAVE. Centro de Pesquisa de Energia Elétrica

– CEPEL, Rio de Janeiro, 106p, 2006. Disponível em: https://simsee.org/simsee/biblioteca/Brasil/NW201203/ManualReferencia_Newave_comentado.pdf. Acessado em 06/12/2020.

CHANG, Yu Teng; LEAHY, Richard M.; PANTAZIS, Dimitrios. Modularity-based graph partitioning using conditional expected models. **Physical Review E - Statistical, Nonlinear, and Soft Matter Physics**, *[S. l.]*, v. 85, n. 1, p. 1–17, 2012. DOI: 10.1103/PhysRevE.85.016109.

CHARRAD, Malika; GHAZZALI, Nadia; BOITEAU, Véronique; NIKNAFS, Azam. NbClust : An R Package for Determining the Relevant Number of Clusters in a Data Set. **Journal of Statistical Software**, *[S. l.]*, v. 61, n. 6, p. 11744–11750, 2014. DOI: 10.18637/jss.v061.i06. Disponível em: http://www.jstatsoft.org/v61/i06/.

CHIESSI, Cristiano M.; MULITZA, Stefan; PÄTZOLD, Jürgen; WEFER, Gerold; MARENGO, José A. Possible impact of the Atlantic Multidecadal Oscillation on the South American summer monsoon. **Geophysical Research Letters**, *[S. l.]*, v. 36, n. 21, p. 1–5, 2009. DOI: 10.1029/2009GL039914.

CHU, Haibo; WEI, Jiahua; QIU, Jun. Monthly streamflow forecasting using EEMD-Lasso-DBN method based on multi-scale predictors selection. **Water (Switzerland)**, *[S. l.]*, v. 10, n. 10, 2018. DOI: 10.3390/w10101486.

CSARDI, Gabor; NEPUSZ, Tamas. The igraph software package for complex network research. **InterJournal Complex Systems**, *[S. l.]*, v. Complex Sy, n. 1695, p. 1695, 2006. Disponível em: http://igraph.sf.net.

DANON, Leon; DÍAZ-GUILERA, Albert; DUCH, Jordi; ARENAS, Alex. Comparing community structure identification. **Journal of Statistical Mechanics: Theory and Experiment**, *[S. l.]*, v. 2005, n. 09, p. P09008–P09008, 2005. DOI: 10.1088/1742-5468/2005/09/P09008. Disponível em: https://iopscience.iop.org/article/10.1088/1742-5468/2005/09/P09008.

DE SOUSA, Fabiano Berardo; ZHAO, Liang. Evaluating and comparing the IGraph community detection algorithms. **Proceedings - 2014 Brazilian Conference on Intelligent Systems, BRACIS 2014**, *[S. l.]*, p. 408–413, 2014. DOI: 10.1109/BRACIS.2014.79.

DONGES, J. F.; ZOU, Y.; MARWAN, N.; KURTHS, J. Complex networks in climate dynamics: Comparing linear and nonlinear network construction methods. **European Physical**

Journal: Special Topics, *[S. l.]*, v. 174, n. 1, p. 157–179, 2009. a. DOI: 10.1140/epjst/e2009-01098-2.

DONGES, J. F.; ZOU, Y.; MARWAN, N.; KURTHS, J. The backbone of the climate network. **Epl**, *[S. l.]*, v. 87, n. 4, 2009. b. DOI: 10.1209/0295-5075/87/48007.

DORCAS WAMBUI, Gachomo; WAITITU, Gichuhi Anthony; WANJOYA, Anthony. The Power of the Pruned Exact Linear Time(PELT) Test in Multiple Changepoint Detection. **American Journal of Theoretical and Applied Statistics**, *[S. l.]*, v. 4, n. 6, p. 581, 2015. DOI: 10.11648/j.ajtas.20150406.30. Disponível em: http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=146&doi=10.11648/j.ajtas.20150406.30.

DUTTA, Riya; MAITY, Rajib. Temporal Networks-Based Approach for Nonstationary Hydroclimatic Modeling and its Demonstration With Streamflow Prediction. **Water Resources Research**, *[S. l.]*, v. 56, n. 8, p. 1–21, 2020. DOI: 10.1029/2020WR027086.

EBERT-UPHOFF, Imme; DENG, Yi. Causal discovery for climate research using graphical models. **Journal of Climate**, *[S. l.]*, v. 25, n. 17, p. 5648–5665, 2012. DOI: 10.1175/JCLI-D-11-00387.1.

EBERT-UPHOFF, Imme; DENG, Yi. Identifying Physical Interactions from Climate Data: Challenges and Opportunities. **Computing in Science and Engineering**, *[S. l.]*, v. 17, n. 6, p. 27–34, 2015. DOI: 10.1109/MCSE.2015.129.

ENFIELD, D. B.; MESTAS-NUÑEZ, A. M.; TRIMBLE, P. J. The Atlantic Multidecadal Oscillation and its Relationship to Rainfall and River Flows in the Continental U.S.A. **Atlantic**, *[S. l.]*, v. 28, n. 10, p. 2077–2080, 2001. DOI: 10.1029/2000GL012745.

ERDMAN, C.; EMERSON, J. W. bcp: an R package for performing a Bayesian analysis of change point problems. **Journal of Statistical Software**, *[S. l.]*, v. 23, n. 3, p. 1–13, 2007. DOI: 10.18637/jss.v023.i03. Disponível em: http://core.kmi.open.ac.uk/download/pdf/6303235.pdf.

FANG, Koren; SIVAKUMAR, Bellie; WOLDEMESKEL, Fitsum M. Complex networks, community structure, and catchment classification in a large-scale river basin. **Journal of Hydrology**, *[S. l.]*, v. 545, p. 478–493, 2017. DOI: 10.1016/j.jhydrol.2016.11.056. Disponível em: http://dx.doi.org/10.1016/j.jhydrol.2016.11.056.

FERREIRA, N. J.; SANCHES, M.; SILVA DIAS, M. A. F. Composição da Zona de

Convergência do Atlântico Sul em Períodos de El Niño e La Niña. **Revista Brasileira de Meteorologia**, *[S. l.]*, v. 19, n. 1, p. 89–98, 2004.

FRITIER, Nicolas; MASSEI, Nicolas; LAIGNEL, Benoit; DURAND, Alain; DIEPPOIS, Bastien; DELOFFRE, Julien. Links between NAO fluctuations and inter-annual variability of winter-months precipitation in the Seine River watershed (north-western France). **Comptes Rendus - Geoscience**, *[S. l.]*, v. 344, n. 8, p. 396–405, 2012. DOI: 10.1016/j.crte.2012.07.004. Disponível em: http://dx.doi.org/10.1016/j.crte.2012.07.004.

FUNG, Robert; CHANG, Kuo Chu. Weighing and Integrating Evidence for Stochastic Simulation in Bayesian Networks. **Machine Intelligence and Pattern Recognition**, *[S. l.]*, v. 10, n. C, p. 209–219, 1990. DOI: 10.1016/B978-0-444-88738-2.50023-3.

GELMAN, Andrew; RUBIN, Donald B. Inference from Iterative Simulation Using Multiple Sequences. **Statistical Science**, *[S. l.]*, v. 7, n. 4, p. 457–472, 1992. DOI: 10.1214/ss/1177011136. Disponível em: http://projecteuclid.org/euclid.ss/1177010123.

GLYMOUR, Clark; CHU, Tianjiao; DANKS, David. **Data driven methods for nonlinear granger causality: Climate Teleconnection Mechanisms**, 2005.

GLYMOUR, Clark; SPIRTES, Peter. An Algorithm for Fast Recovery of Sparse Causal Graphs. **Social Science Computer Review**, *[S. l.]*, v. 9, n. 1, p. 62–72, 1991. Disponível em: https://doi.org/10.1177/089443939100900106.

HALVERSON, M. J.; FLEMING, S. W. Complex network theory, streamflow, and hydrometric monitoring system design. **Hydrology and Earth System Sciences**, *[S. l.]*, v. 19, n. 7, p. 3301–3318, 2015. DOI: 10.5194/hess-19-3301-2015.

HAN, Xudong; OUARDA, Taha B. M. J.; RAHMAN, Ataur; HADDAD, Khaled; MEHROTRA, Rajeshwar; SHARMA, Ashish. A Network Approach for Delineating Homogeneous Regions in Regional Flood Frequency Analysis. **Water Resources Research**, *[S. l.]*, v. 56, n. 3, p. 0–3, 2020. DOI: 10.1029/2019WR025910.

HAYNES, Kaylea; ECKLEY, Idris A.; FEARNHEAD, Paul. Computationally Efficient Changepoint Detection for a Range of Penalties. **Journal of Computational and Graphical Statistics**, *[S. l.]*, v. 26, n. 1, p. 134–143, 2017. DOI: 10.1080/10618600.2015.1116445. Disponível em: https://doi.org/10.1080/10618600.2015.1116445.

HINKLEY, David V. Inference About the Change-Point in a Sequence of Random Variables.

Biometrika, *[S. l.]*, v. 57, n. 1, p. 1, 1970. DOI: 10.2307/2334932. Disponível em: https://www.jstor.org/stable/2334932?origin=crossref.

HUANG, Boyin; MEHTA, Vikram M. Influences of freshwater from major rivers on global ocean circulation and temperatures in the MIT ocean general circulation model. **Advances in Atmospheric Sciences**, *[S. l.]*, v. 27, n. 3, p. 455–468, 2010. DOI: 10.1007/s00376-009-9022-6.

HUBERT, Lawrence; ARABIE, Phipps. Comparing partitions. **Journal of Classification**, *[S. l.]*, v. 2, n. 1, p. 193–218, 1985. DOI: 10.1007/BF01908075.

IVANCIC, Timothy J.; SHAW, Stephen B. Identifying spatial clustering in change points of streamflow across the contiguous U.S. between 1945 and 2009. **Geophysical Research Letters**, *[S. l.]*, v. 44, n. 5, p. 2445–2453, 2017. DOI: 10.1002/2016GL072444.

JAHFER, S.; JAHFER, S.; VINAYACHANDRAN, P. N.; VINAYACHANDRAN, P. N.; NANJUNDIAH, Ravi S.; NANJUNDIAH, Ravi S.; NANJUNDIAH, Ravi S. The role of Amazon river runoff on the multidecadal variability of the Atlantic ITCZ. **Environmental Research Letters**, *[S. l.]*, v. 15, n. 5, 2020. DOI: 10.1088/1748-9326/ab7c8a.

JAHFER, S.; VINAYACHANDRAN, P. N.; NANJUNDIAH, Ravi S. Long-Term impact of Amazon river runoff on northern hemispheric climate. **Scientific Reports**, *[S. l.]*, v. 7, n. 1, p. 1–9, 2017. DOI: 10.1038/s41598-017-10750-y.

JIMÉNEZ-RUANO, Adrián; RODRIGUES MIMBRERO, Marcos; DE LA RIVA FERNÁNDEZ, Juan. Exploring spatial–temporal dynamics of fire regime features in mainland Spain. **Natural Hazards and Earth System Sciences**, *[S. l.]*, v. 17, n. 10, p. 1697–1711, 2017. DOI: 10.5194/nhess-17-1697-2017.

JONES, Payton J.; MA, Ruofan; MCNALLY, Richard J. Bridge Centrality: A Network Approach to Understanding Comorbidity. **Multivariate Behavioral Research**, *[S. l.]*, v. 0, n. 0, p. 1–15, 2019. DOI: 10.1080/00273171.2019.1614898. Disponível em: https://doi.org/10.1080/00273171.2019.1614898.

JOO, Hongjun; LEE, Myungjin; KIM, Jongsung; JUNG, Jaewon; KWAK, Jaewon; KIM, Hung Soo. Stream gauge network grouping analysis using community detection. **Stochastic Environmental Research and Risk Assessment**, *[S. l.]*, v. 8, 2020. DOI: 10.1007/s00477-020-01916-8. Disponível em: https://doi.org/10.1007/s00477-020-01916-8.

KAYANO, Mary T.; ANDREOLI, Rita V. Relations of South American summer rainfall interannual variations with the Pacific Decadal Oscillation. **International Journal of Climatology**, *[S. l.]*, v. 27, n. 4, p. 531–540, 2007. DOI: 10.1002/joc.1417. Disponível em: http://doi.wiley.com/10.1002/joc.1417.

KAYANO, Mary T.; ANDREOLI, Rita V; SOUZA, Rodrigo A. F. De; GARCIA, Sâmia R.; CALHEIROS, Alan J. P. El Niño e La Niña dos últimos 30 anos : diferentes tipos. **Revista Climanalise**, *[S. l.]*, n. Edição Comemorativa de 30 anos do Climanálise, p. 7–12, 2016.

KAYANO, Mary Toshie; ANDREOLI, Rita Valéria; SOUZA, Rodrigo Augusto Ferreira De. El Niño–Southern Oscillation related teleconnections over South America under distinct Atlantic Multidecadal Oscillation and Pacific Interdecadal Oscillation backgrounds: La Niña. **International Journal of Climatology**, *[S. l.]*, v. 39, n. 3, p. 1359–1372, 2019. DOI: 10.1002/joc.5886.

KAYANO, Mary Toshie; CAPISTRANO, Vinicius Buscioli. How the Atlantic multidecadal oscillation (AMO) modifies the ENSO influence on the South American rainfall. **International Journal of Climatology**, *[S. l.]*, v. 34, n. 1, p. 162–178, 2014. DOI: 10.1002/joc.3674.

KELLER, M.; BUSTAMANTE, M.; GASH, J.; DIAS, P. S. **Amazonia and Global Change**. Washington, D. C.: American Geophysical Union, 2009. v. 186 DOI: 10.1029/GM186. Disponível em: http://doi.wiley.com/10.1029/GM186.

KILLICK, R.; FEARNHEAD, P.; ECKLEY, I. A. Optimal detection of changepoints with a linear computational cost. **Journal of the American Statistical Association**, *[S. l.]*, v. 107, n. 500, p. 1590–1598, 2012. DOI: 10.1080/01621459.2012.737745.

KILLICK, Rebecca; ECKLEY, Ia. changepoint: An R Package for changepoint analysis. **Lancaster University**, *[S. l.]*, v. 58, n. 3, p. 1–15, 2013. DOI: 10.1359/JBMR.0301229. Disponível em: https://www.jstatsoft.org/article/view/v058i03/v58i03.pdf.

KIM, Kyunghun; JOO, Hongjun; HAN, Daegun; KIM, Soojun; LEE, Taewoo; KIM, Hung Soo. On complex network construction of rain gauge stations considering nonlinearity of observed daily rainfall data. **Water (Switzerland)**, *[S. l.]*, v. 11, n. 8, 2019. DOI: 10.3390/w11081578.

KRASKOV, Alexander; STÖGBAUER, Harald; GRASSBERGER, Peter. Estimating mutual information. **Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics**, *[S. l.]*, v. 69, n. 6, p. 16, 2004. DOI: 10.1103/PhysRevE.69.066138.

LÈBRE, Sophie; BECQ, Jennifer; DEVAUX, Frédéric; STUMPF, Michael P. H.; LELANDAIS, Gaëlle. Statistical inference of the time-varying structure of gene-regulation networks. *[S. l.]*, 2010.

LEHNER, Flavio; WOOD, Andrew W.; LLEWELLYN, Dagmar; BLATCHFORD, Douglas B.; GOODBODY, Angus G.; PAPPENBERGER, Florian. Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest. **Geophysical Research Letters**, *[S. l.]*, v. 44, n. 24, p. 12,208-12,217, 2017. DOI: 10.1002/2017GL076043.

LI, Yujie; LIANG, Zhongmin; HU, Yiming; LI, Binquan; XU, Bin; WANG, Dong. A multi-model integration method for monthly streamflow prediction: modified stacking ensemble strategy. **Journal of Hydroinformatics**, *[S. l.]*, v. 2, p. 1–10, 2019. DOI: 10.2166/hydro.2019.066. Disponível em: https://iwaponline.com/jh/article/doi/10.2166/hydro.2019.066/70820/A-multimodel-integration-method-for-monthly.

LIMA, C. H. R.; LALL, Upmanu. Climate informed long term seasonal forecasts of hydroenergy inflow for the Brazilian hydropower system. **Journal of Hydrology**, *[S. l.]*, v. 381, n. 1–2, p. 65–75, 2010. a. DOI: 10.1016/j.jhydrol.2009.11.026. Disponível em: http://dx.doi.org/10.1016/j.jhydrol.2009.11.026.

LIMA, Carlos H. R.; LALL, Upmanu. Climate informed monthly streamflow forecasts for the Brazilian hydropower network using a periodic ridge regression model. **Journal of Hydrology**, *[S. l.]*, v. 380, n. 3–4, p. 438–449, 2010. b. DOI: 10.1016/j.jhydrol.2009.11.016. Disponível em: http://dx.doi.org/10.1016/j.jhydrol.2009.11.016.

LIMA, Carlos H. R.; LALL, Upmanu. Climate informed monthly streamflow forecasts for the Brazilian hydropower network using a periodic ridge regression model. **Journal of Hydrology**, *[S. l.]*, v. 380, n. 3–4, p. 438–449, 2010. c. DOI: 10.1016/j.jhydrol.2009.11.016. Disponível em: http://dx.doi.org/10.1016/j.jhydrol.2009.11.016.

LUO, Jiangmei; WANG, Enli; SHEN, Shuanghe; ZHENG, Hongxing; ZHANG, Yongqiang. Effects of conditional parameterization on performance of rainfall-runoff model regarding hydrologic non-stationarity. **Hydrological Processes**, *[S. l.]*, v. 26, n. 26, p. 3953–3961, 2012. DOI: 10.1002/hyp.8420.

MANTUA, Nathan J.; HARE, Steven R.; ZHANG, Yuan; WALLACE, John M.; FRANCIS, Robert C. A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production.

Bulletin of the American Meteorological Society, *[S. l.]*, v. 78, n. 6, p. 1069–1079, 1997. DOI: 10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2.

MEHDIZADEH, Saeid; KOZEKALANI SALES, Ali. A Comparative Study of Autoregressive, Autoregressive Moving Average, Gene Expression Programming and Bayesian Networks for Estimating Monthly Streamflow. **Water Resources Management**, *[S. l.]*, v. 32, n. 9, p. 3001–3022, 2018. DOI: 10.1007/s11269-018-1970-0.

MEILA, Marina. Comparing clusterings by the variation of information. *In*: **Learning Theory and Kernel Machines**. [s.l.] : Springer Berlin Heidelberg, 2003. p. 173–187.

MILLY, P. C. D.; BETANCOURT, Julio; FALKENMARK, Malin; HIRSCH, Robert M.; KUNDZEWICZ, Zbigniew W.; LETTENMAIER, Dennis P.; STOUFFER, Ronald J. Climate change: Stationarity is dead: Whither water management? **Science**, *[S. l.]*, v. 319, n. 5863, p. 573–574, 2008. DOI: 10.1126/science.1151915.

MME. Brazilian Energy Balance 2019 - Year 2018. Rio de Janeiro: Empresa de Pesquisa Energética, p. 292, 2019. Disponível em: https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-377/topico-494/BEN%202019%20Completo%20WEB.pdf. Acessado em 06/10/2020.

MONTAVON, Grégoire; SAMEK, Wojciech; MÜLLER, Klaus Robert. Methods for interpreting and understanding deep neural networks. **Digital Signal Processing: A Review Journal**, *[S. l.]*, v. 73, p. 1–15, 2018. DOI: 10.1016/j.dsp.2017.10.011. Disponível em: https://doi.org/10.1016/j.dsp.2017.10.011.

MUZA, Michel N.; CARVALHO, Leila M. V.; JONES, Charles; LIEBMANN, Brant. Intraseasonal and interannual variability of extreme dry and wet events over southeastern South America and the subtropical Atlantic during austral summer. **Journal of Climate**, *[S. l.]*, v. 22, n. 7, p. 1682–1699, 2009. DOI: 10.1175/2008JCLI2257.1.

NAGARAJAN, Radhakrishnan; SCUTARI, Marco; LÈBRE, Sophie. **Bayesian Networks in R**. [s.l: s.n.]. DOI: 10.1007/978-1-4614-6446-4.

NEWMAN, M. E. J. Analysis of weighted networks. **Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics**, *[S. l.]*, v. 70, n. 5, p. 9, 2004. DOI: 10.1103/PhysRevE.70.056131.

NEWMAN, M. E. J.; GIRVAN, M. Finding and evaluating community structure in networks.

Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, *[S. l.]*, v. 69, n. 2 2, p. 1–15, 2004. DOI: 10.1103/PhysRevE.69.026113.

NNAMCHI, Hyacinth C.; LI, Jianping; KUCHARSKI, Fred; KANG, In Sik; KEENLYSIDE, Noel S.; CHANG, Ping; FARNETI, Riccardo. An equatorial-extratropical dipole structure of the Atlantic Niño. **Journal of Climate**, *[S. l.]*, v. 29, n. 20, p. 7295–7311, 2016. DOI: 10.1175/JCLI-D-15-0894.1.

NOBRE, Paulo; SHUKLA, J. **Variations of sea surface temperature, wind stress, and rainfall over the tropical Atlantic and South AmericaJournal of Climate**, 1996. DOI: 10.1175/1520-0442(1996)009<2464:VOSSTW>2.0.CO;2.

NOWACK, Peer; RUNGE, Jakob; EYRING, Veronika; HAIGH, Joanna D. Causal networks for climate model evaluation and constrained projections. **Nature Communications**, *[S. l.]*, v. 11, n. 1, p. 1–11, 2020. DOI: 10.1038/s41467-020-15195-y. Disponível em: http://dx.doi.org/10.1038/s41467-020-15195-y.

ONS. Streamflow Series Update – 1931–2016 period (in Portuguese). *Technical Report*. Rio de Janeiro: National Electrical System Operator, 2017.

PHAN, Thuc D.; SMART, James C. R.; CAPON, Samantha J.; HADWEN, Wade L.; SAHIN, Oz. Applications of Bayesian belief networks in water resource management: A systematic review. **Environmental Modelling and Software**, *[S. l.]*, v. 85, p. 98–111, 2016. DOI: 10.1016/j.envsoft.2016.08.006. Disponível em: http://dx.doi.org/10.1016/j.envsoft.2016.08.006.

PONS, Pascal; LATAPY, Matthieu. Computing Communities in Large Networks Using Random Walks. *In*: **Download.Springer.Com**. [s.l: s.n.]. p. 284–293. DOI: 10.1007/11569596_31. Disponível em: http://link.springer.com/10.1007/11569596_31.

QIN, Minhua; LI, Dongliang; DAI, Aiguo; HUA, Wenjian; MA, Hedi. The influence of the Pacific Decadal Oscillation on North Central China precipitation during boreal autumn. **International Journal of Climatology**, *[S. l.]*, v. 38, n. January, p. e821–e831, 2018. DOI: 10.1002/joc.5410.

RAND, William M. Objective Criteria for the Evaluation of Clustering Methods. **Journal of the American Statistical Association**, *[S. l.]*, v. 66, n. December 1971, p. 37–41, 1971.

RAVINDRANATH, Arun; DEVINENI, Naresh; LALL, Upmanu; COOK, Edward R.;

PEDERSON, Greg; MARTIN, Justin; WOODHOUSE, Connie. Streamflow Reconstruction in the Upper Missouri River Basin Using a Novel Bayesian Network Model. **Water Resources Research**, *[S. l.]*, v. 55, n. 9, p. 7694–7716, 2019. DOI: 10.1029/2019WR024901.

RAYNER, N. A.; PARKER, D. E.; HORTON, E. B.; FOLLAND, C. K.; ALEXANDER, L. V.; ROWELL, D. P.; KENT, E. C.; KAPLAN, A. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. **Journal of Geophysical Research: Atmospheres**, *[S. l.]*, v. 108, n. 14, 2003. DOI: 10.1029/2002jd002670.

ROCHA, Renan Vieira; DE SOUZA FILHO, Francisco de Assis. Mapping abrupt streamflow shift in an abrupt climate shift through multiple change point methodologies: Brazil case study. **Hydrological Sciences Journal**, *[S. l.]*, p. 02626667.2020.1843657, 2020. DOI: 10.1080/02626667.2020.1843657. Disponível em: https://www.tandfonline.com/doi/full/10.1080/02626667.2020.1843657.

ROCHA, Renan Vieira; SOUZA FILHO, Francisco de Assis De; SILVA, Samiria Maria Oliveira Da. Análise da Relação entre a Precipitação Média do Reservatório Orós, Brasil - Ceará, e os Índices PDO e AMO Através da Análise de Changepoints e Transformada de Ondeletas. **Revista Brasileira de Meteorologia**, *[S. l.]*, v. 34, n. 1, p. 139–149, 2019. DOI: 10.1590/0102-77863340034.

ROLIM, Larissa Zaira Rafael; FILHO, Francisco de Assis de Souza. Shift detection in hydrological regimes and pluriannual low-frequency streamflow forecasting using the hidden Markov model. **Water (Switzerland)**, *[S. l.]*, v. 12, n. 7, 2020. DOI: 10.3390/w12072058.

RUNGE, Jakob et al. Inferring causation from time series in Earth system sciences. **Nature Communications**, *[S. l.]*, v. 10, n. 1, p. 1–13, 2019. a. DOI: 10.1038/s41467-019-10105-3. Disponível em: http://dx.doi.org/10.1038/s41467-019-10105-3.

RUNGE, Jakob; NOWACK, Peer; KRETSCHMER, Marlene; FLAXMAN, Seth; SEJDINOVIC, Dino. Detecting and quantifying causal associations in large nonlinear time series datasets. **Science Advances**, *[S. l.]*, v. 5, n. 11, 2019. b. DOI: 10.1126/sciadv.aau4996.

RYBERG, Karen R.; HODGKINS, Glenn A.; DUDLEY, Robert W. Change points in annual peak streamflows: Method comparisons and historical change points in the United States. **Journal of Hydrology**, *[S. l.]*, n. August, p. 124307, 2019. DOI: 10.1016/j.jhydrol.2019.124307. Disponível em:

https://linkinghub.elsevier.com/retrieve/pii/S002216941931042X.

SANKARASUBRAMANIAN, A.; LALL, Upmanu; SOUZA FILHO, Francisco Assis; SHARMA, Ashish. Improved water allocation utilizing probabilistic climate forecasts: Short-term water contracts in a risk management framework. **Water Resources Research**, *[S. l.]*, v. 45, n. 11, p. 1–18, 2009. DOI: 10.1029/2009WR007821.

SCARSOGLIO, Stefania; LAIO, Francesco; RIDOLFI, Luca. Climate dynamics: A network-based approach for the analysis of global precipitation. **PLoS ONE**, *[S. l.]*, v. 8, n. 8, 2013. DOI: 10.1371/journal.pone.0071129.

SCHEUERELL, Mark D. muti - An R package for computing mutual information Using muti Data discretization muti computes. *[S. l.]*, 2017. DOI: 10.5281/zenodo.439391.

SCUTARI, M. Learning Bayesian Networks with the bnlearn R Package. **Journal of Statistical Software**, *[S. l.]*, v. 35, n. 3, p. 1–22, 2010. DOI: 10.18637/jss.v035.i03.

SCUTARI, Marco; GRAAFLAND, Catharina Elisabeth; GUTIÉRREZ, José Manuel. Who learns better Bayesian network structures: Accuracy and speed of structure learning algorithms. **International Journal of Approximate Reasoning**, *[S. l.]*, v. 115, p. 235–253, 2019. DOI: 10.1016/j.ijar.2019.10.003.

SCUTARI, Marco; NAGARAJAN, Radhakrishnan. Identifying significant edges in graphical models of molecular networks. **Artificial Intelligence in Medicine**, *[S. l.]*, v. 57, n. 3, p. 207–217, 2013. DOI: 10.1016/j.artmed.2012.12.006. Disponível em: http://dx.doi.org/10.1016/j.artmed.2012.12.006.

SHAPIRO, S. S.; WILK, M. B. An Analysis of Variance Test for Normality (Complete Samples). **Biometrika**, *[S. l.]*, v. 52, n. 3/4, p. 591, 1965. DOI: 10.2307/2333709. Disponível em: https://www.jstor.org/stable/2333709?origin=crossref.

SILVEIRA, Cleiton da Silva; ALEXANDRE, Alan Michell Barros; SOUZA FILHO, Francisco de Assis De; VASCONCELOS JUNIOR, Francisco das Chagas; CABRAL, Samuellson Lopes. Monthly streamflow forecast for National Interconnected System (NIS) using Periodic Auto-regressive Endogenous Models (PAR) and Exogenous (PARX) with climate information. **Rbrh**, *[S. l.]*, v. 22, n. 0, 2017. DOI: 10.1590/2318-0331.011715186.

SIVAKUMAR, B.; WOLDEMESKEL, F. M. Complex networks for streamflow dynamics. **Hydrology and Earth System Sciences**, *[S. l.]*, v. 18, n. 11, p. 4565–4578, 2014. DOI:

10.5194/hess-18-4565-2014.

SIVAKUMAR, Bellie. **Chaos in Hydrology**. Dordrecht: Springer Netherlands, 2017. DOI: 10.1007/978-90-481-2552-4. Disponível em: http://link.springer.com/10.1007/978-90-481-2552-4.

SOUZA FILHO, Francisco Assis; LALL, Upmanu. Seasonal to interannual ensemble streamflow forecasts for Ceara, Brazil: Applications of a multivariate, semiparametric algorithm. **Water Resources Research**, *[S. l.]*, v. 39, n. 11, p. n/a-n/a, 2003. DOI: 10.1029/2002WR001373. Disponível em: http://doi.wiley.com/10.1029/2002WR001373.

TAMADDUN, Kazi Ali; KALRA, Ajay; AHMAD, Sajjad. Wavelet analyses of western us streamflow with ENSO and PDO. **Journal of Water and Climate Change**, *[S. l.]*, v. 8, n. 1, p. 26–39, 2017. DOI: 10.2166/wcc.2016.162.

TAMADDUN, Kazi Ali; KALRA, Ajay; AHMAD, Sajjad. Spatiotemporal Variation in the Continental US Streamflow in Association with Large-Scale Climate Signals Across Multiple Spectral Bands. **Water Resources Management**, *[S. l.]*, v. 33, n. 6, p. 1947–1968, 2019. DOI: 10.1007/s11269-019-02217-8.

TAMADDUN, Kazi Ali; KALRA, Ajay; BERNARDEZ, Miguel; AHMAD, Sajjad. Multi-Scale Correlation between the Western U.S. Snow Water Equivalent and ENSO/PDO Using Wavelet Analyses. **Water Resources Management**, *[S. l.]*, v. 31, n. 9, p. 2745–2759, 2017. DOI: 10.1007/s11269-017-1659-9.

TAMADDUN, Kazi; KALRA, Ajay; AHMAD, Sajjad. Identification of streamflow changes across the continental United States using variable record lengths. **Hydrology**, *[S. l.]*, v. 3, n. 2, 2016. DOI: 10.3390/hydrology3020024.

TANG, Chunling; CHEN, Dong; CROSBY, Benjamin T.; PIECHOTA, Thomas C.; WHEATON, Joseph M. Is the PDO or AMO the climate driver of soil moisture in the Salmon River Basin, Idaho? **Global and Planetary Change**, *[S. l.]*, v. 120, p. 16–23, 2014. DOI: 10.1016/j.gloplacha.2014.05.008. Disponível em: http://dx.doi.org/10.1016/j.gloplacha.2014.05.008.

TONGAL, Hakan. Spatiotemporal analysis of precipitation and extreme indices in the Antalya Basin, Turkey. **Theoretical and Applied Climatology**, *[S. l.]*, v. 138, n. 3–4, p. 1735–1754, 2019. DOI: 10.1007/s00704-019-02927-4.

TORRENCE, Christopher; COMPO, Gilbert P. A Practical Guide to Wavelet Analysis. **Bulletin of the American Meteorological Society**, *[S. l.]*, v. 79, n. 1, p. 61–78, 1998. DOI: 10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

TORRENCE, Christopher; WEBSTER, Peter J. Interdecadal changes in the ENSO-monsoon system. **Journal of Climate**, *[S. l.]*, v. 12, n. 8 PART 2, p. 2679–2690, 1999. DOI: 10.1175/1520-0442(1999)012<2679:ICITEM>2.0.CO;2.

VAN DONGEN, S. Performance criteria for graph clustering and Markov cluster experiments. **Technical Report INS-R0012, National Research Institute for Mathematics and Computer Science**, *[S. l.]*, p. 36, 2000. Disponível em: http://dl.acm.org/citation.cfm?id=868979.

VICENTE-SERRANO, Sergio M.; LÓPEZ-MORENO, Juan I.; BEGUERÍA, Santiago; LORENZO-LACRUZ, Jorge; AZORIN-MOLINA, Cesar; MORÁN-TEJEDA, Enrique. Accurate Computation of a Streamflow Drought Index. **Journal of Hydrologic Engineering**, *[S. l.]*, v. 17, n. 2, p. 318–332, 2012. DOI: 10.1061/(ASCE)HE.1943-5584.0000433.

VIZY, Edward K.; COOK, Kerry H. Influence of the Amazon/Orinoco Plume on the summertime Atlantic climate. **Journal of Geophysical Research Atmospheres**, *[S. l.]*, v. 115, n. 21, p. 1–18, 2010. DOI: 10.1029/2010JD014049.

WANG, Yixuan; ZHANG, Ting; CHEN, Xu; LI, Jianzhu; FENG, Ping. Spatial and temporal characteristics of droughts in Luanhe River basin, China. **Theoretical and Applied Climatology**, *[S. l.]*, v. 131, n. 3–4, p. 1369–1385, 2018. DOI: 10.1007/s00704-017-2059-z.

YANG, Tiantian; ASANJAN, Ata Akbari; WELLES, Edwin; GAO, Xiaogang; SOROOSHIAN, Soroosh; LIU, Xiaomang. Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information. **Water Resources Research**, *[S. l.]*, v. 53, n. 4, p. 2786–2812, 2017. DOI: 10.1002/2017WR020482. Disponível em: http://doi.wiley.com/10.1002/2017WR020482.

ZEBIAK, Stephen E. Air–Sea Interaction in the Equatorial Atlantic Region. **Journal of Climate**, *[S. l.]*, v. 6, n. 8, p. 1567–1586, 1993. DOI: 10.1175/1520-0442(1993)006<1567:AIITEA>2.0.CO;2. Disponível em: http://journals.ametsoc.org/doi/10.1175/1520-0442(1993)006%3C1567:AIITEA%3E2.0.CO;2.

ZEILEIS, Achim; KLEIBER, Christian; WALTER, Krämer; HORNIK, Kurt. Testing and

dating of structural changes in practice. **Computational Statistics and Data Analysis**, *[S. l.]*, v. 44, n. 1–2, p. 109–123, 2003. DOI: 10.1016/S0167-9473(03)00030-6.

ZEILEIS, Achim; LEISCH, Friedrich; HOMIK, Kurt; KLEIBER, Christian. strucchange: An R Package for Testing for Structural Change. **Journal of Statistical Software**, *[S. l.]*, v. 7, n. 2, p. 1–38, 2002.

ZHU, Yuxiang; JIANG, Jianmin; HUANG, Changxing; CHEN, Yongqin David; ZHANG, Qiang. Applications of multiscale change point detections to monthly stream flow and rainfall in Xijiang River in southern China, part I: correlation and variance. **Theoretical and Applied Climatology**, *[S. l.]*, v. 136, n. 1–2, p. 237–248, 2019. DOI: 10.1007/s00704-018-2480-y. Disponível em: http://link.springer.com/10.1007/s00704-018-2480-y.

Appendix A – Supplemental files of chapter 5

This section presents the additional figures mentioned in the main article and the description of the interactive plots provided separately[6].

As stated in the main article, the data was split in two periods: calibration and simulation. We adopted a single continuous time window corresponding to 70% of the data for the calibration period and the remaining 30% of the data, continuous or not, for the simulation period. Figure A1 shows the number of years for each combination of phases of AMO and PDO, according to the initial year of the calibration period. The initial year was chosen aiming an equal amount of years for the two periods. We noticed that the temporal disposition of the cold phase of the AMO made impossible the verification of the forecasting performance corresponding to this phase in the simulation period. The figure shows that for the streamflow node that had both AMO and PDO as parents, the network parameters for each phase combination were estimated from a reduced number of data. However, the concomitant use of both indices as parents was not common, occurring only for a reduced number of stations.

Figures A2 and A3 are interactive plots of the Figure 9 presented in the main article in the html format and are supplementary files of this thesis. We decided to provide these plots due to the high number of different network typologies found. Figure A2 allows for the selection of a reduced number of network typologies and zooming in for a specific region, also providing information of each station analyzed when hoovering the mouse along the stations. Figure A3 eases the visualization of the parents of each network typology, highlighting the links with mouse clicks. It is important to emphasize that this result corresponds to the networks obtained using the non-normalized streamflow and using the AMO and PDO phases as discrete nodes. Therefore, as concluded in the article, the different network configurations tested can provide a different set of parents.

Figures A4 to A7 shows the plot of the series of streamflow forecasted for the key stations. These plots reinforce the general good results obtained through the methodology used and enables the analysis of specific aspects. For Furnas station the streamflow forecasted presented a tendency of underestimating the major peaks, but also overestimated the streamflow

[6] Download link: https://drive.google.com/file/d/1RzbBHcNlu_O7CkfALehzH0U822AaFh4t/view?usp=sharing

for a low-flow period between 2013 and 2016. This discrepancy in the low-flow period was also expanded with the increase of the lead time. Sobradinho results presented a tendency of overestimation only for the flows of the period from 1998 to 2016, this behavior was also expanded with the increase of lead time. This result agrees with the change point found in the mean value of the streamflow series for the year of 1994, which imposed a reduction in the mean value. It also reinforces the conclusion that the use of discrete nodes representing the low-frequency SST phases is not an alternative to contemplate the heterogeneity characteristic of the series, despite the occurrence of phase shifts close to the change point location. Itaipu station showed a slight overestimation pattern for the latest years. Tucuruí station presented a steady periodic behavior, thus, the streamflow forecast for this station greatly benefited from the use of a node for each month. The increase of the discrepancies with the increase of the lead time is expected due to the increase in the uncertainty involved and also corroborates the conclusion that the forecast through Likelihood Weighting produces good results when the base network is relative successful in representing the streamflow with a lead time of one month.

Table A1 presents the information of all stations used in this work. The station name can be used to acquire the daily streamflow data from the webservice of ONS (http://aplicam.ons.org.br/hidrologia/Reservatorio.asmx). Data for the indices used is available from the official NOAA database (https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/ and https://www.ncdc.noaa.gov/data-access/marineocean-data/extended-reconstructed-sea-surface-temperature-ersst-v4).

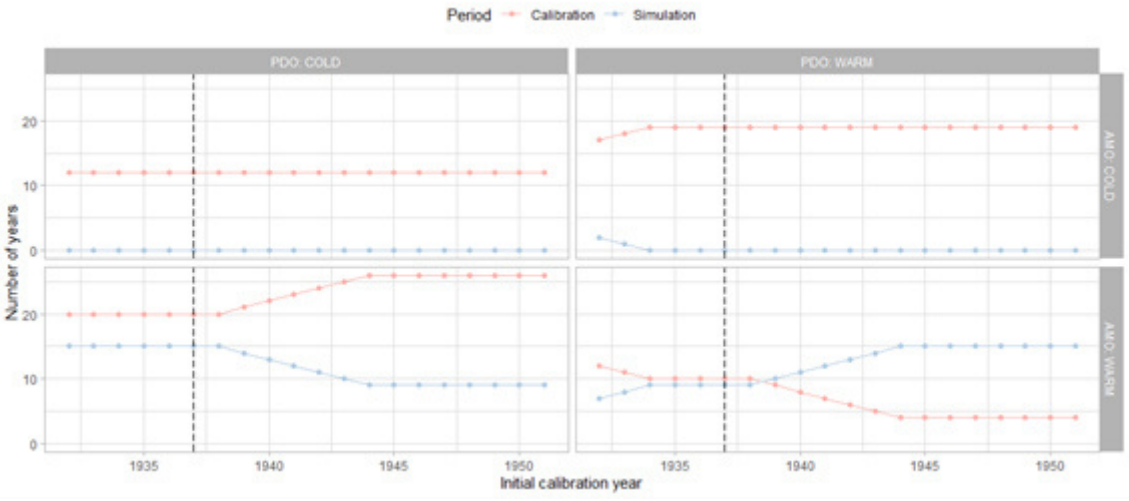


Figure A1. Number of years for each AMO and PDO phase combinations for the calibration and simulation periods according to the initial year of the calibration period. The dotted line represents the initial year selected for the calibration period.

Figure A2. Interactive plot of the spatial dispersion of the network typologies found clustered by similarity (Figure 9 in main article). This result refers to the networks obtained using the original non-normalized streamflow and with lead time of one month. The tooltip shows information of each analyzed station (station name, number, and coordinates). A mouse click in the elements of the caption can be used to display only the desired clusters. The plot also allows zooming. This interactive plot was made using the R packages *ggplot2* and *plotly*.

Figure A3. Interactive plot of the network typologies. This result refers to the networks obtained using the original non-normalized streamflow and with lead time of one month. A mouse click in an node highlights its links, easing the visualization of its parents. The arrow pointing in and out from the same node refers to an autoregressive component. An arc between the index and the station is produced if at least one month has the index as a parent. This interactive plot was made using the R package *visNetwork*.

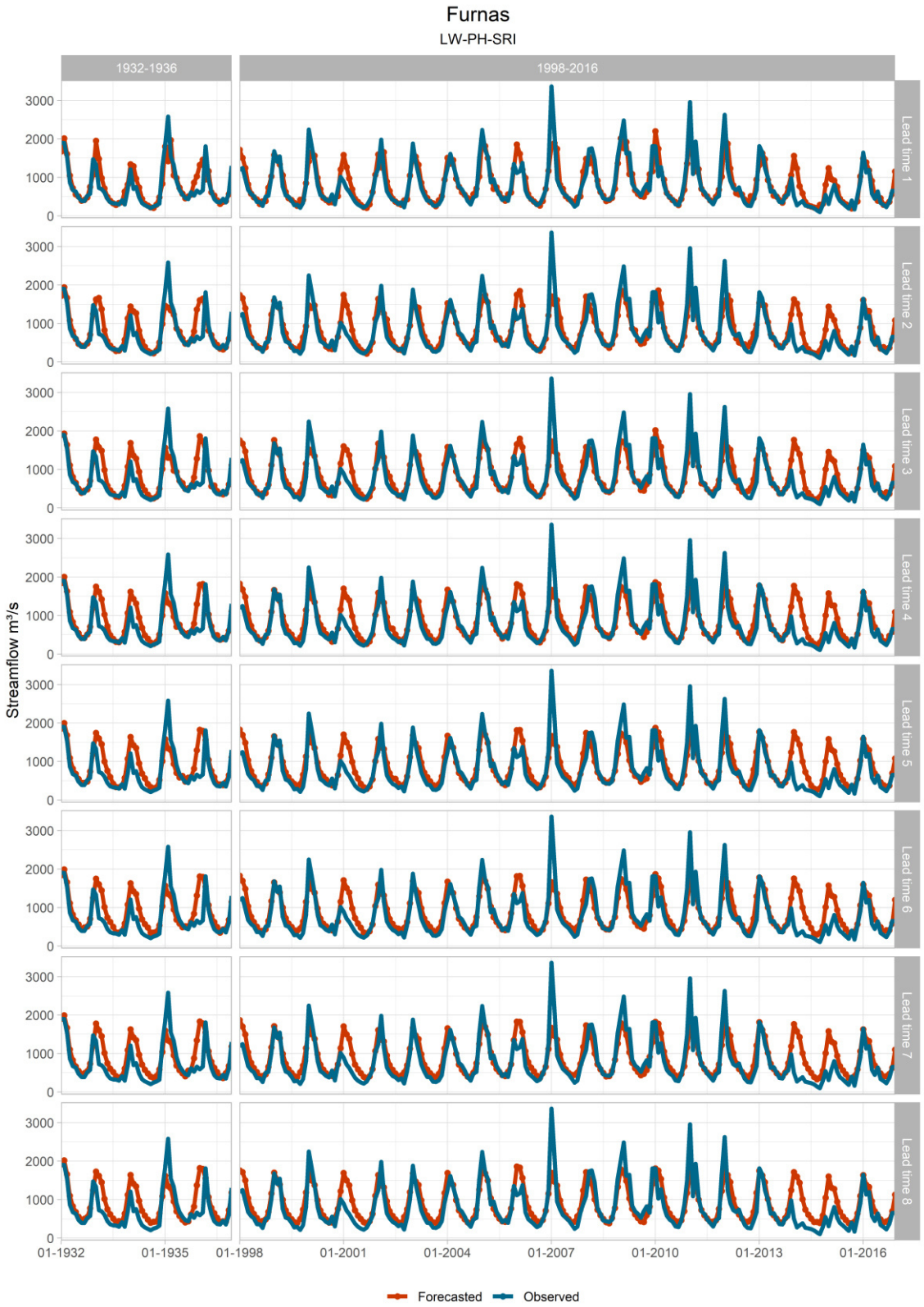


Figure A4. Forecasted and observed streamflow for Furnas station. The subtitle indicates the methodology used to obtain the results presented.

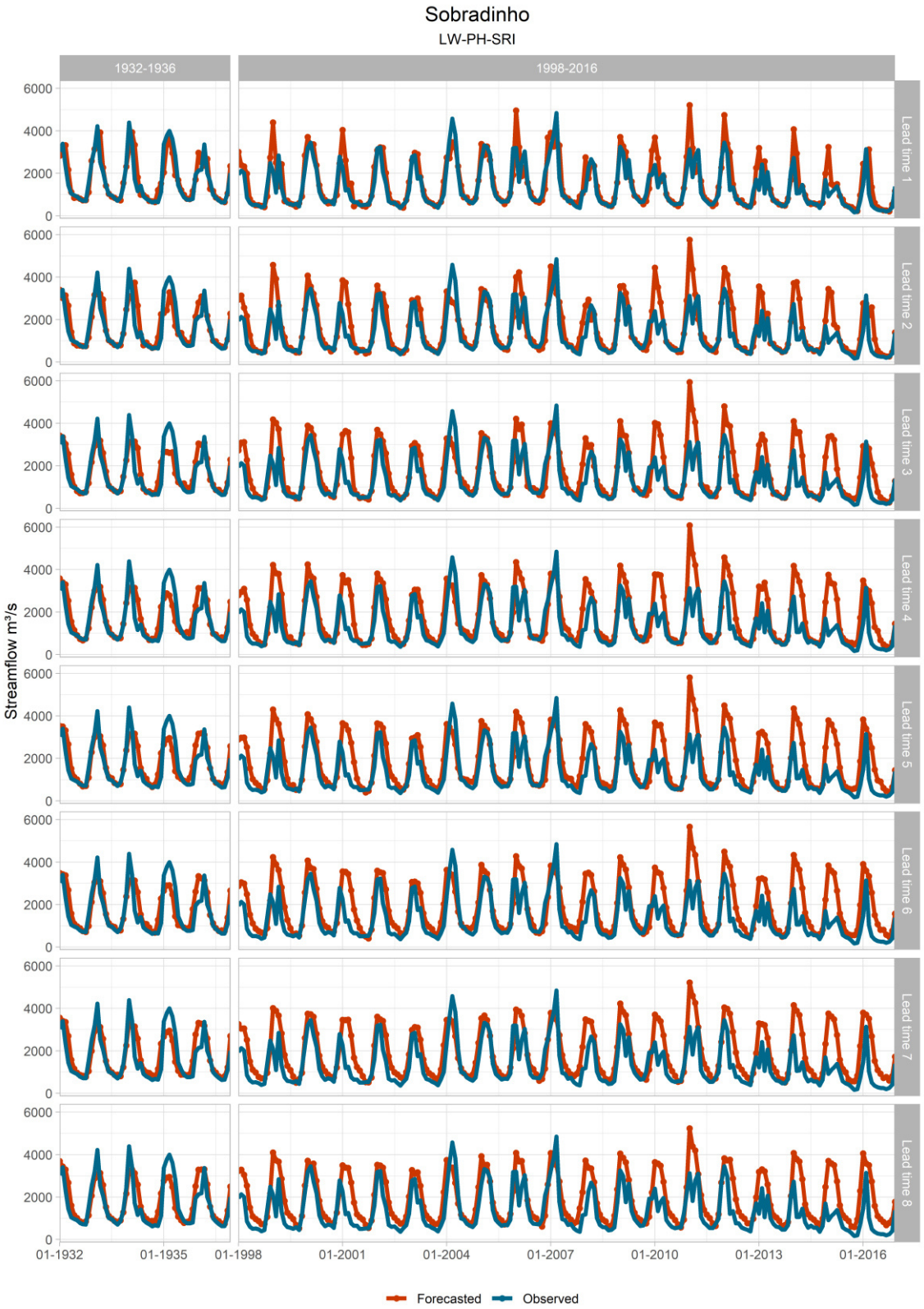


Figure A5. Forecasted and observed streamflow for Sobradinho station. The subtitle indicates the methodology used to obtain the results presented.

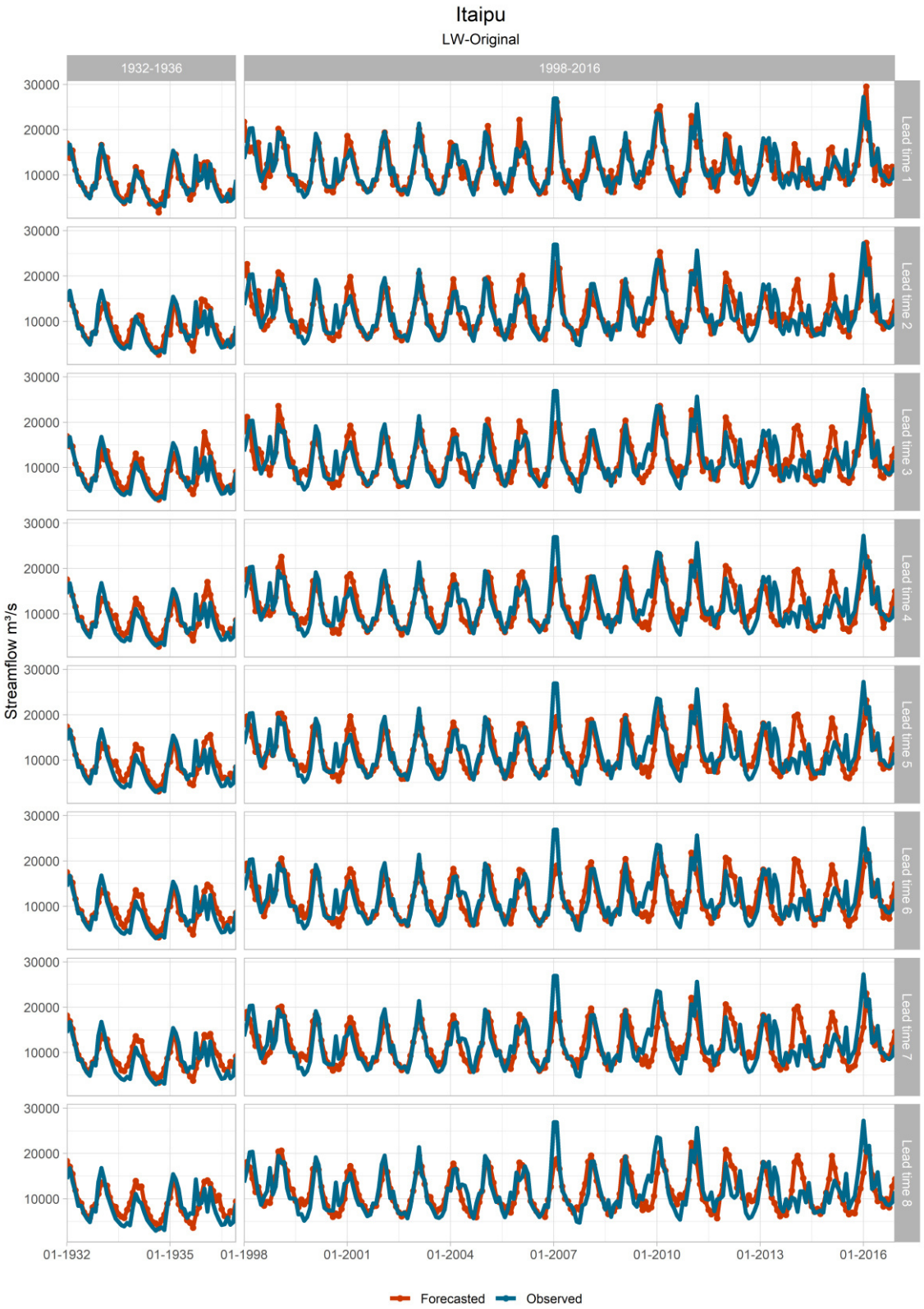


Figure A6. Forecasted and observed streamflow for Itaipu station. The subtitle indicates the methodology used to obtain the results presented.

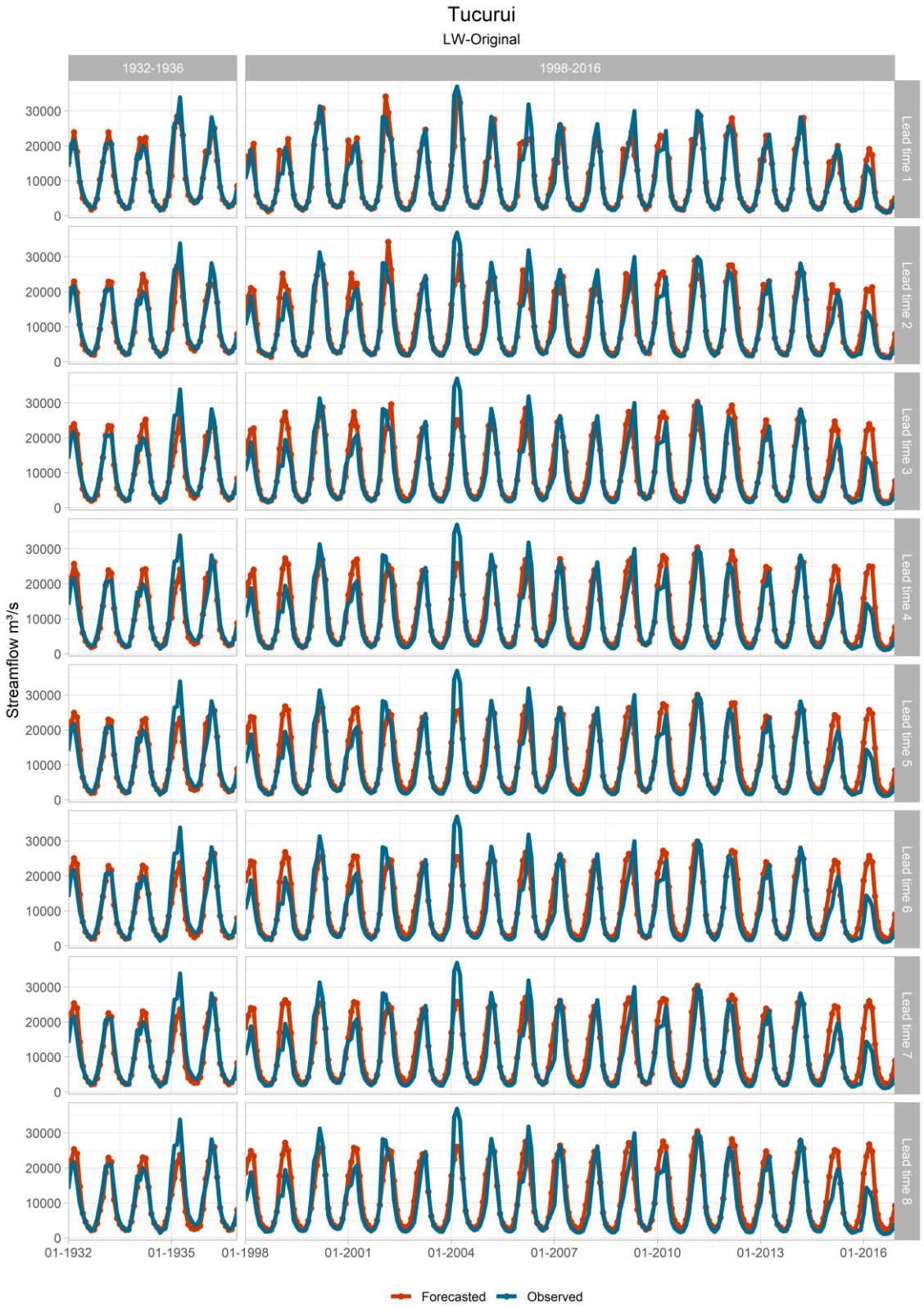


Figure A7. Forecasted and observed streamflow for Tucuruí station. The subtitle indicates the methodology used to obtain the results presented.

Station Number	Station Name	x(lon)	y(lat)	Station Number	Station Name	x(lon)	y(lat)
1	CAMARGOS	-44.6161	-21.3256	164	E.SOUZA	-46.9101	-23.4547
6	FURNAS	-46.3181	-20.6690	168	SOBRADINHO F	-40.8278	-9.4314
14	CACONDE	-46.6242	-21.5767	188	ITAPEBI	-39.5818	-15.9657
17	MARIMBONDO	-49.1953	-20.3034	190	B. ESPERANCA	-43.5673	-6.7502
18	A. VERMELHA	-50.3473	-19.8657	191	CANA BRAVA	-48.1428	-13.4017
24	EMBORCACAO	-47.9847	-18.4529	196	ROSAL	-41.7218	-20.9178
25	NOVA PONTE	-47.6966	-19.1295	197	PICADA	-43.5410	-21.8848
31	ITUMBIARA	-49.0991	-18.4084	201	TOCOS	-44.1253	-22.7491
32	CACH.DOURADA	-49.4931	-18.5025	205	CORUMBA IV	-48.1868	-16.3217
33	SAO SIMAO	-50.4983	-19.0189	206	MIRANDA	-48.0396	-18.9107
34	I.SOLTEIRA	-51.3636	-20.3822	209	CORUMBA I	-48.5301	-17.9890
47	A.A.LAYDNER (JURUMIRIM)	-49.0028	-23.2692	211	FUNIL-GRANDE	-45.0370	-21.1429
61	CAPIVARA	-51.3590	-22.6570	215	BARRA GRANDE	-51.1933	-27.7769
63	ROSANA	-52.8647	-22.5962	216	CAMPOS NOVOS	-51.3326	-27.6011
71	STA CLARA PR	-51.9578	-25.6456	220	MONJOLINHO	-52.7301	-27.3462
72	FUNDAO	-52.0021	-25.7034	237	BARRA BONITA	-48.5344	-22.5196
73	JORDAO	-52.0881	-25.7572	240	PROMISSAO	-49.7831	-21.2961
74	G.B.MUNHOZ	-51.6644	-26.0139	242	NAVANHANDAVA	-50.2011	-21.1189
76	SEGREDO	-52.1139	-25.7900	243	T.IRMAOS	-51.3000	-20.6686
77	SLT.SANTIAGO	-52.6143	-25.6281	245	JUPIA	-51.6311	-20.7793
78	SALTO OSORIO	-53.0093	-25.5380	246	P.PRIMAVERA	-52.9574	-22.4787
92	ITA	-52.3822	-27.2767	247	CACU	-51.1500	-18.5294
93	PASSO FUNDO	-52.7333	-27.5500	251	SERRA FACAO	-47.6750	-18.0459
94	FOZ CHAPECO	-53.0397	-27.1417	253	SAO SALVADOR	-48.2373	-12.8077
98	CASTRO ALVES	-51.3839	-29.0052	254	P.CAVALO	-38.9986	-12.5852
99	ESPORA	-51.8656	-18.6758	255	IRAPE	-42.5747	-16.7381
101	SALTO PILAO	-49.5179	-27.1361	257	PEIXE ANGIC	-48.3865	-12.2337
102	SAO JOSE	-54.8158	-28.1764	259	ITIQUIRA I	-54.8310	-17.0906
111	PASSO REAL	-53.1886	-29.0162	266	ITAIPU	-54.5990	-25.4140
115	G.P.SOUZA	-48.8738	-25.1387	270	SERRA MESA	-48.3049	-13.8333
117	GUARAPIRANGA	-46.7258	-23.6716	271	ESTREITO TOC	-47.4607	-6.5879
119	BILLINGS_PED	-46.6745	-23.7030	273	LAJEADO	-48.3729	-9.7568
120	JAGUARI	-46.0278	-23.1950	275	TUCURUI	-49.6518	-3.8326
121	PARAIBUNA	-45.5983	-23.4103	277	CURUA-UNA	-54.2989	-2.8117
125	STA CECILIA	-43.8391	-22.4826	278	MANSO	-55.7847	-14.8712
130	I. POMBOS	-42.5795	-21.8437	279	SAMUEL	-63.4549	-8.7517
134	SALTO GRANDE	-42.7188	-19.1159	281	PONTE PEDRA	-54.8245	-17.6094
144	MASCARENHAS	-40.9185	-19.5005	283	STA CLARA MG	-51.9578	-25.6456
145	RONDON II	-60.6972	-11.9975	286	QUEBRA QUEIX	-52.5452	-26.6563
149	CANDONGA	-42.8553	-20.2086	287	STO ANTONIO	-63.9524	-8.7966
155	RETIRO BAIXO	-44.7811	-18.8766	291	DARDANELOS	-59.4642	-10.1633
156	TRES MARIAS	-45.2592	-18.2150	294	SALTO	-51.1707	-18.8079
158	QUEIMADO	-47.3233	-16.2106	295	JAURU	-58.7290	-15.2369
160	ALTO TIETÊ	-46.9103	-23.4541	296	GUAPORE	-58.9647	-15.1239

Table A1. Stations information.

Appendix B – Supplemental files of chapter 6

The ARTIVA R package performs its proposed framework through the functions ARTIVAsubnet or ARTIVAnet. The first runs the framework for a "target gene", which in our case refers to a station, where the latter just performs the framework independently for a group of stations. These functions can be run in two ways, for a delimited set of iterations (default) or searching for the convergence of the chains according to the RGELMAN statistics, referred as "PRSF_thres" in the ARTIVA function.

When searching for the convergence the package shows two critical bugs that ends the framework with an error output. First, when the convergence is not reached the algorithm produces an internal error related to the PSRF_CP variable (Figure B1). Second, when the convergence is reached a general variable dimension error is produced (Figure B2). Both errors were not caused by the arguments of the function, since several tests were made, and the output errors remained unchanged. Also, the function runs correctly when using the same arguments but disabling the search for the convergence of the chains (PSRFactor = FALSE).

Therefore, it was not possible to directly obtain the results that produced the convergence of the chains due to the referred bugs, instead the results refer to the algorithm run with the maximum number of iterations (200,000). The explicit assumption here is that for the stations that presented convergence with far less iterations the maximum iteration run would reproduce similar results since the convergence would already occur along the iterations. For the stations that do not converge the reliability of the results could not be assessed and should be treated with caution.

The Table B1 shows details regarding the runtime and the convergence of the stations. The complete parallel computation presented a runtime of 1 day and 15.5 hours. The computer specification was an Intel i7-770 (3.60 GHz), 8 GB RAM and windows 10 was the operational system.

The bugs in the ARTIVA package and the lack of corrections and updates resulted in its removal from the official R repository (CRAN) in July of 2020 (Figure B3), after this work was initiated. The version used was the 1.2.3.

```
[1] "--- !!! WARNING MESSAGE !!! ----"
[1] "Plots will not be saved (see the parameter savePictures)"
[1] "---------------------------------"
[1] "--- !!! WARNING MESSAGE !!! ----"
[1] "Estimation values will not be saved (see the parameter saveEstimations)"
[1] "---------------------------------"
[1] "========================"
[1] "GENERAL INFORMATION"
[1] " *** Name of the analyzed target gene: q286"
[1] " *** Number of different time point measurements: 1020"
[1] " *** Number of repeated measurements: 1"
[1] " *** Number of potential parent gene(s): 5"
[1] "========================"
[1] "ARTIVA PARAMETERS"
[1] " **** Time delay considered in the auto-regressive process: 1"
[1] " **** Minimal length to define a temporal segment: 2 time points"
[1] " **** Maximal number of changepoints (CPs): 1"
[1] " **** Number of CPs at the algorithm initialization: 0"
[1] " **** Number of iterations: 2e+05"
[1] " **** Is PSRF factor calculated? TRUE"
[1] "========================"
[1] "STEP 1: Starting the initialisation procedure"
[1] "STEP 2: Starting the ARTIVA RJ-MCMC procedure"
[1] "Looking for convergence (up to 2e+05 iterations) :"
10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
[1] "Warning convergence was not reached"
[1] "ERROR: objeto 'PSRF_CP' não encontrado"
Time difference of 205.8679 mins
```

Figure B1. First internal error.

```
[1] "--- !!! WARNING MESSAGE !!! ----"
[1] "Plots will not be saved (see the parameter savePictures)"
[1] "---------------------------------"
[1] "--- !!! WARNING MESSAGE !!! ----"
[1] "Estimation values will not be saved (see the parameter saveEstimations)"
[1] "---------------------------------"
[1] "========================"
[1] "GENERAL INFORMATION"
[1] " *** Name of the analyzed target gene: q47"
[1] " *** Number of different time point measurements: 1020"
[1] " *** Number of repeated measurements: 1"
[1] " *** Number of potential parent gene(s): 5"
[1] "========================"
[1] "ARTIVA PARAMETERS"
[1] " **** Time delay considered in the auto-regressive process: 1"
[1] " **** Minimal length to define a temporal segment: 2 time points"
[1] " **** Maximal number of changepoints (CPs): 1"
[1] " **** Number of CPs at the algorithm initialization: 0"
[1] " **** Number of iterations: 2e+05"
[1] " **** Is PSRF factor calculated? TRUE"
[1] "========================"
[1] "STEP 1: Starting the initialisation procedure"
[1] "STEP 2: Starting the ARTIVA RJ-MCMC procedure"
[1] "Looking for convergence (up to 2e+05 iterations) :"
10% 20% 30%
[1] "**** Convergence obtained (after 69000 iterations, PSRF threshold = 1.1) ***"
[1] "RJMCMC iterations for Postdistribution evaluation (50000 iterations)."

[1] "STEP 3: Computing posterior distributions for the changepoints (CPs) number and location"
[1] "ERROR: dim(X) must have a positive length"
Time difference of 213.0203 mins
```

Figure B2. Second internal error.

Package 'ARTIVA' was removed from the CRAN repository.

Formerly available versions can be obtained from the archive.

Archived on 2020-07-05 as email bounced when asked for changes.

A summary of the most recent check results can be obtained from the check results archive.

Please use the canonical form https://CRAN.R-project.org/package=ARTIVA to link to this page.

Figure B3. ARTIVA package removal from CRAN. Print screen taken in 29/01/2021 from the URL https://cran.r-project.org/web/packages/ARTIVA/index.html.

ARTIVA – 1 Changepoint – RGELMAN = 1.1			
Station	Runtime (mins)	Total iterations	Convergence
47	213,0203	69.000	yes
61	184,1128	51.000	yes
63	183,8763	51.000	yes
73	183,9484	51.000	yes
74	182,1699	51.000	yes
76	355,0056	200.000	no
77	355,2802	200.000	no
78	254,2657	200.000	no
99	97,0817	51.000	yes
101	291,0799	200.000	no
102	153,1521	58.000	yes
120	95,6425	51.000	yes
134	119,2241	51.000	yes
144	130,6751	51.000	yes
158	116,6662	51.000	yes
168	126,3162	51.000	yes
188	136,6115	51.000	yes
237	141,2705	51.000	yes
240	145,0770	51.000	yes
242	141,4768	51.000	yes
243	141,1801	51.000	yes
245	84,4074	51.000	yes
246	105,7550	51.000	yes
266	107,4870	51.000	yes
281	67,7273	51.000	yes
286	205,8679	200.000	no

Table B1. Runtime and convergence of ARTIVA parallel runs for each station.

The following figures stands for the networks and time-series plot mentioned in chapter 5.

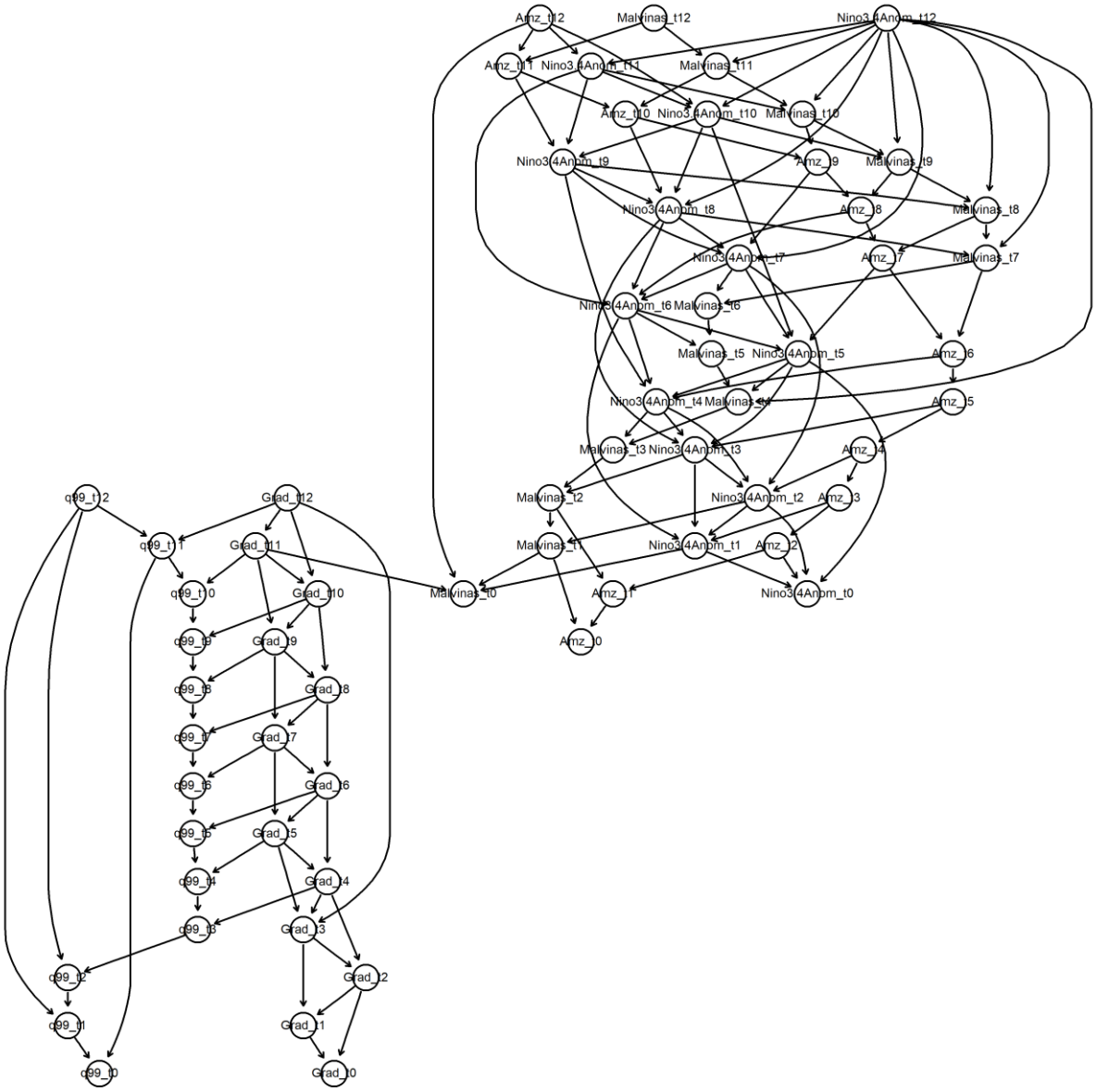


Figure B4. HC network Structure of station 99 (second period)

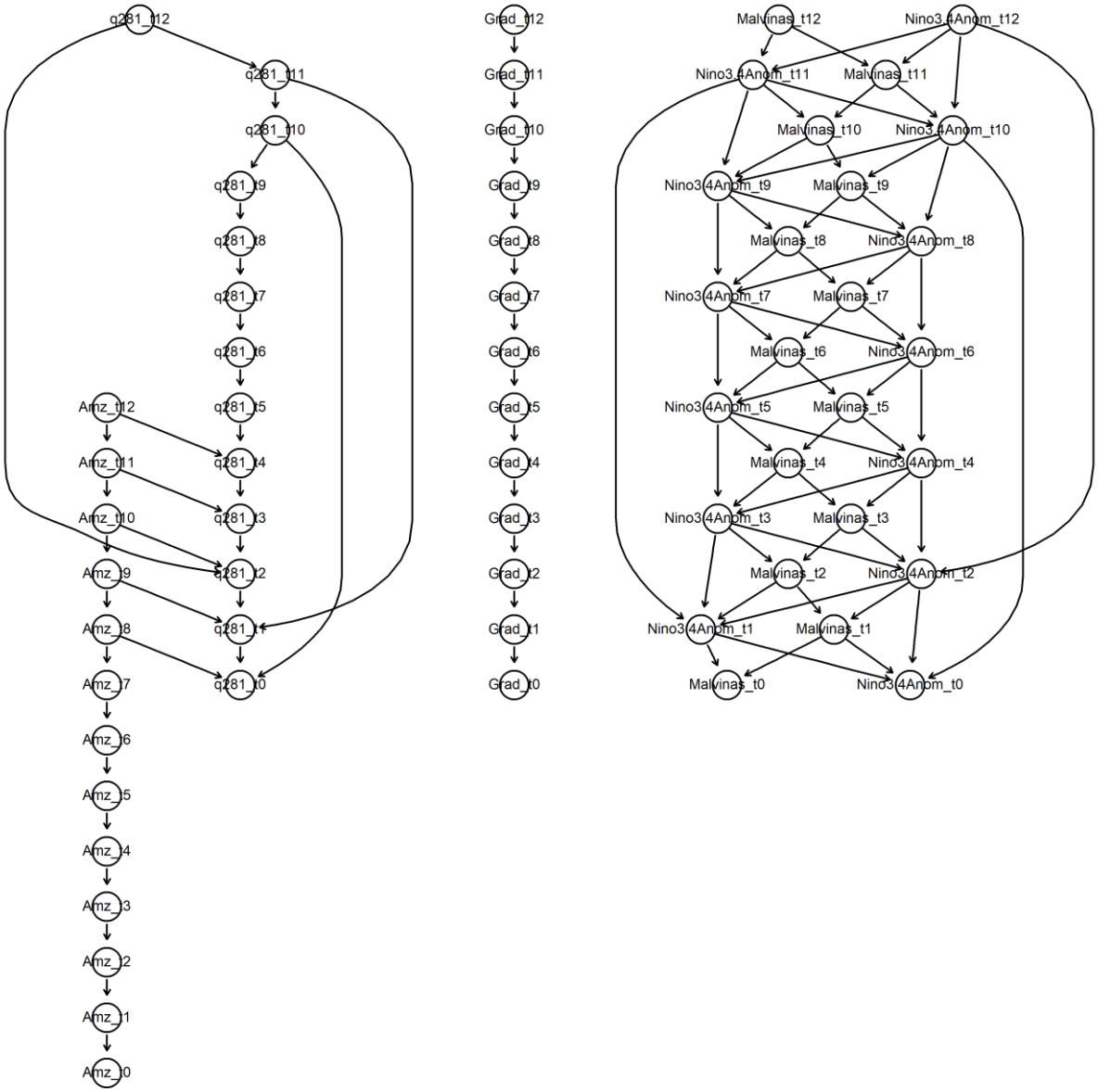


Figure B5. PCMCI network Structure of station 281 (first period)

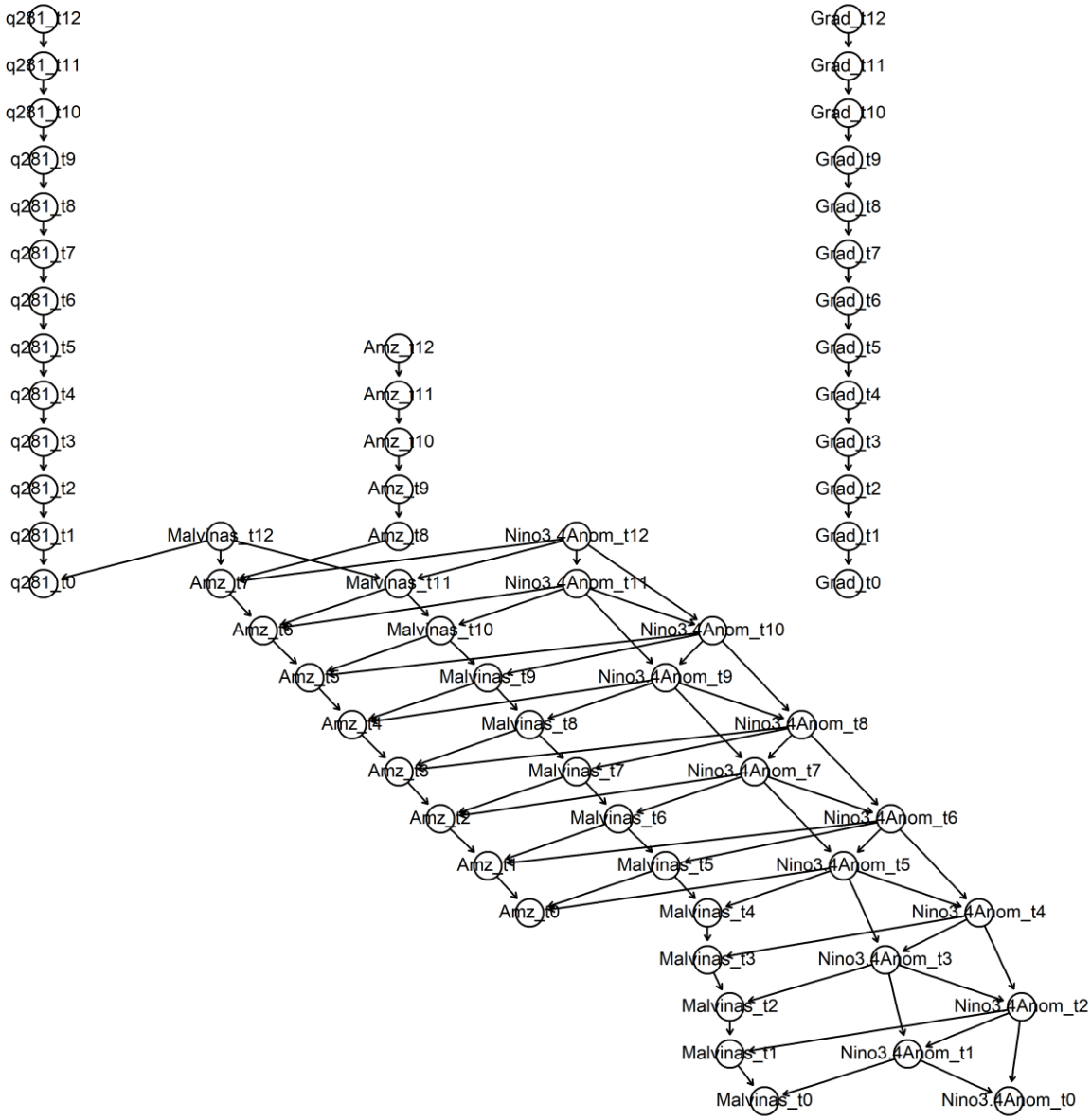


Figure B6. PCMCI network Structure of station 281 (second period)

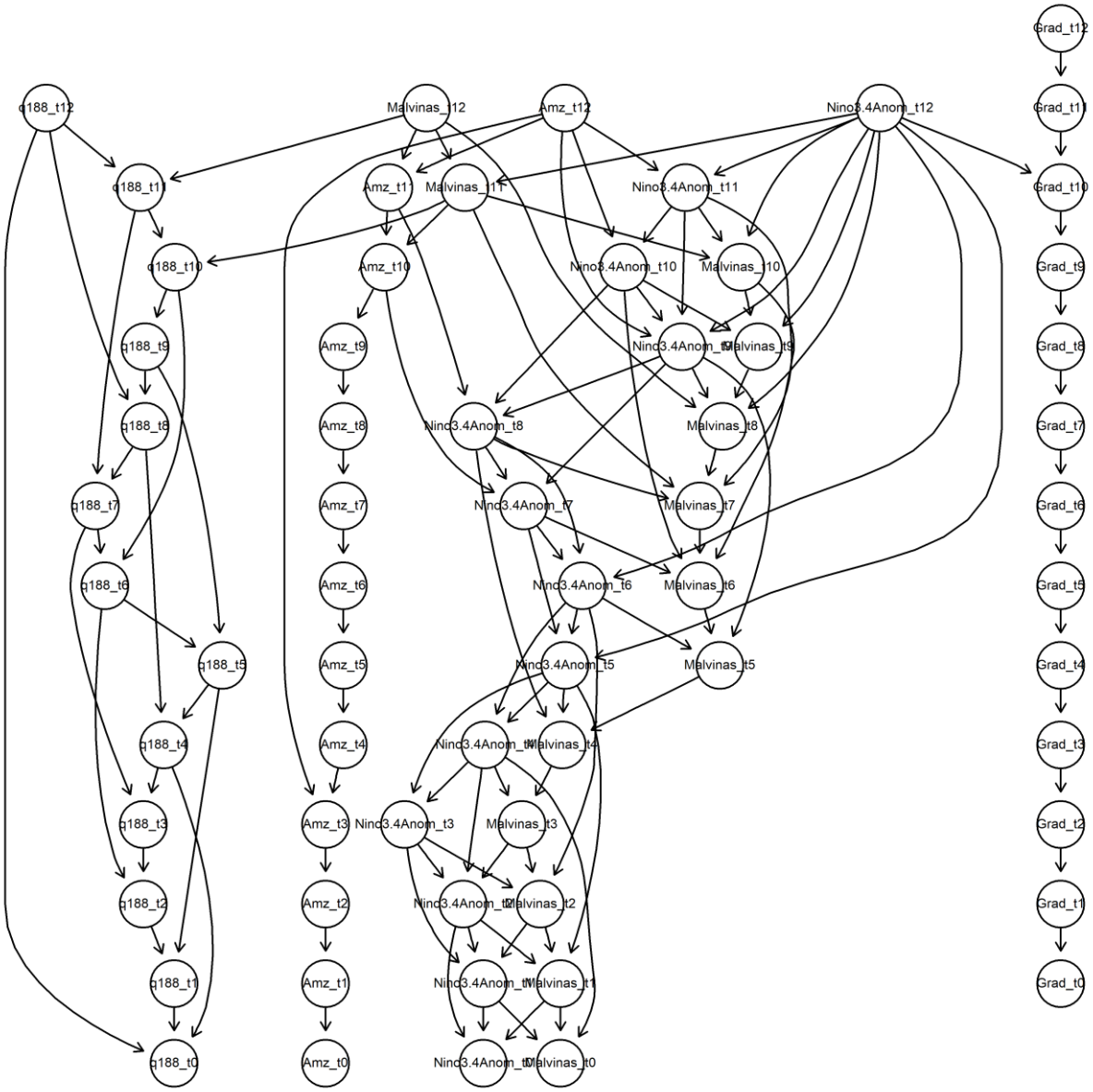


Figure B7. HC network Structure of station 188 (first period)

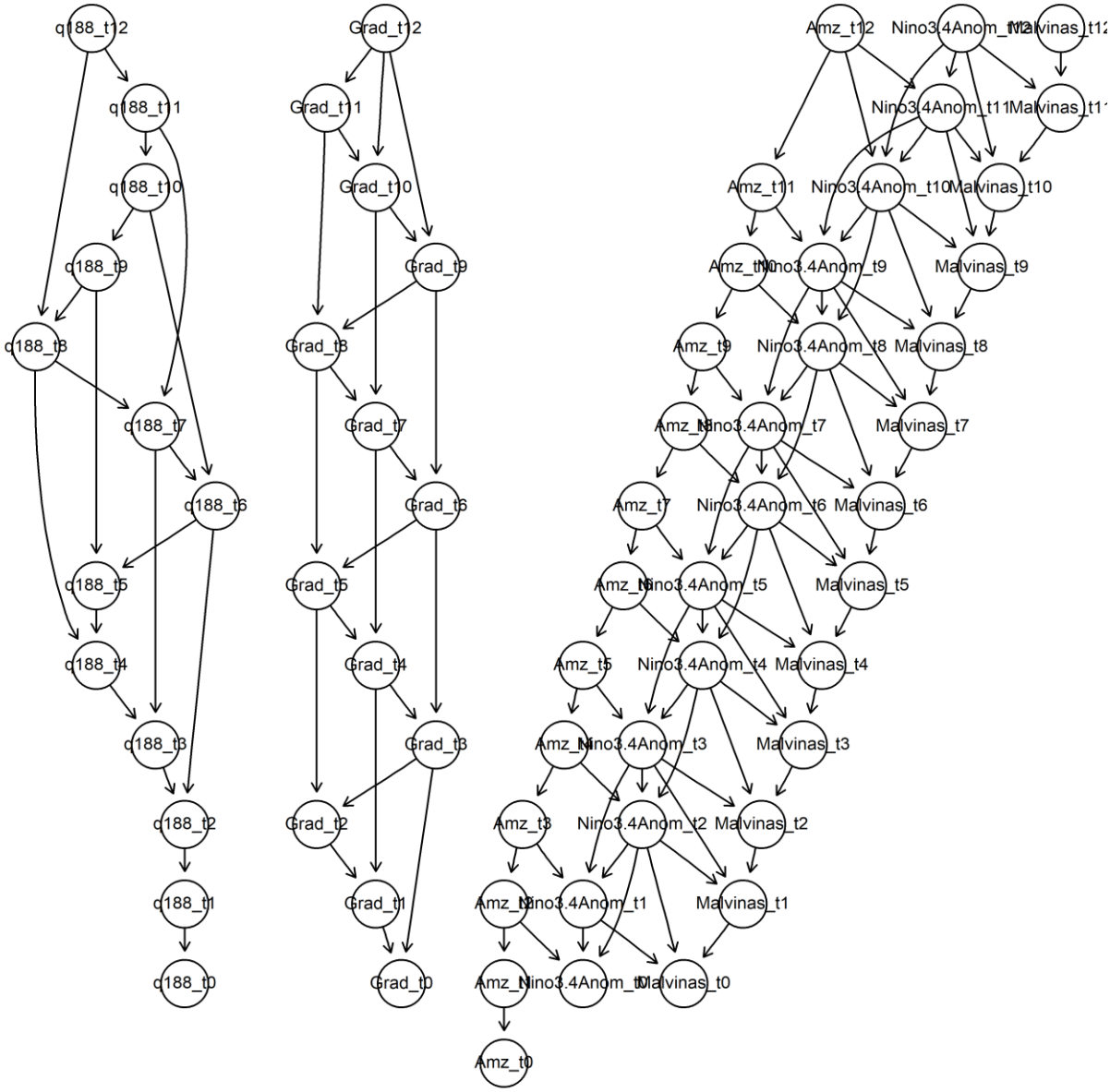


Figure B8. HC network Structure of station 188 (second period)

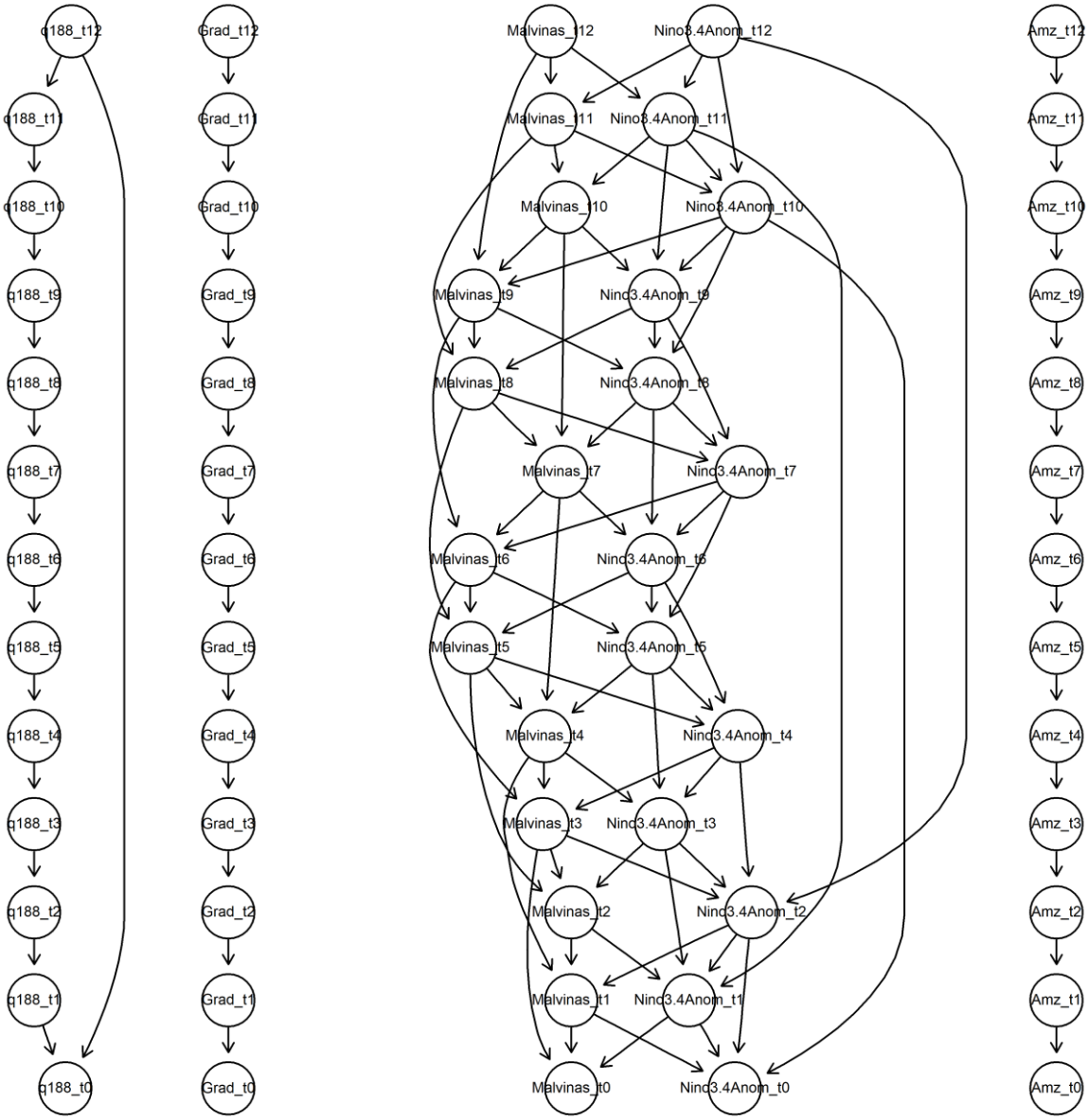


Figure B9. PCMCI network Structure of station 188 (first period)

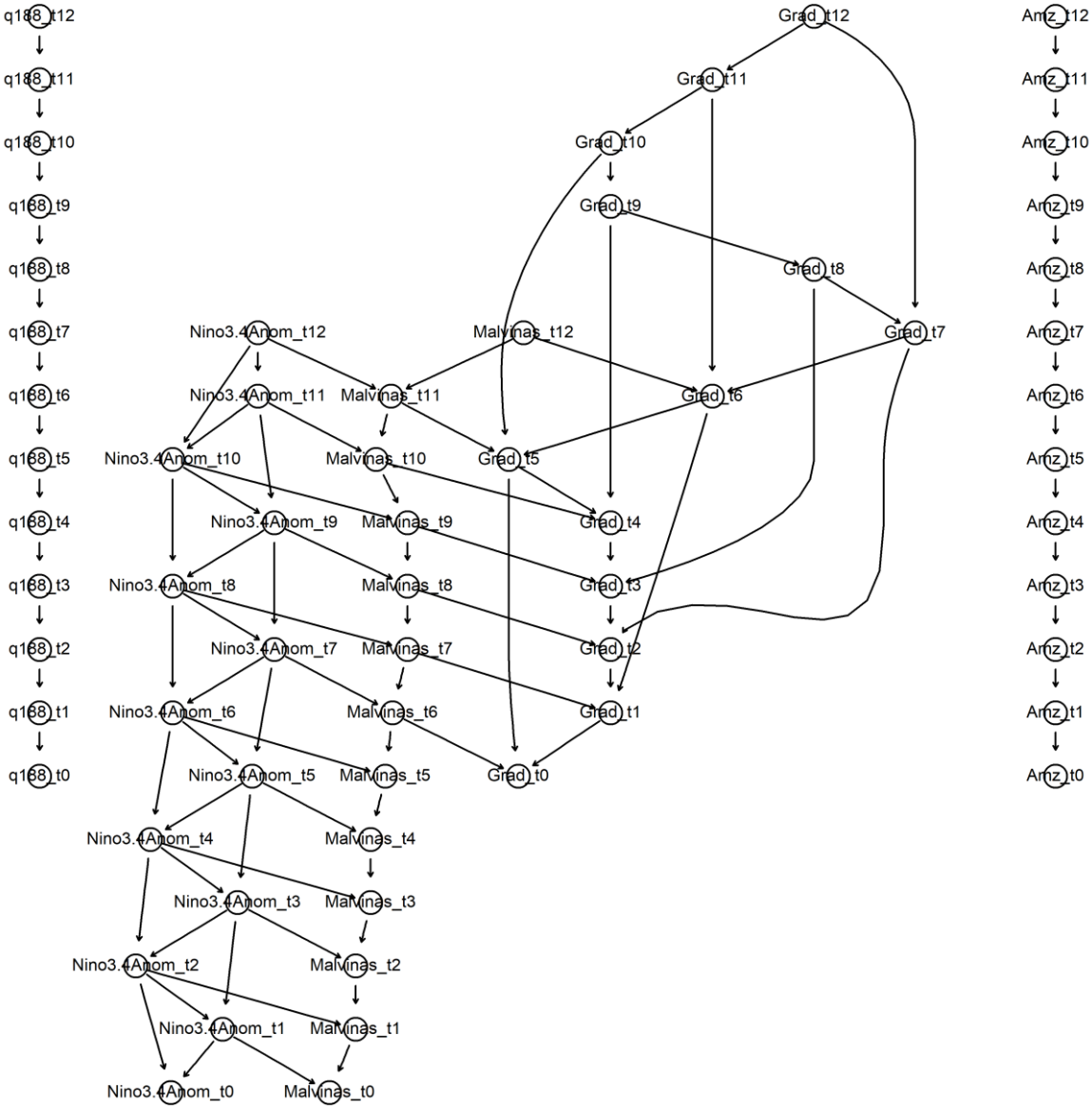


Figure B10. PCMCI network Structure of station 188 (second period)

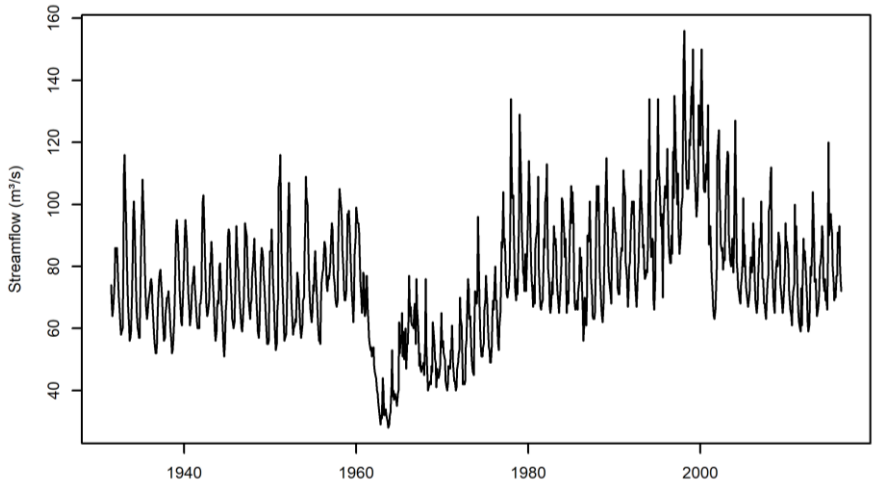


Figure B11. Streamflow time-series of station 281.

Appendix C – Supplemental files of chapter 7

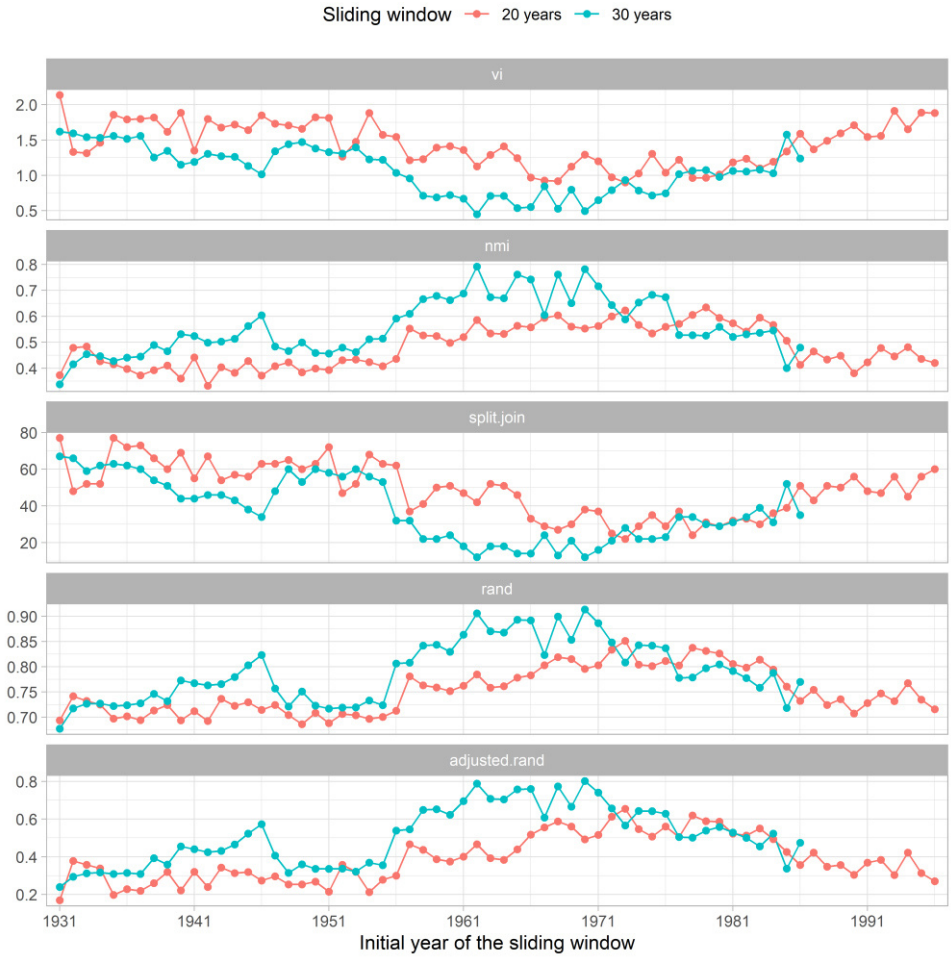


Figure C1. Comparison between the communities detected for the entire series (1931-2015) and for sliding windows of 20 and 30 years through different metrics.

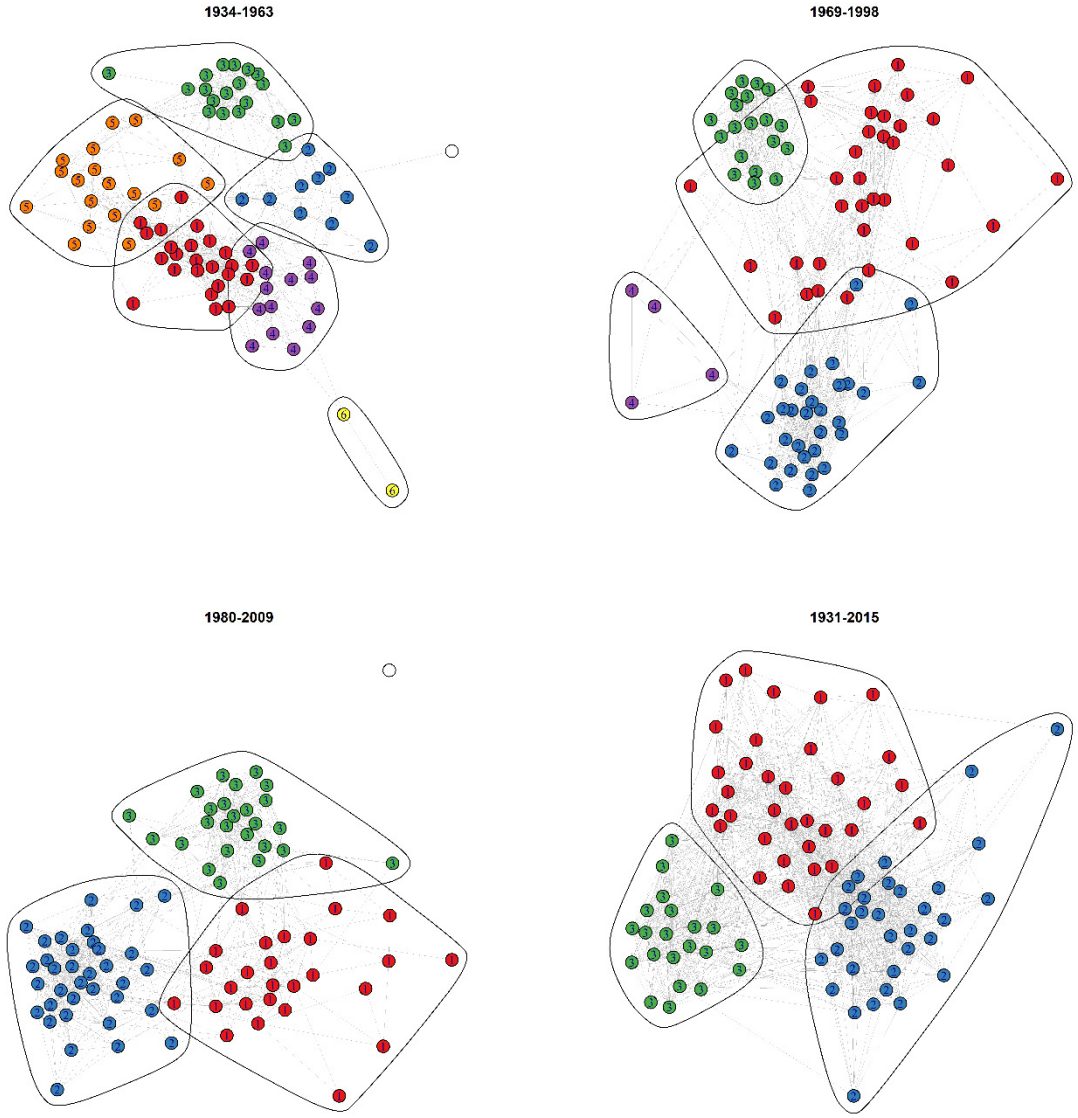


Figure C2. Network structure for the different time-windows. The number inside the nodes stands for the community (cluster) number.

Appendix D – Further information of the streamflow stations and analysis of the streamflow elasticity to climate

This appendix presents further information of the streamflow stations used in this work and streamflow summary statistics. Although not thoroughly analyzed in this work, we find that this complementary information is important to better characterize the data used in this work and may be helpful for future studies. Also, a simple complementary analysis was made regarding the elasticity of streamflow to precipitation and potential evapotranspiration, allowing the visualization of the different empirical elasticity throughout the country.

Table D1 comprises information regarding the drainage area of the stations. This information was extracted from shapefiles provided by ANA, information regarding station 160 (ONS code) was not available.

Station Code		Name	Original Name	Area (km²)
ONS	ANA			
1	3578	CAMARGOS	UHE Camargos	6231,63
6	3460	FURNAS	UHE Furnas	51942,82
14	3434	CACONDE	UHE Caconde	2564,35
17	8042	MARIMBONDO	UHE Marimbondo	118141,42
18	8292	A_VERMELHA	UHE Água Vermelha	139154,89
24	21665	EMBORCACAO	UHE Emborcação	28952,65
25	3794	NOVA_PONTE	UHE Nova Ponte	15326,96
31	8228	ITUMBIARA	UHE Itumbiara	94690,17
32	8398	CACH_DOURADA	UHE Cachoeira Dourada	99699,52
33	8293	SAO_SIMAO	UHE São Simão	171465,51
34	8242	I_SOLTEIRA	UHE Ilha Solteira	377329,04
47	8217	A_A_LAYDNER (JURUMIRIM)	UHE Jurumirim	17924,25
61	8074	CAPIVARA	UHE Capivara	84799,46
63	8107	ROSANA	UHE Rosana	100907,58
71	7663	STA CLARA PR	UHE Santa Clara	3914,60
72	23104	FUNDAO	UHE Fundão	4102,61
73	7662	JORDAO	UHE Barra	4683,52
74	7659	G_B_MUNHOZ	UHE Governador Bento Munhoz da Rocha Neto	30207,57
76	7337	SEGREDO	UHE Governador Ney Aminthas de Barros Braga	34412,89
77	7664	SLT_SANTIAGO	UHE Salto Santiago	43948,49
78	7673	SALTO OSORIO	UHE Salto Osório	45888,69
92	7609	ITA	UHE Itá	44226,51
93	7606	PASSO FUNDO	UHE Passo Fundo	2224,68
94	23054	FOZ CHAPECO	UHE Foz do Chapecó	53181,48
98	23080	CASTRO ALVES	UHE Castro Alves	7768,31
99	23051	ESPORA	UHE Espora	4055,98
101	23127	SALTO PILAO	UHE Salto Pilão	5402,26
102	23064	SAO JOSE	UHE Passo São José	9586,47
111	7558	PASSO REAL	UHE Passo Real	8274,65
115	7687	G_P_SOUZA	UHE Governador Parigot de Souza	1021,88
117	3531	GUARAPIRANGA	Represa Guarapiranga	635,08
119	3454	BILLINGS_PED	Represa Billings	600,33
120	3570	JAGUARI	UHE Jaguari	1316,82
121	3560	PARAIBUNA	UHE Paraibuna	2699,97
125	23128	STA_CECILIA	UHE Santa Cecília	16617,00
130	23103	I_POMBOS	UHE Ilha dos Pombos	32274,35
134	3875	SALTO GRANDE	UHE Salto Grande	2476,72

Station Code ONS	Station Code ANA	Name	Original Name	Area (km²)
144	23087	MASCARENHAS	UHE Mascarenhas	73699,98
145	62829	RONDON II		3288,63
149	3634	CANDONGA	UHE Risoleta Neves	9004,54
155	62828	RETIRO BAIXO		11193,53
156	4213	TRES_MARIAS	UHE Três Marias	50736,26
158	3895	QUEIMADO	UHE Queimado	3652,71
164	131822	E.SOUZA		4797,95
168	22991	SOBRADINHO	UHE Sobradinho	499170,25
188	465	ITAPEBI	UHE Itapebi	67859,88
190	5864	B_ESPERANCA	UHE Boa Esperança	84904,52
191	8521	CANA BRAVA	UHE Cana Brava	58074,38
196	3469	ROSAL	UHE Rosal	1746,87
197	3626	PICADA	UHE Picada	1725,83
201	121795	TOCOS		388,74
205	8252	CORUMBA IV	UHE Corumbá IV	7026,57
206	21664	MIRANDA	UHE Miranda	18008,11
209	8240	CORUMBA I	UHE Corumbá I	27774,32
211	3581	FUNIL-GRANDE	UHE Funil	15720,01
215	7559	BARRA GRANDE	UHE Barra Grande	11916,57
216	7607	CAMPOS NOVOS	UHE Campos Novos	14445,83
220	23138	MONJOLINHO	UHE Alzir dos Santos Antunes	3739,70
237	7922	BARRA_BONITA	UHE Barra Bonita	32949,15
240	8013	PROMISSAO	UHE Promissão	57841,51
242	8063	NAVANHANDAVA	UHE Nova Avanhandava	62474,61
243	8099	T_IRMAOS	UHE Três Irmãos	70973,56
245	8124	JUPIA	UHE Jupiá	476527,73
246	7773	P_PRIMAVERA	UHE Porto Primavera	571692,41
247	23050	CACU	UHE Caçu	12139,28
251	23059	SERRA FACAO	UHE Serra do Facão	10591,93
253	23044	SAO SALVADOR	UHE São Salvador	63750,79
254	795	P_CAVALO	UHE Pedra do Cavalo	53949,73
255	23137	IRAPE	UHE Irapé	15850,89
257	23046	PEIXE ANGIC	UHE Peixe-Angical	125687,27
259	14511	ITIQUIRA I	UHE Itiquira I	5231,21
266	11735	ITAIPU	UHE Itaipu	822904,33
270	8523	SERRA MESA	UHE Serra da Mesa	51352,76
271	23058	ESTREITO	UHE Estreito	285803,38
273	8520	LAJEADO	UHE Luis Eduardo Magalhães	183646,06
275	6540	TUCURUI	UHE Tucuruí	758869,41
277	15527	CURUA-UNA	UHE Curuá-Una	16248,86
278	14602	MANSO	UHE Manso	9392,90
279	17296	SAMUEL	UHE Samuel	14780,50
281	14517	PONTE PEDRA	UHE Ponte de Pedra	4031,64
283	605	STA CLARA MG	UHE Santa Clara	14580,20
286	7624	QUEBRA QUEIX	UHE Quebra Queixo	2664,08
287	23070	STO ANTONIO		976400,57
291	62833	DARDANELOS		15332,61
294	23048	SALTO	UHE Salto	10825,34
295	14588	JAURU	UHE Jauru	2245,73
296	14589	GUAPORE	UHE Guaporé	1344,29

Table D1. Drainage area of the stations. Extracted from the shapefiles provided by ANA. Blank cell refers to unavailable information.

The drainage area varies significantly between the stations. From the boxplot of Figure D1 (left) it was noticeable that mostly station area was lower than 140,000 km² with a median value of 15,850.89 km² and minimum value of 388.7 km². Ten stations presented anomalously higher values (Figure D1 right), ranging from 171,465.5 km² to 976,400.6 km²

with a median of 487,849 km². The station considered outliers through the boxplot standard whiskers were (ONS code): 33, 34, 168, 245, 246, 266, 271, 273, 275, 287. The Sobradinho station (168) large area was one of the possible factors behind the results obtained in the second article.

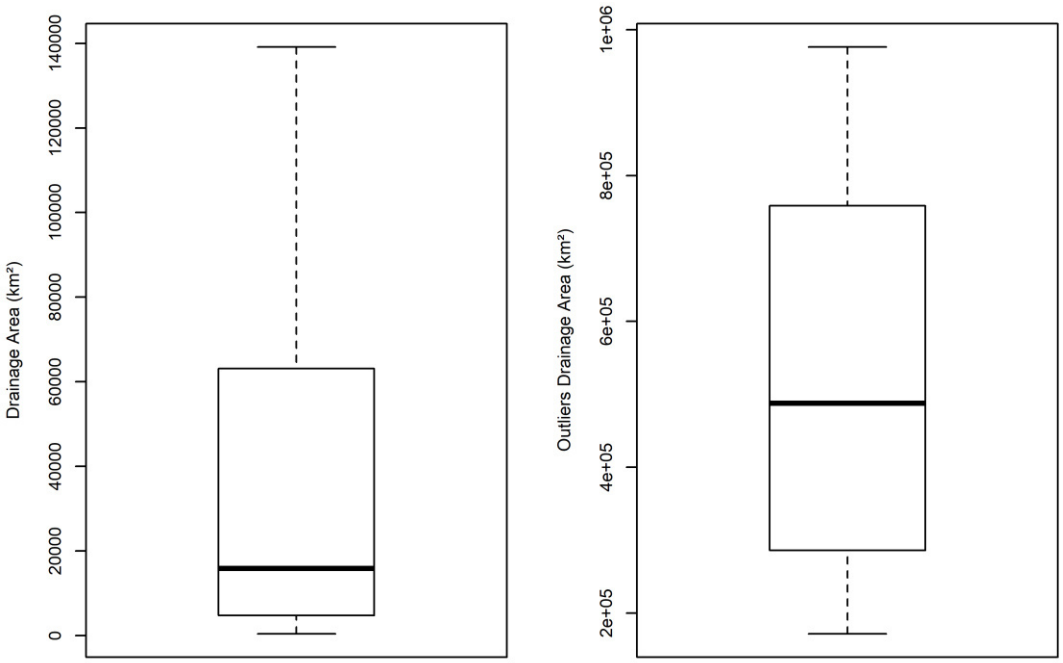


Figure D1. Boxplot of the drainage area of the stations without outliers (left). Boxplot of the outliers (right).

The area great range of variation impose streamflow values in different magnitude scales, as can be seem in the Figure D2 and Table D2. From Figure D2 it was also noticeable the different ranges of variations, resulting in boxplots with a large difference between the first and third quantile and other with a small difference range. These differences were also noticeable from the CV values, which oscillated between 0,12 and 1,53. The stations that presented the largest area also showed the highest streamflow values, as expected.

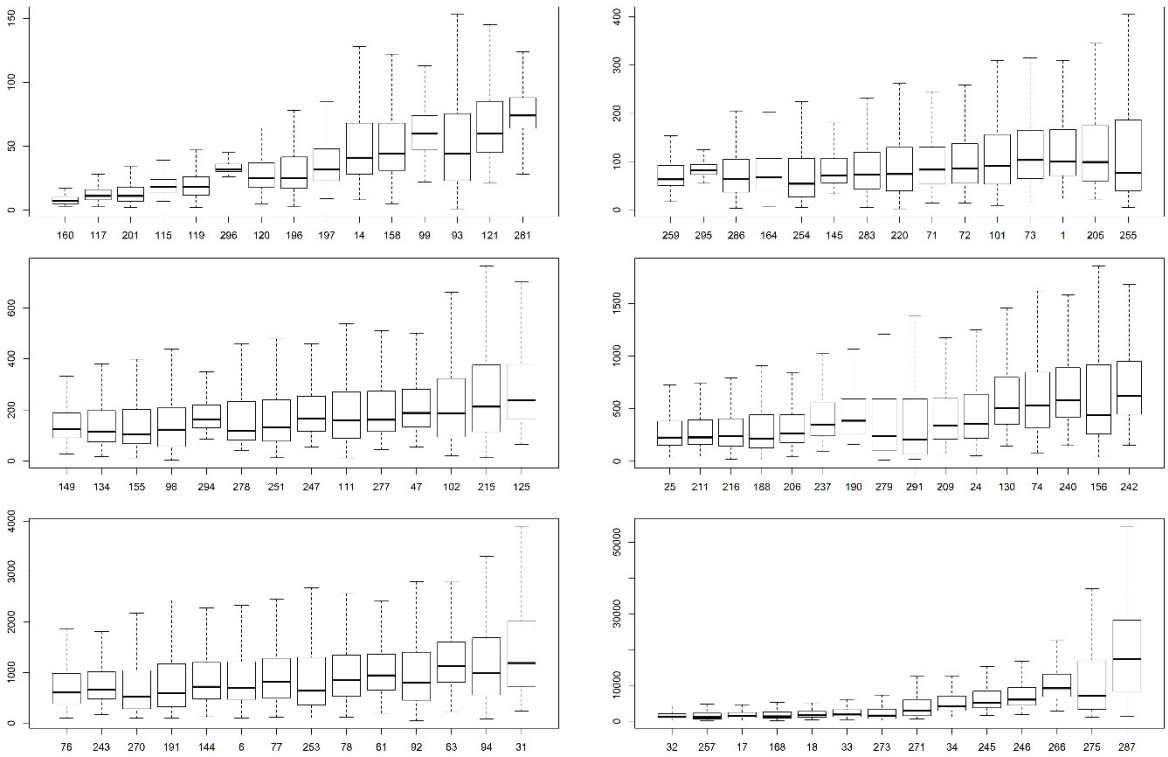


Figure D2. Boxplot without outliers of the streamflow monthly values of the stations ordered by magnitude of the third quantile and presented in groups.

Station Code (ONS)	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	CV	SD
1	24,00	71,00	100,00	130,73	166,25	576,00	0,65	84,43
6	102,00	469,00	696,00	909,42	1214,50	3757,00	0,67	612,05
14	8,00	28,00	41,00	53,27	68,00	235,00	0,67	35,82
17	245,00	972,75	1406,50	1830,07	2430,00	7228,00	0,64	1167,21
18	364,00	1117,25	1619,50	2075,80	2728,00	7938,00	0,62	1293,65
24	51,00	217,75	358,50	477,98	635,50	2329,00	0,75	359,80
25	40,00	151,00	226,00	293,40	381,25	1243,00	0,67	196,94
31	233,00	727,75	1188,00	1531,23	2010,50	6649,00	0,69	1063,16
32	254,00	769,25	1247,50	1610,03	2128,00	7100,00	0,69	1109,70
33	450,00	1221,75	1860,50	2383,85	3196,25	9931,00	0,65	1547,61
34	1212,00	2941,50	4194,00	5271,65	7035,00	20314,00	0,59	3095,17
47	55,00	134,00	188,50	226,86	281,25	1552,00	0,61	137,67
61	192,00	651,75	944,00	1116,90	1362,25	7334,00	0,63	698,82
63	227,00	809,25	1128,00	1319,03	1606,00	9044,00	0,61	807,47
71	14,00	53,00	83,00	104,45	130,00	841,00	0,79	82,18
72	14,00	55,00	86,00	109,56	137,00	880,00	0,79	86,02
73	15,00	65,00	103,00	129,96	165,00	1020,00	0,79	102,10
74	80,00	321,00	532,00	666,48	847,00	5150,00	0,75	498,19
76	94,00	385,25	606,50	770,95	979,00	5893,00	0,74	572,51
77	116,00	502,75	811,50	1023,06	1282,00	8252,00	0,76	773,96
78	119,00	525,75	857,50	1071,20	1342,50	8473,00	0,76	810,15
92	49,00	452,25	804,50	1052,86	1403,00	8292,00	0,82	860,71
93	1,00	23,00	44,00	56,51	75,25	391,00	0,84	47,22
94	79,00	558,75	986,50	1290,75	1687,50	10048,00	0,82	1055,32
98	3,00	57,00	123,00	154,75	210,00	1033,00	0,87	134,32
99	22,00	47,75	60,00	63,19	74,00	179,00	0,32	20,10
101	9,00	53,00	91,00	118,07	156,00	1058,00	0,81	95,40

Station Code (ONS)	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	CV	SD
102	20,00	94,75	187,00	243,04	323,25	1435,00	0,85	207,09
111	12,00	89,00	160,00	209,00	271,00	1318,00	0,81	169,10
115	7,00	14,00	18,00	19,93	24,00	74,00	0,46	9,14
117	3,00	8,00	11,00	12,88	16,00	60,00	0,56	7,18
119	2,00	11,75	18,00	19,97	26,00	99,00	0,58	11,65
120	5,00	18,00	25,00	29,37	37,00	111,00	0,54	15,81
121	21,00	45,00	60,00	69,52	85,00	276,00	0,49	34,25
125	66,00	163,00	239,00	294,92	378,25	1148,00	0,59	173,95
130	146,00	353,00	506,50	623,32	798,25	2579,00	0,58	364,31
134	18,00	75,00	116,00	151,64	197,25	1051,00	0,74	111,46
144	125,00	479,75	720,00	946,07	1200,00	5262,00	0,72	681,33
145	33,00	56,00	71,00	85,98	107,00	366,00	0,49	41,84
149	28,00	91,00	126,00	151,25	188,00	935,00	0,58	88,47
155	12,00	68,00	104,00	152,78	201,25	1024,00	0,83	126,76
156	40,00	262,75	440,50	673,77	915,75	4435,00	0,89	598,71
158	5,00	31,00	44,00	54,27	68,00	294,00	0,65	35,21
160	3,00	5,00	7,00	7,90	10,00	29,00	0,46	3,64
164	6,00	42,00	67,00	79,30	106,00	371,00	0,63	49,80
168	167,00	848,75	1330,00	1861,33	2612,25	12592,00	0,76	1417,71
188	14,00	126,00	215,00	370,56	441,75	4124,00	1,15	427,41
190	161,00	264,00	385,00	456,32	592,00	1907,00	0,53	241,58
191	102,00	324,00	591,50	851,80	1170,25	7012,00	0,89	755,82
196	3,00	17,00	25,00	32,91	41,25	194,00	0,74	24,31
197	9,00	23,00	32,00	38,35	48,00	142,00	0,56	21,40
201	2,00	7,00	11,00	13,99	18,00	115,00	0,72	10,11
205	22,00	60,00	99,00	130,45	174,25	573,00	0,71	92,69
206	42,00	177,75	263,00	340,87	442,25	1427,00	0,66	224,91
209	73,00	209,75	338,50	449,17	598,00	1955,00	0,72	323,39
211	48,00	159,00	227,00	300,03	394,00	1288,00	0,66	197,06
215	13,00	115,00	213,00	277,40	376,00	1807,00	0,82	228,03
216	17,00	144,00	239,50	317,65	404,25	2932,00	0,83	263,50
220	2,00	39,00	74,00	97,22	129,25	702,00	0,86	84,03
237	91,00	241,00	346,00	438,89	556,25	2420,00	0,65	284,07
240	149,00	420,50	582,50	708,88	887,50	3315,00	0,60	425,67
242	153,00	446,75	621,00	756,63	950,00	3748,00	0,61	464,32
243	165,00	477,75	660,50	806,37	1013,25	3761,00	0,61	488,31
245	1649,00	3721,50	5148,50	6414,72	8347,50	23753,00	0,56	3582,61
246	1881,00	4419,50	6015,00	7239,72	9373,25	26596,00	0,53	3813,36
247	54,00	116,00	166,50	194,92	253,25	590,00	0,50	97,50
251	13,00	78,00	131,00	175,74	240,00	834,00	0,74	130,58
253	107,00	354,00	638,50	941,42	1300,25	7377,00	0,88	829,82
254	5,00	27,00	54,00	99,20	106,25	1625,00	1,53	151,52
255	5,00	40,00	76,00	146,34	186,00	1572,00	1,18	172,79
257	185,00	575,00	1096,00	1647,98	2260,75	12561,00	0,93	1536,31
259	17,00	50,00	63,00	73,32	92,00	244,00	0,44	32,27
266	2839,00	6839,25	9197,50	10367,26	13139,25	31630,00	0,46	4798,22
270	94,00	283,00	518,00	761,00	1041,25	6163,00	0,89	681,06
271	585,00	1508,50	2984,00	4174,98	5914,00	22600,00	0,82	3416,09
273	259,00	758,50	1565,50	2390,55	3374,00	15250,00	0,93	2233,02
275	1090,00	3295,75	7116,50	10871,60	16825,50	51539,00	0,84	9161,73
277	45,00	115,75	161,50	212,47	273,25	827,00	0,65	137,34
278	42,00	83,00	119,00	170,52	234,00	727,00	0,71	120,36
279	9,00	99,00	241,00	353,59	592,25	1339,00	0,85	299,60
281	28,00	64,00	74,00	76,16	88,00	156,00	0,26	19,86
283	5,00	43,75	73,00	97,31	119,00	1033,00	0,89	86,77

Station Code (ONS)	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	CV	SD
286	3,00	36,75	63,50	80,54	105,00	621,00	0,81	64,97
287	1407,00	8299,00	17287,00	18925,85	28237,25	54435,00	0,60	11430,99
291	21,00	65,00	203,50	337,01	593,00	1515,00	0,93	314,42
294	85,00	131,00	163,00	180,07	219,00	482,00	0,35	62,99
295	55,00	73,00	82,00	85,54	94,00	154,00	0,20	17,16
296	26,00	30,00	32,00	33,15	36,00	49,00	0,12	4,11

Table D2. Streamflow monthly values summary. CV stands for coefficient of variation and SD for standard deviation.

Streamflow elasticity describes the sensitivity of the changes in streamflow related to changes in a climate variable (SCHAAKE & LIU, 1989[7] *apud* ANDRESSIAN *et al.,* 2016[8]). The comparison of the streamflow elasticity of the stream gauges used in this work provides interestingly information regarding the different streamflow characteristic throughout the country. For example, stations in which variations of streamflow are directly associated to a corresponding precipitation variation indicates a lower influence of groundwater flow. An opposite behavior, combined with a reduced influence of evapotranspiration, indicates a strong influence of groundwater flows and, thus, a strong autocorrelated streamflow time-series is expected.

Andréassian *et al* (2016) states that the estimate of the empirical streamflow elasticity is greatly improved with the combined use of precipitation (P) and potential evapotranspiration (ETP) since the bivariate models showed very superior results compared to the univariate models. In their work the bivariate generalized least squares regression model (GLS) presented the best results, closely followed by the ordinary least squares model (OLS).

The OLS approach was used in this work, the bivariate regression model is given by:

$$\Delta Q_i^{(M)} = e_{Q/P}^{(M)} \cdot \Delta P_i^{(M)} + e_{Q/ETP}^{(M)} \cdot \Delta ETP_i^{(M)} + \omega_i \qquad \text{(Eq. D1)}$$

[7] SCHAAKE, J. C.; LIU, C. Development and application of simple water balance models to understand the relationship between climate and water resources. **New Directions for Surface Water Modelling, Proc. Symp. 3Rd Scientific Assembly Int. Association Hydrological Sciences, Balti**, *[S. l.]*, n. 181), p. 343–352, 1989.

[8] ANDRÉASSIAN, Vazken; CORON, Laurent; LERAT, Julien; LE MOINE, Nicolas. Climate elasticity of streamflow revisited - An elasticity index based on long-term hydrometeorological records. **Hydrology and Earth System Sciences**, *[S. l.]*, v. 20, n. 11, p. 4503–4524, 2016. DOI: 10.5194/hess-20-4503-2016.

where: $\Delta Q_i^{(M)}, \Delta P_i^{(M)}, \Delta ETP_i^{(M)}$ is, respectively, the streamflow, P, and ETP absolute anomaly over the M years, $e_{Q/P}^{(M)}$ and $e_{Q/ETP}^{(M)}$ is the streamflow elasticity related to P and ETP, ω_i is the regression residual and M is the number of years used to calculate long-term average of the variables. Absolute anomaly is calculated by the distance to the mean value over the M years. The streamflow anomaly is given in mm per year (mm/y)

A non-parametric univariate approach (NP) was proposed by Sankarasubramanian *et al* (2001)[9], in which the streamflow elasticity is simply given by the median of the ratio between the streamflow relative anomalies and the climate variable relative anomalies.

The monthly mean precipitation of the basins was obtained by interpolating via inverse distance weighting (IDW) monthly observed precipitation data from stations of the Brazilian National Water Agency (ANA). Penmam-Monteith ETP was calculated for each meteorological station of the National Meteorologic Institute (INMET) and the basins average ETP was also obtained through IDW. Nine streamflow gauges were discarded for this analysis due to the lack of meteorological gauges near its basins or the absence of data for at least 5 complete years.

Annual series of both P and ETP were obtained from the monthly values only for years with complete information. The period of analysis was from 1961 to 2016 due to the INMET time series which only starts in 1961.

The elasticity results from three methods (Univariate and bivariate OLS and NP) illustrates the different streamflow sensibility to P and ETP across Brazil (Figures D3 to D5 and Table D3). Both OLS results indicates a spatial pattern similar to a south-northwest gradient for the streamflow elasticity to P, in which the highest values occurred for the southern stations and the northwestern stations presented the lowest values. NP shows similar results for the southern stations, however, no clear spatial pattern could be observed, with several stations presenting high values of elasticity across Brazil.

The streamflow elasticity to ETP significantly varied across the different methods. The bivariate OLS approach resulted in physically incompatible results for a significant number

---

[9] SANKARASUBRAMANIAN, A.; VOGEL, Richard M.; LIMBRUNNER, James F. Climate elasticity of streamflow in the United States. **Water Resources Research**, *[S. l.]*, v. 37, n. 6, p. 1771–1781, 2001. DOI: 10.1029/2000WR900330.

of stations, presenting positive values. Similarly, the NP also showed positive values for diverse stations. However, the univariate OLS method presented the lowest number of stations with positive elasticity to ETP. A few negative values were also observed for the streamflow elasticity to P. These results derives from the short length of the series for some stations and also to possible impacts of the non-stationarity of the streamflow series.

The univariate OLS results resembles the pattern observed for the streamflow elasticity to P, although presenting higher values for the South region. The remaining results show no clear spatial pattern, however, NP method also shows strong elasticity to ETP for the southern stations.

In conclusion, the streamflow elasticity significant diverged through the different methods, especially for ETP. However, mostly results showed that the South region of brazil presents a strong elasticity to both climate variables. This characteristic is a possible cause behind the worst results obtained for this region in the second article, since an autoregressive based model cannot capture the P and ETP influence in the streamflow behavior. The inclusion of these variables may greatly improve the Bayesian network model performance.

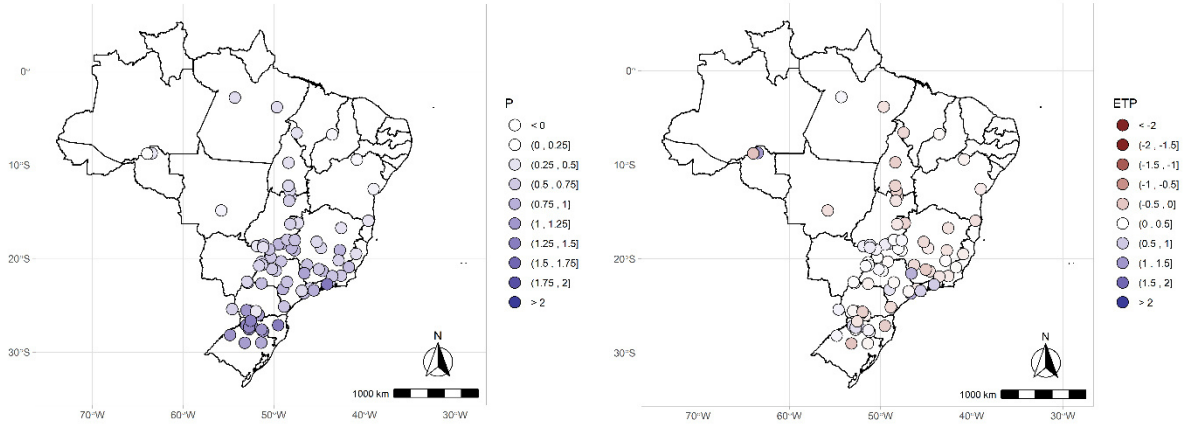


Figure D3. Streamflow elasticity to precipitation (left) and potential evapotranspiration (right) through the bivariate OLS method.

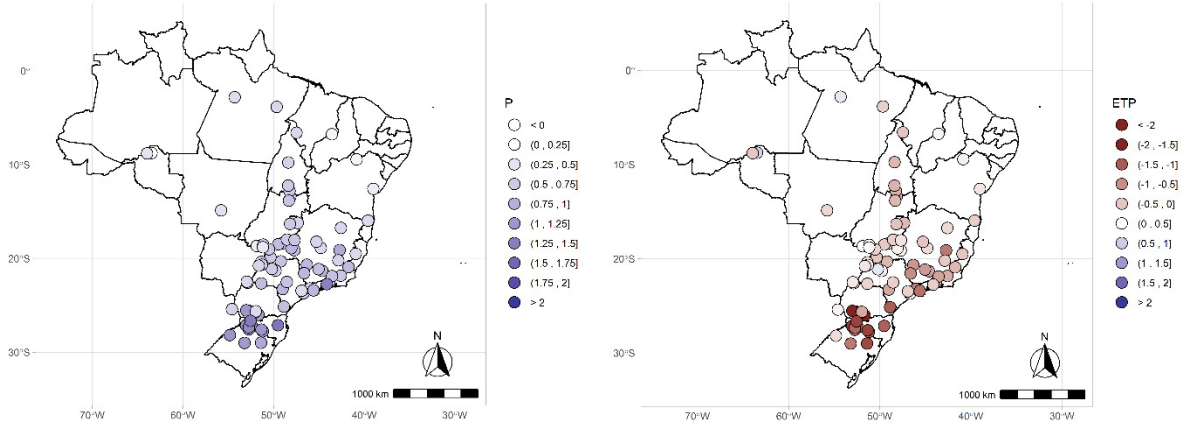


Figure D4. Streamflow elasticity to precipitation (left) and potential evapotranspiration (right) through the univariate OLS method.

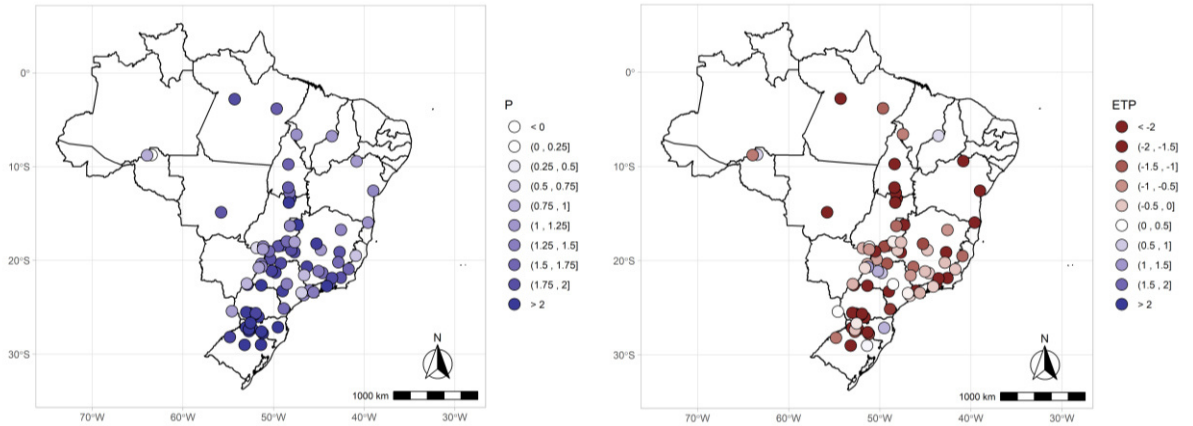


Figure D5. Streamflow elasticity to precipitation (left) and potential evapotranspiration (right) through NP method.

Station	$biOLS\ e_{Q/P}$	$uniOLS\ e_{Q/P}$	$NP\ e_{Q/P}$	$biOLS\ e_{Q/ETP}$	$uniOLS\ e_{Q/ETP}$	$NP\ e_{Q/ETP}$
1	0,63	0,69	1,43	-0,27	-0,82	-0,45
6	0,57	0,66	1,88	-0,32	-0,94	-1,40
14	0,88	0,68	0,63	0,74	-1,01	-0,94
17	0,55	0,57	1,81	-0,05	-0,64	-1,55
18	0,51	0,52	1,73	-0,02	-0,51	-0,82
24	0,65	0,62	1,59	0,11	-0,35	-0,45
25	0,58	0,58	1,79	0,01	-0,11	-2,66
31	0,68	0,62	1,70	0,18	-0,48	-1,61
32	0,67	0,61	1,68	0,17	-0,47	-1,65
33	0,56	0,53	1,45	0,09	-0,39	-2,34
34	0,43	0,43	1,22	0,02	-0,36	-0,71
47	0,74	0,70	1,91	0,27	-0,79	-3,88
61	0,61	0,63	2,23	-0,12	-0,36	-3,40
63	0,60	0,61	2,26	-0,08	-0,32	-2,70
74	0,89	0,91	1,95	-0,11	-2,54	-4,17
76	0,95	0,95	1,90	0,01	-2,63	-2,59
77	1,00	1,01	2,06	-0,07	-2,92	-2,89
78	0,99	1,01	1,98	-0,07	-2,90	-3,01
92	1,12	1,06	2,17	0,48	-1,42	-4,01
93	0,97	0,97	2,05	-0,04	-1,23	-1,60
94	1,14	1,10	2,32	0,36	-1,81	-3,34
98	0,81	0,82	2,21	-0,10	-1,65	0,07
99	0,21	0,16	0,30	0,16	-0,01	-0,69

Station	$biOLS\ e_{Q/P}$	$uniOLS\ e_{Q/P}$	$NP\ e_{Q/P}$	$biOLS\ e_{Q/ETP}$	$uniOLS\ e_{Q/ETP}$	$NP\ e_{Q/ETP}$
101	1,27	1,34	2,86	-0,44	-1,43	0,72
102	1,02	1,01	1,88	0,09	-0,30	-1,24
111	0,94	0,97	2,13	-0,51	-1,19	-2,56
115	0,56	0,60	1,44	-0,30	-1,47	-1,82
117	0,53	0,52	1,18	0,07	-0,51	-0,38
119	0,62	0,54	1,14	0,79	-0,16	-0,30
120	0,48	0,51	1,24	-0,19	-1,12	-2,61
121	0,75	0,68	1,28	0,44	-1,29	-0,68
125	0,50	0,52	1,59	-0,12	-0,62	-2,67
130	0,51	0,54	1,71	-0,18	-0,71	-2,02
134	0,71	0,80	1,67	-0,24	-1,22	-2,07
144	0,30	0,35	0,55	-0,16	-0,50	-1,37
149	0,42	0,43	1,44	-0,02	-0,47	-0,57
155	0,36	0,38	0,98	-0,17	-0,33	-0,53
156	0,34	0,40	1,95	-0,32	-0,50	-1,65
158	0,31	0,41	2,08	-0,37	-0,60	-3,13
164	0,33	0,33	0,45	-0,07	-0,39	-0,11
168	0,07	0,10	0,97	-0,11	-0,13	-2,05
188	0,25	0,35	1,05	-0,28	-0,43	-3,40
190	0,09	0,09	1,01	-0,06	-0,06	0,35
191	0,43	0,54	1,87	-0,29	-0,66	-2,56
196	0,62	0,63	1,52	-0,05	-0,61	-0,50
197	0,56	0,59	1,82	-0,22	-0,82	-2,20
201	1,31	1,23	2,06	0,47	-0,44	-0,45
205	0,43	0,51	1,16	-0,20	-0,53	-1,36
206	0,69	0,69	1,76	0,03	-0,11	-0,46
209	0,52	0,51	1,45	0,03	-0,41	-0,13
211	0,49	0,59	1,29	-0,45	-0,89	-0,55
215	0,97	0,96	2,24	0,10	-1,53	-1,61
216	1,02	1,01	2,11	0,11	-1,73	-2,22
220	1,16	1,13	2,29	0,28	-1,37	-0,56
237	0,55	0,56	1,24	-0,06	-0,42	0,02
240	0,46	0,47	1,89	0,05	0,16	0,57
242	0,49	0,50	1,91	0,07	0,22	0,77
243	0,45	0,46	1,14	0,05	0,20	0,94
245	0,41	0,40	0,84	0,04	-0,23	-0,31
246	0,38	0,37	0,70	0,02	-0,18	-0,68
247	0,50	0,37	1,07	0,25	-0,06	0,02
251	0,57	0,54	0,89	0,09	-0,24	-0,54
253	0,38	0,51	1,42	-0,37	-0,70	-2,38
254	0,12	0,18	1,16	-0,16	-0,22	-5,73
255	0,27	0,37	1,16	-0,26	-0,33	-0,96
257	0,34	0,48	1,60	-0,38	-0,68	-3,27
266	0,45	0,43	0,92	0,13	-0,11	0,01
270	0,46	0,55	2,01	-0,23	-0,63	-2,11
271	0,27	0,37	1,06	-0,33	-0,50	-1,15
273	0,28	0,43	1,54	-0,42	-0,62	-2,81
275	0,30	0,37	1,59	-0,25	-0,41	-1,43
277	0,31	0,32	1,77	0,07	0,21	-5,68
278	0,12	0,28	1,67	-0,27	-0,34	-2,31
279	0,22	-0,14	-0,21	0,98	0,46	0,39
283	0,21	0,36	1,85	-0,41	-0,75	-6,50
286	1,23	1,26	2,10	-0,23	-1,71	-0,19
287	-0,06	0,26	0,76	-0,53	-0,48	-1,21
294	0,28	0,22	0,91	0,19	-0,04	-0,94

Table D3. Streamflow elasticity results. $biOLS$ and $uniOLS$ stands for the bivariate and univariate OLS model, NP stands for the non-parametric method.

Appendix E – Synthesis of the key packages used

This thesis was developed with the use of codes mostly written in the R programming language with few codes also written in Python. In this appendix a short description regarding the purpose of the key packages used is presented in the following tables. The final one presents information of the packages used for data preprocessing, images generation and the performance metrics used.

Description	R package and version	Citation
PELT change point detection for changes in the mean value	*changepoint (2.2.2)*	Killick R, Eckley IA (2014). "changepoint: An R Package for Changepoint Analysis." _Journal of Statistical Software_, *58*(3), 1-19. <URL: http://www.jstatsoft.org/v58/i03/>.
Bai and Perron change point algorithm	*strucchange (1.5-2)*	Achim Zeileis, Friedrich Leisch, Kurt Hornik and Christian Kleiber (2002). strucchange: An R Package for Testing for Structural Change in Linear Regression Models. Journal of Statistical Software, 7(2), 1-38. URL http://www.jstatsoft.org/v07/i02/
		Achim Zeileis, Christian Kleiber, Walter Kraemer and Kurt Hornik (2003). Testing and Dating of Structural Changes in Practice. Computational Statistics & Data Analysis, 44, 109-123.
Pettitt's Test for the median value	*trend (1.1.2)*	Thorsten Pohlert (2020). trend: Non-Parametric Trend Tests and Change-Point Detection. R package version 1.1.2. https://CRAN.R-project.org/package=trend
Wavelet Coherence Analysis	*WaveletComp (1.1)*	Angi Roesch and Harald Schmidbauer (2018). WaveletComp: Computational Wavelet Analysis. R package version 1.1. https://CRAN.R-project.org/package=WaveletComp

Table E1. R packages used in the first article.

Description	Package version	Citation
Bayesian Networks structure learning (Hill Climbing), parameter calculation, Bayesian inference and network structure comparison	*R - bnlearn (4.5)*	Marco Scutari (2010). Learning Bayesian Networks with the bnlearn R Package. Journal of Statistical Software, 35(3), 1-22. URL http://www.jstatsoft.org/v35/i03/.
ARTIVA structure learning and change point detection	*R - ARTIVA (1.2.3)*	S. Lebre and G. Lelandais. (2015). ARTIVA: Time-Varying DBN Inference with the ARTIVA (Auto Regressive TIme VArying) Model. R package version 1.2.3. https://CRAN.R-project.org/package=ARTIVA
PCMCI structure learning	*Python - Tigramite (4.2)*	J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, D. Sejdinovic, Detecting and quantifying causal associations in large nonlinear time series datasets. Sci. Adv. 5, eaau4996 (2019). https://advances.sciencemag.org/content/5/11/eaau4996

Table E2. Packages used in the second and third article.

Description	R Package version	Citation
Mutual Information calculation	*muti (1.0.0)*	Scheuerell, M. D. (2017) muti: An R package for computing mutual information. https://doi.org/10.5281/zenodo.439391
Complex Network Analysis, community detection and network comparison metrics	*igraph (1.2.5)*	Csardi G, Nepusz T: The igraph software package for complex network research, InterJournal, Complex Systems 1695. 2006. http://igraph.org
Bridge Closeness metric	*networktools (1.2.3)*	Payton Jones (2020). networktools: Tools for Identifying Important Nodes in Networks. R package version 1.2.3. https://CRAN.R-project.org/package=networktools

Table E3. R Packages used in the fourth article.

Description	R package and version	Citation
R version used. R Base functions were used to perform the Shapiro-Wilk normality test and to remove seasonality	*3.6.2 (2019-12-12)*	R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Index calculation	*ncdf4 (1.17)*	David Pierce (2019). ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files. R package version 1.17. https://CRAN.R-project.org/package=ncdf4
SRI calculation	*SCI (1.0-2)*	Gudmundsson, L. & Stagge, J. H. (2016). SCI: Standardized Climate Indices such as SPI, SRI or SPEI. R package version 1.0-2
Tukey Ladder of Powers transformation	*rcompanion (2.3.26)*	Salvatore Mangiafico (2020). rcompanion: Functions to Support Extension Education Program Evaluation. R package version 2.3.26. https://CRAN.R-project.org/package=rcompanion
Streamflow performance metrics	*hydroGOF (0.4-0)*	Mauricio Zambrano-Bigiarini. (2020) hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. R package version 0.4-0. URL https://github.com/hzambran/hydroGOF. DOI:10.5281/zenodo.839854
Spatial, network and other plots	*ggplot2 (3.3.3)*	H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016
	igraph (1.2.5)	Csardi G, Nepusz T: The igraph software package for complex network research, InterJournal, Complex Systems 1695. 2006. http://igraph.org
	corrplot (0.84)	Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from https://github.com/taiyun/corrplot
Interactive plots	*plotly (4.9.2.1)*	C. Sievert. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC Florida, 2020.
	visNetwork (2.0.9)	Almende B.V., Benoit Thieurmel and Titouan Robert (2019). visNetwork: Network Visualization using 'vis.js' Library. R package version 2.0.9. https://CRAN.R-project.org/package=visNetwork
Sankey diagram	*networkD3 (0.4)*	J.J. Allaire, Christopher Gandrud, Kenton Russell and CJ Yetman (2017). networkD3: D3 JavaScript Network Graphs from R. R package version 0.4. https://CRAN.R-project.org/package=networkD3

Table E4. Further R packages used.